COUNTING CONGRUENCE SUBGROUPS

Dorian Goldfeld
Alexander Lubotzky
László Pyber

Abstract. Let Γ denote the modular group $SL(2,\mathbb{Z})$ and $C_n(\Gamma)$ the number of congruence subgproups of Γ of index at most n. We prove that $\lim_{n\to\infty} \frac{\log C_n(\Gamma)}{(\log n)^2/\log\log n} = \frac{3-2\sqrt{2}}{4}$. Some extensions of this result for other arithmetic groups are presented as well as a general conjecture.

§0. Introduction

Let k be an algebraic number field, $\mathcal{O}$ its ring of integers, S a finite set of valuations of k (containing all the archimedean ones), and $\mathcal{O}_S = \{x \in k \mid v(x) \geq 0,\ \forall v \notin S\}$. 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

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

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).

The first two authors research is supported in part by the NSF. The third author's Research is supported in part by OTKA T 034878. All three authors would like to thank Yale University for its hospitality.

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

In §2 below, we show that by appealing to a theorem of Linnik [Li1, Li2] on the least prime in an arithmetic progression, the proof can be made unconditional. Following [Lu] we define:

$$\alpha_+(\Gamma) = \overline{\lim}\, \frac{\log C_n(\Gamma)}{\lambda(n)}, \qquad \alpha_-(\Gamma) = \underline{\lim}\, \frac{\log C_n(\Gamma)}{\lambda(n)},$$

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]. It was conjectured by Rademacher that there are only finitely many congruence subgroups of $SL_2(\mathbb{Z})$ of genus zero. This counting problem has a long history. Petersson [Pe, 1974] proved that the number of all subgroups of index n and fixed genus goes to infinity exponentially as $n \to \infty$. Dennin [De, 1975] proved that there are only finitely many congruence subgroups of $SL_2(\mathbb{Z})$ of given fixed genus and solved Rademacher's conjecture. It does not seem possible, however, to accurately count all congruence subgroups of index at most n in $SL_2(\mathbb{Z})$ by using the theory of Riemann surfaces of fixed genus. Here we prove:

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

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 2 is based on the Bombieri-Vinogradov Theorem [Bo], [Da], [Vi], i.e., *the Riemann hypothesis on the average.* The upper bound, on the other hand, is proved by first 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 a number field with Galois group $\mathfrak{g} = \mathrm{Gal}(k/\mathbb{Q})$ and with ring of integers $\mathcal{O}$. Let S be a finite set of primes, and $\mathcal{O}_S$ as above. Assume GRH (generalized Riemann hypothesis) for k and all cyclotomic extensions $k(\zeta_\ell)$ with ℓ a rational prime and ζ_ℓ a primitive ℓ^{th} root of unity. Then*

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

The GRH is needed only for establishing the lower bound. It can be dropped in many cases by appealing to a theorem of Murty and Murty [MM] which generalizes the Bombieri–Vinogradov Theorem cited earlier.

Theorem 4. *Theorem 3 can be proved unconditionally for k if either*

(a) $\mathfrak{g} = \mathrm{Gal}(k/\mathbb{Q})$ *has an abelian subgroup of index at most 4 (this is true, for example, if k is an abelian extension);*

(b) $d = \deg[k : \mathbb{Q}] < 42$.

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 of dimension $d = \dim(G)$ and rank $\ell = rk(G)$. Let $\kappa = |\Phi^+|$ denote the number of positive roots in the root system of G. Letting $R = R(G) = \frac{d-\ell}{2\ell} = \frac{\kappa}{\ell}$, we see that $R = \frac{\ell+1}{2}, \quad (\text{resp.}\ell, \ell, \ell-1, 3, 6, 6, 9, 15)$ 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 3, and suppose that G is a simple Chevalley group scheme. Then*

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

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

(a) *Assuming GRH or the assumptions of Theorem 4;*

$$\alpha_-(G(\mathcal{O}_S)) \geq \frac{\left(\sqrt{R(R+1)} - R\right)^2}{4R^2} \sim \frac{1}{16R^2}.$$

(b) *There exists an absolute constant C such that*

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

Corollary 6. *There exists an absolute constant C such that for $d = 2, 3, \ldots$*

$$(1 - o(1))\frac{1}{4d^2} \ \leq \alpha_-(SL_d(\mathbb{Z})) \leq \alpha_+(SL_d(\mathbb{Z})) \leq C\frac{1}{d^2}.$$

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.

As a byproduct of the proof of Theorem 2 in §7 we obtain the following.

Corollary 7. *The subgroup growth type of $SL_d(\mathbb{Z}_p)$ is at least n^c where $c = (3-2\sqrt{2})d^2 - 2(2-\sqrt{2})$.*

The counting techniques in this paper can be applied to solve a novel extremal problem in multiplicative number theory involving the greatest common divisors of pairs $(p-1, p'-1)$ where p, p' are prime numbers. The solution of this problem does not appear amenable to the standard techniques used in analytic number theory.

Theorem 8. *For $n \to \infty$, let*

$$M(n) = \max\left\{ \prod_{p,p' \in \mathcal{P}} \gcd(p-1,\ p'-1) \;\middle|\; \mathcal{P} = \text{set of different primes where } \prod_{p\in\mathcal{P}} p \le n \right\}.$$

Then we have: $\underline{\lim}\, \frac{\log M(n)}{\lambda(n)} = \overline{\lim}\, \frac{\log M(n)}{\lambda(n)} = \frac{1}{4}$, *(where $\lambda(n) = (\log n)^2/\log\log n$).*

The paper is organized as follows.

In §1, we present some required preliminaries and notation.

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], [Da], [Vi].

In §3, we introduce the notion of a Bombieri set which is the crucial ingredient needed in the proof of the lower bounds. We then use it in §4 and §5 to prove the lower bounds of Theorems 2, 3, 4, and 5. We then turn to the proof of the upper bounds. In §6, we show how the counting problem of congruence subgroups in $SL_2(\mathbb{Z})$ can be completely reduced to an extremal counting problem of subgroups of finite abelian groups; the problem is actually, as one may expect, a number theoretic extremal problem - see §7 and §8 where this extremal problem is solved and the upper bounds of Theorems 2, 3, and 4 are then deduced in §9. In §10 we give the upper bound of Theorem 5. Finally, in §11 we prove Theorem 8.

§1. Preliminaries and notation

Throughout this paper we let

$$\ell(n) = \frac{\log n}{\log\log n}, \qquad \lambda(n) = \frac{(\log n)^2}{\log\log n}.$$

If f and g are functions of n, we will say that f *is small w.r.t.* g if $\lim_{n\to\infty} \frac{\log f(n)}{\log g(n)} = 0$. 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 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

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

The proof of (1.1) follows immediately form the inequalities:

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

Here, we have used 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 with respect to 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 and in its current form is proved in [LS, Proposition 6.1.1].

Lemma 1.1. *("Level versus index"). Let Γ be as before. Then there exists a constant $c > 0$ such that if H is a congruence subgroup of Γ of index at most n, it contains $\Gamma(m)$ for some $m \le cn$, where $m \in \mathbb{Z}$ and by $\Gamma(m)$ we mean $\Gamma(m\mathcal{O}_S)$.*

Corollary 1.2. *Let $\gamma_n(\Gamma) = \sum_{m=1}^{n} s_n(G(\mathcal{O}_S/m\mathcal{O}_S))$, where for a group H, $s_n(H)$ denotes the number of subgroups of H of index at most n. Then we have $\alpha_+(\Gamma) = \overline{\lim}\ \frac{\log \gamma_n(\Gamma)}{\lambda(n)}$ and $\alpha_-(\Gamma) = \underline{\lim}\ \frac{\log \gamma_n(\Gamma)}{\lambda(n)}$.*

Proof. By Lemma 1.1, $C_n(\Gamma) \le \gamma_{cn}(\Gamma)$ for some $c > 0$. It is also clear that $\gamma_n(\Gamma) \le n \cdot C_n(\Gamma)$. Since $c = o(\log n)$ (i.e., c is small w.r.t. n), Corollary 1.2 follows by arguments of the type we have given above. □

The number of elements in a finite set X is denoted by $\#X$ or $|X|$. The set of subgroups of a group G is denoted by $\mathrm{Sub}(G)$.

§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_+. Also, by restriction of scalars (and as we are not worried in Theorem 1 about the precise constants) we can assume $k = \mathbb{Q}$. The proof given here for Theorem 1 will follow the one given in [Lu] (and simplify it a bit). The main new ingredient is the use of a deep result of Linnik [Li1, Li2] 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 a careless choice of the parameters, the interval was *very short*, and the validity of the prime number theorem there is known only modulo GRH.

We introduce some notation which is 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 $p \leq x$ and $p \equiv a(\mod q)$. For $a = 1$, we set $\mathcal{P}(x; q) = \mathcal{P}(x; q, 1)$. We also define $\vartheta(x; q, a) = \sum_{p \in \mathcal{P}(x;q,a)} \log p$.

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, [Li1, Li2]**).** *There exist effectively computable constants $c_0, c_1 > 1$ such that if a and q are relatively prime integers, $q \geq 2$ and $x \geq q^{c_0}$, then*

$$\vartheta(x; q, a) \geq \frac{x}{c_1 q^2 \varphi(q)},$$

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

$$\sum_{p \in X} \log p \sim \frac{1}{c_1} x^{1-\theta(x)},$$

where $0 < \theta(x) \leq \frac{3}{c_0}$. We also define $P = \prod_{p \in X} p$. It follows that

$$\log P \sim \frac{x^{1-\theta(x)}}{c_1}.$$

Let now $\Gamma(P)$ be the corresponding 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, Theorem 6.1]) 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 $q|p-1$, $G(\mathbb{F}_p)$ contains a cyclic group of order q and $\Gamma/\Gamma(P)$ contains a subgroup isomorphic to $(\mathbb{Z}/q\mathbb{Z})^L$ where $L = \#X$.

It now follows from Theorem 2.1 and the choice of X, that

$$L \geq \frac{x^{1-\theta(x)}}{c_1 \log x}.$$

On the other hand, the abelian group $(\mathbb{Z}/q\mathbb{Z})^L$ has $q^{\frac{1}{4}L^2+O(L)}$ subgroups as $L\to\infty$ (cf. [LS, Prop. 15.2] or Proposition 7.1 below). 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:

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

$$\frac{\frac{1}{4}\left(\frac{x^{1-\theta(x)}}{c_1\log x}\right)^2\cdot\frac{1}{c_0}\log x}{\dim(G)^2\frac{1}{c_1^2}\left(x^{1-\theta(x)}\right)^2\Big/(1-\theta(x))\log x} \;=\; \frac{1-\theta(x)}{4c_0(\dim G)^2} \;\geq\; \frac{(1-\frac{3}{c_0})}{4c_0(\dim G)^2}.$$

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].

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

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

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).*

Remark. In all the applications in this paper, we do not really need q to be prime, though it makes the calculations somewhat easier. We could work with q being a "Bombieri number".

Lemma 3.3. *Fix $0 < \rho < \frac{1}{2}$. Then for x sufficiently large, there exists at least one Bombieri prime (relative to x) q 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 primes in the interval. In view of the trivial inequality, $\phi(q) = q - 1 < q$, it immediately follows that

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

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

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

It easily follows that

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

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 $n\le x$. Since $\left(\frac{1}{\log x}-\frac{1}{\log(x+1)}\right)=\frac{\log\left(1+\frac{1}{x}\right)}{\log x\log(x+1}\big)=O\left(\frac{1}{x(\log x)^2}\right)$, the second expression on the right side of the above equation is very small and can be ignored. It remains to estimate the sum $\sum\limits_{n=2}^{x}\vartheta(n;q,1)\frac{1}{n\cdot(\log n)^2}$. This sum can be broken into two parts, the first of which corresponds to $n\le\frac{x}{(\log x)^3}$, which is easily seen to be very small, so can be ignored. We estimate

$$\begin{aligned}\sum_{\frac{x}{(\log x)^3}\le n\le x}\vartheta(n;q,1)\frac{1}{n\cdot(\log n)^2} \;&=\; \sum_{\frac{x}{(\log x)^3}\le n\le x}\frac{n}{\phi(q)}\cdot\frac{1}{n(\log n)^2}\\ &\qquad+\;O\left(\sum_{\frac{x}{(\log x)^3}\le n\le x}\frac{x}{\phi(q)(\log x)^2}\cdot\frac{1}{n(\log n)^2}\right)\\ &=\;\sum_{\frac{x}{(\log x)^3}\le n\le x}\frac{1}{\phi(q)(\log n)^2}\;+\;O\left(\frac{x}{\phi(q)(\log x)^3}\right)\\ &\le\;\frac{3}{2}\frac{x}{\phi(q)(\log x)^2},\end{aligned}$$

which holds for x sufficiently large and where the constant $\frac{3}{2}$ is not optimal. Hence

$$\left|\sum_{p\in\mathcal{P}(x,q)} 1-\frac{\vartheta(x;q,1)}{\log x}\right| \;\le\; \frac{7}{4}\frac{x}{\phi(q)(\log x)^2},$$

say. Since $|\vartheta(x;q,1)-\frac{x}{\phi(q)}|\le\frac{x}{\phi(q)(\log x)^2}$, Lemma 3.4 immediately follows.

□

§4. Proof of the lower bound over $\mathbb{Q}$.

In this section we consider the case of $k = \mathbb{Q}$ and $\mathcal{O} = \mathbb{Z}$.

Fix a real number $0 < \rho_0 < \frac{1}{2}$. It follows from Lemma 3.3 that for $x \to \infty$ there exists a real number ρ which converges to ρ_0, and 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 a Bombieri set that

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

and from Lemma 3.4 that

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

Consider $\Gamma(P) = \ker(G(\mathbb{Z}) \to G(\mathbb{Z}/P\mathbb{Z}))$ 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

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

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

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

But

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

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,

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

where $B(P) \leq G(\mathbb{Z}/P\mathbb{Z})$ is:

$$B(P) = \prod_{p \in \mathcal{P}(x:q)} B(\mathbb{F}_p).$$

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 $\theta \leq t$. Let $0 \leq \sigma \leq 1$.

We will now use Proposition 7.1, a basic result on counting subspaces of finite vector spaces. It follows that $B(P)$ has at least

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

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 $x \to \infty$,

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

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

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

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

as $x\to\infty$.

We may rewrite

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

where

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

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

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

Actually, we choose ρ_0 to be $\sqrt{R(R+1)}-R$, then we can take ρ to be asymptotic to ρ_0 as x is going to infinity. Note that $\frac{\left(\sqrt{R(R+1)}-R\right)^2}{4R^2} < \frac{1}{16R^2}$ holds for all $R>0$. This follows from the easy inequality $\sqrt{R(R+1)}-R\le\frac{1}{2}$. It is also straightforward to see that $\sqrt{R(R+1)}-R$ converges to $\frac{1}{2}$ as $R\to\infty$ hence $\frac{\left(\sqrt{R(R+1)}-R\right)^2}{4R^2}\sim\frac{1}{16R^2}$.

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 5.

§5. Proof of the lower bound for a general number field.

To prove the lower bounds over a general number field we need an extension of the Bombieri–Vinogradov Theorem to these fields, as was obtained by Murty and Murty [MM].

Let us first fix some notations:

Let k be a finite Galois extension of degree d over $\mathbb{Q}$, $\mathfrak{g} = \mathrm{Gal}(k/\mathbb{Q})$, and $\mathcal{O}$ the ring of integers in k. For a rational prime q and $x \in \mathbb{R}$, we will denote by $\tilde{\mathcal{P}}_k(x,q)$ the set of rational primes $p \equiv 1(\mod q)$ where p splits completely in k and $p \leq x$. Let

$$\tilde{\pi}_k(x,q) = \#\tilde{\mathcal{P}}_k(x,q), \qquad \tilde{\nu}_k(x,q) = \sum_{p \in \tilde{\mathcal{P}}_k(x,q)} \log p,$$

and,

$$\tilde{E}_k(x,q) = \tilde{\nu}_k(x,q) - \frac{x}{d\phi(q)}.$$

We shall show that the following theorems follow from Murty and Murty [MM].

Theorem 5.1. *Let k be a fixed finite Galois extension of $\mathbb{Q}$. Assume GRH (generalized Riemann hypothesis) for k and all cyclotomic extensions $k(\zeta_\ell)$ with ℓ a rational prime and ζ_ℓ a primitive ℓ^{th} root of unity. Then for every $0 < \rho < \frac{1}{2}$ and $x \to \infty$, there exists a rational prime q such that*

(a) $\frac{x^\rho}{\log x} \leq q \leq x^\rho$

(b) $|\tilde{\pi}_k(x,q) - \frac{x}{d\phi(q)\log x}| \leq 3\left(\frac{x}{d\phi(q)(\log x)^2}\right)$

(c) $\max_{y \leq x} |\tilde{E}_k(y,q)| \leq \frac{x}{d\phi(q)(\log x)^2}$.

Theorem 5.2. *Theorem 5.1 can be proved unconditionally for k if either*

(a) $\mathfrak{g} = \mathrm{Gal}(k/\mathbb{Q})$ *has an abelian subgroup of index at most 4 (this is true, for example, if k is an abelian extension);*

(b) $d = \deg[k:\mathbb{Q}] < 42$.

Theorem 5.3. *Theorem 5.1 is valid unconditionally for every k with the additional assumption that $0 < \rho < \frac{1}{\eta}$, where η is the maximum of 2 and $d^* - 2$, and where d^* is the index of the largest possible abelian subgroup of $\mathfrak{g} = \mathrm{Gal}(k/\mathbb{Q})$. In particular, we may take $\eta = d^* - 2$ if $d^* \geq 4$ and $\eta = 2$ if $d^* \leq 4$.*

Proof of Theorems 5.1 - 5.3. For any $\epsilon > 0, A > 0$, under the assumptions of Theorem 5.1 or 5.2 (a), Murty and Murty [MM] prove the following Bombieri theorem:

$$\sum_{q \le x^{\frac{1}{2}-\epsilon}} \max_{(a,q)=1} \max_{y \le x} \left| \pi_C(y,q,a) - \frac{|C|}{|G|} \cdot \frac{1}{\phi(q)} \pi(y) \right| \ll \frac{x}{(\log x)^A}. \tag{5.1}$$

Here C denotes a conjugacy class in G, $\pi(y) = \sum_{p \le y} 1$,

$$\pi_C(x,q,a) = \sum_{\substack{p \le x \\ (p,k/\mathbb{Q})=C \\ p \equiv a \pmod q \\ p \text{ unramified in } k}} 1,$$

and $(p, k/\mathbb{Q})$ denotes the Artin symbol.

In fact, under the assumption of the GRH, equation (5.1) holds, but without assuming GRH they showed that (5.1) holds when the sum is over $q < x^{\frac{1}{\eta}-\varepsilon}$ where η is defined as follows: Let

$$d^* = \min_H \max_w [G:H] w(1) \tag{5.2}$$

The minimum here is over all subgroups H of $Gal(k/Q)$ satisfying:

(i) $H \cap C \neq \emptyset$, and

(ii) for every irreducible character w of H and any non-trivial Dirichlet character χ, the Artin L-series $L(s, w \otimes \chi)$ is entire. Then the maximum in (5.2) is over the irreducible characters of such H's.

Now

$$\eta = \begin{cases} d^* - 2 & \text{if } \ d^* \ge 4 \\ 2 & \text{if } \ d^* \le 4 \end{cases}$$

We need their result for the special case when C is the identity conjugacy class. In this case $\frac{|C|}{|G|} = \frac{1}{d}$ and $\pi_C(y,q,1) = \tilde{\pi}_k(y,q)$. So for proving Theorem 5.3 we can take for H an abelian subgroup of smallest index and then H satisfies assumption (i) and (ii). (Recall that abelian groups satisfy (AC) - Artin conjecture, i.e. $L(s, w \otimes \chi)$ are entire – see [CF]).

For Theorem 5.2(a), again take H to be the abelian subgroup of index at most 4. It satisfies (i) and (ii) and this time $\eta = 2$.

For Theorem 5.2(b): Going case by case over all possible numbers $d < 42$, one can deduce by elementary group theoretic arguments that every finite group $\mathfrak{g}$ of order $d < 42$, has an abelian subgroup of index at most 4, unless $d = 24$ and $\mathfrak{g}$ is isomorphic to the symmetric group S_4. But for this group, a highly non-trivial theorem of Tunnell [Tu] asserts that it

satisfies the Artin conjecture. Moreover, every irreducible character of S_4 is of degree at most 4. Thus for $\mathfrak{g} = S_4$ we have $d^* = 4$ and so $\eta = 2$.

The proofs of Theorems 5.1, 5.2 and 5.3 follow now in the same manner as in §3.

Using Theorems 5.1, 5.2, 5.3, we can now prove the lower bounds of Theorem 3 and 4 just as in §4. Note that every prime $p \in \tilde{\mathcal{P}}_k(x,q)$ gives d prime ideals $\pi_1, \ldots, \pi_d$ in $\mathcal{O}$ with $[\mathcal{O} : \pi_i] = p, \pi_i \cap \mathbb{Z} = p\mathbb{Z}$ and $\prod_{i=1}^{d} \pi_i = p\mathcal{O}$. Now let $\mathcal{P}_k(x,q)$ be the set of all prime ideals in $\mathcal{O}$ lying above the primes in $\tilde{\mathcal{P}}_k(x,q)$, and

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

Then

$$\log[\mathcal{O} : P] \sim \frac{x}{\phi(q)} \sim x^{1-\rho}, \qquad L := |\mathcal{P}_k(x,q)| \sim \frac{x}{\phi(q)\log x},$$

and

$$G(\mathcal{O}/P) = \prod_{\pi \in \mathcal{P}_k(x,q)} G(\mathcal{O}/\pi) \simeq \prod_{p \in \tilde{\mathcal{P}}_k(x,q)} G(\mathbb{Z}/p\mathbb{Z})^d.$$

We can now take for every rational prime $p \in \tilde{\mathcal{P}}_k(x,q)$, the Borel subgroup $B(p)$ as in §4 and define:

$$B(P) = \prod_{p \in \tilde{\mathcal{P}}_k(x,q)} B(p)^d.$$

Then $B(P)$ is mapped onto $(\mathbb{Z}/q\mathbb{Z})^{rk(G)\cdot L}$ and

$$\log\big[G(\mathcal{O}/P) : B(P)\big] \sim \frac{\dim(G) - rk(G)}{2} \log P.$$

Thus, by exactly the same computations as in §4, we can show that

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

The lower bounds of Theorems 3, 4, and 5 are now also proved. We now turn to the proof of the upper bounds.

§6. From SL_2 to abelian groups

In this section we show how to reduce the estimation of $\alpha_+(SL_2(\mathbb{Z}))$ to a problem on abelian groups.

Corollary 1.2 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 $m \leq n$. Our first goal 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

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

where K is a nilpotent group of rank at most $\dim G$. Here, the rank of a finite group G is defined to be the smallest integer r such that every subgroup of G is generated by r elements, (see [LS, Window 5, §2]).

Lemma 6.1. *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 $|U|^{3r^2+\ell r}$.*

Proof. See [LS, Corollary 1.3.5].

Corollary 6.2. *$s_n(G(\mathbb{Z}/m\mathbb{Z})) \le m^{f'(\dim G)\log\log\ m} s_n(G(\mathbb{Z}/\overline{m}\mathbb{Z}))$ 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 $L = \pi(H) \le G(\mathbb{Z}/\overline{m}\mathbb{Z})$. So L is of index at most n in $G(\mathbb{Z}/\overline{m}\mathbb{Z})$. 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 6.1, 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, $r \le \dim G$ is the rank of K and $f(r) \le f(\dim G)$ where f is some function depending on r and independent of m (say $f(r) = 3r^2 + r$). Now $|U| \le m^{\dim G}$ and K being nilpotent, is of derived length $O(\log\log|K|)$. We can, therefore, deduce that $s_n(G(\mathbb{Z}/m\mathbb{Z})) \le m^{c\dim G f(\dim G)(\log\log m + \log\dim G)} s_n(G(\mathbb{Z}/\overline{m}\mathbb{Z}))$ for some constant c which proves our claim.

Corollary 1.2 shows us that in order to estimate $\alpha_+(G(\mathbb{Z}))$ one should concentrate on $s_n(G(\mathbb{Z}/m\mathbb{Z}))$ with $m \le n$. Corollary 6.2 implies that we can further assume that m is a product of different primes. So let us now assume that $m = \prod_{i=1}^{t} q_i$ 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 $t \le (1+o(1))\frac{\log m}{\log\log m}$. We can further assume that we are counting only fully proper 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 $1 \le i \le t$, or equivalently the image of H under the projection to $G(\mathbb{Z}/q_i\mathbb{Z})$ is a proper subgroup (see [Lu]). 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 [La, Theorem 2.3])

(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 $q^2, \mathbb{F}_{q^2}^\times$ 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 $\mathbb{F}_{q^2}^\times$ 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 of order at most 120.

In cases B, D there is just one conjugacy class and in case A only boundedly many. Also, 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 inside 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 $|S|$ groups of type A. As every subgroup of $SL_2(\mathbb{Z}/q\mathbb{Z})$ is generated by two elements, H is generated by at most $2\frac{\log m}{\log\log m} \leq 2\frac{\log n}{\log\log n}$ generators. Set $k = [2\frac{\log n}{\log\log n}]$ and chose k generators for $\overline{H}$. By a lemma of Gaschütz (cf. [FJ, Lemma 15.30]) these k generators can be lifted up to give k generators for H. Each generator can be lifted up in at most $|N|$ ways and N is a group of order at most $120^{|S|} \leq 120^t \leq 120^{\frac{\log n}{\log\log n}}$. We, therefore, conclude that given $\overline{H}$ the number of possibilities for H is at most $120^{2(\log n)^2/(\log\log n)^2}$ 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, B_q , the Borel subgroup of $SL_2(\mathbb{Z}/q\mathbb{Z})$, 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 Lemma 6.1 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 $\overline{H}$.

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 6.6 below. This proposition will complete the main reduction.

Lemma 6.3. *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 $H_1/H_1^0 \simeq H_2/H_2^0$ 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 [Su, p 141]. □

Definition 6.4. *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 $V = V_0 \lhd V_1 \lhd \ldots \lhd V_\ell = U$ such that V_i/V_{i-1} is cyclic for every $i = 1, \ldots, \ell$.*

Lemma 6.5. *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 6.4 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 6.6. *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 $F_i = \prod_{j \geq i} D_i$. 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 $|H_{i+1} : L_{i+1}| = |\mathrm{proj}_{F_{i+1}}(H) : \mathrm{proj}_{F_{i+1}}(L)| \leq |H : L| \leq 2^s$ 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. $\square$

Let's now get back to SL_2: Proposition 6.6 implies, in the notations before Lemma 6.3, that when counting subgroups of $L = \prod_{q \in S} B_q/U_q \times \prod_{q \in \overline{S}} D_q$, we can count instead the subgroups of $E = \prod_{q \in S} B_q/U_q \times \prod_{q \in \overline{S}} T_q$ where T_q is the non-split tori in $SL_2(\mathbb{Z}/q\mathbb{Z})$ (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 $X = \prod_{q \in S} B_q \times \prod_{q \in \overline{S}} D_q$ and contains $Y = \prod_{q \in S} U_q \times v \prod_{q \in \overline{S}} \{e\}$. 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 $X_0 = \prod_{q \in S} B_q \times \prod_{q \in \overline{S}} T_q$. 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

$$2^{\log m/\log\log m} = m^{1/\log\log m} \leq n^{1/\log\log n}.$$

As $n \to \infty$, 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:

Let $\mathcal{P}_- = \{q_1, \ldots, q_t\}$ and $\mathcal{P}_+ = \{q'_1, \ldots, q'_{t'}\}$ be two sets of (different) primes and let $\mathcal{P} = \mathcal{P}_- \bigcup \mathcal{P}_+$. Denote

$$f(n) = \sup\{s_r(X) | X = \prod_{i=1}^{t} C_{q_i - 1} \times \prod_{i=1}^{t'} C_{q'_i + 1}\}$$

where the supremum is over all possible choices of $\mathcal{P}_-, \mathcal{P}_+$ and r such that $r \prod_{i=1}^{t} q_i \prod_{j=1}^{t'} q'_j \leq n$, (and C_m denotes the cyclic group of order m).

Corollary 6.7.

$$\alpha_+(SL_2(\mathbb{Z})) = \limsup \frac{\log f(n)}{\lambda(n)}.$$

§7. Counting subgroups of p-groups

In this section we first give some general estimates for the number of subgroups of finite abelian p-groups which will be needed in §8. As an application we obtain a lower bound for the subgroup growth of uniform pro-p-groups (see definitions later).

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 $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_r$ the *layer type* of G. It is clear that this sequence is decreasing.

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 $\mathbb{Z}/p\mathbb{Z}$.

The following holds (see [LS, Proposition 1.5.2]).

Proposition 7.1.

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

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

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

Proposition 7.2. *Let G be an abelian p-group of layer type $\lambda_1 \geq \lambda_2 \ldots \geq \lambda_r$. The number of subgroups H of layer type $\nu_1 \geq \nu_2 \ldots$ is*

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

(In the above expression we allow some of the ν_i to be 0.)

We need the following estimate.

Proposition 7.3.

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

Proof. By Proposition 7.1 we have

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

The lower bound follows in a similar way. □

Corollary 7.4. *Let G be an abelian group of order p^α and layer type $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_r$. Then $|G|^{-1} \prod_{i\geq 1} p^{\lambda_i^2/4} \leq |\mathrm{Sub}(G)| \leq |G|^2 \prod_{i\geq 1} p^{\lambda_i^2/4}$ holds.*

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

On the other hand, for any fixed layer type $\nu_1 \geq \nu_2 \geq \ldots$ 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 $\nu_1 \geq \nu_2 \geq \ldots$ 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 later.

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 $|\mathrm{End}(G)| = \prod_{j,k\geq 1} \gcd(x_j, x_k)$. Noting that $\prod_{j,k\geq 1} \gcd(x_j, x_k) = \prod_{i\geq 1} p^{\lambda_i^2}$ 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.

For the application of the above results to estimating the subgroup growth of $SL_d(\mathbb{Z}_p)$ we have to introduce additional notation .For a group G let G^k denote the subgroup generated by all k-th powers.For odd p a *powerful* p-group G is a p-group with the property that G/G^p is abelian.(In the rest of this section we will always assume that p is odd,the case $p = 2$ requires only slight modifications.) G is said to be *uniformly powerful* (*uniform* ,for short) if it is powerful and the indices $|G^{p^i} : G^{p^{i+1}}|$ do not depend on i as long as $i < e$, where p^e is the exponent of G .

Now let G be a uniform group of exponent p^e , where $e = 2i$, with d generators. Then G^{p^i} is a homocyclic abelian group of exponent p^i and d generators (i.e. it has layer type $d, d, \ldots, d$ with i terms) (see[DDMS]).

Consider subgroups H of G^{p^i} of layer type $\nu, \nu, \ldots, \nu$ (i terms).The number of such subgroups is at least $p^{i\nu(d-\nu)}$ by Proposition 7.3. . The index n of such a subgroup H in G is $p^{di+(d-\nu)i}$. Hence the number of index n subgroups in G is at least n^x where $x = \frac{\nu(d-\nu)}{2d-\nu}$. Substituting $\nu = [d(2-\sqrt{2})]$ we see that x can be as large as $(3-2\sqrt{2})d - (\sqrt{2}-1)$.

Let now U be a uniform pro-p-group of rank d , i.e. an inverse limit of d-generated finite uniform groups G.Then we see that for infinitely many n we have $s_n(G) \geq n^{(3-2\sqrt{2})d-(\sqrt{2}-1)}$.

Now $SL_d(\mathbb{Z}_p)$ is known to have a finite index uniform pro-p- subgroup of rank d^2-1 (see[DDMS]). This proves the following

Proposition 7.5. *$SL_d(\mathbb{Z}_p)$ has subgroup growth of type at least $n^{(3-2\sqrt{2})d^2-2(2-\sqrt{2})}$*

B. Klopsch proved [Kl] that if G is a residually finite virtually soluble minimax group of Hirsch length $h(G)$ then its subgroup growth is of type at least $n^{h(G)/7}$. By using the above argument one can improve this to $n^{(3-2\sqrt{2})h(G)-(\sqrt{2}-1)}$.

§8. Counting subgroups of 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 $SL(2,\mathbb{Z})$.

We will use Propositions 7.2 and 7.3 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 8.1. *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) *$t = \lambda_1 = \lambda_2 = \ldots = \lambda_{r-1}$ (i.e. only λ_r, the last term can be smaller than t).*

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

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

If $\lambda_r = t$ then we also have $\nu_1 = \nu_r$ or $\nu_1 = \nu_r + 1$.

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

$$\left\lceil \frac{\lambda_r}{3} \right\rceil \geq \nu_r \geq \left[\frac{\lambda_r}{3} \right].$$

Proof. Suppose that the maximum is attained for $\{\lambda_i\}, \{\nu_i\}$. Let j be the smallest index such that we have $t > \lambda_j \geq \lambda_{j+1} \geq 1$ (if there is no such j then (i) holds). Assume that $\lambda_{j+1} = \ldots = \lambda_{j+k}$ and $\lambda_{j+k} > \lambda_{j+k+1}$ or $j+k = r$. The condition $\nu_j \geq \nu_{j+k}$ implies that

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

If $\lambda_{j+k} = \nu_{j+k}$ then (by deleting some terms and renumbering the rest) we can clearly replace our sequences by another good 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+k} by $\lambda_{j+k} - 1$ 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 good 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. Part (ii) follows.

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

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

By the claim, replacing ν_j by $\nu_j + 1$ 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 $\nu_r + 1$ we obtain a good pair of sequences).

Suppose now that $\nu_r > \lceil \frac{\lambda_r}{3} \rceil$. This implies $\nu_r \geq \lceil \frac{\lambda_r}{3} \rceil + 1$ 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

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

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 $\nu_r \leq \lceil \frac{\lambda_r}{3} \rceil$ as well.

Finally if $\lambda_r = t$ then (setting $\mu = \nu_{b+1} = \ldots = \nu_r$) the first part of the previous argument establishes $\nu_r \geq [\frac{t}{3}]$. $\square$

Proposition 8.2. *Let $x_1, x_2, \ldots, x_t$ be positive integers such that at most d of the x_i can be equal. Then*

$$\prod_{i=1}^{t} x_i \ \geq \ \left(\frac{t}{ed}\right)^t$$

holds.

Proof. If say, x_1 is the largest among the x_i then $x_1 \geq \frac{t}{d}$. By induction we can assume that $\prod_{i=2}^{t} x_i \geq \left(\frac{t-1}{ed}\right)^{t-1}$ holds. Then

$$\prod_{i=1}^{t} x_i \geq \frac{t}{d}\left(\frac{t-1}{ed}\right)^{t-1} \geq e\left(\frac{t}{ed}\right)\left(\frac{t-1}{ed}\right)^{t-1} \geq e\left(\frac{t}{ed}\right)^t \left(\frac{t-1}{t}\right)^{t-1} =$$

$$= \left(\frac{t}{ed}\right)^t \frac{e}{\left(1 + \frac{1}{t-1}\right)^{t-1}} \geq \left(\frac{t}{ed}\right)^t, \qquad \square$$

as required.

The main result of this section is the following.

Theorem 8.3. *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 equal. 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 $\gamma = \frac{3-2\sqrt{2}}{4}$.*

Proof. We start the proof with several claims.

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

Proof. By Proposition 8.2 we have $\left(\frac{t}{ed}\right)^t \leq n$. This easily implies the claim.

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 $\lambda_1^p \geq \lambda_2^p \geq \dots$ 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_i^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 $\Omega_i(G_p)$ for the largest i such that $\lambda_i^p \geq \frac{\ell(n)}{a(n)}$). Let us call $E = \prod_p E_p$ 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 groups G_p/E_p and hence of G/E can be generated by less than $\frac{\ell(n)}{a(n)}$ elements. Therefore 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. Then E is the subgroup of elements of order dividing v in G. Now v is the product of the exponents of the E_p hence the product of the exponents of the essential layers of G. 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 [Ra] we immediately obtain the following.

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

(ii) The total number of divisors of v is at most $v^{\frac{c}{\log\log v}} \leq \log n^{\frac{ca(n)}{\log\log\log n}}$ 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 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 are equal. Hence altogether ds of the h_i can be equal. Using Proposition 8.2 we obtain that $|G:E| \geq \prod_{i\geq 1} h_i \geq \left(\frac{t}{eds}\right)^t$.

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 7.2 and 7.3 we see that apart from an $n^{o(\ell(n))}$ factor (which we ignore) the number of subgroups R as above is at most

$$\prod_p \prod_{i\geq 1} p^{\nu_i^p(\lambda_i^p - \nu_i^p)} \tag{8.1}$$

for some ν_i^p, λ_i^p where $\{\lambda_i^p\}, \{\nu_i^p\}$ is a good pair of sequences for every p, $\prod_p \prod_{i\geq 1} p^{\lambda_i^p}$ divides f and $\prod_p \prod_{i\geq 1} p^{\nu_i^p}$ divides r. Assuming that fr is fixed together with the upper bound t for all the λ_i^p, let us estimate the value of the expression (8.1).

By Proposition 8.1 a maximal value of an expression like (8.1) is attained for a choice of the λ_i^p, ν_i^p (for the sake of simplicity we use the same notation for the new sequences) such that for every p there are at most 3 different pairs $(p^{\lambda_i^p}, p^{\nu_i^p})$ equal to say

$$(p^t, p^{\mu^p+1}), \quad (p^t, p^{\mu^p}), \quad \text{and} \quad (p^{\tau^p}, p^{\mu_0^p})$$

where $\mu^p \geq \left[\frac{t}{3}\right]$, $\tau^p < t$ and $\lceil \frac{\tau^p}{3} \rceil \geq \mu_0^p \geq \left[\frac{\tau^p}{3}\right] = \mu_1^p$ for all p.

Exchange the pairs equal to the first type for pairs equal to (p^t, p^{μ^p}) and the pairs of the third type for pairs equal to $(p^{\tau^p}, p^{\mu_1^p})$. We obtain an expression like (8.1) (where for every p the $\{\lambda_i^p\}, \{\nu_i^p\}$ still forms a good pair of sequences) such that the ratio of the two expressions is at most

$$\prod_p \prod_{i\geq 1} p^{\lambda_i^p} \leq n \ .$$

If now there are say α^p pairs with $(p^{\lambda_i^p}, p^{\nu_i^p})$ equal to (p^t, p^{μ^p}) then take β^p to be the largest integer with $2^{\beta^p} \leq p^{\alpha^p}$ and set $\beta_1^p = \left[\log_2 p\right]$. (Note that for every p there is at most one pair of the form $(p^{\tau^p}, p^{\mu_1^p})$.)

Consider the expression

$$\prod_{p\geq 1} 2^{\beta^p \mu^p (t-\mu^p)} \ 2^{\beta_1^p \mu_1^p (\tau^p - \mu_1^p)}. \tag{8.2}$$

Its value may be less than that of (8.1) but in this case their ratio is bounded by $(2^{2z})^{t^2} n$ (where z is the number of primes dividing v). Hence this ratio is at most

$$2^{(2+o(1))\ell(n)^2 \ \frac{a(n)\log\log n}{\log\log\log n}} \leq n^{(2+o(1))\ell(n) \ \frac{a(n)}{\log\log\log n}} = n^{o(\ell(n))}.$$

To prove our theorem it is sufficient to bound the value of (8.2) by $n^{(\gamma+o(1))\ell(n)}$.

It is clear that the value of (8.2) is equal to the value of another expression

$$\prod_{k\geq 1} 2^{\nu_k(\lambda_k-\nu_k)} \tag{8.3}$$

which has $\sum_p(\beta^p+1)$ terms and for which $\prod_{k\geq 1} 2^{\lambda_k+\nu_k} \leq f\cdot r$. By the definition of (8.2) it is also clear that for this new pair of sequences $\{\lambda_k\},\{\nu_k\}$ we either have $\lambda_k = t$ and $\nu_k \geq \left[\frac{t}{3}\right]$ or $\lambda_k < t$ and $\nu_k = \left[\frac{\lambda_k}{3}\right]$. This ensures that if $\{\lambda_k\}$ is decreasing then $\{\nu_k\}$ is decreasing as well i.e. our sequences form a good pair.

By Proposition 8.1 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_{a+1} are equal to t and we have $\nu_1 = \nu_2 = \ldots = \nu_b = 1+\nu_{b+1} = \ldots = 1+\nu_a$ for some $b \leq a$.

Consider now the expression

$$\prod_{k\geq 1} 2^{\nu'_k(\lambda'_k-\nu'_k)} \tag{8.4}$$

where

$$t = \lambda'_1 = \ldots = \lambda'_a \quad (\lambda'_{a+1} = 0)$$

and $\nu_a = \nu'_1 = \nu'_2 = \ldots = \nu'_a \quad (\nu'_{a+1} = 0)$.

It easily follows that the value of (8.3) is at most 2^{2t^2} times as large as the value of (8.4) and $2^{2t^2} = n^{o(\ell(n))}$. Hence it suffices to bound the value of (8.4) by $n^{(\gamma+o(n))\ell(n)}$.

To obtain our final estimate denote 2^a by y, $m^{1/t}$ 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 5 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 (8.4) is $y^{\sigma t(t-\sigma t)}$ 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 §4 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 8.4. *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 $|\mathrm{Sub}(G)| \leq n^{(\frac{1}{16}+o(1))\ell(n)}$. This bound is attained if $x_i = t \cdot i$ for all i.*

Combining this result with an earlier remark, we obtain that $n^{(\frac{1}{4}+o(1))\ell(n)}$ 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 8.4 corresponds to the case $r \sim n^{1/3}$ of Theorem 8.3.

§9. End of proofs of Theorems 2, 3, and 4.

Theorem 2 is actually proved now: the lower bound was shown as a special case of $R = R(G) = 1$ in §4. For the upper bound, we have shown in Corollary 6.7 how $\alpha_+(SL_2(\mathbb{Z}))$ is equal to $\limsup \frac{\log f(n)}{\lambda(n)}$ (see Corollary 6.7 for the definition of $f(n)$). But Theorem 8.3 implies, in particular, that $f(n)$ is at most $n^{(\gamma+o(1))\ell(n)}$ where $\gamma = \frac{3-2\sqrt{2}}{4}$. This proves that $\alpha_+(SL_2(\mathbb{Z})) \leq \gamma$ and finishes the proof.

The proof of Theorem 3 is similar, but several remarks should be made: The lower bound was deduced in §5. For the upper bound, one should follow the reductions made in §6. The proof can be carried out in a similar way for $SL_2(\mathcal{O})$ instead of $SL_2(\mathbb{Z})$ but the following points require careful consideration.

1) One can pass to the case that m is an ideal which is a product of different primes π_i's in $\mathcal{O}$, but it is possible that $\mathcal{O}/\pi_i$ is isomorphic to $\mathcal{O}/\pi_j$. Still, each such isomorphism class of quotient fields can occur at most d times when $d = [k : \mathbb{Q}]$.

2) The maximal subgroups of $SL_2(\mathbb{F}_q)$ when $\mathbb{F}_q$ is a finite field of order q (q is a prime power, not necessarily a prime) are the same B, D and A as described in (1), (2), and (3) of §6.

The rest of the reduction can be carried out in a similar way to §6. The final outcome is not exactly as $f(n)$ at the end of §6, but can be reduced to a similar problem when $\tilde{f}(n)$ counts $s_r(X)$ when X is a product of abelian cyclic groups, with a bounded multiplicity. Theorem 8.3 covers also this case and gives a bound to $\tilde{f}(n)$ which is the same as for $f(n)$. Thus $\alpha_+(SL_2(\mathcal{O})) \leq \gamma = \frac{3-2\sqrt{2}}{4}$.

We finally mention the easy fact, that replacing $\mathcal{O}$ by $\mathcal{O}_S$ when S is a finite set of primes (see the introduction) does not change α_+ or α_-. To see this one can either use the fact that for every completion at a simple prime π of $\mathcal{O}$, $G(\mathcal{O}_\pi)$ has polynomial subgroup growth and then use the well known techniques of subgroup growth and the fact that

$$G(\hat{\mathcal{O}}) = G(\hat{\mathcal{O}}_S) \times \prod_{\pi \in S \setminus V_\infty} G(\mathcal{O}_\pi)$$

to deduce that $\alpha(G(\hat{\mathcal{O}})) = \alpha(G(\hat{\mathcal{O}}_S))$.

Another way to see it, is to observe that $G(\hat{\mathcal{O}}_S)$ is a quotient of $G(\hat{\mathcal{O}})$, and, hence, $\alpha_+(G(\mathcal{O})) \geq \alpha_+(G(\mathcal{O}_S))$. On the other hand, the proof of the lower bound for $\alpha(G(\mathcal{O}))$ clearly works for $G(\mathcal{O}_S)$. Theorem 3 is, therefore, now proved, as well as Theorem 4 (since we have not used the GRH for the upper bounds in Theorem 3).

§10. Chevalley groups.

In this section we prove the upper bound of Theorem 5. In view of the results of [Lu] it is sufficient to consider classical groups of large rank. For simplicity of notations we will treat the case $k = \mathbb{Q}$. The general case is similar with minor changes.

We first prove the result for $SL_d(\mathbb{Z})$. To this end, we will use two facts on subgroups of "small" index in $SL_d(q)$:

Proposition 10.1. *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-1} (unless $G = SL_2(9)$).*

(b) *Let H be a subgroup of G of index smaller than $q^{\frac{2}{9}d^2}$. Then V has a sequence of $F[H]$-submodules $\{0\} = V^0 < V^1 < \ldots < V^s = V$ such that:*

(i) *for every $j = 1, \ldots, s$, V^j/V^{j-1} is a simple H-module.*

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

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

Note that (10.1)(b) implies that H can be put in a block form with one large block and the others much smaller.

We will also need a simple number theoretic lemma. Let $\varepsilon > 0$ be a constant such that the product of the first ℓ primes is at least $e^{\varepsilon \ell \log \ell}$ for all $\ell \geq 2$.

Lemma 10.2. *Let $q_1, \ldots, q_t$ be different primes. Let $x_1, \ldots, x_t$ and d be natural numbers such that $x_i \leq d$. If $\prod_{i=1}^{t} q_i^{x_i} \leq m$, then*

$$\sum_{i=1}^{t} x_i \ \leq \ \frac{2}{\varepsilon}\ell(m) + d\sqrt{\log m}\ .$$

Proof. Let t_j $(j = 1, \ldots, d)$ denote the number of indices i for which $x_i \geq j$. Clearly $t = t_1 \geq t_2 \geq \cdots \geq t_d \geq 0$ and $\sum_{j=1}^{d} t_j = \sum_{i=1}^{t} x_i$. It follows that $\prod_{j=1}^{d} e^{\varepsilon t_j \log t_j} \leq \prod_{i=1}^{t} q_i^{x_i} \leq m$ hence $\varepsilon \sum_{j=1}^{d} t_j \log t_j \leq \log m$ holds.

Choose r such that $\log t_r \geq \frac{1}{2}\log\log m > \log t_{r+1}$. Then $\varepsilon \sum_{j=1}^{r} t_j \log t_j \leq \log m$ implies that $\sum_{j=1}^{r} t_j \leq \frac{2}{\varepsilon}\frac{\log m}{\log\log m}$. On the other hand for $j \geq r+1$ we have $t_{r+1} \leq \sqrt{\log m}$. Therefore $\sum_{i=1}^{t} x_i = \sum_{j=1}^{d} t_j \leq \frac{2}{\varepsilon}\ell(m) + d\sqrt{\log m}$ as required. $\square$

Now, our goal is to bound $s_n(SL_d(\mathbb{Z}/m\mathbb{Z}))$ where $m \leq n$ (see Corollary 1.2). By Corollary 6.2, we can assume that m is a product of different primes, $m = \prod_{i=1}^{t} q_i$. Moreover, we count only the fully proper subgroups (see §6) so if H is a subgroup of

$$G = SL_d(\mathbb{Z}/m\mathbb{Z}) = \prod_{i=1}^{t} SL_d(\mathbb{Z}/q_i\mathbb{Z}),$$

of index at most n, we can assume that the projection of H to each factor $SL_d(\mathbb{Z}/q_i\mathbb{Z})$ is a proper subgroup. Thus, by Proposition 10.1(a), the index of H in G is at least $\prod q_i^{d-1}$, so $n \geq m^{d-1}$, i.e., $m \leq n^{\frac{1}{d-1}}$.

Let H now be a subgroup of G and q_i one of the prime divisors of m. The projection of H to $SL_d(\mathbb{Z}/q_i\mathbb{Z})$ will be denoted by $H(q_i)$. We can bring $H(q_i)$ to a block form as in (10.1)(b(i)).

Now if the index of $H(q_i)$ is smaller than $q_i^{\frac{2}{9}d^2}$ there is a large block of dimension $d_i \geq \frac{2}{3}d$. We call the other blocks small. Set $x_i = d - d_i$ or $x_i = d$ if there is no large block. We see that the index of $H(q_i)$ in $SL_d(\mathbb{Z}/q_i\mathbb{Z})$ is at least $q_i^{\frac{2}{9}dx_i}$.

We obtain that $\prod_{i=1}^{t} q_i^{\frac{2}{9}dx_i} \leq n$, that is $\prod_{i=1}^{t} q_i^{x_i} \leq n^{\frac{9}{2d}}$. By Lemma 10.2 this implies

$$\sum_{i=1}^{t} x_i \leq \frac{9}{\varepsilon d}\ell(n) + \sqrt{5d\,\log n},$$

which is less than say $\frac{10}{\varepsilon d}\ell(n)$ for n large enough (and d fixed).

The number of block forms is small, so we can assume we are fixing the block form and count only H with $H(q_i)$ of a given block form. Then $H(q_i)$ is a subgroup of the parabolic subgroup $P(q_i)$ of $SL_d(\mathbb{Z}/q_i\mathbb{Z})$ corresponding to the block form. The number of choices for $P(q_i)$ of a given block form is at most $q_i^{d^2}$ (since GL is flag-transitive) hence the number of choices for $P = \prod P(q_i)$ is small. If $R(q_i)$ denotes the unipotent radical of $P(q_i)$ then it is clear that $H(q_i)R(q_i)/R(q_i)$ is a fully reducible group. The group $R = \prod R(q_i)$ is nilpotent of rank $\leq d^2$. Now H is a subgroup of P and using Lemma 6.1 (again) we see that it is sufficient to count the number of possibilities for the quotient group $\overline{H} = HR/R$ inside $\overline{P} = P/R$. Note that $\overline{P}$ acts faithfully on the direct sum of all the modules V_i^j/V_i^{j-1} in a natural way and $\overline{H}$ acts as a fully reducible subgroup of $\overline{P}$.

Now $\overline{H}$ contains the normal subgroup $N = \prod SL(W_i)$ (where the W_i are the large irreducible modules). By a theorem of Aschbacher-Guralnick [AG] (cf. [LS, Window 1, Theorem 15]), $\overline{H}$ is generated by a subgroup $S \geq N$ such that S/N is solvable, plus one element. Clearly it is sufficient to count the number of choices for S.

Now S acts faithfully on the direct sum of the $\overline{H}$-modules V_i^j/V_i^{j-1}. For every i and j we can choose a sequence of S-submodules of V_i^j/V_i^{j-1} such that the quotients between consecutive submodules are simple S-modules. The number of such choices is small.

Using again Lemma 6.1 we can replace S by a quotient $\overline{S}$ which acts as a fully reducible group on the direct sum of these simple S-modules (as we did before with H).

Note that $\overline{S}$ acts as a subgroup of $GL(W_i)$ containing $SL(W_i)$ on the large modules and acts like an irreducible solvable group on the other "small" modules.

The sum of the dimensions of the small modules over $\mathbb{F}_{q_i}$ is x_i.

We claim that $\overline{S}$ can be generated by $\frac{3}{2}\sum_{i=1}^{t} x_i \ +t$ elements together with $\overline{N} = \prod SL(W_i)$.

Indeed let U be one of the small modules of dimension say y and let K be the kernel of the action of $\overline{S}$ on U. By a result of Kovács-Robinson [KR] the fully reducible linear group $\overline{S}/K$ can be generated by $\frac{3}{2}y$ elements. K acts trivially on U and by Clifford's theorem fully reducibly on all the other $\overline{S}$-modules. Continuing in a similar way (stabilising small modules one by one) we eventually reach a subgroup K_0 of $\overline{S}$ where $\overline{N} \leq K_0 \leq \prod GL(W_i)$ such that $\overline{S}$ can be generated by K_0 and most $\frac{3}{2}\sum x_i$ elements. Since each $GL(W_i)/SL(W_i)$ is cyclic we obtain our claim.

For each small $\overline{S}$-module U we fix a maximal solvable subgroup M of $GL(U)$ containing the image of $\overline{S}$ in $GL(U)$. Since $\dim(U) \leq d$ by [Py, Lemma 3.4] the number of choices for M is at most $|U|^{cd}$ for some absolute constant c.

Altogether the number of choices for all the maximal solvable subgroups M is at most m^{cd} so we can fix them when we count the groups $\overline{S}$.

By the Pálfy-Wolf theorem (cf. [LS, Window 3, Theorem 6]) the order of each such M is at most $|U|^3$. Denote by D the direct product of the groups M (for small modules U) and of the $GL(W_i)$ (for large modules W_i).

As we saw above we can assume that $\overline{S}$ is a subgroup of D containing $\overline{N} = \prod SL(W_i)$. It follows that $|D/\overline{N}| \leq m \cdot \prod q_i^{3x_i} \leq n^{\frac{1}{d-1}} \cdot n^{\frac{27}{2d}}$. On the other hand by the above claim $\overline{S}/\overline{N}$ can be generated by $t + \frac{3}{2}\sum x_i$ elements which is less than $\frac{1}{d-1}\ell(n) + \frac{3}{2}\left(\frac{10}{\varepsilon d}\ell(n)\right)$ for large n.

Hence the number of choices for $\overline{S}$ is at most $n^{\frac{c}{d^2}\ell(n)}$ for some absolute constant c. This completes the proof of Theorem 5 for $SL_d(\mathbb{Z})$.

Remark. The above proof relies on the Classification of Finite Simple Groups (CFSG) via the Aschbacher-Guralnick theorem. However this can be avoided, since all the groups which appear in the proof are linear of bounded dimension and such groups have bounded index subgroups with "known" simple composition factors by a deep but CFSG-free result of Larsen-Pink [LP].

The other classical groups can be handled by essentially the same methods. In each case we need a version of Proposition 10.1. Appropriate bounds for the indices of proper subgroups appear in [KL]. Analogues of Proposition 10.1 (b) appear in [Lie].

Actually the symplectic groups are not considered there, hence it seems appropriate to sketch an argument in this case (along the lines of [Lie]).

For our purposes it is sufficient to consider $G = Sp_{2r}(q)$ for q odd and $r \geq 4$. By a result of Kantor [Ka] if I is an irreducible subgroup of G then its index is at least $q^{\frac{1}{2}r(r+1)}$.

Let $V = V(d,q)$ be the underlying symplectic space (where $d = 2r$). Then V has a standard basis $\{e_1, \ldots, e_r, f_1 \ldots, f_r\}$ with $(e_i, e_j) = (f_i, f_j) = 0$ and $(e_i, f_j) = \delta_{ij}$.

Let K be the stabilizer in $Sp_d(q)$ of a totally isotropic subspace of dimension a. Then

$$|Sp_d(q) : K| = \prod_{j=0}^{a-1}(q^{d-2j} - 1) \Big/ \prod_{j=0}^{a-1}(q^{a-j} - 1) \geq q^{a(d-\frac{3a}{2})} \geq q^{\frac{ad}{4}}.$$

Let H be a subgroup of $Sp_d(q)$ such that $|Sp_d(q) : H| \leq q^{\frac{d^2}{32}}$ and let V^1 be a maximal totally isotropic H-subspace of V. Then $\dim(V^1) \leq \frac{d}{8}$. Now $V^{1\perp}/V^1$ is a direct sum of minimal H-invariant non-degenerate subspaces $W^1, \ldots, W^s$. Assume, say, that W^1 has the largest dimension and let $V^2 > V^1$ be the subspace corresponding to W^1. Let us set $b = d - \dim(W^1)$. An easy calculation shows that the index of H is at least $q^{\frac{bd}{8}}$. Therefore we

have $b \le \frac{d}{4}$. Now, using the above result of Kantor [Ka] we see that H acts on $W = V^2/V^1$ as the full symplectic group $Sp_{d-b}(q)$.

We return to our proof of Theorem 5 in the symplectic case. We can assume that H is a fully proper subgroup of index n in $Sp_d(\mathbb{Z}/m\mathbb{Z})$ where $m \le n$ and m is a product of different primes $m = \prod_{i=1}^{t} q_i$.

Consider H as a subgroup of $SL_d(\mathbb{Z}/m\mathbb{Z})$. Now each projection $H(q_i)$ can be put in a block form such that the largest block has dimension $d_i \ge \frac{3}{4}d$ and on the corresponding large module W_i the group H acts as $Sp(W_i)$.

By the above discussion we have $\prod_{i=1}^{t} q_i^{\frac{1}{8}dx_i} \le n$, hence $t = O\big(\frac{1}{d}\ell(n)\big)$ as $n \to \infty$. Now we replace H by its fully reducible quotient $\overline{H}$ and set $N = \prod Sp(W_i)$. We can finish the proof by the same argument as in the case of $SL_d(\mathbb{Z})$. The remaining classical groups can be handled in essentially the same way.

§11. An extremal problem in elementary number theory.

The counting techniques in this paper can be applied to solve the following extremal problem in multiplicative number theory.

For $n \to \infty$, let

$$M_1(n) = \max\left\{ \prod_{1\le i,j\le t} \gcd(a_i, a_j) \;\middle|\; 0 < t, a_1 < a_2 < \ldots < a_t \in \mathbb{Z},\ \prod_{i=1}^{t} a_i \le n \right\},$$

$$M_2(n) = \max\left\{ \prod_{p,p'\in\mathcal{P}} \gcd(p-1,\ p'-1) \;\middle|\; \mathcal{P} = \text{set of different primes where } \prod_{p\in\mathcal{P}} p \le n \right\}.$$

We shall prove the following theorem which can be considered as a baby version of Theorem 2 (compare also to Theorem 8.3). Note that Theorem 11.1 immediately implies Theorem 8.

Theorem 11.1. *Let* $\lambda(n) = \frac{(\log n)^2}{\log\log n}$. *Then*

$$\underline{\lim}\ \frac{\log M_1(n)}{\lambda(n)} = \overline{\lim}\ \frac{\log M_2(n)}{\lambda(n)} = \frac{1}{4}.$$

Proof. Recall that if $a_1, a_2, \ldots, a_t \in \mathbb{Z}$ and $G = C_{a_1} \times C_{a_2} \times \cdots \times C_{a_t}$ is a direct product of cyclic groups then by §7,

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

and

$$|\mathrm{End}(G)| \;=\; \prod_{1\le i,j\le t} \gcd(a_i,a_j).$$

Proposition 8.4 implies that

$$\overline{\lim}\,\frac{\log M_1(n)}{\lambda(n)} \;\le\; \frac{1}{4}.$$

It is clear that $M_2(n)\le M_1(n)$, so to finish the proof it is enough to obtain a lower bound for $M_2(n)$.

Now, for $x\to\infty$ and $\frac{x^\rho}{\log x}\le q\le x^\rho$ (with $0<\rho<\frac{1}{2}$) choose

$$\mathcal{P}=\mathcal{P}(x,q)=\big\{p\le x \;\big|\; p\equiv 1 \pmod q\big\},$$

to be a Bombieri set relative to x where q is a prime number (Bombieri prime). By Lemma 3.4 we have the asymptotic relation $\#\mathcal{P}(x,q)\sim\frac{x}{\phi(q)\log x}$. In order to satisfy the condition $\prod\limits_{p\in\mathcal{P}} p \;\le n$, we choose $x\sim q\log n$. Without loss of generality, we may choose $q=x^\rho$ for some $0<\rho<\frac{1}{2}$. It follows that

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

Consequently

$$\prod_{p,p'\in\mathcal{P}}\gcd(p-1,\;p'-1)\;\ge\;q^{(\#\mathcal{P})^2}\;\ge\;(x^\rho)^{\frac{(1-\rho)^2(\log n)^2}{(\log\log n)^2}}\sim e^{\frac{\rho(1-\rho)(\log n)^2}{\log\log n}}.$$

Let now ρ go to $\frac{1}{2}$ and the theorem is proved. □

References

[AG] M. Aschbacher, R.M.Guralnick, *Solvable generation of groups and Sylow subgroups of the lower central series*, J. Algebra **77** (1982), 189-201.

[Bo] E. Bombieri, *On the large sieve*, Mathematika **12** (1965), 201-225.

[Bu] L.M. Butler, *A unimodality result in the enumeration of subgroups of a finite abelian group*, Proc. Amer. Math. Soc. **101** (1987), 771-775.

[CF] J.W.S. Cassels, A. Fröhlich, *Algebraic Number Theory, Thompson Book Company*, 1967, pp. 218-230.

[Da] H. Davenport, *Multiplicative Number Theory , Springer-Verlag, GTM* **74**, 1980.

[De] J.B. Dennin, *The genus of subfields of $K(n)$*, Proc. Amer. Math. Soc. **51** (1975), 282-288.

[DDMS] J.D. Dixon M.P.F. du Sautoy, A. Mann, D. Segal, *Analytic Pro-p-Groups, Cambridge University Press London Math. Soc. Lecture Note Series* **157**, 1991.

[FJ] M.D. Fried, M. Jarden, *Field Arithmetic, Springer-Verlag*, 1986.

[Ka] W.M. Kantor, *Permutation representations of the finite classical groups of small degree or rank*, J. Algebra **60** (1979), 158-168.

[KL] P.B. Kleidman, M.W. Liebeck, *The subgroup structure of the finite classical groups, Cambridge University Press, London Math. Soc. Lect. Note Series* **129**, 1990.

[Kl] B. Klopsch, *Linear bounds for the degree of subgroup growth in terms of the Hirsch length*, Bull. London. Math. Soc. **32** (2000), 403-408.

[KR] L. G. Kovács, G.R. Robinson, *Generating finite completely reducible linear groups*, Proc. Amer. Math. Soc. **112** (1991), 357-364.

[La] S. Lang, *Introduction to Modular Forms, Springer-Verlag*, 1976.

[Lan] E. Landau, *Primzahlen, Chelsea Publishing Company*, 1953.

[Lie] M.W.Liebeck, *On graphs whose full automorphism group is an alternating group or a finite classical group*, Proc. London Math. Soc. (3) **47** (1983), 337-362.

[Li1] U.V. Linnik, *On the least prime in an arithmetic progression I. The basic theorem*, Rec. Math. [Mat. Sbornik] N.S. **15(57)** (1944), 139-178.

[Li2] U.V. Linnik, *On the least prime in an arithmetic progression. II. The Deuring-Heilbronn phenomenon*, Rec. Math. [Mat. Sbornik] N.S. **15(57)** (1944), 347-368.

[LP] M. Larsen, R. Pink, *Finite subgroups of algebraic groups*, J. Amer. Math. Soc. to appear.

[LS] A. Lubotzky, D. Segal, *Subgroup growth, Progress in Mathematics, Birkhauser*, 2003.

[Lu] A. Lubotzky, *Subgroup growth and congruence subgroups*, Invent. Math. **119** (1995), 267-295.

[MM] M. Ram Murty, V. Kumar Murty, *A variant of the Bombieri-Vinogradov theorem*, Canadian Math. Soc. Conference Proceedings **7** (1987), 243-272.

[Pe] H. Petersson, *Konstruktionsprinzipien für Untergruppen der Modulgruppe mit einer oder zwei Spitzenklassen*, J. Reine Angew. Math. (1974), 94-109.

[PR] V. Platonov, A. Rapinchuk, *Algebraic Groups and Number Theory, Academic Press*, 1991.

[Py] L. Pyber, *Enumerating finite groups of given order*, Annals of Math. **137** (1993), 203-220.

[Ra] S. Ramanujan, *Highly composite numbers*, Proc. London Math. Soc. (2) **XIV** (1915), 347–409.

[Su] M.Suzuki, *Group Theory 1 ,Springer-Verlag, Grundlehren Math. Viss.* **247**, 1982.

[Tu] J. Tunnell, *Artin's conjecture for representations of octahedral type*, Bull. Amer. Math. Soc. (N.S.) **5** (1981), 173-175.

[Vi] A.I. Vinogradov, *On the density conjecture for Dirichlet L-series*, Izv. Akad. Nauk SSSR Ser. Mat. **29** (1965), 903-934.

D. Goldfeld, Department of Mathematics, Columbia University, NY, NY 10027, USA
E-mail address: goldfeld@columbia.edu

A. Lubotzky, Institute of Mathematics, Hebrew University, Jerusalem 91904, Israel
E-mail address: alexlub@math.huji.ac.il

L. Pyber, A. Renyi Institute of Mathematics, Hungarian Academy of Sciences, P.O. Box 127, H-1364 Budapest ,Hungary
E-mail address: pyber@renyi.hu