ON THE DENSITY OF SOME SEQUENCES OF INTEGERS

P. ERDŐS

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 \( \lim_{n \to \infty} (1/\log n) \sum_{b_i \leq n} 1/b_i \) 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

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

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

\[
\frac{1}{n} \sum_{n^{1-\epsilon} < a_i \leq n} \phi(n; a_i; a_1, a_2, \cdots, a_{i-1}) < \frac{1}{n} \sum_{n^{1-\epsilon} < a_i \leq n} \left[ \frac{n}{a_i} \right] < \sum_{n^{1-\epsilon} < m \log m < n} \frac{c'}{m \log m} = O(\epsilon) + O\left(\frac{1}{n}\right).
\]

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


\(^2\) 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 \( d_1 < d_2 \leq 2d_1 \) 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 \( a_1, a_2, \ldots \) 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 \( \epsilon_1, \epsilon_2, \ldots \) 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 \( n_r^{1-\epsilon_r} \), for \( r = 1, 2, \ldots \).

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

\[
f(m) < n_r \leq m^{1/2} < \frac{m}{\log m}.
\]

Suppose, then, that \( m < n_r^2 \). We have

\[
f(m) < n_{r-1} + M_r(m),
\]

where \( M_r(m) \) denotes the number of integers not exceeding \( m \) which have all their prime factors greater than \( m^{1/2} \). By Brun's\(^8\) method we obtain

\[
M_r(m) < c_1 m \sum_{p \leq m^{\epsilon_r/2}} \left(1 - \frac{1}{p} \right) < c_2 \frac{m}{\epsilon_r^2 \log m},
\]

where \( c_1, c_2 \), denote positive absolute constants. Hence

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

---

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

\[
\frac{c_2}{\varepsilon_r^2} < \frac{1}{2} \psi(n_r^{1-\varepsilon_r}).
\]

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

\[
\sum_{r=1}^{\infty} A(\varepsilon_r, n_r) < \frac{1}{2}.
\]

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 \( m = n_r^{1-\varepsilon_r} \), then

\[
B(m) < m/2.
\]

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

\[
n_r - B(n_r) = o(n_r),
\]

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 \( n_r^{1-\varepsilon_r/2} < t \leq n_r \), and define

\[
(g_r(t)) = g_r(t) = \prod_p p^a,
\]

where the dash indicates that the product is extended over all primes \( p \) with \( p \leq n_r^\varepsilon \), and \( p^a \) is the exact power of \( p \) dividing \( t \).

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

\[
n_r - B(n_r) < n_r^{1-\varepsilon_r/2} + C(n_r),
\]

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

\[
\sum_{r=1}^{\infty} \left\lfloor \frac{N}{p^r} \right\rfloor < \sum_{r=1}^{\infty} \frac{N}{p^r} = \frac{N}{p - 1}.
\]

Hence

\footnote{J. London Math. Soc. vol. 11 (1936) pp. 92–96.}
\[ \prod_{t=1}^{n_r} g_r(t) \leq \prod_{p \leq n_r^2} \phi^{n_r/p - 1} = \exp \left( n_r \sum_{p \leq n_r^2} \frac{\log p}{p - 1} \right) \leq \exp (c_3 \varepsilon n_r \log n_r) = n_r^{c_3 \varepsilon n_r}. \]

Hence \((n_r^{\varepsilon/2})^{C(n_r)} < n_r^{C(n_r)}\), whence

\[ (6) \quad C(n_r) < 2c_3 \varepsilon n_r. \]

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

**Theorem 2.** A necessary and sufficient condition that the \(v_s\) 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 \(b_i\)’s exists and has the value

\[ \lim_{i \to \infty} \lim_{n \to \infty} \frac{1}{n} \sum_{j \leq i} \phi(n; a_i; a_1, \ldots, a_{i-1}). \]

Thus if the density of the \(b_i\)’s exists, we obtain

\[ \lim_{i \to \infty} \lim_{n \to \infty} \frac{1}{n} \sum_{j > i} \phi(n; a_i; a_1, \ldots, a_{i-1}) = 0. \]

This proves the necessity of (1).

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

\[ B(n) = \sum_{a_i \leq n} \phi(n; a_i; a_1, \ldots, a_{i-1}) = \sum_1 + \sum_2 + \sum_3, \]

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

\[ (7) \quad \lim_{\varepsilon \to 0} \lim_{n \to \infty} \frac{1}{n} \sum_3 = 0. \]

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 \(b_i\)’s, and also

\[ \lim_{i \to \infty} \left( \sum_{i \leq j} \frac{1}{a_i} - \sum_{i < i \leq j} \frac{1}{[a_{i1}, a_{i2}]} + \cdots \right). \]

Thus the proof of Theorem 2 will be complete if we are able to prove that
We have
\[ \phi(n; a_i; a_1, \ldots, a_{i-1}) = \phi\left(\frac{n}{a_i}, 1; d_i^{(i)} \cdots \right), \]
where
\[ d_i^{(i)} = \frac{a_i}{(a_i, a_j)}. \]

We shall prove that
\[ \lim_{\epsilon \to 0} \lim_{n \to \infty} \frac{1}{n} \sum_{\substack{A < a_i \leq n^{1-\epsilon} \atop \text{ where }}} \phi'(\frac{n}{a_i} ; 1; d_i^{(i)} \cdots) = 0, \]
where the dash indicates that we retain only those \( d_j^{(0)} \) which satisfy \( d_j^{(0)} < n^2 \). Clearly (8) follows from (9). (Since \( n^2 \to \infty \), not all the \( d_j^{(0)} \) are greater than or equal to \( n^2 \).)

We define \( g_\epsilon(t) \) as before, with \( n \) in place of \( n_r \) and \( \epsilon \) in place of \( \epsilon_r \). It follows from (5) and (6) that it will suffice to prove that
\[ \lim_{\epsilon \to 0} \lim_{n \to \infty} \frac{1}{n} \sum_{\substack{A < a_i \leq n^{1-\epsilon} \atop \text{ where }}} \phi''\left(\frac{n}{a_i} ; 1; d_i^{(i)} \cdots \right) = 0, \]
where \( \phi''(n/a_i; 1; d_i^{(i)} \cdots) \) denotes the number of integers \( m \) satisfying
\[ m \leq \frac{n}{a_i}; \quad m \not\equiv 0 \pmod{d_i^{(i)}}, \quad d_i^{(i)} < n^2; \quad g_\epsilon(m) < n^{s/2}. \]

Consider the integers satisfying (11). They are of the form \( u \cdot v \) where \( u < n^{s/2} \) and all prime factors of \( u \) are less than \( n^s \), \( u \not\equiv 0 \pmod{d_j^{(0)}} \) for \( d_j^{(0)} < n^s \), and all prime factors of \( v \) are greater than \( n^s \). We obtain by Brun's method that the number of integers \( m \leq n/a_i \) with fixed \( u \) does not exceed \( (n/u \cdot a_i > n^{s/2}) \)
\[ c_4 \frac{n}{a_i} \prod_{p \leq n^s} \left(1 - p^{-1}\right). \]

Thus the number \( N_i \) of integers satisfying (11) does not exceed
\[ c_4 \frac{n}{a_i} \sum' \frac{1}{u} \prod_{p < n^s} \left(1 - p^{-1}\right) \geq \phi''\left(\frac{n}{a_i} ; 1; d_i^{(i)} \cdots \right), \]
where the dash indicates that the summation is extended over the
\[ u < n^{\frac{3}{2}}, u \equiv 0 \pmod{d_j^{(0)}}, \quad d_j^{(0)} < n^{\frac{3}{2}} \]
and all prime factors of \( u \) are less than \( n^{\frac{3}{2}} \).

We have to estimate \( \sum N_i \). Put
\[
\lim_{m \to \infty} \frac{1}{m} \phi \left( \frac{m}{a_i}; 1; d_i^{(i)} \ldots \right) = t_i,
\]
where in (14) all the \( d_j^{(0)} \) are considered. (It follows from the definition of the \( d_j^{(0)} \) that they are all less than \( n \). Thus the limit (14) exists.) It follows from our earlier work\(^2\) that
\[
\sum_{a_i > A} t_i = o(1).
\]

Next we estimate \( t_i' \) where
\[
t_i' = \lim_{m \to \infty} \frac{1}{m} \phi \left( \frac{m}{a_i}; 1; d_i^{(i)} \right), \quad d_i^{(i)} < n^{\frac{3}{2}}.
\]

Here we use the following result of Behrend\(^6\)
\[
\lim_{n \to \infty} \frac{1}{n} \phi(n; 1; a_1, \ldots, a_i, b_1, \ldots, b_i)
\geq \lim_{n \to \infty} \frac{1}{n^2} \phi(n; 1; a_1, \ldots, a_i) \cdot \phi(n; 1; b_1, \ldots, b_i) \ldots.
\]

Thus clearly
\[
t_i' \leq t_i \left( \lim_{m \to \infty} \frac{1}{m} \phi(m; 1; x_r)^{-1} \right) = t_i / t_i',
\]
where \( x_r \) runs through the integers from \( n^{\frac{3}{2}} \) to \( n \). It follows from the Sieve of Eratosthenes that the density of integers with \( g_s(m) = k \) equals
\[
\frac{1}{k} \prod_{p < n^{\frac{3}{2}}} (1 - p^{-1}).
\]

Thus clearly
\[
t_i'' \geq \sum_{k < n^{\frac{3}{2}}} \frac{1}{k} \prod_{p \leq n} (1 - p^{-1}) > c_{6e^2}
\]
or
\[
t_i' \leq t_i / c_{6e^2}.
\]

Thus from (15) and (17),

\[(18) \quad \sum_{a_i > A} t_i = o(1).\]

We have by the Sieve of Eratosthenes

\[(19) \quad \ell'_i = \frac{1}{a_i} \sum' \frac{1}{x} \prod_{p \leq a_i^2} (1 - p^{-1})\]

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

\[(20) \quad N_i < c_4 \ell'_i n.\]

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 \(m \leq n\) having a divisor in the interval \((n^{1/2-\epsilon}, n^{1/2})\) is less than \(\eta \epsilon n\) where \(\eta \to 0\) as \(\epsilon \to 0\). But I proved that the density \(c_{e,t}\) of integers having a divisor in \((t, t^{1+\epsilon})\) satisfies

\[\lim_{\epsilon \to 0} \lim_{t \to \infty} c_{e,t} = 0.\]

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 \(d_1 < d_2 \leq 2d_1\) 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 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 Behrend⁶ states that
\[ \lim_{n \to \infty} \frac{1}{\log n} \sum_{a_i \leq n} \frac{1}{a_i} = 0 \]
and I⁷ proved that
\[ \sum \frac{1}{a_i \log a_i} < A \]
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 I² 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 \( d_{i_j} | d_{i_{j+1}} \). 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 \( -f_i | f_j \) 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 \( |ka_i - a_j| \geq 1 \). Is it then true that \( \sum 1/a_i \log a_i \) converges and that \( \lim (1/\log n) \sum_{a_i \leq n} 1/a_i = 0 \)? If the \( a \)'s are all integers the condition \( |ka_i - a_j| \geq 1 \) means that no \( a \) divides any other, and in this case our conjectures are proved by (21) and (22).

III. Let \( a_1 < a_2 < \cdots \leq n \) 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