PROBABILISTIC METHODS IN THE THEORY OF ARITHMETIC FUNCTIONS

By

GUTTI JOGESH BABU

Thesis submitted to the Indian Statistical Institute in partial fulfilment of the requirements for the award of the degree of

Doctor of Philosophy

Calcutta

1973

ACKNOWLEDGEMENTS

Restricted Collection.

I am grateful to Professor C. Radhakrishna Rao, F.R.S., of the Indian Statistical Institute, for providing me the facilities to carry out my work.

Professor Jayanta Kumar Ghosh introduced me to Probabilistic Number Theory and supervised this work. I am indebted to him for the various discussions I had with him and for his going through my work at every stage and offering valuable suggestions. But for his constant encouragement and guidence this work might not have been as fruitful as it turned out to be.

I am thankful to Professors Karalil Matthew Paul and Mahendra Ganapathy Nadkarni for the many useful discussions which I had with them.

I had the pleasure of corresponding with Professor Jonas Kubilius, Vilnius University, U.S.S.R. on a number of problems discussed here. I am grateful to him for his constant encouragement and various suggestions

I thank Professor Walter Philipp, University of Illinois, U.S.A., for the correspondence I had with him on the material of Chapter 2 of this thesis.

I thank Dr. Subrahmanya Natarajan for helping me in preparing the thesis in its final form and Mr. Ramanathapuram Vaidheswaran Ramamoorthy for helping me in various ways during the final stages.

Finally, it is a pleasure to thank Mr. Arun Das for his elegant typing of this thesis.

GUTTI JOGESH BABU

CONTENTS

INTRODUCTION

In Chapter 1, necessary and sufficient conditions, for the existence of the distribution of

$$(1) \qquad \{f_1(F_1(\cdot)), \ldots, f_s(F_s(\cdot))\},$$

are given under very general conditions, where $F_1, \ldots, F_s$ are polynomials taking positive integer values for positive integers and $f_1, \ldots, f_s$ are real-valued additive arithmetic functions. The result is an improvement of a result due to Katai (1969).

In Chapter 2, an alternative proof of W. Philipp's result on weak convergence to Brownian motion, of functionals of arithmetic functions under Lindeberg-Levy type conditions, is given.

Suppose that the density of $\{m : f(F(m)) \varepsilon I\}$ exists and is positive for some bounded interval I, where f is a real-valued additive arithmetic function, F is a polynomial taking positive integer values for positive integers. Then, does $f(F(m))$ have a distribution? This question is answered, in Chapter 3, affirmatively under some restrictions on F and I.

In Chapter 4, the spectrum of the distribution of $\{f(n) - f(n+1), \ldots, f(n+h-1) - f(n+h)\}$ is characterized, whenever the distribution exists, where f is a real-valued additive arithmetic function. In this chapter, it is also shown, under very general conditions, that for any $m \geq 1$,

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

belongs to the spectrum of the distribution of (1), whenever it exists.

In Chapter 5, an attempt is made to characterize real-valued additive arithmetic functions whose distributions are singular. It is shown that if for some $c > 0$,

$$(2) \qquad \sum_{\substack{p>N \\ p\notin A}} \frac{1}{p}[f(p)]^2 = O(N^{-c}) \quad \text{as} \quad N \to \infty,$$

where A is a set of primes such that $\sum_{p\in A} \frac{1}{p} < \infty$ and f is a real-valued additive arithmetic function, then the distribution of $f(m) - f(m+1)$ exists and is singular. Using this it is shown that, if (2) holds, then the distribution of f is singular, whenever it exists. From this, it follows that, every bounded real-valued additive arithmetic function has a singular distribution. Similar results are obtained for $f(F(.))$. Some results concerning the smoothness of the distribution of $f(F(.))$ in terms of the smoothness of the distribution of f are given in this chapter.

In Chapter 6, asymptotic formulae for

$$\frac{1}{n}\,\mathrm{card}\{1 \le m \le n : f(m) > x_n\},$$

for a wide class of real-valued additive arithmetic functions, when $x_n \longrightarrow \infty$ at a certain rate, are obtained. One of the results is that

$$\frac{1}{\log\log n}\log\left[\frac{1}{n}\,\mathrm{card}\{1\le m\le n : \omega(m) > e^x \log\log n\}\right] = (e^x - 1) - xe^x + o(1)$$

as $n \longrightarrow \infty$, uniformly in $x \in (0, 2-\delta)$, where $0 < \delta < 2$ and $\omega(m)$ denotes the number of distinct prime factors of m.

In Chapter 7, necessary and sufficient conditions for a real-valued multiplicative function to have a distribution are obtained. If the distribution H of a multiplicative function g is continuous at zero then, it is shown that, H is absolutely continuous if and only if the distribution of a suitably defined additive arithmetic function is absolutely continuous.

In Chapter 8, an attempt is made to characterize all the distributions which are distributions of additive and multiplicative functions. If f is a real-valued additive arithmetic function having a distribution H and if $f(2^k) = kf(2)$ for all $k \ge 1$, then it is shown that, there exists a discrete infinitely divisible distribution G such that the convolution $H * G$ is infinitely divisible with discrete Levy functions and without Normal factor. Similar results are obtained in terms of characteristic transforms of the distributions of multiplicative arithmetic functions.

In Chapter 9, many results of previous chapters are generalised to additive functions on the set of pairs of positive integers. Necessary and sufficient conditions are given for the existence of distribution (mod 1) of a real-valued additive function on the set of pairs of positive integers.

In Chapter 10, some results of Levin and Fainleib (1968) on integral limit theorems are generalised to additive functions on the set of pairs of positive integers.

CHAPTER 1

DISTRIBUTION OF VALUES OF ADDITIVE ARITHMETIC FUNCTIONS OF INTEGRAL POLYNOMIALS

1. Introduction : An additive arithmetic function f is a complex-valued function on the set of positive integers satisfying

$$f(m.n) = f(m) + f(n)$$

whenever m and n are mutually prime. An additive arithmetic function is called strongly additive if for every prime p

$$f(p^k) = f(p)$$

for all $k \geq 1$.

Katai (1969) proved the following result. Let $f_1, \ldots, f_s$ be real-valued additive arithmetic functions and let $F_1, \ldots, F_s$ be polynomials with integer coefficients satisfying the following conditions :

i) $F_i(m) > 0$ for $m = 1, 2, \ldots$; $i = 1, \ldots, s$.

ii) F_i is not divisible by the square of any irreducible polynomial ; $i = 1, \ldots, s$.

iii) If $i \neq j$, then F_i and F_j are prime to each other; $i, j = 1, \ldots, s$.

Further suppose that

$$f_i(p^k)\ r\ (p^k,\ F_i) \rightarrow 0 \quad \text{as} \quad p \rightarrow \infty \quad \text{for} \quad k = 1,\ldots,D_i-1$$

whenever $D_i \geq 2$; $i = 1,\ldots,s$ and for each $i = 1,\ldots,s$ the following two series are convergent :

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

where the sum is extended over all prime numbers in the increasing order of their magnitude, D_i denotes the degree of F_i, for each positive integer d, $r(d,\ F_i)$ denotes the number of incongruent solutions of the congruence relation

$$F_i(m) \equiv 0 (\text{mod } d),$$

and $f_i^!(p)$ is $f_i(p)$ or 1 according as $|f_i(p)| < 1$ or $|f_i(p)| \geq 1$. Then the s-tuple

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

has a distribution.

Katai proved this result using number-theoretic methods. Recently Galambos (1971a) gave probabilistic proof of the above theorem under an extra assumption that $f_1,\ldots,f_s$ are strongly additive. The result of Galambos (1971a) is true fo any subsequence of integers satisfying his condition K 3. Here we relax the condition (iii) above and show that the convergence of the above two series for each i is <u>necessary</u> and <u>sufficient</u> for the s-tuple

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

to have a distribution. We use probabilistic methods and the results of Novoselov (1966) to prove this result.

2. <u>Notations and definitions</u> : Following are some of the notations and definitions used in this thesis :

Let $\underline{P}$ denote the set of all polynomials F with integer coefficients satisfying the following conditions :

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

P 2. F is not divisible by the square of any irreducible polynomial.

For $F \in \underline{P}$, let D_F denote the degree of F and for each positive integer d let $r(F, d)$ denote the number of incongruent solutions of the congruence relation

$$F(m) \equiv 0 \pmod d .$$

We let $c, c_1, c_2, \ldots$ denote constants and p, q with or without suffixes denote prime numbers. Let $p_1, p_2, \ldots$ be the sequence of all prime numbers in the increasing order of their magnitude. Let j, k, r, s denote non-negative integers and m, n denote positive integers.

Let $f(n), f_1(n), \ldots, f_s(n)$ denote real-valued additive arithmetic functions. Let $N_n \{\ldots\}$ denote the number of positive integers less than or equal to n having the property indicated in $\{\ldots\}$. For any subset A of the natural numbers, let $\overline{D}(A)$ and $\underline{D}(A)$ denote the upper and lower natural density of A respectively; i.e.,

$$\overline{D}(A) = \limsup_{n \to \infty} \frac{1}{n} N_n \{m \in A\}$$

and
$$\underline{D}(A) = \liminf_{n \to \infty} \frac{1}{n} N_n \{m \in A\}$$

We denote by $D(A)$, the common value of $\overline{D}(A)$ and $\underline{D}(A)$ whenever they coincide.

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

We define,

$$A(v,n,f,F) = \sum_{v \le p \le n} \frac{f'(p)\, r(F,p)}{p}$$

$$B(v,n,f,F) = \sum_{v \le p \le n} \left[\frac{\{f'(p)\}^2\, r(F,p)}{p} \right]^{1/2}$$

$$A(n,f,F) = A(1,n,f,F)$$

$$B(n,f,F) = B(1,n,f,F)$$

We say that the s-tuple $\{h_1(n),\ldots,h_s(n)\}$ of real arithmetic functions has a distribution, if

$$\frac{1}{n} N_n \{h_1(m) < c_1,\ldots,h_s(m) < c_s\}$$

tends to an s-dimensional probability distribution function $Q(c_1,\ldots,c_s)$, as $n \to \infty$, at all its continuity points.

FOR ANY OTHER UNEXPLAINED TERMINOLOGY USED IN THE SEQUEL, REFER TO KUBILIUS (1964) OR NOVOSELOV (1966).

3. Main results proved in this Chapter :

Theorem 1.1 : Let $f_1(n),\ldots,f_s(n)$ be real-valued additive arithmetic functions and $F_1,\ldots,F_s$ belong to $\underline{P}$. Suppose

(1.3.1) $\quad f_i(p^k) \quad r(F_i, p^k) \to 0$ as $p \to \infty$ whenever $D_{F_i} \geq 2$ for $k = 1, \ldots, D_{F_i} - 1$. Then the s-tuple $\{f_1(F_1(m)), \ldots, f_s(F_s(m))\}$ has a distribution if and only if

$$(1.3.2) \qquad \sum_p \frac{f_i'(p) \; r(F_i, p)}{p} \quad \text{converges for} \quad i = 1, \ldots, s$$

and

$$(1.3.3) \qquad \sum_p \frac{\{f_i'(p)\}^2 \; r(F_i, p)}{p} \quad \text{converges for} \quad i = 1, \ldots, s.$$

Remark : In Theorem 1.1, if F_i is a product of linear polynomials we can omit the condition (1.3.1).

4. Outline of Novoselov's method : Here we give a brief outline of Novoselov's (1966) method, because our proofs depend heavily on the probability space constructed by him.

Let Z denote the set of all integers. If

$$d(k, \lambda) = \sum_{m=2}^{\infty} \frac{\phi_m(k - \lambda)}{2^{m-1}}$$

where

$$\phi_m(n) = \begin{cases} 0 & \text{if } n \equiv 0 (\text{mod } m), \\ 1 & \text{if } n \not\equiv 0 (\text{mod } m). \end{cases}$$

then d defines a metric on Z. The topology induced by d on Z is same as the topology obtained by taking as a neighbourhood basis at the point a the set of all residue classes with respect to non-zero moduli that contain a. In this manner Z becomes a topological ring S with the usual addition and multiplication, and with a non-discrete topology. Note that S is totally disconnected and totally-bounded, [see Kelley (1955)]. Completing S we get a compact ring $\mathfrak{S}$, whose elements will be called polyadic numbers. On $\mathfrak{S}$, as a compact additive group, there exists a normalized Haar measure μ , [see Halmos (1962)]. This measure is not complete. Its completion which is denoted by P is clearly a probability measure.

If $\underline{d}$ is defined by

$$\underline{d}(x, y) = \sum_{m=2}^{\infty} \frac{\phi_m(x-y)}{2^{m-1}} \qquad \text{for } x, y \in \mathfrak{S}$$

where $\phi_m(x) = \begin{cases} 0 & \text{if there is a } y \in \mathfrak{S} \text{ such that } y.m = x \\ 1 & \text{otherwise} \end{cases}$

then $\underline{d}$ defines a metric on $\mathfrak{S}$ and the topology on $\mathfrak{S}$ is the same as the topology induced by $\underline{d}$. It is easy to see that if a sequence $\{x_n\}$ in $\mathfrak{S}$ tends to zero in the topology of $\mathfrak{S}$

then given any positive integer m, there exists an n_o such that $m|x_n$ for all $n > n_o$. (Here $m|x$ means there is a $y \varepsilon \mathfrak{S}$ such that $m.y = x$). If $x,y \varepsilon \mathfrak{S}$, then by $x \equiv y \pmod m$ we mean $x-y$ is divisible by m. If $x \varepsilon \mathfrak{S}$ and m is a positive integer then we can find an integer k such that $0 \leq k < m$ and $m|x-k$. For, there is a sequence $\{x_n\}$ of integers such that x_n tends to x in $\mathfrak{S}$, that is $x_n - x$ tends to zero in $\mathfrak{S}$. So there is an n_o such that if $n \geq n_o$ then $m|x-x_n$. Now, if k is an integer such that $0 \leq k < m$ and $x_{n_o} \equiv k \pmod m$ then clearly $x \equiv k \pmod m$.

Define $N_k = (k - s(n) + n + 2)\, n!$ if $s(n) \leq k < s(n+1)$, where $s(n) = 1 + 2^2 + \ldots + n^2$.

Note that N_k tends to zero in the topology of $\mathfrak{S}$ and $\frac{N_{k+1}}{N_k}$ tends to 1 as k tends to infinity. This sequence $\{N_k\}$ is fixed throughout. Let $R_k(x)$ be the smallest non-negative residue of x modulo N_k.

Let $\mathfrak{Z}_c$ be the class of all complex-valued functions f on $\mathfrak{S}$ such that $f(R_k(x)) \xrightarrow{P} f(x)$ as $k \to \infty$, where $\xrightarrow{P}$ denote the convergence in P-measure. We say that an arithmatic function $f \varepsilon \mathfrak{Z}_c$ if there is an extension $f(x)$ of $f(n)$ to $\mathfrak{S}$ such that $f(x) \varepsilon \mathfrak{Z}_c$.

$p^k \| x$ means the highest power of p that divides x is equal to k if k is a positive integer, and $p^\infty \| x$ means $p^k | x$ for every $k > 0$.

Some results of Novoselov :

Lemma 1.1 : $\overline{D}\{h(m) \varepsilon A\} = \limsup_{k \to \infty} P\{x : h(R_k(x)) \varepsilon A\}$

$$\underline{D}\{h(m) \varepsilon A\} = \liminf_{k \to \infty} P\{x : h(R_k(x)) \varepsilon A\}$$

for any set A and any complex-valued function h on the set of positive integers.

Lemma 1.2 : If $h_n(x) \varepsilon \mathcal{S}_o$ then the validity of any two of the following conditions

$$h_n(x) \xrightarrow{P} h(x)$$

$$\lim_{n \to \infty} \overline{D}\{|h(m) - h_n(m)| \geq \sigma\} = 0, \quad \text{for all } \sigma > 0,$$

$$h(x) \varepsilon \mathcal{S}_o$$

implies the third.

Lemma 1.3 : If $h_1(x) \varepsilon \mathcal{S}_o$ and $h_2(x) \varepsilon \mathcal{S}_o$, then

1. $a\, h_1(x) + b\, h_2(x) \varepsilon \mathcal{S}_o$ for any complex numbers a and b.
2. $h_1(x) \cdot h_2(x) \varepsilon \mathcal{S}_o$.

3. If $h(m) \in \mathcal{C}$ then $h(m)$ has a distribution.

Proofs of all these lemmas are easy. See Novoselov (1966).

5. Some preliminary results :

Lemma 1.4 : Let $F \in \underline{P}$. Then there exists a p_o such that $p > p_o \Rightarrow r(F, p^k) = r(F, p)$ for any positive integer k. Also

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

$$r(F, p^k) \leq c$$

for some constant c depending only on F, and

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

where r is the number of irreducible factors of F.

For proof see Tanaka (1955).

Lemma 1.5 : Let $F \in \underline{P}$ with $D_F \geq 2$. Then for each $\varepsilon > 0$, there exist $v_o = v_o(\varepsilon)$ and $k = k(\varepsilon)$ such that $v > v_o$ implies

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

as $n \to \infty$.

Proof : Choose k and v_o such that

$$\sum_{v_o \le p} \frac{r(F, p^{D_F})}{p^{D_F}} < \frac{\varepsilon}{2} \quad \text{and} \quad \sum_{p} \frac{r(F, p^k)}{p^k} < \frac{\varepsilon}{2}$$

Let $M > 0$ be such that $M.m^{D_F} \ge F(m)$ for every $m \ge 2$.

If $v > v_o$, then

$$N_n \{p^{D_F} | F(m) \text{ for some } p > v \text{ or } q^k | F(m) \text{ for some } q\}$$

$$\le n \sum_{p \ge v} \frac{r(F, p^{D_F})}{p^{D_F}} + \sum_{Mn > p \ge v} r(F, p^{D_F})$$

$$+ n \sum_{p} \frac{r(F, p^k)}{p^k} + \sum_{p < Mn} r(F, p^k)$$

$$< n\varepsilon + O\left(\frac{n}{\log n}\right) = n\,\varepsilon + o(n).$$

Lemma 1.6 : 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

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

$$F_n(c_n x + d_n) \to V(x) \text{ at all points of continuity,}$$

then

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

For proof see Feller (1966, Chapter VIII, Section 2, Lemma 1).

Lemma 1.7 : Let $F \in \underline{P}$. Let f be any additive arithmetic function such that

$$B(n,f,F) \rightarrow \infty \quad \text{as} \quad n \rightarrow \infty$$

$$f(p)\ r(F,p) = o(B(p,f,F))$$

and

$$f(p^k)\ r(F,p^k) \rightarrow 0 \quad \text{as} \quad p \rightarrow \infty \quad \text{for} \quad k=1,\ldots,D_F-1, \text{ if } D_F \geq 2.$$

Then $\frac{1}{n} N_n \{f(F(m)) < A(n,f,F) + xB(n,f,F)\} \rightarrow \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-y^2/2}\, dy$

as $n \rightarrow \infty$, for all real numbers x.

For proof see Halberstam (1956).

Let f be any additive arithmetic function and $F \in \underline{P}$. Suppose $F(m) = a_t m^t + \ldots + a_0$. Define $F(x) = a_t x^t + \ldots + a_0, x \in \mathfrak{S}$. Clearly $F(x)$ is uniformly continuous on $\mathfrak{S}$.

Define $\quad f_p(x, F) = \sum_{k=1}^{\infty} f(p^k)\ w(F, p^k, x)$

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

It is easy to see that $w(F, p^k, .)$ is measurable and

$$P(x : w(F,p^k,x) = 1) = \frac{r(F,p^k)}{p^k} - \frac{r(F,p^{k+1})}{p^{k+1}},$$ [see Novoselov (1966)]

<u>Lemma 1.8</u> : Let f be any additive arithmetic function. Let $F \in \underline{P}$. Suppose

(1.5.1) $D_F \geq 2$ and $f(p^k)\, r(F,p^k) \to 0$ as $p \to \infty$,

$k = 1,\ldots,D_F-1.$

Then given any $\varepsilon > 0$, there exists $v_o = v_o(\varepsilon)$ such that $v \geq v_o$ implies

$$\sum_{m=1}^{n} \Big(\sum_{p>v} \tilde{f}_p(m,F) - A(v,n,\tilde{f},F)\Big)^2 \leq CnB^2(v,n,\tilde{f},F) + \varepsilon n$$

where $\tilde{f}(p^k) = \begin{cases} f(p^k) & \text{if} \quad k \leq D_F - 1, \\ 0 & \text{otherwise,} \end{cases}$

and C depends only on F.

Remark : If F is a product of linear polynomials we may omit the condition (1.5.1).

Proof of this lemma is similar to Turan-Kubilius inequality, see Kubilius (1964, Lemma 3.1, p.31).

6. Proofs of the theorems :

Proof of Theorem 1.1 : We first consider the case $s = 1$. To simplify the notation we write f for f_1 and F for F_1. It is easy to show $f_p(x, F)$ is continuous almost everywhere, see Novoselov (1966). Hence for any n,

$$\sum_{p \leq n} f_p(x, F) \in \mathcal{S}_c .$$

Let p_o be such that $p > p_o$ implies $r(F, p^k) = r(F, p)$ for each $k \geq 1$. Observe that

$$E[w(F, p^k, x)] = \frac{r(F, p)}{p^k} \left(1 - \frac{1}{p}\right) \text{ if } p > p_o$$

$$E[w(F, p^k, x)\, w(F, p^t, x)] = 0 \text{ if } k \neq t.$$

Since $r(F, d)$ is a multiplicative function by Lemma 1.4, $\{f_p(x, F)\}_{p > p_o}$ are all mutually independent random variables, see Novoselov (1966, p.244).

Now suppose that (1.3.2) and (1.3.3) hold. By Kolmogorov's 3-series theorem [see Halmos (1962) section 46, Theorem E] it follows that $\sum_{p>p_o} f_p(x, F)$ converges almost everywhere.

Hence, $\sum_p f_p(x, F)$ converges almost everywhere.

Define

$$f^*(x,F) = \begin{cases} \sum_p f_p(x, F) & \text{whenever it converges,} \\ 0 & \text{otherwise.} \end{cases}$$

Note that $f^*(m,F) = f(F(m))$ for every natural number m.

To show that $f^*(x,F) \in \mathfrak{F}_o$ it is enough to show, in view of Lemma 1.2, that

$$\lim_{v\to\infty} \overline{D} \{|\sum_{p\geq v} f_p(m,F)| > \sigma\} = 0, \text{ for every } \sigma > 0.$$

This however follows from (1.3.1) and Lemmas 1.5 and 1.8. Hence $f(F(m))$ has a distribution.

Conversely, let U be the distribution of $f(F(m))$. If U is degenerate, then choose p_o^k, $k > 1$ such that $r(F, p_o^k) \neq 0$.

Put $f^*(p_o^k) = f(p_o^k) + 1$

$f^*(p^\beta) = f(p^\beta)$ if $p^\beta \neq p_o^k$

Now it is easy to see that, if f^* is the new additive arithmetic function defined above, the distribution of $f^*(F(m))$ exists and is nondegenerate. So without loss of generality we may assume that U is a nondegenerate probability distribution. From Lemmas 1.6 and 1.7 it follows that

$$\overline{\lim_{n \to \infty}} B(n, f, F) < \infty .$$

By Kolmogorov's 3-series theorem, we have

$$\sum_p \left\{ f_p(x, F) - \frac{f'(p)\ r(F, p)}{p} \right\} \text{ converges a. e.}$$

Define

$$g(x) = \begin{cases} \sum_p \left\{ f_p(x,F) - \frac{f'(p)\ r(F, p)}{p} \right\} & \text{if it converges,} \\ 0 & \text{otherwise.} \end{cases}$$

Let $Q(c) = P\{x : g(x) < c\}$.

By Lemma 1.8 and (1.3.1) it is easy to see that

$$\frac{1}{n} N_n \{[f(F(m)) - A(n, f, F)] < c\} \to Q(c) \quad \text{as} \quad n \to \infty$$

at all continuity points c of Q.

If Q is degenerate, it follows that $A(n,f,F)$ are bounded, since $\frac{1}{n} N_n \{[f(F(m)) - A(n,f,F)] < c\}$ are discrete distributions. Hence there exists a subsequence $\{n_r\}$ of natural numbers such that $A(n_r, f, F) \to b$ as $r \to \infty$ for some b. Thus we conclude that $U(c + b) = Q(c)$ which gives a contradiction, since we assumed that U is non-degenerate. Hence Q is non-degenerate. By Lemma 1.6, it follows that

$$\sum_p \frac{f'(p)\ r(F,\ p)}{p} \quad \text{converges.}$$

This proves Theorem 1.1, when $s = 1$.

Now, it follows, from (1.3.1), (1.3.2), (1.3.3) and what we have proved above that

$$f_1(F_1(m)) \ \varepsilon \ \mathcal{S}_o \ , \ \ldots \ , \ f_s(F_s(m)) \ \varepsilon \ \mathcal{S}_o \ .$$

By Lemma 1.3, for every s-tuple $(t_1,\ldots,t_s)$ of real numbers,

$$t_1\ f_1(F_1(m)) + \ldots + t_s\ f_s(F_s(m)) \ \varepsilon \ \mathcal{S}_o \ .$$

Hence by Cramer-Wold device [see Billingsley (1968, p 48)], the distribution of

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

exists.

The converse part follows easily because, if the s-tuple has a distribution then for each i, $f_i(F_i(m))$ has a distribution. This completes the proof of Theorem 1.1.

A function g on the set of positive integers into the complex numbers is said to be multiplicative if $g(m.n)=g(m)\,g(n)$ whenever $(m,n) = 1$, and $g(1) = 1$.

We prove

<u>Theorem 1.2</u> : Let g be a multiplicative function and $F \in \underline{P}$. Suppose that $D_F \geq 2$, g is real-valued and

$$(g(p^k) - 1)\, r(F, p^k) \to 0 \quad \text{as} \quad p \to \infty, \text{ for } k=1,\ldots,D_F-1.$$

Then $g(F(m)) \in \mathcal{G}$ if the following three series

$$\sum_{\substack{p \\ |g(p)-1|<1}} \frac{[g(p) - 1]\, r(F, p)}{p}$$

$$\sum_{\substack{p \\ |g(p)-1| < 1}} \frac{|g(p) - 1|^2 \, r(F, p)}{p}$$

and

$$\sum_{\substack{p \\ |g(p)-1| \geq 1 \\ g(p)\neq 0}} \frac{1}{p}$$

are convergent. For a positive $g(F(m))\ \varepsilon\ \mathcal{E}_o$ these three conditions are also necessary, provided the distribution function Q of the function $g(F(m))$ is continuous at zero.

We omit the proof of this theorem, since the proof is almost same as that of Proposition 51 of Novoselov (1966, p 251). We mention only that we have to use Theorem 1.1 instead of Erdos-Wintner theorem in the proof.

CHAPTER 2

AN INVARIANCE PRINCIPLE FOR ARITHMETIC FUNCTIONS

1. Introduction : For any new unexplained terminology used in the sequel, refer to Parthasarathy (1967) or Billingsley (1968).

Let $\{f_N\}$ be a sequence of real-valued arithmetic functions. For any $n \geq 1$ we write

$$B^2(N, n) = \sum_{p \leq n} f_N^2(p)/p .$$

Assume

$$B(N) = B(N, N) \to \infty \quad \text{as} \quad N \to \infty .$$

We define random functions $h_N(. , m) \in D[0, 1]$ by

$$h_N(x, m) = [1/B(N)] \sum f_N(p) \left(\delta_p(m) - \frac{1}{p}\right)$$

where the sum is extended over all primes $p \leq N$ satisfying

$$B(N, p) \leq x B^2(N)$$

and $\delta_p(m) = 1$ or 0 according as $p | m$ or $p \nmid m$. We give an alternative proof of Theorem 2.1 below due to W. Philipp (1972). Our proof is probabilistic in nature and does not involve many number-theoretic calculations.

Definition : A sequence $\{X_n\}$ of random elements, taking values in a metric space M, is said to converge weakly to a random element X, taking values in M, if

$$P\{X_n \in A\} \to P\{X \in A\}$$

as $n \to \infty$, whenever A is a Borel set in M such that

$$P\{X \in \partial A\} = 0$$

where ∂A denotes the boundary of A. We denote the weak convergence of $\{X_n\}$ to X by

$$X_n \xrightarrow{D} X.$$

Theorem 2.1 : Let $\{f_N\}$ be a sequence of real-valued arithmetic functions with $B(N) \to \infty$. Suppose that for any $\varepsilon > 0$

$$(2.1.1) \qquad \frac{1}{B^2(N)} \Big[\sum_{|f_N(p)| > \varepsilon B(N)} \frac{f_N^2(p)}{p} \Big] \to 0 \quad \text{as} \quad N \to \infty .$$

Then h_N (. , m) tends weakly to the standard Brownian motion.

Remark : Condition (2.1.1) is the natural analogue of the Lindeberg condition.

As an easy consequence of Theorem 2.1 we have

Corollary 2.1 : Under the hypotheses of Theorem 2.1. we have for $x > 0$.

$$\frac{1}{N} N_N \left\{ \max_{k \leq N} [f_N^{(k)}(m) - A(N,k)] < x B(N) \right\} \to \sqrt{2/\pi} \int_0^x e^{-u^2/2} du$$

and

$$\frac{1}{N} N_N \left\{ \max_{k \leq N} |f_N^{(k)}(m) - A(N,k)| < x B(N) \right\}$$

$$\to \sqrt{1/2\pi} \int_{-x}^{x} \sum_{k=-\infty}^{\infty} (-1)^k \exp\left[- \frac{(u-2kx)^2}{2}\right] du$$

as $N \to \infty$. Here

$$A(N,k) = \sum_{p \leq k} \frac{f_N(p)}{p}$$

and

$$f_N^{(k)}(m) = \sum_{\substack{p \leq k \\ p \mid m}} f_N(p).$$

Remark : This Corollary was proved first when $f_N = f$ for all $N \geq 1$, in Jogesh Babu (1972a), directly without any reference to Theorem 2.1. Also see Philipp (1972).

<u>Theorem 2.2</u> : Let $f(m)$ be a real-valued arithmetic function and F be any integral polynomial such that $F(m) > 0$ for $m = 1, 2, \ldots\ldots$. Let

$$A(n, F) = \sum_{p \leq n} \frac{f(p)\ r(F,p)}{p} \quad \text{and} \quad B^2(n,F) = \sum_{p \leq n} \frac{[f(p)]^2 r(F,p)}{p} .$$

Suppose $B(n,F) \to \infty$ as $n \to \infty$ and $f(p) = o(B(p, F))$ as $p \to \infty$, then for $x > 0$,

$$\lim_{n \to \infty} \frac{1}{n} N_n \left\{ \max_{k \leq n} \frac{f^{(k)}(F(m)) - A(k,F)}{B(n, F)} < x \right\}$$

$$= \frac{2}{\sqrt{2\pi}} \int_0^x e^{-t^2/2}\, dt$$

and

$$\lim_{n \to \infty} \frac{1}{n} N_n \left\{ \max_{k \leq n} \left| \frac{f^{(k)}(F(m)) - A(k,F)}{B(n, F)} \right| \leq x \right\}$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-x}^{x} \sum_{k=-\infty}^{\infty} (-1)^k \exp\left[- \frac{(u - 2k\ x)^2}{2}\right] du.$$

The proof of this result is essentially the same as that of Corollary 2.1 and so is omitted.

§ 2. We need the following Lemma.

Lemma 2.1 : Let f be an arithmetic function such that $D(n) \to \infty$ as $n \to \infty$ and for each $\varepsilon > 0$

$$\frac{1}{D^2(n)} \sum_{|f(p)|>\varepsilon D(n)} \frac{f^2(p)}{p} \to 0 \quad \text{as } n \to \infty,$$

where

$$D(n) = \left[\sum_{p \leq n} \frac{f^2(p)}{p} \right]^{1/2}$$

Then there exists a function $r(n)$ such that

$$\frac{\log r(n)}{\log n} \to 0, \qquad \frac{D(r(n))}{D(n)} \to 1$$

and for each $\varepsilon > 0$

$$\frac{1}{n} N_n \left\{ \max_{r(n) \leq k \leq n} \left| \sum_{r(n) < p \leq k} \left[f(p,m) - \frac{f(p)}{p} \right] \right| > \varepsilon D(n) \right\} \to 0$$

as $n \to \infty$, where

$$f(p, m) = \begin{cases} f(p) & \text{if } p \mid m, \\ 0 & \text{otherwise}. \end{cases}$$

Proof : Let

$$\delta_n = \frac{1}{n} N_n \{ \max_{r(n) \le k \le n} | \sum_{r(n) < p \le k} [f(p,m) - \frac{f(p)}{p}] | > \varepsilon D(n) \}$$

Now, by Chebyshev's inequality

$$\delta_n \le [1/(n \varepsilon^2 D^2(n))] \sum_{m=1}^{n} [\sum_{r(n) < p \le n} | f(p,m) - \frac{f(p)}{p} |]^2$$

$$\le [2/(n \varepsilon^2 D^2(n))] [\sum_{m=1}^{n} (\sum_{r(n)<p\le n} | f(p,m) |)^2 + 2n (\sum_{r(n)<p\le n} \frac{|f(p)|}{p})^2].$$

But

$$\sum_{m=1}^{n} [\sum_{r(n)<p\le n} | f(p,m) |]^2 = \sum_{m=1}^{n} [\sum_{r(n)<p\le n} | f(p,m) |^2 + \sum_{\substack{p \ne q \\ r(n)<p,q\le n}} | f(p,m) f(q,m) |]$$

$$\le n \sum_{r(n)<p\le n} \frac{f^2(p)}{p} + n [\sum_{r(n)<p\le n} \frac{|f(p)|}{p}]^2$$

Now, by hypothesis, there exists $\varepsilon(n)$ such that $\varepsilon(n) > 0$ for all n, $\varepsilon(n) \to 0$ as $n \to \infty$ and

$$\frac{1}{D^2(n)} \sum_{|f(p)| > \varepsilon(n) D(n)} \frac{f^2(p)}{p} < \varepsilon(n).$$

Put $r(n) = n^{\varepsilon(n)}$. Then

$$\Big[\sum_{r(n)<p\le n} \frac{|f(p)|}{p}\Big]^2$$

$$\le 2\Big[\sum_{\substack{r(n)<p\le n \\ |f(p)|\le\varepsilon(n)D(n)}} \frac{|f(p)|}{p}\Big]^2 + 2\Big[\sum_{\substack{r(n)<p\le n \\ |f(p)|>\varepsilon(n)D(n)}} \frac{|f(p)|}{p}\Big]^2$$

$$\le 2\, D^2(n)\, \varepsilon^2(n)\, (-\log \varepsilon(n) + o(1))^2$$

$$+ 2\Big(\sum_{r(n)<p\le n} \frac{1}{p}\Big)\Big[\sum_{\substack{r(n)<p\le n \\ |f(p)|>\varepsilon(n)D(n)}} \frac{f^2(p)}{p}\Big]$$

$$\le 2\, D^2(n)\, [\varepsilon^2(n)(-\log \varepsilon(n) + o(1))^2 - 2\,\varepsilon(n) \log \varepsilon(n)]$$

$$= o(D^2(n))$$

as $n \to \infty$, since $\varepsilon(n) \to 0$. Hence $\delta_n \to 0$ as $n \to \infty$.

Similar argument gives us that

$$[D(r(n)) \,/\, D(n)] \to 1$$

as $n \to \infty$. This completes the proof of Lemma 2.1.

§ 3. Now we need some results of Chapter II of Kubilius (1964), a brief outline of which is given below in the required form.

Let $r(n)$ be the function appearing in the proof of Lemma 2.1. Let

$$E_n = \{1, 2, \ldots, n\}.$$

For $p \leq r(n)$ and $t = 0, 1$, define

$$E(p, t, n) = \{1 \leq m \leq n : \delta_p(m) = t\}$$

Let $\underline{\underline{F}}_n$ be the smallest algebra of sets containing all $E(p, t, n)$. The algebra of sets and the function

$$\frac{1}{n} N_n \{m \in A\}$$

form a finite probability space, and for each $p \leq r(N)$, the function $f_N(p)\, \delta_p(m)$ is a random variable on this space. For each square-free integer $1 \leq k \leq n$, let

$$E_{n,k} = \bigcap_{p \leq r(n)} E(p, \delta_p(k), n).$$

Clearly, for different square-free integers k the sets $E_{n,k}$ have no element in common. Also every $A \in \underline{\underline{F}}_n$ is of the form

$$\bigcup_{k \varepsilon K} E_{n,k}$$

where K is a subset of E_n. Let

$$P(E(p,t,n)) = \begin{cases} 1 - \frac{1}{p}, & \text{if } t = 0 \text{ and } p \leq r(n), \\ \frac{1}{p}, & \text{if } t = 1 \text{ and } p \leq r(n). \end{cases}$$

Let

$$P(A) = \sum_{k \varepsilon K} \prod_{p \leq r(n)} P(E(p, \delta_p(k), n))$$

whenever

$$A = \bigcup_{k \varepsilon K} E_{n,k} \, .$$

It is shown in Kubilius (1964, p.27) that uniformly for all $A \varepsilon \underline{F}_n$,

$$\frac{1}{n} N_n \{m \varepsilon A\} - P(A) = O(\exp[-a \log n/\log r(n)])$$

as $n \to \infty$, where a is a positive constant.

On E_N, define the random variables $\xi^*_{Np} = \xi^*_{Np}(m)$, $(p \leq r(N))$ by putting

$$\xi^*_{Np}(m) = [f_N(p) \, (\delta_p(m) - \frac{1}{p}) \, / \, B(N)].$$

It is easy to see that $\{ \xi^*_{Np} : p \leq r(N) \}$ is a sequence of independent random variables such that for each $p \leq r(N)$

$$P \{ \xi^*_{Np} = - f_N(p) / pB(N) \} = 1 - \frac{1}{p}$$

and

$$P \{ \xi^*_{Np} = f_N(p) (1 - \frac{1}{p}) / B(N) \} = \frac{1}{p} .$$

§ 4. We now turn to the proof of Theorem 2.1.

By using (2.1.1) we can find a sequence $\varepsilon(N)$ of positive real numbers tending to zero such that

$$\sum_{|f_N(p)| > \varepsilon(N)B(N)} \frac{[f_N(p)]^2}{p} < \varepsilon(N) B^2(N).$$

If we put $r(N) = N^{\varepsilon(N)}$, clearly we get as in the proof of Lemma 2.1 that

$$[B(N, r(N)) / B(N)] \to 1$$

as $N \to \infty$. Let for each N, $\{ \xi_{Np} : p \leq N \}$ be a sequence of independent random variables such that

$$P \{ \xi_{Np} = - f_N(p) / p B(N) \} = 1 - \frac{1}{p} ,$$

and

$$P\{ \xi_{Np} = f_N(p) \ (1 - \frac{1}{p}) \ / \ B(N) \} = \frac{1}{p} .$$

Let

$$X_N(x) = \sum_{\substack{p \leq r(N) \\ B^2(N,p) \leq x B^2(N)}} \xi_{Np} \quad , \quad x \varepsilon [o, 1].$$

Clearly $X_N \varepsilon D[o, 1]$. Define Y_N by

$$Y_N(x) = \sum_{p \leq q} \xi_{Np} \quad , \quad \text{if} \quad x B^2(N) = B^2(N, q), \quad q \leq N$$

and for other $x \varepsilon [o, 1]$, define $Y_N(x)$ by linear interpolation. Then $Y_N \varepsilon C[o,1]$. In view of the condition (2.1.1), the sequence $\{Y_N\}$ satisfies the hypotheses of Prohorov's theorem, see Parthasarathy (1967, Theorem 4.1, p.221). Hence

$$Y_N \xrightarrow{D} W'$$

where W' is the Wiener measure on $C[o,1]$. Let d denote the Skhorohod metric on $D[o,1]$. For any $\varepsilon > 0$, we have

$$P\{d(X_N,Y_N) > \varepsilon\} \le P\{\sup_{0\le x\le 1} |Y_N(x) - X_N(x)| > \varepsilon\}$$

$$\le P\{\sup_{p\le r(N)} |\xi_{Np}| > \frac{\varepsilon}{2}\} + P\{\sum_{r(N)<p\le N} |\xi_{Np}| > \frac{\varepsilon}{2}\}$$

$$\le \sum_{p\le r(N)} P\{|\xi_{Np}| > \frac{\varepsilon}{2}\} + (4/\varepsilon^2)\,(1 - [B^2(N,r(N))/B^2(N)])$$

$$\le (4/\varepsilon^2 B^2(N)) \left[\sum_{2|f_N(p)|(p-1)>\varepsilon p B(N)} \frac{f_N^2(p)}{p} + \sum_{p\le N} \frac{f_N^2(p)}{p^2} \right]$$

$$= o(1)$$

as $N \to \infty$ by (2.1.1). Since the Wiener measure of $(D[o,1] - C[o,1])$ is zero,

$$X_N \xrightarrow{D} W \quad \text{as} \quad N \to \infty .$$

It follows from the results of § 3 that, uniformly for all Borel subsets A of $D[o,1]$,

$$(2.4.1) \quad P\{X_N \in A\} = \frac{1}{N} N_N \{h_N(\cdot, m, r(N)) \in A\} + o(1)$$

where

$$h_N(x,m,n) = \frac{1}{B(N)} \sum_{\substack{p \le n \\ B^2(N,p)\le xB^2(N)}} f_N(p)(\delta_p(m) - \frac{1}{p}), \; x \in [o,1].$$

Now if we show, for every $\varepsilon > 0$, that

$$(2.4.2) \quad \frac{1}{N} N_N\left\{ \max_{r(N)<i\leq N} \left| \frac{f_N^{(n)}(m) - f_N^{(r(N))}(m) - A(N,n) + A(N,r(N))}{B(N)} \right| \geq \varepsilon \right\} \to 0$$

as $N \to \infty$, then from (2.4.1) it follows easily that

$h_N(\,\cdot\,, m, N) \xrightarrow{D} W$.

Note that the left hand side of (2.4.2) is not more than

$$\frac{1}{N\varepsilon} \sum_{m=1}^{N} \sum_{r(N)<p\leq N} \left| \frac{1}{B(N)} f_N(p) \left(\delta_p(m) - \frac{1}{p}\right) \right|$$

$$= O((1/\varepsilon\, B(N)) \left(\sum_{r(N)<p\leq N} \frac{|f_N(p)|}{p} \right))$$

But

$$\sum_{r(N)<p\leq N} \frac{|f_N(p)|}{p} \leq \Big[-\varepsilon(N) \log \varepsilon(N) + \frac{1}{B(N)} \sum_{\substack{r(N)<p\leq N \\ |f_N(p)|>\varepsilon(N)B(N)}} \frac{|f_N(p)|}{p}\Big]B$$

$$\leq -B(N)\,\varepsilon(N) \log \varepsilon(N) + (\varepsilon(N))^{\frac{1}{2}} B(N) (-\log \varepsilon(N) + o(1))^{\frac{1}{2}}$$

$$= o(B(N)) \quad \text{as} \quad N \to \infty .$$

Hence (2.4.2) holds. This completes the proof of Theorem [illegible].

CHAPTER 3

ON A CONJECTURE OF ERDOS

1. Introduction : Let f be a real-valued additive arithmetic function. Let $F \varepsilon \underline{P}$. Suppose that the density of $\{m : f(F(m)) \varepsilon I\}$ exists and is positive for some bounded interval I. Then, does $f(F(m))$ have a distribution ? This question is answered in this chapter affirmatively under some restrictions on F and I. In the particular case when $F(m) = m$, Erdos (1947) conjectured that the above question has an affirmative answer. Partial solution to Erdos conjecture is given by Paul (1967).

In the last section, we also give some necessary and sufficient conditions for f to have a distribution, which may help in understanding how the behaviour of f in a small interval determines the existance of the distribution.

Throughout this chapter we assume that $F \varepsilon \underline{P}$ and write $r(d)$ instead of $r(F,d)$ and $f_p(x)$ instead of $f_p(x,F)$.

We define :

$$B(v,n) = \left[\sum_{v < p \leq n} \frac{\{f'(p)\}^2 \, r(p)}{p} \right]^{1/2}$$

$$A(v,n) = \sum_{v<p\leq n} \frac{f'(p)\, r(p)}{p},$$

$$A(n) = A(0, n),$$

and

$$B(n) = B(0, n).$$

2. Main results :

Theorem 3.1 : Suppose f and F satisfy the condition

(3.2.1) $f(p^t)\, r(p^t) \to 0$ as $p \to \infty$ for $t=1,\dots,D_F-1$, if $D_F \geq 2$.

If for some real number a

$$\limsup_{n \to \infty} \frac{1}{n} N_n \{f(F(m)) = a\} > 0,$$

then $f(F(m))$ has a distribution.

Theorem 3.2 : Suppose f and F satisfy the condition (3.2.1). If $f(F(m))$ has a distribution in a bounded non-degenerate interval I and if this distribution is not uniform then $f(F(m))$ has a distribution.

Here, by a distribution on a bounded interval we mean a finite countably additive measure μ on I such that whenever a and b are interior points of I and $\mu(a) = \mu(b) = 0$, then

$$\lim_{n \to \infty} \frac{1}{n} N_n \{f(F(m)) \in (a, b)\}$$

exists and equals μ (a, b).

Theorem 3.3 : Let f be a non-negative additive arithmetic function. Suppose for some $c > 0$,

$$\overline{D} \{ f(F(m)) < c \} > 0.$$

Then

$$\sum_p \frac{f'(p) \, r(p)}{p} < \infty .$$

Theorem 3.4 : Suppose f and F satisfy the condition (3.2.1). If for some bounded interval I,

$$\lim_{n \to \infty} \frac{1}{n} N_n \{ f(F(m)) \in I\} = 1,$$

then f(F(m)) has a distribution.

Remark : Proof of Theorem 3.4, when $F(m) = m$, was supplied by Dr. E.M.Paul during one of our discussions.

3. Preliminary results :

Definition : Let P be a probability measure on the real line. Let F be the smallest closed subset of the real line such that $P(F) = 1$. Then F is called the spectrum of P.

<u>Lemma 3.1</u> : Let. $\sup_n B(n) < \infty$. Suppose f and F satisfy the condition (3.2.1). Then there exists a distribution function Q on the real line such that

$$(3.3.1) \qquad \lim_{n \to \infty} \frac{1}{n} N_n \{f(F(m)) < c + A(n)\} = Q(c)$$

for all continuity points c of Q.

Let the sequence $\{A(n)\}$ be bounded. Let $\theta_1 = \liminf_{n \to \infty} A(n)$ and $\theta_2 = \limsup_{n \to \infty} A(n)$. If $\theta_1 \le \theta \le \theta_2$, then $\{f(F(m)) + \theta : m \ge 1\}$ is contained in the spectrum of Q. If $\{A(n)\}$ is not a convergent sequence, then $\{f(F(m)) : m \ge 1\}$ is a dense subset of the real line.

<u>Proof</u> : Since $\{B(n)\}$ is a convergent sequence,

$$\sum_p \left[f_p(x) - \frac{f'(p)\, r(p)}{p}\right] \quad \text{converges a.e. on } \mathfrak{S} \ [P],$$

by Kolmogorov's three-series theorem. Here and in what follows P stands for the probability measure constructed, on $\mathfrak{S}$, by Novoselov (1966).

For $x \in \mathfrak{S}$, let

$$g(x) = \begin{cases} \sum_p \left[f_p(x) - \frac{f'(p)\, r(p)}{p}\right] & \text{if this converges} \\ 0 & \text{otherwise.} \end{cases}$$

Applying Lemma 1.8 or Turan-Kubilius inequality [Kubilius, 1964, Lemma 3.1, p. 31] for cases $D_F \geq 2$ and $D_F = 1$, we have

$$\lim_{n \to \infty} \frac{1}{n} N_n\{f(F(m)) - A(n) < c\} = P\{x : g(x) < c\} = Q(c) \text{ (say)}$$

for all continuity points c of Q. By Egoroff's theorem we conclude that $[f'(p) \cdot r(p)/p] \to 0$ as $p \to \infty$. Hence $\theta_1 \leq \theta \leq \theta_2$ implies that θ is a limit point of the sequence $\{A(n)\}$.

Let $\theta_1 \leq \theta \leq \theta_2$. Let $\{n_r\}$ be such that $\lim_{r \to \infty} A(n_r) = \theta$. We have,

$$\lim_{r \to \infty} \frac{1}{n_r} N_{n_r}\{f(F(m)) < c\} = Q(c + \theta).$$

Fix $\varepsilon > 0$ and $m \geq 1$. Let $F(m) = q_1^{x_1} \cdots q_n^{x_n}$, where $q_1, \ldots, q_n$ are prime numbers and $x_1, \ldots, x_n$ are positive integers. By Egoroff's theorem, for any $\delta > 0$, we can find a $H \subset \mathfrak{S}$ and an r_o such that $P(H) > 1 - \delta$ and for all $r > r_o$, H is a subset of $\{x : |g(x) - \theta - \sum_{p \leq n_r} f_p(x)| < \varepsilon\}$.

Let $n' = n_{r_1} \geq \max(q_1, \ldots, q_n, n_{r_o})$. Let

$$H' = \{x \in H : x = y.F(m) \text{ and } p \leq n' \text{ implies } p \not| y\}.$$

Then

$$P\{x : |g(x) - \theta - f(F(m))| < \varepsilon\}$$

$$\geq P\{x \in H' : |g(x) - \theta - \sum_{p \leq n'} f_p(x)| < \varepsilon\}$$

$$\geq (1-\varepsilon)\,\eta\left[\frac{r(q_1^{x_1})}{q_1^{x_1}} - \frac{r(q_1^{x_1+1})}{q_1^{x_1+1}}\right]\cdots\left[\frac{r(q_n^{x_n})}{q_n^{x_n}} - \frac{r(q_n^{x_n+1})}{q_n^{x_n+1}}\right]$$

$$> 0$$

where

$$\eta = \prod_{\substack{p \neq q_1, \ldots, q_n \\ p \leq n'}} \left(1 - \frac{r(p)}{p}\right).$$

Thus $f(F(m)) + \theta$ is in the spectrum of Q. This completes the proof of the lemma.

<u>Lemma 3.2</u>: Let $\{X_n\}$ be a sequence of independent purely discrete random variables such that $\sum_n X_n$ converges almost

everywhere. Then the distribution of $\sum_n X_n$ is purely discrete, purely continuous singular or purely absolutely continuous. Moreover if

$$d_n = \sup_{d \text{ real}} P\{X_n = d\}$$

then the distribution of $\sum_n X_n$ is continuous if, and only if,

$$\prod_n d_n = 0.$$

For a proof, see Jessen and Wintner (1935, Theorem 35).

4. Proofs of the main results :

Proofs of Theorems 3.1 and 3.2 : By Lemma 1.7 and the hypotheses it follows, in either case, that

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

Let Q and g be as in the proof of Lemma 3.1.

In case of Theorem 3.1 we have for some sequence $\{n_k\}$ of natural numbers

$$(3.4.1) \qquad \lim_{k \to \infty} \frac{1}{n_k} N_{n_k}\{f(F(m)) = a\} > 0 .$$

It follows easily from (3.3.1) and (3.4.1) that $|A(n_k)| \leq M$ for every $k \geq 1$ and for some M. Let $\theta = \limsup_{k \to \infty} A(n_k)$. So

$$0 < \lim_{n_k \to \infty} \frac{1}{n_k} N_{n_k} \{f(F(m)) = a\} \leq P\{x : g(x) + \theta = a\}.$$

Hence the distribution of $g(x)$ is discrete by Lemma 3.2. Again by Lemma 3.2 we have

$$\sum_{f(p)r(p) \neq 0} \frac{r(p)}{p} < \infty .$$

Thus $\{A(n)\}$ is a convergent sequence. Hence it follows by Theorem 1.1 that $f(F(m))$ has a distribution. This proves Theorem 3.1.

If $\liminf_{n \to \infty} \frac{1}{n} N_n \{f(F(m)) \in I\} > 0$ for some bounded interval I, then it follows easily from (3.3.1) that $|A(n)| \leq M$ for all n and for some $M > 0$.

Let

$$\theta_1 = \liminf_{n \to \infty} A(n)$$

and

$$\theta_2 = \limsup_{n \to \infty} A(n).$$

In case of Theorem 3.2, if $\theta_1 = \theta_2$, then there is nothing to prove. If $\theta_1 < \theta_2$, from the proof of

Lemma 3.1 it follows that $\theta \varepsilon [\theta_1, \theta_2]$ implies θ is a limit point of $\{A(n)\}$ and the distribution function of g is continuous. If a and b are in the interior of I we get for all $\theta \varepsilon [\theta_1, \theta_2]$,

$$\lim_{n \to \infty} \frac{1}{n} N_n \{f(F(m)) \varepsilon (a,b)\} = P\{x : g(x) + \theta \varepsilon (a,b)\}$$

By changing θ continuously we get uniform distribution on I, which contradicts the hypothesis. Hence $\{A(n)\}$ is a convergent sequence. So by Theorem 1.1 we conclude that the distribution of $f(F(m))$ exists. This completes the proof of Theorem 3.2.

<u>Proof of Theorem 3.3</u> : Let p_o be as in Lemma 1.4. Suppose that $\sum_p \frac{f'(p)\, r(p)}{p} = +\infty$, then by Kolmogorov's three-series theorem $\sum_{p>p_o} f_p(x)$ diverges to $+\infty$ almost everywhere. Let for $x \varepsilon \mathfrak{S}$

$$h_n(x) = \sum_{p_o < p \leq n} f_p(x).$$

Note that the distribution of h_n is discrete for each n. Suppose for some c

$$\overline{D} \{ f(F(m)) < c \} > \delta .$$

Let $b > c$ be any common continuity point of the distributio[n] of the functions $h_{p_o+1}, h_{p_o+2}, \ldots$. Then as $n \to \infty$,

$$D\{0 \leq h_n(m) \leq b\} \to 0.$$

Let $\{k'\}$ be a subsequence of $\{N_k\}$ such that

$$\lim_{k' \to \infty} \frac{1}{k'} N_{k'} \{f(F(m)) < b\} = \overline{D}\{m : f(F(m)) < b\} = z > o \quad \text{(say)}$$

By using Lemma 1.5, choose $t > D_F$ such that for all n,

$$N_n \{p^t | F(m) \text{ for some } p\} < \frac{z}{4} n + o(n).$$

Let $\{N_{n'}\}$ be an increasing subsequence of $\{k'\}$ that $k > n'$ implies $p^t | N_k$ if $p \leq n$. Then

$$\frac{1}{N_{n'}} N_{N_{n'}} \{ f(F(m)) < b \}$$

$$\leq P\{x : \sum_{p_o < p \leq n} f_p(R_{n'}(x)) < b \}$$

$$\leq P\{x : \sum_{p_o < p \leq n} f_p(x) < b\} + \frac{z}{2} + o(1)$$

as $n \to \infty$. Hence as $n \to \infty$ the left-hand side converges to z and the right-hand side converges to $\frac{z}{2}$ which is a contradiction, since $z > 0$. Therefore,

$$\sum_p \frac{f'(p)\, r(p)}{p} < \infty .$$

This completes the proof of Theorem 3.3.

Proof of Theorem 3.4 : It follows from Lemma 1.7 that

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

From Lemma 3.1 we get that for every $m \geq 1$, $f(F(m))$ belongs to the closure of I ; i.e. for all $m \geq 1$, $|f(F(m))| < M$ for some M.

Now define a new additive arithmetic function g by

$$g(p^t) = |f(p^t)|$$

Note that for all $m \geq 1$, $g(F(m))$ is positive and $g(F(m)) < 2M$

Hence by Theorem 3.3, we get that $\sum_p \frac{g'(p)\, r(p)}{p} < \infty$. So $\sum_p \frac{f'(p)\, r(p)}{p}$ converges absolutely. Hence $f(F(m))$ has a distribution by Theorem 1.1.

§5. In what follows, we use the terminology introduced by Paul (1962 a).

Let $M_L(S)$ and $M_U(S)$ denote the lower and the upper magnifications of S, where S is any set of positive integer Let $\lambda(S)$ denote the logarithmic density of S, whenever it exists. Let $\overline{P}$ denote the measure on $Z_0^{\aleph_0}$ as introduced in Paul (1962a). (Here Z_0 denotes the set of non-negative integers.)

Theorem 3.5 : The following are equivalent :

i) There is a real number a such that for all $\varepsilon > 0$

$$\overline{P}\{M_L(m : |f(m) - a| < \varepsilon)\} > 0$$

ii) For all $\varepsilon > 0$, $\overline{P}\{\bigcup_{n=1}^{\infty} A_{n,\varepsilon}\} > 0$ where

$$A_{n,\varepsilon} = \{x : |\sum_{r \leq i \leq k} f(p_i^{x_i})| < \varepsilon, \text{for all } r, k \geq n\}$$

iii) f has a distribution.

Proof : Proof of Theorem 1 of Paul (1963) shows that if f has a distribution,

$$\lambda\{m : f(m) \in (c, d)\} = \overline{P}\{M_L(m : f(m) \in (c, d))\}.$$

By Theorem 1 of Paul (1967) we have for every $\varepsilon > 0$

$D\{m : |f(m)| < \varepsilon\} > 0$. Hence,

$$\overline{P}\{M_L(m : |f(m)| < \varepsilon)\} > 0 \quad \text{for all } \varepsilon > 0.$$

This proves the implication (iii) $\Rightarrow$ (i).

i) $\Rightarrow$ (ii) is clear because

$$M_L\{m : f(m) \in (a - \varepsilon, a + \varepsilon)\} \subset \bigcup_{n \geq 1}^{\infty} A_{n, 2\varepsilon}.$$

To prove (ii) $\Rightarrow$ (iii), assume (ii). There exists a N such that $\overline{P}(A_{N,\varepsilon}) > 0$. Hence,

$$\overline{P}\{x : k > N \Rightarrow |\sum_{N<i\leq k} f(p_i^{x_i})| < \varepsilon\} > 0.$$

So,

$$\limsup_{k \to \infty} \overline{P}\{x : |\sum_{N<i\leq k} f(p_i^{x_i})| < \varepsilon\} > 0.$$

Hence there exists a sequence $\{d_i\}$ such that

$$\sum_{i=1}^{\infty} \{f(p_i^{x_i}) + d_i\} \text{ converges a.e. } [\overline{P}].$$

Let $\delta > 0$. Find an n_1 such that

$$\overline{P}(A_{n_1, \delta/2}) = \eta > 0.$$

By Egoroff's theorem choose a measurable set H such that

$\overline{P}(H) > 1 - \frac{\eta}{2}$ and $\sum_{i=1}^{\infty} \{f(p_i^{x_i}) + d_i\}$ converges uniformly on H.

Hence there exists $n > n_1$ such that

$$\overline{P}\{x : \text{for all } k, r \geq n, |\sum_{r<i\leq k} \{f(p_i^{x_i}) + d_i\}| < \frac{\varepsilon}{2}\} > 1 - \frac{\eta}{2}.$$

Hence there exists x such that for all $r, k \geq n$, we have

$$|\sum_{r<i\leq k} f(p_i^{x_i})| < \frac{\varepsilon}{2} \text{ and } |\sum_{r<i\leq k} \{f(p_i^{x_i}) + d_i\}| < \frac{\varepsilon}{2}.$$

Hence for all $r, k \geq n$, $|\sum_{r<i\leq k} d_i| < \varepsilon$. So $\sum d_i$ converges, hence $\sum_{i=1}^{\infty} f(p_i^{x_i})$ converges almost everywhere. Therefore f has a distribution (see Paul, 1963). This completes the proof of Theorem 3.5.

CHAPTER 4

SPECTRUM OF THE DISTRIBUTION OF VALUES OF ARITHMETIC FUNCTIONS

1. <u>Introduction</u> : By the spectrum of a probability measure on the real line we mean the smallest closed subset of the real line whose P-measure is 1.

Let f be a real-valued additive arithmetic function. In this chapter we characterize the spectrum of the distribution of $\{f(n) - f(n+1),\ldots,f(n+h-1) - f(n+h)\}$ whenever the distribution exists, where h is a positive integer. We obtain a theorem of Erdos and Schinzel (1961) as a corollary of one of our theorems Under very general conditions we shall show that for any $m \geq 1$, $\{f_1(F_1(m)),\ldots,f_h(F_h(m))\}$ 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-valued additive arithmetic functions and $F_1,\ldots,F_h \in \underline{P}$.

2. <u>Main results</u> :

<u>Theorem 4.1</u> : Let $F \in \underline{P}$. Suppose f and F satisfy the condition (3.2.1) and suppose that $f(F(m))$ has a distribution. Then $f(F(1))$, $f(F(2)),\ldots,$ all belong to the spectrum of the distribution.

This theorem follows easily from Lemma 3.1 and Theorem 1.1.

Theorem 4.2 : Suppose that the series $\sum_p \frac{[f'(p)]^2}{p}$ is convergent. Then for any positive integer h,

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

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 (4.2.1).

Moreover, if $N_0, N_1, \ldots, N_h$ are positive integers such that for all $i = 0, 1, \ldots, h$,

$$(N_i, (h+1)!) = 1 \text{ and } (N_i, N_j) = 1 \quad (0 \leq i < j \leq h),$$

then

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

is in the spectrum of the distribution of (4.2.1).

Corollary [Erdos and Schinzel, 1961] : Let f be a real-valued additive arithmetic function satisfying the following conditions :

(4.2.2) $$\sum_p \frac{[f'(p)]^2}{p} < \infty .$$

(4.2.3) 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 h real numbers $a_1, \ldots, a_h$ and for any $\varepsilon > 0$, the set

$$\{n \geq 1 : |f(n+i) - f(n+i-1) - a_i| < \varepsilon, \quad i = 1, \ldots, h\}$$

has positive natural density.

<u>Theorem 4.3</u> : Let $f_1, \ldots, f_s$ be real-valued additive arithmetic functions and let $F_1, \ldots, F_s$ be in $\underline{P}$. Suppose

$$f_i(p^k)\ r(F_i, p^k) \to 0 \quad \text{as} \quad p \to \infty \quad \text{for} \quad k = 1, \ldots, D_{F_i} - 1$$

whenever $D_{F_i} \geq 2$. If the distribution of

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

exists, then one can find a k_o such that, the spectrum S of the distribution of (4.2.4) is the closure of the set

$$A = \{ [\sum_{\substack{p^t \| F_i(m) \\ p \leq k}} f_i(p^t); \ i = 1, \ldots, s] : m \geq 1, \ k > k_o \}.$$

<u>Remark</u> : Clearly $A \supset B = \{[f_i(F_i(m)) ; i = 1,\dots,s] ; m \geq 1\}$.

3. <u>Proofs</u> :

<u>Proof of Theorem 4.2</u> : Let $H_{i-1}(n) = f(n+i-1) - f(n+i)$, $i = 1,\dots,h$. First we shall extend the functions H_i to the polyadic domain $\mathfrak{S}$ and show that each $H_i \in \mathfrak{L}_o$.

Let for $x \in \mathfrak{S}$

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

For any prime number p define

$$\bar{f}_{ip}(x) = \sum_{k=1}^{\infty} f(p^k)\, \omega(p^k, x+i), \quad i = 0,1,\dots,h-1.$$

Since the random variables $\{\bar{f}_{ip}(x) : p \text{ prime}\}$ are mutually independent [see Novoselov, 1966] and $\sum_p \frac{(f'(p))^2}{p} < \infty$, by Kolmogorov's three-series theorem, it follows that

$$\sum_p [\bar{f}_{ip}(x) - \frac{f'(p)}{p}] \quad \text{converges a.e. [P] for } i=0,1,\dots$$

Hence

$$\sum_p [\bar{f}_{ip}(x) - \bar{f}_{(i+1)p}(x)] \quad \text{converges a.e. [P] for } i=0,1,\dots$$

Moreover it is easy to see that the random variables

$\{[\bar{f}_{ip}(x) - \bar{f}_{(i+1)p}(x)] : p \text{ prime}\}$ are mutually independent for each $i = 0,1,\ldots,h-1$.

Let

$$g_i(x) = \begin{cases} \Sigma [\bar{f}_{ip}(x) - \bar{f}_{(i+1)p}(x)] & \text{if it converges,} \\ 0 & \text{otherwise.} \end{cases}$$

Clearly g_i is an extension of H_i. By using Turán-Kubilius inequality and Lemma 1.2 we get that $H_i \in \mathcal{S}_o$ and the distribution of H_i is $Q_i(c) = P\{x : g_i(x) < c\}$.

Note that for any h-tuple $(t_o,\ldots,t_{h-1})$ of real numbers, the distribution of $\sum_{i=0}^{h-1} t_i H_i(n)$ is given by $P\{x : \sum_{i=o}^{h-1} t_i g_i(x) < c\}$.

Hence by Cramer-Wold device, we get that the distribution of $\{H_o(n), H_1(n),\ldots,H_{h-1}(n)\}$ is given by

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

Let $0 < \delta < 1$. Since

$$\{[\bar{f}_{op}(x) - \bar{f}_{1p}(x),\ldots, \bar{f}_{(h-1)p}(x) - \bar{f}_{hp}(x)] : p \text{ prime}\}$$

is a sequence of mutually independent random vectors, by

Put $Q_1 = \prod_{\substack{p \leq k \\ p \nmid N}} p$ and $Q_2 = N^2 Q_1$. We have

$$P\{x : \sum_{p \leq k} [\bar{f}_{ip}(x) - \bar{f}_{(i+1)p}(x)] = H_i(n_o) \text{ for } i = 0,\ldots,h-1\}$$

$$= D\{n : \sum_{p \leq k} [\bar{f}_{ip}(n) - \bar{f}_{(i+1)p}(n)] = H_i(n_o) \text{ for } i=0,\ldots,h-1\}$$

$$\geq \frac{1}{Q_2} > 0.$$

In fact, since $(Q_1, N) = 1$, we can find an integer u such that

$$u \equiv n_o \pmod{N^2} \quad \text{and}$$
$$u \equiv 1 \pmod{Q_1} .$$

It is easy to show that for any integer t and any i in $\{0,1,\ldots,h$

$\dfrac{Q_2 t + u + i}{n_o + i}$ is an integer not divisible by any prime $p \leq k$.

Since $k > N^2$, we have

$$\left(\frac{Q_2 t + u + i}{n_o + i}, \; n_o + i \right) = 1.$$

Hence for any t such that $Q_2 t + u > 0$, we get that

$$\sum_{p \leq k} \{ \bar{f}_{ip}(Q_2 t+u) - \bar{f}_{(i+1)p}(Q_2 t+u)\}$$
$$= f(Q_2 t+u+i)-f(Q_2 t+u+i+1), \quad \text{for } i = 0,1,\ldots,h-1.$$

But the density of the positive integers of the form $Q_2t + u$ equal to $\frac{1}{Q_2}$. This proves the first part of Theorem 4.2. Pr of the second part is similar to the above proof. So here we note the following fact. We put

$$N = N_o . N_1 \ldots N_h, \quad Q_1 = \prod_{\substack{p \leq k \\ p \nmid N}} \quad \text{and} \quad Q_2 = (h+1)!\, N^2\, Q_1.$$

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

$$u \equiv 1 \pmod{(h+1)!\, Q_1},$$

$$u \equiv -i + N_i \pmod{N_i^2}, \quad 0 \leq i \leq h.$$

It is easy to see that for every integer t the numbers $\{(Q_2t + u + i)/(i+1)\, N_i\}$, $(i = 1, \ldots, h)$ are integers which are not divisible by any prime $p \leq k$. Also the density of integers $Q_2t + u$ is $\frac{1}{Q_2} > 0$.

This completes the proof of Theorem 4.2.

<u>Proof of the Corollary</u> : Let ε be a positive number and let a sequence a_i $(i = 1, \ldots, h)$ be given. By condition (4.2.3) we can find positive integers $N_o, N_1, \ldots, N_h$ such that

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

$$f(N_o) > c_1 + \max_{1 \leq i \leq h} \{f(i+1) - \sum_{j=1}^{i} a_j\},$$

and $\left| f(N_i) - \{f(N_o) - f(i+1) + \sum_{j=1}^{i} a_j\} \right| < \frac{\varepsilon}{4} \quad (1 \leq i \leq h).$

Hence

(4.3.1) $\quad |f((i+1)N_i) - f(i\, N_{i-1}) - a_i| < \frac{\varepsilon}{2} \quad (1 \leq i \leq h).$

By Theorem 4.2, we see that the set

(4.3.2) $\quad \{n \geq 1 : |f(n) - f(n+1) - f(N_o) + f(2N_1)| < \frac{\varepsilon}{2}, \ldots,$

$$|f(n+h-1) - f(n+h) - f(h\, N_{h-1}) + f((h+1)N_h)| < \frac{\varepsilon}{2}\}$$

has positive density. Hence the corollary follows from (4.3.1) and (4.3.2).

Proof of Theorem 4.3 : We need the following lemma.

Lemma 4.1 : If $h(m)$ and $g(m)$ are integer-valued polynomials having no common factors, then there exists k_1 such that $p > k_1$ implies that there is no m such that $h(m) \equiv 0 \pmod p$ and $g(m) = 0 \pmod p$.

See Katai (1969).

Let $F_i(m) = \prod_{j=1}^{r_i} F_{ij}(m)$, where $\{F_{ij}(m) : j = 1, \ldots, r_i\}$

are irreducible and each $F_{ij} \in \underline{P}$. Such a factorization is possible and is unique. Let

$$\{G_1, \ldots, G_h\} = \{F_{ij} : j = 1, \ldots, r_i, \ i = 1, \ldots, s\}$$

such that G_i and G_j have no common factors if $i \neq j$. By Lemma 4.1 choose a k_1 such that $p > k_1$ implies that there is no m with the property that $G_i(m) \equiv 0 \pmod p$ and $G_j(m) \equiv 0 \pmod p$ $(1 \leq i < j \leq h)$. Let $G_i(x)$ be the continuous extension of $G_i(m)$ to the Novoselov's space $\mathfrak{S}$.

It is easy to see that,

$$\{(m_i | G_i(x) : i = 1, \ldots, h), \ q_1^{t_{1i}} | G_i(x), \ i = 1, \ldots, h), \ldots.$$

$$(q_u^{t_{ui}} | G_i(x), \ i = 1, \ldots, h)\}$$

are independent events if t_{ij} are non-negative integers, $u \geq 1$, $q_i > k_1$, $q_i \neq q_j$ if $i \neq j$ and m_i is not divisible by any prime $p > k_1$ $(i = 1, \ldots, h)$. Since either $F_{ij}(m) \equiv F_{ut}$ or $F_{ij}(m)$ and $F_{ut}(m)$ are mutually prime, we get that

$$\{(m_i | F_i(x), \ i = 1, \ldots, s), \ (q_1^{t_{1i}} | F_i(x), \ i = 1, \ldots, s), \ldots,$$

$$(q_r^{t_{ri}} | F_i(x), \ i = 1, \ldots, s)\}$$

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

Now choose $k_c > k_1$ (by using Lemma 1.4) such that, if $p > k_o$, then, $r(F_i,p^t) = r(F_i,p)$ for $t \geq 1$, $i = 1,\ldots,s$ and $r(F_i,p) < \frac{p}{2s}$, $i = 1,\ldots,s$.

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

$$f^*_{io}(x) = \sum_{\substack{p^k \| F_i(x) \\ p \leq k_o}} f_i(p^k) \qquad (i = 1,\ldots,s).$$

For $p > p_o$ and $i = 1,\ldots,s$, we put for $x \in \mathfrak{S}$

$$f^*_{ip}(x) = \begin{cases} f_i(p^k) & \text{if } p^k \| F_i(x),\ k \geq 1 , \\ 0 & \text{if either } p \nmid F_i(x) \text{ of } p^k | F_i(x) \text{ for all } k \geq 1. \end{cases}$$

By Theorem 1.1 we conclude that $\sum_p \frac{f_i'(p)\, r(F_i,p)}{p}$ and $\sum_p \frac{(f_i'(p))^2\, r(F_i,p)}{p}$ converge. Hence by Kolmogorov's three - series theorem $\sum_{p>k_c} f^*_{ip}(x)$ converges a.e. [P].

Fix a positive real number $\delta < (1/4s)$. By Egoroff's theorem choose $H \subset \mathfrak{S}$ such that $P(H) > 1 - \delta$ and on H, $\sum_{p > k_o} f^*_{ip}(x)$ converges uniformly for $i = 1,\ldots,s$. Now fix $\varepsilon > 0$, $k > k_o$ and $m \geq 1$. Choose $k_2 > k$ such that

$$P\{x : |\sum_{p > k_2} f^*_{ip}(x)| < \varepsilon ; i = 1,\ldots,s\} > 1 - \eta$$

where $\eta = \delta . s$. Then

$$D\{|f_i(F_i(n)) - \sum_{k_o < p \leq k} f^*_{ip}(m) - f^*_{io}(m)| < \varepsilon ; i = 1,\ldots,s\}$$

$$\geq P\{x ; f^*_{io}(x) = f^*_{io}(m), \sum_{k_o < p \leq k} f^*_{ip}(x) = \sum_{k_o < p < k} f^*_{ip}(m)$$

$$f^*_{ip}(x) = 0 \text{ for } k < p \leq k_2 \text{ and } |\sum_{p > k_2} f^*_{ip}(x)| < \varepsilon ; i = 1,\ldots$$

$$\geq (1-\eta) P\{x : f^*_{io}(x) = f^*_{io}(m), i = 1,\ldots,s\}$$

$$\times \prod_{k_o < p \leq k} P\{x : f^*_{ip}(x) = f^*_{ip}(m), i = 1,\ldots,s\}$$

$$\times \prod_{k < p \leq k_2} P\{x : f^*_{ip}(x) = 0, i = 1,\ldots,s\}.$$

Clearly $P\{x : f^*_{ip}(x) = 0, (i=1,\ldots,s)\} = 1 - P\{x : f^*_{ip}(x) \neq 0 \text{ for some } i\}$

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

Suppose $k_o < p \leq k$ and $p^{u_{ij}} \| F_{ij}(m)$ for some $u_{ij} \geq 1$ and for some i, j.

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

$$P\{x : f^*_{ip}(x) = f^*_{ip}(m), \quad i = 1, \ldots, s\} \geq P\{x : p^{u_{ij}} \| F_{ij}(x)\}$$

$$= \frac{r(F_{ij}, p^{u_{ij}})}{p^{u_{ij}}} - \frac{r(F_{ij}, p^{u_{ij}+1})}{p^{u_{ij}+1}} > 0.$$

Let $\phi_i(m) = \prod_{\substack{p^u \| F_i(m) \\ p \leq k_o}} p^u$. Note that

$$P\{x : f_{io}(x) = f_{io}(m), \quad i = 1, \ldots, s\}$$

$$\geq D\{\phi_i(m) | F_i(n) \text{ and } \phi_i(m).p \nmid F_i(n) \text{ for any } p \leq k_o \text{ and for } i = 1, \ldots, s\}$$

$$> 0$$

(since $n=m$ 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 Theorem 4.3.

CHAPTER 5

SINGULARITY AND ABSOLUTE CONTINUITY OF DISTRIBUTIONS OF ADDITIVE ARITHMETIC FUNCTIONS

1. Introduction : It is known that, the distribution of a real-valued additive arithmetic function f, if it exists, is pure ; i.e., it is either discrete, continuous singular or absolutely continuous. It is also known that the distribution of a real-valued additive arithmetic function is discrete if and only if

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

Erdos (1939) has shown that, if f is a real-valued additive arithmetic function satisfying $f(p) = O(p^{-c})$ for all prime numbers p and for some positive constant c, then the distribution of f exists and is singular. [Here and in what follows, singular means discrete or continuous singular.] In this chapter we show that if for some $c > 0$

$$(5.1.1) \qquad \sum_{\substack{p>N \\ p\notin A}} \frac{[f(p)]^2}{p} = O(N^{-c}) \quad \text{as} \quad N \to \infty ,$$

where A is a set of prime numbers such that $\sum_{p\in A} \frac{1}{p} < \infty$, then the distribution of $f(m) - f(m+1)$ exists and is singular.

From this result we shall deduce that if f satisfies (5.1.1) and if f has a distribution, then the distribution of f is singular. In particular, every bounded real-valued additive arithmetic function has a singular distribution. We also obtain similar results for the distribution of values of $f(F(m))$, where F is an integral polynomial taking positive values for $m=1,2,...$

Suppose f has an absolutely continuous distribution. We give some sufficient conditions which ensure that $f(F(m))$ has an absolutely continuous distribution, and we shall give an example to show that, in a sense, this is the best possible result.

Some of the proofs depend on the following observation. Let f be a real-valued additive arithmetic function, having a distribution. Distribution of f is singular (absolutely continuous) if and only if the distribution corresponding to the infinitely divisible characteristic function (see Lukacs, 1970) g given by

$$\log. g(t) = \sum_{|f(p)|\leq 1} \frac{1}{p} [e^{it\, f(p)} - 1 - it\, f(p)]$$

is singular (absolutely continuous).

Let L(X) denote the distribution function corresponding to the random variable X and let $\omega(m)$ denote the number of distinct prime divisors of m.

2. Main results :

Theorem 5.1 : If f is a real-valued additive arithmetic function satisfying (5.1.1), then the distribution of $f(m) - f(m+1)$ exists and is singular.

Theorem 5.2 : If f is a real-valued additive arithmetic function having absolutely continuous distribution, then the distribution of $f(m) - f(m+1)$ is absolutely continuous.

Corollary 5.1 : Suppose that f is a real-valued additive arithmetic function having a distribution. If f satisfies the condition (5.1.1), then the distribution of f is singular.

Corollary 5.2 : The distribution of every bounded real-valued additive arithmetic function is singular. In particular no additive arithmetic function can have uniform distribution.

Theorem 5.3 : Suppose g is any real-valued additive arithmetic function for which there exists a constant K such that

$$(5.2.1) \qquad |g(m) - g(m+1)| < K \quad \text{for } m = 1,2,\ldots$$

Then the distribution of $g(m) - g(m+1)$ exists and is singular.

Theorem 5.4 : Let f be a real-valued additive arithmetic function satisfying

$$(5.2.2)\quad \liminf_{\varepsilon \to 0} [1/(\varepsilon^2 |\log \varepsilon|)] \sum_{|f(p)|<\varepsilon} \frac{[f(p)]^2}{p} > 4.$$

Then the distribution of f, if it exists, is absolutely continuous.

Theorem 5.5 : Let $F \varepsilon \underline{\underline{P}}$. Suppose f and F satisfy (3.2.1) and

$$\liminf_{\varepsilon \to 0} [1/(\varepsilon^2 |\log \varepsilon|)] \sum_{|f(p)|<\varepsilon} \frac{1}{p}[f(p)]^2 \, r(F,p) > 4.$$

Then the distribution of $f(F(m))$, if it exists, is absolutely continuous.

Theorem 5.6 : Let $F \varepsilon \underline{\underline{P}}$. Let f be a real-valued additive arithmetic function satisfying (3.2.1) and let

$$(5.2.3)\quad \sum_{\substack{p>N \\ p \not\varepsilon A}} \frac{1}{p} [f(p)]^2 \, r(F,p) = O(N^{-c}) \quad \text{as} \quad N \to \infty$$

where c is a positive constant and A is a set of primes such that $\sum_{p \varepsilon A} \frac{1}{p} r(F, p) < \infty$.

Then the distribution of $f(F(m))$, if it exists, is singular.

Theorem 5.7 : Under the conditions of Theorem 5.6, $f(F(m)) - f(F(m+1))$ has a singular distribution.

Theorem 5.8 : Let $F \in \underline{P}$. Let f be a real-valued additive arithmetic function such that

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

if $D_F \geq 2$. (This condition can be dropped in case F is a product of linear polynomials.) Let A be a set of primes such that

$$(5.2.4) \quad \sum_{p \in A} \frac{1}{p} < \infty \text{ and } q \notin A \text{ implies either } r(F, q) \neq 0$$

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

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.

Theorem 5.9 : Let $F \in \underline{P}$ and $D_F \geq 2$. Let f be a real-valued additive arithmetic function satisfying (3.2.1). Suppose A is a set of primes satisfying (5.2.4). If the distribution of $f(n) - f(n+1)$ exists and is absolutely continuous then the distribution of $f(F(m)) - f(F(m+1))$ also exists and is absolutely continuous.

3. Preliminary results :

Lemma 5.1 : Let X and Y be two independent random variables. If Y is discrete, then the distribution of $X + Y$ is discrete, continuous singular or absolutely continuous according as X is discrete, continuous singular or absolutely continuous.

For a proof of this well known lemma, see Lukacs (1970, Lemma 3.7.4, p. 57).

Lemma 5.2 : Suppose $\{X_n\}$ is a sequence of discrete and independent random variables and suppose $\{Y_n\}$ is another sequence of discrete and independent random variables, all of them defined on the same probability space. If

$$\sum_n P\{X_n \neq Y_n\} < \infty ,$$

then $\sum_n X_n$ converges almost everywhere and $L(\sum_n X_n)$ is absolutely continuous (singular) if and only if, $\sum_n Y_n$ converges almost everywhere and $L(\sum_n Y_n)$ is absolutely continuous (singular).

Proof : In view of Lemma 5.1, we can assume without loss of generality that

$$\sum_n P\{ X_n \neq Y_n \} < 1.$$

It is easy to see, by Kolmogorov's three-series theorem, that $\sum_n X_n$ converges if and only if, $\sum_n Y_n$ converges. If $L(\sum_n X_n)$ is singular then, there exists a Lebesgue null set M of the real line such that

$$P\{\sum_n X_n \in M\} = 1.$$

So

$$\begin{aligned} P\{\sum_n Y_n \in M\} &\geq P\{\sum_n Y_n \in M \text{ and } X_n = Y_n \text{ for all } n\} \\ &= P\{\sum_n X_n \in M \text{ and } X_n = Y_n \text{ for all } n\} \\ &= P\{X_n = Y_n \text{ for all } n\} \\ &\geq 1 - \sum_n P\{X_n \neq Y_n\} \\ &> 0. \end{aligned}$$

This shows that $L(\sum_n Y_n)$ is singular since, by Lemma 3.2, $L(\sum_n Y_n)$ is pure. This completes the proof of Lemma 5.2.

<u>Lemma 5.3</u> : Let f be a real-valued additive arithmetic function satisfying

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

Let $\{X_p\}$ be a sequence of independent random variables with

$$(5.3.1)\quad \begin{cases} P\{X_p = x\} = P\{X_p = -x\} = (1 - \frac{1}{p})\left[\sum\limits_{\substack{k \\ |f(p^k)|=x}} \frac{1}{p^k}\right] & \text{if } x > 0, \\ P\{X_p = 0\} = 1 - \frac{2}{p} + 2(1 - \frac{1}{p})\left[\sum\limits_{\substack{k \\ f(p^k)=0}} \frac{1}{p^k}\right]. \end{cases}$$

Then $\sum_p X_p$ converges almost everywhere and the distribution of $f(m) - f(m+1)$ is $L(\sum_p X_p)$.

This result is contained in the proof of Theorem 4.2.

<u>Lemma 5.4</u> : For any positive integer $k \geq 2$, there exists a constant c_k such that for $x \geq 3$ we have

$$\sum_{1 \leq m \leq x} k^{\omega(m)} - c_k\, x(\log x)^{k-1} = O(x(\log x)^{k-2}).$$

For a proof, see Kubilius (1964, Lemma 9.2, p.140).

<u>Lemma 5.5</u> : For $n \geq 3$ and $k \geq 2$, we have

$$\sum_{m=1}^{n} \frac{k^{\omega(m)}}{m} = b_k (\log n)^k + O((\log n)^{k-1}),$$

where b_k is a constant depending only on k.

Proof : Summing by parts gives,

$$\sum_{m=1}^{n} \frac{k^{\omega(m)}}{m} = \frac{1}{n} \sum_{m=1}^{n} k^{\omega(m)} + \int_{1}^{n} \frac{1}{x^2} \Big[\sum_{m \le x} k^{\omega(m)} \Big] dx$$

$$= O((\log n)^{k-1}) + c_k \int_{2}^{n} \frac{1}{x} (\log x)^{k-1} dx,$$

by Lemma 5.4

$$= O((\log n)^{k-1}) + \frac{1}{k} c_k (\log n)^k .$$

Lemma 5.6 : Suppose $\{a_p\}$ is a sequence of real numbers satisfying

(5.3.2) $$\sum_{p>N} \frac{a_p^2}{p} = O(N^{-c}) \text{ as } N \to \infty \text{ for some } c.$$

If g is defined by

(5.3.3) $$g(t) = \exp\Big\{ \sum_{p} \frac{1}{p} [e^{ita_p} - 1 - it\, a_p] \Big\},$$

then g is the characteristic function of an infinitely divisible distribution and the distribution corresponding to $|g(t)|^{2k}$ is singular for any positive integer k.

Proof : Since $\{\exp[\frac{1}{p}[e^{it\, a_p} - 1 - it\, a_p]]\}$ is characteristic function of a centred Poisson random variable, since

$$\sum_{p} \frac{a_p^2}{p} < \infty$$

and since

$$h_n(t) = \exp\left\{ \sum_{p \leq n} \frac{1}{p} \left[e^{it\,a_p} - 1 - it\,a_p \right] \right\}$$

is a characteristic function for each n, h_n converges absolutely and uniformly to g in every bounded interval and g is the characteristic function of an infinitely divisible distribution.

Fix an integer $k \geq 1$. Let $\{X_p, Y_q : p, q > 2k\}$ be a set of independent random variables satisfying, for each $p > 2k$,

$$P\{X_p = t\} = P\{Y_p = -t\} = \left(\frac{k}{p}\right)^t \left(1 - \frac{k}{p}\right)$$

for any integer $t \geq 0$.

Note that for any integer $t \geq 0$,

$$P\{X_p + Y_p = -t\} = P\{X_p + Y_p = t\}$$

$$= \left(\frac{k}{p}\right)^t (1 + O(p^{-2}))\, e^{-2k/p} .$$

In view of (5.3.2) and Lemma 5.2, it follows that $\sum_p a_p(X_p + Y_p)$ converges almost everywhere and $L(\sum_p a_p(X_p + Y_p))$ is singular if and only if the distribution corresponding to $|g(t)|^{2k}$ is singular.

Without loss of generality, we can assume $c < 1$, in (5.3.2). Let N be a large integer. Let $m < N^{c/6}$ and

$$m = p_{m'1}^{m1} \cdots p_{m'm''}^{mm''}, \quad mi \geq 1.$$

Consider the set

$$D_{m,N} = \left\{ \sum_{i=1}^{m''} \varepsilon_i \, mi \, a_{p_{m'i}} : \varepsilon_i = +1 \text{ or } -1, \; i=1,\ldots,m'' \right\}.$$

Put $D_N = \bigcup_{m<N^{c/6}} D_{m,N}$. Since there are $2^{\omega(m)}$ sequences $(\varepsilon_1,\ldots,\varepsilon_{m''})$ of $+1$ and -1, and since

$$C(m) = P\{X_{p_{m'i}} + Y_{p_{m'i}} = \varepsilon_i \, mi, \; i=1,\ldots,m'' \text{ and } X_p+Y_p = 0 \text{ if } p \leq N \text{ and } (p,m)=1\}$$

is same for all such sequences, we have

$$P\{ \sum_{p\leq N} a_p(X_p + Y_p) \in D_N \}$$

$$\geq \sum_{m<N^{c/6}} 2^{\omega(m)} C(m)$$

$$\geq \sum_{m<N^{c/6}} 2^{\omega(m)} \frac{k^{\omega(m)}}{m} \prod_{p\leq N} (1+O(p^{-2}))\exp\{-2k \sum_{p\leq N} \frac{1}{p}\}.$$

Since $\sum_{p\leq N} \frac{1}{p} = \log \log N + O(1)$, by Lemma 5.5, it follows that

$$P\left\{ \sum_{p \le N} a_p (X_p + Y_p) \varepsilon D_N \right\} \geq a > 0$$

for some constant a and for all large N. Put $h =$ [redacted]$] + 1$, where $[x]$ denotes the largest integer less than or equal to x. For all sufficiently large N, we have,

$$P\left\{ \left| \sum_{p > N} a_p (X_p + Y_p) \right| > N^{-c/3} \right\}$$

$$\leq P\left\{ \sum_{p > N} |a_p (X_p + Y_p)| > N^{-c/3} \text{ and for all } p > N, |X_p + Y_p| < h+2 \right\} + O\left(\sum_{p > N} p^{-h-2} \right)$$

$$= O\left(\sum_{p > N} \sum_{t=1}^{h+1} t^2 a_p^2 k^t p^{-t} \right) N^{2c/3} + O\left(\sum_{p > N} p^{-2} \right) N^{-h}$$

$$= O(N^{-c/3}) < b\, N^{-c/3} \quad \text{for some } b > 0.$$

So

$$P\left\{ \sum_p a_p (X_p + Y_p) \varepsilon G_N \right\} \geq a - b\, N^{-c/3} > \frac{a}{2} > 0$$

for all sufficiently large N, where

$$G_N = \bigcup_{d \varepsilon D_N} [d - N^{-c/3}, \; d + N^{-c/3}].$$

By Lemma 5.3, the Lebesgue measure of the set G_N is not more than

$$2\, N^{-c/3} \sum_{m < N^{c/6}} 2^{\omega(m)}$$

and so tends to zero as $N \to \infty$. Hence $L(\sum_p a_p$ [illegible] $+ Y_p))$ is singular. This completes the proof of the lemma.

<u>Lemma 5.7</u> : Suppose h is a characteristic function of an infinitely divisible distribution with discrete Levy functions M and N. If the total variations of M and N are finite and h has no Normal factor, then the distribution corresponding to h is discrete.

See Lukacs (1970, p.124).

4. Proofs of the main results :

<u>Proof of Theorem 5.1</u> : By (5.1.1) and Lemma 5.3, we can find a sequence of independent random variables $\{X_p\}$ satisfying (5.3.1) and the distribution of $f(m) - f(m+1)$ is $L(\sum_p X_p)$. In view of Lemma 5.2, we can assume that f is strongly additive and $f(p) = 0$ if $p \in A$.

Let $\{Y_p, Z_q : p, q \text{ primes}\}$ be a set of independent random variables. Let $\{\xi_p\}$ be another sequence of independent random variables defined on the same probability space on which $\{Y_p, Z_q : p, q \text{ primes}\}$ are defined satisfying the following properties :

Distribution of ξ_p is same as that of X_p

$$P(Z_p = -k) = P(Y_p = k) = (1 - \frac{1}{p}) \frac{1}{p^k}, \quad k = 0, 1, \ldots .$$

If $f(p) \neq 0$ then, for $x = 0$, $f(p)$ and $-f(p)$ we have

$$P\{\xi_p = x \text{ and } \xi_p \neq f(p)(Y_p + Z_p)\} = O(p^{-2}).$$

It is not difficult to show the existence of such sequences of random variables, since for any non-negative integer k

$$P\{Y_p + Z_p = -k\} = P\{Y_p + Z_p = k\}$$

$$= \sum_{t=0}^{\infty} P\{Y_p = t + k\} P\{Z_p = -t\}$$

$$= (1 - \frac{1}{p})^2 \frac{1}{p^k} \sum_{t=0}^{\infty} \frac{1}{p^{2t}} = \frac{1}{p^k} e^{-2/p} [1 + O(p^{-2})] > 0.$$

It follows, therefore that

$$P\{\xi_p \neq f(p)(Y_p + Z_p)\} = O(p^{-2}).$$

Since $L(\sum_p X_p) = L(\sum_p \xi_p)$, it follows that $L(\sum_p X_p)$ is singular if and only if, $L(\sum_p f(p)(Y_p + Z_p))$ is singular.

Now the characteristic function h of $L(\sum_p f(p)(Y_p + Z_p))$ is given by

$$h(t) = \left|\exp\left\{\sum_p \sum_{k=1}^{\infty} [e^{it\,k\,f(p)} - 1 - it\,k\,f(p)] \frac{1}{kp^k}\right\}\right|^2$$

$$= \left|\exp\left\{\sum_p \frac{1}{p} [e^{it\,f(p)} - 1 - it\,f(p)]\right\}\right|^2 g(t),$$

where

$$g(t) = |\exp\{\sum_p \sum_{k=2}^{\infty} \frac{1}{kp^k}[e^{it\,k\,f(p)}-1-it\,k\,f(p)]\}|^2.$$

Since $\sum_p \sum_{k=2}^{\infty} \frac{1}{kp^k} < \infty$, g is the characteristic function of a discrete infinitely divisible distribution, by Lemma 5.7.

Since the distribution corresponding to the characteristic function $|\exp\{\sum_p \frac{1}{p}[e^{it\,f(p)}-1-it\,f(p)]\}|^2$ is singular by Lemma 5.6 and since g is the characteristic function of a discrete distribution, it follows, by Lemma 5.1, that the distribution corresponding to the characteristic function h is singular. Hence the distribution of $f(m) - f(m+1)$ is singular. This completes the proof of Theorem 5.1.

Proof of Theorem 5.2 : Let $\{\eta_p\}$ be a sequence of independent random variables with

$$P\{\eta_p = x\} = (1 - \frac{1}{p}) \sum_{\substack{t \\ f(p^t)=x}} \frac{1}{p^t}.$$

It is easy to see by Erdos-Wintner theorem [Kubilius, 1964] that, if f has a distribution, then $\sum_p \eta_p$ converges almost everywhere and the distribution of f coincides with $L(\sum_p \eta_p)$. Let $\{Y_p, Z_p : p, q \text{ primes}\}$ be a set of independent random variables defined on the same probability space on which $\{\eta_p\}$ are defined and satisfy the following conditions :

$$P\{f(p)\, Y_p \neq \eta_p\} = O(p^{-2})$$

$$P\{Y_p = k\} = P\{Z_p = -k\} = \frac{1}{p^k}(1 - \frac{1}{p}),\ k = 0,1,\ldots .$$

Since $L(\sum_p \eta_p)$ is absolutely continuous, by Lemma 5.2, it follows that $L(\sum_p f(p)Y_p)$ is absolutely continuous. Consequently $L(\sum_p f(p)(Y_p+Z_p))$ is absolutely continuous. Again by Lemma 5.2 and from the proof of Theorem 5.1, it follows that the distribution of $f(m) - f(m+1)$ is absolutely continuous. This completes the proof of Theorem 5.2.

Corollary 5.1, now follows easily from Theorems 5.1 and 5.2.

<u>Proof of Corollary 5.2</u> : Since f is bounded, $\sum_p f(p)$ converges absolutely and hence $|f(p)| < 1$ for all sufficiently large p. So for N sufficiently large, we have

$$\sum_{p>N} \frac{f(p)^2}{p} \leq (\sum_{p>N} f(p)^4)^{\frac{1}{2}} (\sum_{p>N} \frac{1}{p^2}) = O((\sum_{p>N} \frac{1}{p^{1+\frac{1}{2}}})^{\frac{1}{2}}).\ N^{-\frac{1}{4}}$$

$$= O(N^{-\frac{1}{4}}).$$

The result now follows from Corollary 5.1.

<u>Proof of Theorem 5.3</u> : If a real-valued additive arithmetic function g satisfies (5.2.1) then by a result of Wirsing(1970) there exists a constant D and a bounded real-valued additive arithmetic function f such that

$$g(m) = D \log m + f(m) \quad \text{for} \quad m=1,2,\ldots$$

Since f is bounded it satisfies the condition (5.1.1). So the distribution of $f(m) - f(m+1)$ is singular. But

$$g(m) - g(m+1) = f(m) - f(m+1) + o(1) \quad \text{as} \quad m \to \infty .$$

Hence the distribution of $g(m) - g(m+1)$ is same as that of $f(m) - f(m+1)$. Consequently the distribution of $g(m)-g(m+1)$ is singular.

<u>Proof of Theorem 5.4</u> : Define a function g by

$$\log g(t) = \sum_p \frac{1}{p} [e^{it\, f'(p)} - 1 - it\, f'(p)].$$

Clearly g is the characteristic function of an infinitely divisible distribution. As in the proof of Theorem 5.1 it is sufficient to show that the distribution corresponding to g is absolutely continuous. Since

$$2|\sin y| \geq |y| \quad \text{if} \quad 2|y| \leq 1,$$

we have for any $\varepsilon > 0$,

$$\frac{1}{|\log 2\varepsilon|} \sum_p \frac{1}{p} \left(\sin \frac{1}{2\varepsilon} f'(p)\right)^2 \geq \frac{1}{16\varepsilon^2 |\log 2\varepsilon|} \Big[\sum_{|f'(p)|<\varepsilon} \frac{1}{p} (f'(p))^2 \Big].$$

So, by (5.2.2) we have

$$\liminf_{\varepsilon \to 0} \frac{1}{|\log \varepsilon|} \sum_p \frac{1}{p} \left(\sin \frac{1}{\varepsilon} f'(p)\right)^2 > \frac{1}{4}.$$

We have

$$\log |g(2u)|^2 = \sum_p \frac{1}{p} \left[e^{i2uf'(p)} + e^{-i2uf'(p)} - 2\right]$$

$$= \sum_p \frac{2}{p} \left[\cos 2uf'(p) - 1\right]$$

$$= -4 \sum_p \frac{1}{p} [\sin uf'(p)]^2.$$

So,

$$\liminf_{u \to \infty} \left\{ -\frac{1}{2} [\log |g(2u)|] / |\log u| \right\} > \frac{1}{4}$$

i.e. $$\liminf_{u \to \infty} \left\{ -[\log |g(u)|] / |\log u| \right\} > \frac{1}{2}.$$

Hence for some $\delta > 0$,

$$|g(u)| = O(|u|^{-\frac{1}{2} + \delta}) \quad \text{as } |u| \to \infty.$$

So g is square-integrable, consequently, by Plancherel's theorem it follows that the distribution corresponding to g

is absolutely continuous. This completes the proof of Theorem 5.4.

Proof of Theorem 5.5 is similar to the proof of Theorem 5.4 and so is omitted.

Proof of Theorem 5.6 : Define

$$a_p = \begin{cases} f(p) & \text{if } p \notin A,\ r(F,p) \neq 0 \text{ and } |f(p)| < 1 \\ 0 & \text{otherwise.} \end{cases}$$

Let $r(F,p) \leq k$ for all p. In view of Theorem 1.1 and Lemma 5.2, it follows that the distribution of $f(F(m))$, if it exists, is singular if and only if the distribution corresponding to the characteristic function

$$h(t) = \exp\left\{ \sum_p [e^{it\, a_p} - 1 - it\, a_p] \frac{r(F,p)}{p} \right\}$$

is singular. Define

$$s(t) = \exp\left\{ \sum_{\substack{p \\ r(F,p)<k}} [e^{it\, a_p} - 1 - it\, a_p][(k-r(F,p))/p] \right\}.$$

Since, $\sum_p \frac{1}{p} a_p^2$ is finite, $s(t)$ defines a characteristic function and $|g(t)|^{2k} = |h(t)\, s(t)|^2$, where $g(t)$ is defined by (5.3.3).

Clearly (5.2.3) implies (5.3.2). So the distribution corresponding to $|g(t)|^{2k}$ is singular by Lemma 5.6 and hence the distribution corresponding to h is singular. This completes the proof of Theorem 5.6.

Proof of Theorem 5.7 : Under the hypotheses, it is not difficult to show, from the proof of Theorem 1.1 and from Lemma 5.2, that the distribution of $f(F(m)) - f(F(m+1))$ exists and is singular if and only if the distribution corresponding to the characteristic function

$$g^*(t) = \exp\{2 \sum_p [e^{it\, a_p} + e^{-it\, a_p} - 2][r^*(F,p)/p]\}$$

is singular, where a_p is as defined in the proof of Theorem 5.6 and $r^*(F,p) \leq r(F,p)$. [$r^*(F,p) < r(F,p)$ for some p if there exist two factors P(m) and Q(m) of F(m) such that $Q(m) \equiv P(m+1)$ for all m.]

From Lemma 5.6 and from the proof of Theorem 5.6 it follows easily that the distribution corresponding to g^* is singular. This completes the proof of Theorem 5.7.

We need the following lemma to prove Theorem 5.8.

Lemma 5.8 : Suppose $0 \leq s(p) < c$ and $\{a_p\}$ is a sequence of real numbers. Then one can find a sequence of independent random variables $\{Y_p : p > 2c\}$ defined on a complete probability space such that, if $a_p \neq 0$

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

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

and if $a_p = 0$

$$P\{ Y_p = 0\} = 1.$$

Also one can find another sequence of independent random variables $\{X_p : p > 2c\}$ defined on the same probability space satisfying

$$P\{X_p = 0\} = 1 - \frac{s(p)}{p}$$

$$P\{X_p = a_p\} = \frac{s(p)}{p}$$

whenever $a_p \neq 0$,

$$P\{X_p = 0\} = 1 \quad \text{if} \quad a_p = 0,$$

and

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

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

<u>Proof of Theorem 5.3</u> : By Lemma 1.4 there exists a constant c such that $r(F, p^k) < c$ for all p and k and

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

Let $\{X_p\}$ be a sequence of independent random variables such that

$$P\{X_p = 0\} = 1 \quad \text{if} \quad f(p) = 0$$

and if $f(p) \neq 0$,

$$P\{X_p = nf(p)\} = \frac{1}{p^n}\left(1 - \frac{1}{p}\right), \ n = 0,1,2,\ldots$$

By Lemmas 5.2 and 5.8 and the results of Chapter 1, if f has an absolutely continuous distribution, it follows that $\sum_p X_p$ converges almost everywhere and its distribution is absolutely continuous. If h is the characteristic function of $\sum_{p>2c} X_p$, then clearly

$$\log h(t) = i\sigma t + \sum_{p>2c} \sum_{k=1}^{\infty} \frac{1}{kp^k}\left[e^{it\,k\,f(p)} - 1 - \frac{it\,k\,f(p)}{1+k^2f^2(p)}\right]$$

for some real number σ. Since

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

by Lemma 5.7 we get that the distribution corresponding to the characteristic function

$$\phi(t) = \exp\left\{ \sum_{\substack{p \notin A \\ p > 2c}} \frac{1}{p} \left[e^{it\, f(p)} - 1 - \frac{it\, f(p)}{1+(f(p))^2} \right] \right\}$$

is absolutely continuous.

Since $f(F(m))$ has a distribution, by Lemmas 5.2, 5.7 and 5.8, as above, we conclude that the distribution of $f(F(m))$ is absolutely continuous if the distribution corresponding to the characteristic function $\underline{g}$ given by

$$\underline{g}(t) = \exp\left\{ \sum_{\substack{p > 2c \\ p \notin A}} \frac{r(F,p)}{p} \left[e^{it\, f(p)} - 1 - \frac{it\, f(p)}{1+(f(p))^2} \right] \right\}$$

is absolutely continuous. Since $\sum_{\substack{p>2c \\ p \notin A}} \frac{1}{p}\left[e^{it\, f(p)} - 1 - \frac{it\, f(p)}{1+(f(p))^2}\right]$

and $\sum_{\substack{p>2c \\ p \notin A}} \frac{r(F,p)}{p} \left[e^{it\, f(p)} - 1 - \frac{it\, f(p)}{1+(f(p))^2}\right]$ converge

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

$$\sum_{\substack{p>2c \\ p \notin A}} \frac{(r(F,p)-1)}{p} \left[e^{it\, f(p)} - 1 - \frac{it\, f(p)}{1+(f(p))^2}\right]$$

converges absolutely and uniformly in every compact interval of the real line. Since $r(F,p) \geq 1$ or $f(p) = 0$ if $p \notin A$, it follows that Λ defined by

$$\Lambda(t) = \exp\left\{ \sum_{\substack{p > 2c \\ p \notin A}} \frac{(r(F,p)-1)}{p} \left[e^{it\, f(p)} - 1 - \frac{it\, f(p)}{1+(f(p))^2} \right] \right\}$$

is a characteristic function. We note that $\underline{g}(t) = \phi(t).\Lambda(t)$.

Since ϕ is the characteristic function of an absolutely continuous distribution, $\underline{g}$ is also the characteristic function of an absolutely continuous distribution. This completes the proof of Theorem 5.8.

Proof of Theorem 5.9 is similar to that of Theorem 5.8 and so is omitted.

<u>Example 5.1</u> : Let f be the strongly additive function defined by

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

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.

<u>Lemma 5.9</u> : If $p \equiv 1 \pmod 4$, the congruence

(5.4.1) $$x^2 \equiv -1 \pmod p$$

has exactly two incongruent solutions. The congruence (5.4.1) has no solution when $p \equiv 3 \pmod 4$.

For a proof see Nagell (1951, p.99, Theorem 5.8).

Now we shall show that the distribution of f exists and is absolutely continuous. The characteristic function of the distribution of f is given by

$$h(u) = \prod_p [1 - \frac{1}{p}[1 - e^{i\,u\,f(p)}]].$$

Now as in Erdos (1939) for $u \neq 0$,

(5.4.2) $$|h(u)| \leq \prod{}' |1 - \frac{1}{p}[1 - \exp(iu(\log\log p)^{-3/2})]|$$

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

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

Then each factor of the product on the right side of (5.4.2) is not more than $1 - \frac{1}{p}$, so that

$$|h(u)| \leq \prod{}'(1 - \frac{1}{p}).$$

Hence $|h(u)| = O(\exp(-\Sigma' \frac{1}{p}))$, where, for each fixed $u \neq 0$, Σ' denotes the sum over those primes which satisfy (5.4.3). By Lemmas 5.9 and 1.4 we get that

$$\sum_{\substack{p \equiv 3 \pmod 4 \\ p \leq x}} \frac{2}{p} = \log\log x + O(1).$$

Hence

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

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 h is integrable and hence it is the characteristic function of an absolutely continuous distribution.

This shows that f has an absolutely continuous distribution, but the distribution of $f(F(m))$ is degenerate when $F(m) = m^2 + 1$.

<u>Remarks</u> : (5.2.4) is satisfied if F has a linear factor. The condition (5.2.4) cannot be omitted in Theorems 5.8 and 5.9, since if (5.2.4) is violated then Example 5.1 shows that the distributions of both f and $f(m) - f(m+1)$ are absolutely continuous, but $f(F(m)) = f(F(m)) - f(F(m+1)) = 0$ for all $m \geq 1$. On the other hand there exist additive arithmetic functions such that both f and $f(F(m))$ have absolutely

continuous distributions even if the condition (5.2.4) is violated. In fact, if f is the strongly additive arithmetic function defined by

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

Let $F(m)$ be any non-constant polynomial taking positive integral values for $m \geq 1$. From Theorem 1.1, we conclude that $f(F(m))$ has a distribution. By using an argument similar to the argument used in Example 5.1, it is not difficult to conclude that the distribution of $f(F(m))$ is absolutely continuous.

CHAPTER 6

PROBABILITIES OF DEVIATIONS OF ADDITIVE ARITHMETIC FUNCTIONS

1. Introduction : In this chapter we first obtain an asymptotic formula for

$$(6.1.1) \qquad \frac{1}{n} N_n \{f(m) > x_n\},$$

for a wide class of additive arithmetic functions, when $x_n \to \infty$ at a certain rate.

Next, we consider a sequence $\{X_n\}$ of random variables with finite means and variances. Let

$$F_n(x) = P\{X_n < x\}.$$

Suppose that the moment generating functions of all F_n exist in a non-degenerate interval. We obtain, under some conditions, an asymptotic formula for $F_n(x_n)$, when $x_n \to \infty$ at a certain rate. From this we deduce that, for any δ with $0 < \delta < 2$,

$$\frac{1}{\log\log x} \log[(1/n) N_n\{\omega(m) > e^x \log\log n\}] = (e^x - 1) - xe^x + o(1)$$

as $n \to \infty$, uniformly in $x \in (0, 2-\delta)$. This gives a fairly good estimate of (6.1.1) when $f(m) = \omega(m)$. Kubilius (1964) obtained a similar result for this case when $x_n = (1+o(1))\log\log n$. We note that our sequence $\{x_n\}$ increases at a rate much faster than that considered by Kubilius.

2. Notations and Definitions : Throughout this chapter f denotes a real-valued additive arithmetic function and $F \in \underline{\underline{P}}$. The following notations are used only in this chapter.

$$A_n = \sum_{p \le n} \frac{f(p)}{p}$$

$$D_n = [\sum_{p^t \le n} \frac{1}{p^t} [f(p^t)]^2]^{\frac{1}{2}}$$

$$D(n,F) = [\sum_{p^t \le n} \frac{1}{p^t} [f(p^t)]^2 \, r(F,p^t)]^{\frac{1}{2}}$$

$$r = r(n) = n^{1/D_n}$$

$$s_p = [\frac{\log r}{\log p}] \quad \text{(i.e., integral part of } \frac{\log r}{\log p})$$

$$\pi(p^t) = \begin{cases} (1-p^{-1})p^{-t} & \text{if } 0 \le t \le s_p \\ p^{-t} & \text{if } t = s_p \end{cases}$$

$$t(p,m) = \begin{cases} k & \text{if } 0 \le k \le s_p \text{ and } p^k \| m \\ 0 & \text{otherwise} \end{cases}$$

$$f_r(m) = \sum_{p \le r} f(p^{t(p,m)})$$

$$\Phi(x) = (2\pi)^{-1/2} \int_{-\infty}^{x} e^{-y^2/2} \, dy$$

η is any positive number less than one.

3. Main results :

Theorem 6.1 : If $\max_{p^t \leq n} |f(p^t)| = O(D_n^{3/4})$ and $D_n \to \infty$ as $n \to \infty$, then we have as $n \to \infty$

$$(1/n)\, N_n \{f(m) - A_n > c\, D_n \sqrt{\log D_n}\}$$

$$= (1+o(1))\, D_n^{-c^2/2} / (c \sqrt{2\pi \log D_n}),$$

where the estimate $o(1)$ is uniform in $c \in (o, \eta)$.

Theorem 6.2 : Let $D_F \geq 2$. If $D(n,F) \to \infty$ and

$$\max_{p^t \leq n^{D_F}} |f(p^t)|\, r(F,p^t) = O((D(n,F))^{3/4})$$

as $n \to \infty$, then we have as $n \to \infty$

$$(1/n)\, N_n \{f(F(m)) - A(n,f,F) > c\, D(n,F) \sqrt{\log D(n,F)}\}$$

$$= (1+o(1)) / ((D(n,F))^{c^2/2}\, c \sqrt{2\pi \log D(n,F)}),$$

where the estimate $o(1)$ is uniform in $c \in (o,\eta)$.

For each n, let X_n be a random variable with mean μ_n and variance σ_n^2. Let

$$F_n(x) = P\{X_n < x\}.$$

Suppose for each n,

$$R_n(z) = \int e^{zx}\, d\, F_n(x)$$

is analytic and does not vanish in the region $|z| < c_1$, for some $c_1 > 0$. Now $K_n(z) = \log R_n(z)$ is analytic in $|z| < c_2 < c_1$. Let $K_n^{(s)}$ denote the s-th derivative of K_n.

Theorem 6.3 : Suppose $\sigma_n \to \infty$ as $n \to \infty$ and $|K_n^{(2)}(z)| = O(\sigma_n^2)$ for all z such that $|z| < c_2$. Then there exists a constant $c_3 \in (0, c_2)$ such that $0 < x < c_3$ implies

$$\frac{1}{\sigma_n^2} \log P\{X_n \geq K_n^{(1)}(x)\} = \frac{K_n(x) - x K_n^{(1)}(x)}{K_n^{(2)}(0)} + o(1)$$

as $n \to \infty$ uniformly in x.

Remark : Chernoff's Theorem (see Bahadur and Ranga Rao, 1960) follows immediately from Theorem 6.3. Results similar to this are also obtained by Feller (1969).

Corollary 6.1 : Let $\omega(m)$ denote the number of distinct prime factors of m. Let $0 < \delta < 2$. Then for $0 < x < 2 - \delta$,

$$\frac{1}{\log\log n} \log[\frac{1}{n} N_n \{\omega(m) > e^x \log\log n\}] = (e^x - 1) - x\, e^x + o(1)$$

as $n \to \infty$, uniformly in x.

The corollary follows from the estimates (9.19),(9.20) and (9.21) of Kubilius (1964).

Remark : Kubilius (1964) obtained asymptotic formula for $(1/n)\ N_n \{\omega(m) < \log\log n + x_n \sqrt{\log\log n}\}$ when $x_n \geq 0$ for all n and $x_n = o(\sqrt{\log\log n})$. Our Corollary 6.1 gives a good estimate for the above, even when $x_n \to \infty$ at a much faster rate.

As Kubilius remarked in his book, proof of Theorem 9.2 of Kubilius (1964) can be adapted for many additive arithmetic functions other than $\omega(m)$, for which the proof is given in his book. Here we remark that Kubilius' proof can be adapted to a sequence of random variables to give the following result.

Theorem 6.4 : Suppose that $\sigma_n \to \infty$ as $n \to \infty$ and that $K_n^{(2)}(z) = O(\sigma_n^2)$ for all z in $\{z : |z| < c_2\}$. If $x_n = o(\sigma_n)$, then

$$P\{X_n < \mu_n + x_n \sigma_n\} = e^{Q_n(x_n)}\ \Phi(-|x_n|(1 + O(\frac{1+|x_n|}{\sigma_n}))$$

for $x_n \leq 0$ and

$$P\{X_n > \mu_n + x_n \sigma_n\} = e^{Q_n(x_n)}\ \Phi(-|x_n|)(1 + O(\frac{1+|x_n|}{\sigma_n}))$$

for $x_n \geq 0$, where $Q_n(x) = \frac{x^2}{2} + a_{1,n}\ x^3 + \ldots$ is a power series in x such that $Q_n(x_n)\,\sigma_n^{-2} \to 0$ as $n \to \infty$.

Remark : It is difficult to compare Theorems 6.3 and 6.4 as the sequence considered in Theorem 6.3 goes to infinity at a rate faster than that in Theorem 6.4. But on the other hand the estimates given in Theorem 6.4 are much sharper than the estimate given in Theorem 6.3.

4. Preliminary results : For simplicity in writing we let $r = r(n)$.

Lemma 6.1 : Let, for each r, $\{Y_{p,r} : p \leq r\}$ be a sequence of independent random variables with probability distributions defined by

$$P\{ Y_{p,r} = x \} = \sum_{f(p^t)=x} \pi(p^t).$$

Then there exists a positive constant a such that

$$\frac{1}{n} N_n \{f_r(m) \varepsilon A \} - P\{ \sum_{p\leq r} Y_{p,r} \varepsilon A\} = O(\exp(-a \log n/\log r))$$

uniformly for all Borel subsets A of the real line.

For a proof, see § 3 of Chapter 2 or Chapter II of Kubilius (1964).

Lemma 6.2 : Let $\{Y_{p,r}\}$ be as in Lemma 6.1. Then

$$P\left\{\sum_{p\leq r} Y_{p,r} - A_r > c\, D_r \sqrt{\log D_r}\right\} = [1+o(1)]\ \frac{D_r^{-c^2/2}}{c\sqrt{2\pi \log D_r}}$$

uniformly in $c \in (0, \eta)$.

Proof : The proof of this lemma is similar to the proof of Theorem 1 of Rubin and Sethuraman (1965). Define

$$G_{p,r}(x) = P\{Y_{p,r} \leq x\}$$

$$a_r = \sqrt{\log D_r} / D_r$$

$$g_{p,r}(c) = \int_{-\infty}^{\infty} \exp(c\, a_r\, y)\ dG_{p,r}(y)$$

$$G_{p,r}(x, c)\ g_{p,r}(c) = \int_{-\infty}^{x} \exp(c\, a_r\, y)\ dG_{p,r}(y)$$

$$m_{p,r}(c) = \int_{-\infty}^{\infty} x\ dG_{p,r}(x, c)$$

$$s^2_{p,r}(c) = \int_{-\infty}^{\infty} x^2\, dG_{p,r}(x, c) - m^2_{p,r}(c)$$

$$b_{p,r}(c) = \int_{-\infty}^{\infty} |x|^3 \exp(c\, a_r\, |x|)\ dG_{p,r}(x)$$

$$h_r(c) = \left[\sum_{p\leq r} (b_{p,r}(c) / g_{p,r}(c))\right] / \left(\sum_{p\leq r} s^2_{p,r}(c)\right)^{3/2}$$

$$G_r(x) = \left[\ \mathop{*}_{p\leq r}\ G_{p,r}(\cdot)\right](x)$$

$$G_r(x,c) = [\underset{p\le r}{*} G_{p,r}(\cdot,c)](x)$$

$$\phi_r(x,c) = G_r\Big(\sum_{p\le r} m_{p,r}(c) + \big(x \sum_{p\le r} s^2_{p,r}(c)\big)^{1/2}, c\Big)$$

Here $*$ denotes the convolution. Thus

$$P\Big\{\sum_{p\le r} Y_{p,r} - A_r > c\, D_r \sqrt{\log D_r}\Big\} = 1 - G_r(c\, D_r \sqrt{\log D_r} + A_r)$$

$$= \prod_{p\le r} g_{p,r}(c) \int_{c D_r \sqrt{\log D_r}+A_r}^{\infty} \exp(-c\, a_r\, y)\, dG_r(y,c)$$

$$= A_r(c) \int_{B_r(c)}^{\infty} \exp(-C_r(c)\, x)\, d\phi_r(x,c)$$

where

$$A_r(c) = \prod_{p\le r} [g_{p,r}(c) \exp(-c\, a_r\, m_{p,r}(c))],$$

$$B_r(c) = [A_r + c\, D_r \sqrt{\log D_r} - \sum_{p\le r} m_{p,r}(c)]/[\sum_{p\le r} s^2_{p,r}(c)]^{1/2},$$

$$C_r(c) = c\, a_r \Big(\sum_{p\le r} s^2_{p,r}(c)\Big)^{1/2}.$$

Now, since $\max_{p^t\le n} |f(p^t)| = O(D_n^{3/4})$,

$$a_r^3\Big(\sum_{p\le r} b_{p,r}(c)\Big) \le ((\log D_r)^{3/2} \Big(\sum_{p\le r} \frac{1}{p^t} |f(p^t)|^3\Big)/D_r^3)$$

$$= O((\log D_r)^{3/2}/D_r^{1/4})$$

$$g_{p,r}(c) = 1+c\, a_r\, m_{p,r}(0) + \frac{c^2}{2}\, a_r^2\, s_{p,r}^2(0) + O(b_{p,r}(c)\, a_r^3)$$

$$= 1 + o(1) \quad \text{uniformly in} \quad p \leq r,$$

$$m_{p,r}(c)\, g_{p,r}(c) = m_{p,r}(0) + c\, a_r\, s_{p,r}^2(0) + \frac{c^2}{2} O(a_r^2\, b_{p,r}(c)),$$

and

$$(s_{p,r}^2(c) + m_{p,r}^2(c))\, g_{p,r}(c) = s_{p,r}^2(0) + O(c\, a_r\, b_{p,r}(c)).$$

By simple calculations, we get

$$\log A_r(c) = -\frac{c^2}{2} \log D_r + o(1)$$

$$B_r(c) = o((\log D_r)^{-1/2}$$

$$C_r(c) = c\sqrt{\log D_r}\,(1 + o(1))$$

$$h_r(c) = o((\log D_r)^{-3/2}).$$

By Berry-Esseen approximation theorem

$$|\phi_r(x,c) - G(x)| \leq K\, h_r(c).$$

Now integrating by parts and using the above Berry-Esseen theorem, we get

$$\int_{B_r(c)}^{\infty} \exp(-x\, C_r(c))\; d\phi_r(x,c)$$

$$= \phi_r(B_r(c),c)\, \exp(-C_r(c)\, B_r(c))$$

$$+ \int_{B_r(c)}^{\infty} C_r(c)\, \phi_r(x,c)\, \exp(-C_r(c)\, x)\, dx$$

$$= e^{-C_r(c)\, B_r(c)}\, G_r(B_r(c), c) + \int_{B_r(c)}^{\infty} C_r(c) e^{-C_r(c)x}\, dG_r(x, c)$$

$$+ O((\log D_r)^{-3/2})$$

$$= \frac{1+o(1)}{c\sqrt{2\pi \log D_r}} + O((\log D_r)^{-3/2})$$

Thus

$$P\{\Sigma Y_{p,r} - A_r > c\, D_r \sqrt{\log D_r}\} = \frac{(1+o(1))}{c\sqrt{2\pi \log D_r}}\, D_r^{-c^2/2}.$$

This completes the proof of the lemma.

5. Proofs of the main results :

Proof of Theorem 6.1 : Since

$$D_n^2 - D_r^2 = \sum_{r<p^t\leq n} \frac{1}{p^t} (f(p^t))^2 = O(D_n^{3/2} \log D_n),$$

we have by Lemma 6.2,

$$P\{\sum_{p\leq r} Y_{p,r} - A_r > c\, D_n \sqrt{\log D_n}\} = \frac{D_n^{-c^2/2}}{c\sqrt{2\pi \log D_n}} (1+o(1))$$

uniformly in $c \in (0, \eta)$. Since $D_n = \log n / \log r$, by Lemma 6.1, we have

(6.5.1) $$N_n\{f_r(m) - A_r > c\, D_n \sqrt{\log D_n}\} = \frac{n\, D_n^{-c^2/2}}{c\sqrt{2\pi \log D_n}} (1+o(1))$$

uniformly in $c \in (0, \eta)$. Let

$$Z(n,c) = N_r\{f(m)-A_n > c\,D_n\sqrt{\log D_n} \text{ and } |f(m)-A_n-f_r(m)+A_r| < \frac{D_n}{\log D_n}\}.$$

By Turan-Kubilius inequality, we have

$$|N_n\{f(m) - A_n > c\,D_n\sqrt{\log D_n}\} - Z(n,c)|$$

$$\leq N_n\{|f(m) - f_r(m)-A_n+A_r| \geq (D_n/\log D_n)\}$$

(6.5.2) $$= O(n(\log D_n)^2\,(D_n^2 - D_r^2)/D_n^2)$$

$$= O(n\,D_n^{-1/2}\,(\log D_n)^3)$$

uniformly in $c \in (o,\eta)$. Clearly

$$N_n\{f_r(m) - A_r > c\,D_n\sqrt{\log D_n} + (D_n/\log D_n)\}$$

(6.5.3) $$\leq Z(n, c)$$

$$\leq N_n\{f_r(m) - A_r > c\,D_n\sqrt{\log D_n} - (D_n/\log D_n)\}$$

By (6.5.1) we have

$$N_n\{f_r(m) - A_r > c\,D_n\sqrt{\log D_n} - (D_n/\log D_n)\}$$

$$= \frac{n\,\exp(-\frac{1}{2}[c-(\log D_n)^{3/2}]^2\,\log D_n)}{(c - \log D_n)^{3/2}\sqrt{2\pi\,\log D_n}}\,(1+o(1))$$

$$= \frac{n\,D_n^{-c^2/2}}{c\sqrt{2\pi\,\log D_n}}\,(1+o(1)).$$

Similarly one can show that (6.5.3) is equal to

$$\frac{n\, D_n^{-c^2/2}}{c\sqrt{2\pi \log D_n}}\,[1 + o(1)]$$

Hence

$$(6.5.4) \qquad Z(n, c) = \frac{n\, D_n^{-c^2/2}}{c\sqrt{2\pi \log D_n}}\,[1 + o(1)].$$

Now the theorem follows from (6.5.2) and (6.5.4).

Proof of Theorem 6.2 is similar to that of Theorem 6.1 and so is omitted.

<u>Proof of Theorem 6.3</u> : Since $K_n^{(2)}(z) = O(\sigma_n^2)$ in $|z| < c_2$, by Cauchy's integral formula, (see Mackey, 1967) we have if $c_4 < c_2$,

$$(6.5.5) \qquad |K_r^{(t)}(z)| \le c_5 \frac{(t-2)!}{c_4^{t-1}}\, 2^{t-1}\, \sigma_n^2$$

for $|z| < (c_4/2)$ and $t = 3, 4, \ldots,$ where c_5 is a constant. If c_3 is sufficiently small, then from (6.5.5) it follows that

$$\sigma_n^2 \le c_6\, |K_n^{(t)}(x)|$$

for all $|x| < c_3$, where c_6 is a constant.

For $x \in (0, c_3)$, define

$$G_n(u) = \frac{1}{R_n(x)} \int_{-\infty}^{u} e^{xy} \, dF_n(y).$$

Put

$$H_n(u) = G_n(\sqrt{K_n^{(2)}(x)} \; u + K_n^{(1)}(x))$$

$$p_n = P\{X_n > K_n^{(1)}(x)\}.$$

Then

$$p_n = R_n(x) \exp(-x K_n^{(1)}(x)) \int_{z>0} e^{-xz} \; H_n(\frac{dz}{\sigma_n})$$

$$= R_n(x) \exp(-x K_n^{(1)}(x)) \int_{y>0} e^{-x\sigma_n y} H_n(dy)$$

where $\sigma_n = \sqrt{K_n^{(2)}(x)}$

$$= R_n(x) \exp(-x K_n^{(1)}(x)) x\sigma_n \int_{y>0} e^{-x\sigma_n y} (H_n(y) - H_n(0)) \, dy$$

by integration by parts. Clearly

$$p_n \leq R_n(x) \exp(-x K_n^{(1)}(x)).$$

By (6.5.5) and the fact that $K_n^{(2)}(0) \to \infty$ as $n \to \infty$, it follows that $H_n(y) \to G(y)$ in distribution as $n \to \infty$. Hence for any $\varepsilon > 0$

$$I_n = x\sigma_n \int_{y} e^{-x\sigma_n y}\,[F_n(y) - F_n(0)]\,dy$$

$$\geq [F_n(\varepsilon) - F_n(0)]\; x\sigma_n \int_{\varepsilon}^{\infty} e^{-x\sigma_n y}\,dy$$

$$= [F_n(\varepsilon) - F_n(0)]\; e^{-x\sigma_n \varepsilon}.$$

Hence

$$\liminf_{n \to \infty} \left\{ \frac{1}{\sigma_n} \log I_n \right\} \geq -x\varepsilon .$$

Since $I_n \leq 1$ for every n, and since ε is arbitrary, it follows that $\frac{1}{\sigma_n} \log I_n = o(1)$ uniformly in $x (0 < x < c_3)$. Hence

$$\lim_{n \to \infty} \left\{ \frac{1}{\sigma_n} \log I_n \right\} = 0.$$

Thus

$$\sigma_n^{-2} \log p_n = \frac{K_n(x) - x K_n^{(1)}(x)}{K_n^{(2)}(0)} + o(1)$$

uniformly in x. This completes the proof of Theorem 6.3.

CHAPTER 7

ON THE DISTRIBUTION OF VALUES OF MULTIPLICATIVE FUNCTIONS

1. Introduction : Bakstys (1969) has given necessary and sufficient conditions for a real-valued multiplicative function g to have non-symmetric distribution. Galambos (1971b) has given necessary and sufficient conditions for a real-valued strongly multiplicative function g to have non-degenerate symmetric distribution under the extra assumption that $g(2) \neq -1$. We shall first modify Galambos' proof to find necessary and sufficient conditions for a multiplicative function to have a non-degenerate symmetric distribution without using his extra assumption. Using the above results we give a partial answer to the following interesting question. Suppose that the density of $\{m : g(m) \in I\}$ exists and is positive, where I is a bounded interval not containing zero, then is it true that g has a distribution? Also we obtain necessary and sufficient conditions for $\frac{1}{n} N_n \{a_n g(m) < c\}$ to tend to a distribution function, where $\{a_n\}$ is a sequence of real numbers and g is a multiplicative function.

2. Notations and Definitions :

For any real number x we write

$$x^* = \begin{cases} x & \text{if } |x| < 1, \\ 1 & \text{otherwise.} \end{cases}$$

Definition : An arithmetic function g is called multiplicative if for all m,n

$$g(m.n) = g(m).g(n)$$

whenever $(m,n) = 1$. g is called strongly multiplicative if for all p and $k \geq 1$,

$$g(p^k) = g(p).$$

3. Main results :

Theorem 7.1 : Let g be a real-valued multiplicative function. In order that g has a non-degenerate symmetric distribution, it is necessary and sufficient that there exists a real number $c > 1$ such that each of the series

$$\sum_{|g(p)| \leq 1/c} \frac{1}{p}, \quad \sum_{|g(p)| \geq c} \frac{1}{p},$$

(7.3.1)

$$\sum_{1/c < |g(p)| < c} \frac{1}{p} \log |g(p)|, \quad \sum_{1/c < |g(p)| < c} \frac{1}{p} \log^2 |g(p)|$$

converges and that either $g(2^k) = -1$ for all k or

$$\sum_{g(p) < 0} \frac{1}{p} = \infty .$$

Theorem 7.2 : If g is a real-valued multiplicative function such that either for some $a \neq 0$

$$\limsup_{n \to \infty} \frac{1}{n} N_n \{|g(m)| = a\} > 0,$$

or $|g(m)| \geq 1$ for all $m \geq 1$ and

$$\limsup_{n \to \infty} \frac{1}{n} N_n \{1 \leq |g(m)| \leq c\} > 0$$

for some $c > 1$, then g has a distribution.

Theorem 7.3 : Let g be a real-valued strongly multiplicative function having a distribution H. Let P be the probability measure corresponding to H and let $P\{0\} \neq 1$. Define a countably additive measure on the real line by

$$Q(B) = \begin{cases} P(B) & \text{if } 0 \notin B, \\ P(B) - P\{0\} & \text{if } 0 \in B, \end{cases}$$

for all Borel sets B. Then the measure Q is pure (i.e., either discrete, continuous singular or absolutely continuous with respect to the Lebesgue measure) and Q is absolutely continuous if and only if the distribution of the additive function f defined by

$$f(p) = \begin{cases} \log |g(p)| & \text{if } g(p) \neq 0, \\ 0 & \text{otherwise}, \end{cases}$$

is absolutely continuous.

Theorem 7.4 : Let g be a real-valued multiplicative function. For the existence of a sequence $\{a_n\}$ of real numbers and a non-degenerate distribution function F such that for each continuity point c of H the limit of $(1/n)N_n\{a_n\, g(m) < c\}$ is $H(c)$ as $n \to \infty$, it is necessary and sufficient that there exist a real number b and a multiplicative function h such that

$$g(m) = m^b\, h(m) \tag{7.3.2}$$

for all $m \geq 1$ and

$$\sum_p \frac{1}{p}\, ((|h(p)| - 1)^*)^2 < \infty . \tag{7.3.3}$$

Remark : Theorem 2 of Bakstys (1969) follows from Theorem 7.4 as a corollary.

4. Preliminary results : The tool developed by Zolotarev(1962) for the investigation of the products of independent random variables can effectively be applied to the distribution problem of multiplicative functions (see Bakstys, 1969 and Galambos, 1971b). In what follows we put $0^{it} = 0$. Let G be a distribution function. Define

$$w_0(t) = \int |x|^{it} \, dG(x),$$

$$w_1(t) = \int |x|^{it} \operatorname{sgn}(x) \, dG(x),$$

where

$$\operatorname{sgn}(x) = \begin{cases} -1 & \text{if } x < 0, \\ 0 & \text{if } x = 0, \\ 1 & \text{if } x > 0. \end{cases}$$

The diagonal matrix

$$W(t) = \begin{bmatrix} w_0(t) & 0 \\ 0 & w_1(t) \end{bmatrix}$$

is called the characteristic transform (c.t.) of G.

Lemma 7.1 (Zolotarev, 1962) : If X and Y are independent random variables with c.t. W' and W'' respectively then the c.t. W of the random variable $X.Y$ is $W'.W''$.

Lemma 7.2 : Let X and Y be independent random variables such that $P\{X.Y = 0\} < 1$. Let X be purely discrete. Then the measure Q, defined by

$$Q(B) = \begin{cases} P\{X.Y \in B\} & \text{if } 0 \notin B, \\ P\{X.Y \in B\} - P\{X.Y = 0\} & \text{if } 0 \in B, \end{cases}$$

for all Lebesgue measurable sets B of the real line, is continuous singular, absolutely continuous or discrete according as the measure Q' is continuous singular, absolutely continuous or discrete, where

$$Q'(B) = \begin{cases} P\{Y \varepsilon B\} & \text{if } 0 \not\varepsilon B, \\ P\{Y \varepsilon B\} - P\{Y = 0\} & \text{if } 0 \varepsilon B, \end{cases}$$

for all Lebesgue measurable sets B of the real line.

<u>Proof</u> : Let $\{x_n\}$ be the set of discontinuity points of X. First note that $P\{X \neq 0\} > 0$. If $x \neq 0$, we have

$$P\{X.Y = x\} = \sum_{x_n \neq 0} P\{Y = \frac{x}{x_n}\} P\{X = x_n\} .$$

It follows that Q is continuous if and only if Q' is continuous.

Let B be any Lebesgue null set. If $0 \not\varepsilon B$, then

$$0 = Q(B) = P\{X.Y \varepsilon B\} = \sum_{x_n \neq 0} P\{Y.x_n \varepsilon B\} P\{X = x_n\} .$$

Hence for all $x_n \neq 0$, $P\{Y.x_n \varepsilon B\} = 0$. Thus

$$Q'(B) = P\{Y \varepsilon B\} = 0.$$

because $P\{X \neq 0\} > 0$ and if A is a Lebesgue null set then $|a| A = \{|a|.b : b \in A\}$ is also a Lebesgue null set for any real number a. If $0 \in B$, then on writing $B = (B - \{0\}) \cup \{0\}$ it follows $Q'(B) = 0$. Hence Q' is absolutely continuous. The fact that if Q' is absolutely continuous then Q is also absolutely continuous can be proved similarly.

Lemma 7.3 : Let H_n be a sequence of distribution functions with corresponding c.t. given by

$$\begin{bmatrix} w_{on}(t) & 0 \\ 0 & w_{1n}(t) \end{bmatrix} .$$

For the convergence of H_n to a distribution function H, at all continuity points of the latter, and for $H_n(0) \rightarrow H(0)$ and $H_n(0+) \rightarrow H(0+)$, it is necessary and sufficient that, for all real t, $w_{on}(t)$ and $w_{1n}(t)$ tend to limits, $w_o(t)$ and $w_1(t)$, continuous at $t = 0$. The limits $w_o(t)$ and $w_1(t)$ uniquely determine H and

$$\begin{bmatrix} w_o(t) & 0 \\ 0 & w_1(t) \end{bmatrix}$$

is the c.t. of H. The limit law is non-degenerate and symmetric if and only if $w_1(t) \equiv 0$ and $w_o(t)$ is not identically zero.

See Galambos (1971b).

<u>Lemma 7.4 (Bakstys, 1969)</u>: A real-valued multiplicative function g has non-symmetric distribution if and only if all the series in (7.3.1) converge, $g(2^k) \neq -1$ for some $k \geq 1$ and

$$\sum_{g(p)<0} \frac{1}{p} < \infty .$$

If

$$\sum_{g(p)\neq 1} \frac{1}{p} < \infty$$

then the distribution of g is discrete. Moreover, if g has non-symmetric distribution then, the two infinite products

$$w_o(t) = \prod_p \left(1 - \frac{1}{p}\right) \left(1 + \sum_{k=1}^{\infty} \frac{|g(p^k)|^{it}}{p^k}\right),$$

and

$$w_1(t) = \prod_p \left(1 - \frac{1}{p}\right) \left(1 + \sum_{k=1}^{\infty} \frac{|g(p^k)|^{it} \operatorname{sgn}(g(p^k))}{p^k}\right),$$

converge uniformly and absolutely in every compact interval of the real line and the c.t. of g is

$$\begin{bmatrix} w_o(t) & 0 \\ 0 & w_1(t) \end{bmatrix}$$

Lemma 7.5 : Let h be a multiplicative function such that $|h(n)| \leq 1$. If

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} h(n) = M(h)$$

exists and is not zero, then

$$(7.4.1) \qquad \sum_{p} \frac{1 - h(p)}{p}$$

converges and for some $k \geq 1$, $h(2^k) \neq -1$. Conversely if (7.4.1) converges then $M(h)$ exists and equals

$$(7.4.2) \qquad \prod_{p} \left(1 - \frac{1}{p}\right) \left(1 + \sum_{k=1}^{\infty} \frac{h(p^k)}{p^k}\right).$$

For a proof see Delange (1961).

Lemma 7.6 : Let h be a multiplicative function such that $|h(n)| \leq 1$. Then

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} h(n) = 0$$

if and only if, either for all real u,

$$\sum_{p} \frac{1}{p} \left[1 - \operatorname{Re}\left(h(p)\, p^{-iu}\right)\right] = \infty ,$$

or there is a real s such that

(7.4.3) $$\sum_p \frac{1}{p}[1 - \mathrm{Re}(h(p)\, p^{-is})] < \infty$$

and $2^{-irs}\, h(2^r) = -1$ for all integers $r \geq 1$.

<u>Lemma 7.7</u> : Let f be a real-valued additive arithmetic function. Suppose there exists a $\delta > 0$ such that for each $t \in [-\delta, \delta]$ there exists a $u(t)$ such that

$$\sum_p \frac{1}{p}[1 - \mathrm{Re}(e^{it\, f(p)}\, p^{-iu(t)})] < \infty .$$

Then there exist a real number a and an additive function f_1 such that

$$f(m) = a \log m + f_1(m)$$

for all $m \geq 1$ and

$$\sum_p \frac{(f_1'(p))^2}{p} < \infty .$$

Proof of this lemma is contained in the proof of the theorem of Ryavec (1970).

<u>Lemma 7.8</u> : If h is a real-valued multiplicative function such that

$$\sum_p \frac{1}{p}((|h(p)| - 1)^*)^2 < \infty ,$$

then there exists a continuous function $\ell(h, \ )$ such that

$$\frac{1}{n}\sum_{m=1}^{n}|h(m)|^{it} = \mathcal{C}(h,t)\exp(it\sum_{p\leq n}\frac{1}{p}(|h(p)|-1)^*)+o(1)$$

uniformly in $t \in [-K,K]$, $K > 0$.

This lemma follows easily from Theorems 1 and 2 of Delange (1963).

5. Proofs of the main results :

Proof of Theorem 7.1 : It is well-known that the density of $\{m : g(m) \neq 0\}$ exists ; it is non-zero if and only if

$$\sum_{g(p)=0}\frac{1}{p} < \infty .$$

Suppose each of the series in (7.3.1) converges and $\sum_{g(p)<0}\frac{1}{p} = \infty$. Then as in Galambos (1971 b), it is not hard to show that g has a non-degenerate symmetric distribution.

Suppose each of the series in (7.3.1) is convergent,

$$\sum_{g(p)<0}\frac{1}{p} < \infty$$

and for each $k \geq 1$, $g(2^k) = -1$.

For each real number t, we define

$$h_o(n,t) = |g(n)|^{it}$$

and

$$h_1(n,t) = |g(n)|^{it}\ \mathrm{sgn}(g(n)).$$

By applying Lemma 7.5 to the multiplicative functions $h_o(n,t)$ and $h_1(n,t)$ we get that $M(h_o(n,t))$ is non-zero for all real numbers t and $M(h_1(n,t)) = 0$ for all real numbers t, because $h_o(2,t) = 1$ for all real numbers t and $h_1(2^k,t) = -1$ for all real numbers t and $k \geq 1$. Hence by Lemma 7.3 it follows, as in Galambos (1971b), that g has a non-degenerate symmetric distribution.

Conversely, if g has a non-degenerate symmetric distribution, then $M(h_o(n,t))$ exists and is non-zero for t in an arbitrarily small, but fixed, interval. Hence for all sufficiently small real numbers t

$$(7.5.1) \qquad \sum_p \frac{1}{p}\left(1 - |g(p)|^{it}\right)$$

is convergent. It is easy to show that the convergence of (7.5.1) is equivalent to the convergence of each of the series in (7.3.1).

Moreover, if $g(2^k) \neq -1$ for some $k \geq 1$, then by modifying the proof of the theorem in Galambos (1971b) in an obvious manner it is not hard to show that

$$\sum_{g(p)<0} \frac{1}{p} = \infty .$$

This completes the proof of Theorem 7.1.

Proof of Theorem 7.2 : Suppose for some $a > 0$,

(7.5.2) $$\limsup_{n \to \infty} \frac{1}{n} N_n\{|g(m)| = a\} > 0.$$

If we define a real-valued additive arithmetic function f by

(7.5.3) $$f(p^k) = \begin{cases} \log |g(p^k)| & \text{if } g(p^k) \neq 0, \\ 0 & \text{otherwise,} \end{cases}$$

then we have

$$\limsup_{n \to \infty} \frac{1}{n} N_n\{ f(m) = \log a\} > 0.$$

Hence by Theorem 3.1, f has a distribution. Thus by Erdos-Wintner theorem, it follows that all the series in (7.3.1) converge. If

$$\sum_{g(p)<0} \frac{1}{p} = \infty ,$$

then by Theorem 7.1 the distribution of g exists and is symmetric. If

$$\sum_{g(p)<0} \frac{1}{p} < \infty ,$$

then by Lemma 7.4 g has non-symmetric distribution, provided $g(2^k) \neq -1$ for some integer $k \geq 1$ and g has symmetric distribution, provided $g(2^k) = -1$ for all $k \geq 1$. The second part of the theorem follows similarly from Theorem 3.3. This completes the proof of Theorem 7.2.

Proof of Theorem 7.3 : If the distribution of g is non-symmetric then by Lemma 7.5 we have $g(2) \neq -1$ and

$$\sum_{g(p) \leq 0} \frac{1}{p} < \infty .$$

Define new strongly multiplicative functions g_1 and g_2 by

$$g_1(p) = \begin{cases} g(p) & \text{if } g(p) \leq 0, \\ 1 & \text{otherwise,} \end{cases}$$

$$g_2(p) = \begin{cases} g(p) & \text{if } g(p) > 0, \\ 1 & \text{otherwise.} \end{cases}$$

Clearly $g = g_1 . g_2$. From Lemma 7.4 we have that the distribution of g_1 is discrete and that the c.t. of the distribution of g is the product of c.t. of the distributions of g_1 and g_2. Hence, in view of Lemmas 7.1 and 7.2, we can assume without loss of generality, that $g(p) > 0$ for all p. The distribution function H of g is absolutely continuous if and only if the distribution function G defined by

$$G(x) = \frac{H(e^x) - H(0)}{1 - H(0)}$$

is absolutely continuous. Note that the distribution function G is pure. Thus Theorem 7.3 is proved when

$$\sum_{g(p)\leq 0} \frac{1}{p} < \infty .$$

Now let

$$\sum_{g(p)<0} \frac{1}{p} = \infty .$$

Since the distribution of g is non-degenerate,

$$\sum_{g(p)=0} \frac{1}{p} < \infty .$$

By Theorem 7.1, the distribution of g is symmetric. Hence again in view of Lemmas 7.1, 7.2 and 7.4, we can assume without loss of generality, that $g(p) \neq 0$ for all p. Since the distribution of g is symmetric, the result follows from the proof of first half of the theorem, if we replace g by $|g|$. This completes the proof of Theorem 7.3.

Proof of Theorem 7.4 : Suppose that there exist a sequence $\{a_n\}$ and a non-degenerate distribution function H such that at each of its continuity point c the limit of $(1/n)\ N_n\{a_n\ g(m) < c\}$ is $H(c)$ as $n \to \infty$. Let $\begin{bmatrix} w_0(t) & 0 \\ 0 & w_1(t) \end{bmatrix}$ be the c.t. of H. Since H is non-degenerate $w_0(0) \neq 0$ and

(7.5.4) $$\sum_{g(p)=0} \frac{1}{p} < \infty .$$

Since $w_0(t)$ is continuous at zero, there exists a $\delta > 0$, such that for each $t \in [-\delta, \delta]$

$$\frac{1}{n} |a_n|^{it} \sum_{m=1}^{n} |g(m)|^{it} \not\to 0$$

as $n \to \infty$, that is

$$\frac{1}{n} \sum_{m=1}^{n} |g(m)|^{it} \not\to 0$$

as $n \to \infty$. Hence by Lemma 7.6, for each t, there exists a(t) such that

(7.5.5) $$\sum_{p} \frac{1}{p} \{1 - \mathrm{Re}(|g(p)|^{it}\ p^{-ia(t)}\ )\} < \infty .$$

Define an additive function f by

$$f(p^k) = \begin{cases} \log |g(p^k)| & \text{if } g(p^k) \neq 0, \\ 0 & \text{otherwise.} \end{cases}$$

By (7.5.4) and (7.5.5) we obtain for all $t \in [-\delta, \delta]$ that

$$\sum_p \frac{1}{p} \{1 - \mathrm{Re}(e^{it\, f(p)}\, p^{-ia(t)})\} < \infty .$$

Hence by Lemma 7.7 there exist a real number b and an additive function f_1 such that

$$f(n) = b \log n + f_1(n)$$

and

$$\sum_p \frac{1}{p} (f_1' (p))^2 < \infty .$$

Now we define a multiplicative function h by

$$h(p^k) = \exp(f_1(p^k))\, \mathrm{sgn}(g(p^k)).$$

Clearly for each $m \geq 1$,

$$g(m) = m^b\, h(m).$$

If $h(p) \neq 0$ and $|f_1(p)| < \frac{1}{2}$, then we have

$$(|h(p)| - 1)^2 = (1 - \exp(f_1(p)))^2 \leq 4|f_1(p)|^2.$$

Since

$$\sum_{h(p)=0} \frac{1}{p} + \sum_{|f_1(p)| > \frac{1}{2}} \frac{1}{p} < \infty ,$$

we get that

$$\sum_p \frac{1}{p} ((|h(p)| - 1)^*)^2 < \infty .$$

To prove the converse, first we show that if h is a real-valued multiplicative function such that

$$\sum_p \frac{1}{p} ((|h(p)| - 1)^*)^2 < \infty , \tag{7.5.6}$$

then there exists a non-degenerate distribution function H such that at each of its continuity point c the limit of $(1/n)\ N_n\{a_n\ h(m) < c\}$ is $H(c)$ as $n \to \infty$, where

$$a_n = \exp\left\{ - \sum_{p \leq n} \frac{1}{p} (|h(p)| - 1)^* \right\} .$$

By Lemma 7.8 the limit of $\frac{1}{n} |a_n|^{it} \sum_{m=1}^{n} |h(m)|^{it}$ exists, say $w_o(t)$, $n \to \infty$ uniformly in every finite interval. It is clear that $w_o(0) \neq 0$ and $w_o(t)$ is continuous at zero. Similarly one can show that the limit of

$$\frac{1}{n} |a_n|^{it} \sum_{m=1}^{n} |h(m)|^{it}\ \mathrm{sgn}(h(m))$$

exists as $n \to \infty$ uniformly in every finite interval. Hence there exists a distribution function H such that $(1/n)N_n\{a_n\, h(m) < c\}$ tends to $H(c)$, at each of its continuity point c, as $n \to \infty$.

Now for any fixed ε, $0 < \varepsilon < 1$ and m such that $\varepsilon n < m < n$, we have

$$\Big(\sum_{m<p\leq n} \frac{1}{p}(|h(p)|-1)^*\Big)^2 \leq \Big(\sum_{m<p\leq n} \frac{1}{p}\Big)\Big(\sum_{m<p\leq n} \frac{1}{p}((|h(p)|-1)^*)^2\Big)$$

$$\leq \Big(\sum_{\varepsilon n<m<n} \frac{1}{p}\Big)\Big(\sum_{\varepsilon n<p\leq n} \frac{1}{p}((|h(p)|-1)^*)^2\Big)$$

$$= (-\log \varepsilon + o(1))\Big(\sum_{\varepsilon n<p\leq n} \frac{1}{p}((|h(p)|-1)^*)^2\Big)$$

$$= o(1)$$

as $n \to \infty$ by (7.5.6). So for any $\varepsilon > 0$, $a_n = a_m(1+o(1))$ uniformly in m such that $\varepsilon n < m < n$. Hence

$$\sum_{m\leq n} |h(m)\, a_m|^{it} = \Big(|a_n|^{it} \sum_{m\leq n} |h(m)|^{it}\Big)\,(1+o(1)) + o(n)$$

and

$$\sum_{m\leq n} |h(m)\, a_m|^{it}\ \operatorname{sgn}(h(m))$$

$$= |a_n|^{it} \sum_{m\leq n} |h(m)|^{it}\ \operatorname{sgn}(h(m))\ (1+o(1)) + o(n)$$

uniformly in $t \in [-K, K]$, for every $K > 0$. So the distribution of $a_m\, h(m)$ exists and equals H.

If $g(m) = m^b\, h(m)$ and if h satisfies (7.5.6), then summing by parts gives

$$\frac{1}{n^{itb}} \sum_{m\le n} |g(m)\, a_m|^{it} = \sum_{m\le n} |h(m)\, a_m|^{it} \left(\frac{m}{n}\right)^{itb}$$

$$= \sum_{m\le n} |h(m)\, a_m|^{it} - \frac{itb}{n^{itb}} \int_1^n \left(\frac{1}{x} \sum_{m\le x} |h(m) a_m|^{it}\right) x^{itb}\, dx$$

$$= nw_o(t) + o(n) - \frac{itb}{n^{itb}} \int_1^n (w_o(t) + o(1))\, x^{itb}\, dx$$

$$= nw_o(t)\left(1 - \frac{itb}{1+itb}\right) + o(n) = \frac{n\, w_o(t)}{1 + itb} + o(n)$$

uniformly in every bounded interval, where $\begin{bmatrix} w_o(t) & 0 \\ 0 & w_1(t) \end{bmatrix}$ is the c.t. of the distribution of $a_m\, h(m)$. Similarly we can show that

$$\frac{1}{n^{itb}} \sum_{m\le n} |g(m)\, a_m|^{it}\, \mathrm{sgn}(g(m)) = \frac{n\, w_1(t)}{1 + itb} + o(n)$$

uniformly in every finite interval. From this it follows, as above, that if G is the distribution of $a_m\, g(m)$, then $(1/n)\, N_n\{a_n\; g(m) < c\}$ tends to $G(c)$ at each of its continuity point c, as $n \to \infty$. This completes the proof of Theorem 7.4.

CHAPTER 8

CHARACTERISTIC FUNCTIONS AND TRANSFORMS OF ARITHMETIC FUNCTIONS

1. Introduction : In this chapter, some results towards the characterization of distributions of real-valued additive and multiplicative arithmetic functions are obtained. It is shown that, if H is distribution of a real-valued additive function f, such that $f(2^k) = k\,f(2)$ for all $k \geq 1$, then there exists a discrete infinitely divisible law G such that $H * G$ has an infinitely divisible characteristic function with discrete Levy functions and without Normal factor. A class of stable laws are given which cannot be distributions of real-valued multiplicative functions g satisfying, for all $k \geq 1$, $g(2^k) = (g(2))^k$. One such is the Gamma distribution. Finally it is shown that no real-valued multiplicative function g satisfying $g(2^k) = (g(2))^k$ for all $k \geq 1$, can have uniform distribution.

Most of these results are suggested by the following well-known fact. If a completely additive arithmetic function f (i.e., $f(p^k) = k\,f(p)$ for all $k \geq 1$ and for all p) has a distribution, then it is infinitely divisible with discrete Levy functions and without Normal factor.

2. Notations and definitions : Let $\underset{=}{A}$ denote the set of all real-valued additive arithmetic functions f satisfying $f(2^k) = k\, f(2)$ for all $k \geq 1$. Let $\underset{=}{M}$ denote the set of all real-valued multiplicative arithmetic functions g satisfying $g(2^k) = (g(2))^k$ for all $k \geq 1$.

Definition : A distribution function H is said to be an M-i.d law if there exist $0 < n_1 < n_2 < \ldots$ such that for each i, the characteristic transform (c.t.) of H can be written as $(W^{(n_i)}(t))^{n_i}$ for some c.t. $W^{(n_i)}(t)$.

3. Preliminary results : If h is an infinitely divisible characteristic function (i.d.c.f.) then it can be written uniquely in the form (see Lukacs, 1970, p.118)

$$\log h(t) = it\, a - \sigma\, t^2 + \int_{-\infty}^{-o} \left(e^{itu} - 1 - \frac{itu}{1+t^2}\right) dM(u)$$

$$+ \int_{+o}^{\infty} \left(e^{itu} - 1 - \frac{itu}{1+t^2}\right) dN(u)$$

where M, N and σ satisfy the following conditions :

i) M and N are non-decreasing in the intervals $(-\infty, o)$ and $(o, +\infty)$ respectively.

ii) The integrals $\int_{-\varepsilon}^{0} u^2 \, dM(u)$ and $\int_{0}^{\varepsilon} u^2 \, dN(u)$ are finite for each $\varepsilon > 0$.

iii) $M(-\infty) = N(+\infty) = 0$.

iv) σ is real and non-negative and a is real.

The functions M and N are called Levy functions and σ is said to be the Normal component. If $\sigma = 0$, we say that h has no Normal factor.

The distribution corresponding to h is purely discrete if and only if, $\sigma = 0$, $\int_{-\infty}^{-0} dM(u) < \infty$, $\int_{+0}^{\infty} dN(u) < \infty$ and both M and N are discrete (Lukacs, 1970, p.124).

<u>Lemma 8.1 (Zolotarev, 1962).</u> The distribution function H is M-i.d. if and only if, its c.t.

$$W(t) = \begin{bmatrix} w_0(t) & 0 \\ 0 & w_1(t) \end{bmatrix}$$

has the form

$$w_0(t) = \alpha_0 \, h_1(t) \, h_2(t) \qquad (8.3.1)$$

$$w_1(t) = \alpha_1 \, h_1(t)/h_2(t) \qquad (8.3.2)$$

where

(a) h_1 is an i.d.c.t., h_2 is an i.d.c.f. of the form

$$\log h_2(t) = \int (e^{itu} - 1)\, dH(u)$$

and

(b) $0 \leq \alpha_0 \leq 1$, $|\alpha_1| \leq \alpha_0 \exp(-2 \int dH)$.

<u>Remark</u> : If the distribution H corresponding to W is an M-i.d., if the Levy functions of h_1 and h_2 in (8.3.1) and (8.3.2) are discrete and if h_1 has no Normal factor, then H is purely discrete. This statement can be established in the same way as Lemma 5.7.

4. <u>Main results</u> :

<u>Theorem 8.1</u> : Suppose $g \in \underline{\underline{M}}$ and g has a distribution with c.t. W. Then there exists a discrete M-i.d. law with c.t. $W^{(1)}$ such that the function $W^{(2)}$ given by

$$W^{(2)}(t) = W(t).\ W^{(1)}(t)$$

is the c.t. of an M-i.d. law. If we write

$$W^{(2)}(t) \doteq \begin{bmatrix} w_0^{(2)}(t) & 0 \\ 0 & w_1^{(2)}(t) \end{bmatrix},$$

then the corresponding characteristic functions $h_1^{(2)}$ and $h_2^{(2)}$ in the representation (8.3.1) and (8.3.2) are without Normal factors and with discrete Levy functions. In particular, log-normal distribution cannot be the distribution of a $g \in \underline{\underline{M}}$.

As is well-known (Lukacs, 1970), the logarithms of the characteristic functions of the stable distributions can be written in the form

$$it\gamma - \lambda\,|t|^{\alpha} \exp\left\{-\,i\tfrac{\pi}{2}\,K(\alpha)\,\beta\,\operatorname{sgn}(t)\right\} \quad \text{if} \quad \alpha \neq 1,$$

$$it\gamma - \lambda\,|t|\left\{1 + \tfrac{2}{\pi}\,\beta\,\log|t|\,\operatorname{sgn}(t)\right\} \quad \text{if} \quad \alpha = 1$$

where $K(\alpha) = 1 - |1-\alpha|$ and $0 < \alpha \leq 2$; $-1 \leq \beta \leq 1$, $-\infty < \gamma < \infty$ and $\lambda > 0$.

Let $\underline{\underline{U}}$ denote the set of all stable laws with $\alpha < 1$, $\gamma = 0$; $\alpha = 1$, $\beta = 0$; or $\alpha > 1$, $\beta = 0$, $\gamma = 0$. It is known that (Zolotarev, 1967), all the distributions in $\underline{\underline{U}}$ are M-i.d. laws and the Levy functions of the characteristic functions h_1, h_2 in the representations (8.3.1) and (8.3.2) of the corresponding characteristic transforms are continuous.

From this and Theorem 8.1, we obtain the following

<u>Corollary 8.1</u> : No distribution in $\underline{\underline{U}}$ can be the distribution of a multiplicative function g in $\underline{\underline{M}}$.

<u>Theorem 8.2</u> : If $f \varepsilon \underline{\underline{A}}$ has a distribution H, then there is a discrete infinitely divisible (i.d.) law G such that $H * G$ is i.d., with discrete Levy functions and without Normal factor. Hence no $f \varepsilon \underline{\underline{A}}$ can have Gamma distribution or any stable distribution. In particular no $f \varepsilon \underline{\underline{A}}$ can have exponential distribution. (Here * denotes the convolution of the distribution functions.)

<u>Theorem 8.3</u> : $g \varepsilon \underline{\underline{M}}$ cannot have uniform distribution.

5. <u>Proofs of the main results</u> :

<u>Proof of Theorem 8.1</u> : Suppose $g \varepsilon \underline{\underline{M}}$ has a non-degenerate distribution. Then the c.t.

$$W(t) = \begin{bmatrix} w_o(t) & 0 \\ 0 & w_1(t) \end{bmatrix}$$

of the distribution H of g is given by (see Bakstys, 1969, and Theorem 7.1)

$$w_0(t) = \prod_p \left(1 - \frac{1}{p}\right)\left(1 + \sum_{k=1}^{\infty} \frac{1}{p^k} |g(p^k)|^{it}\right)$$

and

$$w_1(t) \equiv 0 \qquad \text{or}$$

$$w_1(t) = \prod_p \left(1 - \frac{1}{p}\right)\left(1 + \sum_{k=1}^{\infty} \frac{1}{p^k} |g(p^k)|^{it} \operatorname{sgn}(g(p^k))\right)$$

according as H is symmetric or non-symmetric.

For $p > 2$ we have (see Delange, 1963)

$$\left(1 - \frac{1}{p}\right)\left(1 + \sum_{k=1}^{\infty} \frac{|g(p^k)|^{it}}{p^k}\right) = \exp\left[\sum_{k=1}^{\infty} \frac{C(k,p,t) - 1}{k\, p^k}\right]$$

where

$$C(1,p,t) = |g(p)|^{it}$$

and for $r > 1$,

$$C(r,p,t) = r|g(p^r)|^{it} - \sum_{j=1}^{r-1} C(j,p,t)|\, g(p^{r-j})|^{it} .$$

Define for $p > 2$ and for all real numbers t

$$D(1,p,t) = 0,$$

$$D(2,p,t) = (C(1,p,t))^2,$$

and if $r > 2$.

$$D(r,p,t) = \sum_{k=1}^{r-1} (D(k,p,t) + C(k,p,t))|g(p^{r-k})|^{it}.$$

It is not difficult to show, by induction on $r \geq 2$, that

$$C(r,p,t) + D(r,p,t) = \sum_k a_{k,r} |b_{k,r}|^{it},$$

$$D(r,p,t) = \sum_k d_{k,r} |c_{k,r}|^{it},$$

where the summations are taken over a finite number of values of k depending on r; $a_{k,r}$, $d_{k,r}$ are positive integers, <u>independent of</u> g, such that

$$\sum_k a_{k,r} = 2^{r-1} \quad \text{and} \quad \sum_k d_{k,r} = 2^{r-1} - 1$$

and $b_{k,r}$, $c_{k,r}$ are some products of $\{(g(p^s))^n, 1 \leq s \leq r$ and $n \geq 1\}$. We put

$$d(r,p) = \begin{cases} 0 & \text{if } c_{k,r} = 0 \text{ for all } k, \\ \Sigma'' d_{k,r} & \text{otherwise,} \end{cases}$$

where the sum Σ'' is taken over all k for which $c_{k,r} \neq 0$. Since

$$\sum_{p>2} \sum_{r=2}^{\infty} \frac{2^{r-1}-1}{rp^r} < \infty ,$$

it follows, by Lemma 5.7, that

$$s_1(t) = \exp\Big(\sum_{p>2} \sum_{r=2}^{\infty} \sum_{c_{k,r} \neq 0} \frac{d_{k,r}}{rp^r} \Big(|c_{k,r}|^{it} - 1 - \frac{it \log|c_{k,r}|}{1+(\log|c_{k,r}|)^2} \Big)\Big)$$

$$= \exp\Big(\sum_{p>2} \sum_{r=2}^{\infty} ((D(r,p,t) - d(r,p) - it\, c(r,p))/rp^r)\Big)$$

is a discrete i.d.c.f., where for $r \geq 2$

$$c(r,p) = \begin{cases} 0 & \text{if } c_{r,k}=0 \text{ for all } k, \\ \sum\limits_{\substack{k \\ c_{k,r}\neq 0}} d_{k,r} \dfrac{\log|c_{k,r}|}{1+(\log|c_{k,r}|)^2} & \text{otherwise.} \end{cases}$$

So the distributions corresponding to the characteristic transforms,

$$\begin{bmatrix} s_1(t) & 0 \\ 0 & 0 \end{bmatrix} \quad \text{and} \quad s_1(t) \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

are discrete. Let

$$a(r,p) = \begin{cases} 0 & \text{if all } b_{k,r} = 0 , \\ \sum\limits_{\substack{k \\ b_{k,r} \neq 0}} a_{k,r} & \text{otherwise,} \end{cases}$$

$$b(r,p) = \begin{cases} 0 & \text{if } b_{k,r}=0 \text{ for all } k, \\ \sum\limits_{\substack{k \\ b_{k,r} \neq 0}} a_{k,r} \dfrac{\log|b_{k,r}|}{1+(\log|b_{k,r}|)^2} & \text{otherwise,} \end{cases}$$

and

$$s_2(t) = \sum_{k=1}^{\infty} \frac{|g(2)|^{itk} - 1}{k\ 2^k} .$$

Note that if $g(2) = 0$, then $s_2(t)$ is constant for all t and if $g(2) \neq 0$ then $\exp(s_2(t))$ is an i.d.c.f. Now

$$\log w_0(t) + \log s_1(t) = s_2(t) + \sum_{p>2} \Big[\frac{|g(P)|^{it} - 1}{p}$$

$$+ \sum_{r=2}^{\infty} ((C(r,p,t) + D(r,p,t)-1-d(r,p)- it\ c(r,p))/rp^r)\Big].$$

Since

$$s_3(t) = \sum_{p>2} \sum_{r=2}^{\infty} [(C(r,p,t)+D(r,p,t)-a(r,p)-itb(r,p))/rp^r]$$

$$= \sum_{p>2} \sum_{r=2}^{\infty} \sum_{b_{k,r}\neq 0} \frac{a_{k,r}}{rp^r}\left(|b_{k,r}|^{it} - 1 - \frac{it \log|b_{k,r}|}{1 + (\log|b_{k,r}|)^2}\right)$$

and

$$s_4(t) = \sum_{\substack{p>2 \\ g(p)\neq 0}} \frac{1}{p}\left(|g(p)|^{it} - 1 - \frac{it \log|g(p)|}{1+(\log|g(p)|)^2}\right)$$

converge absolutely and uniformly in every finite interval, so $\exp(s_3(t) + s_4(t))$ is an i.d.c.f. Hence

$$\log w_o(t) + \log s_1(t) = s_2(t) + s_3(t) + s_4(t) - b + ita$$

where

$$a = \sum_{\substack{p>2 \\ g(p)\neq 0}} \frac{1}{p} \frac{\log|g(p)|}{1+(\log|g(p)|)^2} + \sum_{p>2} \sum_{r=2}^{\infty} \frac{b(r,p) - c(r,p)}{rp^r}$$

and

$$b = \sum_{p>2} \sum_{r=2}^{\infty} \frac{d(r,p) - a(r,p) + 1}{rp^r} + \sum_{\substack{p>2 \\ g(p)=0}} \frac{1}{p} .$$

Hence

$$w_o(t)\begin{bmatrix} s_1(t) & 0 \\ 0 & 0 \end{bmatrix} \quad \text{and} \quad w_o(t)s_1(t)\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

are M-i.d. characteristic transforms. This proves the theorem when the distribution of g is symmetric as well as when $g(m) \geq 0$ for all $m \geq 1$.

Now let $g \in \underline{M}$. Suppose the distribution of g is non-symmetric. Define the multiplicative functions $g_1 \in \underline{M}$ and $g_2 \in \underline{M}$ by

$$g_1(p^k) = \begin{cases} g(p^k) & \text{if } g(p^k) \geq 0, \\ 1 & \text{if } g(p^k) < 0, \end{cases}$$

$$g_2(p^k) = \begin{cases} g(p^k) & \text{if } g(p^k) < 0, \\ 1 & \text{if } g(p^k) \geq 0. \end{cases}$$

Note that $g = g_1 \cdot g_2$. Since g has non-symmetric distribution,

$$\sum_{g_2(p)\neq 1} \frac{1}{p} < \infty .$$

Since the distribution of g_2 is discrete by Lemma 7.4, it can be shown, as above, that the theorem holds for g_2 also. By Lemma 7.4, the c.t. of g is given by the product of the characteristic transforms of g_1 and g_2. Hence the theorem follows, as we have already shown the validity of Theorem for g_1.

Theorem 8.2 follows from the proof of Theorem 8.1, on replacing g by e^f.

Proof of Theorem 8.3 : If $g \in \underline{M}$ has uniform distribution then $g(m) \neq 0$ for all $m \geq 1$. Let the interval (a,b) be the spectrum of the distribution of g.

Let us consider first the case when either $a = 0$ or $b = 0$ or $a = -b$. Then the distribution of $|g(m)|$ is uniform and so the additive arithmetic function $\log|g(m)|$ has, by Theorem 7.1, a distribution, with characteristic function

$$\frac{C^{it}}{1 + it}$$

for some $C > 0$. This is an i.d.c.f. with continuous Levy functions. Theorem 8.2 gives a contradiction and we conclude that g cannot have uniform distribution in this case.

Now we suppose that both $a \neq 0$ and $b \neq 0$. If either $0 < a < b$ or $a < b < 0$, then $\log|g(m)|$ is bounded and hence its distribution is singular, by Corollary 5.2. So this case does not arise.

Now we examine the case when $a < 0 < b$ and $b \neq -a$. First let $0 < |a| < b$. Note that the characteristic function of $\log|g(m)|$ is

$$s_5(t) = b^{it}(1 + |\tfrac{a}{b}|^{it+1})(1 + |\tfrac{a}{b}|)^{-1}[1 + it]^{-1}.$$

If $d = |\frac{a}{b}|$, then the characteristic function

$$s_6(t) = \exp(\tfrac{1}{2} \sum_{k=1}^{\infty} d^{2k}(e^{2itk \log d} - 1))$$

is i.d. and the corresponding distribution is discrete. Using Theorem 8.2 we can find an i.d.c.f. s_7 with discrete Levy functions such that $s_5 . s_7$ is an i.d.c.f. with discrete Levy functions. Also note that $s_5 . s_6$ is an i.d.c.f. So on the one hand we have that the Levy functions of $s_5 . s_6 . s_7$ are discrete and on the other, the Levy functions of $s_5 . s_6 . s_7$ are not discrete since, $(1/(1+it))$ is a factor of $s_5 . s_6 . s_7$ and the Levy functions of $(1/(1+it))$ are continuous. This contradiction shows that in this case g cannot have uniform distribution. Similarly we can show that if $0<b<|a|$, then again g cannot have uniform distribution. This completes the proof of Theorem 8.3.

CHAPTER 9

DISTRIBUTION OF VALUES OF ADDITIVE FUNCTIONS ON THE SET OF PAIRS OF POSITIVE INTEGERS

1. Introduction : Delange (1969) defined a density for sets of pairs $[m,n]$ of positive integers and determined it for some sets defined by arithmetical properties. In this chapter we give necessary and sufficient conditions (similar to Erdos-Wintner theorem for additive arithmetic functions) for a real-valued additive function f on the set of pairs of positive integers to have a distribution and generalize some of the results obtained in the previous chapters. We also give necessary and sufficient conditions for a real-valued additive function on the set of pairs of positive integers to have distribution (mod 1). Finally we give necessary and sufficient conditions for

$$\frac{1}{xy} \text{ card } \{[m,n] : m \leq x, n \leq y, f(m,n) - a_{x,y} < c\}$$

to tend to a distribution function as x and y tend to infinity independently, where $\{a_{x,y}\}$ is a set of real numbers.

Instead of pairs one could consider systems of k positive integers, k being any fixed integer > 2. There is no essential difference between the case $k = 2$ and $k > 2$. We restrict our exposition to the former case for the sake of simplicity in writing.

2. Notations and definitions : Throughout this thesis, we let Z_2 to denote the set of all pairs of positive integers. For a subset B of Z_2, $N(B)$ denotes its cardinality.

Definition : A subset E of Z_2 is said to have density $\delta(E)$ if $(1/xy)\, N\{[m,n] \in E : m \leq x,\ n \leq y\}$ tends to $\delta(E)$ as x and y tend to infinity independently.

Definition : A real-valued function f on Z_2 is said to be additive if

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

whenever $(m_1 n_1,\ m_2 n_2) = 1$.

Note that if f is additive then $f(1,1) = 0$. It is clear from the definition that if

$$\alpha(p,n) = \begin{cases} 0 & \text{if } p \nmid n \\ r & \text{if } p^r \| n \qquad (r \geq 1) \end{cases}$$

and

$$f_p(m,n) = f(p^{\alpha(p,m)},\ p^{\alpha(p,n)}),$$

then for all $[m,n] \in Z_2$,

$$f(m,n) = \sum_{p|mn} f_p(m,n).$$

We define for any positive integer r.

$$f(m,n)_r = \sum_{p \le r} f_p(m,n).$$

For any $x \geq 1$ and $y \geq 1$, define

$$A_2(x,y,f) = \sum_{p \le x} \frac{f(p,1)}{p} + \sum_{p \le y} \frac{f(1,p)}{p}$$

$$B_2(x,y,f) = \left[\sum_{p \le x} \frac{(f(p,1))^2}{p} + \sum_{p \le y} \frac{(f(1,p))^2}{p} \right]^{1/2}.$$

With each f we associate an additive function f^+ defined by

$$f^+(p^t,p^k) = \begin{cases} f(p,1) + f(1,p) & \text{if } t > 0 \text{ and } k > 0, \\ f(p,1) & \text{if } t > 0 \text{ and } k = 0, \\ f(1,p) & \text{if } t = 0 \text{ and } k > 0. \end{cases}$$

Definition : A real-valued function g on Z_2 is said to have a distribution, if there is a distribution function Q on the real line such that the density of the set $\{[m,n] : g(m,n) < c\}$ exists and is equal to $Q(c)$ for each continuity point c of Q.

<u>Definition</u> : A real-valued additive function f on Z_2 is said to have distribution (mod 1) if there is a non-decreasing, right continuous function H on the real line such that $H(c) = 0$ if $c < 0$, $H(c) = 1$ if $c \geq 1$ and for all continuity points $a, b \in (0,1)$ of H with $a < b$, the density of $\{[m,n] : a < \{f(m,n)\} < b\}$ exists and equals $H(b) - H(a)$, where $\{x\}$ denotes the fractional part of x.

Let, for any real number x, $\| x \| = \min(\{x\}, 1 - \{x\})$.

We put $e(t) = \exp(2\pi i t)$.

For real numbers a, b, integer k and real-valued additive function f, we write

$$h(f,k,a,b) = \left(\sum_{\substack{j \geq 0 \\ r \geq 0}} \frac{e(kf(2^j, 2^r))}{2^{j(1+2\pi i a)+r(1+2\pi i b)}} \right) \left(\sum_{\substack{j \geq 0 \\ r \geq 0}} \frac{e(kf(3^j, 3^r))}{3^{j(1+2\pi i a)+r(1+2\pi i b)}} \right)$$

<u>Definition</u> : A complex-valued function h defined on Z_2 is said to be multiplicative if $h(1,1) = 1$ and

$$h(m_1 m_2, n_1 n_2) = h(m_1, n_1)\, h(m_2, n_2)$$

whenever $(m_1 n_1, m_2 n_2) = 1$.

<u>Definition</u> : A complex-valued function Z_2 is said to have a mean value if the limit of $(1/xy) \sum_{\substack{m \le x \\ n \le y}} h(m,n)$ exists as x and y tend to infinity independently ; and the limit, if it exists, is called the mean value of h.

3. <u>Main results</u> :

<u>Theorem 9.1</u> : Let f be a real-valued additive function on Z_2. f has a distribution if and only if the three series

$$(9.3.1) \qquad \sum_p \frac{f^*(p,1)}{p}$$

$$(9.3.2) \qquad \sum_p \frac{f^*(1,p)}{p}$$

$$(9.3.3) \qquad \sum_p \frac{1}{p}\left([f^*(p,1)]^2 + [f^*(1,p)]^2\right)$$

converge, where

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

and

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

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

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

Theorem 9.1 was also obtained by Delange **independently.** [Personal communication.]

Theorem 9.2 : Let f be a real-valued additive function on Z_2. Let A be a set of primes such that

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

Let f be such that for all $p \notin A$ both $f(p,1)$ and $f(1,p)$ are non-negative. If there exists positive constants c, δ, and two sequences $\{x_i\}$ and $\{y_i\}$ of positive real numbers such that

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

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

Theorem 9.3 : Let f be a real-valued additive function on Z_2. f has distribution (mod 1) if and only if, for each integer k one of the following conditions is satisfied :

i) For all real t and u,

$$\sum_p \frac{1}{p}(\| k\ f(p,1) - t \log p\|^2 + \| k\ f(1,p) - u \log p\|^2) = \infty .$$

ii) Each of the following three series is convergent.

$$\sum_p \frac{1}{p}(\| k\ f(p,1)\|^2 + \| k\ f(1,p)\|^2)$$

$$\sum_p \frac{1}{p} \| k\ f(p,1)\| \operatorname{sgn}(\tfrac{1}{2} - \{k\ f(p,1)\})$$

$$\sum_p \frac{1}{p} \| k\ f(1,p)\| \operatorname{sgn}(\tfrac{1}{2} - \{k\ f(1,p)\}).$$

In particular f has uniform distribution (mod 1) if and only if for each non-zero integer k, i) holds.

Remark : This result was also obtained by Delange (personal communication) independently under an extra assumption that

$$f(2^j,\ 2^r) = f(2^j, 1) + f(1,\ 2^r)$$

and

$$f(3^j,\ 3^r) = f(3^j, 1) + f(1,\ 3^r)$$

for all $j \geq 0$ and $r \geq 0$.

Theorem 9.4 : Let f be a real-valued additive function on Z_2 having a distribution (mod 1). Let, for each integer k,

$$a_k = |h(f,k,0,0)|.$$

The distribution (mod 1) of f is

a) continuous if and only if

$$N^{-1} \sum_{k=-N}^{N} a_k \exp\left(-2\sum_p \frac{1}{p}\left(\sin^2 \pi k f(p,1) + \sin^2 \pi k f(1,p)\right)\right) \to 0, \text{ as } N \to \infty,$$

b) absolutely continuous with a derivative belonging to the Lebesgue class $L_2[0,1]$ if and only if, the series

$$\sum_{k=-\infty}^{\infty} a_k \exp\left(-4 \sum_p \frac{1}{p}(\sin^2 \pi k f(p,1) + \sin^2 \pi k f(1,p))\right)$$

is convergent.

(In the statement of this theorem it is to be understood

that if a series

$$\sum_p \frac{1}{p} (\sin^2 \pi k f(p,1) + \sin^2 \pi k f(1,p)) = \infty$$

then the corresponding

$$\exp(-2 \sum_p \frac{1}{p} (\sin^2 \pi k f(p,1) + \sin^2 \pi k f(1,p)))$$

is defined to be zero.)

Theorem 9.5 : Let f be a real-valued additive function on Z_2. Suppose there is a sequence $\{\alpha_n\}$ of real numbers and a distribution function H on the real line such that for each of its continuity point c the limit of

$$\frac{1}{n^2} N\{[m,m'] : m \leq n, \ m' \leq n \ \text{ and } \ f(m,m') - \alpha_n < c\}$$

is $H(c)$ as $n \to \infty$. Then there exist two real numbers a and b and an additive function g on Z_2 such that for all $m \geq 1, n \geq 1$

(9.3.4) $\quad f(m,n) = a \log m + b \log n + g(m,n)$

and

(9.3.5) $$\sum_p \frac{[g^*(1,p)]^2}{p} + \sum_p \frac{[g^*(p,1)]^2}{p} < \infty .$$

In this case, we can set $\alpha_n = \alpha'_{n,n}$ + constant + o(1), where

$$\alpha_{x,y} = \sum_{p \leq x} \frac{g^*(p,1)}{p} + \sum_{p \leq y} \frac{g^*(1,p)}{p}$$

and $\alpha'_{x,y} = \alpha_{x,y} + a \log x + b \log y$.

If f satisfy (9.3.4) with g satisfying (9.3.5) then there exists a distribution function G such that at each of its continuity point c, we have that

$$\frac{1}{xy} N\{[m,n] : m \leq x,\ n \leq y,\ f(m,n) - \alpha'_{x,y} < c\}.$$

tends to $G(c)$ as x and y tend to infinity independently.

Remark : This result is a generalization to additive functions on Z_2, of a result due to Elliott and Ryavec (1971).

Let f be a real-valued additive function on Z_2. We define two completely additive arithmetic functions f_1 and f_2 by $f_1(p) = f(p,1)$ and $f_2(p) = f(1,p)$. It is clear from Theorem 9.1 and the Erdos-Wintner theorem that f has a distribution if and only both f_1 and f_2 have diatributions.

Theorem 9.6 : The distribution of f is absolutely continuous if and only if the convolution of the distributions of f_1 and f_2 is absolutely continuous.

Many results of the previous chapters can be generalized without much difficulty to additive and multiplicative functions on Z_2. For example, it can be shown that conditions similar to (1.3.1), (1.3.2) and (1.3.3) are necessary and sufficient for the existence of the distribution of

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

where $f_1, \ldots, f_s$ are additive functions on Z_2 and $F_i, G_i \in \underline{P}$. In this connection we state the following results.

<u>Theorem 9.7</u>: Let f be a real-valued additive function on Z_2. If

$$\limsup_{x \to \infty} (1/x^2)\, N\{[m,n] : m \le x,\ n \le x,\ f(m,n) = a\} > 0$$

for some real number a, then f has a distribution.

Proof of this theorem is similar to that of Theorem 3.1.

<u>Theorem 9.8</u>: Let f be a real-valued additive function on Z_2 having a distribution. Suppose A is a set of primes such that

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

If for some $c > 0$,

$$\sum_{\substack{p>N \\ p \not\in A}} \frac{1}{p} [(f(1,p))^2 + (f(p,1))^2] = O(N^{-c})$$

as $N \to \infty$, then the distribution of f is singular.

In particular if f is bounded then the distribution of f exists and is singular.

Proof of this theorem is similar to that of Corollary 5.1.

4. Preliminary results :

Lemma 9.1 : We have, for any f

$$\sum_{\substack{m \le x \\ n \le y}} | f^{+}(m,n) - A_2(x,y,f)|^2 < C\, xy\, B_2^2(x,y,f),$$

where C is a constant independent of f.

Proof of this lemma is similar to Turan-Kubilius inequality and so is omitted.

Lemma 9.2 : If $B_2(x,y,f) \to \infty$ as $x \to \infty$, $y \to \infty$ independently and $\sup_p \max(|f(p,1)|, |f(1,p)|) \le c_1$ for some positive constant c_1, then

$$\lim_{x \to \infty} (1/x^2)\, N\{[m,n] : m \leq x,\ n \leq x, f(m,n) - A_2(x,x,f) < aB_2(x,x,f)\}$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{a} e^{-y^2/2}\, dy$$

for all real numbers a.

<u>Proof</u> : In what follows, we write $A_2(s)$ and $B_2(s)$ for $A_2(s,s,f)$ and $B_2(s,s,f)$ respectively.

Let $r(s) = s^{1/B_2(s)}$. By hypothesis it follows that $[B_2(r(s))/B_2(s)] \to 1$ as $s \to \infty$. First we shall show that for each real number a,

(9.4.1) $$s^{-2} N\{[m,n] : m \leq s, n \leq s, f^+(m,n)_{r(s)} - A_2(r(s)) < aB_2(r(s))\}$$

$$\to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{a} e^{-u^2/2}\, du\,, \quad \text{as} \quad s \to \infty.$$

Let $E(p,t,s)$ be as defined in § 3 of Chapter 2. Define for any square-free integers k and k' such that $1 \leq k, k' \leq s$

$$E_{s,k,k'} = E_{s,k} \times E_{s,k'}.$$

Clearly for different pairs (k,k'), the sets $E_{n,k,k'}$ have no element in common. Let $\underline{\underline{H}}_s$ be the smallest algebra containing all $E_{s,k,k'}$. Clearly every $E \in \underline{\underline{H}}_s$ is of the form

$$\bigcup_{(k',k)\varepsilon K} E_{s,k',k} ,$$

where K is a subset of $E_s \times E_s$. Define, for $t',t = o,1$ and $p \leq r(s)$,

$$P^*(E(p,t,s) \times E(p,t',s)) = P(E(p,t,s))\, P(E(p,t',s))$$

where P is as defined in § 3 of Chapter 2. Let

$$P^*(E) = \sum_{(k,k')\varepsilon K} \prod_{p\leq r(s)} P^*(E(p,\delta_p(k),s) \times E(p,\delta_p(k'),s))$$

whenever K is a subset of $E_s \times E_s$ and

$$E = \bigcup_{(k,k')\ \varepsilon\ K} E_{s,k,k'} .$$

Following the arguments in Kubilius (1964, p.27), we get, uniformly for all $\mathbf{E}\ \varepsilon\ \underline{\underline{H}}_s$, that

(9.4.2) $\quad s^{-2}\, N\{[m,n]\ \varepsilon\ E\} - P^*(\mathbf{E}) = O(\exp(-c \log s/\log r(s)))$

as $s \to \infty$, where c is a positive constant.

Let, for each $p \leq r(s)$, η_p be the random variable on $E_s \times E_s$ defined by

$$\eta_p(m,n) = f_p^+(m,\ n).$$

Clearly $\{\eta_p : p \leq r(s)\}$ are independent random variables with respect to the measure P^*. It follows from (9.4.2) that

(9.4.3) $$s^{-2} N\{[m,n] : m \leq s, n \leq s, f^+(m,n)_{r(s)}, \varepsilon E\}$$

$$= P^*\{\sum_{p \leq r(s)} \eta_p \varepsilon E\} + O(\exp(-c \log s / \log r(s)))$$

uniformly for all Borel subsets E of the real line.

By Lindeberg-Levy central limit theorem (see Billingsley, 1968) it follows, as in the proof of Theorem 4.2 of Kubilius (1964), that for each real number a,

(9.4.4) $$P^*\{\sum_{p \leq r(s)} \eta_p - A_2(r(s)) < aB_2(r(s))\} \to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{a} e^{-u^2/2} du$$

as $s \to \infty$. Since $[B_2(r(s))/B_2(s)] \to 1$ as $s \to \infty$, from (9.4.3) and (9.4.4) we got that

(9.4.5) $$s^{-2} N\{[m,n] : m \leq s, n \leq s, f^+(m,n)_{r(s)} - A_2(r(s)) < aB_2(s)\}$$

$$\to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{a} e^{-u^2/2} du$$

as $s \to \infty$. By Lemma 9.1 and Chebyshev's inequality, we have for any $\varepsilon > 0$

$$(9.4.6)\quad N\{[m,n] : m,n \le s\ , |f^+(m,n)-A_2(x)-f^+(m,n)_{r(s)}+A_2(r(s))| > \varepsilon B_2\}$$

$$\le C\, s^2(B_2^2(s) - B_2^2(r(s))/(\varepsilon^2\, B_2^2(s))$$

where C is as in the statement of Lemma 9.1. Since $(B_2(s)/B_2(r(s))) \to 1$ as $s \to \infty$, it follows from (9.4.5) and (9.4.6) that for all real numbers a,

$$(9.4.7)\quad (1/s^2)N\{[m,n] : m \le x, n \le x,\ f^+(m,n)-A_2(s) < aB_2(s)\}$$

$$\to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{a} e^{-b^2/2}\, db, \quad \text{as} \quad x \to \infty.$$

Now for any $\varepsilon > 0$ choose an M such that

$$\sum_{p>M} \frac{1}{p^2} < \frac{\varepsilon}{4}.$$

Put

$$L(\varepsilon) = \sum_{\max\{p^t, p^k\} \le M} |f(p^t,p^k) - f^+(p^t,p^k)|.$$

Let $A = \{[m,n] : q^2 | mn$ for some $q > M\}$. We have

$$N\{[m,n] \varepsilon\ A : m \le x,\ n \le y\} \le \sum_{q>M} \{[\frac{x}{q^2}]\, y + x[\frac{y}{q^2}] + [\frac{x}{q}]\,[\frac{y}{q}]\}$$

$$\le \varepsilon xy.$$

We note that if $[m,n] \notin A$, then

$$|f(m,n) - f^{+}(m,n)| \leq L(\varepsilon).$$

Hence for every $\varepsilon > 0$ there exists an $L(\varepsilon)$ such that

$$(9.4.8) \quad N\{[m,n] : m \leq x, \ n \leq y, |f(m,n)-f^{+}(m,n)| > L(\varepsilon)\} < \varepsilon xy.$$

Now (9.4.7) and (9.4.8) give the required result.

Lemma 9.3 : Let $b_1,\ldots,b_r$ be real numbers. For each ε such that $0 < \varepsilon < (1/2)$, there exists a non-zero integer n such that

$$\{n\, b_i\} < \varepsilon , \quad i = 1,\ldots,r.$$

Proof : If $b_1,\ldots,b_r$ are rational numbers, then the lemma follows trivially. Suppose b_1 is irrational. The lemma follows from Kronecker's theorem (see Chandrasekharan, 1968), if $1, b_1,\ldots,b_r$ are linearly independent (i.e., there is no linear relation of the form $\sum_{i=1}^{r} n_i\, b_i = n$, where n, n_i are integers, and $(n_1,\ldots,n_r,n) \neq (0,0,\ldots,0)$). Suppose $1, b_1,\ldots,b_j$ are linearly independent and $b_{j+1},\ldots,b_r$ are linearly dependent on $1, b_1,\ldots,b_j$, i.e., there exist integers a_{is} $(i = j+1, \ldots, r \,;\ s = 0,\ldots,j)$ such that

$$\sum_p \frac{1}{p}(1 - \mathrm{Re}(h(1,p)\, p^{-iu})) = \infty$$

for all real u, then h possesses zero mean value.

If there exist real numbers a and b such that

$$\sum_p \frac{1}{p}(1 - \mathrm{Re}(h(p,1)\, p^{-ia})) < \infty,$$

and

$$\sum_p \frac{1}{p}(1 - \mathrm{Re}(h(1,p)\, p^{-ib})) < \infty,$$

then h possesses zero mean value if

$$(9.4.9) \quad \left(\sum_{j,r\geq 0} \frac{h(2^j, 2^r)}{2^{j(1+ia)+r(1+ib)}} \right) \left(\sum_{j,r\geq 0} \frac{h(3^j, 3^r)}{3^{j(1+ia)+r(1+ib)}} \right)$$

is zero. If (9.4.9) is not zero, then as x and y tend to infinity independently, we have

$$(1/xy) \sum_{\substack{m\leq x \\ n\leq y}} h(m,n) = C\, x^{ia}\, y^{ib}\, L_1(\log x)\, L_2(\log y) + o(1)$$

where C is a non-zero complex number and L_1 and L_2 are functions on the positive real line defined by

$$L_1(t) = \exp(i \sum_{p<e^t} \frac{1}{p} \operatorname{Im} (h(p,1)p^{-ia}))$$

$$L_2(t) = \exp(i \sum_{p<e^t} \frac{1}{p} \operatorname{Im} (h(1,p)p^{-ib}))$$

and for $j = 1,2$,

$$\lim_{t\to\infty} (L_j(\lambda t)/L_j(t)) = 1$$

for all $\lambda > 0$; the convergence is uniform in every finite interval.

<u>Lemma 9.5</u> : A real-valued additive function f on Z_2 has distribution (mod 1) if and only if, for each integer k, there exists a real number b_k such such that the limit of

$$(1/xy) \sum_{\substack{m\le x \\ n\le y}} e(k\, f(m,n))$$

is b_k as x and y tend to infinity independently. Moreover, the limiting distribution, if it exists, is continuous if and only if $N^{-1} \sum_{|k|\le N} |b_k|$ tends to zero as $N \to \infty$ and absolutely continuous with a derivative belonging to the class $L_2[o,1]$, if and only if the series $\sum_{k=-\infty}^{\infty} |b_k|^2$ converges.

Proof of this lemma is similar to the proof of Lemma 2 of Elliott (1971) and so is omitted.

5. Proofs of the main results :

Proof of Theorem 9.1 : Define a sequence $\{X_p\}$ of independent random variables satisfying, for each real number a

$$P\{X_p = a\} = \left(1 - \frac{1}{p}\right)^2 \sum_{f(p^t, p^k) = a} p^{-t-k} .$$

It is easy to see that the convergence of the series (9.3.1), (9.3.2) and (9.3.3) imply that ΣX_p converges almost everywhere, by Kolmogorov's three-series theorem.

Let a be a continuity point of the distribution function

$Q(a) = P\{\Sigma X_p < a\}$ and let $B = \{p : \max(|f(1,p)|, |f(p,1)|) \geq 1\}$

Let $\varepsilon > 0$. Choose M such that

$$\sum_{p > M} \frac{1}{p^2} + \sum_{\substack{p \varepsilon B \\ p > M}} \frac{1}{p} < \frac{\varepsilon}{12} .$$

Let

$$z(x, y, a) = \{[m,n] : m \leq x,\ n \leq y,\ f(m,n) < a\} .$$

Let

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

for any $\delta > 0$ such that a, $a-\delta$ and $a+\delta$ are continuity points of Q, $r > M$ and $x, y \geq 1$, we have

$$N\{z(x,y,a)\} \leq N\{[m,n] : m \leq x, n \leq y, f(m,n)_r < a + \tfrac{\delta}{2}\}$$

(9.5.1) $$+ N\{[m,n] \notin D : m \leq x, m \leq y, |f(m,n) - f(m,n)_r| \geq \tfrac{\delta}{2}\}$$

$$+ N\{[m,n] \in D : m \leq x, n \leq y\}.$$

As ΣX_p is convergent a.e., we can choose r_1, such that for $r > r_1$, we get $P\{|\sum_{p>r} X_p| > \frac{\delta}{2}\} < \frac{\varepsilon}{4}$.

In view of the convergence of (9.3.1) and (9.3.2), we can find $r_2 > r_1$, such that

(9.5.2) $$|A_2(x,y,f^*) - A_2(r,r,f^*)| < \frac{\delta}{4}$$

whenever $x, y, r > r_2$. By Lemma 9.1 and Chebyshev's inequality we have for $x, y, r > r_2$

$$(9.5.3)\quad N\{[m,n]\not\in D: m\le x,\ n\le y, |f(m,n)-A_2(x,y,f^*)-f(m,n)_r+A_2(r,r,f^*)| > \frac{\delta}{4}\}$$

$$< 4\delta^{-2}\, xy\ C(B_2^2(x,y,f^*) - B_2^2(r,r,f^*))$$

where C is as in Lemma 9.1. Since the series (9.3.3) is convergent, there exists $r_o > r_2$ such that the left-hand side of (9.5.3) is less than $\frac{\varepsilon}{4}$, for all $x, y, r > r_o$. Hence it follows from (9.5.2) and (9.5.3) that there exists $r_o > r_1$ such that if $r > r_o$

$$N\{[m,n] \not\in D: m\le x,\ n\le y, |f(m,n)-f(m,n)_r| > \frac{\delta}{2}\} < \frac{1}{4}\,\varepsilon xy$$

for all sufficiently large x and y.

Note that for any $r > 3$, the limit of

$$(1/xy)\ N\{[m,n] : m \le x,\ n \le y,\ f(m,n)_r < a + \frac{\delta}{2}\}$$

is $P\{\sum_{p\le r} X_p < a + \frac{\delta}{2}\}$ as x and y tend to infinity independently. Hence for all sufficiently large x and y, the left-hand side of (9.5.1) is not greater than $xyP\{\sum_p X_p < a + \delta\} + \varepsilon xy$. Similarly one can show that for all sufficiently large x and y the left-hand side of (9.5.1) is not less than

$xyP\{\sum_p X_p < a - \delta\} - \varepsilon xy$. Let θ_a be any limit point of $\{(1/xy)\, N\{z(x,y,a)\} : x, y \geq 1\}$. Since a is a continuity point of Q, it follows that $Q(a) = \theta_a$. Hence the limit of $(1/xy)\, N\{z(x,y,a)\}$ exists and is equal to $Q(a)$ as x and y tend to infinity independently. This proves the sufficiency part of the theorem.

To prove the necessity part, we first prove that there is a $c > 0$ such that

(9.5.4) $$\sum_{\max(|f(1,p)|,\,|f(p,1)|) \geq c} \frac{1}{p} < \infty.$$

If f has a distribution, we can find $M > 0$, $x_0 \geq 1$ and $y_0 \geq 1$ such that $x \geq x_0$ and $y \geq y_0$ imply that

$$N\{[m,n] : m \leq x,\ n \leq y,\ |f(m,n)| < M\} > \frac{3xy}{4}.$$

Suppose (9.5.4) does not hold for any $c > 0$. Then there exists a sequence $\{q_i\}$ of primes such that

$$q_i > x_0, \quad \sum_i \frac{1}{q_i} = \infty, \quad \sum_i \frac{1}{q_i^2} < \frac{1}{4}$$

and $\max(|f(1,q_i)|, |f(q_i,1)|) \to \infty$ as $i \to \infty$. Without loss of generality we can assume that

$$f(1, q_i) \to \infty \quad \text{as} \quad i \to \infty .$$

If we put

$$G(x,y) = \{[m,n] : m \leq x,\ n \leq y,\ |f(m,n)| < M \text{ and } q_i \nmid n \text{ for any } i\}$$

we get

$$N\{G(x,y)\} > (xy/2) \quad \text{for all} \quad x \geq x_o \quad \text{and} \quad y \geq y_o .$$

If for every $a_o \leq x_o$

$$\text{card}\{b \leq y : (a_o, b) \in G(x_o, y)\} < \frac{y}{2} ,$$

then,

$$\text{card}\{G(x_o, y)\} < \frac{x_o\ y}{2} .$$

Hence for each $y \geq y_o$ there exists $a_y \leq x_o$ such that $\text{card}\{[a_y, b] : (a_y, b) \in G(x_o, y)\} > \frac{y}{2}$. So for each $a \leq x_o$, there exists a subsequence of positive integers $\{y_{an}\}$ such that

$$\{y_{an} : a \le x_o \text{ and } n \ge 1\} \supset \{y \ge y_o\}$$

$$\{y_{an} : n \ge 1\} \cap \{y_{a'm} : m \ge 1\} = \emptyset \quad \text{if} \quad a \ne a'$$

and

$$\text{card}\{[a,b] \in G(x_o, y_{an})\} > \frac{y_{an}}{2} \quad \text{for all} \quad n \ge 1.$$

Hence $\text{card}\{B_y\} > \frac{y}{2}$ for all $y \ge y_o$ where

$$B_y = \{b \le y : (a,b) \in G(x_o, y)\} \quad \text{if} \quad y = y_{an}, \; a \le x_o, \; n \ge 1.$$

So

$$(9.5.5) \qquad \limsup_{y \to \infty} \frac{1}{\log y} \left[\sum_{\substack{(a,t) \in G(x_o, y) \\ \text{for some } a \le x_o}} \frac{1}{t} \right] \ge \frac{1}{2}.$$

On the other hand, since $\sum_{f(1,q_i) > 2M} \frac{1}{q_i} = \infty$ and $|f(a,b) - f(a,b')| < 2M$ for every $b \in B_y$, $b' \in B_y$ ($y = y_{an}$ for some $n \ge 1$) and $a_o \le x_o$ it follows from Lemma 3 of Erdos (1938) that for any $\varepsilon > 0$, there is a $y_1 > y_o$ such that $y > y_1$ implies

$$\sum_{b \varepsilon B_y} \frac{1}{b} < \varepsilon \log y.$$

So

$$\limsup_{y \to \infty} \frac{1}{\log y} \left[\sum_{\substack{(a,t) \varepsilon G(x_o, y) \\ \text{for some } a \leq x_o}} \frac{1}{t} \right] = 0, \tag{9.5.6}$$

which contradicts (9.5.5). Hence there is a $c > 0$ satisfying (9.5.4).

Let H be the distribution function of f. If H is degenerate define a new additive function g on Z_2 by

$$g(2^2, 1) = f(2^2, 1) + 1$$

$$g(p^t, p^k) = f(p^t, p^k) \quad \text{if } [p^t, p^k] \neq [2^2, 1].$$

Clearly g has a non-degenerate distribution. So without loss of generality we can assume that H is non-degenerate. It follows from (9.5.4) and Lemmas 9.2 and 1.6 that (9.3.3) is satisfied. Hence

$$\sum_p \left\{ X_p - \frac{1}{p}(f^*(p,1) + f^*(1,p)) \right\} \tag{9.5.7}$$

converges almost everywhere. Let Q be the distribution function of (9.5.7). As in the proof of the sufficient part it can be shown that $(1/xy)\, N\{z(x,y,a + A_2(x,y,f^*))\}$ tends to $Q(a)$ at each continuity point a of Q, as x and y tend to infinity independently. This and the fact f has a distribution imply that the set $\{A_2(x,y,f^*)\}$ is bounded. Hence there exist two sequences $\{x_n\}$ and $\{y_n\}$ tending to infinity as $n \to \infty$ such that

$$\lim_{n\to\infty} A_2(x_n, y_n, f^*) = b$$

for some b. Hence we conclude that $H(a+b) = Q(a)$ for all continuity points a of Q such that $a+b$ is a continuity point of H.

Hence b is the only limit point of $\{A_2(x,y,f^*)\}$. Thus the two series (9.3.1) and (9.3.2) are convergent. This completes the proof of Theorem 9.1.

<u>Proof of Theorem 9.2</u>: Choose M and $k \geq 2$ such that

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

Let

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

Clearly we have for all x and y

$$N\{[m,n] \in B : m \leq x,\ n \leq y\} < \tfrac{1}{4}\delta xy .$$

Hence for all i,

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

Let

$$L = \sum_{p \leq M} \sum_{j=0}^{k} |f(p^j, p^{k-j})| + \sum_{\substack{q \in A \\ q \leq M}} (|f(q,1)| + |f(1,q)|).$$

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 $h(m,n) = h(m,1) + h(1,m) \geq 0$ for all m,n and

$$N\{[m,n] : m \leq x_i,\ n \leq y_i,\ h(m,n) < c + L\} > \tfrac{1}{2}\delta x_i\, y_i$$

for all i.

So we have

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

and

$$\limsup_{n \to \infty} \frac{1}{n} N_n\{h(1,m) < c + 1\} > 0.$$

Since $\sum_{p \varepsilon A} \frac{1}{p} < \infty$, it follows from Theorem 3.3, that

$$\sum_p \frac{f^*(p,1)}{p} + \sum_p \frac{f^*(1,p)}{p} < \infty .$$

Theorem 9.1 now shows that f has a distribution. This completes the proof of Theorem 9.2.

<u>Proof of Theorem 9.3</u> : Since

$$8 \| \theta \|^2 \leq \sin^2 \pi \theta \leq 2\pi^2 \| \theta \|^2$$

and

$$1 - \mathrm{Re}[(e(k\ f(p,1))\ p^{-it}) = 2 \sin^2(\pi(k\ f(p,1) - (t/2\pi) \log p)),$$

we clearly have

$$\sum_p \frac{1}{p}[1 - \mathrm{Re}((e(kf(p,1))\ p^{-it})] < \infty ,$$

if and only if

$$\sum_p \frac{1}{p} \| kf(p,1) - (t/2\pi) \log p \|^2 < \infty .$$

So if for each non-zero integer k either i) or ii) is satisfied, then from Lemmas 9.4 and 9.5 it follows that f has a distribution (mod 1).

Now we prove the converse. Suppose f has a distribution (mod 1) then for each integer k the limit of

$$(9.5.8) \qquad \frac{1}{xy} \sum_{\substack{m \leq x \\ n \leq y}} e(kf(m,n))$$

exists as x and y tend to infinity independently. If the limit of (9.5.8) is zero for some non-zero integer k then either for all real t and u

$$\sum_p \frac{1}{p}(\|kf(p,1) - t \log p\|^2 + \|kf(1,p) - u \log p\|^2) = \infty,$$

or there exist real numbers a and b such that

$$(9.5.9) \quad \sum_p \frac{1}{p}(\|kf(p,1)-a \log p\|^2 + \|kf(1,p)-b \log p\|^2) < \infty$$

and $h(f,k,a,b)$ is zero.

Suppose (9.5.9) holds. Observe that the set of all integers k for which there is a t such that

$$\sum_p \frac{1}{p} \|kf(p,1) - t \log p\|^2 < \infty$$

is a group. (This can be seen by using the inequality

$\|x+y\|^2 \le 2\|x\|^2 + 2\|y\|^2$.) By this and by Lemma 9.3 there exists a non-zero integer r such that $h(f, rk, ra, rb) \neq 0$ and

$$\sum_p \frac{1}{p} (\|krf(p,1) - ra \log p\|^2 + \|krf(1,p) - rb \log p\|^2) < \infty .$$

From this and the fact that

$$(1/xy) \sum_{\substack{m \le x \\ n \le y}} e(krf(m,n))$$

tends to a limit as x and y tend to infinity independently, it follows, as in Delange (1970), that $a = 0$ and $b = 0$. So (ii) holds by Lemma 9.4. If the limit of (9.5.8) is non-zero, then (9.5.9) holds for some a and b. Again, as in Delange (1970), it follows that $a = b = 0$ and the two series

$$\sum_p \frac{1}{p} \operatorname{Im}[e(kf(p,1))]$$

and

$$\sum_p \frac{1}{p} \operatorname{Im}[e(kf(1,p))]$$

converge, which gives the convergence of all the series in (i) and (ii). This completes the proof of Theorem 9.3.

Theorem 9.4 follows easily from Lemma 9.5 as in Elliott (1971).

<u>Proof of Theorem 9.5</u> : Let ϕ be the characteristic function of H. We have for all real t

$$e^{-it\alpha_n}\left[\frac{1}{n^2}\sum_{\substack{m\leq n\\ m'\leq n}} \exp[itf(m,m')]\right] \to \phi(t)$$

as $n \to \infty$. Since $\phi(0) = 1$ and ϕ is continuous at zero, there exists a $\delta > 0$ such that for all $|t| < \delta$

$$\frac{1}{n^2}\sum_{\substack{m\leq n\\ m'\leq n}} \exp[itf(m,m')] \not\to 0$$

as $n \to \infty$. By Lemma 9.4, for all $|t| < \delta$, there exist real numbers $a(t)$, $b(t)$ such that

$$\sum_p \frac{1}{p}[1 - \mathrm{Re}(\exp(itf(p,1))\, p^{-ia(t)}] < \infty$$

and

$$\sum_p \frac{1}{p}[1 - \mathrm{Re}(\exp(itf(1,p))\, p^{-ib(t)}] < \infty .$$

Hence by Lemma 7.7, there exist real numbers a, b and additive arithmetic functions g_1 and g_2 such that

$$\sum_p \frac{[g_i'(p)]^2}{p} < \infty \quad \text{for } i = 1,2$$

for $i = 1,2$ and for all $m,n \geq 1$,

$$f(m,1) = a \log m + g_1(m)$$

$$f(1,n) = b \log n + g_2(n).$$

Now we define an additive function $g(m,n)$ by

$$g(m,n) = g_1(m) + g_2(n) + [f(m,n) - f(m,1) - f(1,n)].$$

Clearly,

$$f(m,n) = a \log m + b \log n + g(m,n)$$

and

$$\sum_p \frac{(g^*(p,1))^2}{p} + \sum_p \frac{(g^*(1,p))^2}{p} = \sum_p \frac{(g_1'(p))^2}{p} + \sum_p \frac{(g_2'(p))^2}{p} < \infty .$$

This establishes the first half of the theorem.

Conversely, suppose $f(m,n) = a \log m + b \log n + g(m,n)$ with $g(m,n)$ satisfying (9.3.5). From the proof of Theorem 9.1 it follows that there is a distribution function Q such that for each of its continuity point c,

$$(1/xy)\, N\{(m,n) : m \le x,\ n \le y : g(m,n) - \alpha_{m,n} < c\}$$

tends to $Q(c)$ as x and y tend to infinity independently. Let ϕ be the characteristic function of Q. Let

$$\rho\,(x,y,t) = (1/xy)\Big(\sum_{\substack{m\le x\\ n\le y}} \exp(it\,g(m,n) - \alpha_{m,n})\Big) - \phi(t).$$

Note that $\rho\,(x,y,t) \longrightarrow 0$, uniformly in every finite interval, as x and y tend to infinity independently. Summing by parts, we obtain

$$\sum_{\substack{m\le x\\ n\le y}} \exp(it[f(m,n) - a \log x - b \log y - \alpha_{m,n}])$$

$$= \sum_{\substack{m\le x\\ n\le y}} \exp(it(g(m,n) - \alpha_{m,n}))$$

$$- \frac{ita}{x^{ita}} \int_1^x u^{ita-1}\Big[\sum_{\substack{m\le u\\ n\le y}} \exp[it(g(m,n) - \alpha_{m,m})]\Big]\, du$$

$$- \frac{itb}{y^{itb}} \int_1^y v^{ita-1}\Big[\sum_{\substack{m\le x\\ n\le v}} \exp[it(g(m,n) - \alpha_{m,n})]\Big]\, dv$$

$$+ \frac{(it)^2 ab}{x^{ita}\ y^{itb}} \int_1^y v^{itb-1}\Big[\int_1^x u^{itb-1}\Big[\sum_{\substack{m\le u\\ n\le v}} \exp(it(g(m,n) - \alpha_{m,n}))\Big] du\Big] dv$$

$$= xy(\rho(x,y,t) + \phi(t)) - \frac{ita}{x^{ita}} \int_1^x y u^{ita} (\rho(u,y,t) + \phi(t))\, du$$

$$- \frac{itb}{y^{itb}} \int_1^y x v^{ita} (\phi(t) + \rho(x,v,t))dv$$

$$- \frac{t^2 ab}{x^{ita}\, y^{itb}} \int_1^y v^{itb} \left(\int_1^x u^{ib}(\phi(t) + \rho(u,v,t))du\right) dv$$

$$= xy\, \phi(t) \left(1 - \frac{ita}{1+ita} - \frac{itb}{1+itb} - \frac{t^2 ab}{(1+ita)(1+itb)}\right) + o(xy)$$

$$= \frac{xy\phi(t)}{(1+ita)(1+itb)} + o(xy)$$

as x and y tend to infinity independently, where the estimate $o(xy)$ is uniform for t in every finite interval.

Note that, in view of (9.3.5) we have, for any fixed ε, $0 < \varepsilon < 1$

$$\left(\sum_{m \leq p \leq x} \frac{g^*(p,1)}{p} \right)^2 \leq \left(\sum_{\varepsilon x < p \leq x} \frac{1}{p} \right) \left(\sum_{\varepsilon x < p \leq x} \frac{1}{p}(g^*(p,1))^2\right)$$

$$= (-\log \varepsilon + o(1)).\, o(1)$$

$$= o(1)$$

as $x \to \infty$ uniformly in m such that $\varepsilon x < m \le x$. So for any fixed $\varepsilon > 0$, $0 < \varepsilon < 1$, we have

(9.5.10) $$\sum_{p \le m} \frac{1}{p} g^*(p,1) = \sum_{p \le x} \frac{1}{p} g^*(p,1) + o(1)$$

uniformly in m such that $\varepsilon x < m < x$, as $x \to \infty$. Similarly for any fixed $\varepsilon > 0$, $0 < \varepsilon < 1$, we have

(9.5.11) $$\sum_{p \le m} \frac{1}{p} g^*(1,p) = \sum_{p \le x} \frac{1}{p} g^*(1,p) + o(1)$$

uniformly in m such that $\varepsilon y < m < y$, as $y \to \infty$.

By (9.5.10) and (9.5.11), for any fixed ε, $0 < \varepsilon < 1$ and for t in any finite interval

$$\Big| \sum_{\substack{m \le x \\ n \le y}} [\exp(it(f(m,n) - \alpha'_{x,y})) - \exp(it(f(m,n) - a \log x - b \log y - \alpha_{m,n}))] \Big|$$

$$\le 2\varepsilon^2 xy + \sum_{\substack{\varepsilon x < m \le x \\ \varepsilon y < n \le y}} |e^{it(\alpha_{m,n} - \alpha_{x,y})} - 1|$$

$$\le xy(2\varepsilon^2 + o(1)\,(1 - \varepsilon)^2) \le 3xy\,\varepsilon^2$$

for all sufficiently large x and y, uniformly in t in any finite interval. From this fact, we have

$$\sum_{\substack{m \le x \\ n \le y}} \exp(it(f(m,n) - \alpha'_{x,y})) = \frac{xy\phi(t)}{(1+ita)(1+itb)} + o(xy)$$

as x and y tend to infinity independently, where the estimate $o(xy)$ is uniform for t in any finite interval.

If G is the distribution function corresponding to the characteristic function $\phi(t)/((1+iat)(1+ibt))$ then at each continuity point c of G, we have the limit of

$$(1/xy)\, N\{[m,n] : m \le x,\ n \le y,\ f(m,n) - \alpha'_{x,y} < c\}$$

is $G(c)$ as x and y tend to infinity independently.

Suppose $\{a_{x,y}\}$ and $\{b_{x,y}\}$ are two sets of real numbers. Suppose G and H are two distribution functions such that

$$\frac{1}{xy} N\{[m,n] : m \le x,\ n \le y \text{ and } f(m,n) - a_{x,y} < c_1\}$$

tends to $G(c_1)$ and

$$\frac{1}{xy} N\{[m,n] : m \le x,\ n \le y \text{ and } f(m,n) - b_{x,y} < c_2\}$$

tends to $H(c_2)$ as x and y tend to infinity independently, at all continuity points c_1 and c_2 of G and H respectively. As in the proof of Lemma 1.6 it follows that $a_{x,y} - b_{x,y}$ tends to a limit as x and y tend to infinity independently. This completes the proof of Theorem 9.5.

Proof of Theorem 9.6 : We give only the outline of the proof, as the detailed proof leads to repetition of argument in Chapter 5. As in Chapter 5 it is possible to construct the random variables $\{X_p, Y_{1p}, Y_{2p}\}$ defined on the same probability space such that

1) $\{Y_{1p}, Y_{2q} : p, q \text{ primes}\}$ is a set of mutually independent random variables,

2) $\sum_p P\{X_p \neq Y_{1p} + Y_{2p}\} < \infty$,

3) the characteristic function of X_p is

$$(1 - \frac{1}{p})^2 [\sum_{j,k \geq 0} (p^{-j-k} \exp(it\, f(p^j, p^k)))]$$

and

4) the characteristic function of Y_{jp} is

$$(1 - \frac{1}{p}) [1 + \sum_{k \geq 1} \frac{1}{p^k} \exp(it\, k\, f_j(p))] \; ; \; j = 1,2.$$

It follows from Theorem 9.1 that the distribution of f is same as the distribution of $\sum_p X_p$ and the distribution of f_j is same as that of $\sum_p Y_{jp}$, $j = 1, 2$. Now the theorem follows from Lemma 5.2 .

CHAPTER 10

INTEGRAL LIMIT THEOREMS FOR ADDITIVE FUNCTIONS ON THE SET OF PAIRS OF POSITIVE INTEGERS

1. Introduction : One of the main problems of the probabilistic number theory is to find conditions an additive arithmetic function f must satisfy if there is to exist a distribution function H such that at each of its continuity points

$$\frac{1}{n} N_n \{f(m) < x\, C_n + K_n\} \rightarrow H(x)$$

as $n \rightarrow \infty$ where K_n and C_n are suitably chosen sequences connected with the function f.

This problem has been solved, for several classes of functions, by several authers. (See Kubilius, 1964 and Levin and Fainleib, 1968.) In this chapter we generalise some of the results of Levin and Fainleib(1968) to additive functions on Z_2.

2. Notations and definitions : Throughout this chapter we let f denote a real-valued additive function satisfying

$$\sum_p \frac{1}{p} [(f^*(p,1))^2 + (f^*(1,p))^2] = \infty .$$

A subset E of Z_2 is said to have density $\delta^*(E)$ in the weak sense if the limit of

$$(1/xy)\ N\{[m,n] \in E : m \leq x,\ n \leq y\}$$

is $\delta^*(E)$ as x and y tend to infinity with any fixed ratio.

A real-valued arithmetic function ϕ is said to be measurable (see Levin and Fainleib, 1968) if there exists a distribution function G such that at each of its continuity points

$$\frac{1}{n} \log n \sum_{\substack{p \leq n \\ \phi(p) \leq x}} 1 \rightarrow G(x) \qquad \text{as} \quad n \rightarrow \infty .$$

A real-valued function $\phi(m,n)$ on Z_2 is said to be weakly measurable if there is a distribution function G such that at each of its continuity points

$$\frac{\log n}{n} \sum_{\substack{p \leq n \\ \phi(p,1)+\phi(1,p)<x}} 1 \rightarrow G(x) \qquad \text{as} \quad n \rightarrow \infty .$$

A real-valued function $\phi(m,n)$ on Z_2 is said to be measurable if both $\phi(m,1)$ and $\phi(1,n)$ are measurable.

Let $\phi(m,n)$ be a weakly measurable function such that $\phi(p^j,p^k) = f(p^j,p^k)$ if $j + k > 1$. We call a normalization of $\{C_n\}$ weakly ϕ-admissible for f if there exists $r = r(n) \to \infty$ such that

$$\frac{\log r(n)}{\log n} \longrightarrow 0$$

and

$$C_n^{-2}\Big[\sum_{r(n)<p\leq n} \frac{1}{p}\big((f(p,1) - \phi(p,1))^2 + (f(1,p) - \phi(1,p))^2\big)\Big] \to 0.$$

as $n \to \infty$.

Let $\phi(m,n)$ be a measurable function such that $\phi(p^j,p^k) = f(p^j,p^k)$ if $j + k > 1$. We call a normalization of $\{C_{x,y}\}$ ϕ-admissible for f if there exist $r_1 = r_1(x) \to \infty$ and $r_2 = r_2(y) \to \infty$ such that

$$\frac{\log r_1(x)}{\log x} \longrightarrow 0, \quad \frac{\log r_2(y)}{\log y} \longrightarrow 0$$

and

$$C_{x,y}^{-2}\Big[\sum_{r_1(x)<p\leq x} \frac{1}{p}(f(p,1)-\phi(p,1))^2 + \sum_{r_2(y)<p\leq y} \frac{1}{p}(f(1,p)-\phi(1,p))^2\Big] \to 0$$

as x and y tend to infinity independently.

3. Main results :

Theorem 10.1 : Let $\phi(m,n)$ be a measurable function such that $\phi(p^j,p^k) = f(p^j,p^k)$ if $j + k > 1$. Then there exist constants $\{K_{x,y}\}$ and a distribution function H such that at each of its continuity points the limit

$$(10.3.1) \quad \frac{1}{xy} N\{[m,n] : m \leq x,\ n \leq y,\ f(m,n) < K_{x,y} + a\, C_{x,y}\}$$

is $H(a)$ as x and y tend to infinity independently with an arbitrary ϕ-admissible normalization $\{C_{x,y}\}$ if and only if there exists a non-decreasing function K, $0 = K(-\infty) = \lim_{u\to -\infty} k(u) < \lim_{u\to\infty} K(u) = K(\infty) < \infty$, such that when $B_{x,y} > 0$ is defined by the equation

$$(10.3.2) \quad \sum_{p\leq x} \frac{1}{p} \frac{f^2(p,1)}{B^2_{x,y} + f^2(p,1)} + \sum_{p\leq y} \frac{1}{p} \frac{f^2(1,p)}{B^2_{x,y} + f^2(1,p)} = K(\infty),$$

we have for all $u \neq 0$, the limit of

$$(10.3.3) \quad \sum_{\substack{p\leq x \\ f(p,1)<uB_{x,y}}} \frac{1}{p} \frac{f^2(p,1)}{B^2_{x,y}+f^2(p,1)} + \sum_{\substack{p\leq y \\ f(1,p)<uB_{x,y}}} \frac{1}{p} \frac{f^2(1,p)}{B^2_{x,y} + f^2(1,p)}$$

is $K(u)$ as x and y tend to infinity independently and the normalization $C_{x,y} = B_{r_1(x), r_2(y)}$ is ϕ-admissible.

When these conditions hold we can take $C_{x,y} = B_{r_1(x), r_2(y)}$. The logarithm of the characteristic function of the limit law is equal to

$$(10.3.4) \qquad \int_{-\infty}^{\infty} \left(e^{itu} - 1 - \frac{itu}{1+u^2}\right) \frac{1+u^2}{u^2} \, dK(u).$$

<u>Theorem 10.2</u>: Let $\phi(m,n)$ be a weakly measurable function such that $\phi(p^j, p^k) = f(p^j, p^k)$ if $j+k > 1$. Then there exist constants $\{K_x\}$ and a distribution function H such that each of its continuity points the limit

$$\frac{1}{xy} N\{[m,n] : m \leq x, \; n \leq y, \; f(m,n) < K_x + a\, C_x\}$$

is $H(a)$ as x and y tend to infinity in a fixed ratio, with an arbitrary weakly ϕ-admissible normalization $\{C_x\}$ if and only if there exists a non-decreasing function K, $0 = K(-\infty) = \lim_{u \to -\infty} K(u) < \lim_{u \to \infty} K(u) = K(\infty) < \infty$, such that when $B_n > 0$ is defined by the formula

$$\sum_{p \leq n} \frac{1}{p} \left(\frac{f^2(p,1)}{B_n^2 + f^2(p,1)} + \frac{f^2(1,p)}{B_n^2 + f^2(1,p)} \right) = K(\infty),$$

we have for all $u \neq 0$

$$K(u) = \lim_{n \to \infty} \left(\sum_{\substack{p \leq n \\ f(p,1) < uB_n}} \frac{1}{p} \frac{f^2(p,1)}{B_n^2 + f^2(p,1)} + \sum_{\substack{p \leq n \\ f(1,p) < uB_n}} \frac{1}{p} \frac{f^2(1,p)}{B_n^2 + f^2(1,p)} \right)$$

and the normalization $C_x = B_{r(x)}$ is weakly ϕ-admissible. When these conditions hold, we can take $C_x = B_{r(x)}$, and

$$K_x = \sum_{\substack{p \leq r(x) \\ |f(p,1)| < C_x}} \frac{f(p,1)}{p} + \sum_{\substack{p \leq r(x) \\ |f(1,p)| < C_x}} \frac{f(1,p)}{p}$$

$$+ \sum_{r(x) < p < x} \frac{1}{p} (f(p,1) + f(1,p) - \phi(1,p) - \phi(p,1))$$

$$+ C_x \left(\int_{|u| \geq 1} \frac{1}{u} dK(u) - \int_{|u| < 1} u dk(u) \right).$$

The logarithm of the characteristic function of the limit law is equal to (10.3.4).

4. Preliminary results :

Lemma 10.1 (Delange, 1969) : Let E be a set of pairs of positive integers. Then the following two conditions are equivalent :

(10.4.1) Limit of $\frac{1}{x^2} N\{[m,n] \varepsilon E : m \leq x, n \leq x\}$ is δ as $x \to \infty$.

(10.4.2) Limit of $\frac{1}{xy} N\{[m,n] \varepsilon E : m \leq x, n \leq y\}$ is δ as x and y tend to infinity with a fixed ratio.

Lemma 10.2 (Levin and Fainleib, 1968) : Let $g(m)$ be a real-valued additive arithmetic function. Let $\phi(m)$ be a measurable function such that $\phi(p^k) = g(p^k)$ for $k > 1$ and let $\{C_n\}$ be a sequence of positive real numbers tending to infinity as $n \to \infty$. Then

$$\frac{1}{n} \sum_{m=1}^{n} \exp\left(it \frac{g(m) - K_n}{C_n}\right)$$

$$= \prod_{p \leq r} \left(1 + \frac{e^{itg(p)/C_n} - 1}{p}\right)$$

$$\times \exp\left(\frac{it}{C_n}\left(A_n^*(g) - A_r^*(g) - A_n^*(\phi) + A_r^*(\phi) - K_n\right)\right)$$

$$+ O\left(\frac{1}{C_n}\left(\sum_{r<p\leq n} \frac{(g(p)-\phi(p))^2}{p}\right)^{1/2}\right) + o(1)$$

uniformly in t in every finite interval, where $A_n(g) = \sum_{p \leq n} \frac{g(p)}{p}$.

Lemma 10.3 : Let E be a set of pairs of prime powers such that

$$(10.4.3) \qquad \sum_{[p^\beta, p^\alpha] \varepsilon E} p^{-\alpha-\beta} < \infty .$$

Let $f(m,n)$ and $g(m,n)$ be two additive functions on Z_2 such that $f(p^\alpha, p^\beta) = g(p^\alpha, p^\beta)$ for all $[p^\alpha, p^\beta] \not\varepsilon E$. Then for each $\varepsilon > 0$ there exists $K = K(\varepsilon)$ such that for all x and y

$$N\{[m,n] : m \leq x, \ n \leq y, \ |f(m,n) - g(m,n)| > K\} < \varepsilon xy .$$

Proof : In view of (10.4.3), there exists an $M = M(\varepsilon)$ such that

$$\sum_{\substack{[p^\alpha, p^\beta] \varepsilon E \\ p^{\alpha+\beta} > M}} p^{-\alpha-\beta} < \varepsilon .$$

Put

$$K = \sum_{p^{\alpha+\beta} \leq M} |f(p^\alpha, p^\beta) - g(p^\alpha, p^\beta)| .$$

Let

$$A = \{[m,n] : p^\alpha \| m \text{ and } p^\beta \| n \text{ for some } [p^\alpha, p^\beta] \varepsilon E \text{ and } p^{\alpha+\beta} > M\} .$$

Clearly,

$$E(A) < xy \sum_{p^{\alpha+\beta} > M} p^{-\alpha-\beta} < \varepsilon\, x\, y.$$

To complete the proof of the lemma, we need to show that for all $[m,n] \notin A$

$$|f(m, n) - g(m, n)| \leq K.$$

Note that

$$f(m,n) - g(m,n) = \sum_{\substack{p^{\alpha} \| m \\ p^{\beta} \| n}} (f(p^{\alpha}, p^{\beta}) - g(p^{\alpha}, p^{\beta})).$$

If $[m,n] \notin A$, then

$$|f(m,n) - g(m,n)| = \Big| \sum_{\substack{p^{\alpha} \| m,\, p^{\beta} \| n \\ [p^{\alpha}, p^{\beta}] \varepsilon E \\ p^{\alpha+\beta} \leq M}} (f(p^{\alpha}, p^{\beta}) - g(p^{\alpha}, p^{\beta}))\Big|$$

$$\leq K.$$

This proves the lemma.

5. Proofs of the main results :

Proof of Theorem 10.1 : We give only an outline of the proof, as the proof is similar to the proof of Theorem 3 of Levin and Fainleib (1968).

In view of Lemma 10.3 it is enough to consider functions of the type $f(m,n) = f(m,1) + f(1,n)$. By Lemma 10.2 we have

$$\frac{1}{xy} \sum_{\substack{m \le x \\ n \le y}} \operatorname{esp}\left[it \frac{f(m,n) - K_{x,y}}{C_{x,y}}\right]$$

$$= \exp(itK'_{x,y}/C_{x,y}) \prod_{p \le r_1(x)} \left(1 + \frac{1}{p}(\exp(itf(p,1)/C_{x,y}) - 1)\right)$$

$$\times \prod_{p \le r_2(y)} \left(1 + \frac{1}{p}(\exp(itf(1,p)/C_{x,y}) - 1)\right)$$

$$+ O\left(\frac{1}{C_{x,y}}\left(\sum_{r_1(x) < p \le x} \frac{1}{p}(f(p,1) - \phi(p,1))^2 + \sum_{r_2(y) < p \le y} \frac{1}{p}(f(1,p) - \phi(1,p))^2\right)\right)$$

uniformly in t when t varies over any finite interval, where

$$K'_{x,y} = -K_{x,y} + \sum_{r_1(x) < p \le x} \frac{1}{p}(f(p,1) - \phi(p,1)) + \sum_{r_2(y) < p \le y} \frac{1}{p}(f(1,p) - \phi(1,p)).$$

The rest of the proof is similar to the proof of Theorem 3 of Levin and Fainleib (1968). We only have to note the following fact and use Lemma 7 of Levin and Fainleib (1968).

<u>Fact</u>: If $\prec$ is a binary relation on $D = \{(m,n) : m \geq 1,\ n \geq 1\}$ defined by

$$(m,n) \prec (m',n') \text{ if } m \leq m' \text{ and } n \leq n'$$

then clearly $(D, \prec)$ is a directed set. Let X be a topological space. Let $\{x_\beta : \beta \in D\}$ be a net in X (see Kelley, 1955) and $x \in X$. Then $\{x_\beta : \beta \in D\}$ converges to x if and only if every subnet of $\{x_\beta : \beta \in D\}$ which is a sequence converges to x.

Proof of Theorem 10.2 is similar to the proof of Theorem 10.1. Here we note only that one has to use Lemma 10.1.

R E F E R E N C E S

BAKSTYS, A. (1969) : On the distribution of values of multiplicative arithmetic functions. Soviet Math. Dokl., vol. 10, no. 4.

BILLINGSLEY, P. (1968) : Convergence of probability measure, Wiley, New York.

DELANGE, H. (1961) : Sur les fonctions arithmetiques multiplicatives. Ann. Scient. Ecole Norm. Sup., vol. 78.

DELANGE, H. (1963) : On a class of multiplicative arithmetic functions. Scripta Mathematica, vol. 26 .

DELANGE, H. (1969) : On some sets of pairs of positive integers. J. Number Theory, vol. 1.

DELANGE, H. (1970) : Sur les fonctions multiplicatives de plusieurs entiers. L'Enseignement mathematique, T.24, face. 3-4.

ELLIOTT, P.D.T.A.(1971) : On the limiting distribution of additive functions (mod 1). Pacific J. Math., vol.38.

ELLIOTT, P.D.T.A. and RYAVEC, C.(1971): The distribution of the values of additive arithmetical functions. Acta mathematica, vol.126.

ERDOS, P. (1938) : On the density of some sequences of numbers III. J. Lond. Math. Soc., vol. 13.

ERDOS, P. (1939) : On the smoothness of the asymptotic distribution of additive arithmetical functions. Amer. J. Math., vol. 61.

ERDOS, P. (1947) : Some remarks and corrections to one of my papers. Bull. Amer. Math. Soc., vol. 53.

ERDOS, P. and SCHINZEL, A. (1961) : Distribution of the values of some arithmetical functions. Acta Arithmetica, vol. 6.

FELLER, W. (1966) : An Introduction to Probability Theory and its Applications, vol. II, Wiley, New York.

FELLER, W. (1969) : Limit theorems for probabilities of large deviations. Z. Wahr. Verw. Geb., vol. 14, no. 1.

GALAMBOS, J. (1971a) : Distribution of additive and multiplicative functions. The theory of arithmetic functions, Lecture notes in mathematics series, no. 251, Springer-Verlag.

GALAMBOS, J. (1971b) : On the distribution of strongly multiplicative functions. Bull. Lond. Math. Soc., vol. 3.

HALASZ, G. (1968) : Über die Mittelwerte Multiplikativer Zahlentheoretischer Functionen. Acta Math. Acad. Sci. Hung., T. 19.

HALBERSTAM, H. (1956) : On the distribution of additive number theoretical functions II. J.Lond.Math.Soc., vol.31.

HALMOS, P.R. (1962) : Measure Theory, D. Van Nostrand, Princeton, New Jersey.

HARTMAN, P. and WINTNER, A. (1942) : On the infinitesimal generators of integral convolutions. Amer.J.Math., vol. 64.

JESSEN, B. and WINTNER, A. (1935) : Distribution functions and the Riemann Zeta function. Transactions. Amer. Math. Soc., vol. 38.

JOGESH BABU, G. (1972a) : On the distribution of additive arithmetical functions of integral polynomials, Sankhyā, Series A, vol. 34, no.4.

JOGESH BABU, G. (1972b) : Some results on the distribution of additive arithmetic functions I. Submitted for publication.

JOGESH BABU, G. (1972c) : Some results on the distribution of additive arithmetic functions II. To appear in Acta Arithmetica, vol. 23, no.4.

JOGESH BABU, G. (1972d) : Some results on the distribution of additive arithmetic functions III. To apear in Acta Arithmetica, vol. 25, no.1.

JOGESH BABU, G. (1972e) : A note on invariance principle for additive functions. To appear in Sankhyā, Series A.

JOGESH BABU, G. (1972f) : Some results on the distribution of the values of multiplicative functions. Communicated by Professor J. Kubilius to Lietuvos matematikos rinkinys.

JOGESH BABU, G. (1972g) : On the characteristic function of the distribution of the values of additive arithmetic functions. Communicated by Professor J. Kubilius to Lietuvos matematikos rinkinys.

JOGESH BABU, G. (1972h) : Some results on the distribution of values of additive functions on the set of pairs of positive integers I. Submitted for publication.

JOGESH BABU, G. (1972i) : Some results on the distribution of values of additive functions on the set of pairs of positive integers II. Submitted for publication.

JOGESH BABU, G. (1972j) : On the class of distributions of values of multiplicative functions. Submitted for publication.

JOGESH BABU, G. (1972k) : On the distribution of additive functions (Mod 1) on the set of pairs of positive integers. Submitted for publication.

KATAI, I. (1969) : On the distribution of arithmetical functions. Acta. Math. Acad. Sci. Hung., T.20.

KELLEY, J.L. (1955) : General Topology, D. Van Nostrand, Princeton, New Jersey.

KUBILIUS, J. (1964) : Probabilistic Methods in the Theory of Numbers, vol. 11, Transl. Math. Mono., Amer.Math.Soc.

LEVIN, B.V. and FAINLEIB, A.S. (1968) : Integral limit theorems for certain classes of additive arithmetic functions. Trans.Moscow Math. Soc., vol.18.

LUKACS, E. (1970): Characteristic Functions, Second Edition, Griffin, London.

MACKEY, G.W. (1967) : Lectures on the Theory of Functions of a Complex Variable, D. Van Nostrand, Princeton, New Jersey.

NAGELL, T. (1951) : Introduction to Number Theory, Wiley, New York.

NOVOSELOV, E.V. (1966) : A new method in probabilistic number theory. Transl. Amer. Math. Soc., Series 2, vol. 52.

PARTHASARATHY, K.R. (1967) : Probability Measures on Metric Spaces, Academic Press, New York.

PAUL, E.M. (1962a) : Density in the light of probability theory I. Sankhyā, Series A, vol. 24.

PAUL, E.M. (1962b) : Density in the light of probability theory II. Sankhyā, Series A, vol. 24.

PAUL, E.M. (1963) : Density in the light of probability theory III. Sankhyā, Series A, vol. 25.

PAUL, E.M. (1967) : Some properties of additive arithmetic functions. Sankhyā, Series A, vol. 29.

PHILIPP, W. (1972) : An invariance principle of additive functions. To appear.

RUBIN, H. and SETHURAMAN, J. (1965) : Probabilities of moderate deviations. Sankhya, Series, A, vol. 27 .

TANAKA, M. (1955) : On the number of prime factors of integers. Japan J. Math., T. 25.

WIRSING, E. (1970) : A characterization of log n as an additive arithmetic function. Istituto Nazionale Di Alta Mathematica Symposia Mathematica, vol. 4.

ZOLOTAREV, V.M. (1962) : On a general theory of multiplication of independent random variables. Soviet Math. Dokl., vol.3.

ZOLOTAREV, V.M. (1967) : On the M-divisibility of stable laws. Theory of Probability and its applications, vol. 12.

A LIST OF SYMBOLS AND THEIR MEANING

$c, c_1, c_2 \ldots$ denote constants.

p, q with or without suffixes, denote prime numbers.

m, n denote positive integers.

(m,n) is the greatest common factor of m and n.

j, k, r, s denote non-negative integers.

$\underline{P}$ is 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 the square of any irreducible polynomial.

D_F is the degree of the polynomial F.

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

$\omega(m)$ is the number of distinct prime divisers.

$N_n\{\ldots\}$ is the number of positive integers $m \leq n$ having the property indicated in $\{\ldots\}$.

For any set A of natural numbers

$$\overline{D}(A) = \limsup_{n \to \infty} \frac{1}{n} N_n \{m \in A\},$$

$$\underline{D}(A) = \liminf_{n \to \infty} \frac{1}{n} N_n \{m \in A\}$$

and $D(A)$ is the common value of $\bar{D}(A)$ and $\underline{D}(A)$, whenever they coincide

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

$n|m$ means n divides m and $n \nmid m$ means n does not divide m. $p^k \| m$ means $p^k | m$ but $p^{k+1} \nmid m$.

$$\delta_p(k) = \begin{cases} 1 & \text{if } p|k, \\ 0 & \text{otherwise} \end{cases}$$

$\mathfrak{S}$ is the space constructed by Novoselov (1966).

$N_k = (k - s(n) + n + 2)\, n!$ if $s(n) \leq k < s(n+1)$,
where $s(n) = 1+2^2+ \dots + n^2$.

$R_k(x)$ is the smallest non-negative residue of $x (\mathrm{mod}\ N_k)$.

$\mathfrak{Z}$ is the class of all complex-valued functions f on positive integers, which admit an extension to $\mathfrak{S}$ and satisfy $f(R_k(x)) \xrightarrow{P} f(x)$ as $k \longrightarrow \infty$, where $\xrightarrow{P}$ denotes the convergence in P-measure.

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

[c]

$$f_p(x) = f_p(x,F) = \sum_{k=1}^{\infty} f(p^k)\ w(F,p^k,x).$$

$$A(v,n) = A(v,n,f,F) = \sum_{v\le p\le n} \frac{1}{p} f'(p)\ r(F,p).$$

$$B(v,n) = B(v,n,f,F) = [\sum_{v\le p\le n} \frac{1}{p} (f'(p))^2\ r(F,p)]^{1/2} .$$

$$A(n) = A(n,f,F) = A(1,n,f,F).$$

$$B(n) = B(n,f,F) = B(1,n,f,F).$$

$$B^2(N,n) = \sum_{p\le n} (f_N^2\ (p)/p).$$

$$B(N) = B(N,N).$$

$$f_N^{(k)}(m) = \sum_{\substack{p\le k\\ p|m}} f_N(p).$$

$$D^2(n) = \sum_{p\le n} \frac{f^2(p)}{p}.$$

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

$$A_n = \sum_{p\le n} \frac{f(p)}{p}.$$

$$D_n^2 = \sum_{p^k\le n} \frac{f^2(p^k)}{p^k}.$$

$L(X)$ is the distribution of a random variable X.

For any real number x, $\{x\}$ is the fractional part of x, $\| x \| = \min(\{x\}, 1 - \{x\})$,

$$x^* = \begin{cases} x & \text{if } |x| < 1, \\ 1 & \text{otherwise,} \end{cases}$$

and

$$\operatorname{sgn}(x) = \begin{cases} -1 & \text{if } x < 0, \\ 0 & \text{if } x = 0, \\ 1 & \text{if } x > 0. \end{cases}$$

$\underline{A}$ = Set of all real-valued additive arithmetic functions f with $f(2^k) = kf(2)$ for all $k \geq 1$.

$\underline{M}$ = Set of all real-valued multiplicative arithmetic functions g with $g(2^k) = (g(2))^k$ for all $k \geq 1$.

Z_2 is the set of pairs $[m,n]$ of positive integers.

For any subset B of Z_2, $N(B)$ denotes the cardinality of the set B.

$$e(t) = \exp(2\pi i t).$$

$$f_p(m,n) = f(p^{\alpha(p,m)}, p^{\alpha(p,n)}),$$

where

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

$$f(m,n)_r = \sum_{p<r} f_p(m,n).$$

$$A_2(x,y,f) = \sum_{p \leq x} \frac{1}{p} f(p,1) + \sum_{p \leq y} \frac{1}{p} f(1,p).$$

$$B_2(x,y,f) = \Big[\sum_{p \leq x} \frac{1}{p} f^2(p,1) + \sum_{p \leq y} \frac{1}{p} f^2(1,p)\Big]^{1/2}$$

$$A_2(x) = A_2\,(x,x,f).$$

$$B_2(x) = B_2\,(x,x,f).$$

For any additive function f, f^+ is the additive function defined by

$$f^+(p^t, p^k) = \begin{cases} f(p,1) & \text{if } t > 0 \text{ and } k = 0, \\ f(1,p) & \text{if } t = 0 \text{ and } k > 0, \\ f(1,p)+f(p,1) & \text{if } t > 0 \text{ and } k > 0. \end{cases}$$

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

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