ACTA ARITHMETICA
XXIX (1976)

# Some results on the distribution of values of additive functions on the set of pairs of positive integers, I

by

G. JOGESH BABU (Eugene, Oreg.)*

**0. Introduction.** In 1969 H. Delange [1] defined a density for sets of pairs $[m, n]$ of positive integers and determined it for some sets defined by arithmetical properties of $m$ and $n$. In this paper we find necessary and sufficient conditions for

$$\{f_1(F_1(m), G_1(n)), \ldots, f_s(F_s(m), G_s(n))\}$$

to have distribution, where $f_1, \ldots, f_s$ are additive functions and $F_i, G_i$ are polynomials with integer coefficients, $F_i(m) > 0$, $G_i(m) > 0$ for all $m \geqslant 1$, $F_i, G_i$ are not divisible by square of any irreducible polynomial and $F_i, G_i, f_i$ satisfy the Condition A given in the next section. We also give some sufficient conditions for $f(F(m), G(n))$ to have absolutely continuous distribution.

**1. Notations and definitions.** $\boldsymbol{P}$ denote the set of all polynomials $F$ with integer coefficients satisfying the following conditions:

P1. $F(m) > 0$ for $m = 1, 2, \ldots$

P2. $F$ is not divisible by square of any irreducible polynomial.

For $F \in \boldsymbol{P}$ let $D_F$ denote the degree of the polynomial $F$. For any positive integer $d$, let $r(F, d)$ denote the number of incongruent solutions in integers of the congruence relation $F(m) \equiv 0 \pmod{d}$.

In the sequel $Z_2$ denotes the set of all pairs of positive integers. $p, q, \ldots$ denote prime numbers. The letters $r, j$ will stand for non-negative integers, $k, d$ for integers and $m, n, s$ for positive integers.

DEFINITION. A real-valued function on $Z_2$ is said to be *additive* if

$$f(m_1 m_2, n_1 n_2) = f(m_1, n_1) + f(m_2, n_2)$$

* The author is currently visiting the University of Oregon, Eugene, Oregon, U.S.A.

whenever $(m_1n_1, m_2n_2) = 1$. Define, for any positive integer $k$,

$$f(m, n)_k = \sum_{p\leqslant k} f[p^{a(p,m)}, p^{a(p,n)}]$$

where

$$a(p, n) = \begin{cases} 0 & \text{if} \quad p \nmid n, \\ r & \text{if} \quad p^r \| n \quad (r \geqslant 1). \end{cases}$$

Let $E$ be a set of pairs $[m, n]$ of positive integers. Let $N(E)$ denote the cardinality of the set $E$. If

$$(1/xy) N\{[m, n] \in E\colon\ m \leqslant x \text{ and } n \leqslant y\}$$

tends to a limit $\delta(E)$ as $x$ and $y$ tend to infinity independently then we say that the set $E$ possesses density $\delta(E)$ (see [1]).

We define for any $x \geqslant 1$, $y \geqslant 1$ and $F, G \in \boldsymbol{P}$

$$A(x, y, f, F, G) = \sum_{p\leqslant x} \frac{1}{p} f(p, 1) r(F, p) + \sum_{p\leqslant y} \frac{1}{p} f(1, p) r(G, p),$$

$$\{B(x, y, f, F, G)\}^2 = \sum_{p\leqslant x} \frac{1}{p} f^2(p, 1) r(F, p) + \sum_{p\leqslant y} \frac{1}{p} f^2(1, p) r(G, p).$$

We say that the $s$-tuples $\{h_1(m, n), \ldots, h_s(m, n)\}$ of real functions, on the pairs of positive integers, have a *distribution* if there is an $s$-dimensional probability distribution function $Q(c_1, \ldots, c_s)$ such that the density of

$$\{[m, n]\colon\ h_1(m, n) < c_1, \ldots, h_s(m, n) < c_s\}$$

exists and equals $Q(c_1, \ldots, c_s)$, for all of its continuity points.

We shall often use the following condition and shall refer to it as Condition A.

Condition A. We say that $F \in \boldsymbol{P}$, $G \in \boldsymbol{P}$ and a real-valued additive function $f$ on $Z_2$ satisfy Condition A if the following hold:

$$f(p^k, 1) r(F, p^k) \to 0 \text{ as } p \to \infty \text{ for } k = 1, \ldots, t_F,$$

$$f(1, p^k) r(G, p^k) \to 0 \text{ as } p \to \infty \text{ for } k = 1, \ldots, t_G$$

and

$$f(p^k, p^j) r(F, p^k) r(G, p^j) \to 0 \text{ as } p \to \infty$$

for $k = 1, \ldots, t_F$ and for $j = 1, \ldots, t_G$, where $t_G = \max(1, D_G - 1)$ and $t_F = \max(1, D_F - 1)$.

Throughout this paper $f, f_1, \ldots, f_s$ denote real-valued additive functions on $Z_2$.



For any additive function $f$ on $Z_2$, let $f^*$ denote the additive function given by

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

and

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

## 2. Results.

Theorem 1. *Let $F_i \in \boldsymbol{P}$, $G_i \in \boldsymbol{P}$ for $i = 1, \ldots, s$. Suppose for each $i = 1, \ldots, s$, $F_i, G_i$ and a real-valued additive function $f_i$ on $Z_2$ satisfy Condition* A. *Then the $s$-tuples*

$$\{f_1(F_1(m), G_1(n)), \ldots, f_s(F_s(m), G_s(n))\}$$

*have a distribution if and only if the following series*

$$\sum_p \frac{1}{p} f_i^*(p, 1) r(F_i, p), \tag{2.1}$$

$$\sum_p \frac{1}{p} f_i^*(1, p) r(G_i, p) \tag{2.2}$$

*and*

$$\sum_p \frac{1}{p} \left([f_i^*(p, 1)]^2 r(F_i, p) + [f_i^*(1, p)]^2 r(G_i, p)\right) \tag{2.3}$$

*converge for $i = 1, \ldots, s$.*

Theorem 2. *$f$ has a distribution if and only if the three series*

$$\sum_p \frac{1}{p} f^*(p, 1), \tag{2.4}$$

$$\sum_p \frac{1}{p} f^*(1, p) \tag{2.5}$$

*and*

$$\sum_p \frac{1}{p} \left([f^*(p, 1)]^2 + [f^*(1, p)]^2\right) \tag{2.6}$$

*converge.*

*Moreover, if $f$ has a distribution then it is continuous if and only if either*

$$\sum_{f(1,p)\neq 0} \frac{1}{p} = \infty \quad \text{or} \quad \sum_{f(p,1)\neq 0} \frac{1}{p} = \infty.$$

This theorem was also obtained by Delange independently (personal communication).

An obvious modification of the proof of Proposition 3 in [6] gives the following

THEOREM 3. *Let $F \in \mathbf{P}$ and $G \in \mathbf{P}$. Suppose $f, F$ and $G$ satisfy Condition A. Let $Q$ be a set of primes such that $\sum_{q \in Q} 1/q < \infty$, and $q \notin Q$ implies either $r(F, q) \neq 0$, or $r(G, q) \neq 0$, or $r(F, q) = 0$ and $f(q, 1) = 0$, or $r(G, q) = 0$ and $f(1, q) = 0$. Suppose $f(m, n)$ and $f(F(m), G(n))$ have distributions. Then the distribution of $f(F(m), G(n))$ is absolutely continuous if the distribution of $f(m, n)$ is absolutely continuous.*

THEOREM 4. *If $\limsup_{x\to\infty} (1/x^2) N\{[m, n]: m \leqslant x,\ n \leqslant x,\ f(m, n) = a\} > 0$ for some real number $a$, then $f$ has a distribution.*

THEOREM 5. *Let*

$$A = \{p: \text{ either } f(p, 1) < 0 \text{ or } f(1, p) < 0\}.$$

*Suppose*

$$\sum_{p \in A} \frac{1}{p} < \infty$$

*and there exist positive constants $c$, $\delta$ and two sequences $\{x_i\}$ and $\{y_i\}$ such that*

$$N\{[m, n]: m \leqslant x_i,\ n \leqslant y_i, f(m, n) < c\} > \delta x_i y_i$$

*for all $i$ and $x_i \to \infty$, $y_i \to \infty$ as $i \to \infty$. Then $f$ has a distribution.*

Results similar to Corollaries 1 and 2 of [7] can be obtained in a similar way for additive functions on $Z_2$.

## 3. Preliminary results.

LEMMA 1 ([9]). *Let $F \in \mathbf{P}$. Then there exists a $p_0$ such that $p > p_0$ implies $r(F, p^k) = r(F, p)$ for any positive integer $k$. Also*

$$r(F, ab) = r(F, a) r(F, b) \quad \text{if} \quad (a, b) = 1$$

*and*

$$r(F, p^k) \leqslant c \quad \text{for some constant } c \text{ depending only on } F.$$

LEMMA 2 ([4]). *Let $F \in \mathbf{P}$ with $D_F \geqslant 2$. Then for each $\varepsilon > 0$, there exist $v_0 = v_0(\varepsilon)$ and $k = k(\varepsilon)$ such that*

$$N\{m \leqslant x:\ p^{D_F} | F(m) \text{ for some } p > v \text{ or } q^k | F(m) \text{ for some } q\} < \varepsilon x + o(x)$$

*as $x \to \infty$ for all $v > v_0$.*



LEMMA 3 ([3], p. 246). *Let $U$ and $V$ be two probability distributions neither of which is concentrated at one point. If for a sequence $\{F_n\}$ of probability distributions and constants $a_n > 0$, $c_n > 0$, $b_n$ and $d_n$,*

$$F_n(a_n x + b_n) \to U(x),$$

$$F_n(c_n x + d_n) \to V(x)$$

*at all points of continuity, then*

$$\frac{c_n}{a_n} \to A \neq 0, \qquad \frac{d_n - b_n}{a_n} \to B.$$

LEMMA 4. *Let $F \in \mathbf{P}$ and $G \in \mathbf{P}$. Let $f$ be any additive function on the pairs of positive integers. Suppose $f, F, G$ satisfy Condition A. Then given any $\varepsilon > 0$, there exist $x_0, y_0$ such that*

$$\sum_{m \leqslant x} \sum_{n \leqslant y} |\tilde{f}(F(m), G(n)) - A(x, y, \tilde{f}, F, G)|^2 \leqslant cxy B^2(x, y, \tilde{f}, F, G) + \varepsilon xy$$

*for all $x \geqslant x_0$ and $y \geqslant y_0$ where*

$$\tilde{f}(p^k, p^j) = \begin{cases} f(p^k, p^j) & \text{if} \quad 0 \leqslant k \leqslant t_F \text{ and } 0 \leqslant j \leqslant t_G, \\ 0 & \text{otherwise} \end{cases}$$

*and $c$ depends only on $F$ and $G$.*

Proof is similar to Turán–Kubilius inequality ([8], Lemma 3.1, p. 31).

LEMMA 5. *Let $F \in \mathbf{P}$ and $G \in \mathbf{P}$. Let $f$ be any real-valued additive function such that $f(p^k, p^j) = 0$ whenever $k + j > 1$. Suppose further we have*

$$B(x, y, f, F, G) \to \infty \quad \text{as} \quad x \to \infty,\ y \to \infty,$$

$$r(G, p) f(1, p) = o(B(1, p, f, F, G))$$

*and*

$$r(F, p) f(p, 1) = o(B(p, 1, f, F, G)) \quad \text{as} \quad p \to \infty.$$

*Then*

$$x^{-2} N\left\{[m, n]: m \leqslant x,\ n \leqslant x,\ \frac{f(F(m), G(n)) - A(x, x, f, F, G)}{B(x, x, f, F, G)} < c\right\} \to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{c} e^{-z^2/2} dz$$

*as $x \to \infty$, for all real numbers $c$.*

Proof is similar to that of Theorem 4.2 of [8].

## 4. Proofs of main results.

Proof of Theorem 1. First we prove this theorem when $s = 1$. For simplicity in writing we drop the subscripts. Let $p_0$ be such that

$$r(F, p^k) = r(F, p) \quad \text{and} \quad r(G, p^k) = r(G, p)$$

for all $k > 1$ and $p > p_0$. Define a sequence $\{X_p\colon p > p_0\}$ of independent random variables such that for each real number $a$ and $p > p_0$

$$P\{X_p = a\} = \sum p^{-k-j} r(F, p^k) r(G, p^j)\, \delta(F, k, p)\, \delta(G, j, p)$$

where the summation is taken over all $k, j \geqslant 0$ such that $f(p^k, p^j) = a$, and

$$\delta(F, k, p) = \begin{cases} 1 - r(F, p)p^{-1} & \text{if} \quad k = 0, \\ 1 - p^{-1} & \text{if} \quad k \geqslant 1, \end{cases}$$

and

$$P\{X_{p_0} = a\} = \text{density of } \{[m, n]\colon f(F(m), G(n))_{p_0} = a\}.$$

Note that $X_{p_0}$ is well defined as the density of the set on the right-hand side above exists.

It is not difficult to check that for any $r > p_0$ for each real number $a$, the density of $\{[m, n]\colon f(F(m), G(n))_r = a\}$ exists and equals $P\{\sum_{p\leqslant r} X_p = a\}$. If (2.1), (2.2) and (2.3) converge, then by Kolmogorov's 3-series theorem $\sum X_p$ converges almost everywhere. Hence by Condition A, Lemma 2 and Lemma 4, it follows that, for each continuity point $c$ of the distribution function $P\{\sum X_p < c\}$, the density of $\{[m, n]\colon f(F(m), G(n)) < c\}$ exists and equals $P\{\sum X_p < c\}$.

To prove the converse part we assume without loss of generality, the distribution of $f(F(m), G(n))$ is non-degenerate. [Otherwise, we can choose a $p_1 > p_0$, $k > 1$, such that $r(F, p_1^k) \neq 0$, and define new additive function $g$ such that

$$g(p_1^k, 1) = f(p_1^k, 1) + 1,$$
$$g(p^j, p^t) = f(p^j, p^t) \quad \text{if} \quad (p^j, p^t) \neq (p_1^k, 1).$$

Obviously $g$ has a non-degenerate distribution.] In view of Condition A and Lemmas 2, 3 and 5 we conclude that $\sup_{x,y} B(x, y, f^*, F, G) < \infty$.

By Kolmogorov's 3-series theorem

$$\sum_p \Big\{X_p - \frac{1}{p}\big(f^*(p, 1) r(F, p) + f^*(1, p) r(G, p)\big)\Big\} \tag{4.1}$$

converges almost everywhere. Let $Q$ denote the distribution of (4.1). It is easy to see in view of Condition A and Lemma 4, that at each continuity point $c$ of $Q$



$$(1/xy) N\{[m, n]\colon m \leqslant x, n \leqslant y,\ f(F(m), G(n)) - A(x, y, f^*, F, G) < c\} \tag{4.2}$$

tends to $Q(c)$ as $x$ and $y$ tend to infinity independently. It follows easily by (4.2), that the set $\{A(x, y, f^*, F, G)\}$ is bounded, since $f(F(m), G(n))$ has distribution, and

$$(1/xy) N\{[m, n]\colon m \leqslant x, n \leqslant y, f(F(m), G(n)) - A(x, y, f^*, F, G) < c\}$$

are discrete distributions. Hence there exist sequence $\{x_n\}$, $\{y_n\}$ such that $x_n \to \infty$, $y_n \to \infty$ as $n \to \infty$ and $\lim_{n\to\infty} A(x_n, y_n, f^*, F, G) = b$ for some $b$. So $P\{\sum X_p < a + b\} = Q(a)$ for all continuity points $a$ of $Q$ such that $a+b$ is a continuity point of $Q$. Consequently $b$ is the only limit point of $\{A(x, y, f^*, F, G)\}$. So the two series (2.1) and (2.2) are convergent. This completes the proof of Theorem 1, when $s = 1$.

Now we consider the case $s > 1$. Find a $p_0$ such that

$$r(F_i, p^k) = r(F_i, p) \quad \text{and} \quad r(G_i, p^k) = r(G_i, p)$$

for all $k \geqslant 1$, $i = 1, \ldots, s$ and $p > p_0$. For each $i = 1, \ldots, s$, define a sequence $\{X_{ip}\colon p \geqslant p_0\}$ of independent random variables with the same domain as follows. For $p > p_0$ and for any real number $a$

$$P\{X_{ip} = a\} = \sum_{\substack{k,j\geqslant 0 \\ f_i(p^k, p^j) = a}} \frac{r(F_i, p^k) r(G_i, p^j)\, \delta(F_i, k, p)\, \delta(G_i, j, p)}{p^{k+j}}$$

and

$$P\{X_{ip_0} = a\} = \text{density of } \{[m, n]\colon f_i(F_i(m), G_i(n))_{p_0} = a\}.$$

If (2.1), (2.2) and (2.3) hold then for each $i = 1, \ldots, s$, $\sum_p X_{ip}$ converges almost everywhere. So for each $s$-tuple $(c_1, \ldots, c_s)$ of real numbers $\sum_p \sum_{i=1}^{s} c_i X_{ip}$ converges almost everywhere. As in the above case, it can be shown that distribution of

$$c_1 f_1(F_1(m), G_1(n)) + \ldots + c_s f_s(F_s(m), G_s(n))$$

exists and is same as the distribution of $\sum_{i=1}^{s} c_i \sum_p X_{ip}$. Hence by Cramer–Wold device ([3], p. 495), the distribution of

$$\{f_1(F_1(m), G_1(n)), \ldots, f_s(F_s(m), G_s(n))\}$$

exists. The converse part follows from the above case. This completes the proof of Theorem 1.

Proof of Theorem 2. To prove that the convergence of the series (2.4), (2.5) and (2.6) is necessary, note that following [2] with necessary modifications, one can show the existence of a $c > 0$ such that

$$\sum_{|f(1,p)|\geqslant c} \frac{1}{p} + \sum_{|f(p,1)|\geqslant c} \frac{1}{p} < \infty.$$

The rest of the proof is similar to that of Theorem 1.

We omit the proof of Theorem 4 as it is similar to the proof of Theorem 1 of [5].

Proof of Theorem 5. Choose $M$ and $k \geqslant 2$ such that

$$\sum_{\substack{q>M\\ q\in A}} \frac{2}{q} + \sum_{p} \frac{k+1}{p^k} + \sum_{p>M} \frac{1}{p^2} < \frac{\delta}{4}.$$

Let

$$B = \{[m, n]: \text{ either } q|mn \text{ for some } q > M \text{ and } q\in A \text{ or } p^k|mn \text{ for some } p \text{ or } p^2|mn \text{ for some } p > M\}.$$

Clearly, we have for all $x$ and $y$

$$N\{[m, n]\in B:\ m \leqslant x,\ n \leqslant y\} < \tfrac{1}{2}\delta xy.$$

Hence, for all $i$,

$$N\{[m, n]\notin B:\ m \leqslant x_i,\ n \leqslant y_i,\ f(m, n) < c\} > (\delta/2) x_i y_i.$$

Let

$$L = \sum_{p\leqslant M,\, l\leqslant k} |f(p^j, p^l)|.$$

If we define an additive function $h$ by

$$h(p^j, p^i) = \begin{cases} f(p^j, p^i) & \text{if } i+j = 1 \text{ and } p\notin A, \\ 0 & \text{otherwise,}\end{cases}$$

then clearly $h(m, n) = h(m, 1) + h(1, n) \geqslant 0$ for all $m, n$ and

$$N\{[m, n]:\ m \leqslant x_i,\ n \leqslant y_i,\ h(m, n) < c+L\} > (\delta/2) x_i y_i$$

for all $i$. So we have

$$\limsup_{n\to\infty} \frac{1}{n} \operatorname{card}\{m \leqslant n:\ h(m, 1) < c+L\} > 0$$

and

$$\limsup_{n\to\infty} \frac{1}{n} \operatorname{card}\{m \leqslant n:\ h(1, m) < c+L\} > 0.$$



Since $\sum_{p\in A} 1/p < \infty$, it follows from Theorem 3 of [4] that the two series

$$\sum_p \frac{1}{p} f^*(p, 1), \qquad \sum_p \frac{1}{p} f^*(1, p)$$

converge. Now the result follows from Theorem 2.

**Acknowledgements.** The author wishes to thank Professor J. K. Ghosh for many useful discussions he had with him during the preparation of this paper.

## References

[1] H. Delange, *On some sets of pairs of positive integers*, J. Number Theory 1 (1969), pp. 261–279.

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

[3] W. Feller, *An Introduction to Probability Theory and its Applications*, vol. 2, New York 1966.

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

[5] — *Some results on the distribution of additive arithmetic functions I*, Indian Statistical Institute, Tech. Report No. Math-Stat /40/71.

[6] — *Some results on the distribution of additive arithmetic functions II*, Acta Arith. 23 (1973), pp. 315–328.

[7] — *Some results on the distribution of additive arithmetic functions III*, ibid. 25 (1973), pp. 39–49.

[8] J. Kubilius, *Probabilistic Methods in the Theory of Numbers*, Transl. Math. Mono., Amer. Math. Soc. 11 (1964).

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

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF OREGON
Eugene, Oregon, USA

*Received on 17. 6. 1974*
*and in revised form on 12. 9. 1974* (588)