Acta Math., 193 (2004), 105–139

Subgroup growth of lattices in semisimple Lie groups

by

ALEXANDER LUBOTZKY and NIKOLAY NIKOLOV

Hebrew University
Jerusalem, Israel

New College
Oxford, England, U.K.

1. Introduction

Let H be a simple real Lie group; thus H is the connected part of $G(\mathbf{R})$ for some simple algebraic group G. Let K be a maximal compact subgroup of H, $X=H/K$ be the associated symmetric space, and let Γ be a lattice in H, i.e., a discrete subgroup of finite covolume in H. The lattice Γ is said to be *uniform* if H/Γ is compact, and *non-uniform* otherwise. We denote by $s_n(\Gamma)$ the number of subgroups of Γ of index at most n. The study of $s_n(\Gamma)$ for finitely generated groups Γ has been a focus of a lot of research in the last two decades (see [LuS] and the references therein). Our first result is a precise (and somewhat surprising) estimate of $s_n(\Gamma)$ for higher-rank lattices.

THEOREM 1. *Assume that* $\mathbf{R}$-rank$(H)\geqslant 2$ *and H is not locally isomorphic to* $D_4(\mathbf{C})$. *Then for every non-uniform lattice Γ in H, the limit*

$$\lim_{n\to\infty}\frac{\log s_n(\Gamma)}{(\log n)^2/\log\log n}$$

exists and equals a constant $\gamma(H)$ which depends only on H and not on Γ. The number $\gamma(H)$ is an invariant which is easily computed from the root system of G.

The theorem shows that different lattices in the same Lie group have some hidden algebraic similarity; a phenomenon which also presents itself as a corollary of Margulis super-rigidity, which implies that H can be reconstructed from each Γ.

Every conjugacy class of subgroups of Γ of index n has size at most n (which is negligible compared to $s_n(\Gamma)$) and defines a unique cover of the Riemannian manifold $M=\Gamma\backslash X$. Hence Theorem 1 is equivalent to the following theorem.

THEOREM 1′. *With the same assumptions on H as in Theorem* 1. *Let M be a non-compact manifold of finite volume covered by X, and let $b_n(M)$ be the number of* *covers of M of degree at most n. Then*

$$\lim_{n\to\infty}\frac{\log b_n(M)}{(\log n)^2/\log\log n}$$

exists, equals $\gamma(H)$ and is independent of M.

In spite of the geometric flavor of its statement, the proof of Theorem 1 (and 1′) is based on a lot of number theory. This is due to the fact that a lattice Γ as in Theorem 1 has two properties:

(i) Γ is an arithmetic lattice by Margulis' arithmeticity theorem;

(ii) Γ has the congruence subgroup property.

Now (i) and (ii) imply that counting subgroups of finite index in Γ comes down to counting congruence subgroups in Γ. In fact, the main result of the current paper is the proof of the upper bound of Conjecture 1 below, which was posed in [GLP] (and one extension of the lower bounds proved there). To describe our results we need more terminology.

Let G be a simple, simply-connected, connected algebraic group defined over a number field k, together with a fixed representation $G\hookrightarrow \mathrm{GL}_{n_0}$.

Let $\mathcal{O}$ be the ring of integers of k. Denote by V_f and V_∞ the set of (equivalence classes of) non-archimedean and archimedean valuations of k, respectively, and set $V=V_f\cup V_\infty$. For a valuation $v\in V$, let k_v denote the completion of k with respect to v, and similarly for $v\in V_f$ define $\mathcal{O}_v$ as the completion of $\mathcal{O}$. Let G_v be the group of k_v-points of $G(-)$.

Fix a finite subset S of valuations of k containing V_∞ and consider $\mathcal{O}_S=\{x\in k \mid v(x)\geqslant 0 \text{ for all } v\notin S\}$, the ring of S-integers of k. Define $\Gamma=G(\mathcal{O}_S):=G(k)\cap \mathrm{GL}_{n_0}(\mathcal{O}_S)$. We assume that $G_S:=\prod_{v\in S}G_v$ is non-compact, so that Γ is an infinite group.

For every non-zero ideal I in $\mathcal{O}_S$, let $\Gamma(I)=\ker(G(\mathcal{O}_S)\to G(\mathcal{O}_S/I))$. A subgroup Δ of Γ is called a *congruence subgroup* if Δ contains $\Gamma(I)$ for some ideal I. Let $C_n(\Gamma)$ be the number of congruence subgroups of Γ of index at most n. Let

$$\alpha_+(\Gamma)=\limsup_{n\to\infty}\frac{\log C_n(\Gamma)}{(\log n)^2/\log\log n}$$

and

$$\alpha_-(\Gamma)=\liminf_{n\to\infty}\frac{\log C_n(\Gamma)}{(\log n)^2/\log\log n}.$$

It was shown in [GLP] that for $\Gamma=\mathrm{SL}_2(\mathbf{Z})$, $\alpha_+(\Gamma)=\alpha_-(\Gamma)=\frac{1}{4}(3-2\sqrt{2})$. A general conjecture was formulated there for the case where G splits over k:

Let $R=R(G)=|\Phi_+|/l$, where Φ_+ is the set of positive roots of the root system corresponding to G and $l=\operatorname{rank}(G)$, and let

$$\gamma(G)=\frac{\left(\sqrt{R(R+1)}-R\right)^2}{4R^2}.$$

Then we have:

CONJECTURE 1. $\alpha_+(\Gamma)=\alpha_-(\Gamma)=\gamma(G)$.

It was shown in [GLP] that, assuming the generalized Riemann hypothesis for Artin L-functions (GRH), indeed $\alpha_-(\Gamma)\geqslant\gamma(G)$, and that without assuming the GRH this still holds if $k/\mathbf{Q}$ is an abelian extension of $\mathbf{Q}$.

In this paper we prove the upper bound in full, and extend the lower bound result of [GLP] to the non-split case. In summary:

THEOREM 2. *Let G be an absolutely simple, connected, simply-connected algebraic group over a number field k. Let Φ_+, l, $R(G)$ and $\gamma(G)$ be the numbers defined above for the split form of G. Then*:

(A) $\alpha_+(\Gamma)\leqslant\gamma(G)$.

(B)(1) *Assuming the GRH we have*

$$\alpha_-(\Gamma)\geqslant\gamma(G):=\frac{\left(\sqrt{R(R+1)}-R\right)^2}{4R^2}.$$

Therefore assuming the GRH it follows that $\alpha_+(\Gamma)=\alpha_-(\Gamma)=\gamma(G)$.

(B)(2) *Moreover,* (B)(1) *is unconditional provided there is a Galois field $K/\mathbf{Q}$ such that G is an inner form*(1) *over K, and either* $\operatorname{Gal}(K/\mathbf{Q})$ *has an abelian subgroup of index at most* 4, *or* $\deg[K:\mathbf{Q}]<42$.

COROLLARY 1. *If G is a Chevalley (split) group and $k=\mathbf{Q}$, then* $\alpha_+(\Gamma)=\alpha_-(\Gamma)=\gamma(G)$. *In particular,*

$$\alpha_{\pm 1}(\mathrm{SL}_d(\mathbf{Z}))=\frac{\left(\sqrt{d(d+2)}-d\right)^2}{4d^2}.$$

So Conjecture 1 is now fully proved, modulo the GRH (and it is unconditionally proved for abelian extensions $k/\mathbf{Q}$). The case of $d=3$ of Corollary 1 was also proved independently by Edhan [E]. The main content of this paper is the proof of Theorem 2 (A). Part (B) is just a small improvement over [GLP].

The extension to arbitrary k-simple G is important when one comes to the study of subgroup growth of lattices in a higher-rank simple Lie group H:

---

(1) This term is explained in §4.

As mentioned above, by Margulis' arithmeticity theorem [Ma] every lattice Γ in H is arithmetic. Moreover, a famous conjecture by Serre [S] asserts that such a group Γ has the 'weak' congruence subgroup property (on the finiteness of the congruence kernel, as presented in §7.1 of [LuS] for example). This conjecture is by now proved, unless H is of type A_n and Γ is a cocompact lattice in H. Now, given H we can analyze the possible G, k and S such that $G(\mathcal{O}_S)$ is a lattice in $H=G(\mathbf{R})^0$. The possibilities are given by Galois cohomology and enable us to prove the following result:

THEOREM 3. *Assuming the GRH and Serre's conjecture, then for every non-compact higher-rank simple Lie group $H=G(\mathbf{R})^0$ and every lattice Γ in H, the limit*

$$\lim_{n\to\infty}\frac{\log s_n(\Gamma)}{(\log n)^2/\log\log n}$$

exists and equals $\gamma(G)$. In particular, it depends only on H and not on Γ.

In fact, we shall prove a more general result about subgroup growth of irreducible lattices in all higher-rank *semisimple groups of characteristic* 0: we refer the reader to §5 and Theorem 11 for definitions and the full statement. The proof also shows that for 'most' lattices in simple Lie groups, the conclusion of Theorem 3 holds unconditionally. In particular, this applies to the cases treated in Theorem 1.

Theorem 2 (A) was proved in [GLP] in the special case when $G=\mathrm{SL}_2$. (For general split G, a partial result was also obtained: $\alpha_+(\Gamma)<C\gamma(G)$ for some absolute constant C.) The proof there had two parts:

(a) a reduction to an extremal problem for abelian groups (§5 in [GLP]);

(b) solving this extremal problem (Theorem 5 in [GLP], restated as Theorem 5 below).

Part (a) used the explicit list of the maximal subgroups of $\mathrm{SL}_2(\mathbf{F}_q)$. Such a detailed description becomes too long for general $G(\mathbf{F}_q)$ with the increase of the Lie rank of G and q.

The main new result in this work relates to part (a) and is the following Theorem 4 (deduced in turn from its more refined version, Theorem 7 from §2 below). We need some additional notation:

Let $X(\mathbf{F}_q)$ be a finite quasisimple group of Lie type X over the finite field $\mathbf{F}_q$ of characteristic $p>3$. For a subgroup H of $X(\mathbf{F}_q)$ let

$$h(H)=\frac{\log\,[X(\mathbf{F}_q):H]}{\log|H^\diamond|},$$

where $H^\diamond$ denotes the maximal abelian quotient of H whose order is coprime to p. Set $h(H)=\infty$ if $|H^\diamond|=1$.

Let $\widetilde{X}$ be the untwisted Lie type corresponding to X (so that $X=\widetilde{X}$, $X={}^2\widetilde{X}$ or $X={}^3\widetilde{X}$, the last case occurring only if $\widetilde{X}=D_4$). Then $\widetilde{X}(-)$ is a group scheme of a split, simple, connected algebraic group. Recall that $R(\widetilde{X})$ is the ratio of the number of positive roots of the root system of $\widetilde{X}$ to its Lie rank as defined before Conjecture 1. Extend the definition of R to twisted Lie types by setting $R(X)=R(\widetilde{X})$.

THEOREM 4. *Given the Lie type X (twisted or untwisted). Then*

$$\liminf_{q\to\infty} \min\{h(H) \mid H \leqslant X(\mathbf{F}_q)\} \geqslant R(X).$$

The line of the proof of Theorem 4 is the following: We need to minimize $h(H)$ among all subgroups of $X(\mathbf{F}_q)$. We first show that among the parabolic subgroups the minimum (when $q\to\infty$) is obtained for the Borel subgroup, and there it is equal to $R(X)$ (see Proposition 3 below). We then show that every H can be replaced by a parabolic subgroup $\mathbf{P}$ with $h(\mathbf{P})\leqslant h(H)+o(1)$. The second step itself is divided into two stages: the case when H is not contained in any parabolic subgroup (the atomic case), and then the general case is reduced to this case. We stress that in this process H is replaced by a parabolic subgroup which does not necessarily contain H (though in many cases it is "natural" and possible to choose some $\mathbf{P}$ containing H).

The proof of Theorem 4 does not depend on the classification of the finite simple groups, we use instead the work of Larsen and Pink [LaP] and Liebeck, Saxl and Seitz [LiSS] (the latter for groups of exceptional type).

Once Theorem 4 is proved, one reduces Theorem 2 (A) again to the same extremal problem on abelian groups solved in [GLP]:

THEOREM 5. (Theorem 5 of [GLP]) *Let d and $R\geqslant 1$ be fixed positive numbers. Suppose that $A=C_{x_1}\times C_{x_2}\times \ldots \times C_{x_t}$ is an abelian group such that the orders $x_1, x_2, \ldots, x_t$ of its cyclic factors do not repeat more than d times each. Suppose that $r|A|^R\leqslant n$ for some positive integers r and n. Then as n tends to infinity we have*

$$s_r(A) \leqslant n^{(\gamma+o(1))\log n/\log\log n},$$

where $\gamma=\left(\sqrt{R(R+1)}-R\right)^2/4R^2$.

A few words about the structure of the rest of the paper:

In §2 we show how the upper bound, i.e., Theorem 2 (A), is proved using Theorem 7 below, of which Theorem 4 is an easy corollary. In §3 we prove Theorem 7. In §4 we use all the previous results and Galois cohomology to prove Theorems 1, 2 (B) and 3. We conclude with some remarks in §5 relating to [BGLM], [LiS] and [MP].

The results of this paper are announced in [GLNP].

2. The upper bound: reduction to Theorem 7

Notation. All logarithms in the paper are in base 2 unless stated otherwise. Put

$$l(n)=\frac{\log n}{\log\log n} \quad \text{and} \quad \lambda(n)=\frac{(\log n)^2}{\log\log n}.$$

For functions f and g of integral argument n, we write $f\sim g$ when $f(n)/g(n)\to 1$ as $n\to\infty$, and write $f\asymp g$ if $\log f\sim\log g$.

For a finite group G we denote by $O_p(G)$ the largest normal p-subgroup of G, and $d(G)$ is the minimal size of a generating set for G.

The (Prüfer) rank of G is defined to be the maximal of the numbers $d(H)$ as H ranges over all the subgroups of G. Note that this use of 'rank' is different from the k-rank of an algebraic group H, which is denoted by $\mathrm{rk}_k(H)$.

A group G is said to be a central product of its subgroups $A, B\leqslant G$, denoted as $G=A\circ B$, if $G=AB$ and $[A,B]=1$.

Put $\delta:=[k:\mathbf{Q}]$.

The reductions. By our assumptions, G is a connected, simply-connected simple algebraic group defined over k. Therefore there exist a finite extension K of k and an absolutely simple group $\overline{G}$ such that $G=\mathbf{R}_{K/k}(\overline{G})$, $G(k)=\overline{G}(K)$ and $G(\mathcal{O}_S)$ is commensurable with $\overline{G}(\overline{\mathcal{O}}_{\overline{S}})$, where $\overline{\mathcal{O}}$ is the ring of integers of K and $\overline{S}$ is the set of valuations of K lying above S. Moreover, the congruence topologies of $G(\mathcal{O}_S)$ and of $\overline{G}(\overline{\mathcal{O}}_{\overline{S}})$ are compatible. So for the purpose of counting congruence subgroups we may replace G by $\overline{G}$, K by k, and thus assume that G is absolutely simple to start with.

Recall that G is simply-connected and G_S is non-compact. Therefore by the strong approximation theorem (Theorem 7.12 of [PR]) the congruence subgroups of Γ correspond to open subgroups of the Cartesian product

$$\prod_{v\in V_f\setminus S} G(\mathcal{O}_v),$$

so we count subgroups of $G(\mathcal{O}_S/I)$ for various ideals $I\lhd\mathcal{O}_S$.

The following result is the generalization of the 'level vs. index' lemma to rings of algebraic integers:

LEMMA 1. ([LuS, Proposition 6.1.1]) *Let H be a subgroup of index n in $\Gamma=G(\mathcal{O}_S)$. Then H contains $\Gamma(m\mathcal{O}_S)$ for some positive integer $m\leqslant c_0 n$, where the constant c_0 depends on G only.*

We shall repeatedly quote results from the paper [GLP]. In particular, Corollary 1.2 together with the argument in §1 there imply that for the upper bound it is enough to prove

$$\limsup_{n\to\infty} \frac{\log s_n(G(\mathcal{O}_S/I_0))}{\lambda(n)} \leqslant \gamma(G),$$

where the ideal $I_0 = m\mathcal{O}_S$ with $m \in \mathbf{N}$ satisfies $m \leqslant c_0 n$.

By Corollary 6.2 of [GLP] we can replace $I_0 = (m)$ above with its divisor $I = \pi_1 \dots \pi_t$, defined to be the product of all the different prime ideal divisors π_i of I_0. Note that the norm of I is at most $c'n^\delta$, where the constant c' depends only on the field k and the algebraic group G. Also $t \leqslant (\delta + o(1))\, l(n)$.

Put

$$G_I := \prod_{i=1}^{t} G(\mathcal{O}_S/\pi_i) \simeq G(\mathcal{O}_S/I).$$

Remark. For a prime ideal π of $\mathcal{O}_S$ belonging to a rational prime p we have that $\mathcal{O}_S/\pi$ is a finite field of bounded degree: at most $\delta = [k:\mathbf{Q}]$ over $\mathbf{F}_p$. Therefore the rank of the group $G(\mathcal{O}_S/\pi)$ is bounded by a function $r = r(\dim G, k)$ of $\dim G$ and δ alone, and independent of π, see Proposition 7 of Window 2 from [LuS].

Now, for a rational prime p which is not coprime to I (i.e., $p \mid m$), let $M(p)$ denote the set of those ideals from $\{\pi_1, \dots, \pi_t\}$ which divide (p). Define

$$G_p := \prod_{\pi \in M(p)} G(\mathcal{O}_S/\pi); \quad \text{and thus } G_I = \prod_{p \mid m} G_p.$$

The strategy of the proof follows several steps, in which we gradually reduce the possibilities for the subgroup H of G_I (each time discounting any contributions less than $n^{o(l(n))}$):

In the first step we fix the projections R_p of H on each G_p. Then we apply the Larsen–Pink theorem to each R_p, which roughly says that R_p resembles an algebraic subgroup. By successive reductions we deal with its unipotent part and then its semisimple part, leaving only the 'torus' (in our case just an abelian p'-group) as a possibility where H can live. This is the point where we are in a position to apply Theorem 5 and finish the proof.

While doing these reductions we need several auxiliary group-theoretic results, and in addition we have to keep track of various numerical constants (in particular the change of the index of H), resulting in considerable notation overload.

Step 1. Let R_p be the projection of $H \leqslant G_I$ on the direct factor G_p. We are assuming that G is absolutely simple, and therefore for almost all rational primes p the group G_p is a product of $|M(p)| \leqslant \delta$ quasisimple groups $G(\mathcal{O}_S/\pi)$, $\pi \in M(p)$. By the remark above, it follows that the rank of G_p is at most $r' := \delta r$. We deduce that there are at most $|G_p|^{r'}$ possibilities for R_p in G_p.

Since $|G_I|=O(m^{\dim G})=O(n^{\delta \dim G})$, it follows that the number of choices for the projections $\{R_p \mid p|m\}$ is at most

$$\prod_p |G_p|^{r'}=|G_I|^{r'}=O(n^{\delta r' \dim G}),$$

which is polynomially bounded in n.

Thus we can assume from now on that the set of projections $\{R_p \mid p|m\}$ is fixed, and estimate the further possibilities for H.

Step 2. At this stage we use the following modification of a theorem by Larsen and Pink [LaP] in [LiP, Corollary 3.1]:

THEOREM 6. (Larsen and Pink) *Let G be a finite subgroup of* $\mathrm{GL}_n(\mathbf{F})$, *where* $\mathbf{F}$ *is a finite field of characteristic p. Then G has a normal subgroup $N \geqslant O_p(G)$ such that*

(1) $[G:N] \leqslant C(n)$, *where C depends on n alone*;

(2) $N/O_p(G)$ *is a central product of an abelian p'-group A and quasisimple groups in* $\mathrm{Lie}^*(p)$.

Here and below $\mathrm{Lie}^*(p)$ denotes the family of finite quasisimple groups of Lie type in characteristic p.

Apply Theorem 6 to each one of the groups R_p: they are linear of degree at most δn_0, where n_0 is the degree of the linear representation of G. Hence there exist normal subgroups $R_p^0 \geqslant R_p^1$ of R_p such that

(1) $[R_p:R_p^0] \leqslant c$, where $c=c(n_0,\delta)$ depends on n_0 and δ only;

(2) $R_p^1=O_p(R_p)$ and $R_p^0/R_p^1=A_p \circ S_p$, where A_p is an abelian p'-group and S_p is quasi-semisimple of characteristic p.

Define $R:=\prod_p R_p$, $R^i:=\prod_p R_p^i$ for $i=0,1$, $S:=\prod_p S_p$ and $A:=\prod_p A_p$.

Step 3. Consider R_p^1. It is a nilpotent group of nilpotency class at most n_0 (by Sylow's theorem every p-group of $\mathrm{GL}_{n_0}(\mathbf{F}_{p^s})$ is conjugate to a group of upper unitriangular matrices), and has rank at most $\max_p \operatorname{rank}(G_p) \leqslant r'$. Lemma 6.1 from [GLP] (also Proposition 1.3.3 in [LuS]) says that given $HR^1/R^1 \leqslant R/R^1$, the number of choices for H is at most

$$|R|^{3(r')^2+n_0 r'} \leqslant (c'n^{\delta})^{r'(3r'+n_0)\dim G}.$$

We are ignoring polynomial contributions to $s_n(G(\mathcal{O}_S/I))$. Therefore from now on we can assume that H contains R^1 and count the possibilities for $\bar{H}=H/R^1$ in $\bar{R}=R/R^1$.

Step 4. The group $\bar{H}$ projects onto each factor $\bar{R}_p:=R_p/R_p^1 \geqslant A_p \circ S_p$ of $\bar{R}$. It follows that the non-abelian composition factors of S_p counted together with their multiplicities all occur among the composition factors of $\bar{H}$. Now $\bar{R}_p/S_p$ is an abelian p'-group extended by a group of order at most c.

CLAIM. *Provided that all primes p are bigger than c, then $\overline{H}$ contains each S_p.*

Proof. It follows the proof of Lemma 4.3 in [LiP]:

Let Z be the center of S. It is enough to show that $\overline{H}Z$ contains S: if so, then $S=S\cap(\overline{H}Z)=(S\cap\overline{H})Z$, and therefore $S=S\leqslant\overline{H}$ because S is perfect. Hence we can assume that $\overline{H}$ contains Z and work modulo Z from now on. Note that S/Z is a direct product of its factors $(ZS_p)/Z$ and that they are semisimple groups over distinct fields.

Consider $\overline{H}^0:=\overline{H}\cap\overline{R}^0\leqslant S\circ A$. Then $\overline{H}/\overline{H}^0\simeq HR^0/R^0$ only has composition factors of order at most c. Therefore each simple factor of S/Z (counted with its multiplicity) occurs among the composition factors of $\overline{H}^0/Z$, and hence among its derived subgroup $((\overline{H}^0)'Z)/Z\leqslant S/Z$.

The order of a group is the product of the orders of its composition factors. It follows that $|((\overline{H}^0)'Z)/Z|\geqslant|S/Z|$, and thus $S=(\overline{H}^0)'Z$, proving the claim. □

So $\overline{H}$ contains S and is thus determined by its image $\widetilde{H}=\overline{H}/S$ in $\widetilde{R}=\overline{R}/S$. Define $\widetilde{R}^0=R^0/S$: a quotient of $A=\prod_p A_p$.

Step 5. In the remaining steps we shall reduce the problem of counting the possibilities for $\widetilde{H}$ in $\widetilde{R}$ to counting subgroups in certain abelian groups E and T (to be defined below).

The key to this reduction is the following generalization of Theorem 4. Recall the number $R(X)$ defined in the introduction for each Lie type $X(-)$ of simple simply-connected algebraic groups over finite fields: $R(X)$ is the number of positive roots of the split form $\widetilde{X}$ of X divided by its rank.

THEOREM 7. *Let $G=X(\mathbf{F}_q)$ be a finite quasisimple group of fixed Lie type X over a finite field $\mathbf{F}_q$ of characteristic $p>3$. There exist a finite set $\mathcal{S}\subseteq\mathbf{Q}[x]$ of non-constant polynomials and constants c_1,c_2,m depending only on X with the following property:*

Suppose that $H\leqslant G$ and that A is an abelian p'-group contained in the centre of $\overline{H}=H/O_p(H)$. Then there exist an abelian p'-group T and a subgroup A_0 of A such that

(1) *A_0 is a homomorphic image of T and $[A:A_0]\leqslant c_1$;*

(2) $\displaystyle\liminf_{\substack{q\to\infty\\ H\leqslant G}}\frac{c_2+\log[G:H]}{\log|T|}\geqslant R(X);$

(3) *the group T is a direct product of at most $m=m(X)$ cyclic groups, each having order $f(q)$ for some $f\in\mathcal{S}$.*

For each prime ideal $\pi\in M(p)$ let R_π^0 be the projection of R_p^0 into the direct factor $G_\pi:=G(\mathcal{O}_S/\pi)$. Then

$$[G_p:R_p^0]\geqslant\prod_{\pi\in M(p)}[G_\pi:R_\pi^0],$$

and A_p is a subdirect product of its projections A_π into the various G_π's.

By our assumptions, G is absolutely simple. Hence for all but finitely many primes π (which we can ignore), G_π is a finite quasisimple group which is a form of the (split) Lie type $\widetilde{X}$ of G. Over a finite field all the forms of $\widetilde{X}$ are quasisplit, and it follows that G_π is $X(\mathcal{O}_S/\pi)$, where X is a (possibly twisted) Lie type corresponding to $\widetilde{X}$. For example, when G has type A_n then G_π is either SL_{n+1} or SU_{n+1} over finite fields. It is important to note that Theorem 7 gives the same constant $R(X)=R(\widetilde{X})$ for all the forms of G. (In the example with A_n above we have $R=\frac{1}{2}(n+1)$.)

Now Theorem 7 applied to $R_\pi^0 \leqslant G_\pi$ for each $\pi \in M(p)$ gives that there is an abelian group T_π and a subgroup $A_{\pi,0}$ of A_π with the stated properties. In particular, T_π maps onto $A_{\pi,0}$, and moreover,

$$[G_\pi : R_\pi^0] \geqslant |T_\pi|^{R(G)-o(1)}.$$

Put $A_{p,0}=\prod_{\pi\in M(p)} A_{\pi,0}$. It follows that $A_{p,0}$ is a homomorphic image of the direct product $T_p:=\prod_{\pi\in M(p)} T_\pi$, and moreover,

$$[G_p : R_p^0] \geqslant |T_p|^{R(G)-o(1)}.$$

Define $\widetilde{R}_p^0=R_p^0/S_p$. Let $E_p \leqslant \widetilde{R}_p^0$ be the image of $A_{p,0}$ under the homomorphism $A_p \twoheadrightarrow R_p^0/S_p=\widetilde{R}_p^0$. We have that $[\widetilde{R}_p^0 : E_p] \leqslant [A_p : A_{p,0}] \leqslant c_1^\delta$ (since $|M(p)| \leqslant \delta$). Also, E_p is a homomorphic image of T_p. Let

$$T=\prod_p T_p \quad \text{and} \quad E=\prod_p E_p.$$

Since $[G_I : R^0]=\prod_p [G_p : R_p^0]$ it now follows that

$$[G_I : R^0] \geqslant |T|^{R(G)-o(1)}.$$

Moreover, for any given rational prime p and prime ideal π of $\mathcal{O}_S$ dividing p, there are at most δ possibilities for the size of the residue field $\mathcal{O}_S/\pi$. Also, there are at most δ prime ideals π dividing p. We conclude that T_p is a product of boundedly (by X and δ) many cyclic groups each having order given by a finite set of polynomials in p. A polynomial of degree $b>0$ cannot take the same value at more than b values of its argument. Therefore there exists a number $d=d(X,\delta)$ such that the abelian group T is a product of cyclic groups C_{x_i} and each integer appears at most d times in the sequence $\{x_i\}$.

Step 6. We need a result which is a slight generalization of Proposition 5.6 from [GLP]. It allows us to pass from $\widetilde{R}$ down to the abelian group E. We postpone its proof to §2.1.

PROPOSITION 1. *Let $D=D_1\times\ldots\times D_s$ be a direct product of finite groups, where each D_i has a normal subgroup E_i of index at most C, and E_i is polycyclic of cyclic length at most r. Assume that the (Prüfer) rank of each D_i is at most r. The number of subgroups $H\leqslant D$ whose intersection with $E=E_1\times\ldots\times E_s$ is a given subgroup $L\leqslant E$ is at most*

$$|D|^{4r}C^{2rs^2}K^s,$$

where $K=K(C)$ is the number of isomorphism classes of groups of order at most C.

Recall that the rank of each R_p is at most r'. Hence $\widetilde{R}_p^0$ is an abelian group of rank at most r'. We apply Proposition 1 to $\widetilde{R}=\prod_p \widetilde{R}_p$ and $E=\prod_p E_p$:

Each R_p/R_p^0 has size at most c and $[\widetilde{R}_p^0:E_p]\leqslant c_1^\delta$. Therefore

$$[\widetilde{R}_p:E_p]\leqslant[R_p:R_p^0][\widetilde{R}_p^0:E_p]\leqslant cc_1^\delta=c_0,$$

say. Thus, given the group $\widetilde{H}\cap E$ the number of choices for $\widetilde{H}$ in $\widetilde{R}$ is at most

$$|\widetilde{R}|^{4r'}c_0^{2r't^2}K(c_0)^t\leqslant n^{4\delta r'\dim G}c_0^{O(l(n)^2)}K^{O(l(n))}=n^{O(\log n/(\log\log n)^2)}=n^{o(l(n))}.$$

Since $[\widetilde{R}:E]\leqslant c_0^t=n^{o(1)}$ it follows that $[\widetilde{R}:E\cap\widetilde{H}]$ and $[\widetilde{R}:\widetilde{H}]$ differ by at most a factor $n^{o(1)}$. So we can restrict ourselves to counting the possibilities for $\widetilde{H}\cap E$. Thus without loss of generality assume that $\widetilde{H}\leqslant E$.

Step 7. To summarize the various reductions so far: we are now counting the possibilities for $\widetilde{H}\leqslant E$, where E is a homomorphic image of $A_0=\prod_p A_{0,p}$, which is in turn an image of T. In turn, $T=C_{x_1}\times\ldots\times C_{x_s}$, where each integer appears at most $d=d(X,\delta)$ times in the sequence $\{x_i\}$.

Let $u=[E:\widetilde{H}]\leqslant[R^0:H]$. Then

$$n\geqslant[G_I:H]=[G_I:R^0][R^0:H]\geqslant|T|^{R(G)-o(1)}u.$$

Hence the number of choices for $\widetilde{H}$ in E is at most

$$s_u(E)\leqslant s_u(A_0)\leqslant s_u(T).$$

Now we can apply Theorem 5 to the group T, with constant $R=R(X)$ and $d=d(X,\delta)$, giving that $s_u(T)\leqslant n^{(\gamma+o(1))l(n)}$.

This proves Theorem 2 (A) modulo Theorem 5 (proved in [GLP]), Theorem 7 (proved in §3) and Proposition 1.

2.1. Proof of Proposition 1

We need the following lemma:

LEMMA 2. *Let $A \leqslant B$ be groups and let C and k be positive integers. The number of subnormal subgroups H of B which contain A and for which there exists a subnormal series*

$$H = H_0 \lhd H_1 \lhd \ldots \lhd H_k \lhd B, \quad \textit{with } [B:H_k] \leqslant C \textit{ and } H_i/H_{i-1} \textit{ cyclic},$$

is at most $[B:A]^k C^d K$, where $d = d(B)$ and K is the number of isomorphism classes of groups of order at most C.

Proof. There are at most K possibilities for the quotient group $U = B/H_k$, and then at most $|U|^d \leqslant C^d$ for the homomorphism $B \to U$, which determines H_k as the kernel. Given H_k there are at most $[H_k:A]^k$ possibilities for $H = H_0$ by Lemma 5.5 of [GLP]. □

Proof of Proposition 1. We follow the proof of Proposition 5.6 from [GLP]:

Let $F_i = D_i \times D_{i+1} \times \ldots \times D_s$ and $L_i = \mathrm{proj}_{F_i} L$. Put $\tilde{L}_{i+1} = L_i \cap F_{i+1}$, so that $\tilde{L}_i \leqslant L_i \leqslant F_i$. Let $H_i = \mathrm{proj}_{F_i} H$. We shall bound the number of possibilities for the sequence $(H, H_2, \ldots, H_s)$.

The number of choices for $H_s \leqslant D_s$ is at most $|D_s|^r$ (because every subgroup of D_i is generated by at most r elements). Now asssume that H_{i+1} is given and consider the possibilities for H_i. Let $X = H_i \cap F_{i+1}$, $Y = \mathrm{proj}_{D_i}(H_i)$ and $Z = H_i \cap D_i$. Then $H_i/X \times Z$ is a subdirect product of Y/Z and H_{i+1}/X, and H_i is thus determined by H_{i+1}, X, Y, Z and an isomorphism $\phi: Y/Z \to H_{i+1}/X$.

Since $\mathrm{rank}(D_i) \leqslant r$ the number of choices for Y, Z and ϕ is at most $|D_i|^r$ each. Notice that the pair of groups $H_{i+1} \geqslant \tilde{L}_{i+1}$ together with the group $X \geqslant \tilde{L}_{i+1}$ satisfies the conditions of Lemma 2: $H_{i+1}/X \simeq Y/Z$, and Y/Z is a section of D_i (so $|Y/(E_i Z \cap Y)| \leqslant C$ and $(E_i Z \cap Y)/Z$ is polycyclic of length $\leqslant r$).

Therefore the number of choices for X is at most

$$[H_{i+1}:\tilde{L}_{i+1}]^r C^{d(H_{i+1})} K.$$

Now $d(H_{i+1}) \leqslant \mathrm{rank}(F_{i+1}) \leqslant sr$ and

$$[H_{i+1}:\tilde{L}_{i+1}] \leqslant [H_{i+1}:L_{i+1}][L_{i+1}:\tilde{L}_{i+1}] \leqslant C^s |D_i|,$$

because $[H_{i+1}:L_{i+1}] \leqslant [H:L] \leqslant C^s$ and $[L_{i+1}:\tilde{L}_{i+1}] = [\mathrm{proj}_{F_{i+1}}(L_i): L_i \cap F_{i+1}] \leqslant |D_i|$.

Thus, given H_{i+1} the number of choices for H_i is at most $|D_i|^{4r} C^{2sr} K$. Multiplying from $i = s$ to $i = 1$ we obtain

$$|D|^{4r} C^{2rs^2} K^s$$

as required. □

2.2. Proof of Theorem 4

Assuming Theorem 7, then with the help of the Larsen–Pink result, Theorem 4 is an easy corollary:

Suppose that H is a subgroup of $G=X(\mathbf{F}_q)$, and let H^0 and $S, A \leqslant H^0/O_p(H)$ be the subgroups given by Larsen–Pink's Theorem 6 above. Recall that by $H^\diamondsuit$ we denote the largest abelian p'-quotient of H.

Let L be the least normal subgroup of H such that H/L is an abelian p'-group. Then by looking at the composition factors of L we see that $O_p(H) \leqslant L$, and then $L/O_p(H)$ must contain S because the latter is a perfect group. Hence $H^\diamondsuit$ is a quotient of H/S, whence $|H^\diamondsuit| \leqslant |A| C(n)$.

Apply Theorem 7 to the group H^0. It follows that for some constants c_1, c_2 and an abelian group T we have

$$\liminf_{\substack{q\to\infty \\ H\leqslant G}} \frac{c_2 + \log\,[G:H^0]}{\log |T|} \geqslant R(X)$$

and $|A| \leqslant |T| c_1$.

Clearly $[G:H] \geqslant [G:H^0]/C(n)$, and together with $|H^\diamondsuit| \leqslant |T| c_1 C(n)$ this easily implies the conclusion of Theorem 4.

3. Proof of Theorem 7: Generalities

Recall that $G=X(\mathbf{F}_q)$ is a finite quasisimple group of Lie type X over a finite field $\mathbf{F}_q$ of characteristic $p>3$, H is a subgroup of G, and A is an abelian p'-group in the centre of $\overline{H}=H/O_p(H)$.

Theorem 7 will follow from the next two propositions:

PROPOSITION 2. *In the situation of Theorem 7 there exist constants $c_1, c_0 > 0$, a finite set $\mathcal{S} \subseteq \mathbf{Q}[x]$ of polynomials (all depending only on the Lie type X), an abelian p'-group T and a parabolic subgroup P of G such that*

(1) *$T \twoheadrightarrow A_0$ for some subgroup A_0 of A of index at most c_1;*

(2) *$c_0[G:H] \geqslant [G:P]$ and $|T| \leqslant c_0 |P^\diamondsuit|$, where $P^\diamondsuit$ denotes the largest abelian p'-image of P;*

(3) *T is a direct product of at most $m=m(X)$ cyclic groups, each having order $f(q)$ for some $f \in \mathcal{S}$.*

PROPOSITION 3. *Let $G=X(\mathbf{F})$ be a quasisimple group of Lie type X over a finite field $\mathbf{F}$ of characteristic bigger than 3. In other words, $X(-)$ is an absolutely simple, connected algebraic group scheme defined over $\mathbf{F}_p$.*

Let $P=P(-)\leqslant X$ be a parabolic subgroup and recall the definition

$$h(H):=\frac{\log\,[G\!:\!H]}{\log|H^{\diamond}|},\quad \textit{where } H\leqslant G.$$

Then

$$\lim_{|\mathbf{F}|\to\infty} h(P(\mathbf{F}))\geqslant R(X),$$

with equality if and only if P is the Borel subgroup of X.

Remark. Note that given the type $X(-)$ (an absolutely simple, connected quasisplit algebraic group defined over $\mathbf{F}_p$), there are several possibilities for its fundamental group, and these give several possibilities for the finite group $G=X(\mathbf{F})$, all of which are covers of the same finite simple group $G/Z(G)$. However, a simple argument shows that once Propositions 2 and 3 are proved for any fixed isogeny version of $X(-)$, they will follow for all the others. Therefore from now on, with one exception, we shall assume that X is simply-connected, and thus that G is the universal covering group of G/Z. The exception is §3.2.2 and the orthogonal group types ($X=B_n$, $X=D_n$ and $X={}^2D_n$), where X will be assumed to be one of the classical groups $\Omega_{2n}^{\pm}$ or Ω_{2n+1}.

Assuming the above propositions the proof of Theorem 7 is straightforward: Let T and P be the groups provided by Proposition 2. Then

$$\frac{c_2+\log\,[G\!:\!H]}{\log|T|}\geqslant\frac{\log\,[G\!:\!P]}{\log|P^{\diamond}|}=h(P),$$

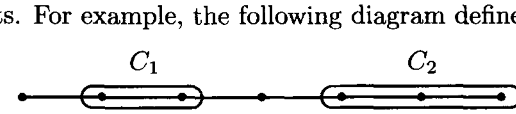
where $c_2=2R\log c_0$. Now Proposition 3 gives that $\liminf_{q\to\infty} h(P)\geqslant R(X)$, and we are done.

3.1. Proof of Proposition 3

Recall that l is the untwisted Lie rank of X and Φ_+ is the set of positive roots. The result is clear if $P=G$.

Case A. Suppose first that X is untwisted Lie type.

$P(-)$ is defined by a subset of the nodes (= the fundamental roots) in the Dynkin diagram of X, which is a disjoint union of maximal connected subsets $C_1, C_2, ..., C_n$, say, of fundamental roots. For example, the following diagram defines a parabolic of $A_7(\mathbf{F})$:



Let $E_i \subseteq \Phi_+$ consist of the positive roots in the span of $r \in C_i$. Then each set $E_i \cup -E_i$ is an irreducible root system with fundamental roots given by C_i and Dynkin diagram which is the connected subgraph defined by C_i.

Put $q = |\mathbf{F}|$. Let L be the Levi factor of P and let M be the greatest normal subgroup of L such that L/M is an abelian p'-group. Hence $P^\diamond \simeq L^\diamond = L/M$. It follows that $P^\diamond \simeq T/T_0$, where T is a maximal split torus contained in L and $T_0 = M \cap T$. Since $X(-)$ is simply-connected, M is a direct product of its simple components, and T_0 is also a torus. The dimension of T_0 is $\sum_{i=1}^n |C_i|$, and therefore

$$[X(\mathbf{F}) : P(\mathbf{F})] \sim q^{|\Phi_+| - \sum_{i=1}^n |E_i|}$$

and

$$\frac{\log |P^\diamond|}{\log q} \sim l - \sum_{i=1}^n |C_i| \quad \text{as } q \to \infty.$$

Notice that since P is proper parabolic, the C_i are proper subsets, and in particular, $l - \sum_{i=1}^n |C_i| > 0$. It follows that

$$\lim_{q\to\infty} h(P) = \frac{|\Phi_+| - \sum_{i=1}^n |E_i|}{l - \sum_{i=1}^n |C_i|}.$$

Let X_i be the split absolutely simple simply-connected group having as Dynkin diagram the connected component C_i. Observe that $|E_i|$ is the number of positive roots of X_i and $|C_i|$ is its rank. It follows that the ratio $R(X_i)$ defined in the introduction is equal to $|E_i|/|C_i|$.

Now it is easy to check that $R(X_i) < R(X) = |\Phi_+|/l$ for every proper non-empty connected subgraph C_i of the Dynkin diagram of X. We now use the following lemma and obvious induction:

LEMMA 3. *Suppose that a, b, c and d are positive real numbers such that $a > b$ and $c > d$. Suppose that $a/c > b/d$. Then $(a-b)/(c-d) > a/c$.*

This shows that

$$\frac{|\Phi_+| - \sum_{i=1}^n |E_i|}{l - \sum_{i=1}^n |C_i|} \geqslant R(X),$$

with equality if and only if $n = 0$, i.e., when P is the Borel subgroup of $X(-)$.

Case B. X is twisted. We assume that the characteristic of $\mathbf{F}$ is bigger than 3, so the corresponding untwisted type $\widetilde{X}$ has Dynkin diagram with single edges, and with the exception of ${}^3D_4$ (which can be treated similarly) $\widetilde{X}$ has a symmetry τ of order 2. Also, $|\mathbf{F}| = q^2$ and $\mathbf{F}$ is a quadratic extension of a field $\mathbf{F}_0$ of order q.

Type of Σ	Root subgroup x_Σ
$A_1=\{w=w^\tau\}$	$\{x_w(t)\mid t\in\mathbf{F}_0\}$
$A_1\times A_1=\{w,u=w^\tau\}$	$\{x_w(t)x_u(t^q)\mid t\in\mathbf{F}\}$
$A_2=\{w,u=w^\tau,u+w\}$	$\{x_w(t)x_u(t^q)x_{u+w}(s)\mid t,s\in\mathbf{F} \text{ and } t+t^q-ss^q=0\}$

Table 1

Then $G=X(\mathbf{F})$ is the group of fixed points in $\widetilde{X}(\mathbf{F})$ under the automorphism $\sigma:=\tau\phi$, where τ is the graph automorphism of $\widetilde{X}(\mathbf{F})$ corresponding to the symmetry τ with the same name, and ϕ is the field automorphism of $\widetilde{X}(\mathbf{F})$ corresponding to the automorphism $x\mapsto x^q$ of $\mathrm{Gal}(\mathbf{F}/\mathbf{F}_0)$.

The type of G is $X={}^2\widetilde{X}\in\{{}^2A_l,{}^2D_l,{}^2E_6\}$. The root subgroups of G correspond to spans Σ of orbits of roots of $\widetilde{X}$ under τ, and are 1-dimensional with the exception of $\Sigma=A_2$, ocurring for ${}^2A_l$ with l even. Table 1 lists the possible root subgroups, and there is a similar parametrization for the diagonal subgroup of G (see [GLS, Tables 2.4 and 2.4.7]).

Observe that (still excluding ${}^3D_4$)

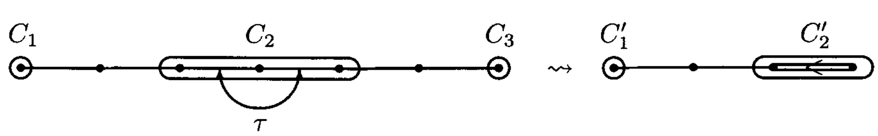
$$|x_\Sigma| \sim \sqrt{\left|\left\{\prod_{r\in\Sigma} x_r(t_r) \,\middle|\, t_r\in\mathbf{F}\right\}\right|},$$

where the right-hand side is computed in $\widetilde{X}(\mathbf{F})$ and the left-hand side x_Σ is a root subgroup of $G=X(\mathbf{F})$. It easily follows that

$$|G|\sim\sqrt{|\widetilde{X}(\mathbf{F})|}.$$

A parabolic P of G is the fixed points $(\widetilde{P})^\sigma$ of a parabolic $\widetilde{P}$ of $\widetilde{X}(\mathbf{F})$ which is defined by a τ-invariant subset of the Dynkin diagram of $\widetilde{X}$.

Here is an example of a parabolic of ${}^2A_7(-)$:



From the above it easily follows that, in the notation of Case A,

$$[G:P]\sim\sqrt{[\widetilde{X}(\mathbf{F}):\widetilde{P}(\mathbf{F})]}\sim q^{|\Phi_+|-\sum_{i=1}^n|E_i|}$$

and

$$|P^\diamond|\sim\sqrt{|\widetilde{P}^\diamond|}\sim q^{l-\sum_{i=1}^n|C_i|}\quad\text{as } q\to\infty.$$

The rest of the proof is the same as in the untwisted case.

Finally, the case $X={}^3D_4$ is similar to the above, with the difference that this time $|\mathbf{F}|=q^3$, $\mathbf{F}_0$ is a subfield of order q and we take cube roots of the corresponding values in the untwisted group $D_4(\mathbf{F})$.

3.2. Proof of Proposition 2

3.2.1. *Reduction to* atomic H. A subgroup H of $G\in\mathrm{Lie}^*(p)$ is called *p-local* if it normalizes a non-trivial p-subgroup of G. We shall use the Borel–Tits theorem, which says that the maximal p-local subgroups of $G\in\mathrm{Lie}(p)$ are parabolic:

THEOREM 8. (Borel–Tits [BT], [GLS, Theorem 3.1.1]) *Let $G\in\mathrm{Lie}^*(p)$ be a finite quasisimple group of Lie type in characteristic p, and let R be a non-trivial p-subgroup of G. Then there is a parabolic subgroup P of G such that $R\leqslant O_p(P)$ and $N_G(R)\leqslant P$.*

We shall distinguish between two cases for H depending on whether H is p-local or not. We refer to the latter case as *atomic*. It is the subject of §§ 3.2.2 and 3.2.3. Assuming that Proposition 2 is proved in the atomic case, we now complete the proof in general. Thus in this section we shall *assume that H is p-local*. Also, since we are not interested in the explicit values of the constants c_0 and c_1, we shall be content to define them recursively from the cases of Proposition 2 for type X having strictly smaller Lie rank l.

Now, by the Borel–Tits Theorem 8 above, we have that H is contained in a proper parabolic P'. Choose P' to be minimal parabolic containing H. Let $U=O_p(P')$ be the unipotent radical of P', and let L be its Levi factor.

Recall that A is an abelian p'-subgroup in the centre of $\overline{H}=H/O_p(H)$. Thus $O_p(H)=H\cap U$, and so $\overline{H}\simeq HU/U$. We can replace H by HU: in this way the index of H in G decreases, while A and $\overline{H}$ stay the same (up to isomorphism). Let H' be the isomorphic image of $\overline{H}$ in $L\simeq P'/U$, and identify A with its isomorphic image in $H'\leqslant L$.

The structure of L is explained in detail in Theorem 2.6.5 of [GLS]:

PROPOSITION 4. *Let $G=X(\mathbf{F})$ be a quasisimple group of Lie type, and let P' be a parabolic subgroup of G with Levi factor L. Define M to be the largest normal subgroup of L such that L/M is an abelian p'-group (so $L/M=(P')^{\diamondsuit}$).*

Then M is a central product of quasisimple groups $L_1, \dots, L_k$ whose types correspond to connected subsets of the Dynkin diagram X of G. When G is universal (i.e., when $X(-)$ is simply-connected), then each L_i is universal and M is in fact the direct product of the L_i.

Moreover there is an abelian p'-subgroup $T=T_L$ of L such that

(1) $L_0:=MT=M\circ T$ is a central product of T and M;

(2) $[L:L_0]\leqslant c_3$ for some constant c_3 depending only on the type X;

(3) T is a direct product of at most $m=m(X)$ cyclic groups whose orders are given by a finite set $\mathcal{A}\subseteq\mathbf{Q}[t]$ of non-constant polynomials in q, depending on X and P only.

Now, let $H_{L_0}=H'\cap L_0$ and $A_{L_0}=A\cap L_0$. Then $[A:A_{L_0}]\leqslant c_3$ and $A_{L_0}\leqslant Z(H_{L_0})$.

Put $H_M=M\cap H_{L_0}$ and $H_T=H_{L_0}\cap T$. Then H_{L_0}/H_M is a quotient of T, so it is abelian, while H_{L_0}/H_T is a quotient of M, so it is perfect. Therefore $H_{L_0}=H_MH_T=H_M\circ H_T$ is a central product of H_M and H_T.

Similarly we have that $A_{L_0}=A_M\circ A_T$, where $A_M=M\cap A_{L_0}$ and $A_T=A_{L_0}\cap T$.

For each direct factor L_i of M, let H_i and A_i be the projections of H_M and A_M, respectively, in L_i. Then A_i is in the centre of H_i, and by the minimality of the parabolic P', each H_i is atomic in L_i.

The atomic case of Proposition 2 applied to $A_i\leqslant H_i\leqslant L_i$ now gives that there exist constants $c(i)$, $i=1,2,...,k$, sets of non-constant polynomials $\mathcal{S}_i\subseteq\mathbf{Q}[x]$ together with an abelian p'-group T_i, and a parabolic P_i of L_i such that

(1) T_i maps onto some subgroup $A_i(0)\leqslant A_i$ of index at most $c(i)$ in A_i;

(2) $|T_i|\leqslant|P_i^{\diamondsuit}|/c(i)$ and $c(i)[L_i:H_i]\geqslant[L_i:P_i]$;

(3) T_i is product of boundedly many cyclic groups each having order $f(q)$ for some $f\in\mathcal{S}_i$.

Put $T'=\prod_{i=1}^k T_i$ and $A(0)=\prod_{i=1}^k A_i(0)\leqslant D:=\prod_{i=1}^k A_i$. We have that A_M is a subdirect product of the A_i. Hence A_M embeds in D, so we can identify A_M as a subgroup of D. Put $A_M(0)=A_M\cap A(0)$. Then

$$[A_M:A_M(0)]\leqslant[D:A(0)]\leqslant\prod_{i=1}^k c(i)=:c_0'.$$

Now $A_M(0)\leqslant A(0)$, and by Pontryagin duality a subgroup of a finite abelian group is also a quotient. Therefore $A_M(0)$ is an image of $A(0)$, hence also an image of T'.

Recall that $A_{L_0}=A_M\circ A_T$. Therefore A_{L_0} is a homomorphic image of $A_M\times A_T$, under a map β, say.

Put $A_0=\beta(A_M(0)\times A_T)$ and $T=T'\times T_L$. Then

$$[A:A_0]\leqslant[A:A_{L_0}][A_{L_0}:A_0]\leqslant c_3c_0'.$$

On the other hand, A_T is a subgroup, hence an image of T_L, and therefore A_0 is an image of $T=T'\times T_L$.

It is clear that T satisfies condition (3) of Proposition 2 for the set of polynomials $\mathcal{S}=\mathcal{A}\cup\mathcal{S}_1\cup\ldots\cup\mathcal{S}_k$. It only remains to define the parabolic P:

$$P := \langle P_1, P_2, \ldots, P_k, B\rangle, \quad \text{where } B \text{ is the Borel subgroup of } G.$$

Then it is easy to see that as $q\to\infty$,

$$[G:P] \sim [G:P'] \prod_{i=1}^{k} [L_i:P_i] \quad \text{and} \quad |P^{\diamond}| \sim |P'/M| \prod_{i=1}^{k} |P_i^{\diamond}|.$$

Also $|T_L| \leqslant |P'/M|\cdot|Z(M)|$ with $|Z(M)|$ bounded by a function of X alone (e.g. $2l$ where l is the untwisted Lie rank of G). Together with

$$[G:H] = [G:P'][L:H'] \geqslant \frac{1}{c_3}[G:P'][L:H_{L_0}]$$

and

$$[L:H_{L_0}] \geqslant \prod_{i=1}^{k} [L_i:H_i]$$

this easily gives that condition (2) in Proposition 2 is satisfied for our choice of P, T and appropriate constant c_0.

This concludes the reduction of Proposition 2 to the atomic case, i.e., when H normalizes no non-trivial p-subgroup of G. This implies that every representation of H over $\mathbf{F}$ is completely reducible.

3.2.2. *The atomic case* I: *Classical groups.* By the remark after Proposition 3, it is enough to prove Proposition 2 for any of the isogeny versions of $X(-)$. In this subsection we consider the case when X is a classical type. Thus we may assume that $G=X(\mathbf{F})$ is one of the classical groups SL_d, Sp_d, SU_d or $\Omega_d^{\pm}$ acting on its associated geometry (V, f) (see Chapter 2 of [KL] for the relevant definitions). Thus V is a vector space of dimension d over the finite field $\mathbf{F}$ with a form $f: V\times V\to\mathbf{F}$ such that one of the following conditions holds:

(a) $f=0$;

(b) f is non-degenerate symmetric or skew-symmetric;

(c) f is non-degenerate Hermitian.

Recall that the characteristic p of $\mathbf{F}$ is assumed to be bigger than 3. In particular, this avoids problems with quadratic forms in characteristic 2.

LEMMA 4. *Suppose that $U \leqslant V$ is an irreducible H-submodule. Then either (U, f) is non-degenerate, or else U is a totally isotropic subspace for f.*

Proof. The assertion is clear in case (a) when f is identically 0. Therefore we can assume that we are in case (b) or (c). Notice that $U^{\perp} := \{v \in V \mid f(u, v) = 0 \text{ for all } u \in U\}$ is an H-submodule of V, and therefore $U \cap U^{\perp}$ is a submodule of U. By the irreducibility of U it follows that either $U \cap U^{\perp} = \{0\}$, in which case U is non-degenerate, or else $U \leqslant U^{\perp}$, i.e., U is totally isotropic. □

The parabolic subgroups of the classical groups are the stabilizers of (chains of) totally isotropic spaces. Therefore the Borel–Tits theorem has the following implication:

The group $H \leqslant G$ is atomic if and only if H stabilizes no non-trivial totally isotropic subspace of V.

In case (a) this means that V is an irreducible H-module. In cases (b) and (c) from Lemma 4 it follows that all irreducible H-submodules of V must be non-degenerate, and then V decomposes as a direct sum

$$V_1 \perp V_2 \perp \ldots \perp V_s$$

of pairwise orthogonal non-degenerate irreducible H-submodules.

Thus we are led to consider the centres of irreducible linear groups preserving a non-degenerate form f. In particular, we have the following lemma:

LEMMA 5. *Let $H \leqslant \mathrm{GL}(V)$ be a finite linear group acting on a vector space V of dimension n over a finite field $\mathbf{F}$. Suppose that*

(1) *H is irreducible over $\mathbf{F}$;*

(2) *H preserves a form $f: V \times V \to \mathbf{F}$ such that one of the following cases holds:*

(a) $f = 0$;

(b) *f is symmetric or skew-symmetric, bilinear and non-degenerate;*

(c) *f is non-degenerate Hermitian (in which case $\mathrm{Aut}(\mathbf{F})$ is assumed to possess an involution σ).*

Then there exists a finite extension E of $\mathbf{F}$ of degree s, say, such that H is isomorphic to a group $H' \leqslant \mathrm{GL}(V')$, where V' is an n/s-dimensional vector space over E and

(1) *H' is absolutely irreducible over E, i.e., $C_{\mathrm{GL}(V')}(H') = E^*$;*

(2) *H' preserves some form $f': V' \times V' \to E$ such that*

(a) $f' = 0$;

(b) *either* (i) *f' is non-degenerate bilinear symmetric or skew-symmetric, or* (ii) *f' is non-degenerate Hermitian and the involution $\sigma' \in \mathrm{Aut}(E)$ fixes $\mathbf{F}$;*

(c) *the form f' is non-degenerate Hermitian, and the involution $\sigma' \in \mathrm{Aut}(E)$ restricts to σ on $\mathbf{F}$.*

COROLLARY 2. *In the situation of Lemma* 5 *above, let* Z *be the centre of* H. *Then* $Z \leqslant E'$, *where the abelian* p'-*group* E' *is defined below for each case*:

(a) $E' := E^*$, *a cyclic group*;

(b)(i) $E' := \{\pm 1\}$;

(b)(ii) and (c) $E' := \{x \in E^* \mid x^{\sigma'} = x^{-1}\}$, *a cyclic group of order* $\sqrt{|E|}+1$.

We delay the proof of Lemma 5 to §3.2.4.

Now return to the problem.

Case (a): $G = \mathrm{SL}_d(q)$. Let $E = \mathrm{End}_{\mathbf{F}H}(V)$ be the splitting field for the irreducible H-module V. Then $s = \dim_{\mathbf{F}} E$ divides the dimension d of V. If $s=1$, take $T=1$.

In case $s>1$, take

$$T = \{x \in E^* \mid \det x = (\mathrm{Norm}_{E/\mathbf{F}_q} x)^{d/s} = 1\} = E^* \cap \mathrm{SL}(V),$$

a cyclic group of order $f_{e,s}(q) = e(q^s-1)/(q-1)$, where $e=(q-1, d/s)$. Again, A is a subgroup, hence a quotient of T. Set $\mathcal{S} = \{f_{e,s}(q) = e(q^s-1)/(q-1) \mid e \text{ and } s>1 \text{ divide } d\}$.

Take $A_0 = A$ and define P to be the stabilizer of the chain

$$\{0\} < U_1 < \ldots < U_{s-1} < V$$

of subspaces U_i with $\dim U_i = di/s$, $i = 1, 2, \ldots, s-1$. If $s=1$ then $P=G$.

Then $\log_q |P^{\diamondsuit}| \sim s-1$ and $\log_q [G:P] \sim \frac{1}{2} d(d-d/s)$.

On the other hand, $H \leqslant \mathrm{End}_E(V) \cap \mathrm{SL}(V, \mathbf{F})$, and therefore

$$\log_q [G:H] \geqslant d^2 - 1 - \left(\frac{d^2}{s} - 1\right) = d\left(d - \frac{d}{s}\right).$$

Thus $[G:H] \geqslant [G:P]$ and $|T|/|P^{\diamondsuit}| = O(1)$ as $q \to \infty$, and we are done.

Case (b): *f is skew-symmetric or symmetric, and G is either* $\mathrm{Sp}_d(q)$ *or* $\Omega_d^{\pm}(q)$. By Lemma 5 the module V decomposes as a sum of irreducible modules

$$(V_1 \oplus \ldots \oplus V_m) \oplus (W_1 \oplus \ldots \oplus W_n),$$

where each V_i has a splitting field E_i and non-degenerate bilinear (symmetric or skew-symmetric) form h_i, say, over E_i preserved by H. On the other hand, each W_j carries a non-degenerate Hermitian form $\bar{h}_j$ over its splitting field K_j. Let V_i' (resp. W_j') denote V_i (resp. W_j) considered as vector space over E_i (resp. K_j) together with its associated non-degenerate form h_i (resp. $\bar{h}_j$).

Let $s_j = [K_j : \mathbf{F}]$. By Lemma 5 (2)(b)(ii), the numbers s_j are even, and K_j has an automorphism σ_j of order 2 fixing $\mathbf{F}$.

Then A acts on each irreducible V_i' as $\{\pm 1\}$, and on each W_j' as $\{x\in K_j^* \mid xx^{\sigma_j}=1\}$, a cyclic group of order $f_j(q)=q^{s_j/2}+1$. Therefore it embeds in

$$\{\pm 1\}^m \times T, \quad \text{where } T := \prod_{j=1}^{n} C_{f_j(q)}.$$

We take $A_0=A\cap T$, where T is as defined above. Set $\mathcal{S}=\{f_1,f_2,\ldots,f_d\}$ and $c_0=2^{\dim X}$, say. We only need to define the parabolic P:

Observe that H embeds in the direct product

$$M := X(V_1')\times\ldots\times X(V_m')\times \mathrm{U}(W_1')\times\ldots\times\mathrm{U}(W_n'),$$

where $X\in\{\mathrm{Sp},\Omega^{\pm}\}$ as appropriate, and $\log_q|T|\sim\frac{1}{2}(s_1+\ldots+s_n)=s$, say.

Let $V_0=V_1\oplus\ldots\oplus V_m$ and $d_i=\dim_{\mathbf{F}_q} W_i'$, $i=1,2,\ldots,n$. Each of the numbers d_i is even. We have that

$$|\mathrm{U}(W_i')|\sim q^{s_i/2(d_i/s_i)^2}=q^{d_i^2/2s_i}.$$

Clearly M is a subgroup of $M':=X(V_0)\times\mathrm{U}(W_1')\times\ldots\times\mathrm{U}(W_n')$.

Let $t=\frac{1}{2}(d_1+\ldots+d_n)$ and consider the chain

$$\{0\}=U_0<U_1<\ldots<U_t<V$$

of t totally isotropic spaces in V, each U_i having codimension 1 in U_{i+1}. Let P be the parabolic in G which is the stabilizer of this chain. Then $|P^\diamondsuit|\sim q^t\geqslant q^s\sim|T|$, and we claim that $|P|\geqslant|M'|$:

It is easy to see that P has a group isomorphic to $X(V_0)$ as a quasisimple component of its Levi factor. Moreover, by its construction the unipotent part of P has dimension at least equal to the number of positive roots in a root system of type D_t, i.e., $t(t-1)$. Hence

$$|P|\geqslant|X(V_0)|\cdot|P^\diamondsuit|q^{t(t-1)}\geqslant|X(V_0)|q^{t^2}.$$

On the other hand, $|M'|=|X(V_0)|\prod_{i=1}^n|\mathrm{U}(W_i')|$. Together with

$$\sum_{i=1}^{n}\log_q|\mathrm{U}(W_i')|\leqslant\sum_{i=1}^{n}\tfrac{1}{4}d_i^2\leqslant t^2$$

this justifies the claim, and we are done.

Case (c): *f is non-degenerate Hermitian and $G=\mathrm{SU}_d(q)$.* The d-dimensional $\mathbf{F}$-vector space V decomposes as an orthogonal sum $V=V_1\perp V_2\perp\ldots\perp V_m$ of irreducible modules V_i. By Lemma 5 there are finite fields E_i such that each V_i is an absolutely irreducible

E_iH-module V_i' which has a non-degenerate Hermitian form h_i over E_i, preserved by H. Thus H embeds in

$$\mathrm{SL}(V)\cap\prod_{i=1}^{m}\mathrm{U}(V_i').$$

Define $G_i=\mathrm{U}(V_i')$ for $i=1,2,\ldots,m$. Let H_i and A_i be the projections of H and A, respectively, into G_i. Then $H\leqslant\mathrm{SL}(V)\cap\prod_{i=1}^m H_i$ and $A\leqslant\mathrm{SL}(V)\cap\prod_{i=1}^m A_i$.

Let $d_i=\dim_{\mathbf{F}} V_i$ and $s_i=\dim_{\mathbf{F}} E_i$. Note that by Lemma 5 (c) all the s_i must be odd. Recall that in this case $\mathbf{F}$ is a quadratic extension of a field $\mathbf{F}_0$ of order q. Let σ_i' be the unique automorphism of order 2 of E_i, and for $i=1,2,\ldots,m$ set

$$T_i=\{x\in E_i^*\mid x^{\sigma_i'}=x^{-1}\},$$

a cyclic group of order $q^{s_i}+1$. Since V_i is absolutely irreducible by Corollary 2 we have that $A_i\leqslant T_i$. Therefore

$$A\leqslant M=\mathrm{SL}(V)\cap\prod_{i=1}^{m}T_i.$$

When $i=m$ and $s_m=1$ set $I=1$; if $s_m>1$ then set

$$I=\{x\in E_m^*\mid x^{\sigma_m'}=x^{-1}\text{ and }\det x=(\mathrm{Norm}_{E_m/\mathbf{F}}(x))^{d_m/s_m}=1\}.$$

It is a cyclic group of order $f_{e,s_m}(q):=e(q^{s_m}+1)/(q+1)$, where $e=(q+1,d/s_m)$. The index of the group $I\mathbf{F}^*\cap T_m$ in T_m is at most d, and therefore by passing to a subgroup of index $\leqslant d$ in A we may assume that

$$A\leqslant M':=\mathrm{SL}(V)\cap(T_1\times\ldots\times T_{m-1}\times(I\circ\mathbf{F}^*)).$$

Let A_0 be the image of A under the projection $\pi: M'\to T_1\times\ldots\times T_{m-1}\times I$. Then we have $|\ker\pi\cap A|\leqslant|\mathbf{F}^*\cap I|\leqslant d$, and so $|A_0|\geqslant|A|/d$. Therefore A_0 is isomorphic to a subgroup of A of index at most d which embeds in

$$T:=T_1\times\ldots\times T_{m-1}\times I.$$

As in case (a) it follows that each H_i has size at most $q^{d_i^2/s_i}$, and hence $\log_q|H|\leqslant(\sum_{i=1}^m d_i^2/s_i)-1$. We set T as above and

$$\mathcal{S}=\left\{q^s+1,\frac{e(q^s+1)}{q+1}\;\middle|\; s\text{ and }e\text{ divide }d,\text{ and }s\text{ is odd}\right\}.$$

Thus T is a product of at most $m\leqslant d$ cyclic groups whose orders are given by polynomials from $\mathcal{S}$, and moreover, $\log_q|T|\sim(\sum_{i=1}^m s_i)-1$. The only thing remaining is to find an appropriate parabolic P satisfying condition (2) of Proposition 2.

Set $v=\sum_{i=1}^m s_i$. Clearly $v\leqslant\sum_{i=1}^m d_i=d=\dim_{\mathbf{F}} V$.

Now, consider the following chain of $\left[\frac{1}{2}v\right]$ totally isotropic spaces in V:

$$\{0\}=U_0<U_1<U_2<\ldots<U_{[v/2]},\tag{1}$$

where each U_i has codimension 1 in U_{i+1}. Let P be the parabolic stabilizing the chain (1).

Now, if $v=1$ then $P=G=\mathrm{SU}_d(q)$, and we are done. Below we assume that $v\geqslant 2$. Then

$$\log_q|P^{\diamond}|\sim 2\left[\tfrac{1}{2}v\right]\geqslant v-1\sim\log_q|T|,$$

and it is easy to see that

$$\log_q|P|\geqslant\tfrac{1}{2}d(d-1)+d-1+\tfrac{1}{2}(d-v)(d-v-1).$$

Now use the following easy result:

LEMMA 6. *Given positive integers d and v, the maximum of the expression*

$$\sum_{i=1}^{m}\frac{d_i^2}{s_i},$$

where $s_i,d_i\in\mathbf{N}$ are subject to $s_i|d_i$, $d=d_1+...+d_m$ and $v=s_1+...+s_m$, is

$$(d-(v-1))^2+v-1,$$

and this maximum is achieved for $d_i=s_i=1$ for all $i=2,3,...,m$.

It follows that $\log_q|H|\leqslant(d-(v-1))^2+v-2$. Thus, in order to prove that $|P|\geqslant|H|$ we need to check that

$$(d-(v-1))^2+v-2\leqslant\tfrac{1}{2}((d+2)(d-1)+(d-v)(d-v-1)),$$

which is in turn equivalent to $(v-2)d\geqslant\frac{1}{2}v(v-3)$, and this inequality holds because $d\geqslant v\geqslant 2$ by our assumption.

This completes the atomic case for the classical groups.

3.2.3. *The atomic case* II: *Exceptional groups.* In this subsection we assume that $G=X(\mathbf{F})$ is a finite quasisimple group of exceptional type in characteristic bigger than 3, so $X\in\{E_6,E_7,E_8,{}^2E_6,{}^3D_4,G_2,F_4\}$ (note that ${}^2D_4$ is not regarded as exceptional since it represents the orthogonal group Ω_8^-).

We shall need some information on centralizers $C_G(x)$ of (non-central) semisimple elements of G. The general structure theory of such centralizers is given in [GLS, Theorem 4.2.2]. In our case, the Lie rank X of G is relatively small (at most 8), so the possibilities for the components of $C_G(x)$ are quite limited. In fact, every such centralizer is contained either in a parabolic, or in a maximal subgroup M of G listed in Tables 5.1 and 5.2 of [LiSS] (the so-called groups of maximal rank).

Recall that in the atomic case A is a subgroup of the centre of H. Provided $|A|$ is big enough (i.e., $|A|>K$ for some constant depending on X only), then A contains a semisimple element x outside the center of G. Then A lies in a maximal torus T' of G and $H \leqslant C_G(x)$. Now, in general, $C_G(x)$ is either contained in a parabolic of G, or else it is contained in a reductive subgroup of *maximal rank* of G, see Theorem 4.2.2 of [GLS]. However, the former possibility is excluded in the atomic case.

The (maximal) subgroups of maximal rank of the exceptional Lie groups have been described by Liebeck, Saxl and Seitz, and the list can be found in Tables 5.1 and 5.2 of [LiSS]. Thus we can assume that $H \leqslant C_G(x) \leqslant M$, where M is one from the lists in Tables 5.1 and 5.2 of [LiSS].

(a) $|M| = O(|B|)$. Observe that if $|M|$ is less than a constant times the order of the Borel subgroup B of G, then we can take the torus $T = T'$ as the required abelian group T and set $A = A_0$: We have $A \leqslant T$, whence A is also an image of T and $|T| \sim q^l$ as $q \to \infty$. Moreover, T is a direct product of at most $l \leqslant 8$ cyclic groups each having order $f_i(q)$, where f_i is from some finite set $\mathcal{S}$ of monic polynomials depending only on the type X of G.

Clearly $H \leqslant M$, and if $|M| \leqslant c_3|B|$ for some constant c_3, then $[G:H] \geqslant [G:B]/c_3$ and $B^{\diamondsuit}$ is isomorphic to the split maximal torus of G, hence $|B^{\diamondsuit}| \sim |T'|$ as $|G| \to \infty$.

Therefore T and B satisfy the requirements in Proposition 2 for appropriate constants c_0 and c_1.

(b) $|M| \geqslant |B|$. The cases where M is larger than the Borel subgroup B are very few: for example, the possibilities for M in [LiSS, Table 5.2] are normalizers of maximal tori and have order bounded by cq^l for some absolute constant c (and l is the Lie rank of G), easily giving that $|M| < |B|$.

By examining Table 5.1 of [LiSS] we list in Table 2 the possibilities for the structure of those M (up to conjugacy). Recall that $q = |\mathbf{F}_0|$, and let d, e and h denote appropriate integers (explicitly defined in [LiSS], but we only need that they are all bounded by an absolute constant). As usual $A.B$ denotes an extension of B by A, and a is a cyclic group of order a. The asymptotic $\sim$ in the last column means that as $q \to \infty$ the quantity tends to the constant specified.

The rest of the argument proceeds on a case-by-case basis:

(1) Suppose that $G = F_4(q)$ and $M = d.B_4(q)$. Thus M is a classical quasisimple group. By the argument in §3.2.2 applied to $H \leqslant M$ we can find groups A_0 and T, and a parabolic P_0 of M, such that the conclusion of Proposition 2 is satisfied for H and P_0 in M. For example, $c[M:H] \geqslant [M:P_0]$, $|T| \leqslant c|P_0^{\diamondsuit}|$, etc. We use the same groups A_0 and T, and we just need to find a parabolic P of $G = F_4(q)$ such that

$$|P_0| = O(|P|) \quad \text{and} \quad |P_0^{\diamondsuit}| = O(|P^{\diamondsuit}|) \quad \text{as } q \to \infty.$$

G	M	$\log_q \lvert M\rvert\sim$	$\log_q \lvert B\rvert\sim$
$F_4(q)$	$d.B_4(q)$	36	28
$E_6(q)$	$h.(D_5(q)\times(q-1)/h).h$	46	42
${}^2E_6(q)$	$h.({}^2D_5(q)\times(q+1)/h).h$	46	42
$E_7(q)$	$e.(E_6(q)\times(q-1)/e).e.2$	79	70
	$e.({}^2E_6(q)\times(q+1)/e).e.2$	79	70
$E_8(q)$	$d.(A_1(q)\times E_7(q)).d$	136	128

Table 2

Now, there are not many possibilities for the parabolic P_0 in $M=B_4(q)$, and clearly if $|P_0|=O(|B|)$ then $P=B$, the Borel subgroup of G will do. It turns out that there is just one parabolic P_0 which fails to have order less than the Borel subgroup, and it is the largest parabolic $P_{\max}$ of M which has order about q^{29}. However, $|P^{\diamond}_{\max}|=O(q)$, and therefore in this case we can take P to be the parabolic of maximal size in G (which has dimension 37 as an algebraic group, and $|P^{\diamond}|\sim q$).

(2) The rest of the cases for M are even simpler: In all of them, M has a subgroup of 'small' (=absolutely bounded) index which is an extension $J\to M\to C\times D$ of a direct product of two groups C and D by a 'small' central subgroup J. By going to a subgroup of small index in H and then factoring J, we may assume that $H\leqslant C\times D$. Moreover, D is a reductive group of rank 1 (either a torus or A_1), and C is one of the simple groups D_5, ${}^2D_5$, E_6, ${}^2E_6$ and E_7 over $\mathbf{F}$.

Let H_C and A_C be the projections of H and A, respectively, into C. If $A_C\neq 1$, then H_C is contained in $N_C(A_C)$, which is a subgroup of maximal rank of C. Therefore $[C:H_C]\geqslant i(C)$, where $i(C)$ is the smallest index of a subgroup of maximal rank of C, and $|H|\leqslant e|C|\cdot|D|/i(C)$ for some absolute constant e. Now the numbers $i(C)$ for $C=E_6$, $C={}^2E_6$ and $C=E_7$ are easy to find from Table 5.1 of [LiSS], and for $C=D_5$ and $C={}^2D_5$ lower bounds for $i(C)$ can be found in [C]. Direct computation then shows that $|H|=o(|B|)$, so we are in the same situation as in case (a).

Therefore we can assume that the projection of $A\leqslant Z(H)$ into C is trivial. It follows that A is a bounded extension of its intersection $A(D)=A\cap D$ with D, which is contained in a 1-dimensional torus T_1.

Thus we can select a subgroup A_0 of small index in A, which is an image of T_1, and for P we take the parabolic of maximal size in G. It is certainly larger than M, and $P^{\diamond}$ is 1-dimensional, i.e., $\log_q|P^{\diamond}|\sim 1$, and so P satisfies the conditions of Proposition 2.

This completes the proof of Proposition 2 in the atomic case.

Theorem 7 has now been proved in full.

3.2.4. *Proof of Lemma* 5. This is well known, but we were unable to find a reference for it in the literature, and we provide the following ad hoc proof.

Recall that an $\mathbf{F}H$-module V is called *absolutely irreducible* if $C_{\mathrm{GL}(V)}(H)=\mathbf{F}^*$, or equivalently, if V stays irreducible over the algebraic closure of $\mathbf{F}$.

Let $E=\mathrm{End}_{\mathbf{F}H}(V)$. By Schur's lemma, E is a finite division ring, and so it is a field. Say that $s=[E:\mathbf{F}]$. Then V becomes a vector space V' over E of dimension n/s, and $G\leqslant \mathrm{GL}(V')$. Moreover, V' is an absolutely irreducible EG-module.

Case (a) is now finished by setting $f'=0$. For cases (b) and (c) we need to work more:

The non-degenerate form defines an antiautomorphism $A\mapsto A^*$ of $\mathrm{End}(V)$ of order 2 given by

$$f(Au,v)=f(u,A^*v)$$

so that A^* is the adjoint of A with respect to f. It is easy to see that E is stable under the adjoint map, and hence it induces an automorphism σ' of E of order at most 2. In case (b), σ' fixes $\mathbf{F}$, while in case (c), $\sigma'|_{\mathbf{F}}=\sigma$. Moreover, as H preserves the form f we have that $g^*=g^{-1}$ for all $g\in H$.

Set $\varepsilon=1$ unless f is skew-symmetric bilinear when we set $\varepsilon=-1$.

LEMMA 7. *In the situation of cases* (b) *and* (c) *there are a non-degenerate form* $f': V'\times V'\to E$ *and an* $\mathbf{F}$*-linear functional* $h: E\to\mathbf{F}$ *such that* $h(x^{\sigma'})=\varepsilon h(x)^{\sigma}$ *and* $f=h\circ f'$. *The form* f' *is bilinear* (*symmetric or skew-symmetric*) *if* $\sigma'=1$, *and Hermitian or skew-Hermitian if* $\sigma'\neq 1$. *More precisely we have*

$$f'(v,u)=\varepsilon f'(u,v)^{\sigma'}. \tag{2}$$

Proof. Fix $v\in V$ and define $h(x):=f(xv,v)$. It satisfies the requirements of the lemma. Now, for any pair of vectors $u,w\in V$ there is a scalar $\lambda(u,w)\in E$ such that

$$f(xu,w)=h(\lambda(u,w)x)\quad \text{for all } x\in E.$$

Then $\lambda(w,u)=\varepsilon\lambda(u,w)^{\sigma'}$.

Let $v_1,v_2,\dots,v_k$ be a basis for $V=V'$ over E (so $k=n/s$). Define f' by

$$f'\left(\sum_{i=1}^{k}\alpha_i v_i,\sum_{j=1}^{k}\beta_j v_j\right)=\sum_{1\leqslant i,j\leqslant k}\lambda(v_i,v_j)\alpha_i\beta_j^{\sigma'}.$$

Then (2) is satisfied, and it is easy to see that $f=h\circ f'$. □

We claim that H preserves the form f':

For a fixed $g\in H$ consider another form $f'': V'\times V'\to E$ defined by

$$f''(u,v)=f'(gu,gv)-f'(u,v).$$

It is of the same type (bilinear or Hermitian) as f', and

$$h\circ f''=f(gu,gv)-f(u,v)=0.$$

Thus $f''(V',V')\subseteq\ker h<E$ giving that $f''=0$. This proves the claim.

To finish the proof of Lemma 5 observe that when $\varepsilon=-1$ and $\sigma'\neq 1$, the form f' is skew-Hermitian. But we may consider instead the form $\mu f'$, where $\mu^{\sigma'}=-\mu$, and this form is Hermitian. (Recall that the characteristic of E is odd, and therefore such $\mu\in E$ always exists.) □

4. The lower bound

In this section we return to the notation from the introduction. Thus G denotes a simple, simply-connected, connected algebraic group defined over a number field k. As explained at the beginning of §2 we can further assume that G is absolutely simple. Fix a linear representation of G, and let Γ be an arithmetic subgroup of G.

The group G is called k-quasisplit if G contains a Borel subgroup defined over k, and G is k-split if it contains a maximal k-torus which is k-split.

Recall that in [GLP] the lower bound from Conjecture 1 was stated and proved for split G. Below we show that with a little modification the same proof applies to the case when G is not necessarily split.

We shall need several basic results from Galois cohomology, which can be found in [PR, §2.2]. Let G_0 be the split form of G (so G_0 is a Chevalley group of type X, say). Given G_0, the possibilities for the k-isomorphism type of G are parametrized by $H^1(\mathrm{Gal}(\bar{k}/k),\mathrm{Aut}_{\bar{k}}(G_0))$, the first cohomology group of the absolute Galois group $\mathrm{Gal}(\bar{k}/k)$ with values in $\mathrm{Aut}_{\bar{k}}(G_0)$, which is usually written as $H^1(k,\mathrm{Aut}_{\bar{k}}(G_0))$.

In turn, $\mathrm{Aut}_{\bar{k}}(G_0)$ is a semidirect product of $\overline{G}=G/Z(G)=G_a$, the adjoint form of G by $\mathrm{Sym}(X)$, the group of symmetries of the Dynkin diagram of X preserving edge lengths:

$$\overline{G}\longrightarrow\mathrm{Aut}_{\bar{k}}(G_0)\longrightarrow\mathrm{Sym}(X).$$

This gives rise to the exact sequence of (non-commutative) cohomology

$$H^1(k,\overline{G})\longrightarrow H^1(k,\mathrm{Aut}_{\bar{k}}(G_0))\xrightarrow{\alpha}H^1(k,\mathrm{Sym}(X)).$$

The group $\mathrm{Gal}(\bar{k}/k)$ acts trivially on $\mathrm{Sym}(X)$, so that the last term is simply the conjugacy classes of homomorphisms of $\mathrm{Gal}(\bar{k}/k)$ into $\mathrm{Sym}(X)$. We observe that when $\mathrm{Sym}(X)$ is non-trivial, it is usually a cyclic group of order 2, with the exception of $X=D_4$, when it is S_3.

The preimage of the trivial homomorphism from $H^1(k,\mathrm{Sym}(X))$ by α inside $H^1(k,\mathrm{Aut}_{\bar{k}}(G_0))$ parametrizes the *inner forms* of G; the rest are called the *outer forms*. Moreover, each fibre of α contains a unique k-quasisplit representative, and for inner forms this is the split form G_0. For example, if k' is a quadratic extension of k, the quasisplit group $\mathrm{SU}_{n+1}(k')$ is an outer k-form (denoted ${}^2A_n$) of $X=A_n$, and the split form is $\mathrm{SL}_{n+1}(k)$. The following proposition (to be used in §5) says that we can always find an extension E of very small degree over k such that G becomes an inner form over E:

PROPOSITION 5. *Let G be an absolutely simple, connected, simply-connected algebraic group over a number field k, and suppose that G is not a form of D_4. Then there exists a Galois field extension E/k such that $[E:k]\leqslant 2$ and G is an inner form over E.*

If G is a form of D_4, then such an E exists with $[E:k]=1,2,3$ or 6, the latter possibility arising only when G is of type ${}^6D_4$.

Proof. This is a consequence of the fact that $\mathrm{Sym}(X)$ is a small group. Let $u\in H^1(k,\mathrm{Aut}_{\bar{k}}(G_0))$. We have to prove the existence of a Galois field E such that the image $\beta\circ a(u)$ in the commutative diagram below is trivial in $H^1(E,\mathrm{Sym}(X))$:

$$\begin{tikzcd} H^1(k,\mathrm{Aut}_{\bar{k}}(G_0)) \arrow[r, "\alpha"] \arrow[d, "a"'] & H^1(k,\mathrm{Sym}(X)) \arrow[d, "b"'] \\ H^1(E,\mathrm{Aut}_{\bar{k}}(G_0)) \arrow[r, "\beta"] & H^1(E,\mathrm{Sym}(X)). \end{tikzcd}$$

Now $\alpha(u)\in H^1(k,\mathrm{Sym}(X))$ is represented by a homomorphism $\mathrm{Gal}(\bar{k}/k)\to\mathrm{Sym}(X)$. Let $Y\leqslant\mathrm{Gal}(\bar{k}/k)$ be the kernel of this homomorphism and let E be the fixed field of Y (so that $Y=\mathrm{Gal}(\bar{k}/E)$). From the definition of E it follows that $b\circ\alpha(u)=1=\beta\circ a(u)$, and we are done. $\square$

Let E be the field given by the above proposition, and suppose that p is a rational prime which splits completely in E. Let π be a prime ideal of the S-integers $\mathcal{O}_S(E)$ of E lying above p, and set $\pi'=\mathcal{O}_S\cap\pi$.

Then

$$\mathcal{O}_S(E)/\pi\simeq\mathcal{O}_S/\pi'\simeq\mathbf{F}_p,$$

and hence $G(\mathcal{O}_S/\pi')=G(\mathbf{F}_p)$ is an inner form of G over $\mathbf{F}_p$.

Let the prime $p\in\mathbf{N}$ be as above. By Lang's theorem each connected algebraic group over a finite field is quasisplit, and so with the strong approximation theorem we conclude that for almost all such p the group Γ maps onto the split Chevalley group $G(\mathbf{F}_p)=X(p)$ of type X over $\mathbf{F}_p$. Notice that these are the same images used to prove the lower bound in [GLP] in the case of *split* G. More precisely, the following theorem is proved there:

THEOREM 9. *Suppose that G is a split Chevalley group, and that k is contained in a Galois field K over $\mathbf{Q}$.*

(i) *Assuming the GRH we have*

$$\alpha_-(\Gamma)\geqslant\frac{\left(\sqrt{R(R+1)}-R\right)^2}{4R^2}.$$

(ii) *Moreover, part* (i) *holds unconditionally if* $\mathrm{Gal}(K/\mathbf{Q})$ *has an abelian subgroup of index at most* 4, *or if* $\deg[K:\mathbf{Q}]<42$.

The proof of Theorem 9 in [GLP] used only the finite images of Γ of the form $G(\mathcal{O}_S/\pi')=G(\mathbf{F}_p)$, where p is a rational prime which splits completely in K. Therefore the same argument proves Theorem 2 (B).

5. Lattices in Lie groups

In this section, H denotes a *semisimple group of characteristic* 0. By this we mean that $H=\prod_{i=1}^r G_i(K_i)$, where for each i, K_i is a local field of characteristic 0 and G_i is a connected simple algebraic group over K_i. The rank of H is defined to be

$$\mathrm{rank}(H)=\sum_{i=1}^{r}\mathrm{rank}_{K_i}(G_i).$$

We assume throughout that none of the factors $G_i(K_i)$ is compact (so that we have $\mathrm{rank}_{K_i}(G_i)\geqslant 1$). Let Γ be an irreducible lattice of H, i.e., for every infinite normal subgroup N of H, the image of Γ in H/N is dense there.

Assume now that $\mathrm{rank}(H)\geqslant 2$. Thus by the Margulis arithmeticity theorem, Γ is an S-arithmetic lattice. More precisely:

THEOREM 10. ([Ma, Theorem 1]) *There exist a number field k, a connected absolutely simple algebraic group G defined over k, and a finite set of valuations S of k containing V_∞, such that H is isomorphic to $G_T=\prod_{v\in T}G_v$ for some set $T\subseteq V$ of valuations of k, and moreover,*

(1) *Γ is the image of some S-arithmetic subgroup of G under the embedding $G(k)\to\prod_{v\in T}G_v$;*

(2) *for all $v\in S\setminus T$, the group G_v is compact.*

Note that the split form of G is uniquely determined by the split form of the simple factors of H, which are necessarily of the same type. We set $\gamma(H):=\gamma(G)$, defined in the introduction for the split form of the algebraic group G.

Since Γ is commensurable with $G(\mathcal{O}_S)$ the two groups have 'roughly the same' subgroup growth. This statement can be made precise, see Proposition 1.11.1 of [LuS]. Passing to the simply-connected cover of G also does not affect the asymptotics of the subgroup growth (see Proposition 1.11.2 of [LuS]), and therefore we can assume that G is in fact simply-connected. As S-rank$(G)=\text{rank}(H)\geqslant 2$, Serre's conjecture (on the finiteness of the congruence kernel of $G(\mathcal{O}_S)$, see [S] and also §7.1 of [LuS] for definition) gives that the congruence subgroup growth of $G(\mathcal{O}_S)$ is asymptotically the same as its subgroup growth.

Now the results of the previous sections (which rely on the GRH at one point: Theorem 2 (B)) imply that

$$\lim_{n\to\infty}\frac{\log s_n(\Gamma)}{(\log n)^2/\log\log n}=\lim_{n\to\infty}\frac{\log C_n(G(\mathcal{O}_S))}{(\log n)^2/\log\log n}=\gamma(G).$$

Thus Theorem 3 is now proved modulo the validity of the generalized Riemann hypothesis for number fields and Serre's conjecture on the finiteness of the congruence kernel. In fact, we have proved more:

THEOREM 11. *Let H be a semisimple group with* $\text{rank}(H)\geqslant 2$. *Assuming the GRH and Serre's conjecture, then for every irreducible lattice Γ of H the limit*

$$\lim_{n\to\infty}\frac{\log s_n(\Gamma)}{(\log n)^2/\log\log n}$$

exists and equals $\gamma(H)$, i.e., it depends only on H and not on Γ.

5.1. Proof of Theorem 1

When H is *simple* and not locally isomorphic to $D_4(\mathbf{C})$, and Γ is a non-uniform lattice (i.e., $\Gamma\backslash H$ is non-compact), we can remove the dependence on the GRH and Serre's conjecture above:

Indeed, then T must consist of a single valuation, and as Γ is non-uniform, G is k-isotropic. Therefore G_v is never compact for any $v\in V$. It follows that $S=T$; in particular, k has only one archimedean valuation. Hence k is either $\mathbf{Q}$ or an imaginary quadratic extension of $\mathbf{Q}$.

Recall that with the exception of $G={}^6D_4$ the extension E given by Proposition 5 has degree at most 3 over k. In that case the Galois closure K of E over $\mathbf{Q}$ is rather small: Its Galois group $\Delta := \mathrm{Gal}(K/\mathbf{Q})$ has subnormal series

$$\Delta \rhd \Delta_1 \rhd \Delta_2,$$

where $[\Delta:\Delta_1]\leqslant 2$, $[\Delta_1:\Delta_2]\leqslant 3$ and Δ_2 is core-free in Δ. An easy group-theoretic argument now gives that $|\Delta|=[K:\mathbf{Q}]$ divides 18 or 8, and for such fields K, Theorem 2 (B)(2) is true unconditionally.

When G is ${}^6D_4$ then E/k may have degree 6 and Galois group S_3, and then $[K:\mathbf{Q}]$ divides 72. The only case not covered by Theorem 2 (B)(2) is when the degree is exactly 72. Indeed, this is the reason that we exclude $D_4(\mathbf{C})$: In this case we must have that k is an imaginary quadratic extension of $\mathbf{Q}$, so H is locally isomorphic to $D_4(\mathbf{C})$. If the form of Γ comes from a form of type ${}^6D_4$ we need to use the GRH. For the other lattices in $D_4(\mathbf{C})$ the result is true unconditionally.

Finally, note that when G is k-isotropic the truth of Serre's conjecture has been verified: see Theorem 9.17 of [PR].

Theorem 1 is now clear.

6. Concluding remarks

Let us relate the results of this paper with those of [BGLM] on one hand, and those of [LiS] and [MP] on the other hand.

Theorem 11 above gives a very precise estimate for the subgroup growth of lattices in higher-rank semisimple groups. By way of contrast, when H is of rank 1 then the type of growth is in general very different: type n^n instead of $n^{\log n/\log\log n}$. (See [LuS, Chapter 7.2] for a detailed discussion; only partial results are known.)

Thus, if $\mathrm{rank}(H)=1$ and $\Gamma\leqslant H$ is a lattice, it is natural to try to study the asymptotic behaviour of $\log s_n(\Gamma)/n\log n$. The following result has been proved recently independently by Liebeck and Shalev and by Müller and Puchta:

THEOREM 12. ([LiS], [MP]) *If $H=\mathrm{PSL}_2(\mathbf{R})$ and Γ is a lattice in H, then*

$$\lim_{n\to\infty}\frac{\log s_n(\Gamma)}{\log n!} = -\chi(\Gamma),$$

where $\chi(\Gamma)$ denotes the Euler characteristic of Γ.

The proof of Theorem 12 relies on the explicit known presentations of lattices in $\mathrm{PSL}_2(\mathbf{R})$ (which are Fuchsian groups). Thus one cannot expect these methods to work for the general groups of rank 1. They still may be extended to the case of groups of rank 1 over non-archimedean local fields of characteristic 0. For such an H every lattice is cocompact and virtually free. The group $H=\mathrm{PSL}_2(\mathbf{Q}_p)$ is an interesting first test case. For some explicit presentatations of lattices there, see [LuW].

We should mention, however, that Theorem 12 in its current form is not true for general lattices in other simple groups of rank 1. Indeed, if $H=\mathrm{PSL}_2(\mathbf{C})$ and Γ is a cocompact subgroup of H, then it follows from Poincaré duality that $\chi(\Gamma)=0$. On the other hand, there exist cocompact lattices in $\mathrm{PSL}_2(\mathbf{C})$ which are mapped onto non-abelian free groups, see [Lu]. For such lattices, clearly

$$\lim_{n\to\infty}\frac{\log s_n(\Gamma)}{\log n!}$$

is positive, if it exists. A similar remark applies to $\mathrm{SO}(n,1)$, when n is odd. (Note that $\mathrm{PSL}_2(\mathbf{C})$ is locally isomorphic to $\mathrm{SO}(3,1)$.)

Recall that with a suitable normalization of the Haar measure on $\mathrm{PSL}_2(\mathbf{R})$, for every lattice Γ in $\mathrm{PSL}_2(\mathbf{R})$ we have $-\chi(\Gamma)=\mathrm{vol}(\mathrm{PSL}_2(\mathbf{R})/\Gamma)$. One may speculate and suggest that for a general lattice Γ in $G=\mathrm{PSL}_2(\mathbf{C})$ (or $G=\mathrm{SO}(n,1)$) the limit

$$\lim_{n\to\infty}\frac{\log s_n(\Gamma)}{\log n!}$$

exists and is proportional to the covolume of Γ in G. This may be a possible way to extend Theorem 12 to more general groups of rank 1.

It is also of interest to relate the results of the current paper to those of [BGLM]. There, the following invariant of a simple Lie group H was studied: For $r\in\mathbf{R}_+$ denote by $\alpha_H(r)$ the number of conjugacy classes of lattices of H of covolume at most r. By a result of Wang this number is finite if H is not $\mathrm{PSL}_2(\mathbf{R})$ or $\mathrm{PSL}_2(\mathbf{C})$. It is proved in [BGLM] that for $H=\mathrm{SO}(d,1)$, $d\geqslant 4$, there exist two positive constants $a(d)$ and $b(d)$ such that

$$a(d)\,r\log r\leqslant \log\alpha_H(r)\leqslant b(d)\,r\log r$$

for all sufficiently large r. It is further conjectured there that for simple Lie groups H of higher rank there exist $a(H)$ and $b(H)$ such that

$$a(H)\frac{(\log r)^2}{\log\log r}\leqslant \log\alpha_H(r)\leqslant b(H)\frac{(\log r)^2}{\log\log r}.$$

The results of the current paper support a stronger conjecture: the limit

$$\lim_{r\to\infty}\frac{\log\alpha_H(r)}{(\log r)^2/\log\log r}$$

exists and equals $\gamma(H)$.

Acknowledgement. The authors thank Inna Korchagina for a helpful discussion on the subgroups of the simple groups of exceptional type. This research was carried out while the second author held a Golda-Meir Postdoctoral Fellowship at the Hebrew University of Jerusalem. The support of the Israel Science Foundation (ISF) and the US–Israel Binational Science Foundation (BSF) is gratefully acknowledged.

References

[BT] BOREL, A. & TITS, J., Éléments unipotents et sous-groupes paraboliques des groupes réductifs, I. *Invent. Math.*, 12 (1971), 95–104.

[BGLM] BURGER, M., GELANDER, T., LUBOTZKY, A. & MOZES, S., Counting hyperbolic manifolds. *Geom. Funct. Anal.*, 12 (2002), 1161–1173.

[C] COOPERSTEIN, B. N., Minimal degree for a permutation representation of a classical group. *Israel J. Math.*, 30 (1978), 213–235.

[E] EDHAN, O., Counting congruence subgroups. M.Sc. Thesis, Hebrew University of Jerusalem, 2003.

[GLNP] GOLDFELD, D., LUBOTZKY, A., NIKOLOV, N. & PYBER, L., Counting primes, groups and manifolds. *Proc. Natl. Acad. Sci. USA*, 101 (2004), 13428–13430.

[GLP] GOLDFELD, D., LUBOTZKY, A. & PYBER, L., Counting congruence subgroups. *Acta Math.*, 193 (2004), 73–104.

[GLS] GORENSTEIN, D., LYONS, R. & SOLOMON, R., *The Classification of the Finite Simple Groups, Number* 3, *Part* I, *Chapter* A. *Almost simple K-groups.* Math. Surveys Monogr., 40.3. Amer. Math. Soc., Providence, RI, 1998.

[KL] KLEIDMAN, P. & LIEBECK, M., *The Subgroup Structure of the Finite Classical Groups.* London Math. Soc. Lecture Note Ser., 129. Cambridge Univ. Press, Cambridge, 1990.

[LaP] LARSEN, M. & PINK, R., Finite subgroups of algebraic groups. To appear in *J. Amer. Math. Soc.*

[LiP] LIEBECK, M. & PYBER, L., Finite linear groups and bounded generation. *Duke Math. J.*, 107 (2001), 159–171.

[LiSS] LIEBECK, M., SAXL, J. & SEITZ, G., Subgroups of maximal rank in finite exceptional groups of Lie type. *Proc. London Math. Soc.*, 65 (1992), 297–325.

[LiS] LIEBECK, M. & SHALEV, A., Fuchsian groups, coverings of Riemann surfaces, subgroup growth, random quotients and random walks. *J. Algebra*, 276 (2004), 552–601.

[Lu] LUBOTZKY, A., Free quotients and the first Betti number of some hyperbolic manifolds. *Transform. Groups*, 1 (1996), 71–82.

[LuS] LUBOTZKY, A. & SEGAL, D., *Subgroup Growth.* Progr. Math., 212. Birkhäuser, Basel, 2003.

[LuW] LUBOTZKY, A. & WEIGEL, T., Lattices of minimal covolume in SL_2 over local fields. *Proc. London Math. Soc.*, 78 (1999), 283–333.

[Ma] MARGULIS, G. A., *Discrete Subgroups of Semisimple Lie Groups.* Ergeb. Math. Grenzgeb., 17. Springer-Verlag, Berlin, 1991.

[MP] MÜLLER, T. W. & PUCHTA, J.-C., Character theory of symmetric groups, subgroup growth of Fuchsian groups and random walks. Preprint.

[PR] PLATONOV, V. P. & RAPINCHUK, A. S., *Algebraic Groups and Number Theory*. "Nauka", Moscow, 1991 (Russian); English translation in Pure Appl. Math., 139. Academic Press, Boston, MA, 1994.

[S] SERRE, J.-P., Le problème des groupes de congruence pour SL_2. *Ann. of Math.*, 92 (1970), 489–527.

ALEXANDER LUBOTZKY
Einstein Institute of Mathematics
The Hebrew University of Jerusalem
IL-91904 Jerusalem
Israel
alexlub@math.huji.ac.il

NIKOLAY NIKOLOV
New College
Oxford University
Oxford OX1 3BN
England, U.K.
nikolay.nikolov@new.ox.ac.uk

Received December 18, 2003
Received in revised form August 12, 2004