TRANSACTIONS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 310, Number 1, November 1988

AN ERDŐS-WINTNER THEOREM FOR DIFFERENCES OF ADDITIVE FUNCTIONS

ADOLF HILDEBRAND

ABSTRACT. An Erdős-Wintner type criterion is given for the convergence of the distributions $D_x(z) = [x]^{-1}\#\{1 \le n \le x: f(n+1) - f(n) \le z\}$, where f is a real-valued additive function. A corollary of this result is that an additive function f, for which $f(n+1) - f(n)$ tends to zero on a set of density one, must be of the form $f = \lambda \log$ for some constant λ. This had been conjectured by Erdős.

1. Introduction. Let f be a real-valued additive function, and for $x \ge 1$ define the distribution functions

$$F_x(z) = \frac{1}{[x]}\#\{1 \le n \le x: f(n) \le z\}. \tag{1.1}$$

About 50 years ago, Erdős and Wintner [**9**] proved their celebrated theorem which states that, as $x \to \infty$, the distributions (1.1) converge weakly towards a limit distribution, if and only if the three series

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

each taken over the sequence of primes in their natural order, converge. This result, and the equally well-known theorem of Erdős and Kac [**8**], were the earliest examples of limit theorems for additive functions, and had a great influence on the development of what is now called probabilistic number theory; see, for example, the books of Kubilius [**14**] and Elliott [**4**, **5**].

The main purpose of this paper is to prove an analogue of the Erdős-Wintner theorem for the differences $f(n+1) - f(n)$ of an additive function, i.e., to give a necessary and sufficient condition for the convergence of the distributions

$$D_x(z) = \frac{1}{[x]}\#\{1 \le n \le x: f(n+1) - f(n) \le z\}. \tag{1.2}$$

Elliott [**3**] proved a result of this type under the assumption that the distributions $D_x(z)$ have bounded mean and variance as $x \to \infty$. The general case, however, had up to now been an open problem (see, e.g., [**1**, p. 110; **5**, p. 334, problem 11]).

THEOREM 1. *The distributions $D_x(z)$ converge weakly towards a limit distribution $D(z)$, as $x \to \infty$, if and only if there exists a real number λ such that the*

Received by the editors August 6, 1987.

1980 *Mathematics Subject Classification* (1985 *Revision*). Primary 11K65; Secondary 11N37, 11N64.

Supported by NSF grant DMS 8640693.

function $h = f - \lambda \log$ satisfies

$$\sum_{|h(p)|\le 1} \frac{|h(p)|^2}{p} < \infty, \qquad \sum_{|h(p)|>1} \frac{1}{p} < \infty. \tag{1.3}$$

If this condition is satisfied, then the characteristic function of the limit distribution $D(z)$ is given by

$$\phi(t) = \prod_p \left(1 - \frac{2}{p} + 2\left(1 - \frac{1}{p}\right) \operatorname{Re} \sum_{m\ge 1} \frac{e^{ith(p^m)}}{p^m}\right). \tag{1.4}$$

An outline of the proof will be given in §2. The sufficiency of conditions (1.3) is in fact known (cf. [**1**, p. 110]), and can be proved in a routine manner, but for completeness we shall give a proof in §3. We shall use the characteristic function method, which has the advantage of yielding, without extra effort, formula (1.4) for the characteristic function of the limit distribution. The proof of the necessity of conditions (1.3) turns out to be much more difficult and forms the core of this paper.

We remark that conditions (1.3) are known to be necessary and sufficient for the convergence of the distributions $F_x(z - \alpha(x))$ with suitable centering constants $\alpha(x)$ (cf. Theorem 7.1 in [**4**]). Thus, the convergence of the distributions $D_x(z)$ is equivalent to that of $F_x(z - \alpha(x))$ with suitable $\alpha(x)$. This is in fact to be expected from probabilistic considerations.

Theorem 1 has an application to the problem of characterizing additive functions as multiples of the logarithm. Beginning with Erdős [**7**], a number of authors have found conditions on an additive function f which imply that f is of the form $f = \lambda \log$ for some constant λ. Erdős [**7**] showed that if f is either monotonic or satisfies $f(n+1) - f(n) = o(1)$, as $n \to \infty$, then $f = \lambda \log$ for some λ. Moreover, he conjectured that the latter condition could be weakened to

$$\lim_{x\to\infty} \frac{1}{x} \sum_{n\le x} |f(n+1) - f(n)| = 0. \tag{1.5}$$

This was later proved by Kátai [**13**] and Wirsing [**15**]. For a survey of these and related results see Chapter 11 in [**6**]. From Theorem 1 we can deduce the following result that had also been conjectured by Erdős [**7**].

COROLLARY. *If f is a real-valued additive function such that $f(n+1) - f(n)$ tends to zero, as n tends to infinity through a set of density one, then there exists a real number λ such that $f = \lambda \log$.*

It is easy to see that condition (1.5) implies that of the corollary. Thus the corollary contains the Kátai-Wirsing result.

To derive the corollary, we note that under the stated condition the distributions $D_x(z)$ converge to the degenerate distribution which has point mass one at the origin. Thus, by the theorem, there exists a real number λ such that (1.3) holds for $h = f - \lambda \log$, and the characteristic function of the limit distribution is given by (1.4). In order for the limit distribution to be the indicated degenerate distribution, this characteristic function has to be identically one. Since each factor in the product (1.4) is, in absolute value, at most 1, and equal to 1 for all t if and only if

$h(p^m) = 0$ for all $m \geq 1$, the function h must be identically zero, i.e., we must have $f = \lambda \log$, as claimed.

2. Outline of the proof. As mentioned in the introduction, we shall use for the proof the characteristic function method. According to the continuity theorem in probability theory the distributions $D_x(z)$, defined in (1.2), converge weakly towards a limit distribution $D(z)$, if and only if the associated characteristic functions

$$\phi_x(t) = \frac{1}{[x]} \sum_{n \leq x} e^{it(f(n+1)-f(n))}$$

converge pointwise towards the characteristic function $\phi(t)$ of $D(z)$. Moreover, if this condition is satisfied, then the convergence of $\phi_x(t)$ is uniform on any bounded t-interval.

The problem therefore essentially reduces to that of finding necessary and sufficient conditions for the existence of the mean value

$$\lim_{x\to\infty} \phi_x(t) = \lim_{x\to\infty} \frac{1}{x} \sum_{n\leq x} e^{it(f(n+1)-f(n))} \tag{2.1}$$

for every t. Since, for fixed t, the function $g(n) = \exp(itf(n))$ is a multiplicative function of modulus 1, one might try to attack this problem by finding necessary and sufficient conditions for the existence of the mean value

$$\lim_{x\to\infty} \frac{1}{x} \sum_{n\leq x} g(n+1)\overline{g(n)}, \tag{2.2}$$

when g is a *general* multiplicative function of modulus 1. As far as the "sufficiency" part is concerned, this approach is successful. In §3 we shall establish the following general result and deduce from it the sufficiency of conditions (1.3) as well as formula (1.4) in Theorem 1.

THEOREM 2. *Let g be a multiplicative function of modulus* 1, *and suppose that, for some real number α, we have*

$$\sum_p \frac{|1-g(p)p^{i\alpha}|^2}{p} < \infty. \tag{2.3}$$

Then the limit (2.2) *exists and is equal to*

$$\prod_p \left(1 - \frac{2}{p} + 2\left(1-\frac{1}{p}\right) \operatorname{Re} \sum_{m\geq 1} \frac{g(p^m)p^{i\alpha m}}{p^m}\right). \tag{2.4}$$

It appears to be much more difficult to obtain *necessary* conditions for the existence of the mean value (2.2), and we have not been able to do so. The following is a plausible, but probably very deep, conjecture. If true, it would easily imply the "necessity" part of Theorem 1.

CONJECTURE. *Let g be a multiplicative function of modulus* 1, *and suppose that the mean value* (2.2) *exists and is nonzero. Then there exists a real number α for which* (2.3) *holds.*

For the application to Theorem 1, this conjecture can be avoided, and a weaker result, namely Theorem 3 below, used instead. The idea is to take full advantage of the fact that, assuming the convergence of the distributions $D_x(z)$, we know the existence of the mean value (2.2) for an entire class of functions $g(n)$ (namely $g(n) = \exp(itf(n)))$, depending on a continuous parameter t, moreover that the convergence in (2.2) is uniform with respect to this parameter on any bounded interval, and finally that the limit function tends to 1, as $t \to 0$.

Roughly speaking, Theorem 3 asserts that if $g(n+1) - g(n)$ is small on average over $n \le x$ (which will be the case if $|g| = 1$ and $\frac{1}{x}\sum_{n\le x} g(n+1)\overline{g(n)}$ is close to 1), then the sum

$$\sum_{p\le x} \frac{|1 - g(p)p^{i\alpha}|^2}{p} \tag{2.5}$$

is either bounded in absolute terms for some (not too large) α, or is very large for $\alpha = 0$. By applying this result with $g(n)n^{i\alpha}$ in place of $g(n)$, we shall deduce that under the same assumptions the sum (2.5) is either bounded for some α in the range $|\alpha| \ll \log x$ or very large for all values α in (roughly) the same range. A precise statement of this variant is given as Theorem 3* in §6.

To derive the "necessity" part of Theorem 1, we shall apply this with $g(n) = e^{itf(n)}$. Assuming the existence of a limit distribution for $D_x(z)$, the quantity

$$\phi_x(t) = \frac{1}{[x]}\sum_{n\le x} e^{itf((n+1)-f(n))} = \frac{1}{[x]}\sum_{n\le x} g(n+1)\overline{g(n)}$$

will be close to 1, provided t is sufficiently small and x sufficiently large. Hence, for such values of t and x, say $|t| \le t_0$ and $x \ge x_0$, the sum

$$\sum_{p\le x} \frac{|1 - e^{itf(p)}p^{i\alpha}|^2}{p} \tag{2.6}$$

will be either absolutely bounded for some $\alpha, |\alpha| \ll \log x$, or very large for all α in this range. However, since the first alternative holds trivially for $t = 0$, it must remain valid by continuity for all t, $|t| \le t_0$. (This continuity argument is crucial for the success of our method, and would not be possible if we would work with a *fixed* multiplicative function, rather than a family of multiplicative functions.) The sums (2.5) therefore are uniformly bounded for $|t| \le t_0$, $x \ge x_0$, with suitably chosen numbers $\alpha = \alpha(t, x)$ satisfying $|\alpha| \ll \log x$. The completion of the proof is then relatively easy. One shows first that for each t, $|t| \le t_0$, there exists a *fixed* number $\alpha = \alpha(t)$ such that the infinite sum $\sum_p |1 - e^{itf(p)}p^{i\alpha}|^2/p$ is bounded, next that $\alpha(t)$ must be of the form $\alpha(t) = \lambda t$ for some real number λ and all rational t, $|t| \le t_0$, and finally that conditions (1.3) must hold for $h = f - \lambda \log$.

THEOREM 3. *There exists an absolute positive constant c with the following property. Let $x \ge 2$ and let g be a multiplicative function of modulus* 1 *and satisfying*

$$\inf_{\alpha\in\mathbf{R}} \left| \sum_{m\ge 0} \frac{g(2^m)}{2^{m(1+i\alpha)}} \right| \ge \frac{1}{4}. \tag{2.7}$$

Suppose that

$$\sup_{x^\delta \le y \le x} \frac{1}{y}\sum_{n\le y} |g(n+1) - g(n)| \le \delta \tag{2.8}$$

holds with some real number δ satisfying

$$(\log x)^{-1/2} \le \delta \le 1/3. \tag{2.9}$$

Then we have either

$$\min_{|\alpha|\le 1/\delta} \sum_{p\le x} \frac{|1-g(p)p^{i\alpha}|^2}{p} < c \tag{2.10}$$

or

$$\sum_{p\le x} \frac{|1-g(p)|^2}{p} \ge \frac{1}{3}\log\left(\frac{1}{\delta}\right). \tag{2.11}$$

An assumption like (2.7) seems to be unavoidable in our approach, but the result is perhaps true without this hypothesis. At any rate, in the application to the proof of Theorem 1, this assumption is easily verified.

The proof of Theorem 3 is based on the arithmetic version of the large sieve, applied with $a_n = g(n)$. Crucial in the proof is the fact that the standard bound provided by the large sieve can be improved, if it is known that $g(n+1) - g(n)$ is small in some average sense. This observation was first made in [**10**], and was used there to derive some consequences of the assumption that $g(n) = g(n+1)$ holds for "most" n in the case g is a completely multiplicative function with values ± 1.

We shall prove Theorem 3 in §§4 and 5, and in §6 we shall deduce from it the "necessity" part of Theorem 1.

3. Proof of Theorem 1: Sufficiency part. As mentioned in the previous section, the proof depends on Theorem 2, which we shall prove first. It suffices to prove the result for $\alpha = 0$; the general case can be deduced from this by setting $g^*(n) = g(n)n^{i\alpha}$ and noting that neither the existence of the limit (2.2), nor its value, are affected if we replace $g(n)$ by $g^*(n)$.

We fix a multiplicative function g of modulus 1 and suppose that (2.3) holds with $\alpha = 0$, i.e. that

$$\sum_p \frac{|1-g(p)|^2}{p} = 2\mathrm{Re}\sum_p \frac{1-g(p)}{p} < \infty. \tag{3.1}$$

It is easily seen that this implies the convergence of the infinite product (2.4) with $\alpha = 0$. For $r \ge 2$, we define "truncated" multiplicative functions g_r by setting

$$g_r(p^m) = \begin{cases} g(p^m) & \text{if } p \le r, \\ 1 & \text{if } p > r. \end{cases}$$

Suppose that $L(u)$ is a slowly oscillating function of modulus 1, i.e., a function satisfying

$$|L| = 1, \qquad \sup_{1\le t\le 2} |L(tu) - L(u)| \to 0 \qquad (u\to\infty).$$

Then, as $x \to \infty$,

$$\begin{aligned}\frac{1}{x}\sum_{n\le x} g(n+1)\overline{g(n)} &= \frac{1}{x}\sum_{n\le x} g(n+1)L(n+1)\overline{g(n)L(n)} + o(1) \\ &= \frac{1}{x}\sum_{n\le x} g_r(n+1)\overline{g_r(n)} + O\left(\frac{1}{x}\sum_{n\le x}|g_r(n) - g(n)L(n)|\right) + o(1).\end{aligned}$$

Thus, to obtain the desired conclusion, it suffices to show that

$$\lim_{x\to\infty}\frac{1}{x}\sum_{n\le x} g_r(n+1)\overline{g_r(n)} = \prod_{p\le r}\left(1-\frac{2}{p}+2\left(1-\frac{1}{p}\right)\mathrm{Re}\sum_{m\ge1}\frac{g(p^m)}{p^m}\right) \tag{3.2}$$

holds for each $r\ge 2$, and further that, with a suitable choice of the function $L(n)$,

$$\overline{\lim_{r\to\infty}}\ \overline{\lim_{x\to\infty}}\ \frac{1}{x}\sum_{n\le x}|g_r(n)-g(n)L(n)| = 0. \tag{3.3}$$

For the proof of (3.2) we define a multiplicative function h_r by

$$h_r(p^m)=\begin{cases} g(p^m)-g(p^{m-1}) & (p\le r),\\ 0 & (p>r),\end{cases}$$

so that $g_r = 1 * h_r$. We have

$$\begin{aligned}\frac{1}{x}\sum_{n\le x} g_r(n+1)\overline{g_r(n)} &= \frac{1}{x}\sum_{d,d'\le x} h_r(d)\overline{h_r(d')}\sum_{\substack{n\le x\\ d\mid n+1\\ d'\mid n}} 1\\ &= \sum_{\substack{d,d'\le x\\ (d,d')=1}}\frac{h_r(d)\overline{h_r(d')}}{d,d'} + O\left(\frac{1}{x}\sum_{d,d'\le x}|h_r(d)h_r(d')|\right).\end{aligned}$$

The error term is bounded by

$$\ll \frac{1}{x}\left(\sum_{d\ge1}|h_r(d)|\left(\frac{x}{d}\right)^{1/3}\right)^2 = x^{-1/3}\prod_{p\le r}\left(1+\sum_{m\ge1}\frac{|h_r(p^m)|}{p^{m/3}}\right)^2,$$

and hence tends to zero as $x\to\infty$. A similar argument shows that, as $x\to\infty$, the main term approaches

$$\begin{aligned}\sum_{\substack{d,d'\ge1\\ (d,d')=1}}\frac{h_r(d)\overline{h_r(d')}}{dd'} &= \prod_{p\le r}\left(1+\sum_{m\ge1}\frac{h_r(p^m)}{p^m}+\sum_{m\ge1}\frac{\overline{h_r(p^m)}}{p^m}\right)\\ &= \prod_{p\le r}\left(1+2\mathrm{Re}\sum_{m\ge1}\frac{h_r(p^m)}{p^m}\right)\\ &= \prod_{p\le r}\left(1+2\mathrm{Re}\sum_{m\ge1}\frac{g(p^m)-g(p^{m-1})}{p^m}\right),\end{aligned}$$

which reduces to the product in (3.2).

To prove (3.3), we apply first Cauchy's inequality, getting

$$
\begin{aligned}
\left(\frac{1}{x}\sum_{n\le x}|g_r(n)-g(n)L(n)|\right)^2 &\le \frac{1}{x}\sum_{n\le x}|g_r(n)-g(n)L(n)|^2 \\
&= \frac{1}{x}\sum_{n\le x} 2\,\mathrm{Re}(1-g(n)L(n)\overline{g_r(n)}) \qquad (3.4)\\
&\le 2\left|1-\frac{L(x)}{x}\sum_{n\le x} g(n)\overline{g_r(n)}\right| + o(1),
\end{aligned}
$$

where in the last relation we used our assumption that the function $L(n)$ is a slowly oscillating function. By a well-known theorem of Halász (cf. Theorem 6.2 in [**4**]) we have, in view of assumption (3.1) and the definition of g_r,

$$
\begin{aligned}
\frac{1}{x}\sum_{n\le x} g(n)\overline{g_r(n)} &= \prod_{r<p\le x}\left(1-\frac{1}{p}\right)\left(1+\sum_{m=1}^{\infty}\frac{g(p^m)}{p^m}\right)+o(1)\\
&= \exp\left\{\sum_{r<p\le x}\frac{g(p)-1}{p}\right\} + O\left(\frac{1}{r}\right)+o(1),
\end{aligned}
$$

as $x\to\infty$. If we now define $L(x)$ by

$$
L(x)=\exp\left\{-i\mathrm{Im}\sum_{r<p\le \max(2r,x)}\frac{g(p)-1}{p}\right\},
$$

which is clearly a slowly oscillating function, then we get

$$
\overline{\lim_{r\to\infty}}\ \overline{\lim_{x\to\infty}}\left|1-\frac{L(x)}{x}\sum_{n\le x}g(n)\overline{g_r(n)}\right| = \overline{\lim_{r\to\infty}}\left|1-\exp\left\{-\mathrm{Re}\sum_{p>r}\frac{1-g(p)}{p}\right\}\right| = 0.
$$

This together with (3.4) proves (3.3), and hence completes the proof of Theorem 2.

PROOF OF THEOREM 1: SUFFICIENCY. Suppose that f is a real-valued additive function and that (1.3) holds for $h=f-\lambda\log$ with some real number λ. We have to show that the mean value (2.1) exists for every real t, is given by the formula (1.4) and is continuous at $t=0$. From (1.3) it readily follows that for every t

$$
\sum_p \frac{|1-e^{itf(p)}p^{-it\lambda}|^2}{p} < \infty.
$$

Thus, applying Theorem 2 with $g(n)=e^{itf(n)}$ and $\alpha=-t\lambda$, we infer that, for every t, the limit (2.1) exists and is equal to the product (1.4). The continuity of the function (1.4) at $t=0$ is a consequence of the relation

$$
\lim_{t\to 0}\mathrm{Re}\sum_p\frac{1-e^{ith(p)}}{p} = \lim_{t\to 0}\frac{1}{2}\sum_p\frac{|1-e^{ith(p)}|^2}{p}=0,
$$

which in turn follows from assumption (1.3).

4. Proof of Theorem 3: Preliminary estimates. Throughout this section we fix a multiplicative function g of modulus 1. We set, for $x \geq 1$,

$$m(x) = \frac{1}{x}\sum_{n\leq x} g(n),$$

$$s(x) = \sum_{p\leq x}\frac{|1-g(p)|^2}{p},$$

$$d(x) = \frac{1}{x}\sum_{n\leq x}|g(n+1)-g(n)|,$$

and denote by $m_\alpha(x)$, $s_\alpha(x)$, and $d_\alpha(x)$ the corresponding quantities with $g(n)$ replaced by $g(n)n^{i\alpha}$. The constants implied in the notations "O" and "$\ll$" are understood to be absolute, and thus, in particular, independent of g.

The first lemma shows that the quantity $d_\alpha(x)$ is, to a large degree, independent of α.

LEMMA 1. *For $x \geq 1$ and real α we have*

$$d_\alpha(x) = d(x) + O\left(\left|\frac{\alpha}{x}\right|\log\left(\left|\frac{x}{\alpha}\right|+2\right)\right).$$

PROOF. We have

$$\begin{aligned} d_\alpha(x) &= \frac{1}{x}\sum_{n\leq x}|g(n+1)(n+1)^{i\alpha} - g(n)n^{i\alpha}| \\ &= \frac{1}{x}\sum_{n\leq x}|g(n+1)-g(n)| + O\left(\frac{1}{x}\sum_{n\leq x}\left|\left(1+\frac{1}{n}\right)^{i\alpha}-1\right|\right) \\ &= d(x) + O\left(\frac{1}{x}\sum_{n\leq x}\min\left(\left|\frac{\alpha}{n}\right|,1\right)\right) \\ &= d(x) + O\left(\left|\frac{\alpha}{x}\right|\log\left(\left|\frac{x}{\alpha}\right|+2\right)\right), \end{aligned}$$

as asserted.

The next lemma implies that sums of the type $s_\alpha(x)$ cannot be small for two values of α, $|\alpha| \leq x$, unless these values are very close to each other.

LEMMA 2. *Let $x \geq 3$ and suppose that the inequality*

$$\sum_{\exp((\log x)^{1/10})<p\leq x}\frac{|1-g(p)p^{i\alpha_j}|^2}{p} \leq K$$

holds for some $K \leq \frac{1}{20}\log\log x$ and two real numbers α_j $(j = 1,2)$ satisfying $|\alpha_j| \leq x$. Then we have

$$|\alpha_1-\alpha_2| \ll e^{2K}/\log x.$$

PROOF. We may assume that x is sufficiently large for otherwise the estimate holds trivially in view of the bound $|\alpha_j| \leq x$. Let $y = \exp((\log x)^{1/10})$ and

$\Delta = \alpha_2 - \alpha_1$. Under the hypothesis of the lemma we have

$$
\begin{aligned}
\operatorname{Re} \sum_{y<p\le x} \frac{1-p^{i\Delta}}{p} &= \frac{1}{2} \sum_{y<p\le x} \frac{|1-p^{i\Delta}|^2}{p} \\
&= \frac{1}{2} \sum_{y<p\le x} \frac{|g(p)p^{i\alpha_1} - g(p)p^{i\alpha_2}|^2}{p} \qquad (4.1)\\
&\le \sum_{y<p\le x} \frac{|g(p)p^{i\alpha_1} - 1|^2}{p} + \sum_{y<p\le x} \frac{|1 - g(p)p^{i\alpha_2}|^2}{p} \le 2K.
\end{aligned}
$$

We express the left-hand side of (4.1) in terms of the zeta function. Using the estimate

$$
\begin{aligned}
\zeta(s) &= \exp\left\{ \sum_p \frac{1}{p^s} + O(1) \right\} \\
&= \exp\left\{ \sum_{p\le e^{1/(\sigma-1)}} \frac{p^{-it}}{p} + O(1) \right\},
\end{aligned}
$$

which is valid for all $s = \sigma + it$ with $\sigma > 1$, we get

$$
\operatorname{Re} \sum_{y<p\le x} \frac{1-p^{i\Delta}}{p} = \log \left| \frac{\zeta(\sigma_1)\zeta(\sigma_2 - i\Delta)}{\zeta(\sigma_2)\zeta(\sigma_1 - i\Delta)} \right| + O(1),
$$

where

$$
\sigma_1 = 1 + \frac{1}{\log x}, \qquad \sigma_2 = 1 + \frac{1}{\log y} = 1 + \frac{1}{(\log x)^{1/10}}.
$$

Since, for $1 < \sigma \le 2$ and $|t| \le 2$,

$$
\log|\zeta(\sigma + it)| = \log \left| \frac{1}{\sigma - 1 + it} \right| + O(1),
$$

we obtain in the case $|\Delta| \le 2$

$$
\begin{aligned}
\operatorname{Re} \sum_{y<p\le x} \frac{1-p^{i\Delta}}{p} &= \log \left| \frac{1 - i\Delta/(\sigma_1 - 1)}{1 - i\Delta/(\sigma_2 - 1)} \right| + O(1) \qquad (4.2)\\
&= \log \left| \frac{1 - i\Delta \log x}{1 - i\Delta (\log x)^{1/10}} \right| + O(1).
\end{aligned}
$$

If $|\Delta| \le (\log x)^{-1/10}$, then (4.1) and (4.2) imply $|\Delta| \ll e^{2K}/\log x$, which is the asserted bound. But if $(\log x)^{-1/10} < |\Delta| \le 2$, then using (4.1), (4.2) and our assumption $K \le \frac{1}{20} \log\log x$, we conclude

$$
\tfrac{9}{10} \log\log x \le 2K + O(1) \le \tfrac{1}{10} \log\log x + O(1),
$$

which is a contradiction, provided x is sufficiently large, as we may assume.

It remains to show that the case $|\Delta| > 2$ cannot occur. Here we use instead of (4.2) the bound

$$
\begin{aligned}
\operatorname{Re} \sum_{y<p\le x} \frac{1-p^{i\Delta}}{p} &\ge \log \frac{\log x}{\log y} + O(1) - \operatorname{Re} \sum_{p\le x} \frac{p^{i\Delta}}{p} - \sum_{p\le y} \frac{1}{p} \\
&= \log \frac{\log x}{(\log y)^2} - \log|\zeta(\sigma_1 - i\Delta)| + O(1),
\end{aligned}
$$

where σ_1 is defined as before. By Vinogradov's bound (cf. [**12**, p. 70])

$$|\varsigma(\sigma+it)| \ll (\log|t|)^{2/3} \qquad (\sigma \geq 1,\ |t| \geq 2)$$

and the definition of y, the last expression is

$$\begin{aligned} &\geq \tfrac{4}{5}\log\log x - \tfrac{2}{3}\log\log|\Delta| + O(1) \\ &\geq \tfrac{2}{15}\log\log x + O(1), \end{aligned}$$

since $2 < |\Delta| = |\alpha_1 - \alpha_2| \leq 2x$ by hypothesis. Combining this with (4.1), we obtain

$$\tfrac{2}{15}\log\log x \leq 2K + O(1) \leq \tfrac{1}{10}\log\log x + O(1),$$

which yields the desired contradiction for sufficiently large x. Hence the lemma is proved.

The next lemma constitutes the principal tool in the proof of Theorem 3. It states roughly that $m(x)$ is close to $g(p)m(x/p)$ on average over primes $p \leq \sqrt{x}$, the error being measured in terms of $d(x)$. A result of this type was first proved in [**10**].

LEMMA 3. *Let $x \geq z^2 \geq 1$. Then we have*

$$\sum_{z<p\leq\sqrt{x}} \frac{1}{p}\left|m(x) - g(p)m\left(\frac{x}{p}\right)\right|^2 \ll \frac{1}{z} + zd(x).$$

PROOF. The estimate can be derived from the arithmetic version of the large sieve. We omit the details, since the argument is essentially the same as in the proof of Lemma 1 in [**10**].

LEMMA 4. *Given $x \geq 1$ there exists a real number $\alpha = \alpha(g;x)$ satisfying*

$$|\alpha| \ll 1/|m(x)|, \tag{4.3}$$

such that for $1 \leq t \leq x^{1/30}$ we have

$$m\left(\frac{x}{t}\right) = m(x)t^{i\alpha} + O\left(\max_{x^{2/3}\leq y\leq x} d(y)^{1/4} + x^{-1/100}\right). \tag{4.4}$$

PROOF. The proof depends on Lemma 3 and the following result, which is Lemma 2 in [**11**]:

Let $Q \geq 2$, $K \geq 10$ and a_p, $Q < p \leq Q^K$, be complex numbers of modulus 1. Let $\phi(t)$ be a complex-valued function defined for $1 \leq t \leq Q^{2K}$, and suppose that for some $\varepsilon > 0$ the inequalities

$$|\phi(t') - \phi(t)| \leq \varepsilon \qquad (1 \leq t \leq t' \leq t(1+Q^{-1/3}) \leq Q^{2K}) \tag{4.5}$$

and

$$\sum_{Q<p\leq Q^K} \frac{1}{p}|\phi(t) - a_p\phi(tp)| \leq \varepsilon\log K \qquad (1 \leq t \leq Q^K) \tag{4.6}$$

hold. Then there exists a real number α satisfying $|\alpha| \leq Q$ and such that

$$\phi(t) = \phi(1)t^{i\alpha} + O(\varepsilon) \qquad (1 \leq t \leq Q), \tag{4.7}$$

where the O-constant is absolute.

We apply this result with $Q = x^{1/30}$, $K = 10$, $a_p = g(p)$ and $\phi(t) = m(x/t)$. The trivial estimate

$$\begin{aligned} m(y(1+h)) &= \frac{1}{1+h} m(y) + O\left(\frac{1}{y} \sum_{y<n\le y(1+h)} 1\right) \\ &= m(y) + O\left(\frac{1}{y} + h\right) \qquad (y \ge 1,\ 0 \le h \le 1) \end{aligned} \tag{4.8}$$

shows that hypothesis (4.5) holds for any ε satisfying $\varepsilon \ge c_1 x^{-1/100}$ with a sufficiently large (absolute) constant c_1. Moreover, by Cauchy's inequality, the left-hand side of (4.6) is

$$\le \left(\sum_{x^{1/30}<p\le x^{1/3}} \frac{1}{p} \left| m\left(\frac{x}{t}\right) - g(p) m\left(\frac{x}{tp}\right)\right|^2 \right)^{1/2} \left(\sum_{x^{1/30}<p\le x^{1/3}} \frac{1}{p}\right)^{1/2},$$

and applying Lemma 3 with x/t in place of x and $z = \min(d(x/t)^{-1/2},\ x^{1/30})$, we get the bound

$$\ll \max_{x^{2/3}\le y\le x} d(y)^{1/4} + x^{-1/60}$$

for $1 \le t \le x^{1/3}$. Therefore, if we put

$$\varepsilon = c_2 \left(\max_{x^{2/3}\le y\le x} d(y)^{1/4} + x^{-1/100}\right)$$

with a sufficiently large constant c_2, then both (4.5) and (4.6) are satisfied, and we obtain from (4.7) the asserted estimate (4.4) with a suitable real number α satisfying $|\alpha| \le x^{1/30}$.

It remains to prove the bound (4.3) for α. We may assume that x is sufficiently large and $|\alpha| \ge 1$, for otherwise (4.3) holds for trivial reasons. Applying (4.4) with $t = 1 + 1/|\alpha| \le 2$ and taking into account (4.8), we obtain

$$m(x) + O\left(\frac{1}{|\alpha|}\right) = m\left(\frac{x}{1+1/|\alpha|}\right) = m(x)\left(1 + \frac{1}{|\alpha|}\right)^{i\alpha} + O(\varepsilon).$$

This relation can only hold if either $|m(x)| \ll 1/|\alpha|$ or $|m(x)| \ll \varepsilon$. In the first case (4.3) follows. In the second case both sides of (4.4) are of order $O(\varepsilon)$. Hence (4.4) remains valid for any real number α, and we can trivially satisfy (4.3) and (4.4) by taking $\alpha = 0$ in this case.

LEMMA 5. *Let $x \ge 2$ and suppose that* (2.8) *holds for some δ satisfying* (2.9). *Then we have, for $x^{3\delta/2} \le y \le x$,*

$$|m(y)| = |m(x)| + O(\delta^{1/5}).$$

PROOF. By (2.8) and (2.9) we have, for $x^{3\delta/2} \le x_1 \le x$,

$$\max_{x_1^{2/3}\le y\le x_1} d(y)^{1/4} + x_1^{-1/100} \ll \delta^{1/4}.$$

Lemma 4 therefore yields

$$|m(y)| = |m(x_1)| + O(\delta^{1/4}),$$

whenever $x^{3\delta/2} \le x_1 \le x$ and $x_1^{29/30} \le y \le x_1$. The result follows by iterating this estimate.

LEMMA 6. *Let $x \geq 2$ and suppose that* (2.8) *holds for some δ satisfying* (2.9). *Suppose further that the function g satisfies* (2.7). *Then we have*

$$\exp\left(-\tfrac{1}{2}s(x)\right) \ll |m(x)| + \delta^{1/5}. \tag{4.9}$$

PROOF. We define the Dirichlet series

$$G(s) = \sum_{n\geq 1} \frac{g(n)}{n^s} = \prod_p \left(1 + \sum_{m\geq 1} \frac{g(p^m)}{p^{ms}}\right).$$

Setting $\sigma = 1 + 1/(\sqrt{\delta}\log x)$, we have

$$|(\sigma-1)G(\sigma)| \gg \left|\frac{G(\sigma)}{\varsigma(\sigma)}\right| = \prod_p \left|\left(1 + \sum_{m\geq 1} \frac{g(p^m)}{p^{m\sigma}}\right)\left(1 - \frac{1}{p^\sigma}\right)\right|.$$

Here the product taken over the primes $p \geq 3$ can be written as

$$\begin{aligned}
&\exp\left\{-\mathrm{Re}\sum_{p\geq 3} \frac{1-g(p)}{p^\sigma} + O(1)\right\} \\
&\quad = \exp\left\{-\mathrm{Re}\sum_{p\leq x^{\sqrt{\delta}}} \frac{1-g(p)}{p} + O(1)\right\} \\
&\quad = \exp\left\{-\frac{1}{2}\sum_{p\leq x^{\sqrt{\delta}}} \frac{|1-g(p)|^2}{p} + O(1)\right\},
\end{aligned}$$

and hence is bounded by

$$\gg \exp\left(-\tfrac{1}{2}s(x)\right).$$

Moreover, in view of assumption (2.7), the factor $p = 2$ in this product is

$$\begin{aligned}
&\geq \left|1 + \sum_{m=1}^{\infty} \frac{g(2^m)}{2^m}\right| - \sum_{m=1}^{\infty}\left(\frac{1}{2^m} - \frac{1}{2^{m\sigma}}\right) \\
&= \left|1 + \sum_{m=1}^{\infty} \frac{g(2^m)}{2^m}\right| + O(\sigma - 1) \\
&\geq \frac{1}{4} + O\left(\frac{1}{\sqrt{\delta}\log x}\right) = \tfrac{1}{4} + O\left(\frac{1}{(\log x)^{3/4}}\right),
\end{aligned}$$

and therefore is bounded away from zero, if x is sufficiently large, as we may assume. Thus, the left-hand side of (4.9) is $\ll |(\sigma-1)G(\sigma)|$, and it remains to show that

$$|(\sigma-1)G(\sigma)| \ll |m(x)| + \delta^{1/5}. \tag{4.10}$$

By partial summation we get

$$|G(\sigma) = \left|\sigma\int_1^\infty m(y)\frac{dy}{y^\sigma}\right| \ll \int_1^\infty |m(y)|\frac{dy}{y^\sigma}.$$

We estimate the integral using the bound

$$|m(y)| \ll |m(x)| + \delta^{1/5}$$

from Lemma 5, if $x^{3\delta/2} \le y \le x$, and the trivial bound $|m(y)| \le 1$ otherwise. This gives

$$
\begin{aligned}
|G(\sigma)| &\ll \int_{x^{3\delta/2}}^{x} (|m(x)| + \delta^{1/5}) \frac{dy}{y^\sigma} + \left(\int_1^{x^{3\delta/2}} + \int_x^\infty \right) \frac{dy}{y^\sigma} \\
&\le \frac{1}{\sigma - 1} (|m(x)| + \delta^{1/5}) + \log x^{3\delta/2} + \frac{x^{1-\sigma}}{\sigma - 1} \\
&= \frac{1}{\sigma - 1} \{ |m(x)| + \delta^{1/5} + \frac{3}{2}\sqrt{\delta} + e^{-1/\sqrt{\delta}} \},
\end{aligned}
$$

which proves (4.10) and hence the lemma.

5. Proof of Theorem 3: Completion. In this section we fix a multiplicative function g of modulus 1 and numbers $x \ge 2$ and $\delta > 0$, for which hypotheses (2.7), (2.8) and (2.9) of Theorem 3 are satisfied. We may assume that $\delta \le \delta_0$ and $x \ge x_0$, where δ_0 and x_0 are arbitrary, but fixed positive constants, for if $\delta > \delta_0$ or $x < x_0$ then the conclusion of the theorem is trivially valid with a sufficiently large constant c.

We suppose that (2.11) fails, so that, with the notation introduced in the previous section,

$$
s(x) < \tfrac{1}{3} \log(1/\delta). \tag{5.1}
$$

We shall show that this implies (2.10) with an appropriate choice of the constant c. More precisely, we shall show that if $\alpha = \alpha(g; x)$ is defined as in Lemma 4, then, assuming $\delta \le \delta_0$ and $x \ge x_0$, we have

$$
|\alpha| \le 1/\delta \tag{5.2}
$$

and

$$
s_\alpha(x) \ll 1, \tag{5.3}
$$

the implied constant being absolute. This clearly implies (2.10).

For the proof of (5.2) we note that by (4.3) $|\alpha| \ll 1/|m(x)|$, while Lemma 6 and (5.1) imply, for $\delta \le \delta_0$,

$$
|m(x)| \gg \exp\left(-\tfrac{1}{2} s(x)\right) > \delta^{1/6}. \tag{5.4}
$$

Combining these two estimates yields (5.2) for sufficiently small δ.

The proof of (5.3) rests mainly on Lemma 3. The basic idea is to replace in the estimate of Lemma 3 $m(x/p)$ by $m(x)p^{i\alpha}$, then factor out $m(x)$ on the left-hand side, and bound $|m(x)|$ from below using (5.4). However, if we would do so immediately, estimating the difference $m(x/p) - m(x)p^{i\alpha}$ by means of Lemma 4, we would get too large an error term. We shall therefore use instead of $m(y)$ a smoother function $\overline{m}(y)$ which satisfies essentially the same estimate as $m(y)$, but for which $\overline{m}(x/p)$ is sufficiently close to $\overline{m}(x)p^{i\alpha}$ for the argument to go through. We define $\overline{m}(y)$ by

$$
\overline{m}(y) = \frac{1}{\log x_1} \int_1^{x_1} m\left(\frac{y}{t}\right) t^{-i\alpha} \frac{dt}{t},
$$

where $x_1 = x^{1/60}$. From estimate (4.4) of Lemma 4 we obtain, in view of the assumptions (2.8) and (2.9),

$$\overline{m}(x) = \frac{1}{\log x_1}\int_1^{x_1}(m(x)+O(\delta^{1/4}))\frac{dt}{t} = m(x)+O(\delta^{1/4}), \tag{5.5}$$

so that by (5.4)

$$|\overline{m}(x)| \gg \exp\left(-\tfrac{1}{2}s(x)\right) > \delta^{1/6}, \tag{5.6}$$

provided $\delta \le \delta_0$. Moreover, if $1 \le y \le x_1$, then

$$\begin{aligned}\overline{m}\left(\frac{x}{y}\right) &= \frac{1}{\log x_1}\int_1^{x_1} m\left(\frac{x}{yt}\right)t^{-i\alpha}\frac{dt}{t} = \frac{y^{i\alpha}}{\log x_1}\int_y^{yx_1} m\left(\frac{x}{t}\right)t^{-i\alpha}\frac{dt}{t}\\ &= \overline{m}(x)y^{i\alpha} + O\left(\frac{\log y}{\log x_1}\max_{1\le t\le yx_1}\left|m\left(\frac{x}{t}\right)\right|\right)\\ &= \overline{m}(x)y^{i\alpha} + O\left(\frac{\log y}{\log x_1}|\overline{m}(x)|\right),\end{aligned} \tag{5.7}$$

where the last estimate follows from Lemma 5, (5.5) and (5.6).

We now show that, for $1 \le z \le x^{1/3}$,

$$\sum_{z<p\le x^{1/3}}\frac{1}{p}\left|\overline{m}(x) - g(p)\overline{m}\left(\frac{x}{p}\right)\right|^2 \ll \frac{1}{z} + z\delta. \tag{5.8}$$

Thus, the function $\overline{m}(y)$ satisfies an estimate of the same type as that of Lemma 3. The proof of (5.8) rests on the inequality

$$\begin{aligned}\left|\overline{m}(x) - g(p)\overline{m}\left(\frac{x}{p}\right)\right|^2 &= \left|\frac{1}{\log x_1}\int_1^{x_1}\left(m\left(\frac{x}{t}\right) - g(p)m\left(\frac{x}{tp}\right)\right)t^{-i\alpha}\frac{dt}{t}\right|^2\\ &\le \frac{1}{\log x_1}\int_1^{x_1}\left|m\left(\frac{x}{t}\right) - g(p)m\left(\frac{x}{tp}\right)\right|^2\frac{dt}{t}.\end{aligned}$$

This gives for the left-hand side of (5.8) the bound

$$\ll \frac{1}{\log x_1}\int_1^{x_1}\sum_{z<p\le x^{1/3}}\frac{1}{p}\left|m\left(\frac{x}{t}\right) - g(p)m\left(\frac{x}{tp}\right)\right|^2\frac{dt}{t},$$

which by Lemma 3 and (2.8) is

$$\ll \frac{1}{z} + z\max_{1\le t\le x_1} d\left(\frac{x}{t}\right) \le \frac{1}{z} + z\delta,$$

as required.

By (5.7) we have, for $p \le x_1$,

$$\begin{aligned}|\overline{m}(x)|^2|1 - g(p)p^{i\alpha}|^2 &\ll \left|\overline{m}(x) - g(p)\overline{m}\left(\frac{x}{p}\right)\right|^2 + \left|\overline{m}\left(\frac{x}{p}\right) - \overline{m}(x)p^{i\alpha}\right|^2\\ &\ll \left|\overline{m}(x) - g(p)\overline{m}\left(\frac{x}{p}\right)\right|^2 + \frac{\log^2 p}{\log^2 x_1}|\overline{m}(x)|^2,\end{aligned}$$

which together with (5.8) yields

$$|\overline{m}(x)|^2 \sum_{z<p\le x_1} \frac{|1-g(p)p^{i\alpha}|^2}{p} \ll \frac{1}{z} + z\delta + |\overline{m}(x)|^2$$

for $1 \le z \le x_1$. If we now restrict z further by

$$1 \le z \le \delta^{-1/3}, \tag{5.9}$$

then this estimate simplifies to

$$\sum_{z<p\le x_1} \frac{|1-g(p)p^{i\alpha}|^2}{p} \ll \frac{1}{z}|\overline{m}(x)|^{-2} + 1,$$

and estimating $|\overline{m}(x)|$ by (5.6), we obtain

$$\sum_{z<p\le x_1} \frac{|1-g(p)p^{i\alpha}|^2}{p} \ll \frac{1}{z}\exp(s(x)) + 1. \tag{5.10}$$

Note that, by (2.9), we have

$$\delta^{-1/3} \le (\log x)^{1/6} \le x^{1/60} = x_1,$$

if x is sufficiently large (as we may assume), so that (5.9) is indeed stronger than the condition $1 \le z \le x_1$.

The remainder of the argument is based on the estimate (5.10). We first apply (5.10) with $z = z_0 = \delta^{-1/3}$. Using the bound (5.6), we obtain

$$\sum_{z_0<p\le x_1} \frac{|1-g(p)p^{i\alpha}|^2}{p} \ll 1 + \frac{\exp(s(x))}{z_0} \ll 1 + \frac{\delta^{-1/3}}{\delta^{-1/3}} \ll 1 \tag{5.11}$$

and hence

$$s_\alpha(x) \ll 1 + \sum_{p\le z_0} \frac{1}{p} \ll \log\log\left(\frac{1}{\delta}\right). \tag{5.12}$$

This bound falls short of the desired bound (5.3), but it is stronger than the bound (5.1) for $s(x)$, we started out with. We shall therefore repeat the above argument with the function $g(n)n^{i\alpha}$ in place of $g(n)$.

We first check that hypotheses (2.7), (2.8) and (5.1), which we have used, remain satisfied for the function $g(n)n^{i\alpha}$. In the case of the first hypothesis, this is clear. Moreover, by Lemma 1 and (5.2) we have, for $x^\delta \le y \le x$,

$$\begin{aligned} d_\alpha(y) &= d(y) + O\left(\left|\frac{\alpha}{y}\right| \log\left(\left|\frac{y}{\alpha}\right| + 2\right)\right) \\ &= d(y) + O\left(\frac{\log x}{x^\delta \delta}\right) = d(y) + O\left(\frac{1}{\log x}\right). \end{aligned}$$

Thus, in view of (2.9), hypothesis (2.8) remains valid for $g(n)n^{i\alpha}$, if we replace δ by $\delta' = 2\delta$, say, and assume that $x \ge x_0$, as we may. Finally, by (5.12), condition (5.1) holds also with $s_\alpha(x)$ in place of $s(x)$ and δ' in place of δ, provided that $\delta \le \delta_0$.

Arguing as before, we therefore conclude that for some real number α' satisfying

$$|\alpha'| \le 1/\delta' = 1/2\delta \tag{5.2$'$}$$

we have

$$\sum_{z<p\leq x_1} \frac{|1-g(p)p^{i(\alpha+\alpha')}|^2}{p} \ll \frac{1}{z}\exp(s_\alpha(x)) + 1, \tag{5.10$'$}$$

under the condition

$$1 \leq z \leq (\delta')^{-1/3} = (2\delta)^{-1/3}. \tag{5.9$'$}$$

Moreover, with $z_0' = (\delta')^{-1/3}$ we have

$$\sum_{z_0'<p\leq x_1} \frac{|1-g(p)p^{i(\alpha+\alpha')}|^2}{p} \ll 1. \tag{5.11$'$}$$

By Lemma 2 we deduce from (5.11) and (5.11)$'$ that

$$|\alpha'| \ll 1/(\log x_1). \tag{5.13}$$

Note that the hypotheses of Lemma 2 are satisfied, if x is sufficiently large, since then

$$z_0' \leq z_0 = \delta^{-1/3} \leq (\log x)^{1/6} \leq \exp(\log x_1)^{1/10}$$

and

$$|\alpha| \leq \frac{1}{\delta} \leq (\log x)^{1/2} \leq x_1, \qquad |\alpha+\alpha'| \leq \frac{1}{\delta} + \frac{1}{\delta'} \leq x_1.$$

In view of (5.13) we can replace α' by 0 in (5.10)$'$, the error being

$$\ll \sum_{p\leq x_1} \frac{|p^{i\alpha'}-1|^2}{p} \ll |\alpha'|^2 \sum_{p\leq x_1} \frac{\log^2 p}{p} \ll |\alpha'|^2 \log^2 x_1 \ll 1.$$

We therefore obtain, for z satisfying (5.9)$'$,

$$\sum_{z<p\leq x_1} \frac{|1-g(p)p^{i\alpha}|^2}{p} \ll \frac{1}{z}\exp(s_\alpha(x)) + 1. \tag{5.14}$$

We now choose z in the interval (5.9)$'$ such that

$$\left| \sum_{z<p\leq x_1} \frac{|1-g(p)p^{i\alpha}|^2}{p} - \frac{1}{2}s_\alpha(x) \right| \leq 2. \tag{5.15}$$

In view of (5.11), this is possible unless $s_\alpha(x) \ll 1$, in which case there is nothing to prove. We then have

$$\frac{1}{2}s_\alpha(x) \leq \sum_{p\leq z} \frac{4}{p} + O(1) \leq 4\log\log(z+2) + O(1),$$

so that

$$\exp(s_\alpha(x)) \ll (\log(z+2))^8. \tag{5.16}$$

Using estimates (5.15) and (5.16) in (5.14) we obtain

$$\frac{1}{2}s_\alpha(x) \ll \frac{\exp(s_\alpha(x))}{z} + ` \ll \frac{(\log(z+2))^8}{z} + 1$$

and hence the desired bound $s_\alpha(x) \ll 1$.

This completes the proof of Theorem 3.

6. Proof of Theorem 1: Necessity. We shall need the following variant of Theorem 3.

THEOREM 3*. *There exists a constant c with the following property. Let g, x and δ satisfy the hypotheses of Theorem 3. Set $K = K(\delta) = \exp(1/2\delta)$. Then we have either*

$$\min_{|\alpha|\le\log(Kx)} s_\alpha(x) < c \tag{6.1}$$

or

$$\min_{|\alpha|\le\log x} s_\alpha(x) \ge \frac{1}{3}\log(1/2\delta). \tag{6.2}$$

PROOF. As in the proof of Theorem 3 we may assume that δ is sufficiently small and x sufficiently large. We apply Theorem 3 to the functions $g(n)n^{i\alpha}$, $|\alpha| \le \log x$. By the hypotheses of Theorem 3*, the function $g(n)$ satisfies (2.7) and (2.8). The functions $g(n)n^{i\alpha}$, $|\alpha| \le \log x$, then satisfy (2.7) as well and, by Lemma 1 and (2.9), also (2.8), provided we replace δ by $\delta' = 2\delta$, say, and assume that x is sufficiently large. Hence, by Theorem 3 we have for each α, $|\alpha| \le \log x$, either

$$\min_{|\alpha'|\le 1/\delta'} s_{\alpha+\alpha'}(x) < c \quad \text{or} \quad s_\alpha(x) \ge \tfrac{1}{3}\log(1/\delta').$$

If the first alternative holds for some α, $|\alpha| \le \log x$, then we obtain (6.1), and if the second alternative holds for all such α, then (6.2) is satisfied. Thus, in any case, either (6.1) or (6.2) must be satisfied, as asserted in Theorem 3*.

PROOF OF THEOREM 1: NECESSITY. We fix a real-valued additive function f and suppose that the distributions (1.2) converge towards a limit distribution. Then the characteristic functions

$$\phi_x(t) = \frac{1}{[x]}\sum_{n\le x} e^{it(f(n+1)-f(n))} \tag{6.3}$$

converge pointwise to the characteristic function of the limit distribution, and the convergence is uniform on every bounded t-interval. In particular, given any $\varepsilon > 0$ there exist numbers $x_0(\varepsilon) \ge 1$ and $t_0(\varepsilon) > 0$ such that

$$|\phi_x(t) - 1| \le \varepsilon \qquad (x \ge x_0(\varepsilon),\ |t| \le t_0(\varepsilon)). \tag{6.4}$$

It is this property that we shall use to obtain conditions (1.3) of the theorem.

For $x \ge 2$ and real α and t we define

$$s_\alpha(t;x) = \sum_{p\le x} \frac{|1 - e^{itf(p)}p^{i\alpha}|^2}{p}.$$

We divide the proof into three steps, given in the form of propositions. The first of these is the most difficult one and depends on Theorem 3*.

PROPOSITION 1. *There exist numbers $x_0 \ge 2$ and $t_0 > 0$ such that if $x \ge x_0$ and $|t| \le t_0$ then $s_\alpha(t;x) \le c$ holds for some real number $\alpha = \alpha(t;x)$ satisfying $|\alpha| \ll \log x$. Here c is the constant of Theorem 3*.*

PROOF. We fix a number $\delta > 0$ sufficiently small to ensure that

$$\tfrac{1}{3}\log(1/2\delta) \ge c + 1, \tag{6.5}$$

where c is the constant of Theorem 3*, and set $K = \exp(1/2\delta)$. We shall apply Theorem 3* with this δ and the functions $g(n) = e^{itf(n)}$. For sufficiently small t we have

$$\sum_{m\geq 1} \frac{|1 - e^{itf(2^m)}|}{2^m} \leq \frac{1}{4}$$

and hence, for any real α,

$$\left|1 + \sum_{m\geq 1} \frac{e^{itf(2^m)}}{2^{m(1+i\alpha)}}\right| \geq \left|1 + \sum_{m\geq 1} \frac{1}{2^{m(1+i\alpha)}}\right| - \frac{1}{4} > \frac{1}{4}.$$

Thus, hypothesis (2.7) holds for $g(n) = e^{itf(n)}$, if t is sufficiently small. To obtain the second hypothesis (2.8), we note that, by Cauchy's inequality,

$$\begin{aligned} &\frac{1}{y} \sum_{n\leq y} |e^{itf(n+1)} - e^{itf(n)}| \\ &\leq \left(\frac{1}{[y]} \sum_{n\leq y} |e^{itf(n+1)} - e^{itf(n)}|^2 \right)^{1/2} \\ &= (2\mathrm{Re}(1 - \phi_y(t)))^{1/2}, \end{aligned}$$

where $\phi_y(t)$ is given by (6.3). By (6.4) this expression is $\leq \delta$, provided that $x \geq x_0(\delta^2/2)$ and $|t| \leq t_0(\delta^2/2)$. Thus (2.8) holds, if $x \geq x_0(\delta^2/2)^{1/\delta}$ and $|t| \leq t_0(\delta^2/2)$. The hypotheses of Theorem 3* are therefore satisfied for the function $g(n) = e^{itf(n)}$ for all sufficiently large x and sufficiently small t, say $x \geq x_0'$ and $|t| \leq t_0'$, and we conclude that for such x and t either

$$\min_{|\alpha|\leq \log(Kx)} s_\alpha(t;x) < c \tag{6.6}$$

or

$$\min_{|\alpha|\leq \log x} s_\alpha(t;x) \geq \tfrac{1}{3}\log(1/2\delta) \tag{6.7}$$

holds.

Now set

$$\psi_x(t) = \min_{|\alpha|\leq \log(Kx)} s_\alpha(t;x).$$

The function $\psi_x(t)$ is obviously continuous in t, uniformly on any finite x-interval, and equal to zero at $t = 0$. Therefore, if we set

$$t_x = \begin{cases} \min\{|t| : \psi_x(t) = c\} & \text{if } \max_{|t|\leq t_0'} \psi_x(t) \geq c, \\ t_0' & \text{otherwise,} \end{cases}$$

then t_x is well defined and satisfies

$$0 < t_x \leq t_0' \tag{6.8}$$

and

$$\psi_x(t) < c \qquad (|t| < t_x). \tag{6.9}$$

Moreover, it is easy to see that t_x is bounded from below on any finite x-interval. We shall show that, with a suitable constant $x_0 \geq x_0'$, we have

$$t_x \geq t_{x/K} \qquad (x \geq x_0). \tag{6.10}$$

This implies

$$t_x \geq t_0 \quad (x \geq x_0) \qquad \text{with } t_0 = \inf_{x_0/K \leq x \leq x_0} t_x > 0.$$

By (6.9) and the definition of $\psi_x(t)$ it follows that

$$\min_{|\alpha| \leq \log(Kx)} s_\alpha(t; x) \leq c \qquad (x \geq x_0, \ |t| \leq t_0),$$

which is the assertion of Proposition 1.

It remains to prove (6.10). In view of (6.8), this holds trivially if $t_x = t_0'$. We may therefore assume that $t_x < t_0'$. Thus, by the definition of t_x, we have, for either $t = t_x$ or $t = -t_x$,

$$\min_{|\alpha| \leq \log(Kx)} s_\alpha(t; x) = \psi_x(t) = c.$$

Hence (6.6) is not satisfied for this value of t, and therefore (6.7) must hold. Now note that for any real α

$$\begin{aligned} s_\alpha(t; x) - s_\alpha\left(t; \frac{x}{K}\right) &= \sum_{x/K < p \leq x} \frac{|1 - e^{itf(p)} p^{i\alpha}|^2}{p} \\ &\leq 4 \sum_{x/K < p \leq x} \frac{1}{p} \leq \frac{1}{2}, \end{aligned}$$

provided x is sufficiently large. Therefore, using (6.7) and (6.5), we deduce

$$\begin{aligned} \psi_{x/K}(t) = \min_{|\alpha| \leq \log x} s_\alpha(t; x/K) &\geq \min_{|\alpha| \leq \log x} s_\alpha(t; x) - 1/2 \\ &\geq (1/3) \log(1/2\delta) - 1/2 \geq c + 1/2. \end{aligned}$$

In view of (6.9), this implies that $t_{x/K} < |t| = t_x$ and hence proves (6.10).

PROPOSITION 2. *Let t_0 be as in Proposition* 1. *Then for each t satisfying $|t| \leq t_0$ there exists a real number $\alpha = \alpha(t)$ such that the series $\sum_p |1 - e^{itf(p)} p^{i\alpha}|^2/p$ is convergent, and its sum bounded uniformly for $|t| \leq t_0$.*

PROOF. Let $\alpha(t; x)$ be defined as in Proposition 1. An application of Lemma 2 shows that if $x \geq x_0$ and $|t| \leq t_0$, then we have, for $x \leq x' \leq x^2$,

$$|\alpha(t; x) - \alpha(t; x')| \ll 1/(\log x).$$

It readily follows from this that the limit $\alpha(t) = \lim_{x \to \infty} \alpha(t; x)$ exists and satisfies

$$\alpha(t) = \alpha(t; x) + O(1/(\log x))$$

uniformly for $x \geq x_0$ and $|t| \leq t_0$. The last relation implies

$$s_{\alpha(t)}(t; x) = s_{\alpha(t; x)}(t; x) + O(1),$$

and since, by Proposition 1, $s_{\alpha(t;x)}(t; x) \leq c$, we obtain

$$\sum_p \frac{|1 - e^{itf(p)} p^{i\alpha(t)}|^2}{p} = \lim_{x \to \infty} s_{\alpha(t)}(t; x) \ll 1$$

uniformly in $|t| \leq t_0$. This proves Proposition 2.

PROPOSITION 3. *There exists a real number λ such that condition* (1.3) *is satisfied for $h = f - \lambda \log$.*

PROOF. This can be deduced from the result of Proposition 2 in a well-known manner; see, e.g., [**2**, p. 295].

With Proposition 3 the proof of the necessity part of Theorem 1 is complete.

REFERENCES

1. J. G. Babu, *Probabilistic methods in the theory of arithmetic functions*, Macmillan, New Delhi, 1978.
2. H. Delange, *Sur les functions arithmétiques multiplicatives*, Ann. Sci. École Norm. Sup. (3) **78** (1961), 273–304.
3. P. D. T. A. Elliott, *On the difference of additive functions*, Mathematika **24** (1977), 153–165.
4. ——, *Probabilistic number theory.* I, Springer-Verlag, New York, 1979.
5. ——, *Probabilistic number theory.* II, Springer-Verlag, New York, 1980.
6. ——, *Arithmetic functions and integer products*, Springer-Verlag, New York, 1985.
7. P. Erdős, *On the distribution function of additive functions*, Ann. of Math. **47** (1946), 1–20.
8. P. Erdős and M. Kac, *The Gaussian law of errors in the theory of additive number-theoretic functions*, Amer. J. Math. **62** (1940), 738–742.
9. P. Erdős and A. Wintner, *Additive arithmetical functions and statistical independence*, Amer. J. Math. **61** (1939), 713–721.
10. A. Hildebrand, *Multiplicative functions at consecutive integers*, Math. Proc. Cambridge Philos. Soc. **100** (1986), 229–236.
11. ——, *Multiplicative functions in short intervals*, Canad. J. Math. **39** (1987), 646–672.
12. A. Karacuba, *Principles of analytic number theory*, Nauka, Moscow, 1975. (Russian)
13. I. Kátai, *On a problem of P. Erdős*, J. Number Theory **21** (1970), 1–6.
14. J. Kubilius, *Probabilistic methods in the theory of numbers*, Transl. Math. Monographs, Vol. 11, Amer. Math. Soc., Providence, R.I., 1964.
15. E. Wirsing, *Characterization of the logarithm as an additive function*, Proc. Sympos. Pure Math., Vol. 20, Amer. Math. Soc., Providence, R.I., 1971, pp. 375–381.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF ILLINOIS, URBANA, ILLINOIS 61801