PACIFIC JOURNAL OF MATHEMATICS
Vol. 48, No. 1, 1973

ON ADDITIVE FUNCTIONS WHOSE LIMITING DISTRIBUTIONS POSSESS A FINITE MEAN AND VARIANCE

P. D. T. A. ELLIOTT

In this paper two characterizations are given of those additive arithmetic functions which possess a limiting distribution with a finite mean and variance. It turns out that the study of such functions fits naturally within the framework of the theory of Lambert series.

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

$$f(ab) = f(a) + f(b)$$

is satisfied. If in addition the relations

$$f(p) = f(p^2) = \cdots$$

hold for each prime power then we say that $f(n)$ is *strongly additive*. For clarity of exposition only we shall confine ourselves to the study of strongly additive functions in this paper.

For each real number $x \geqq 1$ we define the frequency function

$$\nu_x(n; f(n) < z) = x^{-1} \sum_{\substack{n \leqq x \\ f(n) < z}} 1 .$$

If as $x \to \infty$ these frequencies converge to a limiting distribution in the usual probabilistic sense then we say that $f(n)$ has a limiting distribution.

2. THEOREM. *For any (real valued) additive function $f(n)$ the following three propositions are equivalent:*

(i) *$f(n)$ has a limiting distribution with finite mean and variance.*

(ii) *The series*

$$\sum f(p)p^{-1} \text{ and } \sum f^2(p)p^{-1}$$

both converge.

(iii)

$$\limsup_{x \to \infty} x^{-1} \sum_{n \leqq x} f^2(n) < \infty$$

and

$$\lim_{x\to\infty} x^{-1} \sum_{n\leq x} f(n)$$

exist.

REMARK. The equivalence of Propositions (i) and (ii) is exactly what one should expect from the interpretation of $f(n)$ as the sum of independent random variables which take (respective) values $f(p)$ with probability p^{-1} and zero with probability $1 - p^{-1}$. More surprising, perhaps, is the fact that the hypothesis that $f(n)$ be additive improves the otherwise weak conditions (iii) to equivalence with (i). We shall (perhaps surprisingly) appeal to a result concerning Lambert series.

It will be clear that a form of theorem involving complex-valued additive functions could be proved if we confine our attention to the equivalence of Propositions (ii) and (iii).

3. *Proof that* (i) *implies* (ii).

We define the function

$$f^1(p) = \begin{cases} f(p) & \text{if} \quad |f(p)| < 1 \\ 1 & \text{otherwise} . \end{cases}$$

Then the Erdös-Wintner criterion (see for example Kubilius [3] Theorem 4.5 pp. 74–85) asserts that $f(n)$ possesses a limiting frequency (unrestricted) if and only if both of the series

$$\sum f'(p)p^{-1} \text{ and } \sum (f'(p))^2_{p^{-1}}$$

converge. Let $F(z)$ denote the limiting frequency guaranteed by (i). Then for any positive real number B such that $\pm B$ are continuity points of $F(z)$ we see that

$$x^{-1} \sum_{\substack{n\leq x \\ |f(n)|\leq B}} f^2(n) \longrightarrow \int_{|z|\leq B} z^2 dF(z) \, , \quad (x \longrightarrow \infty) \, .$$

Next, for any real $\varepsilon > 0$ there is a number A such that

$$\liminf_{x\to\infty} \nu_x(n; |f(n)| \leqq A) > 1 - \varepsilon \, .$$

From the Erdös-Wintner criterion we see that those primes q_j for which $|f(q_j)| \geqq 1$ are such that the series

$$\sum q_j^{-1}$$

converges. Let us denote the set of these primes by Q.

A straightforward application of the sieve of Eratosthenes shows that those integers which are prime to every q_j have a natural density. In fact we obtain

$$\nu_x(n;\, q_j \nmid n \forall j) \longrightarrow \prod_{i=1}^{\infty}\left(1 - \frac{1}{q_i}\right), \quad (x \longrightarrow \infty) .$$

Set α for this product, and let A be chosen so that the second of our two assertions above holds with $\varepsilon = \alpha/2$. Let the integers n_i run through all those integers n which satisfy both

$$|f(n)| \leqq A \text{ and } q_j \nmid n \forall j .$$

From what we have so far said it is clear that

$$\liminf_{x\to\infty} \nu_x(n;\, n = n_i \leqq x) \geqq \alpha/2 ,$$

and in particular we have

$$\nu_x(n;\, n = n_i \leqq x) \geqq \alpha/4$$

for all $x \geqq x_0$, say.

Consider the sum

$$S_x = \sum_{n_i q_j \leqq x}\sum{}' f(n_i q_j)^2$$

where $'$ denotes that the side condition $2A < |f(q_j)| \leqq B - A$ is to be satisfied.

From these restrictions a typical summand satisfies

$$f^2(n_i q_j) \geqq (|f(q_j)| - A)^2 \geqq \frac{1}{4} f^2(q_j)$$

so that

$$\begin{aligned} S_x &\geqq \frac{1}{4} \sum_{q_j^{-1} \leqq x} f^2(q_j) \sum_{n_i \leqq x q^{-1}} 1 \\ &\geqq \frac{1}{4} \sum_{q_j \leqq x x_0^{-1}} f^2(q_j) \frac{1}{4} \alpha \frac{x}{q_j} \end{aligned}$$

and therefore

$$\begin{aligned} \limsup_{x\to\infty} \sum_{q_j \leqq x} \frac{f^2(q_j)}{q_j} &\leqq \limsup_{x\to\infty} x^{-1} S_x \\ &\leqq \int_{|z| \leqq B} z^2 dF(z) \leqq \int_{-\infty}^{\infty} z^2 dF(z) . \end{aligned}$$

Since these inequalities hold for any sequence of suitable continuity points $\pm B$ which tend (in absolute value) to infinity, we deduce that for any $B > 0, x \geqq 0$

$$\sum_{q_j \leqq x} f^2(q_j) \leqq \int_{-\infty}^{\infty} z^2 dF(z)$$

where

$$2A < |f(q_j)| \leqq B - A$$

so that letting $B \to \infty$ and then x yields

$$\sum_{|f(q_j)| \geqq 2A} \frac{f^2(q_j)}{q_j} < \infty .$$

Moreover,

$$\sum_{1 \leqq |f(p)| \leqq 2A} \frac{|f(p)|}{p} \leqq 2A \sum_{j=1}^{\infty} \frac{1}{q_j} < \infty ,$$

and

$$\sum_{|f(p)|<1} \frac{f^2(p)}{p} < \infty$$

so that altogether the series

$$\sum f^2(p) p^{-1}$$

converges. The convergence of the second series in (ii) follows immediately.

Proof that (ii) *implies* (iii) *and* (i).

We begin with the remark that for any additive function, complex valued or otherwise, the Turan-Kubilius inequality (see for example Kubilius [3] pp. 31-35) asserts that for a suitable positive constant c

$$\sum_{n \leqq x} |f(n) - \sum_{p \leqq x} f(p) p^{-1}|^2 \leqq c \sum_{p \leqq x} |f^2(p)| p^{-1} , \quad (x \geqq 1) .$$

In our present circumstances the sums

$$\sum_{p \leqq x} f(p) p^{-1} \text{ and } \sum_{p \leqq x} f^2(p) p^{-1}$$

are uniformly bounded for all real values of x, so that

$$\begin{aligned} \sum_{n \leqq x} f^2(n) &\leqq 2 \sum_{n \leqq x} \Big(f(n) - \sum_{p \leqq x} f(p) p^{-1} \Big)^2 + 2x \Big(\sum_{p \leqq x} f(p) p^{-1} \Big)^2 \\ &= O(x) , \end{aligned}$$

and

$$\limsup_{x \to \infty} x^{-1} \sum_{n \leqq x} f^2(n) = D < \infty .$$

From the Erdös-Wintner criterion $f(n)$ possesses a limiting distribution $F(z)$, say. For each real number B such that $\pm B$ are continuity points of this limiting distribution, an application of Fatou's lemma yields

$$\int_{|z| \leqq B} z^2 dF(z) \leqq \liminf_{x \to \infty} x^{-1} \textstyle\sum_{n \leqq x} f^2(n) \leqq D .$$

Since B is otherwise arbitrary $F(z)$ has a finite second moment, and hence a finite mean and variance.

This completes the proof of (i).

Furthermore,

$$x^{-1} \sum_{\substack{n \leq x \\ |f(n)| \leq B}} f(n) \longrightarrow \int_{|z| \leq B} z dF(z) , \quad (x \to \infty) ,$$

whilst

$$\limsup_{x \to \infty} x^{-1} \sum_{\substack{n \leq x \\ |f(n)| > B}} |f(n)| \leqq B^{-1} \limsup_{x \to \infty} x^{-1} \sum_{n \leqq x} f^2(n) \leqq B^{-1} D$$

from which it follows trivially that as $x \to \infty$

$$x^{-1} \sum_{n \leqq x} f(n)$$

converges to the mean of $F(z)$.

This completes the proof of (iii).

Proof that (iii) *implies* (ii) (which will complete the proof of the theorem).

As one would expect this part of the proof takes a little more effort since we have to start, so to speak, from scratch. We recall that an additive function $f(n)$ is said to be *finitely distributed* if and only if there are two positive real numbers c_1 and c_2 so that for an unbound sequence of real numbers $x \geqq 1$ we can find at least $k \geqq c_2 x$ integers $1 \leqq a_1 < a_2 < \cdots < a_k \leqq x$ so that

$$|f(a_i) - f(a_j)| \leqq c_1$$

holds for every pair (a_i, a_j), $1 \leqq i, j \leqq k$. This concept was introduced by Erdös [1] who proved

LEMMA 1. *A function $f(n)$ is finitely distributed if and only if there is a constant c_3 and an additive function $g(n)$ so that*

$$f(n) = c_3 \log n + g(n) ,$$

where

$$\sum (g'(p))^2 p^{-1} < \infty .$$

There is an alternative proof, on somewhat different lines, given by Ryavec [4].

In our present circumstances we have

$$x^{-1}\sum_{n\leqq x} f^2(n) \leqq E$$

for all $x \geqq 2$ (say). Thus for any positive real number $A > E^{1/2}$,

$$\nu_x(n; |f(n)| \geqq A) \leqq EA^{-2} < 1\ , \quad (x \geqq 2)\ .$$

It follows from Lemma 1 that $f(n)$ is finitely distributed, and has the form

$$c_3 \log n + g(n)\ .$$

Let π denote the set of primes q_j on which $|g(q_j)| > A$. Let n_i run through those squarefree integers which are prime to each q_j. Since

$$\sum_{q\in\pi} q^{-1}$$

converges, a straightforward application of the sieve of Eratosthenes shows that

$$\nu_x(n; n = n_i \leqq x) \longrightarrow \prod_{i=1}^{\infty}\left(1 - \frac{1}{q_i + 1}\right) = \beta > 0\ , \quad (x \longrightarrow \infty)\ ,$$

say. For each integer n let $\nu(n)$ denote the number of distinct prime divisors of n. We next assume that $c_3 \neq 0$ and obtain a contradiction.

Let c_4 be sufficiently large that the inequality $A\nu(n) \leqq c_4 \log n$ holds for all integers $n \geqq 2$. Then for every real number $x \geqq 2$ we have

$$\begin{aligned} Ex &\geqq \sum_{n_i\leqq x} f^2(n_i) \geqq \sum_{n_i\leqq x} (c_3 \log n_i - A\nu(n_i))^2 \\ &= c_3^2 \sum_{n_i\leqq x} \log^2 n_i + O\Big(\log x \sum_{n\leqq x} \nu(n)\Big)\ . \end{aligned}$$

For all sufficiently large values of x the first of these two terms is

$$(1 + o(1))\beta c_3^2 x \log^2 x$$

whilst the second is at most $O(x \log x \operatorname{loglog} x)$. This clearly yields a contradiction. Hence $c_3 = 0$ and the additive function $f(n)$ satisfies

$$\sum_p (f'(p))^2 p^{-1} < \infty\ .$$

We now argue exactly as in the proof that the existence of a limiting distribution for $f(n)$ which has a finite variance implies that the series

$$\sum f^2(p)p^{-1}$$

converges, and deduce the same result.

It remains to secure the convergence of the series

$$\sum f(p)p^{-1} .$$

(We do not as yet know that a limiting distribution for $f(n)$ exists, although if we set $\alpha_n = \sum_{p \leqq n} f'(p)p^{-1}$ then we do know that $f(n) - \alpha_n$ has a limiting distribution. See, for example, Kubilius [3] Theorem 4.4 pp. 72–74.)

Consider the generating function

$$G(z) = \sum_{n=1}^{\infty} f(n)z^n .$$

If N is any positive integer and z is any complex number then by the Cauchy-Schwarz inequality

$$\begin{aligned} \left| \sum_{N<n \leqq 2N} f(n)z^n \right|^2 &\leqq \sum_{n \leqq 2N} f^2(n) \sum_{N<n \leqq 2N} |z|^{2n} \\ &\leqq EN^2 |z|^{2N} . \end{aligned}$$

It is easily seen that $G(z)$ is defined by an absolutely convergent series if z satisfies $|z| < 1$. By means of the representation

$$f(n) = \sum_{p|n} f(p)$$

we invert the order of summation to obtain:

$$G(z) = \sum_p f(p) \frac{z^p}{1 - z^p} .$$

Since

$$x^{-1} \sum_{n \leqq x} f(n) \longrightarrow A, \ (x \longrightarrow \infty) \ , \ \text{say} \ ,$$

it is readily established that for real values of z

$$G(z) \sim \frac{A}{1 - z} \ \text{as} \ z \longrightarrow 1 - .$$

We now appeal to a Tauberian theorem concerning Lambert series.

LEMMA 2. *Let a_n $n = 1, 2, \cdots$ be a series of real numbers, and define*

$$H(y) = \sum_{n=1}^{\infty} a_n \frac{nye^{-ny}}{1 - e^{-ny}}$$

for positive real values of y. Let $H(y) \to A$ as $y \to 0+$. Let the sum of the a_n be a slowly decreasing function in the sense of Hardy [2] §6.2 pp. 124–125, *that is if $x < y$ are real numbers, so that as $x \to \infty$ and $y \to \infty$ in such a manner that $y/x \to 1$, then*

$$\liminf_{x\to\infty} \sum_{x<n\leqq y} a_n \geqq 0 .$$

Then

$$\sum_{n\leqq x} a_n \longrightarrow A , \quad (x \longrightarrow \infty) .$$

REMARK. If the a_n are allowed to be complex then provided that we replace the condition of slowly decreasing by a condition of slow oscillation viz:

$$\lim_{x\to\infty} \sum_{x<n\leqq y} a_n = 0 ,$$

the same conclusion may be drawn. A proof of this lemma can be found in Hardy [2], Theorem 261, pp. 373-374.

In our present circumstances we set

$$a_n = \begin{cases} f(p)p^{-1} & \text{if} \quad n = p , \\ 0 & \text{otherwise} \end{cases}$$

and have established that

$$H(y) = yG(e^{-y}) \longrightarrow A , \quad (y \longrightarrow 0+) .$$

Moreover,

$$\Bigl(\sum_{x<n\leqq 2x} |a_n|\Bigr)^2 \leqq \sum_{x<n\leqq 2x} f^2(p)p^{-1} \sum_{x<n\leqq 2x} p^{-1} ,$$

so that since the series $\sum f^2(p)p^{-1}$ converges and

$$\sum_{x<n\leqq 2x} \frac{1}{p} = \log\Bigl(\frac{\log 2x}{\log x}\Bigr) + O((\log x)^{-1}) \leqq c_4 < \infty ,$$

we see that the condition of slow decreasing required for an application of Lemma 2 is satisfied.

We deduce that

$$\lim_{x\to\infty} \sum_{p\leqq x} \frac{f(p)}{p} = A = \lim_{x\to\infty} x^{-1} \sum_{n\leqq x} f(n) .$$

Moreover, by (ii) a limiting distribution exists for $f(n)$, which has the finite mean of value A.

This completes the proof of the theorem.

REMARK. The use of the Tauberian theorem in Lemma 2 is very convenient for the study of additive functions. If $f(p)$ assumes complex values the side condition $f(p) = O(\log p)$ will suffice in order for Lemma 2 to be applicable. This is a condition which is satisfied in nearly every case of number theoretical interest.

References

1. P. Erdös, *On the distribution of additive functions*, Ann. Math., **47** (1946), 1-20.
2. G. H. Hardy, *Divergent Series*, Oxford, 1949.
3. J. P. Kubilius, *Probabilistic Methods in the Theory of Numbers*, Amer. Math. Soc. Transl., Vol. 11, 1964.
4. C. A. Ryavec, *A characterization of finitely distributed additive functions*. J. Number Theory, **2** (4) (1970), 393-403.

Received May 22, 1972.

UNIVERSITY OF COLORADO