Algorithmic Invariant Theory of Nonreductive Groups

Tobias Kamke and Gregor Kemper
Technische Universität München, Zentrum Mathematik - M11
Boltzmannstr. 3, 85 748 Garching, Germany
kamke@ma.tum.de, kemper@ma.tum.de

October 18, 2011

Abstract

The main purpose of this paper is to give a survey of algorithms in invariant theory, with emphasis on nonreductive groups and on recent developments. But the article has some novel elements: It contains a new algorithm for computing invariant rings, which works under the condition that the invariant field is the field of fractions of the invariant ring. We also prove that if n is the dimension of the invariant ring, then there exists a separating set of invariants whose size is at most $2n + 1$.

Introduction

Broadly speaking, invariant theory comes into play whenever there is symmetry. This is why invariant theory has applications to areas such as computer vision, material science, geometric classification, molecular dynamics, equivariant dynamical systems, and symmetric differential equations. In these applications, symmetry is given by a group, and invariants serve to parametrize group orbits. In many applications, it is enough to have invariants with sufficiently good separating properties, while in others a generating set of the ring of invariants is required. Calculating the latter is the central problem of algorithmic invariant theory.

The groups whose invariant theory is guaranteed to be well-behaved are the reductive groups. However, the relevance of invariant theory is not limited to this class of groups. This is evident from the observation that groups of (or including) translations occur naturally (and turn out to have well-behaved invariants). In the theory of symmetric differential equations, nonreductive groups have recently gained importance by the article of Gaeta et al. [19]. On the other hand, the computational theory of invariants of nonreductive groups received some impetus from fairly recent papers such as Hubert and Kogan [25], Derksen and Kemper [9], Kamke [27], and Dufresne [14].

This article aims to give a survey of computational invariant theory with an emphasis on nonreductive groups. We present a new method for computing an invariant ring $K[X]^G$, provided that the invariant field $K(X)^G$ is equal to the field of fractions of $K[X]^G$. This condition is guaranteed to be satisfied if G is a unipotent group, but also under different conditions. Therefore the scope of our algorithm is broader than that of the algorithm given in [9], which is limited to unipotent groups (see also Sancho de Salas [42]). Our method also uses a different approach, and does not iterate over a composition series of G. The algorithm produces a finitely generated localization $K[X]_a^G$ of the invariant ring, which can then be fed into a routine for computing the invariant ring itself (which terminates if and only if $K[X]^G$ is finitely generated), or refined so that a representation of the invariant ring as the ring of regular functions on a quasi-affine variety can be constructed.

The article also addresses the topic of separating invariants. We think that this is of interest since separating invariants are a natural concept in particular for nonreductive group (there always exist finitely many of them, which need not be true for generating invariants), and because separating invariants are suitable for many applications. We prove the following result: If n is the Krull dimension of the invariant ring $K[X]^G$, then there exists a separating subset of size at most $2n + 1$. This result appears to be part of the folklore, but we are not aware of a proof in the literature. In fact, we prove the bound $2n + 1$ in a much more general situation, which need not be linked to invariant theory (see Theorem 5.3).

The paper is organized as follows: The first section gives a brief introduction into invariant theory, explaining the central problems and results and introducing the notation. In the second section we introduce an ideal which has come to be known as the Derksen ideal because of its use in Derksen's algorithm [7]. This ideal comes in different guises and provides a core element of other algorithms in invariant theory as well. We introduce the notion of an extended Derksen ideal, which is motivated by the concept of cross-sections from Hubert and Kogan [25] and is very useful for speeding up our above-mentioned algorithm. The third section deals with invariant fields and localizations $K[X]_a^G$ of the invariant ring. We first discuss conditions under which the invariant field is equal to the field of fractions of the invariant ring. Then we present algorithms that compute the invariant field from the Derksen ideal and a localization $K[X]_a^G$ of the invariant ring from an extended Derksen ideal. This is illustrated by an application to an example of Daigle and Freudenburg [6], where $K[X]^G$ is not finitely generated. In Section 4 we come back to the invariant ring $K[X]^G$ itself. The bulk of this section deals with methods to represent $K[X]^G$ as the ring of regular functions on a quasi-affine variety, which by a result of Nagata is always possible, provided that X is normal. We apply these methods to Daigle and Freudenburg's example. Section 5 is about separating invariants. After introducing the concept, we discuss some known results and then prove the above-mentioned upper bound on their number. The final section lists some open problems.

Readers who are interested in invariant theory in general find a huge choice of good introductory texts. Let us just mention the books by Springer [43], Kraft [32], Popov and Vinberg [41], Kraft and Procesi [33], and, more on the computational side, Sturmfels [44] and Derksen and Kemper [8].

Acknowledgment. We thank the anonymous referee for many valuable comments and for pointing out some errors. We also thank Emilie Dufresne for interesting conversations.

1 Invariant theory

Unless specified otherwise, K will always stand for an algebraically closed field. (Much of what we will say also applies to the case that K is just an infinite field, or any field if matters are interpreted scheme-theoretically, so it is for the sake of simplicity that we assume K to be algebraically closed throughout.) Moreover, G will denote a *linear algebraic group* over K. In other words, G is an affine variety (embedded into some K^m) with a group structure, where the multiplication and inversion are given by morphisms $G \times G \to G$ and $G \to G$. Typical examples for linear algebraic groups are the general linear group $\mathrm{GL}_n(K)$, the special orthogonal group $\mathrm{SO}_n(K)$, and the group $\mathrm{U}_n(K)$ of all upper triangular matrices in $K^{n\times n}$ with ones on the diagonal; but also all finite groups appear as linear algebraic groups.

We assume that G acts on an affine variety X over K such that the action is given by a morphism $G \times X \to X$. Then X is called a *G-variety*. The assumption that the action is given by a morphism (i.e., by polynomials) is rather mild in the sense that most actions of algebraic groups that mathematicians usually consider satisfy this assumption. An important special case is the case that $X = K^n$ is affine n-space and the action of G is by linear maps. In this case we usually write V instead of X and call V a *G-module*. We will write $K[X]$ for the *coordinate ring* (also known as the ring of regular functions) of X. The elements of $K[X]$ are functions $X \to K$ given by polynomials. If V is a G-module, $K[V] = K[x_1, \dots, x_n]$ is a polynomial ring in $n = \dim(V)$ indeterminates. If X is an irreducible variety, we will also consider $K(X) = \mathrm{Quot}(K[X])$, the *field of rational functions* on X.

We have a G-action on $K[X]$, given by $\sigma(f) = f \circ \sigma^{-1}$ for $\sigma \in G$ and $f \in K[X]$. The main object of interest in invariant theory is the *invariant ring*

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

Its elements, the *invariants*, may also be defined as regular functions that are constant on all G-orbits in X. Since G acts on $K[X]$ by algebra automorphisms, the invariant ring is a subalgebra of $K[V]$. If V is a G-module, the polynomial ring $K[V]$ is graded by the usual notion of homogeneous polynomials. Since the G-action on $K[V]$ sends homogeneous polynomials to homogeneous polynomials of the same degree, $K[V]^G$ is also graded.

If X is irreducible, the G-action on $K[X]$ extends to an action on $K(X)$. The *invariant field*

$$K(X)^G = \left\{f \in K(X) \mid \sigma(f) = f \text{ for all } \sigma \in G\right\}.$$

is a further object of study in invariant theory.

Example 1.1. The following examples illustrate that in invariant theory the interest often lies with the particular action considered rather than with the group.

(1) Consider the action of $G = \mathrm{GL}_n(K)$ on $V = K^{n\times n}$ by

$$G \times V \to V,\ (\sigma, A) \mapsto \sigma A \sigma^{-1}.$$

Some invariants immediately come to mind: the determinant and the trace of a matrix A. More generally, the functions $a_i\colon K^{n\times n} \to K$ mapping a matrix A to the coefficient of t^{n-i} of the characteristic polynomial $\det(tI_n - A)$ are invariants. It turns out (but requires a proof) that $K[V]^G$ is generated (as a K-algebra) by $a_1, \dots, a_n$. This means that every invariant can be written as a polynomial in the a_i. To express this, we write

$$K[V]^G = K[a_1, \dots, a_n].$$

It is easy to see that there exist no algebraic relations between the a_i, which means that the representation of an invariant in terms of the a_i is unique. While it is true that every a_i is constant on all G-orbits, there exist distinct orbits where all a_i take the same value. Examples are given by the zero matrix and a nonzero nilpotent matrix. (In fact, the orbits are parametrized by the Jordan canonical forms. They contain more information than the a_i, which just encode the eigenvalues with their algebraic multiplicities.)

(2) Consider the orthogonal group $G = \mathrm{O}_2(K) = \left\{A \in K^{2\times 2} \mid A^T A = I_n\right\}$ acting on $V = (K^2)^3 \cong K^6$ by

$$G \times V \to V,\ (A, (v_1, v_2, v_3)) \mapsto (Av_1, Av_2, Av_3).$$

V is just the threefold sum of the natural representation, so it is of no particular representation-theoretic interest. But its invariants are more interesting. Clearly the scalar products

$$f_{i,j}\colon V \to K,\ (v_1, v_2, v_3) \mapsto v_i^T v_j \quad (1 \leqslant i \leqslant j \leqslant 3)$$

are invariants. It can be shown that they generate $K[V]^G$:

$$K[V]^G = K\left[f_{1,1}, f_{1,2}, f_{1,3}, f_{2,2}, f_{2,3}, f_{3,3}\right].$$

There exists an algebraic relation between the $f_{i,j}$:

$$\det \begin{pmatrix} f_{1,1} & f_{1,2} & f_{1,3} \\ f_{1,2} & f_{2,2} & f_{2,3} \\ f_{1,3} & f_{2,3} & f_{3,3} \end{pmatrix} = 0. \tag{1.1}$$

This holds since evaluating the matrix in (1.1) at any $(v_1, v_2, v_3) \in V$ amounts to forming the product of a 3×2-matrix and its transpose, so the determinant is zero. It turns out (but requires a proof) that (1.1) generates the ideal of all relations between the $f_{i,j}$. This means that $K[V]^G$ is isomorphic to the quotient ring of a polynomial ring by a principal ideal, so the structure of $K[V]^G$ is determined. As a field extension of K, $\operatorname{Quot}\left(K[V]^G\right)$ is generated by the $f_{i,j}$, and (1.1) tells us that $f_{1,1}$ is not needed as a generator. So $\operatorname{Quot}\left(K[V]^G\right)$ is isomorphic to a rational function field. Proposition 3.1(b), which we will prove in Section 3, shows that $K(V)^G = \operatorname{Quot}\left(K[V]^G\right)$ holds in this example. ◁

In the above example we have considered the classical problems of invariant theory. These may be described by the following list:

- Hilbert's 14th problem: is $K[X]^G$ finitely generated as a K-algebra?
- If this is the case, find generators.
- If X is irreducible, find generators of the invariant field $K(X)^G$.
- What sort of an algebra is $K[V]^G$? What are its ring-theoretic properties?
- Orbit separation: given two points $p, q \in X$ whose G-orbits are distinct, does there exist an invariant $f \in K[X]^G$ with $f(p) \neq f(q)$?

In order to deal with these problems, it is useful (and indeed inevitable) to distinguish several classes of linear algebraic groups. The following are the classes of groups that are relevant in our context. More detailed information can be found in the literature on algebraic groups, e.g. Humphreys [26].

- A linear algebraic group G is called *unipotent* if for every nonzero G-module V the invariant subspace V^G is nonzero. It follows that with an appropriate choice of a basis, the action on V is given by a homomorphism of G into the group $\mathrm{U}_n(K)$ mentioned above. On the other hand, every linear algebraic group that is isomorphic to a subgroup of $\mathrm{U}_n(K)$ is unipotent (this follows from Humphreys [26, Proposition 15.2 and Theorem 15.3]). A typical example of a unipotent group is the additive group of K, which (in the context of algebraic groups) is written as G_a.
- A linear algebraic group G is called *reductive* if the only connected, normal, unipotent subgroup of G is the trivial group. Readers who are unfamiliar with algebraic groups may find this definition rather technical, so it may be more useful to note that all classical groups (such as GL_n, SL_n, O_n, SO_n, Sp_n) and all finite groups are reductive. Clearly the additive group G_a is not reductive. Every linear algebraic group G has a unique maximal connected, normal, unipotent subgroup (called the unipotent radical and written as $R_u(G)$), and $G/R_u(G)$ is reductive (see Humphreys [26, Section 19.5]).
- A linear algebraic group is called *linearly reductive* if every G-module V is the direct sum of irreducible G-modules. Although the definitions of reductive and linearly reductive groups display no similarities, it turns out that every linearly reductive group is reductive (see Kraft [32, Section II.3.5]). In fact, the two notions coincide if K has characteristic 0. On the other hand, if $p = \operatorname{char}(K)$ is positive, then a linear algebraic group G is linearly reductive if and only if its identity component G^0 is a torus (i.e., isomorphic to a direct product of multiplicative groups $G_m = \mathrm{GL}_1(K)$) and the index $(G : G^0)$ is not divisible by p.

With these notions, the following answer can be given to Hilbert's 14th problem.

Theorem 1.2 (Hilbert [24], Nagata [36], Haboush [22], Popov [40])**.** *The invariant ring $K[X]^G$ is finitely generated for all G-varieties X if and only if G is reductive.*

Notice that this theorem does not make any assertion about individual invariant rings $K[X]^G$ of nonreductive groups G. These may be (and in many examples are) finitely generated. The theorem also leaves the question open for which groups G the invariant ring $K[V]^G$ is finitely generated for all G-modules V. For example, it follows from a result of Weitzenböck [45] that this is true if $R_u(G)$ is isomorphic to G_a and $\operatorname{char}(K) = 0$.

In the sequel, when we talk about algorithms, it is important to be clear about how our mathematical objects are represented. We make the following convention.

Convention 1.3. *We assume that the linear algebraic group G and the G-variety X are given by the following data:*

- *generators of a radical ideal $I_G \subseteq K[t_1, \ldots, t_m]$ in a polynomial ring defining G as an affine variety contained in K^m;*
- *generators of a radical ideal $I_X \subseteq K[x_1, \ldots, x_n]$ in a polynomial ring defining X as an affine variety contained in K^n;*
- *polynomials $A_1, \ldots, A_n \in K[t_1, \ldots, t_m, x_1, \ldots, x_n]$ such that for $\sigma = (\gamma_1, \ldots, \gamma_m) \in G$ and $p = (\xi_1, \ldots, \xi_n) \in X$ we have*

$$\sigma(p) = \Big(A_1(\gamma_1, \ldots, \gamma_m, \xi_1, \ldots, \xi_n), \ldots, A_n(\gamma_1, \ldots, \gamma_m, \xi_1, \ldots, \xi_n)\Big).$$

Notice that this convention makes explicit our assumptions on G, X and the action. It may be remarkable that the multiplication and inversion maps of G do not enter the input data of our algorithms. It should also be noted that although K is assumed to be algebraically closed, all actual computations will be carried out in a subfield that is generated by the coefficients of the polynomials in the input data, so computations are possible.

Assume G and X are given as in Convention 1.3. Then $K[X] = K[x_1, \ldots, x_n]/I_X$, and it is easy to check that for $\sigma = (\gamma_1, \ldots, \gamma_m) \in G$ we have

$$\sigma^{-1}(x_i + I_X) = A_i(\gamma_1, \ldots, \gamma_m, x_1, \ldots, x_n) + I_X. \tag{1.2}$$

2 The Derksen ideal

Invariant theory of finite groups is a separate branch (with its algorithmic side covered by Derksen and Kemper [8, Chapter 3]). In this article we will focus on algorithms in invariant theory of infinite groups. In these algorithms, the so-called Derksen ideal, which we introduce in this section, plays a crucial role. We will also introduce the notion of an *extended Derksen ideal*, which picks up and generalizes the concept of a cross-section from Hubert and Kogan [25]. The (extended) Derksen ideal comes in three guises: algebraic, geometric, and algorithmic. We discuss them here, starting with the algebraic (and most general) notion.

Definition 2.1. *Let G be a group acting on a ring R. (By a ring we mean a commutative ring with unity.) Let $a_1, \ldots, a_n \in R$ be elements, and let $y_1, \ldots, y_n$ be indeterminates.*

(a) The ideal

$$D := \bigcap_{\sigma \in G} \Big(y_1 - \sigma(a_1), \ldots, y_n - \sigma(a_n)\Big) \subseteq R[y_1, \ldots, y_n]$$

in the polynomial ring over R is called the Derksen ideal*. It is clear that D depends not only on R and G but also on the choice of the a_i.*

(b) An ideal $E \subseteq R[y_1, \ldots, y_n]$ is called an extended Derksen ideal *if the following conditions are satisfied:*

(i) $D \subseteq E$,

(ii) E is G-stable (with G acting trivially on the y_i), and

(iii) $R \cap E = \{0\}$.

It is clear that the Derksen ideal D itself is an extended Derksen ideal.

In order to give the (extended) Derksen ideal a geometric interpretation, we assume that $R = K[X]$ is the coordinate ring of an affine variety and that G acts by algebra automorphisms. Then the choice of $a_1, \dots, a_n \in R$ defines a morphism

$$f\colon X \to K^n =: W,\ p \mapsto (a_1(p), \dots, a_n(p)),$$

and $R[y_1, \dots, y_n]$ is the coordinate ring of $X \times W$. Now it is straightforward to check that the Derksen ideal is the vanishing ideal in $K[X \times W]$ of the set

$$\Delta := \{(p,q) \in X \times W \mid \text{there exists } \sigma \in G \text{ such that } f(\sigma(p)) = q\}. \tag{2.1}$$

Since D is a radical ideal, it corresponds to the Zariski closure $\overline{\Delta}$. It often happens that the a_i generate $K[X]$ and are algebraically independent. Then we can identify X and W and obtain

$$\Delta := \{(p,q) \in X \times X \mid \text{there exists } \sigma \in G \text{ such that } \sigma(p) = q\},$$

which is sometimes called the *graph of the action*.

Going back to the general situation where the a_i are not assumed to be algebraically independent or generators of R, we give the following geometric interpretation of an extended Derksen ideal. Let $Y \subseteq W$ be a Zariski-closed subset and consider the set

$$\mathcal{E} := \{(p,q) \in X \times Y \mid \text{there exists } \sigma \in G \text{ such that } f(\sigma(p)) = q\} \subseteq \Delta$$

and its vanishing ideal

$$E := \mathrm{Id}_{K[X \times W]}(\mathcal{E}) \subseteq R[y_1, \dots, y_n].$$

Then E satisfies the conditions (i) and (ii) from Definition 2.1(b). Moreover, it is straightforward to see that (iii) is equivalent to the condition that the union

$$\bigcup_{\sigma \in G} \sigma\left(f^{-1}(Y)\right) \subseteq X$$

of G-translates of the closed subset $f^{-1}(Y) \subseteq X$ is dense in X. (Equivalently, the set of points in X whose orbit passes through $f^{-1}(Y)$ is dense.) This condition is closely related to the concept of cross-sections from Hubert and Kogan [25, Section 3.1], which in fact motivated our definition of extended Derksen ideals. However, cross-sections in the sense of [25] are a more restrictive concept than extended Derksen ideals even in the geometric situation that we are considering here. Notice that there is a lot of freedom of choice for an extended Derksen ideal (even more than for cross-sections in the sense of [25]), with $Y = W$ being one possibility yielding $E = D$.

We now turn to the algorithmic treatment of (extended) Derksen ideals, for which we make the following assumptions: G is a linear algebraic group defined by a radical ideal $I_G \subseteq K[t_1, \dots, t_m]$ as in Convention 1.3, R is a K-algebra, and $a_1, \dots, a_n \in R$ are elements such that there exist polynomials $A_1, \dots, A_n \in R[t_1, \dots, t_m]$ with

$$\sigma^{-1}(a_i) = A_i(\gamma_1, \dots, \gamma_m) \quad \text{for all} \quad \sigma = (\gamma_1, \dots, \gamma_m) \in G. \tag{2.2}$$

Notice that by (1.2), this assumption is satisfied under our standard hypotheses.

Theorem 2.2. *Assume the above notation and hypotheses.*

(a) Let

$$\widehat{D} := \left(I_G \cup \{y_1 - A_1, \dots, y_n - A_n\}\right) \subseteq R[t_1, \dots, t_m, y_1, \dots, y_n]$$

be the ideal in the polynomial ring $R[t_1, \dots, t_m, y_1, \dots, y_n]$ generated by I_G and the polynomials $y_i - A_i$. Then the Derksen ideal is

$$D = R[y_1, \dots, y_n] \cap \widehat{D}.$$

(b) Let $J = (f_1, \ldots, f_r) \subseteq K[y_1, \ldots, y_n]$ be an ideal (or, more generally, $J = (f_1, \ldots, f_r) \subseteq R[y_1, \ldots, y_n]$ a G-stable ideal) and set

$$\widehat{E} := \Big(I_G \cup J \cup \{y_1 - A_1, \ldots, y_n - A_n\}\Big) \subseteq R[t_1, \ldots, t_m, y_1, \ldots, y_n].$$

If

$$R \cap \Big(I_G \cup \{f_1(A_1, \ldots, A_n), \ldots, f_r(A_1, \ldots, A_n)\}\Big) = \{0\} \tag{2.3}$$

(with the round brackets denoting the ideal formed in $R[t_1, \ldots, t_m]$), then

$$E := R[y_1, \ldots, y_n] \cap \widehat{E}$$

is an extended Derksen ideal.

Proof. (a) Let $f \in R[y_1, \ldots, y_n]$ be a polynomial. Clearly $f - f(A_1, \ldots, A_n) \in \widehat{D}$. If $f \in D$, then

$$f\left(A_1(\underline{\gamma}), \ldots, A_n(\underline{\gamma})\right) = f\left(\sigma^{-1}(a_1), \ldots, \sigma^{-1}(a_n)\right) = 0$$

for all $\sigma = (\underline{\gamma}) \in G$, where we write $(\underline{\gamma})$ for $(\gamma_1, \ldots, \gamma_m)$. So $f(A_1, \ldots, A_n) \in (I_G) \subseteq R[t_1, \ldots, t_m]$. This implies $f \in \widehat{D}$.

Conversely, if $f \in \widehat{D}$, then also $f(A_1, \ldots, A_n) \in \widehat{D}$, so $f(A_1, \ldots, A_n) \in (I_G) \subseteq R[t_1, \ldots, t_m]$ (since $f(A_1, \ldots, A_n)$ does not involve the y_i). Hence if $\sigma \in G$, then

$$f\left(\sigma^{-1}(a_1), \ldots, \sigma^{-1}(a_n)\right) = f\left(A_1(\sigma), \ldots, A_n(\sigma)\right) = 0,$$

so $f \in \left(y_1 - \sigma^{-1}(a_1), \ldots, y_n - \sigma^{-1}(a_n)\right) \subseteq R[y_1, \ldots, y_n]$. Since this holds for all $\sigma \in G$, we conclude $f \in D$.

(b) The condition (i) from Definition 2.1(b) follows from $\widehat{D} \subseteq \widehat{E}$ and part (a). To establish condition (iii), let $f \in R \cap E$. Then $f \in \widehat{E}$, so

$$f = \sum_{i=1}^{k} h_i g_i + \sum_{i=1}^{r} \widetilde{h}_i f_i + \sum_{i=1}^{n} \widehat{h}_i \left(y_i - A_i\right) \tag{2.4}$$

with $h_i, \widetilde{h}_i, \widehat{h}_i \in R[\underline{t}, \underline{y}]$ and $g_i \in I_G$. Setting $y_i := A_i$ in this equation yields

$$f = \sum_{i=1}^{k} h_i(\underline{A}) g_i + \sum_{i=1}^{r} \widetilde{h}_i(\underline{A}) f_i(\underline{A}) \in \Big(I_G \cup \{f_1(A_1, \ldots, A_n), \ldots, f_r(A_1, \ldots, A_n)\}\Big),$$

so $f = 0$ by (2.3).

Now we prove condition (ii). Since G is a linear algebraic group embedded in K^m, there exist polynomials $p_1, \ldots, p_m \in K[t_1, \ldots, t_m, s_1, \ldots, s_m]$ in $2m$ indeterminates such that for $\sigma = (\underline{\gamma})$ and $\tau = (\underline{\eta}) \in G$, the ith component of the product $\tau\sigma$ is $p_i(\underline{\eta}, \underline{\gamma})$. Let $\sigma = (\underline{\gamma}) \in G$ and define a homomorphism of $R[\underline{y}]$-algebras

$$\Phi\colon R[\underline{t}, \underline{y}] \to R[\underline{t}, \underline{y}],\ t_i \mapsto p_i(\underline{t}, \underline{\gamma}).$$

Then for $\tau = (\underline{\eta}) \in G$ (and with G acting trivially on the t_i) we have

$$\begin{aligned}\left(\sigma\left(\Phi(A_i)\right)\right)(\underline{\eta}) &= \sigma\left(\Phi(A_i)(\underline{\eta})\right) = \sigma\left(A_i\left(p_1(\underline{\eta}, \underline{\gamma}), \ldots, p_m(\underline{\eta}, \underline{\gamma})\right)\right) \\ &= \sigma\left(A_i(\tau\sigma)\right) = \sigma\left((\tau\sigma)^{-1}(a_i)\right) = \tau^{-1}(a_i) = A_i(\underline{\eta}),\end{aligned}$$

so

$$\sigma\left(\Phi(A_i)\right) - A_i \in (I_G) \subseteq R[\underline{t}]. \tag{2.5}$$

Moreover, for $g \in I_G$ and $\tau = (\underline{\eta}) \in G$ we have

$$\big(\sigma\left(\Phi(g)\right)\big)\left(\underline{\eta}\right) = \big(\Phi(g)\big)\left(\underline{\eta}\right) = g\left(p_1(\underline{\eta}, \underline{\gamma}), \ldots, p_m(\underline{\eta}, \underline{\gamma})\right) = g(\tau\sigma) = 0,$$

so

$$\sigma\left(\Phi(g)\right) \in I_G. \tag{2.6}$$

Let $f \in E$. Then f has a representation as in (2.4), so

$$\begin{aligned} \sigma(f) &= \sigma\left(\Phi(f)\right) \\ &= \sum_{i=1}^{k} \sigma\left(\Phi(h_i)\right) \underbrace{\sigma\left(\Phi(g_i)\right)}_{\underset{(2.6)}{\in} I_G} + \sum_{i=1}^{r} \sigma\left(\Phi(\widetilde{h}_i)\right) \underbrace{\sigma(f_i)}_{\in J} \\ &\quad + \sum_{i=1}^{n} \sigma\left(\Phi(\widehat{h}_i)\right) \Big(\underbrace{y_i - A_i}_{\in \widehat{E}} + \underbrace{A_i - \sigma\left(\Phi(A_i)\right)}_{\underset{(2.5)}{\in} (I_G)}\Big) \in \widehat{E}. \end{aligned}$$

This completes the proof. □

Theorem 2.2 tells us that an (extended) Derksen ideal can be calculated as an *elimination ideal*, i.e., the intersection of an ideal with a polynomial ring in fewer indeterminates. Also the condition (2.3) can be checked by computing an elimination ideal, but this is much easier since no y-variables are involved. We obtain algorithms for computing (extended) Derksen ideals in (at least) two important cases:

(a) In the case of our standard hypotheses, assume that G and X are given as in Convention 1.3 and set $R = K[X]$ and $a_i := x_i + I_X$. Then D can be computed by standard Gröbner basis methods as follows. Form the ideal $\widehat{D}$ in $K[t_1, \ldots, t_m, x_1, \ldots, x_n, y_1, \ldots, y_n]$ generated by I_X, I_G, and the $y_i - A_i$. Choose an $\{x_1, \ldots, x_n, y_1, \ldots, y_n\}$-elimination ordering "$\leqslant$" (e.g., a lexicographical ordering with $x_i < t_j$ and $y_i < t_j$) and compute a Gröbner basis $\mathcal{G}$ of $\widehat{D}$ with respect to "$\leqslant$". Then taking those polynomials from $\mathcal{G}$ which do not involve the t-variables and mapping them into $K[X][y_1, \ldots, y_n]$ (by taking residue classes modulo I_X) produces a generating set of D. Given an ideal J as in Theorem 2.2(b), the checking of (2.3) and the computation of E can be done in the same manner. More details on the computation of elimination ideals can be found in the literature on Gröbner bases (or in Kemper [31, Section 9.2] or [8, Section 1.2]).

(b) Suppose that in the above situation X is irreducible and let R be the field of rational functions $K(X)$. Then D and E can be computed as elimination ideals directly by Theorem 2.2, using $K(X)$ as the coefficient field. Since Gröbner basis computations over such a complicated coefficient field may be hard to implement and tend to be very costly, an alternative is to compute the (extended) Derksen ideal as in the above case, but view the coefficients of the generators as elements of $K(X)$. If the monomial ordering is chosen with the additional property that $s_1 < s_2$ for two monomials s_i in the y-variables implies $ss_1 < s_2$ for all monomials s in the x-variables, then the resulting generating set of D or $E \subseteq K(X)[y_1, \ldots, y_n]$ will even be a Gröbner basis.

The calculation of Gröbner bases in a polynomial ring with many variables ($2n + m$ in the case of (a)) can be very costly and sets a practical limit to the applicability of the algorithms. This is where extended Derksen ideals can be very beneficial: a good choice of the ideal J will result in an effective reduction of the number of variables (see Remark 3.5). Extended Derksen ideals will be used in Section 3, while for the rest of this section we concentrate on the plain Derksen ideal.

The Derksen ideal owes its name to Derksen's algorithm (see [7]), which serves for computing invariant rings of linearly reductive groups. We will present the algorithm now. As a preparation, let us mention that if G is linearly reductive, there exists a unique linear map $\mathcal{R}\colon K[X] \to K[X]^G$ that is constant on all G-orbits and that restricts to the identity map on $K[X]^G$. This map is called the *Reynolds operator*, and it is in fact a homomorphism of modules over $K[X]^G$. In the case of a G-module V, $\mathcal{R}$ preserves the grading of $K[V]$.

Algorithm 2.3 (Derksen's algorithm)**.**

Input: a linearly reductive group G and a G-module V, given as in Convention 1.3.

Output: invariants generating $K[V]^G$ as a K-algebra.

(1) Compute generators of the Derksen ideal D. (An algorithm for this is given in (a) above; the elements $a_i \in K[V]$ that should be used for the formation of the Derksen ideal are also given in (a): they are just the indeterminates generating the polynomial ring $K[V]$.)

(2) Set $y_i := 0$ in all generators of D. Let $f_1, \ldots, f_m \in K[V]$ be the resulting polynomials.

(3) Apply the Reynolds operator to the f_i. Then the $\mathcal{R}(f_i)$ generate $K[V]^G$.

The point about the polynomials f_i computed in step 2 is they were shown by Derksen to generate the so-called *Hilbert ideal* $(K[V]_+^G) \subseteq K[V]$, which by definition is the ideal in $K[V]$ generated by the invariants with zero constant coefficient. From this, a standard argument, which was already known in Hilbert's times, yields that the $\mathcal{R}(f_i)$ generate $K[V]^G$ as a K-algebra.

Example 2.4. It takes 0.13 seconds to use Derksen's algorithm (implemented in the computer algebra system MAGMA [2]) for the verification of the claim in Example 1.1(2) that the invariant ring of $G = \mathrm{O}_2(K)$ is generated by the scalar products. ◁

It is important to note that the Reynolds operator is not part of the input data according to Convention 1.3. In fact, applying the Reynolds operator can be rather tricky and depends on the type of the group. Therefore it is useful to substitute step 3 by

(3') For each $i = 1, \ldots, m$, compute a basis of the space of homogeneous invariants of degree $\deg(f_i)$. Then the union of these bases generates $K[V]^G$.

Before discussing how homogeneous invariants of given degree can be computed, let us mention that the f_i obtained in the algorithm will automatically be homogeneous, so step 3' is correct since $\mathcal{R}$ preserves the grading.

The algorithm for computing homogeneous invariants will later be needed in a more general setting. Let I_X and I_G be as in Convention 1.3. Choose a monomial ordering of $K[t_1, \ldots, t_m, x_1, \ldots, x_n]$ and let $\mathcal{G}$ and $\mathcal{H}$ be Gröbner bases of I_X and I_G, respectively, with respect to the monomial ordering restricted to $K[x_1, \ldots, x_n]$ and $K[t_1, \ldots, t_m]$. Then $\mathcal{G} \cup \mathcal{H}$ is a Gröbner basis of the ideal $(I_X \cup I_G) \subseteq K[t_1, \ldots, t_m, x_1, \ldots, x_n]$ generated by I_X and I_G. This follows since elements of $\mathcal{G}$ and of $\mathcal{H}$ have coprime leading monomials (see [31, Exercise 9.3]). So a polynomial $f \in K[t_1, \ldots, t_m, x_1, \ldots, x_n]$ lies in $(I_X \cup I_G)$ if and only if its normal form $\mathrm{NF}_{\mathcal{G}\cup\mathcal{H}}(f)$ is zero. Since the normal form does not depend on the choice of the Gröbner bases, we will write it as $\mathrm{NF}_{I_X,I_G}(f)$, suppressing the dependence on the chosen monomial ordering. Now let $h \in K[x_1, \ldots, x_n]$ and assume the situation of Convention 1.3. Then

$$\begin{aligned} h + I_X \in K[X]^G \quad &\Longleftrightarrow \quad h(A_1, \ldots, A_n) - h \in (I_X \cup I_G) \subseteq K[t_1, \ldots, t_m, x_1, \ldots, x_n] \\ &\Longleftrightarrow \quad \mathrm{NF}_{I_X,I_G}\left(h(A_1, \ldots, A_n) - h\right) = 0. \end{aligned} \tag{2.7}$$

An important fact is the K-linearity of the normal form map (see [31, Theorem 9.9(b)]). So writing f as a sum of monomials with unknown coefficients and imposing the invariance condition (2.7) yields a system of linear equations for the unknown coefficients. We can now state our algorithm. In most applications, the ideal I' (that is part of the input data, see below) will be equal to $K[x_1, \ldots, x_n]$, and the set S (also part of the input) will consist of all monomials whose degree is equal to or bounded above by some given d. If $X = V$ is a G-module, $I' = K[x_1, \ldots, x_n]$ and S consists of all monomials of degree d, then the algorithm will compute a basis of the space of homogeneous invariants of degree d.

Algorithm 2.5 (Computing spaces of invariants)**.**

Input: a linear algebraic group G, a G-variety X (given as in Convention 1.3), an ideal $I' \subseteq K[x_1, \ldots, x_n]$ with $I_X \subseteq I'$, and a finite subset $S = \{h_1, \ldots, h_l\} \subseteq K[x_1, \ldots, x_n]$.

Output: a basis B of the K-vector space

$$K[X]^G \cap I'/I_X \cap \left\{ \sum_{i=1}^{l} \alpha_i h_i + I_X \mid \alpha_i \in K \right\}.$$

(1) Compute a basis C of the vector space of all $(\alpha_1, \ldots, \alpha_l) \in K^l$ with

$$\sum_{i=1}^{l} \alpha_i \operatorname{NF}_{I_X, I_G} (h_i(A_1, \ldots, A_n) - h_i) = 0 \quad \text{and} \quad \sum_{i=1}^{l} \alpha_i \operatorname{NF}_{I'}(h_i) = 0 \quad \text{for } i = 1, \ldots, l.$$

(See above for the notation and how to compute the normal forms. Computing C amounts to solving a homogeneous system of linear equations for the α_i.)

(2) Compute a basis C' of the vector space of all $(\alpha_1, \ldots, \alpha_l) \in K^l$ with

$$\sum_{i=1}^{l} \alpha_i \operatorname{NF}_{I_X}(h_i) = 0 \quad \text{for } i = 1, \ldots, l.$$

(3) Select a subset $C'' \subseteq C$ such that $C' \cup C''$ is linearly independent and $|C' \cup C''| = |C|$.

(4) Set

$$B := \left\{ \sum_{i=1}^{l} \alpha_i h_i + I_X \mid (\alpha_1, \ldots, \alpha_l) \in C'' \right\}.$$

With this algorithm, the variant of Derksen's algorithm using step 3' can be put into practice. Since Algorithm 2.5 requires the (pre-)computation of much smaller Göbner bases than Algorithm 2.3 and in addition only polynomial arithmetic and linear algebra, the cost of step 3' will be dwarfed by the cost of step 1.

An extension of Derksen's algorithm computes $K[X]^G$ for a G-variety X with G linearly reductive. In fact, one can embed X into a G-module V (see Derksen and Kemper [9, Algorithm 1.2] for an algorithm which does that). Then it is a consequence of the linear reductivity of G that the natural map $K[V] \to K[X]$ stays surjective when restricted to the invariants, so the calculation of $K[X]^G$ reduces to the calculation of $K[V]^G$.

The situation is also rather satisfactory for reductive groups. Kemper [28] gave an algorithm for calculating $K[V]^G$ in the case that G is reductive and V is a G-module. This algorithm uses separating invariants, which will be discussed in Section 5 of this paper, and an important step is the computation of the Derksen ideal. Moreover Derksen and Kemper [9, Algorithm 1.7] reduced the problem of calculating $K[X]^G$ for a G-variety X of a reductive group G to the calculation of some $K[V]^G$. So Theorem 1.2 has been made completely constructive.

However, there remains one gap where the problem of finding an algorithm for computing invariants of reductive groups is still open: if R is a finitely generated K-algebra and G is a reductive group acting on R as in (2.2), then Nagata [36] tells us that R^G is finitely generated. But we do not have an algorithm that computes generators. The point is that R is not assumed to be reduced, so it cannot be interpreted as the coordinate ring of an affine variety. See Kamke [27, Algorithm 2.5] for a solution of this problem in the case that G is finite, and [27, Section 2.2] for a discussion of the difficulties in the general situation.

3 Invariant fields and localized invariant rings

In this section we assume K, G, X, I_X, and I_G to be as introduced in Section 1 and Convention 1.3. In addition, we assume X to be an irreducible variety. We first make some remarks on the difference between $K(X)^G$, the invariant field, and $\operatorname{Quot}\left(K[X]^G\right)$, the field of fractions of the invariant ring. The question whether these two fields coincide is sometimes referred to as the *Italian problem* (see Mukai [34, page 183]). A typical example where $K(X)^G$ and $\operatorname{Quot}\left(K[X]^G\right)$ are different is the action of the multiplicative group G_m on $V = K^2$ by $\sigma(\xi_1, \xi_1) = (\sigma\xi_1, \sigma\xi_2)$ for $\sigma \in G_m$ and $\xi_i \in K$. In this example, $K[V]^{G_m} = K$ but $K(V)^{G_m} = K(x_1/x_2)$.

For every $a \in K(X)^G$, the set

$$Z(a) := \{b \in K[X] \mid ba \in K[X]\} \subseteq K[X]$$

is a nonzero, G-stable ideal. This ideal is important because of the equivalence

$$a \in \operatorname{Quot}\left(K[X]^G\right) \quad\Longleftrightarrow\quad K[X]^G \cap Z(a) \neq \{0\}.$$

If $a = \frac{f+I_X}{g+I_X}$ with $f, g \in K[X]$, $g \notin I_X$, then $Z(a)$ can be computed as the *colon ideal*

$$Z(a) = \left((I_X + (g)) : (f)\right)/I_X \quad \text{with} \quad (I_X + (g)) : (f) := \{h \in K[X] \mid hf \in I_X + (g)\}. \tag{3.1}$$

See [8, page 16] for an algorithm that computes colon ideals. We mention two situations where $K(X)^G = \operatorname{Quot}\left(K[X]^G\right)$ holds.

Proposition 3.1. *The equality* $K(X)^G = \operatorname{Quot}\left(K[X]^G\right)$ *holds if*

(a) the identity component G^0 *is unipotent or*

(b) $K[X]$ *is a factorial ring (= a unique factorization domain), its group of units coincides with* $K \setminus \{0\}$*, and there exists no surjective homomorphism (i.e., group homomorphism that is a morphism of varieties)* $G \to G_m$ *onto the multiplicative group.*

Proof. (a) Let $a \in K(X)^G$ and choose a nonzero $b \in Z(f)$. Since the G-action on $K[X]$ is locally finite (see [8, Lemma A.1.8]), there exists a (finite-dimensional) G-module $V \subseteq K[X]$ with $b \in V$. So $W := Z(a) \cap V$ is also a G-module, and it is nonzero. Since G^0 is unipotent, $W^{G^0} \neq \{0\}$. For $c \in W^{G^0} \setminus \{0\}$, the product $\prod_{\sigma \in G/G^0} \sigma(c)$ is nonzero and lies in $K[X]^G \cap Z(a)$. This means that $a \in \operatorname{Quot}\left(K[X]^G\right)$.

(b) Let $a = b/c \in K(X)^G$ with $b, c \in K[X]$, which we may assume to be coprime. Then for every $\sigma \in G$, the equation $\sigma(b) \cdot c = b \cdot \sigma(c)$ implies that $\sigma(c)$ and c are associated elements, so the hypothesis on the units on $K[X]$ implies that

$$\chi(\sigma) := \frac{\sigma(c)}{c} \in K \setminus \{0\}.$$

It is easy to check that $\chi\colon G \to G_m$ is a group homomorphism. Since the G-action on $K[X]$ is locally finite, there exists a G-module $V \subseteq K[X]$ with $c \in V$. It follows that the G-stable subspace $K \cdot c \subseteq V$ is also a G-module, so χ is a morphism of varieties. By hypothesis, χ is not surjective. Since the image $\chi(G)$ is Zariski-closed (see Humphreys [26, Section 7.4, Proposition B(b)]), it follows that the image is finite. So the product $\prod_{\sigma \in G/\ker(\chi)} \sigma(c)$ is a nonzero element of $K[X]^G \cap Z(a)$. This show that $a \in \operatorname{Quot}\left(K[X]^G\right)$. □

Notice that if the commutator subgroup of G has finite index, then G satisfies the last assumption of Proposition 3.1(b). For example, this holds for the special linear groups, the orthogonal groups, and the special orthogonal groups.

Remark. By Hashimoto [23, Proposition 5.1], the hypothesis on the group of units of $K[X]$ in Proposition 3.1(b) can be dropped. On the other hand, [27, Example 3.15] shows that the hypothesis that $K[X]$ is factorial cannot be replaced by the weaker hypothesis that $K[X]$ is normal. ◁

We now come to the main result of this section. Recall that a Gröbner basis $\mathcal{G}$ is called *reduced* if no monomial of a polynomial in $\mathcal{G}$ is divisible by the leading monomial of another polynomial in $\mathcal{G}$, and if all leading coefficients are 1. For a ring R and an element $a \in R$, $R_a := R[a^{-1}]$ denotes the localization of R with respect to $\{a^i \mid i \in \mathbb{N}_0\}$. Part (a) of the following theorem has appeared in Kemper [29] in a slightly more general form (also see Müller-Quade and Beth [35] and Hubert and Kogan [25, Theorem 2.16]). The case of (not extended) Derksen ideals of part (b) has appeared in the dissertation [27].

Theorem 3.2. *Let K, G, X, I_X, and I_G be as introduced in Section 1 and Convention 1.3, with X irreducible. Set $a_i := x_i + I_X \in K[X]$, and let*

$$D = \bigcap_{\sigma \in G} \big(y_1 - \sigma(a_1), \ldots, y_n - \sigma(a_n)\big) \subseteq K(X)[y_1, \ldots, y_n]$$

be the Derksen ideal and $E \subseteq K(X)[y_1, \ldots, y_n]$ an extended Derksen ideal.

(a) Let $\mathcal{G} \subseteq K(X)[y_1, \ldots, y_n]$ be a reduced Gröbner basis (with respect to an arbitrary monomial ordering) of D. Then $K(X)^G$ is generated as a field extension of K by the coefficients of all polynomials in $\mathcal{G}$.

(b) Assume $K(X)^G = \operatorname{Quot}\big(K[X]^G\big)$ and let $\mathcal{G} \subseteq K(X)[y_1, \ldots, y_n]$ be a reduced Gröbner basis of E. Choose a nonzero invariant $a \in K[X]^G$ such that there exists $k \in \mathbb{N}$ with $a^k g \in K[X][y_1, \ldots, y_n]$ for every $g \in \mathcal{G}$. Then $K[X]_a^G$ is generated as a K-algebra by a^{-1} and the coefficients of all $a^k g$ with $g \in \mathcal{G}$.

Proof. (a) Clearly D is G-stable. Since $\sigma(\mathcal{G})$ is a reduced Gröbner basis for every $\sigma \in G$, it follows from the uniqueness of reduced Gröbner bases (see Becker and Weispfenning [1, Theorem 5.43]) that $\mathcal{G} \subseteq K(X)^G[y_1, \ldots, y_n]$. (Notice that the G-action is only on $K(X)$, which is the field of coefficients over which $\mathcal{G}$ lives.) So the field extension L of K generated by the coefficients of all polynomials in $\mathcal{G}$ is contained in $K(X)^G$.

Conversely, let $b \in K(X)^G$. We can write b as

$$b = \frac{g(a_1, \ldots, a_n)}{h(a_1, \ldots, a_n)} \tag{3.2}$$

with $g, h \in K[y_1, \ldots, y_n]$ polynomials. The G-invariance of b implies $g - bh \in D$. (Recall that g and h are polynomials in the y-variables, while b is an element of $K(X)$.) Using the linearity of the normal form map, this implies

$$0 = \mathrm{NF}_{\mathcal{G}}(g - bh) = \mathrm{NF}_{\mathcal{G}}(g) - b\,\mathrm{NF}_{\mathcal{G}}(h).$$

Assume that $\mathrm{NF}_{\mathcal{G}}(h) = 0$. Then $h \in D \subseteq (y_1 - a_1, \ldots, y_n - a_n)$, so $h(a_1, \ldots, a_n) = 0$, contradicting (3.2). We obtain

$$b = \frac{\mathrm{NF}_{\mathcal{G}}(g)}{\mathrm{NF}_{\mathcal{G}}(h)}.$$

Since g, h, and $\mathcal{G}$ are contained in $L[y_1, \ldots, y_n]$, we also have $\mathrm{NF}_{\mathcal{G}}(g), \mathrm{NF}_{\mathcal{G}}(h) \in L[y_1, \ldots, y_n]$, so the above equation tells us

$$b \in L(y_1, \ldots, y_n) \cap K(X) = L.$$

This shows that $K(X)^G \subseteq L$, and the proof of (a) is complete.

(b) It suffices to show that $K[X]^G$ is contained in the K-algebra A generated by a^{-1} and the coefficients of all $a^k g$ with $g \in \mathcal{G}$. If $b \in K[X]^G$, then the above argument (with $h = 1$ and D substituted by E) yields

$$\mathrm{NF}_{\mathcal{G}}(f) = b\,\mathrm{NF}_{\mathcal{G}}(1) = b,$$

with the last equality following from $1 \notin E$ (see Definition 2.1(b)). Since f and $\mathcal{G}$ are contained in $A[y_1, \ldots, y_n]$ and since all polynomials in $\mathcal{G}$ have leading coefficient 1, we obtain

$$b \in A[y_1, \ldots, y_n] \cap K[X] = A.$$

With this, the proof is complete. □

Remark 3.3. We can also prove the following variant of Theorem 3.2(a):

(a') Assume that $K[X]$ is a factorial ring and that E arises from an ideal $\widehat{E} \subseteq K(X)[t_1, \ldots, t_m, y_1, \ldots, y_n]$ as in Theorem 2.2(b), with $J \subseteq K[y_1, \ldots, y_n]$ satisfying (2.3). Let $\mathcal{G} \subseteq K(X)[y_1, \ldots, y_n]$ be a reduced Gröbner basis of E. Then $K(X)^G$ is generated as a field extension of K by the coefficients of all polynomials in $\mathcal{G}$.

This is a generalization of Theorem 3.7 of Hubert and Kogan [25]. The proof is similar to the one of Theorem 3.2(a), but much more work goes into showing that $\mathrm{NF}_{\mathcal{G}}(h) \neq 0$, which requires the factoriality assumption. We omit the proof. ◁

It is clear that part (a) of the above theorem gives rise to an algorithm for computing the invariant field $K(X)^G$ of a linear algebraic group G. (The computation of a Gröbner basis of the Derksen ideal was discussed in (b) on page 8, and turning this into a reduced Gröbner basis is easy.) Before coming to the algorithm arising from part (b), we will discuss the option (and merits) of using an extended Derksen ideal. The following theorem is similar to Theorem 3.3 of Hubert and Kogan [25].

Theorem 3.4. *In the situation of Theorem 3.2, let r be the maximal dimension of a G-orbit in X (which is attained on a nonempty Zariski-open subset of X, see [31, page 146]). Then there exist $\alpha_{i,j} \in K$ and $\beta_i \in K$ such that the*

$$f_i := \sum_{j=1}^{n} \alpha_{i,j} y_j - \beta_i \in K[y_1, \ldots, y_n] \quad (i = 1, \ldots, r)$$

are algebraically independent and satisfy the condition (2.3) from Theorem 2.2 with $R := K(X)$. So the ideal $E \subseteq R[y_1, \ldots, y_n]$, defined in Theorem 2.2(b), is an extended Derksen ideal.

Remark 3.5. By the special form of the f_i, the number of variables occurring in the ideal $\widehat{E} \subseteq R[t_1, \ldots, t_m, y_1, \ldots, y_n]$ from Theorem 2.2(b) is effectively reduced to $m+n-r$. This brings a huge benefit for the computation of the elimination ideal E. Additionally, the Gröbner basis $\mathcal{G}$ from Theorem 3.2(b) will become smaller, so picking out coefficients will result in a smaller generating set of $K[X]_a^G$.

From the proof of Theorem 3.4 we will see that "most" choices of $\alpha_{i,j}$ and β_i will work. ◁

Proof of Theorem 3.4. It follows from [31, page 145] that for the set $\Delta \subseteq X \times W$ defined in (2.1) we have $\dim(\overline{\Delta}) = \dim(X) + r$. This set corresponds to the Derksen ideal formed in $K[X][\underline{y}]$. For the Derksen ideal $D \subseteq R[\underline{y}]$ from Theorem 2.2, it follows that

$$\dim\big(R[\underline{y}]/D\big) = r.$$

So we can choose r of the y_i such that the corresponding $y_i + D$ are algebraically independent. Alternatively, we may use Noether normalization (see Eisenbud [16, Theorem 13.3]) to choose K-linear combinations

$$z_i = \sum_{j=1}^{n} \alpha_{i,j} y_j + D \in R[\underline{y}]/D \quad (i = 1, \ldots, r)$$

that are algebraically independent. The injective map

$$A := R[z_1, \ldots, z_r] \subseteq R[\underline{y}]/D \hookrightarrow R[\underline{t}, \underline{y}]/\widehat{D}$$

(with $\widehat{D} \subseteq R[\underline{t}, \underline{y}]$ from Theorem 2.2) induces a dominant map $\operatorname{Spec}\left(R[\underline{t}, \underline{y}]/\widehat{D}\right) \to \operatorname{Spec}(A)$. It follows from a theorem of Chevalley (see [31], page 144]) that there exists a nonempty open subset of $\operatorname{Spec}(A)$ that is contained in the image. So there is a nonzero $f \in A$ such that for every prime ideal $P \in \operatorname{Spec}(A)$ with $f \notin P$ there exists $Q \in \operatorname{Spec}\left(R[\underline{t}, \underline{y}]/\widehat{D}\right)$ such that $P = A \cap Q$. We can choose $\beta_1, \ldots, \beta_r \in K$ such that $f(\beta_1, \ldots, \beta_r) \neq 0$, so

$$f \notin P := (z_1 - \beta_1, \ldots, z_r - \beta_r) \in \operatorname{Spec}(A),$$

and it follows that the ideal generated by P in $R[\underline{t}, \underline{y}]/\widehat{D}$ is proper. So with $f_1, \ldots, f_r$ defined as in the statement of the theorem, the ideal

$$\widehat{D} + (f_1, \ldots, f_r) \subseteq R[\overline{t}, \overline{y}]$$

is proper. By way of contradiction, assume that the condition (2.3) from Theorem 2.2 is violated. Then

$$1 \in \Big(I_G \cup \{f_1(A_1, \ldots, A_n), \ldots, f_r(A_1, \ldots, A_n)\}\Big) \subseteq R[\underline{t}].$$

But since $y_i - A_i \in \widehat{D}$, this implies $1 \in \widehat{D} + (f_1, \ldots, f_r)$, a contradiction. This completes the proof. □

Now we can turn Theorem 3.2(b) into an algorithm.

Algorithm 3.6 (Computation of a localization of the invariant ring)**.**

Input: a linear algebraic group G, a G-variety X (given as in Convention 1.3), such that X is irreducible and $K(X)^G = \operatorname{Quot}\left(K[X]^G\right)$. (See Proposition 3.1 for the last assumption.)

Output: invariants $a, b_1, \ldots, b_m \in K[X]^G$ with $a \neq 0$ such that

$$K[X]^G_a = K[a^{-1}, b_1, \ldots, b_m].$$

(1) This step is optional. With r equal to (or less than) the maximal dimension of a G-orbit in X, search for $\alpha_{i,j}, \beta_i \in K$ such that

$$K[x_1, \ldots, x_n] \cap \Big(I_X \cup I_G \cup \big\{\sum_{j=1}^{n} \alpha_{i,j} A_j - \beta_i \mid i = 1, \ldots, r\big\}\Big) \subseteq I_X.$$

(The round brackets denote the ideal formed in $K[t_1, \ldots, t_m, x_1, \ldots, x_n]$.)

(2) With additional indeterminates $y_1, \ldots, y_n$, form the ideal

$$\widehat{E} := \Big(\, I_G \cup \Big\{\, \sum_{j=1}^{n} \alpha_{i,j} y_j - \beta_i \mid i = 1, \ldots, r \,\Big\} \cup \big\{y_i - (A_i + I_X) \mid i = 1, \ldots, n\big\} \,\Big)$$
$$\subseteq K(X)[t_1, \ldots, t_m, y_1, \ldots, y_n].$$

(If step 1 was omitted, set $r = 0$ so the second set of generators is empty.)

(3) Compute a Gröbner basis $\mathcal{G} \subseteq K(X)[y_1, \ldots, y_n]$ (with respect to an arbitrary monomial ordering) of the (extended) Derksen ideal

$$E := K(X)[y_1, \ldots, y_n] \cap \widehat{E}.$$

(See (b) on page 8 on how to do this.)

(4) If necessary, modify $\mathcal{G}$ to turn it into a reduced Gröbner basis.

(5) Let $c_1, \ldots, c_k \in K(X)$ be the coefficients of the polynomials in $\mathcal{G}$. *The remaining steps are concerned with finding a common G-invariant denominator a and multiplying the c_i by a.*

(6) For each $i = 1, \dots, k$, let $c_i = \frac{f_i + I_X}{g_i + I_X}$ with $f_i, g_i \in K[x_1, \dots, x_n]$, and compute the colon ideal $I_i' := (I_X + (g_i)) : (f_i) \subseteq K[x_1, \dots, x_n]$. *(So by (3.1) we have $I_i'/I_X = Z(c_i)$.)*

(7) Form the intersection $I' := \bigcap_{i=1}^k I_i'$. (See [8, page 15] on how to compute intersections of ideals. *I'/I_X will now be the ideal of all $a \in K[X]$ with $ac_i \in K[X]$ for all i.)*

(8) For $d = 0, 1, 2, \dots$, repeat steps 9 and 10.

(9) Let $h_1, \dots, h_l \in K[x_1, \dots, x_n]$ be all monomials of degree $\leqslant d$ and use Algorithm 2.5 to compute a basis B of $K[X]^G \cap I'/I_X \cap \left\{ \sum_{i=1}^l c_i h_i + I_X \mid c_i \in K \right\}$.

(10) If $B \neq \emptyset$, choose $a \in B$ and go to step 11. *(Then $a \in K[X]^G \setminus \{0\}$ and $ac_i \in K[X]$ for all i.)*

(11) Write $a = g + I_X$ with $g \in K[x_1, \dots, x_n]$. For $i = 1, \dots, k$, repeat steps 12 and 13.

(12) Using the extended Buchberger algorithm (see Becker and Weispfenning [1, Section 5.6]), compute a Gröbner basis $\mathcal{G}_i$ of $I_X + (g_i)$ and representations of the elements of $\mathcal{G}_i$ as $K[x_1, \dots, x_n]$-linear combinations of g_i and the generators of I_X.

(13) By computing the normal form $\mathrm{NF}_{\mathcal{G}_i}(f_i g)$ (which will be 0 since $g \in I' \subseteq I_i'$), express $f_i g$ as a $K[x_1, \dots, x_n]$-linear combination of the elements of $\mathcal{G}_i$ and then, using the results of step 12, as a $K[x_1, \dots, x_n]$-linear combination of g_i and the generators of I_X. Let $\widehat{f_i} \in K[x_1, \dots, x_n]$ be the coefficient of g_i in this linear combination. *(Then $\widehat{f_i} g_i + I_X = f_i g + I_X$, so $\frac{\widehat{f_i} + I_X}{a} = c_i$.)*

(14) Set $b_i := \widehat{f_i} + I_X \in K[X]$. Then

$$K[X]_a^G = K\left[a^{-1}, b_1, \dots, b_k\right].$$

Remark 3.7. (a) Often Algorithm 3.6 will produce a lot of unnecessary generators b_i since picking out coefficients of the Gröbner basis elements in step 5 tends to produce an abundance of elements c_i. So cleaning up the resulting generators may be desirable. Step 1 also helps to reduce the number of c_i. Moreover, when choosing the element $a \in K[X]^G$, one may make use of the fact only a power of a needs to be a common denominator of the c_i, not a itself.

(b) The following is a variant of steps 12 and 13 which avoids using the extended Buchberger algorithm: Iterating over $d = 0, 1, 2, \dots$, let $h_1, \dots, h_l \in K[x_1, \dots, x_n]$ be all monomials of degree $\leqslant d$. If there exist $\alpha_1, \dots, \alpha_l \in K$ such that

$$\sum_{j=1}^l \alpha_j \, \mathrm{NF}_{I_X}(h_j g_i) = \mathrm{NF}_{I_X}(f_i g)$$

(this is an inhomogeneous system of linear equations for the α_j), then $\widehat{f_i} := \sum_{j=1}^l \alpha_j h_j$ satisfies $\widehat{f_i} g_i + I_X = f_i g + I_X$. ◁

Let us consider an example. The (in some sense) smallest example known to date of a nonfinitely generated invariant ring was given by Daigle and Freudenburg [6]. So it is tempting to run our algorithm on this example.

Example 3.8. Daigle and Freudenburg's example is an action of the additive group G_a on the polynomial ring $\mathbb{C}[x_1, \dots, x_5]$, which is best given in terms of the nilpotent derivation

$$\mathcal{D} = x_1^3 \frac{\partial}{\partial x_2} + x_2 \frac{\partial}{\partial x_3} + x_3 \frac{\partial}{\partial x_4} + x_1^2 \frac{\partial}{\partial x_5},$$

so

$$\mathbb{C}[x_1, \dots, x_5]^{G_a} = \ker(\mathcal{D}).$$

Converting the action to make it compatible with Convention 1.3 yields an action given by the polynomials

$$A_1 = x_1, \quad A_2 = x_2 + t x_1^3, \quad A_3 = x_3 + t x_2 + \frac{t^2}{2} x_1^3,$$
$$A_4 = x_4 + t x_3 + \frac{t^2}{2} x_2 + \frac{t^3}{6} x_1^3, \quad \text{and} \quad A_5 = x_5 + t x_1^2.$$

It is clear that typical G-orbits are 1-dimensional, so we may choose one linear combination of the A_i that generates an ideal intersecting trivially with $\mathbb{C}[x_1, \ldots, x_n]$. Clearly A_1 does not qualify, but A_2 does. We obtain the ideal

$$\begin{aligned}\widehat{E} &= (y_2, y_1 - A_1, \ldots, y_5 - A_5) \\ &= \Big(y_1 - x_1, y_2, x_2 + t x_1^3, y_3 - x_3 - t x_2 - \frac{t^2}{2} x_1^3, \\ &\quad y_4 - x_4 - t x_3 - \frac{t^2}{2} x_2 - \frac{t^3}{6} x_1^3, y_5 - x_5 - t x_1^2 \Big) \subseteq \mathbb{C}(x_1, \ldots, x_5)[t, y_1, \ldots, y_5]\end{aligned}$$

from Theorem 2.2. Using the lexicographical ordering with $t > y_1 > y_2 > \cdots > y_5$, we compute a Gröbner basis of $\widehat{E}$. The third generator has the leading term $x_1^3 t$, so replacing all other generators by their normal forms with respect to the third generator amounts to substituting $t = -x_2/x_1^3$. This leads to a new generating set of $\widehat{E}$:

$$\widehat{E} = \left(\underline{y_1} - x_1,\ \underline{y_2},\ \underline{t} + \frac{x_2}{x_1^3},\ \underline{y_3} - \frac{2x_1^3 x_3 - x_2^2}{2x_1^3}, \underline{y_4} - \frac{3x_1^6 x_4 - 3x_1^3 x_2 x_3 + x_2^3}{3x_1^6}, \underline{y_5} - \frac{x_1 x_5 - x_2}{x_1} \right),$$

where we have underlined the leading monomials. Now we see that this is already a reduced Gröbner basis, so deleting the third generator yields a reduced Gröbner basis of the extended Derksen ideal E. (Of course, this is no coincidence, but is a consequence of the fact A_2 has degree 1 as a polynomial in t.) Picking out coefficients produces a generating set for the invariant ring:

$$\mathbb{C}(x_1, \ldots, x_5)^{G_a} = \mathbb{C}(f_1, \ldots, f_4).$$

with

$$f_1 = x_1, \quad f_2 = 2x_1^3 x_3 - x_2^2, \quad f_3 = 3x_1^6 x_4 - 3x_1^3 x_2 x_3 + x_2^3, \quad \text{and} \quad f_4 = x_1 x_5 - x_2.$$

So the invariant field is isomorphic to a rational function field. We also see that a power of $a = x_1$ is a common denominator of the polynomials in the Gröbner basis of E. (Again, this is no coincidence, but comes from the fact that a power of x_1 is the coefficient of t in A_2.) Therefore

$$\mathbb{C}[x_1, \ldots, x_5]_{f_1}^{G_a} = \mathbb{C}[f_1^{-1}, f_1, f_2, f_3, f_4].$$

So the localized invariant ring is isomorphic to a localized polynomial ring, which is the simplest possible structure. It seems amazing that in spite of all this simplicity, the invariant ring itself is not finitely generated. ◁

4 Invariant rings of nonreductive groups

This section is a sequel of the previous one, so we continue to assume that G is a linear algebraic group and X is a G-variety which (as a variety) is irreducible. Suppose that we have computed (by Algorithm 3.6 or some other means) a finitely generated subalgebra $A \subseteq K[X]^G$ and a nonzero $a \in A$ such that $K[X]_a^G = A_a$. Then

$$K[X]^G = \{ f \in K[X] \mid \text{ there exists } k \in \mathbb{N} \text{ such that } a^k f \in A \}$$

In general, if $R \subseteq S$ are rings and $I \subseteq R$ is a subset, we define

$$(R : I^\infty)_S := \{f \in S \mid \text{ there exists } k \in \mathbb{N} \text{ such that } g_1 \cdots g_k \cdot f \in R \text{ for all } g_1, \dots, g_k \in I\}$$

which is a subring of S containing R. Since replacing I by the ideal in R generated by I does not change $(R : I^\infty)_S$, I may be assumed to be an ideal in R. In terms of this definition, the above equation reads

$$K[X]^G = (A : \{a\}^\infty)_{K[X]} . \tag{4.1}$$

In Derksen and Kemper [9, Algorithm 2.6] we find a procedure for computing $(R : I^\infty)_S$ in the case that S is a finitely generated domain over a field, $R \subseteq S$ is a finitely generated subalgebra, and $I \subseteq R$ is an ideal. This procedure is a *pseudo-algorithm* in the following sense:

- It terminates after finitely many steps if $(R : I^\infty)_S$ is finitely generated as a K-algebra.
- If $(R : I^\infty)_S$ is not finitely generated, it keeps producing new generators forever.
- While the procedure is running, it cannot be determined whether $(R : I^\infty)_S$ is finitely generated.

Combining this with Algorithm 3.6, we get an algorithm for computing $K[X]^G$ if it is finitely generated and $K(X)^G = \operatorname{Quot}\left(K[X]^G\right)$ holds. Since by Proposition 3.1(a) the latter condition is satisfied if G is unipotent, we can apply this to the unipotent radical $R_u(G)$ (assuming that $R_u(G)$ is known or can be calculated). If $K[X]^{R_u(G)}$ is finitely generated, we can write it as the coordinate ring of an affine variety Y. Then

$$K[X]^G = K[Y]^{G/R_u(G)}.$$

Since the factor group $G/R_u(G)$ is reductive, we have an algorithm (Algorithm 1.7 in [9]) for computing $K[Y]^{G/R_u(G)}$, so we obtain an algorithm for computing $K[X]^G$ in the case that $K[X]^{R_u(G)}$ is finitely generated. However, this is not completely satisfactory since it happens that $K[X]^{R_u(G)}$ is not finitely generated, but $K[X]^G$ is.

But even if $K[X]^G$ is not finitely generated, there is a way to give it a "finite description". In fact, if X is normal (which is always the case if $K[X]$ is a polynomial ring), then by Nagata [37, Chapter V, Proposition 4], $K[X]^G$ is isomorphic to the ring of regular functions on a quasi-affine variety, i.e., a Zariski-open subset of an affine variety. (See Winkelmann [46] for a modern proof and some extensions.) In more concrete terms, this result can be expressed as follows: if X is normal, then there exists a finitely generated subalgebra $A \subseteq K[X]^G$ and an ideal $I \subseteq A$ such that

$$K[X]^G = (A : I^\infty)_{\operatorname{Quot}(A)} . \tag{4.2}$$

The connection with quasi-affine varieties is as follows: if we write A as the coordinate ring of an affine variety Y, then I determines a closed subset $Z \subseteq Y$, and for the affine variety $U := Y \setminus Z$, the ring of regular functions $K[U]$ is isomorphic to $(A : I^\infty)_{\operatorname{Quot}(A)}$ (see [9, Lemma 2.3]). So

$$K[X]^G \cong K[U].$$

It should be noted that rings of regular functions on a quasi-affine variety need not be finitely generated. How can this result be made constructive? While (4.1) and (4.2) look deceitfully similar, they are in fact quite different. But the following result shows that the conversion of (4.1) into (4.2) can be performed if we have enough representations of $K[X]^G$ as (4.1).

Proposition 4.1 (Dufresne [14]). *Assume that X is normal. Let $A \subseteq K[X]^G$ a subalgebra and $a_1, a_2 \in A$ such that*

$$A_{a_i} = K[X]^G_{a_i} \quad \textit{for} \quad i = 1, 2.$$

If the ideal $(a_1, a_2) \subseteq K[X]$ generated by the a_i in $K[X]$ has height at least 2, then

$$K[X]^G = (A : \{a_1, a_2\}^\infty)_{\operatorname{Quot}(A)} .$$

Proof. Let $f \in K[X]^G$. Then there exist $k_i \in \mathbb{N}$ such that $a_i^{k_i} f \in A$. Setting $k := k_1 + k_2 - 1$, we see that every product of f and k of the a_i lies in A, so $f \in (A : \{a_1, a_2\}^\infty)_{\mathrm{Quot}(A)}$.

To prove the converse, we first remark that for every prime ideal $P \in \mathrm{Spec}\,(K[X])$ with $\mathrm{ht}(P) = 1$ there exists $i \in \{1, 2\}$ such that $a_i \notin P$. Therefore

$$K[X]_{a_1} \cap K[X]_{a_2} \subseteq \bigcap_{\substack{P \in \mathrm{Spec}(K[X]) \\ \text{with } \mathrm{ht}(P)=1}} K[X]_P = K[X],$$

where the last equation holds since X is normal (see Eisenbud [16, Corollary 11.4]). Now let $f \in (A : \{a_1, a_2\}^\infty)_{\mathrm{Quot}(A)}$. Then there exists $k \in \mathbb{N}$ such that $a_i^k f \in A \subseteq K[X]$, so $f \in K[X]_{a_1} \cap K[X]_{a_2} = K[X]$. Moreover, $f \in \mathrm{Quot}(A) \subseteq K(X)^G$, so $f \in K[X]^G$. □

Of course Proposition 4.1 only produces a representation of $K[X]^G$ as the ring of regular functions on a quasi-affine variety if A is finitely generated. This motivates the study of the set

$$F_R := \{a \in R \mid R_a \text{ is finitely generated as a } K\text{-algebra}\},$$

where R stands for any K-algebra. It turns out that F_R is always a radical ideal in R (see Onoda and Yoshida [39] or [9, Proposition 2.9]). Following [9], we call F_R the *finite generation ideal* of R. It follows from Theorem 3.2(b) that if $K(X)^G = \mathrm{Quot}\left(K[X]^G\right)$, then the finite generation ideal of $K[X]^G$ is nonzero. But much more is true: in fact, if $R \subseteq A$ is a subalgebra of a finitely generated domain A over a field (or even over a ring), then F_R is nonzero (see Giral [20, Proposition 2.1(b)] or [31, Exercise 10.3]).

Returning to our situation, the question is whether the ideal $(F_{K[X]^G}) \subseteq K[X]$ generated by the finite generation ideal of $K[X]^G$ has height at least 2. Then there exist a_1, a_2 and a finitely generated subalgebra $A \subseteq K[X]^G$ satisfying the hypotheses of Proposition 4.1. Derksen and Kemper [9] gave an affirmative answer for the following situation: if $K[X]$ is a factorial ring and G is connected and unipotent, then $(F_{K[X]^G}) \subseteq K[X]$ has height at least 2. The paper [9] also contains an algorithm (Algorithm 3.9) for computing $K[X]^G$ as a ring of regular functions on a quasi-affine variety in this situation. However, this algorithm seems to be rather impractical: when applying it to the example of Daigle and Freudenburg [6] (see Example 3.8), the Gröbner basis computations quickly become too hard to perform.

Using Proposition 4.1 but choosing the a_i in a more ad hoc fashion is a more promising approach, as the following continuation of Example 3.8 shows.

Example 4.2. We use the notation of Example 3.8. In Example 3.8, we observed that the localization $\mathbb{C}[x_1, \ldots, x_5]_{x_1}^{G_a}$ is finitely generated because A_2 has degree 1 as a polynomial in t, and a power of the invariant x_1 appears as the coefficient of t in A_2. A further invariant with this property comes from the observation that

$$\mathcal{D}(3x_1^3 x_4 - x_2 x_3) = 2x_1^3 x_3 - x_2^2 = f_2.$$

We make use of this observation by "artificially" introducing the further generator $a_0 := 3x_1^3 x_4 - x_2 x_3$ of $\mathbb{C}[x_1, \ldots, x_5]$. The corresponding polynomial A_0 defining the G_a-action on a_0 is

$$A_0 = 3A_1^3 A_4 - A_2 A_3 = a_0 + f_2 t.$$

Now we proceed as in Example 3.8. This time we choose (A_0) as an ideal which intersects trivially with $\mathbb{C}[x_1, \ldots, x_5]$. We obtain the ideal

$$\begin{aligned}\widehat{E} &= (y_0, A_0, y_1 - A_1, \ldots, y_5 - A_5) \\ &= \Big(y_0, a_0 + f_2 t, y_1 - x_1, y_2 - x_2 - t x_1^3, y_3 - x_3 - t x_2 - \frac{t^2}{2} x_1^3, \\ &\qquad y_4 - x_4 - t x_3 - \frac{t^2}{2} x_2 - \frac{t^3}{6} x_1^3, y_5 - x_5 - t x_1^2\Big) \subseteq \mathbb{C}(x_1, \ldots, x_5)[t, y_0, \ldots, y_5]\end{aligned}$$

from Theorem 2.2(b). Using the lexicographical ordering with $t > y_0 > y_1 > \cdots > y_5$, we compute a Gröbner basis of $\widehat{E}$. The second generator has the leading term $f_2 t$, so replacing all other generators by their normal forms with respect to the second generator amounts to substituting $t = -a_0/f_2$. It is clear that this leads to a reduced Gröbner basis of $\widehat{E}$ (with leading monomials y_0, t, y_1, y_2, y_3, y_4, and y_5). It is also clear that the denominators occurring in the Gröbner basis are powers of f_2. Excluding the second generator, we get a Gröbner basis of E. It remains to extract the numerators of the coefficients of this Gröbner basis and express them as polynomials in a small number of invariants. Using the computer algebra system MAPLE [5], we verified that the following invariants suffice:

$$f_1 = x_1, \quad f_2 = 2x_1^3x_3 - x_2^2, \quad f_4 = x_1x_5 - x_2 \quad \text{(as in Example 3.8)},$$
$$f_5 = 2x_1^3x_3x_5 + x_1^2x_2x_3 - x_2^2x_5 - 3x_1^5x_4, \quad \text{and} \quad f_6 = \frac{f_2^3 + (f_2f_4 - f_1f_5)^2}{f_1^6}.$$

(It is confirmed by computation that f_6 is a polynomial and also that $f_3 = f_2f_4 - f_1f_5$ with f_3 as in Example 3.8.) In fact, we get

$$E = \left(y_0, y_1 - f_1, y_2 + \frac{f_2f_4 - f_1f_5}{f_2}, y_3 - \frac{f_1^3f_6}{2f_2^2}, y_4 + \frac{f_6(f_2f_4 - f_1f_5)}{6f_2^3}, y_5 - \frac{f_5}{f_2}\right).$$

So with

$$A := \mathbb{C}[f_1, f_2, f_4, f_5, f_6],$$

Theorem 3.2(b) and the results from Example 3.8 tell us that

$$\mathbb{C}[x_1, \ldots, x_5]_{f_i}^{G_a} = A_{f_i} \quad \text{for } i = 1, 2.$$

Clearly $(f_1, f_2) \subseteq \mathbb{C}[x_1, \ldots, x_5]$ has height 2, so

$$\mathbb{C}[x_1, \ldots, x_5]^{G_a} = (A : \{f_1, f_2\}^\infty)_{\mathrm{Quot}(A)}$$

by Proposition 4.1. To write this as the ring of regular functions on a quasi-affine variety, we need the relations between the generators of A. It is obvious that the relation $f_1^6f_6 - f_2^3 - (f_2f_4 - f_1f_5)^2 = 0$ (derived from the definition of f_6) generates the ideal of relations. It follows that $\mathbb{C}[x_1, \ldots, x_5]^{G_a}$ is isomorphic to the ring of regular functions on

$$\begin{aligned} U = &\left\{(\xi_1, \xi_2, \xi_4, \xi_5, \xi_6) \in \mathbb{C}^5 \mid \xi_1^6\xi_6 - \xi_2^3 - (\xi_2\xi_4 - \xi_1\xi_5)^2 = 0\right\} \\ &\setminus \left\{(\xi_1, \xi_2, \xi_4, \xi_5, \xi_6) \in \mathbb{C}^5 \mid \xi_1 = \xi_2 = 0\right\}. \end{aligned}$$

Notice that Winkelmann [46, Section 4] obtained the same generators f_i and the same quasi-affine variety U, and Dufresne [14] presents this example from a slightly different point of view. In fact, her presentation helped us finding the representations of the Gröbner basis coefficients in terms of the invariants f_i. Since Daigle and Freudenburg [6] showed that $\mathbb{C}[x_1, \ldots, x_5]^{G_a}$ is not finitely generated, we get an example of a quasi-affine variety whose ring of regular functions is not finitely generated. Although we expect that there are simpler examples of this kind known, we are not aware of any. ◁

Further examples where a nonfinitely generated invariant ring is represented as the ring of regular functions on a quasi-affine variety can be found in Dufresne [14].

5 Separating invariants

An article on invariant theory of nonreductive groups should not fail to address the topic of separating invariants. The concept of separating invariants is motivated by the study of group orbits separated by invariants (see Section 1) and by the observation that many applications of invariant theory (e.g., to computer vision, graph theory, orbit space reduction, and geometric classification) only require invariants with good separation properties. Here is the definition:

Definition 5.1. *Assume the situation introduced in Section 1. A subset $S \subseteq K[X]^G$ is called* separating *if the following condition holds for all points $p, q \in X$:*

If there exists $f \in K[X]^G$ with $f(p) \neq f(q)$, then there exists $f \in S$ with $f(p) \neq f(q)$.

(Loosely speaking, this means that S has the same capabilities of separating orbits as all of $K[X]^G$.)

It is clear that a set of generating invariants is always separating. In other words, "separating" weaker than "generating". In the last few years, a trend has emerged to consider separating invariants instead of generating ones.

Example 5.2. Consider the action of the cyclic group G of order 3 on $\mathbb{C}^2$ by scalar multiplications by third roots of unity. The invariant ring is

$$\mathbb{C}[x_1, x_2]^G = \mathbb{C}\Big[\underbrace{x_1^3}_{=:f_1}, \underbrace{x_1^2 x_2}_{=:f_2}, \underbrace{x_1 x_2^2}_{=:f_3}, \underbrace{x_2^3}_{=:f_4}\Big],$$

and the f_i form a minimal set (even a set of minimal size) of generators. However, $S = \{f_1, f_2, f_4\}$ is a separating subset. Indeed, for $p \in V$ with $f_1(p) \neq 0$ we have $f_3(p) = f_2(p)^2/f_1(p)$, and if $f_1(p) = 0$, then also $f_3(p) = 0$.

This example shows that separating sets of invariants can be strictly smaller than generating ones. ◁

As it turns out, separating invariants are in many ways much better behaved than generating ones. This is exemplified by the following results:

Finiteness: Even if $K[X]^G$ is not finitely generated, there exists a finite separating subset (see [8, Theorem 2.3.15]). A proof of a more general result can be found in Kemper [30].

Noether's degree bound: If G is finite and V is a G-module, then there exist homogeneous separating invariants of degree at most $|G|$ (see [8, Corollary 3.9.14]). Notice that Noether's degree bound holds for generating invariants if the characteristic $\mathrm{char}(K)$ does not divide $|G|$ (see Noether [38], Fleischmann [17], Fogarty [18], or [8, Section 3.8]). But it fails badly in the case that $\mathrm{char}(K)$ divides $|G|$ (the *modular case*, see [8, Section 3.9] and the references given there). In contrast, Noether's bound for separating invariants holds independently of the characteristic.

Weyl's polarization theorem: Weyl's polarization theorem, which enables the transfer of invariants from $K[V^n]^G$ (where V^n stands for the direct sum of $n = \dim(V)$ copies of V) to invariants from $K[V^m]^G$ for any m, holds for separating invariants independently of the characteristic of K (see Draisma et al. [11]). By contrast, it holds for generating invariants only in characteristic 0.

Reflection groups: Assume that G is finite and V is a G-module. If there exists a separating subset of size $n = \dim(V)$ (in other words, if there exists a separating subalgebra that is isomorphic to a polynomial ring), then G is generated by reflections (i.e., elements fixing a hyperplane pointwise). This was proved recently by Dufresne [13] and extends results by Chevalley, Shephard, Todd, and Serre.

In this paper we will prove the following result about the number of separating invariants, which quantifies the finiteness result stated above: there exists a separating separating subset $S \subseteq K[X]^G$ of size

$$|S| \leqslant 2 \dim\left(K[X]^G\right) + 1,$$

where $\dim\left(K[X]^G\right)$ denotes the Krull dimension of $K[X]^G$. Notice that $\dim\left(K[X]^G\right)$ is equal to the transcendence degree of $K[X]^G$ over K even if $K[X]^G$ is not finitely generated (see [31, Theorem 5.9 and Exercise 5.3] or Giral [20, Proposition 2.3]), and bounded above by the dimension of X. The bound on the size of a separating set seems to be part of the folklore, and a proof in the case that $X = V$ is a G-module appeared in Dufresne [12, Proposition 5.1.1]. We will formulate and prove the result in the following more general context.

Let K be an infinite field and let X be a nonempty set. We consider the K-algebra

$$\mathrm{Map}(X,K) := \{f\colon X \to K \mid f \text{ is a map}\}$$

and a subset $F \subseteq \mathrm{Map}(X,K)$. Then a further subset $S \subseteq F$ is called *F-separating* if the following condition holds for all points $p,q \in X$:

If there exists $f \in F$ with $f(p) \neq f(q)$, then there exists $f \in S$ with $f(p) \neq f(q)$.

We write

$$\gamma_{\mathrm{sep}}(F) := \min\{|S| \mid S \subseteq F \text{ is } F\text{-separating}\} \in \mathbb{N}_0 \cup \{\infty\}.$$

Moreover, we call functions $f_1,\ldots,f_n \in \mathrm{Map}(X,K)$ *algebraically independent* if $H(f_1,\ldots,f_n) \neq 0$ for all nonzero polynomials $H \in K[x_1,\ldots,x_n]$, and write

$$\mathrm{trdeg}(F) := \sup\{n \in \mathbb{N}_0 \mid \text{ there exist } f_1,\ldots,f_n \in F \text{ which are algebraically independent}\}.$$

Theorem 5.3. *In the above situation, let $F \subseteq \mathrm{Map}(X,K)$ be a subspace and assume that there exists a finite F-separating subset $S \subseteq F$. (Notice that by [30, Theorem 2.1], the last hypothesis is satisfied if F is contained in a finitely generated subalgebra of $\mathrm{Map}(X,K)$.) Then*

$$\gamma_{\mathrm{sep}}(F) \leqslant 2\,\mathrm{trdeg}(F) + 1.$$

Proof. Let $S = \{f_1,\ldots,f_k\} \subseteq F$ be an F-separating subset with k minimal. Set $n := \mathrm{trdeg}(F)$ and assume that $k > 2n+1$. The idea is to find $k-1$ suitable linear combinations of the f_i that are F-separating.

Let $\pi_1,\pi_2\colon X \times X \to X$ be the natural projections and consider the subalgebra

$$A := K\left[f_1 \circ \pi_1,\ldots,f_k \circ \pi_1, f_1 \circ \pi_2,\ldots,f_k \circ \pi_2\right] \subseteq \mathrm{Map}(X \times X, K)$$

and the polynomial algebra $A[t]$. For $i = 1,2$, an algebraically independent subset of $\{f_1 \circ \pi_i,\ldots,f_k \circ \pi_i\} \subseteq A[t]$ has size at most n, since an algebraic relation between the f_j is also an algebraic relation between the $f_j \circ \pi_i$. So an algebraically independent subset of $\{f_1 \circ \pi_1,\ldots,f_k \circ \pi_1, f_1 \circ \pi_2,\ldots,f_k \circ \pi_2, t\} \subseteq A[t]$ has size at most $2n+1$. Therefore $\mathrm{trdeg}(A[t]) \leqslant 2n+1$ (see [31, Proposition 5.10]). It follows that the

$$g_i := t \cdot (f_i \circ \pi_1 - f_i \circ \pi_2) \in A[t] \quad (i = 1,\ldots,k) \tag{5.1}$$

are algebraically dependent, so there exists a nonzero polynomial $H \in K[x_1,\ldots,x_k]$ with

$$H(g_1,\ldots,g_k) = 0. \tag{5.2}$$

Since K is infinite, there exist $\xi_1,\ldots,\xi_k \in K$ with

$$H(\xi_1,\ldots,\xi_k) \neq 0 \quad \text{and} \quad \xi_k \neq 0. \tag{5.3}$$

Set

$$\widetilde{f_i} := \xi_k f_i - \xi_i f_k \in F \quad (i = 1,\ldots,k-1). \tag{5.4}$$

By the minimality of k, the set $\widetilde{S} = \{\widetilde{f_1},\ldots,\widetilde{f_{k-1}}\}$ cannot be F-separating. So there exist points $p,q \in X$ and $f \in F$ such that $f(p) \neq f(q)$ but $\widetilde{f_i}(p) = \widetilde{f_i}(q)$ for $i = 1,\ldots,k-1$. By (5.4), this implies

$$f_i(p) - f_i(q) = \frac{\xi_i}{\xi_k}\left(f_k(p) - f_k(q)\right) \quad (i = 1,\ldots,k-1). \tag{5.5}$$

Since $\{f_1,\ldots,f_k\}$ is F-separating and $f(p) \neq f(q)$, we must have $f_i(p) \neq f_i(q)$ for some i, so by (5.5) this is true for $i = k$. Let $\Psi\colon A[t] \to K$ be the algebra homomorphism sending $f \in A$ to $f(p,q)$ and t to $\xi_k \cdot (f_k(p) - f_k(q))^{-1}$. From (5.1) and (5.5), we get

$$\Psi(g_i) = \xi_k \cdot (f_k(p) - f_k(q))^{-1} \cdot (f_i(p) - f_i(q)) = \xi_i,$$

so (5.2) implies

$$0 = \Psi\left(H(g_1, \dots, g_k)\right) = H\left(\Psi(g_1), \dots, \Psi(g_k)\right) = H(\xi_1, \dots, \xi_k),$$

contradicting (5.3). This shows that the assumption $k > 2n + 1$ was false. □

It is easy to turn the above proof into a constructive version of Theorem 5.3. In fact, given an F-separating subset $S = \{f_1, \dots, f_k\} \subseteq F$ with $k > 2\operatorname{trdeg}(F)+1$, one can find a nonzero polynomial $H \in K[x_1, \dots, x_k]$ satisfying (5.2) by writing H as a linear combination of all monomials of degree bounded above by some $d \in \mathbb{N}$ with unknown coefficients. Setting $H(g_1, \dots, g_k) = 0$ leads to a homogeneous system of linear equations for the unknown coefficients. By increasing d, one will eventually find a nonzero solution. Having found H, it is easy to find $\xi_1, \dots, \xi_k \in K$ satisfying (5.3) by using a (given) injective map $\mathbb{N} \to K$ and specializing each x_i in $x_k \cdot H$ to an image ξ_i of this map such that the result remains nonzero. Then the $\widetilde{f}_i$ defined by (5.4) will form an F-separating subset of size $k-1$. This procedure can be continued until $k \leqslant 2\operatorname{trdeg}(F)+1$. (Values of k that are smaller than $2\operatorname{trdeg}(F)+1$ are possible if a nonzero polynomial H satisfying (5.2) happens to exist even when this is not guaranteed a priori.)

When we apply this to $F = K[V]^G$ with V a G-module, it should be noted that even if we start with a homogeneous separating set, the above procedure will usually destroy the homogeneity.

In order to give readers an idea about the extent of the reduction of the size of a separating set due to Theorem 5.3, we present two tables here. The first one deals with the classical topic of *invariants of binary forms* of some degree (see [8, Example 2.1.2]), and compares the minimal number of generating invariants, written as $\gamma\left(\mathbb{C}[V]^{\mathrm{SL}_2}\right)$, to the upper bound of Theorem 5.3 on the minimal number $\gamma_{\mathrm{sep}}\left(\mathbb{C}[V]^{\mathrm{SL}_2}\right)$ of separating invariants. The values of $\gamma\left(\mathbb{C}[V]^{\mathrm{SL}_2}\right)$ are taken from Dixmier and Lazard [10] and Brouwer and Popoviciu [3, 4].

Binary forms of degree	1	2	3	4	5	6	7	8	9	10
$\gamma\left(\mathbb{C}[V]^{\mathrm{SL}_2}\right)$	0	1	1	2	4	5	30	9	92	106
$2\dim\left(\mathbb{C}[V]^{\mathrm{SL}_2}\right)+1$	1	3	3	5	7	9	11	13	15	17

The second table does the same for the invariants $\mathbb{C}[x_1, \dots, x_n]^{C_n}$ of the cyclic group of order n acting as a cyclic permutation of the x_i. The values of $\gamma\left(\mathbb{C}[V]^{C_n}\right)$ were calculated using MAGMA [2].

n	1	2	3	4	5	6	7	8	9	10	11	12	13
$\gamma\left(\mathbb{C}[V]^{C_n}\right)$	1	2	4	7	15	20	48	65	119	166	348	367	823
$2\dim\left(\mathbb{C}[V]^{C_n}\right)+1$	3	5	7	9	11	13	15	17	19	21	23	25	27

6 Some open problems

We finish this article by listing a few open questions in algorithmic invariant theory.

Hilbert's 14th problem: For which linear algebraic groups G is $K[V]^G$ finitely generated for all G-modules V? A reasonable conjecture may be that this is the case if and only if the unipotent radical $R_u(G)$ has dimension at most 1.

Test for finite generation: Find an algorithm that tests $K[X]^G$ for finite generation, where G is a given a linear algebraic group and X is a G-variety. If $K[X]^G$ is finitely generated, calculate generators. Under reasonable hypotheses, we can compute the invariant ring of the unipotent radical $R_u(G)$ of G as the ring of regular functions on a quasi-affine algebra U. So one may continue to compute invariants of $G/R_u(G)$ acting on U. The first author's dissertation [27] contains algorithms for computing invariants of some groups acting on quasi-affine varieties, but did not succeed in dealing with the case of reductive groups.

Quasi-affine varieties: Write $K[X]^G$ as the ring of regular functions on a quasi-affine variety U. Even if that were possible for general G and X, it would not provide a finite generation test since (to the best of our knowledge) no algorithm is known to test $K[U]$ for finite generation.

Separating invariants: Find an algorithm for computing separating invariants in $K[X]$ for G nonreductive. The above-mentioned proof that there exists a finite separating set is nonconstructive. Some examples of finite separating sets of nonfinitely generated invariant rings were given by Dufresne and Kohls [15] and Dufresne [14].

Nonreduced algebras: Find an algorithm for computing the invariants of a reductive group acting on a finitely generated, nonreduced K-algebra R as in (2.2) (see at the end of Section 2).

Implementations: There exist good implementations of algorithms for invariants of finite groups in MAGMA [2]. Derksen's algorithm is also implemented in MAGMA and in SINGULAR [21]. Moreover, MAGMA has an implementation of the algorithm for computing invariant fields of linear algebraic groups. However, most of the other algorithms mentioned in this paper, including the algorithms for computing $K[X]^G$ according to (4.1), have been implemented at best in an ad hoc fashion, with no implementation available to the public. Neither have any serious efforts been made to optimize these algorithms.

References

[1] Thomas Becker, Volker Weispfenning, *Gröbner Bases*, Springer-Verlag, Berlin, Heidelberg, New York 1993.

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

[3] Andries E. Brouwer, Mihaela Popoviciu, *The invariants of the binary decimic*, J. Symbolic Comput. **45** (2010), 837–843.

[4] Andries E. Brouwer, Mihaela Popoviciu, *The invariants of the binary nonic*, J. Symbolic Comput. **45** (2010), 709–720.

[5] B. Char, K. Geddes, G. Gonnet, M. Monagan, S. Watt, *Maple Reference Manual*, Waterloo Maple Publishing, Waterloo, Ontario 1990.

[6] Daniel Daigle, Gene Freudenburg, *A Counterexample to Hilbert's Fourteenth Problem in Dimension* 5, J. Algebra **221** (1999), 528–535.

[7] Harm Derksen, *Computation of Invariants for Reductive Groups*, Adv. Math. **141** (1999), 366–384.

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

[9] Harm Derksen, Gregor Kemper, *Computing invariants of algebraic group actions in arbitrary characteristic*, Adv. Math. **217** (2008), 2089–2129.

[10] J. Dixmier, D. Lazard, *Minimum number of fundamental invariants for the binary form of degree* 7, J. Symbolic Comput. **6** (1988).

[11] Jan Draisma, Gregor Kemper, David Wehlau, *Polarization of Separating Invariants*, Canad. J. Math. **60** (2008), 556–571.

[12] Emilie Dufresne, *Separating Invariants*, Dissertation, Queen's University, Kingston, Ontario, Canada 2008, `http://hdl.handle.net/1974/1407`.

[13] Emilie Dufresne, *Separating invariants and finite reflection groups*, Adv. Math. **221(6)** (2009), 1979–1989.

[14] Emilie Dufresne, *Finite separating sets and quasi-affine quotients*, preprint, Universität Basel, 2011.

[15] Emilie Dufresne, Martin Kohls, *A Finite Separating Set for Daigle and Freudenburg's Counterexample to Hilbert's Fourteenth Problem*, Comm. Algebra **38** (2010), 3987–3992.

[16] David Eisenbud, *Commutative Algebra with a View Toward Algebraic Geometry*, Springer-Verlag, New York 1995.

[17] Peter Fleischmann, *The Noether Bound in Invariant Theory of Finite Groups*, Adv. in Math. **156** (2000), 23–32.

[18] John Fogarty, *On Noether's Bound for Polynomial Invariants of a Finite Group*, Electron. Res. Announc. Amer. Math. Soc. **7** (2001), 5–7.

[19] Giuseppe Gaeta, Frank D. Grosshans, Jürgen Scheurle, Sebastian Walcher, *Reduction and reconstruction for symmetric ordinary differential equations*, J. Differential Equations **244** (2008), 1810–1839.

[20] José M. Giral, *Krull Dimension, Transcendence Degree and Subalgebras of Finitely Generated Algebras*, Arch. Math. (Basel) **36** (1981), 305–312.

[21] Gert-Martin Greuel, Gerhard Pfister, Hannes Schönemann, *Singular Version 1.2 User Manual*, Reports On Computer Algebra **21**, Centre for Computer Algebra, University of Kaiserslautern, 1998, available at `http://www.mathematik.uni-kl.de/~zca/Singular`.

[22] William J. Haboush, *Reductive Groups are Geometrically Reductive*, Ann. of Math. **102** (1975), 67–83.

[23] Mitsuyasu Hashimoto, *Equivariant total ring of fractions and factoriality of rings generated by semiinvariants*, preprint, see http://arxiv.org/abs/1009.5152, Nagoya University, 2010.

[24] David Hilbert, *Über die Theorie der algebraischen Formen*, Math. Ann. **36** (1890), 473–534.

[25] Evelyne Hubert, Irina A. Kogan, *Rational invariants of an algebraic groups action. Constructing and rewriting*, J. Symb. Comput. **42** (2007), 203–217.

[26] James E. Humphreys, *Linear Algebraic Groups*, Springer-Verlag, 1981.

[27] Tobias Kamke, *Algorithms for the Computation of Invariant Rings*, Dissertation, Technische Universität München, 2009.

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

[29] Gregor Kemper, *The Computation of Invariant Fields and a Constructive Version of a Theorem by Rosenlicht*, Transformation Groups **12** (2007), 657–670.

[30] Gregor Kemper, *Separating Invariants*, J. Symbolic Comput. **44** (2009), 1212–1222.

[31] Gregor Kemper, *A Course in Commutative Algebra*, Graduate Texts in Mathematics **256**, Springer-Verlag, Berlin, Heidelberg 2011.

[32] Hanspeter Kraft, *Geometrische Methoden in der Invariantentheorie*, Vieweg, Braunschweig 1985.

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

[34] Shigeru Mukai, *An introduction to invariants and moduli*, vol. 81 of *Cambridge Studies in Advanced Mathematics*, Cambridge University Press, Cambridge 2003, Translated from the 1998 and 2000 Japanese editions by W. M. Oxbury.

[35] Jörn Müller-Quade, Thomas Beth, *Calculating Generators for Invariant Fields of Linear Algebraic Groups*, in: *Applied Algebra, Algebraic Algorithms and Error-Correcting Codes (Honolulu, HI, 1999)*, Lecture Notes in Comput. Sci. **1719**, pp. 392–403, Springer, Berlin 1999.

[36] Masayoshi Nagata, *Invariants of a Group in an Affine Ring*, J. Math. Kyoto Univ. **3** (1963/1964), 369–377.

[37] Masayoshi Nagata, *Lectures on the Fourteenth Problem of Hilbert*, Tata Institute of Fundamental Research, Bombay 1965.

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

[39] Nobuharu Onoda, Ken-ichi Yoshida, *On Noetherian subrings of an affine domain*, Hiroshima Math. J. **12** (1982), 377–384.

[40] Vladimir L. Popov, *On Hilbert's Theorem on Invariants*, Dokl. Akad. Nauk SSSR **249** (1979), English translation Soviet Math. Dokl. **20** (1979), 1318–1322.

[41] Vladimir L. Popov, Ernest B. Vinberg, *Invariant Theory*, in: N. N. Parshin, I. R. Shafarevich, eds., *Algebraic Geometry IV*, Encyclopaedia of Mathematical Sciences **55**, Springer-Verlag, Berlin, Heidelberg 1994.

[42] Carlos Sancho de Salas, *Invariant theory for unipotent groups and an algorithm for computing invariants*, Proc. London Math. Soc. (3) **81** (2000), 387–404.

[43] Tonny A. Springer, *Invariant Theory*, Lecture Notes in Math. **585**, Springer-Verlag, Berlin, Heidelberg, New York 1977.

[44] Bernd Sturmfels, *Algorithms in Invariant Theory*, Springer-Verlag, Wien, New York 1993.

[45] Roland Weitzenböck, *Über die Invarianten von linearen Gruppen*, Acta Math. **58** (1932), 231–293.

[46] Jörg Winkelmann, *Invariant Rings and Quasiaffine Quotients*, Math. Zeitschrift **244** (2003), 163–174.