JOURNAL OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 24, Number 4, October 2011, Pages 1105–1158
S 0894-0347(2011)00695-4
Article electronically published on April 28, 2011

FINITE SUBGROUPS OF ALGEBRAIC GROUPS

MICHAEL J. LARSEN AND RICHARD PINK

Contents

0. Introduction

Consider a finite subgroup Γ of GL_n over an arbitrary field. What can be said about Γ without further hypothesis? Jordan's theorem [20, p. 114] provides an answer in characteristic zero:

Theorem 0.1. *For every n there exists a constant $J(n)$ such that any finite subgroup of GL_n over a field of characteristic zero possesses an abelian normal subgroup of index $\leq J(n)$.*

The corresponding statement in characteristic $p > 0$ is false. For example, the group $\mathrm{GL}_n(\bar{\mathbb{F}}_p)$ contains arbitrarily large subgroups of the form $\mathrm{SL}_n(\mathbb{F}_{p^r})$ which are simple modulo their center. The problem lies in the existence of unipotent elements of finite order. If all elements of Γ are semisimple, then Γ has order prime to p and can therefore be lifted to characteristic zero, where Jordan's theorem applies.

The following seems to us essentially the best possible generalization of Jordan's theorem to arbitrary characteristic:

Received by the editors August 20, 2010 and, in revised form, September 2, 2010, and January 27, 2011.

2010 *Mathematics Subject Classification.* Primary 20G40.

The first author was partially supported by a Sloan grant and by NSF grants DMS-9727553 and DMS-0800705.

Theorem 0.2. *For every n there exists a constant $J'(n)$ depending only on n such that any finite subgroup Γ of GL_n over any field k possesses normal subgroups $\Gamma_3 \subset \Gamma_2 \subset \Gamma_1$ such that*

(a) $[\Gamma : \Gamma_1] \leq J'(n)$.
(b) *Either* $\Gamma_1 = \Gamma_2$, *or* $p := \mathrm{char}(k)$ *is positive and* Γ_1/Γ_2 *is a direct product of finite simple groups of Lie type in characteristic* p.
(c) Γ_2/Γ_3 *is abelian of order not divisible by* $\mathrm{char}(k)$.
(d) *Either* $\Gamma_3 = \{1\}$, *or* $p := \mathrm{char}(k)$ *is positive and* Γ_3 *is a* p*-group.*

In particular, we have the following special case:

Theorem 0.3. *For any finite simple group Γ possessing a faithful linear or projective representation of dimension $\leq n$ over a field k we have either*

(a) $|\Gamma| \leq J'(n)$, *or*
(b) $p := \mathrm{char}(k)$ *is positive and* Γ *is a group of Lie type in characteristic* p.

Note that the case (a) allows only finitely many isomorphism classes for each value of n.

Without much effort one can deduce Theorem 0.2 from Theorem 0.3. With some work the latter follows in turn from the classification of finite simple groups. The object of this paper is to give a completely independent proof, based on the theory of algebraic groups instead of methods from finite group theory.

There have been several previous generalizations of Jordan's theorem to characteristic p. Brauer and Feit [2] approached the problem using modular representation theory. They showed that Γ possesses an abelian normal subgroup whose index is bounded by a constant depending on n as well as the order of the p-Sylow subgroup $\Gamma_{(p)}$ of Γ. Unfortunately, this bound is exponential in $|\Gamma_{(p)}|$. Theorem 0.2, by contrast, implies the following bound, whose dependence on $|\Gamma_{(p)}|$ is optimal, as one sees by considering finite groups of Lie type of the form $\mathrm{PGL}_2(\mathbb{F}_{p^r})$. (Assuming classification of finite simple groups, Weisfeiler [32] gave the estimate $O(|\Gamma_{(p)}|^7)$.)

Theorem 0.4. *Any finite subgroup of GL_n over a field of characteristic $p > 0$ possesses an abelian normal subgroup of order prime to p and of index $\leq J'(n) \cdot |\Gamma_{(p)}|^3$.*

In the case $\Gamma \subset \mathrm{GL}_n(\mathbb{F}_p)$, Nori ([26, §3]) and Gabber (see [21, Thm. 12.4.1]) proved results essentially equivalent to Theorem 0.2 using ideas from algebraic geometry. Their approach is based on the fact that every subgroup of order p of Γ is the group of $\mathbb{F}_p$-valued points of a one-parameter additive subgroup $\mathbb{G}_a \hookrightarrow \mathrm{GL}_n$. They relate Γ to the algebraic group generated by these. But this method does not generalize to subgroups $\Gamma \subset \mathrm{GL}_n(\mathbb{F}_{p^r})$. Of course, one can embed $\mathrm{GL}_n(\mathbb{F}_{p^r})$ in $\mathrm{GL}_{nr}(\mathbb{F}_p)$ and obtain an estimate of some kind, but the resulting upper bound $J'(nr)$ on the index tends rapidly to infinity with r.

Our proof resembles that of Nori and Gabber in that we approximate Γ by an algebraic subgroup G of GL_n. It differs, however, in several important respects. Instead of building up G from below by multiplying together algebraic subgroups, we cut it down from above by exploiting irregularities in the overall distribution of the elements of Γ in GL_n. We cannot assume *a priori* that the coefficients of Γ can be made to lie in any particular finite field. Rather, such information is one of the things that must be determined from Γ. Whereas Nori and Gabber can ignore all problems associated with small primes, we cannot avoid dealing with their pathologies. In particular, our framework must be flexible enough to accommodate the Suzuki and Ree groups.

Genericity for finite subgroups. Our approach is based on the following observation. Any special property that Γ might have, and that can be recognized by representation-theoretic information, can be expressed by saying that Γ is contained in some proper algebraic subgroup of GL_n. For example, the tautological representation is reducible if and only if Γ lies in a proper parabolic subgroup. Similar characterizations exist for imprimitivity, tensor decomposability, and so on. In all these cases the algebraic subgroups in question form an algebraic family which is indexed by a scheme of finite type over $\mathbf{Spec}\,\mathbb{Z}$.

Let us imagine that we are given such a family of algebraic subgroups, which is closed under intersections, and that G is the smallest one that contains Γ. If the family is large, there are many properties of special subgroups of G which Γ does not have. If it is sufficiently large for a problem at hand, we call Γ a *sufficiently general finite subgroup of* G. We make this concept precise in Section 2, and show how to recognize it by looking at a suitable representation of G. It may be helpful to think of G as a kind of algebraic envelope of Γ, which replaces the Zariski-closure since every finite subgroup is already Zariski-closed.

Let $G_3 \subset G_2 \subset G_1$ denote the unipotent radical, the radical, and the identity component of G. The subgroups Γ_i in Theorem 0.2 will be roughly equal to $\Gamma \cap G_i$. Observe that the index $[G : G_1]$ is bounded, because G belongs to a family over a scheme of finite type over $\mathbf{Spec}\,\mathbb{Z}$. Thus the least accessible part of Γ is the image of $\Gamma \cap G_1$ in G_1/G_2. After replacing G by a simple quotient of G_1/G_2, we are reduced to the case that G is connected simple. Here we have the following fundamental result. The group of fixed points under a Frobenius map F is denoted G^F (compare Section 3), and the derived group is indicated by the suffix $(\ )^{\mathrm{der}}$.

Theorem 0.5. *Let $\mathscr{G} \to \mathbf{Spec}\,\mathbb{Z}$ be the family of connected adjoint groups with a fixed simple root system Φ, and let G denote a geometric fiber of $\mathscr{G}$. Consider a finite subgroup $\Gamma \subset G$. If Γ is sufficiently general, then the characteristic of the base field of G is positive, and there exists a Frobenius map $F : G \to G$ so that $(G^F)^{\mathrm{der}}$ is simple and*

$$\left(G^F\right)^{\mathrm{der}} \subset \Gamma \subset G^F.$$

The proof of this theorem takes up most of this article and is sketched in the outline below. The other theorems above are deduced from it. The following reformulation will be explained in Section 2.

Theorem 0.6. *The family $\mathscr{G} \to \mathbf{Spec}\,\mathbb{Z}$ of connected adjoint groups with a fixed simple root system Φ possesses a representation $\mathscr{G} \to \mathrm{GL}_n$ with the following property. Consider any algebraically closed field k and any finite subgroup Γ of the associated geometric fiber G of $\mathscr{G}$. If every Γ-invariant subspace of k^n is G-invariant, the characteristic of k is positive and there exists a Frobenius map $F : G \to G$ so that $(G^F)^{\mathrm{der}}$ is simple and*

$$\left(G^F\right)^{\mathrm{der}} \subset \Gamma \subset G^F.$$

History. This paper is a lightly revised version of a manuscript which has been circulating since 1998. In particular, the numbering has remained unchanged from the preprint version except at the end of section 4, from (4.6) on. At the time it was written, the quasi-thin case of the classification of finite simple groups had not yet been completed, at least in print. Our Theorem 0.3 gives much less than classification, in that it applies only to linear groups and does not specify the finite set of exceptions for each n. It was our hope that it could nevertheless serve as a viable substitute in a range of applications in group theory and number theory. To some extent, the subsequent history has justified that hope.

In the last dozen years, there have been several results which in one way or another improved on or extended the results of this paper. In particular, Ehud Hrushovski and Frank Wagner [16] have reproved the basic "nonconcentration estimate" (Theorem 4.2) in a very general, model-theoretic framework. (In fact, since the inductive scheme in the proof of Lemma 4.4 in the original manuscript was circular, theirs is the first complete proof in the literature.) Robert Guralnick proved [11] in effect that if $p \geq n-3 \geq 6$ and Γ_3 is trivial, we can bound $[\Gamma:\Gamma_1]$ by the classical Jordan constant $J(n)$ rather than $J'(n)$. Michael Collins determined the optimal values for $J(n)$ [5] and $J'(n)$ [6]. Both Guralnick and Collins made use of the classification.

Outline of the paper. In Section 1 we show that a whole range of constructions for algebraic varieties and algebraic groups can be carried out simultaneously in families. To emphasize our point of view, we use the term *constructible family* for any morphism of separated schemes of finite type over $\mathbf{Spec}\,\mathbb{Z}$. Although we are interested mainly in the set of geometric fibers of such a family, keeping track of the total space as a scheme enables us to bound the complexity of the fibers in a uniform way. We use only the softest general techniques of algebraic geometry such as Noetherian induction and constructibility of images. In fact, our proofs could perhaps have been cast equivalently into the language of model theory, in the spirit of Hrushovski-Pillay [15].

Section 2 discusses the concept of sufficiently general finite subgroups and their basic properties: They can be made arbitrarily large and assumed to be not contained in any nowhere dense subvariety that belongs to a constructible family.

In Section 3 we show that finite groups of Lie type provide examples of sufficiently general finite subgroups and discuss a corollary of Theorem 0.5. This is not used in the rest of the paper.

In Section 4 we consider an algebraic group G and a subvariety X, each of which belongs to a given constructible family. Using multiple induction over other constructible families we derive an upper bound for the size of $\Gamma \cap X$, for any sufficiently general finite subgroup $\Gamma \subset G$. If G is almost simple, this bound reads

$$\big|\Gamma \cap X\big| \leq c \cdot \big|\Gamma\big|^{\frac{\dim X}{\dim G}}, \tag{0.7}$$

where the constant on the right-hand side depends only on the family to which X belongs. Remarkably, this order of magnitude is the same as when G is defined over a finite field $\mathbb{F}_q$ and $\Gamma = G(\mathbb{F}_q)$. This upper bound is our key technical result as far as algebraic geometry is concerned, and it is used systematically in the rest of the paper.

Section 5 is an application to abelian varieties and plays no further role in the paper.

Sections 6 through 11 contain the proof of Theorem 0.5. Here G is connected adjoint with a fixed simple root system and Γ is a sufficiently general finite subgroup. Over the course of the proof we build up a collection of structural features of G which have counterparts for Γ, such as maximal tori, Borel subgroups, root subgroups, and so on. We can even estimate their number and size.

The starting point is Section 6, where we show that centralizers in Γ satisfy a lower bound of the same order of magnitude as the upper bound (0.7). This is an existence theorem, which we will use as a machine to exhibit nontrivial elements and subgroups of Γ with special properties. We also show that every G-conjugacy class meets Γ in a bounded number of Γ-conjugacy classes. This fact will be important for unipotent conjugacy classes later on.

The centralizer estimate is used in Section 7 to analyze maximal toric subgroups of Γ. They are self-centralizing and have bounded index in their normalizer; these facts allow precise counting arguments (which are known already for finite groups of Lie type). We count the maximal toric subgroups inside centralizers of semisimple elements, using the Jordan decomposition, and prove that the number of maximal toric subgroups is equal to the number of unipotent elements of Γ. By estimating the former number from below, we deduce that Γ must contain some regular unipotent elements. This already implies that the characteristic of the base field is positive and, by the way, reproves Jordan's Theorem 0.1.

In Section 8 we consider a Borel subgroup $B \subset G$ containing a regular unipotent element of Γ. Using the preceding results we manage to show that $\Gamma \cap B$ also contains many regular semisimple elements. With this information it is not hard to construct other types of elements, either via centralizers, or as commutators. In this way one finds sufficiently many elements of Γ in the center of the unipotent radical of B. Usually this center is a root group, but in certain nonstandard cases in small characteristics it may be the product of two root groups. In either case we can find a connected unipotent subgroup V in the center of the unipotent radical such that $\Gamma \cap V$ can be identified with a finite field $\mathbb{F}_V$. It will turn out that Γ is essentially a finite group of Lie type over $\mathbb{F}_V$.

This is proved in Sections 9 through 11. The first problem is to translate the internal characterization of $\mathbb{F}_V$ inside Γ into external information on coefficients in suitable representations. This is achieved by showing that the traces of certain elements $\gamma \in \Gamma$ lie in $\mathbb{F}_V$. Varying γ, we can construct global coordinates over $\mathbb{F}_V$ for some algebraic representation of G. Eventually this leads to the desired Frobenius map F on G such that $\Gamma \subset G^F$. Our size estimates then imply that the index is bounded, and Theorem 0.5 follows. Both here and in Section 8 there are a number of additional difficulties in characteristics 2 and 3 if G possesses nonstandard isogenies. But our proof of Theorem 0.5 covers all these cases.

The other theorems mentioned above are proved in Section 12. For further information, see the introductions to the individual sections.

Notation. The cardinality of a set X is denoted by $|X|$. For any group G acting on a set X, the *normalizer* of a subset $Y \subset X$ is

$$N_G(Y) \;:=\; \bigl\{\, g \in G \;\bigm|\; \forall y \in Y\colon\; gy \in Y \,\bigr\}.$$

The simultaneous *centralizer* of Y is

$$G_Y \;:=\; \bigl\{\, g \in G \;\bigm|\; \forall y \in Y\colon\; gy = y \,\bigr\}.$$

For any single element $x \in X$ we abbreviate $G_x := G_{\{x\}}$. Its *orbit* is

$$O_G(x) \ := \ \big\{ \, gx \ \big| \ g \in G \, \big\}.$$

Mostly we will apply this to the action of G on itself by conjugation. In this case, $O_G(x)$ is the *conjugacy class* of x. The *center* is denoted by $Z(G) := G_G$ and the commutator subgroup by G^{der}. In the context of algebraic groups these concepts have an algebro-geometric meaning. The identity component of an algebraic group G is denoted by G° and the adjoint group of a connected reductive group by G^{ad}. The Zariski closure of a subset X of an algebraic variety or scheme is denoted by $\overline{X}$.

The following list summarizes notation which is introduced within the text and, in most cases, retains its meaning over several sections.

Symbol	**Page**	**Description**
$N_G(X)$	1109	normalizer
G_X, G_x	1109	(simultaneous) centralizer
$O_G(x)$	1110	orbit, conjugacy class
$Z(\ )$	1110	center
$(\ )^{\mathrm{der}}$	1110	derived group
$(\ )^\circ$	1110	identity component
$(\ )^{\mathrm{ad}}$	1110	adjoint group
$\overline{(\ )}$	1110	Zariski closure
$\mathscr{X}$, $\mathscr{Y}$, ...	1111	constructible family of algebraic varieties
$\mathscr{S}$, $\mathscr{T}$, ...	1111	base scheme of a constructible family
$\mathscr{X}_s$, $\mathscr{Y}_t$, ...	1111	geometric fiber
$\mathscr{X}_{\mathscr{T}}$	1112	pullback of a constructible family
$\mathscr{G}$	1114, 1127	constructible family of algebraic groups
$\mathscr{H}$	1114	constructible family of algebraic subgroups
$(\ )^n$	1115	n-fold fiber product with itself
$G = \mathscr{G}_s$	1116, 1127	algebraic group
Γ	1116, 1127	finite subgroup of G
$\mathbb{F}_q$	1119	finite field with q elements
Φ	1120, 1127	root system of G
F	1120	Frobenius map on G
q_F	1120	numerical constant attached to F
G^F	1120	fixed points of F
q_Γ	1122, 1127	numerical constant attached to Γ
$X = \mathscr{X}_t$	1123	subvariety of G
$Y = \mathscr{Y}_t$	1123	subvariety of G
k	1127	algebraically closed base field
p	1127	characteristic of k
Λ	1128, 1129, 1143	subset of Γ
c_0	1128	constant in Theorem 6.2
$(\ )^{\mathrm{rss}}$	1129	subset of regular semisimple elements
$(\ )^{\mathrm{un}}$	1129	subset of unipotent elements
Θ	1130	maximal toric subgroup of Γ
Tor_Λ	1131	set of maximal toric subgroups of Γ_Λ°
$\mathrm{Tor}_\Lambda^\natural$	1131	subset of representatives under Γ_Λ°
B	1134	Borel subgroup of G
U	1134	unipotent radical of B

T	1134	maximal torus of B
Φ^+	1134	set of positive roots
$\mathbb{G}_a$	1134	additive group
U_α	1134	root group
α_ℓ, α_s	1135	highest long and short roots
U^{run}	1137	regular unipotent elements in U
V	1138	nontrivial subgroup of $Z(U)$ normalized by B
d	1139	dimension of V
$\mathbb{F}_V$	1139	finite field attached to Γ and V
p^r	1140	order of $\mathbb{F}_V$
p^e	1140	Frobenius twist
ρ, ρ_ℓ, ρ_s	1142	constituents of the adjoint representation
Ψ	1149	root subsystem
$\dot{w}$	1149	longest Weyl group element
$H_{(g)}$	1149	subgroup generated by V and gVg^{-1}
$\mathrm{Rad}_u\, G$	1154	unipotent radical

1. Constructible families

Most constructions in algebraic geometry have a meaning not only for single algebraic varieties but can be carried out in families. That is, the fibers of a morphism $\mathscr{X} \to \mathscr{S}$ of finite type are viewed as forming a family of algebraic varieties $\mathscr{X}_s$, and operating with the total space $\mathscr{X}$ amounts to doing the same with all fibers at the same time in a coherent fashion. The result is then another family, i.e., another morphism of finite type. Now, it is a basic fact of algebraic geometry that many properties of fibers, such as dimension or number of components, vary constructibly over the base. It follows that the numerical invariants arising in such constructions are bounded uniformly in the family. This phenomenon plays a central role in the counting arguments of Section 4. The current section is devoted to establishing the necessary framework for them. All this is basically standard algebraic geometry.

Conventions. We are eventually interested only in questions concerning varieties over algebraically closed fields. Nevertheless, since we aim at statements that are independent of characteristic, we are forced to use the language of schemes (see [9], [12]). For simplicity we assume that all our schemes are separated and of finite type over $\mathbf{Spec}\,\mathbb{Z}$.

By a *variety* we will always mean the set of closed points of a scheme of finite type over an algebraically closed field, with its induced structure of an algebraic variety in the common sense. Note that a scheme and its reduced subscheme determine the same variety. Note also that a variety is not required to be irreducible (compare [1], [17]). Usually, schemes will be denoted by calligraphic letters, varieties by roman letters. For example, the fiber of a morphism of schemes $\mathscr{X} \to \mathscr{S}$ over a geometric point s of $\mathscr{S}$ determines a variety, called simply the geometric fiber above s, and abbreviated by $X := \mathscr{X}_s$.

To clarify our point of view we use the following terminology:

Definition 1.1. A *constructible family* $\mathscr{X} \to \mathscr{S}$ is a morphism of schemes of finite type over $\mathbf{Spec}\,\mathbb{Z}$.

The pullback of such a constructible family by a morphism $\mathscr{T} \to \mathscr{S}$ will be abbreviated $\mathscr{X}_{\mathscr{T}} := \mathscr{X} \times_{\mathscr{S}} \mathscr{T}$.

Definition 1.2. A *morphism* from a constructible family $\mathscr{Y} \to \mathscr{T}$ to a constructible family $\mathscr{X} \to \mathscr{S}$ consists of a morphism $\mathscr{T} \to \mathscr{S}$ and a $\mathscr{T}$-morphism $\varphi\colon \mathscr{Y} \longrightarrow \mathscr{X}_{\mathscr{T}}$.

Definition 1.3. A *constructible family of subvarieties* is a morphism for which $\mathscr{Y} \longrightarrow \mathscr{X}_{\mathscr{T}}$ is a closed embedding.

Thus the geometric points t of $\mathscr{T}$ parametrize a family of morphisms, resp. closed embeddings, of varieties $\varphi_t : \mathscr{Y}_t \to \mathscr{X}_s$, where s denotes the corresponding geometric point of $\mathscr{S}$. Note that we do not rule out the possibility that the same subvariety of $\mathscr{X}_s$ occurs for different t. In fact, to avoid this in our constructions would be quite a burden and without any benefit.

Numerical invariants. One of the main features of constructible families is that numerical invariants of the fibers are uniformly bounded:

Proposition 1.4. *In any given constructible family $\mathscr{X} \to \mathscr{S}$ the dimension and the number of irreducible components of the geometric fibers $\mathscr{X}_s$ are bounded.*

Proof. See [10, Prop. 13.1.7, Cor. 9.7.9]. □

Stratifications. The word stratification normally refers to the decomposition of a scheme into a disjoint union of locally closed subschemes, perhaps satisfying additional hypotheses. In our case we care only about the following property:

Definition 1.5. A *stratification map* is a morphism $\mathscr{S}' \to \mathscr{S}$ which induces a bijection on geometric points.

Various useful scheme-theoretic properties can be attained by pulling back a constructible family via a stratification map:

Proposition 1.6. *For any constructible family $\mathscr{X} \to \mathscr{S}$ there exists a stratification map $\mathscr{S}' \to \mathscr{S}$ such that $\mathscr{X}_{\mathscr{S}'} \longrightarrow \mathscr{S}'$ is flat and its fiber dimension locally constant on $\mathscr{S}'$.*

Proof. See [10, Thm. 11.1.1, Thm. 13.1.3]. □

Fiberwise closure. The process of taking Zariski-closure does not generally commute with taking fibers unless one first pulls everything back by a suitable stratification map. For the closure of the image in a family of morphisms we have:

Proposition 1.7. *For any morphism of constructible families $\varphi\colon \mathscr{Y} \to \mathscr{X}_{\mathscr{T}}$, there exist a stratification map $\mathscr{T}' \to \mathscr{T}$ and a constructible family of subvarieties $\mathscr{Z} \to \mathscr{T}'$ of $\mathscr{X} \to \mathscr{S}$ with the following property. For any geometric point t of $\mathscr{T}$, with t' the corresponding geometric point of $\mathscr{T}'$, we have*

$$\mathscr{Z}_{t'} = \overline{\varphi_t(\mathscr{Y}_t)}.$$

Proof. The image of φ is a constructible subset of $\mathscr{X}_{\mathscr{T}}$ (see [10, Prop. 9.2.6]). Its closure in the total space is constructible by definition, so by [10, Prop. 9.5.3] the points t for which $\varphi_t(\mathscr{Y}_t) = \varphi(\mathscr{Y})_t$ is dense in $\big(\overline{\varphi(\mathscr{Y})}\big)_t$ form a constructible subset of $\mathscr{T}$. This set contains all generic points of $\mathscr{T}$ and therefore some open dense subscheme $\mathscr{U}$. Pulling the morphism φ back to $\mathscr{T}_1 := \mathscr{T} \smallsetminus \mathscr{U}$, by Noetherian induction we already have a stratification map $\mathscr{T}_1' \to \mathscr{T}_1$ and a constructible family of subvarieties $\mathscr{Z}_1 \subset \mathscr{X}_{\mathscr{T}_1'}$ with the desired property over $\mathscr{T}_1$. Putting $\mathscr{T}' := \mathscr{U} \sqcup \mathscr{T}_1'$ and $\mathscr{Z} := \big(\overline{\varphi(\mathscr{Y})} \times_{\mathscr{T}} \mathscr{U}\big) \sqcup \mathscr{Z}_1$ the assertion follows over $\mathscr{T}$. □

Similarly, for the locus of points with given fiber dimension we have:

Proposition 1.8. *For any morphism of constructible families $\varphi\colon \mathscr{Y} \to \mathscr{X}_{\mathscr{T}}$ and any integer $d \geq 0$, there exist a stratification map $\mathscr{T}' \to \mathscr{T}$ and a constructible family of subvarieties $\mathscr{Z} \to \mathscr{T}'$ of $\mathscr{X} \to \mathscr{S}$ with the following property. Take any geometric point t of $\mathscr{T}$ and let t' and s denote the corresponding geometric points of $\mathscr{T}'$, resp. $\mathscr{S}$. Then we have*

$$\mathscr{Z}_{t'} \;=\; \overline{\big\{\, x \in \mathscr{X}_s \;\big|\; \dim(\varphi_t^{-1}(x)) = d \,\big\}}.$$

Proof. The set of points of $\mathscr{X}_{\mathscr{T}}$ where the fiber dimension is d is constructible by [10, Prop. 9.2.6.1]. Using this, one proceeds as in the preceding proof. □

Irreducible components. To decompose the geometric fibers into irreducible components one needs more than a stratification map:

Proposition 1.9. *For any constructible family $\mathscr{X} \to \mathscr{S}$ there exists a constructible family of subvarieties $\mathscr{Y} \to \mathscr{T}$ such that for every geometric point s of $\mathscr{S}$ the subvarieties $\mathscr{Y}_t \subset \mathscr{X}_s$, as t runs through all geometric points of $\mathscr{T}$ above s, are precisely the irreducible components of the geometric fiber $\mathscr{X}_s$.*

Proof. We proceed as in the proof of [10, Thm. 9.7.7]. Consider a generic point η of $\mathscr{S}$. By [10, Cor. 4.6.8], there exists a finite extension K' of its residue field K such that every irreducible component Z_i of $\mathscr{X}_\eta \times_{\mathbf{Spec}\,K} \mathbf{Spec}\,K'$ is geometrically irreducible. Choose a morphism $\mathscr{U} \to \mathscr{S}$ of finite type, where $\mathscr{U}$ is integral with function field K'. For each i let $\mathscr{Z}_i$ denote the Zariski-closure of Z_i in $\mathscr{X}_{\mathscr{U}}$.

By [10, Thm. 9.7.7], the fibers of $\mathscr{Z}_i \to \mathscr{U}$ are geometrically irreducible in a neighborhood of the generic point. Thus after shrinking $\mathscr{U}$ all these fibers are geometrically irreducible. Next, over the generic point, none of these fibers is contained in any other. By [10, Cor. 9.5.2], the same is true over a whole neighborhood, so after shrinking again it is true over all of $\mathscr{U}$. Furthermore, the inclusion $\bigcup_i \mathscr{Z}_i \subset \mathscr{X}_{\mathscr{U}}$ is an equality over the generic point. By [10, Cor. 9.5.2], this remains true in a neighborhood, and so again without loss of generality over all of $\mathscr{U}$. We conclude that the fibers of the different families $\mathscr{Z}_i \to \mathscr{U}$ are precisely the irreducible components of the fibers of $\mathscr{X}_{\mathscr{U}} \longrightarrow \mathscr{U}$.

This solves the problem in a neighborhood of the generic point η. To finish, we apply Noetherian induction to the pullback of $\mathscr{X}$ to a suitable complement in $\mathscr{S}$, take the resulting family of subvarieties, and let $\mathscr{Y} \to \mathscr{T}$ be its disjoint union with all $\mathscr{Z}_i \to \mathscr{U}$. The desired assertion follows. □

Intersections. Since the topological space underlying an algebraic variety is Noetherian, the intersection of any collection of closed subvarieties is already the intersection of a finite number of them. The following result shows that this number is uniformly bounded when both the subvarieties and the ambient variety are allowed to vary in constructible families.

Theorem 1.10. *For any constructible family $\mathscr{X} \to \mathscr{S}$ and any constructible family of subvarieties $\mathscr{Y} \to \mathscr{T}$ there is an integer n with the following property. Consider* *any geometric point s of $\mathscr{S}$ and any collection I of geometric points of $\mathscr{T}$ above s. Then there exists a subset $I' \subset I$ of at most n points such that*

$$\bigcap_{t\in I} \mathscr{Y}_t \;=\; \bigcap_{t\in I'} \mathscr{Y}_t.$$

Proof. Fix d such that the fiber dimension of $\mathscr{X} \to \mathscr{S}$ is everywhere $\leq d$. Then every intersection in question is a variety of dimension $\leq d$. To any such variety Z let us associate the tuple $\underline{r}(Z) := (r_d, \ldots, r_0) \in \mathbb{N}^{d+1}$, where r_i is the number of irreducible components of Z of dimension i. Consider the lexicographical total order on $\mathbb{N}^{d+1}$ defined by $(r_d, \ldots, r_0) < (r'_d, \ldots, r'_0)$ if and only if in the leftmost entry where these tuples differ we have $r_i < r'_i$. It is well known that this makes $\mathbb{N}^{d+1}$ a well-ordered set. Note also that $\underline{r}(Z) < \underline{r}(Z')$ whenever $Z \subsetneqq Z'$.

Now let us assume that the theorem is false. Then for every n there exist geometric points $t_1, \ldots, t_n \mapsto s$ such that $Z := \mathscr{Y}_{t_1} \cap \ldots \cap \mathscr{Y}_{t_n}$ cannot be written as an intersection of fewer terms. Since all intersections of n terms form the constructible family of subvarieties

$$\underbrace{\mathscr{Y} \times_{\mathscr{X}} \ldots \times_{\mathscr{X}} \mathscr{Y}}_{n} \longrightarrow \underbrace{\mathscr{T} \times_{\mathscr{S}} \ldots \times_{\mathscr{S}} \mathscr{T}}_{n}, \tag{1.11}$$

Proposition 1.4 implies that there are only finitely many possibilities for the associated tuple $\underline{r}(Z)$. Let $\underline{r}(n)$ be the maximum of $\underline{r}(Z)$ for all Z which are intersections of n terms but not of fewer terms. We claim that $\underline{r}(n) < \underline{r}(n-1)$. Indeed, for suitable $t_1, \ldots, t_n$ we have

$$\underline{r}(n) = \underline{r}\big(\mathscr{Y}_{t_1} \cap \ldots \cap \mathscr{Y}_{t_n}\big) < \underline{r}\big(\mathscr{Y}_{t_1} \cap \ldots \cap \mathscr{Y}_{t_{n-1}}\big) \leq \underline{r}(n-1),$$

as claimed. Thus the elements $\underline{r}(n)$ form an infinite strictly decreasing sequence, contradicting the fact that $\mathbb{N}^{d+1}$ is well ordered. □

As a consequence, arbitrary intersections of closed subvarieties in a constructible family form a constructible family:

Corollary 1.12. *For any constructible family $\mathscr{X} \to \mathscr{S}$ and any constructible family of subvarieties $\mathscr{Y} \to \mathscr{T}$ there exists another constructible family of subvarieties $\mathscr{Z} \to \mathscr{U}$ with the following property. Consider any geometric point s of $\mathscr{S}$. Then for any nonempty collection I of geometric points of $\mathscr{T}$ above s there exists a geometric point u of $\mathscr{U}$ above s with*

$$\bigcap_{t\in I} \mathscr{Y}_t \;=\; \mathscr{Z}_u.$$

Conversely, every $\mathscr{Z}_u$ is such an intersection.

Proof. With n as in Theorem 1.10, the family (1.11) has the desired property with respect to all intersections of a positive number of terms. □

Families of algebraic groups. For the general theory of algebraic groups and group schemes, see [7], [1], or [17]. Following our general conventions, an algebraic group is always of finite type over an algebraically closed field. In accordance with Definition 1.1 a *constructible family of algebraic groups* is a group scheme $\mathscr{G} \to \mathscr{S}$, where $\mathscr{G}$ and $\mathscr{S}$ are of finite type over $\mathbf{Spec}\,\mathbb{Z}$. Similarly, a *constructible family of algebraic subgroups* of $\mathscr{G} \to \mathscr{S}$ consists of a morphism $\mathscr{T} \to \mathscr{S}$ and a closed subgroup scheme $\mathscr{H} \subset \mathscr{G}_{\mathscr{T}}$. Usually we have in mind constructible families of linear algebraic groups, i.e., of algebraic subgroups of GL_n for some n. But our results also have consequences for abelian varieties; see Section 5.

An *action* of $\mathscr{G} \to \mathscr{S}$ on a constructible family $\mathscr{X} \to \mathscr{T}$ is a morphism $\mu : \mathscr{G} \times_{\mathscr{S}} \mathscr{X} \longrightarrow \mathscr{X}$ satisfying the usual associativity and identity axioms. If $\mathscr{X} \to \mathscr{T}$ is a vector bundle and the action is linear, this is a *constructible family of representations.* The n-fold *fiber product* of $\mathscr{G}$ with itself over $\mathscr{S}$ will be denoted by $\mathscr{G}^n$.

Transporter, Normalizer, Centralizer. In general these fiberwise constructions can be carried out only after a suitable stratification map. We begin with transporters and normalizers:

Proposition 1.13. *Consider a constructible family of algebraic groups $\mathscr{G} \to \mathscr{S}$ which acts on a constructible family $\mathscr{X} \to \mathscr{S}$. Consider constructible families of subvarieties $\mathscr{Y}_1 \to \mathscr{T}$ and $\mathscr{Y}_2 \to \mathscr{T}$ of $\mathscr{X} \to \mathscr{S}$. Then there exist a stratification map $\mathscr{T}' \to \mathscr{T}$ and a constructible family $\mathscr{Z}' \to \mathscr{T}'$ of subvarieties of $\mathscr{G} \to \mathscr{S}$ with the following property. Take any geometric point t of $\mathscr{T}$ and let t' and s denote the corresponding geometric points of $\mathscr{T}'$, resp. $\mathscr{S}$. Then we have*

$$\mathscr{Z}'_{t'} = \left\{\, g \in \mathscr{G}_s \;\middle|\; g\mathscr{Y}_{1,t} \subset \mathscr{Y}_{2,t} \,\right\}.$$

If $\mathscr{Y}_1 = \mathscr{Y}_2$, then $\mathscr{Z}' \to \mathscr{T}'$ is a family of algebraic subgroups, with $\mathscr{Z}'_{t'} = N_{\mathscr{G}_s}(\mathscr{Y}_{1,t})$.

Proof. First we look at a single geometric fiber. We must prove that the right-hand side in the above equality is Zariski-closed in $\mathscr{G}_s$. To do this, note that for every point y the set $\{g \in \mathscr{G}_s \mid gy \in \mathscr{Y}_{2,t}\}$ is Zariski-closed. The transporter is the intersection of these as y runs through $\mathscr{Y}_{1,t}$, so it is closed.

To extend this argument to the family let us first replace $\mathscr{G}$ and $\mathscr{X}$ by their pullbacks to $\mathscr{T}$, after which we may assume $\mathscr{T} = \mathscr{S}$. Let $\mu\colon \mathscr{G} \times_{\mathscr{S}} \mathscr{X} \longrightarrow \mathscr{X}$ be the morphism defining the group action, and consider the subscheme

$$\mu^{-1}(\mathscr{Y}_2) \cap \bigl(\mathscr{G} \times_{\mathscr{S}} \mathscr{Y}_1\bigr) \;\subset\; \mathscr{G} \times_{\mathscr{S}} \mathscr{Y}_1.$$

By [10, Cor. 9.5.2], the points $g \in \mathscr{G}$ over which this inclusion is an equality form a constructible subset $\mathscr{Z}$. In any geometric fiber this is precisely the desired transporter. Since it is closed in every generic fiber, it is a closed subset over some open dense subscheme $\mathscr{U} \subset \mathscr{T}$. We conclude by Noetherian induction, as in the proof of Proposition 1.7. (For other approaches, see [7, Exp.$\mathrm{VI_B}$ §6.1] or [19, §2.6].) □

Applying this to the conjugation action of $\mathscr{G}$ on itself we deduce that the normalizer of an algebraic subgroup which belongs to a constructible family again belongs to a constructible family. One can formulate a similar result for centralizers, but using Corollary 1.12 we can do even better:

Proposition 1.14. *For every constructible family of algebraic groups $\mathscr{G} \to \mathscr{S}$ which acts on a constructible family $\mathscr{X} \to \mathscr{S}$, there exists a constructible family $\mathscr{H} \to \mathscr{T}$ of algebraic subgroups of $\mathscr{G} \to \mathscr{S}$ with the following property. Take any geometric point s of $\mathscr{S}$. Then for any subset $I \subset \mathscr{X}_s$ there exists a point t of $\mathscr{T}$ above s such that*

$$\mathscr{H}_t = \bigl(\mathscr{G}_s\bigr)_I := \left\{\, g \in \mathscr{G}_s \;\middle|\; \forall x \in I\colon\ gx = x \,\right\}.$$

Conversely, every $\mathscr{H}_t$ is such a centralizer.

Proof. Consider the morphism

$$\mathscr{G} \times_{\mathscr{S}} \mathscr{X} \longrightarrow \mathscr{X}, \ (g, x) \mapsto (gx, x).$$

The pullback of the diagonal is a closed subscheme, consisting of all points (g, x) with $gx = x$. Therefore the point stabilizers form a constructible family of algebraic subgroups. By Corollary 1.12, arbitrary intersections of these form again a constructible family, as desired. □

Nonconstructible families. There are collections of algebraic subgroups which cannot be the fibers of any constructible family. For instance, every finite subgroup is algebraic, but if it varies in a constructible family its cardinality is bounded, by Proposition 1.4. A similar phenomenon may happen even when the subgroups are connected. For example, consider the standard torus $\mathbb{G}_m^d$ of dimension $d \geq 2$ over $\mathbf{Spec}\,\mathbb{Z}$ and a constructible family of 1-dimensional subtori T. Then the degree of all projection maps $\mathrm{pr}_i \colon T \to \mathbb{G}_m$ is bounded, leaving only finitely many possibilities for the type of T. It follows that the collection of all 1-dimensional subtori does not form a constructible family.

Similar examples can be obtained by means of Frobenius twisting. For instance, for any algebraically closed field k of characteristic $p > 0$ and any integer $n \geq 0$ consider the algebraic subgroup of $\mathrm{GL}_{3,k}$ consisting of all matrices

$$\begin{pmatrix} 1 & x & x^{p^n} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Here the degree of the projection map to the upper right corner is p^n, which would have to be bounded in any constructible family. Thus even in fixed positive characteristic we have a collection of unipotent subgroups which does not form a constructible family.

2. Genericity for finite subgroups

Traditionally a point on an algebraic variety is called *generic* or *general* if it does not satisfy some nontrivial Zariski-closed condition which remains tacit but is understood to be fixed during the discussion under way. In other words, it has to lie inside an arbitrarily small but fixed Zariski-dense open subset. This subset may depend on choices which have already been made but should not be modified after genericity is invoked. With the advent of schemes this somewhat vague concept was turned into a precise technical term under the name "generic point". But the old point of view retains its usefulness, particularly in the setting we have in mind.

Consider a constructible family of algebraic groups $\mathscr{G} \to \mathscr{S}$ and abbreviate a typical geometric fiber by $G := \mathscr{G}_s$. Consider a finite subgroup $\Gamma \subset G$. If Γ is contained in some previously given algebraic subgroup $H \subset G$ of smaller dimension, we can try to analyze it using induction on $\dim H$. The same applies when H varies in a constructible family of fiberwise nowhere dense algebraic subgroups $\mathscr{H} \to \mathscr{T}$. Even if nothing else is known about this family, we can still conclude from Proposition 1.4 that the number of irreducible components of H is bounded. Thus if $\Gamma \cap H^\circ$ is somehow understood by induction, we obtain a rough qualitative description of Γ itself. This recursive analysis will be carried out in detail in Section 12.

The big remaining problem is to deal with the "generic" case, where Γ is not contained in an algebraic subgroup of smaller dimension over which one has this kind of control. The following terminology serves as a conceptual framework for this. To avoid confusion with the meaning of "generic" in modern algebraic geometry, we use the word "general".

Definition 2.1. Let $\mathscr{G} \to \mathscr{S}$ be a constructible family of algebraic groups, and let $\mathscr{H} \to \mathscr{T}$ be a constructible family of fiberwise nowhere dense algebraic subgroups. A finite subgroup Γ of a geometric fiber $\mathscr{G}_s$ is called *$\mathscr{H}$-general* if and only if for every point t of $\mathscr{T}$ above s we have $\Gamma \not\subset \mathscr{X}_t$.

Note that a closed subgroup H of an algebraic group G is nowhere dense if and only if H does not contain the identity component of G.

In Definition 2.1, the family $\mathscr{H}$ may be complicated to describe and awkward to carry along in our notation. Therefore we will mostly use the following abbreviation.

Metadefinition 2.2. Let $\mathscr{G} \to \mathscr{S}$ be a constructible family of algebraic groups and consider a statement $\mathbf{A}(\Gamma)$ about finite subgroups Γ of a geometric fiber $\mathscr{G}_s$. The following assertions are defined as equivalent:

(a) *For any sufficiently general* Γ we have $\mathbf{A}(\Gamma)$.

(b) There exists a constructible family of fiberwise nowhere dense algebraic subgroups $\mathscr{H} \to \mathscr{T}$ of $\mathscr{G} \to \mathscr{S}$ such that for any geometric point s of $\mathscr{S}$ and any *$\mathscr{H}$-general* finite subgroup $\Gamma \subset \mathscr{G}_s$ we have $\mathbf{A}(\Gamma)$.

To further justify this usage, let us imagine that the collection of all finite subgroups possesses some kind of algebro-geometric structure. For every n the subgroups of order $\leq n$ form a constructible family (see Proposition 2.5 below), but in the limit for $n \to \infty$ the parameter space could perhaps be viewed as infinite dimensional. For any fixed $\mathscr{H}$, the set of $\mathscr{H}$-general finite subgroups should then be an open dense subvariety, and as $\mathscr{H}$ varies, these subvarieties should be cofinal among all open dense subvarieties. With this interpretation our use of the word "general" becomes a direct analogue of the classical one.

Recognizing genericity. For some applications it will be useful to translate the above concept into the language of invariant theory.

Proposition 2.3. *Consider a constructible family of* linear *algebraic groups $\mathscr{G} \to \mathscr{S}$, and a constructible family of fiberwise nowhere dense algebraic subgroups $\mathscr{H} \to \mathscr{T}$.*

(a) *There exist a stratification map $\mathscr{S}' \to \mathscr{S}$ and a constructible family of representations of $\mathscr{G}$ on a vector bundle $\mathscr{W} \to \mathscr{S}'$ with the following property. Consider a geometric point s of $\mathscr{S}$, with corresponding point s' of $\mathscr{S}'$, and a finite subgroup $\Gamma \subset \mathscr{G}_s$. If every Γ-invariant subspace of the fiber $\mathscr{W}_{s'}$ is $\mathscr{G}_s$-invariant, then Γ is $\mathscr{H}$-general.*

(b) *If a faithful representation of $\mathscr{G}$ on a vector bundle $\mathscr{V} \to \mathscr{S}$ is given, then in (a) one can take $\mathscr{S}' := \mathscr{S}$ and*

$$\mathscr{W} := \bigoplus_{i=1}^{r} \mathscr{V}^{\otimes m_i} \otimes (\mathscr{V}^{\vee})^{\otimes n_i}$$

for suitable integers r, m_i, and n_i.

Proof. First we prove (b), using Noetherian induction on $\mathscr{T}$. Consider a generic point θ of $\mathscr{T}$, and let η be the corresponding point of $\mathscr{S}$. Then $\mathscr{H}_\theta$ is a closed algebraic subgroup of $\mathscr{G}_\eta$. It can therefore ([1, Chap. II, Thm. 5.1]) be described as the stabilizer of a subspace $\mathscr{W}'_\theta$ of some tensor space

$$\mathscr{W}_\eta := \bigoplus_{i=1}^{r} \mathscr{V}_\eta^{\otimes m_i} \otimes (\mathscr{V}_\eta^\vee)^{\otimes n_i}.$$

This subspace extends to a vector subbundle $\mathscr{W}'$ over a neighborhood $\mathscr{U}$ of θ in $\mathscr{T}$. Since $\mathscr{H}$ coincides with the stabilizer of $\mathscr{W}'$ at the generic point θ, by [10, Cor. 9.5.2], it does so over a whole neighborhood. Let us shrink $\mathscr{U}$ accordingly. Then for any geometric point t of $\mathscr{U}$ with image s in $\mathscr{S}$, and any subgroup $\Gamma \subset \mathscr{G}_s$, we have $\Gamma \subset \mathscr{H}_t$ if and only if $\mathscr{W}'_t$ is Γ-invariant. Now recall that, by assumption, $\mathscr{H}_t$ is a proper subgroup of $\mathscr{G}_s$. Therefore $\mathscr{W}'_t$ is not $\mathscr{G}_s$-invariant. Thus if every Γ-invariant subspace of $\mathscr{W}_s$ is $\mathscr{G}_s$-invariant, then Γ is $(\mathscr{H} \times_{\mathscr{T}} \mathscr{U})$-general.

Repeating this argument by Noetherian induction, we obtain a finite stratification of $\mathscr{T}$ and for each stratum a vector bundle of the desired form, which detects whether Γ is general with respect to the corresponding subfamily of $\mathscr{H}$. Clearly the direct sum of these vector bundles does the job over all of $\mathscr{T}$, which proves (b).

To prove (a) we will construct a faithful representation of $\mathscr{G}$ on a vector bundle over $\mathscr{S}'$. For this, note first that by assumption any generic fiber of $\mathscr{G} \to \mathscr{S}$ possesses a faithful linear representation $\mathscr{G}_\eta \hookrightarrow \mathrm{GL}_n$. This homomorphism extends automatically to an open neighborhood $\mathscr{U} \subset \mathscr{S}$. Its kernel is a Zariski-closed subgroup scheme of $\mathscr{G} \times_{\mathscr{S}} \mathscr{U}$ whose generic fiber coincides with the identity section. By [10, Cor. 9.5.2], the same is true over a whole neighborhood, so after shrinking $\mathscr{U}$, this representation is faithful. Applying Noetherian induction to the complement $\mathscr{S} \smallsetminus \mathscr{U}$ we find a faithful representation of $\mathscr{G}$ on a vector bundle $\mathscr{V} \to \mathscr{S}'$, where $\mathscr{S}' \to \mathscr{S}$ is a stratification map. Now (b) implies (a). □

Proof of Theorem 0.6. We deduce this from Theorem 0.5. Let $\mathscr{H} \to \mathscr{T}$ be the constructible family of fiberwise nowhere dense algebraic subgroups implicit in Theorem 0.5 via Metadefinition 2.2. Take the representation furnished by Proposition 2.3(b), starting with the adjoint representation of $\mathscr{G}$. Then the assumptions in Theorem 0.6 imply that Γ is $\mathscr{H}$-general, so the desired assertion follows from Theorem 0.5. □

Subvarieties versus subgroups. With equal right, one might have defined the concept of sufficiently general finite subgroups with respect to arbitrary nowhere dense subvarieties instead of subgroups. But this makes no difference:

Proposition 2.4. *Let $\mathscr{G} \to \mathscr{S}$ be a constructible family of algebraic groups, and let $\mathscr{X} \to \mathscr{T}$ be a constructible family of fiberwise nowhere dense subvarieties. Then for any sufficiently general finite subgroup Γ of a geometric fiber $\mathscr{G}_s$ and every point t of $\mathscr{T}$ above s we have $\Gamma \not\subset \mathscr{X}_t$.*

Proof. We first look at the problem for a single fiber. Suppose that $\Gamma \subset X := \mathscr{X}_t$, and put $Y := \bigcap_{\gamma \in \Gamma} \gamma X$. By construction this is a nowhere dense closed subvariety of $\mathscr{G}_s$ which is invariant under left translation by Γ. In other words, Γ is contained in the normalizer N of Y for the left translation action of $\mathscr{G}_s$ on itself, and N is a nowhere dense algebraic subgroup.

In view of Metadefinition 2.2 it suffices to show that these subgroups N form a constructible family. By Corollary 1.12 this is already so for the subvarieties Y. By Proposition 1.13 the same follows for N, as desired. □

General finite subgroups are arbitrarily large. To further illustrate the concept we note the following basic fact:

Proposition 2.5. *Let $\mathscr{G} \to \mathscr{S}$ be a constructible family of algebraic groups of dimension ≥ 1 and fix an integer n. Then any sufficiently general finite subgroup Γ of a geometric fiber $\mathscr{G}_s$ has order $> n$.*

Proof. The individual points on $\mathscr{G}$ are indexed by the tautological family $\mathrm{id}\colon \mathscr{G} \to \mathscr{G}$, so the nonempty finite subsets of $\mathscr{G}_s$ of cardinality $\leq n$ can be indexed by $\mathscr{G}^n$, the nth fiber power of $\mathscr{G}$ relative to $\mathscr{S}$. The condition for a finite subset to be a subgroup is Zariski-closed. Thus the subgroups of order $\leq n$ are the fibers of some constructible family of subgroups of $\mathscr{G} \to \mathscr{S}$. Now Metadefinition 2.2 applies. □

3. Finite groups of Lie type

In this section we show that finite groups of Lie type are sufficiently general in the sense of Metadefinition 2.2, whenever the base field is sufficiently large. This result is intended to clarify the scope of the concept of sufficiently general subgroups, although it will play no further role in this paper. We begin with the following estimate:

Proposition 3.1. *For any connected algebraic group G over a finite field $\mathbb{F}_q$ with q elements, we have*

$$(\sqrt{q}-1)^{2\dim G} \;\leq\; \big|G(\mathbb{F}_q)\big| \;\leq\; (\sqrt{q}+1)^{2\dim G}.$$

Proof. For abelian varieties these bounds are best possible; see [24, §21, Thm. 4]. For connected linear algebraic groups one has the stronger estimate $(q-1)^{\dim G} \leq \big|G(\mathbb{F}_q)\big| \leq (q+1)^{\dim G}$ (compare, e.g., [26, Lemma 3.5]). Every connected algebraic group is an extension of an abelian variety by a connected linear algebraic group ([22]). Lang's theorem implies that every short exact sequence of connected algebraic groups induces a short exact sequence on $\mathbb{F}_q$-valued points. Thus the bounds follow in general. □

Proposition 3.2. *Let $\mathscr{G} \to \mathscr{S}$ be a constructible family of algebraic groups, and $\mathscr{H} \to \mathscr{T}$ a constructible family of fiberwise nowhere dense algebraic subgroups. Then there exists a constant q_0 such that for every finite field $\mathbb{F}_q$ with $q \geq q_0$ elements and every point $s \in \mathscr{S}(\mathbb{F}_q)$ the subgroup $\mathscr{G}_s(\mathbb{F}_q)$ is $\mathscr{H}$-general.*

Proof. For any geometric point t of $\mathscr{T}$ above s we must show that $\mathscr{G}_s(\mathbb{F}_q) \not\subset \mathscr{H}_t$. We cannot apply the estimate in Proposition 3.1 directly to $\mathscr{H}_t$, because this subgroup is not necessarily defined over $\mathbb{F}_q$. Let K be the intersection of all translates of $\mathscr{H}_t$ under powers of the Frobenius Frob_q. This subgroup is defined over $\mathbb{F}_q$ and satisfies $\mathscr{G}_s(\mathbb{F}_q) \cap \mathscr{H}_t = K(\mathbb{F}_q)$. Every Frobenius translate of $\mathscr{H}_t$ is a (possibly different) geometric fiber of the same constructible family $\mathscr{H} \to \mathscr{T}$. Thus although K is the intersection of an indeterminate number of terms, by Corollary 1.12 it is a fiber of a constructible family of algebraic subgroups. Now Proposition 1.4 shows that the index $[K : K^\circ]$ is bounded by some fixed constant c. Abbreviating $G := \mathscr{G}_s$, Proposition 3.1 implies that

$$\frac{|K(\mathbb{F}_q)|}{|G(\mathbb{F}_q)|} \;\leq\; c \cdot \frac{|K^\circ(\mathbb{F}_q)|}{|G^\circ(\mathbb{F}_q)|} \;\leq\; c \cdot \frac{(\sqrt{q}+1)^{2\dim K}}{(\sqrt{q}-1)^{2\dim G}} \;\leq\; \frac{c}{q} \cdot \left(\frac{\sqrt{q}+1}{\sqrt{q}-1}\right)^{2\dim G}.$$

For $q \gg 0$ this is less than 1; hence $\mathscr{G}_s(\mathbb{F}_q) \cap \mathscr{H}_t \subsetneqq \mathscr{G}_s(\mathbb{F}_q)$, as desired. □

Remark. The upper bound used in the above proof can be generalized to the number of points on algebraic subvarieties instead of subgroups. Namely, consider any algebraic variety X over $\mathbb{F}_q$. Using elementary estimates, e.g. stratifying X and realizing each stratum as a quasi-finite covering of an affine space, one easily shows that $|X(\mathbb{F}_q)| \leq c \cdot q^{\dim X}$. Here the constant c is independent of $\mathbb{F}_q$ and can remain fixed as X varies in a given constructible family.

Now suppose that $\mathscr{G} \to \mathscr{S}$ is a constructible family of algebraic groups and $\mathscr{X} \to \mathscr{T}$ a constructible family of subvarieties. Abbreviate $G := \mathscr{G}_s$ and $X := \mathscr{X}_t$. The procedure in the above proof implies a similar upper bound $\big|G(\mathbb{F}_q) \cap X\big| \leq c' \cdot q^{\dim X}$. Combining this with Proposition 3.1, we obtain

$$\big|G(\mathbb{F}_q) \cap X\big| \;\leq\; c'' \cdot \big|G^\circ(\mathbb{F}_q)\big|^{\frac{\dim X}{\dim G}}, \tag{3.3}$$

where the constant c'' depends only on the families $\mathscr{G} \to \mathscr{S}$ and $\mathscr{X} \to \mathscr{T}$. We interpret this inequality as saying that the finite subgroup $G(\mathbb{F}_q)$ is *not concentrated* on any proper closed subvariety which belongs to a constructible family. In the next section we will generalize this to arbitrary sufficiently general finite subgroups in place of $G(\mathbb{F}_q)$.

Simple groups and Frobenius maps. A central role in this article is played by connected simple groups. In this paper, we use *simple* (for linear algebraic groups) in the strong sense of adjoint and absolutely simple. A group will be called *almost simple* if its center is finite and its quotient by its center is simple. To any simple root system Φ one can associate a natural constructible family of split connected simple linear algebraic groups $\mathscr{G} \to \mathbf{Spec}\,\mathbb{Z}$ with root system Φ (see [7, Exp.XXV]). It is necessarily adjoint; in fact, we will stick to adjoint groups as much as possible. Consider a geometric fiber $G = \mathscr{G}_s$ over a field of positive characteristic.

The set of fixed points of any endomorphism $F\colon G \to G$ will be denoted G^F. Any model G_0 of G over a finite field $\mathbb{F}_q$ with q elements corresponds to a so-called *standard Frobenius map* $\mathrm{Frob}_q\colon G \to G$. In local coordinates over $\mathbb{F}_q$ it is given by $x \mapsto x^q$, and its chief defining property is $G_0(\mathbb{F}_q) = G^{\mathrm{Frob}_q}$. An arbitrary isogeny $F\colon G \to G$ is called a *Frobenius map* if and only if some positive power is a standard Frobenius map. If $F^n = \mathrm{Frob}_q$, we set $q_F := \sqrt[n]{q}$. This is a positive real number which depends only on F. It plays the role of the cardinality of a finite field, even when it is an irrational number, as happens for Suzuki and Ree groups. The group of fixed points G^F is finite and is called a *finite group of Lie type.*

Simple groups of Lie type. Keeping the above notation, let m denote the index of the root lattice in the weight lattice of Φ, $\tilde{G}$ the simply connected covering group, and $\tilde{G} \to G$ the covering map. For later use we record some well-known facts (see [3, §11.1, §14.4], [4, §2.9]).

Theorem 3.4. *Assume* $q_F \geq 4$. *Then:*

(a) *The derived group* $(G^F)^{\mathrm{der}}$ *is nonabelian simple.*

(b) *The index* $\left[G^F : (G^F)^{\mathrm{der}}\right]$ *is* $\leq m$.

(c) *The kernel of* $\tilde{G} \to G$ *has order* $\leq m$.

(d) *We have* $(q_F - 1)^{\dim G} < \left|G^F\right| < q_F^{\dim G}$. *Moreover, the order of* G^F *is less than the cube of the order of one of its* p*-Sylow subgroups.*

Genericity. We now prove an analogue of Proposition 3.2 which includes Suzuki and Ree groups.

Proposition 3.5. *Let* $\mathscr{G} \to \mathbf{Spec}\,\mathbb{Z}$ *be the constructible family of connected adjoint groups associated to a simple root system* Φ*, and consider a constructible family of fiberwise nowhere dense algebraic subgroups* $\mathscr{H} \to \mathscr{T}$*. Then there exists a constant* q_0 *such that for any Frobenius map* F *on a geometric fiber* $G := \mathscr{G}_s$ *with* $q_F \geq q_0$ *the finite subgroup* $(G^F)^{\mathrm{der}}$ *is* $\mathscr{H}$*-general.*

Proof. By the classification of isogenies of simple algebraic groups, F is either a standard Frobenius map, or the composite of a fixed basic nonstandard isogeny with a standard Frobenius map. As in the proof of Proposition 3.2 we deduce that the intersection K of all F-power translates of $\mathscr{H}_t$ belongs to a constructible family of algebraic subgroups. This is an F-invariant proper algebraic subgroup, and it remains to show that the ratio $|(G^F)^{\mathrm{der}}| \,\big/\, |(K^\circ)^F|$ becomes arbitrarily large with q_F. By Theorem 3.4 this reduces to bounding $|(K^\circ)^F|$ from above. The following assertion suffices:

$$(\sqrt{q_F} - 1)^{2\dim K^\circ} \;\leq\; \left|(K^\circ)^F\right| \;\leq\; (\sqrt{q_F} + 1)^{2\dim K^\circ}.$$

It is proved with the same methods as Proposition 3.1. The details are left to the reader. □

The above proof required some caution, because the analogue of Proposition 3.5 for a general family of groups is false if one does not know that F is the composite of a standard Frobenius with an isogeny that varies in a constructible family. Indeed, suppose that $G = G_1 \times G_1$ and $F\colon (g, g') \mapsto (g', F_1(g))$, where F_1 is an arbitrary Frobenius map on G_1. Then F^2 is just F_1 on each factor, so F is a nonstandard Frobenius map, where $q_F = \sqrt{q_{F_1}}$ can become arbitrarily large. On the other hand, the fixed points of F are just the fixed points of F_1 on G_1, diagonally embedded into G. Thus G^F is not $\mathscr{H}$-general if $\mathscr{H}$ consists of the diagonal in G.

J. Tilouine pointed out to us that combining Proposition 3.5 with Theorem 0.5 yields the following corollary. It will not be used in the rest of this paper.

Corollary 3.6. *For every simple root system* Φ *there exists a constant* q_0 *with the following property. Consider a connected adjoint group* G *with simple root system* Φ *over an algebraically closed field of positive characteristic, and a finite subgroup* $\Gamma \subset G$*. Assume that there is a Frobenius map* $F_1\colon G \to G$ *with* $q_{F_1} \geq q_0$*, so that* $(G^{F_1})^{\mathrm{der}} \subset \Gamma$*. Then there exists a Frobenius map* $F\colon G \to G$ *so that*

$$\left(G^F\right)^{\mathrm{der}} \;\subset\; \Gamma \;\subset\; G^F\,.$$

Proof. Let $\mathscr{G} \to \mathbf{Spec}\,\mathbb{Z}$ be as above and $\mathscr{H} \to \mathscr{T}$ the constructible family of fiberwise nowhere dense algebraic subgroups implicit in Theorem 0.5. Let q_0 be given by Proposition 3.5. Then $(G^{F_1})^{\mathrm{der}}$ is $\mathscr{H}$-general, hence so is Γ, and the desired assertion follows from Theorem 0.5. □

4. Basic nonconcentration estimate

Consider an arbitrary constructible family of algebraic groups $\mathscr{G} \to \mathscr{S}$, and a geometric fiber $G := \mathscr{G}_s$. The aim of this section is to generalize the inequality (3.3) to arbitrary sufficiently general finite subgroups $\Gamma \subset G$. It may happen that a disproportionately large subgroup of Γ is contained in a proper normal algebraic subgroup $N \triangleleft G$. This is so, for instance, when G and N are defined over $\mathbb{F}_q$ and $\Gamma = G(\mathbb{F}_q) \cdot N(\mathbb{F}_{q^r})$ with r large. Thus a general analogue of the upper bound (3.3) can be expected only in terms of the following quantity. Set

$$q_\Gamma := \sup_N \, |\Gamma \cap N|^{\frac{1}{\dim N}}, \tag{4.1}$$

where N runs through all connected normal algebraic subgroups of $\mathscr{G}_s$. Clearly we have $q_\Gamma = |\Gamma|^{1/\dim \mathscr{G}_s}$ whenever $\mathscr{G}_s$ is connected and almost simple. The following theorem is the main result of this section.

Theorem 4.2. *Consider a constructible family of algebraic groups $\mathscr{G} \to \mathscr{S}$, and a constructible family of subvarieties $\mathscr{X} \to \mathscr{T}$. Then there exists a constant c such that for any sufficiently general finite subgroup Γ of a geometric fiber $\mathscr{G}_s$ and any point t of $\mathscr{T}$ above s we have*

$$\big|\Gamma \cap \mathscr{X}_t\big| \leq c \cdot q_\Gamma^{\dim \mathscr{X}_t}.$$

There is also a variant for Cartesian products:

Theorem 4.3. *Consider a constructible family of algebraic groups $\mathscr{G} \to \mathscr{S}$, a positive integer n, and a constructible family of subvarieties $\mathscr{X} \to \mathscr{T}$ of $\mathscr{G}^n \to \mathscr{S}$. Then there exists a constant c such that for any sufficiently general finite subgroup Γ of a geometric fiber $\mathscr{G}_s$ and any point t of $\mathscr{T}$ above s we have*

$$\big|\Gamma^n \cap \mathscr{X}_t\big| \leq c \cdot q_\Gamma^{\dim \mathscr{X}_t}.$$

The proof of these theorems will occupy the rest of this section.

Proof of Theorem 4.2: The idea. As an easy example let us consider an irreducible curve X in a connected algebraic group G of dimension r, and a sufficiently general finite subgroup $\Gamma \subset G$. Ideally, we would like to find elements $\gamma_1, \ldots, \gamma_{r-1} \in \Gamma$ such that the morphism of algebraic varieties

$$X^r \longrightarrow G,\ (x_1, \ldots, x_r) \mapsto x_1\gamma_1x_2\gamma_2 \cdots \gamma_{r-1}x_r$$

is dominant and quasi-finite. Suppose all its fibers contain $\leq n$ points. By counting points in Γ we deduce

$$\big|\Gamma \cap X\big|^r \leq n \cdot \big|\Gamma\big|.$$

This implies the desired estimate:

$$\big|\Gamma \cap X\big| \leq \sqrt[r]{n|\Gamma|} \leq \sqrt[r]{n} \cdot q_\Gamma.$$

In general, there are two technical problems with this method. First, it may not be possible to cover G by multiplying translates of X. In that case one can show that X is contained in a translate of a proper normal algebraic subgroup and use induction on $\dim G$. The second problem is that the morphism obtained by multiplying subvarieties in G may have fibers of positive and nonconstant dimension. A counting argument as above can still be made to work if the number of points of Γ in such fibers can be bounded. The bound we need here is of the same kind as the original statement. We therefore proceed by induction. Note that since the fibers of the multiplication morphism vary, one is forced to prove the theorem uniformly for a whole constructible family of subvarieties, even if one wants it only for a single X.

Reduction steps. We perform several reductions. First, since the number of irreducible components of $\mathscr{X}_t$ is bounded by Proposition 1.4, it suffices to establish the desired upper bound for the number of points in any one irreducible component. We can then replace the family $\mathscr{X} \to \mathscr{T}$ by that given by Proposition 1.9, that is, assume that $\mathscr{X}_t$ is irreducible. Next, if $\Gamma \cap \mathscr{X}_t$ is nonempty, e.g. contains some element γ, then its cardinality is equal to that of $\Gamma \cap \gamma^{-1}\mathscr{X}_t$. Here the translate $\gamma^{-1}\mathscr{X}_t$ is again a fiber of a constructible family, namely that of all translates $g^{-1}\mathscr{X}_t$, where $g \in \mathscr{G}_s$ lies in the same irreducible component as $\mathscr{X}_t$. Thus we are reduced to the case that all $\mathscr{X}_t$ are contained in the identity component of $\mathscr{G}_s$. After this reduction, we may also replace $\mathscr{G}$ by its identity component, i.e., assume that $\mathscr{G}_s$ is connected.

Using Proposition 1.6 we may stratify the base and assume that the dimensions of $\mathscr{G}_s$ and $\mathscr{X}_t$ are constant. We can then do induction on fiber dimensions. The outermost induction is on $\dim \mathscr{G}_s$, the next one on $d := \dim \mathscr{X}_t$. The theorem is obvious in the zero-dimensional case, since our fibers are already irreducible. So we assume $d > 0$.

For technical reasons, it will be convenient to modify the desired estimate by a certain defect δ and to consider two subvarieties at a time. Specifically, we will prove the following statement for every integer $\delta \geq 0$, while the family $\mathscr{G} \to \mathscr{S}$ is fixed:

Lemma 4.4. *Given constructible families of subvarieties $\mathscr{X} \to \mathscr{T}$ and $\mathscr{Y} \to \mathscr{U}$ of $\mathscr{G} \to \mathscr{S}$ there exists a constant c such that for any sufficiently general finite subgroup Γ of a geometric fiber $\mathscr{G}_s$, every point t of $\mathscr{T}$ above s, and every point u of $\mathscr{U}$ above s, we have*

$$\left|\Gamma \cap \mathscr{X}_t\right| \cdot \left|\Gamma \cap \mathscr{Y}_u\right| \leq c \cdot q_\Gamma^{\dim \mathscr{X}_t + \dim \mathscr{Y}_u + \delta}.$$

For $\delta \geq 2\dim \mathscr{G}_s$, this is automatically true by the definition of q_Γ. So we must prove Lemma 4.4 for fixed $\delta \geq 0$, assuming it for $\delta + 1$ in place of δ. For given δ, we prove the lemma by descending induction on $\dim \mathscr{Y}_u - \dim \mathscr{X}_t$. Without loss of generality, we may assume $\dim \mathscr{X}_t \leq \dim \mathscr{Y}_u$. The condition $\dim \mathscr{Y}_u - \dim \mathscr{X}_t \geq \dim \mathscr{G}_s$ implies that $\mathscr{X}_t$ is a single point and $\mathscr{Y}_u = \mathscr{G}_s$, and in this case the claim is obviously true for every $\delta \geq 0$. We abbreviate $G := \mathscr{G}_s$, $X := \mathscr{X}_t$, and $Y := \mathscr{Y}_u$.

Our induction scheme is nested as follows:

(A) Induction on $\dim G$;
(B) Descending induction on δ;
(C) Descending induction on $\dim Y - \dim X$.

Lemma 4.5. *Suppose that $Y \neq G$.*

(a) *The set of $g \in G$ with $\dim \overline{XgY} = \dim Y$ is a fiber of some constructible family of subvarieties of $\mathscr{G} \to \mathscr{S}$.*
(b) *If $\dim \overline{XgY} = \dim Y$ for every $g \in G$, then X is contained in a translate of a connected normal subgroup $N \lhd G$ of smaller dimension, which is a fiber of some constructible family of algebraic subgroups of $\mathscr{G} \to \mathscr{S}$.*

Proof. Choose any $x \in X$. Since X and Y are irreducible, the relation $\dim \overline{XgY} = \dim Y$ is equivalent to $XgY = xgY$. This in turn means that $g^{-1}x^{-1}Xg$ is contained in the normalizer M of Y for the left translation action of G on itself. By

Proposition 1.13, M belongs to a constructible family of proper algebraic subgroups, and the assumption $Y \neq G$ implies $M \neq G$. The inclusion $g^{-1}x^{-1}Xg \subset M$ is tantamount to the condition that g lies in the transporter from $x^{-1}X$ to M for the conjugation action of G on itself. By Proposition 1.13, this transporter belongs to a constructible family, which proves (a).

For (b), let $M^{\cap}$ denote the intersection of the conjugates gMg^{-1} for all $g \in G$. This is a proper normal subgroup of G. By Corollary 1.12, it belongs to a constructible family of algebraic subgroups. The same holds for the identity component $N := (M^{\cap})^{\circ}$, for instance by Proposition 1.9. Finally, the assumptions in (b) imply $x^{-1}X \subset N$, which proves the claim. □

In the situation of Lemma 4.5(b) we prove Lemma 4.4 as follows. Choose an element $x \in \Gamma \cap X$ if that set is nonempty. Then we have $|\Gamma \cap X| = |\Gamma \cap x^{-1}X|$, and $x^{-1}X \subset N$ belongs to a constructible family. By induction hypothesis (A), we know Theorem 4.2 already in that situation. Since $q_{\Gamma\cap N} \leq q_\Gamma$, Lemma 4.4 is proved in this case.

Thus whenever $Y \neq G$, we may suppose that there exists $g \in G$ with $\dim \overline{XgY} > \dim Y$. That is, in the constructible family of Lemma 4.5(a), we restrict ourselves to that part of the base where the fiber is a proper subvariety of G. Afterwards, if Γ is sufficiently general, by Proposition 2.4 we may choose $\gamma \in \Gamma$ with $\dim \overline{X\gamma Y} > \dim Y$. We set $Z := \overline{X\gamma Y}$ and consider the dominant morphism

$$\varphi\colon\ X \times Y \longrightarrow Z,\ (x,y) \mapsto x\gamma y.$$

Fiber dimensions. The fiber above any point $z \in Z$ is

$$\varphi^{-1}(z) = \left\{ (x, \gamma^{-1}x^{-1}z) \;\middle|\; x \in X \cap zY^{-1}\gamma^{-1} \right\}.$$

Thus its dimension is always $\leq \dim X$, and in the case of equality, the fiber is isomorphic to X. Recall that $\dim Z > \dim Y$. The generic fiber dimension of φ is therefore

$$\dim X + \dim Y - \dim Z < \dim X,$$

and by semicontinuity ([10, Thm. 13.1.3]), all nonempty fibers have dimension at least as great as the generic fiber dimension. For any integer

$$f \in [\dim X + \dim Y - \dim Z, \dim X],$$

let

$$Z_f := \left\{ z \in Z \;\middle|\; \dim \varphi^{-1}(z) = f \right\}.$$

This locally closed subset of Z is open if and only if $f = \dim X + \dim Y - \dim Z$. For all other values, the subset $\varphi^{-1}(Z_f)$ is nowhere dense in $X \times Y$, so its closure has dimension $< \dim X + \dim Y$. Thus for all f we have either

$$\dim \overline{Z_f} + f \leq \dim X + \dim Y - 1 \tag{4.6}$$

or

$$\begin{aligned} &\dim \overline{Z_f} + f = \dim X + \dim Y, \\ &\dim \overline{Z_f} - f = (\dim Z - \dim X) + (\dim Z - \dim Y) > \dim Y - \dim X. \end{aligned} \tag{4.7}$$

Note that by Proposition 1.7 the subvariety Z is a fiber of some constructible family, and by Proposition 1.8 the same is true for the closure $\overline{Z_f}$. For all $z \in Z$, the projection map pr_1 from $X \times Y$ to X induces an isomorphism from $\varphi^{-1}(z)$ to its projection $\mathrm{pr}_1(\varphi^{-1}(z))$, and it follows that the image closures $\overline{\mathrm{pr}_1(\varphi^{-1}(z))}$, as z ranges over Z, are fibers of some constructible family.

Counting arguments. Now we count the points $(x, y) \in (\Gamma \cap X) \times (\Gamma \cap Y)$ by means of the partition

$$(\Gamma \cap X) \times (\Gamma \cap Y) = \Gamma^2 \cap (X \times Y) = \bigcup_{f=\dim X+\dim Y-\dim Z}^{\dim X} \Gamma^2 \cap \varphi^{-1}(Z_f).$$

This gives

$$|\Gamma \cap X| \cdot |\Gamma \cap Y| = \sum_{f=\dim X+\dim Y-\dim Z}^{\dim X} |\Gamma^2 \cap \varphi^{-1}(Z_f)|.$$

The map φ sends $\Gamma^2 \cap \varphi^{-1}(Z_f)$ to $\Gamma \cap Z_f \subset \Gamma \cap \bar{Z}_f$, and each fiber is contained in a set of the form $\Gamma^2 \cap \varphi^{-1}(z)$, which maps injectively under the projection map pr_1 into

$$\Gamma \cap \mathrm{pr}_1(\varphi^{-1}(z)) \subset \Gamma \cap \overline{\mathrm{pr}_1(\varphi^{-1}(z))}.$$

The condition $z \in Z_f$ means that $\dim \overline{\mathrm{pr}_1(\varphi^{-1}(z))} = f$, and we claim

$$|\Gamma \cap Z_f| \cdot |\Gamma \cap \mathrm{pr}_1(\varphi^{-1}(z))| \leq c_f q_\Gamma^{\dim X+\dim Y+\delta},$$

for some constant c_f depending only on f and the original families $\mathscr{X}$ and $\mathscr{Y}$. Indeed, if (4.6) holds, this follows from induction hypothesis (B); and if (4.7) holds, it follows from induction hypothesis (C). Thus,

$$|\Gamma^2 \cap \varphi^{-1}(Z_f)| \leq |\Gamma \cap Z_f| \max_{z \in \Gamma \cap Z_f} |\Gamma \cap \mathrm{pr}_1(\varphi^{-1}(z))| \leq c_f q_\Gamma^{\dim X+\dim Y+\delta}.$$

Summing over f, we obtain

$$|\Gamma \cap X| \cdot |\Gamma \cap Y| \leq q_\Gamma^{\dim X+\dim Y+\delta} \sum_{f=\dim X+\dim Y-\dim Z}^{\dim X} c_f,$$

and Lemma 4.4 follows by induction.

Theorem 4.2 follows immediately by specializing Lemma 4.4 to the case $\delta = 0$, $\mathscr{T} = \mathscr{U}$, $\mathscr{X} = \mathscr{Y}$, and $t = u$. □

Proof of Theorem 4.3. As in the preceding proof, we abbreviate $G := \mathscr{G}_s$ and $X := \mathscr{X}_t$. Let $\pi\colon X \longrightarrow G^{n-1}$ denote the projection map obtained by forgetting the last factor in G^n. We count the points of $\Gamma^n \cap X$ by fibers of π, using induction on n. The case $n = 1$ is just Theorem 4.2.

The fibers of π form a constructible family of subvarieties of $\mathscr{G} \to \mathscr{S}$. Thus, if Γ is sufficiently general, by Theorem 4.2 for every $\underline{\gamma} = (\gamma_1, \ldots, \gamma_{n-1}) \in \Gamma^{n-1}$ we have

$$\left|\Gamma^n \cap \pi^{-1}(\underline{\gamma})\right| \leq c_1 \cdot q_\Gamma^{\dim \pi^{-1}(\underline{\gamma})}. \tag{4.8}$$

Next recall from Proposition 1.4 that $\dim G$ is bounded in the family $\mathscr{G} \to \mathscr{S}$, say by d. For every $0 \leq f \leq d$ put

$$Y_f := \left\{\, \underline{g} \in G^{n-1} \;\middle|\; \dim \pi^{-1}(\underline{g}) = f \,\right\}.$$

By Proposition 1.8, its Zariski-closure $\overline{Y_f}$ belongs to a constructible family of subvarieties of $\mathscr{G}^{n-1} \to \mathscr{S}$. Thus for sufficiently general Γ we have

$$\left|\Gamma^{n-1} \cap Y_f\right| \leq c_{n-1} \cdot q_\Gamma^{\dim \overline{Y_f}}. \tag{4.9}$$

By construction the constants c_1 and c_{n-1} depend only on the families $\mathscr{G} \to \mathscr{S}$ and $\mathscr{X} \to \mathscr{T}$. Now observe that

$$f + \dim \overline{Y_f} \ = \ \dim \overline{X \cap \pi^{-1}(Y_f)} \ \leq \ \dim X. \tag{4.10}$$

Thus we can calculate

$$\begin{aligned} \bigl|\Gamma^n \cap X\bigr| &= \sum_{f=0}^{d} \sum_{\underline{\gamma} \in \Gamma^{n-1} \cap Y_f} \bigl|\Gamma^n \cap \pi^{-1}(\underline{\gamma})\bigr| \\ &\overset{(4.8)}{\leq} \sum_{f=0}^{d} c_1 \cdot q_\Gamma^f \cdot \bigl|\Gamma^{n-1} \cap Y_f\bigr| \\ &\overset{(4.9)}{\leq} \sum_{f=0}^{d} c_1 \cdot c_{n-1} \cdot q_\Gamma^{f + \dim \overline{Y_f}} \\ &\overset{(4.10)}{\leq} (d+1) \cdot c_1 \cdot c_{n-1} \cdot q_\Gamma^{\dim X}, \end{aligned}$$

which is the desired assertion. □

5. Finite subgroups of abelian varieties

In this section we briefly detour to apply Theorem 4.2 to abelian varieties. The results here are not used in the rest of the paper. First we specialize everything to commutative groups. A simplification arises in this case from the fact that all subgroups are normal, and therefore in the definition (4.1) of q_Γ, the supremum is taken over all connected closed subgroups.

Theorem 5.1. *Consider a constructible family of commutative algebraic groups $\mathscr{G} \to \mathscr{S}$ and a constructible family of subvarieties $\mathscr{X} \to \mathscr{T}$. Then there exists a constant c such that for every finite subgroup Γ of a geometric fiber $\mathscr{G}_s$ and every point t of $\mathscr{T}$ above s we have*

$$\bigl|\Gamma \cap \mathscr{X}_t\bigr| \leq c \cdot q_\Gamma^{\dim \mathscr{X}_t}.$$

Proof. By Theorem 4.2 and Metadefinition 2.2, the desired conclusion holds unless Γ is contained in a fiber $\mathscr{H}_u$ of some constructible family of fiberwise nowhere dense algebraic subgroups $\mathscr{H} \to \mathscr{U}$ of $\mathscr{G} \to \mathscr{S}$. Let $\mathscr{Y} \longrightarrow \mathscr{V} := \mathscr{T} \times_{\mathscr{S}} \mathscr{U}$ be the constructible family of subvarieties of $\mathscr{H} \to \mathscr{U}$ which consists of all intersections $\mathscr{Y}_v := \mathscr{X}_t \cap \mathscr{H}_u$, where $v = (t,u) \in \mathscr{V}_s$. By induction on fiber dimension, we may suppose that the theorem holds already for $\mathscr{H}$ and $\mathscr{Y}$. In other words, we have

$$\bigl|\Gamma \cap \mathscr{Y}_v\bigr| \leq c \cdot (q'_\Gamma)^{\dim \mathscr{Y}_v},$$

where c is some constant, and q'_Γ is defined as in (4.1) except that the supremum is extended only over subgroups $N \subset \mathscr{H}_u$. Thus in the case $\Gamma \subset \mathscr{H}_u$ we deduce

$$\bigl|\Gamma \cap \mathscr{X}_t\bigr| = \bigl|\Gamma \cap \mathscr{Y}_v\bigr| \leq c \cdot (q'_\Gamma)^{\dim \mathscr{Y}_v} \leq c \cdot q_\Gamma^{\dim \mathscr{X}_t},$$

as desired. □

For abelian varieties a natural collection of finite subgroups is given by the n-torsion points $\mathscr{A}_s[n]$ for varying n:

Corollary 5.2. *Let $\mathscr{A} \to \mathscr{S}$ be a constructible family of abelian varieties, and $\mathscr{X} \to \mathscr{T}$ a constructible family of subvarieties. Then there exists a constant c such that for every positive integer n, every geometric fiber $\mathscr{A}_s$, and every point t of $\mathscr{T}$ above s we have*

$$\big|\mathscr{A}_s[n] \cap \mathscr{X}_t\big| \leq c \cdot (nn')^{\dim \mathscr{X}_t},$$

where n' is the largest divisor of n which is prime to the residue characteristic at s.

Proof. The connected algebraic subgroups $B \subset \mathscr{A}_s$ are precisely the abelian subvarieties. Thus we have

$$\big|\mathscr{A}_s[n] \cap B\big| = \big|B[n]\big| \leq (nn')^{\dim B}.$$

(See [24, §6] for the part prime to the characteristic, [24, §15] for the rest.) Thus formula (4.1) implies that $q_\Gamma = nn'$. The result follows from Theorem 5.1. □

R. Weissauer pointed out to us that this can also be proved using intersection theory, following the lines of the proof of [25, Prop. 7.7].

Remark. The bound in (3.3) is optimal for subvarieties that are defined over $\mathbb{F}_q$, hence so is the bound in Theorem 5.1. The bound in Corollary 5.2 cannot be improved either, as the following special case shows. Suppose that $\mathscr{A}_s$ is defined over a finite field $\mathbb{F}_q$ and isogenous to a product of supersingular elliptic curves all of whose endomorphisms are defined over $\mathbb{F}_q$. Then it is well known that

$$\mathscr{A}_s(\mathbb{F}_{q^m}) = \mathscr{A}_s[q^m - 1]$$

(cf. [31, Thm. 2(d)]), so we are back in the situation (3.3). We do not know if an improvement is possible in other cases or for other values of n.

6. Orders of conjugacy classes and centralizers

From here to the end of Section 11 we fix a simple root system Φ and let $\mathscr{G} \to \mathbf{Spec}\,\mathbb{Z}$ denote the family of split connected adjoint groups associated to Φ (see [7, Exp. XXV]). We will consider a geometric fiber $G = \mathscr{G}_s$ over an algebraically closed field k of characteristic $p \geq 0$ and a finite subgroup $\Gamma \subset G$. In any quantification of the form "for every sufficiently general Γ", the whole triple (k, G, Γ) is allowed to vary, with Γ being subject to Metadefinition 2.2. The assumption that G is adjoint is irrelevant in this section but will become convenient later on. Recall from (4.1) that in this case $q_\Gamma = |\Gamma|^{1/\dim G}$. We will often use the following reformulation of Proposition 2.5:

Proposition 6.1. *For any fixed constant c, if Γ is sufficiently general, we have $q_\Gamma > c$.*

In this section, we use the results of Section 4 to estimate the size of centralizers in Γ. The main observation is that Theorem 4.2 can be applied not only when $\mathscr{X}$ is the family of centralizers, but also when it is the family of conjugacy classes in $\mathscr{G}$. Thus although Theorem 4.2 gives only an upper bound in each case, the formula

$$\big|\Gamma_\gamma\big| \cdot \big|O_\Gamma(\gamma)\big| = \big|\Gamma\big|$$

implies a lower bound as well and thereby determines both factors to within a multiplicative constant. The following result generalizes this to centralizers of arbitrary subsets:

Theorem 6.2. *There is a constant c_0 depending only on Φ such that for any sufficiently general $\Gamma \subset G$ and any subset $\Lambda \subset \Gamma$ we have*

$$\frac{1}{c_0}\cdot q_\Gamma^{\dim G_\Lambda} \;\le\; \big|\Gamma_\Lambda\big| \;\le\; c_0\cdot q_\Gamma^{\dim G_\Lambda}.$$

Proof. By Theorem 1.10, it suffices to consider centralizers of subsets of cardinality $\le n$, where n depends only on the family $\mathscr{G}\to\mathbf{Spec}\,\mathbb{Z}$, that is, on Φ. So suppose $\Lambda=\{\gamma_1,\ldots,\gamma_m\}$ with $m\le n$. Setting $\gamma_i:=1$ for $m<i\le n$, the centralizer of Λ coincides with the stabilizer of the point $\underline{\gamma}:=(\gamma_1,\ldots,\gamma_n)$ for the diagonal conjugation action on G^n.

Consider the morphism

$$\mathscr{G}\times\mathscr{G}^n \longrightarrow \mathscr{G}^n\times\mathscr{G}^n,\ \big(g,(g_1,\ldots)\big)\mapsto\big((gg_1g^{-1},\ldots),(g_1,\ldots)\big).$$

This may be viewed as a morphism of families from $\mathscr{G}$ to $\mathscr{G}^n$ which is indexed by the second factor $\mathscr{G}^n$. The algebraic stabilizer in G of a point $\underline{g}\in G^n$ is a fiber of this morphism, so it belongs to a constructible family of subvarieties of $\mathscr{G}\to\mathbf{Spec}\,\mathbb{Z}$. On the other hand, the G-orbit of $\underline{g}$ is just the image of this map in the fiber above $\underline{g}$. Thus by Proposition 1.7, the orbit closures form a constructible family of subvarieties of $\mathscr{G}^n$.

Applying Theorem 4.2 to the family of centralizers we find a constant c_1 such that

$$\big|\Gamma_\Lambda\big| = \big|\Gamma_{\underline{\gamma}}\big| = \big|\Gamma\cap G_{\underline{\gamma}}\big| \le c_1\cdot q_\Gamma^{\dim G_{\underline{\gamma}}} \tag{6.3}$$

whenever Γ is sufficiently general. Similarly, applying Theorem 4.3 to the orbit closures $\overline{O_G(\underline{\gamma})}$ we find a constant c_2 such that

$$\big|O_\Gamma(\underline{\gamma})\big| \le \big|\Gamma^n\cap\overline{O_G(\underline{\gamma})}\big| \le c_2\cdot q_\Gamma^{\dim O_G(\underline{\gamma})} \tag{6.4}$$

whenever Γ is sufficiently general. Combining the second estimate with

$$\big|\Gamma_{\underline{\gamma}}\big|\cdot\big|O_\Gamma(\underline{\gamma})\big| \;=\; \big|\Gamma\big| \;=\; q_\Gamma^{\dim G} \;=\; q_\Gamma^{\dim G_{\underline{\gamma}}+\dim O_G(\underline{\gamma})}$$

we obtain

$$\big|\Gamma_\Lambda\big| \ge \frac{1}{c_2}\cdot q_\Gamma^{\dim G_{\underline{\gamma}}}. \tag{6.5}$$

Setting $c_0:=\sup\{c_1,c_2\}$, the theorem follows from (6.3) and (6.5). □

The constant c_0 of Theorem 6.2 will be fixed throughout the rest of the paper. The same kind of argument shows:

Theorem 6.6. *Let $\Gamma\subset G$ be as in Theorem 6.2. Then for every $\gamma\in\Gamma$, the intersection $\Gamma\cap O_G(\gamma)$ consists of at most c_0^2 conjugacy classes of Γ.*

Proof. We will use the estimates (6.3) and (6.4) in the case $n=1$. Without loss of generality, we may assume that γ is the element of $\Gamma\cap O_G(\gamma)$ whose Γ-conjugacy class is smallest. Thus, the total number of Γ-conjugacy classes in $\Gamma\cap O_G(\gamma)$ is no more than

$$\frac{|\Gamma\cap O_G(\gamma)|}{|O_\Gamma(\gamma)|} \;=\; \frac{|\Gamma_\gamma|\cdot|\Gamma\cap O_G(\gamma)|}{|\Gamma|} \;\le\; \frac{c_1c_2\cdot q_\Gamma^{\dim G_\gamma+\dim O_G(\gamma)}}{|\Gamma|} \;=\; c_1c_2 \;\le\; c_0^2. \;\square$$

Remark. Theorem 6.2 implies that Γ_Λ is a sufficiently general subgroup of the algebraic centralizer G_Λ whenever Γ is sufficiently general. Indeed, any constructible family of nowhere dense subvarieties $X \subset G_\Lambda$ can be viewed as a family of subvarieties of $\mathscr{G}$. Thus for suitable c we have

$$\bigl|\Gamma \cap X\bigr| \overset{\text{Thm. }4.2}{\leq} c \cdot q_\Gamma^{\dim X} \leq \frac{c}{q_\Gamma} \cdot q_\Gamma^{\dim G_\Lambda} \overset{\text{Thm. }6.2}{\leq} \frac{cc_0}{q_\Gamma} \cdot \bigl|\Gamma_\Lambda\bigr|.$$

On the other hand, $q_\Gamma > cc_0$ whenever Γ is sufficiently general, by Proposition 6.1. Therefore $\Gamma_\Lambda \not\subset X$, as desired. This behavior allows induction arguments and will be exploited in the following section.

7. Regular semisimple and unipotent elements

Let Φ, $\mathscr{G}$, and $\Gamma \subset G = \mathscr{G}_s$ be as in the preceding section. An element of Γ will be called *semisimple, unipotent, regular*, etc. if and only if it has this property as an element of G. The set of all regular semisimple elements of a subset X is denoted X^{rss}, the set of all unipotent elements X^{un}. Since G^{rss} is open and dense in G, it follows easily from Theorem 4.2 that most elements of a sufficiently general finite subgroup are regular semisimple. It is more difficult to find elements of other types.

It is well known that any finite group of Lie type contains a regular unipotent element ([4, Prop. 5.1.7], [18, §8.4]). In this section, we prove the same assertion for every sufficiently general finite subgroup $\Gamma \subset G$. The idea is to count the elements of Γ in a particular way, breaking them up via their Jordan decomposition and using induction on centralizers of semisimple elements with the help of Theorem 6.2. The resulting formula shows that the number of unipotent elements in Γ is so large that some of them must be regular unipotent.

For finite groups of Lie type such computations were carried out by Steinberg [30, §§14–15] (see also [4, Thm. 3.4.1, Thm. 6.6.1], or [18, Thms. 8.8, 8.14]). Namely, suppose that G lives in positive characteristic, and $F : G \to G$ is a Frobenius map. In two separate calculations, Steinberg shows that the number of unipotent elements in G^F and the number of F-stable maximal tori of G are each equal to the square of the order of a maximal unipotent subgroup of G^F. A more direct proof that the former quantities are equal was given by Lehrer [23, Cor. 1.13]. Our argument resembles Lehrer's approach.

Toric subsets and centralizers. For convenience we call a subset of G *toric* if and only if it is contained in an algebraic torus of G. Any toric subset consists of pairwise commuting semisimple elements, but the converse is not true in general. The induction argument in Theorem 7.8 below employs reduction to the identity components G_Λ° of the centralizers of toric subsets $\Lambda \subset \Gamma$. By construction, each G_Λ° contains a maximal torus of G.

Proposition 7.1. *For any toric subset $\Lambda \subset G$ we have:*

(a) $\Lambda \subset G_\Lambda^\circ$.
(b) *For any semisimple element $s \in G_\Lambda^\circ$ the set $\Lambda \cup \{s\}$ is toric.*
(c) *Any unipotent element of G_Λ lies in G_Λ°.*

Proof. By Noetherian induction, every toric subset Λ has a finite subset Λ' such that $G_\Lambda = G_{\Lambda'}$. Without loss of generality, therefore, we may assume Λ is finite. If $T \subset G$ is a maximal torus containing Λ, we have $\Lambda \subset T \subset G_\Lambda^\circ$, whence (a). Next by induction on $|\Lambda|$ we see that G°_Λ is reductive, since the connected centralizer of a semisimple element in any reductive group is reductive. Thus the center of G°_Λ, and hence Λ itself, is contained in every maximal torus of G°_Λ. As any semisimple element of G°_Λ lies in a maximal torus, this implies (b). Finally, (c) is the assertion of [18, §1.12] if Λ consists of one element. The proof given there applies to all connected reductive groups, so the general case of (c) follows again by induction on Λ. □

Regular semisimple elements. To simplify notation we will abbreviate $\Gamma^\circ_\Lambda := \Gamma \cap G^\circ_\Lambda$.

Proposition 7.2. *Fix any* $0 < \varepsilon < 1$. *If* Γ *is sufficiently general, then for every toric subset* $\Lambda \subset \Gamma$ *we have*

$$1-\varepsilon \;\leq\; \frac{|(\Gamma^\circ_\Lambda)^{\mathrm{rss}}|}{|\Gamma^\circ_\Lambda|} \;\leq\; 1.$$

Proof. Let $g \mapsto \mathrm{Ad}_g$ denote the adjoint representation of $\mathscr{G}$. It is well known that g is regular semisimple if and only if the multiplicity of 1 as a zero of the characteristic polynomial of Ad_g is minimal, i.e., equal to the rank of Φ. This is a Zariski-open condition, so the complement $\mathscr{G} \smallsetminus \mathscr{G}^{\mathrm{rss}}$ is a constructible family of proper closed subvarieties of $\mathscr{G}$.

By Proposition 1.14 the algebraic centralizers G_Λ form a constructible family of algebraic subgroups of $\mathscr{G}$. So by Proposition 1.9 the same is true for their identity components. By construction these are irreducible and contain regular semisimple elements; hence $G^\circ_\Lambda \smallsetminus G^{\mathrm{rss}}$ belongs to a constructible family of subvarieties of strictly smaller dimension. Thus for suitable c we have

$$\left|\Gamma^\circ_\Lambda \smallsetminus G^{\mathrm{rss}}\right| \;\overset{\text{Thm. 4.2}}{\leq}\; c\cdot q_\Gamma^{\dim(G^\circ_\Lambda \smallsetminus G^{\mathrm{rss}})} \;\leq\; \frac{c}{q_\Gamma}\cdot q_\Gamma^{\dim G^\circ_\Lambda} \;\overset{\text{Thm. 6.2}}{\leq}\; \frac{cc_0}{q_\Gamma}\cdot\left|\Gamma^\circ_\Lambda\right|,$$

provided Γ is sufficiently general. This implies that

$$\left|(\Gamma^\circ_\Lambda)^{\mathrm{rss}}\right| \geq \Big(1-\frac{cc_0}{q_\Gamma}\Big)\cdot\left|\Gamma^\circ_\Lambda\right|.$$

Since q_Γ may be assumed arbitrarily large by Proposition 6.1, the desired inequality follows. □

Maximal toric subgroups. We call a subgroup of Γ°_Λ *maximal toric in* Γ°_Λ if it is maximal among the toric subgroups of Γ°_Λ.

Proposition 7.3. *Fix any* $0 < \varepsilon < 1$, *and suppose that* Γ *is sufficiently general. Consider any toric subset* $\Lambda \subset \Gamma$ *and any maximal toric subgroup* $\Theta \subset \Gamma^\circ_\Lambda$. *Then we have*

$$1-\varepsilon \;\leq\; \frac{|\Theta^{\mathrm{rss}}|}{|\Theta|} \;\leq\; 1.$$

In particular Θ *lies in a unique maximal torus of* G *and is a maximal toric subgroup of* Γ.

Proof. By assumption Θ is contained in some maximal torus $T \subset G^\circ_\Lambda$. Since Λ lies in the center of G°_Λ by Proposition 7.1 (a), it is also contained in T. Thus $\Lambda \cup \Theta$ generates a toric subgroup of Γ°_Λ, and the maximality of Θ implies $\Lambda \subset \Theta$.

Next we apply Proposition 7.2 to Θ in place of Λ. It follows that Γ°_Θ contains a regular semisimple element γ. Proposition 7.1 (b) shows that $\Theta \cup \{\gamma\}$ generates a toric subgroup of $\Gamma^\circ_\Theta \subset \Gamma^\circ_\Lambda$, so the maximality of Θ implies $\gamma \in \Theta$. Thus Θ contains a regular semisimple element, and its connected centralizer G_Θ° is a maximal torus of G. Therefore Γ_Θ° is a toric subgroup containing Θ. Again by maximality it must be equal to Θ. This implies the last two assertions, and the estimate is precisely that of Proposition 7.2 for Θ in place of Λ. □

The fact that most elements are regular semisimple can be used to count the number of maximal toric subgroups, as follows. Let Tor_Λ denote the set of all maximal toric subgroups of Γ_Λ°. Let $\mathrm{Tor}_\Lambda^\natural \subset \mathrm{Tor}_\Lambda$ be a system of representatives of Γ_Λ°-conjugacy classes.

Proposition 7.4. *Assume that* Γ *is sufficiently general. Then for any toric subset* $\Lambda \subset \Gamma$ *we have*

$$\sum_{\Theta \in \mathrm{Tor}_\Lambda^\natural} \frac{1}{[N_{\Gamma_\Lambda^\circ}(\Theta) : \Theta]} \;=\; 1.$$

Proof. Since every regular semisimple element of Γ_Λ° lies in a unique maximal toric subgroup $\Theta \subset \Gamma_\Lambda^\circ$, we can count them by looking at all maximal toric subgroups in turn. We find

$$\begin{aligned} \frac{|(\Gamma_\Lambda^\circ)^{\mathrm{rss}}|}{|\Gamma_\Lambda^\circ|} &= \sum_{\Theta \in \mathrm{Tor}_\Lambda} \frac{|\Theta^{\mathrm{rss}}|}{|\Gamma_\Lambda^\circ|} \\ &= \sum_{\Theta \in \mathrm{Tor}_\Lambda^\natural} \frac{|\Theta^{\mathrm{rss}}|}{|N_{\Gamma_\Lambda^\circ}(\Theta)|} \\ &= \sum_{\Theta \in \mathrm{Tor}_\Lambda^\natural} \frac{1}{[N_{\Gamma_\Lambda^\circ}(\Theta) : \Theta]} \cdot \frac{|\Theta^{\mathrm{rss}}|}{|\Theta|}. \end{aligned}$$

For any constant $0 < \varepsilon < 1$, combining the preceding calculation with Propositions 7.2 and 7.3, we find

$$1 - \varepsilon \;\le\; \sum_{\Theta \in \mathrm{Tor}_\Lambda^\natural} \frac{1}{[N_{\Gamma_\Lambda^\circ}(\Theta) : \Theta]} \;\le\; \frac{1}{1 - \varepsilon} \tag{7.5}$$

for any sufficiently general Γ. On the other hand, every Θ is contained in a unique maximal torus $T \subset G$; hence $N_{\Gamma_\Lambda^\circ}(\Theta)/\Theta$ is a subgroup of the associated Weyl group $N_G(T)/T$. The order m of this Weyl group is fixed, and we deduce

$$\sum_{\Theta \in \mathrm{Tor}_\Lambda^\natural} \frac{1}{[N_{\Gamma_\Lambda^\circ}(\Theta) : \Theta]} \;\in\; \frac{1}{m} \cdot \mathbb{Z}. \tag{7.6}$$

Taking $0 < \varepsilon < \frac{1}{m+1}$, the desired equality follows from (7.5) and (7.6). □

Corollary 7.7. *Under the hypotheses of Proposition* 7.4 *we have*

$$\sum_{\Theta \in \mathrm{Tor}_\Lambda} |\Theta| \;=\; |\Gamma_\Lambda^\circ|.$$

Proof.

$$\sum_{\Theta \in \mathrm{Tor}_\Lambda} |\Theta| \;=\; \sum_{\Theta \in \mathrm{Tor}_\Lambda^\natural} \frac{|\Gamma_\Lambda^\circ|}{|N_{\Gamma_\Lambda^\circ}(\Theta)|} \cdot |\Theta| \;=\; |\Gamma_\Lambda^\circ| \cdot \sum_{\Theta \in \mathrm{Tor}_\Lambda^\natural} \frac{1}{[N_{\Gamma_\Lambda^\circ}(\Theta) : \Theta]} \;\overset{\text{Prop. 7.4}}{=}\; |\Gamma_\Lambda^\circ|.$$

□

Unipotent elements. Now we are in a position to give a precise formula for the number of unipotent elements in any sufficiently general Γ. The induction procedure forces us to prove the analogue for all Γ°_Λ as well.

Theorem 7.8. *Assume that Γ is sufficiently general. Then for any toric subset $\Lambda \subset \Gamma$ the number of unipotent elements in Γ°_Λ is equal to the number of maximal toric subgroups in Γ°_Λ.*

Proof. We use induction on $\dim G^\circ_\Lambda$. The starting point is the case that G°_Λ is a maximal torus. Here the assertion is obvious, because a toric subgroup contains precisely one unipotent element, namely the identity. So assume that the assertion is known in dimension $< \dim G^\circ_\Lambda$. Then it holds with $\Lambda \cup \{\gamma\}$ in place of Λ, for any semisimple element $\gamma \in \Gamma^\circ_\Lambda$ outside the center $Z(G^\circ_\Lambda)$.

For any element $g \in G^\circ_\Lambda$ consider the Jordan decomposition $g = su$. If g is in Γ°_Λ, so are s and u, for the following reason. Recall that Γ°_Λ is a finite group. Thus if the base field has characteristic zero, the unipotent part u must be trivial, and both $u = 1$ and $s = g$ are in Γ°_Λ. In characteristic $p > 0$ the Jordan decomposition coincides with the decomposition into the prime-to-p part and the p-part inside Γ°_Λ.

Now we count the elements of Γ°_Λ by separating their semisimple and unipotent parts, in the following way. The second equality uses Proposition 7.1 (b) and (c) and the induction hypothesis:

$$\begin{aligned}
\left|\Gamma^\circ_\Lambda\right| &= \left|\Gamma^\circ_\Lambda \cap Z(G^\circ_\Lambda)\right| \cdot \left|(\Gamma^\circ_\Lambda)^{\mathrm{un}}\right| + \sum_{\substack{s \in \Gamma^\circ_\Lambda \smallsetminus Z(G^\circ_\Lambda)\\ \text{semisimple}}} \left|\Gamma^{\mathrm{un}}_{\Lambda\cup\{s\}}\right| \\
&= \left|\Gamma^\circ_\Lambda \cap Z(G^\circ_\Lambda)\right| \cdot \left|(\Gamma^\circ_\Lambda)^{\mathrm{un}}\right| + \sum_{\substack{s \in \Gamma^\circ_\Lambda \smallsetminus Z(G^\circ_\Lambda)\\ \text{semisimple}}} \left(\sum_{\Theta \in \mathrm{Tor}_{\Lambda\cup\{s\}}} 1\right) \\
&= \left|\Gamma^\circ_\Lambda \cap Z(G^\circ_\Lambda)\right| \cdot \left|(\Gamma^\circ_\Lambda)^{\mathrm{un}}\right| + \sum_{\Theta \in \mathrm{Tor}_\Lambda} \left(\sum_{s \in \Theta \smallsetminus Z(G^\circ_\Lambda)} 1\right) \\
&= \left|\Gamma^\circ_\Lambda \cap Z(G^\circ_\Lambda)\right| \cdot \left|(\Gamma^\circ_\Lambda)^{\mathrm{un}}\right| + \sum_{\Theta \in \mathrm{Tor}_\Lambda} \Big(\left|\Theta\right| - \left|\Theta \cap Z(G^\circ_\Lambda)\right|\Big) \\
&= \left(\sum_{\Theta \in \mathrm{Tor}_\Lambda} \left|\Theta\right|\right) + \left|\Gamma^\circ_\Lambda \cap Z(G^\circ_\Lambda)\right| \cdot \left(\left|(\Gamma^\circ_\Lambda)^{\mathrm{un}}\right| - \sum_{\Theta \in \mathrm{Tor}_\Lambda} 1\right).
\end{aligned}$$

By Corollary 7.7 the first summand on the right-hand side equals the left-hand side. Thus

$$\left|(\Gamma^\circ_\Lambda)^{\mathrm{un}}\right| = \sum_{\Theta \in \mathrm{Tor}_\Lambda} 1,$$

as desired. □

Regular unipotent elements. An element of G is unipotent if and only if its characteristic polynomial in any given faithful representation is a power of $X - 1$. Clearly this is a Zariski-closed condition in any family, so the set of unipotent elements $\mathscr{G}^{\mathrm{un}}$ is a constructible family of subvarieties of $\mathscr{G}$. It is fiberwise irreducible of dimension $\dim G - \operatorname{rank} G$, so Theorem 4.2 implies

$$\left|\Gamma^{\mathrm{un}}\right| \le c \cdot q_\Gamma^{\dim G - \operatorname{rank} G}$$

if Γ is sufficiently general, where c is a constant depending only on Φ. Theorem 7.8 implies the corresponding lower bound:

Proposition 7.9. *For any sufficiently general Γ we have*

$$\left|\Gamma^{\mathrm{un}}\right| \geq \frac{1}{c_0} \cdot q_\Gamma^{\dim G - \operatorname{rank} G}.$$

Proof.

$$\begin{aligned}
\left|\Gamma^{\mathrm{un}}\right| &\overset{\text{Thm. 7.8}}{=} \sum_{\Theta \in \mathrm{Tor}_\emptyset} 1 \\
&= \sum_{\Theta \in \mathrm{Tor}_\emptyset^\natural} \frac{|\Gamma|}{|N_\Gamma(\Theta)|} \\
&= \sum_{\Theta \in \mathrm{Tor}_\emptyset^\natural} \frac{q_\Gamma^{\dim G}}{[N_\Gamma(\Theta) : \Theta] \cdot |\Theta|} \\
&\overset{\text{Thm. 6.2}}{\geq} \sum_{\Theta \in \mathrm{Tor}_\emptyset^\natural} \frac{1}{[N_\Gamma(\Theta) : \Theta]} \cdot \frac{q_\Gamma^{\dim G}}{c_0 \cdot q_\Gamma^{\operatorname{rank} G}} \\
&\overset{\text{Prop. 7.4}}{=} \frac{1}{c_0} \cdot q_\Gamma^{\dim G - \operatorname{rank} G}.
\end{aligned}$$

□

Corollary 7.10. *Any sufficiently general finite subgroup $\Gamma \subset G$ contains a regular unipotent element.*

Proof. It is well known (see, e.g., [18, Theorem 4.13]) that the nonregular unipotent elements of a semisimple group form a nowhere dense closed subset of the set of all unipotent elements. It follows from Proposition 1.8 that the set of nonregular unipotent elements is a constructible family of subvarieties of $\mathscr{G}^{\mathrm{un}}$ of dimension $< \dim G - \operatorname{rank} G$. Thus, by Theorem 4.2 and Proposition 7.9 the number of nonregular unipotent elements of Γ is less than or equal to

$$c \cdot q_\Gamma^{\dim G - \operatorname{rank} G - 1} \;\leq\; \frac{cc_0}{q_\Gamma} \cdot \left|\Gamma^{\mathrm{un}}\right|$$

for some constant c. On the other hand, $q_\Gamma > cc_0$ for sufficiently general Γ, by Proposition 6.1. □

Corollary 7.11. *If $\Gamma \subset G$ is a sufficiently general finite subgroup, the characteristic of the base field of G divides $|\Gamma|$. In particular it is nonzero.*

Proof. The order of any nontrivial unipotent element is a power of the characteristic. □

Remark: Jordan's theorem. At this point, we have the main ingredients necessary to reprove Jordan's classical Theorem 0.1 in a purely algebraic manner. In fact, we can prove a slightly more general result, covering all finite subgroups of $\mathrm{GL}_n(k)$ whose order is not divisible by the characteristic of k. We remark that Jordan's original proof bears some resemblance to ours. Consider a finite subgroup $\Gamma \subset \mathrm{GL}_n(k)$ of order not divisible by $p = \mathrm{char}(k) \geq 0$. Suppose that Γ is contained in a connected algebraic subgroup $H \subset \mathrm{GL}_{n,k}$. If H possesses a simple factor group G, we can identify G with a fiber of the above constructible family $\mathscr{G} \to \mathbf{Spec}\,\mathbb{Z}$, so by Corollary 7.11 the image of Γ in G is contained in an algebraic subgroup $\mathscr{H}_t \subsetneqq G$ which belongs to a constructible family. By Proposition 1.4 the number of connected components of $\mathscr{H}_t$ is bounded. Thus after replacing H by the identity component of the inverse image of $\mathscr{H}_t$, and Γ by a subgroup of bounded index, we have decreased the dimension of H. After fewer than n^2 such steps, the group H is connected solvable, and the rest of the argument is straightforward. For more details, see Section 12.

8. Minimal unipotent elements

Let Φ, $\mathscr{G}$, and $\Gamma \subset G = \mathscr{G}_s$ be as before. The regular unipotent elements of G lie at one end of the range of all types of unipotent elements. At the other end we find the identity, followed by the elements of the center of the unipotent radical of a Borel subgroup of G. For the purposes of this section the latter elements are called *minimal unipotent*. In most cases, they lie in a canonical one-parameter additive subgroup associated to a root of G.

In this section we begin with a regular unipotent element of Γ, whose existence is guaranteed by Corollary 7.10, and manufacture minimal unipotent elements in Γ using centralizers and maximal toric subgroups. We will show that Γ contains a sufficiently large subgroup which consists purely of minimal unipotent elements, and which is (in a natural way) a vector space of dimension one over a finite field. This fact will be exploited later on.

In the following, we consider a Borel subgroup $B \subset G$ and its unipotent radical U. Recall that k denotes the base field of G, and p its characteristic.

Structure of the unipotent radical. Choose any maximal torus $T \subset B$, so that the root system of G with respect to T is Φ. Let $\Delta \subset \Phi^+ \subset \Phi$ denote the subsets of simple, resp. positive, roots for the given Borel subgroup B. To every root $\alpha \in \Phi$ there is associated a *root group* $\mathbb{G}_{a,k} \cong U_\alpha \subset G$ on which T acts through the character α. It is known that every algebraic subgroup of U which is normalized by T is a product of root groups, which is a direct product of algebraic varieties ([17, Prop. 28.1]). In particular, we have

$$U = \prod_{\alpha \in \Phi^+} U_\alpha, \tag{8.1}$$

the product being taken in any order.

Proposition 8.2. *The commutator subgroup of U is*

$$U^{\mathrm{der}} = \prod_{\alpha \in \Phi^+ \smallsetminus \Delta} U_\alpha.$$

Proof. Let U' denote the right-hand side of this equality. The commutator of U_β and U_γ for any two nonproportional roots lies in the product of all $U_{i\beta+j\gamma}$, where i and j are positive integers such that $i\beta + j\gamma$ is a root. The precise formulas are well known; see, for instance, [17, Props. 33.3 (b), 33.4 (b), 33.5 (b)]. It follows immediately that U' is a subgroup of U which contains U^{der}.

To prove equality consider any nonsimple positive root α. Choose a simple root β such that $\alpha - \beta$ is also a positive root. The root system $\Psi := \Phi \cap (\mathbb{Z}\alpha \oplus \mathbb{Z}\beta)$ is then irreducible of rank two. The simple roots in $\Psi^+ := \Psi \cap \Phi^+$ are β and another root γ, and α is a linear combination of these with positive integral coefficients.

Now the formulas of [loc. cit.] show that the commutators $[U_\beta, U_\gamma]$ have nontrivial components in U_α. Since everything is normalized by T, it follows that $U_\alpha \subset U^{\mathrm{der}}$, as desired. □

In the following, we will call (p, Φ) *nonstandard* whenever Φ possesses roots of different lengths whose square length ratio is p. Otherwise it is called *standard.* By the classification of root systems, the nonstandard cases are precisely $(p, \Phi) = (2, B_n)$ and $(2, C_n)$ for $n \geq 2$, as well as $(2, F_4)$ and $(3, G_2)$. Now we can describe the center of U:

Proposition 8.3. *In the standard case we have $Z(U) = U_\alpha$, where α is the highest positive root. In the nonstandard case we have $Z(U) = U_{\alpha_\ell} U_{\alpha_s}$, where α_ℓ is the highest long root and α_s the highest short root.*

Proof. Since $Z(U)$ is a characteristic subgroup of U, it is normalized by T and therefore a product of root groups. Thus we must find all positive roots α such that U_α commutes with U_β for all $\beta \in \Phi^+$. These subgroups always commute when $\alpha + \beta$ is not a root. In particular, we have $U_\alpha \subset Z(U)$ whenever α is the highest positive root.

Suppose that α is not the highest positive root, and consider $\beta \in \Phi^+$ such that $\alpha + \beta$ is a root. Then the root system $\Psi := \Phi \cap (\mathbb{Z}\alpha \oplus \mathbb{Z}\beta)$ is irreducible of rank two. The formulas [17, Props. 33.3 (b), 33.4 (b), 33.5 (b)] show that U_α and U_β commute in the given characteristic p if and only if $|\alpha + \beta|^2 = p \cdot |\alpha|^2 = p \cdot |\beta|^2$. In particular, we must have $(p, \Psi) = (2, B_2)$ or $(3, G_2)$. In that case, moreover, U_α commutes with U_γ for all $\gamma \in \Psi^+ := \Psi \cap \Phi^+$ if and only if α is the highest short root in Ψ.

This rank two analysis shows that there is another candidate for α only when (p, Φ) is nonstandard, and that α must be short. If it is not the highest short root in Φ, then there exists a simple root β with $(\alpha, \beta) < 0$, so that $s_\beta(\alpha) \succ \alpha$ is a higher short root in Ψ. From the rank two case we then know that U_α is not in the center. By contrast, the highest short root in Φ is also the highest short root in Ψ for any β as above. Thus its root group is in the center of U, as desired. □

Normalizers and centralizers. We will need the following information on normalizers and centralizers of minimal unipotent elements. Clearly the nontrivial elements of any root group form a single orbit under T. In the nonstandard case, the fact that α_ℓ and α_s are nonproportional implies that the elements with nontrivial component in both U_{α_ℓ} and U_{α_s} form the unique open T-orbit in $Z(U)$. In all cases, the nontrivial B-invariant subgroups of $Z(U)$ are precisely $Z(U)$ and the root groups inside $Z(U)$.

Proposition 8.4. *Consider a nontrivial B-invariant subgroup $V \subset Z(U)$.*

(a) *The normalizer $N_G(V)$ is a parabolic subgroup of G, and its action on V factors through multiplicative characters corresponding to the roots occurring in V. Its orbits on V therefore coincide with the orbits under T.*

In the following let v denote any element of the open orbit in V.

(b) *The centralizer G_v is connected and equal to the centralizer G_V.*
(c) *The centralizer of G_V in G is equal to V.*
(d) *Any element $g \in G$ with $gvg^{-1} \in V$ lies already in $N_G(V)$.*

Proof. (a) Since $N_G(V)$ contains B, it is a parabolic subgroup of G. In particular it is connected. The kernel of the conjugation action $N_G(V) \to \mathrm{Aut}(V)$ contains the unipotent radical of B, which is also a maximal connected unipotent subgroup of $N_G(V)$. Thus the image of this homomorphism is a connected linear algebraic group without nontrivial connected unipotent subgroups. It is therefore a torus. The corresponding characters of $N_G(V)$ are uniquely determined by their restrictions to T, which are precisely the roots occurring in V. This proves (a).

For (b) recall that U is a maximal connected unipotent subgroup of G. Since it lies inside G_v, it is also a maximal connected unipotent subgroup of G_v. Its normalizer in G is B, so its normalizer in G_v is $B \cap G_v$. Note that the identity component of this intersection is a Borel subgroup of G_v. The fact that any two maximal connected unipotent subgroups of G_v are conjugate under G_v° implies

$$G_v = (B \cap G_v) \cdot G_v^\circ.$$

Now $B \cap G_v = U \cdot T_v$, and T_v is just the intersection of the kernels of the roots occurring in V. It therefore centralizes V.

We claim T_v is connected. Indeed, if $\{\alpha_1, \ldots, \alpha_k\}$ are roots of any adjoint semisimple group G with respect to any maximal torus T, all belonging to the same base, then the intersection of the $\ker \alpha_i$ is connected. This statement is equivalent to the claim that the intersection of the root lattice with the $\mathbb{Q}$-span of $\{\alpha_1, \ldots, \alpha_k\}$ coincides with the $\mathbb{Z}$-span of the same set of roots. Suppose $\beta = a_1\alpha_1 + \cdots + a_k\alpha_k$ lies in the root lattice. For each i, there exists a fundamental dominant coweight α_i^* such that $\alpha_i^*(\alpha_j)$ is 0 or 1 if $i \neq j$ or $i = j$, respectively. Moreover $\alpha_i^*(\beta) \in \mathbb{Z}$. Thus, $a_i \in \mathbb{Z}$. From this general statement, it is clear that if $k = 1$, $\ker \alpha_1$ is connected. In the nonstandard case, it suffices to observe that there exists a base to which α_s and $-\alpha_l$ both belong.

Thus $B \cap G_v$ is connected and equal to $B \cap G_V$. On the one hand this shows that G_v is connected and that $B \cap G_v$ is a Borel subgroup of G_v. On the other hand it shows that V commutes with a Borel subgroup of G_v, and therefore with all of G_v (see [17, Prop. 21.4A]). This proves (b).

For (c) we first determine G_U. To begin with, note that G_U is obviously contained in $N_G(U) = B$. Next, the action of $B = TU$ on U/U^{der} factors through a faithful action of T. Therefore $G_U \subset U$. As the centralizer of any group in itself is just its center, we find $G_U = Z(U)$. Since $G_V \supset U$, this equality implies $G_{G_V} \subset Z(U)$. If $V \neq Z(U)$, and α denotes the root of T on V, the kernel of α acts nontrivially on the other root group in $Z(U)$. This shows that $G_{G_V} \subset V$. The reverse inclusion holds automatically, which proves (c).

Finally, if $gvg^{-1} \in V$, assertion (b) implies $G_v = G_V \subset G_{gvg^{-1}} = gG_vg^{-1}$. This inclusion must be an equality, so g normalizes $G_v = G_V$. Thus by (c) it normalizes V, which proves (d). □

Regular elements in a Borel subgroup. Now we fix a regular unipotent element $u \in \Gamma$, whose existence is guaranteed by Corollary 7.10, and assume $u \in B$. By counting regular unipotent elements in $\Gamma \cap U$ we will find sufficiently many regular semisimple elements in $\Gamma \cap B$. We begin with the following abstract estimate:

Lemma 8.5. *For every positive integer r there is a constant $0 < \varepsilon_r \leq 1$ with the following property. Consider any finite group A and subgroups $A_1, \ldots, A_r$ whose* *union is not A. Then*

$$\big|A \smallsetminus (A_1 \cup \ldots \cup A_r)\big| \;\geq\; \varepsilon_r \cdot \big|A\big|.$$

Proof. Let $\varepsilon_1 := 1/2$ and $\varepsilon_r := \big(\varepsilon_{r-1}/2\big)^r$ for every $r \geq 2$. The lemma is obvious for $r = 1$, so we proceed by induction. Without loss of generality we may suppose that A_r is the smallest of the given subgroups. If $|A_r| \leq (\varepsilon_{r-1}/2) \cdot |A|$, then

$$\begin{aligned}\big|A \smallsetminus (A_1 \cup \ldots \cup A_r)\big| &\geq \big|A \smallsetminus (A_1 \cup \ldots \cup A_{r-1})\big| - \big|A_r\big| \\ &\geq \Big(\varepsilon_{r-1} - \frac{\varepsilon_{r-1}}{2}\Big) \cdot \big|A\big| \\ &\geq \varepsilon_r \cdot \big|A\big|\end{aligned}$$

by the induction hypothesis. Otherwise we have

$$\begin{aligned}\big[A : A_1 \cap \ldots \cap A_r\big] &\leq \big[A : A_1\big] \cdots \big[A : A_r\big] \\ &\leq \Big(\frac{2}{\varepsilon_{r-1}}\Big)^r \\ &= \frac{1}{\varepsilon_r}.\end{aligned}$$

As $A \smallsetminus (A_1 \cup \ldots \cup A_r)$ is nonempty and a union of cosets of $A_1 \cap \ldots \cap A_r$, its proportion is at least ε_r. □

Let U^{run} denote the set of regular unipotent elements in U. By assumption $\Gamma \cap U^{\mathrm{run}}$ is nonempty. In fact:

Lemma 8.6. *If* Γ *is sufficiently general, we have*

$$\big|\Gamma \cap U^{\mathrm{run}}\big| \;\geq\; \varepsilon_{\operatorname{rank} G} \cdot \big|\Gamma \cap U\big|,$$

where ε_r *is defined as in Lemma* 8.5*.*

Proof. An element of U is regular unipotent if and only if, in the decomposition 8.1, the component in each simple root group is nontrivial (see [18, Prop. 4.1, Thm. 4.6]). Let A be the image of $\Gamma \cap U$ in the quotient group U/U^{der}, which is isomorphic to $\mathbb{G}_{a,k}^{\operatorname{rank} G}$ by Proposition 8.2. Then an element of $\Gamma \cap U$ is regular unipotent if and only if its image in A does not lie in a coordinate hyperplane. The proportion of such elements is estimated in Lemma 8.5. □

Lemma 8.7. *If* Γ *is sufficiently general, we have*

$$\big[\Gamma \cap B : \Gamma \cap U\big] \;\geq\; \frac{\varepsilon_{\operatorname{rank} G}}{c_0^3} \cdot q_\Gamma^{\operatorname{rank} G}.$$

Proof. We decompose

$$\begin{aligned}\big[\Gamma \cap B : \Gamma \cap U\big] &= \frac{[\Gamma \cap B : \Gamma_u] \cdot |\Gamma_u|}{|\Gamma \cap U|} \\ &= \frac{|O_{\Gamma \cap B}(u)|}{|\Gamma \cap U^{\mathrm{run}}|} \cdot \frac{|\Gamma \cap U^{\mathrm{run}}|}{|\Gamma \cap U|} \cdot \big|\Gamma_u\big|,\end{aligned}$$

and deal separately with each term on the right-hand side. First recall that all regular unipotent elements of G are conjugate and that each one lies in a unique Borel subgroup (see [18, Thm. 4.6]). Thus any two elements of U^{run} are conjugate and, since B is its own normalizer, any element of G that conjugates one into the other lies in B. Now Theorem 6.6 implies that $\Gamma \cap U^{\mathrm{run}}$ consists of at most c_0^2

conjugacy classes under $\Gamma \cap B$. We may assume without loss of generality that the conjugacy class of u is the largest of these. Then the first term on the right-hand side of 8 is at least $1/c_0^2$.

The second term is bounded below by Lemma 8.6. The third term is at least $q_\Gamma^{\dim G_u}/c_0$ by Theorem 6.2. Since u is regular unipotent, we have $\dim G_u = \operatorname{rank} G$, and the desired estimate follows. □

Now we can find many semisimple elements in $\Gamma \cap B$:

Proposition 8.8. *If Γ is sufficiently general, there exists a maximal torus $T \subset B$ with*

$$\left|\Gamma \cap T\right| \;\geq\; \frac{\varepsilon_{\operatorname{rank} G}}{c_0^3} \cdot q_\Gamma^{\operatorname{rank} G}.$$

Proof. Since $\Gamma \cap U$ is a p-group which is normal in $\Gamma \cap B$ of index prime to p, it possesses a semidirect complement. As B is connected solvable, this complement is contained in a maximal torus of B (cf. [17, Prop. 19.4] or [1, Thm. 10.6(5)]). □

Minimal unipotent elements. Using the preceding estimate we can describe $Z(U)$ as the centralizer of a subset of Γ. Let T be as in Proposition 8.8.

Lemma 8.9. *If Γ is sufficiently general, the centralizer of $O_{\Gamma\cap T}(u)$ in G is $Z(U)$.*

Proof. Clearly the desired centralizer contains $Z(U)$. Assume that it possesses an element $g \in G \setminus Z(U)$. Then, dually, the conjugacy class $O_{\Gamma\cap T}(u)$ is contained in the centralizer U_g. If $f : T \to U$ denotes the conjugation map $t \mapsto tut^{-1}$, this in turn means that $\Gamma \cap T \subset f^{-1}(U_g)$. Here $f^{-1}(U_g)$ is a subvariety of T which belongs to a constructible family since U, g, u, and T do so.

We claim that $f^{-1}(U_g) \neq T$. Indeed, equality would mean that $O_T(u)$ is contained in U_g. Since u is regular unipotent, its image in U/U^{der} lies in the unique open T-orbit. Therefore $O_T(u)$ generates U modulo U^{der}. But U^{der} is the commutator subgroup of U, and U is a nilpotent group, so $O_T(u)$ generates U. Since, by construction, U_g is a proper subgroup of U, it cannot contain $O_T(u)$. This proves the claim.

If $\Gamma \cap T \subset f^{-1}(U_g)$, Theorem 4.2 now shows $|\Gamma \cap T| \leq c \cdot q_\Gamma^{\operatorname{rank} G-1}$ for some constant c. Since the size of $\Gamma \cap T$ is also bounded below by Proposition 8.8, we obtain an upper bound for q_Γ. But this contradicts Proposition 6.1 if Γ is sufficiently general. □

Plugging Lemma 8.9 into Theorem 6.2, we deduce:

Corollary 8.10. *If Γ is sufficiently general, we have*

$$\frac{1}{c_0} \cdot q_\Gamma^{\dim Z(U)} \;\leq\; \left|\Gamma \cap Z(U)\right| \;\leq\; c_0 \cdot q_\Gamma^{\dim Z(U)}.$$

Let us note in passing that one can manufacture minimal unipotent elements also by taking repeated commutators of elements of $O_{\Gamma\cap T}(u)$.

Decomposing further. With Corollary 8.10 we have already constructed many minimal unipotent elements in Γ. In the nonstandard case, we can sometimes specialize further. Namely, suppose that $B \subset G$ is any Borel subgroup and U is its unipotent radical. Consider a B-invariant subgroup $V \subset Z(U)$ such that $\Gamma \cap V \neq \{1\}$. We begin by characterizing V as the centralizer of a subset of Γ, as in Lemma 8.9:

Lemma 8.11. *If Γ is sufficiently general, the centralizer of $\Gamma \cap G_V$ in G is V.*

Proof. Note first that $G_{\Gamma\cap G_V}$ contains G_{G_V}, which equals V by Proposition 8.4 (c). For the reverse inclusion we study the size of $\Gamma \cap G_V$. From Proposition 8.4 (b) we know that $G_V = G_{\Gamma\cap V}$. Thus Theorem 6.2 implies that

$$\big|\Gamma \cap G_V\big| \;=\; \big|\Gamma_{\Gamma\cap V}\big| \;\geq\; \frac{1}{c_0}\cdot q_\Gamma^{\dim G_V} \tag{8.12}$$

whenever Γ is sufficiently general. Assume that there exists an element $g \in G_{\Gamma\cap G_V} \smallsetminus V$. Then g commutes with $\Gamma \cap G_V$; hence, dually, $\Gamma \cap G_V \subset G_g$. The assumption $g \notin V = G_{G_V}$ means that g does not commute with G_V, so that $G_V \not\subset G_g$. Therefore $G_V \cap G_g$ is a proper subgroup of G_V. Since by Proposition 8.4 (b) the latter is connected, we must have $\dim(G_V \cap G_g) < \dim G_V$. Now Theorem 4.2 implies that

$$\big|\Gamma \cap G_V\big| \;=\; \big|\Gamma \cap G_V \cap G_g\big| \;\leq\; c\cdot q_\Gamma^{\dim(G_V\cap G_g)} \;\leq\; \frac{c}{q_\Gamma}\cdot q_\Gamma^{\dim G_V},$$

where c is fixed and Γ is sufficiently general. Comparing this with the lower bound (8.12), we obtain an upper bound for q_Γ. This contradicts Proposition 6.1 if Γ is sufficiently general. Therefore $G_{\Gamma\cap G_V} = V$, as desired. □

Combining Lemma 8.11 with Theorem 6.2, we deduce:

Corollary 8.13. *If Γ is sufficiently general, we have*

$$\frac{1}{c_0}\cdot q_\Gamma^{\dim V} \;\leq\; \big|\Gamma \cap V\big| \;\leq\; c_0\cdot q_\Gamma^{\dim V}.$$

Finding a finite field. In the following we abbreviate $d := \dim V$ and impose:

Assumption 8.14. d is minimal for all possible B and V with $\Gamma \cap V \neq \{1\}$.

Let $\mathbb{F}_V$ denote the image of the group ring $\mathbb{F}_p[N_\Gamma(V)]$ in $\mathrm{End}(V)$. Proposition 8.4 (a) implies that this is a finite $\mathbb{F}_p$-subalgebra of k^d.

Proposition 8.15. *The ring $\mathbb{F}_V$ is a field.*

Proof. If not, we must have $d = 2$, and any zero-divisor $x \in \mathbb{F}_V$ lies in one of the factors of k^2. Decomposing under x we deduce $\Gamma \cap V = (\Gamma \cap U_{\alpha_\ell}) \oplus (\Gamma \cap U_{\alpha_s})$. This contradicts Assumption 8.14. □

Thus $\Gamma \cap V$ is a finite vector space over $\mathbb{F}_V$. Note the general fact:

Lemma 8.16. *Consider a finite nonzero vector space M over a finite field $\mathbb{F}$. Suppose there is a constant n such that*

(a) *the number of $\mathbb{F}^\times$-orbits on $M \smallsetminus \{0\}$ is at most n, and*
(b) *$|M| \geq n^2$.*

Then $\dim_{\mathbb{F}} M = 1$.

Proof. Abbreviate $q := |\mathbb{F}|$ and $r := \dim_{\mathbb{F}} M$, and assume $r \geq 2$. The number of $\mathbb{F}^\times$-orbits on $M \smallsetminus \{0\}$ is then

$$\frac{q^r - 1}{q - 1} \;=\; q^{r-1} + \cdots + 1 \;>\; q^{r-1}.$$

Thus (a) implies $n^2 > q^{2(r-1)} \geq q^r = |M|$, which contradicts (b). □

Now we can prove the following crucial result:

Theorem 8.17. *If Γ is sufficiently general, we have $\dim_{\mathbb{F}_V}(\Gamma \cap V) = 1$ and, in particular,*

$$\frac{1}{c_0} \cdot q_\Gamma^d \;\le\; \bigl|\mathbb{F}_V\bigr| \;\le\; c_0 \cdot q_\Gamma^d.$$

Proof. The second assertion follows from the first together with Corollary 8.13. For the first assertion note that, by the minimality of V, all nontrivial elements of $\Gamma \cap V$ lie in the open T-orbit of V. Therefore they all lie in the same G-conjugacy class. By Theorem 6.6 they fall into at most c_0^2 conjugacy classes under Γ. By Proposition 8.4 (d) two such elements are conjugate under Γ if and only if they are conjugate under $N_\Gamma(V)$. Since $N_\Gamma(V)$ acts through a subgroup of $\mathbb{F}_V^\times$, the number of $\mathbb{F}_V^\times$-orbits on $\Gamma \cap V \smallsetminus \{1\}$ is $\le c_0^2$. On the other hand $\Gamma \cap V$ is arbitrarily large, by Corollary 8.13 and Proposition 6.1. Thus the theorem follows from Lemma 8.16. □

Notation. We fix some notation to be used in the following sections. The order of $\mathbb{F}_V$ is denoted by p^r. If $d = 2$ we suppose that the first component of $\mathbb{F}_V \subset k^2$ corresponds to the action on U_{α_ℓ}, the second to the action on U_{α_s}. As $\mathbb{F}_V$ is a finite field, the two projection maps $\mathbb{F}_V \to k$ must differ by a Frobenius twist. Thus the elements of $\mathbb{F}_V$ have the form (x, x^{p^e}) for a unique integer $0 \le e < r$.

9. Frobenius map

We keep the notation of the preceding sections. As a result of Proposition 8.15 we have associated to any sufficiently general Γ a certain finite field $\mathbb{F}_V$ of characteristic $p > 0$, and by Theorem 8.17 the size of Γ is roughly that expected of a finite group of Lie type over $\mathbb{F}_V$. We will establish that Γ indeed has this form.

Strategy of proof. The first problem is to translate the internal characterization of $\mathbb{F}_V$ within Γ into external information on the coefficients in representations. This is achieved by showing that the traces of certain elements of Γ in a suitable algebraic representation of G lie in $\mathbb{F}_V$. By combining this information for many elements of Γ one can then show that some other algebraic representation descends to $\mathbb{F}_V$ when restricted to Γ. We will give a precise formulation of this intermediate result. By a *model over* $\mathbb{F}_V$ of a k^d-module M we mean an $\mathbb{F}_V$-submodule $M_0 \subset M$ such that the natural map $M_0 \otimes_{\mathbb{F}_V} k^d \longrightarrow M$ is an isomorphism.

Theorem 9.1. *There exists a nontrivial representation σ of G on a k^d-module M of finite type, which belongs to a constructible family of representations of $\mathscr{G}$, such that, for any sufficiently general finite subgroup $\Gamma \subset G$, there exists a Γ-invariant model M_0 of M over $\mathbb{F}_V$.*

Our method to prove this in general depends on knowing Theorem 0.5 already in the case $\operatorname{rank} G = d$, for which we therefore need a different argument. We call this the *basic* case and handle it in Section 10. The general case will be treated in Section 11. In the remainder of this section we show how Theorem 9.1 implies Theorem 0.5 for the given group $\mathscr{G}$.

Construction of Frobenius. Let us view the automorphism group $\operatorname{Aut}_{k^d}(M)$ as an algebraic group over k. If $d = 1$, the model M_0 determines a standard Frobenius map $F\colon \operatorname{Aut}_k(M) \to \operatorname{Aut}_k(M)$ relative to the finite field $\mathbb{F}_V$.

In the case $d = 2$, let $M = M_\ell \oplus M_s$ be the decomposition according to the two factors of k^2. If i_ℓ, $i_s : \mathbb{F}_V \to k$ denote the two projection maps, recall that $i_s = \mathrm{Frob}_{p^e} \circ i_\ell$. Thus the choice of M_0 determines an isomorphism

$$M_s \;\cong\; M_0 \otimes_{\mathbb{F}_V, i_s} k \;\cong\; \big(M_0 \otimes_{\mathbb{F}_V, i_\ell} k\big) \otimes_{k,\mathrm{Frob}_{p^e}} k \;\cong\; M_\ell \otimes_{k,\mathrm{Frob}_{p^e}} k,$$

and hence an isogeny

$$F\colon\; \mathrm{Aut}_k(M_\ell) \longrightarrow \mathrm{Frob}_{p^e}^* \mathrm{Aut}_k(M_\ell) \cong \mathrm{Aut}_k(M_s).$$

Similarly, we have $i_\ell = \mathrm{Frob}_{p^{r-e}} \circ i_s$ and an isogeny

$$F\colon\; \mathrm{Aut}_k(M_s) \longrightarrow \mathrm{Frob}_{p^{r-e}}^* \mathrm{Aut}_k(M_s) \cong \mathrm{Aut}_k(M_\ell).$$

Taken together we find an isogeny F on $\mathrm{Aut}_k(M_s) \times \mathrm{Aut}_k(M_\ell) = \mathrm{Aut}_{k^d}(M)$ whose square is a standard Frobenius map relative to the finite field $\mathbb{F}_V$. In both cases Theorem 9.1 implies, for all $\gamma \in \Gamma$,

$$F \circ \sigma(\gamma) = \sigma(\gamma). \tag{9.2}$$

Lemma 9.3. *If Γ is sufficiently general, then $F(\sigma(G)) = \sigma(G)$.*

Proof. We claim that $\sigma^{-1}(F(\sigma(G)))$ belongs to a constructible family of algebraic subgroups of $\mathscr{G}$. To prove this, consider first the case $d = 1$. The constructibility assumption in Theorem 9.1 means that there is a vector bundle $\mathscr{M}$ on a scheme $\mathscr{T}$ of finite type over $\mathbf{Spec}\,\mathbb{Z}$, and a homomorphism $\sigma\colon \mathscr{G} \times_{\mathbf{Spec}\,\mathbb{Z}} \mathscr{T} \longrightarrow \mathrm{Aut}_{\mathscr{O}_\mathscr{T}}(\mathscr{M})$, such that $M = \mathscr{M}_t$ for some $t \in \mathscr{T}(k)$ with the induced representation of $G = \mathscr{G}(k)$. As $\sigma(G)$ is Zariski-closed in $\mathrm{Aut}(M)$, by Proposition 1.7 it belongs to a constructible family of algebraic subgroups $\mathscr{H}$ of $\mathrm{Aut}_{\mathscr{O}_\mathscr{T}}(\mathscr{M})$. Without loss of generality we may assume that $\mathscr{H}$ is indexed by the same scheme $\mathscr{T}$, so that $\sigma(G) = \mathscr{H}_t$. Let $\mathrm{Frob}_{p^r} : k \to k$ denote the Frobenius map $x \mapsto x^{p^r}$, and $t' \in \mathscr{T}(k)$ the image of t under Frob_{p^r}. The choice of M_0 corresponds to an identification $\mathscr{M}_{t'} = \mathscr{M}_t \otimes_{k,\mathrm{Frob}_{p^r}} k \cong \mathscr{M}_t = M$, resulting in a commutative diagram

$$\begin{tikzcd}
\mathrm{Aut}_k(\mathscr{M}_{t'}) \arrow[r, phantom, "\cong"] & \mathrm{Aut}_k(\mathscr{M}_t) \arrow[r, phantom, "="] & \mathrm{Aut}_k(M) \\
\mathscr{H}_{t'} \arrow[u, phantom, "\cup"] \arrow[r, "\sim"] & F(\mathscr{H}_t) \arrow[u, phantom, "\cup"] \arrow[r, phantom, "="] & F(\sigma(G)). \arrow[u, phantom, "\cup"]
\end{tikzcd}$$

Here the isomorphism $\mathscr{M}_{t'} \cong \mathscr{M}_t$ is indexed by a point on the constructible family $\mathrm{Isom}\,(\mathrm{pr}_1^* \mathscr{M}, \mathrm{pr}_2^* \mathscr{M})$ over $(t, t') \in \mathscr{T} \times \mathscr{T}$. Therefore $F(\sigma(G)) \subset \mathrm{Aut}_k(M)$ belongs to a constructible family of algebraic subgroups, and so does $\sigma^{-1}(F(\sigma(G))) \subset G$, as claimed. In the case $d = 2$ the proof is analogous.

Now (9.2) implies that $\sigma^{-1}(F(\sigma(G)))$ contains Γ. Thus, if Γ is sufficiently general, this subgroup must be equal to G. This implies that $\sigma(G) \subset F(\sigma(G))$, and equality follows from the fact that both sides are irreducible of the same dimension. □

Lemma 9.4. *If Γ is sufficiently general, there exists a Frobenius map $F : G \to G$ with $q_F^d = |\mathbb{F}_V|$, so that $\Gamma \subset G^F$.*

Proof. As σ is a nontrivial representation, and G is simple adjoint, the induced map to the adjoint group $G \to \sigma(G)^{\mathrm{ad}}$ is a totally inseparable isogeny. From Lemma 9.3 and the classification of isogenies of simple adjoint groups (see [28, Thm. 1.7]) we deduce that there is an isogeny $F\colon G \to G$ satisfying $\sigma \circ F = F \circ \sigma$. Moreover, F^d is a standard Frobenius map relative to the field $\mathbb{F}_V$; hence $q_F^d = q_{F^d} = |\mathbb{F}_V|$. On the other hand, the fact that σ is injective on k-valued points and the property (9.2) imply the desired inclusion $\Gamma \subset G^F$. □

Proof of Theorem 0.5. Combining Lemma 9.4 with Theorem 8.17 we find

$$\frac{1}{\sqrt[d]{c_0}} \cdot q_\Gamma \;\leq\; q_F \;\leq\; \sqrt[d]{c_0} \cdot q_\Gamma.$$

Thus using Theorem 3.4 (d) we deduce

$$\big|G^F\big| \;\leq\; q_F^{\dim G} \;\leq\; c_0^{\frac{\dim G}{d}} \cdot q_\Gamma^{\dim G} \;=\; c_0^{\frac{\dim G}{d}} \cdot \big|\Gamma\big|.$$

Therefore the index

$$\big[(G^F)^{\mathrm{der}} : \Gamma \cap (G^F)^{\mathrm{der}}\big] \;\leq\; \big[G^F : \Gamma\big]$$

is bounded, and so is the index of the largest normal subgroup of $(G^F)^{\mathrm{der}}$ that is contained in Γ. On the other hand, if Γ is sufficiently general, Proposition 6.1 says that q_Γ and hence q_F is arbitrarily large. Thus by Theorem 3.4 (a), (b), and (d) the group $(G^F)^{\mathrm{der}}$ is simple and arbitrarily large. Therefore Γ contains $(G^F)^{\mathrm{der}}$; hence $\big(G^F\big)^{\mathrm{der}} \subset \Gamma \subset G^F$, as desired. This finishes the proof of Theorem 0.5 modulo Theorem 9.1. □

Notation: The adjoint representation. We fix some notation to be used in the next two sections. Observe that the Lie algebra of G is a fiber of the vector bundle $\mathrm{Lie}\,\mathscr{G} \to \mathbf{Spec}\,\mathbb{Z}$; hence the adjoint representation of G belongs to a constructible family. The G-invariant subspaces are known completely, by [13], [14], or [28, Prop. 1.11].

If $\dim Z(U) = 1$, there is a unique simple subquotient on which G acts nontrivially. The representation on it is denoted by ρ. The root system being fixed, $\mathrm{Lie}\,G$ is already irreducible whenever $p \gg 0$. Thus by replacing $\mathbf{Spec}\,\mathbb{Z}$ by a stratification consisting of a union of a dense open subset and a finite number of points $\mathbf{Spec}\,\mathbb{F}_p$, we see that ρ belongs to a constructible family of representations.

If $\dim Z(U) = 2$, there are precisely two simple subquotients with nontrivial G-action, one of which contains copies of all long root spaces, the other of all short root spaces. The corresponding representations of G are denoted by ρ_ℓ and ρ_s. Since this case arises in at most one characteristic p for each Φ, these representations form a tautological constructible family over $\mathbf{Spec}\,\mathbb{F}_p \subset \mathbf{Spec}\,\mathbb{Z}$.

If $\dim Z(U) = d = 2$, we also view $\rho := (\rho_\ell, \rho_s)$ as a representation over k^2. If $\dim Z(U) = 2$, but $d = 1$, we let $\rho := \rho_\ell$ if $V = U_{\alpha_\ell}$, and $\rho := \rho_s$ if $V = U_{\alpha_s}$.

10. Traces in the basic case

In this section we prove Theorem 9.1 in the basic case $\mathrm{rank}\,G = d$. So this assumption, as well as the other notation of the preceding sections, will be in force. Note that either $d = 1$ and $\Phi = A_1$, or $d = 2$ and (p, Φ) is $(2, B_2)$ or $(3, G_2)$.

The rank one case. Here we have $G \cong \mathrm{PGL}_2$, and everything can be deduced directly from the following classical theorem of Dickson [8, §§260–261]:

Theorem 10.1. *Consider a field k and a finite subgroup $\Gamma \subset \mathrm{PGL}_2(k)$. Then either*

(a) *the inverse image of Γ in $\mathrm{GL}_2(k)$ acts reducibly on k^2;*

(b) *Γ is a dihedral group;*

(c) $\Gamma \cong A_4$, S_4, A_5*; or*

(d) $p := \mathrm{char}(k)$ *is positive, and after a suitable change of basis we have* $\Gamma = \mathrm{PGL}_2(\mathbb{F}_{p^r})$ *or* $\mathrm{PGL}_2(\mathbb{F}_{p^r})^{\mathrm{der}}$ *for some* $r \geq 1$*, where the latter group is simple.*

The subgroups in (a) through (c) are special: they lie in a Borel subgroup, or in the normalizer of a maximal torus, or have bounded order. Thus any sufficiently general subgroup must be of type (d), which proves Theorem 0.5 in the case $\Phi = A_1$.

However, our method in the cases $(2, B_2)$ or $(3, G_2)$ adapts very easily to the A_1-case as well. For the sake of completeness we therefore include an independent proof based on the ideas of this paper. The reader willing to ignore the existence of Suzuki and Ree groups in characteristics 2 and 3 may skip the rest of this section.

The whole basic case. The main idea is that the Γ-conjugacy classes of elements of $\Gamma \cap B$ contribute sufficiently many elements with trace in $\mathbb{F}_V$. Consider a maximal torus $T \subset B$ which contains many elements of Γ, as in Proposition 8.8, and let $\Lambda \subset \Gamma$ be the set of elements which are conjugate to an element of $\Gamma \cap T^{\mathrm{rss}}$.

Proposition 10.2. *There is a constant* $\varepsilon > 0$ *such that, whenever* Γ *is sufficiently general, we have* $|\Lambda| \geq \varepsilon \cdot |\Gamma|$.

Proof. If both t and gtg^{-1} lie in T^{rss}, we must have $g \in N_G(T)$. Therefore

$$\big|\Lambda\big| = \big[\Gamma : N_\Gamma(T)\big] \cdot \big|\Gamma \cap T^{\mathrm{rss}}\big| = \frac{|\Gamma|}{[N_\Gamma(T) : \Gamma \cap T]} \cdot \frac{|\Gamma \cap T^{\mathrm{rss}}|}{|\Gamma \cap T|}.$$

The denominator in the first fraction is at most the order of the Weyl group of Φ; hence it is bounded. The second ratio can be bounded below by any constant less than 1, using Proposition 7.3. The desired estimate follows. □

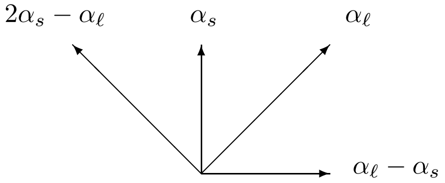
Recall that ρ is a representation over k^d; hence its trace takes values in k^d.

Proposition 10.3. *If* Γ *is sufficiently general, for every* $\gamma \in \Lambda$ *we have* $\mathrm{Tr}\,\rho(\gamma) \in \mathbb{F}_V$.

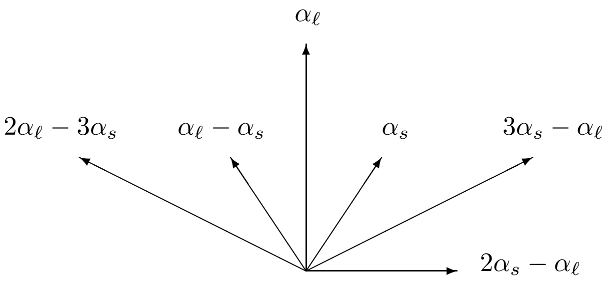
Proof in the rank one case. As the assertion is invariant under conjugation, we may assume $\gamma \in \Gamma \cap T^{\mathrm{rss}}$. Then γ acts on $U \cong \mathbb{G}_a$ and $\mathrm{Lie}\,U$ through the same scalar x, and the construction of $\mathbb{F}_V$ implies $x \in \mathbb{F}_V$. The total trace is $x + 1 + x^{-1}$ if the adjoint representation of $\mathrm{PGL}_{2,k}$ is irreducible, and $x + x^{-1}$ otherwise. It therefore also lies in $\mathbb{F}_V$, as desired. □

The rank two case. This part is more involved. It requires a closer look at $\Gamma \cap B$ from the viewpoint of the geometry of roots, combined with some arithmetic of finite fields. To begin with, note that the positive roots are

Case $(2, B_2)$:



Case $(3, G_2)$:


α_ℓ
$2\alpha_\ell - 3\alpha_s$
$\alpha_\ell - \alpha_s$
α_s
$3\alpha_s - \alpha_\ell$
$2\alpha_s - \alpha_\ell$

Recall that $V = U_{\alpha_\ell} U_{\alpha_s}$. Let W denote the subgroup generated by V and the next two lower root groups. Thus in the case $(2, B_2)$ we set $W := U$; in the case $(3, G_2)$ we set $W := U'$. Then

$$W/V \;\cong\; U_{p\alpha_s - \alpha_\ell} \times U_{\alpha_\ell - \alpha_s}. \tag{10.4}$$

Lemma 10.5. *There exists an element $w \in \Gamma \cap W$ whose component in each factor of W/V is nontrivial.*

Proof. In the case $(2, B_2)$ any regular unipotent element in $\Gamma \cap U$ has this property. In the case $(3, G_2)$ we take a regular unipotent element $u \in \Gamma \cap U$ and an element $\gamma \in \Gamma \cap T$ whose action on U/U' is not scalar. The existence of the latter is guaranteed by Proposition 8.8. The commutator of u and $\gamma u \gamma^{-1}$ then has the desired property, by the formulas in [17, Prop. 33.5 (b)]. □

Next recall from Proposition 8.4 (a) that $N_G(V)$ is a parabolic subgroup. By case analysis we easily find $N_G(V) = B$. By the proof of Proposition 8.8 we have $\Gamma \cap B = (\Gamma \cap U) \rtimes (\Gamma \cap T)$. Thus, as $\Gamma \cap U$ acts trivially on V, the field $\mathbb{F}_V \subset k^2$ is generated by the image of $\Gamma \cap T$. Since the root lattice is generated by α_ℓ and α_s, the torus T acts faithfully on V. It follows that $\Gamma \cap T$ maps isomorphically to a subgroup of $\mathbb{F}_V^\times$. In particular it is cyclic, and we choose a generator γ. Let (x, x^{p^e}) be its image in $\mathbb{F}_V^\times$. Since $|\mathbb{F}_V| = p^r$, the order of x in the multiplicative group is a divisor of $p^r - 1$. We will use the following facts:

Lemma 10.6. *If Γ is sufficiently general, then:*

(a) $p^r \gg 0$.
(b) $e \not\equiv 0 \bmod r$.
(c) *The order of x in $k^\times$ is at least $(p^r - 1)/c_0^2$.*

Proof. (a) follows from Theorem 8.17 and Proposition 6.1. Next recall that $\Gamma \cap T$ is a maximal toric subgroup; hence γ is regular semisimple. Its eigenvalue on the root $\alpha_\ell - \alpha_s$ is x^{1-p^e}, whence (b). Finally, the proof of Theorem 8.17 shows that the image of $\Gamma \cap T$ is a subgroup of $\mathbb{F}_V^\times$ of index at most c_0^2. This implies (c). □

Lemma 10.7. *If Γ is sufficiently general, at least one of the following assertions is true:*

(a) $x^{p^{e+1}-1} = x^{p^n}$ *for some integer $n \geq 0$.*
(b) $x^{p^{e+1}-1} = (x^{1-p^e})^{p^n}$ *for some integer $n \geq 0$.*

Proof. The eigenvalues of γ on $U_{p\alpha_s - \alpha_\ell}$ and $U_{\alpha_\ell - \alpha_s}$ are $x^{p^{e+1}-1}$ and x^{1-p^e}, respectively. Thus if (b) fails, the two eigenvalues of γ on W/V are not Frobenius conjugates of each other. This means that the subring of $\mathrm{End}(W/V)$ generated by the action of γ decomposes as in (10.4). Beginning with the element w of Lemma 10.5 we can therefore manufacture an element of $\Gamma \cap W$ whose image in W/V has a nontrivial component only in $U_{p\alpha_s-\alpha_\ell}$. In other words, we have found an element in $\Gamma \cap (U_{p\alpha_s-\alpha_\ell}V) \smallsetminus V$.

The group $U_{p\alpha_s-\alpha_\ell}V$ is commutative. If (a) fails, the eigenvalue of γ on $U_{p\alpha_s-\alpha_\ell}$ is not a Frobenius conjugate of the eigenvalues on $V = U_{\alpha_\ell}U_{\alpha_s}$. Therefore the subring of the endomorphism ring $\mathrm{End}(U_{p\alpha_s-\alpha_\ell}V)$ generated by the action of γ decomposes, and we can find a nontrivial element in $\Gamma \cap U_{p\alpha_s-\alpha_\ell}$. But then we could have worked from the start with $U_{p\alpha_s-\alpha_\ell}$ in place of V, contradicting Assumption 8.14. □

Next we need the following result about greatest common divisors:

Lemma 10.8. *Consider any prime p and any integers r, a, a', b, $b' \geq 0$.*

(a) *We have*

$$\big(p^r - 1,\ p^a + p^{a'} - p^b\big) \ < \ 2p^{2r/3},$$

unless $p = 2$ and $a \equiv a' \equiv b - 1 \bmod r$.

(b) *We have*

$$\big(p^r - 1,\ p^a + p^{a'} - p^b - p^{b'}\big) \ < \ 2p^{3r/4},$$

unless $\{a \bmod r,\ a' \bmod r\} = \{b \bmod r,\ b' \bmod r\}$.

Proof. For (a) observe that the left-hand side depends only on a, a', b modulo r and is unchanged on replacing these by $a+n$, $a'+n$, $b+n$ for any integer n. There are fewer than $r/3$ values of n modulo r for which the remainder of $a+n$ modulo r is greater than $2r/3$, and likewise for $a'+n$ and $b+n$. Thus there is at least one value of n for which all three remainders are $\leq 2r/3$. Without loss of generality we may therefore assume $0 \leq a, a', b \leq 2r/3$. This implies $\big|p^a + p^{a'} - p^b\big| < 2p^{2r/3}$. As we have $p^a + p^{a'} = p^b$ only if $a = a' = b - 1$ and $p = 2$, this shows (a). The proof of (b) follows the same lines and is left to the reader. □

Lemma 10.9. *If Γ is sufficiently general, we have $2e + 1 = r$.*

Proof. In the situation of Lemma 10.7 (a) the order of x is a divisor of

$$\big(p^r - 1,\ p^n + 1 - p^{e+1}\big).$$

By Lemma 10.8 (a) this is $< 2p^{2r/3}$ unless $p = 2$ and $n \equiv 0 \equiv e$ modulo r. Both of these possibilities are excluded by Lemma 10.6. A similar argument applies in the situation of Lemma 10.7 (b). Here the order of x is a divisor of

$$\big(p^r - 1,\ p^{e+1} + p^{e+n} - 1 - p^n\big).$$

By Lemma 10.8 (b), this is $< 2p^{3r/4}$ unless $\{e + 1 \bmod r,\ e + n \bmod r\} = \{0 \bmod r,\ n \bmod r\}$. The former case is excluded by Lemma 10.6 (a) and (c). As $e \not\equiv 0$ modulo r by Lemma 10.6 (b), we deduce $e + 1 \equiv n$ and $e + n \equiv 0$ modulo r. This implies $2e + 1 \equiv 0$ modulo r, whence the assertion. □

Proof of Proposition 10.3 in the rank two case. It suffices to show that

$$\operatorname{Tr}\rho_s(\gamma^i) = \big(\operatorname{Tr}\rho_\ell(\gamma^i)\big)^{p^e}$$

for every integer i. The weight 0 occurs in both representations with the same multiplicity, namely multiplicity 0 in the case $(2, B_2)$, and multiplicity 1 in the case $(3, G_2)$ (see [28, Prop. 1.11]). Thus we may replace the trace by the sum over all nonzero weights. In ρ_ℓ these are precisely the long roots, in ρ_s the short roots. So the desired formula follows if we can match bijectively long roots β_ℓ and short roots β_s such that $\beta_s(\gamma) = \beta_\ell(\gamma)^{p^e}$. Among positive roots such a matching is given by the following table:

β_ℓ	β_s	$\beta_\ell(\gamma)$	$\beta_s(\gamma)$
α_ℓ	α_s	x	x^{p^e}
$p\alpha_s - \alpha_\ell$	$\alpha_\ell - \alpha_s$	$x^{p^{e+1}-1}$	$x^{1-p^e} = (x^{p^{e+1}-1})^{p^e}$
$2\alpha_\ell - p\alpha_s$	$2\alpha_s - \alpha_\ell$	$x^{2-p^{e+1}}$	$x^{2p^e-1} = (x^{2-p^{e+1}})^{p^e}$

Here the indicated equalities follow from Lemma 10.9, and the last row applies only to the case $(3, G_2)$. The corresponding matching works for negative roots. This finishes the proof of Proposition 10.3. □

From traces to matrices. By looking at traces of suitable products in Γ we will obtain information on all matrix coefficients. First we note the following abstract result:

Lemma 10.10. *Consider a finite group Γ and a subset $\Lambda \subset \Gamma$ satisfying $|\Lambda| \geq \varepsilon \cdot |\Gamma|$ with $\varepsilon > 0$. Consider a positive integer ℓ and set $\varepsilon' := \varepsilon^\ell/2$. Let Ω be the set of all tuples $(\gamma_1, \ldots, \gamma_\ell) \in \Gamma^\ell$ with the property*

$$\Bigg|\bigcap_{i=1}^{\ell} \gamma_i^{-1}\Lambda\Bigg| \;\geq\; \varepsilon' \cdot \big|\Gamma\big|.$$

Then $|\Omega| \geq \varepsilon' \cdot |\Gamma|^\ell$.

Proof. We estimate the number of tuples $(\gamma_1, \ldots, \gamma_\ell, \gamma) \in \Gamma^{\ell+1}$ satisfying $\gamma_i\gamma \in \Lambda$ for all $1 \leq i \leq \ell$. Summing first over $\gamma \in \Gamma$, this number is equal to

$$\big|\Gamma\big| \cdot \big|\Lambda\big|^\ell \;\geq\; \varepsilon^\ell \cdot \big|\Gamma\big|^{\ell+1} \;=\; 2\varepsilon' \cdot \big|\Gamma\big|^{\ell+1}.$$

On the other hand, summing first over $(\gamma_1, \ldots, \gamma_\ell)$ the tuples in Ω contribute at most $|\Omega| \cdot |\Gamma|$. The remaining tuples contribute at most

$$\Big(\big|\Gamma\big|^\ell - \big|\Omega\big|\Big) \cdot \varepsilon' \cdot \big|\Gamma\big| \;\leq\; \varepsilon' \cdot \big|\Gamma\big|^{\ell+1}.$$

Thus all together we find

$$\big|\Omega\big| \cdot \big|\Gamma\big| + \varepsilon' \cdot \big|\Gamma\big|^{\ell+1} \;\geq\; 2\varepsilon' \cdot \big|\Gamma\big|^{\ell+1},$$

whence the lemma. □

Let M be the ring of k^d-linear endomorphisms of the representation space of ρ. In the nonstandard case we have $\dim\rho_\ell = \dim\rho_s$, since $\operatorname{rank} G = 2$ (cf. [28, Prop. 1.11]). Thus in either case M is a ring of matrices of some size $n \times n$ over k^d. Take $\Lambda \subset \Gamma$ and ε as in Proposition 10.2, and let Ω and ε' be as in Lemma 10.10 with $\ell := n^2$.

Lemma 10.11. *If Γ is sufficiently general, there exists a tuple $(\gamma_1, \ldots, \gamma_{n^2}) \in \Omega$ such that the elements $\rho(\gamma_i)$ form a basis of M over k^d.*

Proof. Let X denote the set of tuples $(g_1, \ldots, g_{n^2}) \in G^{n^2}$ for which the elements $\rho(g_i)$ do not form a basis of M over k^d. This is a fiber of a constructible family of Zariski-closed subvarieties of $\mathscr{G}^{n^2}$, since its defining condition can be expressed in terms of the vanishing of certain determinants. It is a proper subvariety, because Burnside's theorem, applied to ρ, respectively to ρ_ℓ and ρ_s, implies that $\rho(G)$ contains a basis of M. Thus Theorem 4.3 implies that

$$\left|\Gamma^{n^2} \cap X\right| \;\le\; c \cdot q_\Gamma^{\dim X} \;\le\; \frac{c}{q_\Gamma} \cdot \left(q_\Gamma^{\dim G}\right)^{n^2} \;=\; \frac{c}{q_\Gamma} \cdot \left|\Gamma\right|^{n^2}$$

for some fixed constant c. Combined with Lemma 10.10 this implies that $\Omega \not\subset X$ if q_Γ is large, as guaranteed by Proposition 6.1. Clearly, any tuple in $\Omega \smallsetminus X$ has the desired property. □

Consider a tuple as in Lemma 10.11, and select any element $\gamma \in \bigcap_{i=1}^{n^2} \gamma_i^{-1}\Lambda$. After replacing each γ_i by $\gamma_i\gamma$, the condition in Lemma 10.11 still holds, and in addition we have $1 \in \bigcap_{i=1}^{n^2} \gamma_i^{-1}\Lambda$. We fix such a tuple and set

$$M_0 \;:=\; \left\{\, m \in M \;\middle|\; \forall 1 \le i \le n^2\colon\; \mathrm{Tr}(\rho(\gamma_i)m) \in \mathbb{F}_V \,\right\}.$$

By construction this defines a model of M as a vector space over $\mathbb{F}_V$. We do not yet worry about its relation with the algebra structure on M, but note that our normalization of the tuple implies that $\mathrm{id} \in M_0$.

Lemma 10.12. *We have $\left|\Gamma \cap \rho^{-1}(M_0)\right| \;\ge\; \varepsilon' \cdot \left|\Gamma\right|$.*

Proof. By construction

$$\begin{aligned}\Gamma \cap \rho^{-1}(M_0) &= \left\{\, \gamma \in \Gamma \;\middle|\; \forall 1 \le i \le n^2\colon\; \mathrm{Tr}(\rho(\gamma_i)\rho(\gamma)) \in \mathbb{F}_V \,\right\} \\ &\overset{\text{Prop. 10.3}}{\supset} \left\{\, \gamma \in \Gamma \;\middle|\; \forall 1 \le i \le n^2\colon\; \gamma_i\gamma \in \Lambda \,\right\} \\ &= \bigcap_{i=1}^{n^2} \gamma_i^{-1}\Lambda.\end{aligned}$$

Thus the desired lower bound follows from the choice of $(\gamma_1, \ldots, \gamma_{n^2})$ and the definition of Ω in Lemma 10.10. □

Next, we consider the left stabilizer

$$\Delta \;:=\; \left\{\, \gamma \in \Gamma \;\middle|\; \rho(\gamma)M_0 = M_0 \,\right\}.$$

Lemma 10.13. *If Γ is sufficiently general, we have $\left[\Gamma : \Delta\right] \;<\; 2/\varepsilon'$.*

Proof. Let $M_0, \ldots, M_h \subset M$ denote the pairwise distinct left Γ-translates of M_0, and let ℓ be the greatest integer less than $2/\varepsilon'$. We must prove $h+1 \le \ell$. So let us assume $h \geq \ell$. We calculate

$$\begin{aligned}
\bigl|\Gamma\bigr| &\geq \left|\Gamma \cap \bigcup_{j=0}^{\ell} \rho^{-1}(M_j)\right| \\
&= \sum_{j=0}^{\ell} \left|\Gamma \cap \rho^{-1}(M_j) \smallsetminus \bigcup_{i=0}^{j-1} \rho^{-1}(M_i)\right| \\
&= \sum_{j=0}^{\ell} \bigl|\Gamma \cap \rho^{-1}(M_j)\bigr| - \sum_{j=0}^{\ell} \left|\Gamma \cap \rho^{-1}(M_j) \cap \bigcup_{i=0}^{j-1} \rho^{-1}(M_i)\right|.
\end{aligned}$$

Here all terms in the first sum are equal to $\bigl|\Gamma \cap \rho^{-1}(M_0)\bigr|$; hence by Lemma 10.12 that sum is $\geq (\ell+1) \cdot \varepsilon' \cdot |\Gamma| \geq 2 \cdot |\Gamma|$. This implies

$$\begin{aligned}
\bigl|\Gamma\bigr| &\leq \sum_{j=0}^{\ell} \left|\Gamma \cap \rho^{-1}(M_j) \cap \bigcup_{i=0}^{j-1} \rho^{-1}(M_i)\right| \\
&\leq \sum_{j=0}^{\ell} \sum_{i=0}^{j-1} \bigl|\Gamma \cap \rho^{-1}(M_j \cap M_i)\bigr|.
\end{aligned}$$

By assumption $M_j \cap M_i$ is contained in a proper k^d-submodule $N \subsetneqq M$. Such submodules are indexed by Grassmannians, so as in the proof of Lemma 10.11 we deduce that $\rho^{-1}(N)$ belongs to a constructible family of Zariski-closed proper subvarieties of $\mathscr{G}$. Thus Theorem 4.2 implies that every term in the last sum is $\leq c \cdot q_\Gamma^{\dim G - 1} \leq |\Gamma| \cdot c/q_\Gamma$ if Γ is sufficiently general. Therefore

$$1 \;\leq\; \frac{\ell(\ell+1)}{2} \cdot \frac{c}{q_\Gamma},$$

so q_Γ is bounded, contrary to Proposition 6.1. □

Proof of Theorem 9.1 in the basic case. Let σ denote the representation of G on M, defined by $\sigma(g)(m) := \rho(g)m\rho(g)^{-1}$. As ρ belongs to a constructible family of representations, so does σ. It is also nontrivial, since ρ is not scalar. It remains to prove $\sigma(\gamma)(M_0) = M_0$ for every $\gamma \in \Gamma$.

Note first that, since $\mathrm{id} \in M_0$, we also have $\rho(\Delta) \subset M_0$. Therefore

$$\begin{aligned}
\Delta \cap \gamma\Delta\gamma^{-1} &\subset \Gamma \cap \rho^{-1}(M_0) \cap \gamma\rho^{-1}(M_0)\gamma^{-1} \\
&= \Gamma \cap \rho^{-1}\bigl(M_0 \cap \rho(\gamma)M_0\rho(\gamma)^{-1}\bigr) \\
&= \Gamma \cap \rho^{-1}\bigl(M_0 \cap \sigma(\gamma)(M_0)\bigr).
\end{aligned}$$

On the one hand, Lemma 10.13 implies that

$$\bigl|\Delta \cap \gamma\Delta\gamma^{-1}\bigr| \;\geq\; \frac{|\Gamma|}{[\Gamma : \Delta]^2} \;>\; \left(\frac{\varepsilon'}{2}\right)^2 \cdot \bigl|\Gamma\bigr|.$$

On the other hand, if M_0 and $\sigma(\gamma)(M_0)$ differ, their intersection is contained in a proper k^d-submodule of M. As in the proof of Lemma 10.13 we deduce that

$$\Bigl|\Gamma \cap \rho^{-1}\bigl(M_0 \cap \sigma(\gamma)(M_0)\bigr)\Bigr| \;\leq\; \frac{c}{q_\Gamma} \cdot \bigl|\Gamma\bigr|.$$

Thus all together we find $(\varepsilon'/2)^2 < c/q_\Gamma$, so q_Γ is bounded, contradicting Proposition 6.1. Therefore $M_0 = \sigma(\gamma)(M_0)$, as desired. □

11. Traces in the general case

In this section we prove Theorem 9.1 in general, assuming that Theorem 0.5 is already known in the basic case $\operatorname{rank} G = d$. Thus throughout this section we assume $\operatorname{rank} G > d$. We keep the notation of Section 9.

The main idea is to analyze the subgroup generated by $\Gamma \cap V$ and its conjugate under any element $\gamma \in \Gamma$ which is in sufficiently general position with respect to V. For this we will first show that the algebraic group $H_{(\gamma)}$ generated by V and $\gamma V \gamma^{-1}$ is almost simple of rank d. We also show that $\Gamma \cap H_{(\gamma)}$ is a sufficiently general finite subgroup of $H_{(\gamma)}$. Thus by Theorem 0.5 in the basic case we know that $\Gamma \cap H_{(\gamma)}$ is a finite group of Lie type over $\mathbb{F}_V$. From this we deduce that for any $1 \neq v \in \Gamma \cap V$, the trace of $v\gamma v \gamma^{-1}$ in a suitable representation of G lies in $\mathbb{F}_V$. Finally, we use this information to show that all matrix coefficients of Γ in another representation of G lie in $\mathbb{F}_V$, as desired.

Subgroups generated by root groups. Fix a maximal torus $T \subset B$ and recall that V is a product of root groups in the center of U. Let $\Psi \subset \Phi$ denote the set of roots which are $\mathbb{Z}$-linear combinations of roots occurring in V. Take $\dot{w} \in N_G(T)$ such that $\dot{w} B \dot{w}^{-1}$ is the Borel subgroup opposite to B. For every $g \in G$ let $H_{(g)}$ denote the algebraic subgroup generated by V and gVg^{-1}.

Proposition 11.1. (a) *Ψ is a simple root system of rank d. If $d = 1$, it has type A_1; otherwise we have $(p, \Psi) = (2, B_2)$ or $(3, G_2)$.*

(b) *$H_{(\dot{w})}$ is connected almost simple with root system Ψ.*

(c) *For every $g \in B\dot{w}B$ the subgroup $H_{(g)}$ is conjugate to $H_{(\dot{w})}$.*

(d) *The complement $G \smallsetminus B\dot{w}B$ is a fiber of a constructible family of proper subvarieties of $\mathscr{G}$.*

Proof. Part (a) follows from the proof of Proposition 8.3. For (b) note that $\dot{w}$ transforms the highest root, resp. the highest short root, to its negative. Thus $\dot{w} V \dot{w}^{-1}$ is the product of root groups associated to the negatives of the roots occurring in V. Clearly $H_{(\dot{w})}$ is contained in the connected almost simple subgroup of G, normalized by T, with root system Ψ. Well-known facts on commutators ([17, Props. 33.3, 33.4, 33.5]) imply equality. For (c) we write $g = b\dot{w}b'$ with $b, b' \in B$ and calculate

$$H_{(g)} \;=\; \big\langle V, b\dot{w}b'Vb'^{-1}\dot{w}^{-1}b^{-1} \big\rangle \;=\; b\big\langle V, \dot{w}V\dot{w}^{-1}\big\rangle b^{-1} \;=\; bH_{(\dot{w})}b^{-1}.$$

Finally, (d) follows from the fact that $B\dot{w}B$ is the big cell in the Bruhat decomposition of G. □

Genericity. Let $\mathscr{H} \to \operatorname{\mathbf{Spec}} \mathbb{Z}$ denote the family of split connected adjoint groups with simple root system Ψ, and let H be its geometric fiber over the field k. For any $g \in B\dot{w}B$ we identify the adjoint group $H_{(g)}^{\mathrm{ad}}$ with H by means of a central isogeny $\pi \colon H_{(g)} \twoheadrightarrow H$.

Proposition 11.2. *The subgroup $H_{(g)}$ belongs to a constructible family of algebraic subgroups of $\mathscr{G}$, and π to a constructible family of homomorphisms.*

Proof. The maximal tori T and the root groups U_α form constructible families of algebraic subgroups of $\mathscr{G} \to \operatorname{\mathbf{Spec}} \mathbb{Z}$. The connected semisimple subgroups associated to a closed root subsystem $\Psi \subset \Phi$ are the Zariski-closures of $T \cdot \prod_{\alpha \in \Psi} U_\alpha$, so by Proposition 1.7 they also form a constructible family. If $\Psi = \{\pm\alpha\}$, where

α is the highest positive root, these subgroups occur as $H_{(g)}$ in any characteristic. For all other cases in the list of Proposition 11.1 (a) the characteristic p is fixed, but there is no further restriction. Thus in these cases the base of the constructible family must be restricted to $\mathbb{F}_p$. □

The following proposition says that the subgroup $\pi(\Gamma \cap H_{(\gamma)})$ is sufficiently general in H for every $\gamma \in \Gamma \cap B\dot{w}B$, provided that Γ is sufficiently general.

Proposition 11.3. *Consider a constructible family $\mathscr{K} \to \mathscr{T}$ of proper algebraic subgroups of $\mathscr{H}$. Assume that Γ is sufficiently general. Then for every element $\gamma \in \Gamma \cap B\dot{w}B$ and every point $t \in \mathscr{T}(k)$ we have $\pi(\Gamma \cap H_{(\gamma)}) \not\subset \mathscr{K}_t$.*

Proof. If $\pi(\Gamma \cap H_{(\gamma)}) \subset \mathscr{K}_t$, then both $\Gamma \cap V$ and $\Gamma \cap \gamma V \gamma^{-1}$ are contained in $\pi^{-1}(\mathscr{K}_t)$. This subgroup belongs to a constructible family of proper subgroups of $H_{(\gamma)}$, by Proposition 11.2. The definition of $H_{(\gamma)}$ shows that not both V and $\gamma V \gamma^{-1}$ can be contained in $\pi^{-1}(\mathscr{K}_t)$. Suppose $V \not\subset \pi^{-1}(\mathscr{K}_t)$. Then $V \cap \pi^{-1}(\mathscr{K}_t)$ is a fiber of a constructible family of proper algebraic subgroups of V. Viewing it as a subgroup of G, Theorem 4.2 implies that

$$\big|\Gamma \cap V\big| \;=\; \big|\Gamma \cap V \cap \pi^{-1}(\mathscr{K}_t)\big| \;\le\; c \cdot q_\Gamma^{\dim(V \cap \pi^{-1}(\mathscr{K}_t))} \;\le\; c \cdot q_\Gamma^{d-1}$$

for some fixed constant c if Γ is sufficiently general. But this contradicts Corollary 8.13, if q_Γ is large, as guaranteed by Proposition 6.1. The analogous arguments apply when $\gamma V \gamma^{-1} \not\subset \pi^{-1}(\mathscr{K}_t)$. □

Algebraic properties of certain traces. The representations ρ, ρ_ℓ, and ρ_s were defined at the end of Section 9. We will need the following information on their traces:

Lemma 11.4. *If $d = 1$, the function $\operatorname{Tr} \rho | H_{(\dot{w})}$ is nonconstant.*

Proof. It suffices to show that the formal character of the restriction is not congruent modulo p to a multiple of the trivial character. The weights are elements of the character space $\mathbb{R}\Phi$, and their restriction to $H_{(\dot{w})}$ is obtained by orthogonal projection to the subspace $\mathbb{R}\Psi$. If $V = U_\alpha$, we will prove that α occurs exactly once in the restriction of the formal character.

Suppose first that α is the highest root in Φ. Then for every $\beta \in \Phi \smallsetminus \{\pm\alpha\}$ we have $|(\beta, \alpha)| < (\alpha, \alpha)$, since otherwise $|\beta|^2 > |\alpha|^2$, which contradicts the fact that α is a longest possible root. Therefore the weight α occurs in the restriction exactly once, as desired. If α is not the highest root in Φ, by Proposition 8.3 we have a nonstandard case and α is the highest short root. Then by definition we have $\rho = \rho_s$, so only short roots occur as nonzero weights β. The same argument then applies. □

Proposition 11.5. *For any element $v \in V$ in the open T-orbit, the function $G \longrightarrow k^d$, $g \mapsto \operatorname{Tr} \rho(vgvg^{-1})$ is nonconstant.*

Proof. We first consider the case $d = 1$. It suffices to verify the assertion on elements of the form $g = \dot{w}t$ with $t \in T \cap H_{(\dot{w})}$. For these the product $vgvg^{-1}$ can be calculated purely inside $H_{(\dot{w})}$. Lifting everything to SL_2, we can compute explicitly. Suppose that

$$v = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \quad \dot{w} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad t = \begin{pmatrix} x & 0 \\ 0 & x^{-1} \end{pmatrix}.$$

Then

$$vgvg^{-1} = \begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix} \cdot \begin{pmatrix}1 & 0\\ x^2 & 1\end{pmatrix} = \begin{pmatrix}1+x^2 & 1\\ x^2 & 1\end{pmatrix}.$$

The trace of this matrix is $2 + x^2$, which is a nonconstant function of x. As any central function on SL_2 is a polynomial in the trace, we deduce that any nonconstant central function on $H_{(\dot{w})}$ remains nonconstant on elements of the above form $vgvg^{-1}$. Thus in this case the desired assertion follows from Lemma 11.4.

In the case $d = 2$ we can avoid explicit calculations by the following argument. If the function is constant, the calculation

$$\operatorname{Tr}\rho\big((tvt^{-1})g(tvt^{-1})g^{-1}\big) \;=\; \operatorname{Tr}\rho\big(v(t^{-1}gt)v(t^{-1}gt)^{-1}\big)$$

for any $t \in T$ shows that this constant value is independent of v in the open orbit. It is therefore attained for all $v \in V$. In particular, the function $g \mapsto \operatorname{Tr}\rho_\ell(v_\ell g v_\ell g^{-1})$ is constant for any $1 \neq v_\ell \in U_{\alpha_\ell}$. But this contradicts what we have just proved in the case $V = U_{\alpha_\ell}$. □

In the following lemma we assume $d = 2$, and let σ_ℓ and σ_s denote the simple subquotients of the adjoint representation of $H_{(g)}$ associated to the long, resp. short, roots. Accordingly, we view $\sigma = (\sigma_\ell, \sigma_s)$ as a representation over k^2.

Lemma 11.6. *If $d = 2$, for any $g \in B\dot{w}B$ we have $\operatorname{Tr}\rho|H_{(g)} = \operatorname{Tr}\sigma$.*

Proof. By Proposition 11.1 (c) it is enough to work with $g = \dot{w}$. As in Lemma 11.4 the assertion depends only on the formal characters. Since we are here in the nonstandard case with $\Psi \neq \Phi$, we must have $(p, \Psi) = (2, B_2)$ and $\Phi = B_n$, C_n, or F_4. Thus we must show that the formal characters are congruent modulo 2.

Consider first the nonzero weights in ρ_ℓ or ρ_s. These are precisely the long (resp. short) roots $\alpha \in \Phi$. Those in Ψ occur on both sides of the desired congruence, so let us suppose $\alpha \notin \Psi$. Write $\alpha = \lambda + \lambda^\perp$ with $\lambda \in \mathbb{R}\Psi$ and $0 \neq \lambda^\perp \in (\mathbb{R}\Psi)^\perp$. Since $\Psi = B_2$, the image of α under the longest Weyl group element of Ψ is $-\lambda + \lambda^\perp$. Its negative $\lambda - \lambda^\perp$ also occurs, is different from α, and has the same restriction to $H_{(\dot{w})}$. The contribution of each such pair is congruent to $0 \bmod 2$, as desired.

For weight 0 we compare multiplicities directly. It turns out that the multiplicity is even on both sides: this results from the general description of the adjoint representation [13] or [28, Prop. 1.11]. The proposition follows. □

Arithmetic properties of traces. Fix a nontrivial element $v \in \Gamma \cap V$.

Proposition 11.7. *If Γ is sufficiently general, we have $\operatorname{Tr}\rho(v\gamma v\gamma^{-1}) \in \mathbb{F}_V$ for every $\gamma \in \Gamma \cap B\dot{w}B$.*

Proof. By Proposition 11.3 the subgroup $\Delta := \pi(\Gamma \cap H_{(\gamma)})$ is sufficiently general in H, so by the basic case of Theorem 0.5 we have $(H^F)^{\mathrm{der}} \subset \Delta \subset H^F$ for some Frobenius map $F \colon H \to H$. In the case $d = 2$ this cannot be a standard Frobenius map, since otherwise Δ contains nontrivial elements of some root group; hence so does Γ, which contradicts Assumption 8.14. Therefore the finite field underlying this Frobenius map is the given field $\mathbb{F}_V \subset k^d$.

In the case $d = 1$ we can therefore choose an identification $H \cong \mathrm{PGL}_{2,k}$ so that $\mathrm{PGL}_2(\mathbb{F}_V)^{\mathrm{der}} \subset \Delta \subset \mathrm{PGL}_2(\mathbb{F}_V)$. The universal covering $\mathrm{SL}_{2,k} \to \mathrm{PGL}_{2,k}$ factors through a unique central isogeny $\varpi \colon \mathrm{SL}_{2,k} \to H_{(\gamma)}$. Since both v and $\gamma v\gamma^{-1}$ are unipotent elements in $\Gamma \cap H_{(\gamma)}$, they lift canonically to elements of $\mathrm{SL}_2(\mathbb{F}_V)$. Therefore $v\gamma v\gamma^{-1} \in \varpi\big(\mathrm{SL}_2(\mathbb{F}_V)\big)$. Now, it is known that every irreducible algebraic representation of $\mathrm{SL}_{2,k}$ can be defined already over $\mathbb{F}_p$. In particular, the traces of $\mathrm{SL}_2(\mathbb{F}_V)$ in any algebraic representation lie in $\mathbb{F}_V$. Applying this fact to the representation $\rho \circ \varpi$, the lemma follows.

In the case $d = 2$ we have $\operatorname{Tr} \rho(v\gamma v\gamma^{-1}) = \operatorname{Tr} \sigma(v\gamma v\gamma^{-1})$ by Lemma 11.6. Thus it suffices to prove $\operatorname{Tr} \sigma(\delta) \in \mathbb{F}_V$ for every $\delta \in \Delta$. Let $\varphi\colon H \to H$ denote any nonstandard isogeny whose square is a standard Frobenius map relative to the prime field $\mathbb{F}_p$. By [28, Prop. 1.11] and the succeeding remarks we know that $\sigma_\ell \cong \sigma_s \circ \varphi$. On the other hand, recall that F^2 is a standard Frobenius map relative to the field $\mathbb{F}_V$ of order p^r. Thus the classification of isogenies ([28, Thm. 1.7]) implies that F differs from φ^r by an automorphism. Furthermore, recall that $r = 2e + 1$ by Lemma 10.9. All together this implies that

$$\sigma_s \circ F \;\cong\; \sigma_s \circ \varphi^{2e+1} \;\cong\; \sigma_\ell \circ \varphi^{2e} \;\cong\; \mathrm{Frob}_{p^e} \circ \sigma_\ell,$$

and similarly

$$\sigma_\ell \circ F \;\cong\; \sigma_s \circ \varphi \circ \varphi^{2e+1} \;\cong\; \mathrm{Frob}_{p^{e+1}} \circ \sigma_s.$$

For elements $\delta \in \Delta \subset H^F$, it follows that $\operatorname{Tr} \sigma_s(\delta) = \operatorname{Tr} \sigma_\ell(\delta)^{p^e}$ and $\operatorname{Tr} \sigma_\ell(\delta) = \operatorname{Tr} \sigma_s(\delta)^{p^{e+1}}$. This implies that $\operatorname{Tr} \sigma(\delta) \in \mathbb{F}_V$, as desired. □

From traces to matrices. The information in Proposition 11.7 involves a quadratic expression in the matrix coefficients of $\rho(\gamma)$. To linearize it, we first pass to the ring E of k^d-linear endomorphisms of the representation space of ρ. This is a direct sum of d matrix rings over k.

On the other hand, our information relates only the element $\rho(v)$ with its conjugates. Thus we can use it to access only the following subquotient. Let $E' \subset E$ be the smallest G-invariant k^d-submodule containing the element $\rho(v)$. Let $E'^\perp$ be the orthogonal complement of E' with respect to the trace form

$$E \times E \longrightarrow k^d,\ (f_1, f_2) \mapsto \operatorname{Tr}(f_1 f_2).$$

The k^d-module $M := E'/E' \cap E'^\perp$ then carries a natural representation of G, which we denote by σ. By construction the trace form induces a nondegenerate symmetric k^d-bilinear pairing $\overline{\operatorname{Tr}}\colon M \times M \to k^d$.

Lemma 11.8. *σ belongs to a constructible family of representations of $\mathscr{G}$.*

Proof. As ρ varies in a constructible family, so does E. The submodule E' can be described as the image of the morphism

$$(\mathbb{G}_a^d \times G)^n \longrightarrow E,\ \big((x_i, g_i)\big) \mapsto \sum_{i=1}^{n} x_i \cdot \rho(g_i v g_i^{-1})$$

for any sufficiently large n. This morphism depends only on v, so it is a fiber of some morphism of constructible families. Every linear subspace is already closed, so Proposition 1.7 shows that E' varies in a constructible family. Its orthogonal complement $E'^\perp$ is characterized by a Zariski-closed condition, so it also varies in a constructible family. Therefore so does $E' \cap E'^\perp$, and by Proposition 1.6 we may assume that the dimensions of all these subspaces are locally constant over the base. Then the quotient space M can be constructed in the family and carries a natural representation of $\mathscr{G}$, as desired. □

Let $m_0 \in M$ denote the image of $\rho(v) \in E'$. Then for every $g \in G$ we have

$$\overline{\operatorname{Tr}}\big(m_0, \sigma(g)(m_0)\big) \;=\; \operatorname{Tr} \rho(vgvg^{-1}). \tag{11.9}$$

Another direct calculation shows that

$$\overline{\mathrm{Tr}}\big(\sigma(g)(m),\sigma(g)(m')\big)=\overline{\mathrm{Tr}}\big(m,m'\big) \tag{11.10}$$

for all m, $m'\in M$ and $g\in G$. Combining (11.9) and (11.10) with Proposition 11.7 we find

$$\begin{aligned}\overline{\mathrm{Tr}}\big(\sigma(\gamma)(m_0),\sigma(\gamma')(m_0)\big)&=\overline{\mathrm{Tr}}\big(m_0,\sigma(\gamma^{-1}\gamma')(m_0)\big)\\&=\mathrm{Tr}\,\rho\big(v(\gamma^{-1}\gamma')v(\gamma^{-1}\gamma')^{-1}\big)\\&\in\mathbb{F}_V\end{aligned} \tag{11.11}$$

for all γ, $\gamma'\in\Gamma$ with $\gamma^{-1}\gamma'\in B\dot{w}B$. Let $M_0\subset M$ be the $\mathbb{F}_V$-subspace generated by the Γ-orbit $O_\Gamma(m_0)$.

Lemma 11.12. *If Γ is sufficiently general, the natural map $M_0\otimes_{\mathbb{F}_V}k^d\longrightarrow M$ is an isomorphism.*

Proof. Suppose first that the map is not surjective. Then the Γ-orbit of m_0 generates a proper k^d-submodule $N\subsetneqq M$. In other words, Γ is contained in the proper subvariety

$$X\;:=\;\big\{\,g\in G\;\big|\;\sigma(g)(m_0)\in N\,\big\}\;\subsetneqq\;G. \tag{11.13}$$

Note that the submodules N are indexed by some Grassmannian and thus by a constructible family, and σ varies in a constructible family of representations by Lemma 11.8. Thus X also belongs to a constructible family. By Proposition 2.4 it cannot contain Γ if Γ is sufficiently general. Therefore the desired map is surjective.

Suppose that the map is not injective. Then we can find elements $\gamma_i\in\Gamma$ for $1\le i\le\ell$ so that the vectors $\sigma(\gamma_i)(m_0)\in M$ are $\mathbb{F}_V$-linearly independent but k^d-linearly dependent. Moreover, we may assume $\ell\le n+1$, where n is the smallest number of generators of M over k^d. As the pairing $\overline{\mathrm{Tr}}$ is nondegenerate, the set of elements

$$\big\{\,m\in M\;\big|\;\forall\,1\le i\le\ell\colon\;\overline{\mathrm{Tr}}(\sigma(\gamma_i)(m_0),m)\in\mathbb{F}_V\,\big\}$$

is then contained in a proper k^d-submodule $N\subsetneqq M$. Now (11.11) implies that $\sigma(\gamma)(m_0)\in N$ for every $\gamma\in\Gamma\cap\bigcap_{i=1}^{\ell}\gamma_iB\dot{w}B$. In other words, these elements γ lie in the subvariety X of (11.13), or equivalently

$$\Gamma\;\subset\;X\cup\bigcup_{i=1}^{\ell}\gamma_i\big(G\smallsetminus B\dot{w}B\big).$$

Each term in this finite union is a proper subvariety of G which belongs to a constructible family. As the number of terms is bounded, the whole union belongs to a constructible family. Thus Proposition 2.4 yields a contradiction if Γ is sufficiently general. Therefore the map in question is injective, and hence an isomorphism, as desired. □

Proof of Theorem 9.1 in the general case. Proposition 11.5 and Formula 11.9 imply that the representation σ is nontrivial. By Lemma 11.8 it varies in a constructible family. By construction the subspace $M_0\subset M$ is Γ-invariant, and by Lemma 11.12 it constitutes a model of M over $\mathbb{F}_V$. This finishes the proof of Theorem 9.1. □

12. Finite subgroups of general linear groups

In this section we prove Theorems 0.1 through 0.4 of the introduction. We begin with the following technical lemma.

Lemma 12.1. *For any constructible family of algebraic groups $\mathscr{G} \to \mathscr{S}$, any constructible family of fiberwise nowhere dense algebraic subgroups $\mathscr{L} \to \mathscr{T}$, and every positive integer N, there exists a constructible family of fiberwise nowhere dense algebraic subgroups $\mathscr{L}_N \to \mathscr{T}_N$ of $\mathscr{G} \to \mathscr{S}$ with the following property. For any finite subgroup Γ of a geometric fiber $\mathscr{G}_s$, if $\left[\Gamma : \Gamma \cap \mathscr{L}_t\right] \leq N$ for some point t of $\mathscr{T}$ above s, then $\Gamma \subset \mathscr{L}_{N,t_N}$ for some point t_N of $\mathscr{T}_N$ above s.*

Proof. The conjugates of $\mathscr{L}$ form a constructible family of subgroups, indexed by the total space $\mathscr{G} \times_{\mathscr{S}} \mathscr{T}$. Therefore the intersections of at most N conjugates also form a constructible family, say $\mathscr{L}' \to \mathscr{T}'$. By Proposition 1.13, after stratifying $\mathscr{T}'$ if necessary, there is a constructible family of subgroups $\mathscr{N} \to \mathscr{T}'$ of $\mathscr{G}$ which is fiberwise the normalizer of $\mathscr{L}'$.

Now consider any $\Gamma \subset \mathscr{G}_s$ with $\left[\Gamma : \Gamma \cap \mathscr{L}_t\right] \leq N$ for some t. By construction there is a point t' of $\mathscr{T}'$ above s with $\mathscr{L}'_{t'} = \bigcap_{\gamma\in\Gamma} \gamma\mathscr{L}_t\gamma^{-1}$. Then $\Gamma \subset \mathscr{N}_{t'}$, and we have

$$\left[\Gamma : \Gamma \cap \mathscr{L}'_{t'}\right] \leq N!.$$

The $N!$-tuples $(n_1, \ldots, n_{N!})$ of sections of $\mathscr{N}$ are indexed by the $N!$-fold fiber product $\mathscr{N}^{N!}$. Thus the union of the translates $n_i\mathscr{L}'$ is a constructible family of subvarieties of $\mathscr{N}$. The condition for a fiber of this family to be a subgroup is Zariski-closed, so the subgroups arising in this way form a constructible family $\mathscr{L}_N \to \mathscr{T}_N$. Clearly it has the desired property vis-à-vis Γ. □

Next we will show that simple quotients of connected linear algebraic groups G vary in a constructible family if the groups G do so. To fix ideas, by a *simple quotient of* G we mean the epimorphism $G \twoheadrightarrow H$ to a simple direct factor of the adjoint group $(G/\operatorname{Rad}_u G)^{\mathrm{ad}}$, where $\operatorname{Rad}_u G$ denotes the unipotent radical of G. We call two simple quotients $f_1 : G \twoheadrightarrow H_1$ and $f_2 : G \twoheadrightarrow H_2$ *equivalent* if there exists an isomorphism $\psi\colon H_1 \xrightarrow{\sim} H_2$ such that $f_2 = \psi \circ f_1$. Note that here we do not allow arbitrary epimorphisms with a simple target group. The reason is that the composite of a simple quotient map $f : G \twoheadrightarrow H$ with an arbitrary Frobenius map on H still constitutes a quotient in the category of algebraic groups, but all these do not form a constructible family.

Lemma 12.2. *Consider a constructible family of connected linear algebraic groups $\mathscr{G} \to \mathscr{S}$. Let $\mathscr{H} \to \mathbf{Spec}\,\mathbb{Z}$ be the constructible family of connected adjoint groups associated to a simple root system Φ. Then there exists a morphism of finite type $\mathscr{S}' \to \mathscr{S}$ and a homomorphism $f : \mathscr{G} \times_{\mathscr{S}} \mathscr{S}' \to \mathscr{H} \times \mathscr{S}'$ such that*

(a) *f is a simple quotient in every geometric fiber, and*
(b) *every simple quotient with root system Φ of any geometric fiber of $\mathscr{G}$ is equivalent to one occurring in f.*

Proof. By Noetherian induction it suffices to prove this over a neighborhood of any fixed generic point η of $\mathscr{S}$. After shrinking $\mathscr{S}$ and passing to a finite covering, we may suppose that all simple quotients of $\mathscr{G}_\eta$ of type Φ can be defined over the residue field $k(\eta)$. Consider one of them, say $f_\eta : \mathscr{G}_\eta \to \mathscr{H}_\eta$. As η is a generic point of $\mathscr{S}$, this morphism extends to some neighborhood. After shrinking $\mathscr{S}$, the extension $f : \mathscr{G} \to \mathscr{H}_{\mathscr{S}}$ remains a homomorphism, as well as surjective (compare Proposition 1.7). Since f_η is a simple quotient, the image of its derivative

$$df_\eta \colon \operatorname{Lie}\mathscr{G}_\eta \twoheadrightarrow \operatorname{Lie}\bigl(\mathscr{G}_\eta / \operatorname{Rad}_u \mathscr{G}_\eta\bigr) \longrightarrow (\operatorname{Lie}\mathscr{H}) \otimes k(\eta)$$

contains all root spaces. This assertion remains true in a neighborhood of η, where f remains a simple quotient, as desired. □

Theorem 12.3. *For every constructible family of linear algebraic groups $\mathscr{G} \to \mathscr{S}$ there exists a constructible family of algebraic subgroups $\mathscr{H} \to \mathscr{T}$ with the following property. For any finite subgroup Γ of a geometric fiber $\mathscr{G}_s$, there exists a point t of $\mathscr{T}$ above s, such that $\Gamma \subset \mathscr{H}_t$ and for every simple quotient $f : \mathscr{H}_t^\circ \twoheadrightarrow H_1$ there is a Frobenius map $F \colon H_1 \to H_1$ with $(H_1^F)^{\mathrm{der}}$ simple and*

$$\bigl(H_1^F\bigr)^{\mathrm{der}} \;\subset\; f\bigl(\Gamma \cap \mathscr{H}_t^\circ\bigr) \;\subset\; H_1^F.$$

Proof. As the dimension of $\mathscr{G}_s^\circ$ is bounded, only finitely many root systems can occur for its simple quotients. Let Φ be one of them, $\mathscr{H}_\Phi \to \mathbf{Spec}\,\mathbb{Z}$ the associated constructible family of connected adjoint groups, and $\mathscr{K}_\Phi \to \mathscr{T}_\Phi$ the constructible family of fiberwise nowhere dense algebraic subgroups given by Theorem 0.5. Thus any $\mathscr{K}_\Phi$-general finite subgroup of a geometric fiber $H_\Phi = \mathscr{H}_{\Phi,t}$ is trapped between $(H_\Phi^F)^{\mathrm{der}}$ and H_Φ^F for some Frobenius map F on H_Φ.

The groups $\mathscr{G}_s^\circ$ form a constructible family $\mathscr{G}^\circ$, for instance by Proposition 1.9. Let $f_\Phi \colon \mathscr{G}^\circ \times_{\mathscr{S}} \mathscr{S}'_\Phi \to \mathscr{H}_\Phi \times \mathscr{S}'_\Phi$ be the homomorphism given by Lemma 12.2. The inverse image of $\mathscr{K}_\Phi$ is a constructible family of fiberwise nowhere dense algebraic subgroups $\mathscr{L}_\Phi$ of $\mathscr{G}^\circ$, and thus of $\mathscr{G}$. Let N be an upper bound for the index $[\mathscr{G}_s : \mathscr{G}_s^\circ]$ in all fibers, and $\mathscr{L}_{\Phi,N}$ the constructible family of fiberwise nowhere dense algebraic subgroups of $\mathscr{G}$ given by Lemma 12.1.

Now consider any finite subgroup Γ of a geometric fiber $\mathscr{G}_s$. If the desired assertion does not hold with $\mathscr{H}_t = \mathscr{G}_s$, there exists a simple quotient $f_\Phi : \mathscr{G}_s^\circ \twoheadrightarrow H_\Phi$ for which the image $f_\Phi\bigl(\Gamma \cap \mathscr{G}_s^\circ\bigr)$ is not $\mathscr{K}_\Phi$-general, i.e., is contained in some fiber of $\mathscr{K}_\Phi$. Then $\Gamma \cap \mathscr{G}_s^\circ$ is contained in a fiber of $\mathscr{L}_\Phi$. Moreover, its index in Γ is at most N, so by Lemma 12.1 the whole group Γ is contained in a fiber of $\mathscr{L}_{\Phi,N}$. By induction on fiber dimension we may assume that the theorem is already proved for $\mathscr{L}_{\Phi,N}$ in place of $\mathscr{G}$. We take the constructible families of algebraic subgroups of $\mathscr{L}_{\Phi,N}$ determined by Theorem 12.3 for all possible Φ, and define $\mathscr{H}$ as the disjoint union of these with the original family $\mathscr{G}$. This family clearly has the desired properties. □

Proof of Theorem 0.2. We apply Theorem 12.3 to the ambient group $\mathrm{GL}_{n,\mathbf{Spec}\,\mathbb{Z}}$. To remain in keeping with the notation in the introduction, we abbreviate a typical geometric fiber of the resulting family $\mathscr{H}$ by $G := \mathscr{H}_t$. By Proposition 1.4 the index $[G : G^\circ]$ is bounded, say $\leq N$. As the dimension is bounded, so are the type and the number of simple quotients of G. Now consider a finite subgroup $\Gamma \subset \mathrm{GL}_n(k)$, where k is any field. Without loss of generality, we may assume k algebraically closed. By Theorem 12.3 we can choose t such that $\Gamma \subset G$, and for every simple quotient $f_i \colon G^\circ \twoheadrightarrow H_i$ there exists a Frobenius map $F \colon H_i \to H_i$ so that $(H_i^F)^{\mathrm{der}}$ is simple and

$$\bigl(H_i^F\bigr)^{\mathrm{der}} \;\subset\; f_i\bigl(\Gamma \cap G^\circ\bigr) \;\subset\; H_i^F.$$

Define

$$\begin{aligned} G_1 &:= G^\circ, & \Gamma_1 &:= (\Gamma \cap G_1)^{\mathrm{der}} \cdot (\Gamma \cap G_2),\\ G_2 &:= \textstyle\bigcap_i \ker f_i, & \Gamma_2 &:= \Gamma \cap G_2,\\ G_3 &:= \mathrm{Rad}_u\, G_1, & \Gamma_3 &:= \Gamma \cap G_3. \end{aligned}$$

We claim that these subgroups have the desired properties. Clearly they are normal subgroups of Γ that are contained in each other. The group Γ_2 is the kernel of the homomorphism $\Gamma \cap G_1 \to \prod_i H_i^F$, and Γ_1 is the kernel of $\Gamma \cap G_1 \to \prod_i \big(H_i^F/(H_i^F)^{\mathrm{der}}\big)$. Let r be an upper bound for the number of simple factors, and let m be an upper bound for the index of their root lattices in their weight lattices. Using Theorem 3.4 (b) we deduce

$$\big[\Gamma : \Gamma_1\big] \;=\; \big[\Gamma : \Gamma \cap G_1\big] \cdot \big[\Gamma \cap G_1 : \Gamma_1\big] \;\leq\; Nm^r \;=:\; J'(n),$$

whence Theorem 0.2 (a). The next subquotient Γ_1/Γ_2 is embedded into the product of noncommutative simple groups $(H_i^F)^{\mathrm{der}}$ and surjects onto each factor. By Goursat's lemma we obtain an isomorphism from Γ_1/Γ_2 to the product of some of the $(H_i^F)^{\mathrm{der}}$. This implies Theorem 0.2 (b). Assertion (c) follows from the fact that Γ_2/Γ_3 is contained in the center of the connected reductive group G_1/G_3. Finally, (d) holds by construction. □

Proof of Theorem 0.4. Consider a finite subgroup $\Gamma \subset \mathrm{GL}_n(k)$, where $p := \mathrm{char}(k) > 0$. Let $\Gamma_3 \subset \Gamma_2 \subset \Gamma_1$ be the subgroups given by Theorem 0.2. Let Z be the maximal abelian normal subgroup of Γ_2 of order prime to p. Being a characteristic subgroup of Γ_2, it is also normal in Γ. In the product

$$\big[\Gamma : Z\big] \;=\; \big[\Gamma : \Gamma_1\big] \cdot \big[\Gamma_1 : \Gamma_2\big] \cdot \big[\Gamma_2 : Z\big],$$

the first factor is $\leq J'(n)$, by Theorem 0.2 (a). The second factor is at most the cube of its p-part, by Theorem 3.4 (d). Thus it suffices to prove the same for the third factor. This term is, in fact, bounded by the square of its p-part, by the following lemma applied to Γ_2 in place of Γ:

Lemma 12.4. *Consider a finite group Γ with a normal Sylow p-subgroup $\Gamma_{(p)}$ and abelian factor group $\Gamma/\Gamma_{(p)}$. Then the maximal abelian normal subgroup $Z \subset \Gamma$ of order prime to p has index $\leq |\Gamma_{(p)}|^2$.*

Proof. Write Γ as a semidirect product $\Gamma_{(p)} \rtimes \Delta$. Then Z can be described as the kernel of the conjugation action $\Delta \to \mathrm{Aut}(\Gamma_{(p)})$. In other words, the factor group Δ/Z acts faithfully on $\Gamma_{(p)}$. Choose a composition series of $\Gamma_{(p)}$ as a group with Δ-action. The successive quotients M_i are elementary abelian p-groups, that is, $\mathbb{F}_p$-vector spaces, with irreducible representations of Δ. As Δ is abelian, each M_i can be viewed as a 1-dimensional vector space over a field $\mathbb{F}_{p^{r_i}}$, on which Δ acts through the multiplicative group $\mathbb{F}_{p^{r_i}}^\times$. Now, since Δ/Z has order prime to p, it still acts faithfully on the product of all M_i. Therefore

$$\big|\Delta/Z\big| \;\leq\; \prod_i \big|\mathbb{F}_{p^{r_i}}^\times\big| \;\leq\; \prod_i \big|M_i\big| \;=\; \big|\Gamma_{(p)}\big|.$$

It follows that

$$\big[\Gamma : Z\big] \;=\; \big|\Gamma_{(p)}\big| \cdot \big|\Delta/Z\big| \;\leq\; \big|\Gamma_{(p)}\big|^2,$$

which proves Lemma 12.4. This also finishes the proof of Theorem 0.4. □

References

[1] Borel, A., *Linear algebraic groups*, Graduate Texts in Math. **126**, New York: Springer (1991). MR1102012 (92d:20001)

[2] Brauer, R., Feit, W., An analogue of Jordan's theorem in characteristic p, *Annals of Math.* (2) **84** (1966), 119–131. MR0200350 (34:246)

[3] Carter, R. W., *Simple Groups of Lie Type*, London: Wiley (1972). MR0407163 (53:10946)

[4] Carter, R. W., *Finite Groups of Lie Type, Conjugacy Classes and Complex Characters*, Chichester: Wiley (1985). MR794307 (87d:20060)

[5] Collins, Michael J., On Jordan's theorem for complex linear groups., *J. Group Theory* **10** (2007), 411–423. MR2334748 (2008g:20106)

[6] Collins, Michael J., Modular analogues of Jordan's theorem for finite linear groups. *J. Reine Angew. Math.* **624** (2008), 143–171. MR2456628 (2009j:20071)

[7] Demazure, M., Grothendieck, A., (Eds.), *Schémas en Groupes I–III*, Séminaire de Géométrie Algébrique du Bois Marie 1962/64, SGA3, Lect. Notes Math. 151–153, Berlin: Springer (1970).

[8] Dickson, L. E., *Linear groups: with an exposition of the Galois field theory*, Leipzig: B. G. Teubner (1901).

[9] Grothendieck, A., Dieudonné, J. A., *Éléments de Géométrie Algébrique I*, EGA1, Berlin: Springer (1971).

[10] Grothendieck, A., *Étude locale des schémas et des morphismes de schémas*, Éléments de Géométrie Algébrique IV, EGA4, *Publ. Math. IHES* **20** (1964), **24** (1965), **28** (1966), **32** (1967). MR0173675 (30:3885); MR0199181 (33:7330); MR0217086 (36:178); MR0238860 (39:220)

[11] Guralnick, R. M., Small representations are completely reducible, *J. Algebra* **220** (1999), 531–541. MR1717357 (2000m:20018)

[12] Hartshorne, R., *Algebraic Geometry*, Graduate Texts in Math. **52**, New York: Springer (1977). MR0463157 (57:3116)

[13] Hiss, G., Die adjungierten Darstellungen der Chevalley-Gruppen, *Arch. Math.* **42** (1984), 408–416. MR756692 (85k:20134)

[14] Hogeweij, G. M. D., Almost Classical Lie Algebras I, *Indagationes Math.* **44** (1982), 441–460. MR683531 (84f:17007)

[15] Hrushovski, E., Pillay, A., Definable subgroups of algebraic groups over finite fields, *J. reine angew. Math.* **462** (1995), 69–91. MR1329903 (97f:20059)

[16] Hrushovski, E., Wagner, F., Counting and dimensions. *Model theory with applications to algebra and analysis.* Vol. 2, 161–176, London Math. Soc. Lecture Note Ser., 350, Cambridge Univ. Press, Cambridge, 2008. MR2436141 (2009k:03042)

[17] Humphreys, J. E., *Linear Algebraic Groups*, Graduate Texts in Math. **21**, New York: Springer (1975), (1981). MR0396773 (53:633)

[18] Humphreys, J. E., *Conjugacy Classes in Semisimple Algebraic Groups*, (Mathematical Surveys and Monographs; v. 43) Providence: AMS (1995). MR1343976 (97i:20057)

[19] Jantzen, J. C., *Representations of Algebraic Groups*, Boston: Academic Press (1987). MR899071 (89c:20001)

[20] Jordan, C., Mémoire sur les équations differentielles linéaires à intégrale algébrique, *J. für Math.* **84** (1878), 89–215.

[21] Katz, N. M., *Gauss Sums, Kloosterman Sums, and Monodromy Groups*, Annals of Math. Studies **116**, Princeton: Princeton Univ. Press (1988). MR955052 (91a:11028)

[22] Kneser, M., Semi-Simple Algebraic Groups, in: *Algebraic Number Theory*, Cassels, J.W.S., Fröhlich, A. (Eds.), London: Academic Press (1967), 250–265. MR0217077 (36:171)

[23] Lehrer, G. I., Rational tori, semisimple orbits and the topology of hyperplane complements, *Comment. Math. Helvetici* **67** (1992), 226–251. MR1161283 (93e:20065)

[24] Mumford, D., *Abelian Varieties*, Oxford: Oxford Univ. Press (1974). MR0282985 (44:219)

[25] Mumford, D., *Geometric Invariant Theory*, Berlin: Springer (1965). MR0214602 (35:5451)

[26] Nori, M. V., On subgroups of $\mathrm{GL}_n(\mathbb{F}_p)$, *Inventiones Math.* **88** (1987), 257–275. MR880952 (88d:20068)

[27] Pink, R., The Mumford-Tate conjecture for Drinfeld modules, *Publ. RIMS, Kyoto University* **33** (1997), 393–425. MR1474696 (98f:11062)

[28] Pink, R., Compact subgroups of linear algebraic groups, *J. Algebra* **206** (1998), 438–504. MR1637068 (99g:20087)

[29] Serre, J.-P., Propriétés galoisiennes des points d'ordre fini des courbes elliptiques, *Inventiones Math.* **15** (1972), 259–331. MR0387283 (52:8126)

[30] Steinberg, R., Endomorphisms of linear algebraic groups, *Mem. Amer. Math. Soc.* **80** (1968). MR0230728 (37:6288)

[31] Tate, J., Endomorphisms of Abelian Varieties over Finite Fields, *Inventiones Math.* **2** (1966), 134–144. MR0206004 (34:5829)

[32] Weisfeiler, B., Post-classification version of Jordan's theorem on finite linear groups, *Proc. Natl. Acad. Sci. USA* **81** (1984), 5278–5279. MR758425 (85j:20041)

Department of Mathematics, Indiana University, Bloomington, Indiana 47405
E-mail address: mjlarsen@indiana.edu

Department of Mathematics, ETH Zürich, CH - 8092 Zürich, Switzerland
E-mail address: pink@math.ethz.ch