ADDITIVE ARITHMETIC FUNCTIONS BOUNDED BY MONOTONE FUNCTIONS ON THIN SETS

By

P. D. T. A. ELLIOTT* and P. ERDŐS

University of Colorado, Boulder

(Received August 2, 1978)

1. An arithmetic function $f(n)$ is said to be *additive* if it satisfies the relation $f(ab) = f(a)+f(b)$ for every pair of coprime positive integers a, b.

In this paper we establish two results to the effect that an additive function which is not too large on many integers cannot often be large on the primes.

If $a_1 < a_1 < \ldots$ is a sequence of positive integers, let $A(x)$ denote the number of such integers which do not exceed x.

THEOREM 1. *Let the additive arithmetic function $f(n)$ satisfy*

$$c_1 \le \frac{|f(a_j)|}{g(a_j)} \le c_2 \tag{1}$$

on a sequence of positive integers $a_1 < a_2 < \ldots$ which satisfies $A(x) \ge \alpha x$ for all $x \ge c_3$. Here $\alpha > 1/2$, c_1, c_2 and c_3 are positive constants, the function $g(x)$ is positive non-decreasing, defined for all real $x \ge 2$.

Then there are further constants B and c so that

(i) $g(x^2) \le Bg(x)$,

(ii) $\displaystyle\sum_{p \le x} \frac{1}{p} \left\| \frac{f(p)}{g(x)} \right\|^2 \le c$

uniformly for all $x \ge 2$. Here

$$\|y\| = \begin{cases} 1 & \text{if} \quad |y| > 1\,, \\ y & \text{if} \quad |y| \le 1\,. \end{cases}$$

* Partially supported by N. S. F. Contract Number MCS 78–04374.

Note that by simple inductive argument $g(x) = O((\log x)^{c_4})$ for some constant $c_4 > 0$, and (ii) then asserts that $f(p)$ is usually $O((\log x)^{c_4})$ so long as p does not exceed x.

The following variant of Theorem 1 has a weaker hypothesis but a weaker conclusion.

THEOREM 2. *Let the definitions of theorem* 1 *be in force save that hypothesis* (1) *is to be replaced by*

$$|f(a_j)| \leq g(a_j). \tag{2}$$

Then conclusion (i) *need not be valid, but there are constants* c, $\beta > 1$ *so that the inequality*

(iii) $\displaystyle\sum_{p \leq x} \frac{1}{p} \left\| \frac{f(p)}{g(x^\beta)} \right\|^2 \leq c$

holds for all $x \geq 2$.

In particular if

$$g(x) = \varepsilon + \max_{n \leq x} |f(n)| > 0$$

then inequality (iii) shows that

$$\sum_{p \leq x} \frac{|f(p)|^2}{p} \leq c g(x^\beta)^2,$$

and letting $\varepsilon \to 0^+$,

$$\sum_{p \leq x} \frac{|f(p)|^2}{p} \leq c \max_{n \leq x^\beta} |f(n)|^2$$

uniformly for $x \geq 2$. Replacing x by $x^{1/\beta}$ we see that so long as $x \geq 2^\beta$ we have

$$\sum_{p \leq x} \frac{|f(p)|^2}{p} \leq \max_{n \leq x} |f(n)|^2 \left(c + \sum_{x^{1/\beta} < p \leq x} \frac{1}{p} \right).$$

For $2 \leq x < 2^\beta$ we obtain a similar result with the sum of the p^{-1} on the right-hand side taken over the range $p \leq 2^\beta$. In either case we reach the

COROLLARY. *There is a positive constant* A *so that every additive arithmetic function* $f(n)$ *satisfies*

$$\sum_{p \leq x} \frac{|f(p)|^2}{p} \leq A \max_{n \leq x} |f(n)|^2.$$

We shall show that the constant A in this corollary may in fact be given an absolute value.

The function $f(n) = \log n$ shows that apart from the value of A this result is best possible.

By taking for $g(x)$ a suitable constant function we obtain from theorem 2 a new proof of the proposition that if an additive function is bounded on a sequence of positive lower density then the series

$$\sum_{|f(p)|>1} \frac{1}{p}, \quad \sum_{|f(p)|\leqq 1} \frac{f^2(p)}{p}$$

both converge.

By setting $g(x) = \varepsilon > 0$ and then letting ε approach zero we see that if an additive function vanishes on a sequence of positive lower density then the series

$$\sum_{f(p)\neq 0} \frac{1}{p}$$

converges. For other proofs of these two propositions, see for example, ELLIOTT [5].

As a further result we show that theorem 1 does not continue to hold if one assumes only that $A(x) \geqq \alpha x$ where $\alpha < 1/2$ is allowed.

THEOREM 3. *Let a real number η be given, $0<\eta<1/2$. Then there is an additive function $f(n)$, a positive non-decreasing function $g(x)$, two positive numbers d_1 and d_2 and a sequence of positive integers $a_1 < a_2 < \ldots$, so that*

(iv) $d_1 < \dfrac{|f(a_j)|}{g(a_j)} < d_2$,

(v) $A(x) \geqq (1/2-\eta)x$

holds for all values of $x \geqq 2$,

(vi) $\limsup_{x\to\infty} \dfrac{g(x^2)}{g(x)} = \infty$.

In fact we shall show that one may arrange that on a suitable unbounded sequence of x-values the ratio $g(x^2)/g(x)$ may increase at least as rapidly as any prescribed rate.

2. For the proof of theorem 1 we need two preliminary results.

LEMMA 1. *The inequality*

$$\sum_{p\leqq x} p \left| \sum_{\substack{n\leqq x \\ p\| n}} d_n - \frac{1}{p} \sum_{n\leqq x} d_n \right|^2 \leqq c_5 x \sum_{n\leqq x} |d_n|^2$$

holds for all complex numbers d_n, and real numbers $x \geqq 2$. Here $p\|n$ means that p divides n but p^2 does not.

PROOF. See ELLIOTT [4], [5].

LEMMA 2. *Let r, x be real numbers, $2 \leq r \leq x$. Define the (strongly-) additive function*

$$f_r(n) = \sum_{p|n,\, p \leq r} f(p),$$

and independent random variables Y_p, one for each prime p not exceeding r, so that

$$Y_p = \begin{cases} f(p) & \text{with probability} \quad \dfrac{1}{p} \\ 0 & \text{with probability} \quad 1 - \dfrac{1}{p} \end{cases}.$$

Then the number of integers n, not exceeding x, for which the inequality $f_r(n) \leq z$ is satisfied may be estimated by

$$xP\left(\sum_{p \leq r} Y_p \leq z\right) + O\left(x \exp\left(-c_6 \frac{\log x}{\log r}\right)\right).$$

Here c_6 and the constant implied in the error term are positive and absolute, and the result holds uniformly in all real z.

PROOF. This result forms the fundament of the ERDÖS–KAC–KUBILIUS method of treating additive arithmetic functions. A proof may be found in KUBILIUS [10].

In order to establish theorem 2 we shall need, besides Lemma 2, the following result.

LEMMA 3. *Let $q_1 < q_2 < \ldots < q_l \leq x$ be a set of prime numbers, and let $n_1 < n_2 < \ldots < n_k \leq x$ be a sequence of positive integers, none of which is divisible by the square of any prime q_i. Define*

$$S = \sum_{i=1}^{l} \frac{1}{q_i}.$$

Suppose that there is no relation of the form $n_i = \lambda n_j$ where the integer λ is made up entirely of primes chosen from amongst the q_i.

Then there is a positive absolute constant c_0 so that the bound

$$S^{1/2} \sum_{r=1}^{k} \frac{1}{n_r} \leq c_0 \log x, \quad x \geq 2$$

is satisfied.

PROOF. A result of this type plays an important rôle in Behrend's treatment of primitive sequences of integers [1], and Erdös' treatment of certain distributional problems concerning additive functions ([8], see also ERDÖS and WINTNER [9]). A proof of lemma 3 may be obtained by modifying that of lemma 5 of ELLIOTT [6].

In theorems 1 and 2 neither the hypotheses nor the conclusions are affected by replacing $f(n)$ and $g(x)$ simultaneously with $\Theta f(n)$ and $\Theta g(x)$ for some real Θ. Without loss of generality we shall therefore assume that $g(2) \geq 2$.

We begin the proof of part (i) of theorem 1, (cf. ELLIOTT [4], [7]).

3. Let x be a real number, $x \geq 2$. Let $b_1 < b_2 < \ldots$, run through the integers n not exceeding x for which one of the inequalities (1) with n in place of a_j fails. Let ε_1, $0 < \varepsilon_1 < 1/2$ be a positive number.

We apply lemma 1 with

$$d_n = \begin{cases} 1 & \textit{if} \quad n = b_j \quad \textit{for some} \quad j, \quad \varepsilon_1 x < n \leq x, \\ 0 & \textit{otherwise}. \end{cases}$$

Then

$$\sum_{p \leq x} p \left| \sum_{\substack{b_j \leq x \\ p \| b_j}} 1 - \frac{1}{p} \sum_{b_j \leq x} 1 \right|^2 \leq c_5 x^2. \tag{3}$$

Let δ be a positive number, to be specified presently, and let $\Lambda = \Lambda(x)$ denote the set of primes p not exceeding x for which the inequality

$$\left| \sum_{\substack{b_j \leq x \\ p \| b_j}} 1 - \frac{1}{p} \sum_{b_j \leq x} 1 \right| > \frac{\delta x}{p}$$

is satisfied. Then according to (3)

$$\sum_{p \leq x,\, p \in \Lambda} \frac{1}{p} < c_5 \delta^{-2}.$$

Consider now a prime p, not exceeding $x^{1/2}$, which does not belong to Λ. Then

$$\sum_{\substack{\varepsilon_1 x < a_i \leq x \\ p \| a_i}} 1 = \sum_{\substack{\varepsilon_1 x < n \leq x \\ p \| n}} 1 - \sum_{\substack{b_i \leq x \\ p \| b_i}} 1 \geq \left[\frac{x}{p}\right] - \left[\frac{x}{p^2}\right] - \left[\frac{\varepsilon_1 x}{p}\right] - \frac{1}{p} \sum_{b_i \leq x} 1 - \frac{\delta_x}{p} \geq$$

$$\geq \frac{x}{p}\left(1 - \frac{1}{\sqrt{x}} - \frac{1}{p} - \varepsilon_1 - (1 - \alpha - \varepsilon_1) - \delta\right) > \frac{x}{p}(\alpha - \varepsilon_2)$$

for any (fixed) $\varepsilon_2 > 0$ provided only that we choose ε_1 and δ to be sufficiently small, x sufficiently large, and p to exceed a certain number depending only upon ε_2. Moreover, the numbers $p^{-1} a_2$ where $p \| a_i$ lie in the interval $(\varepsilon_1 x p^{-}, ^1 x p^{-1}]$.

The number of a_j not exceeding $p^{-1} x$ is

$$A\left(\frac{x}{p}\right) \geq \alpha \frac{x}{p}$$

provided that $x \geq c_3^2$. Since $\alpha+(\alpha-\varepsilon_2)>1$ if ε_2 is fixed at a sufficiently small value (*here we make use of the hypothesis* $\alpha>1/2$) then $p^{-1}a_i = a_j$ must hold for some pair of values i and j, where $\varepsilon_1 x < a_i \leq x$.

In view of the additive property of $f(n)$, and the fact that by construction $(p, a_j) = 1$,

$$f(p) = f(a_i) - f(a_j). \tag{4}$$

In particular

$$|f(p)| \leq 2c_2 g(x). \tag{5}$$

We have now shown that an equation of type (4) holds for all primes p not exceeding $x^{1/2}$ save for a set of primes q which may be estimated by

$$\sum_{q \leq x^{1/2}} q^{-1} \leq c_7 .$$

We can carry out a similar procedure with x replaced by x^2, and obtain

$$f(p) = f(a_i') - f(a_j') \tag{6}$$

where $\varepsilon_1 x^2 < a_i' \leq x^2$, for all primes p not exceeding x save possibly for a set of primes q' which satisfy

$$\sum_{q' \leq x^{1/2}} (q')^{-1} \leq c_7 .$$

Choose a number $\gamma>0$ so small that

$$\sum_{x^{\gamma/2} < p \leq x^{1/2}} = \log\frac{1}{\gamma} + o(1) > 2c_7 .$$

Then there is a prime p lying in the range $x^{\gamma/2} < p \leq x^{1/2}$ so that *both* the equations (4) and (6) hold. Eliminating $f(p)$ from such a pair of equations yields

$$f(a_i) - f(a_j) = f(a_i') - f(a_j'),$$

or, in a more convenient arrangement,

$$f(a_i') = f(a_i) + f(a_j') - f(a_j).$$

Bearing in mind the property (1) of function $f(n)$,

$$c_1 g(a_i') \leq c_2 \left(g(a_i) + g(a_j') + g(a_j)\right),$$

and since $g(\ )$ is non-decreasing

$$c_1 g(\varepsilon_1 x^2) \leq 3c_2 g(x^{2-(\gamma/2)})$$

$$g(x) \leq c_8 g(x^{1-(\gamma/6)})$$

uniformly for $x \geq c_9$. We apply this last inequality k times where the (fixed) integer k is chosen so that $(1-(\gamma/6))^k < 1/2$, to obtain the inequality

$$g(x) \leq c_8^k\, g(x^{1/2}), \quad x \geq c_{10}\,.$$

The validity of inequality (1) in theorem 1 is now clear, with a constant B which depends upon c_1, c_2, α, c_3 and the supremum of $g(x^2)/g(x)$ for $2 \leq x \leq c_{10}^{1/2}$.

4. Returning to the inequality (5) in the previous section we note the estimate

$$\sum_{\substack{p \leq x^{1/2} \\ |f(p)| > 2c_2 g(x)}} \frac{1}{p} \leq c_7\,. \tag{7}$$

which we shall need to establish part (ii) of theorem 1.

Let q run through the primes not exceeding $x^{1/2}$ for which the inequality $|f(q)| > 2c_2 g(x)$ is satisfied, and which are therefore counted in (7). Let μ be a positive number, $0 < \mu < 1$, to be chosen presently.

For any positive integer m_0,

$$\sum_{m=1}^{m_0} \sum_{(\sqrt{x})^{\mu^m} < q \leq (\sqrt{x})^{\mu^{m-1}}} \frac{1}{q} \leq c_7\,.$$

so that if m_0 is fixed at a value so large that $c_7/m_0 < \alpha/4$, then there will be an integer m, $1 \leq m \leq m_0$, and a number $y = (\sqrt{x})^{\mu^{m-1}}$, so that

$$\sum_{y^\mu < q \leq y} \frac{1}{q} < \frac{\alpha}{4}\,. \tag{8}$$

Note that whatever the value of $x \geq 2$, m_0 is fixed and y exceeds a certain fixed (positive) power of x.

Consider those integers a_j not exceeding y for which the inequalities (1) are valid. The number of these which are divisible by the square of a prime $l > l_0$ is at most

$$\sum_{l > l_0} \left[\frac{y}{l^2}\right] < \frac{\alpha\, y}{8}$$

provided that we fix l_0 at a sufficiently large value. The number of a_j which are divisible by a prime-power l^s where $l \leq l_0$ and $s > s_0$ (≥ 2) is not more than

$$\sum_{s > s_0} \sum_{l \leq l_0} \left[\frac{y}{l^s}\right] < c_{11}\, y\, 2^{-s_0} < \frac{\alpha\, y}{8}$$

provided that s_0 is fixed at a sufficiently large value. The number of a_j which are divisible by one of the primes q which are counted in (8) is at most

$$\sum_{y^\mu < q \leq y} \left[\frac{y}{q}\right] < \frac{\alpha y}{4}.$$

Altogether, therefore, there are at least $\alpha y/2$ of the a_j ($\leq y$) remaining on which

$$\left|\sum_{\substack{p|a_j \\ p \leq y^\mu}} f(p)\right| \leq 2c_2 g(x) + c_{12} \leq c_{13} g(x)$$

where the constant c_{12} depends only upon the value of $f(l^s)$ with $l \leq l_0$, $s \leq s_0$. *Note that in this part of the argument we only need the lower bound for $f(a_j)$ which is given in* (1) *in so far as it is embodied in the estimate* (8).

Let the random variables Y_p be defined as in Lemma 2 with x replaced by y and r by y^μ. Let μ be fixed at a value so that the term $O(\exp(-c_5 \mu^{-1}))$ is less than $\alpha/4$. Then we have the lower bound

$$P\left(\left|\sum_{p \leq y^\mu} Y_p\right| \leq c_{13} g(x)\right) > \frac{\alpha}{4}. \tag{9}$$

Let $\varphi(t) = \varphi(x, t)$ denote the characteristic function of the distribution function

$$P\left(g(x)^{-1} \sum_{p \leq y^\mu} Y_p \leq z\right). \tag{10}$$

A straightforward calculation shows that it is

$$\prod_{p \leq y^\mu} \left(1 + \frac{1}{p}\left\{\exp \frac{itf(p)}{g(x)}\right\} - 1\right\}\right).$$

Here, for $3 \leq p \leq y^\mu$, a typical term in the product is

$$\exp\left(\frac{1}{p}\left\{\exp\left(\frac{itf(p)}{g(x)}\right) - 1\right\} + O\left(\frac{1}{p^2}\right)\right)$$

so that

$$|\varphi(t)| \leq c_{14} \exp\left(-2 \sum_{p \leq y^\mu} \frac{1}{p} \sin^2\left(\frac{tf(p)}{2g(x)}\right)\right). \tag{11}$$

We now appeal to a relation between a distribution function $F(z)$ and its characteristic function $\psi(t)$. According to a formula which appears in

CRAMER [2] (it may readily be obtained by convoluting $F(z)$ with a normal law and applying Plancherel's relation)

$$\frac{1}{2h}\int_0^h (F(w+u)-F(w-u))\,du = \tag{12}$$

$$= \frac{1}{2\pi}\int_{-\infty}^{\infty}\left(\frac{\sin t}{t}\right)^2 \cdot e^{-2itw/h}\cdot\psi\left(\frac{2t}{h}\right)dt\,,$$

uniformly for all real w and real $h>0$. We apply this equation with $F(z)$ the distribution function at (10) so that $\psi(t)=\varphi(t)$, $h=2c_{13}\eta$ where $\eta>1$, and $w=0$.

According to our lower bound (9), $F(u)-F(-u)>\alpha/4$ for $u>c_{13}$, so that the left-hand side of (12) in the case at hand exceeds $\alpha/16$; hence

$$\int_{-\infty}^{\infty}\left(\frac{\sin t}{t}\right)^2\left|\varphi\left(\frac{t}{\eta c_{13}}\right)\right|dt>\alpha\pi/8\,.$$

Moreover, the contribution of the range $|t|>t_1=32/\alpha$ to this last integral is not more than

$$2\int_{t_1}^{\infty}\frac{dt}{t^2}\le\frac{\alpha}{16}\,,$$

whilst the range $|t|\le t_0=\alpha/32$ contributes at most

$$2\int_0^{t_0}1\cdot dt\le\frac{\alpha}{16}\,.$$

Altogether

$$\int_{t_0<|t|\le t_1}\left|\varphi\left(\frac{t}{\eta c_{13}}\right)\right|dt>\frac{\alpha}{4}\,.$$

Hence there is a value of τ in the range $t_0\le|\tau|\le t_1$ for which

$$\left|\varphi\left(\frac{\tau}{\eta c_{13}}\right)\right|>\frac{\alpha^2}{128}\,,$$

and so (in view of our upper bound (11))

$$\sum_{p\le y^u}\frac{1}{p}\sin^2\left(\frac{\tau f(p)}{2\eta c_{13}g(x)}\right)\le\frac{1}{2}\log\left(\frac{128c_{14}}{\alpha^2}\right)=c_{15}\,.$$

In view of the inequality

$$|\sin w| \geq \frac{2}{\pi} |w| \quad \text{if} \quad |w| \leq \frac{\pi}{2},$$

we see that on the primes p_1 for which $|f(p_1)| \leq \pi \eta c_{13} g(x)/|\tau| = c_{16} g(x)$ we have

$$\sum_{p_1 \leq y^\mu} \frac{1}{p_1} f(p_1)^2 \leq c_{15} \pi^2 \eta^2 c_{13}^2 |\tau|^{-2} \cdot g(x)^2 .$$

Choose a value of η so that $\pi \eta c_{13} |\tau|^{-1} > 2c_2$ (see (7) and what follows). Then if the inequality $|f(p)| \leq c_{16} g(x)$ fails, for say $p = p_2$, we have $|f(p_2)| > 2c_2 g(x)$ and so

$$\sum_{p_2 \leq y^\mu} \frac{1}{p_2} \leq c_7 .$$

We have now proved that

$$\sum_{p \leq y^\mu} \frac{1}{p} \min \left(1, \frac{|f(p)|}{c_{16} g(x)}\right)^2 \leq c_{17} .$$

Here $y^\mu > x^\Theta$ for some constant Θ, $0 < \Theta < 1$. Replacing x in this last inequality by $x^{1/\Theta}$ gives

$$\sum_{p \leq x} \frac{1}{p} \min \left(1, \frac{|f(p)|}{c_{16} g(x^{1/\Theta})}\right)^2 \leq c_{17}, \quad (x \geq x_0) .$$

Bearing in mind (from (i)) that $g(x^{1/\Theta}) \leq c_{18} g(x)$ for some constant c_{18} and all x sufficiently large, we readily obtain the desired inequality (ii).

This completes the proof of theorem 1.

PROOF OF THEOREM 2. We obtain an inequality which is to replace that of (7) using only the weaker hypothesis $\alpha > 0$ rather than $\alpha > 1/2$.

Consider the integers n not exceeding x which are not divisible by the square of any prime q greater than q_0 and for which $|f(n)| \leq g(x)$. According to the hypothesis of theorem 2 they are in number at least

$$\alpha x - \sum_{q > q_0} \left[\frac{x}{q^2}\right] \geq \alpha x/2 ,$$

provided that q_0 is fixed at a value sufficiently large with respect to α. Let us denote these integers by $m_1 < m_2 < \ldots < m_k$. Indeed, replacing x by w, $2 \leq w \leq x$, in this result, making use of the nondecreasing property of the function $g(\ )$, and integrating by parts, we see that for all $x \geq x_1$,

$$\sum_{r=1}^{k} \frac{1}{m_r} \geq c_{19} \log x .$$

Let $q_1 < q_2 < \ldots q_l \leq x$ run through the primes q greater than q_0 for which $f(q) > 2g(x)$. If there is any equation of the form $m_i = \lambda m_j$ where λ is made up entirely from the primes q then we have $2g(x) < f(\lambda) \leq \leq |f(m_i)| + |f(m_j)| \leq 2g(x)$, which is impossible. The conditions of lemma 3 are therefore satisfied and we see that

$$\sum_{m=1}^{l} \frac{1}{q_m} \leq c_0 \cdot c_{19}^{-2} = c_{20} .$$

Considering those primes on which $f(p) < -2g(x)$ similarly, we deduce that

$$\sum_{\substack{p \leq x \\ |f(p)| > 2g(x)}} \frac{1}{p} \leq 2\left(c_{20} + \sum_{p \leq q_0} \frac{1}{p}\right) = c_{21} .$$

This inequality then replaces that of (7). Here c_{21} depends at most upon α and x_1 (and so the c_3 which appears in the hypothesis involving $A(x)$).

The proof of theorem 2 is now readily completed along the lines of the proof of theorem 1, save that in place of the last step, (i) not being available, we set $\Theta^{-1} = \beta$.

Concerning the corollary we note that the constant c_{12} which occurs in the proof of theorem 1 can be given the value

$$c_{12} = \sum_{l \leq l_0} \sum_{s \leq s_0} |f(l^s)| \leq l_0 s_0 \max_{n \leq y} |f(n)| \leq l_0 s_0 g(x)$$

in this (particular) case. This establishes the inequality

$$\sum_{p \leq x} \frac{|f(p)|^2}{p} \leq A_0 \max_{n \leq x} |f(n)|^2$$

with some absolute constant A_0, for all x absolutely large: we may choose $\alpha = 3/4$. For the remaining values of x such an inequality is trivially valid.

REMARK. It is perhaps worthwhile to mention here that if an additive arithmetic function $f(n)$ satisfies $|f(p)| = h(p)$ where the real function $h(x)$ is a slowly-oscillating function of $\log x$, that is $h(x^y)/h(x) \to 1$ as $x \to \infty$ for each fixed $y > 0$, then for any δ, $0 < \delta < 1$,

$$|f(p)|^2 \sum_{p^\delta < q \leq p} \frac{1}{q} \leq 2 \sum_{q \leq p} \frac{|f(q)|^2}{q} = 2B(p)^2$$

so that

$$\limsup_{p \to \infty} |f(p)| \, B(p)^{-1} \leq \left(\log \frac{1}{\delta}\right)^{-1/2}$$

Letting $\delta \to 0^+$ we see that $f(p) = o(B(p))$, $p \to \infty$. It is now easy to check that the function $f(n)$ belongs to the class H of KUBILIUS ([10] p. 47) and indeed that for any $\varepsilon > 0$

$$\frac{1}{B(N)^2} \sum_{\substack{p \leq N \\ |f(p)| > \varepsilon B(N)}} \frac{f(p)^2}{p} \to 0, \quad (N \to \infty).$$

Hence, defining

$$U(N) = \sum_{p \leq N} \frac{f(p)}{p}$$

one obtains the result

$$\frac{1}{N} \sum_{\substack{n \leq N \\ f(n) - U(N) \leq zB(N)}} \to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-u^2/2}\, du, \quad (N \to \infty);$$

see, for example, KUBILIUS [10] Theorem (4.2), p. 61; or SHAPIRO [11]. This remark would apply to the case when $f(p) = \exp(\sqrt{\log\log p})$.

If $|f(p)| \geq \log p$ then this remark is no longer applicable, but theorems 1 and 2 can still be useful.

PROOF OF THEOREM 3. Let $n_1 < n_2 < \ldots$ be a sequence of real numbers, which for the moment are only restricted by the requirements $n_1 \geq 2$ and $n_{i+1} > \exp(n_i)$, for all $i \geq 1$.

Let B, l be positive numbers, which will be fixed presently in terms of η, and which may be thought of as "large." We write δ for $e^{-1/2}$.

We define the sequence a_j by induction, it is to consist of all positive integers up to and including n_1, and in the interval $n_i < a_j \leq n_{i+1}$ it is to consist of those integers which have at least one but not more than l prime divisors p in the range $n_i^\delta < p \leq n_i^B$, but no prime divisor q in the range $n_{i+1}^\delta < q \leq n_{i+1}$. These prime divisors need not be distinct.

It is clear that the estimate $A(x) \geq x/2 \geq x(1/2 - \eta)$ holds for $1 \leq x \leq n_1$. We assume that it holds for all values of x up to n_i, and prove that it also holds for $n_i < x \leq n_{i+1}$.

Case 1. $n_i < x \leq n_i^B$.

If $\delta l > B$ then for any fixed μ, $\delta < 1 - \mu < 1$,

$$A(x) - A(n_i^{1-\mu}) \geq x - n_i^{1-\mu} - 1 - \sum_{m \leq x}{}_1 1$$

where $\sum_1$ counts those integers m, $n_i^{1-\mu} < m \leq x$, which have no prime divisor p in the range $n_i^\delta < p \leq n_i^B$. For, if an integer $m \leq x \leq n_i^B$ possesses at least l prime divisors which are greater than n_i^δ then $n_i^{\delta l} \leq m \leq n_i^B$, $\delta l \leq B$, which is false.

For $0 < \beta < 1$ let $\psi(x, x^\beta)$ denote the number of integers not exceeding x which are made up of primes not exceeding x^β. In the range $1/2 < \beta < 1$ there is the estimate

$$\psi(x, x^\beta) = [x] - \sum_{x^\beta < p \leq x} \left[\frac{x}{p}\right] = (1 + \log\beta + o(1))\, x$$

as $x \to \infty$. Hence, for any fixed $\varepsilon > 0$,

$$\begin{aligned}\sum_{m \leq x}{}_1 1 &\leq \psi(x, x^{\delta}) - \psi(n_i^{1-\mu}, n_i^{(1-\mu)\delta_1}), \ \delta_1(1-\mu) = \delta, \\ &\leq (1 + \log \delta + \varepsilon) x - (1 + \log \delta_1 - \varepsilon) n_i^{1-\mu} \leq \\ &\leq (1/2 + \varepsilon)(x - n_i^{1-\mu}) + O((\mu + \varepsilon) n_i^{1-\mu}) \leq \\ &\leq (1/2 + 2\varepsilon)(x - n_i^{1-\mu}),\end{aligned}$$

since $x - n_i^{1-\mu} \geq n_i - n_i^{1-\mu} \geq (n_i)/2$ provided that n_1 is sufficiently large. Hence

$$A(x) - A(n_i^{1-\mu}) \geq (1/2 - 3\varepsilon)(x - n_i^{1-\mu})$$

and $A(x) \geq (1/2 - \eta) x$ if $3\varepsilon \leq \eta$, and n_1 is sufficiently large in terms of ε.

Case 2. $n_i^B < x \leq n_{i+1}^{\delta}$.

In this case

$$A(x) - A(n_i^{B/2}) \geq x - n_i^{B/2} - 2 - \sum_{m \leq x}{}_2 1 - \sum_{m \leq x}{}_3 1$$

where $\sum_2$ counts those integers m, $n_i^{B/2} < m \leq x$, which have no prime factor p in the range $n_i^{\delta} < p \leq n_i^B$, and $\sum_3$ counts those integers in the same range which have more than l such factors.

Let $w_2(m)$ denote the number of distinct prime factors p of m which satisfy $n_i^{\delta} < p \leq n_i^{B/4}$, and let $w_3(m)$ denote the number which lie in the larger range $n_i^{\delta} < p \leq n_i^B$.

The well-known method of TURÁN [11] allows one to prove that

$$\sum_{n_i^{B/2} < m \leq x} (w_2(m) - S)^2 = O((x - n_i^{B/2}) S)$$

where

$$S = \sum_{n_i^{\delta} < p \leq n_i^{B/4}} \frac{1}{p} = \log \frac{B}{4\delta} + O\left(\frac{1}{\log n_i}\right).$$

If m is counted in $\sum_2$ then $w_2(m) = 0$, and so

$$\sum_{m \leq x}{}_2 1 = O\left(\frac{x - n_i^{B/2}}{\log B}\right) < \varepsilon(x - n_i^{B/2})$$

if B has a sufficiently large value, depending upon ε.

From the readily obtained estimate

$$\sum_{n_i^{B/2} < m \leq x} w_3(m) = O((x - n_i^{B/2}) \log B)$$

we see that those integers m which have at least $(l+1)$ distinct prime divisors in the interval $(n_i^{\delta}, n_i^B]$ are at most

$$O\left(\frac{\log B}{l}(x-n_i^{B/2})\right)<\varepsilon(x-n_i^{B/2})$$

in number. Those m which have a divisor p^2 with p in the same range are also

$$O\left(\sum_{p>n_i^\delta}\frac{x-n_i^{B/2}}{p^2}\right)+O\left(\frac{n_i^B}{\log n_i}\right)<\varepsilon(x-n_i^{B/2})$$

in number.

Altogether, therefore

$$A(x)-A(n_i^{B/2})>(1-3\varepsilon)(x-n_i^{B/2})$$

and $A(x)\geq(1/2-\eta)x$ once again.

Case 3. $n_{i+1}^\delta<x\leq n_{i+1}$.

We estimate $A(x)-A(n_{i+1}^{\delta/2})$ from below. Everything proceeds as in case 2 save that we must now remove those integers m, $n_{i+1}^{\delta/2}<m\leq x$, which have a prime divisor q in the range $n_{i+1}^\delta<q\leq n_{i+1}$. They are in number at most

$$\sum_{n_{i+1}{}^\delta<q\leq x}\left(\frac{x-n_{i+1}^{\delta/2}}{q}+O(1)\right)<(x-n_{i+1}^{\delta/2})\left(\log\frac{1}{\delta}+O\left(\frac{1}{\log n_{i+1}}\right)\right)+$$
$$+O\left(\frac{x}{\log x}\right)<(1/2+\varepsilon)(x-n_{i+1}^{\delta/2})$$

since $x<2(x-n_{i+1}^{\delta/2})$.

Hence

$$A(x)-A(n_{i+1}^{\delta/2})>(1/2-4\varepsilon)(x-n_{i+1}^{\delta/2})$$

and once again $A(x)\geq(1/2-\eta)x$, this time if $4\varepsilon\leq\eta$.

We fix $\varepsilon=\eta/4$, B in terms of ε, l in terms of B, and n_1 in terms of ε, B, l and so ε, so that the finitely many conditions of the type $n_i^B<n_{i+1}^{\delta/2}$ are satisfied, and our argument works in every case.

The proof of theorem 3 is now readily completed. The additive function $f(n)$ is defined to be $\log n$ if $n\leq n_1$; by $f(p)=\log p$ if $n_i^\delta<p<n_i^B$ for some i, and to be zero otherwise.

The function $g(x)$ is defined to be $\log n_i$ when $n_i<x\leq n_{i+1}$.

An integer a_j which lies in the range $n_i<a_j\leq n_{i+1}$, $i\geq 2$, has between 1 and l prime factors in the range $(n_i^\delta, n_i^B]$, and none greater than n_i^B. Hence

$$\delta\log n_i\leq|f(a_j)|\leq B\log n_i+\sum_{p\leq n_{i-1}{}^B}\log p+\log n_1\leq(2B+1)\log n_i$$

provided that n_i is sufficiently large in terms of n_{i-1}. This will follow from the condition $n_{i+1} > \exp(n_i)$ if n_1 is sufficiently large in terms of B. For such integers a_j we have

$$\delta \leq \frac{|f(a_j)|}{g(a_j)} \leq 2B+1 .$$

Finally, if $x_i = n_{i+1} - 1 \geq 4$, then we have $x_i^2 > n_{i+1}$, and

$$\frac{g(x_i^2)}{g(x_i)} \geq \frac{\log n_{i+1}}{\log n_i} \to \infty, \quad i \to \infty ,$$

which establishes (vi).

References

[1] F. Behrend, On sequences of numbers not divisible one by another, *Journ. London Math. Soc.*, **10** (1935), 42–44.

[2] H. Cramer, *Random Variables and Probability Distributions*, Cambridge, 1970.

[3] P. D. T. A. Elliott, A conjecture of Kátai, *Acta Arithmetica*, **26** (1974), 11–20.

[4] P. D. T. A. Elliott, On connections between the Turán–Kubilius inequality and the Large Sieve: Some applications, *Amer. Math. Soc. Proceedings of Symposia in Pure Math. Vol.* **24** Providence, 1973, pp. 77–82.

[5] P. D. T. A. Elliott, *Probabilistic Number Theory*. Springer, New York, 1980.

[6] P. D. T. A. Elliott, On two conjectures of Kátai, *Acta Arithmetica*, **30** (1976), 35–59.

[7] P. D. T. A. Elliott, On a conjecture of Narkiewicz about functions with non--decreasing normal orders, *Colloquium Math.*, **36** (1976), 289–294.

[8] P. Erdös, On the density of some sequences of numbers III., *Journ. London Math. Soc.*, **13** (1938), 119–127.

[9] P. Erdös, A. Wintner, Additive arithmetical functions and statistical independence, *Amer. Journ. Math.*, **61** (1939), 713–721.

[10] J. Kubilius, *Probabilistic Methods in the Theory of Numbers*, Amer. Math. Soc. Translations of Math. Monographs. 11, Providence, Rhode Island, 1964.

[11] P. Turán, On a theorem of Hardy and Ramanujan, *J. London Math. Soc.*, **9** (1934), 274–276.