A STUDY OF THE MOMENTS OF ADDITIVE FUNCTIONS USING LAPLACE TRANSFORMS AND SIEVE METHODS

Krishnaswami Alladi*

Dedicated to Paul Erdos for his seventieth birthday

§0. Introduction

In this paper we obtain upper bounds, and in some cases even asymptotic estimates, for the moments of additive functions f(n), where n belongs to a subset S of the positive integers $\mathbb{Z}^+$ satisfying some properties to be specified later. We also discuss certain consequences of these estimates. The theorems stated in §2 extend various classical results to such subsets; nevertheless the main point here is the method we employ, which is new. Our method stems from a recent technique due to Elliott [6] who obtained uniform upper bounds for the moments of arbitrary additive functions in the case $S = \mathbb{Z}^+$. We will show that by employing the combinatorial sieve and the bilateral Laplace transform along with some of Elliott's ideas one obtains a substantially improved method. For a class of sets S we can derive similar upper bounds for the absolute moments of all complex valued additive functions. In addition we can evaluate the moments asymptotically provided the f(n) satisfy some conditions. From these asymptotic estimates we get information concerning the distribution function of such f(n), for $n \in S$.

There is a vast literature on the distribution of additive functions and a variety of methods available (see Elliott [5], Vols. I and II). One approach, which is due to Halberstam[13: I, II, III], makes use of the method of moments to determine the limiting distribution of certain additive functions. Despite the difficult calculations it involved, this method had the attraction of being elementary and so capable of wide application. Subsequently this method underwent simplification and refinement by Delange [3], [4].

Being based upon the combinatorial sieve, which is an elementary tool, our method retains the applicability of Halberstam's approach, but without the latter's complications, since our use of the bilateral Laplace transform introduces the necessary simplification.

Supported by National Science Foundation Grant MCS77-18723-AO4 at the Institute for Advanced Study, Princeton, while on leave of absence from The Institute of Mathematical Sciences, Madras, during 1981-82.

Also, uniform upper bounds like those of Elliott [6], can be derived quite easily.

Classical results, some of which are surveyed in §3, mostly relate to the situation $S = \mathbb{Z}^+$. Particular subsets such as

$$S_1 = \{Q(n) \mid n \in \mathbb{Z}^+\}, \quad \text{where } Q(x) \in \mathbf{Z}^+[x]$$

and

$$S_2 = \{p + a \mid p = 2, 3, 5, \ldots \text{ primes}\}, \text{ where } a \in \mathbb{Z}^+$$

have certainly evinced a lot of interest, but the situation regarding general subsets of $\mathbb{Z}^+$ has not been given much attention. It is known that the methods underlying the fundamental results of Erdős-Kac [8] and Kubilius [18] described in §3, can be applied to any set S for which the so-called 'Brun Fundamental Lemma' from Sieve Theory holds. These methods, however, yield little information about moments, for which not much is known apart from certain cases relating to particular sets such as $\mathbb{Z}^+$, S_1 and S_2. Hence our method and results are of interest.

While work was in progress I had several useful discussions with Professors E. Bombieri, P.D.T.A. Elliott, P. Erdős, H.L. Montgomery and A. Selberg. In particular, Prof. Erdős always provided the advice and encouragement I needed. I am grateful to Prof. Montgomery for having suggested on the basis of some earlier work of mine [1], that it would be worthwhile to study in a fairly general setting, the moments of additive functions among subsets of the positive integers. Finally, I had the pleasure of being a visiting member at The Institute for Advanced Study, Princeton, during the period most of this work was done.

§1. Notation and statement of results

Recall that an additive function $f(n)$ is an arithmetical function that satisfies

$$f(mn) = f(m) + f(n) \quad \text{for} \quad (m, n) = 1. \tag{1.1}$$

Similarly a multiplicative (arithmetic) function $g(n)$ is one satisfying

$$g(mn) = g(m)g(n) \quad \text{for} \quad (m, n) = 1. \tag{1.2}$$

Thus additive and multiplicative functions are completely determined by their values on prime powers p^e, $e \in \mathbb{Z}^+$.

For the sake of convenience we concentrate only on strongly additive functions $f(n)$

which are given by

$$f(n) = \sum_{\substack{p \mid n \\ p = \text{prime}}} f(p) \tag{1.3}$$

and subsets S of $\mathbb{Z}^+$ which satisfy conditions (i) and (ii) below. It is possible to apply our method to general additive functions as well and this is briefly indicated in §11. We can treat an even wider class of sets which will then contain S_1 and S_2 as special cases; but due to technical reasons we postpone this discussion to a subsequent paper. We need some notation now before we can state our results.

For any set $S \subseteq \mathbb{Z}^+$ we let $S(x) = S \cap [1, x]$. We associate with S a sequence $A = \{a_n\}_{n \in S}$ of positive numbers called 'weights'. If $a_n \equiv 1$, we simply denote this by $A = 1$. For $d \in \mathbb{Z}^+$ we denote by

$$S_d(x) = \sum_{\substack{n \in S(x) \\ n \equiv 0 (\text{mod } d)}} a_n \tag{1.4}$$

and for convenience write $X = S_1(x)$. We will always assume that $\log x << X << \log x$. The set S is called 'special' if it also satisfies

(i) There exists a multiplicative function $\omega(d)$ such that $0 \leq \omega(p) << 1$ for all primes p, and $S_d(x) << X\omega(d)/d$ for all $d \in \mathbb{Z}^+$.

(ii) There is a constant $c > 0$ for which there corresponds to any $b > 0$ a number $a > 0$ such that

$$\sum_{d \leq X^c/\log^a X} |R_d(x)| <<_b \frac{X}{\log^b X} ,$$

where

$$R_d(x) = S_d(x) - \frac{X\omega(d)}{d} .$$

For a special set S and a strongly additive f we define sums

$$B_1(x) = \sum_{p \leq x} \frac{f(p)\omega(p)}{p} \tag{1.5}$$

and

$$B_k(x) = \sum_{p \le x} \frac{|f(p)|^k \omega(p)}{p}, \quad \text{for} \quad k \ge 2 . \tag{1.6}$$

The class $\mathcal{A} = \mathcal{A}(S)$ is the collection of all strongly additive functions that satisfy the following conditions :

(iii) There exists c_f such that

$$\left\{\max_{p \le x} \frac{|f(p)|}{\log p}\right\} \le c_f \sqrt{B_2(x)} \quad \text{for all} \quad x \ge 1$$

(iv) There exists $\alpha = \alpha(x) \longrightarrow \infty$ with x such that

$$\lim_{x\to\infty} \frac{B_k(x) - B_k(y)}{B_2(x)^{k/2}} = 0, \quad \text{for } k = 1, 2, 3, \ldots ,$$

where $y = x^{1/\alpha}$.

Our first main result is the following :

Theorem 1 : *(a) Let S be a special set, $f \ge 0$ belong to $\mathcal{A}$, and $B_2(x) \longrightarrow \infty$ with x. For real v define*

$$K_x(v) = \sum_{\substack{p \le x \\ f(p) \le v\sqrt{B_2(x)}}} \frac{f^2(p)\omega(p)}{p . B_2(x)} . \tag{1.7}$$

Assume that there is a probability distribution function $K(v)$ such that

$$K_x(v) \longrightarrow K(v) \quad \text{as} \quad x \longrightarrow \infty \text{, almost surely in } v . \tag{1.8}$$

In addition assume that there exists $R > 0$ such that for all x

$$\int_{-\infty}^{\infty} \frac{e^{uv} - 1 - uv}{v^2} \, dK_x(v) \ll 1 \qquad 0 \le u \le R , \tag{1.9}$$

where in (1.9) and in what follows, $(e^{uv} - 1 - uv)/v^2$ is set equal to $u^2/2$ at $v = 0$. Then

$$\lim_{x\to\infty} \frac{1}{XB_2(x)^{k/2}} \sum_{n \in S(x)} a_n \{f(n) - B_1(x)\}^k = m_k \tag{1.10}$$

exists for $k = 0, 1, 2, \ldots$, and is finite.

(b) Let S, f, K_x and K be as in part (a). Define

$$F_x(v) = X^{-1} \sum_{\substack{n \in S(x) \\ f(n) - B_1(x) < v\sqrt{B_2(x)}}} a_n . \tag{1.11}$$

Then there exists a probability distribution $F(v)$ *such that*

$$m_k = \int_{-\infty}^{\infty} v^k dF(v) \tag{1.12}$$

and

$$F_x(v) \longrightarrow F(v) \quad \text{weakly as } x \longrightarrow \infty \ . \tag{1.13}$$

(c) Let S, F, K_x *and* K *be as in part (a). Then for* $|z| < R$

$$L(z) = \exp\left\{\int_{-\infty}^{\infty} \frac{e^{zv}-1-zv}{v^2}\, dK(v)\right\} \tag{1.14}$$

is an analytic function. Furthermore

$$m_k = \frac{d^k}{dz^k} L(z)\Big|_{z=0} \quad \text{for } k = 0, 1, 2, \ldots \ . \tag{1.15}$$

We have the following Corollaries to Theorem 1.

Corollary 1 : *Let* f, S, K_x *be as in Theorem 1(a). Then* $F(v)$ *in* (1.12) *is the Gaussian distribution*

$$G(v) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{v} e^{-u^2/2} du \tag{1.16}$$

if and only if

$$K(v) = \begin{cases} 0 & \text{if } v < 0 \\ 1 & \text{if } v > 0 \end{cases} \tag{1.17}$$

Corollary 2 : *Let* S *be special and* $f \geq 0$ *satisfy*

$$\left\{\max_{p \leq x} f(p)\right\} / \sqrt{B_2(x)} \longrightarrow 0 \text{ as } x \longrightarrow \infty \ . \tag{1.18}$$

Then (1.10), (1.12), *and* (1.13) *hold with* $F(v) = G(v)$.

Our next theorem is for the case when $B_2(x)$ tends to a finite limit.

Theorem 2 : *(a) Let* S *be a special set,* $f \geq 0$ *belong to* $\mathcal{A}$, *and* $B_2(x) \longrightarrow B < \infty$ *as* $x \longrightarrow \infty$. *For real* v *define*

$$K_x^*(v) = \sum_{\substack{p \leq x \\ f(p) \leq v}} \underline{f^2(p)\omega(p)} \tag{1.19}$$

and assume that there is $R^* > 0$ *such that for all* x

$$\int_{-\infty}^{\infty} \frac{e^{uv} - 1 - uv}{v^2}\, dK_x^*(v) \ll 1 \qquad 0 \le u \le R^* . \tag{1.20}$$

Then

$$\lim_{x\to\infty} \frac{1}{X} \sum_{n\in S(x)} a_n\{f(n)-B_1(x)\}^k = m_k^* \tag{1.21}$$

exists for $k = 0, 1, 2, 3, \ldots,$ *and is finite.*

(b) Let S, f, K_x^* *be as in part (a). Define*

$$F_x^*(v) = \frac{1}{X} \sum_{\substack{n\in S(x)\\ f(n)-B_1(x)\le v}} a_n . \tag{1.22}$$

Then there is a probability distribution $F^*(v)$ *such that*

$$m_k^* = \int_{-\infty}^{\infty} v^k dF^*(v) \qquad k = 0, 1, 2, 3, \ldots \tag{1.23}$$

and

$$F_x^*(v) \longrightarrow F^*(v) \quad \text{weakly as } x \longrightarrow \infty . \tag{1.24}$$

Finally

$$m_k^* = \frac{d^k}{dz^k} L^*(z)\Big|_{z=0} , \tag{1.25}$$

where

$$L^*(z) = \prod_p \left(1 + \frac{(e^{zf(p)}-1)\omega(p)}{p}\right) e^{-zf(p)\omega(p)/p} . \tag{1.26}$$

Without imposing any growth conditions upon f, the following result provides a uniform upper bound.

Theorem 3 : *Let* S *be a special set and* f *any complex valued strongly additive function. Then*

$$\sum_{n\in S(x)} a_n |f(n)-B_1(x)|^k \ll_k \begin{cases} X.B_2(x)^{k/2} & \text{for } 0 \le k \le 2 \\ X\{B_2(x)^{k/2}+B_k(x)\} & \text{for } k \ge 2 \end{cases}$$

The implicit constant depends only on S *and* k.

Except for the implicit constant Theorem 3 is best possible. So, from Theorems 1 and 2 we see that the presence of $B_2(x)^{k/2}$ is necessary. By considering a strongly additive f which vanishes on all primes except a single fixed prime, we see that the term $B_k(x)$ cannot be dispensed with. The implicit constant is effectively computable in terms of k and the implicit constants in (i) and (ii).

Note that Theorems 1 and 2 have been stated only for $0 \leq f \in \mathcal{A}$ but then trivially these results hold for similar negative f also. This has been the main limitation with our method so far; the sieve has restricted our discussion of asymptotics to additive functions which do not change sign. On the basis of Theorem 3 and certain ideas of Kubilius to be mentioned in the sequel, it is possible to extend Theorems 1 and 2 to similar real f as well and this is indicated in §11. This is not completely satisfactory since we have to be dependent on Kubilius' method for our derivation. It is therefore desirable to extend our method directly to bring real valued functions within the scope of Theorems 1 and 2.

As to what extent the growth conditions (iii) and (iv) are necessary for the validity of Theorems 1 and 2, remains at present open. These conditions arise naturally in our method. On the basis of earlier results in the subject we feel that such conditions cannot be relaxed considerably. For more on this see §10.

All notation introduced so far will be retained. The $\ll$ and '0' notations are equivalent and will be used interchangeably as is convenient. Implicit constants depend only on S unless otherwise indicated. As usual empty sums mean zero and empty products one. A strongly multiplicative function g is one given by $g(n) = \prod_{p|n} g(p)$. We say g is totally multiplicative, if in addition to (1 . 2) we have $g(p^e) = g(p)^e$ for each p and all $e \in \mathbb{Z}^+$. The Moebius function $\mu(n)$ is the multiplicative function given by $\mu(p) = -1$ and $\mu(p^e) = 0$ for each p, and $e \geq 2$. In addition, p(n) denotes the smallest prime factor of n if $n > 1$, and $p(1) = \infty$. Finally $P(y) = \prod_{p<y} p$.

We shall (as we have done before) refer to probability distributions like F(v) in (1.13) and $F^*(v)$ in (1.24) as limiting distributions or limit laws.

In what follows it will often be necessary to refer to various statements in this paper, such as Theorems, Corollaries and numbered expressions, in the special case $S = \mathbb{Z}^+$, $A = 1$. For convenience such references shall always be indicated by attaching a super-script ', to that statement - e.g.: Theorem 1', or (1.12').

Before proving our theorems it will be appropriate to describe briefly certain earlier approaches and results and compare them with ours. This is done in the next section.

§2. Some earlier approaches - a comparison

The study of additive functions originated in 1917 with the result due to Hardy and Ramanujan [15], that $\nu(n)$, the number of prime divisors of n, is almost always nearly of size log log n. Their proof depended on an inductive procedure used to bound uniformly the quantity

$$\nu_k(x) = \sum_{\substack{1 \leq n \leq x \\ \nu(n)=k}} 1 .$$

In 1934, Turan gave a short proof of this result [22] by showing that

$$\sum_{n \leq x} \{\nu(n) - \log\log x\}^2 \ll x \log\log x \quad , \qquad (2.1)$$

using only the simplest results on primes. This gave the first indication of the probabilistic nature of the problem since (2.1) is essentially an estimate for the second moment of $\nu(n)$.

In 1939 Erdös and Kac [8] established a remarkable distribution result for real strongly additive functions satisfying some mild conditions. More precisely when f is real

$$f(p) = O(1) \text{ and } B_2(x) \longrightarrow \infty \qquad (2.2)$$

they established (1.13') with $F(v) = G(v)$. They considered additive functions as sums of nearly independent random variables, one for each prime p, in the interval $[1, x]$. The Gaussian distribution was a consequence of the Central Limit Theorem which arises naturally when f is compared to a sum of independent random variables. Their ideas provided the impetus for much of the later work in this area.

Motivated by these developments and a remark by Kac [16] (see also §13), Halberstam [13: I] established in 1955 our Corollary 2' for all real strongly additive functions satisfying (2. 2). He actually expanded the left side of (1.10') and carried out the calculations in an elementary manner. This involved great complications but it enabled him to extend his results to S_1 [13: II] and S_2 [13: III], by which time it was realised that (2. 2) could be replaced by the weaker condition (1.18).

Rényi and Turán [19] took an analytic approach. They considered the problem of determining the limiting distribution as equivalent to finding the limit of the characteristic functions (Fourier transforms). The characteristic function is given in terms of the sum

$$\sum_{n \leq x} z^{f(n)} \tag{2.3}$$

with $z = e^{iu}$, u real. This sum of a multiplicative function can be treated by well-known analytic methods provided the $f(p)$ enjoy some regular behaviour. In particular the method applies nicely to $f(n) = \nu(n)$, and for this case they obtained the best estimate for the rate of convergence in (1.13') to the Gaussian distribution. Independently, Selberg [20] demonstrated that asymptotic estimates for (2. 3) with $|z| \leq r$ in the case $f(n) = \nu(n)$, enables one to quickly derive asymptotic estimates for $\nu_k(x)$, where k could vary with x. Previously Sathe following the inductive procedure of Hardy - Ramanujan had obtained similar uniform estimates for $\nu_k(x)$ in a rather complicated fashion. These ideas of Rényi-Turán and Selberg have subsequently been applied to other additive functions also.

All results referred to so far in this section are for the case $B_2(x) \longrightarrow \infty$ with the limiting distribution as the Gaussian law. The first instance of $B_2(x) \longrightarrow \infty$ and limiting distributions other than the Gaussian, was provided by Kubilius [17] who, amongst other things, successfully combined the ideas of Erdös-Kac [8] and Rényi-Turán [19]. He established rather general distribution results in the situation $S = \mathbb{Z}^+$, $A = 1$, for a large class of strongly additive functions which he called the class $\mathcal{H}$. This class comprises of all real f for which $B_2(x) \longrightarrow \infty$ and (iv)' needs to hold only for $k = 2$. By employing techniques from probability theory, such as independent random variables, infinitely divisible distributions, and characteristic functions, he proved the striking result that for the class $\mathcal{H}$, condition (1. 8') is both necessary and sufficient for (1.13') to hold. He even succeeded in determining the characteristic function of $F(v)$ and thus in principle determined F. This method of Kubilius, and that of Erdös-Kac, applies to any set for which 'Brun's Fundamental Lemma' (see §4) holds. It should be noted however, that these methods do not yield information regarding moments. But then, Theorem 1 guarantees in this situation that if one imposes certain conditions upon f, then (1.10) holds in special sets for every integer $k \geq 0$.

If f is a real strongly additive function for which $B_2(x) \longrightarrow B < \infty$, then condition (iv) for $k = 2$ is redundant. So Kubilius [18] observed that here (1.24') holds without imposing any more conditions upon f. If in addition $B_1(x)$ converges as $x \longrightarrow \infty$, then the frequencies

$$\frac{1}{x} \sum_{\substack{n \le x \\ f(n) < v}} 1 \tag{2.4}$$

also converge weakly to a limiting distribution. Previously Erdös [7: I, II, III] had considered such questions and established distribution results unders weaker conditions. His efforts culminated in the celebrated Erdös-Wintner Theorem, which gives necessary and sufficient conditions for the frequencies in (2. 4) to converge; the sufficiency part was obtained by Erdos [7: III], and the necessity jointly established by Erdös-Wintner [9]. But none of these results deal with moments. Theorem 2 on the other hand provides sufficient conditions for (1.21) to hold in special sets, when $B_2(x)$ tends to a finite limit.

In 1962 Kubilius wrote a monograph (English translation [18] in 1964), where he described his methods in detail and compared them to earlier approaches. He mentioned ([18], p.71) that the method of moments which had previously been employed to determine the limiting distribution of additive functions, had been used only in the situation when the limit was the Gaussian distribution. Apparently, Kubilius had overlooked Delange's announcement [4] even though he referred to it in [18]! It is true that Halberstam always made use of conditions such as (2. 2) or (1.18) and in such situations the limiting contribution is Gaussian. Delange, who had previously investigated [2] the moments of $\nu(n)$, showed in [3] that Halberstam's method could be simplified by introducing suitable generating functions and interpreting the quantities arising out of the expansion of the left side of (1.10') in terms of the derivatives of these functions. After Kubilius established his general distribution theorems in 1956 [17], Delange [4] observed that such results could fit in his method also provided one imposed certain conditions upon f (his conditions were slightly stronger than ours). Delange only announced his results and did not give any details since the method was clearly an extension of what he used for the Gaussian case. We have recently come to know that Delange has verified that his generating function method also yields Halberstam's results for polynomials, in which case the limiting distribution is Gaussian.

Realising that a complete generalization of (2. 1) would be useful, Kubilius showed in 1956 (see [5], Vol.1, Ch.4), that for arbitrary complex f

$$\sum_{n \le x} |f(n) - B_1(x)|^2 \ll xB_2(x) \quad . \tag{2.5}$$

Recently Elliott [6] by employing sums similar to (2. 3) derived in an elegant fashion the inequality

$$\sum_{n\leq x} |f(n) - B_1(x)|^k \ll_k \begin{cases} xB_2(x)^{k/2} & \text{for } 0 \leq k \leq 2 \\ x\{B_2(x)^{k/2} + B_k(x)\} & \text{for } k \geq 2 \ . \end{cases} \tag{2.6}$$

and thus extended (2. 5) to arbitrary exponents. The method we describe in this paper is an improvement of this technique due to Elliott.

We interpret certain sums similar to those considered by Elliott, in terms of the bilateral Laplace transform of the distribution of $f(n)$, $n \in S$. We use the combinatorial sieve to estimate this sum involving multiplicative functions, in certain cases. The sieve yields a crucial improvement over Elliott's method for the following reasons : First it permits treatment of a reasonably wide class of sets. The combinatorial sieve also yields in some cases, asymptotic estimates for the Laplace transform, and from this we derive Theorems 1 and 2 without much trouble because the Laplace transform retains some of the elegance of the Fourier transform. In addition we can derive a general upper bound like that in Theorem 3, which incidentally extends (2. 6). We shall point out the exact differences between Eliott's method and ours in §12.

Theorems 1 and 3 are proved in sections 7 and 9 respectively. We omit the proof of Theorem 2 since it is similar to Theorem 1, and in fact simpler. We shall in sections 3, 4, 5 and 6 establish the necessary preliminaries.

§3. The bilateral Laplace transform

The aim of this section is to prove

Lemma 1 : *(a) Let* $\varphi_x(v)$ *be a sequence of probability distributions for which there is* $R_0 > 0$ *such that*

$$\int_{-\infty}^{\infty} e^{uv} d\varphi_x(v) \ll 1 \text{ for } -R_0 \leq u \leq R_0 \ . \tag{3.1}$$

Then

$$\int_{-\infty}^{\infty} |v|^k d\varphi_x(v) \ll \frac{k!}{R_0^k} \quad \text{for } k = 0, 1, 2, \ldots \tag{3.2}$$

If in addition to (3. 1) we have

$$\lim_{x\to\infty} \int_{-\infty}^{\infty} e^{uv} d\varphi_x(v) = \ell(u) < \infty \text{ for } -R_0 \le u \le 0 \quad . \tag{3.3}$$

where the convergence is uniform in u, then

$$\lim_{x\to\infty} \int_{-\infty}^{\infty} v^k d\varphi_x(v) = \mu_k \tag{3.4}$$

exists and is finite for $k = 0, 1, 2, \ldots$.

(b) *Let* $\varphi_x(v)$ *satisfy (3. 1) and (3. 3). Then there is a probability distribution* $\varphi(v)$ *such that*

$$\mu_k = \int_{-\infty}^{\infty} v^k d\varphi(v) \quad \text{for} \quad k = 0, 1, 2, 3, \ldots \tag{3.5}$$

and

$$\varphi_x(v) \longrightarrow \varphi(v) \quad \text{weakly as } x \longrightarrow \infty \quad . \tag{3.6}$$

(c) *Let* $\varphi_x(v)$ *satisfy (3. 1) and (3. 3). Then* $\ell(u)$ *can be extended to an analytic function in* $|z| < R_0$. *Therefore*

$$\mu_k = \frac{d^k}{dz^k} \ell(z)\Big|_{z=0} \qquad k = 0, 1, 2, 3, \ldots \quad . \tag{3.7}$$

Proof of part (a) : The assertions are trivial for $k = 0$, so assume $k \in \mathbb{Z}^+$.
Note that

$$|\theta^k|/k! \le e^{\theta} + e^{-\theta} \quad \text{for all } k \in \mathbb{Z}^+ \text{ and real } \theta \tag{3.8}$$

Therefore by (3. 8) and (3. 1) we get

$$\frac{R_0^k}{k!} \int_{-\infty}^{\infty} |v|^k d\varphi_x(v) \le \int_{-\infty}^{\infty} (e^{R_0 v} + e^{-R_0 v}) d\varphi_x(v) \ll 1$$

and this proves (3. 2).

From (3. 2) we deduce that

$$\int_{|v|\ge T} d\varphi_x(v) \le T^{-2n} \int_{-\infty}^{\infty} v^{2n} d\varphi_x(v) \ll_{n,R_0} T^{-2n} \qquad n = 1, 2, \ldots \quad . \tag{3.9}$$

By the Cauchy-Scwartz inequality, (3. 1) and (3. 9) we have

$$\int_{|v|\ge T} e^{uv} d\varphi_x(v) \le \left(\int_{-\infty}^{\infty} e^{2uv} d\varphi_x(v)\right)^{1/2} \left(\int_{|v|\ge T} d\varphi_x(v)\right)^{1/2} \ll_{n,R_0} T^{-n} \tag{3.10}$$

for $|u| \leq R_0/2$. Similarly

$$\int_{|v|\geq T} |v|^k d\varphi_x(v) \leq \left(\int_{-\infty}^{\infty} v^{2k} d\varphi_x(v)\right)^{1/2} \left(\int_{|v|\geq T} d\varphi_x(v)\right)^{1/2} \ll_{n,k,R_0} T^{-n} . \tag{3.11}$$

for all integers k, n, ≥ 1 .

From (3. 3) we see that given any $\varepsilon > 0$, there exists x_0 such that for all x, x' $\geq$ x_0 we have

$$\left|\int_{-\infty}^{\infty} e^{uv} d\varphi_x(v) - \int_{-\infty}^{\infty} e^{uv} d\varphi_{x'}(v)\right| < \varepsilon \quad \text{for } -R_0 \leq u < 0 . \tag{3.12}$$

Therefore from (3.10) and (3.12) we deduce that

$$\left|\int_{-T}^{T} e^{uv} d\varphi_x(v) - \int_{-T}^{T} e^{uv} d\varphi_{x'}(v)\right| \ll_{n,R_0} \varepsilon + T^{-n}, \text{ for } -\frac{R_0}{2} \leq u \leq 0. \tag{3.13}$$

Observe that if u, v are real numbers such that uv is bounded then

$$\left(\frac{e^{uv}-1}{u}\right)^k = v^k + O_k(|uv^{k+1}|) , \quad k = 1, 2, 3, \ldots . \tag{3.14}$$

Given a negative number u close to zero, we choose T in (3.13) such that $T|u|=R_0/2k$. With this choice of T, we see from (3.13), (3.14), (3. 2) and (3.11) that

$$\left|\int_{-\infty}^{\infty} v^k d\varphi_x(v) - \int_{-\infty}^{\infty} v^k d\varphi_{x'}(v)\right| \ll_{k,n,R_0} \frac{\varepsilon + T^{-n}}{u^k} + u, \quad k = 1, 2, 3, . \tag{3.15}$$

In (3.15) we choose $n = 2k$, and $u^k = \sqrt{\varepsilon}$. Then the right side of (3.15) can be arbitrarily small. Thus for each $k \in \mathbb{Z}^+$, the sequence of k^{th} moments of $\varphi_x(v)$ is a Cauchy sequence. Hence this sequence converges to a value μ_k, giving (3. 4). That proves part (a).

<u>Proof of part (b)</u> : By the theorem of Frechet and Shohat [11], it follows from (3. 4) that there exists a probability distribution $\varphi(v)$ such that (3. 5) holds, and that there is a subsequence φ_{x_j}, j = 1, 2, 3, . . . , which converges weakly to $\varphi(v)$. From (3. 2) we see that $\mu_k \ll k!/R_0^k$ and so $\sum_k |\mu_k| t^k/k!$ has a non-zero radius of convergence. Therefore (see Feller [10], p. 224 and p. 487) these moments determine a unique probability distribution, and so (3. 6) holds.

<u>Proof of part (c)</u> : From (3. 1), (3. 3) and (3. 4) we can see that ℓ extends to an analytic function in $|z|<R_0$. To prove (3. 7) we now observe that, instead of (3.13) we can write

$$\left|\int_{-T}^{T} e^{uv}\varphi_x(v) - \ell(u)\right| \ll_{n,R_0} \varepsilon + T^{-n} \quad \text{for} \quad -\frac{R_0}{2} \leq u \leq 0. \tag{3.16}$$

This follows from (3. 3) and (3.10). If we combine (3.16) with (3.11) and (3.14) we get

$$\left|\int_{-\infty}^{\infty} v^k d\varphi_x(v) - \sum_{j=0}^{k} \ell(ju)\binom{k}{j}\frac{(-1)^{j-k}}{u^k}\right| \ll_{k,n,R_0} \frac{\varepsilon+T^{-n}}{u^k}, \quad k \in \mathbb{Z}^+. \tag{3.17}$$

We choose n, T, u as in part (b), and let $\varepsilon \longrightarrow 0$. Then the right side of (3. 17) tends to zero, with ε. Also, since $\ell(z)$ is analytic, we have

$$\sum_{j=0}^{k} \ell(ju)\binom{k}{j}\frac{(-1)^{j-k}}{u^k} \longrightarrow \frac{d^k}{dz^k}\ell(z)\Big|_{z=0} \quad (\text{as} \quad u \longrightarrow 0). \tag{3.18}$$

because the sum in (3.18) is the k^{th}-difference of $\ell(z)$, with arguments as multiples of u. So (3. 7) for $k \in \mathbb{Z}^+$ follows from (3.18). Since (3. 7) is trivial for $k = 0$, this proves part (c), and completes the proof of Lemma 1.

§4. The combinatorial sieve

The earliest sieve method, now known as the Eratosthenes-Legendre sieve, rests on the use of the identity

$$\sum_{d|n} \mu(d) = \begin{cases} 1 & \text{if} \quad n = 1 \\ 0 & \text{if} \quad n > 1. \end{cases}$$

Let $S(x,y) = \{n \leq x \mid p(n) \geq y\}$. Then we have

$$\sum_{n \in S(x,y)} a_n = \sum_{n \in S(x)} a_n \sum_{d|(n,P(y))} \mu(d). \tag{4.1}$$

While rearranging the right hand side one has to consider error terms arising out of all the divisors of P(y), and these are too numerous if y is large. In such a situation the method becomes unwieldy.

To compensate for this the combinatorial sieve considers functions χ_1, χ_2 which have the properties

$$\chi_1(1) = \chi_2(1) = 1 \tag{4.2}$$

and

$$\sum_{d|n} \mu(d)\chi_2(d) \leq \sum_{d|n} \mu(d) \leq \sum_{d|n} \mu(d)\chi_1(d) \quad \text{for } n = 1,2,3,\ldots \tag{4.3}$$

The χ_1 are chosen to vanish quite often on the divisors of $P(y)$, in order to keep the number of error terms in check. The sum in (4. 3) involving χ_1 when used in conjunction with (4. 1) provides an upper bound, while the sum involving χ_2 supplies a lower bound. If for a suitable choice of χ_1, χ_2, these bounds turn out to be close, then one has an asymptotic estimate for the quantity in (4. 1).

Viggo Brun, who first introduced this idea, treated the sum (4. 1) in a large number of cases by suitable choices of χ_1. Subsequently, Brun's method has been considerably improved and it is known (see Halberstam and Richert [14], p. 83) that for special sets the combinatorial sieve yields

$$\sum_{n\in S(x,y)} a_n = X \prod_{p\le y}\left(1-\frac{\omega(p)}{p}\right)\{1+\eta(x,y)\},\quad 1\le y\le x, \qquad (4.4)$$

where $\eta(x, y)$ is a bounded function that tends to zero as $\log x/\log y = \alpha$ tends to infinity. In particular from (4. 4) we see that

$$\sum_{n\in S(x,y)} a_n \ll X\prod_{p\le y}\left(1-\frac{\omega(p)}{p}\right),\quad 1\le y\le x, \qquad (4.5)$$

holds uniformly. An estimate such as (4. 4), which is asymptotic for large α, is known in sieve theory as "Brun's Fundamental Lemma".

We are going to employ the combinatorial sieve in a new way to estimate sums of certain multiplicative functions $g(n)$, for $n \in S$. Our method rests on the following crucial inequalities.

Lemma 2 : *Let $g^*(n)$ be a multiplicative function satisfying $0 \le g^*(n) \le 1$ for all $n \in \mathbb{Z}^+$. Let χ_1, χ_2 satisfy (4. 2) and (4. 3). Then for all positive integers n we have*

$$\sum_{d|n}\mu(d)\chi_2(d)g^*(d) \le \sum_{d|n}\mu(d)g^*(d) \le \sum_{d|n}\mu(d)\chi_1(d)g^*(d).$$

Proof : We only prove one of these inequalities, the other being similar. We may assume without loss of generality that n is square free. We let

$$\sigma(n) = \sum_{d|n}\mu(d)\chi_1(d) \quad . \qquad (4.6)$$

Then

$$\sigma(1) = 1 \quad \text{and} \quad \sigma(n) \ge 0 \text{ for } n > 1 \ . \qquad (4.7)$$

By Moebius inversion and (4. 6) we have

$$\sum_{d|n}\mu(d)g^*(d)\chi_1(d) = \sum_{d|n} g^*(d) \sum_{\delta|d}\mu(\tfrac{d}{\delta})\sigma(\delta) = \sum_{\delta|n}\sigma(\delta) \sum_{t|n/\delta}\mu(t)g^*(t\delta). \quad (4.8)$$

We separate the term corresponding to $\delta = 1$ in (4. 8) and arrive at

$$\sum_{d|n}\mu(d)g^*(d)\chi_1(d) = \sum_{d|n}\mu(d)g^*(d) + \sum_{1<\delta|n}\sigma(\delta)g^*(\delta) \sum_{t|n/\delta}\mu(t)g^*(t)$$
$$\geq \sum_{d|n}\mu(d)g^*(d)$$

because of (4. 7) and the fact that

$$g(n) = \sum_{d|n}\mu(d)g^*(d) \geq 0 \quad \text{for all } n \in \mathbb{Z}^+. \quad (4.9)$$

This proves Lemma 2.

Let $g(n)$ be a given strongly multiplicative function such that $0 \leq g(n) \leq 1$ for all $n \in \mathbb{Z}^+$. We associate with g, the function g^* as in (4. 9). Then

$$g_y(n) \overset{def}{=} \prod_{\substack{p|n \\ p\leq y}} g(p) = \sum_{d|(n,P(y))}\mu(d)g^*(d). \quad (4.10)$$

Instead of the sum in (4. 1) we now consider

$$\sum_{n\in S(x)} a_n g_y(n) = \sum_{n\in S(x)} a_n \sum_{d\,|\,(n,\,P(y))}\mu(d)g^*(d). \quad (4.11)$$

Our idea is to use Lemma 2 to treat (4.11).

The error terms that will arise here are bounded by

$$\sum_{d\_x,d|P(y)} |\chi_i(d)||R_d(x)|g^*(d)$$

and this is smaller than the sum

$$\sum_{d\leq x,d|P(y)} |\chi_i(d)||R_d(x)|$$

which bounds the error terms arising out of (4. 1) and (4. 3). Our main term is

$$X \prod_{p\leq y}\left(1 - \frac{g^*(p)\omega(p)}{p}\right)$$

and this is

$$\geq X \prod_{p\leq y}\left(1 - \frac{\omega(p)}{p}\right),$$

which is the main term in (4. 4). Thus by using Lemma 2 and following the derivation of

(4. 4) by the combinatorial sieve (see [14], Ch. 2, for details), we arrive at

Lemma 3 : *Let S be a special set and $g(n)$, $g^*(n)$, $g_y(n)$ as in (4. 9) and (4.10), where $0 \le g \le 1$. Then for $0 \le y \le x$ we have*

$$\sum_{n\in S(x)} a_n g_y(n) = X \prod_{p\le y}\left(1 - \frac{g^*(p)\omega(p)}{p}\right)\{1 + O(\eta(x,y))\}, \qquad (4.12)$$

where $\eta(x,y)$ is as in (4. 4). Since $g(n) \le g_y(n)$, we have from (4.12)

$$\sum_{n\in S(x)} a_n g(n) << X \prod_{p\le x}\left(1 - \frac{g^*(p)\omega(p)}{p}\right) . \qquad (4.13)$$

Remarks : If we choose $g(n)$ in Lemma 3 to satisfy, $g(p) = 0$ for $p < y$ and $g(p) = 1$ for $p > y$, then (4.12) and (4.13) correspond to (4. 4) and (4. 5) respectively. The usefulness of Lemma 3 lies in the unformity with respect to all strongly multiplicative g, satisfying $0\le g\le 1$. Estimate (4.12) will be used to establish Theorems 1 and 2, whereas inequality (4.13)will be employed in the proof of Theorem 3. The idea is that from estimates such as these for multiplicative functions, one can extract information concerning additive functions.

§5. The sums $T_u(x)$ and $T_u(x,y)$.

For a strongly additive function f, we define

$$f_y(n) = \sum_{\substack{p|n\\ p\le y}} f(p) . \qquad (5.1)$$

Next, for real u, we let

$$T_u(x,y) = T_u(x,y,f,S) = \sum_{n\in S(x)} a_n\, e^{u\{f_y(n)-B_1(y)\}/\sqrt{B_2(y)}} \qquad (5.2)$$

and set $T_u(x,x) = T_u(x)$. Suppose S is special and $f > 0$. We consider two cases.

Case 1 : $u \le 0$.

Here $0 \le g(n) \le 1$, where g is the strongly multiplicative function

$$g(n) = g(n,y) = e^{uf(n)/\sqrt{B_2(y)}} , \quad n = 1, 2, \ldots . \qquad (5.3)$$

Note that (5. 3), (5. 1), and (4.10) give

$$g_y(n) = e^{uf_y(n)/\sqrt{B_2(y)}} \qquad (5.4)$$

Therefore by (5. 2), (5. 3), (5. 4) and Lemma 2 we have

$$T_u(x,y) = e^{-uB_1(y)/\sqrt{B_2(y)}} . X . \prod_{p\le y} \left(1 - \frac{g^*(p)\omega(p)}{p}\right)\{1 + O(\eta(x,y))\}\ (u \le 0) \quad (5.5)$$

and

$$T_u(x) \ll e^{-uB_1(x)/\sqrt{B_2(x)}} . X . \prod_{p\le x} \left(1 - \frac{g^*(p)\omega(p)}{p}\right)(u \le 0) . \quad (5.6)$$

Case 2 : $u > 0$.

Consider strongly multiplicative functions

$$h(n) = h(n,y) = e^{uf(n)/\sqrt{B_2(y)}}, \quad h_y(n) = e^{uf_y(n)/\sqrt{B_2(y)}}, \quad (5.7)$$

and note that these are always ≥ 1. Let h^* be the unique function satisfying

$$h(n) = \sum_{d|n} h^*(d) .$$

Then h^* is multiplicative, ≥ 0, and $h^*(p^e) = 0$ for all p, and $e \ge 2$. Also

$$h_y(n) = \sum_{d|(n,P(y))} h^*(d) . \quad (5.8)$$

So from (i) and (5. 8) we see that

$$\sum_{n\in S(x)} a_n h_y(n) = \sum_{n\in S(x)} a_n \sum_{d|(n,P(y))} h^*(d) = \sum_{d\le x, d|P(y)} h^*(d) S_d(x)$$

$$\ll X \sum_{d\le x, d|P(y)} \frac{h^*(d)\omega(d)}{d} \le X \prod_{p\le y} \left(1 + \frac{h^*(p)\omega(p)}{p}\right) . \quad (5.9)$$

Therefore from (5. 7), (5. 9) and (5. 2) we deduce that

$$T_u(x,y) \ll e^{-uB_1(y)/\sqrt{B_2(y)}}\ X \prod_{p\le y} \left(1 + \frac{h^*(p)\omega(p)}{p}\right) (y \le x, u > 0) . \quad (5.10)$$

If we put $y = x$ in (5.10), we get an upper bound for $T_u(x)$.

Observe that inequalities (5. 6) and (5.10) are similar, although they were derived differently.

§6. Auxilliary upper bounds.

The bounds supplied by Lemma 4 below will be useful in the sequel for two reasons. First, while proving Theorems 1 and 2 it will enable us to estimate the moments of $f(n)$ from those of $f_y(n)$. Next, in the proof of Theorem 3, Lemma 4 will be used to control the contribution due to large values of $f(p)$, which have to be considered separately.

Lemma 4 : *(a) Let S be a special set. Let E be an arbitrary set of primes and f^* an strongly additive function ≥ 0 for which $f^*(p) = 0$ if $p \notin E$-equivalently $f^*(n) = \sum_{p|n, p\in E} f^*(p)$ Also for each non-negative integer k define*

$$B_k^*(x) = \sum_{p\leq x,\, p\in E} \frac{f^*(p)^k \omega(p)}{p} . \tag{6.1}$$

Then

$$\sum_{n\in S(x)} a_n f^*(n)^k \ll_k X\, B_k^*(x) B_0^*(x)^{k-1} .$$

(b) Let S be special, and $E = \{p \mid y<p\leq x\}$, where y is as in (iv). Let $f \in \mathcal{A}$ and $f^*(n) = \sum_{p|n,\, p\in E} f(p)$, -equivalently $f^*(n) = f(n) - f_y(n)$. Then

$$\sum_{n\in S(x)} a_n f^*(n)^k = o(X\, B(x)^{k/2}) \quad \text{as } x \longrightarrow \infty, \text{ for } k = 1, 2, \ldots .$$

Proof of part (a) : The assertion is trivial for $k = 0$, so assume $k \in \mathbb{Z}^+$. By expansion and reaarangement we have

$$\sum_{n\in S(x)} a_n f^*(n)^k = \sum_{n\in S(x)} a_n \left(\sum_{p|n,\, p\in E} f^*(p)\right)^k =$$

$$= \sum_{\substack{j_1+j_2+\cdots+j_r=k \\ j_i \in \mathbb{Z}^+}} \sum_{\substack{p_1,\ldots p_r\in E \\ \text{distinct primes}}} f(p_1)^{j_1}\ldots f(p_r)^{j_r} \sum_{\substack{n\in S(x) \\ n\equiv 0 (\text{mod } p_1\cdots p_r)}} a_n . \tag{6.2}$$

In (6. 2) the outer summation on the extreme right is over all ordered partitions (compositions) of k. From (i), (6. 1) and (6. 2) we get

$$\sum_{n\in S(x)} a_n f^*(n)^k \ll X \sum_{\substack{j_1+\cdots+j_r=k \\ j_i \in \mathbb{Z}^+}} \sum_{\substack{p_1,\ldots,p_r\in E \\ \text{distinct primes}}} \frac{f^*(p_1)^{j_1}\ldots f^*(p_r)^{j_r}\omega(p_1)\ldots\omega(p_r)}{p_1\cdots p_r}$$

$$\leq X \sum_{\substack{j_1+\cdots+j_r=k \\ j_i \in \mathbb{Z}^+}} B_{j_1}^*(x)\ldots B_{j_r}^*(x). \tag{6.3}$$

By the Holder-Minkowski inequality

$$B_j^*(x) \leq B_k^*(x)^{j/k} B_0^*(x)^{1-j/k}, \quad \text{for } 1 \leq j \leq k. \tag{6.4}$$

Note that in (6. 3) we have $r \leq k$. Lemma 4 (a) follows from (6. 3) and (6. 4) - the implicit constant being given in terms of the number of compositions of k, and the constant in (i).

Proof of part (b) : In this case also expressions (6. 2) and (6. 3) hold, because according to our notation $f(p) = f^*(p)$ for $p \in E$. So by (6. 1), (1. 5), (1. 6) and (iv) we get

$$B_j^*(x) = B_j(x) - B_j(y) = o(B_2(x)^{j/2}), \quad j \in \mathbb{Z}^+ . \tag{6.5}$$

Part (b) follows from (6. 3) and (6. 5), because $j_1 + \dots + j_r = k$. Lemma 4 is proved.

§7. Proof of Theorem 1

Let f and S satisfy the hypothesis of Theorem 1, and x, y, α as in (iv). With these choices, we let

$$F_{x,y}(v) = \frac{1}{X} \sum_{\substack{n \in S(x) \\ f_y(n) - B_1(y) < v\sqrt{B_2(y)}}} a_n \quad . \tag{7.1}$$

From (5. 2) and (7. 1) we get

$$\int_{-\infty}^{\infty} e^{uv} dF_{x,y}(v) = \frac{T_u(x,y)}{X} \quad . \tag{7.2}$$

So, using (5. 5) we see that as $x \longrightarrow \infty$, the expression in (7. 2) is

$$\sim e^{-uB_1(y)/\sqrt{B_2(y)}} \prod_{p \le y} \left(1 - \frac{g^*(p)\omega(p)}{p}\right), \quad \text{for } u \le 0, \tag{7.3}$$

where the value of $g^*(p)$ in (7. 3) is given by (5. 3) and (4. 9) to be

$$g^*(p) = 1 - g(p) = 1 - e^{uf(p)/\sqrt{B_2(y)}} \quad . \tag{7.4}$$

Condition (i) along with (7. 4) implies that there exists $R_1 > 0$ such that

$$0 \le \frac{g^*(p)\omega(p)}{p} \le \frac{1}{2} \quad \text{for } -R_1 \le u \le 0 . \tag{7.5}$$

Furthermore, condition (iii) guarantees that there exists $R_2 > 0$ such that

$$R_2 \; f(p) \le \frac{1}{3} \log p\sqrt{B_2(y)} \; , \quad \text{for} \quad p \le y \; . \tag{7.6}$$

We now choose

$$R_0 = \min(R, R_1, R_2), \tag{7.7}$$

where R is as in the hypothesis of Theorem 1. We will estimate the expression in (7. 2) for $|u| \le R_0$ by making use of the estimate

$$\log(1+t) = t + O(t^2), \qquad \text{for} \quad t \geq -1/2 . \tag{7.8}$$

We begin by observing that

$$\log\left\{\prod_{p\leq y}\left(1 + \frac{(e^{uf(p)/\sqrt{B_2(y)}}-1)\omega(p)}{p}\right)\right\} = \sum_{p\leq y} \frac{(e^{uf(p)/\sqrt{B_2(y)}}-1)\omega(p)}{p} +$$

$$O\left(\sum_{p\leq y}\left\{\frac{(e^{uf(p)/\sqrt{B_2(y)}}-1)\omega(p)}{p}\right\}^2\right) = \sum\nolimits_1 + O(\sum\nolimits_2) \quad \text{for } -R_0 \leq u \leq R_0 . \tag{7.9}$$

Estimate (7. 9), for $u \geq 0$ follows from (7. 8), and for $u \leq 0$ is a consequence of (7. 8) and (7. 5). We shall bound $\sum_2$ first. For this we choose Y tending to infinity with y such that

$$B_2(Y) \log Y = o(B_2(y)), \tag{7.10}$$

and decompose

$$\sum\nolimits_2 = \sum_{p\leq Y} + \sum_{Y\leq p\leq y} = \sum\nolimits_3 + \sum\nolimits_4 \quad \text{respectively} \tag{7.11}$$

Note that condition (iii) implies that $f(p) \leq c_f \log p\sqrt{B_2(p)}$. So from (i) and (7.10) we get

$$\sum\nolimits_3 = \sum_{p\leq Y} \frac{o(1)}{p^2} = o(1) \qquad (y \longrightarrow \infty;\ -R_0 \leq u \leq R_0). \tag{7.12}$$

On the other hand from (7. 6) and (7. 7) we deduce that

$$\sum\nolimits_4 \ll \sum_{p\leq Y} \frac{1}{p^{3/2}} = o(1) \quad (y \longrightarrow \infty;\ -R_0 \leq u \leq R_o). \tag{7.13}$$

Thus estimates (7. 9), (7.10), (7.12) and (7.13) yield

$$\log\left\{\prod_{p\_y}\left(1 + \frac{(e^{uf(p)/\sqrt{B_2(y)}}-1)\omega(p)}{p}\right)\right\} = \sum\nolimits_1 + o(1), \quad (y \longrightarrow \infty;\ -R_0 \leq u \leq R_0). \tag{7.14}$$

By expansion of $\sum_1$ and (1. 7) we have

$$\sum\nolimits_1 = \frac{uB_1(y)}{\sqrt{B_2(y)}} + \sum_{p\leq y}\sum_{k=2}^{\infty} \frac{u^k f(p)^k \omega(p)}{k!\,p.B_2(y)^{k/2}} = \frac{uB_1(y)}{\sqrt{B_2(y)}} + \int_{-\infty}^{\infty} \frac{e^{uv}-1-uv}{v^2}\, dK_y(v). \tag{7.15}$$

So by (7. 3), (7. 4), (7.14) and (7.15), the quantity in (7. 2) is

$$\sim\exp\left\{\int_{-\infty}^{\infty} \frac{e^{uv}-1-uv}{v^2}\, dK_y(v)\right\} \qquad -R_0 \leq u \leq 0. \tag{7.16}$$

The quantities $(e^{uv}-1-uv)/v^2$, treated as functions of v, for $v \in [0,\infty)$, and parametrized by u, for $u \in [-R_0,0]$, form an equicontinuous family, which is uniformly bounded.

Since $K_y(v) = 0$ for $v < 0$, it follows from hypothesis of Theorem 1, and by a result on convergence of measures (see Feller [10], p. 245) that

$$\lim_{y \to \infty} \int_{-\infty}^{\infty} \frac{e^{uv}-1-uv}{uv} dK_y(v) = \int_{-\infty}^{\infty} \frac{e^{uv}-1-uv}{v^2} dK(v) = L(u) \quad \text{uniformly for } -R_0 \le u \le 0. \tag{7.17}$$

Therefore combining (7.16), (7.17), (7. 3) and (7. 2) we arrive at

$$\lim_{x \to \infty} \int_{-\infty}^{\infty} e^{uv} dF_{x,y}(v) = L(u) \quad \text{uniformly for } -R_0 \le u \le 0 . \tag{7.18}$$

On the other hand, by (5.10) and (7. 2) we have

$$\int_{-\infty}^{\infty} e^{uv} dF_{x,y}(v) << e^{-uB_1(y)/\sqrt{B_2(y)}} \prod_{p \le y} \left(1 + \frac{h^*(p)\omega(p)}{p}\right), \text{ for } 0 \le u \le R_0, \tag{7.19}$$

where

$$h^*(p) = e^{uf(p)/\sqrt{B_2(y)}} - 1. \tag{7.20}$$

In this case, by (7.14), (7.15), (7.20), (7.19) and (1. 9) we arrive at

$$\int_{-\infty}^{\infty} e^{uv} dF_{x,y}(v) << 1, \quad \text{uniformly for } 0 \le u \le R_0 \tag{7.21}$$

Thus by setting $F_{x,y}(v) = \varphi_x(v)$ we see from (7.18) and (7.21) that φ_x satisfies the conditions of Lemma 1, part (a). Consequently part (a) of Theorem 1 holds with $f_y(n)$ in the place of $f(n)$. Similarly part (b) of Theorem 1 for $f_y(n)$ also follows from part (b) of Lemma 1.

From (1. 9) we see by expansion that

$$\sum_{k=2}^{\infty} \frac{u^k}{k!} \left(\int_0^{\infty} v^{k-2} dK_x(v)\right) << 1 , \quad \text{for } 0 \le u \le R.$$

Hence

$$0 \le \int^{\infty} v^{k-2} dK_x(v) << \frac{k!}{R^k}, \quad \text{for } k = 0, 1, 2, \ldots \tag{7.22}$$

So by the teorem of Frechet-Shohat [11], (1. 8) and (7.22) we deduce that there is a sequence $x_i \to \infty$ such that

$$0 \le \lim_{x_i \to \infty} \int_0^{\infty} v^{k-2} dK_{x_i}(v) = \int_0^{\infty} v^{k-2} K(v) << \frac{k!}{R^k}, \text{ for } k = 0, 1, 2, \ldots$$

Therefore the series

$$\sum_{k=2}^{\infty} \frac{z^k}{k!} \int_0^{\infty} v^{k-2} dK(v)$$

represents an analytic function in $|z| < R$ and so (1.14) follows. Finally (1.15) is a consequence of (1.14), (7.18) and part (c) of Lemma 1. Thus we have established Theorem 1 for $f_y(n)$ and from this want to show that Theorem 1 holds for $f(n)$. Since the moments determine the limiting distribution uniquely in this case, it suffices to check (1.10) for $f(n)$. For this Lemma 4(b) will be useful.

Consider the decomposition

$$f(n) - B_1(x) = \{f_y(n) - B_1(y)\} + \{B_1(y) - B_1(x)\} + \{f(n) - f_y(n)\} \tag{7.23}$$

and identify the quantities on the right of (7.23) with those of Lemma 4 (b) as follows:

$$E = \{p \mid y \le p \le x\}, \quad f(n) - f_y(n) = f^*(n), \quad B_1(x) - B_1(y) = B_1^*(x).$$

The multinomial expansion and (7.23) yield

$$\sum_{n\in S(x)} a_n (f(n) - B_1(x))^k = \sum_{n\in S(x)} a_n (f_y(n) - B_1(y))^k +$$

$$\sum_{n\in S(x)} a_n \sum_{\substack{r_1+r_2+r_3=k \\ r_1 \le k-1}} \{f_y(n) - B_1(y)\}^{r_1} (-B_1^*(x))^{r_2} f^*(n)^{r_3}. \tag{7.24}$$

Theorem 1 for $f_y(n)$ implies that the first term on the right of (7.24) is

$$m_k(1 + o(1)).X.B_2(y)^{k/2}. \tag{7.25}$$

By the Cauchy-Schwartz inequality the second term on the right of (7.24) is

$$0\left(\sum_{\substack{r_1+r_2+r_3=k \\ r_1 \le k-1}} \left\{\sum_{n\in S(x)} a_n (f_y(n) - B_1(y))^{2r_1}\right\}^{1/2} \left\{\sum_{n\in S(x)} B_1^*(x)^{2r_2} a_n f^*(n)^{2r_3}\right\}^{1/2}\right) \tag{7.26}$$

By virtue of Lemma 4 (b), and Theorem 1 for $f_y(n)$, we deduce that the quantity in (7.26) is

$$\le \sum_{r_1+r_2+r_3=k,\, r_1 \le k-1} 0(X.B_2(y)^{r/2})\ o(B_2(x)^{(r_2+r_3)/2}) = o(XB_2(x)^{k/2}) \tag{7.27}$$

because $r_2 + r_3 \ge 1$. So from (7.24), (7.25), (7.26) and (7.27) we arrive at

$$\sum_{n\in S(x)} a_n \{f(n) - B_1(x)\}^k = m_k.X.B_2(y)^{k/2} + o(X.B_2(x)^{k/2}). \tag{7.28}$$

Now (1.10) follows from (7.28) because $B_2(x) \sim B_2(y)$, as $x \longrightarrow \infty$.

Theorem 1 is proved.

Remarks : The above proof shows that there is a one-to-one correspondence between the probability distributions K(v) in (1. 8) and F(v) in (1.13). This correspondence is set up by the moments m_k which satisfy (1.12) and (1.15). In particular the moments corresponding to K(v) in (1.17) are

$$m_k = \begin{cases} \dfrac{k!}{(\frac{k}{2})!\,2^{k/2}} & \text{if } k \text{ is even} \\ 0 & \text{if } k \text{ is odd} \end{cases} \tag{7.29}$$

and these are the moments of the Gaussian distribution G(v) in (1.16), establishing Corollary 1. Corollary 2 is a special case of the sufficiency part of Corollary 1.

§8. Proof of Theorem 3

We begin with

Lemma 5 : *Let S be special, $f > 0$ be strongly additive, and $\delta > 0$. Let E be the set of primes $\leq x$, for which $f(p) \geq \delta\sqrt{B_2(x)}$. Then*

$$B_0^*(x) = \sum_{\substack{p\leq x\\ p\in E}} \frac{\omega(p)}{p} \leq \delta^{-2} .$$

Proof : We have

$$\sum_{\substack{p\leq x\\ p\in E}} \frac{\omega(p)}{p} \leq \sum_{\substack{p<x\\ p\in E}} \frac{\omega(p)f^2(p)}{p\,\delta^2 B_2(x)} \leq \frac{1}{\delta^2 B_2(x)} \sum_{p\leq x} \frac{f^2(p)\omega(p)}{p} = \delta^{-2}$$

as claimed.

Proof of Theorem 3 : Let $k \geq 0$ be an integer and $f > 0$ strongly additive. For special S, we let E be as in Lemma 5 with $\delta = 1$. Define

$$f^*(n) = \sum_{\substack{p\mid n\\ p\in E}} f(p), \quad \bar{f}(n) = \sum_{\substack{p\mid n\\ p\notin E}} f(p) \tag{8.1}$$

and

$$B_1^*(x) = \sum_{\substack{p\leq x\\ p\in E}} \frac{f(p)\omega(p)}{p}, \quad \bar{B}_1(x) = \sum_{\substack{p\leq x\\ p\notin E}} \frac{f(p)\omega(p)}{p} . \tag{8.2}$$

If 'a' and 'b' are arbitrary complex numbers, note that

$$|a+b|^k \ll_k |a|^k + |b|^k. \tag{8.3}$$

In particular from (8. 1), (8. 2) and (8. 3) we have

$$\sum_{n\in S(x)} a_n |f(n)-B_1(x)|^k \ll_k \sum_{n\in S(x)} a_n |\overline{f}(n)-\overline{B}_1(x)|^k +$$

$$+ \sum_{n\in S(x)} a_n |f^*(n)-B_1^*(x)|^k \ll_k \sum\nolimits_5 + \sum\nolimits_6 \text{ respectively} \tag{8.4}$$

To estimate $\sum_5$, the ideas of §§5 and 7 apply. We start with estimates (5. 6) and (5.10), and observe also that (1. 9) holds trivially for $\overline{f}$. So the arguments underlying (7.16) and (7.21) show that the expression in (7. 2) (suitably defined for $\overline{f}$) with $y = x$ is true. Consequently by (3. 2) of Lemma 1 (a) we obtain

$$\sum\nolimits_5 \ll_k X.B_2(x)^{k/2} . \tag{8.5}$$

We have omitted repeating the details in the derivation of (8. 5).

To estimate $\sum_6$, observe first that (8. 3) yields

$$\sum\nolimits_6 \ll_k \sum_{n\in S(x)} a_n\{f^*(n)^k+B_1^*(x)^k\} \ll_k X.B_1^*(x)^k + \sum_{n\in S(x)} a_n f^*(n)^k. \tag{8.6}$$

From (6. 4) and Lemma 5 we get

$$B_1^*(x)^k \leq B_k(x)B_0^*(x)^{k-1} \leq B_k(x). \tag{8.7}$$

Similarly by Lemma 4 (a) and Lemma 5

$$\sum_{n\in S(x)} a_n f^*(n)^k \ll_k X.B_k(x)B_0^*(x)^{k-1} \ll_k X.B_k(x). \tag{8.8}$$

Theorem 3 for $f \geq 0$ and k a non-negative integer, follows from inequalities (8. 4) through (8. 8).

Suppose f is a real valued strongly additive function. We decompose f into

$$f = f^+ - f^- , \tag{8.9}$$

where f^+ and f^- are strongly additive functions generated by

$$f^+(p) = \max\ (0,f(p)), \quad f^-(p) = -\min\ (0,f(p)). \tag{8.10}$$

Theorem 3 for real f follows from (8. 3), (8. 9) and (8.10), because f^+ and f^- are non-negative strongly additive functions. If f is complex valued, we break it into its real and imaginary parts and once again apply (8. 3). Thus we can establish Theorem 3 for arbitrary complex f, and all non-negative integers k. To pass from non-negative integral exponents k to arbitrary non- negative exponents, we simply have to apply the Holder-

Minkowski inequality and (8. 1) suitably. The proof of Theorem 3 is complete.

§9. Examples

First we give examples of special sets S.

Eg 1: $S = \mathbb{Z}^+$, $A = 1$.

Eg 2: An arithmetic progression $an+b$ of positive integers, where $(a, b) = 1$, and $A = 1$.

Eg 3: The set S of positive integers for which $\lambda(n) = 1$ (or $\lambda(n) = -1$), with $A = 1$, where $\lambda(n) = (-1)^{\Omega(n)}$, $\Omega(n)$ being the number of prime divisors of n counted with multiplicity.

In the above examples the sets S have constant (non-zero density) and $A = 1$. In the next examples S will have zero density and the $\{a_n\}$ are unbounded.

Eg 4: Let $\mathcal{P}$ be a set of primes such that

$$\sum_{p \le x, p \in \mathcal{P}} 1 = \varkappa \sum_{p \le x} 1 + O(xe^{-\sqrt{\log x}}), \quad 0 < \varkappa < 1. \tag{9.1}$$

Let S_3 be the multiplicative semigroup generated by $\mathcal{P}$; that is

$$S_3 = \{n \in \mathbb{Z}^+ \,|\, p|n \longrightarrow p \in \mathcal{P}\}.$$

For $n \in S_3$, let $a_n = \varkappa^{-\Omega(n)}$. Then S_3 with these weights $\{a_n\}$ is special. Note that S_3 is the collection of integers relatively prime to the elements of $\bar{\mathcal{P}}$. So by (4.13'), with g chosen to be the characteristic function of S_3 we have

$$S_3(x) \ll X \prod_{\substack{p \le x \\ p \in \bar{\mathcal{P}}}} (1 - \frac{1}{p}) \ll \frac{x}{(\log x)^{1-\varkappa}} .$$

Hence S_3 is of zero density.

Eg 5: Let S_4 be the set of integers representable as a sum of two squares and $a_n = r(n)$, the number of such representations. Then S_4 is special. This example is related to example 5 in the case $\mathcal{P} = \{p \,|\, p \equiv 1 \pmod 4\}$, and $\varkappa = 1/2$.

An obvious way to construct $\mathcal{P}$ satisfying (9. 1) is to consider the set of all primes lying in one of several arithmetic progressions. In example 4, one may also think of $\mathbb{Z}^+$ as the underlying special set with weights a_n defined to be the totally multiplicative function given by

$$a_p = \varkappa^{-1} \text{ if } p \in \mathcal{P}, \; a_p = 0 \text{ if } p \notin \mathcal{P} .$$

This suggests that in examples 1,2 and 3, the condition $A = 1$ could be relaxed somewhat. For instance, with $S = \mathbb{Z}^+$ we could consider any sequence of weights $\{a_n\}$ that are totally multiplicative, for which

$$\sum_{n \leq x} a_n = cx + O(xe^{-\sqrt{\log x}}), \quad c > 0.$$

Also, one can construct special sets by suitable combinations of the above examples. Although in the above cases the weights turned out to be multiplicative, this condition is not required for the validity of our results.

Next we consider certain functions belonging to the class $\mathcal{A}(S)$. We begin by observing that the validity of condition (iv) for $k = 2$ implies that there is $\alpha' < \alpha$ such that $\alpha' \longrightarrow \infty$ with x and

$$\{B_2(x) - B_2(y')\} \log \alpha' = o(B_2(x)). \tag{9.2}$$

So by the Cauchy-Schwartz inequality and (9. 2) we have

$$|B_1(x) - B_1(y')| \leq \sqrt{\{B_2(x)-B_2(y')\} \log \alpha'} = o(\sqrt{B_2(x)}. \tag{9.3}$$

Suppose we have

$$\{\max_{p \leq x} f(p)\} \sqrt{B_2(x)} << 1. \tag{9.4}$$

Then for integers $k \geq 2$

$$B_k(x) - B_k(y') = \sum_{y' \leq p \leq x} \frac{|f(p)|^k \omega(p)}{p} << \sqrt{B_2(x)^{k-2}(B_2(x)-B_2(y'))}$$

$$= o(B_2(x)^{k/2}). \tag{9.5}$$

Hence from (9. 3) and (9. 5) we see that if f satisfies (9. 4), then the validity of condition (iv) for all $k \in \mathbb{Z}^+$ is a consequence of the truth that statement for $k = 2$. Delange [4] assumed (9. 4) (which is stronger than our condition (iii)) and thus required condition (iv) only for $k = 2$. The nice feature of Kubilius' method is that one requires (iv) only for $k = 2$, and simultaneously can dispense with assumptions such as Delange's (9. 4), or our condition (iii).

The functions f satisfying (1.18) of Corollary 2 have their limiting distribution as the Gaussian law. The following example due to Kubilius fits into the conditions of our Theorem 1 and provides an instance where $B_2(x) \longrightarrow \infty$ and the limit law $F(v) \neq G(v)$. For this, consider a set of primes E for which

$$E(x) \sim \frac{x}{\log x . \log \log x} .$$

Next let f be a strongly additive function given by

$$f(p) = \begin{cases} \log \log p & \text{for} \quad p \in E \\ 0 & \text{for} \quad p \notin E . \end{cases} \tag{9.6}$$

It is easily checked that f satisfies (9. 4) and belongs to $\mathcal{A}(\mathbb{Z}^+)$, with $A = 1$. The function $K(v)$ in (1. 8') is

$$K(v) = \begin{cases} 0 & \text{if} \quad v \leq 0 \\ \frac{v^2}{2} & \text{if} \quad 0 \leq v \leq \sqrt{2} \\ 1 & \text{if} \quad v \leq \sqrt{2} \end{cases} \tag{9.7}$$

and hence by Corollary 1, the limiting distribution is not Gaussian.

In this example of Kubilius, (9. 4) holds, and so assumption (1. 9) in Theorem 1 is redundant, since the $K_x(v)$ are all supported inside a fixed compact interval. This example can be modified slightly so that (9. 4) does not hold, in which case one has to check that (1. 9) is true. For instance, on a subset $E_1 \leq E$ for which $E_1(x) = 0(\sqrt{x})$ we can set

$$f(p) = \log p . \log \log p, \; p \in E_1$$

and let $f(p)$ be given by (9. 6) when $p \notin E_1$. Here condition (iii) holds when asymptotic equality as $x \longrightarrow \infty$. Or we may consider a subset $E_2 \subseteq E$ of larger size, say $E_2(x) \sim x/(\log x)(\log \log \log x)^2$, and set

$$f(p) = \log \log p . \log \log \log p, \; p \in E_2 .$$

Here again we let $f(p)$ be as in (9. 6) for $p \notin E_2$. In both these cases we have

$$\sum_{\substack{p \leq x \\ p \in E_i}} \frac{f(p)^k}{p} = o(B_2(x)^{k/2}), \quad i = 1, 2, \quad \text{and} \quad k \in \mathbb{Z}^+ .$$

Therefore, $K(v)$ is as in (9. 7). Here the $K_x(v)$ are not supported in a compact interval and so (1. 9) is not obvious. But it can be checked by some computation.

There are many examples of functions fitting the hypotheses of Theorem 2 - such as $-\log\{\varphi(n)/n\}$, where $\varphi(n)$ is Euler's function. Here also one can construct examples where verification of (1.20) is non-trivial; we omit repeating ideas that have just been considered above.

§10. Necessity of the hypotheses

In §9 we considered functions satisfying our growth conditions (iii), (iv) and (1. 9). But are such conditions necessary for the validity of our results ? If not, to what extent can they be relaxed ? Although we do not answer these questions here, we shall briefly give reasons which indicate that such conditions are perhaps necessary.

Condition (iii) enters naturally to keep the right hand side of (7.19) bounded for $0 \leq u \leq R$. If there was a sequence of primes p_j for which $f(p_j)/(\log p_j)\sqrt{B_2(p_j)} \longrightarrow \infty$, then $1 + \frac{h^*(p_j)}{p_j}$ would also tend to infinity for every $u > 0$. On the other hand $e^{-uf(p_j)\omega(p_j)/p_j\sqrt{B_2(p_j)}}$ could not compensate for the size of $1 + \frac{h^*(p_j)}{p_j}$, and so the right side of (7.19) will not remain bounded for $0 \leq u \leq R_0$. Similarly, (1. 9) is also required to keep the right sides of (7. 3) and (7.19) bounded for $0 \leq u \leq R_0$.

One may remark at this point that it is not necessary to bound for instance the right side of (7.19) since we are only after $X^{-1}T_u(x, y)$, which could be much smaller in size, when $h^*(p)$ is large. But then Lemma 5 shows that not many values of $f(p)$ can be large. In fact, most $f(p)$, for $p \leq y$ are $<<\sqrt{B_2(y)}$ in size, in which case $h^*(p)$ will be bounded for $0 \leq u \leq R_0$. Thus the right and left sides of (7.19) will be comparable, and so not much is lost in trying to bound the right side.

So, these growth conditions guarantee that $X^{-1}T_u(x, y)$ remains bounded for $|u| \leq R_0$; in other words, they keep the bilateral Laplace transforms bounded. One may also note that for the limit of the moments to exist, it is not necessary to have the sequence of bilateral Laplace transforms bounded for any u. But then, we do require the moments to determine the limiting distribution uniquely. This is often achieved by showing that the moments do not grow too fast. In such cases it is quite possible that the bilateral Laplace transform will be uniformly bounded in compact intervals of u.

Finally, condition (iv) is a natural requirement to bridge the gap between $f_y(n)$ and $f(n)$. It seems unlikely that one could directly establish results of our type for $f(n)$, without considering $f_y(n)$. For, asymptotic formulae like our (4.12) (or similar ones considered by Erdos-Kac [8] and Kubilius [17],[18]) which were employed in the proofs, hold only when $\alpha \longrightarrow \infty$ with x.

Recall that the method of Kubilius shows that for the class $\mathcal{H}$ of strongly additive functions, (1. 8) is both necessary and sufficient for (1.13) to hold. The functions of the class $\mathcal{A}$ for which $B_2(x) \longrightarrow \infty$, all belong to the class $\mathcal{H}$. Thus it is clear that (1. 8) is necessary for the validity of Theorem 1.

So on the basis of classical results we feel that our growth conditions cannot be weakened considerably. It might be worthwhile to pursue this aspect further. Perhaps they could be replaced by weaker conditions of an average type, but it seems unlikely that one could do away with any of these assumptions.

The assumption that f is strongly additive has been made mainly for convenience and can be removed in most situations quite easily. We briefly indicate how one may use our method to tackle general additive functions also, for special sets, where $\omega(d) \ll A_0^{\Omega(d)}$ for some $A_0 > 1$.

Suppose $\hat{f}$ denotes a generic complex additive function and $f(n) = \sum_{p|n} \hat{f}(p)$ the strongly additive function generated by $\hat{f}$. Then $\tilde{f} = \hat{f} - f$ is an additive function that vanishes on all primes. So

$$\tilde{f}(n) = \sum_{p^e || n, e \geq 2} \tilde{f}(p^e) \quad ,$$

where $p^e || n$ means that $p^e | n$ and $p^{e+1} \nmid n$. Note that

$$\sum_{e=2}^{\infty} \sum_{p} \frac{\omega(p^e)}{p^e} < \infty \quad . \tag{10.1}$$

Therefore the method underlying Lemma 4 (a) gives

$$\sum_{n \in S(x)} a_n |\tilde{f}(n)|^k \ll X\tilde{B}_k(x) \tag{10.2}$$

where

$$\tilde{B}_k(x) = \sum_{\substack{p^e \leq x \\ e \geq 2}} \frac{|\tilde{f}(p^e)|^k \ (p^e)}{p^e} \quad , \qquad k = 0, 1, 2, \ldots \quad .$$

Similarly, define

$$\hat{B}_1(x) = \sum_{p^e \leq x} \frac{\hat{f}(p \ )\omega(p \ )}{p^e} \quad \text{and} \quad B_k(x) = \sum_{p^e \leq x} \frac{|f(p^e)|^k \omega(p^e)}{} \quad \text{for} \quad k \geq 2$$

and consider the decomposition

$$\hat{f}(n) - \hat{B}_1(x) = (f(n) - B_1(x)) + O(|\tilde{f}(n)| + \tilde{B}_1(x)) . \tag{10.3}$$

Then from (8. 3), (10.1), (10.2), (10.3) and method of proof of Theorem 3, we can show that

$$\sum_{n\in S(x)} a_n|\hat{f}(n) - \hat{B}_1(x)| \ll_k \begin{cases} X\ \hat{B}_2(x)^{k/2} & \text{for} \quad 1 \le k \le 2 \\ X\ \{\hat{B}_2(x)^{k/2} + \hat{B}_k(x)\} & \text{for} \quad k > 2 \end{cases} \tag{10.4}$$

for all $\hat{f} \geq 0$, and from this the truth of (10.4) for all complex $\hat{f}$ follows.

Similarly by enlarging the clause $\mathcal{A}$ suitably to admit certain additive functions, the restriction that f is strongly additive can be removed in Theorem 1. From (10.3) we obtain, as in (7.24), by multinomial theorem

$$\sum_{n\in S(x)} a_n(\hat{f}(n)-\hat{B}_1(x))^k = \sum_{n\in S(x)} a_n(f(n)-B_1(x))^k + \\ + O\left(\sum_{n\in S(x)} a_n \sum_{\substack{r_1+r_2+r_3=k \\ r_1\le k-1}} (f(n)-B_1(x))^{r_1}|\tilde{f}(n)|^{r_2}\tilde{B}_1(x)^{r_3}\right). \tag{10.5}$$

An estimate on the size of the first term on the right of (10.5) is provided by Theorem 1. The contribution due to the last term on the right of (10.5) is negligible, since $B_2(x) \longrightarrow \infty$. The argument is similar to what we had in (7.27).

§11. Extension to real valued functions

The statements of Theorems 1 and 2 are restricted to functions $f \geq 0$ of the class $\mathcal{A}$. Obviously these results hold for similar f which are ≤ 0. The classical results mentioned in §2 do not involve such a restriction and so it is desirable to bring real valued $f \in \mathcal{A}$ within the scope of Theorems 1 and 2.

It is the way our method was set up that forced this restriction. More precisely we are able to estimate $T_u(x, y)$ satisfactorily only when the underlying multiplicative functions have their range in either $[0, 1]$ or in $[1,\infty]$. In order to treat real valued $f \in \mathcal{A}$ we need similar estimates even when these multiplicative functions have their range in $[0, \infty]$; this at present seems difficult. However, we do have an 'indirect' argument based upon our method and that of Kubilius which enables us to remove the restriction $f \geq 0$ in Theorems 1 and 2. Previously I had noticed this argument only when the limiting distributions were continuous but Professor Elliott has recently pointed out to me that this continuity condition is unnecessary; my earlier argument has to be modified slightly. We outline the idea briefly, below.

Lemma 6 : *Let $\varphi_x(v)$ be a sequence of probability distributions and $\varphi(v)$ a probability* *distribution such that*

$$\varphi_x(v) \to \varphi(v) \quad \textit{weakly in } v, \textit{ as } x \to \infty . \tag{11.1}$$

In addition, suppose that

$$\int_{-\infty}^{\infty} v^k \, d\varphi_x(v) \ll_k 1, \qquad \forall\, x, k \geq 1 . \tag{11.2}$$

Then

$$\lim_{x\to\infty} \int_{-\infty}^{\infty} v^k \, d\varphi_x(v) = \int_{-\infty}^{\infty} v^k \, d\varphi(v), \quad \textit{for } k = 1, 2, 3, \ldots \tag{11.3}$$

Proof : First by (11.2) and the theorem of Frechet-Shohat, we note that there is a sub-sequence $\{x_l\}$ such that (11.3) holds if the limit is taken over this sub-sequence. In particular all the moments of φ exist and are finite. Observe that (11.2) also yields

$$\int_{|v|\geq T} |v^k| \, d\varphi_x(v) \leq \frac{1}{T^k} \int_{|v|\geq T} v^{2k} \, d\varphi_x(v) \ll_k T^{-k}, \quad \forall\, x, T, k \geq 1. \tag{11.4}$$

Since all moments of φ are finite, (11.4) holds with φ in place of φ_x .

We let

$$\varphi_x(v) = \varphi(v) + e_x(v) , \tag{11.5}$$

and choose T_1 , $T_2 > 1$ so that $-T_1$ and T_2 are points of continuity for φ . Then by (11.4) and (11.5) we have

$$\begin{aligned}\int_{-\infty}^{\infty} v^k \, d\varphi_x(v) &= \int_{-T_1}^{T_2} v^k \, d\varphi_x(v) + 0_k(T_1^{-k} + T_2^{-k}) \\ &= \int_{-T_1}^{T_2} v^k \, d\varphi(v) + \int_{-T_1}^{T_2} v^k \, de_x(v) + 0_k(T_1^{-k} + T_2^{-k}) \\ &= \int_{-\infty}^{\infty} v^k \, d\varphi(v) + \int_{-T_1}^{T_2} v^k \, de_x(v) + 0_k(T_1^{-k} + T_2^{-k}) .\end{aligned} \tag{11.6}$$

We use integration-by-parts on the second integral of the right hand side of (11.6) and get

$$\int_{-T_1}^{T_2} v^k \, de_x(v) = v^k \, de_x(v) \Big|_{-T_1}^{T_2} - \int_{-T_1}^{T_2} e_x(v) k v^{k-1} dv . \tag{11.7}$$

By our assumption on T_1, T_2 and (11.1) we see that for each fixed T_1, T_2 the quantity in (11.7) goes to zero as $x \to \infty$. Thus from (12.6) we get

$$\overline{\lim} \int_{\infty}^{\infty} v^k \, d\varphi_x(v) \leq \int_{-\infty}^{\infty} v^k \, d\varphi(v) + 0_k(T_1^{-k} + T_2^{-k}) \tag{11.8}$$

with a similar lower bound for $\underline{\lim}$. Since $T_1, T_2 > 1$ could be arbitrarily large, Lemma 6 follows from (11.8).

Suppose $f \in \mathcal{A}$ is real valued and satisfies (1. 9). We write $f = f^+ - f^-$ and apply estimates (4.13) and (5.10) on f^+ and f^-. Then by the ideas underlying the proof of Theorem 1 and (8. 3) we get

$$\sum_{n \in S(x)} a_n |f(n) - B_1(x)|^k \ll_k XB_2(x)^{k/2}, \quad k = 1, 2, 3, \ldots \quad . \quad (11.9)$$

Suppose f satisfies (1. 8) and $B_2(x) \to \infty$. Then by the method of Kubilius, we see (1.13) will hold. This is to be compared with (11.1) of Lemma 6. Then (11.9) is the analogue of (11.2) and so (11.3) in Lemma 6 shows that (1.10) holds; consequently all other assertions of Theorem 1 hold as well.

The above argument is unsatisfactory for the following reasons. Although based upon our method, we have to lean upon the method of Kubilius and so it is an 'indirect proof'; the chances of a direct argument seem, at present, remote. We have also gone against the general philosophy of this paper by estimating the moments asymptotically after evaluating the weak limit of the distribution functions first. Our aim here has been to draw information about the distribution functions from a study of their moments. At any rate it is good to know that Theorems 1 and 2 could be extended to admit real valued $f \in \mathcal{A}$. In addition, our arguments here are interesting for their own sake and have other uses also. We shall in fact make crucial use of Lemma 6 later, while investigating the distribution of additive functions in S_1 and S_2.

§12. Elliott's method

In §§0 and 2 it was pointed out that that the sieve and the bilateral Laplace transform combine to yield an improvement of the method employed by Elliott in [6] to derive (2. 6). Since no description of Elliott's ideas was given there we feel it is appropriate to do so now and also mention certain similarities and other essential differences between his technique and ours.

Elliott begins by considering for $f \geq 0$ the sum of $T_z(x)$ with complex z and uses a contour integral, as in (3.19), to bound the moments suitably, in the situation $S = \mathbb{Z}^+$, $A = 1$. He notices that $f \geq 0$ implies $|T_z(x)| \leq T_u(x)$, where $u = \mathrm{Re}(z)$. Since he was not after asymptotics, he had no requirement for the analysis of bilateral Laplace transforms.

His treatment of Case 1 (see §5) for T_u is simpler compared to ours for two reasons :

(1) He was interested only in upper bounds and so required only the weaker estimate (5. 6').

(2) For $S = \mathbb{Z}^+$, $A = 1$ one does not encounter certain difficulties one experiences with subsets S of $\mathbb{Z}^+$ and their associated weights $\{a_n\}$.

Thus he had no need for the powerful combinatorial sieve method. In fact, he takes care of Case 1 by using a result of Hall [12], which provides for him the required bound (5. 6').

Our treatment of Case 2 in this paper is similar to Elliott's who also employs the function $h^*(d)$ to derive (5.10') with $y = x$. Later, when we consider a larger class of sets satisfying a condition weaker than our (i) we will have to deal carefully with sums

$$\sum h(d)\,|R_d(x)|\,\omega(d)$$

for large divisors d. This is the situation for example with sets S_1 and S_2. The problem will then take a different point of view, for Case 2 will become the more complicated one since Case 1 can be controlled by the combinatorial sieve!

Using bounds for $T_u(x)$ in an interval $|u| \leq r$, Elliott arrives at (11.9') when $f(p) << \sqrt{B_2(x)}$. With regard to large values of $f(p)$ he first obtains a uniform upper bound for

$$\sum_{n \leq x} \nu(n,E)^k$$

by induction on k, where

$$\nu(n,E) = \sum_{p|n,\, p\in E} 1 \quad .$$

Then he observes that for any strongly additive function $f \geq 0$

$$f^*(n)^k = \left(\sum_{p|n,\, p\in E} f(p)\right)^k \leq \left(\sum_{p|n,\, p\in E} f(p)^k\right)\nu(n,E)^{k/(k-1)},$$

and employs what we have called Lemma 5 to care of this contribution. Such an approach when applied to special sets would lead us to consider subsets of S comprising of multiples of products of primes $p_1, \ldots, p_j$, $j \leq k$ and determine whether such subsets satisfy (i). This being a little awkward to handle, we prefer to obtain directly a bound as in Lemma 4 (a) and hence our treatment of large values of f(p) is slightly different.

As is customary in such problems, Elliott observes that it suffices to establish (2. 6') for $f \geq 0$, because (8. 3) will enable one to extend such a result first to real valued f, and then to all complex additive functions.

Finally, Elliott also discusses certain consequences and applications of (2. 6). In particular, he obtains a necessary and sufficient condition for the mean

$$\frac{1}{x} \sum_{n \leq x} |f(n)|^k$$

of an additive function f(n) to be bounded. By suitably modifying his arguments we can discuss similar consequences of Theorem 3. Our exposition has already become long, and so we refer to the reader to [6] for the basic ideas in this direction.

§13 Turán's thesis

Inequality (2. 1) was established by Turán in his Ph.D. thesis [21] (Hungarian). It was also published in [22] since it provided a simple proof of the Hardy-Ramanujan theorem, that $\nu(n)$ is almost always nearly log log n in size. In his thesis using analytic tools he also showed that

$$\sum_{1 \leq n \leq x} 2^{r\nu(n)} = c_r x(\log x)^{2^r - 1} \left\{ 1 + 0\left(\frac{1}{\log x}\right)\right\} \tag{13.1}$$

holds uniformly for $-1/2 \leq r \leq 1/2$. By choosing r appropriately in (13.1) he gave yet another proof of the Hardy-Ramanujan result. Probably he considered this last proof somewhat complicated since he derived it only in his thesis and did not publish it elsewhere.

On the basis of (2. 1) Kac suggested, first in a letter to Turán (see [5], Vol. 2, p.18), and later in [16], that one should try to estimate the quantity

$$x^{-1} \sum_{n \leq x} \{\nu(n) - \log \log x\}^k, \quad k = 1, 2, 3, \ldots \quad . \tag{13.2}$$

Turán felt that this could be done on the basis of the method underlying (2. 1) but did not go through the effort since he saw no use for it. It appears that he had Halberstam's method in mind and did not carry out the computations owing to the complications. An asymptotic estimate for the quantity in (13.2) was first provided by Delange [2] who used generating functions and analytic methods.

Later Turán along with Rényi [19] considered the sum (2. 3) for $z = e^{iu}$ and by interpreting it in terms of Fourier transforms established (1.13') for $f(n) = \nu(n)$ with a best estimate for the rate of convergence. If one does not care about error terms but only wants to check (1.13'), then such a result for $\nu(n)$ follows from an asymptotic estimate for (13.2) which Turán could have deduced from (13.1). We outline this argument briefly below.

For real u consider the expression

$$\sum_{n \le x} e^{u\left(\frac{\nu(n) - \log\log x}{\sqrt{\log\log x}}\right)} = e^{-u\sqrt{\log\log x}} \sum_{n \le x} e^{u\nu(n)/\sqrt{\log\log x}},$$

and estimate this asymptotically using (13.1) with r chosen to make

$$2^r = e^{u/\sqrt{\log\log x}}.$$

This estimate can be interpreted as an estimate for the bilateral Laplace transform. Then Lemma 1 shows, after suitable identification, that the expression in (13.2) is equal to $\{m_k + o(1)\}(\log\log x)^{k/2}$, where m_k is as in (7.29), and hence the limiting distribution is Gaussian.

Turán could have estimated (13.2) this way and therefore anticipated the results of Delange [2], Erdős-Kac [8] and Halberstam [13]. Maybe then he would have felt motivated to publish (13.1) separately!

References

1. Alladi, K. The distribution of $\nu(n)$ in the sieve of Eratosthenes, Quart. J. Math. Oxford (2), 33(1982), 129-148

2. Delange, H. Sur le nombre des diviseurs premiers de n, C.R.Acad. Sci. Paris, 237 (1953), 542-544.

3. Delange, H. Sur un théorème d'Erdős et Kac, Acad. Roy. Belg. Bull. Cl. Sci., 5, 42 (1956), 130-144.

4. Delange, H. Sur les fonctions arithmetiques fortement additives, C.R.Acad. Sci. Paris, 244 (1957), 2122-2124.

5. Elliott, P.D.T.A. Probabilistic Number Theory, Vol. 1 and 2, Grundelehren 239-240, Springer-Verlag, Berlin, New York (1980).

6. Elliott, P.D.T.A. High power analogues of the Turan-Kubilius inequality and an application to number theory, Can. J. Math., 32 (1980), 893-907.

7. Erdős, P. On the density of some sequences of numbers I, II, III, J. London Math. Soc., 10 (1935), 120-135; 12 (1937), 7-11; 13 (1938), 119-127.

8. Erdős, P. & Kac, M. The Gaussian law of errors in the theory of additive number theoretic functions, Amer. J. Math., 62 (1940), 738-742.

9. Erdős, P & Wintner, A. Additive arithmetical functions and statistical independence, Amer. J. Math., 61 (1939), 713-721.

10. Feller, W. An Introduction to Probability Theory and Its Applications, Vol. 2, John Wiley, New York (1966).

11. Frechet, M. & Shohat, J. A proof of the generalised second limit theorem in the theory of probability, Trans. Amer. Math. Soc., 33 (1931), 533-543.

12. Hall, R. R. Halving an estimate obtained from Selberg's upper bound method, Acta Arith., 25 (1974), 347-351.

13. Halberstam, H. On the distribution of additive number theoretic functions I, II, III, J. London Math. Soc., 30 (1955), 43-53; 31 (1956), 1-14; 31 (1956), 14-27.

14. Halberstam, H. & Richert, H. -E. Sieve Methods, Academic Press, London, New York (1974).

15. Hardy, G. H. & Ramanujan, S. On the normal number of prime factors of a number n, Quart. J. Math., Oxford, 48 (1917), 76-92.

16. Kac, M. Probability methods in some problems of analysis and number theory, Bull. Amer. Math. Soc., 55 (1949), 641-665.

17. Kubilius, J. Probabilistic methods in the theory of numbers, Uspekhi Math. Nauk., N.S. 11 (1956), 2 (68), 31-66 = Amer. Math. Soc., Translations 19 (1962), 47-85.

18. Kubilius, J. Probabilistic methods in the theory of numbers, Amer. Math. Soc. Translations of Mathematical Monographs, 11, Providence (1964).

19. Renyi, A. & Turan, P. On a theorem of Erdos-Kac, Acta Arith., 4 (1958), 71-84.

20. Selber, A. Note on a paper by L. G. Sathe, J. Indian Math. Soc., 18 (1954), 83-87.

21. Turan, P. Az egesz szamok primosztoinak szamarol, Mat. Lopok., 41 (1934), 103-130.

22. Turan, P. On a theorem of Hardy and Ramanujan, J. London Math. Soc., 9 (1934), 274-276.

School of Mathematics,
Institute for Advanced Study,
Princeton, New Jersey, U.S.A.

MATSCIENCE,
Institute of Mathematical Sciences,
Madras 600 113, INDIA.