Annals of Mathematics
Vol. 47, No. 1, January, 1946

ON THE DISTRIBUTION FUNCTION OF ADDITIVE FUNCTIONS

By P. Erdös

(Received February 24, 1945)

Let $f(m)$ be a real valued number theoretic function. We say that it is *additive* if for $(m_1, m_2) = 1, f(m_1 \cdot m_2) = f(m_1) + f(m_2)$. Denote by $N(f; c, n)$ the number of integers $m \leqq n$ for which $f(m) \leqq c$. The function $\psi(c)$ is called the *distribution function* of $f(m)$ if $\psi(-\infty) = 0$, $\psi(\infty) = 1$ and for every $-\infty < c < \infty$

$$\psi(c) = \lim_{n\to\infty} \frac{N(f; c, n)}{n}.$$

Clearly $\psi(c)$ is non-decreasing.

Wintner and I[1] showed that a necessary and sufficient condition for the existence of a distribution function is that both

$$1) \qquad \sum_p \frac{f'(p)}{p} < \infty$$

$$2) \qquad \sum_p \frac{(f'(p))^2}{p} < \infty$$

where $f'(p) = f(p)$ for $|f(p)| \leqq 1$ and $f'(p) = 1$ otherwise.

It can be noted that the existence of the distribution function does not depend on the values $f(p^\alpha)$, $\alpha > 1$, and that the existence of the distribution function cannot be destroyed by the behavior of $f(p)$ on a sequence of primes r where $\sum 1/r_i < \infty$.

Let us now assume that $\sum_p (f'(p))^2/p$ diverges. The distribution function of course does not exist. We define $F(m)$ by $F(m) = f(m) - [f(m)]$. ($[a]$ denotes the greatest integer $\leqq a$.) We shall prove the following

Theorem I. *Let $f(p) \to 0$ as $p \to \infty$ and assume that $\sum_p (f'(p))^2/p = \infty$. Then the distribution function of $F(m)$ is x. In other words the density of integers m for which*

$$F(m) \leqq c$$

equals c. (To obtain Theorem I, it is of course sufficient to assume that $f(p) \to 0 \pmod 1$).

The proof of Theorem 1 depends on methods similar to those used in our joint paper with Kac[2] (we will refer to this paper as I), and on a result of Berry.[3] The proof will be given later.

[1] Amer. Journal of Math. (1939) Vol. 61, p. 713–721.
[2] Ibid. (1940) Vol. 62, p. 738–742.
[3] Trans. Amer. Math. Soc. (1941) Vol. 49, p. 122–136.

If we do not assume that $f(p) \to 0$, the situation becomes rather complicated. First it is clear that the distribution function of $F(m)$ does not have to be x. Put $f(p) = \frac{1}{2}$, $f(p^\alpha) = 0$. Then it follows from the prime number theorem that $\psi(x) = \frac{1}{2}$, $0 \leqq x \leqq \frac{1}{2}$, $\psi(x) = 1$, $\frac{1}{2} \leqq x \leqq 1$. If $f(p) = \alpha$, α irrational, it can be shown that the distribution function of $F(m)$ is again x. The proof is not easy and we do not discuss it here.

It can be conjectured that $F(m)$ always has a distribution function. This if true must be very deep, since it contains the prime number theorem.[3a]

Next we assume that $\sum_p (f'(p))^2/p < \infty$ and $\sum f'(p)/p$ diverges. Then we have

THEOREM II. *Put*

$$\varphi(m) = f(m) - \sum_p \frac{f'(p)}{p}.$$

Then $f(m)$ has a distribution function, and the distribution function is continuous and strictly increasing in $(-\infty, +\infty)$.

We can prove the following slightly stronger

THEOREM III. *Let $f(m)$ be additive. Assume that a constant c exists such that if we put $f(m) - c \log m = \psi(m)$, $\psi(m)$ will satisfy the conditions of Theorem II; then*

$$\varphi(m) = f(m) - c \log m - \sum_p \frac{\psi'(p)}{p} = f(m) - \sum_{p \leqq n} \frac{f(p)}{p} + c + o(1)$$

has a distribution function.

Theorem III is essentially identical with Theorem II. The converse of Theorem III is probably true, i.e., that if $f(m) - \sum_p (f(p))/p$ has a distribution function, then

$$f(m) = c \log m + \varphi(m), \qquad \sum_p \frac{\varphi'(p)^2}{p} < \infty.$$

At present we can prove this only if $f(p) > 0$.

We omit the proof of Theorem II since it is similar to that given in a previous paper.[4] In a previous paper[4] we proved that a necessary and sufficient condition for the continuity of a distribution function is that $\sum_{f(p) \neq 0} 1/p$ diverges. (We of course assumed that $\sum_p f'(p)/p$ and $\sum_p (f'(p))^2/p$ converge.) We can prove the following more general

THEOREM IV. *Let $f(m)$ be an additive function such that $\sum_{f(p) \neq 0} 1/p$ diverges. Then to every ϵ there exists a δ such that if $a_1 < a_2 < \cdots < a_x \leqq n$ is a sequence of integers with $| f(a_i) - f(a_j) | < \epsilon$ then $x < \delta n$ for n sufficiently large.*

[3a] *Added in proof.* An example of Wintner shows that this conjecture is false. His example in fact shows that not even $1/n \sum_{m=1}^{n} F(m)$ exists. See *The Theory of Measure in Arithmetical Semigroups*, 1944, p. 48, II bis.

[4] London Math. Soc. Journal (1938) Vol. 13, p. 119–127. The result in question is a special case of a result of P. Lévy.

We shall deduce Theorem IV from

THEOREM V. *Let the additive function be such that there exist two constants* c_1 *and* c_2 *and infinitely many* n, *so that there exists* $a_1 < a_2 < \cdots < a_x \leqq n$, $x > c_1 n$, $|f(a_i) - f(a_j)| < c_2$. *Then there exists a constant* c *such that if we write*

$$f^+(p) = f(p) - c \log p,$$

$\sum_p (f'^+(p)/p)^2$ *converges.*

In other words, if for many integers the values of $f(m)$ are close together, then $f(m)$ is almost equal to $c \log m$. If $f(m)$ satisfies the conditions of Theorem V we shall say that it is *finitely distributed*.

The converse is also true. In fact if

$$f(p) = c \log p + f^+(p), \qquad \sum_p \frac{(f'^+(p))^2}{p} < \infty,$$

then for every $c_1 < 1$ there exists c_2 such that for every n there exists a sequence $a_1 < a_2 < \cdots < a_x \leqq n$, $x > c_1 n$, $|f(a_i) - f(a_j)| < c_2$.

From Theorem V, we shall deduce the following two results:

1) if $f(n+1) \geqq f(n)$ for all n, then $f(n) = c \log n$.

2) $f(n+1) - f(n) \to 0$ $f(p) \neq 0$ for infinitely many p, then $f(n) = c \log n$.

The following result probably holds, but I cannot prove it: Assume that $f(n+1) - f(n) < c_1$ for all n. Then

$$f(n) = c \log n + \varphi(n), \; |\varphi(n)| < c_2 \text{ for all } n.$$

The converse is clearly true.

I also conjecture the following results:

1) if $f(n+1) \geqq f(n)$ for almost all n (i.e., all n except for a sequence of density 0), then $f(n) = c \log n$

2) if $f(n+1) - f(n) \to 0$ when n runs through a sequence of density 1 then $f(n) = c \log n$.

We shall give the proof of Theorems IV and V in full detail.

By applying the law of the iterated logarithm we obtain the following

THEOREM VI. *Let* $f(m)$ *be an additive function* $|f(p)| < c$, $\sum_p (f(p)/p)^2 = \infty$. *Let*

$$\sum_{p \leqq n} \frac{f(p)}{p} = A_n, \qquad \sum_{p \leqq n} \frac{(f(p))^2}{p} = B_n.$$

Denote further by $N_\epsilon^+(d, n)$ *the number of integers* $m \leqq n$ *such that for at least one* $u > d$

$$\sum_{\substack{p \mid m \\ p \leqq u}} f(p) > A_u + (1 + \epsilon)\sqrt{2B_u \log \log B_u}.$$

Put

$$\limsup_{n \to \infty} \frac{1}{n} N_\epsilon^+(d, n) = U^+(d).$$

Then

$$\lim_{d\to\infty} U^+(d) = 0.$$

Similarly if we denote by $N_\epsilon^-(d, n)$ the number of integers $m \leqq n$ such that for at least one $u > d$

$$\sum_{\substack{p\mid m\\ p<u}} f(p) > A_u + (1-\epsilon)\sqrt{2B_u \log\log B_u}$$

then $\lim_{n\to\infty} \frac{1}{n} N_\epsilon(d, n) = 1$, *for every* d.

If we apply Theorem VI for $f(p) = 1$ we obtain the following results: The density of integers m which have divisors d with

$$\nu(d) > \log\log d + (1-\epsilon)\sqrt{2\log\log d \log\log\log\log d}$$

is 1. ($\nu(n)$ denotes the number of different prime factors of n.) To every ϵ and η there exists a d_0 such that the density of integers m having at least one divisor $d > d_0$ with

$$\nu(d) > \log\log d + (1+\epsilon)\sqrt{2\log\log d \log\log\log\log d}$$

is $< \eta$. We can express these results roughly by stating that for almost all integers $m = p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_k^{\alpha_k}$ we have for large k

$$e^{e^{k(1-\epsilon)}} < p_k < e^{e^{k(1+\epsilon)}}.$$

We omit the proof since it is very similar to that of I.

In a previous paper[5] I proved the following results: Let $f(m)$ be an additive function such that $\sum_p f(p)/p$ converges and $\sum_{f(p)\neq 0} 1/p = \infty$. Then the density of integers for which $f(m+1) > f(m)$ is $\frac{1}{2}$. I also showed that this holds for $\nu(m)$ and $d(n)$. By using the results of I, together with the method used in that paper, we obtain the following

THEOREM VII. *Let $f(m)$ be an additive function with $|f(p)| < c$, $\sum_{f(p)\neq 0} 1/p = \infty$. Then the density of integers with $f(m+1) \geqq f(m)$ is $\frac{1}{2}$. If $\sum_{f(p)\neq 0} 1/p < \infty$, then the density of integers with $f(m+1) > f(m)$ is $< \frac{1}{2}$ and equals the density of integers with $f(m) < f(m+1)$.*

The function $f(m) = \log m$ shows that Theorem VII does not hold for all additive functions. But it very likely holds under very much more general conditions than $|f(p)| < c$. I did not even succeed thus far in making a plausible guess for a necessary and sufficient condition. Does Theorem VII hold for $f(p) = (\log p)^\alpha$, $\alpha \neq 1$? Also does it hold for $f(p) = p$? If Theorem VII is true in this case, we can easily show that the density of the integers, for which the greatest prime factor of $n + 1$ is greater than the greatest prime factor of n, is $\frac{1}{2}$. At present I cannot decide these questions. We can prove, though, that if $f(p) = (\log p)^\alpha$, $\alpha \neq 1$ then $f(n)/(\log n)^\alpha$ has a distribution function.

[5] Journal Cambridge Phil. Soc. (1936), Vol. 32, p. 530–540.

Let $f(n)$ be additive. Let $m \to \infty$, $n \to \infty$, $n - m \to \infty$. Consider

$$\frac{1}{n-m}\sum_{k=m}^{n} f(k) = M(m, n).$$

Hartmann and Wintner[6] proved that a necessary and sufficient condition that $\lim M(m, n)$, $m \to \infty$, $n \to \infty$, $n - m \to \infty$ exist is that $f(n)$ be uniformly bounded. Consider now

$$\frac{1}{n-m}\sum_{k=m}^{n} f(k) - \frac{1}{n}\sum_{k=1}^{n} f(k) = M(m, n) - M(1, n).$$

If $f(n) = c \log n + \varphi(n)$, $|\varphi(n)|$ uniformly bounded, then $M(m, n) - M(1, n) \to 0$. It can be shown that the converse is true. In other words, if $M(m, n) - M(1, n) \to 0$ then $f(n) = c \log n + \varphi(n)$, $|\varphi(n)| < c$.

Clearly we could formulate several questions of this type, e.g., involving almost periodic properties of $f(n)$. We discuss one more such problem:

Theorem VIII. *Let* $f(n) = c \log n + \varphi(n)$, *(*$|\varphi(n)|$ *uniformly bounded* $\sum_{\varphi(p)\neq 0} 1/p = \infty$*). Let* $m \to \infty$, $n \to \infty$, $n - m \to \infty$. *Then the number of integers in the interval* (m, n) *for which* $f(x + 1) > f(x)$ *equals*

$$\tfrac{1}{2}(n - m) + o(n - m).$$

We omit the proof since it is very similar to that used in a previous paper.[7]

Is the converse of Theorem VIII true? At present we cannot answer this question.

One final remark. Let $n \to \infty$, $m \to \infty$, $n - m/\log \log \log n \to \infty$. Then for Euler's φ function $\lim M(m, n) = \lim M(1, n)$. A similar result holds for Theorem VIII. In fact the distribution function of $\varphi(n)$ in (m, n) is the same as its distribution function for $(1, n)$. The same result holds for $\sigma(n)$.[8] It can be shown that the condition $n - m/\log \log \log n \to \infty$ is the best possible.

Analogous questions can be asked for $\nu(n)$ and $d(n)$, but so far the results here are very unsatisfactory.

Proof of Theorem I. First we introduce some notations:

1) $f_{u,v}(m) = \sum_{\substack{u \leq p \leq v \\ p \mid m}} f(p)$.

2) $\eta_{u,v} = \max |f(p)|$, $u \leqq p \leqq v$ ($\eta \to 0$ if $u \to \infty$).

3) $A_{u,v} = \sum_{u \leq p \leq v} \frac{f(p)}{p}$, $\qquad B_{u,v} = \sum_{u \leq p \leq v} \frac{(f(p))^2}{p}$.

4) $\epsilon_{u,v} = \frac{\eta_{u,v}}{\sqrt{2B_{u,v}}}$.

[6] Duke Journal (1942) Vol. 9, p. 112–119.

[7] Journal Cambridge Phil. Soc. (1936) Vol. 32, p. 530–540.

[8] P. Erdös, London Math. Soc. Journal (1935) Vol. 10, p. 128–131.

5) $\delta_{x,y,u,v}$ denotes the density of integers m for which

$$A_{u,v} + x\sqrt{2B_{u,v}} \leqq f_{u,v}(m) \leqq A_{u,v} + y\sqrt{2B_{u,v}}$$

6) $D_{c,u,v}$ denotes the density of integers for which

$$f_{u,v}(m) \leqq c \pmod 1.$$

7) $K_x(m) = \prod_{\substack{p^\alpha \| m \\ p \leqq x}} p^\alpha.$

8) $D_v(a_i)$ denotes the density of integers with $K_u(m) = a_i$ and

$$f_{u,v}(m) \leqq c \pmod 1)$$

9) $A(u, v, c, n)$ denotes the number of integers $m \leqq n$ with

$$f_{(u,v)}(m) \leqq c \pmod 1).$$

To shorten the proof we will refer to I wherever our proof follows I closely.

LEMMA 1. *We have*

$$\delta_{x,y,u,v} - \frac{1}{\pi^{1/2}} \int_u^v e^{-x^2}\, dx < 3.76\epsilon_{u,v}.$$

PROOF. Lemma 1 is an immediate consequence of a result of Berry (just as Lemma 1 of I is a consequence of the central limit theorem).

LEMMA 2. *Let $u \to \infty$ and $v > u$ sufficiently large. Then*

$$D_{c,u,v} \to c.$$

PROOF. Let $r \to \infty$ sufficiently slowly. We obtain from the central limit theorem that

$$\delta_{r,\infty,u,v} + \delta_{r,-\infty,u,v} \sim 2\frac{1}{\pi^{1/2}} \int_r^\infty e^{-x^2}\, dx \to 0, \qquad \text{as } r \to \infty. \tag{1}$$

Consider now the integers $z, z + 1, \cdots z + k$ in the interval

$$(A_{u,v} - r\sqrt{2B_{u,v}},\, A_{u,v} + r\sqrt{2B_{u,v}}) \qquad (\text{Since } B_{u,v} \to \infty,\, k \to \infty).$$

Then by (1)

$$\sum_{i=0}^{k-1} D_i \leqq D_{c,u,v} \leqq \sum_{i=0}^{k} D_i + o(1)$$

where D_i denotes the density of integers for which

$$z + i \leqq f_{u,v}(m) \leqq z + i + c.$$

By Lemma 1 we have

$$D_i = \frac{1}{\pi^{1/2}} \int_{\mu_i}^{\nu_i} e^{-x^2}\, dx + c\,\epsilon_{u,v}, \qquad |c| < 3.76$$

where μ and ν are determined by

$$A_{u,v} + \mu_i\sqrt{2B_{u,v}} = z + i, \qquad A_{u,v} + \nu\sqrt{2B_{u,v}} = z + i + c.$$

Thus a simple calculation shows that $\nu_i - \mu_i \to 0$ and hence we easily obtain that $D_i \to 0$. Thus

$$D_{c,u,v} = \sum_{i=0}^{k-1} D_i + o(1).$$

Now determine λ_i from $A_{u,v} + \lambda_i \sqrt{2B_{u,v}} = z + i + 1$. Then it is easy to see that

$$\frac{\int_{u_i}^{\nu_i} e^{-x^2}\,dx}{\int_{u_i}^{\lambda_i} e^{-x^2}\,dx} \to c. \tag{2}$$

Clearly from (1)

$$\sum_{i=0}^{k-1} \frac{1}{\pi^{1/2}} \int_{u_i}^{\lambda_i} e^{-x^2}\,dx = \frac{1}{\pi^{1/2}} \int_{-r}^{+r} e^{-x^2}\,dx + o(1) \to 1.$$

Hence finally from (2)

$$D_{c,u,v} = \sum_{i=0}^{k-1} D_i + o(1) \to c, \qquad \text{q.e.d.}$$

LEMMA 3.

$$\lim D_{c,1,v} = c.$$

PROOF. As in the proof of Lemma 2 we can show that if v tends to infinity sufficiently quickly

$$\lim D_v(a_i) = \frac{c}{a_i} \prod_{p \leq u} \left(1 - \frac{1}{p}\right).$$

Now trivially

$$\sum_i \frac{1}{a_i} = \frac{1}{\prod_{p \leq u} \left(1 - \frac{1}{p}\right)}.$$

Thus we obtain

$$\lim_{v \to \infty} D_{c,1,v} = \lim \sum_{a_i} D_v(a_i) = c,$$

which completes the proof of Lemma 3.

From now on the proof is practically identical with that of I. Let t_v tend to infinity sufficiently slowly and put $n = v^{t_v}$.

LEMMA 4. *The number of integers $m \leq n$ for which $K_v(m) = a_i < v^{\sqrt{t_v}} = n^{1/\sqrt{t_v}}$ equals*

$$(1 + o(1)) \frac{e^{-\gamma} n}{a_i \log v}.$$

PROOF. This is Lemma 3 of I. (γ is Euler's constant).

LEMMA 5. *The number of integers $m \leq n$ for which $K(m) > n^{1/\sqrt{t_v}}$ is $o(n)$.*

PROOF. Lemma 4 of I.

LEMMA 6.

$$\lim_{n\to\infty} \frac{1}{n} A(1, v, c, n) = c.$$

PROOF. Follows from Lemmas 3, 4, and 5 as in I Lemma 5 followed from Lemmas 1, 3, and 4.

LEMMA 7. *Let $t_v \to \infty$ sufficiently slowly then for every ϵ and η and sufficiently large n the number of integers $m \leqq n$ for which*

$$| f_{(1,v)}(m) - f(m) | > \epsilon$$

is less than ηn.

PROOF. We evidently have

$$\sum_{m=1}^{n} | f_{1,v}(m) - f(m) | \leqq n \sum_{p \geqq v}^{n} \frac{| f(p) |}{p} \tag{3}$$
$$\leqq n \max_{v \leqq p \leqq n} | f(p) | \sum_{p=v}^{n} \frac{1}{p} < 2n \max_{v \leqq p \leqq n} | f(p) | \log t_v < \epsilon\eta n$$

since $f(p) \to 0$ only if $t_v \to \infty$ sufficiently slowly. Clearly (3) implies Lemma 7.

Now we can prove Theorem I. It follows from Lemma 6 that it suffices to show that for every η the number of integers $m \leqq n$ for which

$$f(m) \leqq c \qquad \text{and } f_{(1,v)}(m) > c \qquad (\text{mod } 1) \tag{4}$$

or

$$f(m) > c \qquad \text{and } f_{(1,v)}(m) \leqq c \qquad (\text{mod } 1) \tag{5}$$

is $< 2\eta n$ for sufficiently large n. We split the integers satisfying (4) or (5) into two classes. In the first class are the integers for which

$$| f(m) - f_{1,v}(m) | > \epsilon.$$

By Lemma 7 the number of integers of the first class is $< \eta n$. For the integers of the second class we have

$$c - \epsilon \leqq f_{1,v}(m) \leqq c + \epsilon.$$

Hence by Lemma 6, the number of the integers of class 2 is $\leqq \eta n$ too, (i.e. these numbers belong to $A(1, v, c + \epsilon, n) - A(1v, c - \epsilon, n)$) and this completes the proof of Theorem I.

Now we prove Theorem V. First we prove that the condition of Theorem V implies that the values of $f(m)$ are finitely distributed. Put

$$f(p) = c \log p + f^{+}(p), \qquad \sum \frac{(f^{+}(p))^2}{p} < \infty.$$

For simplicity we assumed that $(f^+(p)) \leqq 1$ and $f^+(p^\alpha) = f^+(p)$ (in other words we do not have to consider $f^{+\prime}(p)$). Put

$$f^+(m) = \sum_{p \mid m} f^+(p).$$

Denote

$$A_n = \sum_{p \leqq n} \frac{f^+(p)}{p}.$$

We have

$$\sum_{m=1}^{n} (f^+(m) - A_n)^2 = \sum_{m=1}^{n} (f^+(m))^2 - 2A_n \sum_{m=1}^{n} f^+(m) + nA_n^2.$$

Now

$$\sum f^+(m) = \sum_{p \leqq n} \left[\frac{n}{p}\right] f^+(p) = n \sum_{p \leqq n} \frac{f^+(p)}{p} + 0\ (\sum_{p \leqq n} f^+(p)) = nA_n + 0\left(\frac{n}{\log n}\right).$$

Since $(f^+(p)) \leqq 1$. Further

$$\sum_{m=1}^{n} (f^+(m))^2 = \sum_{\substack{pq \leqq n \\ p \neq q}} \left[\frac{n}{pq}\right] f^+(p) f^+(q) + \sum_{p \leqq n} \left[\frac{n}{p}\right] (f^+(p))^2$$

$$= \sum_{\substack{pq \leqq n \\ p \neq q}} \frac{n}{pq} f(p) f(q) + n \sum_{p \leqq n} \frac{(f^+(p))^2}{p} + 0\left(\frac{n \log \log n}{\log n}\right)$$

$$= n\left(\sum_{p \leqq n} \frac{f^+(p)}{p}\right)^2 - n \sum_{p > \sqrt{n}} \frac{f^+(p)}{p} \sum_{n/p < q \leqq n} \frac{f^+(q)}{q} + 0(n).$$

Now

$$\sum_{p > \sqrt{n}} \frac{f^+(p)}{p} \sum_{n/p < q \leqq n} \frac{f^+(q)}{q} = \sum\nolimits_1 + \sum\nolimits_2 + \cdots$$

where in $\sum_r$

$$n^{1-1/2^r} < p \leqq n^{1-1/2^{r+1}}.$$

For the p in $\sum_r$

$$\sum_{n/p < q \leqq n} \frac{f^+(q)}{q} < \epsilon(\log \log n - \log \log n^{1-1/2^{r+1}}) + o(1) < \epsilon r,$$

since $|f^+(q)| < c$ and from $\sum (f^+(q))^2/q < \infty$ it follows that for every $\delta > 0$, $\sum_{|f^+(q)| > \delta} |f^+(q)|/q$ converges.
Also

$$\sum_{p = n^{1-1/2^r}}^{n^{1-1/2^{r+1}}} \frac{f^+(p)}{p} < \frac{c}{2^r}.$$

Thus

$$\sum_r < \frac{c\epsilon r}{2^r},$$

hence finally

$$\sum_{p \geqq \sqrt{n}} f^+(p) \sum_{n/p < q \leqq n} f^+(q) < \sum_r \frac{c\epsilon r}{2^r} < \epsilon. \tag{6}$$

Thus

$$\sum_{m=1}^{n} (f^+(m))^2 = nA_n^2 + O(n).$$

Hence

$$\sum_{m=1}^{n} (f^+(m) - A_n)^2 < c_1 n$$

which shows that for every $c_2 < 1$ there exists a c_3 such that for more than $c_2 n$ integers $m \leqq n$

$$| f^+(m) - A_n | < c_3 .$$

Hence for more than $c_2 n$ integers $m \leqq n$

$$| f(m) - c \log n - A_n | < c_4$$

which completes the proof.

The proof of the sufficiency will be very much harder. We have to distinguish several cases:

CASE 1. There exist two sequences of primes p_i, q_i, with

$$\lim p_i/q_i = c, 1 < c < \infty, f(p_i) - f(q_i) \to \infty, \sum 1/p_i = \sum 1/q_i = \infty.$$

We shall show that in Case 1 $f(m)$ is not finitely distributed. Assume that $a_1 < a_2 < \cdots < a_x \leqq n$, $| f(a_i) - f(a_j) | < c_1$. We shall show $x = o(n)$.

Choose $k_1 = k_1(\epsilon)$ large. Then the number of integers $\leqq n$ divisible by a $p_i q_i$, $p_i > k_1$ does not exceed

$$\sum_{p_i > k_1} \frac{n}{p_i q_i} > n \sum_{i > k_1} \frac{1}{i^2} < \epsilon n.$$

Let r be a large number which will be determined later. Choose l_1 so that

$$r < \sum_{p_i \epsilon (k_1, l_1)} \frac{1}{p_i} < r + 1, \min f(p_i) - f(q_i) > c_1 .$$

Suppose the interval (k_{i-1}, l_{i-1}) has already been defined. We define (k_i, l_i) as follows:

$$\min_{p_i, q_i \epsilon (k_i, l_i)} (f(p_i) - f(q_i)) > 2 \max_{p_i, q_i \epsilon (k_{i-1}, l_{i-1})} (f(p_i) - f(q_i)), \tag{7}$$

$$r < \sum_{p_i \epsilon (k_i, l_i)} \frac{1}{p_i} < r + 1.$$

Clearly these conditions can be easily fulfilled. Of course we have

$$cr(1-\epsilon) < \sum \frac{1}{q} < c(r+1)(1+\epsilon)$$

ϵ denotes a small number which can be made arbitrarily small, and which does not have to be the same in different cases.

Let $i \leqq j$ where j will be determined later. Denote respectively by $d_{p,i}(x)$ and $d_{q,i}(x)$ the number of divisors of x among the p's and among the q's in (k_i, l_i). It follows from the method of Turán[9] that if r is sufficiently large then for all integers $m \leqq n$, with the possible exception of ϵn integers

$$r(1-\epsilon) < d_{p,i}(m) < r(1+\epsilon);\; cr(1-\epsilon) < d_{q,i}(m) < cr(1+\epsilon).$$

Let now $a_1 < a_2 < \cdots < a_x \leqq n$ be a sequence of integers with

$$[f(a_u) - f(a_v)] < c_1 \text{ or } D < f(a_u) < D + c_1 .$$

Assume that no a_u is divisible by a $p_k q_k$ and that for $i \leqq j$

$$(8) \qquad r(1-\epsilon) < d_{p,i}(a_u) < r(1+\epsilon);\; cr(1-\epsilon) < d_{q,i}(a_u) < cr(1+\epsilon).$$

The number of a's which do not satisfy this condition is less than $(j+1)\,\epsilon n$. Now we define the new sequence $b_1^{(i)} \cdots b_{y_i}^{(i)}$ $i = 1, 2, \cdots j$ as follows: Let p be any prime in (k_i, l_i). If $p \mid a_u$, consider all the integers of the form $(a_u/p)q$. Thus we obtain the sequence $b_1^{(i)}, \cdots b_{y_i}^{(i)}$. All the b's are clearly $\leqq ((1/c) + \epsilon)n$. Next we show

$$(9) \qquad y_i > x\left(\frac{1}{c} - \epsilon\right),\; i = 1, 2, \cdots j.$$

From (8) it follows that to each a_u correspond at least $r(1-\epsilon)$, $b_v^{(i)}$'s and each $b_v^{(i)}$ occurs at most $cr(1+\epsilon)$ times. This proves (9).

Now we prove that these j sequences are disjoint. In other words

$$b_u^{(i_1)} \neq b_v^{(i_2)}, \qquad i_1 < i_2 \leqq j, \qquad u \leqq y_{i_1}, \qquad v \leqq y_{i_2}.$$

It will suffice to show that

$$f(b_u^{(i_1)}) \neq f(b_v^{(i_2)}).$$

Clearly

$$b_u^{(i_1)} = a_u \frac{q_1}{p_1}, \qquad p_1 \epsilon (k_{i_1}, l_{i_1}), \qquad b_v^{(i_2)} = a_v \frac{q_2}{p_2}, \qquad p_2 \epsilon (k_{i_2}, l_{i_2}).$$

Thus

$$f(b_v^{i_2}) < D + c_1 - \min_{p \epsilon (k_{i_2}, l_{i_2})} (f(p) - f(q)),$$

$$f(b_u^{(i_1)}) > D - \max_{p \epsilon (k_{i_1}, l_{i_1})} [f(p) - f(q)].$$

[9] Ibid. (1934) Vol. 9, p. 274–276. See also ibid (1936) Vol. 11, p. 125–133.

Now by definition

$$\min_{p\in(k_{i_2},l_{i_2})}(f(p)-f(q)) > 2\max_{p\in(k_{i_1},l_{i_1})}(f(p)-f(q)) > \max_{p\in(k_{i_1},l_{i_1})}(f(p)-f(q))+c_1.$$

Thus

$$f(b_v^{(i_2)}) > f(b_u^{i_1}) \quad i.e. \quad b_v^{i_2} \neq b_u^{i_1}.$$

Hence

$$y_1+y_2+\cdots y_j < \left(\frac{1}{c}+\epsilon\right)n < n.$$

Thus

$$jx\left(\frac{1}{c}-\epsilon\right) < n, \qquad x < \frac{2cn}{j} < \epsilon n$$

for sufficiently large j, this completes the proof of case 1.

Clearly the case $f(p_i) - f(q_i) \to -\infty$ can be dealt with similarly.

CASE II. Next suppose that there exist sequences p_i, q_i with $\lim p_i/q_i = c$, $1 < c < \infty$ $f(p_i) - f(q_i) \to 0$, $\sum 1/p_i = \sum 1/q_i = \infty$.

Let $a_1^{(1)} < a_2^{(1)} < \cdots < a_{x_1}^{(1)} \leqq n$, n large, $x_1 > c_2 n$, $|f(a_i) - f(a_j)| < c_1$. We will show that this assumption leads to a contradiction. In other words $f(n)$ is not finitely distributed.

Define (k_1, l_1) as before. We can assume as before that a_u divides no $p_i q_i$ and that

$$r(1-\epsilon) < d_{p,1}(a_u) < r(1+\epsilon);\ cr(1-\epsilon) < d_{q,1}(a_u) < cr(1+\epsilon).$$

As previously if $p_i \mid a_u$ we consider $(a_u/p_i)q_i$, $p_i \in (k_1, l_1)$. Thus we obtain a new sequence of integers $a_1^{(2)} < a_2^{(2)} < \cdots < a_{x_2}^{(2)}$ where as in case 1 $x_2 > x_1(1/c - \epsilon)$ and all the $a_i^{(2)}$ are less than $n(1/c + \epsilon)$. Repeat the same process for the $a_i^{(2)}$ etc. We repeat this operation j times (j large) and order all the $a_i^{(r)}$, $i \leqq x_r$, $r \leqq j$ in a sequence $b_1 < b_2 < \cdots < b_s < n$. Clearly for every $m > n/c^j$ the number of the b's $\leqq m$ is $> c_3 m$. Also since $f(p_i) - f(q_i) \to 0$ we have (for sufficiently large k_1)

$$|f(b_i) - f(b_j)| < 2c_1.$$

We can assume that

$$\sum_{|f(t)|>2c_1}\frac{1}{t} = \infty \quad (t \text{ prime}).$$

For if not, since $\sum (f'(p))^2/p$ diverges, we would immediately obtain from I that $f(m)$ is not finitely distributed.

Without loss of generality we can thus assume that $\sum_{f(t)>2c_1} 1/t = \infty$. Let A be large (its dependence on c_3 will be indicated later). We choose j so large that

$$\sum_{\substack{f(t)>2c_1\\ t<c^j}}\frac{1}{t} > A.$$

This choice of j is obviously possible, if we only took r to be sufficiently large. Since $f(t) > 2c_1$ the quotient of two b's can not be an integer composed entirely of the t's. We will show that this leads to a contradiction. The proof will be similar to that used in a previous paper.[10]

Denote by $d(m)$ the number of divisors of m among the b's greater than n/c^j. Clearly

$$\sum_{m=1}^{n} d(m) = \sum_{b_i > n/c^j} \left[\frac{n}{b_i}\right] > c_4 jn,$$

where c_4 depends on c_2. The quotient of two b's is never composed only of the t's therefore the quotient of two integers of the form m/b_i is also never composed entirely of the t's. Also $m/b_i < c^j (m \leqq n)$. Denote now by $d^+(m)$ the maximum number of divisors of m which are $\leqq c^j$ and the quotient of no two is entirely composed of t's. We clearly have

$$d(m) \leqq d^+(m).$$

Thus

$$\sum_{m=1}^{n} d^+(m) \geqslant c_4 jn. \tag{10}$$

From here on the proof follows very closely that of Lemma 2 of my paper "On the density of some sequences of numbers", London Math. Soc. Journal, (1937) Vol. XII p. 9–10, so that it will be sufficient to give only an outline of the argument.

We have to show that (10) is false. Put

$$\sum_{m=1}^{n} d^+(m) = \sum\nolimits_1 + \sum\nolimits_2$$

where $\sum_1$ is extended over the m less than A divisors among the t's, and $\sum_2$ is extended over the other m's. On p. 10 of the above article I prove that for sufficiently large A

$$\sum_1 < \epsilon jn, \qquad \sum_2 < \epsilon jn$$

which proves that (10) is false. This contradiction shows that $f(m)$ is not finitely distributed in case 2.

Case 3. There exists a sequence p_i and q_i with $p_i/q_i \to c$, $1 < c < \infty$ $f(p_i) - f(q_i) \to d$, $1 < d < \infty$, $\sum 1/p_i = \sum 1/q_i = \infty$.

Suppose that $f(m)$ is finitely distributed, then

$$\varphi(m) = f(m) + \frac{d}{\log c} \log m$$

is clearly also finitely distributed, and

$$\varphi(p_i) - \varphi(q_i) \to 0.$$

[10] London Math. Soc. Journal (1937) Vol. 9, p. 7–11.

If $\sum(\varphi(p'))^2/p = \infty$ then we are back in case 2 and $\varphi(m)$ cannot be finitely distributed. Thus

$$\sum \frac{\varphi(p')^2}{p} < \infty$$

which completes the proof of Theorem V.

Now we prove Theorem IV. If the conditions of Theorem IV are satisfied $f(m)$ is certainly finitely distributed, thus there exists a constant c such that

$$f(m) = c \log m + \varphi(m), \sum \frac{\varphi'(p)^2}{p} < \infty.$$

Assume first $c = 0$. Then Theorem IV would follow from Theorem II, but since we suppressed the proof of Theorem II we will give the proof of this case. Let then $f(n)$ be an additive function with

$$\sum \frac{(f'(p))^2}{p} < \infty. \tag{11}$$

We shall show that to every ϵ there exists a δ such that if $a_1 < a_2 < \cdots < a_x \leqq n$ is such that

$$D < f(a_i) < D + \delta \tag{12}$$

then $x < \epsilon n$, and this will prove Theorem IV for $c = 0$.

The proof follows very closely the argument used in my paper "On the density of some sequences of numbers III" London Math. Soc. Journal (1938) Vol. 13 p. 119–127. Put

$$f_k(m) = \sum_{\substack{p \mid m \\ p \leqq p_k}} f(p), \ k \text{ large}.$$

Let $f(m)$ satisfy (11). We assume for sake of simplicity that $| f(p) | < c$.

Lemma 8. *The number of integers $m \leqq n$ for which*

$$\left| f(m) - f_k(m) - \sum_{p > p_k}^{p \leqq n} \frac{f(p)}{p} \right| > \delta$$

is less than ϵn for k sufficiently large $[\delta = \delta(\epsilon), k = k(\epsilon)]$.

We have

$$\sum_{m=1}^{n} \left(f(m) - f_k(m) - \sum_{p > p_k}^{p \leqq n} \frac{f(p)}{p} \right)^2$$

$$= \sum_{m=1}^{n} | f(m) - f_k(m) |^2 - 2A_{k,n} \left| \sum_{m=1}^{n} f(m) - f_k(m) \right| + nA_{k,n}^2,$$

$$\left(A_{k,n} = \sum_{p > p_k}^{n} \frac{f(p)}{p} \right).$$

Now

$$\sum_{m=1}^{n} | f(m) - f_k(m) | = \sum_{p > p_k}^{p \leq n} \left[\frac{n}{p}\right] f(p) = nA_{k,n} + o(n).$$

and by (11) and (6) for k sufficiently large

$$\begin{aligned}\sum_{m=1}^{n} | f(m) - f_k(m) |^2 &= \sum_{p > p_k}^{p \leq n} \left[\frac{n}{pq}\right] f(p)f(q) + \sum_{p > p_k}^{p \leq n} \left[\frac{n}{p}\right] (f(p))^2 \\ &\leqq \sum_{\substack{p_k < p, q; pq \leq n \\ p \neq q}} \frac{n}{pq} f(p)f(q) + \eta n + o(n) \\ &= n \left[\sum_{p > p_k}^{p \leq n} \frac{f(p)}{p}\right]^2 - n \sum_{p > \sqrt{n}} \frac{f(p)}{p} \sum_{n/p < q \leq n} \frac{f(p)}{p} + \eta n + o(n) \\ &= nA_{k,n}^2 + 2\eta n + o(n) (\eta \to 0 \text{ as } k \to \infty)\end{aligned}$$

Thus

$$\sum_{m=1}^{n} (f(m) - f_k(m) - A_{k,n})^2 < 2\eta n + o(n)$$

which proves the lemma (with $\epsilon = 3\eta/\delta^2$).

LEMMA 9. *The number of integers with*

$$c - \eta < f_k(m) < c + \eta$$

is $< \epsilon n$ for large k.

The proof follows Lemma 2 of "Density III" London Math. Soc. Journal Vol. 13 (1930) p. 124.

We now split the integers satisfying (12) into two classes. In class I are the integers with $D - A_{k,n} - \delta < f_k(m) < D - A_{k,n} + \delta$. And in the second class are the other integers. By Lemma 9, the number of integers of the first class is $< \epsilon n$, and by Lemma 8 the same holds for the integers in class II, this completes the proof.

REMARK. The above proof shows that in Theorem II the distribution function is continuous if it exists. This is the most difficult part of the proof of Theorem II.

Assume next $c \neq 0$. We then have

$$f(m) = c \log m + \varphi(m), \qquad \sum \frac{(\varphi'(p))^2}{p} < \infty.$$

Thus $\varphi(p) \to 0$ except for a sequence q, with $\sum 1/q < \infty$.

We prove the following Lemma: *Let p be the primes where possibly a sequence q with $\sum 1/q < \infty$ has been omitted. Then to every c_1 there exists a $c_2 > 1$ such that if $a_1 < a_2 < \cdots < a_x \leqq n$, $x > c_1 n$, n sufficiently large, then there exist a_i, a_j, p_i, p_j with*

$$\frac{a_i}{p_i} = \frac{a_j}{p_j}, \qquad \frac{a_i}{p_i} \not\equiv 0 \pmod{p_i}, \qquad \frac{a_j}{p_j} \not\equiv \bmod p_j, \qquad \frac{p_i}{p_j} < c_2.$$

Choose k large. Then there clearly are at least $nc_1/2$ integers a_i not divisible by any $p_i^2 > k$, and $q > k$, we consider only these a's. Choose l large

$$\sum_{k<p<l} \frac{1}{p} = \log\log l - \log\log k + 0(1).$$

We can assume (by the method of Turán)[9] that for at least $nc_1/4$ a's, for sufficiently large l

$$(13) \qquad (1-\epsilon)\log\log l < \quad \nu_p(a_i) <, \qquad (1+\epsilon)\log\log l$$

where $\nu_p(a_i)$ denotes the number of divisors of a_i among the p in (k, l). Clearly we can also assume that for all but $nc_1/8$ of the a's, $c_3 n < a_i$ (c_3 small). We consider only the a_i satisfying these conditions. Let us denote them by $a_1 < a_2 c \cdots < a_{x_1}$, $x_1 > nc_1/8$.

Consider all the integers of the form

$$\frac{a_i}{p}, \qquad k < p < l, \qquad a_i \equiv 0 \pmod{p}.$$

Denote these integers by $b_1, b_2, \cdots b_y$. Clearly

$$c_3 \frac{n}{l} < b_i < \frac{n}{k}.$$

Suppose that if we have

$$\frac{a_i}{p_i} = \frac{a_j}{p_j}, \qquad p_i < p_j < (1+\epsilon)p_j .$$

Then clearly $b_i p$ can be an a only if p lies in some interval $(u_{b_i}, u_{b_i}(1+\epsilon))$, $u_{b_i}(1+\epsilon) < n/b_i$.

Thus

$$\sum_{i=1}^{x_1} \nu_p(a_i) < \sum_{i=1}^{y} \pi[u_{b_i}(1+\epsilon)] - \pi(u_{b_i}) < c\epsilon \sum_{i=1}^{y} \frac{n}{b_i \log n/b_i}$$

$$< c\epsilon \sum_{z=c_3(n/l)}^{n} \frac{n}{z \log \frac{n}{z}} < c_2 \epsilon \log\log l,$$

since a simple calculation shows that

$$\sum_{z=n/u}^{u} \frac{n}{z \log \frac{n}{z}} \sim \log\log u.$$

But by (13)

$$\sum_{i=1}^{x_1} \nu_p(a_i) > \frac{c_1}{16} \log\log l.$$

This contradiction establishes the lemma.

Assume now that (12) holds with $x > cn$. By our lemma we obtain

$$\frac{a_i}{p_i} = \frac{a_j}{p_j}, \qquad p_i > (1 + c_1)p_j, \qquad \frac{a_i}{p_i} \not\equiv 0 \pmod{p_i},$$

$$\frac{a_j}{p_j} \not\equiv 0 \pmod{p_j} \, |\varphi(p_i)| < \epsilon, \, |\varphi(p_j)| < \eta.$$

But then we have

$$f(a_j) - f(a_i) = f(p_i) - f(p_j)$$
$$= c(\log p_i - \log p_j) + \varphi(p_i) - \varphi(p_j) > c \log (1 + c_1) - 2\eta > \delta,$$

which contradicts (12). This completes the proof of Theorem IV.

REMARK: It is easy to see that our lemma would not hold for every sequence of primes with $\sum 1/p_i = \infty$.

We can state Theorem IV roughly as follows: *If* $\sum_{f(p) \neq 0} 1/p = \infty$ *the distribution function tries to be continuous whether it exists or not.*

We state a few results without proof.

THEOREM IX. *Let* $f(m)$ *be additive.* $|f(p)| < c$, $\sum f^2(p)/p = \infty$. *K any given number. Then* $f(x)$, $f(x + 1)$, $\cdots f(x + k)$ *are independent and have Gaussian distribution.*

The proof is very similar to that of I.

THEOREM X. *Assume that*

$$f(m) = c \log m + \varphi(m), \qquad \sum \frac{(\varphi'(p))^2}{p} < \infty.$$

Then $g(m) = f(m + 1) - f(m)$ *has a distribution function. The distribution function is continuous if and only if* $\sum_{\varphi(p) \neq 0} 1/p = \infty$ *and then the integers with* $g(m) \geqq 0$ *have density* $1/2$.

If $\sum_{\varphi(p) \neq 0} 1/p < \infty$, denote by $D(x)$ the distribution function of $g(x)$. Then

$$\lim_{x \to 0, x > 0} D(x) = \lim_{x \to 0, x < 0} D(x).$$

$D(x) = 0$ for $x < 0$ holds if and only if $\varphi(m) = 0$, $f(m) = c \log m$.

We omit the proof since it is similar to that used in a previous paper.[11]

Now we prove

THEOREM XI. *Assume* $f(m + 1) \geqq f(m)$. *Then* $f(m) = c \log m$.

Let m be odd, $m < n < 2n$. Then

$$f(m) \leqq f(n) \leqq f(2n) = f(2) + f(n).$$

Thus $f(m)$ is finitely distributed. Hence

$$f(m) = c \log m + \varphi(m), \qquad \sum \frac{\varphi'(p)^2}{p} < \infty.$$

---

[11] Ibid. (1938) Vol. 13, p. 119–127.

Then by Theorem X if $\varphi(m)$ is not identically 0, we have for infinitely many m,

$$\varphi(m+1) - \varphi(m) < -\delta$$

which contradicts $f(m+1) \geqq f(m)$. Thus $\varphi(m) = 0$ for all m, which completes the proof.

By a much more complicated argument we can prove THEOREM XII. *Let $f(m)$ be additive and assume that there exists a constant c_1 and infinitely many n_i such that for every n_i there exists a sequence $a_1 < a_2 < \cdots < a_x \leqq n_i$, with $f(a_1) \leqq f(a_2) \leqq \cdots \leqq f(a_x)$, $x > c_1 n$. Then*

$$f(n) = c \log n + \varphi(n), \quad \sum_{\varphi(p) \neq 0} \frac{1}{p} < \infty.$$

It is easy to see that the converse of Theorem XII is true.

We do not give the proof of Theorem XII.

THEOREM XIII. *Let $f(m)$ be additive, $f(m+1) - f(m) \to 0$, then $f(m) = c \log m$.*

Denote by $P_1 < P_2 < \cdots$ the primes and their powers. Put

$$c = \limsup \frac{f(P_i)}{\log P_i}.$$

Denote by $p_1 < p_2 < \cdots$ the primes. First we assume that for infinitely many primes p_i there exists an α_i such that

$$\frac{f(p_i^{\alpha_i})}{\log p_i^{\alpha_i}} > c.$$

Put $p_i^{\alpha_i} = Q_i$ and order the Q_i by

$$\frac{f(Q_1)}{\log Q_1} \geqq \frac{f(Q_2)}{\log Q_2} \geqq \cdots$$

since $\lim f(Q_i)/\log Q_i = c$ and $f(Q_1)/\log Q_1 > c$ somewhere we must have an inequality, assume that the first inequality occurs for Q_j. Put

$$N_k = Q_1 \cdot Q_2 \cdots Q_k, \qquad k > j, \qquad c_1 = \frac{f(Q_1)}{\log Q_1},$$

$$c_k = \frac{f(Q_k)}{\log Q_k}, \qquad c < c_k < c_1 \quad \text{for } k \geqq j.$$

We evidently have

$$f(N_k) - f(N_k - 1) \geqq c_k \log N_k + (c_1 - c_2) \log Q_1$$

$$- c_k \log (N_k - 1) > \delta, \text{ for a fixed } \delta > 0$$

which contradicts $f(m+1) - f(m) \to 0$.

Assume next that there are only a finite number (perhaps none) of p_i with

$$\frac{f(p_i^{\alpha_i})}{\log p_i^{\alpha_i}} > c.$$

Denote these primes by $p_1, p_2, \cdots p_j, p_i^{\alpha_i} = Q_i, i = 1, 2, \cdots j$. Assume first $j \neq 0$. In this case we clearly have

$$-\infty < c < \infty.$$

Put $N_j = Q_1Q_2 \cdots Q_j$. Define n_i by

$$(n_i, N_j) = 1, \qquad \frac{f(n_i)}{\log n_i} \geqq \frac{f(m)}{\log m} \text{ for all } m < n_i, \qquad (m, N_j) = 1.$$

Clearly there exist infinitely many such n_i. Let n_i be large. Let r be the least prime $> N_j$. Choose u so that

$$n_iN_j - u \equiv 0 \pmod{r}, \text{ but } n_iN_j - u \equiv 0 \pmod{r^2}.$$

Clearly $u < 2r$. Put

$$c_i = \frac{f(n_i)}{\log n_i}, \qquad c_i \leqq c.$$

Also $n_iN_j - u = rx$, $x < n_i$, $(r, x) = 1$, hence clearly

$$f(rx) = f(r) + f(x) \leqq c_i \log rx.$$

Hence finally

$$f(n_iN_j) - f(n_iN_j - u) > \delta, \text{ for some fixed } \delta > 0.$$

This contradicts $f(n + 1) - f(n) \to 0$.

Thus we can assume that for all Q_i,

$$\frac{f(Q_i)}{\log Q_i} \leqq c. \quad (c = +\infty \text{ is possible now}).$$

Define n_i by

$$\frac{f(n_i)}{\log n_i} \geqq \frac{f(m)}{\log m}, \qquad m \leqq n_i.$$

Now for the last time we again distinguish two cases. First there exists a Q_i with

$$\frac{f(Q_i)}{\log Q_i} < c.$$

Clearly $\limsup (f(n_i))/\log n_i = c$. Put $Q_i = q_i^\alpha$ determine u by $n_i - u \equiv 0 \pmod{Q_i}$, $n_i - u \not\equiv 0 \pmod{q^{\alpha+1}}$, we can choose u to be $< 2Q_i$. A simple computation shows that if n_i is large

$$f(n_i) - f(n_i - u) > \delta, \text{ for a fixed } \delta > 0,$$

this contradicts $f(n + 1) - f(n) \to 0$.

Thus for all Q_i, $f(Q_i))/\log Q_i = c$, or $f(m) = c \log(m)$; this completes the proof.

It seems likely that if

$$\frac{1}{n} \sum_{m=1}^{n} | f(m + 1) - f(m) | \to 0$$

$f(m) = c \log m$.

THEOREM XIV. *Let $f(n)$ be additive and suppose that*

$$M(m, n) - M(1, n) \to 0, n \to \infty, m \to \infty, n - m \to \infty$$

then $f(n) = c \log n + \varphi(n)$, $\varphi(n)$ uniformly bounded.

If $f(n)$ satisfies the condition of the theorem we certainly must have

$$|f(n) - M(1, n)| < c \tag{14}$$

Let n_i satisfy

$$f(n_i) \geqq f(m) \text{ for } m \leqq n_i .$$

We can assume that there exist infinitely many such n_i (if need be, we replace $f(m)$ by $-f(m)$). We obtain from (14) that there exist $c_1 n_i$ integers $m \leqq n_i$ with $f(m) \geqq f(n_i) - c_2$. Thus $f(m)$ is finitely distributed, hence

$$f(n) = c \log n + \varphi(n), \qquad \sum \frac{(\varphi'(p))^2}{p} < \infty,$$

and a simple calculation shows that $\varphi(n)$ must be uniformly bounded.

Before concluding we consider additive functions which are not necessarily real valued. We state without proof the following results:

THEOREM IV′. *Let $f(m)$ be a complex valued additive function such that $\sum_{f(p) \neq 0} 1/p$ diverges. Then to every ϵ there exists a δ such that if $a_1 < a_2 < \cdots < a_x \leqq n$ is a sequence of integers with $(|f(a_i)| - |f(a_j)|) < \epsilon$ then $x < \delta n$ for n sufficiently large.*

THEOREM V′. *Let $f(m)$ be additive, complex valued, and such that there exist two constants c_1 and c_2 and infinitely many n, so that there exist $a_1 < a_2 < \cdots < a_x \leqq n$, $x > c_1 n$, $(|f(a_1)| - |f(a_j)|) < c_2$. Then there exists a constant c such that if we write*

$$f^+(p) = f(p) - c \log p, f^+(p) = g(p) + ih(p)$$

we have

$$\sum_p \frac{(g'(p))^2}{p} < \infty, \qquad \sum_{p \leqq n} \frac{(h'(p))^2}{p} < c_3 \sum_{p \leqq n} \frac{g'(p)}{p}. \tag{15}$$

We can of course interchange $g(p)$ and $h(p)$.

It can be shown that if (15) is satisfied, the condition of Theorem V is also satisfied, in other words (15) is a necessary and sufficient condition.

UNIVERSITY OF MICHIGAN.