Mathematisches
Forschungsinstitut
Oberwolfach

Member of
Leibniz
Association

Oberwolfach
Preprints

Number 2024-13

Jitendra Bajpai
Daniele Dona

A CFSG-Free Explicit Jordan's Theorem over Arbitrary Fields

November 2024

ISSN 1864-7596

Mathematisches Forschungsinstitut Oberwolfach gGmbH (MFO)
Schwarzwaldstrasse 9-11
77709 Oberwolfach-Walke
Germany
https://www.mfo.de

DOI 10.14760/OWP-2024-13

A CFSG-FREE EXPLICIT JORDAN'S THEOREM OVER ARBITRARY FIELDS

JITENDRA BAJPAI AND DANIELE DONA

Abstract. We prove a version of Jordan's classification theorem for finite subgroups of $\mathrm{GL}_n(K)$ that is at the same time quantitatively explicit, CFSG-free, and valid for arbitrary K. This is the first proof to satisfy all three properties at once. Our overall strategy follows Larsen and Pink [24], with explicit computations based on techniques developed by the authors and Helfgott [2, 3], particularly in relation to dimensional estimates.

Contents

1. Introduction

Results about the structure of subgroups of $\mathrm{GL}_n(\mathbb{C})$ have been known for a long time, at least since Jordan proved the following result [23, Thm. 40].

Theorem 1.1 (Jordan's theorem). *Let Γ be a finite subgroup of $\mathrm{GL}_n(\mathbb{C})$. Then there is a normal abelian subgroup $A \trianglelefteq \Gamma$ of index bounded by a constant $J(n)$ depending only on n.*

Since then, $J(n)$ has been bounded explicitly. A bound of the form $e^{O(n^2/\log n)}$ is given in [22, Thm. 14.12], based on ideas of Frobenius and Blichfeldt, and it does not use the Classification of Finite Simple Groups (CFSG). With the aid of CFSG, Collins [9, Thm. A] proved the bound $(n+1)!$ for $n \geq 71$, which is tight in general. We refer the reader for an exposition on Jordan's theorem to a recent survey by Breuillard [6].

Date: November 18, 2024.
2020 *Mathematics Subject Classification.* Primary: 20G07, 20G15; Secondary: 14A10.
Key words and phrases. Jordan's theorem, subgroup structure, groups of Lie type, algebraic groups.

Theorem 1.1 is false if we replace $\mathbb{C}$ by a field of positive characteristic. Nevertheless, there exist results about the structure of finite subgroups of $\mathrm{GL}_n(K)$ that generalize Jordan's theorem. Here we prove one such result.

Theorem 1.2. *Let K be any field, and let Γ be a finite subgroup of $\mathrm{GL}_n(\overline{K})$. Then there are $\Gamma_3 \trianglelefteq \Gamma_2 \trianglelefteq \Gamma_1 \trianglelefteq \Gamma$, each of them normal inside Γ, such that*

(a) $|\Gamma/\Gamma_1| \leq J'(n) := n^{n^{2^{23}n^{10}}}$*;*
(b) either $\Gamma_1 = \Gamma_2$, or $\mathrm{char}(K) = p > 0$ and Γ_1/Γ_2 is a product of finite simple groups of Lie type of characteristic p;
(c) Γ_2/Γ_3 is abelian of size not divisible by $\mathrm{char}(K)$;
(d) either $\Gamma_3 = \{e\}$, or $\mathrm{char}(K) = p > 0$ and Γ_3 is a p-group.

A version of Theorem 1.2 without any explicit expression for $J'(n)$ was proved by Larsen and Pink [24], without relying on CFSG. With the use of CFSG, Collins [10] showed that we can take $J'(n) = (n+2)!$ for $n \geq 71$. Our goal is to have at the same time an explicit $J'(n)$ in the statement and a CFSG-free proof. The present paper is the first to have both properties for an arbitrary field K.

Our interest in the question stems from our previous work on *dimensional estimates*. These tools were developed first in [24, §4] in a non-explicit form, and then used in several papers in the context of Babai's conjecture. Most recently, the authors in joint work with Helfgott gave explicit dimensional estimates in order to achieve sharper diameter bounds for untwisted classical groups [2, 3], and a natural question was whether the rest of the techniques of [24] could be made equally explicit. In the present paper, we follow the procedure in [24], sharpening and cleaning the route taken by them through the strategies involved in the proofs of [2, 3].

1.1. Outline of the strategy.

The first three sections collect some preliminary facts: Section 2 concerns varieties, Section 3 deals with linear algebraic groups in general, and Section 4 focuses on almost simple groups. Some of the definitions and properties are standard, some are taken from [2, 3], and some are new although in line with the spirit of those papers.

Section 5 deals with dimensional estimates. A *dimensional estimate* is a bound of the form $|A \cap V(K)| \leq C|A^C|^{\dim(V)/\dim(G)}$, where G is an algebraic group over K, $V \subseteq G$ is a subvariety, $A \subseteq G(K)$ is a finite subset, and C is some constant depending only on the data of G and V (but not on A and K). Such an estimate appeared first in [24, Thm. 4.2], with $A = \Gamma$ a subgroup and with a non-explicit C. The bounds in [2, Thm. 4.4] and [3, Thm. 1.1] have instead an explicit C and hold for A a generating set of $G(K)$. Our task in this section is to show that the assumption $\langle A \rangle = G(K)$ can be weakened, so that we may have estimates for $A = \Gamma$ with explicit C. The section plays the role of [24, §4] through its main result (Theorem 5.3), and of [24, §6] through its corollary on centralizers (Corollary 5.6).

The goal of Section 6 is to prove an explicit version of [24, Thm. 0.5], following the path laid out in [24, §§7–11]. Its main result (Theorem 6.7) shows that, for any G almost simple and any $\Gamma \leq G(\overline{K})$, either $[G^F, G^F] \leq \Gamma \leq G^F$ for some appropriate endomorphism F, where G^F is a group of Lie type and its commutator is simple, or Γ is trapped in some substructure: either $|\Gamma|$ is bounded in terms of the rank of G, or $|\Gamma| \leq H(\overline{K})$ for some proper subgroup $H \lneq G$ of smaller dimension and bounded degree.

The path of Section 6 is articulated in several steps. Starting from the assumption that Γ is not trapped as above, we first find regular unipotent elements in Γ, thus incidentally proving that $\mathrm{char}(K)$ must be positive (Proposition 6.2)[1]. Then we find a variety V of minimal unipotent elements, representing the finite field $\mathbb{F}_q$ that "correctly determines" Γ (Proposition 6.3): when at the end $[G^F, G^F] \leq \Gamma \leq G^F$, F will be either the Frobenius map with respect to $\mathbb{F}_q$ or a twist of that map.

For now $\mathbb{F}_q$ is a good model only for the minimal unipotent elements of Γ, meaning that $\Gamma \cap V(\overline{K}) \simeq \mathbb{F}_q$ (as abelian groups). The final step of Section 6 is to prove that $\mathbb{F}_q$ is a good model for the whole Γ. We do so in two stages: first, in Propositions 6.4–6.5 we achieve our goal under some conditions on Γ (Assumption A1–A2) and on V ($\dim(V) = \mathrm{rk}(G)$); then, we use that partial case and stronger conditions on Γ (Assumption B1–B2) to complete the proof without any hypothesis on V. The general case is Theorem 6.7.

In Section 7 we complete the proof of Theorem 1.2. In rough terms, the case of $[G^F, G^F] \leq \Gamma \leq G^F$ gives rise to (b), whereas the case of $|\Gamma|$ small gives rise to (a). The descent from $\Gamma \leq G(\overline{K})$ to $\Gamma \leq H(\overline{K})$ with $\dim(H) < \dim(G)$ can be repeated until we reach either one of the other cases or $\dim(H) = 0$ (in which case $|\Gamma|$ is small again, thanks to the bound on $\deg(H)$). Since H is not necessarily almost simple, at every stage we need to take quotients by the unipotent radical and by the centre: the former is a p-group, whence (d), and the latter is abelian, whence (c).

2. Varieties

In this section we collect basic properties about varieties, morphisms, and degrees.

2.1. **Basic nomenclature.** We go over some standard terms, whose definition in the literature can vary.

A *variety*[2] V in n-dimensional affine space $\mathbb{A}^n$ is defined by a set of s equations of the form $P_i(x_1, \ldots, x_n) = 0$ (for $1 \leq i \leq s$), where all P_i are polynomials and s is any non-negative integer. V is *defined over* a field K if the coefficients of all P_i belong to K. The *set of points* $V(K)$ is

$$V(K) = \{(k_1, \ldots, k_n) \in K^n : P_i(k_1, \ldots, k_n) = 0 \ \ (1 \leq i \leq s)\}.$$

Two varieties V, W are equal if and only if the ideals generated by their defining polynomials inside the ring $\overline{K}[x_1, \ldots, x_n]$ have the same radical, which by the Nullstellensatz holds if and only if $V(\overline{K}) = W(\overline{K})$.

Let V, W be defined by polynomials $\mathcal{P} = \{P_i\}_{i \leq s}$ and $\mathcal{Q} = \{Q_j\}_{j \leq t}$. The *(set-theoretic, or reduced) intersection* $V \cap W$ is the variety defined by $\mathcal{P} \cup \mathcal{Q}$, and the *union* $V \cup W$ is defined by $\{P_i Q_j\}_{i \leq s, j \leq t}$.

The *Zariski topology* is the topology whose closed sets are the sets $V(\overline{K})$ for all varieties V. The affine space $\mathbb{A}^n(\overline{K})$ is Noetherian under this topology. The *Zariski*

[1] At this stage, a CFSG-free proof of Theorem 1.1 with explicit $J(n)$ is already within reach. We do not bother doing so, since [22, Thm. 14.12] already does that and with a better $J(n)$ than ours.

[2] For us a variety is affine and closed, not necessarily irreducible nor connected nor pure-dimensional.

closure $\overline{S}$ of a set $S \subseteq \mathbb{A}^n(\overline{K})$ is the smallest set of the form $V(\overline{K})$ containing S. We call V itself the Zariski closure of S.

A variety V is *irreducible* if it is not equal to any union $V_1 \cup V_2$ with $V_1 \not\subseteq V_2$ and $V_2 \not\subseteq V_1$. Every V can be uniquely decomposed into a finite union of irreducible varieties not contained in each other, called the *irreducible components* of V. The *dimension* $\dim(V)$ of an irreducible variety V is the largest d for which we can write a chain of irreducible proper subvarieties $V_0 \subsetneq V_1 \subsetneq \ldots \subsetneq V_d = V$. For V non-irreducible, $\dim(V)$ is the largest of the dimensions of its irreducible components. A variety is *pure-dimensional* when all its components have the same dimension.

A *morphism* $f : \mathbb{A}^n \to \mathbb{A}^m$ defined over K is an m-tuple of polynomials f_i on n variables whose coefficients belong to K. A morphism $f : X \to Y$ for $X \subseteq \mathbb{A}^n, Y \subseteq \mathbb{A}^m$ is the restriction of a morphism $g : \mathbb{A}^n \to \mathbb{A}^m$ such that $g(x) \in Y(\overline{K})$ for every $x \in X(\overline{K})$, and we write $f = g|_X$. We have $g_1|_X = g_2|_X$ if and only $g_1(x) = g_2(x)$ for all $x \in X(\overline{K})$. For a morphism $f : X \to Y$ and a subvariety $V \subseteq Y$ defined by polynomials $P_i(y_1, \ldots, y_n)$, the *preimage* $f^{-1}(V)$ is the variety defined by the polynomials $P_i(f_1(\vec{x}), \ldots, f_n(\vec{x}))$ and by the polynomials defining X. The image $f(X(\overline{K}))$ need not be the set of points of a variety, though it is a *constructible set* (Chevalley; see [27, §I.8, Cor. 2 to Thm. 3]), meaning a finite union of intersections $U \cap W$, where U is open and W is closed.

2.2. **Degrees.** Let $V \subseteq \mathbb{A}^n$ be a pure-dimensional variety over K with $\dim(V) = d$. The *degree* $\deg(V)$ of V is the number of points in the set $(V \cap L)(\overline{K})$, where L is a generic $(n-d)$-dimensional affine subspace of $\mathbb{A}^n$ (by [12, §II.3.1.2, Thm.] the definition makes sense and $\deg(V)$ is finite).

We can extend the definition of degree to general varieties V: $\deg(V)$ is the sum of the degrees of the pure-dimensional parts of V. The bound $\deg(V_1 \cup V_2) \leq \deg(V_1) + \deg(V_2)$ holds for any V_1, V_2 directly by definition. If V is the union of irreducible components V_i, then $\deg(V) = \sum_i \deg(V_i)$.

By a generalization of *Bézout's theorem* due to Fulton and Macpherson, as in [15, Ex. 8.4.6], [30, (2.26)], or [12, §II.3.2.2, Thm.], for V_1, V_2 pure-dimensional we have

$$\deg(V_1 \cap V_2) \leq \deg(V_1) \deg(V_2). \tag{2.1}$$

By our definition of degree, (2.1) holds for V_1, V_2 not necessarily pure-dimensional as well.

If V is defined by a single polynomial equation $P = 0$ with $\deg(P) > 0$, then $\deg(V) \leq \deg(P)$, and equality holds if P has no repeated factors. By Bézout, if V is defined by many equations $P_i = 0$, then $\deg(V) \leq \prod_i \deg(P_i)$.

Following [3], for a morphism $f : \mathbb{A}^n \to \mathbb{A}^m$ given by an m-tuple of polynomials f_i, we define the *maximum degree* of f to be $\mathrm{mdeg}(f) := \max_i \deg(f_i)$. For a morphism $f : X \to Y$, we define $\mathrm{mdeg}(f)$ to be the minimum of $\mathrm{mdeg}(g)$ over all $g : \mathbb{A}^n \to \mathbb{A}^m$ with $g|_X = f$.

We can bound the degree of images and preimages of varieties.

Lemma 2.1. *Let $V \subseteq \mathbb{A}^n$ and $W \subseteq \mathbb{A}^m$ be varieties and $f : V \to \mathbb{A}^m$ a morphism. Then*

(a) $\deg(\overline{f(V)}) \leq \deg(V)\mathrm{mdeg}(f)^{\dim(\overline{f(V)})}$, *and*

(b) $\deg(f^{-1}(W)) \leq \deg(V) \deg(W)\mathrm{mdeg}(f)^{\dim(\overline{f(V)})}$.

Proof. See [2, Lem. 2.3] and [3, Lem. 2.4]. □

An important result of algebraic geometry is that, given an irreducible variety V and a map $f : V \to \mathbb{A}^n$, all fibres $f^{-1}(x)$ have dimension $\geq \dim(V) - \dim(\overline{f(V)})$, with equality holding for a generic fibre. Below is a quantitative version of this statement.

Proposition 2.2. *Let V be an irreducible variety and let $f : V \to \mathbb{A}^n$ be a morphism. Call $u := \dim(V) - \dim(\overline{f(V)})$. Then for every $x \in f(V)$ we have $\dim(f^{-1}(x)) \geq u$. Moreover, there is a proper subvariety $Z \subsetneq \overline{f(V)}$ for which*

$$\deg(Z) \leq \mathrm{mdeg}(f)^{\dim(\overline{f(V)})-1} \deg(V)$$

and for which every $x \in f(V)$ with $\dim(f^{-1}(x)) > u$ is contained in Z.

Proof. See [27, §I.8] and [3, Prop. 3.2]. □

2.3. **Degree of intersections.** We show now how to bound the degree of intersections of varieties. One can always repeatedly apply Bézout, but when the number of intersecting varieties is large (or infinite) the naïve bound thus obtained may be unmanageable. The results of this subsection can be seen as explicit versions of [24, Thm. 1.10, Cor. 1.12]. Our technique is already essentially contained in [2, §3.1] and [3, §2.6] (where it was used for a different purpose, namely escape from subvarieties), although we tweak it to make it slightly more general and suited to our needs.

Lemma 2.3. *For any $D \geq 1$, define the function f_D as follows: for any variety $X \subseteq \mathbb{A}^n$, if $\{X_j\}_j$ is the finite collection of irreducible components of X, set*

$$f_D(X) := \sum_j \deg(X_j) D^{\dim(X_j)}.$$

Then, for any two varieties $Y, Z \subseteq \mathbb{A}^n$ with $\deg(Z) \leq D$, we have $f_D(Y \cap Z) \leq f_D(Y)$.

Proof. For any variety X, if we partition the collection of its components $\{X_j\}_j$ into two subsets, say for simplicity $\{X_j\}_{j \leq J}$ and $\{X_j\}_{j>J}$, and consider their unions $X_{\leq J}$ and $X_{>J}$, we clearly have $f_D(X) = f_D(X_{\leq J}) + f_D(X_{>J})$. Thus, it is sufficient to prove the result for Y irreducible.

If $Y = Y \cap Z$ then $f_D(Y \cap Z) = f_D(Y)$, and if $Y \cap Z = \emptyset$ then $f_D(Y \cap Z) = 0$. In both cases we are done, so assume otherwise. We must have $0 \leq \dim(Y \cap Z) \leq \dim(Y) - 1$. If $\{X_j\}_j$ is the collection of irreducible components of $Y \cap Z$, by Bézout

$$\begin{aligned} f_D(Y \cap Z) = \sum_j \deg(X_j) D^{\dim(X_j)} &\leq D^{\dim(Y)-1} \sum_j \deg(X_j) = D^{\dim(Y)-1} \deg(Y \cap Z) \\ &\leq D^{\dim(Y)-1} \deg(Y) \deg(Z) \leq D^{\dim(Y)} \deg(Y) = f_D(Y), \end{aligned}$$

proving the result. □

Corollary 2.4. *Let $\{Z_i\}_{i \in I}$ be a (not necessarily finite) collection of varieties inside $\mathbb{A}^n$, with $\dim(Z_i) \leq d$ and $\deg(Z_i) \leq D$ for all $i \in I$. Let $Z = \bigcap_{i \in I} Z_i$. Then*

(a) $Z = \bigcap_{i \in I'} Z_i$ *for some* $I' \subseteq I$ *with* $|I'| \leq 1 + (d+1)D^{d+1}$,

(b) $\deg(Z) \leq D^{d+1}$, *and*

(c) for any $Y \subseteq \mathbb{A}^n$ *with* $\dim(Y) = d'$ *and* $\deg(Y) = D'$, *calling* $\hat{d} = \min\{d', \dim(Z_i)\}$, *we have* $\deg(Y \cap Z) \leq D'D^{\hat{d}+1}$.

Proof. Since $\mathbb{A}^n$ is Noetherian, there is a finite subset of I (say $\{1, 2, \ldots, J\}$, after renaming the indices) with $\bigcap_{i \leq J} Z_i = Z$. Choose the smallest such subset, and from now on we may assume that $I = \{1, 2, \ldots, J\}$. We reorder I as follows. Choose Z_1 arbitrarily, and assume that we are done ordering up to some $j < J$. If $Z_{(j)} := \bigcap_{i \leq j} Z_i$, there is an irreducible component X of $Z_{(j)}$ not contained in Z, otherwise we would contradict the minimality of I. Among such components, choose one X having largest dimension, and then choose Z_{j+1} such that $X \subsetneq Z_{j+1}$.

By the minimality of I we must have $Z_1 = Z_{(1)} \supsetneq Z_{(2)} \supsetneq \ldots \supsetneq Z_{(J)} = Z$. Moreover, by the ordering chosen above, we have the following property: there are indices $1 = i_{d+1} \leq i_d \leq i_{d-1} \leq \ldots \leq i_1 \leq i_0 = J$ such that, if $i_{d'+1} < j \leq i_{d'}$, the number of d'-dimensional components in $Z_{(j)}$ is strictly smaller than in $Z_{(j-1)}$, and if $j \geq i_{d'}$ the d'-dimensional components in $Z_{(j)}$ are the same as those of Z.

Now, $f_D(Z_{(j)}) \leq f_D(Z_{(i)})$ whenever $j \geq i$ by Lemma 2.3. This allows us to give an upper bound on the number $n(d', j)$ of irreducible components of dimension d' inside $Z_{(j)}$. In fact, since we have the bounds

$$\sum_{\substack{X \text{ irr.comp.of } Y \\ \dim(X)=d'}} \deg(X) D^{d'} \leq f_D(Y) \leq \deg(Y) D^{\dim(Y)}$$

valid for any Y by definition, and since $\deg(Z_i) \leq D$ for all i, we obtain

$$\begin{aligned} n(d', j) = \sum_{\substack{X \text{ irr.comp.of } Z_{(j)} \\ \dim(X)=d'}} 1 &\leq \sum_{\substack{X \text{ irr.comp.of } Z_{(j)} \\ \dim(X)=d'}} \deg(X) \leq f_D(Z_{(j)}) D^{-d'} \\ &\leq f_D(Z_{(1)}) D^{-d'} = f_D(Z_1) D^{-d'} \leq D^{d-d'+1} \end{aligned}$$

for all d', j. Moreover, by our ordering of the indices, $n(d', j) < n(d', j-1)$ whenever $i_{d'+1} < j \leq i_{d'}$. Therefore, for every d' we have $i_{d'} - i_{d'+1} \leq n(d', i_{d'+1}) \leq D^{d-d'+1}$, and the bound

$$J = 1 + \sum_{d'=0}^{d} (i_{d'} - i_{d'+1}) \leq 1 + \sum_{d'=0}^{d} D^{d-d'+1} \leq 1 + (d+1) D^{d+1}$$

proves (a).

Using again Lemma 2.3,

$$\deg(Z) = \sum_{X \text{ irr.comp.of } Z} \deg(X) \leq f_D(Z) \leq f_D(Z_1) \leq D^{d+1},$$

proving (b). We have $\dim(Y \cap Z_1) \leq \hat{d}$ and $\deg(Y \cap Z_1) \leq D'D$, so

$$\deg(Y \cap Z) = \sum_{X \text{ irr.comp.of } Y \cap Z} \deg(X) \leq f_D(Y \cap Z) \leq f_D(Y \cap Z_1) \leq D'D^{\hat{d}+1},$$

proving (c). □

3. Linear algebraic groups

In this section we define linear algebraic groups, fix the relative notations, and use the tools of Section 2 to prove some general facts that will be used later. We shall discuss the special case of almost simple groups in more depth in Section 4.

3.1. **Definition and basic properties.** The space of $n \times n$ matrices Mat_n is the affine space $\mathbb{A}^{n^2}$ endowed with a *multiplication map*, i.e. a morphism $\cdot : \mathbb{A}^{n^2} \times \mathbb{A}^{n^2} \to \mathbb{A}^{n^2}$ defined by the usual matrix multiplication. Clearly, $\mathrm{mdeg}(\cdot) = 2$. The *general linear group* GL_n is commonly defined as an open set inside Mat_n. We shall however need to work with GL_n as a (Zariski-closed) variety, so we define instead

$$\mathrm{GL}_n := \{x \in \mathrm{Mat}_{n+1} : \quad x_{i,n+1} = x_{n+1,i} = 0 \quad (1 \leq i \leq n), \quad \det(x|_{n\times n}) \cdot x_{n+1,n+1} = 1\},$$

where $x_{i,j}$ is the (i,j)-th entry of x and $x|_{n\times n}$ is the restriction to the $n \times n$ upper left corner of x. The *determinant* is the morphism $\det : \mathrm{GL}_n \to \mathbb{A}^1$ given by taking the determinant of the $n \times n$ upper left corner of x (we just write $\det(x)$ for simplicity); it has maximum degree $\mathrm{mdeg}(\det) = n$. The *inversion map* is the morphism ${}^{-1} : \mathrm{GL}_n \to \mathrm{GL}_n$ sending (x, y) to $(\mathrm{adj}(x)y, \det(x))$, where the *adjugate* $\mathrm{adj}(x)$ is defined by $\mathrm{adj}(x)_{i,j} = (-1)^{i+j} M_{j,i}$, with $M_{j,i}$ being the (j,i)-th minor of x.

We occasionally work with a specified n-dimensional K-vector space L, and denote by $\mathrm{Mat}(L)$ and $\mathrm{GL}(L)$ respectively the space Mat_n and the group GL_n over K.

A *(linear) algebraic group*[3] G is a subvariety of GL_n closed under the multiplication and inversion maps [4, p. 51], and in this case we write $G \leq \mathrm{GL}_n$. The maximum degree of these maps may change when restricted to G, and the degree of G and its maps may also differ depending on how we choose to represent G. We call H an *(algebraic) subgroup* of G, and write $H \leq G$, if it is both a subvariety of G and an algebraic group (not necessarily defined over the same field K). A *normal subgroup* $H \trianglelefteq G$ is a subgroup for which $\varphi(H) = H$ for all automorphisms $\varphi : G \to G$ of the form $\varphi(g) = xgx^{-1}$ for some fixed $x \in G(\overline{K})$; a *characteristic subgroup* $H \blacktriangleleft G$ is a normal subgroup for which the above holds for all automorphisms φ, not only the inner ones.

One can define the *Lie algebra of* G by endowing the tangent space of G at e with a Lie algebra structure: see [4, §3.5] or [21, §9.1] for details. We use the conventional fraktur notation, such as $\mathfrak{gl}_n$ and $\mathfrak{g}$, to denote Lie algebras. We have $\dim(G) = \dim(\mathfrak{g})$ for all G.

We now define several objects inside G. The *identity component* G^{o} is the connected component of the identity of G. If $G = G^{\mathrm{o}}$, or equivalently if G is connected, then G is irreducible [26, Cor. 1.35]. For any element $g \in G(\overline{K})$, any set $\Lambda \subseteq G(\overline{K})$, and any variety $V \subseteq G$, their *centralizers* in G are

$$C_G(g) := \{x \in G : gx = xg\}, \qquad C_G(\Lambda) := \bigcap_{g\in\Lambda} C_G(g), \qquad C_G(V) := C_G(V(\overline{K})).$$

Every centralizer is an algebraic subgroup of G. The *centre* of G is defined as $Z(G) := C_G(G)$, and we have $Z(G) \blacktriangleleft G$.

[3] Since we work exclusively in the affine space, there is no need to distinguish between "algebraic groups" and "linear algebraic groups". See [21, §8.6].

A *torus* T of G is a subgroup of G isomorphic to the product of m copies of GL_1 for some $m \geq 1$ [4, §8.5]; T is a *maximal torus* if it has maximal m among all tori of G. A *Cartan subgroup* of G is a subgroup of the form $C_G(T)$ for some maximal torus T. If G is connected then all maximal tori are conjugate [4, Cor. 11.3(1)], and the Cartan subgroups of G have the same dimension, which is called the *rank* $\mathrm{rk}(G)$ of G. Throughout the rest of the paper, when dealing with an algebraic group G we use

$$d = \dim(G), \qquad D = \deg(G), \qquad r = \mathrm{rk}(G), \qquad \iota = \mathrm{mdeg}(^{-1}\colon G \to G). \tag{3.1}$$

Let $G \leq \mathrm{GL}_n \subseteq \mathrm{Mat}_{n+1}$ be defined over K. An element $g \in G(K)$ is *unipotent* if $(g|_{n\times n} - \mathrm{Id}_n)^m = 0$ for some $m \geq 1$, and it is *semisimple* if it is conjugate to some diagonal matrix in $G(\overline{K})$ [4, §4.1]. There are unique elements g_s, g_u, respectively semisimple and unipotent, such that $g = g_s g_u = g_u g_s$ [4, p. 81, Cor. 1(1)]. An element $g \in G$ is *regular* if $\dim(C_G(g_s)) = \mathrm{rk}(G)$. We use the notation

$$G^{\mathrm{rss}} = \{g \in G : g \text{ regular and semisimple}\}, \qquad G^{\mathrm{un}} = \{g \in G : g \text{ unipotent}\},$$
$$G^{\mathrm{run}} = \{g \in G : g \text{ regular and unipotent}\}, \qquad G^{\mathrm{irr}} = \{g \in G : g \text{ not regular}\}.$$

We adopt the shorthand $V^{\mathrm{rss}} = V \cap G^{\mathrm{rss}}$ for varieties $V \subseteq G$, and $X^{\mathrm{rss}} = X \cap G^{\mathrm{rss}}(\overline{K})$ for subsets $X \subseteq G(\overline{K})$ (and similarly for the other notations).

For G connected, a *Borel subgroup* B is a maximal connected solvable subgroup of G. Its unipotent part $U = B^{\mathrm{un}}$ is a maximal connected unipotent subgroup of G. If $\mathcal{B}$ is the collection of Borel subgroups of G, then

$$R(G) = \left(\bigcap_{B\in\mathcal{B}} B\right)^{\mathrm{o}}, \qquad\qquad R_u(G) = R(G)^{\mathrm{un}}$$

are respectively the *radical* and the *unipotent radical* of G [4, §11.21]. The definitions above are invariant under automorphisms of G, meaning in particular that $R_u(G) \trianglelefteq G$. A connected G is called *reductive* if $R_u(G) = \{e\}$, and *semisimple* if $R(G) = \{e\}$. If G is reductive, Cartan subgroups and maximal tori coincide [4, §13.17, Cor. 2(c)], so $\mathrm{rk}(G)$ is the dimension of any maximal torus. If G is also semisimple, we have further properties on the sets of regular, semisimple, and unipotent elements: G^{rss} is open and dense [20, §2.5], G^{un} is closed and irreducible of dimension $\dim(G) - \mathrm{rk}(G)$ [20, §4.2], G^{run} is nonempty [20, §4.5], and G^{irr} is closed and proper (since the set of regular elements is open and dense [20, §1.4]).

The degree of many objects above can be bounded effectively.

Lemma 3.1. *Let $G \leq \mathrm{GL}_n$ be a connected algebraic group with $d = \dim(G)$, $D = \deg(G)$, and $\iota = \mathrm{mdeg}(^{-1})$, defined over a field K.*

(a) For any maximal torus T of G, $\deg(T) \leq D$.
(b) For any Borel subgroup B of G, $\deg(B) \leq D$.
(c) For any $H \leq G$ connected unipotent, $\deg(H) \leq \dim(H)^{\dim(H)}$.
(d) For any maximal connected unipotent subgroup U of G, $\deg(U) \leq \min\{D, d^d\}$.
(e) For the unipotent radical $R_u(G)$ of G, $\deg(R_u(G)) \leq \min\{(nD)^{d+1}, d^d\}$.

Proof. Each of the subgroups T, B, U as above is respectively of the form $(G\cap T')^{\mathrm{o}}, (G\cap B')^{\mathrm{o}}, (G \cap U')^{\mathrm{o}}$ for T', B', U' of the same type in GL_n [4, Prop. 11.14(2)]. All Borel subgroups B' of GL_n are conjugate to each other [4, §11.1], and similarly for T', U' [4, Cor. 11.3]. Thus, it is enough to choose T', B', U' whose degrees we can bound well. Choose respectively the diagonal maximal torus T', the group B' of upper triangular matrices, and the subgroup U' of B' whose elements have all diagonal entries equal to 1; hence, T', B', U' are intersections of GL_n with varieties of degree 1. The bound in (a) and (b) and the first bound in (d) all follow from Bézout.

By definition, $R_u(G)$ is the unipotent part of the radical $R(G) = \left(\bigcap_{B\in\mathcal{B}} B\right)^{\mathrm{o}}$. The unipotent part is defined by equations of degree $\leq n$ since, by what we said about Borel subgroups of GL_n, $R(G)$ lies in a conjugate of the set of upper triangular matrices. Furthermore, the Borel subgroups of G have degree bounded by (b). Then, for some varieties Z_i of degree $\leq n$, Corollary 2.4(b)–(c) yields

$$\begin{aligned}\deg(R_u(G)) = \deg\left(\left(\bigcap_{B\in\mathcal{B}} B\right)^{\mathrm{o}} \cap \bigcap_{i\in I} Z_i\right) &\leq \deg\left(\left(\bigcap_{B\in\mathcal{B}} B\right)^{\mathrm{o}}\right)\cdot n^{\dim(R(G))+1}\\ &\leq D^{\dim(B)+1} n^{\dim(R(G))+1} \leq (nD)^{d+1},\end{aligned}$$

which gives the first bound in (e).

Now let $H \leq G$ be connected unipotent. By definition H is also solvable, so there is some $z \in \mathrm{GL}_n(\overline{K})$ for which $H' = zHz^{-1}$ is upper triangular by the Lie-Kolchin theorem [4, Cor. 10.5]. H' is still unipotent of dimension $\dim(H)$, so we can write $H = \overline{U_1 U_2 \ldots U_{\dim(H)}}$ for some 1-dimensional irreducible subgroups U_i each generated by a unipotent matrix $u_i \in H'$ (see for instance [21, §7.5], which is more general and forgoes the Zariski closure at the expense of lengthening the product to $2\dim(H)$ factors). For some $z_i \in \mathrm{GL}_n(\overline{K})$ we can write $u_i = z_i v_i z_i^{-1}$ with v_i in Jordan normal form. If $\mathcal{J}_i$ is the set of indices j for which $(v_i)_{j,j+1}$ is nonzero, then the group $V_i = z_i^{-1} U_i z_i$ is the variety defined by $x_{j,j} = 1$ for all j, $x_{j_1,j_1+1} = x_{j_2,j_2+1}$ for all $j_1, j_2 \in \mathcal{J}_i$, and $x_{j,k} = 0$ everywhere else. Therefore $\deg(V_i) = 1$, and using the morphism $f : V_1 \times \ldots \times V_{\dim(H)} \to G$ defined by

$$f(x_1, \ldots, x_{\dim(H)}) = z^{-1} \cdot z_1 x_1 z_1^{-1} \cdot \ldots \cdot z_{\dim(H)} x_{\dim(H)} z_{\dim(H)}^{-1} \cdot z$$

we conclude that $\overline{f(V_1 \times \ldots \times V_{\dim(H)})}$ has degree $\leq \dim(H)^{\dim(H)}$ by Lemma 2.1(a). This object is H itself, so we obtain (c).

Finally, both U and $R_u(G)$ are connected unipotent, so (c) implies the second bounds in (d) and (e). □

We can also bound the degree of the centralizer of any set in $G(\overline{K})$. This is an easy application of Corollary 2.4, although a more elementary argument relying on the fact that we intersect linear varieties would also be sufficient.

Corollary 3.2. *Let $G \leq \mathrm{GL}_n$ be an algebraic group defined over K, with $d = \dim(G)$ and $D = \deg(G)$, and let $\Lambda \subseteq G(\overline{K})$. Then $C_G(\Lambda) = C_G(\Lambda')$ for some $\Lambda' \subseteq \Lambda$ of size $|\Lambda'| \leq d+1$, and $\deg(C_G(\Lambda)) \leq D$.*

Proof. For any $x \in G(\overline{K})$, the centralizer $C_G(x)$ is defined as the set of $g \in G$ with $gx = xg$, which yields a finite number of equations of degree 1. Thus $C_G(x)$ is the intersection of G with varieties $Z_{i,x}$ of degree 1, for some set of indices i. In turn, $C_G(\Lambda)$ is the intersection of $C_G(\lambda)$ for all $\lambda \in \Lambda$. Apply Corollary 2.4(a) to the collection of

$Z_{i,\lambda}$ to obtain the bound on $|\Lambda'|$, and Corollary 2.4(c) to G and the $Z_{i,\lambda}$ to obtain the bound on $\deg(C_G(\Lambda))$. □

3.2. **Escaping from a subgroup.** The tools from [2, 3] that we are going to use in Section 5 rely on the procedure called *escape from subvarieties*, which first appeared in [14]. To produce dimensional estimates for a set A of generators of $G(K)$, we need to be able to say that for any proper subvariety $V \subsetneq G$ there is some $g \in A^k$ with $g \notin V(\overline{K})$, where k is bounded appropriately in terms of $\deg(V)$.

Here we start with a different object, namely a subgroup Γ, and a weaker hypothesis, namely that we escape from algebraic subgroups. Thus, we have to prove that escaping from subgroups is enough to escape from every subvariety as well, up to paying a price in degree bounds. The following result shows the contrapositive statement: if Γ is large enough and is trapped in a subvariety, then it is trapped in a subgroup as well.

Lemma 3.3. *Let G be a linear algebraic group defined over K. Let $\Gamma \leq G(\overline{K})$, and let $V \subsetneq G$ be a proper subvariety. Assume that $\Gamma \subseteq V(\overline{K})$.*

Then, either $|\Gamma| \leq \deg(V)^{\dim(V)+1}$, or there is a proper algebraic subgroup $H \lneq G$ with $\Gamma \leq H(\overline{K})$ and $\deg(H) \leq \deg(V)^{\dim(V)+1}$.

Proof. For every proper subvariety $W \subsetneq G$, the stabilizer $\mathrm{Stab}(W) = \{g \in G : Wg = W\}$ is a proper algebraic subgroup of G.

We build a sequence of V_i with the following properties: $\Gamma \subseteq V_i(\overline{K})$, $V_i = \bigcap_{\gamma \in S_i} V\gamma$ for some $S_i \subseteq \Gamma$, and $V_i \subsetneq V_{i-1}$. We stop constructing the sequence when either $\Gamma \leq \mathrm{Stab}(V_i)(\overline{K})$ or $\dim(V_i) = 0$. The starting point is $S_0 = \{e\}$ and $V_0 = V$. To construct V_{i+1}, assume that we have V_i as above, and suppose that there is some $\gamma \in \Gamma \setminus \mathrm{Stab}(V_i)(\overline{K})$. Then $V_i \cap V_i\gamma \subsetneq V_i$ and $\Gamma = \Gamma\gamma \subseteq V_i(\overline{K})\gamma$, so we can choose $S_{i+1} = S_i \cup S_i\gamma$ and obtain $V_{i+1} = V_i \cap V_i\gamma$ accordingly. The V_i are never empty because they contain Γ, thus since the affine space is Noetherian we will eventually reach some zero-dimensional V_i (unless we stopped because $\Gamma \subseteq \mathrm{Stab}(V_i)(\overline{K})$).

By the above, we obtained that either $\Gamma \subseteq V_i(\overline{K})$ with $\dim(V_i) = 0$, or $\Gamma \leq H(\overline{K})$ for $H = \mathrm{Stab}(V_i)$; in either case, $V_i = \bigcap_{\gamma \in S_i} V\gamma$ for some $S_i \subseteq \Gamma$. Each $V\gamma$ has the same dimension and degree as V, so $\deg(V_i) \leq \deg(V)^{\dim(V)+1}$ by Corollary 2.4(b); if $\Gamma \subseteq V_i(\overline{K})$ and $\dim(V_i) = 0$, then $|\Gamma| \leq \deg(V_i)$ and we are done. We can rewrite H as

$$H = \mathrm{Stab}(V_i) = \bigcap_{w \in V_i(\overline{K})} w^{-1}V_i = \bigcap_{(w,\gamma) \in V_i(\overline{K}) \times S_i} w^{-1}V\gamma,$$

so again by Corollary 2.4(b) we get $\deg(H) \leq \deg(V)^{\dim(V)+1}$. □

3.3. **Quotients.** Even if a linear algebraic group G and a normal algebraic subgroup $H \trianglelefteq G$ are naturally defined as varieties, one may not always be able to see the quotient G/H as a variety in any obvious way. Below, we explain how to do so, following [4, §6].

Let $H \trianglelefteq G$ be both defined over K. A *(geometric) quotient* of G by H is a pair (π, W) made of an affine variety W and a surjective morphism $\pi : G \to W$, both defined over K, satisfying the following universal property: if $\alpha : G \to Z$ is a morphism constant on H-orbits, there is a unique morphism $\beta : W \to Z$ such that $\alpha = \beta \circ \pi$, and if α is a K-morphism of K-varieties then so is β.

By [4, Thm. 6.8], under the conditions above on G and H, a geometric quotient always exists and is unique, and W is also an algebraic group defined over K. We use the notation G/H for W, and we say that G/H is the *quotient* of G by H (forgoing π). In this case, the quotient G/H also coincides with the *categorical quotient* [4, §6.16], so we ignore the distinction.

By [4, Prop. 6.4(b)], $\dim(G/H)=\dim(G)-\dim(H)$. Furthermore, since G/H is an algebraic group, we can represent it as a variety inside GL_m for some possibly large m. The result below gives an upper bound for m and $\deg(G/H)$, at the expense of possibly defining G/H over $\overline{K}$ (which simplifies the matter without being a problem for us later).

Proposition 3.4. *Let $G\leq\mathrm{GL}_n\subseteq\mathrm{Mat}_{n+1}$ be an algebraic group defined over a field K, with $d=\dim(G)$ and $D=\deg(G)$. Let $H\trianglelefteq G$ be a normal subgroup, defined over K by polynomials of degree $\leq\Delta$ (excluding the ones defining G itself).*

Then G/H is an algebraic group. More precisely, there are algebraic groups $\hat{H}\trianglelefteq\hat{G}\leq\mathrm{GL}_{2n+3}$ and $Q\leq\mathrm{GL}_m$ and a morphism $\hat{\beta}:\hat{G}\to Q$ (possibly over $\overline{K}$) such that

(a) there is a morphism of algebraic groups $\hat{G}\to G$ of maximal degree 1 with rational inverse (so in particular $\hat{G}(\overline{K})\simeq G(\overline{K})$ as abstract groups), and the same holds for its restriction $\hat{H}\to H$,

(b) $Q\simeq\hat{G}/\hat{H}$, meaning that Q satisfies the aforementioned universal property (possibly over $\overline{K}$), and

(c) the following quantitative bounds hold:

$$M=\binom{n^2+\Delta}{\Delta}\leq(n^2+\Delta)^{\min\{n^2,\Delta\}}, \qquad m\leq 2^{2M},$$

$$\deg(\hat{G})\leq M2^{M+n^2+4}\Delta D, \qquad \deg(\hat{H})\leq M2^{M+n^2+4}\Delta\deg(H),$$

$$\deg(Q)\leq M^{d+1}2^{M+n^2+d+5}\Delta^{d+1}D, \qquad \mathrm{mdeg}(\hat{\beta})\leq 2M\Delta.$$

Proof. We follow the path laid out in [4], pasting together parts of the proofs of [4, §1.9 and Thms. 5.1–5.6–6.8]. We work everywhere over $\overline{K}$ for simplicity.

By our definition of GL_n in Section 3.1, we are already keeping track of $\det(g)^{-1}$ in the definition of $g\in G$. We will now use a new $\hat{G}$ to keep track of one more determinant (to be defined in the future through some function F) and of the entire g^{-1}. In brief, using $\oplus$ to denote the diagonal join of matrices, if an element of $\mathrm{GL}_n\subseteq\mathrm{Mat}_{n+1}$ is of the form $a\oplus\det(a)^{-1}$, we pass to $a\oplus\det(a)^{-1}\oplus a^{-1}\oplus\det(a)\oplus y^{-1}\oplus y$ inside $\mathrm{GL}_{2n+3}\subseteq\mathrm{Mat}_{2n+4}$, with y defined using F.

Let $R_1:=\overline{K}[x_{11},x_{12},\ldots,x_{n+1,n+1}]$ and $R_2:=\overline{K}[x_{11},x_{12},\ldots,x_{2n+2,2n+2}]$, and fix $F\in R_2$ to be chosen later. Call π_{ij} the restriction map sending $x\in\mathrm{Mat}_{2n+4}$ to the square submatrix whose corners are the (i,i)-th, (i,j)-th, (j,i)-th, and (j,j)-th entries (with $i\leq j$). Define

$$\begin{aligned}\mathcal{O}:=&([1,n+1]\times[n+2,2n+2])\cup([n+2,2n+2]\times[1,n+1])\\&\cup([n+2,2n+1]\times\{2n+2\})\cup(\{2n+2\}\times[n+2,2n+1])\\&\cup([1,2n+2]\times\{2n+3\})\cup(\{2n+3\}\times[1,2n+2])\\&\cup([1,2n+3]\times\{2n+4\})\cup(\{2n+4\}\times[1,2n+3]),\end{aligned}$$

$$X := \left\{ x \in \mathrm{Mat}_{2n+4} : \begin{array}{l} x_{ij} = 0 \quad ((i,j) \in \mathcal{O}), \\ \pi_{1,n}(x) \cdot \pi_{n+2,2n+1}(x) = \mathrm{Id}_n, \\ x_{n+1,n+1} \cdot x_{2n+2,2n+2} = 1, \\ F(\pi_{1,2n+2}(x)) \cdot x_{2n+3,2n+3} = 1, \\ x_{2n+3,2n+3} \cdot x_{2n+4,2n+4} = 1 \end{array} \right\}.$$

The variety $\hat{G} := X \cap \pi_{1,n+1}^{-1}(G)$ is an algebraic group $\hat{G} \leq \mathrm{GL}_{2n+3} \subseteq \mathrm{Mat}_{2n+4}$, since by construction $\det(\pi_{1,2n+3}(x)) \cdot x_{2n+4,2n+4} = 1$. If $F(g \oplus g^{-1}) \neq 0$ for all $g \in G$, then there is a rational map $G \to \hat{G}$ that is the inverse of $\pi_{1,n+1}|_{\hat{G}}$, yielding a group isomorphism $\hat{G}(\overline{K}) \simeq G(\overline{K})$ and proving (a). Let $f_1, \ldots, f_k \in R_1$ be the polynomials defining H in G, i.e.

$$H = \{g = (g_{ij})_{i,j \leq n+1} \in G : \forall k' \leq k\, (f_{k'}(g_{11}, g_{12}, \ldots, g_{n+1,n+1}) = 0)\}.$$

By hypothesis $\deg(f_i) \leq \Delta$, and by construction the same polynomials define $\hat{H} := X \cap \pi_{1,n+1}^{-1}(H)$ in $\hat{G}$. The group isomorphism $\hat{H}(\overline{K}) \simeq H(\overline{K})$ follows from restricting $\pi_{1,n+1}|_{\hat{G}}$ and its inverse rational map. By Bézout and Lemma 2.1(b), we get

$$\deg(\hat{G}) \leq \deg(G) \cdot 2^{n^2+2}(\deg(F) + 1), \qquad \deg(\hat{H}) \leq \deg(H) \cdot 2^{n^2+2}(\deg(F) + 1).$$

We shall pass to $\hat{G}, \hat{H}$ at the end of our procedure, whereas for now we use the original G, H for our constructions.

Let L be the $\overline{K}$-vector space spanned by the set $\mathcal{L}$ of all monomials of degree $\leq \Delta$ in the variables x_{ij} $(i, j \leq n+1)$; then $\dim(L) = |\mathcal{L}| = M$ with M as in the statement. Each f_k is an element of L. For every $g \in G(\overline{K})$, let $\rho_g : R_1 \to R_1$ be the right translation by g, namely $\rho_g(f)(x) = f(xg)$. Then ρ_e is the identity map, and $\rho_{g_1 g_2} = \rho_{g_2} \circ \rho_{g_1}$. The latter property implies that the elements $\rho_g(f_{k'})$ for all $k' \leq k$ and all $g \in G(\overline{K})$ span a $\overline{K}$-vector subspace $L' \leq L$ that is invariant under the transformations ρ_g. Fix a basis $\{\ell_i\}_{i \leq M}$ of L whose first $\dim(L')$ members form a basis of L'. Every ℓ_i is a polynomial in R_1 of degree $\leq \Delta$, and by construction there are polynomials $\tilde{f}_{i,j} \in R_1$ of degree $\leq \Delta$ for which $\rho_g(\ell_i) = \sum_{j \leq M} \tilde{f}_{i,j}(g)\ell_j$.

Therefore, putting together the facts above, we obtain the following: ρ_g is a linear transformation of L, i.e. $\rho_g \in \mathrm{Mat}(L)$, the resulting morphism $\rho^{(L)} : G \to \mathrm{Mat}(L)$ given by $\rho^{(L)}(g) = \rho_g$ has $\mathrm{mdeg}(\rho^{(L)}) \leq \Delta$, and $\rho^{(L)}$ naturally restricts to $\rho : G \to \mathrm{Mat}(L')$ by taking the upper left $(\dim(L') \times \dim(L'))$-corner of the matrix, thus giving again $\mathrm{mdeg}(\rho) \leq \Delta$.

If $I \subseteq \overline{K}[G]$ is the ideal of functions vanishing on H, the set $W = L' \cap I$ generates I since L' contains all the $f_{k'}$. We have $\dim(W) \leq \dim(L') \leq M$. The proof of [4, Thm. 5.1] shows that $H = \{g \in G : \rho_g(W) = W\}$.

Next, let $E = \bigwedge^{\dim(W)}(L')$. Then ρ induces a map $\alpha : G \to \mathrm{Mat}(E)$, with $\mathrm{mdeg}(\alpha) \leq \dim(W)\mathrm{mdeg}(\rho) \leq M\Delta$. Furthermore, [4, Thm. 5.1] shows that there is a 1-dimensional subspace $E' \subseteq E$ such that $H = \{g \in G : \alpha(g)(E') = E'\}$. Finally, [4, Thm. 5.6] shows that there is some subspace $V \subseteq \mathfrak{gl}(E)$ such that the rational map $\beta : G \to \mathrm{Mat}(V)$ given by $\beta(g)(v) = \alpha(g)v\alpha(g)^{-1}$ has the property that $H = \mathrm{Ker}(\beta)$; [4, Thm. 6.8] then shows that $\beta(G)$ is a closed subgroup of $\mathrm{Mat}(V)$ with $\beta(G) \simeq G/H$. We have $\dim(E) = \binom{\dim(L')}{\dim(W)} < 2^{\dim(L')} \leq 2^M$ and $\dim(V) \leq \dim(E)^2 < 2^{2M}$.

Our only problems with the construction of β are that β is a rational map and not a morphism, and that it goes to $\mathrm{Mat}_{\dim(V)}$ rather than to $\mathrm{GL}_{\dim(V)}$. To be clear on the second issue, since $(\rho_g)^{-1} = \rho_{g^{-1}}$, the maps $\rho^{(L)}, \rho, \alpha, \beta$ do indeed send elements of G to invertible matrices, but we are missing the inverse of the determinant in the lower right corner (needed in our definition of GL_n from Section 3.1). These two problems are why $\hat{G}$ comes into play: we build a new $\hat{\beta}$ that has the same behaviour as β, but is also a morphism to $\mathrm{GL}_{\dim(V)}$ because the inverses of g and of $\det(\beta(g))$ can be taken directly from the extra variables of $\hat{G}$.

By construction, every element $\hat{g} \in \hat{G}$ is of the form $\hat{g} = b \oplus y_1 \oplus y_2$ with $y_2 = F(b) = y_1^{-1}$, $b = g \oplus g^{-1}$, and $g \in G \leq \mathrm{GL}_n$ (so that in turn $g = a \oplus \det(a)^{-1}$ for some invertible $a \in \mathrm{Mat}_n$). Define the morphism

$$\gamma : \pi_{1,2n+2}(\hat{G}) \to \mathrm{Mat}_{\dim(V)}, \quad \gamma(\pi_{1,2n+2}(\hat{g}))(v) = \alpha(\pi_{1,n+1}(\hat{g})) \cdot v \cdot \alpha(\pi_{n+2,2n+2}(\hat{g})).$$

Since $\rho_{g^{-1}} = (\rho_g)^{-1}$ we have $\rho(g^{-1}) = \rho(g)^{-1}$ and $\alpha(g^{-1}) = \alpha(g)^{-1}$. Therefore $\alpha(\pi_{n+2,2n+2}(\hat{g})) = \alpha(\pi_{1,n+1}(\hat{g}))^{-1}$ for all $\hat{g} \in \hat{G}$.

Choose $F \in R_2$ to be the polynomial $F(x) = \det(\gamma(x))$. By construction, for all $\hat{g} \in \hat{G}$ we have $F(g \oplus g^{-1}) = \det(\gamma(g \oplus g^{-1})) = \det(\beta(g))$. In particular $F(g \oplus g^{-1}) \neq 0$ for all $g \in G$ (because β sends g to an invertible matrix), so by what we said before we obtain (a). Finally, define the morphism $\hat{\beta} : \hat{G} \to \mathrm{GL}_{\dim(V)} \subseteq \mathrm{Mat}_{\dim(V)+1}$ by $\hat{\beta}(\hat{g}) = \gamma(\pi_{1,2n+2}(\hat{g})) \oplus y_1$. As $\gamma(\pi_{1,2n+2}(\hat{g})) = \beta(g)$ and $y_1 = \det(\beta(g))^{-1}$, the construction of $\hat{\beta}$ coincides with that of β with the addition of the extra entry for the inverse of the determinant. Hence we have again a closed subgroup $Q := \hat{\beta}(\hat{G}) \simeq \hat{G}/\hat{H}$ of $\mathrm{GL}_{\dim(V)}$, but by passing through γ we are now only working with polynomials, and we have (b).

To obtain (c), it remains to compute degrees. We have

$$\begin{aligned}
\deg(F) &\leq 2\mathrm{mdeg}(\alpha)\dim(E) \leq M2^{M+1}\Delta,\\
\deg(\hat{G}) &\leq \deg(G) \cdot 2^{n^2+2}(\deg(F)+1) \leq M2^{M+n^2+4}\Delta D,\\
\deg(\hat{H}) &\leq \deg(H) \cdot 2^{n^2+2}(\deg(F)+1) \leq M2^{M+n^2+4}\Delta \deg(H),\\
\mathrm{mdeg}(\hat{\beta}) &\leq \mathrm{mdeg}(\gamma) \leq 2\mathrm{mdeg}(\alpha) \leq 2M\Delta,\\
\deg(Q) &\leq \deg(\hat{G})\mathrm{mdeg}(\hat{\beta})^{\dim(G)} \leq M^{d+1}2^{M+n^2+d+5}\Delta^{d+1}D,
\end{aligned}$$

where the last line used Lemma 2.1(a). □

4. Almost simple groups

In this section we restrict our focus to a special class of algebraic groups, the almost simple groups, with particular emphasis on the adjoint ones. We recall their classification and, for any such G, describe several related objects that will be fundamental in our main proof: the Weyl group $\mathcal{W}$ of G, the Steinberg endomorphisms $F : G(\overline{\mathbb{F}_p}) \to G(\overline{\mathbb{F}_p})$ in positive characteristic, and a certain representation ρ of G taken from [28].

4.1. **Almost simple and adjoint groups.** A connected linear algebraic group G is *almost simple*[4] if it is non-abelian and has no connected normal linear algebraic subgroups except $\{e\}$ and G. If G is almost simple then it is also semisimple. If not, we would have $R(G) = G$, and in particular G would be connected solvable, so that by [4, Thm. 10.6(1)] either $G^{\mathrm{un}} = \{e\}$ or $G^{\mathrm{un}} = G$. Then, [4, Thm. 10.6(2)] implies in the first case that G is a torus, and in the second case that $\dim(G) \leq 1$, contradicting in both cases the fact that G is not abelian (in the latter case by [21, §20]).

To every connected semisimple algebraic group G one can associate a *Dynkin diagram*, a finite multigraph that is connected if and only if G is almost simple. The connected Dynkin diagrams are completely classified, which allows us to characterize the corresponding groups by types: the possibilities are the *classical types* $A_{r\geq 1}, B_{r\geq 2}, C_{r\geq 3}, D_{r\geq 4}$ (where r is the same as the rank of the corresponding group), and the *exceptional types* E_6, E_7, E_8, F_4, G_2. For G connected, an *isogeny* is a surjective morphism $\varphi : G \to H$ with finite kernel. Every connected almost simple algebraic group is uniquely determined by its Dynkin diagram and isogeny type.

See [8, §1.11], [21, §32 and App.], [25, §9], or [26, §24] for more details about Dynkin diagrams, isogenies, almost simple groups, and their classification. Here we only spend a few words on roots and weights. The identification between groups and diagrams passes through a *root system* Φ, defined for groups in [21, §16.4] and for diagrams as in [21, App.] and put in relation to each other in [21, §27]. The *roots* are (finitely many) elements spanning a vector space that can be identified with $\mathbb{R} \otimes_{\mathbb{Z}} X(T)$, with $X(T)$ the character group of a maximal torus T, and the *weights* are the elements of the vector space whose inner products with the roots are all integers. Roots and weights span two lattices, Λ_r and Λ_w, with $\Lambda_r \leq X(T) \leq \Lambda_w$ and with $[\Lambda_w : \Lambda_r]$ finite. The groups Λ_w/Λ_r and $\Lambda_w/X(T)$ are called the *fundamental groups* of Φ and G, respectively (see [21, §31.1 and App.]). For a fixed Dynkin diagram, isogeny types and fundamental groups for G are in 1-to-1 correspondence (here we really need equality of fundamental groups, not just isomorphism: see [21, §32.1, Thm.]). The group Λ_w/Λ_r is given in [21, App. (A.9)] or [25, Table 9.2]; we shall only need the uniform cruder bound

$$|\Lambda_w/\Lambda_r| \leq r + 1, \tag{4.1}$$

where r is the rank of G.

Among the (finitely many) connected almost simple groups G with the same Dynkin diagram, i.e. isogenous to each other, there are two extremes: the *simply connected* group with $X(T) = \Lambda_w$ and the *adjoint* group with $X(T) = \Lambda_r$ [25, Def. 9.14]. For every G of a given type, there are isogenies from the corresponding simply connected group to G and from G to the adjoint group [25, Prop. 9.15]. As in [24], we work almost exclusively with the adjoint groups.

[4] There are many different names for these objects. Some authors call these groups *simple*, as in [8, §1.11], [21, p. 168], and [25, p. xiv], but they are not necessarily simple as abstract groups. Some call them *quasi-simple*, as in [29, Ex. 8.1.12(4a)], keeping closer to the convention of abstract groups. Authors that put emphasis on the field of definition use *geometrically almost simple*, as in [26, Def. 19.7], or *absolutely almost simple*, as in [7, §5], and they may drop the "almost", as in [17]. Our work is mostly on the algebraic closure $\overline{K}$, and we deal with finite simple groups as well, so we choose to adopt the term *almost simple* as in [4, Prop. 14.10(3)].

Another definition independent of the concepts above is the following: a semisimple group G is *adjoint* if and only if its centre $Z(G)$ is trivial [26, §17.g]. For G reductive the *adjoint representation* of G is the morphism Ad_G defined in [26, §10.d] as follows:

$$\mathrm{Ad}_G : G \to \mathrm{GL}(\mathfrak{g}), \qquad g \mapsto \mathrm{Ad}_{G,g}, \qquad \text{with } \mathrm{Ad}_{G,g} : \mathfrak{g} \to \mathfrak{g}, \qquad x \mapsto gxg^{-1}.$$

We abbreviate $\mathrm{Ad}_G(G)$ as G^{ad}; it is isomorphic to $G/Z(G)$, which is an adjoint group by [26, Cor. 17.62(e)]. If $G \leq \mathrm{GL}_n$ has $\mathrm{mdeg}(^{-1}) = \iota$, it is clear that

$$\mathrm{mdeg}(\mathrm{Ad}_G) \leq \iota + 1 \leq n + 1. \tag{4.2}$$

In fact, in concrete terms, the action of G on $\mathfrak{g}$ is given by conjugation in Mat_{n+1}, so it has maximal degree $\leq \iota+1$; up to a change of basis (not affecting mdeg), Mat_{n+1} is the sum of the subspace $\mathfrak{g}$ and a complement of it, and the restriction of the conjugation map to $\mathfrak{g}$ has degree $\leq \iota + 1$ too. The bound $\iota \leq n$ comes from using the adjugate map to define inverses.

From any G connected we can naturally descend to an adjoint quotient group. Taking the quotient here is essentially what gives (c) and (d) in the main theorem.

Lemma 4.1. *Let $G \leq \mathrm{GL}_n$ be a connected algebraic group with $d = \dim(G)$, $D = \deg(G)$, and $\iota = \mathrm{mdeg}(^{-1})$, defined over a field K.*

There is some characteristic subgroup $Y \trianglelefteq G$, defined by the polynomials defining G and by some additional polynomials over $\overline{K}$ of degree $\leq (\iota + 1)d^d$, such that

$$G/Y \simeq (G/R_u(G))^{\mathrm{ad}} \simeq (G/R_u(G))/Z(G/R_u(G)).$$

Moreover there are algebraic groups $\hat{G}, \hat{Y}$, there is a morphism

$$\lambda : \hat{G} \to G, \qquad \lambda|_{\hat{Y}} : \hat{Y} \to Y, \qquad \mathrm{mdeg}(\lambda) = \mathrm{mdeg}(\lambda|_{\hat{Y}}) = 1,$$

having rational inverse, so that $G(\overline{K}) \simeq \hat{G}(\overline{K})$ and $Y(\overline{K}) \simeq \hat{Y}(\overline{K})$, and there is a morphism $\hat{\beta} : \hat{G} \to \hat{G}/\hat{Y} \leq \mathrm{GL}_m$ satisfying

$$\mathrm{mdeg}(\hat{\beta}) \leq 2(n^2 + (\iota + 1)d^d)^{n^2}(\iota + 1)d^d, \qquad m \leq 2^{2(n^2+(\iota+1)d^d)^{n^2}}.$$

Proof. The group Y can be defined as the set of $y \in G$ such that $x^{-1}y^{-1}xy \in R_u(G)$ for all $x \in G$. Unipotent radicals and centres are preserved by automorphisms, so Y is characteristic. For any fixed $x \in G$, let $f_x : G \to G$ be defined by $f_x(y) = x^{-1}y^{-1}xy$, so that $\mathrm{mdeg}(f_x) \leq \iota + 1$. By Lemma 3.1(e) and [18, Prop. 3], $R_u(G)$ is defined by polynomials (over the algebraic closure $\overline{K}$) of degree $\leq d^d$. Thus, Y is defined by the polynomials of G and by polynomials of degree $\leq (\iota + 1)d^d$.

Applying Proposition 3.4, we obtain the various assertions on $\hat{G}, \hat{Y}, \lambda, \hat{\beta}, m$. □

4.2. **Classification of almost simple adjoint groups.** The almost simple adjoint groups can all be defined very explicitly by taking quotients, identity components, and preimages of subgroups in appropriate matrix spaces. See [25, Table 9.2] for a list of groups for each type, where the simply connected and adjoint extremes are highlighted. The adjoint groups for each classical type are: PGL_{r+1} for type A_r, SO_{2r+1} for type B_r, PCSp_{2r} for type C_r, and $\mathrm{P}((\mathrm{CO}_{2r}^{+})^{\circ})$ for type D_r. As for the exceptional types, the groups $E_6^{\mathrm{ad}}, E_7^{\mathrm{ad}}, E_8^{\mathrm{ad}}, F_4^{\mathrm{ad}}, G_2^{\mathrm{ad}}$ can be obtained via the adjoint representation on their own Lie algebras; note that for types E_8, F_4, G_2 there is a unique almost simple group of that type.

One can use the definition itself of each adjoint group G above and combine it with Lemmas 2.1(a)–4.1 to find a concrete way to write $G \leq \mathrm{GL}_m$ with bounds on m and $\deg(G)$. However, every adjoint group G is also the image of the adjoint representation $\mathrm{Ad}_{\tilde{G}}$ for any $\tilde{G}$ almost simple with the same Dynkin diagram as G. Thus, we can choose some suitable $\tilde{G}$ and define the adjoint groups as subgroups of matrices with more convenient bounds.

Proposition 4.2. *Every connected almost simple adjoint algebraic group G of dimension d and rank r can be written as a linear algebraic group $G \leq \mathrm{GL}_d$ with $\deg(G) \leq (2r)^{2^{16}r^2}$.*

Proof. Every connected almost simple adjoint G is the image of $\mathrm{Ad}_{\tilde{G}} : \tilde{G} \to \mathrm{GL}(\tilde{\mathfrak{g}})$ for any $\tilde{G}$ with the same Dynkin diagram as G. We only need to choose a suitable $\tilde{G}$.

We start with the classical types A_r, B_r, C_r, D_r. Going through the possibilities in [8, §1.11] or [25, Table 9.2] for each type, we may choose the following (Ω_C, Ω_D below are fixed constant matrices).

Type A_r: $\tilde{G} = \mathrm{SL}_{r+1} = \{x \in \mathrm{Mat}_{r+1} : \det(x) = 1\}$.

Type B_r: $\tilde{G} = G = \mathrm{SO}_{2r+1}$.

Type C_r: $\tilde{G} = \mathrm{Sp}_{2r} = \{x \in \mathrm{Mat}_{2r} : x^\top \Omega_C x = \Omega_C\}$.

Type D_r: $\tilde{G} = \mathrm{SO}^+_{2r} = \{x \in \mathrm{Mat}_{2r} : \det(x) = 1, x^\top \Omega_D x = \Omega_D\}$.

A uniform degree bound for every $\tilde{G}$ as above is $\deg(\tilde{G}) \leq (2r+1)2^{(2r+1)^2}$ (slightly better bounds appear in [2, Table 1]). By Lemma 2.1(a), (4.2), and the bound $\dim(G) \leq 2r^2+r$, we conclude that

$$\deg(G) \leq \deg(\tilde{G})\mathrm{mdeg}(\mathrm{Ad}_{\tilde{G}})^{\dim(G)} \leq (2r+1)2^{(2r+1)^2}(2r+2)^{2r^2+r} < 2^{17r^2}r^{2r^2+r+1}$$

for G of classical type.

We pass now to the exceptional types E_6, E_7, E_8, F_4, G_2. We may act as above and use any $\tilde{G}$, for which $\deg(\tilde{G})$ is bounded by some absolute constant, but it is possible to describe G very explicitly. Every G adjoint is the automorphism group of its own finite-dimensional Lie algebra, and in some cases of other algebras too: as already observed in [3, §6.1], this means that we can define G via quadratic equations

$$\sum_{i,j} \lambda_{ija} g_{bi} g_{cj} - \sum_k \lambda_{bck} g_{ka} = 0$$

for all triples (a,b,c), where we have specified a basis $\{e_i\}_i$ for the non-associative algebra with $g(e_i) = \sum_j g_{ij} e_j$ and $e_i \circ e_j = \sum_k \lambda_{ijk} e_k$. Hence, we conclude that $\deg(G) \leq 2^{\dim(G)^3}$, which for the exceptional types may be bounded uniformly by $(2r)^{2^{16}r^2}$, say. As this bound works for the classical types too, we obtain the result. □

Remark 4.3. A few observations about the result above.

(a) To bound $\deg(G)$, one could write G directly as a quotient and use Proposition 3.4. However, this route would give $G \leq \mathrm{Mat}_m$ with $m \leq C^{r^2}$ and $\deg(G) \leq C^{r^4}$ for some absolute constant C. Passing afterwards through the adjoint representation would give $G \leq \mathrm{GL}_d$, but $\deg(G)$ would grow even more.

(b) In order to deal with a specific computation later, i.e. (6.19), we now bound $\deg(G)$ more tightly for G connected almost simple adjoint of type A_1, B_2, G_2. For $G = \mathrm{PGL}_2$ and $\tilde{G} = \mathrm{SL}_2$, we use $\deg(G) \leq \deg(\tilde{G})\mathrm{mdeg}(\mathrm{Ad}_{\tilde{G}})^{\dim(G)}$ and obtain $\deg(G) \leq 16$. For $G = \tilde{G} = \mathrm{SO}_5$, [5, Thm. 1.1] yields $\deg(G) \leq 384$. For G of type G_2, the proof of Proposition 4.2 already shows the inequality $\deg(G) \leq 2^{7^3}$.

(c) We give references for several algebras that can be used for the groups G of exceptional type. The Lie algebras can be constructed rather explicitly from [19, §3]; even more explicit resources are the multiplication table of $\mathfrak{g}_2$ from [16, §22], and the complete bases of all five types from [13, App. A.1] (based on the aforementioned references). Moreover, as pointed out before, for types G_2, F_4, E_8 there is only one connected almost simple G per type; thus, other constructions not using Lie algebras will also give the same algebraic group. For G_2 we can use the octonion algebra, given in [31, §§4.3.2–4.3.4–4.4.3] in three bases with different conditions on the characteristic. For F_4 we can use the Albert algebra, or a quadratic and a cubic form based on its construction, as in [31, §4.8]. For E_8 there are no smaller representations than the Lie algebra $\mathfrak{e}_8$.

4.3. **Weyl groups.** For G connected and T a maximal torus, the *Weyl group* $\mathcal{W}$ is the quotient of the normalizer of T by the centralizer of T inside G [8, §1.9]. The construction is independent from the choice of T, so we can refer to $\mathcal{W}$ as the Weyl group of G. For G almost simple, $\mathcal{W}$ is a finite group determined by the Dynkin diagram: in fact, it can be equivalently defined using reflections of the root system [4, §14.7]. Below we present a list of $\mathcal{W}$ and their sizes, taken from [31, §2.8.4, (3.22), (3.31), (3.39)]; the notation on group extensions $A \times B$, $A^{\cdot}B$, $A : B$ and on abelian groups p^{m+n} is as in [31, §1.6], which follows [11, p. xx].

Type A_r:	$\mathcal{W} \simeq \mathrm{Sym}(r+1)$,	$\lvert\mathcal{W}\rvert = (r+1)!$,
Type B_r:	$\mathcal{W} \simeq \mathrm{Sym}(2) \wr \mathrm{Sym}(r)$,	$\lvert\mathcal{W}\rvert = 2^r r!$,
Type C_r:	$\mathcal{W} \simeq \mathrm{Sym}(2) \wr \mathrm{Sym}(r)$,	$\lvert\mathcal{W}\rvert = 2^r r!$,
Type D_r:	$\mathcal{W} \simeq H$ with $[\mathrm{Sym}(2) \wr \mathrm{Sym}(r) : H] = 2$,	$\lvert\mathcal{W}\rvert = 2^{r-1} r!$,
Type E_6:	$\mathcal{W} \simeq \mathrm{GO}_6^-(2) \simeq \mathrm{U}_4(2) : 2$,	$\lvert\mathcal{W}\rvert = 51840$,
Type E_7:	$\mathcal{W} \simeq \mathrm{GO}_7(2) \times 2 \simeq \mathrm{Sp}_6(2) \times 2$,	$\lvert\mathcal{W}\rvert = 2903040$,
Type E_8:	$\mathcal{W} \simeq 2^{\cdot}\mathrm{GO}_8^+(2) \simeq 2^{\cdot}\Omega_8^+(2) : 2$,	$\lvert\mathcal{W}\rvert = 696729600$,
Type F_4:	$\mathcal{W} \simeq \mathrm{GO}_4^+(3) \simeq 2^{1+4} : (\mathrm{Sym}(3) \times \mathrm{Sym}(3))$,	$\lvert\mathcal{W}\rvert = 1152$,
Type G_2:	$\mathcal{W} \simeq \mathrm{Dih}_6 \simeq \mathrm{Sym}(3) \times 2$,	$\lvert\mathcal{W}\rvert = 12$.

As a uniform cruder bound for all $|\mathcal{W}|$ at once, we shall use

$$|\mathcal{W}| \leq (2r)^r, \tag{4.3}$$

where r is the rank of G.

4.4. **Steinberg endomorphisms.** In this subsection we work over the field $\overline{\mathbb{F}_p}$.

Let G be a linear algebraic group $G \leq \mathrm{GL}_n$ defined over $\overline{\mathbb{F}_p}$, and let q be a power of p. The *Frobenius map (with respect to* $\mathbb{F}_q$*)* is the morphism $F_q : G \to G$ given by raising each entry of g to the q-th power. A *Steinberg endomorphism* [25, Def. 21.3] is an automorphism of abstract groups $F : G(\overline{\mathbb{F}_p}) \to G(\overline{\mathbb{F}_p})$ such that there are q and m for which F^m is (the map on $\overline{\mathbb{F}_p}$-points induced by) the Frobenius map of G with respect to $\mathbb{F}_q$.

This terminology comes from [25], but is not unanimously adopted in the literature: [24] uses the locution "Frobenius map" to refer to what we call a Steinberg endomorphism, and "standard Frobenius map" to refer to what we call a Frobenius map.

A Steinberg endomorphism is not necessarily a morphism of linear algebraic groups. The classification of Steinberg endomorphisms of connected almost simple groups is known: see [25, Thm. 22.5], which however is restricted to the simply connected case.

For a Steinberg endomorphism F of G, we denote by G^F the subgroup of $G(\overline{\mathbb{F}_p})$ of points fixed by F. The group G^F is finite whenever G is almost simple [25, Thm. 21.5]. The classification of all possible G^F for any connected almost simple G and any Steinberg endomorphism F is given in [8, §1.19] and in [25, §22.2].

4.5. **The representation ρ of G adjoint.** Finally, we must study a particular representation of G. We follow Pink's notation in [28].

Let G be connected almost simple adjoint defined over a field K, and let $\tilde{G}$ be the corresponding simply connected group. There is a natural isogeny $\pi : \tilde{G} \to G$, which induces a linear transformation $d\pi : \tilde{\mathfrak{g}} \to \mathfrak{g}$; the two Lie algebra have the same dimension, and the same is true for the kernel $\mathfrak{z}$ and cokernel $\mathfrak{z}^*$ of $d\pi$. We can write $\tilde{\mathfrak{g}} = \mathfrak{z} \oplus \bar{\mathfrak{g}}$ and $\mathfrak{g} = \bar{\mathfrak{g}} \oplus \mathfrak{z}^*$, so that $\bar{\mathfrak{g}}$ is the subalgebra whose copies inside $\tilde{\mathfrak{g}}$ and $\mathfrak{g}$ are identifiable with $\tilde{\mathfrak{g}}/\mathfrak{z}$ and $d\pi(\tilde{\mathfrak{g}})$ respectively; when restricted to these two copies of $\bar{\mathfrak{g}}$, the map $d\pi$ is given by some $\bar{\pi} \in \mathrm{GL}(\bar{\mathfrak{g}})$.

As we already showed in (4.2), the adjoint representation $\mathrm{Ad}_G : G \to \mathrm{GL}(\mathfrak{g})$ has $\mathrm{mdeg}(\mathrm{Ad}_G) \leq \iota + 1$. Following [28, (1.3)–(1.4)], we can define $\kappa : \mathrm{Hom}(\mathfrak{g}, \tilde{\mathfrak{g}}) \to \mathrm{End}(\mathfrak{g})$ by $\kappa(f) = \mathrm{Id}_{\mathfrak{g}} + d\pi \circ f$, and a morphism $\widetilde{\mathrm{Ad}}_G : G \to \mathrm{Hom}(\mathfrak{g}, \tilde{\mathfrak{g}})$ so that $\mathrm{Ad}_G = \kappa \circ \widetilde{\mathrm{Ad}}_G$. Therefore, $\mathrm{Ad}_{G,g}(v) = v + (d\pi \circ \widetilde{\mathrm{Ad}}_G(g))(v)$ for every $g \in G$ and $v \in \mathfrak{g}$; knowing $\mathrm{mdeg}(\mathrm{Ad}_G)$ and using the fact that $d\pi : \mathfrak{z} \oplus \bar{\mathfrak{g}} \to \bar{\mathfrak{g}} \oplus \mathfrak{z}^*$ acts as the zero map on $\mathfrak{z}$ and as $\bar{\pi}$ on $\bar{\mathfrak{g}}$, we obtain $\mathrm{mdeg}(\widetilde{\mathrm{Ad}}_G) \leq \iota + 1$.

Now, following [28, Prop. 1.10], let $\hat{\mathfrak{g}} = \mathfrak{z} \oplus \bar{\mathfrak{g}} \oplus \mathfrak{z}^*$, let $di : \tilde{\mathfrak{g}} \to \hat{\mathfrak{g}}$ be the natural inclusion, and let $d\omega : \hat{\mathfrak{g}} \to \mathfrak{g}$ be the map induced by $d\pi$: namely, $d\omega$ acts as the zero map, as $\bar{\pi}$, and as the identity on $\mathfrak{z}, \bar{\mathfrak{g}}, \mathfrak{z}^*$ respectively. Define the representation $\hat{\rho} : G \to \mathrm{GL}(\hat{\mathfrak{g}})$ by $\hat{\rho}(g) = \mathrm{Id}_{\hat{\mathfrak{g}}} + di \circ \widetilde{\mathrm{Ad}}_G(g) \circ d\omega$. By the discussion above, $\mathrm{mdeg}(\hat{\rho}) \leq \iota + 1$. Furthermore, $\hat{\rho}$ is defined over K, since this is true for all the objects defined so far.

By [28, Prop. 1.11(b)], if G does not have a non-standard isogeny then $\bar{\mathfrak{g}}$ is the unique simple G-submodule of $\mathfrak{g}$ and the unique simple quotient G-module of $\tilde{\mathfrak{g}}$, and the restriction of $\hat{\rho}$ on $\bar{\mathfrak{g}}$ is irreducible and non-constant. By [28, Prop. 1.11(c)], if G has a non-standard isogeny then there is a unique simple G-submodule $\bar{\mathfrak{g}}_s$ of $\mathfrak{g}$ and there is a unique simple quotient G-module $\bar{\mathfrak{g}}_\ell$ of $\tilde{\mathfrak{g}}$, and the restrictions of $\hat{\rho}$ on them are irreducible, non-constant, and not equivalent to each other; moreover, we can decompose $\hat{\mathfrak{g}}$ so that $\bar{\mathfrak{g}}_s$ and $\bar{\mathfrak{g}}_\ell$ are transversal direct summands of $\hat{\mathfrak{g}}$ (see the graphs in [28, Prop. 1.11(c)]). The restrictions of $\hat{\rho}$ to $\bar{\mathfrak{g}}, \bar{\mathfrak{g}}_s, \bar{\mathfrak{g}}_\ell$ in the various cases are denoted by

$\alpha^G, \alpha_s^G, \alpha_\ell^G$. Up to a change of basis, which does not affect $\mathrm{mdeg}(\hat{\rho})$, we see again that the restrictions $\alpha^G, \alpha_s^G, \alpha_\ell^G$ have maximal degree $\leq \iota + 1$.

It is time to define the representation ρ of G, as given in [24, p. 1142]. Let U be the unipotent part of a Borel subgroup B of G; up to isomorphism, U is independent from the choice of B. By [4, Prop. 8.3], the centre $Z(U)$ is of the form

$$Z(U) = \begin{cases} U_\alpha & \begin{array}{l}\text{if all roots have the same length,}\\ \alpha = \text{highest positive root,}\end{array} \\ & \\ U_{\alpha_\ell} U_{\alpha_s} & \begin{array}{l}\text{if roots have different lengths,}\\ \alpha_\ell = \text{highest long root,}\\ \alpha_s = \text{highest short root,}\end{array} \end{cases} \tag{4.4}$$

so in particular $\dim(Z(U)) \in \{1, 2\}$. During the proof of Proposition 6.3, we will fix a B-invariant subgroup $V \leq Z(U)$ of minimal dimension among those with $|\Gamma \cap V(\overline{K})| > 1$; based on that choice, we shall define

$$\rho := \begin{cases} \alpha^G & \text{if } Z(U) = V = U_\alpha, \\ (\alpha_\ell^G, \alpha_s^G) & \text{if } Z(U) = V = U_{\alpha_\ell} U_{\alpha_s}, \\ \alpha_\ell^G & \text{if } Z(U) = U_{\alpha_\ell} U_{\alpha_s} \text{ and } V = U_{\alpha_\ell}, \\ \alpha_s^G & \text{if } Z(U) = U_{\alpha_\ell} U_{\alpha_s} \text{ and } V = U_{\alpha_s}. \end{cases} \tag{4.5}$$

In all cases, ρ is a representation over K, and when $\dim(V) = 2$ it can be seen as a representation over K^2 as well. By what we argued so far, we can bound the maximal degree of ρ.

Proposition 4.4. *Let G be a connected almost simple adjoint linear algebraic group defined over a field K, with $d = \dim(G)$ and $\iota = \mathrm{mdeg}(^{-1})$, and let ρ be defined as in (4.5) (for an appropriate choice of V). Then $\rho : G \to \mathrm{GL}_m$ is a representation defined over K, and as such we have*

$$m \leq d, \qquad\qquad \mathrm{mdeg}(\rho) \leq \iota + 1 \leq d.$$

Proof. The bound we obtained above for $\alpha^G, \alpha_s^G, \alpha_\ell^G$ becomes directly $\mathrm{mdeg}(\rho) \leq \iota + 1$. We started the construction from a representation $G \leq \mathrm{GL}_d$ as in Proposition 4.2, so $\iota \leq d - 1$ as well. Comparing [28, Prop. 1.11] with the values of $|\Lambda_w / \Lambda_r|$ for each type (readable from [21, App. (A.9)] or [25, Table 9.2]), we see that $\dim(\bar{\mathfrak{g}}_s) + \dim(\bar{\mathfrak{g}}_\ell) \leq \dim(\mathfrak{g})$. □

5. Dimensional estimates

In this section we deal with dimensional estimates. Our work relies on the explicit estimates for $|A \cap V(K)|$ contained in [3], which we first adapt so as to replace a set A generating $G(K)$ with a subgroup $\Gamma \leq G(\overline{K})$ not necessarily equal to $G(K)$. This first result is given in Theorem 5.3. Then we provide a few additional variations that will be useful later.

We start with the following lemma, which allows us to grow in dimension using a generic element and the almost simplicity of G.

Lemma 5.1. *Let $G \leq \mathrm{GL}_n$ be an almost simple linear algebraic group over a field K. Let $\iota = \mathrm{mdeg}(^{-1})$, the maximum degree of the inversion map. Let V, V' be subvarieties of G defined over $\overline{K}$, with $\dim(V) < \dim(G)$ and $\dim(V') > 0$.*

Then, for every $g \in G(\overline{K})$ outside a variety $W = \{x \in \mathrm{Mat}_n : F(x) = 0\}$ with $\deg(F) \leq 1 + \min\{\iota, \dim(V)\}$ and $G \not\subseteq W$, the variety $\overline{VgV'}$ has dimension $> \dim(V)$.

Proof. See [3, Lem. 4.1]. □

Now we write our main inductive step for Theorem 5.3.

Proposition 5.2. *Let $G \leq \mathrm{GL}_n$ be a connected almost simple linear algebraic group of rank r with $\dim(G) = d$ and $\deg(G) = D$, defined over a field K. Let $\Gamma \leq G(\overline{K})$ be a finite group. Then at least one of the following holds:*

(a) $|\Gamma| \leq (2dD)^{d+1}$;

(b) $\Gamma \leq H(\overline{K})$ for some subgroup $H < G$ with $\dim(H) < d$ and $\deg(H) \leq (2dD)^{d+1}$;

(c) for any irreducible subvariety $V \subsetneq G$ defined over $\overline{K}$ with $0 < \dim(V) = d' < d$, there is an integer $\ell \leq d - d' + 1$ and there are proper subvarieties $F_1, \ldots, F_{\ell-1}$ of V and a variety $E \subsetneq \mathrm{Mat}_n$ with

$$|\Gamma \cap V(\overline{K})|^{\ell} \leq |\Gamma| \cdot \prod_{j=1}^{\ell-1} |\Gamma \cap F_j(\overline{K})| + \ell |\Gamma \cap E(\overline{K})| |\Gamma \cap V(\overline{K})|^{\ell-1}$$

and E not containing V, and satisfying the following properties as well:

$$\sum_{j=1}^{\ell-1} \dim(F_j) = \ell d' - d, \quad \deg(F_j) \leq 2^d \ell^{d-1} \deg(V)^{j+1}, \quad \deg(E) \leq (2\ell)^d \deg(V)^{\ell}. \tag{5.1}$$

Proof. The statement and proof are very similar to [3, Prop. 4.4], which differs from our result in that in lieu of Γ it has a set A that generates $G(K)$. The fact that $\langle A \rangle = G(K)$ was only needed in order to provide an escape argument, and obviously here we do not always have $\langle \Gamma \rangle = \Gamma = G(K)$. Hence, we have to replace the escape argument, which we are able to do by relying on Lemma 3.3.

Let us assume that Γ does not satisfy the conditions (a) and (b). Then, following Lemma 3.3 we get that $\Gamma \not\subseteq W(\overline{K})$ for any subvariety W of G with $\deg(W) \leq 2dD$. We shall show that Γ satisfies condition (c).

As in the statement, let $V \subsetneq G$ be any irreducible subvariety defined over $\overline{K}$ with $0 < \dim(V) = d' < d$. Fix an integer ℓ and elements $g_1, \ldots, g_{\ell-1} \in G(\overline{K})$ (to be chosen soon), and define

$$V_1 = V, \qquad\qquad V_{j+1} = \overline{V_j g_j V} \qquad (1 \leq j < \ell).$$

For each $1 \leq j \leq \ell$, let $d'_j = \dim(V_j)$ and $D_j = \deg(V_j)$. By Lemma 2.1(a), we have $D_j \leq \deg(V)^j j^{d'_j}$. We will show that the g_j can be chosen so that they belong to Γ and that we have $d'_{j+1} > d'_j$ at every step.

If G is a connected almost simple linear algebraic group, by Lemma 5.1 we get $d'_{j+1} > d'_j$ for any choice of g_j outside a variety W_j with $\deg(W_j) \leq 1 + \min\{\iota, d'_j\}$ and $G \not\subseteq W_j$. Consider the subvariety $G \cap W_j$ of G, which is a proper subvariety of dimension $< d$ and of degree

$$\deg(G \cap W_j) \leq D \deg(W_j) \leq D(d+1) \leq 2dD.$$

Then, by what we said before, $\Gamma \nsubseteq (G \cap W_j)\,(\overline{K})$.

Hence, we can choose an element g_j inside Γ with $d'_{j+1} > d'_j$ at each step j, and let ℓ be the least index such that $V_\ell = G$. Clearly, $\ell \leq d - d' + 1$, and $g_1, \ldots, g_{\ell-1} \in \Gamma$. The rest of the proof follows, mutatis mutandis, the steps in the proof of [3, Prop. 4.4]. □

Theorem 5.3. *Let $G \leq \mathrm{GL}_n$ be a connected almost simple linear algebraic group of rank r with $\dim(G) = d$ and $\deg(G) = D$, defined over a field K, and let $\Gamma \leq G(\overline{K})$ be a finite subgroup. Then at least one of the following holds:*

(a) $|\Gamma| \leq (2dD)^{d+1}$;

(b) $\Gamma \leq H(\overline{K})$ for some subgroup $H < G$ with $\dim(H) < d$ and $\deg(H) \leq (2dD)^{d+1}$;

(c) for any proper subvariety $V \subsetneq G$, we have

$$|\Gamma \cap V(\overline{K})| \leq C|\Gamma|^{\dim(V)/d} \quad \textit{with} \quad C \leq (2d\deg(V))^{d^{\dim(V)}}. \tag{5.2}$$

Proof. The statement and proof are almost as in [3, Thm. 4.5]. The inductive step here is provided by Proposition 5.2, which introduces cases (a) and (b). After replacing A with the group Γ in the conclusion, we can ignore the exponent C_2 since $\Gamma = \langle\Gamma\rangle$, and the computation of C_1, which is denoted by C here, is the same. □

The dimensional estimate of Theorem 5.3 can be extended to cover more general G. If one is not concerned about the correct exponent for $|\Gamma|$, this is quite easy to achieve in the group G^k (the direct product of k copies of G).

Corollary 5.4. *Let $G \leq \mathrm{GL}_n$ be a connected almost simple linear algebraic group of rank r with $\dim(G) = d$ and $\deg(G) = D$, defined over a field K, and let $\Gamma \leq G(\overline{K})$ be a finite subgroup. Then at least one of the following holds:*

(a) $|\Gamma| \leq (2dD)^{d+1}$;

(b) $\Gamma \leq H(\overline{K})$ for some subgroup $H < G$ with $\dim(H) < d$ and $\deg(H) \leq (2dD)^{d+1}$;

(c) for any $k \geq 1$ and any proper subvariety $V \subsetneq G^k$, we have

$$|\Gamma^k \cap V(\overline{K})| \leq k(2d\deg(V))^{d^{d-1}}|\Gamma|^{k-\frac{1}{d}}.$$

Proof. Apply the "crumbling" lemma given as [3, Lem. 4.3], so that

$$|\Gamma^k \cap V(\overline{K})| \leq k|\Gamma|^{k-1}|\Gamma \cap V'(\overline{K})|$$

for some subvariety $V' \subsetneq G$ with $\deg(V') \leq \deg(V)$, and then apply Theorem 5.3 to bound $|\Gamma \cap V'(\overline{K})|$. □

If we want to achieve the correct exponent for $|\Gamma^k \cap V(\overline{K})|$, we can set up an appropriate induction on both k and $\dim(V)$. We do so in full generality in the following lemma, and then we show below an application to orbits that will yield a useful lower bound for centralizers of subsets of Γ.

Lemma 5.5. *Let $G \leq \mathrm{GL}_n$ be a connected almost simple linear algebraic group of rank r with $\dim(G) = d$, $\deg(G) = D$ and $\mathrm{mdeg}(^{-1}) = \iota$, defined over a field K, and let $\Gamma \leq G(\overline{K})$ be a finite subgroup. Then at least one of the following holds:*

(a) $|\Gamma| \leq (2dD)^{d+1}$*;*
(b) $\Gamma \leq H(\overline{K})$ *for some subgroup* $H < G$ *with* $\dim(H) < d$ *and* $\deg(H) \leq (2dD)^{d+1}$*;*
(c) for any $k \geq 1$ *and any proper subvariety* $V \subsetneq G^k$*, we have*

$$|\Gamma^k \cap V(\overline{K})| \leq (2d\deg(V))^{kd^{\dim(V)}} |\Gamma|^{\frac{\dim(V)}{d}}.$$

Proof. Assume that (a) and (b) do not hold. We prove the inequality in (c) by applying a double induction on k and $\dim(V)$. The base case $k = 1$ follows from Theorem 5.3, and the other base case $\dim(V) = 0$ is vacuously true (since $|\Gamma^k \cap V(\overline{K})| \leq \deg(V)$). Assume that (c) holds for all pairs (k', V') that either have $k' < k$ and $\dim(V') \leq \dim(V)$ or have $k' \leq k$ and $\dim(V') < \dim(V)$. We may assume that V is irreducible, since the bound is more than linear in $\deg(V)$.

Consider the projection map $\pi : V \to G^{k-1}$ defined by

$$(x_1, \cdots, x_{k-1}, x_k) \mapsto (x_1, \cdots, x_{k-1}).$$

Since V is irreducible, by [3, Lem. 4.2]

$$|\Gamma^k \cap V(\overline{K})| \leq |\Gamma^{k-1} \cap \overline{\pi(V)}(\overline{K})||\Gamma^k \cap W(\overline{K})| + |\Gamma^k \cap E(\overline{K})|, \tag{5.3}$$

where $W = \pi^{-1}(y) \subsetneq G^k$ for some point $y \in \Gamma^{k-1} \cap \pi(V)$ with $\dim(W) = \dim(V) - \dim(\overline{\pi(V)})$, and where $E = \pi^{-1}(Z) \subsetneq V$ for some subvariety $Z \subsetneq \overline{\pi(V)}$ with $\deg(Z) \leq \mathrm{mdeg}(\pi)^{\dim(\overline{\pi(V)})-1} \deg(V)$. The projection π has $\mathrm{mdeg}(\pi) = 1$, so by Lemma 2.1(a) we obtain the bounds

$$\deg(\overline{\pi(V)}) \leq \deg(V), \qquad \deg(W) \leq \deg(V), \qquad \deg(E) \leq \deg(V)^2.$$

We have also $\dim(E) < \dim(V)$ since V is irreducible, so by induction on the dimension

$$|\Gamma^k \cap E(\overline{K})| \leq (2d\deg(V)^2)^{kd^{\dim(E)}} |\Gamma|^{\frac{\dim(E)}{d}} \leq (2d\deg(V))^{2kd^{\dim(V)-1}} |\Gamma|^{\frac{\dim(V)-1}{d}}, \tag{5.4}$$

and by induction on k

$$|\Gamma^{k-1} \cap \overline{\pi(V)}(\overline{K})| \leq (2d\deg(V))^{(k-1)d^{\dim(\overline{\pi(V)})}} |\Gamma|^{\frac{\dim(\overline{\pi(V)})}{d}}. \tag{5.5}$$

We now estimate $|\Gamma^k \cap V(\overline{K})|$ for all three different possibilities:

(1) $\dim(\overline{\pi(V)}) = 0$.
(2) $\dim(W) = 0$.
(3) $1 \leq \dim(W), \dim(\overline{\pi(V)}) \leq \dim(V) - 1$.

Consider first $\dim(\overline{\pi(V)}) = 0$. Since V is irreducible, $\overline{\pi(V)}$ is irreducible, therefore $\overline{\pi(V)} = \{y\}$ for some $y = (y_1, \cdots, y_{k-1})$. Hence $V = \pi^{-1}(y)$, which means that V is in fact a variety of the form $\{y_1\} \times \cdots \times \{y_{k-1}\} \times V$ sitting inside a copy of G itself. The result follows from Theorem 5.3.

Now consider $\dim(W) = 0$. In this case $\dim(\overline{\pi(V)}) = \dim(V)$, and W is a set of $\deg(W)$ points so $|\Gamma^k \cap W(\overline{K})| \leq \deg(W)$. Therefore, following (5.3) along with (5.4) and (5.5), for $k \geq 1$, $\dim(V) \geq 1$, and $d \geq 3$, we write that

$$\begin{aligned}|\Gamma^k \cap V(\overline{K})| &\leq (2d\deg(V))^{(k-1)d^{\dim(V)}+1} |\Gamma|^{\frac{\dim(V)}{d}} + (2d\deg(V))^{2kd^{\dim(V)-1}} |\Gamma|^{\frac{\dim(V)-1}{d}} \\ &\leq (2d\deg(V))^{kd^{\dim(V)}-2} |\Gamma|^{\frac{\dim(V)}{d}} + (2d\deg(V))^{kd^{\dim(V)}-1} |\Gamma|^{\frac{\dim(V)}{d}} \\ &\leq (2d\deg(V))^{kd^{\dim(V)}} |\Gamma|^{\frac{\dim(V)}{d}}.\end{aligned}$$

Finally consider the case $1 \leq \dim(W), \dim(\overline{\pi(V)}) \leq \dim(V) - 1$. Applying induction on the dimension, we know that

$$|\Gamma^k \cap W(\overline{K})| \leq (2d\deg(V))^{kd^{\dim(W)}} |\Gamma|^{\frac{\dim(W)}{d}}. \tag{5.6}$$

Following (5.5) and (5.6), we find that

$$\begin{aligned} |\Gamma^{k-1} \cap \overline{\pi(V)}(\overline{K})||\Gamma^k \cap W(\overline{K})| &\leq (2d\deg(V))^{kd^{\dim(W)}+(k-1)d^{\dim(\overline{\pi(V)})}} |\Gamma|^{\frac{\dim(W)+\dim(\overline{\pi(V)})}{d}} \\ &\leq (2d\deg(V))^{2kd^{\dim(V)-1}} |\Gamma|^{\frac{\dim(V)}{d}}, \end{aligned} \tag{5.7}$$

since $\dim(W) + \dim(\overline{\pi(V)}) = \dim(V)$. Using (5.4) and (5.7) in (5.3), we get

$$\begin{aligned} |\Gamma^k \cap V(\overline{K})| &\leq 2(2d\deg(V))^{2kd^{\dim(V)-1}} |\Gamma|^{\frac{\dim(V)}{d}} \\ &\leq (2d\deg(V))^{kd^{\dim(V)}} |\Gamma|^{\frac{\dim(V)}{d}}, \end{aligned}$$

since $d \geq 3$ and $\dim(V) \geq 1$. □

Thanks to the previous lemma, we are able to obtain not only an upper bound but also a lower bound for centralizers. This is due essentially to the *orbit-stabilizer theorem*: if a finite group G acts on a set X and $C_G(x), Gx$ are respectively the stabilizer and the orbit of a point $x \in X$ under this action, then $|C_G(x)||Gx| = |G|$. If G is an algebraic group, as in our case, in certain cases an alternative version of the theorem still holds. If G is a connected semisimple algebraic group acting on itself by conjugation, $C_G(x)$ is the centralizer as defined in Section 3.1, and $\mathrm{Cl}_G(x)$ (the *conjugacy class* of x) is the image of G under the map $\sigma_x : G \to G$ given by $\sigma_x(g) = gxg^{-1}$, then

$$\dim(C_G(x)) + \dim(\overline{\mathrm{Cl}_G(x)}) = \dim(G) \tag{5.8}$$

(see for instance [20, §1.5]). Moreover, (5.8) still holds if G acts by conjugation on G^k as follows: $(g, (x_1, \cdots, x_k)) \mapsto \left(gx_1g^{-1}, \cdots, gx_kg^{-1}\right)$. In fact, we can see $C_G(x)$ as the fibre of x through $\sigma_x : G \to G^k$, and then (5.8) follows from fundamental results of algebraic geometry (see [27, §I.8], used also in [3, §4] and when we say $\dim(W) = \dim(V) - \dim(\overline{\pi(V)})$ in the proof of Lemma 5.5) and from the fact that every fibre through σ_x must have the same dimension (as they are images of each other through appropriate conjugation maps).

Corollary 5.6. *Let $G \leq \mathrm{GL}_n$ be a connected almost simple linear algebraic group of rank r with $\dim(G) = d$, $\deg(G) = D$ and $\mathrm{mdeg}(^{-1}) = \iota$, defined over a field K, and let $\Gamma \leq G(\overline{K})$ be a finite subgroup. Then at least one of the following holds:*

(a) $|\Gamma| \leq (2dD)^{d+1}$;
(b) $\Gamma \leq H(\overline{K})$ for some subgroup $H < G$ with $\dim(H) < d$ and $\deg(H) \leq (2dD)^{d+1}$;
(c) for any subset $\Lambda \subseteq \Gamma$, the centralizer $C_G(\Lambda)$ satisfies the bounds

$$\frac{1}{\varphi(d-d')}|\Gamma|^{\frac{d'}{d}} \leq |\Gamma \cap C_G(\Lambda)(\overline{K})| \leq \varphi(d')|\Gamma|^{\frac{d'}{d}}$$

where $d' = \dim(C_G(\Lambda))$ and $\varphi(x) = (2dD(\iota + 1))^{x(d+1)d^x}$.

This result corresponds to [24, Thm. 6.2]. We are more pedantic with the constants, using the function $\varphi(x)$ instead of a unique constant c_0 as in [24], since otherwise the tower in the bound of Theorem 1.2(a) has one more floor (due to the factorial in (6.16)).

Proof. Our strategy is to use the upper bounds coming from the previous theorems on both centralizers and conjugacy classes, and transform the upper bound for the latter into a lower bound for the former via the orbit-stabilizer theorem and (5.8). Since we have a set Λ instead of a single element x we must work in the direct product G^k, which is why we needed Lemma 5.5 and we had to explain why (5.8) works for the action of G on G^k.

Assume that (a) and (b) do not hold. By Corollary 3.2 we may assume that $\Lambda = \{\gamma_1, \cdots, \gamma_k\} \subseteq G(\overline{K})$ with $k \leq d+1$, and also $\deg(C_G(\Lambda)) \leq D$. Call $\vec{\lambda} = (\gamma_1, \cdots, \gamma_k) \in G^k$. G acts on G^k by conjugation as follows: $(g, (g_1, \cdots, g_k)) \mapsto \left(gg_1g^{-1}, \cdots, gg_kg^{-1}\right)$. Under this action $C_G(\vec{\lambda}) = C_G(\Lambda)$, and following Theorem 5.3(c)

$$|\Gamma \cap C_G(\Lambda)(\overline{K})| \leq C_1 |\Gamma|^{\frac{\dim(C_G(\Lambda))}{d}}, \qquad C_1 \leq (2d \deg(C_G(\Lambda)))^{d^{\dim(C_G(\Lambda))}}, \tag{5.9}$$

so that $C_1 \leq \varphi(\dim(C_G(\Lambda)))$. Similarly, Lemma 5.5 gives

$$|\Gamma^k \cap \overline{\mathrm{Cl}_G(\vec{\lambda})}(\overline{K})| \leq C_2 |\Gamma|^{\frac{\dim(\overline{\mathrm{Cl}_G(\vec{\lambda})})}{d}}, \qquad C_2 \leq (2d \deg(\overline{\mathrm{Cl}_G(\vec{\lambda})}))^{kd^{\dim(\overline{\mathrm{Cl}_G(\vec{\lambda})})}}. \tag{5.10}$$

By (5.8) we have $\dim(C_G(\Lambda)) + \dim(\overline{\mathrm{Cl}_G(\vec{\lambda})}) = d$, and by Lemma 2.1(a) $\deg(\overline{\mathrm{Cl}_G(\vec{\lambda})}) \leq D(\iota+1)^{\dim(\overline{\mathrm{Cl}_G(\vec{\lambda})})}$, so that $C_2 \leq \varphi(\dim(\overline{\mathrm{Cl}_G(\vec{\lambda})}))$.

To make (5.10) into a lower bound, we now need to pass to stabilizers and orbits in Γ itself. The equality $C_\Gamma(\Lambda) = \Gamma \cap C_G(\Lambda)(\overline{K})$ is trivial by definition, and since $\Lambda \subseteq \Gamma$ we also have the inclusion $\mathrm{Cl}_\Gamma(\vec{\lambda}) \subseteq \Gamma^k \cap \overline{\mathrm{Cl}_G(\vec{\lambda})}(\overline{K})$. By the orbit-stabilizer theorem then

$$|\Gamma \cap C_G(\Lambda)(\overline{K})| = \frac{|\Gamma|}{|\mathrm{Cl}_\Gamma(\vec{\lambda})|} \geq \frac{1}{C_2}|\Gamma|^{1-\frac{\dim(\overline{\mathrm{Cl}_G(\vec{\lambda})})}{d}} = \frac{1}{C_2}|\Gamma|^{\frac{\dim(C_G(\Lambda))}{d}}. \tag{5.11}$$

We get the desired inequalities in (c) by combining (5.9) and (5.11). □

6. Finding a simple group of Lie type

The present section is devoted to giving an explicit version of [24, Thm. 0.5]. In brief, we are in a situation where Γ is a finite subgroup of $G(\overline{K})$, where G is already assumed to be connected almost simple adjoint. Under these conditions, we shall conclude that either Γ is not "sufficiently general" (meaning that either $|\Gamma|$ is bounded or Γ is contained in $H(\overline{K})$ with H of bounded degree and strictly smaller dimension) or Γ is, up to a small index, a finite simple group of Lie type in the same characteristic as K (meaning that $[G^F, G^F] \leq \Gamma \leq G^F$ for some Steinberg endomorphism F).

All the results of this section work under the same assumptions, except that we may need to use two different explicit versions of the "sufficiently general" condition. We collect the assumptions here and reference the appropriate choices throughout the section to make the statements less cumbersome.

Assumption. $G \leq \mathrm{GL}_n$ is a connected almost simple adjoint linear algebraic group of rank r defined over an arbitrary field K, with $d = \dim(G)$, $D = \deg(G)$, and $\iota = \mathrm{mdeg}(^{-1})$. $\Gamma \leq G(\overline{K})$ is a finite subgroup, either satisfying

(Assumption A1) $|\Gamma| > (2dDrn\iota)^{(2dDr\iota)^{10d^4}}$, and

(Assumption A2) $\Gamma \not\leq H(\overline{K})$ for any linear algebraic subgroup $H \lneq G$ with

$$\dim(H) < d \quad \text{and} \quad \deg(H) \leq (2dD)^{4d},$$

or satisfying

(Assumption B1) $|\Gamma| > (2dDrn\iota)^{(2dDr\iota)^{11d^4}}$, and

(Assumption B2) $\Gamma \not\leq H(\overline{K})$ for any linear algebraic subgroup $H \lneq G$ with

$$\dim(H) < d \quad \text{and} \quad \deg(H) \leq (2dDr)^{4d^2}.$$

The reader should not be alarmed by the salad of letters appearing in the bound for $|\Gamma|$: what is really important is the number of exponential floors we use, and we include all the parameters with the sole intention of covering the many ways in which the assumption plays its role in the rest of the paper. In the final theorem, everything can be expressed just in terms of the rank.

6.1. **Finding the finite field $\mathbb{F}_q$.** The first goal is to find the "correct" field $\mathbb{F}_q$ that underlies the structure of Γ: if the final objective is to have $[G^F, G^F] \leq \Gamma \leq G^F$ for some Steinberg endomorphism F (as defined in Section 4.4), we need to know to what field the endomorphism is referred.

First, we shall find a regular unipotent element $u \in \Gamma$. Its existence proves incidentally that $\mathrm{char}(K) \neq 0$, since otherwise any nontrivial unipotent element generates an infinite group. Then we shall use u to find a unipotent algebraic subgroup V for which $\Gamma \cap V(\overline{K})$ is isomorphic (as a group) to $\mathbb{F}_q$. The group V is essentially made of "minimal" elements: for instance, if $G = \mathrm{PGL}_n$ and u is upper triangular, V is the variety of unipotent elements having their only nonzero off-diagonal entry in the upper right corner; then, the values taken in that entry by elements of Γ form the field $\mathbb{F}_q$.

Following [24, p. 1129], we call $\Lambda \subseteq G(\overline{K})$ a *toric subset* if it is contained in an algebraic torus of G. The preliminary result below shows that centralizers of toric subsets of Γ contain mostly regular semisimple elements of Γ.

Lemma 6.1. *Let $G, n, r, K, d, D, \iota, \Gamma$ be as in Assumption A1–A2. Then, for any toric subset $\Lambda \subseteq \Gamma$ we have*

$$|(\Gamma \cap C_G(\Lambda)^{\mathrm{o}})^{\mathrm{rss}}| \geq \left(1 - \frac{1}{2(2r)^r}\right) |\Gamma \cap C_G(\Lambda)^{\mathrm{o}}|.$$

Proof. We follow [24, Prop. 7.2]. For $g \in G(\overline{K})$, by (4.2) the matrix $\mathrm{Ad}_{G,g} \in \mathrm{Gl}(\mathfrak{g})(\overline{K})$ given by the adjoint representation has entries of degree $\leq \iota + 1$ in the entries of g. We know that g is not regular semisimple when the characteristic polynomial of Ad_g (call it $p_g(t)$) is divisible by $(t-1)^{r+1}$; using Hasse derivatives, this condition becomes $p_g^{(r)}(1) = 0$, which is an equation of degree $\leq d(\iota + 1)$ in the entries of g. Let V be the variety of such g.

Since Λ is toric, let T be a maximal torus containing Λ. Every torus is connected [4, §8.5] and its elements commute with all elements of Λ by definition, therefore $C_G(\Lambda)^{\mathrm{o}} \supseteq$

T. Since regular semisimple elements are dense in T (see [20, §2.3]), the intersection $V\cap T$ is proper inside T, so $V\cap C_G(\Lambda)^{\circ}\subsetneq C_G(\Lambda)^{\circ}$ as well, and in particular $\dim(V\cap C_G(\Lambda)^{\circ})\leq\dim(C_G(\Lambda)^{\circ})-1$.

We also have $\deg(V\cap C_G(\Lambda)^{\circ})\leq\deg(V)\deg(C_G(\Lambda))\leq d(\iota+1)D$ by Corollary 3.2. By Assumption A1–A2, Theorem 5.3(c), and Corollary 5.6(c),

$$\begin{aligned}|\Gamma\cap(V\cap C_G(\Lambda)^{\circ})(\overline{K})|&\leq(2d^2D(\iota+1))^{d^{d-1}}|\Gamma|^{\frac{\dim(C_G(\Lambda)^{\circ})-1}{d}}\\&\leq\frac{1}{(2dD(d+1))^{(d+1)d^{d+1}}\cdot 2D(2r)^r}|\Gamma|^{\frac{\dim(C_G(\Lambda)^{\circ})}{d}}\\&\leq\frac{1}{2D(2r)^r}|\Gamma\cap C_G(\Lambda)(\overline{K})|\leq\frac{1}{2(2r)^r}|\Gamma\cap C_G(\Lambda)^{\circ}(\overline{K})|,\end{aligned}$$

using the obvious equality $\dim(C_G(\Lambda))=\dim(C_G(\Lambda)^{\circ})$ and the fact that by Corollary 3.2 we have $|\Gamma\cap C_G(\Lambda)(\overline{K})|\leq\deg(C_G(\Lambda))|\Gamma\cap C_G(\Lambda)^{\circ}(\overline{K})|\leq D|\Gamma\cap C_G(\Lambda)^{\circ}(\overline{K})|$. □

Now we prove that there is a regular unipotent $u\in\Gamma$.

Proposition 6.2. *Let G,n,r,K,d,D,ι,Γ be as in Assumption A1–A2. Then $|\Gamma^{\mathrm{run}}|\geq\frac{1}{2}|\Gamma^{\mathrm{un}}|\geq 1$ and $\mathrm{char}(K)\neq 0$.*

Proof. We follow the proof leading to [24, Cor. 7.10]; we skip the theoretical details, focusing on the explicit bounds.

Fix any toric subset $\Lambda\subseteq\Gamma$, let $\mathcal{T}(\Lambda)$ be the set of all maximal toric subgroups inside $\Gamma\cap C_G(\Lambda)^{\circ}(\overline{K})$, and pick a set $\mathcal{S}(\Lambda)\subseteq\mathcal{T}(\Lambda)$ of representatives of $(\Gamma\cap C_G(\Lambda)^{\circ}(\overline{K}))$-conjugacy classes of elements of $\mathcal{T}(\Lambda)$. By Lemma 6.1, we prove that

$$1-\frac{1}{2(2r)^r}\leq\sum_{\Theta\in\mathcal{S}(\Lambda)}\frac{1}{[N_{\Gamma\cap C_G(\Lambda)^{\circ}(\overline{K})}(\Theta):\Theta]}\leq\frac{1}{1-\frac{1}{2(2r)^r}};\tag{6.1}$$

furthermore, each denominator in the sum above divides the size of the Weyl group of G, so by (4.3) the sum in (6.1) must be equal to 1. This implies

$$\sum_{\Theta\in\mathcal{T}(\Lambda)}|\Theta|=|\Gamma\cap C_G(\Lambda)^{\circ}(\overline{K})|,\qquad\sum_{\Theta\in\mathcal{T}(\Lambda)}1=|(\Gamma\cap C_G(\Lambda)^{\circ}(\overline{K}))^{\mathrm{un}}|.\tag{6.2}$$

For all $\Theta\in\mathcal{T}(\Lambda)$, we also see from the proof of [24, Prop. 7.3] that

$$\Theta=\Gamma\cap C_G(\Theta)^{\circ}(\overline{K}),\qquad\dim(C_G(\Theta)^{\circ})=r.\tag{6.3}$$

We can apply the second equality of (6.2) to $\Lambda=\emptyset$, for which $C_G(\emptyset)^{\circ}=G^{\circ}=G$, then the first equality of (6.3), then the upper bound of Corollary 5.6 to $C_G(\Theta)\supseteq C_G(\Theta)^{\circ}$, and finally the second equality of (6.3) and the fact that the sum in (6.1) is equal to 1. We conclude that

$$|\Gamma^{\mathrm{un}}|\overset{(6.2)}{=}\sum_{\Theta\in\mathcal{T}(\emptyset)}1=\sum_{\Theta\in\mathcal{S}(\emptyset)}\frac{|\Gamma|}{|N_{\Gamma}(\Theta)|}\overset{(6.3)}{=}\sum_{\Theta\in\mathcal{S}(\emptyset)}\frac{|\Gamma|}{[N_{\Gamma}(\Theta):\Theta]\cdot|\Gamma\cap C_G(\Theta)^{\circ}|}$$

$$\overset{\text{Cor. 5.6}}{\geq}\sum_{\Theta\in\mathcal{S}(\emptyset)}\frac{|\Gamma|^{1-\frac{\dim(C_G(\Theta)^{\circ})}{d}}}{[N_{\Gamma}(\Theta):\Theta]\cdot(2dD(\iota+1))^{(d+1)d^{d+1}}}\overset{(6.1)-(6.3)}{=}\frac{|\Gamma|^{1-\frac{r}{d}}}{(2dD(\iota+1))^{(d+1)d^{d+1}}}.\tag{6.4}$$

Now we bound $|\Gamma^{\mathrm{un}} \setminus \Gamma^{\mathrm{run}}| = |\Gamma \cap (G^{\mathrm{un}} \cap G^{\mathrm{irr}})(\overline{K})|$. As mentioned in Section 3.1, G^{un} is an irreducible variety of dimension $d-r$, and by Corollary 2.4(c) it has degree

$$\deg(G^{\mathrm{un}}) = \deg(G \cap \{x \in \mathrm{Mat}_{n+1} : (x|_{n\times n} - \mathrm{Id}_n)^n = 0\}) \leq Dn^{d+1}.$$

Then, $G^{\mathrm{un}} \cap G^{\mathrm{irr}}$ is a proper subvariety of G^{un} [20, §4.13], and it can be defined as the set of $g \in G^{\mathrm{un}}$ for which, if $f : G \times G^{\mathrm{un}} \to G^{\mathrm{un}}$ is the map given by $f(x,y) = xyx^{-1}$, the fibre $f^{-1}(g)$ is larger than the generic one [20, §1.4]. By Proposition 2.2, there is some variety Z with $G^{\mathrm{un}} \cap G^{\mathrm{irr}} \subseteq Z \subsetneq G^{\mathrm{un}}$ (so in particular $\dim(Z) < d-r$) and $\deg(Z) \leq (\iota+2)^{d-r-1}D^2 n^{d+1}$. Then, Theorem 5.3(c) gives

$$|\Gamma^{\mathrm{un}} \setminus \Gamma^{\mathrm{run}}| \leq |\Gamma \cap Z(\overline{K})| \leq \left(2d(\iota+2)^{d-r-1}D^2 n^{d+1}\right)^{d^{d-r-1}} |\Gamma|^{1-\frac{r+1}{d}}, \tag{6.5}$$

since the alternatives (a) and (b) of Theorem 5.3 do not hold by Assumption A1–A2. Combining (6.4) and (6.5) with the condition on $|\Gamma|$ coming from Assumption A1, we obtain the claim on Γ^{run}.

Finally, the existence of a regular unipotent element implies that $\mathrm{char}(K)$ divides its order, so $\mathrm{char}(K) \neq 0$ (see [24, Cor. 7.11]). □

We conclude the subsection by finding the desired algebraic subgroup V of minimal unipotent elements. In the following, the normalizer $N_\Gamma(V) = \{\gamma \in \Gamma : \gamma V = V\gamma\}$ is a subgroup of $\Gamma \leq G(\overline{K}) \leq \mathrm{GL}_n(\overline{K})$ with $\mathrm{char}(K) = p$, so it makes sense to talk about the group ring $\mathbb{F}_p[N_\Gamma(V)]$.

Proposition 6.3. *Let $G, n, r, K, d, D, \iota, \Gamma$ be as in Assumption A1–A2.*

Then $\mathrm{char}(K) = p > 0$, there is some $q = p^e$, and there is some abelian unipotent subgroup $V = V^{(G)} \leq G$ of dimension $1 \leq \dim(V) \leq \min\{2, r\}$ and degree $\deg(V) \leq D$ such that $\Gamma \cap V(\overline{K}) \simeq \mathbb{F}_q$ (as abelian groups), $\mathbb{F}_q \leq K^{\dim(V)}$ (as $\mathbb{F}_p$-algebras), $\mathbb{F}_q \simeq \mathbb{F}_p[N_\Gamma(V)]$ (as rings), and

$$\frac{1}{\varphi(d-1)}|\Gamma|^{\frac{\dim(V)}{d}} \leq |\Gamma \cap V(\overline{K})| = q \leq \varphi(2)|\Gamma|^{\frac{\dim(V)}{d}}, \quad \varphi(x) = (2dD(\iota+1))^{x(d+1)d^x}.$$

The function $\varphi(x)$ above is the same as in Corollary 5.6.

Proof. We follow the proof leading to [24, Thm. 8.17], skipping the theoretical details.

By Proposition 6.2, $\mathrm{char}(K) = p > 0$ and there exists some $u \in \Gamma^{\mathrm{run}}$. Pick a Borel subgroup B of G so that its unipotent part U contains u: $\Gamma \cap U(\overline{K})$ is a normal Sylow p-subgroup of $\Gamma \cap B(\overline{K})$, so by the Schur–Zassenhaus theorem $\Gamma \cap B(\overline{K})$ is a semidirect product of $\Gamma \cap U(\overline{K})$ and the corresponding quotient; then [21, §19.4, Prop. (a)] guarantees that there is a maximal torus T in B such that

$$\Gamma \cap B(\overline{K}) = (\Gamma \cap U(\overline{K})) \rtimes (\Gamma \cap T(\overline{K})). \tag{6.6}$$

First we prove that $Z(U)$ is a centralizer, following the steps of [24, Lem. 8.9]. T is a maximal torus of G as well [21, §21.3, Cor. A], and T is the centralizer of itself [21, §26.2, Cor. A(b)], so the lower bound of Corollary 5.6(c) yields

$$|\Gamma \cap T(\overline{K})| \geq \frac{1}{(2dD(\iota+1))^{(d-r)(d+1)d^{d-r}}}|\Gamma|^{\frac{r}{d}}, \tag{6.7}$$

since Assumption A1–A2 forbids the other cases. Now, if $f : T \to U$ is the conjugation map $f(t) = tut^{-1}$, then for any given element $g \in C_G(f(\Gamma \cap T(\overline{K})))$ we get $\Gamma \cap T(\overline{K}) \subseteq f^{-1}(C_U(g))(\overline{K})$, and in particular

$$|\Gamma \cap T(\overline{K})| \leq |\Gamma \cap f^{-1}(C_U(g))(\overline{K})|. \tag{6.8}$$

On the other hand,

$$\begin{aligned}\deg(f^{-1}(C_U(g))) &\leq \deg(T)\deg(C_U(g))\mathrm{mdeg}(f)^{\dim(C_U(g))} \\ &\leq \deg(T)\deg(U)\mathrm{mdeg}(f)^{\dim(U)} \leq D^2(\iota+1)^{d-r}\end{aligned}$$

using Lemma 2.1(b), the inequality $\deg(C_U(g)) \leq \deg(U)$ (proved as in Corollary 3.2), Lemma 3.1(a)–(d), and $\dim(U) \leq d-r$ (see for instance [20, §4.2]). Therefore, calling $r' = \dim(f^{-1}(C_U(g)))$, we have

$$|\Gamma \cap f^{-1}(C_U(g))(\overline{K})| \leq (2dD^2(\iota+1)^{d-r})^{dr'}|\Gamma|^{\frac{r'}{d}} \tag{6.9}$$

by Theorem 5.3(c). If $r' < \dim(T) = r$, Assumption A1 becomes incompatible with (6.7)–(6.8)–(6.9), so we must have $r' = \dim(T)$. This implies $f^{-1}(C_U(g)) = T$ since T is irreducible by definition, and thus that $f(T) \subseteq C_U(g)$ for any $g \in C_G(f(\Gamma \cap T(\overline{K})))$. Since U is also irreducible and $f(T)$ is made of regular elements, the image $f(T)/[U,U]$ of $f(T) \subseteq U$ in the quotient $U/[U,U]$ is the quotient's unique open T-orbit (see the description in terms of root subgroups in [20, §4.1]). Thus $f(T)/[U,U]$ generates $U/[U,U]$ as an algebraic group, and because U is nilpotent similarly $f(T)$ generates U. Hence, since $f(T) \subseteq C_U(g)$ and the latter is an algebraic group too, we have $U = C_U(g)$, or in other words $g \in Z(U)$. By our choice of g then $C_G(f(\Gamma \cap T(\overline{K}))) \subseteq Z(U)$, and since the reverse inclusion is trivial we conclude that $Z(U)$ is a centralizer. In particular, by Corollary 5.6(c) we obtain

$$\frac{1}{\varphi(d-1)}|\Gamma|^{\dim(Z(U))/d} \leq |\Gamma \cap Z(U)(\overline{K})| \leq \varphi(2)|\Gamma|^{\dim(Z(U))/d} \tag{6.10}$$

with $\varphi(x)$ as in the statement, using the fact that $1 \leq \dim(Z(U)) \leq 2$ by (4.4).

Now we search for our V. The group $Z(U)$ is nontrivial since U is nilpotent, and every nontrivial B-invariant subgroup $V \subseteq Z(U)$ (including $V = Z(U)$ itself) is connected with $V = C_G(V) = C_G(C_G(V))$ by [24, Prop. 8.4]. As $\dim(Z(U)) \geq 1$, (6.10) and Assumption A1 imply that there is some $v \in (\Gamma \cap Z(U)(\overline{K})) \setminus \{e\}$. From now on, fix a $V = V^{(G)}$ of minimal dimension among all possible B-invariant subgroups of $Z(U)$ with $|\Gamma \cap V(\overline{K})| > 1$, where B also runs among all possible Borel subgroups of G. We warn the reader that, although here we just write V instead of $V^{(G)}$, the dependence on G is important in later subsections.

By what we said before, such a V exists and $\dim(V) \geq 1$; V is abelian since $Z(U)$ contains it, which implies also $\dim(V) \leq \min\{2, r\}$. Moreover, as V is a centralizer, Corollary 3.2 yields $\deg(V) \leq D$ and Corollary 5.6 yields

$$\frac{1}{\varphi(d-1)}|\Gamma|^{\dim(V)/d} \leq |\Gamma \cap V(\overline{K})| \leq \varphi(2)|\Gamma|^{\dim(V)/d} \tag{6.11}$$

with $\varphi(x)$ as in the statement. Combining again the above with Assumption A1, in which the bound is larger than $\varphi(d-1)^{5d}$, we have $|\Gamma \cap V(\overline{K})| \geq \varphi(d-1)^4$. Recall now that $\mathrm{char}(K) = p$, and consider the ring $\mathbb{F}_p[N_\Gamma(V)]$: by [24, Prop. 8.4(a)-8.15], it is in fact a finite field $\mathbb{F}_q$ and at the same time a finite $\mathbb{F}_p$-subalgebra of $K^{\dim(V)}$, and $\Gamma \cap V(\overline{K})$ is a nontrivial finite-dimensional $\mathbb{F}_q$-vector space. Then we plug [24, Thm. 6.6] and the bound $|\Gamma \cap V(\overline{K})| \geq \varphi(d-1)^4$ into [24, Lem. 8.16], and we obtain that the dimension of $\Gamma \cap V(\overline{K})$ as a vector space over $\mathbb{F}_q$ is 1. The various claims on $\mathbb{F}_q$ and V follow from this conclusion and from (6.11). □

6.2. **Finding the simple group** $[G^F, G^F]$**.** Now that we have identified the correct field $\mathbb{F}_q$, a copy of which occurs inside Γ as a subgroup $\Gamma \cap V(\overline{K})$ of minimal unipotent elements, we must show that $\mathbb{F}_q$ is a good choice throughout the whole Γ. More rigorously, as said before, our goal is now to prove that $[G^F, G^F] \leq \Gamma \leq G^F$ for some Steinberg endomorphism F with respect to $\mathbb{F}_q$; additionally, $[G^F, G^F]$ shall be a finite simple group of Lie type.

The procedure is covered in [24, §§9–10–11]. In this subsection we follow the theoretical proof with just enough details to allow the reader to understand the process, and focus instead on the quantitative bookkeeping. Overall, the results here are arranged differently than in [24]. The correspondence is in general terms as follows:

$$\begin{tikzcd}
\textbf{Larsen–Pink:} & & \textbf{This paper:} \\
\text{Thm. 9.1, basic case, proved in §10} \arrow[d, Rightarrow] \arrow[r, leftrightarrow] & {} & \text{Prop. 6.4, under Assumption A1–A2} \arrow[dd, Rightarrow, "\text{(Prop. 6.5)}"] \\
\text{Thm. 0.5, basic case, proved in §9} \arrow[d, Rightarrow] & & \\
\text{Thm. 9.1, general case, proved in §11} \arrow[d, Rightarrow] \arrow[r, leftrightarrow] & {} & \text{Prop. 6.6, under Assumption B1–B2} \arrow[d, Rightarrow, "\text{(Prop. 6.5)}"] \\
\text{Thm. 0.5, general case, proved in §9} \arrow[r, leftrightarrow] & {} & \text{Thm. 6.7, under Assumption B1–B2}
\end{tikzcd}$$

The fact that the basic case of [24, Thm. 9.1] is used to prove the general case of the same result is the reason why we have two versions of the quantitative assumptions.

We start by showing that, under the weaker quantitative Assumption A1–A2 and the stronger condition $\dim(V) = r$, we can build a representation of G onto a module that, up to changing the base field from K to $\mathbb{F}_q$, is also Γ-invariant. One can imagine that this fact gets us close to saying $\Gamma = G(\mathbb{F}_q)$; the last statement is too strong, but it shall be relaxed to saying $[G^F, G^F] \leq \Gamma \leq G^F$ instead.

Proposition 6.4. *Let $G, n, r, K, d, D, \iota, \Gamma$ be as in Assumption A1–A2. Since Proposition 6.3 holds, let $p, q = p^e, V = V^{(G)}$ be as given therein. Assume that* $\dim(V) = r$.

Then, there is a $K^{\dim(V)}$-module M, there is a nontrivial representation $\sigma : G \to \mathrm{GL}(M)$ with $\mathrm{mdeg}(\sigma) \leq d^2$ (σ defined over K), and there is an $\mathbb{F}_q$-submodule M_0 of M with $M_0 \otimes_{\mathbb{F}_q} K^{\dim(V)} \simeq M$ and $\sigma(\gamma)(M_0) = M_0$ for every $\gamma \in \Gamma$.

Proof. The proof follows chiefly [24, §10].

The process of constructing V involves fixing B, U, T, respectively a Borel subgroup, its unipotent radical, and a maximal torus T satisfying (6.6), so let them be fixed here as well. Since $\dim(V) = r$ we must have $\dim(Z(U)) = r$ (by construction $V \leq Z(U) \leq T$), which can happen only for $r = 1, 2$ by (4.4). If $\Lambda \subseteq \Gamma$ is the set of elements conjugate to an element of $(\Gamma \cap T(\overline{K}))^{\mathrm{rss}}$, then

$$(6.12)\quad |\Lambda| = \frac{|(\Gamma \cap T(\overline{K}))^{\mathrm{rss}}|}{|\Gamma \cap T(\overline{K})|} \cdot \frac{1}{[N_\Gamma(T) : \Gamma \cap T(\overline{K})]} \cdot |\Gamma| \geq \left(1 - \frac{1}{2(2r)^r}\right) \cdot \frac{1}{|\mathcal{W}|} \cdot |\Gamma| \geq \frac{|\Gamma|}{2|\mathcal{W}|}$$

using Lemma 6.1 on $T = C_G(T) = C_G(T)^\circ$.

Let ρ be given as in (4.5). We know that $\mathrm{Tr}(\rho(\Gamma)) \subseteq K^r$ by definition of ρ, and that by Proposition 6.3 $\mathbb{F}_q$ is a subalgebra of K^r and isomorphic to $\mathbb{F}_p[N_\Gamma(V)]$ as a ring. We want to prove that $\mathrm{Tr}(\rho(\Lambda)) \subseteq \mathbb{F}_q$. The case $r = 1$ is immediate (see [24, p. 1143]), so assume $r = 2$.

By (6.6) and the fact that $\Gamma \cap U(\overline{K})$ acts trivially on $V \subseteq Z(U)$, $\Gamma \cap T(\overline{K})$ acts faithfully by conjugation on V and therefore maps isomorphically to a subgroup of $\mathbb{F}_q^* = (\mathbb{F}_p[N_\Gamma(V)])^*$. Thus $\Gamma \cap T(\overline{K})$ is cyclic, generated by some element γ that maps to some $(x, x^{p^f}) \in \mathbb{F}_q^*$ (recall that $\mathbb{F}_q$ is also a subalgebra of K^2). Fix such γ, x, f. Under Assumption A1–A2, one can prove that $2f + 1 = e$: this is not a surprise, since the cases with $\dim(Z(U)) = r = 2$ are those of type B_2, G_2 with characteristic $p = 2, 3$ respectively, and the automorphisms $(\cdot)^{p^f}$ with $q = p^{2f+1}$ provide the twists that yield Suzuki–Ree groups. The proof is contained in [24, Lemmas 10.5 to 10.9]; here we just sum up the computational details.

There exists some $u \in \Gamma^{\mathrm{run}}$ by Proposition 6.2; this fact and [24, Lem. 8.6] imply

$$|(\Gamma \cap U(\overline{K}))^{\mathrm{run}}| \geq \frac{1}{2^{k_r}} |\Gamma \cap U(\overline{K})|, \qquad\qquad k_r = \sum_{i=0}^{r-1} \frac{r!}{i!},$$

(so $k_2 = 4$) and then by [24, Prop. 8.7], (6.6), and Assumption A1 we conclude that

$$|\Gamma \cap T(\overline{K})| = [\Gamma \cap B(\overline{K}) : \Gamma \cap U(\overline{K})] \geq \frac{1}{2^{k_r}(2dD(\iota + 1))^{3(d+1)d^{d+1}}} |\Gamma|^{\frac{r}{d}} > 1.$$

The subgroup $T' \leq T$ yielding a scalar action on $U/[U, U]$ has dimension 1, so the bound above and Theorem 5.3 imply that there is some element $\gamma' \in \Gamma \cap T(\overline{K})$ whose action on $U/[U, U]$ is non-scalar. We also need the size of the subgroup of $\mathbb{F}_q^*$ to which $\Gamma \cap T(\overline{K})$ maps isomorphically to be at least $(q-1)/\varphi(d-1)^2$ (true by Assumption A1–A2 and [24, Thm. 6.6]), and we need the bound $(q-1)/\varphi(d-1)^2 \geq 2q^{3/4}$ (true by Assumption A1 and Proposition 6.3). The existence of u and γ', the minimality of V as defined inside the proof of Proposition 6.3, and the two bounds above force the equality $2f + 1 = e$. A case-by-case analysis for the possible groups G and the corresponding roots then gives $\mathrm{Tr}(\rho_s(\gamma^i)) = \mathrm{Tr}(\rho_\ell(\gamma^i))^{p^f}$ for all i, yielding the case $r = 2$ of the claim $\mathrm{Tr}(\rho(\Lambda)) \subseteq \mathbb{F}_q$.

It is time to define the objects that we look for. By Proposition 4.4, ρ is a representation over K of the group G in a space of dimension $\leq d$, and it has $\mathrm{mdeg}(\rho) \leq d$; the same can be said for ρ seen as a representation over $K^{\dim(V)}$. Let $M \leq \mathrm{GL}_d(K^{\dim(V)})$ be the ring of $K^{\dim(V)}$-linear transformations of the representation space of ρ. Call $X \subseteq G^{d^2}$ the set of $(g_i)_{i \leq d^2}$ for which $(\rho(g_i))_{i \leq d^2}$ does not span M as a $K^{\dim(V)}$-vector subspace of $(K^{\dim(V)})^{d^2}$. The set X is a proper subvariety of G^{d^2} (thus of dimension $< d^3$) defined as the set of zeros of the $(\dim(M) \times \dim(M))$-minors of the matrix whose rows are the $\rho(g_i)$. By Corollary 2.4(b) we have

$$\deg(X) \leq (\mathrm{mdeg}(\rho)\dim(M))^{d^3} \leq d^{3d^3},$$

so Corollary 5.4 gives $|\Gamma^{d^2} \cap X(\overline{K})| \leq d^2(2d^{3d^3+1})^{d^{d-1}}|\Gamma|^{d^2-\frac{1}{d}}$. On the other hand, if we call $\Omega \subseteq \Gamma^{d^2}$ the set of $(\gamma_i)_{i \leq d^2}$ for which

$$\left|\bigcap_{i=1}^{d^2} \gamma_i^{-1}\Lambda\right| \geq \frac{|\Gamma|}{2(2|\mathcal{W}|)^{d^2}},$$

then by (6.12) and [24, Lem. 10.10] we also have $|\Omega| \geq |\Gamma|^{d^2}/2(2|\mathcal{W}|)^{d^2}$. By (4.3), Assumption A1, and the two bounds above, there must be some $(\gamma_i)_{i \leq d^2} \in \Omega \setminus X$. Fix such a tuple, fix $\gamma_0 \in \bigcap_{i=1}^{d^2} \gamma_i^{-1}\Lambda$, and define

$$M_0 := \{m \in M : \mathrm{Tr}(\rho(\gamma_i\gamma_0)m) \in \mathbb{F}_q \quad (1 \leq i \leq d^2)\}.$$

This is an $\mathbb{F}_q$-vector space contained in M that by definition of X spans the whole M over $K^{\dim(V)}$. Now, [24, Lem. 10.12] shows that $|\Gamma \cap \rho^{-1}(M_0)(\overline{K})| \geq |\Gamma|/2(2|\mathcal{W}|)^{d^2}$. In turn, this implies that the left stabilizer $\Delta := \{\gamma \in \Gamma : \rho(\gamma)M_0 = M_0\}$ is large, in the sense that $[\Gamma : \Delta] \leq 4(2|\mathcal{W}|)^{d^2}$: if it were not, for some left Γ-translates M_0', M_0'' of M_0 we would have a large intersection $|\Gamma \cap \rho^{-1}(M_0' \cap M_0'')(\overline{K})|$, meaning at least $|\Gamma|/(4(2|\mathcal{W}|)^{d^2})^2$ in size (see the proof of [24, Lem. 10.13] for details), but this is not possible because every proper submodule $N \subsetneq M$ must have

$$\deg(\rho^{-1}(N)) \leq Dd^d \qquad |\Gamma \cap \rho^{-1}(N)(\overline{K})| \leq (2Dd^{d+1})^{d^d}|\Gamma|^{1-\frac{1}{d}} \tag{6.13}$$

by Lemma 2.1(b) and Theorem 5.3(c), contradicting the previous lower bound by (4.3) and Assumption A1. Let $\sigma : G \to \mathrm{GL}(M)$ be the representation defined by

$$\sigma(g)(m) = \rho(g)m\rho(g)^{-1}. \tag{6.14}$$

By definition of ρ, we have $\mathrm{mdeg}(\sigma) \leq d^2$ (as ρ comes from the adjoint representation $\mathrm{mdeg}(\sigma)$ can be shown to be linear in d, but the improvement is negligible), and σ is defined over K. We need to show that $M_0 = \sigma(\gamma)(M_0)$ for every $\gamma \in \Gamma$. We know that $\rho(\Delta) \subseteq M_0$, and the large index of Δ implies that

$$|\Gamma \cap \rho^{-1}(M_0 \cap \sigma(\gamma)(M_0))| \geq |\Delta \cap \gamma\Delta\gamma^{-1}| \geq \left(\frac{1}{4(2|\mathcal{W}|)^{d^2}}\right)^2 |\Gamma|,$$

while if $M_0 \neq \sigma(\gamma)(M_0)$ we would get an upper bound on $|\Gamma \cap \rho^{-1}(M_0 \cap \sigma(\gamma)(M_0))(\overline{K})|$ like in (6.13), contradicting again Assumption A1. Therefore $M_0 = \sigma(\gamma)(M_0)$, proving the result. □

From the Γ-invariant module M_0 we build the desired Steinberg endomorphism in Proposition 6.5 below. We do not assume any condition on $\dim(V)$, so that we may be able to apply the result again later in the paper.

Proposition 6.5. *Let $G, n, r, K, d, D, \iota, \Gamma$ be as in Assumption A1–A2. Since Proposition 6.3 holds, let $p, q = p^e, V = V^{(G)}$ be as given therein. Assume that there exist σ, M, M_0 satisfying the conclusions of Proposition 6.4 (unlike in that result, we do not necessarily assume that $\dim(V) = r$).*

Then, there is a Steinberg endomorphism $F : \mathrm{Aut}_K(M) \to \mathrm{Aut}_K(M)$ (where M is seen as a module over K, instead of over $K^{\dim(V)}$) such that $[G^F, G^F] \leq \Gamma \leq G^F$ with $[G^F, G^F]$ simple.

Proof. We follow [24, §9], skipping the theoretical details and focusing on explicit bounds.

By construction M is a $K^{\dim(V)}$-module, where $\dim(V) \in \{1, 2\}$ by (4.4). As σ in nontrivial and G is almost simple adjoint, $\mathrm{Ad}_{\sigma(G)} \circ \sigma$ is a totally inseparable isogeny on its image, hence injective (as follows from its definition, see [28, p. 448]), and since all maps are defined over K then σ must be injective on K-points.

As M is a module and $\sigma(G) \leq \mathrm{GL}(M)$ (over K), we may define our desired Steinberg endomorphism $F : \mathrm{GL}(M) \to \mathrm{GL}(M)$ as a linear transformation of the K-vector space $\mathrm{Mat}(M)$. Therefore, as F amounts to a change of basis, we shall have $\mathrm{mdeg}(F) = 1$ when F is seen as an isomorphism of the variety $\mathrm{GL}(M)$ onto itself.

The details of the choice of F are contained in [24, p. 1141]. If $\dim(V) = 1$, we take F to be the Frobenius map $F : \mathrm{Aut}_K(M) \to \mathrm{Aut}_K(M)$ with respect to $\mathbb{F}_q \subseteq K$. If $\dim(V) = 2$, we can write $M = M_\ell \oplus M_s$ and $\mathrm{Aut}_{K^2}(M) = \mathrm{Aut}_K(M_\ell) \times \mathrm{Aut}_K(M_s)$ and take isogenies on the two components so that their product is a map F whose square is the Frobenius map with respect to $\mathbb{F}_q \subseteq K^2$. By construction, $\sigma^{-1}(F(\sigma(G))) \leq G$, and the conclusion of Proposition 6.4 gives $F(\sigma(\gamma)) = \sigma(\gamma)$ for all $\gamma \in \Gamma$, meaning that $\Gamma \subseteq \sigma^{-1}(F(\sigma(G)))$.

The bounds $\mathrm{mdeg}(\sigma) \leq d^2$ and $\mathrm{mdeg}(F) = 1$ give by Lemma 2.1(a)–(b)

$$\deg(\sigma^{-1}(F(\sigma(G)))) \leq D^2 \mathrm{mdeg}(\sigma)^{2d} \mathrm{mdeg}(F)^d \leq D^2 d^{4d}.$$

Hence, Assumption A2 forces $\sigma^{-1}(F(\sigma(G))) = G$. The map F becomes naturally an isogeny $F : G \to G$ (see [24, Lem. 9.4], which uses [28, Thm. 1.7]) for which $F^{\dim(V)}$ is the Frobenius map with respect to $\mathbb{F}_q$. On one hand we have $\Gamma \subseteq G^F$, using the injectivity of σ on K-points. On the other hand, Assumption A1 and [24, Thm. 3.4] imply that $[G^F, G^F]$ is simple with index $\leq r + 1$ inside G^F by (4.1), and that

$$\frac{|\Gamma|}{2\varphi(d-1)^{\frac{d}{\dim(V)}}} \leq (q^{\frac{1}{\dim(V)}} - 1)^d \leq |G^F| \leq q^{\frac{d}{\dim(V)}} \leq \varphi(2)^{\frac{d}{\dim(V)}} |\Gamma| \tag{6.15}$$

by Proposition 6.3. Let $H_1 := \Gamma \cap [G^F, G^F]$, and let H_2 be the normal core of H_1 inside $[G^F, G^F]$. The upper bound in (6.15) gives

$$\begin{aligned} |[G^F, G^F]/H_2| &\leq [[G^F, G^F] : H_1]! \leq [G^F : \Gamma]! \leq \left(\varphi(2)^{\frac{d}{\dim(V)}}\right)! \\ &\leq \left((2dD(\iota+1))^{2(d+1)d^3}\right)! \leq (2dD\iota)^{(2dD\iota)^{5d^4}}, \end{aligned} \tag{6.16}$$

and the lower bound of (6.15) gives

$$[G^F, G^F] \geq \frac{|G^F|}{r+1} \geq \frac{|\Gamma|}{2(r+1)(2dD(\iota+1))^{d^{d+1}}}. \tag{6.17}$$

Putting together (6.16), (6.17), Assumption A1, and the simplicity of $[G^F, G^F]$, we are forced to have $H_2 = [G^F, G^F]$, concluding that $\Gamma \supseteq [G^F, G^F]$. □

Next we prove the existence of a Γ-invariant module M_0 as in Proposition 6.4, this time without the assumption $\dim(V) = r$. The construction relies on finding a smaller subgroup $H \leq G$ for which $\dim(V^{(H)}) = \mathrm{rk}(H)$, and use the field and the representation resulting from H to build the representation of G. Passing from G to H worsens the conditions on Γ, hence we need to impose the stronger Assumption B1–B2 on $\Gamma \leq G(\overline{K})$ to ensure that a second group $\Delta \leq H(\overline{K})$ related to Γ satisfies Assumption A1–A2.

Proposition 6.6. *Let $G, n, r, K, d, D, \iota, \Gamma$ be as in Assumption B1–B2. Since Proposition 6.3 holds, let $p, q = p^e, V = V^{(G)}$ be as given therein.*

Then, there is a $K^{\dim(V)}$-module M, there is a nontrivial representation $\sigma : G \to \mathrm{GL}(M)$ with $\mathrm{mdeg}(\sigma) \leq d^2$ (σ defined over K), and there is an $\mathbb{F}_q$-submodule M_0 of M with $M_0 \otimes_{\mathbb{F}_q} K^{\dim(V)} \simeq M$ and $\sigma(\gamma)(M_0) = M_0$ for every $\gamma \in \Gamma$.

Proof. If $\dim(V) = r$ we are done by Proposition 6.4, so we may assume $\dim(V) < r$. We follow [24, §11]. The first step is to find a suitable $H \leq G$ for which $\dim(V^{(H)}) = \mathrm{rk}(H)$ and apply Propositions 6.4–6.5 to it.

By [21, §§28.3–28.5], we can decompose G as a disjoint union

$$G = \coprod_{w \in \mathcal{W}} BwB = B\dot{w}B \sqcup \coprod_{w \neq \dot{w}} \overline{BwB},$$

where $B\dot{w}B$ is open and dense. If X is the union over $w \neq \dot{w}$ on the right-hand side, by Lemmas 2.1(a)–3.1(b) and (4.3) we have

$$\deg(X) \leq (|\mathcal{W}| - 1)\deg(B)^2 2^d \leq ((2r)^r - 1)D^2 2^d. \tag{6.18}$$

Hence, by Assumption B1–B2 and Theorem 5.3(c) applied to X, there is some $\gamma \in \Gamma \cap (B\dot{w}B)(\overline{K})$. From now on, fix such a γ. Call $H_{(\gamma)}$ the algebraic group generated by V and $\gamma V \gamma^{-1}$: $H_{(\gamma)}$ is connected almost simple of type A_1, B_2, G_2, and is in fact the product of root subgroups normalized by T [24, Prop. 11.1(a)–(b)]. Thus, we can write $H_{(\gamma)}$ as the image of the map from $T \times T$ to G given by $f(t_1, t_2) = t_1 g_1 t_1^{-1} t_2 g_2 t_2^{-1}$ for some appropriate g_1, g_2, which by Lemma 2.1(a) gives $\deg(H_{(\gamma)}) \leq D^2(2\iota + 2)^2$. Write also π for the adjoint representation $\mathrm{Ad}_{H_{(\gamma)}}$ and H for its image $H^{\mathrm{ad}}_{(\gamma)}$: the map $\pi : H_{(\gamma)} \to H$ has $\mathrm{mdeg}(\pi) \leq \iota + 1$ by (4.2), and every fibre has size ≤ 2 (checking the types case by case). Most notably, H is a connected almost simple adjoint group of a type for which $\dim(V^{(H)}) = \mathrm{rk}(H)$, so Propositions 6.4–6.5 apply to H, provided that we also have a suitable finite group inside $H(\overline{K})$.

Define $\Delta := \pi(\Gamma \cap H_{(\gamma)}(\overline{K}))$. By the bound on fibre sizes and Proposition 6.3, we have

$$|\Delta| \geq \frac{1}{2}|\Gamma \cap H_{(\gamma)}(\overline{K})| \geq \frac{1}{2}|\Gamma \cap V(\overline{K})| \geq \frac{1}{2(2dD(\iota+1))^{d^{d+1}}}|\Gamma|^{\frac{\dim(V)}{d}},$$

and if Γ satisfies Assumption B1 then Δ must satisfy Assumption A1. Now, let L be any algebraic subgroup of H that is proper (so $\dim(L) < \dim(H)$ automatically) and that satisfies

$$\deg(L) \leq (2\dim(H)\deg(H))^{4\dim(H)} \leq 2^{14000}, \tag{6.19}$$

where we used $\dim(H) \leq 10$ and $\deg(H) \leq 2^{7^3}$ by Remark 4.3(b). Since L is proper, by definition of $H_{(\gamma)}$ we cannot have $V \subseteq \pi^{-1}(L)$ and $\gamma V\gamma^{-1} \subseteq \pi^{-1}(L)$ at the same time. Suppose that $V \not\subseteq \pi^{-1}(L)$. Lemma 2.1(b) implies that $\deg(V \cap \pi^{-1}(L)) \leq 2^{14000}D(\iota+1)^2$. Combining Assumption B1 with Theorem 5.3(c) and Proposition 6.3,

$$\begin{aligned}|\Gamma \cap (V \cap \pi^{-1}(L))(\overline{K})| &\leq (2d \cdot 2^{14000}D(\iota+1)^2)^{d^2}|\Gamma|^{\frac{\dim(V)-1}{d}}\\ &< \frac{|\Gamma|^{\frac{\dim(V)}{d}}}{(2dD(\iota+1))^{d^{d+1}}} \leq |\Gamma \cap V(\overline{K})|,\end{aligned}$$

implying that $\Gamma \cap V(\overline{K}) \not\subseteq \pi^{-1}(L)(\overline{K})$. Analogously, $\gamma V\gamma^{-1} \not\subseteq \pi^{-1}(L)$ gives $\Gamma \cap \gamma V\gamma^{-1}(\overline{K}) \not\subseteq \pi^{-1}(L)(\overline{K})$. In either case, we obtain $\Delta \not\subseteq L(\overline{K})$, which means that Δ satisfies Assumption A2. Hence, we can apply Proposition 6.5 to Δ and H, we have $[H^F, H^F] \leq \Delta \leq H^F$ for some Steinberg endomorphism with $[H^F, H^F]$ simple.

Finally, by Assumption B1 and (6.11) there is some $v \in (\Gamma \cap V(\overline{K})) \setminus \{e\}$, which we fix. If γ runs through the elements of $\Gamma \cap (B\dot{w}B)(\overline{K})$ the group $H_{(\gamma)}$ may change, but all such $H_{(\gamma)}$ are conjugate by [24, Prop. 11.1(c)]; therefore H, F do not depend on the choice of γ. The proof of [24, Prop. 11.7] then shows that $\mathbb{F}_q$ and the field underlying the map F are the same field, and that, for ρ as in (4.5), $\mathrm{Tr}(\rho(v\gamma v\gamma^{-1})) \in \mathbb{F}_q$ for all $\gamma \in \Gamma \cap (B\dot{w}B)(\overline{K})$ (and v fixed).

Now we use the result above about traces to define the appropriate σ, M, M_0. Let $\bar{M}$ be the ring of $K^{\dim(V)}$-linear transformations of the representation space of ρ, and let $\bar{M}'$ be the smallest G-invariant $K^{\dim(V)}$-submodule of $\bar{M}$ containing $\rho(v)$. Just as in the proof of Proposition 6.4, there is a representation $\bar{\sigma} : G \to \mathrm{GL}(\bar{M})$ given as in (6.14), which in particular is defined over K and satisfies $\mathrm{mdeg}(\bar{\sigma}) \leq d^2$. Its restriction $\bar{\sigma}' : G \to \mathrm{GL}(\bar{M}')$ is again defined over K, because $\bar{\sigma}$ and $\bar{M}'$ are, and has $\mathrm{mdeg}(\bar{\sigma}') \leq d^2$. Since $\bar{\sigma}(g)$ is conjugation by $\rho(g)$ it preserves traces, so we may take the quotient $M := \bar{M}'/(\bar{M}' \cap \bar{M}'^{\perp})$ (the orthogonal complement is taken with respect to the trace form) and still obtain a representation $\sigma : G \to \mathrm{GL}(M)$ with $\mathrm{mdeg}(\sigma) \leq d^2$. Moreover σ is nontrivial by [24, Prop. 11.5, (11.9)], and it is defined over K because $\bar{M}'^{\perp}$ is. Finally, let $m_0 \in M$ be the element corresponding to $\rho(v) \in \bar{M}'$, and call M_0 the $\mathbb{F}_q$-submodule generated by the orbit $O_\Gamma(m_0)$. By construction, $\sigma(\gamma)(M_0) = M_0$ for any $\gamma \in \Gamma$.

It remains to prove that $M_0 \otimes_{\mathbb{F}_q} K^{\dim(V)} \simeq M$. See the proof of [24, Lem. 11.12]; we only present the computational details. Let N be a proper $K^{\dim(V)}$-submodule of M, and define

$$X(N) := \{g \in G : \sigma(g)(m_0) \in N\}$$

as in [24, (11.13)]. The variety $X(N)$ is proper inside G, and it is the preimage of N through the map $\sigma(\cdot)(m_0)$, so

$$\deg(X(N)) \leq \deg(G)\deg(N)\,\mathrm{mdeg}(\sigma)^{\dim(N)} \leq Dd^{2d}$$

by Lemma 2.1(b). If the natural map $M_0 \otimes_{\mathbb{F}_q} K^{\dim(V)} \to M$ is not surjective then $\Gamma \subseteq X(N)(\overline{K})$ for some N, and if it is not injective then

$$\Gamma \subseteq X(N)(\overline{K}) \cup \bigcup_{i=1}^{\ell} \bigcup_{w \neq \dot{w}} (\gamma_i BwB)(\overline{K}),$$

where ℓ is at most one more than the dimension of M as a $K^{\dim(V)}$-vector space. Recalling the degree bound of (6.18), in both cases we obtain that Γ is contained in a variety of degree $\leq 2^{d+r+2}r^r D^2 d^{2d} \leq (2dDr)^{2d+r+1}$; applying Lemma 3.3, we contradict either Assumption B1 or B2, so the map must be an isomorphism, and we are done. □

Combining Propositions 6.5 and 6.6, we reach the finite simple group $[G^F, G^F]$ that we are looking for. Below we write a self-contained statement, but the conditions on G and Γ coincide with Assumption B1–B2.

Theorem 6.7. *Let $G \leq \mathrm{GL}_n$ be a connected almost simple adjoint group of rank r defined over K, with $d = \dim(G)$, $D = \deg(G)$, and $\iota = \mathrm{mdeg}(^{-1})$. Let $\Gamma \leq G(\overline{K})$ be finite. Assume the following:*

(a) $|\Gamma| > (2dDrn\iota)^{(2dDr\iota)^{11d^4}}$;

(b) $\Gamma \not\leq H(\overline{K})$ for any subgroup $H \lneq G$ with $\dim(H) < d$ and $\deg(H) \leq (2dDr)^{4d^2}$.

Then $\mathrm{char}(K) = p > 0$, there is some $q = p^e$ such that $\mathbb{F}_q$ is an $\mathbb{F}_p$-subalgebra of either K or K^2, and there is a Steinberg endomorphism $F : G \to G$ such that either F or F^2 is the Frobenius map with respect to $\mathbb{F}_q$, with

$$[G^F, G^F] \leq \Gamma \leq G^F$$

and with $[G^F, G^F]$ simple.

Proof. Since the conditions on G and Γ are exactly the ones in Assumption B1–B2, we obtain Proposition 6.6. Then we apply Proposition 6.5, whose hypothesis is the weaker Assumption A1–A2, and the result follows. □

7. Proof of the main theorem

The main result of the previous section is that, for G almost simple adjoint and Γ sufficiently general, we have $[G^F, G^F] \leq \Gamma \leq G^F$ for some Steinberg endomorphism such that $[G^F, G^F]$ is a finite simple group of Lie type. This assertion lies at the the core of (b) in Theorem 1.2. In the current section we build the rest of the theorem around this initial core.

The proof relies mainly on a descent process: if Γ is not sufficiently general (meaning that it violates either (a) or (b) in Theorem 6.7), then it is trapped in a smaller subgroup, for some definition of "smaller"; repeating the procedure enough times, eventually the subgroup becomes zero-dimensional, giving a bound on $|\Gamma|$ in terms of n only. Being connected almost simple adjoint is not preserved at every step though, so we need a few more steps in between. If G is an algebraic group, we first pass to a small-index connected subgroup G° (and the effect on Γ is absorbed by (a) in the main theorem), then we get to a reductive group via taking quotient by $R_u(G)$ (which is absorbed by (d)), then to a semisimple group via quotient by $Z(G)$ (which is absorbed by (c)). A semisimple group is a product of almost simple pieces, which gives either a product of finite simple groups as in (b), up to a small index contributing to (a), or a descent as before. The potential case of $\dim(G)=0$ and bounded $|\Gamma|$ is again dealt with by (a).

Let us specify what parameter we use to track the descent. Let $\{G_i\}_{i\in\mathcal{I}(G)}$ be the collection of the connected almost simple adjoint factors of the semisimple group $(G^{\mathrm{o}}/R_u(G^{\mathrm{o}}))^{\mathrm{ad}}$; the set $\mathcal{I}(G)$ could be empty, and the G_i could be defined over $\overline{K}$. Define

$$d_{\mathrm{ad}}(G) := \sum_{i\in\mathcal{I}(G)} \dim(G_i) = \dim((G^{\mathrm{o}}/R_u(G^{\mathrm{o}}))^{\mathrm{ad}}).$$

Trivially $d_{\mathrm{ad}}(G)\leq\dim(G)$, and if G is almost simple then equality holds. It is also easy to show that if $H\leq G$ then $d_{\mathrm{ad}}(H)\leq d_{\mathrm{ad}}(G)$. We set up an induction using $d_{\mathrm{ad}}(G)$.

Lemma 7.1. *Let $G\leq\mathrm{GL}_n$ be an algebraic group of rank r defined over K, with $d=\dim(G)$, $D=\deg(G)$, and $\iota=\mathrm{mdeg}(^{-1})$. Let $\Gamma\leq G(\overline{K})$ be finite.*

Assume $\mathcal{I}=\mathcal{I}(G)\neq\emptyset$ (i.e. $d_{\mathrm{ad}}(G)\geq 1$). Then at least one of the following holds:

(a) $\Gamma\leq H(\overline{K})$ for some subgroup $H\lneq G$ with $d_{\mathrm{ad}}(H)<d_{\mathrm{ad}}(G)$ and

$$\deg(H)\leq(2dn\iota)^{(2dn\iota)^{2^{20}d^4r^2+2dn^2}}D^{d+1};$$

(b) $\mathrm{char}(K)=p>0$, and for every $i\in\mathcal{I}$ there is a Steinberg endomorphism $F_i:G_i\to G_i$ with

$$[G_i^{F_i},G_i^{F_i}]\leq\pi_i(\Gamma\cap G^{\mathrm{o}}(\overline{K}))\leq G_i^{F_i}$$

and with $[G_i^{F_i},G_i^{F_i}]$ simple.

Proof. Since GL_1 is abelian we may assume $n\geq 2$, or else $\mathcal{I}=\emptyset$. Take $Y=Y(G^{\mathrm{o}})$ and the corresponding $\hat{G},\hat{Y},\lambda,\hat{\beta},m$ as in Lemma 4.1. Since $\hat{G}(\overline{K})\simeq G^{\mathrm{o}}(\overline{K})$, we may consider $\Gamma\cap G^{\mathrm{o}}(\overline{K})\leq\hat{G}(\overline{K})$. The quotient $\hat{G}/\hat{Y}$ (possibly defined over $\overline{K}$) is connected adjoint, so it is isomorphic to the direct product of the G_i: the isomorphism can be taken to be the adjoint representation $\mathrm{Ad}_{\hat{G}/\hat{Y}}$, whose image sits in $\mathrm{GL}(\mathfrak{x})$ (where $\mathfrak{x}$ is the Lie algebra of $\hat{G}/\hat{Y}$) and which has maximum degree $\leq m+1$ by (4.2). Then, up to a change of basis of $\mathfrak{x}$ (which does not affect the degree), the projection to any almost simple factor has maximum degree ≤ 1.

For every $i\in\mathcal{I}$, call $\pi_i:\hat{G}\to G_i$ the natural epimorphism given by composing the quotient map $\hat{\beta}$, the adjoint representation $\mathrm{Ad}_{\hat{G}/\hat{Y}}$, the change of basis of $\mathfrak{x}$, and the projection to G_i. By the facts above,

$$\mathrm{mdeg}(\pi_i)\leq 2(n^2+(\iota+1)d^d)^{n^2}(\iota+1)d^d\cdot\left(2^{2(n^2+(\iota+1)d^d)^{n^2}}+1\right)\leq(2dn\iota)^{(2dn\iota)^{2dn^2}}. \tag{7.1}$$

Furthermore, by Proposition 4.2 we have the following bounds on the parameters of the G_i depending only on $d=\dim(G)=\dim(\hat{G})$ and $r=\mathrm{rk}(G)=\mathrm{rk}(\hat{G})$:

$$\mathrm{rk}(G_i)\leq r\leq d,\qquad\dim(G_i)\leq d,\qquad\iota|_{G_i}\leq n|_{G_i}\leq d,\qquad\deg(G_i)\leq(2r)^{2^{16}r^2}. \tag{7.2}$$

For each $i\in\mathcal{I}$, apply Theorem 6.7 to the finite group $\pi_i(\Gamma\cap G^{\mathrm{o}}(\overline{K}))$ contained in $G_i(\overline{K})$. If conditions (a)–(b) are both satisfied for all i, we obtain case (b) of the present result. Assume then that either (a) or (b) in Theorem 6.7 is violated for some i. As a matter of fact, we may rewrite case (a) to look like an instance of case (b): to do so, interpret $\pi_i(\Gamma\cap G^{\mathrm{o}}(\overline{K}))$ as the zero-dimensional algebraic group H defined as the union $\bigcup_{\gamma\in\pi_i(\Gamma\cap G^{\mathrm{o}}(\overline{K}))}\{\gamma\}$. Hence, in both cases, $\Gamma\cap G^{\mathrm{o}}(\overline{K})\leq\pi_i^{-1}(H)(\overline{K})$ for some algebraic subgroup $H\leq G_i$ such that either $\dim(H)<\dim(G_i)$ and $\deg(H)\leq(2d)^{2^{19}d^2r^2}$, or $\dim(H)=0$ and $\deg(H)\leq(2d)^{(2d)^{2^{20}d^4r^2}}$.

Set $P:=\overline{\lambda(\pi_i^{-1}(H))}$, which is an algebraic group. As G^{o} is irreducible we have $P\lneq G^{\mathrm{o}}\leq G$ and $\dim(P)<d$, while $d_{\mathrm{ad}}(H)\leq\dim(H)<\dim(G_i)=d_{\mathrm{ad}}(G_i)$ implies that $d_{\mathrm{ad}}(P)<d_{\mathrm{ad}}(G)$. Lemma 2.1(a)–(b) gives $\deg(P)\leq D\deg(H)\mathrm{mdeg}(\pi_i)^d$. Let $L=\bigcap_{\gamma\in\Gamma}\gamma P\gamma^{-1}$. Again $d_{\mathrm{ad}}(L)\leq d_{\mathrm{ad}}(P)<d_{\mathrm{ad}}(G)$, and Corollary 2.4(b) yields $\deg(L)\leq\deg(P)^d$.

Since Γ is finite, $\Gamma L=\bigcup_{\gamma\in\Gamma}\gamma L$ is a variety. Moreover, the normalizer of L in G contains Γ, so ΓL is an algebraic subgroup of G containing Γ. We have $(\Gamma L)^{\mathrm{o}}=L^{\mathrm{o}}$, giving $d_{\mathrm{ad}}(\Gamma L)=d_{\mathrm{ad}}(L)<d_{\mathrm{ad}}(G)$. It remains to bound $\deg(\Gamma L)$. By definition, $\Gamma\cap L(\overline{K})$ is the normal core of $\Gamma\cap P(\overline{K})=\Gamma\cap G^{\mathrm{o}}(\overline{K})$ inside Γ, but $\Gamma\cap G^{\mathrm{o}}(\overline{K})\trianglelefteq\Gamma$ already, therefore

$$|\Gamma L(\overline{K})/L(\overline{K})|=|\Gamma/(\Gamma\cap L(\overline{K}))|=|\Gamma/(\Gamma\cap G^{\mathrm{o}}(\overline{K}))|\leq|G/G^{\mathrm{o}}|\leq D.$$

Hence, using (7.1),

$$\deg(\Gamma L)\leq D\deg(L)\leq D^{d+1}\deg(H)^d\mathrm{mdeg}(\pi_i)^{d^2}\leq(2dn\iota)^{(2dn\iota)^{2^{20}d^4r^2+2dn^2}}D^{d+1},$$

so the algebraic subgroup ΓL satisfies the conditions in (a). □

With the induction step of Lemma 7.1 at hand, we can prove the main theorem.

Proof of Theorem 1.2. If Γ is abelian then the result holds by taking $\Gamma_2=\Gamma$ and Γ_3 its Sylow p-subgroup. Thus we may assume Γ non-abelian, and thus $n\geq 2$.

Start with $G=\mathrm{GL}_n$, and apply Lemma 7.1 to it. If case (a) of the lemma holds, repeat the process with the subgroup of G found in this way, and repeat this step until either $\mathcal{I}(G)=\emptyset$ or we reach case (b). Since $d_{\mathrm{ad}}(G)$ strictly decreases at each step, the process ends in at most d steps. By the bound on $\deg(H)$ inside Lemma 7.1(a) and the natural bounds on $\dim(\mathrm{GL}_n),\deg(\mathrm{GL}_n),\mathrm{rk}(\mathrm{GL}_n),\mathrm{mdeg}(^{-1})$ in terms of n, at the last step we have $\deg(G)\leq n^{n^{(2^{23}-1)n^{10}}}$.

Let $Y=Y(G^{\mathrm{o}})$ be as in Lemma 4.1. Since $Y\,\underline{\blacktriangleleft}\,G^{\mathrm{o}}\,\underline{\blacktriangleleft}\,G$ and $R_u(G^{\mathrm{o}})\,\underline{\blacktriangleleft}\,G^{\mathrm{o}}\,\underline{\blacktriangleleft}\,G$, we have $Y\,\underline{\blacktriangleleft}\,G$ and $R_u(G^{\mathrm{o}})\,\underline{\blacktriangleleft}\,G$. Call $\Gamma_2:=\Gamma\cap Y(\overline{K})$ and $\Gamma_3:=\Gamma\cap R_u(G^{\mathrm{o}})(\overline{K})$: we have $\Gamma_3\trianglelefteq\Gamma_2\trianglelefteq\Gamma$ and $\Gamma_3\trianglelefteq\Gamma$.

By construction, $Y/R_u(G^{\mathrm{o}})$ is the centre of $G^{\mathrm{o}}/R_u(G^{\mathrm{o}})$; in particular, its finite subgroup Γ_2/Γ_3 is contained in a torus of the reductive group $G^{\mathrm{o}}/R_u(G^{\mathrm{o}})$, thus it is abelian of order not divisible by $\mathrm{char}(K)$. Since $R_u(G^{\mathrm{o}})$ is unipotent, there is a central normal series whose quotients are isomorphic to algebraic subgroups of $\mathbb{G}_a$ [26, Prop. 14.21]; hence, since $\Gamma_3\leq R_u(G^{\mathrm{o}})$ is finite, either it is trivial (if $\mathrm{char}(K)=0$) or it is a p-group (if $\mathrm{char}(K)=p>0$).

It remains to deal with Γ/Γ_2. If $\mathcal{I}(G)=\emptyset$, by definition $G^{\mathrm{o}}/R_u(G^{\mathrm{o}})$ is equal to its own centre, so we set $\Gamma_1:=\Gamma_2$ and the quotient Γ/Γ_2 has size $\leq|G/G^{\mathrm{o}}|\leq\deg(G)$. Assume from now on that we have fallen into case (b) of Lemma 7.1. Therefore, $\mathrm{char}(K)=p>0$ and $(\Gamma\cap G^{\mathrm{o}}(\overline{K}))/\Gamma_2$ is a subgroup of the nonempty product $R=\prod_{i\in\mathcal{I}}\pi_i(\Gamma\cap G^{\mathrm{o}}(\overline{K}))$, for which we know that

$$\prod_{i\in\mathcal{I}}[G_i^{F_i},G_i^{F_i}]\leq R\leq\prod_{i\in\mathcal{I}}G_i^{F_i}$$

and that each factor $[G_i^{F_i},G_i^{F_i}]$ is a finite simple group of Lie type of characteristic p. The commutator $[R,R]$ must then be equal to the product of the $[G_i^{F_i},G_i^{F_i}]$ and, by a folklore consequence of Goursat's lemma (see [1, Prop. 3.3]), if we call $\Gamma_1 := [\Gamma\cap G^{\mathrm{o}}(\overline{K}),\Gamma\cap G^{\mathrm{o}}(\overline{K})]\Gamma_2$ then Γ_1/Γ_2 must be isomorphic to the product of some of the $[G_i^{F_i},G_i^{F_i}]$. By construction $\Gamma_1\trianglelefteq\Gamma$, since $\Gamma\cap G^{\mathrm{o}}$, its commutator, and Γ_2 are all normal in Γ. Finally,

$$|\Gamma/\Gamma_1|\leq|G/G^{\mathrm{o}}|\prod_{i\in\mathcal{I}}|G_i^{F_i}/[G_i^{F_i},G_i^{F_i}]|\leq\deg(G)(r+1)^d\leq n^{n^{2^{23}n^{10}}}$$

by (4.1). □

Acknowledgements

The second author was funded by a Young Researcher Fellowship from the HUN-REN Alfréd Rényi Institute of Mathematics, and by the Leibniz Fellowship 2405p from the Mathematisches Forschungsinstitut Oberwolfach (MFO).

The authors would like to thank the MFO for providing a wonderful environment for the development of this article.

References

[1] L. Babai and A. Seress. On the diameter of permutation groups. *European J. Combin.*, 13(4):231–243, 1992. 38

[2] J. Bajpai, D. Dona, and H. A. Helfgott. Growth estimates and diameter bounds for classical Chevalley groups. `arXiv:2110.02942`, 2021. 1, 2, 5, 10, 16

[3] J. Bajpai, D. Dona, and H. A. Helfgott. New dimensional estimates for subvarieties of linear algebraic groups. *Vietnam J. Math.*, 52(2):479–518, 2024. 1, 2, 4, 5, 10, 16, 19, 20, 21, 22, 23

[4] A. Borel. *Linear Algebraic Groups*, volume 126 of *Graduate Texts in Mathematics*. Springer, New York (USA), second enlarged edition, 1991. 7, 8, 9, 10, 11, 12, 14, 17, 19, 25

[5] M. Brandt, J. Bruce, T. Brysiewicz, R. Krone, and E. Robeva. The degree of $\mathrm{SO}(n,\mathbb{C})$. In G. G. Smith and B. Sturmfels, editors, *Combinatorial algebraic geometry*, volume 80 of *Fields Institute Communications*, pages 229–246. Springer, New York (USA), 2017. 17

[6] E. Breuillard. An exposition of Jordan's original proof of his theorem on finite subgroups of $\mathrm{GL}_n(\mathbb{C})$. *Model Theory*, 2:429–447, 2023. 1

[7] E. Breuillard, B. Green, and T. Tao. Approximate subgroups of linear groups. *Geom. Funct. Anal.*, 21(4):774–819, 2011. 14

[8] R. W. Carter. *Finite groups of Lie type: Conjugacy classes and complex characters*. John Wiley & Sons, Chichester (UK), reprinted edition, 1993. 14, 16, 17, 18

[9] M. J. Collins. On Jordan's theorem for complex linear groups. *J. Group Theory*, 10(4):411–423, 2007. 1

[10] M. J. Collins. Modular analogues of Jordan's theorem for finite linear groups. *J. Reine Angew. Math.*, 624:143–171, 2008. 2

[11] J. H. Conway, R. T. Curtis, S. P. Norton, R. A. Parker, and R. A. Wilson. *ATLAS of Finite Groups: Maximal Subgroups and Ordinary Characters for Simple Groups*. Clarendon Press, Oxford (UK), 1985. 17

[12] V. I. Danilov and V. N. Shokurov. *Algebraic Geometry I: Algebraic Curves, Algebraic Manifolds and Schemes*, volume 23 of *Encyclopaedia of mathematical sciences*. Springer-Verlag, Berlin, 1998. 4

[13] D. Dona. A sum-bracket theorem for simple Lie algebras. *J. Algebra*, 631:658–694, 2023. 17

[14] A. Eskin, S. Mozes, and H. Oh. On uniform exponential growth for linear groups. *Invent. Math.*, 160(1):1–30, 2005. 10

[15] W. Fulton. *Intersection theory*, volume 2 of *Ergebnisse der Mathematik und ihrer Grenzgebiete : a series of modern surveys in mathematics. Folge 3*. Springer, Berlin, 1984. 4

[16] W. Fulton and J. Harris. *Representation theory: a first course*, volume 129 of *Graduate Texts in Mathematics*. Springer, New York (USA), 2004. 17

[17] S. Garibaldi. What is... a linear algebraic group? *Notices Amer. Math. Soc.*, 57(9):1125–1126, 2010. 14

[18] J. Heintz. Definability and fast quantifier elimination in algebraically closed fields. *Theoret. Comput. Sci.*, 24(3):239–277, 1983. 15

[19] R. B. Howlett, L. J. Rylands, and D. E. Taylor. Matrix generators for exceptional groups of Lie type. *J. Symbolic Comput.*, 31(4):429–445, 2001. 17

[20] J. E. Humphreys. *Conjugacy classes in semisimple algebraic groups*, volume 43 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 1995. 8, 23, 26, 27, 28

[21] J. E. Humphreys. *Linear algebraic groups*, volume 21 of *Graduate Texts in Mathematics*. Springer-Verlag, New York (USA), fourth edition, 1995. 7, 9, 14, 19, 27, 33

[22] I. M. Isaacs. *Character theory of finite groups*, volume 69 of *Pure and Applied Mathematics*. Academic Press, New York (USA), 1976. 1, 3

[23] C. Jordan. Mémoire sur les équations différentielles linéaires à intégrale algébrique. *J. Reine Angew. Math.*, 84:89–215, 1878. In French. 1

[24] M. J. Larsen and R. Pink. Finite subgroups of algebraic groups. *J. Amer. Math. Soc.*, 24(4):1105–1158, 2011. 1, 2, 5, 14, 18, 19, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34

[25] G. Malle and D. Testerman. *Linear algebraic groups and finite groups of Lie type*, volume 133 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge (UK), 2011. 14, 15, 16, 18, 19

[26] J. S. Milne. *Algebraic groups*, volume 170 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2017. 7, 14, 15, 37

[27] D. Mumford. *The red book of varieties and schemes*, volume 1358 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, expanded edition, 1999. Includes the Michigan lectures (1974) on curves and their Jacobians, With contributions by Enrico Arbarello. 4, 5, 23

[28] R. Pink. Compact subgroups of linear algebraic groups. *J. Algebra*, 206(2):438–504, 1998. 13, 18, 19, 32

[29] T. A. Springer. *Linear algebraic groups*. Modern Birkhäuser Classics. Birkhäuser Boston, Boston (USA), reprint of the second edition, 2008. 14

[30] W. Vogel. *Lectures on results on Bezout's theorem. Notes by D. P. Patil*, volume 74 of *Lect. Math. Phys., Math., Tata Inst. Fundam. Res.* Springer, Berlin; Tata Inst. of Fundamental Research, Bombay, 1984. 4

[31] R. A. Wilson. *The Finite Simple Groups*, volume 251 of *Graduate Texts in Mathematics*. Springer, London (UK), 2009. 17

Department of Mathematics, Kiel University, 24118 Kiel, Germany
Email address: jitendra@math.uni-kiel.de

Hun-Ren Alfréd Rényi Institute of Mathematics, 1053 Budapest, Hungary
Email address: dona@renyi.hu