Finite Subgroups of Algebraic Groups

by

Michael J. Larsen*

Department of Mathematics
Indiana University
Bloomington, IN 47405
U. S. A.

larsen@iu-math.math.indiana.edu

Richard Pink

Fakultät für Mathematik und Informatik
Universität Mannheim
D-68131 Mannheim
Germany

pink@math.uni-mannheim.de

September 30, 1998

Abstract

Generalizing a classical theorem of Jordan to arbitrary characteristic, we prove that every finite subgroup of GL_n over a field of any characteristic p possesses a subgroup of bounded index which is composed of finite simple groups of Lie type in characteristic p, a commutative group of order prime to p, and a p-group. While this statement can be deduced from the classification of finite simple groups, our proof is self-contained and uses methods only from algebraic geometry and the theory of linear algebraic groups. We believe that our results can serve as a viable substitute for classification in a range of applications in various areas of mathematics.

*partially supported by a Sloan grant and by NSF DMS-97-27553

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 [16] 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 $\mathrm{GL}_n(\overline{\mathbb{F}}_p)$ contains arbitrarily large subgroups of the form $\mathrm{SL}_n(\mathbb{F}_{p^r})$ which are simple modulo 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:

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 [30] 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 ([22] §3) and Gabber (see [17] 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 accomodate 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* $\to \mathbf{Spec}\,\mathbb{Z}$ *be the family of connected adjoint groups with a fixed simple root system* Φ*, and let* G *denote a geometric fiber of* . *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* $F\colon G \to G$ *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* $\to \mathbf{Spec}\,\mathbb{Z}$ *of connected adjoint groups with a fixed simple root system* Φ *possesses a representation* $\to \mathrm{GL}_n$ *with the following property. Consider any algebraically closed field* k *and any finite subgroup* Γ *of the associated geometric fiber* G *of . If every* Γ*-invariant subspace of* k^n *is* G*-invariant, the characteristic of* k *is positive and there exists a Frobenius map* $F\colon G \to G$ *so that* $(G^F)^{\mathrm{der}}$ *is simple and*

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

Applications: Our Theorem 0.3 gives much less than the classification of finite simple groups in that it applies only to linear groups and does not specify the finite set of exceptions for each n. Nevertheless, we believe that our results can serve as a viable substitute in a range of applications in group theory and number theory. For example, this is clearly so in Weisfeiler's work on strong approximation for finitely generated subgroups of algebraic groups [29]. This direction is pursued further in the paper [25] by the second author. We plan to discuss other group theoretic implications in a sequel.

The problem that originally motivated this paper arises in the theory of Drinfeld modules. By analogy with the construction of Tate modules of an abelian variety, one attaches a compatible system of Galois representations to a Drinfeld module or, more generally, a t-motive over a global field. These representations take values in GL_n over various completions of the ring $A := \mathbb{F}_p[t]$ or a finite extension thereof. We would like to prove an analogue of Serre's theorem [26] on the adelic openness of the image of Galois for such a system. The openness at any finite number of places has already been settled for Drinfeld modules by the second author in [23]. To analyze the adelic image, one must first study the reductions modulo all maximal ideals $\mathfrak{p} \subset A$. Although the residue fields $A/\mathfrak{p}$ have the same characteristic, they are unbounded in size. Thus one needs a systematic way of accounting for subgroups of $\mathrm{GL}_n(\mathbb{F}_{p^r})$ as r varies. This is provided by our results above.

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 [14].

Section 2 discusses the concept of sufficiently general finite subgroups and their basic properties: They can be made arbitrarily large, and supposed 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 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 non-trivial 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 non-standard 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 non-standard 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.

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

$$N_G(Y) \;:=\; \{\; g \in G \;\mid\; \forall y \in Y\colon\; gy \in Y \;\}.$$

The simultaneous *centralizer* of Y is

$$G_Y \;:=\; \{\; g \in G \;\mid\; \forall y \in Y\colon\; gy = y \;\}.$$

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

$$O_G(x) \;:=\; \{\; gx \;\mid\; g \in G \;\}.$$

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 $Z(G) := G_G$, the commutator subgroup 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 G°, the adjoint group of a connected reductive group G^{ad}. The Zariski closure of a subset X of an algebraic variety or scheme is denoted $\overline{X}$.

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

Symbol	**Page**	**Description**
$N_G(X)$	7	normalizer
G_X, G_x	7	(simultaneous) centralizer
$O_G(x)$	7	orbit, conjugacy class
$Z(\ )$	7	center
$(\ )^{\mathrm{der}}$	7	derived group
$(\ )^\circ$	7	identity component
$(\ )^{\mathrm{ad}}$	7	adjoint group
$\overline{(\ )}$	7	Zariski closure
$,\ ,\ldots$	9	constructible family of algebraic varieties
$,\quad ,\ldots$	9	base scheme of a constructible family
$_s,\ _t,\ldots$	9	geometric fiber
	9	pullback of a constructible family

	12, 27	constructible family of algebraic groups
	12	constructible family of algebraic subgroups
$(\ )^n$	13	n-fold fiber product with itself
$G = {}_s$	14, 27	algebraic group
Γ	14, 27	finite subgroup of G
$\mathbb{F}_q$	17	finite field with q elements
Φ	18, 27	root system of G
F	19	Frobenius map on G
q_F	19	numerical constant attached to F
G^F	19	fixed points of F
q_Γ	20, 27	numerical constant attached to Γ
$X = {}_t$	22	subvariety of G
k	27	algebraically closed base field
p	27	characteristic of k
Λ	28, 29, 44	subset of Γ
c_0	28	constant in Theorem 6.2
$(\ )^{\mathrm{rss}}$	29	subset of regular semisimple elements
$(\ )^{\mathrm{un}}$	29	subset of unipotent elements
Θ	30	maximal toric subgroup of Γ
Tor_Λ	31	set of maximal toric subgroups of Γ°_Λ
$\mathrm{Tor}^\natural_\Lambda$	31	subset of representatives under Γ°_Λ
B	34	Borel subgroup of G
U	34	unipotent radical of B
T	34	maximal torus of B
Φ^+	34	set of positive roots
$\mathbb{G}_a$	34	additive group
U_α	34	root group
α_ℓ, α_s	35	highest long and short roots
U^{run}	37	regular unipotent elements in U
V	39	non-trivial subgroup of $Z(U)$ normalized by B
d	40	dimension of V
$\mathbb{F}_V$	40	finite field attached to Γ and V
p^r	40	order of $\mathbb{F}_V$
p^e	40	Frobenius twist
ρ, ρ_ℓ, ρ_s	43	constituents of the adjoint representation
Ψ	50	root subsystem
$\dot{w}$	50	longest Weyl group element
$H_{(g)}$	50	subgroup generated by V and gVg^{-1}
$\mathrm{Rad}_u\, G$	56	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 $\to$ of finite type are viewed as forming a family of algebraic varieties $_s$, and operating with the total space 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 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 [7], [9]). 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 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], [12]). Usually, schemes will be denoted by calligraphic letters, varieties by roman letters. For example the fiber of a morphism of schemes $\to$ over a geometric point s of determines a variety, called simply the geometric fiber above s, and abbreviated by $X := _s$.

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

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

The pullback of such a constructible family by a morphism $\to$ will be abbreviated $:= \times$.

Definition 1.2 *A* morphism *from a constructible family* $\to$ *to a constructible family* $\to$ *consists of a morphism* $\to$ *and a morphism* $\varphi :$ $\longrightarrow$.

Definition 1.3 *A* constructible family of subvarieties *is a morphism for which* $\longrightarrow$ *is a closed embedding.*

Thus the geometric points t of parametrize a family of morphisms, resp. closed embeddings, of varieties $\varphi_t : _t \to _s$, where s denotes the corresponding geometric point of . Note that we do not rule out the possibility that the same subvariety of $_s$ 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* $\to$ *the dimension and the number of irreducible components of the geometric fibers* $_s$ *are bounded.*

Proof. See [8] Prop. 13.1.7, Cor. 9.7.9. **q.e.d.**

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* $' \to$ *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* $\to$ *there exists a stratification map* $' \to$ *such that* $, \longrightarrow '$ *is flat and its fiber dimension locally constant on* $'$*.*

Proof. See [8] Th. 11.1.1, Th. 13.1.3. **q.e.d.**

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* $\varphi :$ $\to$ *, there exist a stratification map* $' \to$ *and a constructible family of subvarieties* $\to '$ *of* $\to$ *with the following property. For any geometric point* t *of* *, with* t' *the corresponding geometric point of* $'$*, we have*

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

Proof. The image of φ is a constructible subset of (see [8] Prop. 9.2.6). Its closure in the total space is constructible by definition, so by [8] Prop. 9.5.3 the points t for which $\varphi_t(_t) = \varphi(\ )_t$ is dense in $\left(\overline{\varphi(\ )}\right)_t$ form a constructible subset of . This set contains all generic points of and therefore some open dense subscheme . Pulling the morphism φ back to $_1 :=$ $\smallsetminus$, by noetherian induction we already have a stratification map $_1' \to _1$ and a constructible family of subvarieties $_1 \subset _{1'}$ with the desired property over $_1$. Putting $' :=$ $\sqcup\, _1'$ and $:= \left(\overline{\varphi(\ )} \times \ \right) \sqcup _1$ the assertion follows over . **q.e.d.**

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

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

$$_{t'} = \overline{\left\{\, x \in _s \;\middle|\; \dim(\varphi_t^{-1}(x)) = d \,\right\}}.$$

Proof. The set of points of where the fiber dimension is d is constructible by [8] Prop. 9.2.6.1. Using this one proceeds as in the preceding proof. **q.e.d.**

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

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

Proof. We proceed as in the proof of [8] Th. 9.7.7. Consider a generic point η of . By [8] Cor. 4.6.8 there exists a finite extension K' of its residue field K such that every irreducible component Z_i of $_\eta \times_{\mathbf{Spec}\, K} \mathbf{Spec}\, K'$ is geometrically irreducible. Choose a morphism → of finite type where is integral with function field K'. For each i let $_i$ denote the Zariski-closure of Z_i in .

By [8] Th. 9.7.7 the fibers of $_i \to$ are geometrically irreducible in a neighborhood of the generic point. Thus after shrinking all these fibers are geometrically irreducible. Next, over the generic point none of these fibers is contained in any other. By [8] Cor. 9.5.2 the same is true over a whole neighborhood, so after shrinking again it is true over all of . Furthermore, the inclusion $\bigcup_i {}_i \subset$ is an equality over the generic point. By [8] Cor. 9.5.2 this remains true in a neighborhood, and so again without loss of generality over all of . We conclude that the fibers of the different families $_i \to$ are precisely the irreducible components of the fibers of ⟶ .

This solves the problem in a neighborhood of the generic point η. To finish, we apply noetherian induction to the pullback of to a suitable complement in , take the resulting family of subvarieties, and let → be its disjoint union with all $_i \to$. The desired assertion follows. **q.e.d.**

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 → and any constructible family of subvarieties → there is an integer n with the following property. Consider any geometric point s of and any collection I of geometric points of above s. Then there exists a subset $I' \subset I$ of at most n points, such that*

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

Proof. Fix d such that the fiber dimension of → 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 $Z := {}_{t_1} \cap \ldots \cap {}_{t_n}$ cannot be written as an intersection of fewer terms. Since all intersections of n terms form the constructible family of subvarieties

$$\underbrace{\times \ \ldots \times}_{n} \longrightarrow \underbrace{\times \ \ldots \times}_{n}, \tag{1.11}$$

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

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

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. **q.e.d.**

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

Corollary 1.12 *For any constructible family $\to$ and any constructible family of subvarieties $\to$ there exists another constructible family of subvarieties $\to$ with the following property. Consider any geometric point s of . Then for any non-empty collection I of geometric points of above s there exists a geometric point u of above s with*

$$\bigcap_{t \in I} {}_t \ = \ {}_u.$$

Conversely, every ${}_u$ 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. **q.e.d.**

Families of algebraic groups: For the general theory of algebraic groups and group schemes see [5], [1], or [12]. 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 $\to$, where and are of finite type over $\mathbf{Spec}\,\mathbb{Z}$. Similarly, a *constructible family of algebraic subgroups* of $\to$ consists of a morphism $\to$ and a closed subgroup scheme $\subset$. 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 $\to$ on a constructible family $\to$ is a morphism $\mu\colon \times \longrightarrow$ satisfying the usual associativity and identity axioms. If $\to$ is a vector

bundle and the action is linear, this is a *constructible family of representations*. The n-fold *fiber product* of with itself over will be denoted ^{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* $\to$ *which acts on a constructible family* $\to$. *Consider constructible families of subvarieties* $_1 \to$ *and* $_2 \to$ *of* $\to$. *Then there exist a stratification map* $' \to$ *and a constructible family* $' \to$ $'$ *of subvarieties of* $\to$ *with the following property. Take any geometric point* t *of* *and let* t' *and* s *denote the corresponding geometric points of* $'$*, resp.* . *Then we have*

$$ {}'_{t'} = \{\, g \in {}_s \;\mid\; g_{1,t} \subset {}_{2,t} \,\}. $$

If $_1 = {}_2$*, then* $' \to$ $'$ *is a family of algebraic subgroups, with* ${}'_{t'} = N_{\,s}({}_{1,t})$.

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 $_s$. To do this note that for every point y the set $\{g \in {}_s \mid gy \in {}_{2,t}\}$ is Zariski-closed. The transporter is the intersection of these as y runs through $_{1,t}$, so it is closed.

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

$$ \mu^{-1}({}_2) \cap \big(\ \times\ {}_1\big) \ \subset\ \ \times\ {}_1. $$

By [8] Cor. 9.5.2 the points $g \in$ over which this inclusion is an equality form a constructible subset . 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 $\subset$. We conclude by noetherian induction, as in the proof of Proposition 1.7. (For other approaches, see [5] Exp.VI$_{\rm B}$ §6.1 or [15] §2.6.) **q.e.d.**

Applying this to the conjugation action of 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* $\to$ *which acts on a constructible family* $\to$ *, there exists a constructible family* $\to$ *of algebraic subgroups of* $\to$ *with the following property. Take any geometric point* s *of* . *Then for any subset* $I \subset {}_s$ *there exists a point* t *of* *above* s *such that*

$$ {}_t = \big({}_s\big)_I := \{\, g \in {}_s \;\mid\; \forall x \in I\colon\ gx = x \,\}. $$

Conversely, every $_t$ *is such a centralizer.*

Proof. Consider the morphism

$$\times \quad \longrightarrow \quad , \ (g,x) \mapsto (gx,x).$$

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. **q.e.d.**

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 $\mathrm{pr}_i : T \to \mathbb{G}_m$ 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 using Frobenius twist. 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 non-trivial 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 $\quad \to \quad$, and abbreviate a typical geometric fiber by $G := {}_s$. 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 $\quad \to \quad$. 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 $\to$ be a constructible family of algebraic groups, and $\to$ a constructible family of fiberwise nowhere dense algebraic subgroups. A finite subgroup Γ of a geometric fiber $_s$ is called* -general *if and only if for every point t of above s we have $\Gamma \not\subset {}_t$.*

Of course, this notion depends on the family , whose description may be complicated and awkward to carry along. Therefore we will mostly use the following abbreviation.

Metadefinition 2.2 *Let $\to$ be a constructible family of algebraic groups, and consider a statement $\mathbf{A}(\Gamma)$ about finite subgroups Γ of a geometric fiber $_s$. 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 $\to$ of $\to$ such that for any geometric point s of and any -general *finite subgroup $\Gamma \subset {}_s$ 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, but in the limit for $n \to \infty$ the parameter space could perhaps be viewed as infinite dimensional. For any fixed , the set of -general finite subgroups should then be an open dense subvariety, and as 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 $\to$, and a constructible family of fiberwise nowhere dense algebraic subgroups $\to$.*

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

(b) If a faithful representation of on a vector bundle → is given, then in (a) one can take ′ := and

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

for suitable integers r, m_i, and n_i.

Proof. First we prove (b), using noetherian induction on . Consider a generic point θ of , and let η be the corresponding point of . Then $_\theta$ is a closed algebraic subgroup of $_\eta$. It can therefore ([1] Chap. II Thm. 5.1) be described as the stabilizer of a subspace $'_\theta$ of some tensor space

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

This subspace extends to a vector subbundle ′ over a neighborhood of θ in . Since coincides with the stabilizer of ′ at the generic point θ, by [8] Cor. 9.5.2 it does so over a whole neighborhood. Let us shrink accordingly. Then for any geometric point t of with image s in , and any subgroup $\Gamma \subset {}_s$, we have $\Gamma \subset {}_t$ if and only if ${}_t{}'$ is Γ-invariant. Now recall that, by assumption, ${}_t$ is a proper subgroup of ${}_s$. Therefore ${}_t{}'$ is not ${}_s$-invariant. Thus if every Γ-invariant subspace of ${}_s$ is ${}_s$-invariant, then Γ is (×)-general.

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

To prove (a) we will construct a faithful representation of on a vector bundle over ′. For this note first that by assumption any generic fiber of → possesses a faithful linear representation ${}_\eta \hookrightarrow \mathrm{GL}_n$. This homomorphism extends automatically to an open neighborhood ⊂ . Its kernel is a Zariski-closed subgroup scheme of × whose generic fiber coincides with the identity section. By [8] Cor. 9.5.2 the same is true over a whole neighborhood, so after shrinking this representation is faithful. Applying noetherian induction to the complement ∖ we find a faithful representation of on a vector bundle → ′, where ′ → is a stratification map. Now (b) implies (a). **q.e.d.**

Proof of Theorem 0.6: We deduce this from Theorem 0.5. Let → be the constructible family of fiberwise nowhere dense algebraic subgroups implicit in 0.5. Take the representation furnished by Proposition 2.3 (b), starting with the adjoint representation of . Then the assumptions in 0.6 imply that Γ is -general, so the desired assertion follows from Theorem 0.5. **q.e.d.**

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

Proof. We first look at the problem for a single fiber. Suppose that $\Gamma \subset X := {}_t$, and put $Y := \bigcap_{\gamma\in\Gamma} \gamma X$. By construction this is a nowhere dense closed subvariety of $_s$ 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 $_s$ 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. **q.e.d.**

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

Proposition 2.5 *Let* $\to$ *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* $_s$ *has order* $> n$.

Proof. The individual points on are indexed by the tautological family id: $\to$, so the non-empty finite subsets of $_s$ of cardinality $\leq n$ can be indexed by ^{n}. 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 $\to$. Now Metadefinition 2.2 applies. **q.e.d.**

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 [20] §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., [22] Lemma 3.5). Every connected algebraic group is an extension of an abelian variety by a connected linear algebraic group ([18]). 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. **q.e.d.**

Proposition 3.2 *Let* $\to$ *be a constructible family of algebraic groups, and* $\to$ *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* $s \in (\mathbb{F}_q)$ *the subgroup* $_s(\mathbb{F}_q)$ *is -general.*

Proof. For any geometric point t of above s we must show $_s(\mathbb{F}_q) \not\subset _t$. We cannot apply the estimate 3.1 directly to $_t$, because this subgroup is not necessarily defined over $\mathbb{F}_q$. Let K be the intersection of all translates of $_t$ under powers of Frobenius Frob_q. This subgroup is defined over $\mathbb{F}_q$ and satisfies $_s(\mathbb{F}_q) \cap _t = K(\mathbb{F}_q)$. Every Frobenius translate of $_t$ is a (possibly different) geometric fiber of the same constructible family $\to$. 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 $G := _s$, Proposition 3.1 implies

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

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

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 $|X(\mathbb{F}_q)| \leq c \cdot q^{\dim X}$. 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 $\to$ is a constructible family of algebraic groups and $\to$ a constructible family of subvarieties. Abbreviate $G := _s$ and $X := _t$. 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 $\to$ and $\to$. 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. Recall that a non-trivial connected algebraic group is called *simple* (resp. *almost simple*) if and only if it possesses no non-trivial (resp. no non-trivial connected) proper normal algebraic subgroup. To any simple root system Φ one can associate a natural constructible family of split connected simple linear algebraic groups $\to \mathbf{Spec}\,\mathbb{Z}$ with root system Φ (see [5] Exp.XXV). It is necessarily adjoint; in fact, we will stick to adjoint groups as much as possible. Consider a geometric fiber $G = _s$ 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 $F : G \to G$ 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 called a *finite group of Lie type*.

Simple groups of Lie type: Keeping the above notations, let m denote the index of the root lattice in the weight lattice of Φ, and let $\tilde{G} \to G$ be the universal covering. 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 non-abelian 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 its p-Sylow subgroup.

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

Proposition 3.5 *Let $\to \mathbf{Spec}\,\mathbb{Z}$ 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 $\to$. Then there exists a constant q_0 such that for any Frobenius map F on a geometric fiber $G := {}_s$ with $q_F \geq q_0$ the finite subgroup $(G^F)^{\mathrm{der}}$ is -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 non-standard isogeny with a standard Frobenius map. As in the proof of 3.2 we deduce that the intersection K of all F-power translates of ${}_t$ belongs to a constructible family of algebraic subgroups. This is an F-invariant proper algebraic subgroup, and it remains to show that the ratio $|(G^F)^{\mathrm{der}}| \,/\, |(K^\circ)^F|$ 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. **q.e.d.**

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 non-standard 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 -general, if 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 $\to \mathbf{Spec}\,\mathbb{Z}$ be as above and $\to$ 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 -general, hence so is Γ, and the desired assertion follows from Theorem 0.5. **q.e.d.**

4 Basic Nonconcentration Estimate

Consider an arbitrary constructible family of algebraic groups $\to$, and a geometric fiber $G := {}_s$. 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 $N \lhd G$. 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 ${}_s$. Clearly we have $q_\Gamma = |\Gamma|^{1/\dim {}_s}$ whenever ${}_s$ 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 $\to$, and a constructible family of subvarieties $\to$. Then there exists a constant*

c such that for any sufficiently general finite subgroup Γ *of a geometric fiber* $_s$ *and any point* t *of above* s *we have*

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

There is also a variant for cartesian products:

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

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

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

$$X^r \longrightarrow G,\ (x_1, \ldots, x_r) \mapsto x_1\gamma_1 x_2\gamma_2 \cdots \gamma_{r-1}x_r$$

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 the fibers can be bounded. The bound we need here is of the same kind as the original statement. We therefore proceed by induction. We shall do induction on $\dim X$ and another quantity, to be explained below. 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 $_t$ 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 $\to$ by that given by Proposition 1.9, that is, assume that $_t$ is irreducible. Next, if $\Gamma \cap {}_t$ is non-empty, e.g. contains some element γ, then its cardinality is equal to that of $\Gamma \cap \gamma^{-1}{}_t$. Here the translate $\gamma^{-1}{}_t$ is again a fiber of a constructible family, namely that of all

translates $g^{-1}\ _t$ where $g \in\ _s$ lies in the same irreducible component as $\ _t$. Thus we are reduced to the case that all $\ _t$ are contained in the identity component of $\ _s$. After this reduction, we may also replace by its identity component, i.e., assume that $\ _s$ is connected.

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

It will be convenient to modify the desired estimate by a certain defect δ and then perform descending induction on δ. That is, we will prove the following statement for every integer $\delta \geq 0$, while the family $\ \to\ $ is fixed:

Lemma 4.4 *For any constructible family of subvarieties $\ \to\ $ of $\ \to\ $ there exists a constant c such that for any sufficiently general finite subgroup Γ of a geometric fiber $\ _s$ and every point t of above s we have*

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

For $\delta \geq \dim\ _s$ 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.

A sequence of subvarieties: Abbreviate $G :=\ _s$ and $X :=\ _t$. We will inductively construct a certain sequence of irreducible subvarieties $X = X_1, X_2, \ldots$ of G. At each step we will show that the next subvariety is a fiber of some constructible family of subvarieties of $\ \to\ $, but to ease notation we avoid giving that family a name. Furthermore, the sequence of $d_i := \dim X_i$ will be strictly increasing, so the number of steps is at most $\dim G$. Thus at each of this bounded number of steps we are permitted to impose a new condition on Γ as in Metadefinition 2.2.

Suppose that X_i has already been constructed. If $X_i = G$, we can stop and jump to the end of the proof. Otherwise we consider the subvariety $\overline{XgX_i} \subset G$ for arbitrary $g \in G$.

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

(a) The set of $g \in G$ with $\dim \overline{XgX_i} = \dim X_i$ is a fiber of some constructible family of subvarieties of $\ \to\ $.

(b) If $\dim \overline{XgX_i} = \dim X_i$ for every $g \in G$, then X is contained in a translate of a connected normal subgroup $N \triangleleft G$ of smaller dimension, which is a fiber of some constructible family of algebraic subgroups of $\ \to\ $.

Proof. Choose any $x \in X$. Since X and X_i are irreducible, the relation $\dim \overline{XgX_i} = \dim X_i$ is equivalent to $XgX_i = xgX_i$. This in turn means that $g^{-1}x^{-1}Xg$ is contained in the normalizer N_i of X_i for the left translation action of G on itself. By Proposition 1.13 N_i belongs to a constructible family of proper algebraic subgroups, and the assumption $X_i \neq G$ implies $N_i \neq G$. The inclusion $g^{-1}x^{-1}Xg \subset N_i$ is tantamount to the condition that g lies in the transporter

from $x^{-1}X$ to N_i 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 N_i' denote the intersection of the conjugates gN_ig^{-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 $N := N_i'^{\circ}$, for instance by Proposition 1.9. Finally, the assumptions in (b) imply $x^{-1}X \subset N$, which proves the claim. **q.e.d.**

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 non-empty. Then we have $|\Gamma \cap X| = |\Gamma \cap x^{-1}X|$, and $x^{-1}X \subset N$ belongs to a constructible family. By induction on $\dim G$ 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 $X_i \neq G$, we may suppose that there exists $g \in G$ with $\dim \overline{XgX_i} > \dim X_i$. 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 $\dim \overline{X\gamma X_i} > \dim X_i$. Set $Y := \overline{X\gamma X_i}$ and consider the dominant morphism

$$\varphi\colon\ X \times X_i \longrightarrow Y,\ (x, x_i) \mapsto x\gamma x_i.$$

We will decide about the continuation of the sequence $X_1, \ldots, X_i$ according to the following considerations.

Fiber dimensions: The fiber above any point $y \in Y$ is

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

Thus its dimension is always $\leq d = \dim X$, and in the case of equality the fiber is isomorphic to X. Set $e := \dim Y$ and recall that this is $> d_i = \dim X_i$. The generic fiber dimension of φ is then $d + d_i - e < d$, and by semicontinuity ([8] Th. 13.1.3) all non-empty fibers have dimension $d + d_i - e \leq f \leq d$. For any such f put

$$Y_f := \left\{\ y \in Y\ \middle|\ \dim \varphi^{-1}(y) = f\ \right\}.$$

This is a locally closed subset of Y which is open if and only if $f = d+d_i-e$. For all other values the subset $\varphi^{-1}(Y_f)$ is nowhere dense in $X \times X_i$, so its closure has dimension $< d + d_i$. Thus for these f we deduce

$$\dim \overline{Y_f} < d + d_i - f. \tag{4.6}$$

For the following arguments note also that by Proposition 1.7 the subvariety Y is a fiber of some constructible family, and by Proposition 1.8 the same is true for the closure $\overline{Y_f}$.

Counting arguments: Now we count the points in $(\Gamma \cap X) \times (\Gamma \cap X_i)$ by fibers of φ. We have the following bounds:

Lemma 4.7 *(a) For all $f < d$ and $y \in \Gamma \cap Y_f$ we have*

$$\left|\Gamma^2 \cap \varphi^{-1}(y)\right| \leq c \cdot q_\Gamma^f.$$

(b) For all $y \in \Gamma \cap Y_d$ we have

$$\big|\Gamma^2 \cap \varphi^{-1}(y)\big| = \big|\Gamma \cap X\big|.$$

(c) For all $f > d + d_i - e$ we have

$$\big|\Gamma \cap \overline{Y_f}\big| \le c \cdot q_\Gamma^{d+d_i-f+\delta}.$$

Here the constant c depends only on the constructible families that are tacitly carried along, and so ultimately only on the original families $\to$ and $\to$.

Proof. The assertion (a) follows from the assumption that Theorem 4.2 holds in fiber dimensions smaller than d. Assertion (b) follows directly from the isomorphism $\varphi^{-1}(y) \cong X$. Assertion (c) follows from the inequality 4.6 and the assumption that Lemma 4.4 holds with $\delta + 1$ in place of δ:

$$\big|\Gamma \cap \overline{Y_f}\big| \le c \cdot q_\Gamma^{\dim \overline{Y_f}+\delta+1} \le c \cdot q_\Gamma^{d+d_i-f+\delta}.$$

q.e.d.

Now choose $f \le d$ such that $|\Gamma^2 \cap \varphi^{-1}(Y_f)|$ is maximal. Then we have

$$\begin{aligned} \big|\Gamma \cap X\big| \cdot \big|\Gamma \cap X_i\big| &\le (d+1) \cdot \big|\Gamma^2 \cap \varphi^{-1}(Y_f)\big| \\ &= (d+1) \cdot \sum_{y \in \Gamma \cap Y_f} \big|\Gamma^2 \cap \varphi^{-1}(y)\big|. \end{aligned}$$

In the case $f = d$ this is

$$\le (d+1) \cdot \big|\Gamma \cap X\big| \cdot c \cdot q_\Gamma^{d_i+\delta}$$

by Lemma 4.7 (b) and (c), so that

$$\frac{|\Gamma \cap X_i|}{q_\Gamma^{d_i}} \le (d+1) \cdot c \cdot q_\Gamma^{\delta}. \tag{4.8}$$

In the case $d + d_i - e < f < d$ we similarly find

$$\frac{|\Gamma \cap X|}{q_\Gamma^{d}} \cdot \frac{|\Gamma \cap X_i|}{q_\Gamma^{d_i}} \le (d+1) \cdot c^2 \cdot q_\Gamma^{\delta}, \tag{4.9}$$

using Lemma 4.7 (a) and (c). In both these cases we stop our sequence of subvarieties at X_i. In the remaining case $f = d + d_i - e$ we set $X_{i+1} := Y$, and consequently $d_{i+1} := \dim X_{i+1} = e$. Lemma 4.7 (a) implies

$$\frac{|\Gamma \cap X|}{q_\Gamma^{d}} \cdot \frac{|\Gamma \cap X_i|}{q_\Gamma^{d_i}} \le (d+1) \cdot c \cdot \frac{|\Gamma \cap X_{i+1}|}{q_\Gamma^{d_{i+1}}}. \tag{4.10}$$

End of the proof: Suppose that the sequence of subvarieties stops at X_i. Then in all earlier steps the formula 4.10 applies, so by induction we have

$$\left(\frac{|\Gamma \cap X|}{q_\Gamma^{d}}\right)^i \le \big((d+1) \cdot c\big)^{i-1} \cdot \frac{|\Gamma \cap X_i|}{q_\Gamma^{d_i}}. \tag{4.11}$$

There are three possible ways that the sequence can have stopped at X_i. If it stopped as above with $f = d$, the formulas 4.11 and 4.8 imply

$$\left(\frac{|\Gamma \cap X|}{q_\Gamma^d}\right)^i \leq \big((d+1)\cdot c\big)^i \cdot q_\Gamma^\delta,$$

and therefore

$$\big|\Gamma \cap X\big| \leq (d+1)\cdot c \cdot q_\Gamma^{d+\frac{\delta}{i}} \leq (d+1)\cdot c \cdot q_\Gamma^{d+\delta},$$

as desired. If it stopped with $d + d_i - e < f < d$, we reach the same conclusion using formula 4.9 in place of 4.8. At last, there remains the possibility that the sequence stopped with $X_i = G$. Then we have

$$\frac{|\Gamma \cap X_i|}{q_\Gamma^{d_i}} = \frac{|\Gamma|}{q_\Gamma^{\dim G}} \leq 1,$$

so by formula 4.11 we obtain

$$\big|\Gamma \cap X\big| \leq \big((d+1)c\big)^{\frac{i-1}{i}} \cdot q_\Gamma^d \leq \sup\{(d+1)c, 1\} \cdot q_\Gamma^{d+\delta}.$$

This finishes the proof of Lemma 4.4 in all cases, and thus of Theorem 4.2.

q.e.d.

Proof of Theorem 4.3: As in the preceding proof we abbreviate $G := {}_s$ and $X := {}_t$. Let $\pi : X \longrightarrow G^{n-1}$ 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 $\quad \to \quad$. 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

$$\big|\Gamma^n \cap \pi^{-1}(\underline{\gamma})\big| \leq c_1 \cdot q_\Gamma^{\dim \pi^{-1}(\underline{\gamma})}. \tag{4.12}$$

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

$$Y_f := \big\{\ \underline{g} \in G^{n-1} \ \big|\ \dim \pi^{-1}(\underline{g}) = f\ \big\}.$$

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

$$\big|\Gamma^{n-1} \cap Y_f\big| \leq c_{n-1} \cdot q_\Gamma^{\dim \overline{Y_f}}. \tag{4.13}$$

By construction the constants c_1 and c_{n-1} depend only on the families $\quad \to \quad$ and $\quad \to \quad$. Now observe that

$$f + \dim \overline{Y_f} \ = \ \dim \overline{X \cap \pi^{-1}(Y_f)} \ \leq \ \dim X. \tag{4.14}$$

Thus we can calculate

$$\begin{aligned} \big|\Gamma^n \cap X\big| \quad &= \quad \sum_{f=0}^{d} \sum_{\underline{\gamma} \in \Gamma^{n-1} \cap Y_f} \big|\Gamma^n \cap \pi^{-1}(\underline{\gamma})\big| \\ &\overset{4.12}{\leq} \quad \sum_{f=0}^{d} c_1 \cdot q_\Gamma^f \cdot \big|\Gamma^{n-1} \cap Y_f\big| \end{aligned}$$

$$\begin{aligned} &\overset{4.13}{\leq} \quad \sum_{f=0}^{d} c_1 \cdot c_{n-1} \cdot q_\Gamma^{f+\dim \overline{Y_f}} \\ &\overset{4.14}{\leq} \quad (d+1) \cdot c_1 \cdot c_{n-1} \cdot q_\Gamma^{\dim X}, \end{aligned}$$

which is the desired assertion. **q.e.d.**

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.

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

$$\left|\Gamma \cap {}_t\right| \leq c \cdot q_\Gamma^{\dim {}_t}.$$

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

$$\left|\Gamma \cap {}_v\right| \leq c \cdot (q'_\Gamma)^{\dim {}_v},$$

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

$$\left|\Gamma \cap {}_t\right| = \left|\Gamma \cap {}_v\right| \leq c \cdot (q'_\Gamma)^{\dim {}_v} \leq c \cdot q_\Gamma^{\dim {}_t},$$

as desired. **q.e.d.**

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

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

$$\left|{}_s[n] \cap {}_t\right| \leq c \cdot (nn')^{\dim {}_t},$$

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

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

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

(See [20] §6 for the part prime to the characteristic, §15 for the rest.) Thus formula 4.1 implies $q_\Gamma = nn'$. The result follows from Theorem 5.1. **q.e.d.**

R. Weissauer pointed out to us that this can also be proved using intersection theory, following the lines of the proof of [21] 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 ${}_s$ 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

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

(cf. [28] 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 $\to \mathbf{Spec}\,\mathbb{Z}$ denote the family of split connected adjoint groups associated to Φ (see [5] Exp. XXV). We will consider a geometric fiber $G = {}_s$ 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 is the family of centralizers, but also when it is the family of conjugacy classes in . 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 $\to \mathbf{Spec}\,\mathbb{Z}$, 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

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

This may be viewed as a morphism of families from $\times^n$ to ^{n} which is indexed by the second factor ^{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 $\to \mathbf{Spec}\,\mathbb{Z}$. 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 ^{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. **q.e.d.**

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

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

q.e.d.

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 . Thus for suitable c we have

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

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, , and $\Gamma \subset G = {}_s$ 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, [13] §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 [27] §§14–15 (see also [4] Thms. 3.4.1 and 6.6.1, or [13] 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 [19] 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. 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 Λ 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 [13] §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. **q.e.d.**

Regular semisimple elements: To simplify notation we will abbreviate $\Gamma_\Lambda^\circ := \Gamma \cap G_\Lambda^\circ$.

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

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

Proof. Let $g \mapsto \mathrm{Ad}_g$ denote the adjoint representation of . 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 $\smallsetminus^{\mathrm{rss}}$ is a constructible family of proper closed subvarieties of .

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

$$\big|\Gamma_\Lambda^\circ \smallsetminus G^{\mathrm{rss}}\big| \;\overset{4.2}{\leq}\; c \cdot q_\Gamma^{\dim(G_\Lambda^\circ \smallsetminus G^{\mathrm{rss}})} \;\leq\; \frac{c}{q_\Gamma} \cdot q_\Gamma^{\dim G_\Lambda^\circ} \;\overset{6.2}{\leq}\; \frac{cc_0}{q_\Gamma} \cdot \big|\Gamma_\Lambda^\circ\big|,$$

provided Γ is sufficiently general. This implies

$$\big|(\Gamma_\Lambda^\circ)^{\mathrm{rss}}\big| \geq \Big(1 - \frac{cc_0}{q_\Gamma}\Big) \cdot \big|\Gamma_\Lambda^\circ\big|.$$

Since q_Γ may be assumed arbitrarily large by Proposition 6.1, the desired inequality follows. **q.e.d.**

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 $\Theta \subset \Gamma_\Lambda^\circ$. 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. **q.e.d.**

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 $\mathrm{Tor}^\natural_\Lambda \subset \mathrm{Tor}_\Lambda$ 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*

$$\sum_{\Theta \in \mathrm{Tor}^\natural_\Lambda} \frac{1}{[N_{\Gamma^\circ_\Lambda}(\Theta) : \Theta]} \;=\; 1.$$

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

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

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

$$1 - \varepsilon \;\le\; \sum_{\Theta \in \mathrm{Tor}^\natural_\Lambda} \frac{1}{[N_{\Gamma^\circ_\Lambda}(\Theta) : \Theta]} \;\le\; \frac{1}{1-\varepsilon} \tag{7.5}$$

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

$$\sum_{\Theta \in \mathrm{Tor}^\natural_\Lambda} \frac{1}{[N_{\Gamma^\circ_\Lambda}(\Theta) : \Theta]} \;\in\; \frac{1}{m} \cdot \mathbb{Z}. \tag{7.6}$$

Taking $0 < \varepsilon < \frac{1}{m+1}$, the desired equality follows from 7.5 and 7.6. **q.e.d.**

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

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

Proof.

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

q.e.d.

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 $\dim G_\Lambda^\circ$. 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 $<\dim G_\Lambda^\circ$. Then it holds with $\Lambda\cup\{\gamma\}$ in place of Λ, for any semisimple element $\gamma\in\Gamma_\Lambda^\circ$ outside the center $Z(G_\Lambda^\circ)$.

For any element $g\in G_\Lambda^\circ$ 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 prime-to-p part and 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:

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

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

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

as desired. **q.e.d.**

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 $^{\mathrm{un}}$ is a constructible family of subvarieties of $.$ It is fiberwise irreducible of dimension $\dim G - \operatorname{rank} G$, so Theorem 4.2 implies

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

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.

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

q.e.d.

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

Proof. The set of non-regular unipotent elements is a constructible family of subvarieties of $^{\mathrm{un}}$ of dimension $< \dim G - \operatorname{rank} G$. Thus by Theorem 4.2 and Proposition 7.9 the number of non-regular 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. **q.e.d.**

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 non-zero.*

Proof. The order of any non-trivial unipotent element is a power of the characteristic. **q.e.d.**

Remark: Jordan's theorem: At this point, we have the main ingredients necessary to reprove Jordan's theorem 0.1 in a purely algebraic manner. In fact, Jordan's original proof bears some resemblance to ours. Consider a finite subgroup $\Gamma \subset \mathrm{GL}_n(k)$ of order not divisible by $p = \mathrm{char}(k)$. 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 $\to \mathbf{Spec}\,\mathbb{Z}$, so by Corollary 7.11 the image of Γ in G is contained in an algebraic subgroup $_t \subsetneq G$ which belongs to a constructible family. By Proposition 1.4 the number of connected components of $_t$ is bounded. Thus after replacing H by the identity component of the inverse image of $_t$, 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, , and $\Gamma \subset G = {}_s$ 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 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 ([12] 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, U_γ for any two non-proportional 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 [12] 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 non-simple 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 non-trivial components in U_α. Since everything is normalized by T, it follows that $U_\alpha \subset U^{\mathrm{der}}$, as desired. **q.e.d.**

In the following, we will call (p, Φ) *non-standard* 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 non-standard 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 non-standard 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 [12] 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 non-standard, and that α must be short. If it is not the highest short root in Φ, 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. **q.e.d.**

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

Proposition 8.4 *Consider a non-trivial 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 non-trivial 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 is therefore connected and also centralizes V. 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 [12] 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 non-trivially on the other root group in $Z(U)$. This shows $G_{G_V} \subset V$. 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). **q.e.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 non-empty and a union of cosets of $A_1 \cap \ldots \cap A_r$, its proportion is at least ε_r. **q.e.d.**

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

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

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

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 non-trivial (see [13] Prop. 4.1, Th. 4.6). Let A be the image of $\Gamma \cap U$ in $U/U^{\mathrm{der}} \cong \mathbb{G}_{a,k}^{\mathrm{rank}\, G}$. 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. **q.e.d.**

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

$$\begin{aligned} \big[\Gamma \cap B : \Gamma \cap U\big] &= \frac{[\Gamma \cap B : \Gamma_u] \cdot |\Gamma_u|}{|\Gamma \cap U|} \\ &= \frac{|O_{\Gamma\cap B}(u)|}{|\Gamma \cap U^{\mathrm{run}}|} \cdot \frac{|\Gamma \cap U^{\mathrm{run}}|}{|\Gamma \cap U|} \cdot \big|\Gamma_u\big|, \end{aligned} \tag{8.8}$$

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 [13] Th. 4.6). Thus any two elements of U^{run} are conjugate and, since B is its own normalizer, any element of G which 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 8.8 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. **q.e.d.**

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

Proposition 8.9 *If Γ is sufficiently general, there exists a maximal torus $T \subset B$ with*

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

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. [12] Prop. 19.4). **q.e.d.**

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 8.9.

Lemma 8.10 *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 $g \in G \smallsetminus Z(U)$. 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 $\Gamma \cap T \subset f^{-1}(U_g)$. 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 8.9, we obtain an upper bound for q_Γ. But this contradicts Proposition 6.1, if Γ is sufficiently general. **q.e.d.**

Plugging Lemma 8.10 into Theorem 6.2, we deduce:

Corollary 8.11 *If Γ is sufficiently general, we have*

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

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 8.11 we have already constructed many minimal unipotent elements in Γ. In the non-standard 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 8.10:

Lemma 8.12 *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 $G_V = G_{\Gamma\cap V}$. Thus Theorem 6.2 implies

$$\big|\Gamma \cap G_V\big| \;=\; \big|\Gamma_{\Gamma\cap V}\big| \;\geq\; \frac{1}{c_0} \cdot q_\Gamma^{\dim G_V} \tag{8.13}$$

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

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

where c is fixed and Γ is sufficiently general. Comparing this with the lower bound 8.13, 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. **q.e.d.**

Combining Lemma 8.12 with Theorem 6.2, we deduce:

Corollary 8.14 *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 8.15 *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 8.16 *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 8.15. **q.e.d.**

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

Lemma 8.17 *Consider a finite non-zero 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

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

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

Now we can prove the following crucial result:

Theorem 8.18 *If Γ is sufficiently general, we have $\dim_{\mathbb{F}_V}(\Gamma \cap V) = 1$ and, in particular,*

$$\frac{1}{c_0} \cdot q_\Gamma^d \;\leq\; \left|\mathbb{F}_V\right| \;\leq\; c_0 \cdot q_\Gamma^d.$$

Proof. The second assertion follows from the first together with Corollary 8.14. For the first assertion note that, by the minimality of V, all non-trivial 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 $\leq c_0^2$. On the other hand $\Gamma \cap V$ is arbitrarily large, by Corollary 8.14 and Proposition 6.1. Thus the theorem follows from Lemma 8.17. **q.e.d.**

Notations: We fix some notation to be used in the following sections. The order of $\mathbb{F}_V$ is denoted 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 $0 \leq e < r$.

9 Frobenius Map

We keep the notations of the preceding sections. As a result of Proposition 8.16 we have associated to any sufficiently general Γ a certain finite field $\mathbb{F}_V$ of characteristic $p > 0$, and by Theorem 8.18 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 non-trivial representation σ of G on a k^d-module M of finite type, which belongs to a constructible family of representations of , 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 .

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

$$\begin{array}{ccccc} \mathrm{Aut}_k({}_{t'}) & \cong & \mathrm{Aut}_k({}_t) & = & \mathrm{Aut}_k(M) \\ \cup & & \cup & & \cup \\ {}_{t'} & \xrightarrow{\sim} & F({}_t) & = & F(\sigma(G)) \end{array}$$

Here the isomorphism ${}_{t'} \cong {}_t$ is indexed by a point on the constructible family $\mathrm{Isom}(\mathrm{pr}_1^*\ , \mathrm{pr}_2^*\ )$ over $(t, t') \in \ \times\ $. 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 $\sigma(G) \subset F(\sigma(G))$, and equality follows from the fact that both sides are irreducible of the same dimension. **q.e.d.**

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 non-trivial 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 [24] Thm. 1.7) we deduce that there is an isogeny $F : G \to G$ 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 $\Gamma \subset G^F$. **q.e.d.**

Proof of Theorem 0.5: Combining Lemma 9.4 with Theorem 8.18 we find

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

Thus using Theorem 3.4 (d) we deduce

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

Therefore the index

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

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 $\left(G^F\right)^{\mathrm{der}} \subset \Gamma \subset G^F$, as desired. This finishes the proof of Theorem 0.5 modulo Theorem 9.1. **q.e.d.**

Notations: 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 $\mathrm{Lie} \to \mathbf{Spec}\,\mathbb{Z}$; hence the adjoint representation of G belongs to a constructible family. The G-invariant subspaces are known completely, by [10], [11], or [24] Prop. 1.11.

If $\dim Z(U) = 1$, there is a unique simple subquotient on which G acts non-trivially. The representation on it is denoted ρ. The root system being fixed, $\mathrm{Lie}\,G$ is already irreducible whenever $p \gg 0$. Thus by separating the remaining characteristics we see that ρ belongs to a constructible family of representations.

If $\dim Z(U) = 2$, there are precisely two simple subquotients with non-trivial 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 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 notations 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 [6] §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 characteristic 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 8.9, 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. **q.e.d.**

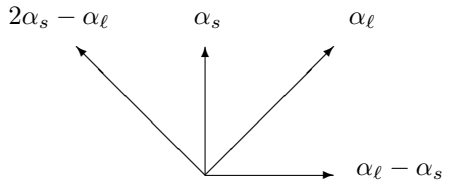
Recall that ρ is a representation over k^d; hence its trace takes values in k^d.

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

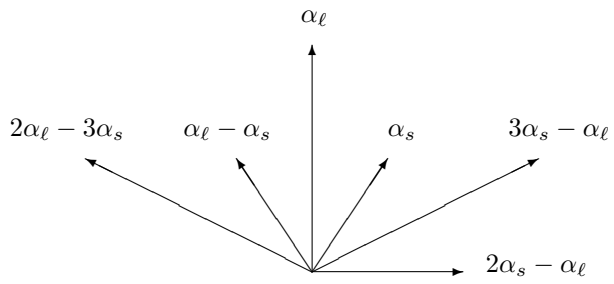
Proof in the rank one case: As the assertion is invariant under conjugation, we may assume $\gamma \in \Gamma \cap T^{\mathrm{rss}}$. Then γ acts on $U \cong \mathbb{G}_a$ and $\operatorname{Lie} U$ through the same scalar x, and the construction of $\mathbb{F}_V$ implies $x \in \mathbb{F}_V$. The total 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. **q.e.d.**

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 $(3, G_2)$:



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 non-trivial.*

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 8.9. The commutator of u and $\gamma u \gamma^{-1}$ then has the desired property, by the formulas in [12] Prop. 33.5 (b). **q.e.d.**

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 8.9 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) $e \not\equiv 0 \mod r$.

(c) The order of x in $k^\times$ is at least $(p^r - 1)/c_0^2$.

Proof. (a) follows from Theorem 8.18 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 8.18 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). **q.e.d.**

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 non-trivial 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 non-trivial 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 8.15. **q.e.d.**

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

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

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

(b) We have

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

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

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 less 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 $\left|p^a + p^{a'} - p^b\right| < 2p^{2r/3}$. 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. **q.e.d.**

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

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

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 $\{e + 1 \mod r,\ e + n \mod r\} = \{0 \mod r,\ n \mod r\}$. 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. **q.e.d.**

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

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

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 [24] Prop. 1.11). Thus we may replace the trace by the sum over all non-zero 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 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. **q.e.d.**

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*

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

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. **q.e.d.**

Let M be the ring of k^d-linear endomorphisms of the representation space of ρ. In the non-standard case we have $\dim \rho_\ell = \dim \rho_s$, since $\operatorname{rank} G = 2$ (cf. [24] 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 $^{n^2}$, 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

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

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

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 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

$$M_0 \;:=\; \big\{\; m \in M \;\big|\; \forall 1 \leq i \leq n^2 \colon\; \operatorname{Tr}(\rho(\gamma_i)m) \in \mathbb{F}_V \;\big\}.$$

By construction this defines a model of M as 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 $\mathrm{id} \in M_0$.

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

Proof. By construction

$$\begin{aligned}
\Gamma \cap \rho^{-1}(M_0) \quad &= \quad \big\{\; \gamma \in \Gamma \;\big|\; \forall 1 \leq i \leq n^2 \colon\; \operatorname{Tr}(\rho(\gamma_i)\rho(\gamma)) \in \mathbb{F}_V \;\big\} \\
&\overset{10.3}{\supset} \quad \big\{\; \gamma \in \Gamma \;\big|\; \forall 1 \leq i \leq n^2 \colon\; \gamma_i\gamma \in \Lambda \;\big\} \\
&= \quad \bigcap_{i=1}^{n^2} \gamma_i^{-1}\Lambda.
\end{aligned}$$

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. **q.e.d.**

Next, we consider the left stabilizer

$$\Delta \;:=\; \{\; \gamma \in \Gamma \;\mid\; \rho(\gamma)M_0 = M_0 \;\}.$$

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

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 $h+1 \leq \ell$. So let us assume $h \geq \ell$. We calculate

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

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

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

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 . 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. **q.e.d.**

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 non-trivial, 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

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

On the one hand Lemma 10.13 implies

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

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

$$\Big|\Gamma \cap \rho^{-1}\big(M_0 \cap \sigma(\gamma)(M_0)\big)\Big| \;\leq\; \frac{c}{q_\Gamma} \cdot \big|\Gamma\big|.$$

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. **q.e.d.**

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 notations 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 .

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 ([12] 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. **q.e.d.**

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

Proposition 11.2 *The subgroup $H_{(g)}$ belongs to a constructible family of algebraic subgroups of , 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 $\to \mathbf{Spec}\,\mathbb{Z}$. 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 $\mathbb{F}_p$. **q.e.d.**

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

Proof. If $\pi(\Gamma \cap H_{(\gamma)}) \subset {}_t$, then both $\Gamma \cap V$ and $\Gamma \cap \gamma V\gamma^{-1}$ are contained in $\pi^{-1}({}_t)$. 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 $\pi^{-1}({}_t)$. Suppose $V \not\subset \pi^{-1}({}_t)$. Then $V \cap \pi^{-1}({}_t)$ is a fiber of a constructible family of proper algebraic subgroups of V. Viewing it as a subgroup of G, Theorem 4.2 implies

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

for some fixed constant c, if Γ is sufficiently general. But this contradicts Corollary 8.14, if q_Γ is large, as guaranteed by Proposition 6.1. The analogous arguments apply when $\gamma V\gamma^{-1} \not\subset \pi^{-1}({}_t)$. **q.e.d.**

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 non-constant.*

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 non-standard case and α is the highest short root. Then by definition we have $\rho = \rho_s$, so only short roots occur as non-zero weights β. The same argument then applies. **q.e.d.**

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 non-constant.*

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 non-constant function of x. As any central function on SL_2 is a polynomial in the trace, we deduce that any non-constant central function on $H_{(\dot{w})}$ remains non-constant 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 $V = U_{\alpha_\ell}$. **q.e.d.**

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 non-standard 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 non-zero 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 $0 \mod 2$, 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 [10] or [24] Prop. 1.11. The proposition follows. **q.e.d.**

Arithmetic properties of traces: Fix a non-trivial 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 non-trivial elements of some root group; hence so does Γ, which contradicts Assumption 8.15. 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 non-standard isogeny whose square is a standard Frobenius map relative to the prime field $\mathbb{F}_p$. By [24] 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 ([24] 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

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

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 $\operatorname{Tr}\sigma(\delta) \in \mathbb{F}_V$, as desired. **q.e.d.**

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 non-degenerate 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 .*

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 , as desired. **q.e.d.**

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

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

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

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

for all $\gamma, \gamma' \in \Gamma$ 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

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

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 $1 \leq i \leq \ell$ 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 $\ell \leq n+1$, where n is the smallest number of generators of M over k^d. As the pairing $\overline{\mathrm{Tr}}$ is non-degenerate, the set of elements

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

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 $\gamma \in \Gamma \cap \bigcap_{i=1}^{\ell} \gamma_i B \dot{w} B$. 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. **q.e.d.**

Proof of Theorem 9.1 in the general case: Proposition 11.5 and Formula 11.9 imply that the representation σ is non-trivial. 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. **q.e.d.**

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 $\to$, any constructible family of fiberwise nowhere dense algebraic subgroups $\to$, and every positive integer N, there exists a constructible family of fiberwise nowhere*

dense algebraic subgroups $_N \to {}_N$ *of* $\to$ *with the following property. For any finite subgroup* Γ *of a geometric fiber* $_s$*, if* $\big[\Gamma : \Gamma \cap {}_t\big] \le N$ *for some point* t *of above* s*, then* $\Gamma \subset {}_{N,t_N}$ *for some point* t_N *of* $_N$ *above* s*.*

Proof. The conjugates of form a constructible family of subgroups, indexed by the total space $\times$. Therefore the intersections of at most N conjugates also form a constructible family, say $' \to '$. By Proposition 1.13, after stratifying $'$ if necessary, there is a constructible family of subgroups $\to '$ of which is fiberwise the normalizer of $'$.

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

$$\big[\Gamma : \Gamma \cap {}_{t'}'\big] \le N!.$$

The $N!$-tuples $(n_1, \dots, n_{N!})$ of sections of are indexed by the $N!$-fold fiber product $^{N!}$. Thus the union of the translates $n_i\ '$ is a constructible family of subvarieties of . 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 $_N \to {}_N$. Clearly it has the desired property vis-à-vis Γ. **q.e.d.**

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 $f_1 \colon G \twoheadrightarrow H_1$ and $f_2 \colon G \twoheadrightarrow H_2$ *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 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* $\to$ *. Let* $\to \mathbf{Spec}\,\mathbb{Z}$ *be the constructible family of connected adjoint groups associated to a simple root system* Φ*. Then there exists a morphism of finite type* $' \to$ *and a homomorphism* $f \colon \times\ ' \to \times\ '$ *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 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 . After shrinking and passing to a finite covering, we may suppose that all simple quotients of $_\eta$ of type Φ can be defined over the residue field $k(\eta)$. Consider one of them, say $f_\eta \colon {}_\eta \to {}_\eta$. As η is a generic point of , this morphism extends to some neighborhood. After shrinking the extension $f : \to$ remains a homomorphism, as well as surjective (compare Proposition 1.7). Since f_η is a simple quotient, the image of its derivative

$$df_\eta \colon \operatorname{Lie}{}_\eta \twoheadrightarrow \operatorname{Lie}\big({}_\eta/\operatorname{Rad}_u {}_\eta\big) \longrightarrow (\operatorname{Lie}\ ) \otimes k(\eta)$$

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

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

$$\big(H_1^F\big)^{\mathrm{der}} \;\subset\; f\big(\Gamma \cap {}_t{}^\circ\big) \;\subset\; H_1^F.$$

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

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

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

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 by $G := {}_t$. By Proposition 1.4 the index $[G : G^\circ]$ is bounded, say $\le N$. 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 $f_i : G^\circ \twoheadrightarrow H_i$ there exists a Frobenius map $F\colon H_i \to H_i$ so that $(H_i^F)^{\mathrm{der}}$ is simple and

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

Define

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

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 m an upper bound for the index of their root lattices in their weight lattices. Using Theorem 3.4 (b) we deduce

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

whence 0.2 (a). The next subfactor Γ_1/Γ_2 is embedded into the product of non-commutative 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 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. **q.e.d.**

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 $\le J'(n)$, by 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 p-Sylow 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 $\le |\Gamma_{(p)}|^2$.*

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 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

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

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. **q.e.d.**

References

[1] Borel, A., *Linear algebraic groups*, GTM 126, New York etc.: Springer (1991).

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

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

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

[5] 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 etc.: Springer (1970).

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

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

[8] Grothendieck, A., *Études locale des schémas et des morphismes de schémas*, Élements de Géométrie Algébrique IV, EGA4, *Publ. Math. IHES* **20** (1964), **24** (1965), **28** (1966), **32** (1967).

[9] Hartshorne, R., *Algebraic Geometry*, GTM **52**, New York etc.: Springer (1977).

[10] Hiss, G., Die adjungierten Darstellungen der Chevalley-Gruppen, *Arch. Math.* **42** (1984), 408–416.

[11] Hogeweij, G. M. D., Almost Classical Lie Algebras I, *Indagationes Math.* **44** (1982), 441 - 460.

[12] Humphreys, J. E., *Linear Algebraic Groups*, GTM **21**, New York etc.: Springer (1975), (1981).

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

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

[15] Jantzen, J. C., *Representations of Algebraic Groups*, Boston etc.: Academic Press (1987).

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

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

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

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

[20] Mumford, D., *Abelian Varieties*, Oxford: Oxford Univ. Press (1974).

[21] Mumford, D., *Geometric Invariant Theory*, Berlin etc.: Springer (1965).

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

[23] Pink, R., The Mumford-Tate conjecture for Drinfeld modules, *Publ. RIMS, Kyoto University* **33** (1997), 393–425.

[24] Pink, R., Compact subgroups of linear algebraic groups, *J. Algebra* **206** (1998), 438–504.

[25] Pink, R., Strong approximation for Zariski dense subgroups over arbitrary global fields, *in preparation.*

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

[27] Steinberg, R., Endomorphisms of linear algebraic groups, *Mem. Amer. Math. Soc.* **80** (1968).

[28] Tate, J., Endomorphisms of Abelian Varieties over Finite Fields, *Inventiones Math.* **2** (1966), 134–144.

[29] Weisfeiler, B., Strong approximation for Zariski dense subgroups of semisimple algebraic groups, *Annals of Mathematics* **120** (1984), 271–315.

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