ON THE DIVISIBILITY PROPERTIES OF SEQUENCES OF INTEGERS

By P. ERDÖS *and* A. SÁRKÖZI†

[Received 13 March 1970]

In memory of H. DAVENPORT

Let $a_1 < a_2 < \dots$ be a sequence of integers denoted by A. Put $A(x) = \sum_{a_i<x} 1$. If no a_i divides any other then A is called a *primitive sequence*. It is well known and easy to see that, for a primitive sequence, $\max A(x) = [\frac{1}{2}(x+1)]$. Besicovitch (**1**) constructed a primitive sequence of positive upper density and Behrend and Erdös (**1**) proved that every primitive sequence has lower density 0. Davenport and Erdös (**1**) proved that if A has positive upper logarithmic density then there is an infinite subsequence $(a_{i_j})_{j=1,\dots}$ of A such that $a_{i_j} | a_{i_{j+1}}$. Erdös (**2**) proved that, if we assume that no a_i divides the product of two others, then‡

$$\pi(x) + \frac{c_1 x^{\frac{2}{3}}}{(\log x)^2} < \max A(x) < \pi(x) + \frac{c_2 x^{\frac{2}{3}}}{(\log x)^2},$$

where the maximum is to be taken over all sequences no term of which divides the product of two others.

These results led us to consider the question: assuming that no a_i divides the sum of two others, how large can $\max A(x)$ be? In this form the question can be reduced to an old problem. Denote by $r_k(x)$ the maximum number of integers not exceeding x which do not contain an arithmetic progression of k terms. Then it is easy to see that

$$r_3([\tfrac{1}{3}x]) \leqslant \max A(x) \leqslant r_3(x) \leqslant 3r_3([\tfrac{1}{3}x]) + 1.$$

Further, a well-known result of Roth (**3**) states that

$$r_3(x) < \frac{cx}{\log\log x}.$$

These facts lead us to modify our condition slightly. We say that a sequence A has *property* P if no term a_i divides the sum of two larger terms. We believe that if A has property P then

$$\max A(x) = [\tfrac{1}{3}x] + 1. \qquad (1)$$

† Supported in part by NSF Grant No. GP-9661.

‡ $\pi(x)$ denotes the number of primes not exceeding x and C, c, c_1, c_2, … denote suitable positive absolute constants not necessarily the same at each occurrence.

It is perhaps surprising that we could not prove (1). To show that $\max A(x) \geqslant [\frac{1}{3}x]+1$ is easy—it suffices to let A be the $[\frac{1}{3}x]+1$ greatest integers not exceeding x. Szemerédi has proved (oral communication) that if $A(x) > [\frac{1}{3}x]+1$ then there are three distinct terms a_i, a_j, a_l such that $a_i \mid (a_j+a_l)$ but $(a_j+a_l)/a_i \neq 2$. We do not give the proof of Szemerédi.

We prove the following

THEOREM. *Let the infinite set A satisfy property* P. *Then A has density* 0.

Before we prove our theorem we make a few remarks. Our theorem is best possible in the following sense. Let $f(x)$ be an increasing function tending to infinity as slowly as we please. Then there exists a sequence A having property P such that $A(x_v) > x_v/f(x_v)$ for a sequence x_v tending to infinity.

To see this, let $y_1 < y_2 < \ldots$ be a sequence tending to infinity sufficiently fast. Our sequence A consists of all integers a satisfying $y_i < a < \frac{3}{2}y_i$ and $a \equiv 1(\mathrm{mod}(2y_{i-1})!)$, $i = 2, 3, \ldots$. It is easy to see that A has property P and, in fact, that $a_i+a_j \equiv 0 \pmod{a_r}$ implies $a_i+a_j = 2a_r$. Further, if the sequence (y_i) tends to infinity fast enough we evidently have

$$A(\tfrac{3}{2}y_i) > \frac{y_i}{2.(2y_{i-1})!} > \frac{y_i}{f(y_i)},$$

which proves our assertion.

Despite this counter-example it seems to us that our theorem can be improved. Probably, if A satisfies P then $\sum 1/a_i$ is convergent and in fact $\sum 1/a_i < c$ where c is an absolute constant. Also, probably, $A(x) < x^{1-c_1}$ for infinitely many x.

Let $a_i = p_i^2$ where p_i is the ith prime congruent to 3 (mod 4). This gives an example of a sequence A with property P for which $A(x) > cx^{\frac{1}{2}}/\log x$ for every x. We have not been able to do better.

A similar situation prevails with a different problem. Let A have *property* P′ if no a_i is the sum of distinct terms of A. Erdös proved (**4**) that, if A has property P′, then $A(x) = o(x)$. This result is best possible, but $\sum 1/a_i < 103$ and $A(x) < x^{1-c}$ for infinitely many x; further, there is a sequence with property P′ such that $A(x) > x^{1-c_1}$ for every x.

Denote by $p(n)$ the least, and by $P(n)$ the greatest, prime factor of n. To prove our theorem we need two lemmas.

LEMMA 1. *Let l be an integer, $x > x_0(l)$, $a_1 < \ldots < a_k \leqslant x$, $k > c_1x$. There exists d, where $d < l^{c_2}$, $P(d) \leqslant l$, $c_2 = c_2(c_1)$, such that the number of a_i of the form dt with $p(t) > l$ is greater than $c_3x/d\log l$ for some suitably small positive constant c_3.*

Proof. Put†

$$f_l(m) = \prod (p^\alpha : p^\alpha \| m, p \leqslant l).$$

We evidently have, by the theorem of Mertens,

$$\prod_{m=1}^{x} f_l(m) = \prod_{p \leqslant l} p^{[x/p]+[x/p^2]+\dots} < \prod_{p \leqslant l} p^{x/(p-1)} = \exp\left(x \sum_{p \leqslant l} \frac{\log p}{p-1}\right) < \exp(c_4 x \log l). \tag{2}$$

Denote by N the number of integers m less than x for which $f_l(m) \geqslant l^{c_2}$. From (2) we have

$$l^{c_2 N} < l^{c_4 x},$$

or

$$N < \frac{c_4 x}{c_2} < \frac{c_1 x}{2} \tag{3}$$

for sufficiently large c_2.

From (3) and the inequality $k > c_1 x$ it follows that, for at least $\frac{1}{2}c_1 x$ indices i, we have

$$f_l(a_i) < l^{c_2}. \tag{4}$$

Thus, if our lemma were not true, we should have (by (4) and the theorem of Mertens)

$$\tfrac{1}{2}c_1 x < \sum_{P(d) \leqslant l} \frac{c_3 x}{d \log l} < \frac{c_3 x}{\log l} \prod_{p \leqslant l}\left(1 + \frac{1}{p-1}\right) < c_5 c_3 x,$$

which is false for sufficiently small c_3. This contradiction proves the lemma.

LEMMA 2. *Let $l > l_0(c, k)$, and let $t_1 < \dots < t_r \leqslant y$ be a sequence of integers satisfying*

$$p(t_i) > l,\ r > cy/\log l.$$

Then there are k terms t_i which are pairwise relatively prime.

The proof is very simple. Denote by q the least prime greater than l. Clearly, for any z, the t_i satisfying $z < t \leqslant z+q$ are pairwise relatively prime. Since $r > cy/\log l$ there is a z for which there are at least $[cl/\log l]$ terms t_i in $(z, z+q)$ (and these are relatively prime). Further, for $l > l_0(c,k)$, $cl \log l > k$, which completes the proof of the lemma.

One could pose here the following extremal problem. Denote by $f(y; l, k)$ the largest value of r for which there is a sequence $t_1 < \dots < t_r \leqslant y$, with $p(t_i) > l$ for each i, such that no k of the t_i are pairwise relatively prime. Our guess is that $f(y; l, k)$ is obtained as follows. Let $p_{l+1}, \dots, p_{l+k-1}$ be the first $k-1$ primes greater than l, and let $A(y; l, k-1)$

† $p^\alpha \| m$ means $p^\alpha | m$, $p^{\alpha+1} \nmid m$.

denote the number of distinct integers not exceeding y of the form $p_{l+i}t$, with $1 \leqslant i \leqslant k-1$, $p(t) > l$. We conjecture that

$$f(y;\, l, k) = A(y;\, l, k-1); \tag{5}$$

but this has been proved only in a few special cases.

Now we are ready to prove our theorem. We show that if A has property P it must have density 0. For if not, there are infinitely many integers x_i satisfying

$$A(x_i) > c_1 x_i. \tag{6}$$

Now let $l = l(c_1)$ be sufficiently large but fixed and independent of the x_i. By Lemma 1, for every x_i there exists d_i such that

(7) $d_i < l^{c_2}$, $P(d_i) \leqslant l$, and the number of terms in A of the form

$$d_i t_s, \quad \text{with} \quad t_s < x_i/d_i, \quad p(t_s) > l \quad \text{is greater than} \quad \frac{c_3 x_i}{d_i \log l}.$$

Since the number of x_i is infinite and the number of d_i is finite (in fact, less than l^{c_2}) there are infinitely many x_i for which the same d_i satisfies (7). We now show that this leads to a contradiction.

Choose two values x_i, $x_{i'}$ for which the same d_i satisfies (7) and which satisfy

$$2^k > \frac{10}{c_3}, \quad x_{i'} > x_i^{2k}. \tag{8}$$

Apply Lemma 2 to the integers t_s in (7). If $l > l_0(c_1, k)$ then we can assume that there are k integers t_s $(1 \leqslant s \leqslant k)$ which are pairwise relatively prime and for which

$$d_i t_s \in A, \quad t_s < \frac{x_i}{d_i}, \quad p(t_s) > l \quad (s = 1, \ldots, k), \quad P(d_i) \leqslant l. \tag{9}$$

Now observe that d_i satisfies (7) for $x_{i'}$. Thus there are integers T_u $(1 \leqslant u \leqslant r)$ satisfying

$$d_i T_u \in A, \quad r > \frac{c_3 x_{i'}}{d_i \log l}, \quad p(T_u) > l. \tag{10}$$

Now we show that (8), (9), and (10) lead to a contradiction. Since A has property P we have

$$T_{u_1} + T_{u_2} \not\equiv 0 \pmod{t_s} \quad (1 \leqslant u_1 < u_2 \leqslant r,\ 1 \leqslant s \leqslant k). \tag{11}$$

From (11) it follows that the T_u lie in at most $\frac{1}{2}t_s$ residue classes mod t_s, and, since $(t_i, t_j) = 1$ $(1 \leqslant i < j \leqslant k)$, the T_u lie in at most

$$\frac{1}{2^k} \prod_{s=1}^{k} t_s$$

residue classes mod $\prod_{s=1}^{k} t_s$. It immediately follows from the sieve of Eratosthenes that, for sufficiently large $x_{i'}$, there are at most

$$(1+o(1)) \frac{x_{i'}}{d_i \prod_{s=1}^{k} t_s} \prod_{p \leqslant l} \left(1 - \frac{1}{p}\right) < \frac{10 x_{i'}}{d_i \prod_{s=1}^{k} t_s \log l}$$

T_u in any residue class mod $\prod_{s=1}^{k} t_s$. Thus, by (8), the number of T_u is less than

$$\frac{10 x_{i'}}{d_i 2^k \log l} < \frac{c_3 x_{i'}}{d_i \log l},$$

which contradicts (10) and hence our theorem is proved.

REFERENCES

1. H. Halberstam and K. F. Roth, *Sequences*, Ch. V (Oxford University Press, 1966).
2. P. Erdös, 'On sequences of integers no one of which divides the product of two others and on some related problems', *Izv. Inst. Mat. i Mech. Tomsk* 2 (1938) 74–82.
3. K. F. Roth, 'On certain sets of integers', *J. London Math. Soc.* 28 (1953) 104–9.
4. P. Erdös, 'Remarks on number theory, III. Some problems on additive number theory', *Mat. Lapok* 13 (1962) 28–38 (in Hungarian).

University of California
Los Angeles, California, U.S.A.

Mathematical Institute
University of Budapest
Budapest, Hungary