ACTA ARITHMETICA
XV (1968)

On a problem of P. Erdös and S. Stein

by

P. Erdös and E. Szemerédi (Budapest)

The system of congruences

$$a_i(\operatorname{mod} n_i), \quad n_1 < \ldots < n_k \tag{1}$$

is called a *covering system* if every integer satisfies at least one of the congruences (1). An old conjecture of P. Erdös states that for every integer c there is a covering system with $n_1 = c$. Selfridge and others settled this question for $c \leqslant 8$. The general case is still unsettled and seems difficult.

A system (1) is called *disjoint* if every integer satisfies at most one of the congruences (1). It is trivial that in a disjoint system we must have

$$(n_i, n_j) > 1 \quad \text{and} \quad \sum_{i=1}^{k} 1/n_i \leqslant 1.$$

It is known that a disjoint system can never be covering [2] and that for a disjoint system we have [3]

$$\sum_{i=1}^{k} \frac{1}{n_i} \leqslant 1 - \frac{1}{2^k}. \tag{2}$$

(2) is easily seen to be best possible.

Denote by $f(x)$ the largest value of k for which there exists a disjoint system (1) satisfying $n_k \leqslant x$. P. Erdös and S. Stein conjectured that $f(x) = o(x)$.

The main purpose of this paper will be to prove this conjecture. In fact, we prove the following

Theorem 1. *For every $\varepsilon > 0$ if $x > x_0(\varepsilon)$ we have* ($c_1, c_2, \ldots$ denote suitable positive constants)

$$\frac{x}{\exp\left((\log x)^{1/2+\varepsilon}\right)} < f(x) < \frac{x}{(\log x)^{c_1}}. \tag{3}$$

The proof of the lower bound we obtained with the help of S. Stein [3]. First we outline the proof of the lower bound in (3) leaving some details to the reader.

Let p_r be the least prime greater than $\exp\left((\log x)^{1/2}\right)$, $n_1 < \ldots < n_k$ are the squarefree integers not exceeding x the greatest prime factor of which is p_r. Put

$$n_j = p_{i_1} \ldots p_{i_l} p_r, \quad p_{i_1} < \ldots < p_{i_l} < p_r.$$

Let

$$a_j \equiv 0 \pmod{p_{i_1}}, \quad a_j \equiv p_{i_{s-1}} \pmod{p_{i_s}}, \quad 1 < s \leqslant l, \tag{4}$$
$$a_j \equiv p_{i_l} \pmod{p_r}.$$

The congruences (4) determine a_j uniquely $(\bmod\, n_j)$. It is easy to see that the system $a_j (\bmod\, n_j)$, $1 \leqslant j \leqslant k$, is disjoint. Clearly k equals $\psi_1(x/p_r, p_r)$ where $\psi_1(u, v)$ denotes the number of squarefree integers not exceeding u all whose prime factors do not exceed v. It easily follows from the results of de Bruijn and others [1] that for $x > x_0(\varepsilon)$

$$\psi_1(x/p_r, p_r) > \frac{x}{\exp\left((\log x)^{1/2+\varepsilon}\right)},$$

which proves the lower bound in (3).

The proof of the upper bound will be considerably more difficult. Let $N = \{n_1 < \ldots < n_k \leqslant x\}$ be an arbitrary sequence of integers. Denote by $g_N(d)$ the largest j for which there are j n's the greatest common divisor of any two of which is d. ($g_N(1)$ is thus the largest integer for which there are $g_N(1)$ n's which are pairwise relatively prime.)

Now we prove the following

LEMMA 1. *Assume that the system* (1) *is disjoint. Then we have for every* $d \geqslant 1$

$$g_N(d) \leqslant d. \tag{5}$$

Assume that (5) is not satisfied for a certain d and assume that the greatest common divisor of any two of the integers $n_{i_1}, \ldots, n_{i_{d+1}}$ is d. We show that the congruences

$$a_{i_j} (\bmod\, n_{i_j}), \quad 1 \leqslant j \leqslant d+1, \tag{6}$$

cannot be disjoint. To see this put $n_{i_j} = d m_{i_j}$, $1 \leqslant j \leqslant d+1$, where any two of the m's are relatively prime. By the box principle, there are two integers $1 \leqslant j_1 < j_2 \leqslant d+1$ satisfying $a_{i_{j_1}} \equiv a_{i_{j_2}} (\bmod\, d)$, but then the congruences $a_{i_{j_1}} (\bmod\, d)$ and $a_{i_{j_2}} (\bmod\, d)$ have a common solution, or the system (6) is not disjoint, which proves (5) and the lemma.

Denote $A_N(x) = \sum_{n_l \leqslant x} 1$. Put $F(x) = \max A_N(x)$ where the maximum is taken over all the sequences N which satisfy (5) for every $d \geqslant 1$. By Lemma 1 we have

$$F(x) \geqslant f(x). \tag{6}$$

Now we prove

THEOREM 2. *Let $c_3 > 0$ be sufficiently small and c_2 sufficiently large. Then*

$$\frac{x}{(\log x)^{c_2}} < F(x) < \frac{x}{(\log x)^{c_3}}. \tag{7}$$

Theorem 2 and Lemma 1 prove the upper bound in (3) and this completes the proof of Theorem 1.

It is quite possible that $f(x) < x/\exp(\log x)^{c_4}$ for some $c_4 > 0$, but the lower bound in (7) shows that the method used in this paper cannot give $f(x) < x/(\log x)^{c_2}$.

To prove Theorem 2 we need some lemmas.

LEMMA 2. *The number of integers $n \leqslant x$ divisible by the square of a prime $p > \log x$ is $o(x/\log x)$.*

The number of these integers is clearly less than

$$\sum_{p > \log x} \frac{x}{p^2} = o\left(\frac{x}{\log x}\right)$$

which proves the lemma.

LEMMA 3. *Put $n = \prod_{i=1}^{k} p_i^{a_i}, p_1 < \ldots < p_k$. Let $c_3 > 0$ be sufficiently small. All but $o\left(x/(\log x)^{c_3}\right)$ integers $n \leqslant x$ have a prime factor p_j satisfying*

$$p_j > (\log x)^{10} \prod_{i=1}^{j-1} p_i^{a_i} \quad ((\log x)^{10} = T_1). \tag{8}$$

A well known theorem of Hardy and Ramanujan [4] states that for a sufficiently small $c_3 > 0$ for all but $o\left(x/(\log x)^{c_3}\right)$ integers $n \leqslant x$ we have

$$\sum_{i=1}^{k} a_i < (1 + \tfrac{1}{10}) \log\log x. \tag{9}$$

Hence we clearly can assume that n satisfies (9) and

$$x/\log x < n \leqslant x. \tag{10}$$

Denote by p_r the greatest prime factor of n which is less than $\log x$. By Lemma 2 we can assume that $a_{r+l} = 1$ for all $1 \leqslant l \leqslant k-r$. Further since n satisfies (9) we evidently have

$$\prod_{i=1}^{r} p_i^{a_i} < (\log x)^{2\log\log x} = T_2. \tag{11}$$

If (8) fails to hold for every $r < j \leqslant k$ we have from (11)

$$p_{r+1} < T_1 T_2, \quad p_{r+2} < T_1^2 T_2^2 \tag{12}$$

and by induction with respect to i (using (11) and (12))

$$p_{r+i} < (T_1 T_2)^{2^{i-1}}. \tag{13}$$

Hence finally from (13) and (9) by a simple calculation ($\exp z = e^z$)

$$p_k < (T_1 T_2)^{2^{k-1}} < \exp(2^{(1+1/10)\log\log x} \log 2 \cdot \log T_1 T_2) < x^{1/(\log\log x)^2}. \tag{14}$$

From (14), (11) and (9) we obtain

$$n < T_2 p_k^{2\log\log x} < x^{1/2}$$

which contradicts (10) and hence Lemma 3 is proved.

Now we are ready to prove the upper bound in (7). Let $n_1 < \ldots < n_r \leqslant x$ be a sequence of integers which satisfies (5) for all $d \geqslant 1$. Assume that

$$r \geqslant x/(\log x)^{c_3}. \tag{15}$$

We shall show that (15) leads to a contradiction. First of all if (15) holds then by Lemma 3 we can assume that for at least $r/2$ n_i's there is a d_i so that $d_i | n_i$ and all prime factors of n_i/d_i are greater than $d_i(\log x)^{10}$. If d_i has these properties we say that d_i corresponds to n_i. Now we prove the simple but crucial

Lemma 4. *There is at least one d which corresponds to at least $x/d(\log x)^5$ values of n_i.*

From (15) and what we just stated it follows that at least one d_i ($1 \leqslant d_i \leqslant x$) corresponds to more than $r/2 > x/2(\log x)^{c_3}$ n_i's. Thus if our lemma would be false we would have

$$\frac{x}{2(\log x)^{c_3}} < \frac{r}{2} \leqslant \frac{x}{(\log x)^5} \sum_{d=1}^{x} \frac{1}{d} = o\left(\frac{x}{\log x}\right),$$

an evident contradiction for $c_3 < 1$, which proves Lemma 4.

Let now d be an integer which satisfies Lemma 4 and let $n_1 < \ldots < n_s \leqslant x$, $s > x/d(\log x)^5$ be the n's to which d corresponds. Put

$$n_i = dv_i, \quad 1 \leqslant i \leqslant s, \quad v_i \leqslant \frac{x}{d}, \quad s > \frac{x}{d(\log x)^5}, \tag{16}$$

where all prime factors of v_i are greater than $d(\log x)^{10}$. Let $v_{i_1}, \ldots, v_{i_t}$ be a maximal set of v's which are pairwise relatively prime. We evidently have by (5)

$$d \geqslant g_N(d) \geqslant t \tag{17}$$

since $(n_{i_{j_1}}, n_{i_{j_2}}) = d, 1 \leqslant j_1 < j_2 \leqslant t$. Now we show that (16) and (17) contradict each other and this will complete the proof of the upper bound in (7).

Let $q_1 < \ldots < q_z$ be the set of prime factors of $\prod_{r=1}^{t} v_{i_r}$. Clearly

$$z < t \log x \tag{18}$$

since every $m \leqslant x$ has fewer than $\log x$ distinct prime factors. The maximality property of $v_{i_1}, \ldots, v_{i_t}$ implies that every v is divisible by at least one of the q's. Thus by (16), (18) and $q_1 > d(\log x)^{10}$ we evidently have

$$\frac{x}{d(\log x)^5} < s < \frac{x}{d} \sum_{i=1}^{z} \frac{1}{q_i} < \frac{x}{d} \cdot \frac{t \log x}{q_1} < \frac{x}{d} \cdot \frac{t}{d(\log x)^9},$$

or $t > d(\log x)^4$ which contradicts (17) and completes our proof. Thus as stated previously Theorem 1 is also proved.

To complete the proof of Theorem 2 we outline the proof of the lower bound in (7), leaving many of the details to the reader. Let n be squarefree, put $n = p_1 \ldots p_k, p_1 < \ldots < p_k$. Denote by N the set of all integers n for which

$$p_i < \prod_{j=1}^{i-1} p_j, \quad p_1 = 3, \quad p_2 = 5, \tag{19}$$

holds for every prime factor p_i, $i \geqslant 3$.

Now we show that the sequence N satisfies (5) for every $d \geqslant 1$.

To see this let $n_{i_1} < \ldots < n_{i_s}, s = g_N(d)$ be a maximal set of n's the greatest common divisor of any two of which is d. Write now $n_{i_j} = dv_j$. By (19) each v_j must have a prime factor less than d and since we must have $(v_{j_1}, v_{j_2}) = 1, 1 \leqslant j_1 < j_2 \leqslant s$, we clearly have

$$s = g_N(d) \leqslant \pi(d) < d$$

which proves that the sequence N satisfies (5) for every $d \geqslant 1$. To complete the proof of Theorem 2 we only have to show that for sufficiently large c_2 ($n_i \epsilon N$ satisfies (19))

$$N(x) = \sum_{n_i \leqslant x} 1 > \frac{x}{(\log x)^{c_2}}. \tag{20}$$

We only outline the proof of (20). Let $x^{1/2} < a_1 < \ldots < a_k < x^{3/4}$ be the sequence of squarefree integers $\equiv 0 (\bmod 3, 5, 7, 11)$ so that if

p_i and p_{i+1} are two consecutive prime factors of $a_j, p_{i+1} > 11$, then $p_{i+1} < p_i^{5/4}$. It is immediate that the a's satisfy (19) and it is not hard to prove that

$$\sum_{j=1}^{k} \frac{1}{a_j} > \frac{1}{(\log x)^{c_4}}. \tag{21}$$

It is immediate that the integers of the form

$$a_j p, \quad p < x/a_j, \quad (p, a_j) = 1, \tag{22}$$

also satisfy (19). From (21) we obtain that the number of integers of the form (22) is less than ($\nu(a_j)$ denotes the number of prime factors of a_j)

$$\frac{1}{\log x} \sum_{j=1}^{k} \left(\pi\left(\frac{x}{a_j}\right) - \nu(a_j) \right) > \frac{x}{(\log x)^{c_3}}. \tag{23}$$

The factor $1/\log x$ in (23) comes from the fact that an integer $n \leqslant x$ can be represented in the form $a_j p$ at most $\nu(n) < \log x$ times. (23) clearly implies (20), and thus the proof of Theorem 2 is complete.

References

[1] N. G. de Bruijn, *On the number of positive integers $< x$ and free of prime factors $> y$*, Nederl. Acad. Wetensch. Proc. Ser. A (Indag. Math.) 54 (1951), pp. 54-60.

[2] This result was first proved by L. Mirsky and D. Neuman, see P. Erdös, *Egy kongruenciarendszerekről szóló problémáról*, Mat. Lapok 3 (1952), pp. 122-128, see also S. K. Stein, *Unions of arithmetic sequences*, Math. Annalen 134 (1958-59), pp. 289-294.

[3] P. Erdös, *Számelméleti megjegyzések IV*, Mat. Lapok 13 (1962), pp. 241-243.

[4] S. Ramanujan, *Collected papers* (1927), pp. 262-275.

Reçu par la Rédaction le 16. 2. 1967