SIMPLE GROUPS ARE SCARCE

LARRY DORNHOFF

Even gross limits on the frequency of occurrence of finite simple groups seem to be lacking in the mathematical literature. The following theorem is offered as a first step in this direction.

I am indebted to E. L. Spitznagel for discussions leading to the formulation of this problem.

THEOREM 1. *Let $r(x)$ denote the number of integers $n \leqq x$ such that every group of order n has a normal Sylow subgroup. Then* $\lim_{x\to\infty}(r(x)/x)=1$.

The proof requires only Sylow's theorem from group theory, plus the following two results from number theory. $[x]$ denotes the greatest integer in the number x.

THEOREM 2 [1, p. 356]. *Let $f(n)$ denote the number of prime factors of n (distinct or not). For any $\delta>0$, there exists x_0 such that, for all $x \geqq x_0$, the number of integers $n \leqq x$ satisfying $|f(n)-\log\log n| > (\log\log n)^{1/2+\delta}$ is less than δx.*

THEOREM 3 [1, p. 351]. *There is a constant B such that for any $\delta>0$, there exists x_0 such that $x \geqq x_0$ implies*

$$\left| \sum_{p \leq x} \frac{1}{p} - \log\log x - B \right| < \delta.$$

(*Here the letter p is restricted to run through primes only.*)

PROOF OF THEOREM 1. Let $\epsilon>0$ be assigned. We wish to find X_0 such that for all $X \geqq X_0$, $r(X) \geqq (1-\epsilon)X$. Choose X_0 so large that the following conditions are all satisfied:

(1) If $x \geqq (X_0)^{1/2}-1$, then the number of integers $n \leqq x$ with more than $\log\log n + (\log\log n)^{3/4}$ prime factors is $<\epsilon x/4$.

(2) If $x \geqq X_0$, then

$$\frac{\log\log x - \log\log\log x - 4\log 2}{\log\log x + (\log\log x)^{3/4} + 1} > 1 - \frac{\epsilon}{4}\cdot$$

(3) If $x \geqq X_0$, then $\log x > (1-\epsilon/4)^{-1}\cdot(4/\epsilon)$.

(4) If $x \geqq \log X_0$, then

Received by the editors March 9, 1967.

$$\left|\sum_{p\leq x}\frac{1}{p}-\log\log x-B\right|<\log 2.$$

(5) If $x\geqq (X_0)^{1/2}$, then $(1-\epsilon/4)x<[x]$.

Now choose $X\geqq X_0$ and let $\mathfrak{J}$ denote the set of integers $n\leqq X$ such that every group of order n has a normal Sylow subgroup. We shall show that the order of $\mathfrak{J}$ is at least $(1-\epsilon)X$, as required.

Let $\mathcal{P}$ be the set of all primes p satisfying $(\log X)^2<p\leqq X^{1/2}$. If p is in $\mathcal{P}$, let M_p denote the set of multiples of p less than or equal to X; thus $M_p=\{p, 2p, \cdots, [X/p]p\}$. Then $1+p$ divides at most X/p^2 members of M_p, $1+2p$ divides at most $X/2p^2$ members of M_p, etc. We find that at most

$$\frac{X}{p^2}\sum_{n\leq X/p^2}\frac{1}{n}<\frac{X}{p^2}\left(\log\frac{X}{p^2}+1\right)$$

members of M_p are divisible by a number of form $1+kp$, $k>0$.

Let M_p' consist of those numbers n in M_p such that n has at most $\log\log n+(\log\log n)^{3/4}+1$ prime divisors. By (1), the order of M_p' is at least $(1-\epsilon/4)[X/p]$. Hence, using Sylow's theorem, the order of $M_p'\cap\mathfrak{J}$ is at least $(1-\epsilon/4)[X/p]-(X/p^2)(\log(X/p^2)+1)$.

Any integer $n\leqq X$ is in M_p' for at most $\log\log X+(\log\log X)^{3/4}+1$ values of p. Using this fact, (5), (3), and (4) successively, we see that the order of $\mathfrak{J}$ is at least

$$\frac{\sum_{p\in\mathcal{P}}\left\{\left(1-\frac{\epsilon}{4}\right)\left[\frac{X}{p}\right]-\frac{X}{p^2}\left(\log\frac{X}{p^2}+1\right)\right\}}{\log\log X+(\log\log X)^{3/4}+1}$$

$$\geqq\left(1-\frac{2\epsilon}{4}\right)\frac{\sum_{p\in\mathcal{P}}\frac{X}{p}\left(1-\frac{1}{1-\epsilon/4}\frac{\log X}{p}\right)}{\log\log X+(\log\log X)^{3/4}+1}$$

$$\geqq\left(1-\frac{3\epsilon}{4}\right)X\frac{\sum_{p\in\mathcal{P}}\frac{1}{p}}{\log\log X+(\log\log X)^{3/4}+1}$$

$$\geqq\left(1-\frac{3\epsilon}{4}\right)X\frac{\log\log X^{1/2}-\log\log(\log X)^2-2\log 2}{\log\log X+(\log\log X)^{3/4}+1}.$$

Now $\log\log X^{1/2}=\log\log X-\log 2$ and $\log\log(\log X)^2=\log\log\log X+\log 2$, so the order of $\mathfrak{J}$ is at least $(1-\epsilon)X$ by (2). Q.E.D.

This proof is based on the "relatively large" primes between $(\log X)^2$ and $X^{1/2}$. A proof cannot be based on the "very large" primes greater than $X^{1/2}$. The reason why is contained in Theorem 4 below, which may be of independent number-theoretic interest.

In the following, $\pi(x)$ denotes the number of primes less than or equal to x. We use the fact that $\sum_{n\leq x}(1/n)-\log x$ approaches a constant limit (Euler's constant) as $x\to\infty$ and also the prime number theorem $\lim_{x\to\infty}(\pi(x)/(x/\log x))=1$.

THEOREM 4. *If $p(x)$ is the number of integers $n\leqq x$ with a prime divisor greater than $n^{1/2}$, then* $\lim_{x\to\infty}(p(x)/x)=\log 2$.

PROOF. Let $\epsilon>0$ be assigned. We shall show that there exists N_0 such that for all integers $N\geqq N_0$,

$$\left|\,(p(2N)-p(N))-N\log 2\,\right|<\epsilon N.$$

This is sufficient. We may assume $\epsilon<\frac{1}{2}$.

We note that $\lim_{x\to 1^+}((1/(x-1))/(1/\log x))=1$, but $x/(x-1)>1/\log x$ for $x>1$. Hence we may choose $k>1$ with $((k-1)/k)\cdot(1/\log k)>1-\epsilon/6$ and $k<1+\epsilon/48$. We also choose τ such that $\tau/(k-1)=\epsilon/6$.

Note that by the prime number theorem, as $x\to\infty$, $\pi(kx)-\pi(x)$ is asymptotic to

$$\frac{kx}{\log kx}-\frac{x}{\log x}=\frac{x}{\log x}\left(k\cdot\frac{\log x}{\log k+\log x}-1\right),$$

and so asymptotic to $(k-1)x/\log x$.

Choose N_0 sufficiently large so that

(6) For $M\geqq(N_0)^{1/2}/k$, $(k-1-\tau)M/\log M\leqq\pi(kM)-\pi(M)\leqq(k-1+\tau)M/\log M$.

(7) For $N\geqq N_0$, $\pi(2N)-\pi(N)<(\epsilon\cdot\log 2/6)N$.

(8) If $N\geqq N_0$ and $s=[\log_k N^{1/2}]$, then $1-\epsilon/6<(s-1)/s$.

(9) With s as in (8), we have $\left|1/s+1/(s+1)+\cdots+1/(2s-1)-\log 2\right|<\epsilon\cdot\log 2/6$.

Let N be any integer greater than or equal to N_0, and define integers L and T by $k^L\leqq N^{1/2}$, $k^{L+1}>N^{1/2}$, $k^T\leqq(2N)^{1/2}$, $k^{T+1}>(2N)^{1/2}$. If M is any number $M\geqq N^{1/2}$, consider primes p satisfying $M<p\leqq kM$. There are $\pi(kM)-\pi(M)$ such primes. If p is such, then the number of multiples of p between N and $2N$ is at least $[N/kM]$ and not more than $N/M+1$.

By choosing $M=k^t\cdot(2N)^{1/2}$, $t=0,1,\cdots,T-1$, and using (6), we see that

$$(10)\quad p(2N) - p(N) \geqq \sum_{t=0}^{T-1} (k - 1 - \tau) \frac{k^t \cdot (2N)^{1/2}}{\log k^t \cdot (2N)^{1/2}} \left(\frac{N}{k^{t+1}(2N)^{1/2}} - 1 \right).$$

By choosing $M = k^t \cdot N^{1/2}$, $t = 0, 1, \cdots, L$, we also see that

$$(11)\quad \begin{aligned} p(2N) - p(N) &\leqq \sum_{t=0}^{L} (k - 1 + \tau) \frac{k^t \cdot (N)^{1/2}}{\log k^t \cdot (N)^{1/2}} \left(\frac{N}{k^t \cdot (N)^{1/2}} + 1 \right) \\ &\quad + \pi(2N) - \pi(N). \end{aligned}$$

Using (6) and cancellation, (10) becomes

$$\begin{aligned} p(2N) - p(N) &\geqq \frac{k - 1 - \tau}{k - 1} \cdot \frac{k - 1}{k} N \cdot \sum_{t=0}^{T-1} \frac{1}{\log(2N)^{1/2} + t \log k} \\ &\quad - (\pi(2N) - \pi(N)). \end{aligned}$$

Using (7) and the definitions of τ and k, we get

$$(12)\quad \begin{aligned} p(2N) - p(N) &\geqq \left(1 - \frac{2\epsilon}{6}\right) \log k \cdot N \cdot \sum_{t=0}^{T-1} \frac{1}{\log(2N)^{1/2} + t \log k} \\ &\quad - \frac{\epsilon \cdot \log 2}{6} N. \end{aligned}$$

Now $k^T \leqq (2N)^{1/2}$. Pick k_0 such that $k_0^T = (2N)^{1/2}$. Since $k_0 \geqq k$, we may replace $\log k$ by $\log k_0$ in the summation of (12) and preserve the inequality. We get

$$\begin{aligned} p(2N) - p(N) &\geqq \left(1 - \frac{2\epsilon}{6}\right) \log k \cdot N \cdot \sum_{t=0}^{T-1} \frac{1}{(T + t) \log k_0} - \frac{\epsilon \cdot \log 2}{6} N \\ &= \left(1 - \frac{2\epsilon}{6}\right) \frac{\log k}{\log k_0} \cdot N \cdot \sum_{t=0}^{T-1} \frac{1}{T + t} - \frac{\epsilon \cdot \log 2}{6} N. \end{aligned}$$

Since $k^{T+1} > (2N)^{1/2}$, we use (8) to see that

$$\frac{\log k}{\log k_0} > \frac{(1/(T + 1)) \log (2N)^{1/2}}{(1/T) \log (2N)^{1/2}} = \frac{T}{T + 1} > 1 - \frac{\epsilon}{6} \cdot$$

By (9),

$$\sum_{t=0}^{T-1} \frac{1}{T + t} > \log 2 - \frac{\epsilon \cdot \log 2}{6} = \log 2 \left(1 - \frac{\epsilon}{6}\right).$$

Therefore we conclude that

$$(13)\quad \begin{aligned} p(2N) - p(N) &> (1 - 4\epsilon/6) \log 2 \cdot N - (\epsilon \cdot \log 2/6) N \\ &> (1 - \epsilon) \log 2 \cdot N. \end{aligned}$$

Now $k-1+\tau<2(k-1-\tau)$. Hence, using (11), (6), and the facts $k<1+\epsilon/48$, $\log 2>\frac{1}{2}$, we easily get

$$p(2N)-p(N) \leqq \sum_{t=0}^{L-1}(k-1+\tau)\,\frac{k^t\cdot N^{1/2}}{\log k^t\cdot N^{1/2}}\cdot\frac{N}{k^t\cdot N^{1/2}} + 3(\pi(2N)-\pi(N)) + \frac{\epsilon\cdot\log 2}{6}N$$
$$\leqq \frac{k-1+\tau}{k-1}\cdot\frac{k-1}{k}\cdot kN\cdot\sum_{t=0}^{L-1}\frac{1}{\log N^{1/2}+t\log k} + \frac{4\epsilon\cdot\log 2}{6}N.$$

$k^L \leqq N^{1/2}$. Hence $L\log k \leqq \log N^{1/2}$, and the inequality is preserved if we replace $\log N^{1/2}$ by $L\log k$. Therefore

$$p(2N)-p(N) \leqq \left(1+\frac{\epsilon}{6}\right)\left(1+\frac{\epsilon}{48}\right)\frac{\log k}{\log k}N\cdot\sum_{t=0}^{L-1}\frac{1}{L+t} + \frac{4\epsilon\cdot\log 2}{6}N$$
$$\leqq \left(1+\frac{\epsilon}{48}\right)\left(1+\frac{\epsilon}{6}\right)N\log 2 + \frac{4\epsilon\cdot\log 2}{6}N$$
$$< (1+\epsilon)N\log 2.$$

This together with (13) gives the result.

ADDED IN PROOF. P. T. Bateman has kindly pointed out that Theorem 4 is a special case of several results mentioned in Math. Rev. **34** (1967), #5770.

REFERENCE

1. G. H. Hardy and E. M. Wright, *An introduction to the theory of numbers*, 4th ed., Clarendon Press, Oxford, 1960.

YALE UNIVERSITY