The functions $f(m)$ and $\phi(m)$, where $\phi(m) > 0$, are called additive and multiplicative respectively if they are defined for non-negative integers $m$ and if, for $(m_1, m_2) = 1$,

$$f(m_1 m_2) = f(m_1) + f(m_2),$$

$$\phi(m_1 m_2) = \phi(m_1) \cdot \phi(m_2).$$

The first question is under what conditions does the density of the integers for which $f(m)$ [or $\phi(m)$] is not less than $c$ exist, for any given $c$. If we denote this density by $\psi(c)$, the second question is, under what conditions is $\psi(c)$ a continuous function of $c$. We shall call the function $\psi(c)$ the distribution function of $f(m)$.

Since the logarithm of a multiplicative function is additive, it will be sufficient to consider additive functions only.

So far as I know, the first paper on this subject is due to Schoenberg†, who proved (among other results) that $\phi(m)/m$, where $\phi(m)$ is Euler’s

---

* Received 30 August, 1937; read 18 November, 1937.
function has a continuous distribution function. Later Davenport \* proved the same for \( \sigma(m)/m \), where \( \sigma(m) \) denotes the sum of the divisors of \( m \), i.e. he proved that the density of the abundant numbers exists. Some time ago Schoenberg published some new and general results \( \dagger \), which included all previously known results. He proved the following theorems: Let an additive function \( f(m) \) satisfy the condition that

\[
\sum_{p} \frac{|f(p)|}{p}
\]

converges, where \( ||x|| = \min(1, |x|) \). Then

1. The distribution function of \( f(m) \) exists.

2. If \( f(m) \) satisfies the supplementary condition that there exists an infinite sequence of primes \( p_1, p_2, \ldots \) with \( f(p_\mu) \neq f(p_\nu) \) for \( \mu \neq \nu \) and such that \( \sum_{\nu=1}^{\infty} \frac{1}{p_\nu} \) diverges, then the distribution function is continuous.

3. If, on the other hand, \( \sum_{f(p)=0} \frac{1}{p} \) converges, the distribution function is purely discontinuous.

In his proofs Schoenberg used the theory of Fourier transforms.

Independently of Schoenberg I have proved by elementary methods the following results \( \ddagger \).

(i) Let the additive function \( f(m) \) satisfy the following conditions:

1. \( f(m) \geq 0 \).
2. \( f(p_1) \neq f(p_2) \).
3. \( \sum_{p} \frac{|f(p)|}{p} \) converges.

Then the distribution function of \( f(m) \) exists. Implicitly I also proved that the distribution function is continuous.

(ii) If \( f(m) \geq 0 \) and \( \sum_{p} \frac{|f(p)|}{p} \) diverges, then, for every \( c, f(m) > c \) for almost all \( m \).\( \S \)

\( \ddagger \) P. Erdős, "On the density of some sequences of numbers", Journal London Math. Soc., 10 (1935), 120–125. This paper will be referred to as I.
\( \S \) This result is also proved in Schoenberg’s paper previously quoted.
In a second paper* (referred to in the following as II) I have proved the following results:

(i) Let the additive function $f(m)$ satisfy the following conditions:

$$f(m) \geq 0,$$

$$\sum_p \frac{|f(p)|}{p} \text{ converges;}$$

then the distribution function of $f(m)$ exists.

(ii) If $f(m)$ satisfies the following supplementary condition:

$$\sum_{p|f(p)| \neq 0} \frac{1}{p} \text{ diverges},$$

then the distribution function is continuous. This result is not stated explicitly. This result together with the third result of Schoenberg gives a necessary and sufficient condition for the continuity of the distribution function in the case $f(m) \geq 0$.

In the present paper I prove the following generalization of Schoenberg's and my own results:

(i) Let the additive function $f(m)$ satisfy the following conditions:

(a) $$\sum_p \frac{|f(p)|'}{p} \text{ converges,}$$

where $|f'|$ denotes $f(p)$ for $|f(p)| \leq 1$ and 1 for $|f(p)| > 1$,

(b) $$\sum_p \frac{|f(p)|^2}{p} \text{ converges;}$$

then the distribution function of $f(m)$ exists.

(ii) If the additive function satisfies the supplementary condition

(c) $$\sum_{p|f(p)| \neq 0} \frac{1}{p} \text{ diverges;}$$

then the distribution function is continuous.

(iii) If $$\sum_{p|f(p)| \neq 0} \frac{1}{p} \text{ converges,}$$ the distribution function is purely discontinuous.

It is easy to see that this result contains the result of Schoenberg as well as my own [except (ii) of I].
The proof is elementary and very similar to the argument used in I and II.

First suppose that \( \sum_{f(p) \neq 0} \frac{1}{p} \) converges. This case is settled as in II.

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

Denote by \( a(m) \) the greatest \( a_i \) contained in \( m \). Since \( \sum_{f(p) \neq 0} \frac{1}{p} \) converges, an application of the sieve of Eratosthenes shows 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

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

Finally, since \( \sum 1/a_i \) converges, the density of the integers for which \( f(m) \geq c \) is equal to \( \prod_{f(p) \neq 0} \left(1 - \frac{1}{p}\right) \sum_{a_i} \frac{1}{a_i} \); thus the distribution function exists. It is clear that its points of discontinuity are the values \( f(a_i) \). Thus it is purely discontinuous, and this proves (iii).

Let us now suppose that \( \sum_{f(p) \neq 0} \frac{1}{p} \) diverges.

We denote by \( N(f; c, d) \) the number of positive integers not exceeding \( n \) for which

\[
c \leq f(m) \leq d,
\]

where \( c, d \) are given constants [when \( d = \infty \) we write \( N(f; c) \)].

As in I and II it is sufficient to consider the special case \( f(p^n) = f(p) \) for any \( a \), so that

\[
f(m) = \sum_{p|m} f(p).
\]

Consider also the function

\[
f_k(m) = \sum_{p|m, p < k} f(p).
\]

We show that \( N(f_k; c)n \) tends to a limit. For, if we denote by \( A_1, A_2, \ldots, A_i \) the integers whose prime factors are not greater than \( k \) and for which also \( f_k(A) \geq c \), we obtain the integers \( m \leq n \) for which \( f_k(m) \geq c \).
by taking all the multiples of $A_1, A_2, \ldots, A_t$ not exceeding $n$. Hence $N(f_k; c)/n$ tends to a limit.

To prove the existence of $N(f; c)/n$ it is sufficient to show that for every $\varepsilon > 0$ a $k_0$ exists so great that, for every $k > k_0$ and $n > n(\varepsilon)$,

$$|N(f; c) - N(f_k; c)|/n < \varepsilon.$$  

This will be the case if the number of integers $m \leq n$ for which $f_k(m) < c$ and $f(m) \geq c$ or $f_k(m) \geq c$ and $f(m) < c$ is less than $\varepsilon n$.

We require three lemmas.

**Lemma 1.** Let the additive function $f(m)$ satisfy the conditions (a) and (b). The number of integers $m \leq n$ for which

$$|f(m) - f_k(m)| > \varepsilon$$

is then less than $\frac{1}{2} \varepsilon n$ for $k > k_0(\varepsilon, \delta)$ and $n > n_0(k, \epsilon, \delta)$.

**Proof.** We divide the integers $m \leq n$ for which $|f(m) - f_k(m)| > \varepsilon$ into two classes. In the first class are the integers divisible by a prime $p > k$ with $|f(p)| \geq 1$, and in the second class all other integers. From (b) it follows that $\Sigma \frac{1}{p}$ converges; hence the number of integers $m \leq n$ of the first class is less than

$$\sum_{p \geq k} \frac{n}{p} < \{\varepsilon n$$

for sufficiently large $k$.

For the integers of the second class we evidently have

$$\sum_{m=1}^{n} (|f(m) - f_k(m)|)^2$$

$$\leq \sum_{p \geq k} f(p)^2 \left[ \frac{n}{p} \right] + 2 \sum_{p \geq k} f(p) f(q) \left[ \frac{n}{pq} \right]$$

$$< \sum_{p \geq k} \frac{n f(p)^2}{p} + 2 \sum_{p \geq k} \frac{n f(p) f(q)}{pq} + 2 \sum_{p \geq k} \frac{f(p) f(q)}{|f(p) f(q)|},$$

where $\Sigma'$ means that the summation is extended only over the $m$'s of the second class.

Now

$$2 \sum_{p \geq k} \frac{f(p) f(q)}{pq} \leq \left( \sum_{p \geq k} \frac{f(p)}{p} \right)^2 + 2 \sum_{p \geq k} \frac{f(p)}{p} \sum_{q \geq k} \frac{f(q)}{q}.$$
but, from (a) and (b),

\[ \left| \sum_{n^2 \leq q \leq k \atop (f(q) < 1)} \frac{f(q)}{q} \right| < \eta \]

for any fixed \( \eta > 0 \), and all \( n' \) if \( k \) is sufficiently large, so that

\[ 2 \sum_{p > q > k \atop pq \leq n} \frac{f(p)f(q)}{pq} < \eta^2 + 2\eta \sum_{n > p > v_n} \frac{1}{p} < c\eta, \tag{1} \]

since

\[ \sum_{n > p > v_n} \frac{1}{p} < c. \]

(The \( c \)'s denote absolute constants, not necessarily the same.)

Thus finally from (b), (1) and from the fact that the number of integers of the form \( pq \) not exceeding \( n \) is \( o(n) \), we get

\[ \sum_{m=1}^{n} \left[ f(m) - f_k(m) \right]^2 < \frac{1}{2} c\delta^2 n + \delta \eta n + c(n) < \frac{1}{2} c\delta^2 n. \]

Thus the number of integers of the second class is also less than \( \frac{1}{4} \varepsilon n \); and the lemma is proved.

**Lemma 2.** Let the additive function \( f(m) \) satisfy (a), (b), and (c), then for every \( \varepsilon > 0 \) there exists a \( \delta > 0 \) such that

\[ N(f; c-\delta, c+\delta) < \varepsilon n. \]

**Proof.** We divide the integers \( m \leq n \) for which \( c-\delta < f(m) < c+\delta \) into two classes, putting in the first 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 \( \frac{1}{4} \varepsilon n \). For the integers of the second class,

\[ c - 2\delta \leq f_k(m) \leq c + 2\delta; \]

hence we see that Lemma 2 will be proved if we can show that the number of integers \( m \leq n \), for which \( c - 2\delta < f_k(m) < c + 2\delta \), is less than \( \frac{1}{4} \varepsilon n \) for sufficiently large \( k = k(\varepsilon) \) say.

Since \( \sum \frac{1}{f(p) \neq 0 \atop p} \) diverges, we may suppose without loss of generality that \( \sum \frac{1}{p} \) diverges.
ON THE DENSITY OF SOME SEQUENCES OF NUMBERS: III. 125

We now denote
(1) by \( q_i \) the primes less than or equal to \( k \), for which \( f(q_i) > 4 \beta \),
(2) by \( r_i \) the other primes less than or equal to \( k \),
(3) by \( a_i \) the square-free integers composed of primes less than or equal to \( k \) for which \( c-2 \delta < f(a) < c+2 \delta \),
(4) by \( \beta_1, \beta_2, \ldots \) the square-free integers composed of the \( q_i \),
(5) by \( \gamma_1, \gamma_2, \ldots \) the square-free integers composed of the \( r_i \),
(6) by \( d_q(m) \) the number of divisors of \( m \) among the \( a_i \),
(7) by \( d_r(m) \) the number of divisors of \( m \) among the \( \gamma_i \),
(8) by \( d_s(m) \) the number of divisors of \( m \) among the square-free integers composed of primes less than or equal to \( k \).

Now choose \( \delta \) so small and \( k \) so large that
\[
\sum \frac{1}{q_i} > B = B(\epsilon),
\]
where \( B \) is sufficiently large. This is evidently possible since \( \sum \frac{1}{p} \) diverges.

We then prove
**Lemma** 3. \( \sum \frac{1}{a_i} \leq \epsilon^3 \log k. \)

*Proof.* We evidently have
\[
\sum_{i=1}^{M} d_q(l) = \sum_{a_i} \left[ \frac{M}{a_i} \right] > \frac{M}{a_i} - M.
\]

We write
\[
\sum_{i=1}^{M} d_q(l) = \Sigma_1 + \Sigma_2,
\]
where \( \Sigma_1 \) contains the \( i \)'s having less than \( B \) divisors amongst the \( q_i \), and \( \Sigma_2 \) all the other \( i \)'s.

Then
\[
\Sigma_1 < 2^B \sum_{i=1}^{M} d_q(l) = 2^B \sum_{r_i} \left[ \frac{M}{r_i} \gamma_i \right] \leq M 2^B \prod_{r_i} \left( 1 + \frac{1}{r_i} \right) = M 2^B \prod_{q_i} \left( 1 + \frac{1}{q_i} \right)
\]
\[
\leq \frac{c M 2^B \log k}{e^B} < \epsilon^3 M \log k
\]
for sufficiently large \( B = B(\epsilon) \) say.

* This Lemma is proved in II.
We now estimate $\Sigma_2$.
Let $l$ be an integer of $\Sigma_2$; then, if $\beta = q_1 q_2 \ldots q_x$, $r = r_1 r_2 \ldots r_y$,
\[ l = \beta \gamma l, \]
where $x \geq B$ and $l$ is composed of primes greater than $k$ and the factors of $\beta \gamma$.

We estimate $d_\gamma (l)$ as follows.
Any $a \mid l$ is of the form $a = \beta_i \gamma_j$, where $\beta_i \mid \beta$, $\gamma_j \mid \gamma$.
The $\beta_i$'s belonging to the same $\gamma_j$ cannot divide one another, for if we had $a_1 = \beta_1 \gamma$, $a_2 = \beta_2 \gamma$, and $\beta_1 \mid \beta_2$, then
\[ 4\delta > f(a_2) - f(a_1) = f(\beta_2) - f(\beta_1) > 4\delta, \]
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 $\left( \frac{x}{\lfloor \frac{1}{2} x \rfloor} \right)$ elements.

Further, from Stirling's formula
\[ (2\pi)^n n^{n+1} e^{-n} < n! \leq (2\pi)^n n^{n+1} e^{-n} e^1, \]
we easily deduce that
\[ \left( \frac{x}{\lfloor \frac{1}{2} x \rfloor} \right) \leq \frac{2^x}{x!} \leq \frac{2^x}{B!}, \]
so that
\[ d_\gamma (l) \leq \frac{2^{x+y}}{B^2} \leq \frac{d_\gamma (l)}{B^2}. \]

Hence
\[ \Sigma_2 < \sum_{l=1}^M d_\gamma (l) \leq \frac{\sum_{l=1}^M d_\gamma (l)}{B^2} \leq \frac{M}{B^2} \prod_{1 < k} \left( 1 + \frac{1}{p} \right) \leq \frac{e \log k}{B^2} < e^a M \log k \]
for sufficiently large $B$.
Finally, from (1),
\[ \Sigma \frac{1}{a_q} < 2e^a \log k + 1 < e^a \log k; \]
and so Lemma 3 is proved.

We now prove Lemma 2, as follows.

---

We divide the integers $m \leq n$ for which $c-2\delta < f_k(m) < c+2\delta$ into two classes. In the first class are the integers for which $m$ is divisible by a square greater than $1/e^4$, and in the second class the other integers. The number of integers of the first class is evidently less than or equal to

$$\sum_{\tau > 1/e^4} \frac{n}{\tau^2} < c e^2 n.$$ 

The number of integers of the second class we estimate as follows. We write $K(m) = \prod_{p | m} p$. Since $c-2\delta \leq f_k(m) = f[K(m)] \leq c+2\delta$, $K(m)$ is evidently an $a$. 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 $\mu = p_1^{\alpha_1} p_2^{\alpha_2} \ldots$, $m$ is divisible by $p_1^{\alpha_1+1} p_2^{\alpha_2} \ldots$. Therefore $\mu < 1/e^4$. 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

$$cn \prod_{p \leq k} \left(1 - \frac{1}{p}\right) \sum_{a_i < 1/e^4} \frac{1}{a_i}.$$

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

$$cn \prod_{p \leq k} \left(1 - \frac{1}{p}\right) \sum_{a_i < 1/e^4} \frac{1}{a_i} \sum_{a_i \mu > 1/e^4} \frac{1}{\mu} < cn e^2 \log \frac{1}{e^4} < \frac{1}{2} en;$$

this proves Lemma 2.

We now prove the existence of the distribution function of $f(m)$. We divide the integers not exceeding $n$ satisfying the two conditions

$$f_k(m) < c, \quad f(m) \geq c$$

into two classes. In the first class we put the integers $m$ for which $f(m) > c + \delta$. For these, $f(m)-f_k(m) > \delta$, and so, from Lemma 1, their number is less than $\frac{1}{2} en$. In the second class, we put the integers for which $f(m) \leq c + \delta$. Their number is less than $\frac{1}{2} en$ from Lemma 2. Similarly for the $m$ for which $f_k(m) \geq c$, $f(m) < c$. Thus the existence of the distribution function is proved.

It is evident that the distribution function is a non-increasing function of $c$, and so its continuity is an immediate consequence of Lemma 2. This completes the proof of our result.

University of Manchester.