COUNTING CONGRUENCE SUBGROUPS

Dorian Goldfeld
Alexander Lubotzky
La'szlo' Pyber

§0. Introduction

Let k be an algebraic number field, $\mathcal{O}$ its ring of integers, S a finite set of valuations of k (including all the archimedean ones), and $\mathcal{O}_S = \big\{ x \in k \; \big| \; v(x) \geq 0, \; \forall v \notin S \big\}$. Let G be a semisimple, simply connected, connected algebraic group defined over k with a fixed embedding into GL_d. Let $\Gamma = G(\mathcal{O}_S) = G \cap GL_d(\mathcal{O}_S)$ be the corresponding S-arithmetic group. We assume that Γ is an infinite group.

For every non-zero ideal I of $\mathcal{O}_S$ let $\Gamma(I) = \mathrm{Ker}\big(\Gamma \to GL_d(\mathcal{O}_S/I)\big)$. A subgroup of Γ is called a congruence subgroup if it contains $\Gamma(I)$ for some I. For $n > 0$, define

$$C_n(\Gamma) = \#\big\{\text{congruence subgroups of } \Gamma \text{ of index at most } n\big\}.$$

Theorem 1. *There exist two positive real numbers α_- and α_+ such that for all sufficiently large positive integers n*

$$n^{\frac{\log n}{\log\log n}\alpha_-} \; \leq \; C_n(\Gamma) \; \leq \; n^{\frac{\log n}{\log\log n}\alpha_+}.$$

This theorem is proved in [Lu], although the proof of the lower bound presented there requires the prime number theorem on arithmetic progressions in an interval where its validity depends on the GRH (generalized Riemann hypothesis for arithmetic progressions). In §2 below, we show that by appealing to a theorem of Linnik [Li] on the least prime in an arithmetic progression, the proof can be made unconditional.

Typeset by $\mathcal{A}\mathcal{M}\mathcal{S}$-TEX

Following [Lu] we define:

$$\alpha_+(\Gamma) = \limsup_n \frac{\log C_n(\Gamma)}{\lambda(n)},$$
$$\alpha_-(\Gamma) = \liminf_n \frac{\log C_n(\Gamma)}{\lambda(n)},$$

where $\lambda(n) = \frac{(\log n)^2}{\log\log n}$.

It is not difficult to see that α_+ and α_- are independent of both the choice of the representation of G as a matrix group, as well as independent of the choice of S. Hence $\alpha_\pm$ depend only on G and k. The question whether $\alpha_+(\Gamma) = \alpha_-(\Gamma)$ and the challenge to evaluate them for $\Gamma = SL_2(\mathbb{Z})$ and other groups were presented in [Lu]. Here we prove:

Theorem 2. $\alpha_+(SL_2(\mathbb{Z})) = \alpha_-(SL_2(\mathbb{Z})) = \frac{3}{4} - \frac{1}{\sqrt{2}} = 0.0428932\dots$

We believe that $SL_2(\mathbb{Z})$ represents the general case and we expect that $\alpha_+ = \alpha_-$ for all groups.

The proof of the lower bound in Theorem 1 is based on the Bombieri-Vinogradov Theorem [Bo], [Vi], [Da], i.e., *the Riemann hypothesis on the average.* The upper bound, on the other hand, is proved by reducing the problem to a counting problem for subgroups of abelian groups and then solving that extremal counting problem.

We will, in fact, show a more remarkable result: the answer is independent of $\mathcal{O}$!

Theorem 3. *Let k be any number field, $\mathcal{O}$ its ring of integers, S a finite set of primes, and $\mathcal{O}_S$ as above. Then*

$$\alpha_+(SL_2(\mathcal{O}_S)) = \alpha_-(SL_2(\mathcal{O}_S)) = \frac{3}{4} - \frac{1}{\sqrt{2}}.$$

We conjecture that for every Chevalley group scheme G, the upper and lower limiting constants, $\alpha_\pm(G(\mathcal{O}_S))$, depend only on G and not on $\mathcal{O}$. In fact, we have a precise conjecture, for which we need to introduce some additional notation. Let G be a Chevalley group scheme, d its dimension, ℓ, its rank, and $\kappa = |\Phi^+|$ the number of positive roots in the root system of G. Letting $R = R(G) = \frac{d-\ell}{2\ell}$, we see that

$$R = \frac{\ell+1}{2}, \quad (\text{resp.,} \qquad)$$

if G is of type A_ℓ (resp. $B_\ell, C_\ell, D_\ell, G_2, F_4, E_6, E_7, E_8$).

Conjecture. *Let $k, \mathcal{O}$, and S be as in Theorem 2, and suppose that G is a simple Chevalley group scheme. Then*

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

The conjecture reflects the belief that "most" subgroups of $H = G(\mathbb{Z}/m\mathbb{Z})$ lie between the Borel subgroup B of H and the unipotent radical of B. Our proof covers the case of SL_2 and we are quite convinced that this will hold in general. For general G, we do not have such an in depth knowledge of the subgroups of $G(\mathbb{F}_q)$ as we do for $G = SL_2$, yet we can still prove:

Theorem 4. *With $k, \mathcal{O}$, and S as in Theorem 3, G a simple Chevalley group scheme of dimension d and rank ℓ, and $R = R(G) = \frac{d-\ell}{2\ell}$, then:*

(a) $\alpha_-(G(\mathcal{O}_S)) \geq \frac{\left(R-\sqrt{R(R+1)}\right)^2}{4R^2}$.

(b) *There exists an absolute constant C such that $\alpha_+(G(\mathcal{O}_S)) \leq$???.*

Corollary 5. *There exists an absolute constant C such that for $d = 2, 3, \ldots$*

$$??? \leq \alpha_-(SL_d(\mathbb{Z})) \leq \alpha_+(SL_d(\mathbb{Z})) \leq C\frac{1}{d^2}.$$

This greatly improves the upper bound $\alpha_+(SL_d(\mathbb{Z})) < \frac{5}{4}d^2$ implicit in [Lu] and settles a question asked there.

The paper is organized as follows.

In §1, we present some require preliminaries and notations.

In §2, we prove the lower bound of Theorem 1. As shown in [Lu] this depends essentially on having uniform bounds on the error term in the prime number theorem along arithmetic progressions. The choice of parameters in [Lu] needed an estimate on this error term in a domain in which it is known only modulo the GRH. We show here that by a slight modification of the proof and an appeal to a result of Linnik the proof will be unconditional. Still, if one is interested in good lower bounds on $\alpha_-(\Gamma)$, better estimates on the error terms are needed. To obtain unconditional results (independent of the GRH), we will use the Bombieri–Vinogradov Theorem [Bo], [Vi], [Da].

In §3, we introduce the notion of Bombieri set which is the crucial ingredient needed in the proof of the lower bounds. We then use it in §4 to prove the lower bounds of Theorems 2, 3, and 4.

We then turn to the proof of the upper bounds. In §5, we present a reduction lemma (Proposition 5.1) which plays an important role in several steps of the proofs. We then show in §6, how the counting problem of congruence subgroups in $SL_2(\mathbb{Z})$ can be completely reduced to an extremal counting problem in a finite abelian group; the problem is actually, as one may expect, a number theoretic extremal problem - see §7 where this extremal problem is solved and the upper bounds of Theorems 1 and 2 are then deduced in §8. Finally, in §9 we give the upper bound of Theorem 3.

§1. Preliminaries and notations

Throughout this paper we let $\ell(n)$ denote $\log n/\log\log n$ and $\lambda(n) = (\log n)^2/\log\log n$. If f and g are functions of n, we will say that f *is small w.r.t.* g if $\lim\limits_{n\to\infty} \frac{\log f(n)}{\log g(n)} = 0$. We say that f is *small* if f is *small* with respect to $n^{\ell(n)}$. Note that if f is small, then multiplying $C_n(\Gamma)$ by f will have no effect in the estimate on the estimates of $\alpha_+(\Gamma)$ or $\alpha_-(\Gamma)$. We may, and we will, ignore factors which are small.

Note also that if $\varepsilon(n)$ is a function of n which is smaller than n
(i.e., $\log\varepsilon(n) = o(\log n)$) then:

$$\overline{\lim}\frac{\log C_{n\varepsilon(n)}(\Gamma)}{\lambda(n)} = \alpha_+(\Gamma) \tag{1.1}$$

and

$$\underline{\lim}\frac{\log C_{n\varepsilon(n)}(\Gamma)}{\lambda(n)} = \alpha_-(\Gamma). \tag{1.2}$$

Indeed, to prove (1.1) it suffices to show that $\overline{\lim}\frac{\log C_{n\varepsilon(n)}(\Gamma)}{(\log n)^2/\log\log n} \leq \alpha_+(\Gamma)$. Now

$$\begin{aligned}\overline{\lim}\frac{\log C_{n\varepsilon(n)}(\Gamma)}{\lambda(n)} &= \overline{\lim}\frac{\log C_{n\varepsilon(n)}(\Gamma)}{\lambda(n\varepsilon(n))}\cdot\frac{\lambda(n\varepsilon(n))}{X(n)}\\ &\leq \alpha_+(\Gamma)\cdot 1\\ &= \alpha_+(\Gamma).\end{aligned}$$

The inequality follows from the fact that $\overline{\lim}\frac{\lambda(n\varepsilon(n))}{\lambda(n)} = 1$ which is an immediate consequence of the assumption that $\varepsilon(n)$ is small w.r.t. n. A similar argument proves (1.2).

It follows that we can, and we will sometimes indeed, enlarge n a bit when evaluating $C_n(\Gamma)$, again without influencing α_+ or α_-. Similar remarks apply if we divide n by $\varepsilon(n)$ provided $\varepsilon(n)$ is bounded away from 0.

The following lemma is proved in [Lu] in a slightly weaker form. Here, when we care about the constant appearing in the exponents, we need the more precise formulation:

Lemma 1.1. *("Level versus index" - (see* [LS], *). Let Γ be as before. Then there exists a function $\varepsilon(n)$ small w.r.t. n., such that if H is a congruence subgroup of Γ of index at most n, it contains $\Gamma(m)$ for some $m \leq n\varepsilon(n)$.*

Corollary 1.2. *Let $\gamma_n(\Gamma) = \sum\limits_{m=1}^{n} s_n(G(\mathbb{Z}/m\mathbb{Z})$, where for a group H, $s_m(H)$ denotes the number of subgroups of H of index at most m. Then $\alpha_+(\Gamma) = \limsup \log \gamma_n(\Gamma)/\lambda(n)$ and $\alpha_-(\Gamma) = \liminf \log \gamma_n(\Gamma)/\lambda(n)$.*

Proof of Corollary. By the Lemma, $C_n(\Gamma) \leq \gamma_{n\varepsilon(n)}(\Gamma)$. It is also clear that $\gamma_n(\Gamma) \leq n \cdot C_n(\Gamma)$. So, an argument as above implies the corollary. □

§2. Proof of Theorem 1

Before proving the theorem, a remark is in order (see also [Lu]): we may change S, as long as $\Gamma = G(\mathcal{O}_S)$ is infinite without changing α_- or α_+. We may, therefore, enlarge S to contain all the primes above some set of rational primes S_0. We can then use restriction of scalar to get a group $\tilde{G}$ over $\mathbb{Q}$, with $\tilde{G}(\mathbb{Z}_{S_0}) = G(\mathcal{O}_S)$. It suffices, therefore, to prove the theorem for $k = \mathbb{Q}$. Now, for simplifying the notations we assume $S = \{\infty\}$ and so $\Gamma = G(\mathbb{Z})$. For an integer m we denote $G(\mathbb{Z}/m\mathbb{Z})$ the image of Γ in $GL_d(\mathbb{Z}/m\mathbb{Z})$. If $m = p$ is a prime, G may be considered, with the possible exceptions of finitely many primes, as a group over $\mathbb{F}_p$. By the Strong Approximation Theorem (see [PR,]) the image of Γ is indeed the $\mathbb{F}_p$-points of this group.

The proof of the upper bound is already given in [Lu] and we do not reproduce that here. The proof of the lower bound in Theorem 1 will follow the footsteps of the proof given in [Lu]; in fact, it will actually simplify it a bit. The main new ingredient is the use of a deep result of Linnik, [Li] giving an estimate for the number of primes in a *short interval* of an arithmetic progression. A result of that kind was also used in [Lu], but because of an uncareful choice of the parameters, the interval was *very short*, and the validity of the prime number theorem there is known only module GRH.

We introduce some notations which are needed here and for the next section. Let a, q be relatively prime integers with $q > 0$. For $x > 0$, let $\mathcal{P}(x; q, a)$ be the set of primes p with $p \le x$ and $p \equiv a(\mod q)$. For $a = 1$, we set $\mathcal{P}(x; q) = \mathcal{P}(x; q, 1)$. We also define $\nu(x; q, a) = \sum_{p \in \mathcal{P}(x;q,a)} \log p$ and $\nu(x; q) = \nu(x; q, 1)$.

If $f(x)$, $g(x)$ are arbitrary functions of a real variable x, we say $f(x) \sim g(x)$ as $x \to \infty$ if

$$\lim_{x \to \infty} \frac{f(x)}{g(x)} \;=\; 1.$$

Theorem 2.1 (Linnik, [Li]**).** *There exist effectively computable constants* $c_0, c_1 > 1$ *such that if* a *and* q *are relatively prime integers,* $q \ge 2$ *and* $x \ge q^{c_0}$*, then*

$$\nu(x; q, a) \ge \frac{x}{c_1 q^2 \varphi(k)}$$

where φ *is the Euler function.*

Let now x be a large number and q a prime with $q \sim x^{1/c_0}$. Let X be a subset of $\mathcal{P}(x; q)$ satisfying

$$\sum_{p \in X} \log p \sim \frac{x}{c_1 q^2 \varphi(q)} \sim \frac{1}{c_1} x^{1 - \frac{3}{c_0}}.$$

We also define $P = \prod_{p \in X} p$. It follows from Theorem 2.1 that

$$\log P \sim \frac{x}{c_1 x^{3/c_0}}.$$

Let now $\Gamma(P)$ be the principal congruence subgroup. It is of index approximately $P^{\dim G}$ in Γ and by Strong Approximation, $\Gamma/\Gamma(P) = \prod_{p \in \mathcal{P}(x;q)} G(\mathbb{F}_p)$, where G is considered as a group defined over $\mathbb{F}_p$. (This can be done for almost all p's and we can ignore the finitely many exceptions). Moreover, by a Theorem of Lang (see [PR,]) G is quasi-split over $\mathbb{F}_p$, which implies that G has a split one dimensional torus, so $G(\mathbb{F}_p)$ has a subgroup isomorphic to $\mathbb{F}_p^\times$. The latter is a cyclic group of order $p - 1$.

Since $q|p-1$, it follows that $G(\mathbb{F}_p)$ contains a cyclic group of order q and $\Gamma/\Gamma(P)$ contain a subgroup isomorphic to $(\mathbb{Z}/q\mathbb{Z})^L$ where $L = \#X$.

It now follows from Theorem 2.1 and the choice of X, that $L \sim \frac{x}{c_1 q^2 \varphi(x) \log x}$. On the other hand, the abelian group $(\mathbb{Z}/q\mathbb{Z})^L$ has $q^{\frac{1}{4}L^2 + \mathcal{O}(L)}$ subgroups as $L \to \infty$ (cf. [LS,]). Consequently, Γ has at least $q^{\frac{1}{4}L^2+O(L)}$ subgroups of index at most $P^{\dim(G)}$.

Taking logarithms, we compute:

$$\frac{\log(\#\text{subgroups})}{(\log(\text{index}))^2/\log\log(\text{index})} \geq \frac{(\frac{1}{4}L^2 + O(L))\log q}{(\log(P^{\dim(G)}))^2/\log\log(P^{\dim G})} \geq$$

$$\frac{\left(\frac{x}{c_1(x^{1/c_0})^3 \log x}\right)^2 \cdot \frac{1}{c_0}\log x}{\dim(G)^2 \frac{1}{c_1^2}(x^{1-\frac{3}{c_0}})^2/(1-\frac{3}{x_0})\log x} = \frac{(1-\frac{3}{c_0})}{c_0(\dim G)^2}.$$

This finishes the proof of the lower bound with $\alpha_- = \frac{c}{(\dim G)^2}$ for some constant c.

When one is interested in better estimates on α_-, Linnik's result is not sufficient. We show, however, in the next two sections, that the Bombieri–Vinogradov Theorem, *Riemann hypothesis on the average*, suffices to get lower bounds on α_- which are as good as can be obtained using GRH (though the construction of the appropriate congruence subgroup is probabilistic and not effective).

§3. Bombieri Sets.

Let a, q be relatively prime integers with $q > 0$. For $x > 0$ let

$$\vartheta(x;q,a) \;=\; \sum_{\substack{p \leq x \\ p \equiv a \pmod q}} \log p,$$

where the sum ranges over rational prime numbers p. Define the error term

$$E(x;q,a) = \vartheta(x;q,a) - \frac{x}{\phi(q)},$$

where $\phi(q)$ is Euler's function. Then Bombieri proved the following deep Theorem [Bo], [Da] (see also Vinogradov [Vi]).

Theorem 3.1. *(Bombieri) Let $A > 0$ be fixed. Then there exists a constant $c(A) > 0$ such that*

$$\sum_{q \le \frac{\sqrt{x}}{(\log x)^A}} \max_{y \le x} \max_{(a,q)=1} \big| E(y; q, a) \big| \;\le\; c(A) \cdot \frac{x}{(\log x)^{A-5}}$$

as $x \to \infty$.

This theorem shows that the error terms $\max\limits_{(a,q)=1} E(x; q, a)$ behave as if they satisfy the Riemann hypothesis in an averaged sense.

Definition 3.2. *Let x be a large positive real number. A* **Bombieri prime** *(relative to x) is a prime $q \le \sqrt{x}$ such that the set $\mathcal{P}(x, q)$ of primes $p \le x$ with $p \equiv 1 \pmod q$ satisfies*

$$\max_{y \,\le\, x} |E(y; q, 1)| \;\le\; \frac{x}{\phi(q)(\log x)^2}.$$

We call $\mathcal{P}(x, q)$ a **Bombieri set** *(relative to x).*

Lemma 3.3. *Fix $0 < \rho < \frac{1}{2}$. Then for x sufficiently large, there exists at least one Bombieri set $\mathcal{P}(x, q)$ (relative to x) with q a Bombieri prime in the interval*

$$\frac{x^\rho}{\log x} \le q \le x^\rho.$$

Proof:. Assume that

$$\max_{y \,\le\, x} |E(y; q, 1)| \;>\; \frac{x}{\phi(q)(\log x)^2}$$

for all primes $\frac{x^\rho}{\log x} \le q \le x^\rho$, i.e., that there are no such Bombieri sets in the interval. In view of the trivial inequality, $\phi(q) = q - 1 < q$, it immediately follows that

$$\sum_{\frac{x^\rho}{\log x} \le q \le x^\rho} \max_{y \,\le\, x} \big| E(y; q, 1) \big| \;>\; \frac{x}{(\log x)^2} \sum_{\frac{x^\rho}{\log x} \le q \le x^\rho} \frac{1}{q} \;>\; \frac{x \cdot (\log\log x)^2}{2\rho \cdot (\log x)^3},$$

say, for sufficiently large x. This follows from the well known asymptotic formula [Lan] for the partial sum of the reciprocal of the primes

$$\sum_{q \,\le\, Y} \frac{1}{q} \;=\; \log\log Y + b + O\left(\frac{1}{\log Y}\right)$$

as $Y \to \infty$. Here b is an absolute constant. This contradicts Theorem 3.1 with $A \ge 8$ provided x is sufficiently large. $\square$

Lemma 3.4. *Let $\mathcal{P}(x,q)$ be a Bombieri set. Then for x sufficiently large*

$$\left|\#\{\mathcal{P}(x,q)\} \; - \; \frac{x}{\phi(q)\log x}\right| \; \le \; 3\left(\frac{x}{\phi(q)(\log x)^2}\right).$$

Proof. We have

$$\begin{aligned}\sum_{p\in\mathcal{P}(x,q)} 1 \; &= \; \sum_{n=2}^{x} \frac{\vartheta(n;q,1)-\vartheta(n-1;q,1)}{\log n}\\ &= \sum_{n=2}^{x} \vartheta(n;q,1)\Big(\frac{1}{\log(n)} - \frac{1}{\log(n+1)}\Big) + \frac{\vartheta(x;q,1)}{\log([x]+1)}\\ &= \sum_{n=2}^{x} \vartheta(n;q,1)\frac{\log\left(1+\frac{1}{n}\right)}{\log n\log(n+1)} + \frac{\vartheta(x;q,1)}{\log x} + \vartheta(x;q,1)\left(\frac{1}{\log x} - \frac{1}{\log(x+1)}\right).\end{aligned}$$

By the property of a Bombieri set, we have the estimate $|\vartheta(n;q,1) - \frac{n}{\phi(q)}| \le \frac{x}{\phi(q)(\log x)^2}$, for $n \le x$. It easily follows that

$$\begin{aligned}\left|\sum_{p\in\mathcal{P}(x,q)} 1 - \frac{\vartheta(x;q,1)}{\log x}\right| \; &\le \; \sum_{n=2}^{x} \vartheta(n;q,1)\frac{1}{n\cdot(\log n)^2} + \vartheta(x;q,1)\left(\frac{1}{\log x} - \frac{1}{\log(x+1)}\right)\\ &\le \; 3\left(\frac{x}{\phi(q)(\log x)^2}\right).\end{aligned}$$

□

§4. Proof of the lower bound in Theorems 2, 3, 4.

Fix a real number $0 < \rho < \frac{1}{2}$. It follows from Lemma 3.3 that for every $x \to \infty$ there exists a prime number $q \sim x^\rho$ such that $\mathcal{P}(x,q)$ is a Bombieri set.

Define

$$P = \prod_{p\,\in\,\mathcal{P}(x,q)} p.$$

It is clear from the definition of Bombieri set that

$$\log P \; \sim \; \frac{x}{\phi(q)} \sim x^{1-\rho}.$$

Consider $\Gamma(P)$ which is of index at most $P^{\dim(G)}$ in Γ. Note that for every subgroup $H/\Gamma(P)$ in $\Gamma/\Gamma(P)$ there corresponds a subgroup H in Γ of index at most $P^{\dim(G)}$ in Γ.

By strong approximation

$$\Gamma/\Gamma(P) = G\left(\mathbb{Z}/P\mathbb{Z}\right) = \prod_{p\in\mathcal{P}(x,q)} G(\mathbb{F}_p).$$

It follows from Lang's Theorem [PR] that G is quasi–split over $\mathbb{F}_p$, which implies that G has a split one–dimensional torus, i.e., $G(\mathbb{F}_p)$ has a subgroup isomorphic to $\mathbb{F}_p^\times$. The latter is a cyclic group of order $p-1$. Since $q|(p-1)$, it follows that $G(\mathbb{F}_p)$ contains a cyclic group of order q and $\Gamma/\Gamma(P)$ contains a subgroup isomorphic to $(\mathbb{Z}/q\mathbb{Z})^L$ where

$$L = \#\{\mathcal{P}(x,q)\}.$$

Note that by Lemma 3.4

$$L \sim \frac{x}{\phi(q)\log x} \sim \frac{x^{1-\rho}}{\log x}.$$

We shall need a basic lemma on counting subspaces of a vector space defined over a finite field (see [Lb], p. xxx).

Lemma 4.1. *Let q be a fixed prime. Let $V = \mathbb{F}_q^\ell$ be an ℓ dimensional vector space over $\mathbb{F}_q$. For $1 \le r \le \ell$, define*

$$\Phi(q,\ell,r) = \#\{\text{subspaces of } V \text{ of dimension } r\}.$$

Then

(i) $q^{r\cdot(\ell-r)} \le \Phi(q,\ell,r) \le q^{r\cdot(\ell-r+1)}$,

(ii) $\max\limits_{1\le r\le\ell}\Phi(q,\ell,r)$ *is obtained for* $r = \left[\frac{\ell}{2}\right]$, *and in this case,*

$$\Phi(q,\ell,r) \; = \; q^{\frac{1}{4}\ell^2+O(\ell)}, \qquad (as\ \ell\to\infty).$$

By Lemma 4.1, the elementary abelian group $\left(\mathbb{Z}/q\mathbb{Z}\right)^L$ has

$$q^{\frac{1}{4}L^2+O(L)}$$

subgroups as $L \to \infty$. Consequently, Γ has $q^{\frac{1}{4}L^2+O(L)}$ subgroups of index at most $P^{\dim(G)}$.

Taking logarithms, we compute

$$\begin{aligned}
\frac{\log\Big(\#\{\text{of subgroups}\}\Big)}{\big(\log(\text{index})\big)^2/\log\log(\text{index})} &\geq \frac{\big(\frac{1}{4}L^2+O(L)\big)\log q}{\big(\log\big(P^{\dim(G)}\big)\big)^2/\log\log\big(P^{\dim(G)}\big)} \\
&\gtrsim \frac{\frac{1}{4}\Big[\Big(\frac{x^{1-\rho}}{\log x}\Big)^2 \;+\; O\Big(\frac{x^{1-\rho}}{\log x}\Big)\Big]\cdot\rho\log x}{(\dim(G))^2x^{2-2\rho}/\big(\log(\dim(G))+(1-\rho)\log x\big)} \\
&\geq \frac{\frac{1}{4}\rho(1-\rho)}{(\dim(G))^2}
\end{aligned}$$

as $x \to \infty$.

Choosing $\rho = \frac{1}{2}$, we have proved the lower bound with $\alpha_- = \frac{1}{16(\dim(G))^2}$.

This proof follows the proof of [Lu1] where the Bombieri set and q replace the choice of m there. The reader is referred to [Lu1] for the proof of the upper bound.

We now show how to improve the lower bound $\alpha_- = \frac{1}{16(\dim(G))^2}$ and turn our attention to Theorems 2, 3, 4. We first consider the case when we are working over $\mathbb{Q}$ instead of a general number field.

As in section 3, let $x \to \infty$, fix a positive $\rho < \frac{1}{2}$, and let $\mathcal{P}(x,q)$ be a Bombieri set where $q \sim x^\rho$ is a prime number. We know such a Bombieri set exists by Lemma 3.3. Define, as before,

$$P = \prod_{p\in\mathcal{P}(x,q)} p,$$

where

$$\log P \;\sim\; \frac{x}{\phi(q)} \sim x^{1-\rho}.$$

Let $B(p)$ denote the Borel subgroup of upper triangular matrices in $G(\mathbb{F}_p)$. Then

$$\log\Big(\#\big\{B(p)\big\}\Big) \;\sim\; \frac{\dim(G)+\mathrm{rk}(G)}{2}\,\log p.$$

But

$$\log\Big(\#\big\{SL_r(\mathbb{F}_p)\big\}\Big) \;\sim\; \dim(G)\,\log p.$$

It immediately follows that (for $p \to \infty$)

$$\log \big[G(\mathbb{F}_p) : B(p)\big] \;\sim\; \frac{\dim(G) - \mathrm{rk}(G)}{2}\,\log p,$$

and, therefore,

$$\log\left(G(\mathbb{Z}/P\mathbb{Z} : B(P)\right) \sim \frac{\dim(G) - \mathrm{rk}(G)}{2}\,\log P.$$

Now $B(p)$ is mapped onto ${\mathbb{F}_p^{\times}}^{\mathrm{rk}(G)}$ and, hence, is also mapped onto $(\mathbb{Z}/q\mathbb{Z})^{\mathrm{rk}(G)}$ since $\#\{F_p^{\times}\} = p-1$ and $p \equiv 1 \pmod q$. So $B(P)$ is mapped onto

$$(\mathbb{Z}/q\mathbb{Z})^{\mathrm{rk}(G)\cdot L}$$

where

$$L = \#\{\mathcal{P}(x,q)\} \sim \frac{x}{\phi(q)\log x} \sim \frac{x^{1-\rho}}{\log x}.$$

For a real number θ, define $\lceil\theta\rceil$ to be the smallest integer t such that $\theta \le t$. Let $0 \le \sigma \le 1$. It follows from Lemma 4.1 that $B(P)$ has at least

$$q^{\sigma(1-\sigma)\mathrm{rk}(G)^2 L^2 + O(L)}$$

subgroups of index equal to

$$q^{\lceil \sigma\cdot\mathrm{rk}(G)\cdot L\rceil}\cdot\Big[G(\mathbb{Z}/P\mathbb{Z}) : B(P)\Big].$$

Hence, for $x \to \infty$,

$$\begin{aligned}\log\Big(\#\big\{\text{of subgroups}\big\}\Big) &= \Big(\sigma(1-\sigma)\mathrm{rk}(G)^2L^2 + O(\mathrm{rk}(G)\cdot L)\Big)\log q\\ \sim\;\; &\sigma(1-\sigma)\mathrm{rk}(G)^2\frac{x^{2-2\rho}}{(\log x)^2}\cdot\rho\log x,\end{aligned}$$

while

$$\begin{aligned}\log(\text{index}) &= \lceil \sigma\cdot\mathrm{rk}(G)\cdot L\rceil\cdot\log q + \frac{1}{2}\big(\dim(G)-\mathrm{rk}(G)\big)\log P\\ &\sim \mathrm{rk}(G)\sigma\frac{x^{1-\rho}}{\log x}\rho\log x \;+\; \frac{1}{2}\big(\dim(G)-\mathrm{rk}(G)\big)x^{1-\rho}\\ &= \Big(\sigma\cdot\rho\cdot\mathrm{rk}(G) + \frac{1}{2}\big(\dim(G)-\mathrm{rk}(G)\big)\Big)x^{1-\rho},\end{aligned}$$

and

$$\log\log(\text{index}) \sim (1-\rho)\log x.$$

We compute

$$\frac{\log\Big(\#\{\text{of subgroups}\}\Big)}{\big(\log(\text{index})\big)^2/\log\log(\text{index})} \sim \frac{\sigma(1-\sigma)\cdot\operatorname{rk}(G)^2\cdot\rho\,\frac{x^{2-2\rho}}{\log x}}{\Big(\sigma\cdot\rho\cdot\operatorname{rk}(G)+\frac{1}{2}\big(\dim(G)-\operatorname{rk}(G)\big)\Big)x^{1-\rho}\Big/(1-\rho)\log x}$$

$$\sim \frac{\sigma(1-\sigma)\rho(1-\rho)\cdot\operatorname{rk}(G)^2}{\Big(\big(\sigma\rho-\frac{1}{2}\big)\cdot\operatorname{rk}(G)+\frac{1}{2}\dim(G)\Big)^2}$$

as $x\to\infty$.

We may rewrite

$$\frac{\sigma(1-\sigma)\rho(1-\rho)\cdot\operatorname{rk}(G)^2}{\Big(\big(\sigma\rho-\frac{1}{2}\big)\cdot\operatorname{rk}(G)+\frac{1}{2}\dim(G)\Big)^2} = \frac{\sigma(1-\sigma)\rho(1-\rho)}{(\sigma\rho+R)^2}$$

where

$$R=\frac{\dim(G)-\operatorname{rk}(G)}{2\cdot\operatorname{rk}(G)}.$$

Now, for fixed R, it is enough to choose σ,ρ so that

$$\frac{\sigma(1-\sigma)\rho(1-\rho)}{(\sigma\rho+R)^2}$$

is maximized. This occurs when

$$\rho=\sigma=\sqrt{R(R+1)}-R,$$

in which case we get

$$\frac{\sigma(1-\sigma)\rho(1-\rho)}{(\sigma\rho+R)^2} = \frac{\Big(R-\sqrt{R(R+1)}\Big)^2}{4R^2}.$$

In the special case when $R=1$, we obtain the lower bound of Theorem 2. For a simple Chevalley group scheme over $\mathbb{Q}$, this gives the lower bound in Theorem 4.

§5. A reduction Lemma

Corollary 14 shows us that in order to give an upper bound on $\alpha_+(\Gamma)$ it suffices to bound $s_n(G(\mathbb{Z}/m\mathbb{Z}))$ when $m \leq n$. The goal of this section is to show that we can further assume that m is a product of different primes. To this end denote $\overline{m} = \prod p$ where p runs through all the primes dividing m.

We have an exact sequence

$$1 \to K \to G(\mathbb{Z}/m\mathbb{Z}) \xrightarrow{\pi} G(\mathbb{Z}/m\mathbb{Z}) \to 1$$

where K is a nilpotent group of rank at most $\dim G$ (see [ ]).

The following result will give us the desired reduction and will serve us later a few more times:

Lemma 4.1. *Let $1 \to K \to U \xrightarrow{\pi} L \to 1$ be an exact sequence of finite groups, where K is a solvable group of derived length ℓ and of rank at most r. Then the number of supplements to K in U (i.e., of subgroups H of U for which $\pi(H) = L$) is bounded by $|U|^{\ell f(r)}$ where f is some function depending on r.*

We postpone the proof of Proposition 4.1, deducing first the desired reduction:

Corollary 4.2. $s_n(G(\mathbb{Z}/m\mathbb{Z})) \leq m^{f'(\dim G)\log\log\ m} s_n(G(\mathbb{Z}/\overline{m}\mathbb{Z}))$ *where* $f'(\dim G)$ *depends only on* $\dim G$.

Proof. Let H be a subgroup of index at most n in $G(\mathbb{Z}/m\mathbb{Z})$ and denote $L = \pi(H) \leq G(\mathbb{Z}/\overline{m}\mathbb{Z})$. So L is of index at most n in $G(\mathbb{Z}/m\mathbb{Z})$. Let $U = \pi^{-1}(L)$, so every subgroup H of $G(\mathbb{Z}/m\mathbb{Z})$ with $\pi(H) = L$ is a subgroup of U. Given L (and hence also U) we have the exact sequence $1 \to K \to U \xrightarrow{\pi} L \to 1$ and by Lemma 4.1, the number of H in U with $\pi(H) = L$ is at most $|U|^{\ell f(r)}$ where ℓ is the derived length of K, $r \leq \dim G$ is the rank of K and $f(r) \leq f(\dim G)$ is independent of m. Now $|U| \leq m^{\dim G}$. By a result of Glasby ([ ]) if K is a finite solvable group, then its derived length is at most $c \log\log|K|$ for some absolute constant c. We can, therefore, deduce that $s_n(G(\mathbb{Z}/m\mathbb{Z})) \leq m^{c\dim G f(\dim G)\log\dim(G)} s_n(G(\mathbb{Z}/m\mathbb{Z}))$ which proves our claim.

We now turn to the proof of Lemma 4.1. The main part of the proof is the following.

Proposition 4.3. *Let $1 \to K \to U \xrightarrow{\pi} L \to 1$ be as in Proposition 4.1 and assume in addition that K is abelian, so $\ell = 1$. Then the number of H in U with $\pi(H) = L$ is bounded by $|U|^{f(r)}$.*

Proof of (4.3). Let $H \leq U$ be a subgroup with $\pi(H) = L$. Look at $H \cap K$. The number of possibilities for $H \cap K$ is at most $|K|^r \leq |U|^r$. As K is abelian, $H \cap K \lhd K$ and we can divide everything by $H \cap K$ to assume that $H \cap K = \{e\}$, i.e. H is a complement of K in U.

Assume first that U is a p-group. Let C be the centralizer of K in U, so $K \leq C$ as K is abelian. The quotient U/C acts faithfully on K and hence $U/C \hookrightarrow \mathrm{Aut}(K)$. Now, as $\mathrm{rank}(K) \leq r$, it is known that $\mathrm{rank}(\mathrm{Aut}(K)) \leq 4r^2$ (see [ ]). Thus, if $H \cap C$ is also given, the number of choices for H is at most $|U|^{4r^2}$ (as H is generated mod $H \cap C$ by at most $4r^2$ elements from U). We are left with the need to bound the number of possibilities for $H \cap C$.

Claim. $H \cap C$ is a complement to K in C. Indeed, if $c \in C$, $c = kh$ with $k \in K$ and $h \in H$ (as H is a complement to K in U). But $K \leq C$, so $k \in C$ and so $h \in H \cap C$.

We deduce that we need to bound the number of complements to K in C. This is equal to $|\mathrm{Hom}(C/K, K)| = |\mathrm{Hom}(A, K)|$ where A is the commutator quotient of C/K. Now, A and K are abelian groups and as such $|\mathrm{Hom}(A, K)| = |\mathrm{Hom}(K, A)|$ (see [ ]). Hence $|\mathrm{Hom}(A, K)| \leq |A|^r \leq |U|^r$ and Proposition 4.3 is proven under the assumption that U is a p-group.

Assume now that U is solvable: Let $p_1, \dots, p_s$ be the primes dividing $|U|$ and $U_1, \dots, U_s$ a Sylow system in U, i.e., each U_i is a p_i-Sylow subgroup of U and $U_iU_j = U_jU_i$ for every i and j. Note that K is abelian, hence has a unique p_i-Sylow K_i subgroup for each p_i, and $K_i \leq U_i$.

Now, if $H_1, \dots, H_s$ is a Sylow system for H, then there exists $U_1, \dots, U_s$ a Sylow system for U, with $H_i \leq U_i$ for every $i = 1, \dots, s$. As there are at most $|U|$ Sylow systems (since they are conjugate in U), we can fix one and assume $H_i \leq U_i$ $\forall i$. We claim that $U_i = K_iH_i$. Indeed, we assumed (as we could) that H is a complement to K in U, hence $K_i \cap H_i = \{e\}$ and clearly $|K_i||H_i| = |U_i|$ - so the claim is proved.

We can now use the case of p-groups proved before to deduce that the number of choices for H_i is at most $|U_i|^{f_1(r)}$ and so the number of choices for H is at most

$\prod\limits_{i=1}^{s} |U_i|^{f_1(r)} = |U|^{f_1(r)} \cdot |U|^{r+1}$.

We pass now to the general case (where we will use twice the (CFSG):

By a well known result of Aschbacher and Guralnick ([AG]) H is generated by some maximal solvable subgroup S of it plus one element $h \in H$. Now, SK is solvable and S is a complement of K in SK. As K is abelian, it is easy to see that SK is a maximal solvable subgroup of U. Now by [ ] a group U has at most $|U|^c$ maximal solvable subgroups for some absolute constant c. For each one of them we can apply the previous case (when we assume that U is solvable) to deduce, altogether, that the number of choices for H is bounded by $|U|^{f(r)}$ for a suitable $f(r)$. This ends the proof of Proposition 4.3. □

Lemma 4.1 is now deduced from 4.3 by induction on ℓ: So (4.3) is just the case $\ell = 1$. Assume it is true for $\ell - 1$. Divide U by K_1'-the commutator subgroup of K. This gives us the sequence $1 \to K/K' \to U/K' \xrightarrow{\tilde{\pi}} L \to 1$. By (4.3) the number of $\tilde{H}$ in U/K' with $\tilde{\pi}(\tilde{H}) = L$ is at most $(U/K')^{f(r)}$. Now, if H is a supplement to K in U (i.e., $\pi(H) = L$) look at $H \cdot K'$. When taken $\mod K'$, this is a supplement to K/K' in U/K' hence one of $(U/K')^{f(r)}$ possibilities. Now, H is a supplement to K' in $H \cdot K'$ and by the induction hypothesis there are at most $|H \cdot K'|^{(\ell-1)f(r)}$ such (since the derived length of K' is $\ell - 1$). So, altogether the number of possibilities for H is at most

$$|U|^{(\ell-1)f(r)} \cdot |U/K'|^{f(r)} \leq |U|^{\ell f(r)}$$

and Lemma 4.1 is proved. □

§6. From SL_2 to abelian groups

Corollary 1.4 shows us that in order to estimate $\alpha_+(G(\mathbb{Z}))$ one should concentrate on $s_n(G(\mathbb{Z}/m\mathbb{Z}))$ with $m \leq n$. Corollary 4.2 implies that we can further assume that m is a product of different primes. So let us now assume that $m = \prod\limits_{i=1}^{t} q_i$ where the q_i are different primes and so, $G(\mathbb{Z}/m\mathbb{Z}) \simeq \prod G(\mathbb{Z}/q_i\mathbb{Z})$ and $t \leq (1 + o(1))\frac{\log m}{\log\log m}$. We can further assume that we are counting only essential subgroups of $G(\mathbb{Z}/m\mathbb{Z})$, i.e., subgroups H which do not contain $G(\mathbb{Z}/q_i\mathbb{Z})$ for any $1 \leq i \leq t$, or equivalently the image of H under the projection to $G(\mathbb{Z}/q_i\mathbb{Z})$ is a proper subgroup. Thus H is contained in $\prod_{i=1}^{t} M_i$ where M_i is a maximal subgroup of $G(\mathbb{Z}/q_i\mathbb{Z})$.

Let us now specialize to the case $G = SL_2$, and let q be a prime.

Maximal subgroups of $SL_2(\mathbb{Z}/q\mathbb{Z})$ are conjugate to one of the following three subgroups (see [ ])

(1) $B = B_q$-the Borel subgroup of all upper triangular matrices in SL_2.

(2) $D = D_q$ -a dihedral subgroup of order $2(q+1)$ which is equal to $N(T_q)$ the normalizer of a non-split torus T_q. The group T_q is obtained as follows: Let $\mathbb{F}_{q^2}$ be the field of order $q^2, (\mathbb{F}_{q^2})^*$ acts on $\mathbb{F}_{q^2}$ by multiplication. The latter is a 2-dimensional vector space over $\mathbb{F}_q$. The elements of norm 1 in $(\mathbb{F}_{q^2})^*$ induce the subgroup T_q of $SL_2(\mathbb{F}_q)$.

(3) $A = A_q$-a subgroup of $SL_2(\mathbb{Z}/q\mathbb{Z})$ which is isomorphic to Alt(5) -the alternating group on 5 letters.

The number of conjugates of every subgroup is small, so it suffices to count only subgroups of $SL_2(\mathbb{Z}/m\mathbb{Z})$ whose projection to $SL_2(\mathbb{Z}/q\mathbb{Z})$ (for $q|m$) is either B, D, or A.

Let $S \subseteq \{q_1 \dots, q_t\}$ be the subset of the prime divisors of m for which the projection of H is in A_{q_i} and $\overline{S}$ the complement to S. Let $\overline{m} = \prod_{q \in \overline{S}} q$ and $\overline{H}$ the projection of H to $SL_2(\mathbb{Z}/\overline{m}\mathbb{Z})$. So $\overline{H}$ is a subgroup of index at most n in $SL_2(\mathbb{Z}/\overline{m}\mathbb{Z})$ and the kernel N from $H \to \overline{H}$ is inside a product of $|S|$ copies of Alt(5). As every subgroup of Alt(5) is generated by two elements, N is generated by at most $2\frac{\log m}{\log\log n} \leq 2\frac{\log n}{\log\log n}$ generators, which are all from a group of order $60^{|S|} \leq 60^t \leq 60^{\frac{\log n}{\log\log t}}$. We, therefore, conclude that given $\overline{H}$ the number of possibilities for H is at most $60^{2(\log n)^2/(\log\log n)^2}$ which is small w.r.t. $n^{\ell(n)}$.

We can, therefore, assume that $S = \phi$ and all the projections of H are either into groups of type B or D.

Now, if B_q is the Borel subgroup of $SL_2(\mathbb{Z}/q\mathbb{Z})$ it has a normal unipotent cyclic subgroup U_q of order q. Let now S be the subset of $\{q_1, \dots, q_t\}$ for which the projection is in B and $\overline{S}$-the complement. Then $H \leq \prod_{q \in S} B_q \times \prod_{q \in \overline{S}} D_q$. Let $\overline{H}$ be the projection of H to $\prod_{q \in S} B_q/U_q \times \prod_{q \in \overline{S}} D_q$. The kernel is a subgroup of the cyclic group

$U = \prod_{q \in S} U_q$. By Proposition 4.1 we know that given $\overline{H}$, there are only few possibilities for H. We are, therefore, led to counting subgroups in $L = \prod_{q \in S} B_q/U_q \times \prod_{q \in \overline{S}} D_q$. Let E now be the product $\prod_{q \in S} B_q/U_q \times \prod_{q \in \overline{S}} T_q$ and for a subgroup H of L we denote $H \cap E$ by $\overline{M}$.

Our next goal will be to show that given $\overline{H}$ in E, the number of possibilities for H is small. To this end we formulate first two easy lemmas, which will be used in the proof of Proposition 5.4 below. This Proposition will do for us the main reduction.

Lemma 5.1. *Let H be a subgroup of $U = U_1 \times U_2$, for $i = 1, 2$ denote $H_i = \pi_i(H)$ where π_i is the projection from U to U_i, and $H_i^0 = H \cap U_i$. Then:*

(i) H_i^0 is normal in H_i and $H_i/H_1^0 \simeq H_2/H_2^0$ with an isomorphism φ induced by the inclusion of $H/(H_1^0 \times H_2^0)$ as a subdirect product of H_1/H_1^0 and $H_2 H_2^0$,

(ii) H is determined by:

(a) H_i for $i = 1, 2$

(b) H_i^0 for $i = 1, 2$

(c) the isomorphism φ from H_1/H_1^0 to H_2/H_2^0.

Proof. See [ ]. □

Definition 5.2. *Let U be a group and V a subnormal subgroup of U. We say that V is co-poly-cyclic in U of co-length ℓ if there is a sequence $V = V_0 \lhd V_1 \lhd \ldots \lhd V_\ell = V$ such that V_i/V_{i-1} is cyclic for every $i = 1, \ldots, \ell$.*

Lemma 5.3. *Let U be a group and F a subgroup of U. The number of subnormal co-poly-cyclic subgroups V of U containing F and of co-length ℓ is at most $[U : F]^\ell$.*

Proof. For $\ell = 1$, V contains $[U, U]F$ and so it suffices to prove the lemma for the abelian group $\overline{U} = U/[U, U]F$ and $\overline{F} = \{e\}$. For an abelian group $\overline{U}$, the number of subgroups V with $\overline{U}/V$ cyclic is equal, by Pontrjagin duality to the number of cyclic subgroups. This is clearly bounded by $|\overline{U}| \leq [U : F]$. If $\ell > 1$, then by induction the number of possibilities for V_1 as in Definition 5.2 is bounded by $[U : F]^{\ell-1}$. Given V_1, the number of possibilities for V is at most $[V_1 : F] \leq [U : F]$ by the case $\ell = 1$. Thus, V has at most $[U : F]^\ell$ possibilities. □

Proposition 5.4. *Let $D = D_1 \times \ldots \times D_s$ where each D_i is a finite dihedral group with a cyclic subgroup T_i of index 2. Let $T = T_1 \times \ldots \times T_s$, so: $[D : T] = 2^s$. The number of subgroups H of D whose intersection with T is a given subgroup L of T is at most $|D|^8 2^{2s^2}$.*

Proof. Denote $F_i = \prod\limits_{j \geq i} D_i$. We want to count the number of subgroups H of D with $H \cap T = L$. Let $L_i = \mathrm{proj}_{F_i}(L)$ i.e., the projection of L to F_i, and $\tilde{L}_{i+1} = L_i \cap F_{i+1}$, so $\tilde{L}_{i+1} \subseteq L_{i+1}$. Let H_i be the projection of H to F_i. Given H, the sequence $(H_1 = H, H_2, \ldots, H_s)$ is determined and, of course, vice versa. We will actually prove that the number of possibilities for $(H_1, \ldots, H_s)$ is at most $|D|^8 2^{2s^2}$.

Assume now that H_{i+1} is given. What is the number of possibilities for H_i? Well, H_i is a subgroup of $F_i = D_i \times F_{i+1}$ containing L_i, whose projection to F_{i+1} is H_{i+1} and its intersection with F_{i+1}, which we will denote by X, contains $\tilde{L}_{i+1}$. By Lemma 5.2, H_i is determined by H_{i+1}, X, Y, Z and φ where Y is the projection of H_i to D_i, $Z = H_i \cap D_i$ and φ is an isomorphism from Y/Z to H_{i+1}/X. Now, every subgroup of the dihedral group is generated by two elements and so the number of possibilities for Y and Z is at most $|D_i|^2$ each, and the number of automorphisms of Y/Z is also at most $|D_i|^2$.

Let us now look at X : X is a normal subgroup of H_{i+1} with H_{i+1}/X isomorphic to Y/Z, so it is meta-cyclic. Moreover, X contains $\tilde{L}_{i+1}$. So by Lemma 5.3, the number of possibilities for X is at most $[H_{i+1} : \tilde{L}_{i+1}]^2$.

Now $[H_{i+1} : \tilde{L}_{i+1}] \leq [H_{i+1} : L_{i+1}][L_{i+1} : \tilde{L}_{i+1}]$. We know that $[H_{i+1} : L_{i+1}] = [\mathrm{proj}_{F_{i+1}} H] \leq |H : L| \leq 2^s$ and $[L_{i+1} : \tilde{L}_{i+1}] = [\mathrm{proj}_{F_{i+1}}(L_i) : F_{i+1} \cap L_i] \leq |D_i|$. So, $[H_{i+1} : \tilde{L}_{i+1}] \leq 2^s \cdot |D_i|$.

Altogether, given H_{i+1} (and L and hence also L_i's and $\tilde{L}_i$'s) the number of possibilities for H_i is at most $|D_i|^8 2^{2s}$. Arguing, now by induction we deduce that the number of possibilities for $(H_1, \ldots, H_s)$ is at most $|D|^8 2^{2s^2}$ as claimed. □

Let's now get back to SL_2: Proposition 5.4 implies, in the notations before Lemma 5.1, that when counting subgroups of $L = \prod\limits_{q \in S} B_q/U_q \times \prod\limits_{q \in \overline{S}} D_q$, we can count instead the subgroups of $E = \prod\limits_{q \in S} B_q/U_q \times \prod\limits_{q \in \overline{S}} T_q$ where T_q is the non-split tori in S_q (so T_q is a cyclic group of order $q+1$ while B_q/U_q is a cyclic group of order $q-1$).

A remark is needed here: Let H be a subgroup of index at most n in $SL_2(\mathbb{Z}/m\mathbb{Z})$ which is contained in $X = \prod\limits_{q\in S} B_q \times \prod\limits_{q\in \overline{S}} D_q$ and contains $Y = \prod\limits_{q\in S} U_q \times \prod\limits_{q\in \overline{S}} \{e\}$. By our analysis in this section, these are the groups which we have to count in order to determine $\alpha_+(SL_2(\mathbb{Z}))$. We proved that for counting them, it suffices for us to count subgroups of X_0/Y where $X_0 = \prod\limits_{q\in S} B_q \times \prod\limits_{q\in \overline{S}} T_q$. Note though that replacing H with its intersection with X_0, may enlarge the index of H in $SL_2(\mathbb{Z}/m\mathbb{Z})$. But the factor is at most $2^{\log m/\log\log m} = m^{1/\log\log m} \le n^{1/\log\log n}$. As $n \to \infty$ with n this factor is small with respect to n. By the remark made in §1, we can deduce that our original problem is now completely reduced to the following extremal problem on counting subgroups of finite abelian groups:

§7. Counting abelian groups

The aim of this section is to solve a somewhat unusual extremal problem concerning the number of subgroups of abelian groups. The result we prove is the crucial ingredient in obtaining a sharp upper bound for the number of congruence subgroups of $SL(2,\mathbb{Z})$.

For an abelian p-group G, we denote by $\Omega_i(G)$ the subgroup of elements of order dividing p^i. Then $\Omega_i(G)/\Omega_{i-1}(G)$ is an elementary abelian group of order say p^{λ_i} called the i-th *layer* of G. We call the sequence $\lambda_1 \ge \lambda_2 \ge \ldots \ge \lambda_r$ the *layer type* of G.

Denote by $\begin{bmatrix}\lambda\\ \nu\end{bmatrix}_p$ the p-binomial coefficient, that is, the number of ν dimensional subspaces of a λ-dimensional vector space over $\mathbb{Z}/p\mathbb{Z}$.

It is easy to see that the following holds.

Proposition 6.1.

(i) $p^{\nu(\lambda-\nu)} \le \begin{bmatrix}\lambda\\ \nu\end{bmatrix}_p \le p^{\nu}\cdot p^{\nu(\lambda-\nu)}$.

(ii) $\max \begin{bmatrix}\lambda\\ \nu\end{bmatrix}_p$ *is attained for* $\nu = [\frac{\lambda}{2}]$ *in which case* $\begin{bmatrix}\lambda\\ \nu\end{bmatrix}_p = p^{\frac{1}{4}\lambda^2+0(\lambda)}$ *holds as* $\lambda \to \infty$.

We need the following well-known formula (see[Bu]).

Proposition 6.2. *Let G be an abelian p-group of layer type $\lambda_1 \geq \lambda_2 \dots$. The number of subgroups H of layer type $\nu_1 \geq \nu_2 \dots$ is*

$$\prod_{i\geq 1} p^{\nu_{i+1}(\lambda_i-\nu_i)} \begin{bmatrix} \lambda_i - \nu_{i+1} \\ \nu_i - \nu_{i+1} \end{bmatrix}_p . \qquad \square$$

We need the following estimate.

Proposition 6.3.

$$\prod_{i\geq 1} p^{\nu_i(\lambda_i-\nu_i)} \leq \prod_{i\geq 1} p^{\nu_{i+1}(\lambda_i-\nu_i)} \begin{bmatrix} \lambda_i - \nu_{i+1} \\ \nu_i - \nu_{i+1} \end{bmatrix}_p \leq p^{\nu_1} \prod_{i\geq 1} p^{\nu_i(\lambda_i-\nu_i)}.$$

Proof. By Proposition 6.1 we have

$$\begin{aligned} &\prod_{i\geq 1} p^{\nu_{i+1}(\lambda_i-\nu_i)} \begin{bmatrix} \lambda_i - \nu_{i+1} \\ \nu_i - \nu_{i+1} \end{bmatrix}_p \\ &\leq \prod_{i\geq 1} p^{\nu_i(\lambda_i-\nu_i)} \cdot p^{(\nu_i-\nu_{i+1})((\lambda_i-\nu_{i+1})-(\nu_i-\nu_{i+1}))} \cdot p^{(\nu_i-\nu_{i+1})} \\ &= p^{\nu_1} \prod_{i\geq 1} p^{\nu_{i+1}(\lambda_i-\nu_i)} \cdot p^{(\nu_i-\nu_{i+1})(\lambda_i-\nu_i)} = p^{\nu_1} \prod_{i\geq 1} p^{\nu_i(\lambda_i-\nu_i)} \end{aligned}$$

The lower bound follows in a similar way. □

Corollary 6.4. *Let G be an abelian group of order p^α and layer type $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_r$. Then $|G|^{-1} \prod_{i\geq 1} p^{\lambda_i^2/4} \leq |Sub(G)| \leq |G|^2 \prod_{i\geq 1} p^{\lambda_i^2/4}$ holds.*

Proof. Considering subgroups H of layer type $[\frac{\lambda_1}{2}] \geq [\frac{\lambda_2}{2}] \geq \dots$ we obtain that $|\mathrm{Sub}(G)| \geq \prod_{i\geq 1} p^{[\frac{\lambda_i}{2}](\lambda_i-[\frac{\lambda_i}{2}])} \geq p^{-\nu} \prod_{i\geq 1} p^{\lambda_i^2/4}$ which implies the lower bound.

On the other hand, for any fixed layer type $\nu_1 \geq \nu_2 \geq \dots$ the number of subgroups H with this layer type is at most

$$p^{\nu_1} \prod_{i\geq 1} p^{\nu_i(\lambda_i-\nu_i)} \leq |G| \prod_{i\geq 1} p^{\lambda_i^2/4}.$$

The number of possible layer types $\nu_1 \geq \nu_2 \geq \dots$ of subgroups of G is bounded by the number of partitions of the number α hence it is at most $2^\alpha \leq |G|$. This implies our statement. □

Let us make an amusing remark which will not be needed in the rest of the paper.

If G is an abelian p-group of the form $G = C_{x_1} \times C_{x_2} \times \ldots \times C_{x_t}$ then it is known (see [LS]) that $|\mathrm{End}(G)| = \prod_{j,k} gcd(x_j, x_k)$. Noting that $\prod_{j,k} gcd(x_j, x_k) = \prod_{i \geq 1} p^{\lambda_i^2}$ we obtain that

$$|G|^{-1}|\mathrm{End}(G)|^{\frac{1}{4}} \leq |\mathrm{Sub}(G)| \leq |G|^2|\mathrm{End}(G)|^{\frac{1}{4}}.$$

These inequalities clearly extend to arbitrary finite abelian groups G.

Propositions 6.2 and 6.3 will be used in conjunction with the following simple (but somewhat technical) observations.

Let us call a pair of sequences of integers $\{\lambda_i\}, \{\nu_i\}$ *good* if $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_r \geq 1, \nu_1 \geq \nu_2 \geq \ldots \geq \nu_r \geq 1$ and $\lambda_i \geq \nu_i$ for $i = 1, 2, \ldots, r$.

Proposition 6.5. *Let α, t be fixed positive integers. Consider good pairs of sequences $\{\lambda_i\}, \{\nu_i\}$ such that $\sum_{i \geq 1}(\lambda_i + \nu_i) \leq \alpha$ and $\lambda_1 \leq t$.*

Under these assumptions the maximal value of the expression $\sum_{i \geq 1} \nu_i(\lambda_i - \nu_i)$ is also attained by a pair of sequences $\{\lambda_i\}, \{\nu_i\}$ such that

(i) $t = \lambda_1 = \lambda_2 = \ldots = \lambda_{r-1}$

(ii) for some $0 \leq b \leq r-1$ we have

$$\nu_1 = \nu_2 = \ldots = \nu_b = 1 + \nu_{b+1} = \ldots = 1 + \nu_{r-1}.$$

(iii) We have $\nu_i \geq [\frac{t}{3}]$ except possibly for $i = r$ if $\lambda_r < t$ in which case we have $[\frac{\lambda_r}{3}] \geq \nu_r \geq [\frac{\lambda_r}{3}]$.

Proof. Suppose that the maximum is attained for $\{\lambda_i\}, \{\nu_i\}$. Let j be the smallest index such that we have $t > \lambda_j \geq \lambda_{j+1} > 0$. Assume that $\lambda_{j+1} = \ldots = \lambda_{j+k}$ and $\lambda_{j+k} \geq \lambda_{j+k+1}$ or $j + k = r$. The condition $\nu_j \geq \nu_{j+1}$ implies that

$$\begin{aligned}&\nu_j((\lambda_j + 1) - \nu_j) + \nu_{j+1}((\lambda_{j+1} - 1) - \nu_{j+1})\\ &\geq \nu_j(\lambda_j - \nu_j) + \nu_{j+1}(\lambda_{j+1} - \nu_{j+1}).\end{aligned}$$

If $\lambda_{j+k} = \nu_{j+k}$ then we can clearly replace our sequences by another pair for which $\sum_{i \geq 1} \lambda_j$ is strictly smaller and $\sum_{i \geq 1} \nu_i(\lambda_i - \nu_i)$ is the same. Otherwise, replacing λ_j by $\lambda_j + 1$ and λ_{j+1} by $\lambda_{j+1} - 1$ we obtain a good pair of sequences for which $\{\lambda_i\}$ is lexicographically strictly greater and for which $\sum_{i\geq 1}\nu_i(\lambda_i-\nu_i)$ is at least as large (hence maximal).

It is clear that by repeating these two types of moves we eventually obtain a pair $\{\lambda_i\}, \{\nu_i\}$ satisfying (i).

Now set $\beta = \nu_1 + \nu_2 + \ldots + \nu_{r-1}$. Then

$$\sum_{i\geq 1}\nu_i(\lambda_i - \nu_i) = t\beta - (\nu_1^2 + \ldots + \nu_{r-1}^2) + \nu_r(\lambda_r - \nu_r).$$

It is clear that if the value of such an expression is maximal, then the difference of any two of the ν_j with $j \leq r-1$ is at most 1. This proves (ii).

Let us assume now that $\lambda_r < t$. Suppose that $\mu = \nu_{r-1} = \ldots = \nu_{b+1} < [\frac{t}{3}]$. This implies that $\mu \leq [\frac{t}{3}] - 1$ and hence $3\mu < t-2$.

We claim that $\mu(t-\mu) < (\mu+1)((t-1)-(\mu+1))$. This reduces to

$$\begin{aligned}
&\mu(t-\mu) < (\mu+1)(t-\mu) - 2(\mu+1)\\
&2\mu + 2 < t - \mu\\
&\text{and } 3\mu < t-2 \text{ which is true.}
\end{aligned}$$

By the claim replacing ν_j by ν_j+1 and λ_j by $t-1$ for $b+1 \leq j \leq r-1$ we obtain a good pair of sequences for which $\sum_{i\geq 1}\nu_i(\lambda_i-\nu_i)$ is strictly greater, a contradiction.

Hence we have $\nu_{r-1} \geq [\frac{t}{3}] \geq [\frac{\lambda_r}{3}]$. Using this a similar argument establishes that $\nu_r \geq [\frac{\lambda_r}{3}]$ (note that if $\nu_r < [\frac{\lambda_r}{3}]$ then replacing λ_r by $\lambda_r - 1$ and ν_r by ν_{r+1} we obtain a good pair of sequences.

Suppose now that $\nu_r > |\frac{\lambda_r}{3}|$. This implies $\nu_r \geq |\frac{\lambda_r}{3}| + 1$ and hence $3\nu_r > \lambda_r + 2$.

We claim that $\nu_r(\lambda_r - \nu_r) < (\nu_r - 1)((\lambda_r+1) - (\nu_r - 1))$. This reduces to

$$\begin{aligned}
&\nu_r(\lambda_r-\nu_r) \leq (\nu_r - 1)(\lambda_r - \nu_r) + 2(\nu_r - 1)\\
&\quad\lambda_r - \nu_r < 2\nu_r - 2\\
&\quad\lambda_r + 2 < 3\nu_r \qquad \text{which is true.}
\end{aligned}$$

By the claim replacing ν_r by $\nu_r - 1$ and λ_r by $\lambda_r + 1$ we obtain a pair of good sequences for which $\sum_{i\geq 1}\nu_i(\lambda_i-\nu_i)$ is strictly greater, a contradiction.

Hence we have $\nu_r \leq [\frac{\lambda r}{3}]$ as well.

Finally if $\lambda_r = t$ then the first part of the previous argument establishes $\nu_r \geq [\frac{t}{3}]$. □

The main result of this section is the following.

Theorem 6.6. *Let d be a fixed integer ≥ 1. Let n, r be positive integers. Let G be an abelian group of the form $G = C_{x_1} \times C_{x_2} \times \ldots \times C_{x_t}$ where at most d of the x_i can be the same. Suppose that $r|G| \leq n$ holds. Then the number of subgroups R of order $\leq r$ in G is at most $n^{(\gamma+o(1))\ell(n)}$ where $\gamma = \frac{3}{4} - \frac{1}{\sqrt{2}}$.*

Proof. We start the proof with several claims.

Claim 1. $t \leq (1 + o(1))\ell(n)$.

Proof. This follows from $t! \leq n$.

Claim 2. In proving the theorem, we may assume that $t \geq \gamma\ell(n)$.

Proof. For otherwise, every subgroup of G can be generated by $\gamma\ell(n)$ elements hence $|\mathrm{Sub}(G)| \leq |G|^{\gamma\ell(n)} \leq n^{\gamma\ell(n)}$.

Now let $a(n)$ be a monotone increasing function which goes to infinity sufficiently slowly. For example, we may set $a(n) = \log\log\log\log n$.

Let G_p denote the Sylow p-subgroup of G and let $\lambda_1^p \geq \lambda_2^p \geq \ldots$ denote the layer type of G_p. Altogether the layers of the G_p comprise the layers of G_j. We call such a layer essential if its dimension λ_1^p is at least $\frac{\ell(n)}{a(n)}$. Clearly the essential layers in G_p correspond to the layers of a certain subgroup E_p of G_p (which equals $\Omega_j(G_p)$ for some j). Let us call $E = \prod E_p$ the *essential subgroup* of G.

Claim 3. Given $E \cap R$ we have at most $n^{o(\ell(n)}$ (i.e., a small number of) choices for R.

Proof. It is clear from the definitions that every subgroup of the quotient group G/E can be generated by less than $\frac{\ell(n)}{a(n)}$ elements. Hence the same is true for $R/R \cap E$. This implies the claim.

By Claim 3, in proving the theorem, it is sufficient to consider subgroups R of E.

Let v denote the exponent of E. It is clear from the definitions that we have $v^{\ell(n)/a(n)} \leq n$, hence $v \leq (\log n)^{a(n)}$. Using well-known estimates of number theory [ ] we immediately obtain the following.

Claim 4. (i) the number z of different primes dividing v is at most $\frac{a(n)\log\log n}{\log\log\log n}$.

(ii) The total number of divisors of v is at most $\log n^{\frac{Ca(n)}{\log\log\log n}}$ for some constant $c > 0$.

Claim 5. $|G : E| \geq (\log n)^{(1+o(1))t}$.

Proof. Consider the subgroup $E_i = E \cap C_{x_i}$. It follows that E_i is the subgroup of elements of order dividing v in C_{x_i}. Set $e_i = |E_i|$ and $h_i = x_i/e_i$. It is easy to see that $E = \prod_{i\geq 1} E_i$, hence $|G : E| = \prod_{i\geq 1} h_i$.

By Claim 4(ii) for the number s of different values of the numbers e_i we have $s = (\log n)^{o(1)}$. We put the numbers x_i into s blocks of size $t_1, \dots, t_s$ according to the value of e_i.

By our condition on the x_i it follows that at most d of the numbers h_i corresponding to a given block of size say t_j are the same. Hence for some non-negative integers $t_{j1}, t_{j2}, \dots, t_{jd}$ with $\sum_{r\geq 1} t_{jk} = t_j$ the product of the h_i corresponding to this block is at least $\prod_{k\geq 1} t_{jk}!$. Therefore $|G : E| \geq \prod_{j,k} t_{jk}!$. Such a product is the smallest if the t_{jk} differ by at most 1 (given their number sd and their sum t) in which case we have $t_{jk} \geq [\frac{t}{sd}]$.

Hence $|G : E| \geq \prod_{j,k} t_{j,k}! \geq \prod_{j,k} (\frac{t_{jk}}{e})^{t_{jk}} \geq \prod_{j,k} (\frac{1}{e}[\frac{t}{sd}])^{t_{jk}} = (\frac{1}{e}[\frac{t}{sd}])^t$.

Since $sd = (\log n)^{o(1)}$ and by Claim 2 $t \geq \gamma \frac{\log n}{\log\log n}$ we obtain that $|G : E| \geq (\log n)^{(1+o(1))t}$ as required.

Let us now choose a group G and a number r as in the theorem for which the number of subgroups $R \leq E$ of order dividing r is maximal. To complete the proof it is clearly sufficient to show that this number is at most $n^{(\gamma+o(1))\ell(n)}$.

Denote the order of the corresponding essential subgroup E by f and the index $|G : E|$ by m.

Using Propositions 6.2 and 6.3 we see that apart from an $n^{o(\ell(n))}$ factor (which we ignore) the number of subgroups R as above is at most

(1) $\prod_p \prod_{i\geq 1} p^{\nu_i^p(\lambda_i^p - \nu_i^p)}$ for some ν_i^p, λ_i^p where $\prod_p \prod_{i\geq 1} p^{\lambda_i^p}$ divides f and $\prod_p \prod_{i\geq 1} p^{\nu_i^p}$ divides r for all p. Assuming that such an f and r are fixed together with the upper bound t for all the λ_i^p, let us estimate the value of the expression (1).

By Proposition 6.5 a maximal value of (1) is attained for a choice of the λ_i^p, ν_i^p such that for any given p there are at most 3 different pairs $(p^{\lambda_i^p}, p^{\nu_i^p})$ equal to say $(p^{\kappa_1^p}, p^{\mu_1^p}), (p^{\kappa_2^p}, p^{\mu_2^p})$ and $(p^{\kappa_3^p}, p^{\mu_3^p})$.

If there are say α_j^p pairs with $(p^{\lambda_i^p}, p^{\lambda_i^p})$ equal to $(p^{\kappa_j^p}, p^{\mu_j^p})$ then take β_j^p to be the largest number such that $2^{\beta_j^p} \leq p^{a_j^p}$.

Consider the expression

$$(2) \qquad \prod_p \prod_{j=1}^3 2^{\beta_j^p \mu_j^p (\kappa_j^p - \mu_j^p)}.$$

Its value is less than that of (1) but by definition their ratio is bounded by $(2^{3z})^{t^2}$ (where z is the number of primes dividing v). Hence this ratio is $\leq 8^{\ell(n)^2 \cdot \frac{a(n) \log\log n}{\log\log\log n}} \leq n^{2\ell(n) \frac{a(n)}{\log\log\log n}} = n^{o(\ell(n))}$. To prove our theorems it is sufficient to bound the value of (2) by $n^{(\gamma + o(1))\ell(n)}$.

It is clear that the value of the expression (2) is equal to the value of another expression

$$(3) \qquad \prod_{k \geq 1} 2^{v_k(\lambda_k - \nu_k)} \text{ which has } \sum_p \sum_{j=1}^3 \beta_j^p \text{ terms}$$

and for which $\prod_{r \geq 1} 2^{\lambda_k} \leq f$, $\prod_{k \geq 1} 2^{\nu_k} \leq \nu$.

By Proposition (3) such an expression attains its maximal value for some sequences $\{\lambda_k'\}, \{\nu_k'\}$ such that all but one of the λ_k', say λ_r' are equal to t and we have
$\nu_1' = \nu_2' = \ldots = \nu_b' = 1 + \nu_{b+1}' = \ldots = 1 + \nu_{r-1}'$ for some $b \leq r - 1$.

Consider now the expression

$$(4) \qquad \prod_{k \geq 1} 2^{\nu_k''(\lambda_k'' - \nu_k'')}$$

where

$$t = \lambda_1'' = \ldots = \lambda_{r-1}'' \quad (\lambda_r'' = 0)$$

and $\min(\nu_1', \nu_a') = \nu_1'' = \nu_2'' = \ldots = \nu_{r-1}''(v_r'' = 0)$.

It easily follows that the value of (3) is at most 2^{2t^2} times as large as the value of (4) and $2^{2t^2} = n^{o(\ell(n))}$. Hence it suffices to bound the value of (4) by $n^{(\gamma + o(n))\ell(n)}$.

To obtain our final estimate denote 2^a by y, $\log_t m$ by w (where $m = |G : E|$) and set $x = y \cdot w$.

For some constants between 0 and 1 we have $y = x^\rho$ and $\nu_1'' = \sigma t$. Then $w = x^{1-\rho} = y^{\frac{1-\rho}{\rho}}$.

We have $n \geq m.f.r \geq w^t \cdot y^t \cdot y^{\sigma t}$ hence $\log n \geq t \cdot \log y(1 + \sigma + \frac{1-\rho}{\rho})$.

By Claim 3 we have $w \geq (\log n)^{(1+o(1))}$ hence $(1 + o(1)) \log\log n \leq \log w = \frac{1-\rho}{\rho} \log y$. Therefore

$$\frac{(\log n)^2}{\log\log n} \geq \frac{t^2(\log y)^2(1 + \sigma + \frac{1-\rho}{\rho})^2}{(\frac{1-\rho}{\rho}\log y)}(1+o(1)) = (1+o(1))t^2 \log y(1+\sigma+\frac{1-\rho}{\rho})^2 \cdot (\frac{\rho}{1-\rho}).$$

The value of (4) is $2^{\sigma t(1-\sigma t)}$ which as we saw is an upper bound for the number of subgroups R (ignoring an $n^{o(\ell(n))}$ factor). Hence

$$\begin{aligned}
&\frac{\log(\text{number of subgroups } R)}{(\frac{(\log n)^2}{\log\log n})}\\
&\leq (1 + o(1)) \frac{t^2\sigma(1-\sigma)\log y}{t^2 \log y(1 + \sigma - \frac{1-\rho}{\rho})^2(\frac{\rho}{1-\rho})}\\
&= (1 + o(1)) \frac{\sigma(1-\sigma)(\frac{1-\rho}{\rho})}{(1 + \sigma + \frac{1-\rho}{\rho})^2} = (1 + o(1)) \frac{\sigma(1-\sigma)\rho(1-\rho)}{1 + \rho\sigma)^2}.
\end{aligned}$$

As observed in section 3 the maximum value of $\frac{\sigma(1-\sigma)\rho(1-\rho)}{1+\rho\sigma)^2}$ is γ. The proof of the theorem is complete. □

By using a similar but simpler argument, one can also show the following

Proposition 6.7. *Let G be an abelian group of order n of the form $G = C_{x_1} \times C_{x_2} \times \ldots \times C_{x_t}$ where $x_1 > x_2 > \ldots > x_t$. Then $|Sub(G)| \leq n^{(\frac{1}{16}+o(1))\ell(n)}$. This bound is attained if $C_{x_i} = t \cdot i$ for all i.*

Combining this result with an earlier remark, we obtain that $n^{(\frac{1}{16}+o(1))\ell(n)}$ is the maximal value of $\prod_{i,j} gcd(x_i, x_j)$ where the x_i are different numbers whose product is at most n.

Note that $|\mathrm{Sub}(G)|$ is essentially the number of subgroups R of order $[\sqrt{|G|}]$ (see [Bu]) for a strong version of this assertion. Hence Proposition 6.7 corresponds to the case $r \sim \sqrt{n}$ of Theorem 6.6.

§9. $SL_d(\mathbb{Z})$ as $d \to \infty$.

In this section we prove the upper bound of Theorem 3. To this end, we will use two facts on subgroups of "small" index in $SL_d(q)$:

Proposition 7.1. *Let F be a finite field of order q, $V = F^d$ and $G = SL(V) = SL_d(F)$.*

(a) Every proper subgroup of G is of index at least q^d.

(b) Let H be a subgroup of G of index smaller than $q^{\frac{2}{9}d^2}$. Then V has a sequence d of $F[H]$-submodules $\{0\} = V_0 < V_1 < \ldots < V_s = V$ such that:

(i) for every $i = 1, \ldots, s$, V_i/V_{i-1} is a simple H-module.

(ii) There exists $i \in \{1, \ldots, s\}$ such that $W = V_i/V_{i-1}$ has dimension at least $\frac{2}{3}d$ and the induction of H on W contains $SL(W)$.

Proof. (a) is well known - see [ ]. (b) is proved in [ ]. □

Note that (7.1)(b) implies that H can be put in a block form with one large block and the others are much smaller.

Now, our goal is to bound $s_n(SL_d(\mathbb{Z}/m\mathbb{Z}))$ where $m \leq n$ (see Proposition). By Corollary 4.2, we can assume m is a product of primes $m = \prod_{i=1}^{t} q_i$. Moreover, we count only the essential subgroup (see) so if H is a subgroup of $U = SL_d(\mathbb{Z}/m\mathbb{Z}) = \prod_{i=1}^{t} SL_d(\mathbb{Z}/q_i\mathbb{Z})$, we can assume the projection of H to each factor $SL_d(\mathbb{Z}/q_i\mathbb{Z})$ is a proper subgroup. Thus, by Proposition 7.1(a), the index of H in U is at least $\prod q_i^d$, so $n \geq m^d$, i.e., $m \leq n^{1/d}$.

Let H now be a subgroup of U and q one of the prime divisors of m. The projection of H to $SL_d(\mathbb{Z}/q\mathbb{Z})$ will be denoted by $H(q)$. We can bring $H(q)$ to a block form as in (7.1)(b(i)).

The number of block forms is small, so we can assume we are fixing the block form and count only H with $H(q)$ of a given block form. Moreover, we can use (again) Proposition 4.1 to assume that $H(q)$ is fully reducible. Indeed, the kernel of the projection to the blocks (i.e., replacing the representation by its semisimplification) is a unipotent group of bounded rank - so up to a small factor the number of the fully reducible groups is the significant one.

Now, if the index of $H(q)$ is smaller than $q^{\frac{2}{g}d^2}$ there is a large block of dimension at least $\frac{2}{3}d$, on which $H(q)$ acts as a subgroup containing $SL(W)$ (and inside $GL(W)$). If there is no such block, we say that all the blocks are small.

If there is a large block, then we can assume that H acts on it by exactly $SL(W)$ (since the number of subgroups between $SL(W)$ and $GL(W)$ is small). It is also clear that in this case $SL(W)$ is a direct factor of H and for counting purposes we can consider H as inside $GL_{d/3}(\mathbb{F}_q)$. We can, therefore, immediately consider the worst case situation and assume that all the blocks of $H(q)$ are small, and so the index of $H(q)$ is at least $q^{\frac{2}{g}d^2}$. Hence the index of H is at least $m^{\frac{2}{g}d^2}$ so $m^{\frac{2}{g}d^2} \leq n$.

We now count the number of all subgroups H of $U = SL_d(\mathbb{Z}/m\mathbb{Z})$. By Aschbacher-Guralnick ([ ]) H is generated by a solvable subgroup plus one element. It suffices, therefore, to count the solvable ones. Every such solvable subgroup is within a maximal solvable, and the number of the latter is small ($\leq |U|^C$ -see [ ]). So we can fix a maximal solvable subgroup S of U and count its subgroup. Such S is equal to $\prod\limits_{q|m} S(q)$ where $S(q)$ is the projection of S to $SL_d(\mathbb{Z}/m\mathbb{Z})$. Moreover, $S(q)$ is mapped onto its "semisimplification" $\overline{S(q)}$, which is a fully reducible solvable subgroup of $SL_d(\mathbb{Z}/m\mathbb{Z})$, with a kernel $N(q)$. So we get an exact sequence:

$$1 \to N = \prod_q N(q) \to S = \prod_q S(q) \to \overline{S} = \prod_q \overline{S(q)} \to 1.$$

Now, N is nilpotent of rank $\leq d^2$ ([ ]) so using again Proposition 4.1, it suffices to count the number of subgroups of $\overline{S}$. By the Palfy-Wolf Theorem ([ ]), $\overline{S(q)}$ is of order at most q^{3d} and so $|\overline{S}| \leq m^{3d}$.

Claim: $\mathrm{rank}(\overline{S}) \leq \frac{3}{2}d\frac{\log m}{\log\log m} + d^2 + 1$.

Proof. By ([ ]) it suffices to prove that if p divides $|\overline{S}|$ and if P is the p-Sylow-subgroup of $\overline{S}$ then $\mathrm{rank}(P) \leq \frac{3}{2}d\frac{\log m}{\log\log m} + d^2$.

To this end, note that the projection of P to $H(q)$ is fully reducible if $q \neq p$ and as such it is known that its rank is at most $\frac{3}{2}d$. While, if $q = p$, the rank is bounded by d^2. This proves the claim as m has at most $\frac{\log m}{\log\log m}$ prime divisors.

Putting all this together we deduce that (up to small factors) the number of subgroups of $SL_d(\mathbb{Z}/m\mathbb{Z})$ is bounded by

$$\begin{aligned}|\overline{S}|^{\mathrm{rank}(\overline{S})} &\leq m^{3d(\frac{3}{2}d\frac{\log m}{\log\log m}+d^2)}\\ &\leq n^{\frac{g}{2d^2}3d^{\frac{3}{2}d}(\frac{g}{2d^2}\frac{\log n}{\log\log n}+d^2)} \leq n^{(9^3/2^3+o(1))}\ell(n)\end{aligned}$$

and the upper bound of Theorem 3 is proven. □

References

[Lb] A. Lubotzky, D. Segal, *Subgroup growth.*

[Bu] L.M. Butler, *A unimodality result in the enumeration of subgroups of a finite abelian group*, Proc. AMS **101**, 771-775.