Lecture Notes in Mathematics

Edited by A. Dold and B. Eckmann

1052

Number Theory

New York 1982

Edited by D. V. Chudnovsky, G. V. Chudnovsky, H. Cohn and M. B. Nathanson

Lecture Notes in Mathematics

1052

Number Theory

A Seminar held at the Graduate School and University Center of the City University of New York 1982

Edited by D.V. Chudnovsky, G.V. Chudnovsky, H. Cohn and M.B. Nathanson

Springer-Verlag
Berlin Heidelberg New York Tokyo 1984

Editors

David V. Chudnovsky
Gregory V. Chudnovsky
Department of Mathematics, Columbia University
New York, NY 10027, USA

Harvey Cohn
Department of Mathematics, C.U.N.Y. City College
New York, NY 10031, USA

Melvyn B. Nathanson
Department of Mathematics, Rutgers – The State University
Newark, NJ 07102, USA

AMS Subject Classification (1980): 10-06

ISBN 3-540-12909-X Springer-Verlag Berlin Heidelberg New York Tokyo
ISBN 0-387-12909-X Springer-Verlag New York Heidelberg Berlin Tokyo

Library of Congress Cataloging in Publication Data. Main entry under title: Number theory. (Lecture notes in mathematics; 1052) 1. Numbers, Theory of–Addresses, essays, lectures. I. Chudnovsky, D. (David), 1947-. II. New York Number Theory Seminar (1982–1983) III. Series: Lecture notes in mathematics (Springer-Verlag); 1052.
QA3.L28 no. 1052 [QA241] 510s [512'.7] 84-1360
ISBN 0-387-12909-X (U.S.)

Printing and binding: Beltz Offsetdruck, Hemsbach/Bergstr.
2146/3140-543210

INTRODUCTION

In January, 1982 four number theorists - David and Gregory Chudnovsky, Harvey Cohn, and Melvyn B. Nathanson - organized the New York Number Theory Seminar. The Seminar met weekly during the spring, 1982 semester at the Graduate School and University Center of the City University of New York at 11 West 42 Street in Manhattan. This volume contains expanded texts of the lectures delivered in the Seminar. The Seminar continued in the 1982-83 academic year, and the reports presented in this second year will be published in a subsequent volume.

The organizers hope that the New York Number Theory Seminar will provide a continuing opportunity to discuss recent results in the higher arithmetic, and that the publication of the annual proceedings will contribute to research in number theory.

TABLE OF CONTENTS

K. ALLADI , Moments of Additive Functions and Sieve Methods 1

H. COHEN and H.W. LENSTRA, JR. , Heuristics on Class Groups 26

D.V. CHUDNOVSKY and G.V. CHUDNOVSKY , Padé and Rational Approximations to Systems of Functions and their Arithmetic Applications 37

D.V. CHUDNOVSKY and G.V. CHUDNOVSKY , Padé Approximations to Solutions of Linear Differential Equations and Applications to Diophantine Analysis 85

J. DIAMOND , p- adic Gamma Functions and their Applications . . 168

J.C. LAGARIAS and A.M. ODLYZKO , New Algorithms for Computing $\pi(x)$ 176

J. LEPOWSKY and M. PRIMC , Standard Modules for Type One Affine Lie Algebras 194

C. MORENO , Some Problems of Effectivity in Arithmetic, Geometry and Analysis 252

M.B. NATHASON , The Exact Order of Subsets of Additive Bases 273

C.F. OSGOOD , Effective Bounds on the Diophantine Approximations of Algebraic Functions, and Nevanlinna Theory 278

P. SARNAK , Additive Number Theory and Maass Forms 286

Moments of additive functions and sieve methods

Krishnaswami Alladi*

§1. Introduction

We shall report here our recent work on the estimation of moments of additive functions $f(n)$, for values n restricted to certain subsets S of the positive integers. Although of relatively recent origin, additive functions have been an object of intense study in the past few decades because a number of impressive distribution results could be established using a variety of techniques. Most research however relates to the set $\mathbb{Z}^+$ of all positive integers and the problem concerning subsets S of $\mathbb{Z}^+$ has received little attention. The purpose of this exposition is to describe a new method that we employ, which enables us, amongst other things, to extend various classical results to certain subsets S; besides, our method may have other implications as well. As far as we know, it is the first occasion when the sieve method has been used in such moment problems and this results in a satisfactory treatment of a wide class of sets S.

So far the main interest in employing the moment method to study the distribution of additive functions has been due to its entirely elementary nature, but this in turn led to several tedious calculations. Our approach eliminates much of this difficulty by using the machinery of bilateral Laplace transforms.

Our paper is basically divided into two parts. Up to §4 we discuss several classical results and remark on their merits and limitations. From §5 to §10 we describe our method and results and compare these with earlier approaches. We state our results in §§8 and 9. Finally in §10 we briefly discuss limitations in our technique and indicate directions for further work and progress.

Our method is an improvement of a recent technique due to Elliott [6] who obtained uniform upper bounds for the moments of arbitrary additive functions, in the situation $S = \mathbb{Z}^+$. Later, in §9, we shall point out similarities and differences between Elliott's approach and ours.

We only discuss the main ideas here and not give details of proofs. A more complete treatment of our method can be found in [1]. For now, we conclude this section by collecting some notations and conventions.

Recall that an additive function $f(n)$ is an arithmetical function that satisfies $f(mn) = f(m) + f(n)$ for integers m, n with g.c.d. $(m,n) = 1$. Similarly a

---

*Talk given at the New York Number Theory Seminar on April 12, 1982.

multiplicative function g satisfies $g(mn) = g(m)g(n)$ for $(m,n) = 1$. Thus additive and multiplicative functions are completely determined by their values on prime powers p^e, $e \in \mathbb{Z}^+$. For simplicity we concentrate here on strongly additive functions f which satisfy

$$f(n) = \sum_{\substack{p|n \\ p=\text{prime}}} f(p) \,. \tag{1.1}$$

Similarly a strongly multiplicative function g is given by

$$g(n) = \prod_{p|n} g(p) \,. \tag{1.2}$$

As usual emply sums equal zero and empty products, one.

The sets S we discuss will satisfy some conditions imposed upon the quantity

$$S_d(x) = \sum_{\substack{n\leq x,\, n\in S \\ n\equiv 0 \pmod d}} a_n \,, \tag{1.3}$$

where $\{a_n\}$ is a sequence of 'weights' attached to S. The weights are ≥ 0 and we write

$$S_d(x) = \frac{X\omega(d)}{d} + R_d(x) \,, \tag{1.4}$$

where $X = S_1(x)$. We require that

$$\omega(d) \text{ is multiplicative, } \geq 0 \text{ and } \omega(p) \text{ is bounded} \,. \tag{1.5}$$

In addition we need

$$\text{"average of } |R_d(x)| \text{ is in some sense small"} \,. \tag{1.6}$$

We shall make (1.6) more precise in the sequel.

With S as above we associate with each strongly additive f, the sums

$$A(x) = A_f(x) = \sum_{p\leq x} \frac{f(p)\omega(p)}{p} \tag{1.7}$$

and

$$B_k(x) = \sum_{p\leq x} \frac{|f(p)|^k \omega(p)}{p} \,, \quad \text{for } k \geq 0 \,. \tag{1.8}$$

As usual the $\ll$ and 'O' notation are equivalent and will be used interchangeably as is convenient. Unless indicated otherwise, implicit constants are absolute, or depend at most upon S and this will be clear from the context. The Moebius function $\mu(n)$ is the multiplicative function given by $\mu(p) = -1$ for each prime p, and $\mu(p^e) = 0$ for all p and $e \geq 2$. We denote by $p(n)$ the smallest prime factor of n if $n > 1$, and $p(1) = \infty$. Finally $P(y) = \prod_{p \leq y} p$. Further notation will be introduced when needed.

§2. Classical distribution results

Despite the fact that prime numbers have been the focus of attention in Number Theory since Greek antiquity, the first significant results on $\nu(n)$, the number of prime divisors of n, were established only in the twentieth century. More precisely Hardy and Ramanujan [13] observed in 1917 that $\nu(n)$ is almost always nearly $\log\log n$ in size. This was really the starting point for the study of additive functions although the subject took shape only two decades later.

Intensive study began in 1934 when P. Turán [19] showed that

$$\sum_{n \leq x} (f(n) - A(x))^2 \ll xA(x) \tag{2.1}$$

holds for strongly additive functions f satisfying $0 \leq f(p) \ll 1$. In particular from (2.1) with $f(n) = \nu(n)$, the Hardy-Ramanujan result follows. But more importantly, for the first time, the role of probability theory in the study of additive functions could be perceived because (2.1) is essentially an estimate for the second moment (variance) of f.

This led Turán's Hungarian colleague Erdös to an active study of the mean values of additive functions [7; I, II, III]. The efforts of Erdös culminated in two major theorms which clearly set the study of additive functions on a firm probabilistic foundation. One of these established jointly with Wintner [9], provided necessary and sufficient conditions for the frequencies

$$\frac{1}{x} \sum_{\substack{n \leq x \\ f(n) \leq v}} 1$$

of an additive function to converge weakly (in v) as $x \to \infty$. The second, established jointly with Kac [8], considered the value $f(n)$, $1 \leq n \leq x$, in terms of sums of nearly independent random variables, one for each prime $p \leq x$. When compared to a sum of independent random variables, it showed under the influence of the Central Limit Theorem that

$$\lim_{x \to \infty} F_x(v) = G(v) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{v} e^{-u^2/2}\, du \tag{2.2}$$

for all real strongly additive f for which

$$f(p) = O(1) , \quad \text{and} \quad B_2(x) \to \infty \quad \text{with} \quad x , \tag{2.3}$$

where

$$F_x(v) = \frac{1}{x} \sum_{\substack{n \le x \\ f(n) - A(x) \le v\sqrt{B(x)}}} 1 \tag{2.4}$$

(In establishing (2.2) they made crucial use of a certain sieve estimate due to Brun; further light on this will be shed in §3.) Thus the study of additive functions (alternatively Probabilistic Number Theory), gained momentum.

For additive functions taking certain simple values on primes, Rényi and Turán [18] demonstrated that an alternate elegant treatment is possible. More precisely with $z = e^{iu}$, u real, they noticed that the sum of multiplicative functions

$$\sum_{n \le x} z^{f(n)} \tag{2.5}$$

could be estimated asymptotically by familiar analytic techniques. The sum in (2.5) is related to the characteristic function (Fourier transform) of the distribution $F_x(v)$ and hence leads to the determination of the limit of $F_x(v)$ as $x \to \infty$. In particular for $\nu(n)$ they used (2.5) to explicitly calculate the rate of convergence in (2.2).

The next major advance was made by Kubilius in 1956 [15], who successfully combined all of the above ideas (see also [16] for a more elaborate treatment). By employing several tools from probability theory such as infinitely divisible distributions, independent random variables and characteristic functions, he obtained necessary and sufficient conditions for the frequencies $F_x(v)$ in (2.4) to converge weakly (in v) as $x \to \infty$, for all functions of a certain class $\mathcal{H}$. This class comprised of all strongly additive functions for which $B_2(x) \to \infty$ with x and

$$\frac{B_2(x)}{B_2(y)} \to 1 \quad \text{as} \quad x \to \infty \quad \text{for some} \quad \alpha = \alpha(x) = \frac{\log x}{\log y} \to \infty . \tag{2.6}$$

He produced the first major examples of additive functions f with the limit in (2.2) different from the Gaussian distribution. Kubilius was able to determine the limit $\varphi(v)$ of the characteristic functions $\varphi_x(v)$ of $F_x(v)$ and thus, in principle, determined the weak limit of F_x.

In a somewhat different direction, it was realised soon after Turán established (2.1), that it would be useful to generalise it to arbitrary strongly additive functions. This was done by Kubilius in 1956 who showed that

$$\sum_{n\leq x} |f(n)-A(x)|^2 \ll xB_2(x) \tag{2.7}$$

holds uniformly for all such f. Special cases of (2.6) were used by Erdős-Kac, Erdős-Wintner and Kubilius to establish their distribution results. There will be more on this in the sequel.

§3. Brun's sieve in Probabilistic Number Theory

First we sketch the ideas due to Erdős-Kac.

For each prime p define a function

$$\rho_p(n) = \begin{cases} 1 & \text{if } p \mid n \\ 0 & \text{if } p \nmid n . \end{cases}$$

Then every strongly additive function can be written as

$$f(n) = \sum_p f(p)\rho_p(n) .$$

The functions ρ_p have the property that in the interval $[1,x]$

$$\rho_p(n) = \begin{cases} 1 & \text{with probability } \sim 1/p \\ 0 & \text{with probability } \sim (1-\frac{1}{p}) \end{cases}, \quad \text{for } p = o(x) .$$

Also, for $p = o(x)$, the $\rho_p(n)$ are nearly independent in $[1,x]$. The idea is to compare f(n) with the sum

$$X = \sum_p x_p$$

of infinite independent random variables defined by

$$x_p = \begin{cases} f(p) & \text{with probability } 1/p \\ 0 & \text{with probability } 1-\frac{1}{p} . \end{cases}$$

If $B_2(y) \to \infty$ with y, it follows that the random variable

$$X_y = \sum_{p\leq y} x_p \tag{3.1}$$

has mean and variance asymptotically equal to $A(y)$ and $\sqrt{B(y)}$ respectively. Furthermore if $f(p) = O(1)$, then by the Central Limit Theorem the random variable

$$\{X_y - A(y)\}/\sqrt{B(y)} \tag{3.2}$$

has a distribution function $\longrightarrow G(v)$ as $y \to \infty$.

It is only natural to compare the quantity in (3.2) with

$$\{f_y(n)-A(y)\}/\sqrt{B(y)}\ , \tag{3.3}$$

where $f_y(n)$ is the truncated additive function

$$f_y(n) = \sum_{\substack{p|n \\ p\le y}} f(p)\ . \tag{3.4}$$

While comparing (3.2) and (3.3), Erdös and Kac required an estimate for the quantity

$$\Phi(x,y) = \sum_{\substack{n\le x \\ p(n)>y}} 1\ .$$

On a probabilistic basis one expects that

$$\Phi(x,y) \sim x \prod_{p\le y} \left(1-\frac{1}{p}\right)\ , \quad \text{as} \quad x\to\infty\ . \tag{3.5}$$

In fact they observed that the random variables in (3.2) and (3.3) have asymptotically the same distribution function (the Gaussian) if (3.5) is true. Burn's sieve method shows that (3.5) holds so long as

$$\alpha = \log x/\log y \to \infty \quad \text{with} \quad x\ .$$

(In fact (3.5) is false if $\alpha \not\to \infty$ with x!)

To complete the transition from $f_y(n)$ to $f(n)$ they used Turán's inequality (2.1) to deduce that for the function $\bar{f} = f - f_y$

$$\sum_{n\le x} \{\bar{f}(n)-A_{\bar{f}}(x)\}^2 = o(xB_2(x)) \tag{3.6}$$

holds for a suitable choice of y and $\alpha\to\infty$, on the basis of (2.3).

Although it was implicit in the work of Erdös-Kac that their method applies more generally to sets $S \subseteq \mathbb{Z}^+$ where the analogue of Brun's estimate (3.5) holds, this was carried out only by Kubilius some years later. In doing so Kubilius used the method of characteristic functions and thus avoided the use of the Central Limit Theorem. This enabled him, amongst other things, to consider additive functions of growth faster than in (2.3) and therefore could obtain significant results with limiting distributions other than the Gaussian.

Consider a set S for which (1.4) and (1.5) hold. Assume also the following precise version of (1.6): For each b there exists c such that

$$\sum_{d\le X/\log^c X} |R_d(x)| \ll_b \frac{X}{\log^b X}\ . \tag{3.7}$$

It is known by the Combinatorial Sieve (i.e. Brun's sieve) method (see Halberstam-Richert [12], p. 83) that

$$S(x,y) = \sum_{\substack{n\leq x, n\in S \\ p(n)>y}} a_n \sim X \prod_{p\leq y} \left(1 - \frac{\omega(p)}{p}\right), \text{ if } \alpha \to \infty \text{ with } x . \qquad (3.8)$$

Therefore consider independent random variables

$$x_p = \begin{cases} f(p) & \text{with probability } \omega(p)/p \\ 0 & \text{with probability } 1 - \frac{\omega(p)}{p} . \end{cases}$$

Also let

$$F_{x,y}(v) = \frac{1}{x} \sum_{\substack{n\leq x, n\in S \\ f_y(n) - A(y) < v\sqrt{B_2(y)}}} a_n \text{ and set } F_{x,x}(v) = F_x(v) . \qquad (3.9)$$

Kubilius' first step was to show that if the distribution function of the random variable in (3.2) tends to a weak limit $F(v)$ as $y \to \infty$, then $F_{x,y}(v)$ also tends weakly to $F(v)$ as $x \to \infty$, provided $\alpha \to \infty$. This was a consequence of (3.8). In order to determine $F(v)$ he noted that it suffices to compute the limit of the characteristic functions φ_y of the random variables in (3.2). By combining the idea of infinitely divisible distributions with the independence of x_p he showed that

$$\lim_{y\to\infty} \varphi_y(v) = \exp\left\{\int_{-\infty}^{\infty} \frac{e^{iuv} - 1 - iuv}{u^2}\, dK(u)\right\} \qquad (3.10)$$

provided

$$\lim_{x\to\infty} \frac{1}{B_2(x)} \sum_{\substack{p\leq x \\ f(p)\leq u\sqrt{B_2(x)}}} \frac{f^2(p)\omega(p)}{p} = K(u) \text{ almost surely in } u . \qquad (3.11)$$

Ideally we want to make the transition from $f_y(n)$ to $f(n)$ by appealing to an upperbound of the sort

$$\sum_{n\leq x, n\in S} (f(n) - A_f(x))^2 a_n = o(XB_2(x)) , \text{ as } x \to \infty . \qquad (3.12)$$

But (3.12) is not at all easy to establish generally. If the $f(p)$ do not grow fast, namely, if

$$\{\max_{p\leq x} f(p)\}/\sqrt{B_2(x)} \to 0 \text{ as } x \to \infty , \qquad (3.13)$$

then (3.12) can be shown to be true provided we choose y and $\alpha \to \infty$ optimally. The limiting distribution then, is Gaussian.

In order to get different limiting distributions Kubilius needed to consider larger values of $f(p)$. More precisely he required the expression on the left of (3.13) to be bounded away from zero. In such a case, an analogue of (2.1) for S is insufficient to establish (3.12). It is for this reason that he required an inequality as general as (2.7) but then this is simply not true for subsets S of $\mathbb{Z}^+$. When $S = \mathbb{Z}^+$ there is no problem. So in this situation Kubilius used (3.10) and showed that (3.11) was necessary and sufficient for the weak limit of $F_x(v)$ to exist, for functions f of the class $\mathcal{H}$. The growth condition (2.6) arises naturally out of (2.7) for then

$$x^{-1} \sum_{n \le x} (\bar{f}(n) - A_{\bar{f}}(x))^2 \ll \{B_2(x) - B_2(y)\} = o(B_2(x)) . \tag{3.14}$$

On the other hand for subsets S Kubilius could only treat the Gaussian case satisfactorily, because then it was possible to prove (3.12) by choosing $y, \alpha \to \infty$ optimally, whereas one does not yet know how to establish an analogue of (3.14) for $n \in S$.

The methods of Erdős-Kac and Kubilius yield no information about the moments of f although they make crucial use of an upper bound for the "second moment" of $f(n) - f_y(n)$, about $A_{\bar{f}}(x)$. For some reason the estimation of moments has until now remained out of the sphere of the powerful methods of probabilistic number theory and sieve theory. But then our technique brings these moment estimates within the control of these sophisticated tools. Before we describe this, we first review some early work on the moments of additive functions, in the next section.

§4. The method of moments

Motivated by Turán's inequality (2.1) Mark Kac suggested, first in a letter to Turán (see Elliott [5], Vol. 2, p. 18) and later in an address to the American Mathematical Society [14] that one ought to try and estimate asymptotically the expression

$$\frac{1}{x(\log\log x)^{k/2}} \sum_{n \le x} (\nu(n) - \log\log x)^k \quad \text{for } k = 1, 2, 3, \ldots . \tag{4.1}$$

With the notation $F_x(v)$ in (2.4) now applied to $\nu(n)$, the quantity in (4.1) is

$$\int_{-\infty}^{\infty} v^k dF_x(v) , \quad k = 1, 2, 3, \ldots .$$

Kac's idea was that evaluating the limits of these moments would lead to the determination of the weak limit of $F_x(v)$, which in this case $G(v)$.

Turan felt he could do this on the basis of the method underlying (2.1), One has to expand the quantity in (4.1) using the Binomial Theorem and estimate the various terms using familiar results on primes. For some reason Turán did not carry this out; perhaps he suspected the great complications that these computations would involve.

The first to successfully compute the moments in (4.1) was Delange [2], but it was only after the Erdös-Kac Theorem was proved. Delange interpreted the terms arising out of the expansion of (4.1) in terms of the coefficients of a suitable generating function, and employed analytic tools to evaluate these coefficients.

But in 1955 Halberstam attacked this problem in a purely elementary manner. He in fact succeeded in evaluating the moments for all functions satisfying (2.3). More precisely for these functions he showed [11; I] that

$$\lim_{x\to\infty} \frac{1}{xB_2(x)^{k/2}} \sum_{n\le x} (f(n)-A(x))^k = \int_{-\infty}^{\infty} v^k dG(v) = m_k \qquad k=1,2,3,\ldots \quad (4.2)$$

Consequently $F_x(v) \to G(v)$ as $x \to \infty$.

In order to prove (4.2) Halberstam also needed to establish first a similar result for $f_y(n)$ and then had to make the transition from $f_y(n)$ to $f(n)$ by showing that

$$\sum_{n\le x} |f_y(n)-f(n)|^k = o(xB(x)^{k/2}), \quad k=1,2,\ldots \qquad (4.3)$$

for a suitable choice of y and $\alpha \to \infty$ with x. This was no problem for functions satisfying (2.3). Since his method was elementary he [11;II,III] could successfully apply it to sets such as

$$S^{(1)} = \{p+a \mid p = \text{prime}\}, \quad \text{where} \quad a \in \mathbb{Z}^+, \qquad (4.4)$$

or

$$S^{(2)} = \{Q(n) \mid n \in \mathbb{Z}^+\}, \quad \text{where} \quad Q(x) \in \mathbb{Z}^+[x],$$

by which time it was realised that (2.3) could be replaced by the weaker condition (3.13).

After Halberstam's proof of (4.2), Delange [3] noticed that it could be established also by his analytic method involving generating functions. In fact, after Kubilius announced his general distribution results, Delange [4] observed that similar results could be obtained by his method also provided one imposed further growth conditions upon f; namely

$$\{\max_{p\le x} f(p)\} \ll \sqrt{B_2(x)}, \qquad (4.5)$$

in addition to the growth condition (2.6) of the Kubilius class $\mathcal{H}$. Thus Delange was the first to realise that limiting distributions other than the Gaussian could be obtained by the moment method. The use of generating functions and their dependence on the Euler products, restricted the kind of sets that Delange could treat. For instance while considering arithmetic progressions, he needed various analytic and multiplicative properties of Dirichlet L-series.

The elementary approach of Halberstam makes use of (4.3) which is not obvious for general subsets S, or when $f(p)$ is large. In fact Halberstam used (4.3) only when conditions such as (3.13) held, and so the limiting distribution was always Gaussian. It is thus not known for instance, whether for sets $S^{(1)}$ and $S^{(2)}$, limiting distributions other than the Gaussian arise as a consequence of the moment method. Barban showed by means of a trick (see [5], Vol. 1, p. 173, Vol. 2, p. 29) which avoided (3.12), that analogues of Kubilius' distribution theorems hold for $S^{(1)}$. But for $S^{(2)}$ the only limiting distribution known with $B_2(x) \to \infty$, is the Gaussian distribution.

We shall show by means of our method that for $S^{(1)}$ it is possible to obtain limiting distributions other than the Gaussian with $B_2(x) \to \infty$, by evaluating the moments asymptotically. (In this regard $S^{(2)}$ still remains a problem when $\deg Q \geq 2$.) We make use of the bilateral Laplace transform to retain the elegance of Delange's generating functions. But then, since our method is based upon the combinatorial sieve, an elementary tool, it retains the wide applicability to sets $S \subseteq \mathbb{Z}^+$ as in Halberstam's approach. Without much further ado, we proceed to describe this technique.

§5. Multiplicative functions and bilateral Laplace transforms

In order to estimate the moments of a strongly additive function $f(n)$ for $n \in S$ we begin by considering its moment generating function, namely, its bilateral Laplace transform. So let $F_{x,y}(v)$ be as in (3.9). Now define

$$T_u(x,y,f,S) = T_u(x,y) = \int_{-\infty}^{\infty} e^{uv} dF_{x,y}(v)$$

$$= \frac{e^{-uA(y)/\sqrt{B_2(y)}}}{X} \sum_{n\leq x,\, n\in S} a_n e^{uf_y(n)/\sqrt{B_2(y)}}, \qquad (5.1)$$

where u is real. We set $T_u(x,x) = T_u(x)$.

Ideally we would like to extract the moments of f from $T_u(x)$ in the following fashion:

$$\int_{-\infty}^{\infty} v^k dF_x(v) = \frac{d^k}{du^k} T_u(x)\Big|_{u=0}.$$

The following lemma shows how this could be done rigorously.

Lemma 1: Let F_x be a sequence of probability distributions. Suppose there is $R > 0$ such that

$$\int_{-\infty}^{\infty} e^{uv} dF_x(v) \ll 1 \quad \text{for } -R \le u \le R . \tag{5.2}$$

Then

$$\int_{-\infty}^{\infty} v^k dF_x(v) \ll \frac{k!}{R^k} \quad \text{for } k = 1,2,3,\dots . \tag{5.3}$$

If in addition to (5.2) we have

$$\lim_{x \to \infty} \int_{-\infty}^{\infty} e^{uv} dF_x(v) = \ell(u) , \quad \text{uniformly for } -R \le u \le 0 \tag{5.4}$$

then

$$\lim_{x \to \infty} \int_{-\infty}^{\infty} v^k dF_x(v) = \mu_k \quad \text{exists for } k = 1,2,3,\dots .$$

Also there is a probability distribution $F(v)$ such that

$$\mu_k = \int_{-\infty}^{\infty} v^k dF(v) , \quad k = 1,2,3,\dots$$

and

$$\lim_{x \to \infty} F_x(v) = F(v) \quad \text{(weakly in } v) .$$

If $\ell(u)$ can be extended analytically in $|z| \le R$ then

$$\mu_k = \frac{d^k}{dz^k} \ell(z) \Big|_{z=0} . \quad //$$

For a proof of this lemma see [1].

We want to apply this lemma to the sums $T_u(x)$ and $T_u(x,y)$. We study closely the situation $f \ge 0$. The strongly multiplicative function

$$g(n) = e^{uf(n)/\sqrt{B(x)}} \quad \text{is always } > 0 . \tag{5.5}$$

Two cases now arise:

Case 1: $u \le 0 \Rightarrow 0 < g(n) \le 1$.

Here we shall use the combinatorial sieve to obtain upper bounds for $T_u(x)$ for all such g, and also estimate this sum asymptotically for certain g. This will be interpreted in terms of (5.2) and (5.4) in Lemma 1. More on this in §6.

Case 2: $u > 0 \Rightarrow g(n) \ge 1$.

Here we shall obtain upper bounds for $T_u(x)$ by various methods. These methods will depend on the kind of information imposed upon $R_d(x)$, and are described in §7. As in Case 1, these bounds will be related to (5.2) in Lemma 1.

We now proceed to describe how we treat Case 1.

§6. A new application of the combinatorial sieve

Let $g(n)$ be a strongly multiplicative function satisfying $0 \leq g(n) \leq 1$. We shall view the quantity

$$\sum_{\substack{n\leq x \\ n\in S}} a_n g(n) \tag{6.1}$$

as the residual amount after a sieve process.

We begin with the sum $S_1(x)$ containing weights a_n for each $n \in S$. For the prime 2, we remove from the weights a_n corresponding to even $n \in S$, the amount $a_n(1-g(2))$, so that these weights shrink to

$$a_n^{(1)} = \begin{cases} a_n g(2) & \text{if } n \in S \text{ is even} \\ a_n & \text{if } n \in S \text{ is odd}. \end{cases}$$

Let $p_1, p_2, \ldots, p_r, \ldots$ be the sequence of primes in increasing order. Our sieve will be described inductively as follows. Assume that the sieve has been employed on $S_1(x)$, up to the prime p_{k-1}, and that the original weights a_n have shrunk to $a_n^{(k-1)}$. Then corresponding to the prime p_k we shrink the weights $a_n^{(k-1)}$ in the following fashion:

$$a_n^{(k)} = \begin{cases} a_n^{(k-1)} g(p_k) & \text{if } n \in S \text{ is } \equiv 0 \pmod{p_k} \\ a_n^{(k-1)} & \text{if } n \in S \text{ is } \not\equiv 0 \pmod{p_k}. \end{cases} \tag{6.2}$$

Thus for $n \in S$ which are multiples of p_k, the amount $a_n^{(k-1)}(1-g(p_k))$ is removed from the previous weights $a_n^{(k-1)}$. So if this sieve process is carried out until the last prime $p_\ell \leq x$, we have

$$\sum_{\substack{n\leq x \\ n\in S}} a_n g(n) = \sum_{\substack{n\leq x \\ n\in S}} a_n^{(\ell)}. \tag{6.3}$$

It is easy to see that this is equivalent to saying

$$\sum_{\substack{n\leq x \\ n\in S}} a_n g(n) = \sum_{\substack{n\leq x \\ n\in S}} a_n \sum_{d|n} \mu(d) g^*(d), \tag{6.4}$$

where g^* is the strongly multiplicative function defined by

$$g^*(p) = 1 - g(p) \quad \text{for each } p. \tag{6.5}$$

The expression in (6.4) describes the 'inclusion-exclusion procedure' in our sieve.