ON THE DENSITY OF SOME SEQUENCES OF NUMBERS (II)

PAUL ERDÖS*


The functions $f(m)$ and $\phi(m)$ are called additive and multiplicative respectively if they are defined for non-negative integers $m$, and if, for $(m_1, m_2) = 1$,\n$$f(m_1 m_2) = f(m_1) + f(m_2), \quad \phi(m_1 m_2) = \phi(m_1) \phi(m_2).$$

In my paper “On the density of some sequences of numbers†” I proved the following:

**Theorem.** Let the additive function $f(m)$ satisfy the following conditions:

1. $f(m) \geq 0$,
2. $f(p_1) \neq f(p_2)$ if $p_1$, $p_2$ are different primes.

Further let $N(f; c, d)$ denote the number of positive integers $m$ not exceeding $n$, for which\n$$c \leq f(m) \leq d,$$
where $c$, $d$ are given constants; when $d = \infty$, write $N(f; c)$ for $N(f; c, \infty)$. Then $N(f; c)/n$ tends to a limit as $n \to \infty$.

I shall now prove that condition (2) is superfluous. Just as in (1), it is sufficient to consider the case when $f$ is such that $f(p) = f(p^a)$, for any positive integer $a$. I use throughout the notation of (1).

The case in which $\sum f(p) / p$ diverges may be settled just as in (1).

Suppose then that $\sum f(p) / p$ is convergent.

First take the case in which $\sum 1 / f(p)$ converges. Denote by $a_1, a_2, \ldots$ the integers composed of the primes $p$ for which $f(p) \neq 0$. Evidently\n$$\sum 1 / a_i = \prod_{f(p) = 0} \frac{1}{1 - (1/p)}$$
converges.

---

* Received 6 June, 1936; read 18 June, 1936.
Let us denote by \(a(m)\) the greatest \(a_i\) contained in \(m\). Since \(\sum \frac{1}{\prod_{f(p) \neq 0} p}\) converges, it easily follows from the sieve of Eratosthenes that the density of integers not divisible by any \(p\), with \(f(p) \neq 0\), is equal to \(\prod_{f(p) \neq 0} \left(1 - \frac{1}{p}\right)\). Hence the density of the integers \(m\) for which \(a(m) = a_i\) is

\[
\frac{1}{a_i} \prod_{f(p) \neq 0} \left(1 - \frac{1}{p}\right).
\]

Finally, since \(\sum 1/a_i\) converges, the density of the integers for which \(f(m) \geq c\) is equal to

\[
\prod_{f(p) \neq 0} \left(1 - \frac{1}{p}\right) \sum_{a_i > c} \frac{1}{a_i}.
\]

And so the theorem holds.

Take next the case in which \(\sum \frac{1}{\prod_{f(p) \neq 0} p}\) diverges. The proof is similar to that of (I). We require the same lemmas, and nothing is to be altered except that Lemma 1 of (I) must be proved without using the hypothesis \(f(p_1) \neq f(p_2)\).

**LEMMA 1 of (I).** We can find a positive number \(\delta\) such that, for all sufficiently large \(n\),

\[N(f; c, c + \delta) < en.\]

The new proof requires two lemmas. The first is the same as Lemma 2 of (I), namely:

**LEMMA 1.** Let \(f_k(m) = \sum_{p \leq m} f(p)\), where \(p_k\) denotes the \(k\)-th prime.

Then the number of integers \(m \leq n\), for which

\[f(m) - f_k(m) > \delta,
\]

is less than \(\frac{1}{2}en\) for sufficiently large \(k = k(c)\).

The proof of this did not involve the hypothesis \(f(p_1) \neq f(p_2)\). Now we split the integers \(m \leq n\) for which \(c \leq f(m) \leq c + \delta\) into two classes, putting in the first class those for which \(f(m) - f_k(m) > \delta\), and in the second class the others. By Lemma 1, the number of integers of the first class is less than \(\frac{1}{2}en\). For the integers of the second class,

\[c - \delta \leq f_k(m) \leq c + \delta;\]
hence we see that Lemma 1 of (I) will be proved if we can show that the number of integers \( m < n \) for which \( c - \delta \leq f_k(m) \leq c + \delta \) is less than \( \frac{1}{k}n \) for sufficiently large \( k = k(\varepsilon) \).

We now denote

1. by \( g_i \) the primes less than or equal to \( k \) for which \( f(g_i) > 2\delta \),
2. by \( r_i \) the other primes less than or equal to \( k \),
3. by \( a_i \) the squarefree integers composed of primes less than or equal to \( k \) for which \( c - \delta \leq f(a) \leq c + \delta \),
4. by \( \beta_1, \beta_2, \ldots \) the squarefree integers composed of the \( g_i \),
5. by \( \gamma_1, \gamma_2, \ldots \) the squarefree integers composed of the \( r_i \),
6. by \( d_s(m) \) the number of divisors of \( m \) among the \( a_i \),
7. by \( d_r(m) \) the number of divisors of \( m \) among the \( r_i \),
8. by \( d_s(m) \) the number of divisors of \( m \) among the squarefree integers composed of primes less than or equal to \( k \),
9. by \( c_1, c_2, c_3 \) absolute constants.

Now choose \( \delta \) so small and \( k \) so great that

\[
\sum \frac{1}{g_i} > A = A(\varepsilon),
\]

where \( A \) is sufficiently large. This is possible since \( \sum \frac{1}{\alpha} \) diverges.

We then prove*

**Lemma 2.** \[ \sum \frac{1}{a_i} \leq \varepsilon^2 \log k. \]

We evidently have

\[
\sum_{i=1}^{M} d_s(l) = \sum_{a_i} \left[ \sum_{a_i}^{M} \right] > \sum_{a_i} M - M.
\]

We write

\[
\sum_{i=1}^{M} d_s(l) = \sum_{1} + \sum_{2}.
\]

* The proof runs similarly to that of Behrend, "On sequences of numbers not divisible one by another", *Journal London Math. Soc.*, 10 (1935), 42–44.
where $\Sigma_1$ contains the $l$'s having less than $A$ divisors among the $q_i$, and $\Sigma_2$ all the other $l$'s. Then
\[ \Sigma_1 < 2^A \sum_{i=1}^{M} d_i(l) = 2^A \sum_{n} \left[ \frac{M}{\gamma_i} \right] \leq M 2^A \prod_{n} \left( 1 + \frac{1}{r_i} \right) = M 2^A \frac{\prod_{p \leq k} \left( 1 + \frac{1}{p} \right)}{\prod_{q_i} \left( 1 + \frac{1}{q_i} \right)} \leq c M \frac{2^A \log k}{e^A} \leq \epsilon^3 M \log k, \]
for sufficiently large $A = A(\epsilon)$.

We now estimate $\Sigma_2$. Let $l$ be an integer of $\Sigma_2$, then, if $\beta = q_1 q_2 \ldots q_x, \gamma = r_1 r_2 \ldots r_y$, we have
\[ l = \beta \gamma t, \]
where $x \geq A$ and $t$ is composed of primes greater than $k$ and the factors of $\beta \gamma$.

We estimate $d_i(l)$ as follows. Any $\alpha | l$ is of the form $\alpha = \beta_i \gamma_i$, where $\beta_i | \beta, \gamma_i | \gamma$. The $\beta_i$'s belonging to the same $\gamma_i$ cannot divide one another, for if we had $\alpha_1 = \beta_1 \gamma_1, \alpha_2 = \beta_2 \gamma_1$, and $\beta_1 | \beta_2$, then
\[ 2\delta \geq f(\alpha_2) - f(\alpha_1) = f(\beta_2) - f(\beta_1) > 2\delta, \]
an evident contradiction. From a theorem of Sperner* it follows immediately that a set of divisors of the product $q_1 q_2 \ldots q_x$, of which no one is divisible by any other, has at most $\left( \frac{x}{[\frac{1}{2}x]} \right)$ elements.

Further, from Stirling's formula
\[ (2\pi)^{n^{n+1} e^{-n} < n! \leq (2\pi)^{n^{n+1} e^{-n} e^{1/n}}, \]
we easily deduce that
\[ \left( \frac{x}{[\frac{1}{2}x]} \right) \leq \frac{2x}{x!} \leq \frac{2x}{A^x}, \]
so that
\[ d_i(l) \leq \frac{2^{x+y}}{A^x} \leq \frac{d_k(l)}{A^x}. \]
Hence
\[ \Sigma_2 < \sum_{i=1}^{M} d_i(l) \leq \frac{\sum_{i=1}^{M} d_k(l)}{A^x} \leq \frac{M \prod_{p \leq k} \left( 1 + \frac{1}{p} \right)}{A^x} \leq \frac{c M \log k}{A^x} < \epsilon^3 M \log k \]
for sufficiently large $A$.

Finally, from (1), we have

$$\sum \frac{1}{a_i} < 2c^2 \log k + 1 < c^2 \log k,$$

and so Lemma 2 is proved.

We now prove our main theorem.

We split the integers $m < n$ for which $c - \delta \leq f_k(m) \leq c + \delta$ into two classes. In the first class are the integers for which $m$ is divisible by a square greater than $1/\epsilon^4$, and in the second class the other integers. The number of integers of the first class is evidently less than or equal to

$$\sum_{r < 1/\epsilon^4} \frac{n}{r^2} < c_4 \epsilon^8 n.$$

The number of integers of the second class we estimate as follows. We write $K(m) = \Pi_{p < k} p_i^{\alpha_i}$ since $c - \delta < f_k(m) = \pi[K(m)] < c + \delta$, $K(m)$ is evidently an $a_i$. The integers $m$ of the second class for which $K(m) = a_i$ are of the form $a_i\mu t$, where $\mu$ is composed of the prime factors of $a_i$ and $t$ is composed of primes greater than $k$; $m$ is divisible by a square greater than or equal to $\mu$, for, if $\mu = p_1^{2\alpha_1} p_2^{2\alpha_2} \ldots p_1^{2\alpha_{k-1}} \ldots$, $m$ is divisible by

$$p_1^{2\alpha_1} p_2^{2\alpha_2} \ldots p_1^{2\alpha_{k-1}} \ldots.$$

Thus $\mu < 1/\epsilon^4$. Hence it easily follows from the sieve of Eratosthenes that the number of integers $m$ of the second class for which $K(m) = a_i$ is less than or equal to

$$\frac{1}{a_i} \left( c_2 n \Pi_{p < k} \left( 1 - \frac{1}{p} \right) \left( 1 - \frac{1}{\mu} \right) \sum_{\mu < 1/\epsilon^4} \frac{1}{\mu} \right).$$

Hence the number of the integers of the second class is less than or equal to

$$c_2 n \Pi_{p < k} \left( 1 - \frac{1}{p} \right) \sum_{a_i < 1/\epsilon^4} \frac{1}{a_i} \sum_{\mu < 1/\epsilon^4} \frac{1}{\mu} < c_3 n \epsilon^2 \log \frac{1}{\epsilon^4} < \frac{1}{\epsilon^4} n;$$

hence the result.

Similar results hold for multiplicative functions, since, if $\phi(m)$ is multiplicative, $\log \phi(m)$ is additive. Hence we find that, if $\phi(m) > 1$, $N(\phi; c)/n$ tends to a limit as $n \to \infty$.

The University,
Manchester.