ON THE DENSITY OF SOME SEQUENCES OF INTEGERS

P. ERDÖS

Let $a_1 < a_2 < \cdots$ be any sequence of integers such that no one divides any other, and let $b_1 < b_2 < \cdots$ be the sequence composed of those integers which are divisible by at least one a. It was once conjectured that the sequence of b's necessarily possesses a density. Besicovitch[1] showed that this is not the case. Later Davenport and I[2] showed that the sequence of b's always has a logarithmic density, in other words that $\lim_{n\to\infty} (1/\log n) \sum_{b_i \leqq n} 1/b_i$ exists, and that this logarithmic density is also the lower density of the b's.

It is very easy to see that if $\sum 1/a_i$ converges, then the sequence of b's possesses a density. Also it is easy to see that if every pair of a's is relatively prime, the density of the b's equals $\prod(1-1/a_i)$, that is, is 0 if and only if $\sum 1/a_i$ diverges. In the present paper I investigate what weaker conditions will insure that the b's have a density. Let $f(n)$ denote the number of a's not exceeding n. I prove that if $f(n) < cn/\log n$, where c is a constant, then the b's have a density. This result is best possible, since we show that if $\psi(n)$ is any function which tends to infinity with n, then there exists a sequence a_n with $f(n) < n \cdot \psi(n)/\log n$, for which the density of the b's does not exist. The former result will be obtained as a consequence of a slightly more precise theorem. Let $\phi(n; x; y_1, y_2, \cdots, y_n)$ denote generally the number of integers not exceeding n which are divisible by x but not divisible by $y_1, \cdots, y_n$. Then a necessary and sufficient condition for the b's to have a density is that

$$\lim_{\epsilon\to 0} \limsup_{n\to\infty} \frac{1}{n} \sum_{n^{1-\epsilon} < a_i \leqq n} \phi(n; a_i; a_1, a_2, \cdots, a_{i-1}) = 0. \tag{1}$$

The condition (1) is certainly satisfied if $f(n) < cn/\log n$, since

$$\frac{1}{n} \sum_{n^{1-\epsilon} < a_i \leqq n} \phi(n; a_i; a_1 \cdots a_{i-1}) < \frac{1}{n} \sum_{n^{1-\epsilon} < a_i \leqq n} \left[\frac{n}{a_i}\right]$$
$$< \sum_{n^{1-\epsilon} < m \log m < n} \frac{c'}{m \log m} = O(\epsilon) + O\left(\frac{1}{n}\right).$$

Received by the editors April 28, 1947, and, in revised form, September 5, 1947.

[1] Math. Ann. vol. 110 (1934–1935) pp. 336–341.

[2] Acta Arithmetica vol. 2.

As an application of the condition (1) we shall prove that the set of all integers m which have two divisors d_1, d_2 satisfying $d_1 < d_2 \leqq 2d_1$ exists. I have long conjectured that this density exists, and has value 1, but have still not been able to prove the latter statement.

At the end of the paper I state some unsolved problems connected with the density of a sequence of positive integers.

THEOREM 1. *Let $\psi(n) \to \infty$ as $n \to \infty$. Then there exists a sequence $a_1 < a_2 < \cdots$ of positive integers such that no one of them divides any other, with $f(n) < n\psi(n)/\log n$, and such that the sequence of b's does not have a density.*

PROOF. We observe first that the condition that one a does not divide another is inessential here, since we can always select a subsequence having this property, such that every a is divisible by at least one a of the subsequence. The condition on $f(n)$ will remain valid, and the sequence of b's will not be affected.

Let $\epsilon_1, \epsilon_2, \cdots$ be a decreasing sequence of positive numbers, tending to 0 sufficiently rapidly, and let $n_r = n_r(\epsilon_r)$ be a positive integer which we shall suppose later to tend to infinity sufficiently rapidly. We suppose that $n_r^{1-\epsilon_r} > n_{r-1}$ for all r. We define the a's to consist of all integers in the interval $(n_r^{1-\epsilon_r}, n_r)$ which have all their prime factors greater than $n_r^{\epsilon_r^2}$, for $r = 1, 2, \cdots$.

We have first to estimate $f(m)$, the number of a's not exceeding m. Let r be the largest suffix for which $n_r^{1-\epsilon_r} \leqq m$. If $m \geqq n_r^2$, then clearly

$$f(m) < n_r \leqq m^{1/2} < \frac{m}{\log m} \cdot$$

Suppose, then, that $m < n_r^2$. We have

$$f(m) < n_{r-1} + M_\epsilon(m),$$

where $M_\epsilon(m)$ denotes the number of integers not exceeding m which have all their prime factors greater than $m^{\epsilon_r^2/2}$. By Brun's[3] method we obtain

$$M_\epsilon(m) < c_1 m \sum_{p \leqq m^{\epsilon_r^2/2}} (1 - p^{-1}) < c_2 \frac{m}{\epsilon_r^2 \log m},$$

where c_1, c_2, denote positive absolute constants. Hence

$$f(m) < n_{r-1} + c_2 \frac{m}{\epsilon_r^2 \log m} < \frac{n\psi(m)}{\log m}$$

[3] P. Erdös and M. Kac, Amer. J. Math. vol. 62 (1940) pp. 738–742.

provided $n_r(\epsilon_r)$ is sufficiently large. It will suffice if

$$\frac{c_2}{\epsilon_r^2} < \frac{1}{2}\psi(n_r^{1-\epsilon_r}).$$

We have now to prove that the sequence of b's (the multiples of the a's) have no density. Denote by $A(\epsilon, n)$ the density of the sequence of all integers which have at least one divisor in the interval $(n^{1-\epsilon}, n)$. In a previous paper[4] I proved that $A(\epsilon, n)\to 0$ if $\epsilon\to 0$ and $n\to\infty$ independently. Thus if $\epsilon\to 0$ and $n\to\infty$ sufficiently fast, we have

$$\sum_{r=1}^{\infty} A(\epsilon_r, n_r) < \frac{1}{2}\cdot \tag{2}$$

Denote the number of b's not exceeding m by $B(m)$. It follows from (2) that if $n_r\to\infty$ sufficiently rapidly, and $m=n_r^{1-\epsilon_r}$, then

$$B(m) < m/2. \tag{3}$$

This proves that the lower density of the b's is at most 1/2.

Next we show that the upper density of the b's is 1, and this will complete the proof of Theorem 1. It suffices to prove that

$$n_r - B(n_r) = o(n_r), \tag{4}$$

in other words that the number of integers up to n_r which are not divisible by any a is $o(n_r)$. Consider any integer t satisfying $n_r^{1-\epsilon_r/2}<t\leqq n_r$, and define

$$(g_{\epsilon_r}(t)) = g_r(t) = \prod_p{}' p^{\alpha},$$

where the dash indicates that the product is extended over all primes p with $p\leqq n_r^{\epsilon_r^2}$, and p^α is the exact power of p dividing t.

If $g_r(t)<n^{\epsilon_r/2}$, then t is divisible by an a, since $t/g_r(t)>n^{1-\epsilon_r}$ and $t/g_r(t)$ has all its prime factors greater than $n_r^{\epsilon_r^2}$, and so is an a. Hence

$$n_r - B(n_r) < n_r^{1-\epsilon_r/2} + C(n_r), \tag{5}$$

where $C(n_r)$ denotes the number of integers $t\leqq n_r$ for which $g_r(t)\geqq n_r^{\epsilon_r/2}$. We recall that the exact power of a prime p dividing $N!$ is

$$\sum_{\nu=1}^{\infty}\left[\frac{N}{p^\nu}\right] < \sum_{\nu=1}^{\infty}\frac{N}{p^\nu} = \frac{N}{p-1}\cdot$$

Hence

---

[4] J. London Math. Soc. vol. 11 (1936) pp. 92–96.

$$\prod_{t=1}^{n_r} g_r(t) \leqq \prod_{p \leqq n\epsilon_r^2} p^{n_r/p-1} = \exp\left(n_r \sum_{p \leqq n\epsilon_r^2} \frac{\log p}{p-1}\right.$$
$$< \exp(c_3\epsilon_r^2 n_r \log n_r) = n_r^{c_3\epsilon_r^2 n_r}.$$

Hence $(n_r^{\epsilon_r/2})^{C(n_r)} < n_r^{c_3\epsilon_r^2 n_r}$, whence

(6) $\quad C(n_r) < 2c_3\epsilon_r n_r.$

Substituted in (5), this proves (4), provided that $n_r^{\epsilon_r} \to \infty$, which we may suppose to be the case. This completes the proof of Theorem 1.

THEOREM 2. *A necessary and sufficient condition that the b's shall have a density is that* (1) *shall hold.*

PROOF. The necessity is easily deduced from an old result. Davenport and I[2] proved that the logarithmic density of the b's exists and has the value

$$\lim_{i\to\infty} \lim_{n\to\infty} \frac{1}{n} \sum_{j \leqq i} \phi(n; a_j; a_1, \cdots, a_{j-1}).$$

Thus if the density of the b's exists, we obtain

$$\lim_{i\to\infty} \lim_{n\to\infty} \frac{1}{n} \sum_{j > i} \phi(n; a_j; a_1, \cdots, a_{j-1}) = 0.$$

This proves the necessity of (1).

The proof of the sufficiency is much more difficult. We have

$$B(n) = \sum_{a_i \leqq n} \phi(n; a_i; a_1, \cdots, a_{i-1}) = \sum\nolimits_1 + \sum\nolimits_2 + \sum\nolimits_3,$$

where $\sum_1$ is extended over $a_i \leqq A$, $\sum_2$ over $A < a_i \leqq n^{1-\epsilon}$, $\sum_3$ over $n^{1-\epsilon} < a_i \leqq n$. Here $A = A(n)$ will be chosen later to tend to infinity with n. By the hypothesis (1) we have

$$\lim_{\epsilon\to 0} \limsup_{n\to\infty} \frac{1}{n} \sum\nolimits_3 = 0. \tag{7}$$

It follows from the earlier work[2] that if $A = A(n)$ tends to infinity sufficiently slowly, then $(1/n)\sum_1$ has a limit, this limit being the logarithmic density of the b's, and also

$$\lim_{j\to\infty}\left(\sum_{i \leqq j} \frac{1}{a_i} - \sum_{i_1 < i_2 \leqq j} \frac{1}{[a_{i_1}, a_{i_2}]} + \cdots\right).$$

Thus the proof of Theorem 2 will be complete if we are able to prove that

$$(8)\qquad \frac{1}{n}\sum\nolimits_2 = \lim_{\epsilon\to 0}\ \limsup_{n\to\infty}\frac{1}{n}\sum_{A<a_i\leqq n^{1-\epsilon}}\phi(n;\, a_i;\, a_1,\, \cdots,\, a_{i-1}) = 0.$$

We have

$$\phi(n;\, a_i;\, a_1,\, \cdots,\, a_{i-1}) = \phi\left(\frac{n}{a_i},\ 1;\, d_1^{(i)}\cdots\right),$$

where

$$d_j^{(i)} = \frac{a_j}{(a_i,\, a_j)}\,\cdot$$

We shall prove that

$$(9)\qquad \lim_{\epsilon\to 0}\ \limsup_{n\to\infty}\frac{1}{n}\sum_{A<a_i\leqq n^{1-\epsilon}}\phi'\left(\frac{n}{a_i};\ 1;\, d_1^{(i)}\cdots\right) = 0$$

where the dash indicates that we retain only those $d_j^{(i)}$ which satisfy $d_j^{(i)}<n^{\epsilon^2}$. Clearly (8) follows from (9). (Since $n^{\epsilon^2}\to\infty$, not all the $d_j^{(i)}$ are greater than or equal to n^{ϵ^2}.)

We define $g_\epsilon(t)$ as before, with n in place of n_r and ϵ in place of ϵ_r. It follows from (5) and (6) that it will suffice to prove that

$$(10)\qquad \lim_{\epsilon\to 0}\ \limsup_{n\to\infty}\frac{1}{n}\sum_{A<t_i\leqq n^{1-\epsilon}}\phi''\left(\frac{n}{a_i};\ 1;\, d_1^{(i)}\cdots\right) = 0,$$

where $\phi''(n/a_i;\, 1;\, d_1^{(i)}\cdots)$ denotes the number of integers m satisfying

$$(11)\quad m\leqq\frac{n}{a_i};\qquad m\not\equiv 0\ (\mathrm{mod}\ d_j^{(i)}),\quad d_j^{(i)}<n^{\epsilon^2};\qquad g_\epsilon(m)<n^{\epsilon/2}.$$

Consider the integers satisfying (11). They are of the form $u\cdot v$ where $u<n^{\epsilon/2}$ and all prime factors of u are less than n^{ϵ^2}, $u\not\equiv 0\ (\mathrm{mod}\ d_j^{(i)})$ for $d_j^{(i)}<n^{\epsilon^2}$, and all prime factors of v are greater than n^{ϵ^2}. We obtain by Brun's method[3] that the number of integers $m\leqq n/a_i$ with fixed u does not exceed $(n/u\cdot a_i>n^{\epsilon/2})$

$$(12)\qquad c_4\frac{n}{a_iu}\prod_{p<n^{\epsilon^2}}(1-p^{-1}).$$

Thus the number N_i of integers satisfying (11) does not exceed

$$(13)\qquad c_4\frac{n}{a_i}\sum\nolimits'\frac{1}{u}\prod_{p<n^{\epsilon^2}}(1-p^{-1})\geqq\phi''\Big(\frac{n}{a_i};\, 1;\, d_1^{(i)}\cdots\Big),$$

where the dash indicates that the summation is extended over the $u<n^{\epsilon/2}$, $u\not\equiv 0 \pmod{d_j^{(i)}}$, $d_j^{(i)}<n^{\epsilon^2}$ and all prime factors of u are less than n^{ϵ^2}.

We have to estimate $\sum N_i$. Put

$$\lim_{m\to\infty}\frac{1}{m}\,\phi\left(\frac{m}{a_i};\,1;\,d_1^{(i)}\cdots\right)=t_i, \tag{14}$$

where in (14) all the $d_j^{(i)}$ are considered. (It follows from the definition of the $d_j^{(i)}$ that they are all less than n. Thus the limit (14) exists.) It follows from our earlier work[2] that

$$\sum_{a_i>A} t_i = o(1). \tag{15}$$

Next we estimate t_i' where

$$t_i' = \lim_{m\to\infty}\frac{1}{m}\,\phi\,(\frac{m}{a_i};\,1;\,d_j^{(i)}),\qquad d_j^{(i)} < n^{\epsilon^2}.$$

Here we use the following result of Behrend[5]

$$\lim_{n\to\infty}\frac{1}{n}\,\phi(n;\,1;\,a_1,\cdots,a_i,\,b_1,\cdots,b_i)$$
$$\geqq \lim_{n\to\infty}\frac{1}{n^2}\,\phi(n;\,1;\,a_1,\cdots,a_i)\cdot\phi(n;\,1;\,b_1,\cdots,b_i)\cdots.$$

Thus clearly

$$t_i' \leqq t_i\left(\lim_{m\to\infty}\frac{1}{m}\,\phi(m;\,1;\,x_{i'})\right)^{-1} = t_i/t_i'', \tag{16}$$

where x_i runs through the integers from n^{ϵ^2} to n. It follows from the Sieve of Eratosthenes that the density of integers with $g_\epsilon(m)=k$ equals

$$\frac{1}{k}\prod_{p<n^{\epsilon^2}}(1-p^{-1}).$$

Thus clearly

$$t_i'' \geqq \sum_{k<n^{\epsilon^2}}\frac{1}{k}\prod_{p\leqq n}(1-p^{-1}) > c_5\epsilon^2$$

or

$$t_i' \leqq t_i/c_5\epsilon^2. \tag{17}$$

[5] Bull. Amer. Math. Soc. vol. 54 (1948) pp. 681–684.

Thus from (15) and (17),

$$\sum_{a_i>A} t_i = o(1). \tag{18}$$

We have by the Sieve of Eratosthenes

$$t_i' = \frac{1}{a_i}\sum{}' \frac{1}{x} \prod_{p<n^{\epsilon^2}} (1 - p^{-1}) \tag{19}$$

where the dash indicates that $x \not\equiv 0 \pmod{d_j^{(i)}}$ $d_j^{(i)} < n^{\epsilon^2}$ and all prime factors of x are less than n^{ϵ^2}. Comparing (13) and (19) we obtain

$$N_i < c_4 t_i' n. \tag{20}$$

Thus finally from (10) and (18) we obtain $\sum_{a_i>A} N_i = o(n)$ which proves (10) and completes the proof of Theorem 2.

THEOREM 3. *The density of integers having two divisors d_1 and d_2 with $d_1 < d_2 < 2d_1$ exists.*

PROOF. Define a sequence $a_1, a_2, \cdots$ of integers as follows: An integer m is an a if m has two divisors d_1 and d_2 with $d_1 < d_2 < 2d_1$, but no divisor of m has this property. To prove Theorem 3 it will be sufficient to show that the multiples of the a's have a density. Thus by Theorem 2 we only have to show that (1) is satisfied. We shall only sketch the proof.

Clearly the a's are of the form xy, where $x < y < 2x$. Thus it will be sufficient to show that the number of integers $m \leqq n$ having a divisor in the interval $(n^{1/2-\epsilon}, n^{1/2})$ is less than ηn where $\eta \to 0$ as $\epsilon \to 0$. But I proved that the density $c_{\epsilon,t}$ of integers having a divisor in $(t, t^{1+\epsilon})$ satisfies

$$\lim_{\epsilon\to 0} \lim_{t\to\infty} c_{\epsilon,t} = 0.$$

A similar argument will prove the above result, and so complete the proof of Theorem 3.

It can be shown that the density of integers having two divisors d_1 and d_2 with $d_1 < d_2 \leqq 2d_1$ and either d_1 or d_2 a prime exists and is less than 1. This result is not quite trivial, since if we denote by $a_1 < a_2 < \cdots$ the sequence of those integers having this property and such that no divisor of any a has this property, then $\sum 1/a_i$ diverges.

We now state a few unsolved problems.

I. Besicovitch[1] constructed a sequence $a_1 < a_2 < \cdots$ of integers such that no a divides any other, and the upper density of the a's is positive. A result of Behrend[6] states that

$$\lim \frac{1}{\log n} \sum_{a_i \leqq n} \frac{1}{a_i} = 0 \tag{21}$$

and I[7] proved that

$$\sum \frac{1}{a_i \log a_i} < A \tag{22}$$

where A is an absolute constant. It follows from the last two results that the lower density of the a's must be 0. In fact Davenport and I[2] proved the following stronger result: Let $d_1 < d_2 < \cdots$ be a sequence of integers of positive logarithmic density, then there exists an infinite subsequence $d_{i_1} < d_{i_2} < \cdots$ such that $d_{i_j} | d_{i_{j+i}}$. Let now $f_1 < f_2 < \cdots$ be a sequence of positive lower density. Can we always find two numbers f_i and f_j with $-f_i \nmid f_j$ and so that $[f_i | f_j]$ also belongs to the sequence? This would follow if the answer to the following purely combinatorial conjecture is in the affirmative: Let c be any constant and n large enough. Consider $c2^n$ subsets of n elements. Then there exist three of these subsets B_1, B_2, B_3 such that B_3 is the union of B_1 and B_2.

II. Let $a_1 < a_2 < \cdots$ be a sequence of real numbers such that for all integers k, i, j we have $|ka_i - a_j| \geqq 1$. Is it then true that $\sum 1/a_i \log a_i$ converges and that $\lim (1/\log n) \sum_{a_i < n} 1/a_i = 0$? If the a's are all integers the condition $|ka_j - a_s| \geqq 1$ means that no a divides any other, and in this case our conjectures are proved by (21) and (22).

III. Let $a_1 < a_2 < \cdots \leqq n$ be any sequence of integers such that no one divides any other, and let $m > n$. Denote by $B(m)$ the number of b's not exceeding m. Is it true that

$$\frac{B(m)}{m} > \frac{1}{2} \frac{B(n)}{n} ?$$

It is easy to see that the constant 2 can not be replaced by any smaller one. (Let the a's consist of a_1 and $n = a_1$, $m = 2a_1 - 1$.)

I was unable to prove or disprove any of these results.

SYRACUSE UNIVERSITY

---

[6] J. London Math. Soc. vol. 10 (1935) pp. 42–44.

[7] Ibid. vol. 10 (1935) pp. 126–128.