JOURNAL OF NUMBER THEORY **18**, 178–191 (1984)

On the Number of Groups of a Given Order

M. RAM MURTY*

Department of Mathematics, McGill University, 805 Sherbrooke St. West, Montreal, Quebec H3A 2K6, Canada

AND

V. KUMAR MURTY

Department of Mathematics, Harvard University, Cambridge, Massachusetts 02138

Communicated by H. Stark

Received February 17, 1981

Letting $G(n)$ denote the number of nonisomorphic groups of order n, it is shown that for square-free n, $G(n) \leqslant \varphi(n)$ and $G(n) \leqslant (\log n)^c$ on a set of positive density. Letting $F_k(x)$ denote the number of $n \leqslant x$ for which $G(n) = k$, it is shown that $F_2(x) = O(x(\log_4 x)/(\log_3 x)^2)$, where $\log_r x$ denotes the r-fold iterated logarithm.

1. INTRODUCTION

Let $G(n)$ denote the number of nonisomorphic groups of order n. Gallagher [2] has shown that

$$\log G(n) \ll n^{2/3}(\log n)^2.$$

The problem of obtaining sharp estimates, in general, seems to be very difficult. Sims [5] has conjectured that

$$\log G(n) \ll (\log n)^3, \tag{1.1}$$

and has shown that (1.1) is true if we restrict our attention to solvable groups of order n. We shall show that for square-free n,

$$G(n) \leqslant \varphi(n),$$

* Supported in part by NSF Grant MCS 77-18723 A03

0022-314X/84 $3.00

where φ is Euler's totient function, using the well-known fact that groups of square-free order are supersolvable. Moreover, we show that for n belonging to a set of positive density,

$$G(n) \leqslant (\log n)^c,$$

is true.

The distribution of the values of $G(n)$ is a more intricate question. Let $F_k(x)$ denote the number of $n \leqslant x$ for which $G(n) = k$. Erdös [1] showed that

$$F_1(x) = \frac{(1 + o(1))\, xe^{-\gamma}}{\log_3 x},$$

where γ is Euler's constant and $\log_r x = \log(\log_{r-1} x)$, $\log_1 x = \log x$. We shall show that

$$F_2(x) = O\left(\frac{x \log_4 x}{(\log_3 x)^2}\right).$$

2. Groups of Square-Free Order

We now derive a bound for $G(n)$, when n is square-free. We refer to a group and its isomorphism class interchangeably.

THEOREM 2.1. *Let n be square-free. Then,*

$$G(n) \leqslant \prod_{p \mid n} (p - 1, n),$$

where (a, b) denotes the greatest common divisor of a and b.

Proof. Let q be the largest prime divisor of n, and G a group of order n. Since G is supersolvable, we have by [4, 7.2.19] that the q-Sylow subgroup S (say) is normal and G is a semidirect product of a group H of order (n/q) and S. Thus, an upper bound for $G(n)$, is obtained by counting the number of homomorphisms $\theta: H \to \mathrm{Aut}(S)$. For each prime $p \mid n$, choose a p-Sylow subgroup of H and let x_p be a generator of this p-Sylow subgroup. As H is generated by these Sylow subgroups, θ is determined by the $\theta(x_p)$. Since $\mathrm{Aut}(S)$ is cyclic of order $(q - 1)$, and $\theta(x_p)^p = 1$, the number of solutions for $\theta(x_p)$ is $(q - 1, p)$. Therefore, there are at most

$$\prod_{p \mid (n/q)} (q - 1, p) = \prod_{p \mid n} (q - 1, p)$$

possibilities for θ. As there are $G(n/q)$, possibilities for H, we obtain

$$G(n) \leqslant G\left(\frac{n}{q}\right) \prod_{p \mid n} (q-1, p) = G\left(\frac{n}{q}\right) (q-1, n)$$

and the result of the theorem follows by induction.

COROLLARY 2.2. *For square-free n,*

$$G(n) \leqslant \varphi(n).$$

This theorem allows us to deduce that there are very "few" groups of square-free order. More precisely, Higman [3] and Sims [5] have shown that,

$$G(2^n) \geqslant 2^{cn^3},$$

for some constant $c > 0$. Therefore, if $t = [\log x/\log 2]$, then,

$$\sum_{n \leqslant x} G(n) \geqslant G(2^t) \geqslant 2^{ct^3} \geqslant x^{c_1(\log x)^2},$$

for some constant $c_1 > 0$. Now by Corollary 2.2,

$$\sum_{n \leqslant x} \mu^2(n)\, G(n) \leqslant \sum_{n \leqslant x} \varphi(n) = O(x^2),$$

and so we deduce

COROLLARY 2.3. *Groups of square-free order are scarce.*

Let us write,

$$f(n) = \prod_{p \mid n} p^{v_p(n)},$$

where $v_p(n)$ is the number of primes $q \mid n$, with $q \equiv 1 \pmod{p}$. Then, for n square-free,

$$f(n) = \prod_{p \mid n} (p-1, n).$$

In the next sections, we shall determine the average orders of $\log f(n)$ and $\log^2 f(n)$. This will allow us to deduce that for n belonging to a set of positive density,

$$G(n) \ll (\log n)^2.$$

3. The Average Order of $\log f(n)$

We need some preliminary lemmas.

LEMMA 3.1. *Let p be a prime. Then*

$$\sum_{\substack{q<z \\ q\equiv 1(p)}}' \frac{1}{q} = \frac{\log\log z}{p-1} + O(1),$$

where the constant implied is absolute. If furthermore, $p < (\log z)^c$ (where c is an arbitrary constant) then

$$\sum_{\substack{q<z \\ q\equiv 1(p)}}' \frac{1}{q} = \frac{\log\log z}{p-1} + O\left(\frac{\log p}{p}\right).$$

Proof. The result follows easily from the Siegel–Walfisz theorem and partial summation, and the Brun–Titchmarsh theorem.

LEMMA 3.2. *Let p be a prime. Then,*

$$\sum_{\substack{q<z \\ q\equiv 1(p)}}' \frac{1}{q} \leqslant \frac{C\log\log z}{p-1} + O\left(\frac{\log p}{p}\right),$$

where C is an absolute constant.

Proof. See Erdös [1, p. 76].

THEOREM 3.3.

$$\sum_{n\leqslant x}' \log f(n) = cx\log\log x + O(x\log_3 x),$$

where $c = \sum_p \log p/p(p-1)$.

Proof.

$$\sum_{n\leqslant x}' \log f(n) = \sum_{p\leqslant x}' (\log p)\, T_p(x) = \Sigma_1 + \Sigma_2 + \Sigma_3 \qquad \text{(say)},$$

where $T_p(x) = \sum_{n\leqslant x,\, p\mid n} \nu_p(n) = \sum_{q\leqslant (x/p),\, q\equiv 1(p)} [x/pq]$, and in Σ_1, $p < \log_2 x$, in Σ_2, $(\log_2 x) < p < \log x$, and in Σ_3, $p > \log x$. Trivially,

$$T_p(x) \leqslant \sum_{t\leqslant (x/p^2)}' \frac{x}{p(pt+1)} = O\left(\frac{x\log x}{p^2}\right).$$

Therefore,

$$\Sigma_3 = O(x).$$

For $p < \log x$, we note that,

$$\log_2 \frac{x}{p} = \log_2 x + O\left(\frac{\log_2 x}{\log x}\right),$$

So that by Lemma 3.1,

$$T_p(x) = \frac{x \log_2 x}{p(p-1)} + O\left(\frac{x}{p}\right).$$

This gives,

$$\Sigma_1 = cx \log_2 x + O(x \log_3 x),$$

where

$$c = \sum_p \log p/p(p-1).$$

Finally, using Lemma 3.2 in Σ_2, we see

$$\Sigma_2 \ll \sum_p{}' x\left(\frac{\log_2 x + \log p}{p}\right)\frac{\log p}{p},$$

where the dash on the sum indicates, $(\log_2 x) < p < \log x$. It is now easily seen that

$$\Sigma_2 = O(x),$$

which completes the proof of the theorem.

4. The Average Order of $\log^2 f(n)$

Define

$$T_{p,q}(x) = \sum_{\substack{n \leqslant x \\ p \mid n \\ q \mid n}} v_p(n)\, v_q(n).$$

Lemma 4.1. *If* $p < (\log x)^3$, *then*

$$T_{p,p}(x) = \frac{x(\log_2 x)^2}{p(p-1)^2} + T_p(x) + O\left(\frac{x(\log_2 x)\log p}{p^3}\right).$$

Proof. We have, using Lemma 3.2 and the definition of $T_p(x)$, that

$$T_{p,p}(x) = \sum_{\substack{q\equiv 1(p)\\ r\equiv 1(p)}} \left[\frac{x}{pqr}\right] - \sum_{q\equiv 1(p)} \left[\frac{x}{pq^2}\right] + \sum_{q\equiv 1(p)} \left[\frac{x}{pq}\right]$$

$$= \sum_{\substack{qr<x/p\\ q\equiv 1(p)\\ r\equiv 1(p)}} \frac{x}{pqr} + T_p(x) + O\left(\frac{x\log_2 x}{p^3}\right).$$

Now, as

$$\left(\sum_{\substack{q<\sqrt{x/p}\\ q\equiv 1(p)}} \frac{1}{q}\right)^2 \leqslant \sum_{\substack{qr<x/p\\ r\equiv 1(p)\\ q\equiv 1(p)}} \frac{1}{qr} \leqslant \left(\sum_{\substack{q<x/p\\ q\equiv 1(p)}} \frac{1}{q}\right)^2,$$

we get

$$T_{p,p}(x) = \frac{x}{p}\cdot\frac{(\log_2(x/p))^2}{(p-1)^2} + T_p(x) + O\left(\frac{x(\log_2 x)\log p}{p^3}\right)$$

$$= \frac{x(\log_2 x)^2}{p(p-1)^2} + T_p(x) + O\left(\frac{x(\log_2 x)\log p}{p^3}\right),$$

since $p < (\log x)^3$. This is the desired result.

LEMMA 4.2.

$$\sum_{p\leqslant x} T_{p,p}(x)(\log p)^2 = c_1 x(\log_2 x)^2 + O(x\log_2 x),$$

where $c_1 = \sum_p (\log p)^2/p(p-1)^2$.

Proof. We have by Lemma 4.1,

$$\sum_{p<(\log x)^2} T_{p,p}(x)(\log p)^2 = x(\log_2 x)^2 \sum_{p<(\log x)^2} \frac{(\log p)^2}{p(p-1)^2} + O(x\log_2 x)$$

$$= c_1 x(\log_2 x)^2 + O(x\log_2 x).$$

Since $v_p(n) \leqslant \log n/\log p$, we have by a previous estimate,

$$T_{p,p}(x) \leqslant \sum_{n\leqslant x} \left(\frac{\log x}{\log p}\right) v_p(n) \leqslant \left(\frac{\log x}{\log p}\right) T_p(x) \leqslant \frac{x(\log x)^2}{p^2\log p}.$$

Hence, as

$$\sum_{p>(\log x)^2} \frac{x(\log x)^2}{p^2(\log p)} (\log p)^2 \ll x(\log x)^2 \sum_{p>(\log x)^2} \frac{\log p}{p^2} \ll x,$$

we get the result of the lemma.

LEMMA 4.3. *If $p \neq q$, then,*

$$T_{p,q}(x) \leqslant x(\log x)^2/p^2q \log q.$$

Proof. Since $v_q(n) \leqslant \log x/\log q$,

$$T_{p,q}(x) \leqslant \sum_{\substack{n \leqslant x \\ q \mid n \\ p \mid n}} \left(\frac{\log x}{\log q}\right) v_p(n) \leqslant \left(\frac{\log x}{\log q}\right) \sum_{\substack{r \equiv 1(p) \\ r \leqslant x/pq}} \left[\frac{x}{pqr}\right]$$

$$\leqslant \left(\frac{\log x}{\log q}\right) T_p\left(\frac{x}{q}\right) T_p\left(\frac{x}{q}\right) \leqslant \frac{x(\log x)^2}{p^2q \log q},$$

as desired.

Now,

$$\sum_{n \leqslant x} \log^2 f(n) = \sum_{p,q \leqslant x} T_{p,q}(x)(\log p)(\log q),$$

and by Lemma 4.3,

$$\sum_{q \leqslant x} (\log q) \sum_{\substack{x > p > (\log x)^3 \\ p \neq q}} (\log p)\, T_{p,q}(x) \leqslant \sum_{\substack{q \leqslant x \\ p > (\log x)^3}} (\log p)(\log q) \frac{x(\log x)^2}{p^2q \log q}$$

$$\leqslant x(\log x)^2 \left(\sum_{q \leqslant x} \frac{1}{q}\right)\left(\sum_{p > (\log x)^3} \frac{\log p}{p^2}\right)$$

$$\ll x.$$

In order to obtain the average order of $\log^2 f(n)$, we may therefore assume that $p < (\log x)^3$ and $q < (\log x)^3$. We therefore need to evaluate,

$$\sum_{\substack{p < (\log x)^3 \\ q < (\log x)^3 \\ p \neq q}} (\log p)(\log q)\, T_{p,q}(x).$$

LEMMA 4.4. *If $p \neq q$, $p, q < (\log x)^3$, then,*

$$T_{p,q}(x) = \frac{x(\log_2 x)^2}{pq(p-1)(q-1)} + O\left(\frac{x \log_2 x \log pq}{p^2q^2}\right) + O\left(\frac{x(\log p)(\log q)}{p^2q^2}\right)$$

$$+ v_p(q)\, O\left(\frac{x \log_2 x}{pq^2}\right) + v_q(p)\, O\left(\frac{x \log_2 x}{p^2q}\right).$$

Remark. Note that the latter two error terms do not appear unless $q \equiv 1 (\mathrm{mod}\, p)$ or $p \equiv 1 (\mathrm{mod}\, q)$.

Proof. We have if $p \not\equiv 1 \pmod q$ or $q \not\equiv 1 \pmod p$,

$$T_{p,q}(x) = \sum_{\substack{r \equiv 1 \pmod p \\ s \equiv 1 \pmod q}} \left\lfloor \frac{x}{pq[r,s]} \right\rfloor.$$

If $q \equiv 1 \pmod p$, then,

$$T_{p,q}(x) = \sum_{\substack{r \equiv (\mathrm{mod}\, p) \\ s \equiv 1 (\mathrm{mod}\, q)}} \left\lfloor \frac{x}{pq[r,s]} \right\rfloor - \sum_{s \equiv 1 (\mathrm{mod}\, q)} \left\lfloor \frac{x}{pq^2 s} \right\rfloor + \sum_{s \equiv 1 (\mathrm{mod}\, q)} \left\lfloor \frac{x}{pqs} \right\rfloor.$$

A similar result holds if $p \equiv 1 \pmod q$. In these last two cases, the error is easily estimated to be

$$v_p(q)\, O\left(\frac{x \log_2 x}{pq^2}\right) + v_q(p)\, O\left(\frac{x \log_2 x}{p^2 q}\right).$$

The sum,

$$\begin{aligned}
&\sum_{\substack{r \equiv 1 \pmod p \\ s \equiv 1 \pmod q}} \left\lfloor \frac{x}{pq[r,s]} \right\rfloor \\
&\quad = \sum_{\substack{r \equiv 1 \pmod p \\ s \equiv 1 \pmod q}} \left\lfloor \frac{x}{pqrs} \right\rfloor - \sum_{r \equiv 1 \pmod{pq}} \left\lfloor \frac{x}{pqr^2} \right\rfloor + \sum_{r \equiv 1 \pmod{pq}} \left\lfloor \frac{x}{pqr} \right\rfloor \\
&\quad = \frac{x}{pq} \sum_{\substack{r \equiv 1 \pmod p \\ s \equiv 1 \pmod q \\ rs < x/pq}} \frac{1}{rs} + O\left(\sum_{\substack{r \equiv 1 \pmod p \\ r < x^{1/2}}} \pi\left(\frac{x}{pqr}, q, 1\right)\right) \\
&\qquad + O\left(\sum_{\substack{s \equiv 1 \pmod q \\ s < x^{1/2}}} \pi\left(\frac{x}{pqs}, p, 1\right)\right) + O\left(\frac{x \log_2 x}{p^2 q^2}\right).
\end{aligned}$$

Using Lemma 3.1 and the Brun–Titchmarsh inequality, this becomes

$$\begin{aligned}
&\sum_{\substack{r \equiv 1 \pmod p \\ s \equiv 1 \pmod q}} \left\lfloor \frac{x}{pq[r,s]} \right\rfloor \\
&\quad = \frac{x(\log_2 x)^2}{pq(p-1)(q-1)} + O\left(\frac{x(\log_2 x)(\log pq)}{p^2 q^2}\right) + O\left(\frac{x(\log p)(\log q)}{p^2 q^2}\right),
\end{aligned}$$

as desired.

We can now prove

THEOREM 4.5.

$$\sum_{n \leqslant x} \log^2 f(n) = c_2 x(\log_2 x)^2 + O(x \log_2 x),$$

where

$$c_2 = c^2 + \sum_p \frac{(\log p)^2}{p(p-1)^2}\left(1 - \frac{1}{p}\right).$$

Proof. We have, by Lemma 4.2,

$$\sum_{n \leqslant x} \log^2 f(n) = c_1 x(\log_2 x)^2 + \sum_{\substack{p \neq q \\ p,q \leqslant x}} T_{p,q}(x)(\log p)(\log q) + O(x \log_2 x).$$

We have already seen that we need only consider

$$\sum_{\substack{p,q < (\log x)^3 \\ p \neq q}} T_{p,q}(x)(\log p)(\log q).$$

By Lemma 4.4, this is

$$\sum_{\substack{p,q < (\log x)^3 \\ p \neq q}} T_{p,q}(x)(\log p)(\log q)$$

$$= \sum_{\substack{p,q < (\log x)^3 \\ p \neq q}} \frac{x(\log_2 x)^2 (\log p)(\log q)}{pq(p-1)(q-1)} + O\left(\sum_{p,q} \frac{x(\log_2 x)(\log p)^2 (\log q)^2}{p^2 q^2}\right)$$

$$+ O\left(\sum_{p,q} v_p(q) \frac{x(\log_2 x)(\log p)(\log q)}{pq^2}\right)$$

$$+ O\left(\sum_{p,q} v_q(p) \frac{x(\log_2 x)(\log p)(\log q)}{p^2 q}\right).$$

The first error term is easily seen to be

$$O(x \log_2 x).$$

The penultimate error term is

$$O\left(\sum_{q \equiv 1 (\mathrm{mod}\, p)} \frac{x(\log_2 x)(\log p)(\log q)}{pq^2}\right)$$

$$= O\left(\sum_p \sum_{t=1}^{\infty} \frac{x(\log_2 x)(\log p)(\log pt)}{p^3 t^2}\right)$$

$$= O(x \log_2 x).$$

The error term corresponding to $p \equiv 1 (\mathrm{mod}\, q)$ is handled similarly. Finally,

$$\sum_{\substack{p,q < (\log x)^3 \\ p \neq q}} \frac{x(\log_2 x)^2 (\log p)(\log q)}{pq(p-1)(q-1)}$$

$$= \left\{ c^2 - \sum_p \frac{(\log p)^2}{p^2(p-1)^2} + O\left(\frac{1}{(\log x)^3}\right) \right\} x (\log_2 x)^2,$$

which completes the proof.

We deduce

COROLLARY 4.6. *Let*

$$S = \{ n \leqslant x : n \text{ square-free}, G(n) < (\log n)^2 \}.$$

Then

$$|S| \geqslant \tfrac{2}{5} x.$$

Proof. We see easily that $|S| \geqslant ((6/\pi^2) - (c_2/4))x$ and a simple computation gives the result.

Remark. The weaker result, $|S| \geqslant \frac{1}{5} x$ can be deduced from Theorem 3.3.

5. THE CASE $G(n) = 2$

In this section, we shall show that $F_2(x)$ is much smaller than $F_1(x)$. We begin with a group-theoretic lemma.

LEMMA 5.1. *$G(n) = 2$ if and only if*

(i) *$n = 2p$, p prime, or*

(ii) *$n = p_1 p_2 m$, $(p_1 m, \varphi(p_1 m)) = 1$, $(p_2 m, \varphi(p_2 m)) = 1$, $p_2 \equiv 1 (\mathrm{mod}\, p_1)$, or*

(iii) *$n = p^2 m$, $(pm, \varphi(pm)) = 1$, $(p+1, m) = 1$.*

Proof. Suppose first $G(n) = 2$. If n is even, then writing $n = 2^a m$, with m odd, we observe that m must be a prime power. For if p and q are distinct primes dividing m we can use the dihedral groups of orders $2p$ and $2q$ to construct two nonabelian groups of order n. This would force $G(n) \geqslant 3$. Writing then $m = p^b$, we note that $0 < a \leqslant 2$, $b \leqslant 2$ as $G(q^3) = 5$ for any prime q. If $a = 2$ clearly $b = 0$, for otherwise we would have two abelian groups of order n and at least one nonabelian group of order n. If $a = 1$ then $b = 1$ as $G(2q) = 2$ for any prime q. So we deduce (i).

Now suppose n is odd. If n is square-free, then it must have exactly one pair of prime divisors (p_1, p_2) such that $p_2 \equiv 1 \pmod{p_1}$. It must have at least one such pair, otherwise $(n, \varphi(n)) = 1$ and by a well-known result of Burnside, this implies $G(n) = 1$. Furthermore, it cannot have more than one such pair as $G(pq) = 2$ if $q \equiv 1 \pmod{p}$. This shows that n is of the form (ii).

Finally, if n is odd and not square-free, it must clearly have exactly one squared prime factor as $G(q^2) = 2$. If $n = p^2 m$, with m square-free, then $G(n) = 2$ implies all groups of order n are abelian. So by a known result (see Scott [4, p. 217]), (iii) holds.

For the converse, we note that (i) easily gives $G(n) = 2$, and the cited result in Scott [4] shows that (iii) implies all groups of order n are abelian so that $G(n) = 2$, in this case. In case (ii), we have that n is square-free. Any group G of order n must therefore be supersolvable. In particular, it has a subgroup H of order m which, by the congruence conditions, must be cyclic and normal in G. Furthermore, G is the semidirect product of H and a group of order $p_1 p_2$. Since $p_1 p_2 \nmid \varphi(m)$, this semidirect product is in fact a direct product, so that $G(n) = G(p_1 p_2) = 2$.

Remark. An alternate proof can be deduced from the proof of Theorem 2.1 and by noting that $(p_1 - 1)$ homomorphisms lead to isomorphic groups.

LEMMA 5.2. *The number of $n \leqslant x$ satisfying* (i) $n = 2p$, p *is prime is* $O(x/\log x)$.

Proof. Obvious.

LEMMA 5.3. *Let* $p < (\log_2 x)/(\log_3 x)$. *Then the number of* $n \leqslant x$, $n \equiv 0 \pmod{p}$ *and having no prime divisor* $\equiv 1 \pmod{p}$ *is* $\ll x/p(\log_3 x)^A$.

Proof. We have by Brun's sieve that for $z = x^{1/2}$

the number of positive integers in question is $\ll x/p \prod_{q \equiv 1(p), q < z} (1 - 1/q) \ll x/p \exp(-(\log_2 z/p - 1) + O(1))$, by Lemma 3.1. As $p < (\log_2 x)/\log_3 x$, this is

$$p \ll x/p(\log_3 x)^A \qquad \text{for any} \quad A > 0.$$

LEMMA 5.4. *The number of $n \leqslant x$ satisfying* (iii) *is* $o(x/(\log_3 x)^2)$.

Proof. First suppose $p < (\log_3 x)^3$. Then, the number of $n \leqslant x$ satisfying (iii) $\leqslant \sum_{p < (\log_3 x)^3} (x/p)(1/(\log_3 x)^3) = o(x/(\log_3 x)^2)$, by Lemma 5.3. If $p > (\log_3 x)^3$, then the number is at most $\leqslant \sum_{p > (\log_3 x)^3} x/p^2 \ll x/(\log_3 x)^3$.

We now focus our attention on the integers satisfying (ii). These integers have a distinguished pair (p, q) of prime divisors, with $q \equiv 1 \pmod{p}$. Let us write $A_{pq}(x)$ to be those $n \leqslant x$ satisfying (ii) divisible by pq, $q \equiv 1 \pmod{p}$.

LEMMA 5.5.

$$\sum_{p>\log x}{}' A_{pq}(x) \ll \frac{x}{\log_2 x}.$$

Proof. Clearly,

$$\sum_{p>\log x}{}' A_{pq}(x) \leqslant \sum_{p>\log x}{}' \sum_{t=1}^{[x/p]}{}' \frac{x}{p(pt+1)} \ll (x \log x) \sum_{p>\log x}{}' \frac{1}{p^2} \ll \frac{x}{\log_2 x},$$

as desired.

LEMMA 5.6.

$$\sum_{p<\log_2 x/\log_3 x}{}' A_{pq}(x) \ll \frac{x}{(\log_3 x)^2}.$$

Proof. The number of $n \leqslant x/pq$ which are not divisible by any prime $\equiv 1 \pmod p \ll (x/pq)\exp(-(\log_2 x)/(p-1))$, by Brun's sieve. Therefore, summing over p in the stated range gives by Lemma 3.2,

$$\sum_{p<\log_2 x/\log_3 x}{}' A_{p,q}(x) \ll \sum_{p<\log_2 x/\log_3 x}{}' \sum_{\substack{q\equiv 1(p)\\ q<x}}{}' \frac{px}{q(\log_2 x)^2}$$

$$\ll x/(\log_3 x)^2,$$

as desired.

LEMMA 5.7. *Let* $p < (\log_2 x)^{1-\varepsilon}$. *Then the number of* $n \leqslant x$, $n \equiv 0 \pmod p$, *having no prime divisor* $\equiv 1 \pmod p$ *is* $o(x/(\log_2 x)^2)$ *uniformly in* p.

Proof. See Erdös [1, p. 77].

THEOREM 5.8.

$$F_2(x) = O\left(\frac{x \log_4 x}{(\log_3 x)^2}\right).$$

Proof. Lemmas 5.5 and 5.6 show that we can assume

$$\frac{\log_2 x}{\log_3 x} < p < (\log x).$$

Moreover, by Lemma 5.7, we can assume that those integers satisfying (ii) have all their prime factors $>(\log_2 x)^{1-\varepsilon}$. By Brun's sieve, the number of

integers $n \leqslant x$ divisible by p and having all their prime factors $>(\log_2 x)^{1-\varepsilon}$ is $\ll x/p \log_3 x$. Summing over p in the range,

$$\frac{\log_2 x}{\log_3 x} < p < \log_2 x,$$

we deduce

$$\sum_p{}' A_{pq}(x) \ll \sum_p{}' \frac{x}{p \log_3 x} \ll \frac{x \log_4 x}{(\log_3 x)^2},$$

where the dash on the sum indicates p is in the stated range. Here we have used the well-known formula,

$$\sum_{p \leqslant x} \frac{1}{p} = \log\log x + B + O\left(\frac{1}{\log x}\right).$$

Again, by a sieve argument it is seen that the number of $n \leqslant x$ divisible by pq and having all the prime factors $>(\log_2 x)^{1-\varepsilon}$ is $\ll (x/pq) \cdot (1/\log_3 x)$. Hence,

$$\sum_{\log_2 x < p < \log x} A_{pq}(x) \ll \sum_{p > \log_2 x} \frac{x \log_2 x}{p^2 \log_3 x} \ll \frac{x}{(\log_3 x)^2}.$$

This completes the proof.

6. Concluding Remarks

An examination on the proof of the upper bound for $F_2(x)$ reveals that the main contribution is coming from those primes p satisfying

$$(\log_2 x)^{1-\varepsilon} < p < \log_2 x.$$

This should not be the case, though we are unable to show it. Ignoring the primes in this range, our proof can be modified to show

Conjecture 6.1.

$$F_2(x) = \frac{(1 + o(1))\, x e^{-\gamma}}{(\log_3 x)^2}.$$

The corresponding question for $k = 3$ is more difficult. An arithmetical description of those n such that $G(n) = k$ becomes more complicated as k increases. One can show (though not easily),

$$F_3(x) = O\left(\frac{x}{\log_3 x}\right).$$

It would be important to have asymptotic formulas for both F_2 and F_3. In the general case, we conclude by stating the following:

Problem. Determine an asymptotic formula for $F_k(x)$.

ACKNOWLEDGMENTS

We would like to thank Professors P. Erdös and J. Dixon for their useful suggestions.

REFERENCES

1. P. ERDÖS, Some asymptotic formulas in number theory, *J. Indian Math. Soc.* **12** (1948), 75–78.
2. P. GALLAGHER, Counting finite groups of given order, *Math. Z.* **102** (1967), 236–237.
3. G. HIGMAN, Enumerating p-groups, I. Inequalities, *Proc. London Math. Soc.* **10** (1960), 24–30.
4. W. SCOTT, "Group Theory," Prentice–Hall, Englewood Cliffs, N.J., 1964.
5. C. SIMS, Enumerating p-groups, *Proc. London Math. Soc.* **15** (1965), 151–166.