ACTA ARITHMETICA
XXIII (1973)

Some results on the distribution of additive arithmetic functions, II

by

G. JOGESH BABU (Calcutta)

Introduction. Let f be a real-valued additive arithmetic function. In this paper we characterize the spectrum of the distribution of $\{f(n)-f(n+1),\ldots,f(n+h-1)-f(n+h)\}$ whenever the above distribution exists, where h is a positive integer. We obtain a theorem of Erdös and A. Schinzel [3] as a corollary of one of our propositions. Under very general conditions we shall show that, for any $m\geqslant 1$, $\{f_1(F_1(m)),\ldots,f_h(F_h(m))\}$ belongs to the spectrum of the distribution of $\{f_1(F_1(n)),\ldots,f_h(F_h(n))\}$ if it exists, where $f_1,\ldots,f_h$ are real additive arithmetic functions and $F_1,\ldots,F_h$ are positive integer-valued polynomials. In the last section we give a sufficient condition for an additive arithmetic function to have a singular distribution. Finally we shall show, under fairly general conditions on F, that if the distributions of $f(n)$ and $f(F(m))$ exist (F is an integer-valued polynomial) and if the distribution of $f(n)$ is absolutely continuous, then the distribution of $f(F(m))$ is also absolutely continuous. At the end we shall give an example to show that this is the best possible result.

Notations and definitions. Define,

$P=\{F\colon F$ is an integer-valued polynomial of degree $\nu_F\geqslant 1$ which is not divisible by the square of any irreducible polynomial and $F(m)>0$ for $m=1,2,\ldots\}$.

Let $r(F,d)$ denote the number of incongruent solutions of the congruence relation $F(m)\equiv 0\pmod d$.

$p,q,\ldots$ denote prime numbers.

$\sum_p$ denote the sum over prime numbers.

Put

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

Results.

PROPOSITION 1. *Suppose that the series*

$$\sum_p \frac{[f'(p)]^2}{p}$$

is convergent. For any positive integer h,

$$\{f(n)-f(n+1), \ldots, f(n+h-1)-f(n+h)\} \tag{1}$$

has a distribution and for any $n_0 \geq 1$, the vector $\{f(n_0)-f(n_0+1), \ldots, f(n_0+h-1)-f(n_0+h)\}$ belongs to the spectrum of the distribution of (1).

Moreover, if $N_0, N_1, \ldots, N_h$ are positive integers such that for all $i = 0, 1, \ldots, h$, $(N_i, (h+1)!) = 1$ and $(N_i, N_j) = 1$ $(0 \leq i < j \leq h)$, then

$$\{f(N_0)-f(2N_1), f(2N_1)-f(3N_2), \ldots, f(hN_{h-1})-f((h+1)N_h)\}$$

is in the spectrum of the distribution of (1).

COROLLARY (Erdös and A. Schinzel [3]). *Let $f(n)$ be an additive arithmetic function satisfying the following conditions:*

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

2. *There is a number c_1 such that, for any integer $M > 0$, the set of numbers $f(N)$, where $(N, M) = 1$, is dense in (c_1, ∞).*

Then for any given sequence of h real numbers $a_1, a_2, \ldots, a_h$ and $\varepsilon > 0$ the set $\{n \geq 1: |f(n+i)-f(n+i-1)-a_i| < \varepsilon, i = 1, \ldots, h\}$ has positive natural density.

PROPOSITION 2. *Suppose that $F_1, \ldots, F_s$ belong to* **P**. *Suppose*

$$f_i(p^k) r(F_i, p^k) \to 0$$

as $p \to \infty$ for $k = 1, \ldots, \nu_{F_i}-1$ whenever $\nu_{F_i} \geq 2$. If, moreover, the distribution of

$$\{f_1(F_1(m)), \ldots, f_s(F_s(m))\} \tag{2}$$

exists, then one can find a K_0 such that the spectrum S of the distribution of (2) *is the closure of the set*

$$A = \Big\{\Big(\sum_{\substack{p^l \| F_1(m) \\ p \leq k}} f_1(p^l), \sum_{\substack{p^l \| F_2(m) \\ p \leq k}} f_2(p^l), \ldots, \sum_{\substack{p^l \| F_s(m) \\ p \leq k}} f_s(p^l)\Big): m \geq 1, k > k_0\Big\}.$$

Remark 1. Clearly $A \supset B = \big\{\big(f_1(F_1(m)), \ldots, f_s(F_s(m))\big): m \geq 1\big\}$.

Proofs.

Proof of Proposition 1. Let $H_{i-1}(n) = f(n+i-1)-f(n+i)$, $i = 1, \ldots, h$. We extend the functions H_i to the polyadic domain (see Novoselov, [8]) and show that for each i, $H_i \in \mathfrak{H}_0$ ([8]), proceeding as follows.

Let

$$\omega(p^k, x) = \begin{cases} 1 & \text{if} \quad p^k \| x, \\ 0 & \text{otherwise.} \end{cases}$$

For any prime number p define

$$f_{ip}(x) = \sum_{k=1}^{\infty} f(p^k)\,\omega(p^k, x+i), \qquad i = 0, 1, \ldots, h-1.$$

Since the random variables $\{f_{ip}(x):\ p \text{ is a prime}\}$ are mutually independent ([8]) and

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

by Kolmogorov's three series theorem, it follows that

$$\sum_p \left\{ f_{ip}(x) - \frac{f'(p)}{p} \right\}$$

converges almost everywhere for $i = 0, 1, \ldots, h$. Hence

$$\sum_p \{f_{ip}(x) - f_{(i+1)p}(x)\}$$

converges a.e. for $i = 0, 1, \ldots, h-1$. Moreover, it is easy to see that the random variables $[\{f_{ip}(x) - f_{(i+1)p}(x)\}:\ p$ is a prime$]$ are mutually independent random variables for each $i = 0, 1, \ldots, h-1$.

Let

$$g_i(x) = \begin{cases} \sum \{f_{ip}(x) - f_{(i+1)p}(x)\} & \text{if it converges,} \\ 0 & \text{otherwise.} \end{cases}$$

Clearly $g_i(x)$ is an extension of $H_i(n)$. By using the Turán–Kubilius inequality ([5]), it is easy to show that $H_i(n) \in \mathfrak{H}_0$ ([8]) and the distribution of $H_i(n)$ is $Q_i(c) = P\{x:\ g_i(x) < c\}$.

Note that for any h-tuple $(t_0, \ldots, t_{h-1})$ of real numbers the distribution of $\sum_{i=0}^{h-1} t_i H_i(n)$ is given by

$$\mathrm{P}\left\{x:\ \sum_{i=0}^{h-1} t_i g_i(x) < c\right\}.$$

Hence by the Cramer–Wold device ([4]) we find that the distribution of $\{H_0(n), H_1(n), \ldots, H_{n-1}(n)\}$ is given by

$$Q(c_0, \ldots, c_{h-1}) = \mathrm{P}\{x:\ g_i(x) < c_i,\ i = 0, \ldots, h-1\}.$$

Let $0 < \delta < 1$. Since

$$\{(f_{0p}(x)-f_{1p}(x), \ldots, f_{(h-1)p}(x)-f_{hp}(x)): p \text{ is a prime}\};$$

is a sequence of mutually independent random variables, by using Egoroff's theorem one can find a $H \subset \mathfrak{S}$ such that $P(H) > 1-\delta$ and $\sum_{p \leqslant r} \{f_{ip}(x) - -f_{(i+1)p}(x)\}$ converges uniformly on H for $i = 0, 1, \ldots, h-1$.

Now fix a positive integer n_0 and a real number $\varepsilon > 0$. Let $N = n_0(n_0+1)\ldots(n_0+h)$. Let k be any integer greater than N^2 and such that for $x \in H$

$$\Big|\sum_{p>k}\{f_{ip}(x)-f_{(i+1)p}(x)\}\Big| < \varepsilon \quad \text{for} \quad i = 0, \ldots, h-1.$$

Hence

$$\mathrm{P}\Big\{x: \Big|\sum_{p>k}[f_{ip}(x)-f_{(i+1)p}(x)]\Big| < \varepsilon \text{ for } i = 0, \ldots, h-1\Big\} > 1-\delta.$$

Now the density of

$$\{n \geqslant 1: |f(n+i-1)-f(n+i)-f(n_0+i)| < \varepsilon, i = 1, 2, \ldots, h\}$$

is greater than or equal to

$$\mathrm{P}\Big\{x: \Big|\sum_{p>k}[f_{(i-1)p}(x)-f_{ip}(x)]\Big| < \varepsilon \text{ and}$$
$$\sum_{p\leqslant k}[f_{(i-1)p}(x)-f_{ip}(x)] = f(n_0+i-1)-f(n_0+i) \text{ for } i = 1, \ldots, h\Big\} \geqslant$$

$$(1-\delta)\mathrm{P}\Big\{x: \sum_{p\leqslant k}[f_{(i-1)p}(x)-f_{ip}(x)] = f(n_0+i-1)-f(n_0+i);\ i = 1, \ldots, h\Big\}.$$

Put

$$P = \prod_{\substack{p\leqslant k\\ p\nmid N}} p, \qquad Q = N^2P.$$

$$\mathrm{P}\Big\{x: \sum_{p\leqslant k}[f_{ip}(x)-f_{(i+1)p}(x)] = f(n_0+i)-f(n_0+i+1);\ i = 0, \ldots, h-1\Big\}$$
$$= \text{Density of } \Big\{n \geqslant 1: \sum_{p\leqslant k}[f_{ip}(n)-f_{(i+1)p}(n)] = f(n_0+i)-f(n_0+i+1);$$
$$i = 0, 1, \ldots, h-1\Big\} \geqslant \frac{1}{Q} > 0.$$

In fact, since $(P, N) = 1$, we can find an l such that

$$l \equiv n_0 \pmod{N^2} \quad \text{and} \quad l \equiv 1 \pmod{P}.$$

It is easy to show that, for any integer t,

$$\frac{Qt+l+i}{n_0+i}, \quad i = 0, 1, \ldots, h,$$

is an integer not divisible by any prime $p \leqslant k$. Since $k > N^2$, we have

$$\left(\frac{Qt+l+i}{n_0+i},\ n_0+i\right) = 1.$$

Hence for any t such that $Qt+l > 0$, we get

$$\sum_{p\leqslant k}' \{f_{ip}(Qt+l) - f_{(i+1)p}(Qt+l)\} = f(Qt+l+i) - f(Qt+l+i+1),$$
$$i = 0, 1, \ldots, h-1.$$

But the density of the positive integers of the form $Q\varepsilon+l$ is equal to $1/Q$. This proves the first part of Proposition 1. The proof of the second part of Proposition 1 is similar to the above proof. So here we only note the following fact:

We put

$$N = N_0 N_1 \ldots N_h, \quad P = \prod_{\substack{p\leqslant k \\ p\nmid N}} p \quad \text{and} \quad Q = (h+1)!\, N^2 P.$$

Since $(N_i, (h+1)!) = 1$ for $i = 0, \ldots, h$ and $(N_i, N_j) = 1$ $(0 \leqslant i \leqslant j \leqslant h)$, it follows from the Chinese Remainder Theorem that there exists a number l satisfying the congruence relations

$$l \equiv 1 \left(\operatorname{mod} (h+1)!\,P\right),$$
$$l \equiv -i + N_i \,(\operatorname{mod} N_i^2) \quad (0 \leqslant i \leqslant h).$$

It is easy to see that for every integer t the numbers

$$\{(Qt+l+i)/(i+1)N_i\}, \quad i = 1, \ldots, h,$$

are integers which are not divisible by any prime $p \leqslant k$. Also the density of the integers $Qt+l$ is $1/Q > 0$.

This completes the proof of Proposition 1.

Proof of the Corollary. Let ε be a positive number and let a sequence a_i $(i = 1, \ldots, h)$ be given. By condition 2 we can find positive integers $N_0, N_1, \ldots, N_h$ such that

$$(N_i, (h+1)!) = 1 \ (i = 0, \ldots, h), \quad (N_i, N_j) = 1 \ (0 \leqslant i < j \leqslant h),$$

$$f(N_0) > c_1 + \max_{1\leqslant i\leqslant h} \left\{f(i+1) - \sum_{j=1}^{i} a_j\right\}$$

and

$$\left|f(N_i) - \left\{f(N_0) - f(i+1) + \sum_{j=1}^{i} a_j\right\}\right| < \varepsilon/4 \quad (1 \leqslant i \leqslant h).$$

Hence

$$(3) \qquad \left|f\left((i+1)N_i\right) - f(iN_{i-1}) - a_i\right| < \varepsilon/2 \quad (1 \leqslant i \leqslant h).$$

By Proposition 1, we have

$$(4)\quad \{n \geqslant 1: |f(n)-f(n+1)-f(N_0)+f(2N_1)| < \varepsilon/2, \ldots,$$
$$|f(n+h-1)-f(n+h)-f(hN_{h-1})+f((h+1)N_h)| < \varepsilon/2\}$$

has positive density. Hence the corollary follows from (3) and (4).

Proof of Proposition 2. We need the following two lemmas.

LEMMA 1. *If $h(m)$ and $g(m)$ are integer-valued polynomials having no common factors, then there exists a k_1 such that $p > k_1$ implies that there is no m such that $h(m) \equiv 0 \pmod p$ and $g(m) \equiv 0 \pmod p$.*

LEMMA 2. *If $F \in \mathbf{P}$, then there exists a k such that $p > k$ implies*

$$r(F, p^l) = r(F, p) \quad \text{for all } l \geqslant 1.$$

Also there exists a constant c such that $r(F, p^l) \leqslant c$ for all p and l.

For proofs of these lemmas see [9].

Let $F_i(m) = \prod_{j=1}^{l_i} F_{ij}(m)$, where $\{F_{ij}(m): j = 1, \ldots, l_i\}$ are irreducible and each $F_{ij} \in \mathbf{P}$. Such a factorization is possible and is unique.

Let $\{G_1, \ldots, G_h\} = \{F_{ij}: j = 1, \ldots, l_i, i = 1, \ldots, s\}$ such that G_i and G_j have no common factors if $i \neq j$. By Lemma 1 choose a k_1 such that $p > k_1$ implies that there is no m such that $G_i(m) \equiv 0 \pmod p$ and $G_j(m) \equiv 0 \pmod p$ $(1 \leqslant i < j \leqslant h)$. Let $G_i(x)$ be the continuous extension of $G_i(m)$ to Novoselov's space $\mathfrak{S}$.

It is easy to see that

$$\{(m_i \mid G_i(x);\ i = 1, \ldots, h), (p_1^{t_{1i}} \mid G_i(x), i = 1, \ldots, h), \ldots,$$
$$(p_r^{t_{ri}} \mid G_i(x),\ i = 1, \ldots, h)\}$$

are independent events if t_{ij} are non-negative integers, $r \geqslant 1, p_i > k_1$, $p_i \neq p_j$ if $i \neq j$ and m_i is not divisible by any prime $p > k_1$ $(i = 1, \ldots, h)$. Since either $F_{ij}(m) \equiv F_{rl}(m)$ or $F_{ij}(m)$ and $F_{rl}(m)$ are mutually prime, we infer that

$$\{(m_i \mid F_i(x), i = 1, \ldots, s), (p_1^{t_{1i}} \mid F_i(x), i = 1, \ldots, s), \ldots,$$
$$(p_l^{t_{li}} \mid F_i(x), i = 1, \ldots, s)\}$$

are independent events on Novoselov's space if $l \geqslant 1, t_{ij} \geqslant 0$, $p_i > k_1$, $p_i \neq p_j$ if $i \neq j$ and m_i is not divisible by any prime $p > k_1$ for any $i = 1, \ldots, s$.

Now choose $k_2 > k_1$ (by using Lemma 2) such that, if $p > k_2$, then

$$r(F_i, p^t) = r(F_i, p) \quad \text{for} \quad t \geqslant 1, i = 1, \ldots, s$$

and

$$(F_i, p) < p/2s, \quad i = 1, \ldots$$

We now show that $A \subset S$. Let

$$f_{i0}(x) = \sum_{\substack{p^k \| F_i(x) \\ p \leqslant k_0}} f_i(p^k), \quad i = 1, \ldots, s.$$

For $p > p_0$ and $i = 1, \ldots, s$, we put

$$f_{ip}(x) = \begin{cases} f_i(p^k) & \text{if } p^k \| F_i(x), k \geqslant 1, \\ 0 & \text{if either } p \nmid F_i(x) \text{ or } p^k | F_i(x) \text{ for all } k \geqslant 1. \end{cases}$$

By Theorem 2 of [1], we conclude that

$$\sum_p \frac{f_i'(p) r(F_i, p)}{p} \quad \text{and} \quad \sum_p \frac{(f_i'(p))^2 r(F_i, p)}{p}$$

converge. Hence by Kolmogorov's three series theorem $\sum_{p > k_0} f_{ip}(x)$ converges a.e.

Fix a positive real number $\delta < 1/4s$. By Egoroff's theorem choose $H \subset \mathfrak{S}$ such that $\mathrm{P}(H) > 1 - \delta$ and, on H, $\sum_{p > k_0} f_{ip}(x)$ converges uniformly for $i = 1, \ldots, s$.

Now fix $\varepsilon > 0$, $k > k_0$ and $m \geqslant 1$. Choose $k_2 > k$ such that

$$\mathrm{P}\Big\{x \colon \Big|\sum_{p > k_2} f_{ip}(x)\Big| < \varepsilon;\ i = 1, \ldots, s\Big\} > 1 - \eta \quad \text{where} \quad \eta = \delta s.$$

Let $\mathrm{D}\{\ldots\}$ denote the natural density of integers satisfying the conditions mentioned in $\{\ldots\}$.

$$\mathrm{D}\Big\{\Big|f_i(F_i(n)) - \sum_{k_0 < p \leqslant k} f_{ip}(m) - f_{i0}(m)\Big| < \varepsilon, i = 1, \ldots, s\Big\}$$

$$\geqslant \mathrm{P}\Big\{x \colon f_{i0}(x) = f_{i0}(m), \sum_{k_0 < p \leqslant k} f_{ip}(x) = \sum_{k_0 < p \leqslant k} f_{ip}(m)$$

$$\text{and } \Big|\sum_{p > k_2} f_{ip}(x)\Big| < \varepsilon,\ i = 1, \ldots, s\Big\}$$

$$\geqslant (1 - \eta)\mathrm{P}\{x \colon f_{i0}(x) = f_{i0}(m), i = 1, \ldots, s\} \times$$

$$\times \prod_{k_0 < p \leqslant k} \mathrm{P}\{x \colon f_{ip}(x) = f_{ip}(m), i = 1, \ldots, s\} \times \prod_{k < p \leqslant k_2} \mathrm{P}\{f_{ip}(x) = 0, i = 1, \ldots, s\}.$$

Clearly

$$\mathrm{P}\{f_{ip}(x) = 0, i = 1, \ldots, s\} = 1 - \mathrm{P}\{x \colon f_{ip}(x) \neq 0 \text{ for some } i\}$$

$$= 1 - \sum_{i=1}^{s} \frac{r(F_i, p)}{p} \geqslant \frac{1}{2} \quad \text{if} \quad p > k_0.$$

Suppose $k_0 < p \leqslant k$ and $p^{l_{ij}} \| F_{ij}(m)$ for some $l_{ij} \geqslant 1$ and for some (i, j).

In this case by the definition of k_0, we have clearly

$$\begin{aligned}\mathrm{P}\{x\colon f_{ip}(x)=f_{ip}(m),\ i=1,\dots,s\} &\geqslant \mathrm{P}\{x\colon p^{l_{ij}}\parallel F_{ij}(x)\}\\ &=\frac{r(F_{ij},p^{l_{ij}})}{p^{l_{ij}}}-\frac{r(F_{ij},p^{l_{ij}+1})}{p^{l_{ij}+1}}>0.\end{aligned}$$

Let $\Phi_i(m)=\prod_{\substack{p^l\parallel F_i(m)\\ p\leqslant k_0}} p^l$. Note that

$$\mathrm{P}\{x\colon f_{i0}(x)=f_{i0}(m),\ i=1,\dots,s\}$$
$$\geqslant \mathrm{D}\{\Phi_i(m)\mid F_i(n) \text{ and } \Phi_i(m)p\nmid F_i(n) \text{ for any } p\leqslant k_0 \text{ and for } i=1,\dots,s\}$$
$$>0 \quad \text{(since } n=m \text{ is a solution of the above relations).}$$

So $A\subset S$. Hence $B\subset A\subset S$. Clearly B is dense in S. This completes the proof of Proposition 2.

Absolute continuity of the distributions of $f(m)$ and $f(F(m))$.

Remark 2. Let f be the strongly additive arithmetic function defined by

$$f(p)=\begin{cases}0 & \text{if} \quad p\leqslant e^e,\\ \dfrac{1}{(\log\log p)^{3/2}} & \text{if} \quad p>e^e.\end{cases}$$

Let $F(m)$ be any polynomial taking positive integral values for $m\geqslant 1$. From Theorem 1 of [1] we can conclude that $f(F(m))$ has a distribution. Since

$$\sum_{p\leqslant n}\frac{r(F,p)}{p}=r\log\log n+O(1)$$

(see [9]) where r is the number of distinct irreducible factors of F.

Following an argument similar to the argument given in [2] it is not difficult to conclude that the distribution of $f(F(m))$ is absolutely continuous.

Remark 3. Let f be any real-valued additive arithmetic function having a distribution. Suppose that there exist sequences of real numbers g_N, l_N, s_N and a constant b such that $g_N l_N\to 0$, $l_N\to\infty$.

$$\frac{1}{g_N^2}\left\{\sum_{p>s_N}\frac{\{f'(p)\}^2}{p}+\left(\sum_{p>s_N}\frac{f'(p)}{p}\right)^2\right\}\to 0 \quad \text{as} \quad N\to\infty$$

and there exist positive integers $m_1,\dots,m_{l_N}$ composed of primes $p\leqslant s_N$ such that $\dfrac{1}{\log s_N}\sum_{i=1}^{l_N}\dfrac{1}{m_i}\geqslant b$ for all sufficiently large N. Then the distribution of f is singular.

This fact can be proved as follows. Without loss of generality we can assume that f is strongly additive and $|f(p)| < 1$. We write every positive integer $m = m'm''$, where m' is composed of primes $p \leqslant s_N$ and m'' of primes $p > s_N$. The density of integers $m = m'm''$ such that $m' = m_i$ for some $i = 1, \ldots, l_N$ is

$$\sum_{i=1}^{l_N} \frac{1}{m_i} \prod_{p \leqslant s_N} \left(1 - \frac{1}{p}\right) \sim \frac{e^{-\gamma}}{\log s_N} \sum_{i=1}^{l_N} \frac{1}{m_i} \geqslant e^{-\gamma} b, \tag{5}$$

where γ is Euler's constant.

For $x \in \mathfrak{S}$ and any prime p, put

$$f_p(x) = \begin{cases} f(p) & \text{if} \quad p|x, \\ 0 & \text{otherwise}. \end{cases}$$

Since f has a distribution, $\sum_p f_p(x)$ converges almost everywhere ([8]) and

$$\mathrm{D}\{n\colon f(n) < c\} = \mathrm{P}\Big\{x\colon \sum_p f_p(x) < c\Big\}.$$

Clearly

$$\mathrm{P}\Big\{x\colon \Big|\sum_{p > s_N} f_p(x)\Big| > g_N\Big\} \leqslant \frac{1}{g_N^2}\Big\{\sum_{p > s_N} \frac{f^2(p)}{p} + \Big(\sum_{p > s_N} \frac{f(p)}{p}\Big)^2\Big\} \to 0 \tag{6}$$

$$\text{as} \quad N \to \infty.$$

Consider the open intervals $\big(f(m_i) - g_N, f(m_i) + g_N\big)$, $i = 1, \ldots, l_N$. By (5) and (6)

$$\mathrm{P}\Big\{x\colon \sum_p f_p(x) \in \bigcup_{i=1}^{l_N} \big(f(m_i) - g_N, f(m_i) + g_N\big)\Big\}$$

$$\geqslant be^{-\gamma} - \frac{1}{g_N^2}\Big(\sum_{p > s_N} \frac{f(p)^2}{p} + \Big(\sum_{p > s_N} \frac{f(p)}{p}\Big)^2\Big) \geqslant \frac{be^{-\gamma}}{2}$$

for all sufficiently large N.

And the sum of the lengths of these l_N intervals is less than or equal to $2g_N l_N$. Hence it follows that the distribution of $f(m)$ cannot be absolutely continuous. Hence it is singular.

PROPOSITION 3. *Let $F \in \boldsymbol{P}$. Let f be a real-valued additive arithmetic function such that*

$$f(p^k) r(F, p^k) \to 0 \quad \textit{as} \quad p \to \infty \quad \textit{for} \quad k = 1, \ldots, \nu_F - 1,$$

if $\nu_F \geqslant 2$. (This condition can be dropped if F is a product of linear polynomials.)

Let Q be a set of primes such that

$$(7)\qquad \sum_{p\in Q}\frac{1}{p}<\infty \text{ and } q\notin Q \text{ implies either } r(F,q)\neq 0$$

$$\text{or } r(F,q)=0 \text{ and } f(q)=0.$$

If $f(m)$ and $f(F(m))$ have distributions, then the distribution of $f(F(m))$ is absolutely continuous if the distribution of $f(m)$ is absolutely continuous.

Proof. By Lemma 2 there exists a constant c such that $r(F,p^k)<c$ for all p and k and

$$r(F,p^k)=r(F,p)\quad \text{for all } k \text{ if } p>c.$$

Without loss of generality we can assume that f is strongly additive.

LEMMA 3. *If $\{X_n\}$ is a sequence of independent discrete random variables and $\{Y_n\}$ is another sequence of independent discrete random variables such that $\sum_n \mathrm{P}\{X_n\neq Y_n\}<\infty$, then $\sum_n X_n$ converges almost everywhere and its distribution function is absolutely continuous iff $\sum Y_n$ converges almost everywhere and its distribution is absolutely continuous.*

The proof of this lemma is well known [10].

LEMMA 4. *Suppose that $0\leqslant s(p)<c$ and $\{a_p\}$ is a sequence of real numbers. Then one can find a sequence of independent random variables $\{Y_p\colon p>2c\}$ defined on a complete probability space $(\Omega,\mathfrak{A},\mathrm{P})$ such that*

$$\mathrm{P}\{Y_p=0\}=1-\frac{s(p)}{p},$$

$$\mathrm{P}\{Y_p=na_p\}=\left(\frac{s(p)}{p}\right)^n\left(1-\frac{s(p)}{p}\right),\quad n=1,2,\ldots$$

and another sequence of independent random variables $\{X_p\colon p>2c\}$ defined on the same probability space $(\Omega,\mathfrak{A},\mathrm{P})$ such that

$$\mathrm{P}\{X_p=0\}=1-\frac{s(p)}{p},\quad \mathrm{P}\{X_p=a_p\}=\frac{s(p)}{p},$$

and

$$\sum_{p>2c}\mathrm{P}\{X_p\neq Y_p\}<\infty.$$

The proof of this lemma is easy and so is omitted.

LEMMA 5. *Suppose that h is the characteristic function of an infinitely divisible distribution with the Levy function M. If the total variation of M is finite and M is discrete, then the distribution corresponding to h is discrete.*

(See [6], p. 124.)

Now we prove Proposition 3. Let $\{X'_p \colon p > 2c\}$ be a sequence of independent random variables such that

$$\mathrm{P}\{X'_p = 0\} = 1 - \frac{1}{p}$$

and

$$\mathrm{P}\{X'_p = f(p)\} = \frac{1}{p}.$$

By Lemma 3 and from the results of [1], if f has an absolutely continuous distribution, it follows that $\sum_{p>2c} X'_p$ converges almost everywhere and its distribution function is absolutely continuous.

By Lemmas 3 and 4 one can find a sequence $\{X_p\}$ of independent random variables such that

$$\mathrm{P}\{X_p = 0\} = 1 - \frac{1}{p},$$

$$\mathrm{P}\{X_p = nf(p)\} = \frac{1}{p^n}\left(1 - \frac{1}{p}\right), \quad n = 1, 2, \ldots$$

$\sum_{p>2c} X_p$ converges almost everywhere and its distribution is absolutely continuous. If $h(t)$ is the characteristic function of $\sum_{p>2c} X_p$, then clearly

$$\log h(t) = i\gamma' t + \sum_{p>2c} \sum_{k=1}^{\infty} \left(e^{itkf(p)} - 1 - \frac{itkf(p)}{1 + k^2 f^2(p)}\right) \frac{1}{kp^k}$$

for some γ'. Since

$$\sum_{p \in Q} \frac{1}{p} + \sum_{p>2c} \sum_{k=2}^{\infty} \frac{1}{kp^k} < \infty,$$

by Lemma 5 we infer that the distribution function corresponding to the characteristic function

$$\varphi(t) = \exp\left\{\sum_{\substack{p \notin Q \\ p>2c}} \left(e^{itf(p)} - 1 - \frac{itf(p)}{1 + \left(f(p)\right)^2}\right) \frac{1}{p}\right\}$$

is absolutely continuous. From now on we write $r(p)$ for $r(p, F)$.

Now suppose that $\{Y'_p \colon p > 2c\}$ is a sequence of independent random variables such that

$$\mathrm{P}\{Y'_p = 0\} = 1 - \frac{r(p)}{p} \quad \text{and} \quad \mathrm{P}\{Y'_p = f(p)\} = \frac{r(p)}{p}.$$

Since $f(F(m))$ has a distribution, $\sum_{p>2c} Y'_p$ converges almost everywhere [1] and the distribution function of $f(F(m))$ is absolutely continuous if the distribution function of $\sum_{p>2c} Y'_p$ is absolutely continuous. Again, by Lemmas 3, 4 and 5 as above, we conclude that the distribution function of $\sum_{p>2c} Y'_p$ is absolutely continuous if the distribution function corresponding to the characteristic function $g(t)$ given by

$$g(t) = \exp\left\{\sum_{\substack{p>2c \\ p\nmid Q}} \left(e^{itf(p)} - 1 - \frac{itf(p)}{1+(f(p))^2}\right)\frac{r(p)}{p}\right\}$$

is absolutely continuous.

Since

$$\sum_{\substack{p>2c \\ p\nmid Q}} \left(e^{itf(p)} - 1 - \frac{itf(p)}{1+(f(p))^2}\right)\frac{1}{p}$$

and

$$\sum_{\substack{p>2c \\ p\nmid Q}} \left(e^{itf(p)} - 1 - \frac{itf(p)}{1+(f(p))^2}\right)\frac{r(p)}{p}$$

converge absolutely and uniformly in every compact interval of the real line,

$$\sum_{\substack{p>2c \\ p\nmid Q}} \left(e^{itf(p)} - 1 - \frac{itf(p)}{1+(f(p))^2}\right)\frac{(r(p)-1)}{p}$$

converges absolutely and uniformly in every compact interval of the real line. Since $r(p) \geqslant 1$ or $f(p) = 0$ if $p \nmid Q$, it follows that

$$l(t) = \exp\left\{\sum_{\substack{p>2c \\ p\nmid Q}} \left(e^{itf(p)} - 1 - \frac{itf(p)}{1+f^2(p)}\right)\frac{(r(p)-1)}{p}\right\}$$

is a characteristic function. We note that $g(t) = \varphi(t)\cdot l(t)$.

Since $\varphi(t)$ is a characteristic function of an absolutely continuous distribution, $g(t)$ is also a characteristic function of an absolutely continuous distribution. This completes the proof of Proposition 3.

(7) holds for many polynomials. In fact, if F has a linear factor, then condition (7) obviously holds. (7) is not a necessary condition, as is evident from Remark 2. But Proposition 3 is the best possible in the sense that if condition (7) is omitted then the conclusion of the proposition is not necessarily true.

Example. Let f be the strongly additive arithmetic function defined by

$$f(p) = \begin{cases} \dfrac{1}{(\log\log p)^{3/2}} & \text{if} \quad p > e^e \text{ and } p \equiv 3 \pmod 4, \\ 0 & \text{otherwise}. \end{cases}$$

Let $F(m)$ be the polynomial m^2+1.

The following lemma shows that $f(F(m)) = 0$ for all m and hence $f(F(m))$ has a degenerate distribution.

Lemma 6. *If p is a prime $\equiv 1 \pmod 4$, the congruence*

$$x^2 \equiv -1 \pmod p \tag{8}$$

has exactly two incongruent solutions. The congruence (8) *has no solution when p is a prime $\equiv 3 \pmod 4$.*

See [7], p. 99, Theorem 58.

Now we shall show that the distribution of $f(m)$ exists and is absolutely continuous.

We need the following

Lemma 7. *If $F \in P$ and the number of distinct factors of F is k, then*

$$\sum_{p \leqslant x} \frac{r(F,p)}{p} = k \log\log x + O(1).$$

See [9].

The characteristic function of the distribution function of $f(m)$ is given by

$$L(u) = \prod_p \left(1 - \frac{1-e^{iuf(p)}}{p}\right).$$

Now as in [2] for $u \neq 0$

$$|L(u)| \leqslant \prod{}' \left|1 - \frac{1-\exp\{iu(\log\log p)^{-3/2}\}}{p}\right| \tag{9}$$

where the product $\prod'$ for each fixed $u \neq 0$, is taken over those primes which satisfy the following conditions:

$$p > e^e, \quad p \equiv 3 \pmod 4 \quad \text{and} \quad 3\pi < 4|u(\log\log p)^{-3/2}| < 5\pi. \tag{10}$$

Now each factor of the product on the right of (9) is less than $1 - \dfrac{1}{p}$; so that

$$|L(u)| \leqslant \prod{}' \left(1 - \frac{1}{p}\right).$$

Hence

$$|L(u)| = O\left(\exp\left(-\sum{}' 1/p\right)\right),$$

where, for each fixed $u \neq 0$, $\sum'$ denotes the sum over those primes which satisfy (10). By Lemmas 6 and 7 we get

$$\sum_{\substack{p\equiv 3(\bmod 4)\\ p\leqslant x}} 2/p = \log\log x + O(1).$$

Hence

$$|L(u)| = O(\{\exp(-c|u|^{2/3})\}),$$

where

$$c = \frac{1}{2}\left(\frac{4}{\pi}\right)^{2/3}\left(\frac{1}{3^{2/3}} - \frac{1}{5^{2/3}}\right) > 0.$$

So $L(u)$ is integrable and hence $L(u)$ is the characteristic function of an absolutely continuous distribution function.

Acknowledgements. The author wishes to record his deep gratitude to Prof. J. K. Ghosh for the many useful discussions he had with him during the preparation of this paper. Thanks are also due to Prof. E. M. Paul for some useful conversations regarding this subject.

References

[1] G. J. Babu, *On the distribution of additive arithmetical functions of integral polynomials*, Sankhyā, Series A, 34 (1972), pp. 323–334.

[2] P. Erdös, *On the smoothness of the asymptotic distribution of additive arithmetical functions*, Amer. J. Math. 61 (1939), pp. 722–725.

[3] P. Erdös and A. Schinzel, *Distribution of the values of some arithmetical functions*, Acta Arith. 6 (1961), pp. 473–485.

[4] W. Feller, *An Introduction to Probability Theory and Its Applications*, vol. II, New York—London 1961.

[5] J. Kubilius, *Probabilistic methods in the theory of numbers*, Translations of Math. Monographs, vol. 11, Providence 1964.

[6] E. Lukacs, *Characteristic Functions*, 2nd ed., London 1970.

[7] T. Nagell, *Introduction to Number Theory*, New York 1951.

[8] E. Novoselov, *A new method in probabilistic number theory*, Amer. Math. Soc. Trans. 52 (1966), pp. 217–275.

[9] M. Tanaka, *On the number of prime factors of integers*, Japan J. Math. 25 (1955), pp. 1–20.

[10] H. G. Tucker, *On continuous singular infinitely divisible distribution functions*, Ann. Math. Statis. 35 (1964), pp. 330–335.

INDIAN STATISTICAL INSTITUTE

Received on 15. 1. 1972 (250)