THE METHOD OF TRIGONOMETRICAL SUMS IN THE THEORY OF NUMBERS

THE METHOD OF TRIGONOMETRICAL SUMS IN THE THEORY OF NUMBERS

By I. M. VINOGRADOV

Translated from the Russian, revised and annotated by

K. F. ROTH, B.A., Ph.D., and ANNE DAVENPORT, M.A.

University College London

INTERSCIENCE PUBLISHERS LTD., LONDON
INTERSCIENCE PUBLISHERS, INC., NEW YORK

Preface by the Translators

Since 1934 the analytic theory of numbers has been largely transformed by the work of Vinogradov. This work, which has led to remarkable new results, is characterized by its supreme ingenuity and great power.

Vinogradov has expounded his method and its applications in a series of papers and in two monographs [1]), which appeared in 1937 and 1947. The present book is a translation of the second of these monographs, which incorporated the improvements effected by the author during the intervening ten years.

The text has been carefully revised and to some extent rewritten. The more difficult arguments have been set out in greater detail. Notes have been added, in which we mention the more important changes made and comment on the subject-matter; we hope these will be of assistance or of interest to the reader. In particular, in the Notes on Chapter VI, we mention a simplification effected by Hua in 1948.

We are greatly indebted to Professor Davenport, who has given substantial help throughout. We are also grateful to Professor Mordell and to Dr. G. L. Watson for a number of helpful comments.

[1]) *A new method in the analytic theory of numbers* [Travaux de l'Institut mathématique Stekloff, volume X (1937), Moscow and Leningrad]. *The method of trigonometrical sums in the theory of numbers* [Travaux de l'Institut mathématique Stekloff, volume XXIII (1947)].

CONTENTS

* These, and the notes to the individual chapters, have been added to the translation.

NOTATION

Throughout the book, n denotes a positive integer greater than 1, and

$$\nu = \frac{1}{n}.$$

θ always denotes some number satisfying $-1 \leqq \theta \leqq 1$.

c (and similarly $c_1, \ldots$) denotes a positive constant.

ε (and similarly $\varepsilon_1, \ldots$) denotes an arbitrarily small positive number.

If F and G denote functions of certain variables, and $G \geqq 0$, the notations

$$F = O(G) \quad \text{and} \quad F \ll G$$

both mean that there exists a positive constant c such that $|F| \leqq cG$. The constant c may well depend on certain other parameters (e.g. on n), but the meaning will be plain from the context.

The notation $G \gg F$ means the same as $F \ll G$, but will only be used when F and G are both non-negative.

For any real number x we denote by $[x]$ the integral part of x and by [1] $\{x\}$ the fractional part of x. We denote by $\|x\|$ the distance of x from the nearest integer, so that

$$\|x\| = \min(\{x\}, 1 - \{x\}) = \min |x - m|,$$

the last minimum being taken over all integers m.

If A and B are real numbers satisfying $0 \leqq B - A \leqq 1$, the notation

$$A < z < B \pmod 1$$

means that

$$h + A < z < h + B$$

[1] Occasionally, where there is no risk of confusion, { } are used as ordinary brackets.

for some integer h. Similarly for inequalities in which one or both of the signs $<$ are replaced by $\leq$.

If m is any positive integer, $\tau(m)$ denotes the number of divisors of m, $\Omega(m)$ denotes the number of prime factors of m, and $\varphi(m)$ is Euler's function: the number of positive integers $m' \leq m$ with $(m', m) = 1$.

We use the abbreviations

$$e(\alpha) = e^{2\pi i\alpha}$$

for any real number α, and

$$e_q(a) = e\left(\frac{a}{q}\right)$$

for any integers a and q with $q > 0$.

INTRODUCTION

One of the most important problems in the theory of numbers is that of establishing regularities of various kinds in the distribution of the values of a function $f(x_1, \ldots, x_r)$ of one or more variables. We shall consider only those values of the function which correspond to the integer points $(x_1, \ldots, x_r)$ of r dimensional space belonging to some given set Ω. This set may consist either of all the integer points of the space, or of those integer points which satisfy certain conditions; for example those points whose coordinates satisfy certain inequalities, or those points whose coordinates are primes, and so on.

The problem just formulated in such general terms can assume very different special forms, according to the kind of restrictions imposed both on the function $f(x_1, \ldots, x_r)$ and on the set Ω. We single out three problems which are of great importance in the theory of numbers. These problems resemble one another in their formulation, and moreover the method which we shall use for their solution is substantially the same. I discovered this method in 1934 and gave the first systematic exposition of it in 1937. The method was later considerably revised and simplified, and the results refined. The present book contains a new and improved exposition of my method and of its application to the three problems.

We proceed now to a more detailed description of the three problems in question. We shall give a short account of their origin, and shall mention the methods which existed for their treatment before the discovery of my method in 1934.

1. A very important problem is that of the distribution of the values of the exponential function

$$f(x_1, \ldots, x_r) = e(F(x_1, \ldots, x_r)),$$

where $F(x_1, \ldots, x_r)$ is a real function. The essence of this problem is to establish an upper bound for the absolute value of the sum

$$S = \sum_{\Omega} f(x_1, \ldots, x_r) = \sum_{\Omega} e(F(x_1, \ldots, x_r))$$

extended over the integer points $(x_1, \ldots, x_r)$ in Ω, it being understood that the number T of such points is finite. As the absolute value of each term in the sum is 1, and the number of terms is T, we have for $|S|$ the trivial estimate

$$|S| \leqq T.$$

The sign of equality holds if and only if all the values of the function $F(x_1, \ldots, x_r)$ are integers, or differ by integral amounts. However, for very wide classes of functions $F(x_1, \ldots, x_r)$ and sets Ω, it proves to be possible to establish for $|S|$ an upper bound very much more precise than the trivial one just indicated. This is of the form

$$|S| \leqq T\gamma,$$

where γ tends to zero as the number T of points in the set Ω increases to infinity, even though the function $F(x_1, \ldots, x_r)$ may simultaneously undergo a change of form. The factor γ, distinguishing such a bound from the trivial one, may be called the "factor of reduction".

Let us consider in detail sums of the form

$$\sum_{x} e(F(x)),$$

where the summation is extended over all integers x in some interval $Q \leqq x \leqq Q + P$, or over a subset of these integers. Such sums are special cases of the general sum S with $r = 1$.

Much attention has been given to sums of the form

(1) $$S = \sum_{x=0}^{q-1} e_q(\Phi(x)), \quad \Phi(x) = a_n x^n + \ldots + a_1 x,$$

where $q > 0$ and $(a_n, \ldots, a_1, q) = 1$. The simplest non-trivial sum of this form, namely

$$\sum_{x=0}^{q-1} e_q(ax^2),$$

was evaluated by Gauss [1] and is called a Gaussian sum. The more general sum (1) was investigated by Mordell [2], who obtained the estimate

$$S \ll q^{1-\nu} \qquad \left(\nu = \frac{1}{n}\right)$$

when q is a prime. For the case when q is not restricted to primes, L. K. Hua [3] proved that

$$S \ll q^{1-\nu+\varepsilon}.$$

This last inequality cannot be substantially improved; there are infinitely many values of q for which all sums of the form

$$(2) \qquad \sum_{x=0}^{q-1} e_q(ax^n), \quad (a, q) = 1,$$

are equal to $q^{1-\nu}$ (see Lemma 4 of Chapter II).

It is also possible to estimate sums of the form

$$\sum_{\substack{x=0 \\ (x,q)=1}}^{q-1} e_q(f(x)) \text{ and } \sum_{\substack{x=0 \\ (x,q)=1}}^{q-1} \chi(f(x)),$$

where

$$f(x) = a_n x^n + \ldots + a_1 x + a_{-1} x' + \ldots + a_{-m} x'^m.$$

Here the a's are integers, χ is a non-principal character to the modulus q, and x' is defined by $xx' \equiv 1 \pmod{q}$. We shall not be concerned with such sums in this book, and refer the reader to the literature [4, 5, 6].

It is considerably more difficult to estimate sums of the general form

$$(3) \qquad S = \sum_{x=Q}^{Q+P-1} e(F(x)), \qquad F(x) = \alpha_n x^n + \ldots + \alpha_1 x,$$

where Q and P are integers ($P > 0$) and $\alpha_n, \ldots, \alpha_1$ are real. The first general method for estimating such sums was given by H. Weyl, and consequently the sums are called "Weyl sums". The estimate found by Weyl's method depends on an approximation to the highest coefficient α_n of the polynomial $F(x)$ by a rational fraction. Let

$$\alpha_n = \frac{a}{q} + \frac{\theta t}{q^2}, \text{ where } (a, q) = 1,\ q > 0,\ t \geqq 1.$$

Then Weyl's method leads to the estimate [7] $|S| \leqq P\gamma$, with

(4) $\gamma \ll P^{\varepsilon}(P^{-1} + tq^{-1} + tP^{-n+1} + qP^{-n})^{\varrho}$, where $\varrho = \dfrac{1}{2^{n-1}}$

and ε can be taken to be arbitrarily small.

In order to see more clearly the degree of precision of this estimate, consider the special case when $q \leqq P$ and $t = 1$. Then (4) leads to the estimate $|S| \leqq P\gamma'$, where

(5) $$\gamma' \ll q^{-\varrho'},\ \varrho' = \varrho - \varepsilon.$$

Now let $P \to \infty$ and at the same time $q \to \infty$ in accordance with some law. Then the factor of reduction γ' tends to zero, and the rapidity with which it does so depends on the size of ϱ'. If n is large then ϱ', being about $\dfrac{1}{2^{n-1}}$, is small, and the estimate (5) is then comparatively weak.

In Chapter VI of this book, a new estimate for Weyl sums is obtained by means of my method, and this estimate allows one to replace the exponent ϱ' in (5) by the number

$$\varrho_1 = \frac{1}{3(n-1)^2 \log 12n(n-1)}.$$

This tends to zero as $n \to \infty$ very much more slowly than $\dfrac{1}{2^{n-1}}$, and consequently for large n the new estimate (5) is much more precise than the former one.

Successful variants of my method for estimating Weyl sums were given by van der Corput (in letters to me in June 1936) and by U. V. Linnik [8] (in 1942). But in the present book I confine myself to a variant similar to one which I have expounded in certain papers.

The sums (2) are special cases of Weyl sums with $Q = 0$, $P = q$ and $F(x) = ax^n/q$. As we have already seen, there are infinitely many such sums for which $S = q^{1-\nu}$. This shows that in the estimate (5) the number ϱ' cannot be replaced by any number greater than $\nu = 1/n$.

It is a plausible conjecture that the estimate (5) holds with ϱ' replaced by $\nu - \varepsilon$, and more generally one might conjecture that the exponent ϱ in (4) could be replaced by $\nu - \varepsilon$. A proof or disproof of this conjecture would be very desirable.

But even if such a conjecture were proved, we should still be far from having a completely satisfactory solution of the problem of estimating Weyl sums. The following simple considerations illustrate this. Let s be any one of the numbers $1, \ldots, n$, and let Q and P and all the coefficients of $F(x)$ other than α_s be given. Let α_s vary between 0 and 1, so that the sum (3) becomes a function of α_s, say $S(\alpha_s)$. Then

$$(6) \qquad \int_0^1 |S(\alpha_s)|^2 \, d\alpha_s$$

$$= \sum_{x_1=Q}^{Q+P-1} \sum_{x=Q}^{Q+P-1} e\big(F(x_1) - \alpha_s x_1^s - F(x) + \alpha_s x^s\big) \int_0^1 e\big(\alpha_s(x_1^s - x^s)\big) d\alpha_s = P,$$

since the integral

$$\int_0^1 e\big(\alpha_s(x_1^s - x^s)\big) d\alpha_s$$

is 1 if $x_1 = x$ and 0 otherwise (see Lemma 4 of Chapter I). The equation (6) shows that, if $0 < \lambda < \frac{1}{2}$, the estimate

$$(7) \qquad |S(\alpha_s)| \leqq P^{1-\lambda}$$

is true for all α_s between 0 and 1 except possibly those lying in a finite number of intervals of total length $\leqq P^{2\lambda-1}$, which tends to 0 as $P \to \infty$. Roughly speaking, (7) holds for almost all sums $S(\alpha_s)$.

Using the results of Chapter VI, it is possible to deduce other important conclusions concerning the distribution of the absolute value of the sum (3). For example, the following can be proved. Suppose $n \geqq 11$ and let Q and P be fixed. Denote the sum (3) by $S(\alpha_n, \ldots, \alpha_1)$. Then the estimate

$$S(\alpha_n, \ldots, \alpha_1) \ll P^{0.975}$$

is true for all points $(\alpha_n, \ldots, \alpha_1)$ in the n dimensional cube

$0 \leqq \alpha_n \leqq 1, \ldots, 0 \leqq \alpha_1 \leqq 1$ except possibly for points lying in a finite number of regions of total volume less than

$$P^{-0.125n^2}.$$

The more complete elucidation of the regions containing points $(\alpha_n, \ldots, \alpha_1)$ for which $|S(\alpha_n, \ldots, \alpha_1)|$ is abnormally large represents a most important task in the problem of estimating Weyl sums.

The methods used to find estimates for Weyl sums can also be applied to sums of the form

$$S = \sum_{x=Q}^{Q+P-1} e(F(x)), \tag{8}$$

where Q and P are integers ($P > 0$) and where the nth derivative of the function $F(x)$ satisfies an inequality of the form

$$\frac{1}{A} \leqq \frac{F^{(n)}(x)}{n!} \leqq \frac{c}{A} \qquad (A \geqq 2)$$

in the interval $Q \leqq x \leqq Q + P$. One sum of this kind has a very important application to the question of the distribution of the primes [9]. In the special case when $n = 2$, sums of the above form are of great importance for the problem of the number of integer points in a given region in the plane or in space, for example the region $x^2 + y^2 \leqq r^2$ or the region $x^2 + y^2 + z^2 \leqq r^2$; and a method for estimating such sums was found independently by van der Corput [10] and myself [11]. Van der Corput further showed that by imposing some additional restrictions it is possible to combine his method with Weyl's method so as to improve the result. He also obtained estimates for the general sums (8) by using Weyl's method [12].

In Chapter VI we apply my method to the sums (8) when $n > 11$ and $P \ll A \ll P^{2+2\nu}$, and obtain, subject to these conditions, the new estimate

$$S \ll P^{1-\varrho}, \qquad \varrho = \frac{1}{3n^2 \log 125n}. \tag{9}$$

This represents, for large n, an improvement over earlier results similar to that already mentioned in the case of Weyl sums.

Naturally the question arises whether a further improvement on (9) is possible, that is, whether ϱ can be replaced by a larger number. In view of the very general character of the sums S in (8), it is difficult to make any specific conjectures. Perhaps it will be best to confine ourselves to an important special case. Consider the sum

$$S(t) = \sum_{x=P_0+1}^{P_0+P} e^{it\log x},$$

where P_0 and P are integers satisfying $\frac{1}{2}P_0 \leqq P \leqq P_0$, and t is any number satisfying $P_0{}^{n-2} \leqq t \leqq P_0{}^{n-1}$. (This sum occurs as the example to Theorem 2b of Chapter VI.) Here we have

$$2\pi F(x) = t\log x,\ 2\pi F^{(n)}(x) = \frac{(-1)^{n-1}n!t}{nx^n},$$

and consequently

$$\frac{t}{2\pi n(3P)^n} \leqq (-1)^{n-1}\frac{F^{(n)}(x)}{n!} \leqq \frac{t}{2\pi nP^n}.$$

Putting

$$A = \frac{2\pi n(3P)^n}{t},\ l = 3^n,$$

we have $P \ll A \ll P^2$ and

$$\frac{1}{A} \leqq (-1)^{n-1}\frac{F^{(n)}(x)}{n!} \leqq \frac{l}{A} \text{ for } P_0 + 1 \leqq x \leqq P_0 + P.$$

Thus the conditions which we imposed on the sum (8) are satisfied. Further, we have

$$\int_{P_0{}^{n-2}}^{P_0{}^{n-1}} |S(t)|^2 dt = \sum_{x_1=P_0+1}^{P_0+P}\ \sum_{x=P_0+1}^{P_0+P} \int_{P_0{}^{n-2}}^{P_0{}^{n-1}} e^{it(\log x_1 - \log x)} dt.$$

But the integral

$$\int_{P_0{}^{n-2}}^{P_0{}^{n-1}} e^{it(\log x_1 - \log x)}\, dt$$

has the value $P_0{}^{n-1} - P_0{}^{n-2}$ if $x_1 = x$, and is

$$\ll |\log x_1 - \log x|^{-1}$$

if $x_1 \neq x$. Putting $x_1 = x + u$ in the latter case, we have

$$|\log x_1 - \log x| = \left|\log\left(1 + \frac{u}{x}\right)\right| \gg \frac{|u|}{P}, \text{ or } |\log x_1 - \log x|^{-1} \ll \frac{P}{|u|}.$$

It is now easy to deduce that

$$\int_{P_0^{n-2}}^{P_0^{n-1}} |S(t)|^2 dt - P(P_0^{n-1} - P_0^{n-2}) \ll P \sum_{u=1}^{P} \frac{P}{u} \ll P^2 \log P,$$

whence

$$\int_{P_0^{n-2}}^{P_0^{n-1}} |S(t)|^2 dt = PP_0^{n-1} + O(P_0^{n-1} \log P). \tag{10}$$

The equation (10) shows that the number ϱ in the estimate (9) cannot be replaced by a number $\varrho' > \frac{1}{2}$. Moreover the same equation shows that if $0 < \lambda < \frac{1}{2}$ the estimate

$$|S(t)| \ll P^{1-\lambda} \tag{11}$$

holds for all values of t in the interval $P_0^{n-2} \leqq t \leqq P_0^{n-1}$ except possibly for those lying in a finite number of intervals whose total length is at most $P^{n+2\lambda-2}$, and so is negligible in comparison with $P_0^{n-1} - P_0^{n-2}$ for large P. Thus, in a sense, the estimate (11) holds for almost all sums $S(t)$.

Finally, we consider sums of the form

$$\Sigma\, e(F(x)), \tag{12}$$

where $F(x)$ is a real function and the summation is extended only over a subset of the integers in the interval $Q \leqq x < Q + P$. All the sums with which we shall be concerned will have $\gg P^{1-\varepsilon}$ terms. It must not be thought that an estimate for such a sum is inevitably worse than, or no better than, that which would be obtained if summation were extended over all integers of the interval. For example, if in the sum (2) we restrict x to those integers in the interval $0 \leqq x < q$ which are relatively prime to q, the sum is always $\ll q^{\frac{1}{2}}$, whereas the sum (2) itself can be equal to $q^{1-\nu}$ for infinitely many values of q, as we have already mentioned.

Sums of the form

$$\sum_{p \leqq N} e(F(p)), \tag{13}$$

where p runs through primes, constitute a particularly interesting special case of the sums (12). In Chapter IX we show how the estimation of such sums, in the simplest case when $F(p) = \alpha p$, can be reduced to a straightforward application of my method. This is effected by means of the following identity (or certain generalizations of it):

$$\Phi(1) + \sum_{\sqrt{N} < p \leqq N} \Phi(p) = \sum \sum_{dm \leqq N} \mu(d)\Phi(md). \tag{14}$$

Here d runs through products of primes (including the "empty" product 1) not exceeding $\sqrt{N}$, and m runs through positive integers. The identity (14) has been known for a long time, and can be very easily deduced from the Sieve of Eratosthenes. Closely related in some respects to the identity (14) is the famous identity of Euler:

$$\sum_{m=1}^{\infty} \frac{1}{m^s} = \prod_{p} \left(1 - \frac{1}{p^s}\right)^{-1} = \left(\sum_{d=1}^{\infty} \frac{\mu(d)}{d^s}\right)^{-1}, \tag{15}$$

where $s = \sigma + it$ and $\sigma > 1$. This identity and its generalization for the L-functions were later fundamental for the theory of the distribution of the primes, created by the work [13] of Dirichlet, Riemann, Hadamard, de la Vallée Poussin, Hardy and Littlewood and others.

In fact the identity (15) can be deduced from (14), by taking $\Phi(m)$ to be m^{-s} and making $N \to \infty$. We should mention here that it was by modifying the idea of the Sieve of Eratosthenes that Viggo Brun was led to his well known method, which has made it possible to solve a series of very subtle problems in the theory of the distribution of the primes.

2. Closely related to the problem discussed in § 1 is the problem of the distribution of the fractional part

$$f(x_1, \ldots, x_r) = \{F(x_1, \ldots, x_r)\}$$

of a real function $F(x_1, \ldots, x_r)$. We again suppose that the number T of points in the set Ω is finite. Each value of $f(x_1, \ldots, x_r)$ satisfies $0 \leqq f(x_1, \ldots, x_r) < 1$. It transpires that, for very wide classes of functions F and sets Ω, every number α in the interval $0 < \alpha < 1$ is approached by values of $f(x_1, \ldots, x_r)$. In fact it can be proved that the distance from α to the nearest value of $f(x_1, \ldots, x_r)$ does not exceed γ, where the number $\gamma = \gamma(T)$ does not depend on α and tends to zero as $T \to \infty$, even though there may at the same time be a variation in the form of the function F. Moreover in very general cases it is also possible to establish a considerable measure of uniformity in the distribution of the values of $f(x_1, \ldots, x_r)$. This uniformity is expressed by the fact that, for any δ with $0 < \delta < 1$, the number H of values of $f(x_1, \ldots, x_r)$ satisfying $0 \leqq f(x_1, \ldots, x_r) < \delta$ is approximately proportional to δ; more precisely,

$$H = T\delta + O(T\gamma_1),$$

where γ_1 is independent of δ and tends to zero as $T \to \infty$. In what follows we shall consider only the case when the function f is $f(x) = \{F(x)\}$, and x takes all integral values in some interval $Q \leqq x < Q + P$, or a certain subset of these values. This is a special case, with $r = 1$, of the general problem just stated.

The simplest problem to treat is that of the distribution of the values of the function

$$f(x) = \left\{\frac{\Phi(x)}{q}\right\}, \quad \Phi(x) = a_n x^n + \ldots + a_1 x,$$

where $q > 1$, $(a_n, \ldots, a_1, q) = 1$, and x runs through a complete set of residues to the modulus q. Here, using the estimate stated earlier for sums of the form (1), it can be proved that the number H of values of x for which $f(x)$ lies in the interval $0 \leqq f(x) < \delta$ is given by the asymptotic formula

$$H = q\delta + O(q^{1-\nu+\varepsilon_0}).$$

A more difficult problem is that of the distribution of the values of the function

$$f(x) = \{F(x)\}, \quad F(x) = \alpha_n x^n + \ldots + \alpha_1 x, \tag{16}$$

where $\alpha_n, \ldots, \alpha_1$ are real and x runs through the integers of the interval $Q \leqq x < Q + P$ (Q and P being integers). The first general solution of this problem was given by H. Weyl, who used for this purpose his own estimates for the sums (3) which bear his name. However, Weyl's results were relatively crude. Later the problem was extended and generalized, mainly in the work of van der Corput and Koksma, who used better estimates for the sums (3) and (8), obtained by Weyl's method. The most precise result obtained by the application of Weyl's method to the question of the distribution of the values of the function (16) can be formulated as follows. Let

$$\alpha_n = \frac{a}{q} + \frac{\theta}{q^2}, \ (a, q) = 1, \ 0 < q < P^n.$$

Then the number H of numbers in the sequence

$$f(x) = \{F(x)\}, \quad x = Q, \ldots, Q + P - 1$$

which satisfy $0 \leqq f(x) < \delta$ is given by[7]

$$H = P\delta + O(P\gamma),$$

where

$$\gamma = P^{\varepsilon}(P^{-1} + q^{-1} + qP^{-n})^{\sigma}, \quad \sigma = \frac{1}{2^{n-1}}.$$

In Chapter VIII, my method is applied to the estimation of the error in the preceding formula, the result being an improvement of a kind similar to that mentioned earlier in connection with Weyl sums.

Further, in Chapter V we obtain also a very precise estimate for the distance from any proper fraction α to the nearest number in the sequence

$$f(x) = \{F(x)\}, \ x = 1, \ldots, [q_l^{2\lambda}],$$

where $F(x)$ is as in (16), and l is one of the numbers $n, \ldots, 1$, and $\lambda = 1/l$, and q_l is determined by

$$\alpha_l = \frac{a_l}{q_l} + \frac{\theta}{q_l^2}, \ (a_l, q_l) = 1, \ q_l > 0.$$

Finally, in Chapter XI, it is shown how my method can be used to investigate the distribution of the values of the function

$$f(p) = \{\alpha p\},$$

where p runs through primes $\leqq N$.

3. Of special interest are the laws of distribution of the values of functions $f(x_1, \ldots, x_r)$ which take integral values for points $(x_1, \ldots, x_r)$ of the set Ω. Here the question arises how often a given integer N is represented by the function $f(x_1, \ldots, x_r)$ at points of the set Ω. In other words, what can be said about the number of solutions $I(N)$ of the indeterminate equation

(17) $$f(x_1, \ldots, x_r) = N?$$

In some cases we aim only at establishing the inequality $I(N) > 0$, which shows that (17) is soluble; in other cases it proves to be possible to find an asymptotic formula for $I(N)$ or even an exact formula for $I(N)$.

We now discuss in more detail the distribution of the values of the function

$$f(x_1, \ldots, x_r) = x_1^n + \ldots + x_r^n,$$

where it is supposed that the set Ω consists of all points $(x_1, \ldots, x_r)$ of r dimensional space with non-negative $x_1, \ldots, x_r$.

Here it can very easily be proved that if $r \leqq n$ there is an infinite sequence of positive integers N for which the equation (17), that is the equation

(18) $$x_1^n + \ldots + x_r^n = N,$$

is insoluble. In fact, let N_0 be a sufficiently large positive integer. If (18) is soluble for some $N \leqq N_0$, then all the numbers $x_1, \ldots, x_r$ occurring in a solution will be found among the numbers

(19) $$0, 1, \ldots, [N_0^{\nu}].$$

Therefore, making $x_1, \ldots, x_r$ in the sum $x_1^n + \ldots + x_r^n$ run independently through the numbers (19), we find among the $[N_0^{\nu} + 1]^r$ sums all numbers $N \leqq N_0$ for which the equation (18) is soluble. Here those N for which (18) is soluble with unequal values of $x_1, \ldots, x_r$ will occur at least $r!$ times (since $x_1, \ldots, x_r$ occur in $r!$ different ways). The number of values of N for which (18) is soluble, subject to the condition that $x_1, \ldots, x_r$ are not all different, is $\ll N_0^{\nu(r-1)}$. Therefore the number K of all $N \leqq N_0$ for which (18) is soluble will satisfy

$$K < \frac{1}{r!}(N_0^{\nu} + 1)^r + O(N_0^{\nu(r-1)})$$

$$\leqq \frac{1}{n!}(N_0^{\nu} + 1)^n + O(N_0^{1-\nu}) < 0.6N_0$$

for sufficiently large N_0. This means that for more than $0.4N_0$ numbers $N \leqq N_0$ the equation (18) is insoluble, and this proves our assertion.

What are the values of r for which (18) is soluble for every $N \geqq 0$, or at least for all $N \geqq c_0$, where c_0 is sufficiently large? Lagrange [15] proved that the equation

$$x_1^2 + \ldots + x_4^2 = N$$

is always soluble in non-negative integers $x_1, \ldots, x_4$. In 1770 Waring asserted that for every $n > 2$ there exists $r = r(n)$ such that for every integer $N \geqq 0$ the equation (18) is soluble in non-negative integers $x_1, \ldots, x_r$. This assertion became known as Waring's Problem. It was first proved by Hilbert in 1909. His method was of a somewhat special character, and since it led to very large values for r it is now almost forgotten.

In order to give greater clarity to the subsequent exposition we introduce the symbol $G(n)$ to denote the integer with the following property: there exists some c such that for every integer $N \geqq c$ the equation (18) is soluble for $r = G(n)$, but there does not exist any c_1 such that (18) is soluble for every integer $N \geqq c_1$ when $r = G(n) - 1$. From what has been said above it follows that $G(n)$ exists and that $G(n) > n$ for every n.

In 1919 Hardy and Littlewood developed a new method for the solution of Waring's Problem, which is incomparably more general and exact than that of Hilbert. These scholars found an upper bound for $G(n)$ of the form

$$G(n) \leqq n2^{n-2}h, \tag{20}$$

where $h \to 1$ as $n \to \infty$. Moreover, for

$$r \geqq (n-2)2^{n-1} + 5, \tag{21}$$

Hardy and Littlewood gave for the first time the asymptotic formula [16] for $I(N)$:

$$I(N) = \frac{(\Gamma(1+\nu))^r}{\Gamma(r\nu)} N^{r\nu-1}\mathfrak{S} + O(N^{r\nu-1-c}), \tag{22}$$

where $\mathfrak{S} = \mathfrak{S}(n, r, N)$ is the "singular series", the meaning of which is explained in Chapter II below. Hardy and Littlewood also proved that, if (21) holds, $\mathfrak{S} \gg 1$. The most recent refinements of the method of Hardy and Littlewood, due to L. K. Hua [17], allow one to replace the numbers on the right of (20) and (21) by $2^n + 1$.

In Chapter IV my method is applied to the investigation of $G(n)$, and gives the upper bound

$$G(n) < 3n (\log n + 11)$$

instead of (20). As $n \to \infty$ this is of the order $n \log n$, and consequently is not much larger than the lower bound $n + 1$ established above. As far as the asymptotic formula (22) is concerned, its validity will be proved only for

$$r \geqq [10n^2 \log n]$$

(Chapter VII). It seems probable that by a further development of my method (or perhaps in some other way) the order of magnitude of this lower bound for r might be brought down nearer to n.

Another interesting problem is that of the distribution of the values of the function

$$f(p_1, \ldots, p_r) = p_1^n + \ldots + p_r^n,$$

where $p_1, \ldots, p_r$ run through the primes.

As early as 1742, there arose from the correspondence of Goldbach with Euler the so-called "Goldbach's Problem", which is the conjecture that every integer greater than 1 is the sum of not more than three odd primes. According to this conjecture, any even number greater that 2 must be representable as the sum of two primes. In 1919, V. Brun, when endeavouring to use his method (mentioned above) to prove the latter conjecture, showed that every positive even number is representable as the sum of two numbers, each of which is a product of not more than 9 primes. Later the number 9 was reduced to 4, but the attempt to prove Goldbach's conjecture for even numbers in this way did not succeed. In 1930, L. G. Schnirelmann, by supplementing Brun's method with arguments of his own concerning the density of a sequence of positive integers, proved [18] that every integer greater than 1 is representable as the sum of a bounded number of primes; later the bound was brought down to 67.

In 1923 Hardy and Littlewood indicated a method for solving Goldbach's Problem for odd N which is similar in its nature to the method which these scholars created for the solution of Waring's Problem. They established, conditionally on a certain hypothesis, an asymptotic formula for the number $I(N)$ of representations of N in the form

$$N = p_1 + p_2 + p_3,$$

where p_1, p_2, p_3 are primes. From this asymptotic formula the validity of Goldbach's conjecture for all sufficiently large odd N would follow trivially. The hypothesis underlying Hardy and Littlewood's work is the validity of a theorem, as yet unproved, concerning the zeros of Dirichlet's L-functions. However, by the beginning of 1937, a method was worked out by Page [19] and Estermann [20] which allows one to deduce an asymptotic formula for that part of the integral for $I(N)$ which corresponds to the so-called basic intervals (see Chapter X). This method is applicable not only to Goldbach's Problem but to similar more general problems. A series of such problems was solved towards the beginning of 1937: it was proved that every sufficiently large integer N is representable in the form $N = p' + p'' + x^2$ (p', p'' primes, x a positive integer), and that every sufficiently large odd N is representable as $N = p_1 + p_2 + p_3p_4$ (p_1, p_2, p_3,

p_4 primes), and so on. But for the solution of Goldbach's Problem for odd N it was necessary to have non-trivial estimates for sums of the form (13), with $F(p) = \alpha p$, that is, for the sum

$$\sum_{p \leqq N} e(\alpha p),$$

for all values of α in $0 \leqq \alpha \leqq 1$ not belonging to the basic intervals.

The general method which I found in 1937 for estimating the sums (13) allowed me to solve at last Goldbach's Problem for odd numbers, and also opened up a broad road to the solution of other very diverse analogous problems, for example Waring's Problem for primes, that is, the problem of the representation of an integer N in the form

$$N = p_1^n + \ldots + p_r^n.$$

In Chapter X we restrict ourselves to the detailed solution of Goldbach's Problem for odd numbers. For a treatment of the more general question we refer the reader to the excellent monograph of L. K. Hua [3].

In conclusion, I wish to express my gratitude to K. K. Mardjanichvili who carefully read through the manuscript of this book and drew my attention to a number of oversights.

REFERENCES

1. Landau, E. *Vorlesungen über Zahlentheorie* (Hirzel, Leipzig, 1927; reprinted by Chelsea, New York), vol. I, 153—156.
2. Mordell, L. J. On a sum analogous to a Gauss's sum, *Quart. J. of Math.* (Oxford), 3 (1932), 161—167.
3. Hua, Loo-Keng. *Additive theory of prime numbers*, Trudy Mat. Inst. Steklov, 22 (1947). [Russian with English summary.]
4. Vinogradov, I. M. Sur la distribution des résidus et des non-résidus des puissances, *Journal Physico-Math. Soc. Univ. Perm.* No. 1, (1918), 94—96. See also *Trans. American Math. Soc.*, 29 (1927), 209—217, 218—226.
5. Vinogradov, I. M. An improvement of the estimation of sums with primes, *Izvestiya Akad. Nauk SSSR*, ser. mat., 7 (1943), 17—34.
6. Davenport, H. On character sums in finite fields, *Acta Math.*, 71 (1939), 99—121.
7. Vinogradov, I. M. Analytical proof of a theorem on the distribution of the fractional parts of an integral polynomial, *Izvestiya Akad. Nauk SSSR*, 21 (1927), 567—578.
8. Linnik, U. V. On Weyl's sums, *Doklady Akad. Nauk SSSR*, 34 (1942), 184—186.

9. TCHUDAKOFF, N. G. On zeros of Dirichlet's L-functions, *Recueil Math. (Mat. Sbornik)*, 1 (43) (1936), 591—601.

10. CORPUT, J. G. VAN DER. Zahlentheoretische Abschätzungen mit Anwendung auf Gitterpunktprobleme, *Math. Zeitschr.*, 17 (1923), 250—259.

11. VINOGRADOV, I. M. On the distribution of the fractional parts of functions of two variables, *Izvestiya Leningrad Polytechn. Univ.*, 33 (1927), 31—52.

12. CORPUT, J. G. VAN DER. Zahlentheoretische Abschätzungen mit Anwendung auf Gitterpunktprobleme II, *Math. Zeitschr.*, 28 (1928), 301—310.

13. TITCHMARSH, E. C. *The theory of the Riemann zeta-function* (Oxford, 1951).

14. BRUN, V. Über das Goldbachsche Gesetz und die Anzahl der Primzahlpaare, *Archiv for Math. og Naturvidenskab*, 34 (1915), No. 8.

15. LANDAU, E. *Vorlesungen über Zahlentheorie*, vol. I, 106—109.

16. *ibid.*, 243—275.

17. HUA, LOO-KENG. On Waring's Problem, *Quart. J. of Math.* (Oxford) 9 (1938), 199—202.

18. SCHNIRELMAN, L. G. On additive properties of numbers, *Izvestiya Don Polytech. Univ. Novo Cherkash*, 14 (1930), 3—28. See also Landau, E. *Über einige neuere Fortschritte der additiven Zahlentheorie* (Cambridge, 1937), Kap. 2.

19. PAGE, A. On the number of primes in an arithmetic progression, *Proc. London Math. Soc.* (2), 39 (1935), 116—141.

20. ESTERMANN, T. Proof that every large integer is the sum of two primes and a square, *Proc. London Math. Soc.* (2), 42 (1937), 501—516.

NOTE ON VINOGRADOV'S METHOD

It is not always easy to recognize the unity of idea underlying the various forms which Vinogradov's method assumes, and the following remarks (though necessarily sketchy) may help the reader.

The fundamental principle of the method is that it is possible to estimate effectively sums of the form

$$\sum_u \sum_v e(\alpha uv)$$

(and certain similar but more elaborate sums) under very varied conditions of summation on u and v. In the applications of the method, the variables u and v often arise as functions of a large number of other variables. Generally speaking, it is possible to estimate a sum of the above form provided that the values assumed by u and v are distributed with a certain measure of regularity.

Three simple estimates for sums of the above form occur as Lemmas 10a, 10b, 10c of Chapter I. But not all the applications of the method are based on these particular estimates.

The relevance of exponential sums of the above general form to particular problems in the theory of numbers is usually far from obvious. Even where such sums occur, or can be introduced, it may be exceedingly difficult to prove that the values assumed by the variables u and v have sufficient regularity of distribution to lead to a useful estimate.

There are essentially four applications of the method in the present book, namely in Chapters IV, V, VI, IX. There is a certain distinction which can be drawn between the applications in Chapters IV and V and those in Chapters VI and IX. In the former, the arithmetical problems under consideration are such as to admit of treatment by any one of a variety of constructions. The constructions used by Vinogradov are so designed as to lead to exponential sums of the general kind mentioned above, and to ensure the necessary regularity of distribution of the variables. To achieve this end it is necessary to use methods which, in relation to the problem itself, appear to be distinctly artificial.

The position is different in Chapters VI and IX, the former of which is devoted to Weyl sums and the latter to the sum $\Sigma e(\alpha p)$ extended over primes p. Here the exponential sum is prescribed, and the difficulty lies in making its estimation dependent on that of other sums of the general type mentioned above. To do this, very ingenious subdivisions and transformations are employed. The general line of argument is easily visible in Chapter IX, and was also visible in Vinogradov's earlier and less elaborate treatment of Weyl sums. In the present Chapter VI it is somewhat obscured by the other devices which have been superimposed on the original main idea in order further to improve the final result. But the technique of summation over y in Lemma 6 of Chapter VI represents a generalized form of the fundamental principle mentioned earlier.

CHAPTER I

General Lemmas

In this chapter we give some general lemmas, which will be applied in later chapters. Lemmas which are obvious or well known are given without proof. Lemmas 10, 15 and 16 are original, and it is the application of these that constitutes the distinguishing feature of the method of this book.

LEMMA 1a (*Cauchy's inequality*). *Let r be a positive integer and let $x_1, \ldots, x_r, y_1, \ldots, y_r$ be real. Then*

$$(x_1y_1 + \ldots + x_ry_r)^2 \leqq (x_1^2 + \ldots + x_r^2)(y_1^2 + \ldots + y_r^2).$$

LEMMA 1b (*Hölder's inequality and the inequality of the arithmetic and geometric means*). *Let r be a positive integer, and let $m > 1$ and $x_1, \ldots, x_r$ be non-negative real numbers. Then*

$$(x_1+\ldots+x_r)^m \leqq r^{m-1}(x_1^m+\ldots+x_r^m) \text{ and } r^r x_1\ldots x_r \leqq (x_1+\ldots+x_r)^r.$$

Proof. We may suppose that $r \geqq 2$. Let h be the arithmetic mean of the numbers $x_1, \ldots, x_r$. Among these numbers there is one $\leqq h$ and one $\geqq h$. Suppose $x_1 \leqq h \leqq x_2$. In the interval $0 \leqq z \leqq \min(h - x_1, x_2 - h)$ the function $(x_1 + z)^m + (x_2 - z)^m$ decreases, and therefore

$$x_1^m + x_2^m \geqq h^m + (x_1 + x_2 - h)^m.$$

Further, since $(x_1 - h)(x_2 - h) \leqq 0$, we have

$$x_1x_2 \leqq h(x_1 + x_2 - h).$$

Denoting $x_1 + x_2 - h, x_3, \ldots, x_r$ in any order by $y_2, \ldots, y_r$, we see that

$$x_1^m + \ldots + x_r^m \geqq h^m + y_2^m + \ldots + y_r^m \text{ and } x_1\ldots x_r \leqq hy_2\ldots y_r,$$

where the arithmetic mean of the numbers $y_2, \ldots, y_r$ is again equal to h. Repeating the argument, we find that

$$y_2^m + \ldots + y_r^m \geqq h^m + z_3^m + \ldots + z_r^m \text{ and } y_2 \ldots y_r \leqq hz_3 \ldots z_r,$$

$$\cdot \quad \cdot \quad \cdot$$

$$v_{r-1}^m + v_r^m \geqq h^m + h^m \text{ and } v_{r-1} v_r \leqq h^2.$$

It follows that

$$x_1^m + \ldots + x_r^m \geqq rh^m \text{ and } x_1 \ldots x_r \leqq h^r,$$

which proves the lemma.

LEMMA 2. *Let η and m be positive integers and let $T_1, \ldots, T_\eta$ be non-negative real numbers. Then*

$$\left(\sum_{s=1}^{\eta} 2^{-s} T_s\right)^m \leqq \sum_{s=1}^{\eta} T_s^m.$$

Proof. By Hölder's inequality (see Lemma 1b) we have

$$\left(\sum_{s=1}^{\eta} 2^{-s} T_s\right)^m = 2^{-m}\left(T_1 + \sum_{s=2}^{\eta} 2^{-s+1} T_s\right)^m \leqq T_1^m + \left(\sum_{s=2}^{\eta} 2^{-s+1} T_s\right)^m.$$

Repeating the argument, we see that

$$\left(\sum_{s=2}^{\eta} 2^{-s+1} T_s\right)^m \leqq T_2^m + \left(\sum_{s=3}^{\eta} 2^{-s+2} T_s\right)^m,$$

$$\cdot \quad \cdot \quad \cdot$$

$$\left(\sum_{s=\eta-1}^{\eta} 2^{-s+\eta-2} T_s\right)^m \leqq T_{\eta-1}^m + \left(2^{-\eta+\eta-1} T_\eta\right)^m,$$

whence the result.

LEMMA 3. *Let r be a positive integer, and suppose that $N > 0$. Let $K_r(N)$ denote the number of solutions of the inequality*

$$x_1^n + \ldots + x_r^n \leqq N$$

in positive integers $x_1, \ldots, x_r$. Then

$$K_r(N) = T_r N^{r\nu} - \theta r N^{r\nu-\nu},$$

where $\theta \geqq 0$ (and $|\theta| \leqq 1$, as always) and

$$T_r = \frac{(\Gamma(1+\nu))^r}{\Gamma(1+r\nu)}.$$

Proof. Obviously

$$K_1(N) = N^\nu - \theta', \text{ where } \theta' \geqq 0,$$

and the lemma is therefore true for $r = 1$.

We now apply the method of induction. Let us assume that for some $r \geqq 1$ the lemma is true for $K_r(N)$. We have

$$K_{r+1}(N) = \sum_{0<x\leqq N^\nu} K_r(N-x^n) = T_r \sum_{0<x\leqq N^\nu} (N-x^n)^{r\nu} - \theta'' r N^{r\nu},$$

where $\theta'' \geqq 0$. Now

$$\sum_{0<x\leqq N^\nu} (N-x^n)^{r\nu} < \int_0^{N^\nu} (N-x^n)^{r\nu} dx < \sum_{0\leqq x\leqq N^\nu} (N-x^n)^{r\nu},$$

whence

$$\sum_{0<x\leqq N^\nu} (N-x^n)^{r\nu} = \int_0^{N^\nu} (N-x^n)^{r\nu} dx - \theta''' N^{r\nu}.$$

Hence, noting that $T_r \leqq 1$, we have

$$K_{r+1}(N) = T_r \int_0^{N^\nu} (N-x^n)^{r\nu} dx - \theta(r+1)N^{r\nu}$$
$$= T' N^{(r+1)\nu} - \theta(r+1)N^{r\nu},$$

where

$$T' = \nu T_r \int_0^1 (1-z)^{r\nu} z^{\nu-1} dz = T_r \frac{\nu\Gamma(1+r\nu)\Gamma(\nu)}{\Gamma(1+r\nu+\nu)} = T_{r+1},$$

and $\theta \geqq 0$. Consequently the lemma is also true for $K_{r+1}(N)$.

LEMMA 4. *Let N be an integer and let*

$$I = \int_0^1 e(N\alpha)d\alpha \qquad (e(\xi) = e^{2\pi i\xi}).$$

Then

$$I = \begin{cases} 1 \text{ if } N = 0, \\ 0 \text{ otherwise.} \end{cases}$$

LEMMA 5. *Let m be a positive integer, let a be an integer, and let*

$$S = \sum_{z=0}^{m-1} e_m(az) \qquad \left(e_m(t) = e\left(\frac{t}{m}\right)\right).$$

Then

$$S = \begin{cases} m \text{ if } a \text{ is divisible by } m, \\ 0 \text{ otherwise.} \end{cases}$$

LEMMA 6. *Let M and N be integers with $M < N$, and let α be a non-integral real number. Then*

$$\left| \sum_{x=M}^{N} e(\alpha x) \right| \leqq \frac{1}{2 \| \alpha \|}.$$

Proof. We have

$$\left| \sum_{x=M}^{N} e(\alpha x) \right| = \left| \frac{e(\alpha(N+1)) - e(\alpha M)}{e(\alpha) - 1} \right|$$
$$\leqq \frac{2}{2 |\sin \pi\alpha|} \leqq \tfrac{1}{2} \| \alpha \|^{-1}.$$

LEMMA 7. *Suppose $\tau \geqq 1$. Then every real number α can be represented in the form*

$$\alpha = \frac{a}{q} + z, \text{ where } (a, q) = 1,\ 0 < q \leqq \tau,\ |z| < \frac{1}{q\tau}.$$

Proof. Expanding α as a continued fraction, we can take a/q to be the convergent with the greatest denominator not exceeding τ.

LEMMA 8*a*. *Let*

$$\Phi(y) = \frac{ay + \psi(y)}{q},$$

where $(a, q) = 1$. Let f and q' be integers with $0 < q' \leqq q$. Let y take the values $f, f+1, \ldots, f + q' - 1$, and suppose that for these values of y the function $\psi(y)$ takes real values, the difference between the greatest and least of which does not exceed λ ($\lambda > 0$).

(I) *Suppose $U \geqq 1$, and let*

$$\Omega = \sum_{y} \min \left(U, \frac{1}{2 \| \Phi(y) \|} \right),\ \Omega_0 = \sum_{y} \min \left(U^2, \frac{1}{4 \| \Phi(y) \|^2} \right).$$

Then

$$\Omega < (\lambda + 3)U + q \log q$$

and

$$\Omega_0 < (\lambda + 3)U^2 + 2qU.$$

(II) *Suppose $V > 0$, and let T be the number of values of y for which*

(1) $$\| \Phi(y) \| \leqq Vq^{-1}.$$

Then

$$T < \lambda + 2 + 2V.$$

Proof. Putting $y = f + z$ we have

$$\| \Phi(y) \| = \left\| \frac{az + \delta(z)}{q} \right\|, \text{ where } \delta(z) = af + \psi(f + z).$$

For a suitably chosen B, we have $B \leqq \delta(z) \leqq B + \lambda$ for $z = 0, 1, \ldots, q' - 1$. Putting $\beta = \{B\}$, and denoting by the letter ϱ the least non-negative residue of $az + [B]$ to the modulus q, we obtain

$$\| \Phi(y) \| = \left\| \frac{\varrho + \sigma(\varrho)}{q} \right\|, \text{ where } \beta \leqq \sigma(\varrho) \leqq \beta + \lambda.$$

(I) If $q \leqq \lambda + 3$ the estimates for Ω and Ω_0 are trivial, so we can suppose that $\lambda < q - 3$.

The values of ϱ consist of some or all of the numbers

$$0, 1, \ldots, q - 1.$$

Put $q_0 = [\beta + \lambda + 1]$, so that $0 < q_0 < q$, and note that

$$0 \leqq \sigma(\varrho) < q_0$$

for every value of ϱ.

For the values of y for which

$$\varrho = 0 \text{ and } q - q_0, \ldots, q - 1$$

(if they occur), we take the term U in the sum defining Ω and the term U^2 in the sum defining Ω_0. The number of the above values is $q_0 + 1 \leqq \lambda + 3$, and in this way we obtain the first term in each estimate.

For the values

$$\varrho = 1, 2, \ldots, q - q_0 - 1,$$

we have

$$\| \Phi(y) \| = \left\| \frac{\varrho + \sigma(\varrho)}{q} \right\| \geqq \frac{s}{q},$$

where $s = \varrho$ if $\varrho + \sigma(\varrho) < \frac{1}{2}q$ and $s = q - q_0 - \varrho$ if $\varrho + \sigma(\varrho) \geqq \frac{1}{2}q$. Plainly $1 \leqq s < \frac{1}{2}q$, and each value of s occurs for at most two values of ϱ.

Hence to complete the estimate for Ω we can take

$$2\sum_{s=1}^{[\frac{1}{2}(q-1)]}\frac{q}{2s} < q\sum_{s=1}^{[\frac{1}{2}(q-1)]}\log\frac{2s+1}{2s-1} \leqq q\log q.$$

To complete the estimate for Ω_0 we can take

$$2\sum_{s=1}^{\infty}\min(U^2, \tfrac{1}{4}q^2s^{-2}) < 2\int_0^{\infty}\min(U^2, \tfrac{1}{4}q^2s^{-2})ds = 2qU.$$

(II) For $q < \lambda + 2 + 2V$ the assertion is obvious; for $q \geqq \lambda + 2 + 2V$ the assertion follows from the fact that the inequality (I) can hold only for

$$\varrho = 0, \ldots, [V] \text{ and } q - [\beta + \lambda + V], \ldots, q - 1.$$

LEMMA 8*b*. *Let*

$$\alpha = \frac{a}{q} + \frac{\theta}{q^2}, \textit{ where } (a, q) = 1.$$

Suppose that $2 \leqq q \leqq W$, $1 < W_0 \leqq W$; *and put*

$$S = \sum_{0<z\leqq W_0}\min\left(\frac{W}{z}, \frac{1}{2||\alpha z||}\right).$$

Then

$$S < (W_0 + \tfrac{7}{2}q + 16Wq^{-1})\log W.$$

Proof. We dissect the sum S into parts according to the scheme

$$S = \sum_{0<z\leqq\frac{1}{2}q} + \sum_{\frac{1}{2}q<z\leqq\frac{3}{2}q} + \ldots + \sum_{(j_0-\frac{1}{2})q<z\leqq W_0}.$$

For values of z occurring in the first sum, let ϱ denote the least non-negative residue of az to the modulus q. Since

$$\alpha z = \frac{az + \theta z/q}{q},$$

and $0 < z \leqq \frac{1}{2}q$, we have

$$||\,\alpha z\,|| \geqq (s - \tfrac{1}{2})/q,$$

where

$$s = \begin{cases} \varrho & \text{if } \varrho \leqq \frac{1}{2}q, \\ q - \varrho & \text{if } \varrho > \frac{1}{2}q. \end{cases}$$

Thus the first sum does not exceed

$$q \sum_{0<s\leqq\frac{1}{2}q} \frac{1}{s-\frac{1}{2}} < 2q + q \sum_{1<s\leqq\frac{1}{2}q} \log \frac{(s-\frac{1}{2})+\frac{1}{2}}{(s-\frac{1}{2})-\frac{1}{2}}$$
$$\leqq 2q + q \log \tfrac{1}{2}q$$
$$< 3q \log q.$$

To the remaining sums we apply Lemma 8a, (I). Each sum is of the kind considered there with $\lambda = 1$ and with $U = W/(j-\frac{1}{2})q$, where $j = 1, \ldots, j_0$. Thus the sum in question is less than $q \log q + 4W/(j-\frac{1}{2})q$.

Finally, therefore,

$$S < 3q \log q + \sum_{j=1}^{j_0} \left(q \log q + \frac{4W}{(j-\frac{1}{2})q} \right)$$
$$< (3q + j_0 q) \log q + \frac{4W}{q} \left(2 + \sum_{j=2}^{j_0} \log \frac{(j-\frac{1}{2})+\frac{1}{2}}{(j-\frac{1}{2})-\frac{1}{2}} \right)$$
$$< (3q + W_0 + \tfrac{1}{2}q) \log q + 4Wq^{-1}(2 + \log(\tfrac{1}{2} + W_0 q^{-1}))$$
$$< (W_0 + \tfrac{7}{2}q) \log W + 16Wq^{-1} \log W.$$

LEMMA 8c. *Let P and m be integers, and suppose that $P > 1$, $m > 0$, $s > 1$; let k be a real number $\geqq 1$; and suppose that*

$$\alpha = \frac{am}{q} + \frac{\theta m}{q^2}, \text{ where } (a, q) = 1 \text{ and } 0 < q < P^s.$$

Let y run through at most P consecutive integers, and let H denote the number of values of y which satisfy

$$\| \alpha y \| \leqq kP^{1-s}. \tag{2}$$

Then

$$H < (3m + 2kqP^{1-s})(Pq^{-1} + 1).$$

Proof. Let $(m, q) = d$, $m = dm_1$, $q = dq_1$. Then

$$\alpha = \frac{am_1}{q_1} + \frac{\theta m_1/q}{q_1}.$$

We can apply Lemma 8a, (II) to any sequence of q_1 or fewer consecutive values of y, with $V = kP^{1-s}q_1$ and $\lambda = m_1$. The number of values of y satisfying (2) in such a set will therefore be less than $m_1 + 2 + 2kP^{1-s}q_1$. Hence the number of such values of y in a set of P consecutive integers will be less than

$$\begin{aligned}(m_1+2+2kP^{1-s}q_1)(Pq_1^{-1}+1) &\leqq (3m_1+2kP^{1-s}q_1)(Pq_1^{-1}+1)\\ &= (3m_1d+2kP^{1-s}q_1d)(Pq_1^{-1}d^{-1}+d^{-1})\\ &\leqq (3m+2kP^{1-s}q)(Pq^{-1}+1).\end{aligned}$$

LEMMA 9. *Let N and $Y > 0$ be integers, and suppose that $A \geqq 2\beta \geqq 2$. Let y run through the values $N, \ldots, N+Y-1$, and for these values of y let the function $\Phi(y)$ take real values, subject to the condition that*

$$\frac{1}{A} \leqq \Phi(y+1) - \Phi(y) \leqq \frac{\beta}{A} \text{ for } N \leqq y \leqq N+Y-2.$$

(I) *Suppose that $U \geqq 1$, and put*

$$S = \sum_{y=N}^{N+Y-1} \min\left(U^2, \frac{1}{4\,\|\Phi(y)\|^2}\right).$$

Then

$$S < [Y\beta A^{-1}+1](2U^2+2AU).$$

(II) *Suppose that $W \geqq 1$ and let H be the number of values of y satisfying*

$$\|\Phi(y)\| \leqq WA^{-1}.$$

Then

$$H < [Y\beta A^{-1}+1](2W+1).$$

Proof. For a given real α and a given integer h, there cannot exist more than one value of y satisfying the inequalities

$$\alpha + h < \Phi(y) \leqq \alpha + A^{-1} + h. \tag{3}$$

Therefore the number T of values of y satisfying

$$\alpha < \Phi(y) \leqq \alpha + A^{-1} \pmod 1 \tag{4}$$

is equal to the number of values of h for which the inequalities (3) are soluble. Consequently $T \leqq h_2 - h_1 + 1$, where h_2 is the greatest and h_1 the least of these numbers h. But obviously

$$\Phi(N) \leqq \alpha + A^{-1} + h_1, \quad \alpha + h_2 < \Phi(N+Y-1),$$

whence we deduce that

$$h_2 - h_1 - A^{-1} < \Phi(N+Y-1) - \Phi(N) \leqq \beta A^{-1}(Y-1),$$
$$T < \beta A^{-1}(Y-1) + A^{-1} + 1 \leqq Y\beta A^{-1} + 1.$$

(I) For $0 < \{\Phi(y)\} \leqq \frac{1}{2}$ there is an s in the sequence $s = 0, \ldots, [\frac{1}{2}A]$ such that y satisfies the condition (4) with $\alpha = sA^{-1}$; on the other hand for $\{\Phi(y)\} > \frac{1}{2}$ and for $\{\Phi(y)\} = 0$ there is an s in the same sequence such that y satisfies the condition (4) with $\alpha + A^{-1} = 1 - sA^{-1}$. In both cases we have $\| \Phi(y) \| \geqq sA^{-1}$. Therefore

$$S < [Y\beta A^{-1} + 1]\left(2U^2 + \sum_{s=1}^{\infty} \min\left(2U^2, \frac{A^2}{2s^2}\right)\right)$$
$$\leqq [Y\beta A^{-1} + 1](2U^2 + 2AU).$$

(II) The assertion follows from the fact that an interval of length $2WA^{-1}$ can be covered by $[2W + 1]$ intervals of length $< A^{-1}$.

LEMMA 10*a*. *Let* $(a, q) = 1$, $q \geqq 1$. *Let*

$$S = \sum_{x=0}^{q-1} \sum_{y=0}^{q-1} \xi(x)\eta(y)\, e_q(axy),$$

and suppose that

$$\sum_{x=0}^{q-1} | \xi(x) |^2 = X_0, \qquad \sum_{y=0}^{q-1} | \eta(y) |^2 = Y_0.$$

Then

$$| S | \leqq (X_0 Y_0 q)^{\frac{1}{2}}.$$

Proof. Using Cauchy's inequality (Lemma 1a), we have

$$| S |^2 \leqq X_0 \sum_{x=0}^{q-1} \left| \sum_{y=0}^{q-1} \eta(y) e_q(axy) \right|^2$$
$$= X_0 \sum_{x=0}^{q-1} \sum_{y_1=0}^{q-1} \sum_{y=0}^{q-1} \eta(y_1)\overline{\eta(y)} e_q(ax(y_1 - y)).$$

For given y_1 and y, summation over x gives $q\,|\eta(y)|^2$ if $y_1 = y$ and zero if $y_1 \neq y$ (Lemma 5). Hence $| S |^2 \leqq X_0 Y_0 q$.

LEMMA 10*b*. *Let M, X, N, Y be integers, with $X > 0$, $Y > 0$. Let*

$$\Phi(y) = \frac{ay + \psi(y)}{q}, \text{ where } (a, q) = 1,\ q > 0.$$

Let y run through the values $y = N, \ldots, N + Y - 1$; and suppose that when y runs through any q consecutive values in this set the *difference between the greatest and least values of the function* $\psi(y)$ *does not exceed* λ $(\lambda > 0)$. *Let*

$$S = \sum_{x=M}^{M+X-1} \sum_{y=N}^{N+Y-1} \xi(x)\eta(y)e(x\Phi(y)),$$

and put

$$\sum_{x=M}^{M+X-1} |\xi(x)|^2 = X_0, \quad \sum_{y=N}^{N+Y-1} |\eta(y)| = Y_1, \quad \max |\eta(y)| = \eta.$$

Then

$$|S| \leqq \left(X_0 Y_1 \eta((2\lambda + 6)X + 3q)[Yq^{-1} + 1]\right)^{\frac{1}{2}}.$$

Proof. By Cauchy's inequality, we have

$$|S|^2 \leqq X_0 \sum_{x=M}^{M+X-1} \left| \sum_{y=N}^{N+Y-1} \eta(y)e(x\Phi(y)) \right|^2.$$

We compare the sum on the right with

$$S' = \sum_{x_1=M}^{M+X-1} \sum_{x_2=-X+1}^{X-1} \left| \sum_{y=N}^{N+Y-1} \eta(y)e((x_1 + x_2)\Phi(y)) \right|^2.$$

In the latter sum, there will be at least X pairs x_1, x_2 for which $x_1 + x_2$ has a given value x in the set $x = M, \ldots, M + X - 1$. Hence

$$|S|^2 \leqq X_0 X^{-1} S'.$$

Now

$$S' = \sum_{y_1=N}^{N+Y-1} \sum_{y=N}^{N+Y-1} \eta(y_1)\overline{\eta(y)} \sum_{x_1=M}^{M+X-1} \sum_{x_2=-X+1}^{X-1} e\Big((x_1 + x_2)(\Phi(y_1) - \Phi(y))\Big).$$

Applying Lemma 6 to the summations over x_1 and x_2, we obtain

$$S' \leqq \sum_{y_1=N}^{N+Y-1} |\eta(y_1)| \sum_{y=N}^{N+Y-1} \eta \min\left(2X^2, \frac{1}{4\,\|\Phi(y) - \Phi(y_1)\|^2}\right).$$

The Y values of y in the inner sum can be split into at most $[Yq^{-1} + 1]$ sets, each consisting of at most q consecutive values of y. To each such set we apply Lemma 8a, (I), with $\Phi(y) - \Phi(y_1)$ in place of $\Phi(y)$ and $X\sqrt{2}$ in place of U, and with the same λ. The sum over at most q consecutive values of y is of the form Ω_0, and is therefore less than

$$2(\lambda + 3)X^2 + 2qX\sqrt{2}.$$

Hence

$$S' \leqq \eta(2(\lambda + 3)X^2 + 2qX\sqrt{2})[Yq^{-1} + 1] \sum_{y_1=N}^{N+Y-1} | \eta(y_1) |$$
$$\leqq \eta((2\lambda + 6)X^2 + 3qX)[Yq^{-1} + 1]Y_1.$$

In view of the relation between S and S', this proves the result.

LEMMA 10c. *Let M, X, N, Y be integers, with $X > 0$, $Y > 0$. Suppose that $A \geqq 2\beta \geqq 2$. Let y run through the values $y = N, \ldots, N + Y - 1$, and for these values of y let the function $\Phi(y)$ take real values subject to the condition that*

$$\frac{1}{A} \leqq \Phi(y + 1) - \Phi(y) \leqq \frac{\beta}{A}.$$

Let

$$S = \sum_{x=M}^{M+X-1} \sum_{y=N}^{N+Y-1} \xi(x)\eta(y)e(x\Phi(y)),$$

and put

$$\sum_{x=M}^{M+X-1} | \xi(x) |^2 = X_0, \qquad \sum_{y=N}^{N+Y-1} | \eta(y) | = Y_1, \qquad \max | \eta(y) | = \eta.$$

Then

$$| S | \leq (X_0 Y_1 \eta(4X + 3A)[Y\beta A^{-1} + 1])^{\frac{1}{2}}.$$

Proof. The proof of the preceding lemma applies, down to the inequality

$$S' \leqq \sum_{y_1=N}^{N+Y-1} | \eta(y_1) | \sum_{y=N}^{N+Y-1} \eta \min \left(2X^2, \frac{1}{4\, ||\Phi(y) - \Phi(y_1)\, ||^2}\right).$$

The inner sum satisfies the hypotheses of Lemma 9, if we take $\Phi(y) - \Phi(y_1)$ in place of $\Phi(y)$ and $X\sqrt{2}$ in place of U in that lemma. Hence

$$S' \leqq \eta(4X^2 + 2AX\sqrt{2})[Y\beta A^{-1} + 1] \sum_{y_1=N}^{N+Y-1} \left| \eta(y_1) \right|$$
$$\leqq \eta(4X^2 + 3AX)[Y\beta A^{-1} + 1]Y_1.$$

In view of the relation between S and S', this proves the result.

LEMMA 11 (*Fourier series*). *Let $F(x) = P(x) + iQ(x)$ be a periodic function of x with period 1, and suppose that the interval $0 < x \leqq 1$ can be split up into a finite number of intervals, such* *that the real functions $P(x)$ and $Q(x)$ are continuous and monotonic in the interior of each. Suppose further that*

$$F(x) = \tfrac{1}{2}(F(x+0) + F(x-0))$$

at each point of discontinuity of the function. Then

$$F(x) = \tfrac{1}{2}a_0 + \sum_{m=1}^{\infty} (a_m \cos 2\pi mx + b_m \sin 2\pi mx),$$

where

$$a_m = 2\int_0^1 F(\xi) \cos 2\pi m\xi d\xi, \qquad b_m = 2\int_0^1 F(\xi) \sin 2\pi m\xi d\xi.$$

LEMMA 12. *Let r be a positive integer, and let α, β, Δ be real numbers satisfying*

$$0 < \Delta < \tfrac{1}{2}, \qquad \Delta \leqq \beta - \alpha \leqq 1 - \Delta.$$

Then there exists a periodic function $\psi(x)$, with period 1, satisfying

(i) $\psi(x) = 1$ *in the interval* $\alpha + \frac{1}{2}\Delta \leqq x \leqq \beta - \frac{1}{2}\Delta$,
(ii) $\psi(x) = 0$ *in the interval* $\beta + \frac{1}{2}\Delta \leqq x \leqq 1 + \alpha - \frac{1}{2}\Delta$,
(iii) $0 \leqq \psi(x) \leqq 1$ *in the remainder of the interval* $\alpha - \frac{1}{2}\Delta \leqq x \leqq 1 + \alpha - \frac{1}{2}\Delta$,
(iv) $\psi(x)$ *has an expansion in Fourier series of the form*

$$\psi(x) = \beta - \alpha + \sum_{m=1}^{\infty} (a_m \cos 2\pi mx + b_m \sin 2\pi mx),$$

where

$$\begin{aligned} |a_m| &\leqq 2(\pi m)^{-1}, & |b_m| &\leqq 2(\pi m)^{-1}, \\ |a_m| &\leqq 2(\beta - \alpha), & |b_m| &\leqq 2(\beta - \alpha), \\ |a_m| &< \frac{2}{\pi m}\left(\frac{r}{\pi m\Delta}\right)^r, & |b_m| &< \frac{2}{\pi m}\left(\frac{r}{\pi m\Delta}\right)^r. \end{aligned}$$

Proof. Let $\psi_0(x)$ be the periodic function of period 1, defined by
$\psi_0(x) = 1$ in the interval $\alpha < x < \beta$,
$\psi_0(x) = 0$ in the interval $\beta < x < 1 + \alpha$,
$\psi_0(x) = \frac{1}{2}$ for $x = \alpha$ and $x = \beta$.

Expanding this function in a Fourier series, we obtain

$$\psi_0(x) = \tfrac{1}{2}a_{0,0} + \sum_{m=1}^{\infty} (a_{m,0} \cos 2\pi mx + b_{m,0} \sin 2\pi mx).$$

where

$$a_{0,0} = 2\int_{\alpha}^{\alpha+1} \psi_0(x)dx = 2(\beta - \alpha),$$

$$a_{m,0} = 2\int_{\alpha}^{\beta} \cos 2\pi mx dx = \frac{1}{m\pi} (\sin 2\pi m\beta - \sin 2\pi m\alpha),$$

$$b_{m,0} = 2\int_{\alpha}^{\beta} \sin 2\pi mx dx = \frac{1}{m\pi} (\cos 2\pi m\alpha - \cos 2\pi m\beta).$$

Define δ by $2r\delta = \Delta$. Define the functions $\psi_1(x), \ldots, \psi_r(x)$, all of period 1, by the recurrence relation

$$\psi_\varrho(x) = \frac{1}{2\delta}\int_{-\delta}^{\delta} \psi_{\varrho-1}(x+z)dz, \quad \varrho = 1, \ldots, r.$$

We shall prove by induction on ϱ that

(i) $\psi_\varrho(x) = 1$ in the interval $\alpha + \varrho\delta < x < \beta - \varrho\delta$,

(ii) $\psi_\varrho(x) = 0$ in the interval $\beta + \varrho\delta < x < 1 + \alpha - \varrho\delta$,

(iii) $0 \leqq \psi_\varrho(x) \leqq 1$ in the intervals $\alpha - \varrho\delta \leqq x \leqq \alpha + \varrho\delta$ and $\beta - \varrho\delta \leqq x \leqq \beta + \varrho\delta$,

(iv) $\psi_\varrho(x)$ has the Fourier series expansion

$$\psi_\varrho(x) = \beta - \alpha + \sum_{m=1}^{\infty} (a_{m,\varrho} \cos 2\pi mx + b_{m,\varrho} \sin 2\pi mx),$$

where

$$a_{m,\varrho} = \frac{1}{m\pi} (\sin 2\pi m\beta - \sin 2\pi m\alpha) \left(\frac{\sin 2\pi m\delta}{2\pi m\delta}\right)^\varrho,$$

$$b_{m,\varrho} = \frac{1}{m\pi} (\cos 2\pi m\alpha - \cos 2\pi m\beta) \left(\frac{\sin 2\pi m\delta}{2\pi m\delta}\right)^\varrho.$$

Suppose these four properties hold for $\psi_{\varrho-1}(x)$. It follows at once that the first three hold for $\psi_\varrho(x)$, since $\psi_\varrho(x)$ is the average of $\psi_{\varrho-1}(x)$ in the interval $(x - \delta, x + \delta)$. As regards the fourth property, we have

$$a_{m,\varrho} = 2\int_0^1 \left\{ \frac{1}{2\delta} \int_{-\delta}^{\delta} \psi_{\varrho-1}(\xi + z)dz \right\} \cos 2\pi m\xi d\xi$$

$$= \frac{1}{\delta}\int_{-\delta}^{\delta} dz \int_0^1 \psi_{\varrho-1}(\xi + z) \cos 2\pi m\xi d\xi$$

$$= \frac{1}{\delta}\int_{-\delta}^{\delta} dz \int_0^1 \psi_{\varrho-1}(\xi) \cos 2\pi m(\xi - z) d\xi$$

$$= \frac{1}{2\delta}\int_{-\delta}^{\delta} \{a_{m,\varrho-1} \cos 2\pi mz + b_{m,\varrho-1} \sin 2\pi mz\}dz = a_{m,\varrho-1}\frac{\sin 2\pi m\delta}{2\pi m\delta},$$

and the analogous relation holds for $b_{m,\varrho}$.

Putting $\psi(x) = \psi_r(x)$, and recalling that $2r\delta = \Delta$, the results stated in the enunciation follow at once.

LEMMA 13 (*van der Corput's Lemma*). *Let M and M_1 be integers with $M < M_1$, and let $f(x)$ be a twice differentiable real function defined in the interval $M \leqq x \leqq M_1$ and satisfying*

$$0 \leqq f'(x) \leqq \tfrac{1}{2}, \ f''(x) \geqq 0.$$

Then, taking either both the $+$ signs or both the $-$ signs, we have

$$\sum_{x=M}^{M_1} e(\pm f(x)) = \int_M^{M_1} e(\pm f(x))dx + 2\theta.$$

Proof. It suffices to consider the case of the $+$ signs. We apply Lemma 11 to the function $F(\xi)$ of period 1, defined for $0 < \xi < 1$ by $F(\xi) = e(f(x + \xi))$ and for $\xi = 0$ by

$$F(0) = \tfrac{1}{2}\{e(f(x)) + e(f(x + 1))\}.$$

Putting $\xi = 0$ in the Fourier series for $F(\xi)$, we obtain

$$\tfrac{1}{2}\{e(f(x)) + e(f(x + 1))\} = \tfrac{1}{2}a_0 + \sum_{m=1}^{\infty} a_m,$$

where

$$\tfrac{1}{2}a_0 = \int_0^1 e(f(x + \xi))d\xi,$$

$$a_m = \int_0^1 e(f(x + \xi)) \{e(m\xi) + e(-m\xi)\} d\xi.$$

Summing the above result for $x = M, \ldots, M_1 - 1$, we have

$$\sum_{x=M}^{M_1} e(f(x)) - \tfrac{1}{2}e(f(M)) - \tfrac{1}{2}e(f(M_1)) - \int_M^{M_1} e(f(\xi))d\xi$$
$$= \sum_{m=1}^{\infty} \int_M^{M_1} \{e(m\xi) + e(-m\xi)\}e(f(\xi))d\xi.$$

We can suppose without loss of generality that the common value of these two expressions is real and non-negative, since this can always be ensured by considering $f(x) + \lambda$ instead of $f(x)$, for a suitable real number λ. On integration by parts, the last series becomes

$$\sum_{m=1}^{\infty} \left(-\frac{1}{2\pi im}\right) \int_M^{M_1} \{e(m\xi) - e(-m\xi)\} de(f(\xi)).$$
$$= -\sum_{m=1}^{\infty} \frac{1}{2\pi im} \int_M^{M_1} \frac{f'(\xi)}{m + f'(\xi)}\, de(m\xi + f(\xi))$$
$$- \sum_{m=1}^{\infty} \frac{1}{2\pi im} \int_M^{M_1} \frac{f'(\xi)}{m - f'(\xi)}\, de(-m\xi + f(\xi)).$$

Since the value of the whole expression is real, we can replace the first sum by

$$\sum_{m=1}^{\infty} \left(-\frac{1}{2\pi m}\right) \int_M^{M_1} \frac{f'(\xi)}{m + f'(\xi)}\, d \sin 2\pi(m\xi + f(\xi))$$

and the second sum by a similar expression.

The function $f'(\xi)/(m+f'(\xi))$ is non-negative and non-decreasing in the interval of integration. Hence, by the second mean value theorem,

$$\int_M^{M_1} \frac{f'(\xi)}{m + f'(\xi)}\, d \sin 2\pi(m\xi + f(\xi))$$
$$= \frac{f'(M_1)}{m + f'(M_1)} \{\sin 2\pi(mM_1 + f(M_1)) - \sin 2\pi(mM_0 + f(M_0)\},$$

where $M < M_0 < M_1$. The absolute value of the last expression does not exceed

$$\frac{\frac{1}{2}}{m + \frac{1}{2}} \cdot 2 = \frac{2}{2m + 1}.$$

Similarly for the integral in the second sum.

Finally, allowing for the two terms $\frac{1}{2}e(f(M))$ and $\frac{1}{2}e(f(M_1))$, we see that the difference between the sum and the integral in the enunciation does not exceed, in absolute value,

$$\frac{1}{2}+\frac{1}{2}+\sum_{m=1}^{\infty}\frac{1}{2\pi m}\left(\frac{2}{2m+1}+\frac{2}{2m-1}\right)$$
$$=1+\frac{2}{\pi}\sum_{m=1}^{\infty}\frac{1}{m}\left(\frac{m}{2m-1}-\frac{m}{2m+1}\right)=1+\frac{2}{\pi}<2.$$

LEMMA 14*a*. *Let* $P \geqq 1$, *let* z *be real, and let*

$$I=\int_0^P e(zx^n)dx.$$

Then

$$|I| \leqq \begin{cases} P & \text{if } |z| \leqq P^{-n}, \\ |z|^{-\nu}\sqrt{2} & \text{if } |z| > P^{-n}. \end{cases}$$

Proof. The first of the two results is obvious, so we can suppose that $z > P^{-n}$. Making the change of variable $2zx^n = u$, we obtain

$$I = U + iV,$$

where

$$U=\int_0^\sigma \psi(u)\cos \pi u du, \quad V=\int_0^\sigma \psi(u)\sin \pi u du,$$

and

$$\sigma = 2zP^n > 2, \quad \psi(u)=\nu(2z)^{-\nu}u^{\nu-1}.$$

We express U as

$$\int_0^{\frac{1}{2}}+\int_{\frac{1}{2}}^{\frac{3}{2}}+\cdots+\int_{k-\frac{1}{2}}^{\sigma}\psi(u)\cos \pi u\, du,$$

where $k=[\sigma+\frac{1}{2}]$. Since $\psi(u)$ is positive and decreasing, the first integral is positive, and the remaining integrals are alternately negative and positive and decrease in absolute value. Hence

$$|U| \leqq \max\left(\int_0^{\frac{1}{2}}\psi(u)du, \int_{\frac{1}{2}}^{\frac{3}{2}}\psi(u)du\right),$$

whence

$$|U| \leqq \int_0^1 \psi(u) du = (2z)^{-\nu}.$$

Similarly, on expressing V as

$$\int_0^1 + \int_1^2 + \ldots + \int_l^\sigma \psi(u) \sin \pi u \, du,$$

we see that the same conclusion holds for $|V|$. Finally,

$$|I| \leqq \left((2z)^{-2\nu} + (2z)^{-2\nu}\right)^{\frac{1}{2}} < z^{-\nu}\sqrt{2}.$$

LEMMA 14*b*. *Let* $N \geqq 2$, *and let* $\log N = r$. *Let* z *be real and let*

$$I(z) = \int_2^N \frac{e(zx)}{r} dx, \quad J(z) = \int_2^N \frac{e(zx)}{\log x} dx.$$

Then

$$I(z) \ll r^{-1} \min(N, |z|^{-1}) \text{ for all } z,$$
$$J(z) \ll r^{-1} \min(N, |z|^{-1}) \text{ for } |z| \leq N^{-\frac{1}{2}}.$$

Proof. For the integral $I(z)$ the result is trivial. The estimate Nr^{-1} for $J(z)$ is also trivial. It suffices, therefore, to prove that if $N^{-1} \leq z \leqq N^{-\frac{1}{2}}$ then

$$\int_2^N \frac{e(zx)}{\log x} dx \ll z^{-1} r^{-1}.$$

Now

$$\int_2^{\sqrt{(2N)}} \frac{dx}{\log x} \ll N^{\frac{1}{2}} r^{-1} \ll z^{-1} r^{-1},$$

and

$$\int_{\sqrt{(2N)}}^N \frac{e(zx)}{\log x} dx \ll \frac{1}{\log \sqrt{(2N)}} \cdot \frac{1}{z} \ll z^{-1} r^{-1}$$

by the second mean value theorem. Hence the result.

LEMMA 15. *Let* $h, \ldots, l, n$ *be* g *fixed integers satisfying* $0 < h < \ldots < l < n$, *and let* $k \geqq 1$ *be an integer. For each* $t = 1, 2, \ldots, k$ *let there be given a non-empty set* S_t *of "points", not necessarily distinct, where a "point" means a set of* g *integers* $(u_h, \ldots, u_l, u_n)$.

Suppose that all points $(U_{t,h}, \ldots, U_{t,n})$ *of the set* S_t *satisfy*

$$U_{t,h} \ll Mp_t^h, \ldots, U_{t,n} \ll Mp_t^n,$$

where $M \geqq 1$, $p_1 > 1$, *and*

$$1 < p_t \ll (p_{t-1})^{1-\nu} \text{ for } t = 2, \ldots, k.$$

Suppose further that for any intervals of lengths

$$Mp_t^{h(1-\nu)}, \ldots, Mp_t^{n(1-\nu)},$$

the number of points of the set S_t *whose coordinates fall respectively into these intervals is at most* Φ_t.

Consider all possible selections of k *points, one from each of the sets* $S_1, \ldots, S_k$. *For such a selection, denote the point selected from* S_t *by* $(U_{t,h}, \ldots, U_{t,n})$. *Put*

$$U_h = U_{1,h} + \ldots + U_{k,h}, \ldots, U_n = U_{1,n} + \ldots + U_{k,n}.$$

Then, for every selection, we have

$$U_h \ll Mp_1^h, \ldots, U_n \ll Mp_1^n. \tag{1}$$

Also, for any given integers $z_h, \ldots, z_l, z_n$, *the number* $\psi(z_h, \ldots, z_l, z_n)$ *of selections of points* $(U_{t,h}, \ldots, U_{t,n})$ *for which*

$$U_h = z_h, \ldots, U_l = z_l, \; U_n = z_n \tag{2}$$

satisfies

$$\psi(z_h, \ldots, z_l, z_n) \ll \Phi_1 \ldots \Phi_k. \tag{3}$$

Proof. The inequality (1) follows immediately from the hypotheses of the lemma, since p_1 is the greatest of $p_1, \ldots, p_k$. It is to be understood, of course, that k is bounded.

We have to prove the inequality (3). We can write the equations (2) as

$$U_{1,r} = z_r - (U_{2,r} + \ldots + U_{k,r}),$$

where r takes the values $h, \ldots, l, n$. The sum in brackets is always $\ll Mp_2^r \ll Mp_1^{r(1-\nu)}$. Thus, for given values of $z_h, \ldots, z_n$, the coordinates of the point

$$(U_{1,h}, \ldots, U_{1,n})$$

lie in given intervals whose lengths are respectively

$$\ll Mp_1^{h(1-\nu)}, \ldots, \ll Mp_1^{n(1-\nu)}.$$

It follows from the hypotheses that the number of possible choices for the point $(U_{1,h}, \ldots, U_{1,n})$ is $\ll \Phi_1$.

Having chosen this point, we can write the equations (2) as

$$U_{2,r} = (z_r - U_{1,r}) - (U_{3,r} + \ldots + U_{k,r}),$$

where r takes the values $h, \ldots, l, n$. The sum in brackets is always $\ll Mp_3^r \ll Mp_2^{r(1-\nu)}$. Hence, for given $z_h, \ldots, z_n$, and given $U_{1,h}, \ldots, U_{1,n}$, the number of possible choices for the point $(U_{2,h}, \ldots, U_{2,n})$ is $\ll \Phi_2$.

Continuing in this way, we find that the number of possible selections of k points, satisfying (2), from the k sets $S_1, \ldots, S_k$ is

$$\ll \Phi_1 \ldots \Phi_k.$$

This proves (3).

LEMMA 16. *Let $p = RH$, where $R > 1$, $H > 1$. Suppose that*

$$-p \leq X_1 < Y_1,\ Y_1 + R \leqq X_2 < Y_2, \ldots, Y_{n-1} + R \leqq X_n < Y_n \leqq p.$$

Let $v_1, \ldots, v_n$ take all integral values in the intervals

$$X_1 < v_1 \leqq Y_1, \ldots, X_n < v_n \leqq Y_n.$$

Let E denote the number of sets $v_1, \ldots, v_n$ for which the sums

$$v_1 + \ldots + v_n, \ldots, v_1^n + \ldots + v_n^n$$

fall into any given intervals whose lengths are respectively

$$p^{1-\nu}, \ldots, p^{n(1-\nu)}.$$

Then

$$E \ll H^{\frac{1}{2}n(n-1)} p^{\frac{1}{2}(n-1)}.$$

Proof. Instead of the given intervals of lengths

$$p^{1-\nu}, \ldots, p^{n(1-\nu)}$$

we consider intervals whose lengths are respectively

$$1,\ p, \ldots, p^{n-1}.$$

Each such interval is shorter than the corresponding interval of the given set, except for the last interval which is of the same length. We shall prove that the number of sets of values of

$v_1, \ldots, v_n$ for which the sums in question fall into these shorter intervals is

$$\ll H^{\frac{1}{2}n(n-1)}.$$

This will imply the result stated; for we have then to multiply this number by

$$[p^{1-\nu}+1][p^{1-2\nu}+1]\ldots[p^{1-n\nu}+1] \\ \ll p^{n-\nu(1+2+\ldots+n)} = p^{\frac{1}{2}(n-1)}.$$

Let $v_1, \ldots, v_n$ and $v_1', \ldots, v_n'$ be two sets of values for which the sums

$$s_k = v_1^k + \ldots + v_n^k,\ s_k' = v_1'^k + \ldots + v_n'^k$$

both lie in given intervals whose lengths, for $k = 1, \ldots, n$, are respectively $1, p, \ldots, p^{n-1}$. Then

(4) $$|s_k - s_k'| \leqq p^{k-1} \text{ for } k = 1, \ldots, n.$$

Let $\sigma_1, \ldots, \sigma_n$ be the elementary symmetric functions of $v_1, \ldots, v_n$ and similarly for $\sigma_1', \ldots, \sigma_n'$. Since all the values of the variables v_j and v_j' lie between $-p$ and p, we obviously have

(5) $$s_k \ll p^k,\ s_k' \ll p^k,\ \sigma_k \ll p^k,\ \sigma_k' \ll p^k$$

for $k = 1, \ldots, n$.

The power sums $s_1, \ldots, s_n$ and the elementary symmetric functions $\sigma_1, \ldots, \sigma_n$ are connected by the well known formulae of Newton:

$$s_k - \sigma_1 s_{k-1} + \sigma_2 s_{k-2} - \ldots + (-1)^{k-1}\sigma_{k-1}s_1 + (-1)^k k\sigma_k = 0 \\ (k = 1, \ldots, n).$$

These formulae, together with the inequalities (4) and (5), imply that

(6) $$\sigma_j - \sigma_j' \ll p^{j-1} \text{ for } j = 1, \ldots, n.$$

Since $\sigma_1 = s_1$ and $\sigma_1' = s_1'$, this is obviously true when $j = 1$. Also, if (6) holds for $j < k$, we have

$$\sigma_j s_{k-j} - \sigma_j' s_{k-j}' = (\sigma_j - \sigma_j')s_{k-j} + \sigma_j'(s_{k-j} - s_{k-j}') \ll p^{k-1}$$

for $j < k$. It follows from Newton's formulae that

$$k\,|\,\sigma_k - \sigma_k'\,| \leqq |\,s_k - s_k'\,| + |\,\sigma_1 s_{k-1} - \sigma_1' s_{k-1}'\,| + \ldots + |\,\sigma_{k-1} s_1 - \sigma_{k-1}' s_1'\,| \ll p^{k-1},$$

which proves (6) for $j = k$. Hence, by induction, (6) holds for $j = 1, \ldots, n$.

Now

$$(v - v_1) \ldots (v - v_n) = v^n - \sigma_1 v^{n-1} + \ldots + (-1)^n \sigma_n.$$

Hence, if v is any integer satisfying $|\,v\,| \leqq p$, we have

$$(v - v_1) \ldots (v - v_n) - (v - v_1') \ldots (v - v_n') \ll p^{n-1}$$

by (6). In particular, putting $v = v_n'$, we have

$$(v_n' - v_1) \ldots (v_n' - v_n) \ll p^{n-1}.$$

Since $|\,v_n' - v_j\,| \geqq R$ for $j < n$ by the conditions imposed on the intervals $X_j < v_j \leqq Y_j$ of the enunciation, it follows that

$$v_n - v_n' \ll \frac{p^{n-1}}{R^{n-1}} = H^{n-1}.$$

This shows that there are $\ll H^{n-1}$ possible values for v_n. When v_n is fixed, the same argument can be applied to the variables $v_1, \ldots, v_{n-1}$ and the sums

$$v_1 + \ldots + v_{n-1}, \ldots, v_1^{n-1} + \ldots + v_{n-1}^{n-1},$$

and shows that there are $\ll H^{n-2}$ possible values for v_{n-1}; and so on. Finally, the number of possible sets of values for $v_1, \ldots, v_n$ is

$$\ll H^{n-1} \,.\, H^{n-2} \ldots 1 = H^{\frac{1}{2}n(n-1)}.$$

Lemma 17. *Let k and l be fixed positive integers, and let $\tau_k(m)$ denote the number of solutions of the equation $x_1 \ldots x_k = m$ in positive integers (so that, in particular, $\tau_2(m) = \tau(m)$, the number of divisors of m). Then*

(i) $\tau_k(m) \ll m^{\varepsilon}$;

(ii) $\displaystyle\sum_{0<m\leqq z} \tau(m) = z\,(\log z + 2E - 1) + O(z^{\frac{1}{2}})$;

(iii) $\displaystyle\sum_{0<m\leqq z} (\tau_k(m))^l \ll z(\log z + 1)^{k^l-1}$;

where E denotes Euler's constant.

Proof. The results (i) and (ii) are well known. An elementary proof of (iii) was given by C. Mardjanichvili in Doklady Akad. Nauk SSSR, 22 (1939), No. 7.

NOTES ON CHAPTER I

Many of the lemmas in this chapter represent well known techniques in the analytic theory of numbers.

Lemma 4 is of course the fundamental principle of the Hardy-Littlewood method. In the original memoirs of Hardy and Littlewood, the principle was used in a superficially different form, namely that if $f(x) = \sum_{0}^{\infty} a_n x^n$ for $|x| < 1$, then

$$r^N a_N = \frac{1}{2\pi}\int_0^{2\pi} f(re^{i\alpha})e^{-iN\alpha}d\alpha$$

for $0 < r < 1$. It was Vinogradov who recognized that it is technically simpler to work with finite exponential sums instead of with power series.

Lemma 7 is one of the simplest results on Diophantine approximation, and is due to Dirichlet. For a proof not depending on a knowledge of continued fractions, see Hardy and Wright, *An introduction to the theory of numbers* (Oxford, 1945), Theorem 36 (p. 30).

The use of Fourier series, as in Lemma 12, is essentially that introduced into the analytic theory of numbers by H. Weyl in his great memoir in *Math. Annalen*, 77 (1916), 313—352, which also contains his estimate for Weyl sums.

Lemma 15 represents a generalization of a technique introduced by Hardy and Littlewood in the sixth memoir of their famous series "On some problems of Partitio Numerorum", [*Math. Zeitschrift*, 23 (1925), 1—37].

It may help the reader to appreciate the underlying idea if we examine a particular case, which is essentially that of Hardy and Littlewood. Suppose $g = 1$, so that the numbers $h, \ldots, l, n$

reduce to a single number, say n. Take the sets S_t, for $t=1, \ldots, k$, to consist of the nth powers of the distinct integers between (say) p_t and $2p_t$, where $p_t = p^{(1-\nu)^{t-1}}$. We can take M to be 1. The number Φ_t in the lemma is the number of nth powers of integers between p_t and $2p_t$ which lie in an interval of length $p_t^{n(1-\nu)} = p_t^{n-1}$, and so is bounded. The essential point of the lemma is that the numbers

$$U_n = x_1^n + \ldots + x_k^n, \text{ where } p_t < x_t < 2p_t,$$

are almost distinct. That is, the number of sets $x_1, \ldots, x_k$ for which U_n assumes a given value is bounded (assuming k to be fixed). Thus in the original form of Hardy and Littlewood, the method is one for constructing distinct sums of k nth powers, and it is used for this purpose in connection with Waring's Problem, as for instance in Lemma 1 of Chapter IV.

Vinogradov's more general form of the principle is applied later in the book (in Chapters V and VI) in conjunction with Lemma 16.

Lemma 16 shows that the sums of the powers, up to the nth, of n variables which run through well-separated intervals are (in a sense) independently and uniformly distributed. Lemma 15 will be used to facilitate the repeated application of Lemma 16.

Vinogradov's proof of Lemma 16 has been replaced in the present text by a simpler proof due to Hua [*Quart. J. of Math.* (Oxford), 20 (1949), 48—61]. Vinogradov's formulation of the lemma was slightly more general. Instead of the sums

$$v_1 + \ldots + v_n, \ldots, v_1^n + \ldots + v_n^n$$

there stood the sums

$$\varkappa_1 v_1 + \ldots + \varkappa_n v_n, \ldots, \varkappa_1 v_1^n + \ldots + \varkappa_n v_n^n,$$

where each $\varkappa$ is $+1$ or -1. The proof given in the text is also easily adapted to give this slightly more general result. For we may assume, without loss of generality, that

$$\varkappa_1 = \ldots = \varkappa_r = +1, \ \varkappa_{r+1} = \ldots = \varkappa_n = -1,$$

where the last equation will not occur if $r = n$. Let $v_1, \ldots, v_n$ and $v_1', \ldots, v_n'$ be two sets of values for which the sums

$$\varkappa_1 v_1^k + \ldots + \varkappa_n v_n^k \text{ and } \varkappa_1 v_1'^k + \ldots + \varkappa_n v_n'^k$$

both lie in given intervals whose lengths, for $k = 1, \ldots, n$, are respectively $1, p, \ldots, p^{n-1}$. Then, writing

$$s_k = v_1^k + \ldots + v_r^k + v_{r+1}'^k + \ldots + v_n'^k,$$
$$s_k' = v_1'^k + \ldots + v_r'^k + v_{r+1}^k + \ldots + v_n^k,$$

we have

$$| s_k - s_k' | \leqq p^{k-1} \text{ for } k = 1, \ldots, n.$$

With this new definition of s_k and s_k' (and the corresponding definition of σ_k and σ_k') the proof goes through as before, apart from obvious changes.

CHAPTER II

The Investigation of the Singular Series in Waring's Problem

In the present chapter we establish some properties of the "singular series" $\mathfrak{S}$, defined below, which will be used in later chapters (IV and VII). This series was first discovered and investigated by Hardy and Littlewood.

Notation in this chapter. In the present chapter we suppose $n \geqq 3$ and use the following notations.

For $(a, q) = 1$, $q > 0$ we put

$$S(a, q) = \sum_{x=0}^{q-1} e_q(ax^n). \tag{1}$$

For integral $q > 0$, integral N and fixed positive integral r, we denote by the symbol $M(q) = M(q, N, r)$ the number of solutions of the congruence

$$x_1{}^n + \dots + x_r{}^n \equiv N \pmod{q},$$

when $x_1, \dots, x_r$ run independently through complete sets of residues to the modulus q. Further, letting a run through a reduced set of residues to the modulus q, we put

$$A(q) = A(q, N, r) = q^{-r} \sum_a \{S(a, q)\}^r e_q(-aN). \tag{2}$$

We define $\mathfrak{S}$ by

$$\mathfrak{S} = \mathfrak{S}(N, r) = \sum_{q=1}^{\infty} A(q, N, r) \tag{3}$$

(provided the infinite series converges).

By the letter p we denote a prime. It is well known that for $p > 2$ and any integer $s > 0$ there exists a primitive root g to the modulus p^s, that is, a number g whose order with respect to the modulus p^s is $\varphi(p^s)$. Every number in a reduced set of residues to the modulus p^s is $\equiv g^b \pmod{p^s}$ for a unique b with $0 \leqq b < \varphi(p^s)$.

For $p = 2$ there is no primitive root (mod 2^s) if $s > 2$, but there is a number g (for example, 5) whose order with respect to the modulus 2^s is $\frac{1}{2}\varphi(2^s)$, and every number of the form $4m + 1$ in a reduced set of residues to the modulus 2^s is $\equiv g^b \pmod{2^s}$ for a unique b with $0 \leqq b < \frac{1}{2}\varphi(2^s)$.

By the letter τ we denote the exponent to which p enters into the canonical factorization of the number n. We put

$$\gamma = \begin{cases} \tau + 1 & \text{for } p > 2, \\ \tau + 2 & \text{for } p = 2. \end{cases}$$

We write

$$\psi(p) = \psi(p, N, \gamma) = \sum_{s=0}^{\infty} A(p^s, N, \gamma) \tag{4}$$

(provided the infinite series converges).

LEMMA 1. *For any integers $q_1, \ldots, q_k$ which are relatively prime in pairs, we have*

$$S(a_1, q_1) \ldots S(a_k, q_k) = S(a_1 Q_1 + \ldots + a_k Q_k,\ q_1 \ldots q_k),$$

where $$Q_s = q_1 \ldots q_k q_s^{-1} \text{ for } s = 1, \ldots, k.$$

Proof. We have

$$\begin{aligned} S(a_1, q_1) \ldots S(a_k, q_k) &= \sum_{x_1=0}^{q_1-1} \ldots \sum_{x_k=0}^{q_k-1} e\left(\left(\frac{a_1}{q_1} + \ldots + \frac{a_k}{q_k}\right)(Q_1 x_1 + \ldots + Q_k x_k)^n\right) \\ &= \sum_{x_1=0}^{q_1-1} \ldots \sum_{x_k=0}^{q_k-1} e_{q_1 \ldots q_k}\left((a_1 Q_1 + \ldots + a_k Q_k)(Q_1 x_1 + \ldots + Q_k x_k)^n\right), \end{aligned}$$

and this proves the lemma, since $Q_1 x_1 + \ldots + Q_k x_k$ runs through a complete set of residues to the modulus $q_1 \ldots q_k$.

LEMMA 2. *For any integers $q_1, \ldots, q_k$, relatively prime in pairs, we have*

$$A(q_1) \ldots A(q_k) = A(q_1 \ldots q_k).$$

Proof. If $a_1, \ldots, a_k$ run through the reduced sets of residues to the moduli $q_1, \ldots, q_k$, we have, by Lemma 1,

$$A(q_1)\ldots A(q_k)$$
$$=(q_1\ldots q_k)^{-r}\sum_{a_1}\ldots\sum_{a_k}\{S(a_1,q_1)\ldots S(a_k,q_k)\}^r e\left(-\left(\frac{a_1}{q_1}+\ldots+\frac{a_k}{q_k}\right)N\right)$$
$$=(q_1\ldots q_k)^{-r}\sum_{a_1}\ldots\sum_{a_k}\{S(a_1Q_1+\ldots+a_kQ_k,\ q_1\ldots q_k)\}^r e\left(-\frac{a_1Q_1+\ldots+a_kQ_k}{q_1\ldots q_k}N\right).$$

This proves the lemma, since $a_1Q_1+\ldots+a_kQ_k$ runs through a reduced set of residues to the modulus $q_1\ldots q_k$.

LEMMA 3. *We have*

$$|S(a,p)|\leqq(\delta-1)p^{\frac{1}{2}},\ \textit{where}\ \delta=(n,\ p-1).$$

Proof. It is well known that if $(z,\ p)=1$ the congruence $x^n\equiv z\pmod p$ is soluble if and only if the index of z (that is, the number b for which $z\equiv g^b\pmod p$) is a multiple of δ, and moreover when it is soluble the congruence has δ solutions. Therefore in the case $\delta=1$ we have, by Lemma 5 of Chapter I, $S(a,p)=0$. In the case $\delta>1$ we obtain, using the same lemma,

$$|S(a,p)|=\left|1+\sum_{m=0}^{\delta-1}\sum_{z=1}^{p-1}e\left(\frac{m\operatorname{ind}z}{\delta}\right)e_p(az)\right|$$
$$=\left|\sum_{m=1}^{\delta-1}\sum_{z=1}^{p-1}e\left(\frac{m\operatorname{ind}z}{\delta}\right)e_p(az)\right|$$
$$\leq\left\{(\delta-1)\sum_{m=1}^{\delta-1}\sum_{z_1=1}^{p-1}\sum_{z=1}^{p-1}e_\delta(m(\operatorname{ind}z_1-\operatorname{ind}z))e_p(a(z_1-z))\right\}^{\frac{1}{2}}.$$

Hence, collecting together for each $t=1,\ldots,p-1$ the terms satisfying the condition $z_1\equiv tz\pmod p$, we have

$$|S(a,p)|^2\leq(\delta-1)\sum_{m=1}^{\delta-1}\sum_{t=1}^{p-1}\sum_{z=1}^{p-1}e_\delta(m\operatorname{ind}t)e_p(a(t-1)z)$$
$$=(\delta-1)\sum_{m=1}^{\delta-1}\left(p-1-\sum_{t=2}^{p-1}e_\delta(m\operatorname{ind}t)\right).$$

As t runs from 1 to $p-1$, the index of t assumes all values $(\mathrm{mod}\ \delta)$ equally often. Hence the last expression is

$$(\delta-1)\sum_{m=1}^{\delta-1}p=(\delta-1)^2p.$$

LEMMA 4. *Let α be an integer, and suppose that $1 < \alpha \leqq n$ and $(n, p) = 1$. Then*

$$S(a, p^\alpha) = p^{\alpha-1}.$$

Proof. Transforming the sum $S(a, p^\alpha)$ by the substitution

$$x = p^{\alpha-1}\xi + z,$$

where ξ and z run independently through the values

$$\xi = 0, \ldots, p-1, \; z = 0, \ldots, p^{\alpha-1}-1,$$

we reduce the general term of the sum to

$$e\left(\frac{az^n}{p^\alpha} + \frac{anz^{n-1}\xi}{p}\right),$$

since $2(\alpha - 1) \geqq \alpha$. The sum of the terms corresponding to any z not divisible by p is equal to zero. Therefore the sum $S(a, p^\alpha)$ is equal to the sum of those of its terms for which z is a multiple of p, i.e. x is a multiple of p. But the number of such terms is equal to $p^{\alpha-1}$ and each of them is 1.

LEMMA 5. *Let α be an integer, $\alpha > n$. Then*

$$S(a, p^\alpha) = p^{n-1}S(a, p^{\alpha-n}).$$

Proof. By the definition of τ, we have $n \geqq 2^\tau \geqq \tau + 1$, whence $\alpha \geqq \tau + 2$. Transforming the sum $S(a, p^\alpha)$ by the substitution

$$x = p^{\alpha-\tau-1}\xi + z,$$

where ξ and z run independently through the values

$$\xi = 0, \ldots, p^{\tau+1}-1, \; z = 0, \ldots, p^{\alpha-\tau-1}-1,$$

we reduce the general term of the sum to

$$e\left(\frac{az^n}{p^\alpha} + \frac{anz^{n-1}\xi}{p^{\tau+1}}\right),$$

on recalling that p^τ divides n and $\alpha \geqq \tau + 2$. The sum of the terms corresponding to any z not divisible by p is equal to zero, since n is not divisible by $p^{\tau+1}$. Thus the sum $S(a, p^\alpha)$ is equal to the sum of those of its terms for which z is a multiple of p, i.e. for which x is a multiple of p. Hence

$$S(a, p^\alpha) = \sum_{x_1=0}^{p^{\alpha-1}-1} e_{p^\alpha}(ap^n x_1{}^n) = \frac{p^{\alpha-1}}{p^{\alpha-n}} S(a, p^{\alpha-n}).$$

LEMMA 6. *We have*

$$|S(a, q)| < n^{n^6} q^{1-\nu} \text{ and } A(q) \ll q^{1-r\nu}.$$

Proof. We confine ourselves to the proof of the first inequality as the second follows immediately from it. Let

$$q = p_1^{\alpha_1} \dots p_k^{\alpha_k}$$

be the canonical factorization of the number q. We apply Lemma 1, putting $q_s = p_s^{\alpha_s}$ and defining $a_1, \dots, a_k$ (as is always possible) by the congruence $a \equiv a_1 Q_1 + \dots + a_k Q_k \pmod q$, with the Q's of Lemma 1. Then, defining the symbol $T(a, q)$ by the equation $S(a, q) = q^{1-\nu} T(a, q)$, we obtain

$$T(a, q) = T(a_1, p_1^{\alpha_1}) \dots T(a_k, p_k^{\alpha_k}).$$

If $1 \leqq \alpha \leqq n$ and $(n, p) = p$ we have

$$|T(a, p^\alpha)| = p^{-\alpha(1-\nu)} |S(a, p^\alpha)| \leqq p^{\alpha\nu} \leqq p \leqq n;$$

if $\alpha = 1$ and $(n, p) = 1$ we have, by Lemma 3,

$$|T(a, p^\alpha)| < p^{-(1-\nu)} n p^{\frac{1}{2}} \leqq n p^{-\frac{1}{6}};$$

if $1 < \alpha \leqq n$ and $(n, p) = 1$ we have, by Lemma 4,

$$|T(a, p^\alpha)| = p^{-\alpha(1-\nu)} p^{\alpha-1} = p^{\alpha\nu-1} \leqq 1.$$

Thus for $1 \leqq \alpha \leqq n$ we have

$$|T(a, p^\alpha)| \leqq \begin{cases} n \text{ if } p \leqq n^6, \\ 1 \text{ if } p > n^6. \end{cases}$$

These last inequalities are valid also for $\alpha > n$, since, by Lemma 5, if $\alpha > n$ we have

$$T(a, p^\alpha) = p^{-\alpha(1-\nu)} p^{n-1} S(a, p^{\alpha-n}) = T(a, p^{\alpha-n}),$$

so that each $T(a, p^\alpha)$ with $\alpha > n$ is equal to some $T(a, p^\beta)$ with $\beta \leqq n$.

Multiplying together the factors $T(a, p^\alpha)$ for the various prime power factors p^α of q which do not exceed n^6, we obtain

$$|T(a,q)| \leqq n^{n^6},$$
$$|S(a,q)| \leqq n^{n^6} q^{1-\nu}.$$

LEMMA 7. *If $r \geqq 4n$, and N is any integer, the congruence*

$$x_1^n + \ldots + x_r^n \equiv N \pmod{p^\gamma}$$

has a solution in which not all of $x_1, \ldots, x_r$ are divisible by p.

Proof. It will be sufficient to prove the solubility of the congruence

$$(5) \qquad x_1^n + \ldots + x_t^n \equiv N \pmod{p^\gamma}$$

for some value of $t \leqq 4n - 1$, assuming that $(N, p) = 1$ and $0 < N < p^\gamma$. For if N is divisible by p then $N - 1$ is not, and we can solve the congruence for N by adding a term 1^n to a solution of the congruence for $N - 1$.

If $p = 2$, then $N < p^\gamma = 2^{\tau+2} \leqq 4n$. The congruence (5) will be soluble with $t = N$ since, for example, it is possible to take $x_1, \ldots, x_t$ to be all 1.

If $p > 2$, we let $t = t(N)$ be the least value of t for which the congruence (5) is soluble. Let g be a primitive root $\pmod{p^\gamma}$, and determine b by

$$N \equiv g^b \pmod{p^\gamma}.$$

Let v be the least non-negative residue of b to the modulus n, so that $0 \leqq v < n$. Suppose N_1 is another number which gives the same value of v as N. Then

$$N_1 \equiv g^{b_1} \equiv g^{b+kn} \pmod{p^\gamma},$$

so that

$$N_1 \equiv N z^n \pmod{p^\gamma}, \text{ where } z = g^k.$$

Thus (5) is equivalent to

$$(x_1 z)^n + \ldots + (x_t z)^n \equiv N_1 \pmod{p^\gamma}.$$

It follows that $t(N) = t(N_1)$.

We divide all values of N with $0 < N < p^\gamma$ and $(N, p) = 1$ into classes, placing in the same class those with the same value of $t(N)$. As we have just seen, the number m of classes satisfies $m \leqq n$. Take the least number from each such class, and denote these numbers, arranged in increasing order, by

$$N_1, \ldots, N_m.$$

Obviously $N_1 = 1$, since 1 is the least representative of its set, and is also the least positive integer. Here, since $1 = 1^n$, we have $t(N_1) = 1 = 2 - 1$. We prove by induction that $t(N_j) \leqq 2j - 1$ for $j = 1, 2, \ldots, m$. Let this inequality be true for $N_1, \ldots, N_h$. The number N_{h+1} is the least representative of its set, and moreover one of the numbers $N_{h+1} - 1$, $N_{h+1} - 2$ is not divisible by p, and consequently belongs to one of the sets with the representatives $N_1, \ldots, N_h$. Hence $t(N_{h+1}) \leqq 2h - 1 + 2 = 2(h+1) - 1$. In particular, we have $t(N_m) \leqq 2m - 1 \leqq 2n - 1 < 4n - 1$ and this proves the lemma for the case $p > 2$.

LEMMA 8. *If the congruence*

$$y^n \equiv a \pmod{p^\gamma}$$

is soluble with y not divisible by p, then for any integer $s > \gamma$ the congruence

$$x^n \equiv a \pmod{p^s}$$

is also soluble.

Proof. There is always a non-negative integer b satisfying

$$a \equiv y^n g^b \pmod{p^s}. \tag{6}$$

In particular, $g^b \equiv 1 \pmod{p^\gamma}$; therefore b is a multiple of $p^{\gamma-1}(p-1)$, and since $\gamma \geqq \tau + 1$ we can put $b = p^\tau(p-1)b_1$, where b_1 is an integer. We can replace the exponent b in the congruence (6) by the new exponent

$$b + kp^{s-1}(p-1) = p^\tau(p-1)(b_1 + kp^{s-1-\tau}),$$

where k is any integer, without changing the significance of the congruence. Let $n = p^\tau n_1$, so that $(n_1, p) = 1$. The number k can be so chosen that $b_1 + kp^{s-1-\tau}$ shall be a multiple of n_1. Then the new exponent takes the form $p^\tau n_1 h = nh$, where h is an integer, and the congruence (6) becomes

$$y^n g^{nh} \equiv a \pmod{p^s}.$$

This proves the lemma.

LEMMA 9. *Suppose that $s > \gamma$ and $r \geqq 4n$. Then*

$$M(p^s, N, r) \geqq p^{(s-\gamma)(r-1)}.$$

Proof. By Lemma 7, there exists an integer y, not divisible by p, and integers $y_2, \ldots, y_r$ such that

$$y^n \equiv N - y_2{}^n - \ldots - y_r{}^n \pmod{p^\gamma}.$$

If $x_2, \ldots, x_r$ are any numbers congruent respectively to $y_2, \ldots, y_r \pmod{p^\gamma}$, the congruence

$$x^n \equiv N - x_2{}^n - \ldots - x_r{}^n \pmod{p^s}$$

is soluble, by Lemma 8. There are $p^{s-\gamma}$ choices for each x_i to the modulus p^s, whence the result.

LEMMA 10. *Let m be any positive integer. Then*

$$\sum_{q|m} A(q, N, r) = m^{-(r-1)} M(m, N, r).$$

Proof. We have

$$mM(m) = \sum_{a=0}^{m-1} \sum_{x_1=0}^{m-1} \ldots \sum_{x_r=0}^{m-1} e\left(\frac{a}{m}(x_1{}^n + \ldots + x_r{}^n - N)\right),$$

by Lemma 5 of Chapter I. We collect together the terms for which (a, m) has the same value, and denote this value by m/q. For each particular q, these terms arise from values of a for which $a/m = b/q$ and b runs through a reduced set of residues to the modulus q. Hence,

$$\begin{aligned} mM(m) &= \sum_{q|m} \sum_{b} \sum_{x_1=0}^{m-1} \ldots \sum_{x_r=0}^{m-1} e\left(\frac{b}{q}(x_1{}^n + \ldots + x_r{}^n - N)\right) \\ &= \sum_{q|m} \sum_{b} \left(\frac{m}{q}\right)^r (S(b, q))^r e_q(-bN) \\ &= m^r \sum_{q|m} A(q, N, r). \end{aligned}$$

LEMMA 11. *For $r \geqq 2n + 1$ the two series $\mathfrak{S}(N, r)$ and $\psi(p, N, r)$, defined by* (3) *and* (4), *are absolutely convergent, and moreover*

$$\mathfrak{S}(N, r) = \prod_p \psi(p, N, r),$$

where p runs through all primes.

Proof. The absolute convergence of the series $\mathfrak{S}(N, r)$ and $\psi(p, N, r)$ follows from Lemma 6. Further, for $\xi > 2$ we have, by Lemma 2,

$$\prod_{p \leq \xi} \psi(p, N, r) = \prod_{p \leqq \xi} \left\{ \sum_{s=0}^{\infty} A(p^s, N, r) \right\} = \sum_{q \leqq \xi} A(q, N, r) + \sum_{q > \xi}{}' A(q, N, r),$$

where the last sum contains only such $A(q, N, r)$ as correspond to numbers q which are not divisible by any $p > \xi$. In view of the absolute convergence of $\mathfrak{S}(N, r)$, the last sum tends to zero as $\xi \to \infty$, and the preceding sum tends to $\mathfrak{S}(N, r)$.

LEMMA 12. *For $r \geqq 4n$ we have*

$$\mathfrak{S}(N, r) \gg 1.$$

Proof. By Lemmas 9 and 10, we have, for $s > \gamma$,

$$\sum_{q \mid p^s} A(q, N, r) \geqq p^{-s(r-1)+(s-\gamma)(r-1)} = p^{-\gamma(r-1)}.$$

Taking the limit as $s \to \infty$, we obtain

$$\psi(p, N, r) \geqq p^{-\gamma(r-1)}.$$

Further, by Lemma 6, we have

$$\psi(p, N, r) - 1 = \sum_{s=1}^{\infty} A(p^s, N, r) \ll \sum_{s=1}^{\infty} p^{s(1-r\nu)} \ll p^{-3},$$

since $1 - r\nu \leqq 1 - 4n\nu = -3$. Hence, for a sufficiently large c, we have, for any $p \geqq c$,

$$\psi(p, N, r) > 1 - p^{-2}.$$

Therefore, applying Lemma 11, we obtain

$$\begin{aligned} \mathfrak{S}(N, r) &= \prod_{p < c} \psi(p, N, r) \prod_{p \geqq c} \psi(p, N, r) \\ &\geqq \prod_{p < c} p^{-\gamma(r-1)} \prod_{p \geqq c} (1 - p^{-2}) \\ &\gg 1. \end{aligned}$$

NOTES ON CHAPTER II

The work in this chapter is comparatively straightforward, and calls for little comment. The results are all due to Hardy and Littlewood, and were published in the fourth memoir of their series "On some problems of Partitio Numerorum" [*Math. Zeitschrift*, 12 (1922), 161—188].

Hardy and Littlewood defined, for every prime p, a number γ_p as the least positive integer γ for which there exists a positive number $h = h(n, p)$ such that $\psi(p, N, r) \geqq h$ for all N and all $r \geqq \gamma$. They then defined $\Gamma(n)$ as the greatest value of γ_p for all primes p. They proved that $\mathfrak{S}(N, r)$ has a positive lower bound depending only on n provided that $r \geqq \max(\Gamma(n), 4)$, and that $\Gamma(n) \leqq 4n$.

They also proved that $G(n) \geqq \Gamma(n)$. This inequality is often more precise than the well known inequality $G(n) \geqq n + 1$ proved in the Introduction to this book.

In a later paper, the eighth of the series [*Proc. London Math. Soc.* (2), 28 (1928), 518—542], Hardy and Littlewood proved a result implicit in their earlier work, namely that $\Gamma(n)$ is (except when $n = 4$) the least number r such that every arithmetical progression contains an infinity of numbers which are sums of at most r positive integral nth powers. They also evaluated $\Gamma(n)$ for many values of n. Nevertheless the behaviour of $\Gamma(n)$ for large n is still to a considerable extent unknown.

CHAPTER III

The Contribution of the Basic Intervals in Waring's Problem

Let N_0 be a large positive integer, and let $W(N_0)$ denote the number of representations of N_0 in the form

$$N_0 = x_1^n + \ldots + x_r^n, \tag{1}$$

where $x_1, \ldots, x_r$ are positive integers. We can represent $W(N_0)$ by a definite integral, extended over an interval of length 1, as in (2) below. The object of the present chapter is to investigate the contribution to this integral of the *basic intervals*, defined below. We obtain in Lemma 4 an asymptotic expression for this contribution, which will be used in the later Chapters (IV and VII) which deal with Waring's Problem. The ideas underlying the investigation are due to Hardy and Littlewood.

Notation in this chapter. We suppose $n \geqq 3$, and we denote by r a fixed integer satisfying $r \geqq 2n + 1$.

Let N be a sufficiently large positive integer, and let N_0 be any positive integer not exceeding N.

Let $P = [N^\nu]$, where $\nu = 1/n$ as always. We observe that the positive integers x_j in any representation of N_0 in the form (1) necessarily satisfy $x_j \leqq P$. Define τ by

$$\tau = 2nP^{n-1}.$$

Let β be any fixed number satisfying

$$\tfrac{1}{4} \leqq \beta \leqq 1 - \nu.$$

By Lemma 4 of Chapter I we can express the number $W(N_0)$ of representations of N_0 in the form (1) by the following integral:

$$W(N_0) = \int_{-1/\tau}^{1-1/\tau} (L(\alpha))^r e(-N_0\alpha)\, d\alpha, \tag{2}$$

where

$$L(\alpha) = \sum_{x=1}^{P} e(\alpha x^n).$$

For any integers a, q satisfying

$$(a, q) = 1,\ 0 \leqq a < q,\ 0 < q \leqq P^{\beta},$$

we define a *basic interval* consisting of the real numbers α given by

$$\alpha = \frac{a}{q} + z, \text{ where } |z| \leqq \frac{1}{q\tau}.$$

It is easily seen that two basic intervals which correspond to different pairs of values of a and q do not overlap. For if a, q and a_1, q_1 are two different pairs, we have

$$\left|\frac{a}{q} - \frac{a_1}{q_1}\right| \geqq \frac{1}{qq_1} > \frac{1}{q\tau} + \frac{1}{q_1\tau},$$

since $\tau > P^{n-1} > P > q + q_1$.

The intervals which remain after the removal from the interval $-1/\tau \leqq \alpha \leqq 1 - 1/\tau$ of all the basic intervals will be called the *supplementary* intervals.

We denote by $W^*(N_0)$ the contribution of all the basic intervals to the integral (2) for $W(N_0)$. Thus

$$(3)\quad \begin{cases} W^*(N_0) = \sum\limits_{q \leqq P^{\beta}} \sum\limits_{a} \int_{-1/q\tau}^{1/q\tau} \left(L\left(\frac{a}{q} + z\right)\right)^r e\left(-\left(\frac{a}{q} + z\right) N_0\right) dz \\ \qquad = \sum\limits_{q \leqq P^{\beta}} \sum\limits_{a} W^*(N_0, a, q), \qquad \text{say.} \end{cases}$$

Lemma 1. *If α belongs to the basic interval corresponding to a, q, and $\alpha = \dfrac{a}{q} + z$, then*

$$(4)\quad (L(\alpha))^r = (q^{-1}S(a, q)I(z))^r + O(q^{-1}\{\min(P, |z|^{-\nu})\}^{r-1}),$$

where

$$S(a, q) = \sum_{x=1}^{q} e_q(ax^n),\ I(z) = \int_0^P e(zx^n)dx.$$

Proof. We make the change of variable $x = qt + s$ in the sum $L(\alpha)$, where s takes the values $0, \ldots, q-1$, and for given s the variable t runs through the integers of the interval

$$-sq^{-1} < t \leqq (P-s)q^{-1}. \tag{5}$$

We obtain

$$L(\alpha) = \sum_{s=0}^{q-1} \sum_t e\left(\frac{as^n}{q} + z(qt+s)^n\right) = \sum_{s=0}^{q-1} e_q(as^n) D_s(z),$$

where

$$D_s(z) = \sum_t e\big(z(qt+s)^n\big).$$

The function $f(t) = |z|(qt+s)^n$ satisfies the conditions of Lemma 13 of Chapter I in the interval (5), since $f''(t) > 0$ and

$$0 < f'(t) \leqq n|z|\, qP^{n-1} \leqq n\tau^{-1}P^{n-1} = \tfrac{1}{2}.$$

Hence, allowing for the fact that the endpoints of the interval (5) are not necessarily integers, we obtain

$$\begin{aligned} D_s(z) &= \int_{-s/q}^{(P-s)/q} e\big(z(qt+s)^n\big)dt + 4\theta \\ &= q^{-1}I(z) + 4\theta. \end{aligned}$$

Thus

$$L(\alpha) = q^{-1}S(a, q)I(z) + 4\theta' q. \tag{6}$$

By Lemma 14a of Chapter I and Lemma 6 of Chapter II we have

$$q^{-1}S(a, q)I(z) \ll q^{-\nu} \min\,(P, |z|^{-\nu}).$$

The estimate on the right here is larger than the estimate of error in (6), ignoring the factor 4, since

$$Pq^{-\nu} \geqq q^{\frac{1}{1-\nu}}q^{-\nu} \geqq q$$

and

$$|z|^{-\nu}q^{-\nu} \geqq \tau^{\nu} > P^{\nu(n-1)} = P^{1-\nu} \geqq q.$$

Hence, raising the equation (6) to the power r, we obtain

$$(L(\alpha))^r = (q^{-1}S(a, q)I(z))^r + O\big(q^{-\nu(r-1)}(\min(P, |z|^{-\nu}))^{r-1}q\big).$$

The exponent of q on the right is $-\nu(r-1)+1 \leqq -2\nu n + 1 = -1$, whence the result (4).

LEMMA 2. *The contribution $W^*(N_0, a, q)$ of a single basic interval satisfies*

(7) $$W^*(N_0, a, q) = (q^{-1}S(a, q))^r R(N_0) e_q(-aN_0) + O(q^{-1}P^{r-n-1}),$$

where

(8) $$R(N_0) = \int_{-\infty}^{\infty} (I(z))^r e(-zN_0)dz.$$

Proof. By (3) and Lemma 1, we have

$$W^*(N_0, a, q) = (q^{-1}S(a, q))^r \int_{-1/q\tau}^{1/q\tau} (I(z))^r e\left(-\left(\frac{a}{q}+z\right)N_0\right) dz$$
$$+ O\left\{q^{-1}\int_0^{\infty} (\min(P, z^{-\nu}))^{r-1}dz\right\}.$$

The last error term is

$$O\left\{q^{-1}\int_0^{P^{-n}} P^{r-1}dz + q^{-1}\int_{P^{-n}}^{\infty} z^{-r\nu+\nu}dz\right\} = O(q^{-1}P^{r-n-1}).$$

If we extend the interval of integration from $-1/q\tau \leqq z \leqq 1/q\tau$ to the whole interval from $-\infty$ to ∞, we obtain (7). The error introduced in this operation is

$$O\left(q^{-\nu r}\int_{1/q\tau}^{\infty} z^{-\nu r}dz\right)$$
$$= O(q^{-\nu r}(q\tau)^{r\nu-1}) = O(q^{-1}P^{r-n+1-\nu r}),$$

whence the result, since $1-\nu r < -1$.

LEMMA 3. *The integral $R(N_0)$, defined by (8), satisfies*

(9) $$R(N_0) = \frac{(\Gamma(1+\nu))^r}{\Gamma(r\nu)} N_0^{r\nu-1} + O(P^{r-n-\nu}),$$

provided $N_0 \geqq \frac{1}{2}N$.

Proof. Let N_1 be a positive integer less than N_0. Then

$$R(N_0) - R(N_1) = \int_{-\infty}^{\infty} (I(z))^r (e(-zN_0) - e(-zN_1))\,dz$$

$$\ll \int_0^{\infty} (\min(P, z^{-\nu}))^r (N_0 - N_1) z\,dz \ll (N_0 - N_1) P^{r-2n}.$$

We apply Lemma 2 with N_1 in place of N_0, in the special case $a = 0$, $q = 1$. This gives

$$W^*(N_1, 0, 1) = R(N_1) + O(P^{r-n-1}).$$

Using the result just proved, and assuming that $N_0 - N_1 = O(P^{n-\nu})$, we have

$$W^*(N_1, 0, 1) = R(N_0) + O(P^{r-n-\nu}).$$

Since $W^*(N_1, 0, 1)$ is the contribution of the interval $-1/\tau \leqq \alpha \leqq 1/\tau$ to the integral for $W(N_1)$, this implies

$$W(N_1) = R(N_0) + \int_{1/\tau}^{1-1/\tau} (L(\alpha))^r e(-\alpha N_1)\,d\alpha + O(P^{r-n-\nu}).$$

We take $N_1 = N_0 - N' - N''$, and let N' and N'' assume the values $1, \ldots, M$, where $M = [P^{n-\nu}]$. Summing over N' and N'', and applying Lemma 6 of Chapter I in the summation over N' and N'' in the integral on the right, we obtain

$$\sum_{N'=1}^{M} \sum_{N''=1}^{M} W(N_0 - N' - N'')$$

$$= M^2 R(N_0) + O\left(\int_{1/\tau}^{1-1/\tau} P^r \, ||\alpha||^{-2} d\alpha \right) + O(M^2 P^{r-n-\nu})$$

$$= M^2 R(N_0) + O(M^2 P^{r-n-\nu}),$$

since

$$P^r \tau = O(P^{r+n-1}) = O(M^2 P^{r-n-1+2\nu}) = O(M^2 P^{r-n-\nu}).$$

The sum over N' and N'' represents the number of representations of N_0 as $N' + N'' + x_1^n + \ldots + x_r^n$. Hence, summing over N'', the value of the sum, in the notation of Lemma 3 of Chapter I, is

$$\sum_{N'=1}^{M} \Big(K_r(N_0 - N' - 1) - K_r(N_0 - N' - M - 1) \Big).$$

We note that $N_0 - N' - M - 1 \geqq N_0 - 2M - 1 > 0$. Hence the last expression is

$$r\nu \frac{(\Gamma(1+\nu))^r}{\Gamma(1+r\nu)} M^2 N_0^{r\nu-1} + O(M^3 N_0^{r\nu-2}),$$

since $O(MN_0^{r\nu-\nu}) = O(M^3 N_0^{r\nu-2})$. Substituting this in the previous result, and dividing by M^2, we obtain (9).

LEMMA 4. *The contribution $W^*(N_0)$ of all the basic intervals to the integral* (2) *for $W(N_0)$ is given by*

$$(10)\quad W^*(N_0) = \frac{(\Gamma(1+\nu))^r}{\Gamma(r\nu)} N_0^{r\nu-1} \sum_{q \leqq P^\beta} A(q, N_0, r) + O(P^{r-n-\nu}),$$

provided $N_0 \geqq \frac{1}{2}N$.

Proof. By (3) and Lemmas 2 and 3, we have

$$W^*(N_0) = \sum_{q \leqq P^\beta} \sum_a W^*(N_0, a, q)$$

$$= \sum_{q \leqq P^\beta} \sum_a (q^{-1}S(a, q))^r R(N_0) e_q(-aN_0) + O\Big(\sum_{q \leqq P^\beta} \sum_a q^{-1} P^{r-n-1}\Big)$$

$$= \frac{(\Gamma(1+\nu))^r}{\Gamma(r\nu)} N_0^{r\nu-1} \sum_{q \leqq P^\beta} A(q, N_0, r)$$

$$+ O\Big(\sum_{q \leqq P^\beta} \sum_a (q^{-\nu r} P^{r-n-\nu} + q^{-1} P^{r-n-1})\Big).$$

For each q there are at most q values for a, and the result follows on noting that

$$\sum_{q \leqq P^\beta} q^{1-\nu r} P^{r-n-\nu} = O(P^{r-n-\nu}),$$

$$\sum_{q \leqq P^\beta} P^{r-n-1} = O(P^{r-n-1+\beta}) = O(P^{r-n-\nu}).$$

NOTES ON CHAPTER III

This chapter has been rewritten, mainly in order to separate the different arguments used and so assist those readers to whom they may be unfamiliar.

The basic intervals of Vinogradov correspond in principle to Hardy and Littlewood's "major arcs". Hardy and Littlewood dissected the circle $x = e^{2\pi i\alpha}$, or rather a smaller concentric circle, into "Farey arcs" as follows. Consider all fractions a/q with $(a, q) = 1$, $0 \leqq a < q$, $0 < q \leqq \tau$, arranged in increasing order. If a_1/q_1 and a_2/q_2 are the neighbouring fractions to a/q on the left and on the right, the Farey arc surrounding a/q is defined by

$$\frac{a + a_1}{q + q_1} \leqq \alpha \leqq \frac{a + a_2}{q + q_2}.$$

It is easily proved that the length of each part of this interval, to the left and to the right of a/q, lies between $\frac{1}{2}(q\tau)^{-1}$ and $(q\tau)^{-1}$ (see, for example, Hardy and Wright, § 3.8).

Hardy and Littlewood took τ to be P^{n-1}, which is practically the same as Vinogradov's definition. They called a Farey arc a major or minor arc according as it arises from a fraction a/q with $q \leqq P$ or with $q > P$. The classification was varied in different ways by later workers on Waring's Problem, and the exact demarcation is not usually of vital importance.

The major arcs, or basic intervals, provide the main term in the asymptotic formula for the number of representations. Their treatment does not give rise to any very serious difficulties compared with the problems presented by the minor arcs, or supplementary intervals. The asymptotic formula (10) for the contribution of the basic intervals is valid for $r \geqq 2n + 1$, as we have seen in this chapter, and a better error term could be found if anything were to be gained by doing so.

The device used in the proof of Lemma 3 in order to deal with the integral $R(N_0)$ in (8) is by no means essential, but enables one to avoid some rather elaborate calculations. For a direct treatment, see Landau, *Math. Zeitschrift*, 31 (1929), 319—338.

CHAPTER IV

An Estimate for G(n) in Waring's Problem

The object of the present chapter is to establish a certain estimate for the number $G(n)$ in Waring's Problem. As explained in the Introduction, Waring's Problem is the problem of determining, for each positive integer $n \geqq 3$, a number r such that every positive integer N can be represented in the form

$$N = x_1^n + \ldots + x_r^n \tag{1}$$

with non-negative integers $x_1, \ldots, x_r$. The more fundamental problem is that of representing not *every* N but *every sufficiently large* N, that is, every N exceeding some number depending on n. We denote by $G(n)$ the least value of r with the property that every sufficiently large positive integer is representable in the form (1). The result of this chapter will be that

$$G(n) < n(3 \log n + 11). \tag{2}$$

The method used for the proof is one that can be applied to more general questions, for example that of representing a large integer N by a sum of values of a polynomial. But we shall not consider such questions here.

The number of representations of N in the form (1) can be expressed by a definite integral, as in Hardy and Littlewood's solution of Waring's Problem. The essential idea for the proof of (2) lies in the introduction into this integral of two factors, $S(\alpha)$ and $Q(\alpha)$, which enable one to estimate the contribution of the supplementary intervals more effectively than was hitherto possible. The construction of these factors is carried out by imposing restrictions of a special kind on some of the x's in a representation such as (1).

Notation in this chapter. We suppose $n \geqq 3$, and retain most of the notation of Chapter III. Thus N is arbitrarily large, $P = [N^{\nu}]$, and $\tau = 2nP^{n-1}$. We retain the same division of the interval $-1/\tau \leqq \alpha \leqq 1 - 1/\tau$ into basic and supplementary intervals, but specialize the number β determining the classification by taking $\beta = \frac{1}{4}$. We take $r = 4n$, so that we can appeal to the results of Chapter II.

LEMMA 1. *Let*

$$(3) \qquad P_1 = [\tfrac{1}{4}P],\ P_2 = [\tfrac{1}{2}P_1^{1-\nu}], \ldots,\ P_k = [\tfrac{1}{2}P_{k-1}^{1-\nu}].$$

Let ξ_s, for $s = 1, \ldots, k$, run through the integers of the interval

$$P_s \leqq \xi_s < 2P_s.$$

Then, for fixed k and sufficiently large P, the numbers

$$u = \xi_1^n + \ldots + \xi_k^n$$

are all distinct, and all lie between $(\frac{1}{5}P)^n$ and $(\frac{1}{2}P)^n$.

Proof. We note first that the definition (3) ensures that $(2P_{s+1})^n \leq P_s^{n-1}$. Plainly $u \geqq \xi_1^n \geqq P_1^n > (\frac{1}{5}P)^n$. Also

$$\begin{aligned} u &\leq (2P_1 - 1)^n + (2P_2 - 1)^n + \ldots + (2P_k - 1)^n \\ &\leqq (2P_1 - 1)^n + P_1^{n-1} + \ldots + P_{k-1}^{n-1} \\ &< (2P_1 - 1)^n + 2P_1^{n-1} < (2P_1)^n \leqq (\tfrac{1}{2}P)^n, \end{aligned}$$

provided P is sufficiently large.

Now suppose that

$$\xi_1^n + \xi_2^n + \ldots + \xi_k^n = \eta_1^n + \eta_2^n + \ldots + \eta_k^n,$$

where the η's satisfy the same inequalities as the ξ's. Suppose $\xi_1 \neq \eta_1$, say $\eta_1 > \xi_1$. Then

$$\eta_1^n - \xi_1^n < \xi_2^n + \ldots + \xi_k^n.$$

But

$$\eta_1^n - \xi_1^n > n\xi_1^{n-1} \geqq nP_1^{n-1},$$

and, by the argument used above,

$$\xi_2^n + \ldots + \xi_k^n < (2P_2)^n \leqq P_1^{n-1}.$$

This contradiction proves that $\xi_1 = \eta_1$. Obviously a similar argument proves that $\xi_2 = \eta_2$, and so on.

Further notation. Let u run through the integers defined in Lemma 1, and let

$$S(\alpha) = \sum_u e(\alpha u). \tag{4}$$

The number U of these integers u is given by

$$U = P_1 P_2 \dots P_k,$$

and obviously U is a trivial upper bound for $S(\alpha)$. We have

$$U \gg P^{1+(1-\nu)+\cdots+(1-\nu)^{k-1}}.$$

The exponent here is

$$\frac{1-(1-\nu)^k}{1-(1-\nu)} = n(1-(1-\nu)^k).$$

We choose k to be the least integer for which

$$n(1-\nu)^k < \tfrac{1}{12},$$

so that

$$U \gg P^{n-\frac{1}{12}}. \tag{5}$$

The explicit value of k is

$$k = \left[\frac{\log 12n}{-\log(1-\nu)} + 1\right]. \tag{6}$$

Let $P_0 = [P^{\frac{1}{2}}]$. Define a set of integers u_0 in the same way as the integers u, but with P_0 in place of P and k_0 in place of k. These integers all lie between $(\frac{1}{5}P_0)^n$ and $(\frac{1}{2}P_0)^n$, and their number U_0 satisfies

$$U_0 \gg P_0^{\,n(1-(1-\nu)^{k_0})}.$$

We choose k_0 to be the least integer for which

$$n(1-\nu)^{k_0-1} < \tfrac{1}{6},$$

and have

$$U_0 \gg P^{\frac{1}{2}n-\frac{1}{12}(1-\nu)}. \tag{7}$$

The explicit value of k_0 is

$$(8) \qquad k_0 = \left[\frac{\log 6n}{-\log(1-\nu)} + 2\right].$$

Let v run through the primes in the interval $\frac{1}{2}P_0 \leqq v \leqq P_0$, and let V be the number of these primes. We have

$$(9) \qquad V \gg P_0/\log P_0 \gg P^{\frac{1}{2}-\varepsilon}.$$

Define $Q(\alpha)$ by

$$(10) \qquad Q(\alpha) = \sum_v \sum_{u_0} e(\alpha v^n u_0).$$

The trivial upper bound for $Q(\alpha)$ is VU_0.

LEMMA 2. *If α belongs to any supplementary interval, then*

$$Q(\alpha) \ll U_0^{\frac{1}{2}} P^{\frac{1}{4}n+\frac{3}{8}+\varepsilon}.$$

Proof. By Lemma 7 of Chapter I, each α can be represented as

$$\alpha = \frac{a}{q} + z, \quad \text{where} \quad (a, q) = 1, \quad 0 < q \leqq P_0{}^n, \quad |z| < q^{-1}P_0{}^{-n}.$$

The fact that α does not belong to any basic interval means that if $q \leqq P^{\frac{1}{4}}$ then $|z| \geqq q^{-1}\tau^{-1}$.

Case 1. *Suppose that $q \leqq P^{\frac{1}{4}}$.* We have

$$(11) \qquad \tau^{-1} \leqq q\,|z| < P_0{}^{-n}.$$

Put $v = qt + s$, where t, s are integers and $0 \leqq s < q$. We have

$$e(\alpha v^n u_0) = e\left\{u_0\left(\frac{as^n}{q} + z(qt+s)^n\right)\right\} = e\big(u_0\Phi_s(t)\big), \text{ say.}$$

Define $\xi(x)$ to be 1 if x is one of the numbers u_0, and 0 otherwise. Define $\eta_s(t)$ to be 1 if $qt + s$ is one of the numbers v, and 0 otherwise. Then

$$Q(\alpha) = \sum_{s=0}^{q-1} \sum_x \sum_t \xi(x)\eta_s(t)\, e\big(x\Phi_s(t)\big),$$

where the summations are over $(\frac{1}{5}P_0)^n < x < (\frac{1}{2}P_0)^n$, $(\frac{1}{2}P_0 - s)q^{-1} \leqq t \leqq (P_0 - s)q^{-1}$.

We apply Lemma 10c of Chapter I to the sum over x and t. We have

$$\Phi_s(t+1) - \Phi_s(t) = nqz(qt + q\theta + s)^{n-1},$$

and the absolute value of this lies between $1/A$ and $2^{n-1}/A$, where $1/A = nq\,|z|\,(\frac{1}{2}P_0)^{n-1}$. By (11), noting that $\tau \ll P^{n-1} \ll P_0^{2(n-1)}$, we have $P_0 \ll A \ll P_0^{n-1}$. The hypothesis $A \geqq 2\beta \geqq 2$ of the lemma is satisfied, since $\beta = 2^{n-1}$ here. In the notation of the lemma,

$$X < P_0^n,\ Y \leqq P_0 q^{-1},\ X_0 = U_0,\ \eta = 1,\ Y_1 \leqq Y.$$

In the conclusion, we can omit the term $Y\beta A^{-1}$, since $Y\beta A^{-1} \ll (P_0 q^{-1})P_0^{-1} \ll 1$. Hence, for each s,

$$\sum_x \sum_t \xi(x)\eta_s(t)e(x\Phi_s(t)) \ll (X_0 Y_1(X + A))^{\frac{1}{2}}$$
$$\ll (U_0 P_0 q^{-1} P_0^n)^{\frac{1}{2}}.$$

Thus

$$Q(\alpha) \ll q(U_0 q^{-1} P_0^{n+1})^{\frac{1}{2}} \ll (U_0 P^{\frac{1}{2}n+\frac{3}{4}})^{\frac{1}{2}}.$$

Case 2. *Suppose that* $P^{\frac{1}{4}} < q \leqq P_0^n$. If $v^n \equiv y \pmod q$, we have

$$e(\alpha u_0 v^n) = e\left(u_0 \frac{ay + \psi(v)}{q}\right),$$

where

$$\psi(v) = qzv^n \ll q(qP_0^n)^{-1}P_0^n = 1.$$

Let $\varrho(y)$ denote the number of primes v with $\frac{1}{2}P_0 \leqq v \leqq P_0$ which satisfy $v^n \equiv y \pmod q$, and enumerate these primes (if there are any) as $v_1(y), \ldots, v_{\varrho(y)}(y)$. Since v is restricted to prime values, we have $(v, q) = 1$ except when v is a factor of q. The number of solutions of $x^n \equiv y \pmod q$, $0 \leqq x < q$ is $O(q^\varepsilon)$ when $(y, q) = 1$, and the number of factors of q is $O(q^\varepsilon)$. Hence

$$\varrho(y) \ll (P_0 q^{-1} + 1)q^\varepsilon.$$

Let ϱ be the greatest value of $\varrho(y)$, and define $\eta_j(y)$ to be 1 if $j \leqq \varrho(y)$ and 0 otherwise. Then

$$Q(\alpha) = \sum_x \sum_{y=0}^{q-1} \sum_{j=1}^{\varrho(y)} \xi(x) e\left(\frac{x}{q}\left(ay + \psi(v_j(y))\right)\right)$$
$$= \sum_{j=1}^{\varrho} \sum_x \sum_{y=0}^{q-1} \xi(x)\eta_j(y) e\left(\frac{x}{q}(ay + \psi_j(y))\right),$$

where $\psi_j(y) = \psi(v_j(y))$. The meaning of $\xi(x)$ and the interval of summation for x are the same as in Case 1.

We apply Lemma 10*b* of Chapter I to the sum over x and y. In the notation of that lemma, we have

$$X < P_0^n,\ Y = q,\ \lambda \ll 1,\ X_0 = U_0,\ \eta = 1.$$

Also

$$Y_1 = \sum_{y=0}^{q-1} \eta_j(y) \leqq \min(P_0, q).$$

Hence, for each j,

$$\sum_x \sum_y \xi(x)\eta_j(y) e\left(\frac{x}{q}(ay + \psi_j(y))\right) \ll (U_0 \min(P_0, q) P_0^n)^{\frac{1}{2}}.$$

Thus

$$Q(\alpha) \ll \varrho(U_0 \min(P_0, q)P_0^n)^{\frac{1}{2}}$$
$$\ll (P_0 q^{-1} + 1) q^{\varepsilon}(U_0 \min(P_0, q)P_0^n)^{\frac{1}{2}}$$
$$\ll q^{\varepsilon}(U_0 P_0^n(P_0^2 q^{-1} + P_0))^{\frac{1}{2}} \ll P_0^{n\varepsilon}(U_0 P_0^{n+\frac{3}{2}})^{\frac{1}{2}}.$$

THEOREM. $G(n) < n(3 \log n + 11)$.

Proof. Let $I(N)$ denote the number of representations of N as

$$x_1^n + \ldots + x_r^n + u + u' + v^n u_0,$$

where u' satisfies the same conditions as u. Then $I(N)$ is the number of representations of N as a sum of $r + 2k + k_0$ nth powers of positive integers satisfying certain conditions. We shall prove that $I(N) > 0$ for all sufficiently large N, from which it will follow that

$$G(n) \leqq r + 2k + k_0.$$

Here $r = 4n$, and k and k_0 are given by (6) and (8). On noting that

$$-\log(1-\nu) = \log\frac{(n-\frac{1}{2})+\frac{1}{2}}{(n-\frac{1}{2})-\frac{1}{2}} > \frac{1}{n-\frac{1}{2}},$$

we obtain

$$\begin{aligned} G(n) &< 4n + 4 + (n-\tfrac{1}{2})(2\log 12n + \log 6n) \\ &= 4n + 4 + (n-\tfrac{1}{2})(3\log n + 6.7\ldots) \\ &< n(3\log n + 11), \end{aligned}$$

since $2\log 12 + \log 6 = 6.7\ldots$

For the proof that $I(N) > 0$ we start from the expression

$$I(N) = \int_{-1/\tau}^{1-1/\tau} (L(\alpha))^r Q(\alpha)\,(S(\alpha))^2 e(-N\alpha)\,d\alpha.$$

Consider first the contribution $I^{**}(N)$ which the supplementary intervals make to this integral. By Lemma 2 and the trivial estimate $|L(\alpha)| \leqq P$, we have

$$I^{**}(N) \ll P^r U_0^{\frac{1}{2}} P^{\frac{1}{4}n+\frac{3}{8}+\varepsilon} \int_0^1 |S(\alpha)|^2\,d\alpha.$$

By Lemma 1 the numbers u in the definition of $S(\alpha)$ are distinct; hence

$$\int_0^1 |S(\alpha)|^2\,d\alpha = U.$$

Thus

$$I^{**}(N) \ll UU_0^{\frac{1}{2}} P^{r+\frac{1}{4}n+\frac{3}{8}+\varepsilon}.$$

This implies

(12) $$I^{**}(N) \ll VU_0U^2P^{r-n-\frac{1}{25}\nu},$$

since

$$VU_0^{\frac{1}{2}}U \gg P^{\frac{5}{4}n+\frac{3}{8}+\frac{1}{24}\nu-\varepsilon}$$

by (5), (7) and (9).

Now consider the contribution $I^*(N)$ which the basic intervals make to the integral for $I(N)$. This can be expressed as

$$I^*(N) = \sum_{N_0} W^*(N_0),$$

where $N_0 = N - v^n u_0 - u - u'$, and summation is over all the values of v, u_0, u, u'. Here $W^*(N_0)$ has the meaning of Chapter III, and by Lemma 4 of that Chapter we have

$$I^*(N) = C_n \sum_{N_0} N_0^{rv-1} \sum_{0<q\leqq P^{\frac{1}{4}}} A(q, N_0, r) + O(VU_0U^2P^{r-n-v}),$$

where C_n is a positive number depending only on n. Further, by Lemma 6 of Chapter II, on noting that $1 - rv = -3$, we have

$$\sum_{q>P^{\frac{1}{4}}} A(q, N_0, r) \ll \sum_{q>P^{\frac{1}{4}}} q^{-3} \ll P^{-\frac{1}{2}}.$$

Hence

$$(13) \qquad I^*(N) = C_n \sum_{N_0} N_0^{rv-1} \mathfrak{S}(N_0, r) + O(VU_0U^2P^{r-n-v}).$$

The number of numbers N_0 is VU_0U^2, and each of them is

$$> N - P_0^n(\tfrac{1}{2}P_0)^n - (\tfrac{1}{2}P)^n - (\tfrac{1}{2}P)^n \gg P^n.$$

Further, by Lemma 12 of Chapter II we have $\mathfrak{S}(N_0, r) \gg 1$. Hence the first term on the right of (13) is $\gg VU_0U^2P^{r-n}$. It now follows from (12) and (13) that $I(N) > 0$ for sufficiently large N.

NOTES ON CHAPTER IV

This chapter has been expanded somewhat compared with the original. The expansion was most necessary in the proof of Case 2 of Lemma 2. The original proof is exceedingly concise, and is formulated without the variable j, as if $\psi(v)$ were a function of y.

The reader who is not familiar with work on Waring's Problem may find the following considerations helpful. We are representing a large number N as a sum of nth powers, some of which are restricted in various ways. The number of sets of values over which all the variables in the representation can range is $P^rVU_0U^2$. It follows from the work of Chapter III that the contribution of the basic intervals to the integral for the number of representations is expressible by an asymptotic formula if $r \geqq 4n$; and for this it is immaterial what restrictions are imposed on the variables other than $x_1, \ldots, x_r$. The main term in the asymptotic expression is of the order $VU_0U^2P^{r-n}$. Hence, in order that the method shall succeed, it is necessary to have an estimate for the contribution of the supplementary intervals which "saves" more than P^n in comparison with the trivial estimate. In the present treatment, this saving comes from two sources. In the first place, the use of the identity $\int_0^1 |S(\alpha)|^2 d\alpha = U$ saves an amount U, which with the present choice of k is about $P^{n-\frac{1}{12}}$. This accounts for all but $P^{\frac{1}{12}}$ of what is needed, but nevertheless a new idea is required to save this small additional amount without introducing too many further n th powers. The additional saving is effected by the use of $Q(\alpha)$. The estimate obtained for $Q(\alpha)$ in Lemma 2 is essentially

$$P^{\frac{1}{4}n-\frac{1}{24}(1-\nu)+\frac{1}{4}n+\frac{3}{8}},$$

as compared with the trivial estimate VU_0, which is about

$$P^{\frac{1}{2}+\frac{1}{2}n-\frac{1}{12}(1-\nu)}.$$

Thus a saving is effected of about $P^{\frac{1}{12}+\frac{1}{24}\nu}$.

There was an earlier form of the method, given by Vinogradov in 1934, which led to the result $G(n) < 6n \log n + O(n)$. This already contained two factors corresponding to $S(\alpha)$ and $Q(\alpha)$, but the treatment of $Q(\alpha)$ was not quite so effective.

For results on $G(n)$ for individual small values of n, see G. L. Watson, *J. London Math. Soc.*, 26 (1951), 153–156 [a simple proof that $G(3) \leqq 7$]; H. Davenport, *Annals of Math.*, 40 (1939), 731—747 [proof that $G(4) = 16$ and that a modified number $G^*(n)$ satisfies $G^*(4) \leqq 14$]; H. Davenport, *American J. of Math.*, 64 (1942), 199—207 [proof that $G(5) \leqq 23$ and $G(6) \leqq 36$].

Lemma 1, which is the essential basis for the use of $S(\alpha)$, is due to Hardy and Littlewood. It is practically the same as the special case of Lemma 15 of Chapter I which was discussed in the Notes on that Chapter. However, Hardy and Littlewood did not themselves use the exponential sum $S(\alpha)$. They used the lemma in their estimation of $G_1(n)$, which is defined as the least r for which the numbers not representable by r nth powers have zero density. This enabled them to improve their previous estimate for $G(n)$, but the improvement was not significant for large n.

CHAPTER V

Approximation by the Fractional Parts of the Values of a Polynomial

Let

$$f(x) = a_h x^h + \dots + a_j x^j + a_n x^n$$

be a polynomial with real coefficients, where $h, \dots, j, n$ are positive integers in increasing order. In the present chapter my methods are applied to the problem of finding integers z such that the fractional part of $f(z)$ approximates to a given real number A. The result will be expressed in terms of a rational approximation a/q to a particular coefficient a_l of $f(x)$. We shall determine an exponent $\varrho = \varrho(h, \dots, n)$ such that, provided q is sufficiently large, there always exist integers z and v satisfying

$$| f(z) - v - A | < q^{-\varrho} \text{ and } 0 < z < q^{2/l}.$$

It would be possible to treat in a similar way certain more general questions, for example that which arises when $f(x)$ is replaced by a function which is not a polynomial but which is in certain respects similar to a polynomial.

Notation in this chapter. We denote by g the number of the exponents $h, \dots, j, n$ and by D their sum, so that

$$D = h + \dots + j + n.$$

We shall suppose that $n \geqq 5$. We put $\nu = 1/n$ as always, and

$$\lambda = 1/l,$$

where l (as already stated) is a particular one of $h, \dots, n$.

Suppose that

$$a_l = \frac{a}{q} + \frac{\theta}{q^2}, \text{ where } (a, q) = 1 \text{ and } q > c_0(n),$$

where $c_0(n)$ will be determined later.

Let $c_1, \ldots, c_g$ be constants satisfying

$$0 < c_1 < \ldots < c_g < \tfrac{1}{2}, \qquad \begin{vmatrix} c_1^{h-1}, & \ldots, & c_g^{h-1} \\ & \ldots & \\ c_1^{n-1}, & \ldots, & c_g^{n-1} \end{vmatrix} \neq 0.$$

Let

(1) $$p = [q^\lambda] \text{ and } p_t = [p^{(1-\nu)^{t-1}}] \text{ for } t = 1, \ldots, k,$$

where k will be chosen later and is bounded in terms of n. Let

(2) $$X_{t,s} = [p_t c_s], \qquad \xi_t = [p_t^{1-\varepsilon}]$$

for $t = 1, \ldots, k$ and $s = 1, \ldots, g$.

LEMMA 1. *Let the numbers $m_{t,s}$, where $t = 1, \ldots, k$ and $s = 1, \ldots, g$ be integers satisfying $0 < m_{t,s} \leqq M$, where M is a positive integer. Let $x_{t,s}$ denote a variable which assumes all integral values in the interval*

(3) $$X_{t,s} < x_{t,s} \leqq X_{t,s} + \xi_t.$$

Let $U_{t,r}$, for $t = 1, \ldots, k$ and $r = h, \ldots, j, n$, denote the functions of these variables defined by

(4) $$U_{t,r} = m_{t,1} x_{t,1}^r + \ldots + m_{t,g} x_{t,g}^r.$$

Let U_r, for $r = h, \ldots, j, n$ be defined by

(5) $$U_r = U_{1,r} + \ldots + U_{k,r}.$$

For any particular values of $m_{1,1}, \ldots, m_{k,g}$, let $\psi(z_h, \ldots, z_n)$ denote the number of sets of values of the x's for which

(6) $$U_h = z_h, \ldots, U_n = z_n.$$

Then if q (and therefore p) is sufficiently large, we have

(7) $$\psi(z_h, \ldots, z_n) \ll \frac{M^{gk}(p_1 \ldots p_k)^{g-\nu D}}{m_{1,1} \ldots m_{k,g}} = \Phi, \text{ say.}$$

Proof. We obtain the result by an appeal to Lemma 15 of Chapter I. The hypotheses of that lemma are satisfied, and it follows that

$$\psi(z_h, \ldots, z_n) \ll \Phi_1 \ldots \Phi_k, \tag{8}$$

where Φ_t (for $t = 1, \ldots, k$) denotes an upper bound for the number of sets of values of the x's for which

$$U_{t,h}, \ldots, U_{t,n} \tag{9}$$

lie in any given intervals whose lengths are respectively

$$Mp_t^{h(1-\nu)}, \ldots, Mp_t^{n(1-\nu)}. \tag{10}$$

It remains only to estimate Φ_t. We do this by first estimating the number Φ_t' of sets of values of the x's for which the numbers (9) lie in any given intervals whose lengths are respectively

$$Mp_t^{h-1}, \ldots, Mp_t^{n-1}. \tag{11}$$

These intervals are shorter than the corresponding intervals just mentioned, except for the last, which is of the same length.

Since we are now concerned only with a single value of t we can omit the suffix t from the variables x and m for the time being. Let $x_1, \ldots, x_g$ and $x_1 + \zeta_1, \ldots, x_g + \zeta_g$ be two sets of the variables for which the numbers (9) lie in intervals of the lengths (11). Then

$$m_1((x_1 + \zeta_1)^h - x_1^h) + \ldots + m_g((x_g + \zeta_g)^h - x_g^h) = \theta_h M p_t^{h-1},$$
$$\ldots$$
$$m_1((x_1 + \zeta_1)^n - x_1^n) + \ldots + m_g((x_g + \zeta_g)^n - x_g^n) = \theta_n M p_t^{n-1}.$$

Here

$$\begin{aligned}(x_s + \zeta_s)^h - x_s^h &= h\zeta_s(x_s + \theta\zeta_s)^{h-1}\\ &= h\zeta_s(p_t c_s)^{h-1}\beta_{h,s}, \text{ say,}\end{aligned}$$

and similarly with the other exponents in place of h, where, in view of (1), (2) and (3), the numbers $\beta_{h,1}, \ldots, \beta_{n,g}$ will be arbitrarily near to 1 if p is sufficiently large. The system of equations now takes the form

$$m_1 c_1^{h-1}\beta_{h,1}\zeta_1 + \ldots + m_g c_g^{h-1}\beta_{h,g}\zeta_g = \theta_h M h^{-1},$$
$$\ldots$$
$$m_1 c_1^{n-1}\beta_{n,1}\zeta_1 + \ldots + m_g c_g^{n-1}\beta_{n,g}\zeta_g = \theta_n M n^{-1}.$$

Regarding these as g linear equations for the g unknowns $m_1\zeta_1, \ldots, m_g\zeta_g$, we obtain

$$m_s\zeta_s \ll M \text{ for } s = 1, \ldots, g.$$

Hence the number of possibilities for $\zeta_1, \ldots, \zeta_g$ is

$$\ll \frac{M^g}{m_1 \ldots m_g}.$$

Restoring the suffix t, we therefore have

$$\Phi_t' \ll \frac{M^g}{m_{t,1} \ldots m_{t,g}}. \tag{12}$$

To deduce an estimate for Φ_t we have merely to split up the intervals of the lengths (10) into intervals whose lengths do not exceed those specified in (11). We obtain

$$\Phi_t \ll p_t^{1-h\nu} \ldots p_t^{1-n\nu}\Phi_t' \ll p_t^{g-D\nu}\Phi_t'. \tag{13}$$

The result (7) now follows from (8), (12) and (13).

LEMMA 2. *Let $T = T(m_{1,1}, \ldots m_{k,g})$ denote the sum*

$$T = \sum_{y=1}^{p} \left| \sum\nolimits_0 e(a_h y^h U_h + \ldots + a_n y^n U_n) \right|, \tag{14}$$

where $\sum_0$ denotes a sum extended over all integers $x_{1,1}, \ldots, x_{k,g}$ satisfying (3), and $U_h, \ldots, U_n$ are defined by (4) and (5). Then

$$T \ll p\Xi\left(\frac{M^{g(k+1)}(p_1 \ldots p_k)^{-D\nu+g\varepsilon}p^{D-1}}{m_{1,1} \ldots m_{k,g}}\right)^{\frac{1}{4}}, \tag{15}$$

where

$$\Xi = (\xi_1 \ldots \xi_k)^g. \tag{16}$$

Proof. We note first that Ξ is the number of possibilities for all the variables $x_{1,1}, \ldots, x_{k,g}$ by (3); so that $p\Xi$ is the trivial upper bound for T.

It follows from (14), by Cauchy's inequality, that

$$T^2 \leq p \sum_{y=1}^{p} \sum\nolimits_0' \sum\nolimits_0 e\Big(a_h y^h(U_h' - U_h) + \ldots + a_n y^n(U_n' - U_n)\Big),$$

where Σ_0' denotes a summation similar to Σ_0 extended over variables x', and $U_h', \ldots, U_n'$ are similarly defined in terms of these variables. We can write the last inequality as

(17) $$T^2 \leqq p \sum_{y=1}^{p} \sum_{u_h} \ldots \sum_{u_n} \chi(u_h, \ldots, u_n) e(a_h y^h u_h + \ldots + a_n y^n u_n),$$

where $\chi(u_h, \ldots, u_n)$ denotes the number of sets x and x' for which

$$U_h' - U_h = u_h, \ldots, U_n' - U_n = u_n.$$

In terms of the function $\psi(z_h, \ldots, z_n)$ of Lemma 1, we have

$$\chi(u_h, \ldots, u_n) = \sum_{z_h} \ldots \sum_{z_n} \psi(z_h, \ldots, z_n)\psi(z_h + u_h, \ldots, z_n + u_n).$$

Obviously

$$\sum_{z_h} \ldots \sum_{z_n} \psi(z_h, \ldots, z_n) = \Xi;$$

hence, by Lemma 1,

$$\chi(u_h, \ldots, u_n) \ll \Xi\Phi.$$

Also

$$\sum_{u_h} \ldots \sum_{u_n} \chi(u_h, \ldots, u_n) = \Xi^2,$$

hence

(18) $$\sum_{u_h} \ldots \sum_{u_n} \chi^2(u_h, \ldots, u_n) \ll \Xi^3\Phi.$$

We note also that since $U_h \ll Mp^h, \ldots, U_n \ll Mp^n$ by (2), (3), (4), (5), the function $\chi(u_h, \ldots, u_n)$ must vanish outside a region of the form

(19) $$|u_h| \leqq c^{(h)}Mp^h, \ldots, |u_n| \leqq c^{(n)}Mp^n.$$

Applying Cauchy's inequality to (17), and using (18), we obtain

$$\begin{aligned} T^4 &\ll p^2\Xi^3\Phi \sum_{u_h} \ldots \sum_{u_n} \left| \sum_{y=1}^{p} e(a_h y^h u_h + \ldots + a_n y^n u_n) \right|^2 \\ &= p^2\Xi^3\Phi \sum_{u_h} \ldots \sum_{u_n} |S(u_h, \ldots, u_n)|^2, \end{aligned}$$

say, where the summation for $u_h, \ldots, u_n$ is over the region (19).

We now use a device which was first introduced in the proof of Lemma 10b of Chapter I. Let u_l' be a variable running through the integers of the interval $|u_l'| \leqq 2c^{(l)}Mp^l$. For any integer u in the same interval of (19) as u_l, the number of solutions of $u_l' + u_l = u$ is at least $c^{(l)}Mp^l$. Hence

$$T^4 \ll \frac{p^2\Xi^3\Phi}{Mp^l}\sum_{u_h}\ldots\sum_{u_l'}\sum_{u_l}\ldots\sum_{u_n}\left|S(u_h, \ldots, u_l' + u_l, \ldots, u_n)\right|^2.$$

We now interchange the order of summation. Let Σ_1 denote summation over all the variables $u_h, \ldots, u_n$ other than u_l. Such a summation comprises $\ll M^{g-1}p^{D-l}$ terms. Then

$$T^4 \ll \frac{p^2\Xi^3\Phi}{Mp^l}\sum_{y_1=1}^{p}\sum_{y=1}^{p}\Sigma_1\left|\sum_{u_l}\sum_{u_l'} e\left(a_l(y_1{}^l - y^l)(u_l' + u_l)\right)\right|$$

$$\ll \frac{p^2\Xi^3\Phi}{Mp^l}M^{g-1}p^{D-l}\sum_{y_1=1}^{p}\sum_{y=1}^{p}\left|\sum_{u_l} e(a_l(y_1{}^l - y^l)u_l)\right|^2$$

$$\ll \Xi^3\Phi M^{g-2}p^{D-2l+2}\sum_{y_1=1}^{p}\sum_{y=1}^{p}\min\left\{M^2p^{2l}, \frac{1}{4\,\|a_l(y_1{}^l - y^l)\|^2}\right\},$$

by Lemma 6 of Chapter I.

For a particular y_1, the difference $y_1{}^l - y^l$ runs through certain integers in an interval of length $p^l \leqq q$. Hence we can replace the sum over y_1 and y by

$$p\sum_{z}\min\left\{M^2p^{2l}, \frac{1}{4\,\|a_lz\|^2}\right\},$$

where z runs through an interval of length q. Here

$$a_lz = \frac{az + \theta q^{-1}z}{q}.$$

We can now apply Lemma 8a (I) of Chapter I, with

$$q' = q,\ \lambda = 1,\ U = Mp^l.$$

It follows that the above sum over z is

$$\ll M^2p^{2l} + Mp^lq \ll M^2p^{2l}.$$

Finally, we obtain

$$T^4 \ll \Xi^3 \Phi M^{g-2} p^{D-2l+2} p M^2 p^{2l},$$

$$T \ll \Xi p \left(\frac{\Phi M^g p^D}{p\Xi}\right)^{\frac{1}{4}}.$$

Substituting for Φ from (7) and noting that

$$\Xi \gg (p_1 \ldots p_k)^{g(1-\varepsilon)}$$

by (16) and (2), we obtain the result stated.

THEOREM. *There exists $c_0(n)$ with the following property. Let $f(x)$ be a polynomial of the form under consideration, and let a/q be a rational approximation*[1] *to a_l with $q > c_0(n)$. Then, for any real A, there exist integers z, v satisfying*

(20) $$| f(z) - v - A | < q^{-\varrho}, \qquad 0 < z < q^{2/l},$$

where

(21) $$\varrho = \frac{\log D}{4ngl (\log D + 1) \log (D \log D + D)}.$$

Proof. We take ϱ to be any fixed positive number, and have to prove the result when ϱ has the value given in (21). Let $\Delta = q^{-\varrho}$, and define α, β by $\alpha + \frac{1}{2}\Delta = \beta - \frac{1}{2}\Delta = A$. Let $\psi(z)$ be the function of Lemma 12 of Chapter I, where the positive integer r, which we shall replace by R to avoid confusion, will be fixed later. Then $\psi(z) = 0$ except when z is congruent (mod 1) to a number in the interval $A - \Delta \leqq z \leqq A + \Delta$. If we can prove that for sufficiently large q there exists a positive integer $z < q^{2/l}$ with $\psi(f(z)) > 0$, the conclusion will follow.

By the lemma just referred to, we have

$$\psi(z) = \Delta + \psi_0(z),$$

where

$$\psi_0(z) = \sum_{m=1}^{\infty} (A_m \cos 2\pi mz + B_m \sin 2\pi mz),$$

and the coefficients A_m, B_m satisfy

$$A_m \ll F(m), \ B_m \ll F(m),$$

where

[1] in the sense defined at the beginning of the chapter.

$$F(m) = \begin{cases} \Delta \text{ if } m \leqq \Delta^{-1}, \\ \Delta^{-R} m^{-R-1} \text{ if } m \geqq \Delta^{-1}, \end{cases}$$

provided R is bounded independently of q, as it will be. It will suffice to prove the existence of a positive integer $z < q^{2/l}$ for which $\psi_0(f(z)) > -\Delta$.

We consider the sums

$$S_{t,s}(y) = \sum_{x_{t,s}} \psi_0(f(yx_{t,s})), \tag{22}$$

where $x_{t,s}$ runs through the integers indicated in (3). We observe that, for $y = 1, \ldots, p$,

$$yx_{t,s} \leqq p(p_t c_s + p_t^{1-\varepsilon}) < p(\tfrac{1}{2}p + p^{1-\varepsilon}) < p^2 \leqq q^{2/l}.$$

If $\psi_0(f(z)) \leqq -\Delta$ for all positive integers $z < q^{2/l}$, we should have

$$S_{t,s}(y) \leqq -\Delta \xi_t$$

for $t = 1, \ldots, k$ and $s = 1, \ldots, g$.

Let

$$H = \sum_{y=1}^{p} \left| \prod_{t=1}^{k} \prod_{s=1}^{g} S_{t,s}(y) \right|. \tag{23}$$

Then we should have

$$H \geqq p\Delta^{kg}(\xi_1 \ldots \xi_k)^g = p\Delta^{kg}\Xi.$$

Our object is now to prove that the contrary is true, namely that

$$H < p\Delta^{kg}\Xi, \tag{24}$$

for sufficiently large q. This will establish the Theorem.

By (22) and the expansion of ψ_0, we have

$$S_{t,s}(y) \ll \sum_{m_{t,s}=1}^{\infty} F(m_{t,s}) \left| \sum_{x_{t,s}} e\{m_{t,s} f(yx_{t,s})\} \right|.$$

Hence, by (23),

$$H \ll \sum_{m_{1,1}=1}^{\infty} \cdots \sum_{m_{k,g}=1}^{\infty} F(m_{1,1}) \ldots F(m_{k,g}) T(m_{1,1}, \ldots, m_{k,g}), \tag{25}$$

where T is the sum of Lemma 2. Plainly $\sum_{m=1}^{\infty} F(m) \ll 1$.

Let $M = [q^{\varrho(1+\varepsilon_1)}]$, and let $R = [(k+1)g\varepsilon_1^{-1} + 1]$. Then $M > \Delta^{-1}$ and

$$(M\Delta)^{-R} \ll (q^{-\varrho\varepsilon_1})^R \ll q^{-(k+1)g\varrho}.$$

Hence

$$\sum_{m>M} F(m) \ll (M\Delta)^{-R} \ll q^{-(k+1)g\varrho} \tag{26}$$

and

$$\sum_{m\leq M} F(m)m^{-\frac{1}{4}} \ll \sum_{m\leq\Delta^{-1}} \Delta m^{-\frac{1}{4}} + \sum_{m>\Delta^{-1}} \Delta^{-R}m^{-R-\frac{5}{4}} \tag{27}$$
$$\ll \Delta^{\frac{1}{4}} \ll \Delta M^{\frac{3}{4}}.$$

The trivial estimate for T is $p\Xi$, and it follows from (26) that the contribution to (25) of the terms for which any one of $m_{1,1}, \ldots, m_{k,g}$ exceeds M is

$$\ll p\Xi q^{-(k+1)g\varrho} \ll p\Xi\Delta^{kg}q^{-g\varrho}.$$

For any fixed value of ϱ this is of lower order than the right hand side of (24).

There remains the contribution of the sum over $m_{1,1}, \ldots, m_{k,g}$ which do not exceed M. By Lemma 2 and (27), this is

$$\ll p\Xi\left(M^{g(k+1)}(p_1\ldots p_k)^{-D\nu+g\varepsilon}p^{D-1}\right)^{\frac{1}{4}}\left(\sum_{m\leq M} F(m)m^{-\frac{1}{4}}\right)^{kg}$$
$$\ll p\Xi\Delta^{kg}\{M^{4gk+g}(p_1\ldots p_k)^{-D\nu+g\varepsilon}p^{D-1}\}^{\frac{1}{4}}.$$

It suffices now to choose k and ϱ so that the expression in brackets tends to 0 as $q \to \infty$. We have

$$p_1\ldots p_k \gg p^{n(1-(1-\nu)^k)}, \quad M = [q^{\varrho(1+\varepsilon_1)}], \quad p = [q^\lambda],$$

so that the expression in brackets is $\ll q^\alpha$, where

$$\alpha = (4k+1)g\varrho + (1-\nu)^kD\lambda - \lambda + \varepsilon',$$

where ε' is arbitrarily small with ε and ε_1. We require α to be negative, which will be the case (for sufficiently small ε') if

$$\varrho\{(4k+1)g\} - \lambda\{1-(1-\nu)^kD\}$$

is negative. Thus we can take ϱ to be any number less than

$$\frac{\lambda\{1-(1-\nu)^kD\}}{g(4k+1)}. \tag{28}$$

We choose k to be the least integer for which

$$D(1-\nu)^k < \frac{1}{\log D + 1},$$

or in other words

$$k = \left[\frac{\log (D \log D + D)}{-\log(1-\nu)} + 1\right].$$

Since $-\log(1-\nu) > \nu/(1-\frac{1}{2}\nu)$, we have

$$k < n(1-\tfrac{1}{2}\nu) \log (D \log D + D) + 1.$$

The expression (28) is greater than

$$\frac{\log D}{4ngl(\log D + 1)\left((1-\frac{1}{2}\nu)\log(D \log D + D) + \frac{5}{4}\nu\right)}.$$

Since $D \geqq 5$, we have

$$\tfrac{1}{2}\log(D \log D + D) \geqq \tfrac{1}{2} \log(5 \log 5 + 5) = \tfrac{1}{2}(2.56\ldots) > \tfrac{5}{4}.$$

Hence the last expression is greater than the value of ϱ in the theorem, and the condition is satisfied.

An example. To illustrate the theorem, take

$$f(x) = \alpha x^n + x\sqrt{2},$$

where $n \geqq 5$ and α is real. Taking $l = 1$ and $a_l = \sqrt{2}$, we can choose a/q to be any convergent to $\sqrt{2}$. Since the continued fraction for $\sqrt{2}$ has bounded partial quotients, we can find, for any large N, a convergent for which $q \leqq N^{\frac{1}{2}} \ll q$. As $D = n + 1$ and $g = 2$, the theorem tells us that if N is sufficiently large there exist integers z and v such that

$$|\alpha z^n + z\sqrt{2} - v - A| < N^{-\varrho_0}, \qquad 0 < z < N,$$

where

$$\varrho_0 = \frac{\log (n + 1)}{16n\left(\log (n + 1) + 1\right) \log \left((n + 1) \log (n + 1) + n + 1\right)}.$$

NOTES ON CHAPTER V

This chapter has been considerably revised and expanded. The need for revision was greatest in the proof of what is now Lemma 2; the account in the original is difficult to follow because of the numerous misprints and the extreme conciseness.

The aim of this chapter, that of showing that the inequalities (20) are soluble, is achieved by showing that there is some integer z for which $|\psi_0(z)| < \Delta$. To effect this it is enough to estimate a sum of products, such as H in (23). The main idea lies in the construction of H which is so designed as to lead to a type of exponential sum which can be effectively estimated. These sums, in accordance with the general principle of Vinogradov's method, are of the type $\sum\sum e(\alpha uv)$, where one of the variables u, v arises from the variables $x_{t,s}$ and the other from the variable y.

The estimate obtained for Φ_t in the proof of Lemma 1 represents a result of a similar nature to Lemma 16 of Chapter I. The choice of the intervals (3) for the variables $x_{t,s}$ ensures firstly that these variables, for fixed t and $s = 1, \ldots, g$, lie in "well-spaced intervals", and secondly that the orders of magnitude of these variables diminish as t takes the values $1, \ldots, k$ in the manner required for Lemma 15 of Chapter I. The aim in this choice of intervals for the $x_{t,s}$ is to ensure that $U_h, \ldots, U_n$, defined by (4) and (5), are regularly distributed in the sense expressed by (7).

This regularity of distribution is needed to estimate the sum on the right of (17), which is a generalized form of a sum of the type $\sum\sum e(\alpha uv)$. The regularity in question is conveniently expressed by (18), and enables one to majorize T^4 by a sum which actually is of the form $\sum\sum e(\alpha uv)$.

CHAPTER VI

Estimates for Weyl Sums

In the present chapter my method is applied to prove two theorems which give estimates for sums of the form

$$\sum_{x=M+1}^{M+P} e(f(x)).$$

The first theorem relates to the case when $f(x)$ is a polynomial of degree $n+1$, and the estimate obtained depends on a rational approximation to any one of the coefficients of the powers of x higher than the first in the polynomial $f(x)$. The second theorem relates to the case when $f(x)$ is a real function which, in the interval of summation, behaves in a certain sense like a polynomial of degree n.

The same method can also be applied when n is not necessarily constant, but we do not consider this case here.

Notation in this chapter. We suppose $n \geqq 11$ and put

$$h = n + 2, \qquad b = [\tfrac{5}{4}n + \tfrac{1}{2}].$$

We denote by k a fixed integer greater than n, and write

$$\sigma = (1-\nu)^k.$$

We denote by

$$f(x) = a_{n+1}x^{n+1} + \ldots + a_1 x$$

a polynomial of degree $n+1$ with real coefficients.

Further notation. The following notation and terminology will be used in Lemmas 1 to 5 below. (Lemma 5 plays an important part in the proof of Theorems 1 and 2, and is of interest in itself.)

For any real number $R > 1$ and any integer g we denote by $L(R; g)$ any sum of the form

$$L(R; g) = \Sigma\, e(f(x)), \tag{1}$$

extended over all integers x in an interval

(2) $$gR + R' < x \leqq gR + R'',$$

where $0 \leqq R' < R'' \leqq R$. The precise values of R' and R'' in such a sum will not be important. The integer g will be called the *indicator* of the interval (2) or of the sum (1), relative to R.

A product of b sums, each of the above type, with the same R but with various indicators g and various values of R' and R'', can be written as a multiple sum:

(3) $$L(R; g_1) \ldots L(R; g_b) = Z(R; g_1, \ldots, g_b),$$

where

(4) $$Z(R; g_1, \ldots, g_b) = \sum_{x_1, \ldots, x_b} e(f(x_1) + \ldots + f(x_b)),$$

the summation being extended over all integral points $(x_1, \ldots, x_b)$ in the *box* in b dimensional space defined by

(5) $$g_jR + R_j' < x_j \leqq g_jR + R_j'' \qquad (j = 1, \ldots, b).$$

We denote this box by $B(R; g_1, \ldots, g_b)$, and call $g_1, \ldots, g_b$ the indicators of the box, relative to R. Such a box will be described as "a box of type $B(R)$".

If the sums on the left of (3) are all the same, the box defined by (5) is of a special kind, since g_j, R_j' and R_j'' are then independent of the value of j. Such a special box will be called a *diagonal box*, since one of its diagonals is a segment of the line $x_1 = \ldots = x_b$. The symbol $B^\dagger$ will be reserved for diagonal boxes.

A set of b integers $g_1, \ldots, g_b$ will be called a *proper* set if among them there exist n integers, every pair of which differ by at least 2. In the contrary case the set will be called *improper*. A box of type $B(R)$ or a sum of type $Z(R)$ will be said to be proper or improper according as its indicators constitute a proper or improper set of integers. The definition is relative to the number R, but this number will always be given explicitly in the description of the type of the box or sum.

Finally, it will be convenient to use the notation

$$\sum^{A} T$$

to denote a sum of $\ll A$ terms typified by T. In writing an inequality of the form

$$\sum^{A} T \ll \sum^{C} T,$$

it will sometimes be tacitly assumed that the sum on the right is suitably chosen, in a manner which will be clear from the context. Thus we may write

$$\sum^{A} T \ll C^{-1} \sum^{AC} T,$$

where C is a positive integer, meaning that the sum on the right is chosen to be the sum obtained from that on the left by repeating each term C times.

LEMMA 1a. *Let $p = RH$, where $R > 1$, $H > 1$. Let $B^{\dagger}(p)$ be a diagonal box, and let x_0 be an integer in the corresponding interval, so that $(x_0, \ldots, x_0)$ is a point in $B^{\dagger}(p)$. Let $B(R) = B(R; g_1, \ldots, g_b)$ be any proper box contained in $B^{\dagger}(p)$, and let $Z(R)$ be the corresponding sum. Then*

$$|Z(R)|^2 = \sum_{U_1^*, \ldots, U_{n+1}^*} e(X_1 U_1^* + \ldots + X_{n+1} U_{n+1}^*),$$

where

$$X_r = X_r(x_0) = \frac{1}{r!} f^{(r)}(x_0) \quad for\ r = 1, \ldots, n+1,$$

and the summation is extended over a set S^* of integral points $(U_1^*, \ldots, U_{n+1}^*)$, not necessarily distinct, in $n+1$ dimensional space. This set of points has the properties:

(i) *the coordinates of every point of S^* satisfy*

$$U_1^* \ll p, \ldots, U_{n+1}^* \ll p^{n+1};$$

(ii) *given any n intervals of lengths*

$$p^{1-\nu}, \ldots, p^{n(1-\nu)},$$

the number of points of S^* whose first n coordinates $U_1^*, \ldots, U_n^*$

fall respectively into these intervals is

$$\ll R^{2b-n} H^{\frac{1}{2}n(n-1)} p^{\frac{1}{2}(n-1)}.$$

Proof. As in (3), we have

$$Z(R) = Z(R; g_1, \ldots, g_b) = L(R; g_1) \ldots L(R; g_b),$$

where

$$L(R; g_j) = \sum_{x_j} e(f(x_j)),$$

summed over all integers x_j in an interval of the form (5). In each sum $L(R; g_j)$ we put $x_j = x_0 + v_j$. Then

$$L(R; g_j) = \sum_{v_j} e(f(x_0 + v_j))$$
$$= \sum_{v_j} e(f(x_0) + X_1 v_j + \ldots + X_{n+1} v_j^{n+1}),$$

where the summation is over all integers v_j in the interval

(6) $$-x_0 + g_j R + R_j' < v_j \leqq -x_0 + g_j R + R_j'',$$

and where $X_1, \ldots, X_{n+1}$ are as defined in the enunciation.

Since $g_1, \ldots, g_b$ is a proper set of integers, we can suppose without loss of generality (on permuting the sums $L(R; g_j)$) that

(7) $$g_1 + 1 < g_2, g_2 + 1 < g_3, \ldots, g_{n-1} + 1 < g_n.$$

We have

$$| Z(R) |^2 = | L(R; g_1) \ldots L(R; g_b) |^2$$
$$= \sum_{v_1, \ldots, v_{2b}} e(X_1 U_1^* + \ldots + X_{n+1} U_{n+1}^*),$$

where $v_{b+1}, \ldots, v_{2b}$ run through the same intervals as $v_1, \ldots, v_b$ respectively, and where

$$U_r^* = v_1^r + \ldots + v_b^r - v_{b+1}^r - \ldots - v_{2b}^r \quad (r = 1, \ldots, n+1).$$

Let S^* denote the set of all integral points $(U_1^*, \ldots, U_{n+1}^*)$ in $n+1$ dimensional space, arising from all sets of values of $v_1, \ldots, v_{2b}$ specified above. These points are not necessarily distinct; each point occurs as often as there are values of $v_1, \ldots, v_{2b}$ which give rise to it. The expression for $| Z(R) |^2$ is of the form stated in the enunciation, and it remains to prove that the set S^* has the properties (i) and (ii).

Since $(x_0, \ldots, x_0)$ is in $B^\dagger(p)$, and the box $B(R)$ defined by (5) is contained in $B^\dagger(p)$, all the intervals (6) for the variables v_j are contained in the interval $-p \leqq v_j \leqq p$. Hence $U_r^* \ll p^r$ for $r = 1, \ldots, n+1$, and the property (i) holds.

To prove (ii), we write

$$U_r^* = V_r^* + W_r^* \text{ for } r = 1, \ldots, n,$$

where

$$V_r^* = v_1^r + \ldots + v_n^r,\ W_r^* = v_{n+1}^r + \ldots + v_b^r - v_{b+1}^r - \ldots - v_{2b}^r.$$

By (6) and (7), the variables $v_1, \ldots, v_n$ run through the integers in n intervals, contained in the interval $-p \leqq v_j \leqq p$, such that every two intervals are separated by a gap of length $\geqq R$. We now appeal to Lemma 16 of Chapter I. This asserts that the number of sets of values of $v_1, \ldots, v_n$ for which $V_1^*, \ldots, V_n^*$ fall into any n given intervals whose lengths are respectively

$$p^{1-\nu}, \ldots, p^{n(1-\nu)}$$

is

$$\ll H^{\frac{1}{2}n(n-1)} p^{\frac{1}{2}(n-1)}.$$

The number of possible values for each of $v_{n+1}, \ldots, v_{2b}$ is $\leqq R + 1 \ll R$. For each set of values $v_{n+1}, \ldots, v_{2b}$, the number of sets of values of $v_1, \ldots, v_n$ for which $V_1^* + W_1^*, \ldots, V_n^* + W_n^*$ fall into given intervals of the specified lengths is as estimated above. The property (ii) of the set S^* now follows.

Lemma 1b. *If $R^n \ll H^{\frac{1}{2}n(n-1)} p^{\frac{1}{2}(n-1)}$, then Lemma 1a remains true even if the box $B(R)$ is improper.*

Proof. The proof is the same as before, but instead of appealing to Lemma 16 of Chapter I (which is no longer applicable) we take the trivial upper bound R^{2b} for the total number of points in the set S^*. By the hypothesis of the present lemma,

$$R^{2b} \ll R^{2b-n} H^{\frac{1}{2}n(n-1)} p^{\frac{1}{2}(n-1)},$$

and the conclusion follows.

Remark. Lemma 1b is of course trivial, and is formulated merely for convenience of reference later.

LEMMA 2. *For $H > 1$ and any H_0, the number $N(H)$ of improper sets of integers $g_1, \ldots, g_b$ satisfying*

$$H_0 < g_j \leqq H_0 + H \qquad (j = 1, \ldots, b)$$

can be estimated by

$$N(H) \ll H^{n-1}. \tag{8}$$

Proof. For any such set $g_1, \ldots, g_b$, let G_1 be the least number in the set; G_2 the least number in the set which is greater than $G_1 + 1$; G_3 the least number in the set which is greater than $G_2 + 1$; and so on. This defines a sequence $G_1, G_2, \ldots$, terminating say at a number G_m which has the property that every number in the set is $\leqq G_m + 1$. Since the set is improper, we must have $m < n$.

There are at most $[H] + 1$ possibilities for each G, and therefore at most $(H + 1)^{n-1}$ possibilities for $G_1, \ldots, G_m$. When $G_1, \ldots, G_m$ are known, each g must have one of the values

$$G_1,\ G_1 + 1,\ G_2,\ G_2 + 1, \ldots, G_m,\ G_m + 1.$$

These values number $2m \leqq 2n - 2$, and therefore the number of possible improper sets $g_1, \ldots, g_b$ satisfying the given inequalities is at most

$$(H + 1)^{n-1}(2n - 2)^b,$$

whence the result.

LEMMA 3. *Let $\eta \geqq 3$ be an integer. Any diagonal box of type $B^\dagger(p)$ can be subdivided into boxes of types*

$$B(2^{-2}p),\ B(2^{-3}p), \ldots, B(2^{-\eta+1}p),$$

each of which is a proper box of its type, together with other boxes of the type

$$B(2^{-\eta}p),$$

which may be improper. This can be done in such a way that the number $B^{(s)}$ of boxes of the type $B(2^{-s}p)$ satisfies

$$B^{(s)} \ll 2^{s(n-1)} \text{ for } s = 2, \ldots, \eta. \tag{9}$$

Proof. A diagonal box of type $B^{\dagger}(p)$ is defined by b inequalities of the form

$$gp + p' < x_j \leqq gp + p'' \qquad (j = 1, \ldots, b),$$

where $0 \leqq p' < p'' \leqq p$. This can be subdivided into at most 2^{2b} boxes of type $B(2^{-2}p)$ by dividing each of the intervals $gp < x_j \leqq (g+1)\,p$ into four equal parts. Each box will be of the form

$$g_j 2^{-2}p + p_j' < x_j \leqq g_j 2^{-2}p + p_j'' \quad (j = 1, \ldots, b),$$

where

$$0 \leqq p_j' < p_j'' \leqq 2^{-2}p,$$

and each g_j has one of the values $4g$, $4g+1$, $4g+2$, $4g+3$. Some of these boxes may be proper and some improper [1]); we denote the proper boxes by $B'(2^{-2}p)$ and the improper boxes by $B''(2^{-2}p)$.

We subdivide each of the improper boxes $B''(2^{-2}p)$ into at most 2^b boxes of type $B(2^{-3}p)$, and classify these as $B'(2^{-3}p)$ and $B''(2^{-3}p)$ according as they are proper or improper of their type. We continue the process until we have subdivided the improper boxes of type $B''(2^{-\eta+1}p)$ into boxes of type $B(2^{-\eta}p)$, some of which will be proper and some improper.

We now have the situation described in the enunciation, and it remains only to prove the estimate (9). The indicators of the various boxes of type $B(2^{-s+1}p)$ are integers, each of which is one of the integers $2^{s-1}g + g'$, where $g' = 0, \ldots, 2^{s-1} - 1$. Hence, by Lemma 2 with $H = 2^{s-1}$, the number of improper boxes $B''(2^{-s+1}p)$ is $\ll 2^{(s-1)(n-1)}$. For $s - 1 = 2, \ldots, \eta - 1$, each such improper box is divided into a bounded number ($\leqq 2^b$) of boxes $B'(2^{-s}p)$ and $B''(2^{-s}p)$. Hence the estimate (9) is valid for $s = 3, \ldots, \eta$. Also the total number of boxes of type $B(2^{-2}p)$ is at most 2^{2b}, and therefore the estimate (9) is trivially true when $s = 2$.

[1]) In point of fact, since $n \geqq 11$, all boxes of type $B(2^{-s}p)$ will be improper for the first few values of s.

LEMMA 4. *Let p_1 be an arbitrarily large positive integer, and let*

$$p_t = p_1^{(1-\nu)^{t-1}}, \qquad \eta_t = [\nu \log_2 p_t]$$

for $t = 1, 2, \ldots, k+1$. Let $L(p_t) = L(p_t; g)$ be any sum of the type (1). Then, for $t = 1, \ldots, k$ we have

$$(10) \qquad \begin{cases} | L(p_t) |^{2b(k+h-t+1)} \\ \ll \sum_{s=2}^{\eta_t} \sum^{M(t,s)} | Z(2^{-s}p_t) |^2 \, | L(p_{t+1}) |^{2b(k+h-t)}, \end{cases}$$

where

$$(11) \qquad M(t, s) = 2^{-s(2b-2n)(k+h-t)+2sn} p_t^{2b\nu(k+h-t)},$$

and where

(i) *each sum of type $Z(2^{-s}p_t)$ on the right of (10) is extended over a box of type $B(2^{-s}p_t)$ contained in the diagonal box of type $B^\dagger(p_t)$ corresponding to $(L(p_t))^b$;*

(ii) *if $s < \eta_t$ the sum $Z(2^{-s}p_t)$ is proper;*

(iii) *the integer g implicit in the definition of each $L(p_{t+1})$ on the right of (10) depends on s and on the particular term in the sum of $\ll M(t, s)$ terms;*

(iv) *the diagonal box of type $B^\dagger(p_{t+1})$ corresponding to each $(L(p_{t+1}))^b$ on the right of (10) is contained in the diagonal box $B^\dagger(p_t)$ corresponding to $(L(p_t))^b$.*

Proof. For convenience of writing we put $l = k + h - t$ and omit the suffix t from η_t in the proof. We regard $(L(p_t))^b$ as a multiple sum over a diagonal box of type $B^\dagger(p_t)$, and we subdivide this box in accordance with Lemma 3. Corresponding to this subdivision we have

$$(L(p_t))^b = \sum_{s=2}^{\eta} \sum^{2^{-s}2^{sn}} Z(2^{-s}p_t),$$

where the boxes for the sums $Z(2^{-s}p_t)$ are contained in the box $B^\dagger(p_t)$ and are proper if $s < \eta$. Repeating the term corresponding to each box 2^s times, we can write

$$(L(p_t))^b \ll \sum_{s=2}^{\eta} 2^{-s} \sum^{2^{sn}} | Z(2^{-s}p_t) |.$$

To the sum on the right we apply Lemma 2 of Chapter I, followed by Hölder's inequality. Thus

$$|L(p_t)|^{2b(l+1)} \ll \sum_{s=2}^{\eta} \left(\sum^{2^{sn}} |Z(2^{-s}p_t)|\right)^{2(l+1)}$$

$$\ll \sum_{s=2}^{\eta} 2^{sn(2(l+1)-1)} \sum^{2^{sn}} |Z(2^{-s}p_t)|^{2(l+1)}$$

$$\ll \sum_{s=2}^{\eta} \sum^{2^{2sn(l+1)}} |Z(2^{-s}p_t)|^2 \, |Z(2^{-s}p_t)|^{2l}.$$

Now consider an individual one of the sums $Z(2^{-s}p_t)$. This factorizes, in accordance with (3), as

$$Z(2^{-s}p_t) = L(2^{-s}p_t;\ g_1) \ldots L(2^{-s}p_t;\ g_b).$$

By the inequality of the arithmetic and geometric means,

$$b^b \mid Z(2^{-s}p_t) \mid \leqq \left(\sum_{j=1}^{b} |L(2^{-s}p_t;\ g_j)|\right)^b,$$

whence, by Hölder's inequality,

$$| Z(2^{-s}p_t) |^{2l} \ll \sum_{j=1}^{b} | L(2^{-s}p_t;\ g_j) |^{2bl}.$$

Each sum $L(2^{-s}p_t;\ g_j)$ on the right is extended over an interval of length $\leqq 2^{-s}p_t$. We split up this interval into intervals of length $\leqq p_{t+1}$, each of the form

$$gp_{t+1} + p' < x \leqq gp_{t+1} + p'',$$

where g takes certain consecutive integral values, and

$$0 \leqq p' < p'' \leqq p_{t+1}.$$

The number of such intervals is at most

$$\frac{2^{-s}p_t}{p_{t+1}} + 2 = 2^{-s}p_t^{\nu} + 2 \ll 2^{-s}p_t^{\nu},$$

since

$$2^s \leqq 2^{\eta} \leqq p_t^{\nu}.$$

Accordingly, each sum $L(2^{-s}p_t;\ g_j)$ splits up into $\ll 2^{-s}p_t^{\nu}$ sums each of which is of the type $L(p_{t+1})$. Applying Hölder's inequality, we obtain

$$| Z(2^{-s}p_t) |^{2l} \ll \sum_{j=1}^{b} (2^{-s}p_t^{\nu})^{2bl-1} \sum^{2^{-s}p_t^{\nu}} | L(p_{t+1}) |^{2bl}$$
$$\ll \sum^{(2^{-s}p_t^{\nu})^{2bl}} | L(p_{t+1}) |^{2bl}.$$

Each $L(p_{t+1})$ here is extended over an interval which is contained in one of the b intervals defining the box of type $B(2^{-s}p_t)$ corresponding to the sum $Z(2^{-s}p_t)$ on the left. The latter box is contained in the original diagonal box $B^{\dagger}(p_t)$. Hence the diagonal box of type $B^{\dagger}(p_{t+1})$ corresponding to each $(L(p_{t+1}))^b$ is contained in the original diagonal box $B^{\dagger}(p_t)$ corresponding to $(L(p_t))^b$.

We now have

$$| L(p_t) |^{2b(l+1)} \ll \sum_{s=2}^{\eta} \sum^{M(t,s)} | Z(2^{-s}p_t) |^2 | L(p_{t+1}) |^{2bl}.$$

where

$$M(t, s) = 2^{2sn(l+1)}(2^{-s}p_t^{\nu})^{2bl} = 2^{-s(2b-2n)l+2sn}p_t^{2bl\nu}.$$

This is the result asserted, and the statements (i), (ii), (iii), (iv) are obvious from the proof.

LEMMA 5. *Let p_1 be an arbitrarily large positive integer and let*

$$T_1 = \sum_{x=M+1}^{M+p_1} e(f(x)),$$

where M is any integer. Let

$$p_t = p_1^{(1-\nu)^{t-1}} \text{ and } \eta_t = [\nu \log_2 p_t]$$

for $t = 1, \ldots, k+1$. Then

$$T_1^{2b(k+h)} \ll \sum_{s_1=2}^{\eta_1} \ldots \sum_{s_k=2}^{\eta_k} \sum^{J} K(s_1, \ldots, s_k), \tag{12}$$

where

$$J = p_1^{2b(k+h)}(p_1 \ldots p_k)^{-2b}p_{k+1}^{-2bh}\, 2^{(2n-2h(b-n))(s_1+\cdots+s_k)},$$

and

$$(13) \qquad K(s_1, \ldots, s_k) = \sum_{x_1, \ldots, x_{n+1}} e(X_1 x_1 + \ldots + X_{n+1} x_{n+1}).$$

In the last formula, $X_r = \frac{1}{r!} f^{(r)}(x_0)$, *where* x_0 *is an integer which depends on the particular term* $K(s_1, \ldots, s_k)$ *in question. The summation on the right of* (13) *is extended over the points* $(x_1, \ldots, x_{n+1})$ *of a set* S *of integral points, not necessarily distinct, in* $n + 1$ *dimensional space. This set* S *depends on the particular* $K(s_1, \ldots, s_k)$ *in question, but not on the coefficients of* $f(x)$. *All points of the set* S *satisfy*

$$(14) \qquad |x_1| \leqq cp_1, \ldots, |x_{n+1}| \leqq cp_1^{n+1};$$

and for any given n *integers* $z_1, \ldots, z_n$, *the number of points* $(x_1, \ldots, x_{n+1})$ *of* S *for which*

$$x_1 = z_1, \ldots, x_n = z_n$$

is $\ll \psi$ *where*

$$\psi = (p_1 \ldots p_k)^{2b-\frac{1}{2}(n+1)} p_{k+1}{}^{2bh} 2^{(-2b+\frac{1}{2}n(n+1))(s_1+\cdots+s_k)}.$$

Proof. We can suppose that the sum T_1 is of the type $L(p_1)$, since it can always be split into two sums of this type. We apply Lemma 4 with $t = 1$ to $|L(p_1)|^{2b(k+h)}$, giving

$$|T_1|^{2b(k+h)} \ll \sum_{s_1=2}^{\eta_1} \sum^{M(1, s_1)} |Z(2^{-s_1} p_1)|^2 |L(p_2)|^{2b(k+h-1)}.$$

Next we apply Lemma 4 with $t = 2$ to each term $|L(p_2)|^{2b(k+h-1)}$ on the right, and continue this process, finishing with an application of Lemma 4 with $t = k$. The result can be expressed in the form

$$(15) \qquad |T_1|^{2b(k+h)} \ll \sum_{s_1=2}^{\eta_1} \ldots \sum_{s_k=2}^{\eta_k} \sum^{M(s_1, \ldots, s_k)} K(s_1, \ldots, s_k),$$

where

$$(16) \qquad \begin{cases} M(s_1, \ldots, s_k) = M(1, s_1) \ldots M(k, s_k) \\ \qquad = \prod_{t=1}^{k} 2^{-s_t(2b-n)(k+h-t)+2ns_t} p_t{}^{2b\nu(k+h-t)}, \end{cases}$$

and

(17) $\quad K(s_1, \ldots, s_k) = |Z(2^{-s_1}p_1)|^2 \ldots |Z(2^{-s_k}p_k)|^2 \, |L(p_{k+1})|^{2bh}.$

It is plain, from Lemma 4 and the above construction, that

(i) if $s_t < \eta_t$ the corresponding sum $Z(2^{-s_t}p_t)$ occurring on the right of (17) is proper;

(ii) if $B^\dagger(p_{k+1})$ denotes the diagonal box corresponding to $(L(p_{k+1}))^b$ in any $K(s_1, \ldots, s_k)$, then there exist diagonal boxes of types $B^\dagger(p_1), \ldots, B^\dagger(p_k)$ such that, for $t = 1, \ldots, k$, $B^\dagger(p_t)$ contains both the diagonal box $B^\dagger(p_{k+1})$ and the box $B(2^{-s_t}p_t)$ corresponding to $Z(2^{-s_t}p_t)$.

For the rest of the proof (apart from a mere verification relating to $M(s_1, \ldots, s_k)$) we shall be concerned with the form of one particular $K(s_1, \ldots, s_k)$ in the sum on the right of (15). Thus we can think of $s_1, \ldots, s_k$ as fixed, and the Z's and $L(p_{k+1})$ as particular sums of the specified types. We take x_0 to be any integer in the interval of summation for the sum $L(p_{k+1})$. Then the point $(x_0, \ldots, x_0)$ is in the diagonal box $B^\dagger(p_{k+1})$, and hence is in the diagonal boxes $B^\dagger(p_1), \ldots, B^\dagger(p_k)$ mentioned in (ii) above.

We first investigate an individual sum $Z(2^{-s_t}p_t)$ in the factorization of $K(s_1, \ldots, s_k)$. Suppose first that $s_t < \eta_t$, so that by (i) above the sum is a proper sum. We apply Lemma la to this sum, taking

$$p = p_t, \; R = 2^{-s_t}p_t, \; H = 2^{s_t},$$

and x_0 as just specified. The condition $R > 1$ is amply satisfied, since $p_t > p_t^\nu \geqq 2^{\eta_t} > 2^{s_t}$ as we saw in the proof of Lemma 4. It follows from Lemma la that

$$(18) \quad \begin{cases} |Z(2^{-s_t}p_t)|^2 \\ = \displaystyle\sum_{U_{t,1}^*, \ldots, U_{t,n+1}^*} e(X_1U_{t,1}^* + \ldots + X_{n+1}U_{t,n+1}^*), \end{cases}$$

the summation being extended over an $n+1$ dimensional set of points S_t^*. This set of points is such that

(a) the coordinates of every point of S_t^* satisfy

(19) $$U_{t,1}^* \ll p_t, \ldots, U_{t,n+1}^* \ll p_t^{n+1};$$

(b) the number of points of S_t^* whose first n coordinates fall respectively into any given intervals of lengths

$$p_t^{1-\nu}, \ldots, p_t^{n(1-\nu)}$$

is $\ll \Phi_t$, where

$$\begin{aligned} \Phi_t &= R^{2b-n} H^{\frac{1}{2}n(n-1)} p^{\frac{1}{2}(n-1)} \\ &= (2^{-s_t} p_t)^{2b-n} (2^{s_t})^{\frac{1}{2}n(n-1)} p_t^{\frac{1}{2}(n-1)} \\ &= p_t^{2b-\frac{1}{2}(n+1)} 2^{(-2b+\frac{1}{2}n(n+1))s_t}. \end{aligned} \tag{20}$$

Suppose next that $s_t = \eta_t$. By Lemma 1b, the same conclusions hold, provided we can verify that

$$R^n \ll H^{\frac{1}{2}n(n-1)} p^{\frac{1}{2}(n-1)}.$$

This condition, in the present application, is equivalent (since $s_t = \eta_t$) to

$$(2^{-\eta_t} p_t)^n \ll (2^{\eta_t})^{\frac{1}{2}n(n-1)} p_t^{\frac{1}{2}(n-1)},$$

or

$$p_t^{\frac{1}{2}(n+1)} \ll 2^{\frac{1}{2}n(n+1)\eta_t},$$

and so is satisfied since $p_t^\nu \ll 2^{\eta_t}$ by the definition of η_t. Thus an expression of the form (18), with (a) and (b) above, holds for every factor $|Z(2^{-s_t} p_t)|^2$ in the factorization (17) of the particular K under consideration.

We now turn to the last factor on the right of (17). We have

(21) $$|L(p_{k+1})|^{2bh} = \sum_{y_1, \ldots, y_{2bh}} e(X_1 Y_1^* + \ldots + X_{n+1} Y_{n+1}^*),$$

where

$$Y_r^* = y_1^r + \ldots + y_{bh}^r - y_{bh+1}^r - \ldots - y_{2bh}^r$$

for $r = 1, \ldots, n+1$, and where $y_j = x_j - x_0$ and x_j runs through the interval of summation for $L(p_{k+1})$. As x_0 is in this interval, it is plain that $|y_j| \leq p_{k+1}$. We can rewrite (21) as

(22) $$|L(p_{k+1})|^{2bh} = \sum_{Y_1^*, \ldots, Y_{n+1}^*} e(X_1 Y_1^* + \ldots + X_{n+1} Y_{n+1}^*),$$

the summation being over a set S^* of points, not necessarily distinct, in $n+1$ dimensional space. The number of points in S^* is $\leqq p_{k+1}^{2bh}$, and each point satisfies

$$Y_1^* \ll p_{k+1}, \ldots, Y_{n+1}^* \ll p_{k+1}^{n+1}. \tag{23}$$

Multiplying together the expressions found in (18) and (22) for all the factors on the right of (17), we obtain

$$\begin{cases} K(s_1, \ldots, s_k) \\ = \sum_1 \ldots \sum_k \sum\limits_{Y_1^*, \ldots, Y_{n+1}^*} e(X_1(U_1 + Y_1^*) \\ + \ldots + X_{n+1}(U_{n+1} + Y_{n+1}^*)), \end{cases} \tag{24}$$

where

$$U_j = U_{1,j}^* + \ldots + U_{k,j}^* \text{ for } j = 1, \ldots, n+1, \tag{25}$$

and $\sum_t$ denotes a summation for the point

$$(U_{t,1}^*, \ldots, U_{t,n+1}^*)$$

over the set S_t^* mentioned after (18). It will be appreciated that the set S_t^* (for $t = 1, \ldots, k$) depends on the value of s_t as well as on t, but we regard $s_1, \ldots, s_k$ as fixed for the present. The summation for $Y_1^*, \ldots, Y_{n+1}^*$ is over the set S^*, as before.

Now let S_t denote the set of n dimensional points derived from the $n+1$ dimensional set S_t^* by omitting the last coordinate of every point. It is understood that each point is counted as often as it arises, and therefore the number of points in S_t is the same as the number in S_t^*. The statement (b) above gives an estimate for the number of points of S_t in intervals of the prescribed lengths. We apply Lemma 15 of Chapter I to the sets $S_1, \ldots, S_k$, taking the numbers $h, \ldots, l$ to be $1, \ldots, n-1$ without exception (so that $g = n$) and taking $M = 1$. It follows from (a) and (b) above that the hypotheses of that lemma are satisfied. Hence, for any n given integers $z_1', \ldots, z_n'$ the number of selections of points, one from each of the $n+1$ dimensional sets $S_1^*, \ldots, S_k^*$, for which

$$U_1 = z_1', \ldots, U_n = z_n',$$

is $\ll \Phi$, where

$$\Phi = \Phi_1 \ldots \Phi_k,$$

Φ_t being defined in (20).

Let S denote the $n+1$ dimensional set of points

$$(U_1 + Y_1^*, \ldots, U_{n+1} + Y_{n+1}^*),$$

where $U_1, \ldots, U_{n+1}$, defined in (25), arise from every selection of points one from each of the sets $S_1^*, \ldots, S_k^*$, and $Y_1^*, \ldots, Y_{n+1}^*$ represents every point of the set S^*. Then the expression (24) for $K(s_1, \ldots, s_k)$ is an expression of the form (13), summed over points $(x_1, \ldots, x_{n+1})$ of S. All points of S satisfy (14), by (19), (25), (23). The number of points of S satisfying

$$x_1 = z_1, \ldots, x_n = z_n,$$

for n given integers $z_1, \ldots, z_n$ is $\ll \psi$ where

$$\begin{aligned}\psi &= (p_{k+1})^{2bh}\Phi \\ &= (p_{k+1})^{2bh} \prod_{t=1}^{k} p_t^{2b-\frac{1}{2}(n+1)} 2^{(-2b+\frac{1}{2}n(n+1))s_t},\end{aligned}$$

as asserted in the enunciation.

We have now proved all that was asserted, apart from the value of J. It therefore remains only to verify that $M(s_1, \ldots, s_k)$, given in (16), satisfies

$$M(s_1, \ldots, s_k) \ll J$$

for all choices of $s_1, \ldots, s_k$. Since $b > n$ and $k - t + h \geqq h$, we have

$$\begin{aligned}M(s_1, \ldots, s_k) &\leqq \prod_{t=1}^{k} 2^{-s_t(2b-2n)h+2s_t n} p_t^{2b\nu(k-t+h)} \\ &= 2^{(2n-2h(b-n))(s_1+\cdots+s_k)} \left(\prod_{t=1}^{k} p_t^{\nu(k-t+h)}\right)^{2b}.\end{aligned}$$

Now $p_t^{\nu} = p_t/p_{t+1}$, hence

$$\begin{aligned}\prod_{t=1}^{k} p_t^{\nu(k-t+h)} &= \left(\frac{p_1}{p_2}\right)^{k+h-1} \left(\frac{p_2}{p_3}\right)^{k+h-2} \cdots \left(\frac{p_k}{p_{k+1}}\right)^{h} \\ &= p_1^{k+h}(p_1 \ldots p_k)^{-1} p_{k+1}^{-h}.\end{aligned}$$

This gives the value of J, and the proof is complete.

LEMMA 6. *Let m and P be positive integers, and let ϱ' be a real* *number satisfying* $0 < \varrho' < 1$. *Let*

$$S_1 = \sum_{x=1}^{P} e(mf(x)).$$

Let r *be one of the numbers* $n+1, n, \ldots, 2$, *and let the coefficient* a_r *of* $f(x)$ *satisfy*

$$a_r = \frac{a}{q} + \frac{\theta}{q^2}, \text{ where } (a, q) = 1 \text{ and } 0 < q < P^{r+\varrho'}.$$

Then

$$(26) \begin{cases} S_1^{2b(k+h)} \ll HP^{\frac{1}{2}n(n+1)-1+(n+1)\varrho'} \int_0^1 \ldots \int_0^1 |T_1^*|^{2b(k+h)} d\alpha_1 \ldots d\alpha_n \\ \qquad + (P^{1-\varrho'})^{2b(k+h)}, \end{cases}$$

where

$$T_1^* = T_1^*(\alpha_1, \ldots, \alpha_n) = \sum_{x=1}^{P} e(\alpha_1 x + \ldots + \alpha_n x^n + ma_{n+1}x^{n+1}),$$

and

$$H = (m + qP^{-r-\varrho'+1})(Pq^{-1} + 1).$$

Proof. *Let* $Y = [P^{1-\varrho'}]$. Define $S_1^*(y)$, for $y = 0, 1, \ldots, Y-1$, by

$$S_1^*(y) = \sum_{x=1}^{P} e(m(f(y + x) - f(y))).$$

Then, putting $x' = y + x$,

$$|S_1^*(y)| = \left|\sum_{x'=1+y}^{P+y} e(mf(x'))\right| = |S_1| + 2\theta y.$$

Hence, summing over y and dividing by Y, we have

$$|S_1| = Y^{-1} \sum_{y=0}^{Y-1} |S_1^*(y)| + \theta' Y.$$

Therefore, by Hölder's inequality,

$$|S_1|^{2b(k+h)} \ll \left(Y^{-1} \sum_{y=0}^{Y-1} |S_1^*(y)|\right)^{2b(k+h)} + Y^{2b(k+h)}$$

$$(27) \qquad \ll Y^{-1} \sum_{y=0}^{Y-1} |S_1^*(y)|^{2b(k+h)} + Y^{2b(k+h)}.$$

By Taylor's theorem,

$$mf(y+x) - mf(y) = Y_1 x + \ldots + Y_{n+1} x^{n+1},$$

where

$$(28) \qquad Y_j = Y_j(y) = \frac{m}{j!} f^{(j)}(y) = \binom{n+1}{j} a_{n+1} m y^{n+1-j} + \ldots + \binom{j+1}{j} a_{j+1} m y + a_j m,$$

and in particular

$$Y_{n+1} = m a_{n+1}.$$

Hence

$$S_1^*(y) = \sum_{x=1}^{P} e(Y_1 x + \ldots + Y_n x^n + m a_{n+1} x^{n+1}).$$

Let Ω_y denote the small region in n dimensional space round the point $(Y_1, \ldots, Y_n)$ consisting of points $(\alpha_1, \ldots, \alpha_n)$ satisfying

$$|\alpha_1 - Y_1| \leqq \tfrac{1}{2} P^{-1-\varrho'}, \ldots, |\alpha_n - Y_n| \leqq \tfrac{1}{2} P^{-n-\varrho'}.$$

Since $e(\beta) = 1 + O(|\beta|)$ for $\beta \ll 1$, we have

$$e(Y_1 x + \ldots + Y_n x^n) = e(\alpha_1 x + \ldots + \alpha_n x^n) + O(P^{-\varrho'})$$

for any point $(\alpha_1, \ldots, \alpha_n)$ in Ω_y and for $x = 1, \ldots, P$. Hence

$$S_1^*(y) = T_1^*(\alpha_1, \ldots, \alpha_n) + O(P^{1-\varrho'})$$

for any such point, where T_1^* is the sum defined in the enunciation. It follows from this that

$$|S_1^*(y)|^{2b(k+h)} \ll |T_1^*(\alpha_1, \ldots, \alpha_n)|^{2b(k+h)} + Y^{2b(k+h)}$$

for $(\alpha_1, \ldots, \alpha_n)$ in Ω_y.

Let $\tau(\Omega_y) = \tau(\Omega_y; \alpha_1, \ldots, \alpha_n)$ denote the characteristic function of the region Ω_y (mod 1), that is, the function which is 1 if $\alpha_1, \ldots, \alpha_n$ differ by integral amounts from the coordinates of a point in Ω_y, and 0 otherwise. We note that T_1^* is a periodic function of $\alpha_1, \ldots, \alpha_n$ with period 1 in each variable. Integrating the last relation with respect to $\alpha_1, \ldots, \alpha_n$ over Ω_y, we obtain

$$V\,|S_1^*(y)|^{2b(k+h)} \ll \int_{\Omega_y} |T_1^*|^{2b(k+h)} d\alpha_1 \ldots d\alpha_n + V Y^{2b(k+h)},$$

where V denotes the volume of Ω_y. Since

$$V = P^{-\frac{1}{2}n(n+1)-n\varrho'},$$

we obtain

$$|S_1^*(y)|^{2b(k+h)} \ll P^{\frac{1}{2}n(n+1)+n\varrho'} \int_0^1 \cdots \int_0^1 |T_1^*|^{2b(k+h)} \tau(\Omega_y) d\alpha_1 \ldots d\alpha_n + Y^{2b(k+h)}.$$

Substituting in (27), we have

$$|S_1|^{2b(k+h)} \ll P^{\frac{1}{2}n(n+1)+n\varrho'} \int_0^1 \cdots \int_0^1 |T_1^*|^{2b(k+h)} Y^{-1} \sum_{x=0}^{Y-1} \tau(\Omega_y) d\alpha_1 \ldots d\alpha_n + Y^{2b(k+h)}.$$

The desired result (26) will now follow, in view of the definition of Y, provided we can prove that

$$\sum_{y=0}^{Y-1} \tau(\Omega_y) \ll H = (m + qP^{-r-\varrho'+1})(Pq^{-1} + 1),$$

for every point $(\alpha_1, \ldots, \alpha_n)$. This is tantamount to proving that the number of values of y for which Ω_y contains a point whose coordinates differ from $\alpha_1, \ldots, \alpha_n$ by integral amounts is $\ll H$.

Let Ω_y and Ω_{y_0} correspond to two such values of y. Then, by the definition of Ω_y, we have

$$Y_n(y) - Y_n(y_0) = C_n + O(P^{-n-\varrho'}),$$
$$\ldots$$
$$Y_{r-1}(y) - Y_{r-1}(y_0) = C_{r-1} + O(P^{-r+1-\varrho'}),$$

where $C_n, \ldots, C_{r-1}$ are integers. It suffices to prove that if y_0 is fixed, these relations admit only of $\ll H$ values for y. By (28), the relations are

$$\binom{n+1}{1} a_{n+1} m(y - y_0) = C_n + O(P^{-n-\varrho'}),$$

$$\binom{n+1}{2} a_{n+1} m(y^2 - y_0^2) + \binom{n}{1} a_n m(y - y_0) = C_{n-1} + O(P^{-n+1-\varrho'}),$$
$$\ldots$$
$$\binom{n+1}{n-r+2} a_{n+1} m(y^{n-r+2} - y_0^{n-r+2}) + \ldots + \binom{r}{1} a_r m(y - y_0) = C_{r-1} + O(P^{-r+1-\varrho'}).$$

If we multiply the second equation by $2!$, the third by $3!$, and so on, we obtain equations in which the first term on the left is always divisible by the left hand side of the first equation, the quotient on dividing the first term of the sth equation by the left hand side of the first equation being

$$n \ldots (n-s+2) \frac{y^s - y_0^s}{y - y_0},$$

which is $O(P^{s-1})$. Hence we can subtract this multiple of the first equation from the subsequent equations, and so remove the first term from each of these subsequent equations, without disturbing the various error terms. When this is done, the equations other than the first become

$$2\binom{n}{1} a_n m(y - y_0) = C_{n-1}' + O(P^{-n+1-\varrho'}),$$

$$3!\binom{n}{2} a_n m(y^2 - y_0^2) + 3!\binom{n-1}{1} a_{n-1} m(y - y_0) = C_{n-2}' + O(P^{-n+2-\varrho'}),$$

$$\cdots$$

$$(n-r+2)!\binom{n}{n-r+1} a_n m(y^{n-r+1} - y_0^{n-r+1}) + \ldots$$

$$+ (n-r+2)!\binom{r}{1} a_r m(y - y_0) = C_{r-1}' + O(P^{-r+1-\varrho'}),$$

where $C_{n-1}', \ldots, C_{r-1}'$ are integers. After multiplying the second and subsequent equations here by suitable positive integers, depending only on the position of the equation in the series, a similar situation arises, and we can again remove the first term from the second and subsequent equations, on modifying the integers on the right. Finally, after $n-r+1$ such operations, there remains one equation of the form

$$Da_r m(y - y_0) = C + O(P^{-r+1-\varrho'}),$$

where C is an integer and D is a positive integer depending only on n and r (which can actually be taken to be $(1!2! \ldots (n-r+2)!)^r$, though this is not important).

To estimate the number of values of y satisfying such a relation, for fixed y_0, we appeal to Lemma 8c of Chapter I. The relation is equivalent to

$$\| Dma_r(y - y_0) \| < c' P^{1-r-\varrho'}.$$

We have, by hypothesis,

$$Dma_r = \frac{Dma}{q} + \frac{Dm\theta}{q^2},$$

where

$$(a, q) = 1, \; 0 < q < P^{r+\varrho'}.$$

Also y runs through $Y < P$ consecutive integers, and $r + \varrho' > 1$. Hence the lemma is applicable, with $s = r + \varrho'$ and Dm in place of m, and asserts that the number of values of y satisfying the relation is

$$\ll (m + qP^{1-r-\varrho'})(Pq^{-1} + 1).$$

This proves the required result.

Theorem I. *Let P and m be positive integers. Let*

$$f(x) = a_{n+1}x^{n+1} + \ldots + a_1 x,$$

where $a_{n+1}, \ldots, a_1$ are real, and let

$$S_1 = \sum_{x=1}^{P} e(mf(x)).$$

Let r be one of the numbers $n + 1, \ldots, 2$, and suppose that

$$a_r = \frac{a}{q} + \frac{\theta}{q^2}, \textit{ where } (a, q) = 1 \textit{ and } q > 0.$$

Then

$$S_1 \ll P^{1-\varrho'} m^{2\varrho/\tau}, \tag{29}$$

where

$$\varrho = \frac{\tau}{3n^2 \log (12n(n+1)/\tau)} \textit{ and } \varrho' = \varrho\left(1 + \frac{\nu}{30}\right).$$

Here τ is defined in terms of q and P by

(i) $q = c_1 P^{\tau}$ *for* $1 < q \leqq c_1 P$,

(ii) $\tau = 1$ *for* $c_1 P \leqq q \leqq c_2 P^{r-1}$,

(iii) $q = c_2 P^{r-\tau}$ *for* $c_2 P^{r-1} \leqq q < c_2 P^r$, *where c_1, c_2 are any selected positive constants* (*e.g.* $c_1 = c_2 = 1$). *Thus* $0 < \tau \leqq 1$ *always.*

Proof. We shall apply Lemma 6 to the sum S_1, and then use Lemma 5 in order to estimate the multiple integral on the right of (26). We first determine k as a function of n by putting

$$k = \left[\frac{\log(3n(n+1)/\tau)}{-\log(1-\nu)} + 1\right],$$

which ensures that

$$\sigma = (1-\nu)^k < \frac{\tau}{3n(n+1)}. \tag{30}$$

We note that, since

$$(n-\tfrac{1}{2})^{-1} < -\log(1-\nu) < (n-1)^{-1},$$

we have

$$(n-1)\log\bigl(3n(n+1)/\tau\bigr) < k < (n-\tfrac{1}{2})\log\bigl(3n(n+1)/\tau\bigr) + 1. \tag{31}$$

We apply Lemma 5 to the sum T_1^* of Lemma 6, so that in Lemma 5 we have to replace p_1 by P, M by 0, and take $f(x) = ma_{n+1}x^{n+1} + \alpha_n x^n + \ldots + \alpha_1 x$. We obtain

$$|T_1^*|^{2b(k+h)} \ll \sum_{s_1=2}^{\eta_1} \ldots \sum_{s_k=2}^{\eta_k} \sum^{J} K^*(s_1, \ldots, s_k), \tag{32}$$

where

$$K^*(s_1, \ldots, s_k) = \sum_{x_1, \ldots, x_{n+1}} e(X_1x_1 + \ldots + X_{n+1}x_{n+1}),$$

and the summation here is over a set of points S as described in Lemma 5. For the polynomial just mentioned, we have

$$\begin{aligned}
X_1 &= \alpha_1 + \binom{2}{1}\alpha_2x_0 + \ldots + \binom{n}{1}\alpha_nx_0^{n-1} + \binom{n+1}{1}a_{n+1}mx_0^{n},\\
X_2 &= \alpha_2 + \ldots + \binom{n}{2}\alpha_nx_0^{n-2} + \binom{n+1}{2}a_{n+1}mx_0^{n-1},\\
&\ldots\\
X_n &= \alpha_n + \binom{n+1}{n}a_{n+1}mx_0,\\
X_{n+1} &= ma_{n+1}.
\end{aligned}$$

Thus

$$X_1x_1 + \ldots + X_{n+1}x_{n+1} = A_1\alpha_1 + \ldots + A_n\alpha_n + A_{n+1}ma_{n+1},$$

where $A_1, \ldots, A_{n+1}$ are given by

$$\begin{aligned} A_1 &= x_1, \\ A_2 &= \binom{2}{1} x_0x_1 + x_2, \\ &\ldots \\ A_n &= \binom{n}{1} x_0^{n-1}x_1 + \ldots + x_n, \\ A_{n+1} &= \binom{n+1}{1} x_0^n x_1 + \ldots + \binom{n+1}{n} x_0x_n + x_{n+1}. \end{aligned}$$

The A's are always integers, and $A_1 = \ldots = A_n = 0$ only if $x_1 = \ldots = x_n = 0$.

We have, therefore,

$$\int_0^1 \ldots \int_0^1 K^*(s_1, \ldots, s_k) d\alpha_1 \ldots d\alpha_n$$

$$= \sum_{x_1, \ldots, x_{n+1}} \int_0^1 \ldots \int_0^1 e(A_1\alpha_1 + \ldots + A_n\alpha_n + A_{n+1}ma_{n+1}) d\alpha_1 \ldots d\alpha_n,$$

and by Lemma 4 of Chapter I this does not exceed the number of points $(x_1, \ldots, x_{n+1})$ in the set S for which

$$x_1 = \ldots = x_n = 0.$$

By Lemma 5, this number is $\ll \psi$ (with $p_1 = P$). Thus, by (32),

$$\int_0^1 \ldots \int_0^1 | T_1^* |^{2b(k+h)} d\alpha_1 \ldots d\alpha_n \ll \sum_{s_1=2}^{\eta_1} \ldots \sum_{s_k=1}^{\eta_k} J\psi.$$

Here

$$J\psi = p_1^{2b(k+h)} (p_1 \ldots p_k)^{-\frac{1}{2}(n+1)} 2^{-\lambda(s_1 + \ldots + s_k)},$$

where

$$\begin{aligned} \lambda &= -2n + 2h(b-n) + 2b - \tfrac{1}{2}n(n+1) \\ &= -2n + (2n+4)(b-n) + 2b - \tfrac{1}{2}n(n+1) \\ &= 2(n+3)b - \tfrac{1}{2}n(5n+13) \\ &> 0, \end{aligned}$$

since $b \geqq \frac{1}{4}(5n-1)$. Also

$$p_1 \ldots p_k = p_1^{\,n(1-\sigma)}.$$

Hence

$$J\psi \ll p_1^{\,2b(k+h)-\frac{1}{2}n(n+1)+\frac{1}{2}n(n+1)\sigma}\, 2^{-(s_1+\cdots+s_k)},$$

whence, summing over $s_1, \ldots, s_k$, we have

$$\int_0^1 \cdots \int_0^1 | T_1^* |^{2b(k+h)} d\alpha_1 \ldots d\alpha_n \ll p_1^{\,2b(k+h)-\frac{1}{2}n(n+1)+\frac{1}{2}n(n+1)\sigma}.$$

Applying Lemma 6, we obtain

$$(33) \qquad | S_1 |^{2b(k+h)} \ll H P^{-1+(n+1)\varrho'+2b(k+h)+\frac{1}{2}n(n+1)\sigma} + (P^{1-\varrho'})^{2b(k+h)}.$$

Since $\varrho' > 0$, we have

$$H = (m + qP^{-\tau-\varrho'+1})(Pq^{-1} + 1) \ll \begin{cases} mPq^{-1} \\ m \\ mqP^{-\tau+1} \end{cases}$$

respectively, in the three cases in the enunciation. Hence

$$(34) \qquad H \ll mP^{1-\tau},$$

by the definitions of τ. Further, since $n \geqq 11$,

$$\varrho \leqq \frac{\tau}{3n^2 \log 12n(n+1)} \leqq \frac{\tau}{3n^2\,(2 \log 12 + \log 11)} < \frac{\tau}{22n^2}.$$

Hence

$$\begin{aligned}(n+1)\varrho' = (n+1)\varrho\left(1+\frac{\nu}{30}\right) &< \frac{\tau}{22}\,\nu(1+\nu)\left(1+\frac{\nu}{30}\right) \\ &\leqq \frac{\tau}{22}\,\nu\left(1+\frac{1}{11}\right)\left(1+\frac{1}{330}\right) \\ &< \frac{1}{18}\,\nu\tau.\end{aligned}$$

From these results, and the inequality for σ in (30), we have

$$\begin{aligned}&-\tau + \tfrac{1}{2}n(n+1)\sigma + (n+1)\varrho' \\ &< -\tau + \tfrac{1}{6}\tau + \tfrac{1}{18}\nu\tau = -\tfrac{5}{6}\tau\,(1 - \tfrac{1}{15}\nu).\end{aligned}$$

Hence, using (34) and the last inequality, we deduce from (33)

$$| S_1 |^{2b(k+h)} \ll P^{2b(k+h)-\frac{5}{6}\tau(1-\nu/15)}\, m + (P^{1-\varrho'})^{2b(k+h)}.$$

Thus to establish the desired result (29), we need only prove that

$$(35) \qquad \frac{\frac{5}{6}\tau(1-\frac{1}{15}\nu)}{2b(k+h)} > \varrho' \text{ and } \frac{1}{2b(k+h)} < \frac{2\varrho}{\tau}.$$

Recalling that $h = n+2$ and $b = [\frac{5}{4}n + \frac{1}{2}]$, and using the upper bound for k from (31), we have

$$\begin{aligned} 2b(k+h) &< 2(\tfrac{5}{4}n + \tfrac{1}{2})\Big((n - \tfrac{1}{2}) \log \big(3n(n+1)/\tau\big) + 1 + n + 2\Big) \\ &= \tfrac{5}{2}(n + \tfrac{2}{5})(n - \tfrac{1}{2}) \left(\log \big(3n(n+1)/\tau\big) + \frac{n+3}{n-\frac{1}{2}}\right) \\ &< \tfrac{5}{2}(1 + \tfrac{2}{5}\nu)(1 - \tfrac{1}{2}\nu)n^2 \log \big(12n(n+1)/\tau\big) \\ &= \tfrac{5}{2}(1 + \tfrac{2}{5}\nu)(1 - \tfrac{1}{2}\nu)\frac{\tau}{3\varrho}, \end{aligned}$$

since

$$\frac{n+3}{n-\frac{1}{2}} \leqq \frac{11+3}{11-\frac{1}{2}} = \frac{4}{3} < \log 4.$$

Thus to prove the first inequality in (35), it is enough to show that

$$(1 - \tfrac{1}{15}\nu)(1 + \tfrac{2}{5}\nu)^{-1}(1 - \tfrac{1}{2}\nu)^{-1} > 1 + \tfrac{1}{30}\nu,$$

and in fact we have

$$(1 - \tfrac{1}{15}\nu) - (1 + \tfrac{1}{30}\nu)(1 + \tfrac{2}{5}\nu)(1 - \tfrac{1}{2}\nu) = (\tfrac{1}{5} + \tfrac{1}{300})\nu^2 + \tfrac{1}{150}\nu^3 > 0.$$

To prove the second of the inequalities (35), we have to show that

$$2b(k+h) > \tfrac{3}{2}n^2 \log \big(12n(n+1)/\tau\big).$$

By (31),

$$2b(k+h) > 2(\tfrac{5}{4}n - \tfrac{1}{4})\big((n-1) \log (3n(n+1)/\tau) + n + 2\big).$$

Hence it suffices if

$$2(\tfrac{5}{4}n - \tfrac{1}{4})(n-1) > \tfrac{3}{2}n^2 \text{ and } 2(\tfrac{5}{4}n - \tfrac{1}{4})(n+2) > \tfrac{3}{2}n^2 \log 4.$$

Both these are satisfied for $n \geqq 11$ (the second since $\log 4 < \frac{5}{3}$) and the proof of Theorem I is now complete.

Example. Suppose that in the polynomial $f(x)$ one of the coefficients $a_{n+1}, \ldots, a_2$ is $\sqrt{2}$. Then, since the continued fraction for $\sqrt{2}$ has bounded partial quotients, we can satisfy the requirement that

$$\sqrt{2} = \frac{a}{q} + \frac{\theta}{q^2}$$

with a value of q with $c_1 P < q < P$, where c_1 is a suitable positive constant. Thus we have the second of the three cases in the enunciation, and $\tau = 1$. The theorem implies that

$$\sum_{x=1}^{P} e\big(mf(x)\big) \ll P^{1-\varrho(1+\nu/30)} m^{2\varrho},$$

where

$$\varrho = \frac{1}{3n^2 \log \big(12n(n+1)\big)}.$$

LEMMA 7. *Let N and P be integers, P being large and positive. Let $F(x)$ be a real function defined for $N \leqq x \leqq N + P$, and having in this interval a continuous $(n+1)$th derivative satisfying*

$$\frac{1}{A_0} \leqq \left| \frac{F^{(n+1)}(x)}{(n+1)!} \right| \leqq \frac{c'}{A_0},$$

where

$$P \ll A_0 \ll P^{2+2\nu}.$$

Let

$$S_2 = \sum_{x=N+1}^{N+P} e\big(F(x)\big).$$

Let ϱ satisfy $0 < \varrho < 1$, and let

$$p_1 = [A_0^{(1-\varrho)/(n+1)}].$$

Then

$$(36)\quad \begin{cases} S_2^{2b(k+h)} \\ \ll p_1^{-2b(k+h)+\frac{1}{2}n(n+1)} P^{2b(k+h)-1+2\nu+(n+1)\varrho} \int_0^1 \cdots \int_0^1 | T_2^* |^{2b(k+h)} \, d\alpha_1 \ldots d\alpha_n \\ \quad + (P^{1-\varrho})^{2b(k+h)}, \end{cases}$$

where

$$T_2^*(\alpha_1, \ldots, \alpha_n) = \sum_{x=1}^{P} e(\alpha_1 x + \ldots + \alpha_n x^n).$$

Proof. In principle, the argument is similar to that in the proof of Lemma 6. Let $Y = [P^{1-\varrho}]$. Define $S_2^*(y)$ by

$$S_2^*(y) = \sum_{x=1}^{p_1} e(F(y+x) - F(y)).$$

Then, noting that $0 < p_1 < P$, we have

$$S_2 = p_1^{-1} \sum_{y=N}^{N+P-p_1} \sum_{x=1}^{p_1} e(F(y+x)) + O(p_1),$$

and therefore

$$| S_2 | \leqq p_1^{-1} \sum_{y=N}^{N+P-p_1} | S_2^*(y) | + O(p_1).$$

Hence, by Hölder's inequality,

$$(37) \begin{cases} | S_2 |^{2b(k+h)} \\ \ll p_1^{-2b(k+h)} P^{2b(k+h)-1} \sum_{y=N}^{N+P-p_1} | S_2^*(y) |^{2b(k+h)} + (P^{1-\varrho})^{2b(k+h)}, \end{cases}$$

on noting that $p_1 < P^{1-\varrho}$.

By Taylor's theorem, in view of the condition on $F^{(n+1)}(x)$, we have

$$F(y+x) - F(y) = Y_1 x + \ldots + Y_n x^n + \theta \frac{c'}{A_0} p_1^{n+1}$$

for $1 \leq x \leq p_1$, where

$$Y_j = Y_j(y) = \frac{1}{j!} F^{(j)}(y).$$

We note that

$$\frac{c'}{A_0} p_1^{n+1} \ll A_0^{-1} A_0^{1-\varrho} = A_0^{-\varrho} \ll P^{-\varrho},$$

so that

$$e\left(\theta \frac{c'}{A_0} p_1^{n+1}\right) = 1 + O(P^{-\varrho}).$$

Define the region Ω_y in n dimensional space to consist of the points $(\alpha_1, \ldots, \alpha_n)$ satisfying

$$|\alpha_1 - Y_1| \leqq \tfrac{1}{2}p_1^{-1}P^{-\varrho}, \ldots, |\alpha_n - Y_n| \leqq \tfrac{1}{2}p_1^{-n}P^{-\varrho}.$$

Then for any point $(\alpha_1, \ldots, \alpha_n)$ in Ω_y we have

$$S_2^*(y) = T_2^*(\alpha_1, \ldots, \alpha_n) + O(p_1 P^{-\varrho}),$$

whence

$$|S_2^*(y)|^{2b(k+h)} \ll |T_2^*|^{2b(k+h)} + (p_1 P^{-\varrho})^{2b(k+h)}.$$

Integrating over Ω_y, the volume of which is $p_1^{-\frac{1}{2}n(n+1)}P^{-n\varrho}$, we obtain

$$|S_2^*(y)|^{2b(k+h)}$$
$$\ll p_1^{\frac{1}{2}n(n+1)}P^{n\varrho}\int_0^1 \ldots \int_0^1 |T_2^*|^{2b(k+h)}\tau(\Omega_y)d\alpha_1\ldots d\alpha_n + (p_1P^{-\varrho})^{2b(k+h)},$$

$\tau(\Omega_y)$ being the characteristic function of Ω_y (mod 1). Substituting in (37), we see that the desired result (36) follows, provided we can prove that

$$\sum_{y=N}^{N+P-p_1} \tau(\Omega_y) \ll P^{2\nu+\varrho}$$

for each point $(\alpha_1, \ldots, \alpha_n)$.

Let Ω_y and Ω_{y_0} be two regions each containing a point whose coordinates differ from $\alpha_1, \ldots, \alpha_n$ by integral amounts. Taking only the nth coordinates, it follows that

$$\| Y_n(y) - Y_n(y_0) \| \leqq p_1^{-n}P^{-\varrho}.$$

Now

$$p_1^{-n}P^{-\varrho} \ll P^{2\nu+\varrho}A_0^{-1},$$

for

$$A_0 p_1^{-n} \ll A_0^{1-n(1-\varrho)/(n+1)} \ll (P^{2+2\nu})^{1-n(1-\varrho)/(n+1)}$$

and

$$(2+2\nu)\left(\frac{1+n\varrho}{n+1}\right) = 2\nu + 2\varrho.$$

Thus

$$\| Y_n(y) - Y_n(y_0) \| \ll P^{2\nu+\varrho}A_0^{-1}.$$

We have to estimate the number of integers y satisfying this inequality, with $N \leqq y \leqq N + P - p_1$.

Put $\Phi(y) = Y_n(y) - Y_n(y_0)$. Then, for $N \leqq y < N + P - p_1$,

$$\Phi(y+1) - \Phi(y) = Y_n'(y_1) = \frac{1}{n!} F^{(n+1)}(y_1)$$

for some y_1 with $N \leqq y_1 \leqq N + P - p_1$. Hence (taking $F^{(n+1)}(y)$ to be positive without loss of generality)

$$\frac{1}{A} \leqq \Phi(y+1) - \Phi(y) \leqq \frac{\beta}{A},$$

where $A = A_0/(n+1)$, $\beta = c'$. The function $\Phi(y)$ satisfies the hypotheses of Lemma 9 of Chapter I, with $Y = P - p_1 + 1$, $W \ll P^{2\nu+\varrho}$ (and a different value of N). Hence, by (II) of that lemma, the number of values of y is

$$\ll (PA_0^{-1} + 1)W \ll W \ll P^{2\nu+\varrho}.$$

This proves the desired result.

THEOREM 2a. *Let N and P be integers, P being large and positive. Let $F(x)$ be a real function, defined for $N \leqq x \leqq N + P$, and having in that interval a continuous $(n+1)$th derivative, satisfying*

$$\frac{1}{A_0} \leqq \left| \frac{F^{(n+1)}(x)}{(n+1)!} \right| \leqq \frac{c'}{A_0},$$

where

$$P \ll A_0 \ll P^{2+2\nu}.$$

Let

$$S_2 = \sum_{x=N+1}^{N+P} e(F(x)).$$

Then

$$S_2 \ll P^{1-\varrho}, \textit{ where } \varrho = \frac{1}{3n^2 \log(125n)}. \tag{38}$$

Proof. This is similar to that of Theorem I, but uses Lemma 7 in place of Lemma 6. Put

$$k = \left[\frac{\log(6n+6)}{-\log(1-\nu)} + 1 \right],$$

so that

(39) $$\sigma = (1-\nu)^k < (6n+6)^{-1}.$$

We have

(40) $$k < (n-\tfrac{1}{2})\log(6n+6)+1.$$

Applying Lemma 5 to the sum $T_2^*(\alpha_1, \ldots, \alpha_n)$ of Lemma 7, i.e. taking $a_1 = \alpha_1, \ldots, a_n = \alpha_n$, $a_{n+1} = 0$ in the $f(x)$ of Lemma 5, we obtain

$$|T_2^*|^{2b(k+h)} \ll \sum_{s_1=2}^{\eta_1} \ldots \sum_{s_k=2}^{\eta_k} \sum^{J} K(s_1, \ldots, s_k),$$

where

$$K(s_1, \ldots, s_k) = \sum_{x_1, \ldots, x_{n+1}} e(X_1 x_1 + \ldots + X_n x_n),$$

and the summations are as explained in Lemma 5.

In view of the present form of $f(x)$, we have

$$X_1 = \alpha_1 + \binom{2}{1}\alpha_2 x_0 + \ldots + \binom{n}{1}\alpha_n x_0^{n-1},$$

$$X_2 = \qquad \alpha_2 + \ldots + \binom{n}{2}\alpha_n x_0^{n-2},$$

$$\ldots$$

$$X_n = \qquad \alpha_n.$$

We note that (as before)

$$X_1 x_1 + \ldots + X_n x_n = A_1\alpha_1 + \ldots + A_n\alpha_n,$$

where $A_1, \ldots, A_n$ are integers and are all 0 if and only if $x_1, \ldots, x_n$ are all 0. Hence

$$\int_0^1 \ldots \int_0^1 K(s_1, \ldots, s_k)\,d\alpha_1 \ldots d\alpha_n \ll \psi.$$

It follows that

$$\int_0^1 \ldots \int_0^1 |T_2^*|^{2b(k+h)} d\alpha_1 \ldots d\alpha_n$$
$$\ll \sum_{s_1=2}^{\eta_1} \ldots \sum_{s_k=2}^{\eta_k} J\psi$$
$$\ll p_1^{2b(k+h)-\frac{1}{2}n(n+1)+\frac{1}{2}n(n+1)\sigma},$$

as in the proof of Theorem I. Applying this result in Lemma 7, we obtain

$$S_2^{2b(k+h)} \ll p_1^{\frac{1}{2}n(n+1)\sigma} P^{2b(k+h)-1+2\nu+(n+1)\varrho} + (P^{1-\varrho})^{2b(k+h)}.$$

Recalling that $p_1 = [A_0^{(1-\varrho)/(n+1)}]$ and $A_0 \ll P^{2+2\nu}$, we have by (39),

$$p_1^{\frac{1}{2}n(n+1)\sigma} \ll p_1^{n/12} \ll A_0^{\frac{1}{12}n/(n+1)} \ll P^{1/6}.$$

Further,

$$(n+1)\varrho = \frac{\nu(\nu+1)}{3\log(125n)} < \frac{\nu(1+\frac{1}{11})}{3(3\log 11)} < \tfrac{1}{12}\nu.$$

Thus the preceding estimate for S_2 implies

$$S_2^{2b(k+h)} \ll P^{2b(k+h)-\frac{5}{6}+\frac{25}{12}\nu} + (P^{1-\varrho})^{2b(k+h)}.$$

Comparing this with the desired result (38), we see that it remains only to prove that

$$\tfrac{5}{6} - \tfrac{25}{12}\nu > 2b(k+h)\varrho.$$

Using (40), we have

$$\frac{\frac{5}{6}-\frac{25}{12}\nu}{2b(k+h)} \geqq \frac{\frac{5}{6}(1-\frac{5}{2}\nu)}{2(\frac{5}{4}n+\frac{1}{2})(n-\frac{1}{2})(\log(6n+6)+\frac{4}{3})}$$
$$> \frac{1}{3n^2(1+3.1\nu)(\log n + 3.22)},$$

since

$$(1-2.5\nu)(1+3.1\nu) - (1+\tfrac{2}{5}\nu)(1-\tfrac{1}{2}\nu) = 0.7\nu - 7.55\nu^2 > 0$$

for $n \geqq 11$, and

$$\log(6n+6) \leqq \log n + \log \tfrac{72}{11} < \log n + 1.88.$$

Also

$$\begin{aligned} &\log(125n) - (1+3.1\nu)(\log n + 3.22)\\ &= 1.60\ldots - (3.1\nu)(\log n + 3.22)\\ &> 1.60 - 3.1(\log 11 + 3.22)/11 > 0. \end{aligned}$$

Hence

$$\frac{\frac{5}{6}-\frac{25}{12}\nu}{2b(k+h)} > \frac{1}{3n^2\log(125n)} = \varrho.$$

THEOREM 2*b*. *Let N and P be integers with P large and positive. Let $F(x)$ satisfy the hypotheses of Theorem 2a, but with the stronger condition*

$$P \ll A_0 \ll P^{2+\nu}$$

on A_0. Let $\Phi(x)$ be monotonic for $N \leqq x \leqq N+P$, and suppose

$$\max |\Phi(x)| \ll \Phi_0.$$

Then, if $1 \leqq P_1 \leqq P$ and

$$S(P_1) = \sum_{x=N+1}^{N+P_1} \Phi(x) e(F(x)),$$

we have

$$S(P_1) \ll \Phi_0 P^{1-\varrho}, \text{ where } \varrho = \frac{1}{3n^2 \log(125n)}.$$

Proof. Let

$$S_2(P_0) = \sum_{x=N+1}^{N+P_0} e(F(x))$$

for integers P_0 satisfying $1 \leqq P_0 \leqq P_1$. If $P_0 \geqq A_0^{1/(2+2\nu)}$, the hypotheses of Theorem 2a are satisfied for $S_2(P_0)$, with P_0 in place of P. Hence, subject to this condition, we have

$$S_2(P_0) \ll P_0^{1-\varrho} \ll P^{1-\varrho}.$$

If $P_0 < A_0^{1/(2+2\nu)}$, the same result follows from the trivial estimate

$$S_2(P_0) \ll P_0 \ll P^{(2+\nu)/(2+2\nu)} \ll P^{1-\varrho},$$

since $\varrho < \nu/(2+2\nu)$.

The result now follows on expressing $S(P_1)$ in terms of the sums $S_2(P_0)$ for $1 \leqq P_0 \leqq P_1$, and applying partial summation.

Example. To illustrate Theorem 2b, consider the sum

$$S = \sum_{x=P+1}^{P+P_1} x^{-s},$$

where $s = \sigma + it$, $\sigma > 0$, $t > 1$, and

$$1 \leqq P_1 \leqq P, \ t^{1/n} \leqq P \leqq t^{1/(n-1)}.$$

We have $\Phi(x) = x^{-\sigma}$ and

$$2\pi F(x) = -t \log x.$$

Hence

$$\left| \frac{F^{(n+1)}(x)}{(n+1)!} \right| = \frac{t}{2\pi(n+1)x^{n+1}}.$$

The hypotheses of Theorem 2b are satisfied if we take

$$A_0 = 2\pi(n+1)(2P)^{n+1}t^{-1},\ c' = 2^{n+1},$$

since then $P \ll A_0 \ll P^2$. Also we can take $\Phi_0 = P^{-\sigma}$. Hence, if $n \geq 11$, we have

$$S \ll P^{1-\sigma-\varrho},\ \text{where}\ \varrho = \frac{1}{3n^2 \log (125n)}.$$

NOTES ON CHAPTER VI

This chapter has been considerably expanded, and the different steps towards the proofs of Theorems 1 and 2a have been separated into lemmas. The language of "b dimensional boxes" has been introduced to facilitate the discussion of multiple sums, and to enable us to make clear a number of points which are not explicit in the original.

The main ideas in the proof of Theorem 1 may be summarized as follows. Lemma 6 shows that the estimation of a Weyl sum can be made to depend on the estimation of

$$\int_0^1 \cdots \int_0^1 |T_1^*|^{2s}\, d\alpha_1 \ldots d\alpha_n = N^*(P, s),\ \text{say}.$$

This estimation is required to be of the form

$$N^*(P, s) \ll P^{2s-\frac{1}{2}n(n+1)+\delta},$$

where δ is in a certain sense sufficiently small [1]).

The above multiple integral does not exceed the corresponding integral $N(P, s)$, formed with the term $m\alpha_{n+1}x^{n+1}$ omitted from

[1]) The precision of Theorem 1 will depend on being able to choose δ sufficiently small without choosing s too large.

T_1^* (i.e. taking $a_{n+1} = 0$), and this integral has a simple arithmetical interpretation: it is the number of solutions of

$$x_1 + \ldots + x_s = y_1 + \ldots + y_s,$$
$$\ldots$$
$$x_1^n + \ldots + x_s^n = y_1^n + \ldots + y_s^n,$$

in integers satisfying $1 \leqq x_i \leqq P$, $1 \leqq y_i \leqq P$. An estimate for this number of solutions gives a measure of the regularity of distribution of the power sums on one side of the equations. So far it has not proved possible to obtain an effective estimate by a direct argument.

The treatment applied to obtain an estimate for $N^*(P, s)$, and in effect for $N(P, s)$, is to majorize $| T_1^* |^{2s}$ by sums of the type K of Lemma 5. The variables $x_1, \ldots, x_{n+1}$ in these sums K arise as power sums, and the construction is such as to ensure that these have the regularity of distribution expressed by the final clause of Lemma 5. This in turn gives an estimate for the multiple integral corresponding to each K.

Lemmas 1 to 4 represent steps towards Lemma 5. The tools for the production of well-distributed power-sums are found in Lemmas 16 and 15 of Chapter I. Lemma 1a represents the application of Lemma 16 of Chapter I to multiple exponential sums of a special kind; and Lemma 15 of Chapter I is applicable to certain products of such sums in the manner seen in the proof of Lemma 5.

Considerable difficulty arises, however, from the fact that Lemma 1a is applicable only to sums of a rather special kind: those over proper boxes. This difficulty is surmounted by the use of Lemma 3, which shows how a box can be subdivided into boxes, some proper and some improper, in such a way that the total volume of the improper boxes is very small relative to that of the original box. When this subdivision is applied to sums, the contribution of the sums extended over improper boxes is small.

Lemma 4 prepares the way for the production of sums of the type (17) in the proof of Lemma 5. To each such sum Vinogradov then applies Lemma 15 of Chapter I and Lemma 1a of Chapter VI. The essential idea in the latter is, of course, Lemma 16 of Chapter I.

It has been shown by Hua (*Quart. J. of Math.* (Oxford), 20 (1949), 48—61) that Lemma 5 is not essential for the estimation of $N(P, s)$, and consequently not essential for the results of this Chapter. The main innovation is that the integration over the unit cube is carried out at an earlier stage. We proceed to outline what the effect of this idea would be if we conserved the present notation and started from Lemma 4.

Applying Lemma 4 with $f(x)$ replaced by $f^*(x) = \alpha_n x^n + \ldots + \alpha_1 x$ (i.e. writing $a_{n+1} = 0$, $a_n = \alpha_n, \ldots, a_1 = \alpha_1$), we obtain on integrating each side of (10) over the unit hypercube $0 \leqq \alpha_1 \leqq 1, \ldots, 0 \leqq \alpha_n \leqq 1$,

$$N(p_t, 2b(k+h-t+1))$$

$$\ll \sum_{s=2}^{\eta_t} \sum^{M(t,s)} \int_0^1 \cdots \int_0^1 |Z(2^{-s}p_t)|^2 |L(p_{t+1})|^{2b(k+h-t)} d\alpha_1 \ldots d\alpha_n.$$

Now

$$\int_0^1 \cdots \int_0^1 |Z(2^{-s}p_t)|^2 |L(p_{t+1})|^{2b(k+h-t)} d\alpha_1 \ldots d\alpha_n$$

is the number of solutions of

$$x_1' + \ldots + x_b' + x_1 + \ldots + x_l = y_1' + \ldots + y_b' + y_1 + \ldots + y_l,$$
$$\ldots$$
$$x_1'^n + \ldots + x_b'^n + x_1^n + \ldots + x_l^n = y_1'^n + \ldots + y_b'^n + y_1^n + \ldots + y_l^n,$$

where $l = 2b(k+h-t)$ and where $(x_1', \ldots, x_b')$ and $(y_1', \ldots, y_b')$ lie in $B(2^{-s}p_t)$ and the variables x and y lie in the interval of summation of $L(p_{t+1})$. In view of the invariance of the above equations under a simultaneous translation of all the variables, we may assume that $1 \leqq x_j \leqq p_{t+1}$ and $1 \leqq y_j \leqq p_{t+1}$ $(j = 1, \ldots, l)$. For any given values of $x_1', \ldots, x_b', y_1', \ldots, y_b'$, the equations are of the form

$$x_1 + \ldots + x_l = y_1 + \ldots + y_l + C_1,$$
$$\ldots$$
$$x_1^n + \ldots + x_l^n = y_1^n + \ldots + y_l^n + C_n.$$

By Cauchy's inequality, the number of solutions of these equations in $x_1, \ldots, x_l, y_1, \ldots, y_l$ does not exceed the number of solutions of

$$x_1 + \ldots + x_l = y_1 + \ldots + y_l,$$
$$\ldots$$
$$x_1^n + \ldots + x_l^n = y_1^n + \ldots + y_l^n,$$

and this number is $N(p_{t+1}, l)$.

Now the number of possibilities for the variables of the type x', y' can be estimated directly by means of Lemma 16 of Chapter I applied to n of these $2b$ variables (unless $s = \eta$, in which case the trivial estimate must be taken). Thus we obtain an estimate for $N(p_t, 2b(k + h - t + 1))$ in terms of $N(p_{t+1}, 2b(k+h-t))$. Successive application of this reduction formula for $t = 1, \ldots, k$ leads to an estimation of $N(P, s)$.

Actually Hua's method differs from that of this chapter in some other respects (although the ideas are similar) and the inequalities he obtains are slightly more precise than those of this chapter. In place of Lemma 4 of this chapter, Hua uses a result which is very similar but more appropriate to his method. He obtains, by the argument outlined above, the reduction formula

$$N(P, s) \ll (\log P)^2 \, P^{2n-\frac{1}{2}(n+1)+2(s-n)/n} \, N(P, s-n)$$

for $s \geqq \frac{1}{4}n(n+1) + n$. Repeated application of this reduction formula shows that for any integers s, l satisfying $s \geqq \frac{1}{4}n(n+1) + nl$ we have

$$N(P, s) \ll (\log P)^{2l} P^{2s-\frac{1}{2}n(n+1)+\delta},$$

where

$$\delta = \tfrac{1}{2}n(n+1)(1-\nu)^l.$$

This estimate is of the kind sought in connection with Weyl sums, and leads to slightly more precise results than the corresponding estimate of this chapter.

CHAPTER VII

The Asymptotic Formula in Waring's Problem

Let $W(N)$ denote the number of representations of a positive integer N in the form

$$N = x_1{}^n + \ldots + x_r{}^n,$$

where $x_1, \ldots, x_r$ are positive integers. Hardy and Littlewood gave an asymptotic formula for $W(N)$ as $N \to \infty$, and established its validity when n and r are fixed and $r \geqq (n-2)2^{n-1} + 5$. The main term in the asymptotic formula is

$$\frac{(\Gamma(1+\nu))^r}{\Gamma(r\nu)} N^{r\nu-1} \mathfrak{S}(N, r),$$

an expression which we have already encountered in Chapter III. We recall that $\mathfrak{S}(N, r) \gg 1$ if $r \geqq 4n$, by Lemma 12 of Chapter II. In order that the asymptotic formula should be valid, one needs an estimate of lower order than $N^{r\nu-1}$ for the error term. As mentioned in the Introduction, such an estimate was found by Hardy and Littlewood on the supposition that r satisfies the inequality stated above.

The object of the present chapter is to prove that, provided $n \geqq 12$, the asymptotic formula of Hardy and Littlewood for $W(N)$ is valid for

$$r \geqq [10\, n^2 \log n].$$

The proof is based on the improved estimates for Weyl sums found in the preceding chapter.

Notation in this chapter. We suppose $n \geqq 12$ and retain the symbols b, h, σ with the same meanings as in the preceding chapter, namely

$$b = [\tfrac{5}{4}n + \tfrac{1}{2}], \quad h = n + 2, \quad \sigma = (1-\nu)^k.$$

We denote by k a positive integer, depending only on n, to which a value will be assigned later.

LEMMA 1. *Let $p_1 > 1$ be an integer. Let $a_n, \ldots, a_1$ be fixed positive integers with $a_n > 0$, and let*

$$f(x) = a_n x^n + \ldots + a_1 x.$$

Let

$$T_1(\alpha) = \sum_{x=1}^{p_1} e(\alpha f(x)).$$

Then

$$\int_0^1 | T_1(\alpha) |^{2b(k+h)} d\alpha \ll p_1^{2b(k+h)-n+\frac{1}{2}n(n+1)\sigma}.$$

Proof. We use Lemma 5 of Chapter VI with $a_{n+1} = 0$. This gives

$$T_1^{2b(k+h)} \ll \sum_{s_1=2}^{\eta_1} \ldots \sum_{s_k=2}^{\eta_k} \sum^{J} K(s_1, \ldots, s_k; \alpha),$$

where

$$K(s_1, \ldots, s_k; \alpha) = \sum_{x_1} \ldots \sum_{x_n} \psi(x_1, \ldots, x_n) e(\alpha(x_1 X_1 + \ldots + x_n X_n)),$$

the summation being over

$$-cp_1 \leqq x_1 \leqq cp_1, \ldots, -cp_1^n \leqq x_n \leqq cp_1^n.$$

We also have the estimate $\psi(x_1, \ldots, x_n) \ll \psi$, where

$$J\psi \ll p_1^{2b(k+h)-\frac{1}{2}n(n+1)+\frac{1}{2}n(n+1)\sigma} 2^{-(s_1+\cdots+s_k)}.$$

We recall that $X_j = f^{(j)}(x_0)/j!$, where x_0 is some integer. All the X_j are integers, and $X_n = a_n > 0$.

Now

$$\int_0^1 e(\alpha(x_1 X_1 + \ldots + x_n X_n)) d\alpha$$

is 1 if

$$x_1 X_1 + \ldots + x_n X_n = 0 \tag{1}$$

and is 0 otherwise. When $x_1, \ldots, x_{n-1}$ are given, the equation (1) admits at most one value for x_n. Hence the number of possible sets of values of $x_1, \ldots, x_n$ which satisfy (1) is

$$\ll p_1 p_1^2 \ldots p_1^{n-1} = p_1^{\frac{1}{2}n(n-1)}.$$

It follows that

$$\int_0^1 K(s_1, \ldots, s_k;\ \alpha)\,d\alpha \ll p_1^{\frac{1}{2}n(n-1)}\psi,$$

whence

$$\int_0^1 |\,T_1(\alpha)\,|^{2b(k+h)} \ll \sum_{s_1=2}^{\eta_1} \ldots \sum_{s_k=2}^{\eta_k} p_1^{\frac{1}{2}n(n-1)} J\psi$$

$$\ll p_1^{2b(k+h)-n+\frac{1}{2}n(n+1)\sigma} \sum_{s_1=2}^{\eta_1} \ldots \sum_{s_k=2}^{\eta_k} 2^{-s_1-\cdots-s_k}$$

$$\ll p_1^{2b(k+h)-n+\frac{1}{2}n(n+1)\sigma}.$$

LEMMA 2. *Let*

(2) $$\alpha = \frac{a}{q} + \frac{\theta}{q^2},\ \textit{where}\ (a, q) = 1,\ p_1^{1-\nu} < q \leq 2np_1^{n-1}.$$

Let p_1 be a positive integer, and let

$$T_1 = \sum_{x=1}^{p_1} e(\alpha x^n).$$

Then

$$T_1 \ll p_1^{1-\varrho},$$

where

(3) $$\varrho = (3n(n-1)\log(12n^2))^{-1}.$$

Proof. The result is an immediate deduction from Theorem I of Chapter VI. In that theorem we take m to be 1, and replace $n+1$ throughout by n. We take $f(x) = \alpha x^n$ and $r = n$. Because of the restrictions on q in (2), we have either the first or the second of the three cases in the theorem. In the second case the parameter τ is 1. In the first case the parameter τ is $\geqq 1 - \nu$, since $q > p_1^{1-\nu}$. Also ϱ' can be replaced by the smaller number ϱ. Hence the conclusion follows with

$$\varrho = \frac{1-\nu}{3(n-1)^2 \log\left(12n(n-1)(1-\nu)^{-1}\right)}$$

$$= \left(3n(n-1)\log(12n^2)\right)^{-1}.$$

THEOREM. *If $r \geqq [10n^2 \log n]$, the number $W(N)$ of representations of N as the sum of r nth powers of positive integers satisfies*

$$W(N) = \frac{(\Gamma(1+\nu))^r}{\Gamma(r\nu)} N^{r\nu-1} \mathfrak{S}(N, r) + O(N^{r\nu-1-\nu^2}).$$

Proof. Let $p_1 = [N^\nu]$ and $\tau = 2np_1^{n-1}$. Let

$$T_1(\alpha) = \sum_{x=1}^{p_1} e(\alpha x^n).$$

Then, as in Chapter III, but with $T_1(\alpha)$ for $L(\alpha)$ and p_1 for P,

$$W(N) = \int_{-1/\tau}^{1-1/\tau} (T_1(\alpha))^r e(-N\alpha) d\alpha. \tag{4}$$

We divide the interval of integration into basic intervals and supplementary intervals. The basic intervals comprise all α of the form

$$\alpha = \frac{a}{q} + z, \text{ where } (a, q) = 1,\ 0 < q \leqq p_1^{1-\nu},\ |z| \leqq \frac{1}{q\tau}.$$

This is the same definition as in Chapter III, except that β has now the special value $1 - \nu$ (and p_1 replaces P). Any α in a supplementary interval is representable (not necessarily uniquely) as

$$\alpha = \frac{a}{q} + z, \text{ where } (a, q) = 1,\ p_1^{1-\nu} < q \leqq \tau,\ |z| \leqq \frac{1}{q\tau}. \tag{5}$$

We have

$$W(N) = W^*(N) + W^{**}(N),$$

where $W^*(N)$ is the contribution of the basic intervals to the integral (4) and $W^{**}(N)$ the contribution of the supplementary intervals.

Since $r \geqq [10n^2 \log n] > 2n + 1$, we can appeal to the results of Chapter III for the contribution of the basic intervals. By Lemma 4 of that Chapter, we have

$$W^*(N) = \frac{(\Gamma(1+\nu))^r}{\Gamma(r\nu)} N^{r\nu-1} \sum_{0<q\leqq p_1^{1-\nu}} A(q, N, r) + O(N^{r\nu-1-\nu^2}).$$

By Lemma 6 of Chapter II, noting that $r \geqq 2n + 2$, we have

$$\sum_{q > p_1^{1-\nu}} A(q, N, r) \ll \sum_{q_1 > p_1^{1-\nu}} q^{1-(2n+2)\nu} \ll p_1^{-2\nu(1-\nu)} \ll N^{-\nu^2}.$$

Hence

$$W^*(N) = \frac{(\Gamma(1+\nu))^r}{\Gamma(r\nu)} N^{r\nu-1} \mathfrak{S}(N, r) + O(N^{r\nu-1-\nu^2}).$$

We have now to estimate the contribution $W^{**}(N)$ of the supplementary intervals to the integral (4). By Lemma 1, with $f(x) = x^n$, we have

$$\int_0^1 | T_1(\alpha) |^{2b(k+h)} d\alpha \ll p_1^{2b(k+h)-n+\frac{1}{2}n(n+1)\sigma}.$$

Further, any α in a supplementary interval is representable in the form (5). It follows that the hypotheses of Lemma 2 are satisfied, and therefore if $r \geqq 2b(k + h)$ we have

$$| T_1(\alpha) |^{r-2b(k+h)} \ll p_1^{(r-2b(k+h))(1-\varrho)}$$

in the supplementary intervals, where ϱ is given by (3). Thus

$$W^{**}(N) \ll p_1^{r-n-\lambda},$$

where

$$\lambda = -\tfrac{1}{2}n(n+1)\sigma + (r - 2b(k+h))\varrho.$$

It remains only to prove that $W^{**}(N) \ll N^{r\nu-1-\nu^2}$, that is, that $\lambda \geqq \nu$. The number k is still at our disposal, and we take

$$k = \left[\frac{\log(0.6n^2 \log 12n^2)}{-\log(1-\nu)} + 1\right],$$

thus ensuring that

$$\sigma \leqq (0.6n^2 \log 12n^2)^{-1}.$$

The function $\log(0.6 \log 12n^2) - 0.8 \log n$ is decreasing for $n \geqq 12$ since its derivative is

$$\frac{1}{n}\left(\frac{2}{\log 12 + 2\log n} - 0.8\right),$$

and $\log 12 + 2 \log n \geqq 3 \log 12 > 7$. The value of the function when $n = 12$ is about -0.5. Hence $\log (0.6 \log 12n^2) < 0.8 \log n$, and therefore

$$k < \frac{1 - \frac{1}{2}\nu}{\nu} (2.8 \log n) + 1.$$

Also

$$2b \leqq \tfrac{5}{2}n + 1,\ h = n + 2,\ r \geqq 10n^2 \log n - 1.$$

Hence

$$\lambda > -\frac{n(n+1)}{1.2n^2 \log 12n^2} + \frac{10n^2 \log n - 1 - (\frac{5}{2}n + 1)((n - \frac{1}{2})(2.8 \log n) + n + 3)}{3n(n-1) \log 12n^2},$$

or

$$\begin{aligned}(6n(n-1) \log 12n^2)\lambda &> -5(n^2 - 1) + 20n^2 \log n - 2 \\ &\qquad - (5n + 2)((n - \tfrac{1}{2})(2.8 \log n) + n + 3) \\ &= (6n^2 + 1.4n + 2.8) \log n - (10n^2 + 17n + 3).\end{aligned}$$

Since $n \geqq 12$ and $(\log 12)^{-1} < 0.41$, the last expression is

$$> (1.9n^2 - 5.6n + 1.5) \log n > 1.5n(n-1)\log n.$$

On the other hand, $6n(n-1) \log 12n^2 \leqq 18n(n-1)\log n$. Hence

$$\lambda > \frac{1.5}{18} = \frac{1}{12} \geqq \nu.$$

This proves the result.

NOTES ON CHAPTER VII

This closely follows the original. The reader will recall (see Notes on Chapter IV) that it is necessary to save more than $p_1{}^n$ on the integral for the number of representations. The saving effected by Lemma 1, on $2b(k+h)$ nth powers, is $n - \frac{1}{2}n(n+1)\sigma$. The choice of k is such that $\frac{1}{2}n(n+1)\sigma$ is about $\frac{1}{2.4 \log n}$, and then $2b(k+h)$ is about $5n^2 \log n$. The rest of the necessary saving is effected by the inequality of Lemma 2, which saves about $\frac{1}{6n^2 \log n}$ on each additional nth power. It will be seen that the constant 10 in the final result is of no special significance (for large n).

CHAPTER VIII

The Distribution of the Fractional Parts of the Values of a Polynomial

In the present chapter, the estimate found for Weyl sums in Theorem I of Chapter VI is applied to the proof of an asymptotic formula for the number of fractions in the sequence

$$\{f(x)\}, \text{ where } x = 1, \ldots, P,$$

which are less than a given number between 0 and 1.

We shall restrict ourselves to the case in which $f(x)$ is a polynomial with real coefficients. We shall suppose, as usual, that the degree of $f(x)$ is fixed, though the method is also applicable if the degree increases slowly at the same time as P increases and the form of $f(x)$ varies.

THEOREM. *Suppose $n \geqq 11$, and let*

$$f(x) = a_{n+1}x^{n+1} + \ldots + a_1 x$$

be a polynomial with real coefficients. Let s be one of the numbers $n+1, \ldots, 2$ and suppose that

$$a_s = \frac{a}{q} + \frac{\theta}{q^2}, \text{ where } (a, q) = 1,\ q > 0.$$

Then the number T of integers x with $1 \leqq x \leqq P$ for which

$$\{f(x)\} \leqq \gamma \quad (0 < \gamma < 1)$$

is given by

$$T = \gamma P + O(P^{1-\varrho}),$$

where

$$\varrho = \frac{\tau}{3n^2 \log (12n(n+1)/\tau)}$$

and τ is defined in terms of P and q as follows:

$$\begin{aligned} q &= c_1 P^{\tau} && \textit{if } 1 < q \leqq c_1 P, \\ \tau &= 1 && \textit{if } c_1 P \leqq q \leqq c_2 P^{s-1}, \\ q &= c_2 P^{s-\tau} && \textit{if } c_2 P^{s-1} \leqq q \leqq c_3 P^s. \end{aligned}$$

Here c_1, c_2, c_3 are positive constants which can be chosen arbitrarily.

Proof. Put $\Delta = P^{-\varrho}$. We can obviously suppose that $\Delta < \frac{1}{4}$, since otherwise the theorem asserts no more than the trivial estimate $T = O(P)$.

Let A, B be any two real numbers with $0 \leqq B - A \leqq 1 - 2\Delta$. We apply Lemma 12 of Chapter I, with

$$r = 1, \ \alpha = A - \tfrac{1}{2}\Delta, \ \beta = B + \tfrac{1}{2}\Delta.$$

By that lemma, there exists a function $\psi(z)$ with period 1, whose value is

$$\begin{aligned} &1 \text{ if } A \leqq z \leqq B \pmod 1, \\ &0 \text{ if } B + \Delta \leqq z \leqq 1 + A - \Delta \pmod 1, \end{aligned}$$

and otherwise between 0 and 1. This function has the Fourier series expansion

$$\psi(z) = (B - A + \Delta) + \sum_{m=1}^{\infty} (a_m \cos 2\pi m z + b_m \sin 2\pi m z),$$

where

$$\begin{aligned} &a_m \ll m^{-1}, \ b_m \ll m^{-1} \text{ if } m \leqq \Delta^{-1}, \\ &a_m \ll \Delta^{-1} m^{-2}, \ b_m \ll \Delta^{-1} m^{-2} \text{ if } m \geqq \Delta^{-1}. \end{aligned}$$

Let $T(A, B)$ denote the number of integers x with $1 \leqq x \leqq P$ for which $A \leqq f(x) \leqq B \pmod 1$. Then, by the above,

$$(1) \qquad T(A, B) \leqq \sum_{x=1}^{P} \psi(f(x)) \leqq T(A - \Delta, \ B + \Delta).$$

Now

$$(2) \quad \sum_{x=1}^{P} \psi(f(x)) = P(B - A + \Delta) + \sum_{m=1}^{\infty} (a_m S_m' + b_m S_m''),$$

where

$$S_m = S_m' + i S_m'' = \sum_{x=1}^{P} e(m f(x)).$$

We apply Theorem I of Chapter VI to the sum S_m. This gives

$$S_m \ll m^{2\varrho/\tau} P^{1-\varrho-\varrho/(30n)},$$

where τ and ϱ are as defined in the enunciation. Hence

$$\sum_{m=1}^{\infty} (a_m S_m' + b_m S_m'')$$

$$\ll P^{1-\varrho-\varrho/(30n)} \sum_{m \leqq \Delta^{-1}} m^{-1+2\varrho/\tau} + P^{1-\varrho-\varrho/(30n)} \sum_{m > \Delta^{-1}} \Delta^{-1} m^{-2+2\varrho/\tau}.$$

Since $2\varrho/\tau < 1$, the last expression is

$$\ll P^{1-\varrho-\varrho/(30n)} \Delta^{-2\varrho/\tau} + P^{1-\varrho-\varrho/(30n)} \Delta^{-1} \Delta^{1-2\varrho/\tau}.$$

Now $\Delta = P^{-\varrho}$ and $\tau \leqq 1$ always. Thus

$$2\varrho = \frac{2\tau}{3n^2 \log\left(12n(n+1)/\tau\right)} \leqq \frac{2\tau}{3n^2 \log 12n(n+1)} < \frac{\tau}{30n},$$

whence

$$2\varrho^2\tau^{-1} < \frac{\varrho}{30n}.$$

It follows that the previous expression is $\ll P^{1-\varrho}$.

Substituting in (2), we have

$$\sum_{x=1}^{P} \psi(f(x)) = P(B-A) + O(P^{1-\varrho}).$$

By (1), this is an upper bound for $T(A, B)$ and a lower bound for $T(A-\Delta, B+\Delta)$. In particular, using this upper bound with A and B replaced by $A-\Delta$ and A, we obtain $T(A-\Delta, A) \ll P\Delta$, and in a similar way $T(B, B+\Delta) \ll P\Delta$. Hence

$$T(A, B) = P(B-A) + O(P^{1-\varrho}).$$

This result has been proved for any real numbers A, B satisfying $0 \leqq B-A \leqq 1-2\Delta$. If $\gamma \leqq 1-2\Delta$ we can take $A=0$ and $B=\gamma$ and obtain the desired result. If $\gamma > 1-2\Delta$ then $1-\gamma < 2\Delta < 1-2\Delta$, and we can take $A=\gamma$ and $B=1$ and obtain the desired result by subtracting the interval $(\gamma, 1)$ from the interval $(0, 1)$.

NOTES ON CHAPTER VIII

The reader will observe that there is a certain analogy between the relationship which this chapter bears to Chapter V and the relationship which the preceding chapter bears to Chapter IV. In Chapter V the problem was that of finding *some* integer x for which $f(x)$ is near to a given real number, and therefore x could be restricted to integers of some special kind. Similarly in Chapter IV the problem was that of representing N in any one way as a sum of nth powers, and again the variables could be restricted in any way. In the last two chapters, on the other hand, the problems concern *all* integers from 1 to P, and the only known method of attacking them is by employing an estimate for Weyl sums.

CHAPTER IX

Estimates for the Simplest Trigonometrical Sums with Primes

In this chapter I apply my method to obtain estimates for sums of the form

$$\sum_{p \leqq N} e(\alpha p),$$

where α is real and p runs through the primes. These estimates depend on rational approximations to the number α.

It may be noted that the same methods can be applied to sums of the form

$$\sum_{p \leqq N} e(\alpha f(p)),$$

where $f(p)$ is a polynomial of higher degree than the first, the estimates then depending on rational approximations to the coefficient of some power of p higher than the first; or more generally where $f(p)$ is a function which in a certain sense approximates closely to a polynomial.

The same method also makes it possible to estimate certain purely arithmetical sums in which the variable is a prime, for example sums of the form

$$\sum_{p \leqq N} \chi(p+k),$$

where χ is a non-principal character to the modulus q and $(k, q)=1$. More generally, by combining my method with that of certain English mathematicians, it is possible also to estimate sums of the form

$$\sum_{p \leqq N} \chi(f(p)),$$

where $f(p)$ is a polynomial whose values are integers.

Finally, it may be noted that in all such questions it is possible to replace the variable p, running through primes, by a variable running through some other sequence of positive integers, e.g. the sequence formed by the primes in a given arithmetical progression, or the sequence of integers a for which $\mu(a)\chi(a)$ has a given value, or sequences consisting of products of various general forms.

Notation in this chapter. The letter p will always denote a prime, unless otherwise stated. N will be an arbitrarily large positive integer, and we write

$$r = \log N.$$

LEMMA 1. *Suppose we are given any three increasing sequences of positive integers. Let u assume all values u_1u_2, where u_1 runs through all integers of the first sequence and u_2 runs independently through all integers of the second sequence. Let v run through the integers of the third sequence. Suppose that*

$$1 < U < N, \ U \ll U' \ll U.$$

Suppose that

$$\alpha = \frac{a}{q} + \frac{\theta}{q^2}, \text{ where } (a, q) = 1, \ 1 < q < N.$$

Let

$$S = \sum_{U < u \leqq U'} \sum_{v \leqq N/u} e(\alpha uv).$$

Then

$$S \ll Nr^2(q^{-1} + qN^{-1} + U^{-1} + UN^{-1})^{\frac{1}{2}}.$$

Proof. Let $\xi(z)$ denote the number of representations of an integer z as u_1u_2, so that $0 \leqq \xi(z) \leqq \tau(z)$. Then

$$S = \sum_{U < z \leqq U'} \xi(z) \sum_{v \leqq N/z} e(\alpha zv),$$

where z takes all integral values in the interval specified. By Lemma 17 of Chapter I,

$$\sum_{z \leqq U'} (\tau(z))^2 \ll Ur^3.$$

Hence, by Cauchy's inequality,

$$\begin{aligned} S^2 &\ll Ur^3 \sum_{U<z\leqq U'} \Big| \sum_{v\leqq N/z} e(\alpha z v)\Big|^2 \\ &= Ur^3 \sum_{U<z<U'} \sum_{v\leqq N/z} \sum_{v'\leqq N/z} e\big(\alpha z(v-v')\big). \end{aligned}$$

By Lemma 6 of Chapter I, noting that z, for given v and v', runs through all the integers of a certain interval of length $\ll U$, we have

$$S^2 \ll Ur^3 \sum_{v<N/U} \sum_{v'<N/U} \min\left(U, \frac{1}{\|\alpha(v-v')\|}\right).$$

For each value of v we split the summation over v' into $\ll NU^{-1}q^{-1}+1$ intervals each of length $\leqq q$. For such an interval, the sum over v' is of the type Ω considered in Lemma 8a of Chapter I, with $\lambda = 1$. Hence each such sum is $\ll U + q\log q$. Thus

$$\begin{aligned} S^2 &\ll Ur^3(NU^{-1})(NU^{-1}q^{-1}+1)(U+q\log q) \\ &\ll N^2r^4(q^{-1}+qN^{-1}+U^{-1}+UN^{-1}), \end{aligned}$$

since $\log q < \log N = r$. This gives the result.

Lemma 2. *Let P be a positive integer. Let z run through any finite set of positive integers, and let $f(z)$ be an arbitrary function of z. Let*

$$S' = \sum_{(z,\,P)=1} f(z), \qquad S_d = \sum_{z\equiv 0 \pmod d} f(z).$$

Then

$$S' = \sum_{d|P} \mu(d) S_d. \tag{1}$$

Moreover, if $f(z) \geqq 0$ for the values of z under consideration, and if m is an even positive integer, then

$$S' \leqq \sum_{\substack{d|P \\ \Omega(d)\leqq m}} \mu(d) S_d. \tag{2}$$

Proof. We have

$$\sum_{d|P} \mu(d)S_d = \sum_{d|P} \sum_{z\equiv 0 \pmod d} \mu(d) f(z).$$

The coefficient of $f(z)$ is

$$\sum_{d/(P,z)} \mu(d),$$

which is 1 if $(P, z) = 1$ and 0 otherwise. This proves (1).

For the second result, it suffices to prove that if $n = (P, z) > 1$, then

$$\sum_{\substack{d/n \\ \Omega(d) \leqq m}} \mu(d) \geqq 0$$

for any even positive integer m. If $\Omega(n) = t$, we need consider only the case $m < t$. Collecting together the divisors d of n with the same number of prime factors, the above sum is

$$1 - \binom{t}{1} + \binom{t}{2} - \ldots + \binom{t}{m}.$$

It is easily proved (e.g. by induction on t) that the value of this sum is

$$\binom{t-1}{m} \geqq 0.$$

THEOREM. 1 *Let*

$$\alpha = \frac{a}{q} + \frac{\theta}{q^2}, \textit{ where } (a, q) = 1 \textit{ and } 1 < q < N.$$

Let

$$S = \sum_{p \leqq N} e(\alpha p).$$

Then

$$S \ll N r^{\frac{9}{2}} \left((q^{-1} + qN^{-1})^{\frac{1}{2}} + H^{-1} \right),$$

where

$$H = \exp\left(\tfrac{1}{2}\sqrt{r}\right).$$

Proof. Let P denote the product of all primes not exceeding $\sqrt{N}$. The positive integers $z \leqq N$ satisfying $(z, P) = 1$ consist of 1 and the primes p satisfying $\sqrt{N} < p \leqq N$. In the notation of Lemma 2, putting $f(z) = e(\alpha z)$, we have, therefore,

$$S' = e(\alpha) + \sum_{\sqrt{N} < p \leqq N} e(\alpha p) = S + O(\sqrt{N}).$$

Thus, by the identity of that lemma,

$$S = \sum_{d|P} \mu(d) S_d + O(\sqrt{N})$$
$$= \sum_{d|P} \mu(d) \sum_{m \leqq N/d} e(\alpha d m) + O(\sqrt{N}).$$

We divide the interval $0 < m \leqq N$ into subintervals of the form

$$M < m \leqq M', \text{ where } M < M' \leqq 2M;$$

and put

$$S(M) = \sum_{d|P} \mu(d) \sum_{\substack{M < m \leqq M' \\ m \leqq N/d}} e(\alpha d m).$$

The subdivision can be effected with $\ll r$ subintervals, and consequently S is a sum of $\ll r$ sums of the form $S(M)$, together with an error $O(\sqrt{N})$.

We consider firstly any sum $S(M)$ for which $M \geqq H$. By Lemma 6 of Chapter I,

$$S(M) \ll \sum_{d \leqq N/M} \min\left(\frac{N}{d}, \frac{1}{2 \,\|\, \alpha d \,\|}\right),$$

where now we can allow d to run through all positive integers $\leqq N/M$. The above sum is of the form considered in Lemma 8b of Chapter I, with $W = N$, $W_0 = N/M$. Hence

$$S(M) \ll (NM^{-1} + q + Nq^{-1}) \log N \ll Nr(H^{-1} + qN^{-1} + q^{-1}).$$

The sum of $\ll r$ such sums therefore satisfies the inequality enunciated in the theorem.

We consider secondly any sum $S(M)$ for which $M < H$. We deal first with those values of d in $S(M)$ which are composed entirely of primes $\leqq H^2$. We shall prove that the number of such values of d is $\ll NM^{-1}H^{-1}$. For suppose d is itself greater than $NM^{-1}H^{-1}$. The number $\varkappa$ of prime factors of d satisfies

$$H^{2\varkappa} \geqq d > NM^{-1}H^{-1} > NH^{-2},$$

whence, by the definition of H,

$$(2\varkappa + 2)\tfrac{1}{2}\sqrt{r} > r, \text{ or } \varkappa > \sqrt{r} - 1.$$

Hence for any such d we have $\tau(d) > 2^{\sqrt{r-1}}$, and the number of possibilities for d is

$$\ll \sum_{d \leqq N/M} \frac{\tau(d)}{2^{\sqrt{r}}} \ll NM^{-1} r 2^{-\sqrt{r}} \ll NM^{-1}H^{-1},$$

on using Lemma 17 of Chapter I and noting that $r2^{-\sqrt{r}} < H^{-1}$ since r is large and $\sqrt{e} < 2$. This proves that the number of possible values of d of the kind under consideration is $\ll NM^{-1}H^{-1}$, and their contribution, say $S_0(M)$, to $S(M)$ is $\ll NH^{-1}$. The sum of $\ll r$ such contributions satisfies the inequality enunciated.

Finally, we have to consider

$$S(M) - S_0(M) = \sum_{M < m \leqq M'} \sum_{d \leqq N/m} \mu(d) e(\alpha dm),$$

where $M < H$ and d runs through divisors of P which have at least one prime factor $> H^2$. We write the last expression as

$$\sum_k S_k'(M) - \sum_k S_k''(M),$$

where

$$S_k'(M) = \sum_{M < m \leqq M'} \sum_{d \leqq N/m} e(\alpha dm),$$

d running now through divisors of P which have exactly k prime factors $> H^2$ and for which $\mu(d) = 1$, and where $S_k''(M)$ has a similar definition but with $\mu(d) = -1$. Also $k = 1, 2, \ldots$, and the greatest value of k is $\ll r$, since a number not exceeding N has $\ll r$ prime factors. It will be enough to consider $S_k'(M)$, as the same treatment applies to $S_k''(M)$.

We compare $S_k'(M)$ with the sum $T_k(M)$ defined by

$$T_k(M) = \sum_{M < m \leqq M'} \sum\sum_{pt \leqq N/m} e(\alpha ptm),$$

where p runs through primes satisfying $H^2 < p \leqq \sqrt{N}$, and t runs through divisors of P with $\mu(t) = -1$, which have exactly $k - 1$ prime factors $> H^2$. Each number pt for which $(p, t) = 1$ is one of the possible values of d in the sum for $S_k'(M)$, and each such value of d arises k times in the sum T_k, namely with p as any one of the k prime factors $> H^2$ of d. The number of terms in the sum T_k for which $(p, t) > 1$ is obviously

$$\ll M \sum_{\substack{p^2 u \leqq N/M \\ p > H^2}} \sum 1 \ll M \sum_{p > H^2} NM^{-1} p^{-2} \ll NH^{-2}.$$

Hence

$$S_k'(M) = \frac{1}{k} T_k(M) + O(NH^{-2}). \tag{3}$$

To estimate the sum T_k we split the interval $H^2 < p \leqq \sqrt{N}$ into $\ll r$ intervals of the form $Q < p \leqq Q'$, where $Q < Q' \ll Q$. Let $T_k(M, Q)$ be the part of the sum $T_k(M)$ corresponding to such an interval of primes p. Then

$$T_k(M, Q) = \sum_{M < m \leqq M'} \sum_{Q < p \leqq Q'} \sum_{mpt \leqq N} e(\alpha mpt).$$

To this we apply Lemma 1 above, taking $u = mp$ and $v = t$, where t is restricted in the manner already stated. Since U in that lemma corresponds to MQ here, we obtain

$$T_k(M, Q) \ll Nr^2(q^{-1} + qN^{-1} + M^{-1}Q^{-1} + MQN^{-1})^{\frac{1}{2}}.$$

Now $H^2 \leqq Q \leqq \sqrt{N}$, so that $M^{-1}Q^{-1} < H^{-2}$ and $MQN^{-1} < HN^{-\frac{1}{2}} < H^{-2}$. Hence

$$T_k(M, Q) \ll Nr^2((q^{-1} + qN^{-1})^{\frac{1}{2}} + H^{-1}),$$

whence, summing over $\ll r$ values of Q,

$$T_k(M) \ll Nr^3((q^{-1} + qN^{-1})^{\frac{1}{2}} + H^{-1}).$$

By (3), the same holds for $S_k'(M)$, and similarly for $S_k''(M)$, with a factor $1/k$. Next, summing over k and noting that $\Sigma 1/k \ll \log r \ll r^{\frac{1}{2}}$, we obtain

$$S(M) - S_0(M) \ll Nr^{\frac{7}{2}}((q^{-1} + qN^{-1})^{\frac{1}{2}} + H^{-1}).$$

Finally, summation over $\ll r$ values of M gives the result stated.

LEMMA 3. *If $x > 2$ then*

$$\sum_{p \leqq x} \frac{1}{p} = \log \log x + c + O\left(\frac{1}{x}\right),$$

$$\prod_{p \leqq x} \left(1 - \frac{1}{p}\right) = \frac{c_0}{\log x} + O\left(\frac{1}{(\log x)^2}\right),$$

where c, c_0 are constants and $c_0 > 0$.

For a proof see, for example, A. E. Ingham, *The distribution of prime numbers* (Cambridge, 1932), 22—24.

LEMMA 4. *Define b and b_1 by*

$$b = \exp(r^{1-\frac{3}{2}\varepsilon}), \; b_1 = \exp(r^{1-2\varepsilon}).$$

Suppose that

$$0 < q < b_1, \; 0 \leqq l < q, \; (l, q) = 1, \\ U > 0, \; W \geqq b.$$

Let T denote the number of numbers of the form $qx + l$ which are not divisible by any prime $\leqq b_1$ and which satisfy

(4) $$U < qx + l \leqq U + W.$$

Then

$$T \ll W \frac{(rq)^{2\varepsilon}}{rq}.$$

Proof. Let $p_1, \ldots, p_\sigma$ be the primes not exceeding b_1 which do not divide q, and let $P = p_1 \ldots p_\sigma$, so that $\Omega(P) = \sigma$. Let

$$m = 2[2 \log r + 1].$$

We take $f(z) = 1$ in Lemma 2, and let z run through all integers of the form $qx + l$ in the interval (4). By (2),

$$T \leqq \sum_{\substack{d|P \\ \Omega(d) \leqq m}} \mu(d) S_d,$$

where now S_d is the number of integers $qx + l$ which satisfy (4) and are divisible by d. Since $(d, q) = 1$ we have obviously

$$S_d = \frac{W}{dq} + \theta_d.$$

Thus

$$T \ll \frac{W}{q} \left| \sum_{\substack{d|P \\ \Omega(d) \leqq m}} \frac{\mu(d)}{d} \right| + \sum_{s=0}^{m} \binom{\sigma}{s},$$

the second sum on the right arising by counting the number of divisors d of P with $\Omega(d) = s$. As regards this second sum, we have

$$\sum_{s=0}^{m} \binom{\sigma}{s} < \sum_{s=0}^{m} \sigma^s \ll \sigma^{m+1} \ll b_1^{m+1},$$

and, since $W \geqq b$ and $q < b_1$,

$$b_1^{m+1} < b_1^{5\log r} < \frac{b}{b_1 r} < \frac{W}{qr}.$$

Hence it remains only to prove that

$$\sum_{\substack{d|P \\ \Omega(d) \leq m}} \frac{\mu(d)}{d} \ll \frac{1}{r} (rq)^{2\varepsilon}.$$

For the sum without the restriction $\Omega(d) \leqq m$ we have, using the second result of Lemma 3,

$$\sum_{d|P} \frac{\mu(d)}{d} = \prod_{\substack{p \leqq b_1 \\ p \nmid q}} \left(1 - \frac{1}{p}\right) \ll \frac{\log (q+1)}{\log b_1},$$

and this satisfies the required inequality, since $\log b_1 = r^{1-2\varepsilon}$.

Further, we have

$$\sum_{\substack{d|P \\ \Omega(d) > m}} \frac{1}{d} = \sum_{s=m+1}^{\sigma} \sum_{\substack{d|P \\ \Omega(d) = s}} \frac{1}{d}$$

$$\leqq \sum_{s=m+1}^{\sigma} \frac{1}{s!} \left(\frac{1}{p_1} + \ldots + \frac{1}{p_\sigma}\right)^s.$$

Now, by the first result of Lemma 3,

$$1/p_1 + \ldots + 1/p_\sigma < \log\log p_\sigma + c + 1,$$

and since $\log p_\sigma \leqq \log b_1 = r^{1-2\varepsilon}$ we can replace the last expression by $\log r$. Thus, since $s! > (s/e)^s$, we have

$$\sum_{\substack{d|P \\ \Omega(d)>m}} \frac{1}{d} < \sum_{s=m+1}^{\infty} \left(\frac{e \log r}{s}\right)^s < \sum_{s=m+1}^{\infty} \left(\frac{3}{4}\right)^s \ll \frac{1}{r},$$

on noting that $m > 4 \log r$ and $e < 3$ and $4 \log \frac{4}{3} > 1$.

The last result completes the proof of Lemma 4.

THEOREM 2a. *Let* $b_1 = \exp(r^{1-2\varepsilon})$. *Suppose that*

$$0 < q \leqq \exp(r^{\varepsilon}),\ (a, q) = 1,$$

and let

$$S = \sum_{N-A<p\leqq N} e_q(ap),$$

where

$$N/b_1 \leqq A < N.$$

Then

$$S \ll \frac{A(rq)^{5\varepsilon}}{r\sqrt{q}}.$$

Proof. Let $b = \exp(r^{1-\frac{3}{2}\varepsilon})$ as before, and $b_0 = \exp(r^{1-\varepsilon})$. (Thus $b_1 < b < b_0$.) Let P denote the product of all those primes $\leqq b_0$ which do not divide q.

We apply (1) of Lemma 2, with $f(z) = e_q(az)$, letting z run through those integers of the interval $N - A < z \leqq N$ which are relatively prime to q. This gives

$$\sum_{\substack{N-A<z\leqq N \\ (z,\, Pq)=1}} e_q(az) = \sum_{\substack{d|P \\ d\leqq N}} \mu(d) S_d, \tag{5}$$

where

$$S_d = \sum_{\substack{\frac{N-A}{d} < m \leqq \frac{N}{d} \\ (m,q)=1}} e_q(adm). \tag{6}$$

We first obtain an estimate for the right hand side of (5), and afterwards investigate the difference between the left hand side of (5) and the sum S of the theorem.

We have

$$S_d = \sum_{\substack{l=0 \\ (l,q)=1}}^{q-1} Z_d(l) e_q(al),$$

where $Z_d(l)$ denotes the number of integers z satisfying

(7) $\quad N - A < z \leqq N,\ z \equiv l \pmod{q},\ z \equiv 0 \pmod{d}.$

Since $(d, q) = 1$, we have

$$Z_d(l) = \frac{A}{dq} + O(1).$$

Hence

$$S_d = \frac{A}{dq} \sum_{\substack{l=0 \\ (l,q)=1}}^{q-1} e_q(al) + O(q)$$

$$= \frac{A}{dq}\, \mu(q) + O(q),$$

by a well known result. It will be convenient to modify the last formula slightly; putting $Z_d = Z_d(1)$ we have

(8) $$S_d = \mu(q) Z_d + O(q).$$

We shall use this in the right hand side of (5) if $d \leqq N^{\frac{4}{5}}$.

If $d > N^{\frac{4}{5}}$, we have

$$b_0^{\Omega(d)} \geqq d > N^{\frac{4}{5}} = \exp\left(\tfrac{4}{5}r\right),$$

whence $\Omega(d) > \frac{4}{5} r^{\varepsilon}$, and therefore

$$\tau(d) = 2^{\Omega(d)} > \exp\left(\tfrac{4}{5}(\log 2) r^{\varepsilon}\right) > \exp\left(0.55 r^{\varepsilon}\right).$$

Hence

$$\sum_{\substack{d|P \\ N^{\frac{4}{5}} < d \leqq N}} |S_d| \ll \sum_{N^{\frac{4}{5}} < d \leqq N} \ \sum_{\frac{N-A}{d} < m \leqq \frac{N}{d}} \tau(d) \exp\left(-0.55 r^{\varepsilon}\right)$$

$$\ll \exp\left(-0.55 r^{\varepsilon}\right) \sum_{m < N^{\frac{1}{5}}} \sum_{d} \tau(d),$$

where d now runs through all integers in an interval $d_1 < d \leqq d_1 + h$, where $h \leqq A/m$ and $d_1 + h \leqq N/m$. By Lemma 17 of Chapter I,

$$\sum_{d_1 < d \leq d_1+h} \tau(d) = h \log (d_1 + h) + d_1 \log(1 + h/d_1) + (2E - 1)h + O((d_1 + h)^{\frac{1}{2}})$$
$$\ll Ar/m + A/m + A/m + O((N/m)^{\frac{1}{2}})$$
$$\ll Ar/m + O((N/m)^{\frac{1}{2}}).$$

Hence

$$\sum_{\substack{d|P \\ N^{\frac{4}{5}} < d \leq N}} |S_d| \ll \exp(-0.55r^\varepsilon) \sum_{m < N^{\frac{1}{5}}} \left(\frac{Ar}{m} + \left(\frac{N}{m}\right)^{\frac{1}{2}}\right)$$
$$\ll \exp(-0.55r^\varepsilon)(Ar^2 + N^{\frac{3}{5}})$$
$$\ll \frac{A}{r\sqrt{q}},$$

since $N^{\frac{3}{5}} < A$ and $\sqrt{q} < \exp(\frac{1}{2}r^\varepsilon)$.

The same estimate applies to

$$\sum_{\substack{d|P \\ N^{\frac{4}{5}} < d \leq N}} Z_d,$$

since Z_d, by its definition in (7) (with $l = 1$) obviously satisfies

$$Z_d \leq \sum_{\frac{N-A}{d} < m \leq \frac{N}{d}} 1.$$

Hence, by (8) and the results just proved, we have

$$\sum_{\substack{d|P \\ d \leq N}} \mu(d)S_d = \mu(q) \sum_{\substack{d|P \\ d \leq N}} \mu(d)Z_d + O(N^{\frac{4}{5}}q) + O\left(\frac{A}{r\sqrt{q}}\right).$$

From Lemma 2, letting z run through the integers of the interval $N - A < z \leq N$ which are $\equiv 1 \pmod q$, and taking $f(z) = 1$, we see that

$$\sum_{\substack{d|P \\ d \leq N}} \mu(d)Z_d$$

is exactly equal to the number of the above integers z which are relatively prime to P. By Lemma 4, with $l = 1$, $U = N - A$,

$W = A$, this number is

$$\ll \frac{A\,(rq)^{2\varepsilon}}{rq}.$$

This completes the estimation of the right hand side of (5), and we have now proved that

$$\sum_{\substack{N-A<z\leqq N\\(z,\,Pq)=1}} e_q(az) \ll \frac{A\,(rq)^{2\varepsilon}}{r\sqrt{q}}.$$

The condition $(z, Pq) = 1$ is equivalent to the condition that every prime factor of z is greater than b_0. The number D of integers z in the above sum which are not square free satisfies

$$D \ll \sum_{b_0<p\leqq\sqrt{N}} \left(\frac{A}{p^2}+1\right) \ll \frac{A}{b_0} + \sqrt{N} \ll \frac{A}{r\sqrt{q}},$$

and so can be neglected. Let

(9) $$H_k = \sum_{N-A<z_k\leqq N} e_q(az_k),$$

where z_k runs through the numbers which are products of k distinct primes all greater than b_0. Then, since $N - A > b_0$, the sum S of the enunciation of the Theorem is H_1, and by the above results we have

$$\sum_{k=1}^{k_0} H_k \ll \frac{A\,(rq)^{2\varepsilon}}{r\sqrt{q}},$$

where $k_0 \ll r$ as usual. It remains to find an estimate for H_k when $k \geqq 2$. We shall prove that

(10) $$H_k \ll \frac{A\,(rq)^{4\varepsilon}}{kr\sqrt{q}}$$

for $k \geqq 2$, and this will give the derived estimate for S, since

$$\sum 1/k \ll \log r \ll (rq)^{\varepsilon}.$$

The sum H_k defined in (9) is related in an obvious way to the sum L_k defined by

$$L_k = \sum_{N-A<pv\leqq N}\sum e_q(apv), \tag{11}$$

where p runs through primes greater than b_0 and v runs through products of $k-1$ distinct primes all greater than b_0. The number of terms for which $(p, v) > 1$ satisfies the same estimate as D above, and the other terms give every term in the sum H_k each exactly k times. Hence

$$H_k = \frac{1}{k}L_k + O\left(\frac{A}{r\sqrt{q}}\right)$$

and to prove (10) it suffices to prove that

$$L_k \ll \frac{A(rq)^{4\varepsilon}}{r\sqrt{q}} \tag{12}$$

for $k \geqq 2$.

In the sum L_k, we have

$$b_0 < p < Nb_0^{-k+1}$$

since $v > b_0^{k-1}$. We split up this interval into $\ll r$ intervals of the form

$$Y < p \leqq Y', \text{ where } 2Y \leqq Y' \leqq 3Y; \tag{13}$$

and we then further split up each of these intervals into $\ll Yb^{-1}$ intervals of the form

$$U < p \leqq U + W, \text{ where } b \leqq W \leqq 2b. \tag{14}$$

Consider the part of L_k corresponding to primes p in one of these last intervals. The error introduced if we replace the condition of summation on v, which is

$$\frac{N-A}{p} < v \leqq \frac{N}{p},$$

by the simpler condition

$$\frac{N-A}{U} < v \leqq \frac{N}{U},$$

is

$$\ll \sum_p \left(\frac{NW}{Up} + 1\right) \ll W\left(\frac{NW}{U^2} + 1\right) = \frac{WA}{U}\left(\frac{NW}{AU} + \frac{U}{A}\right).$$

Now

$$\frac{NW}{AU} \ll b_1 \frac{W}{U} \ll \frac{b_1 b}{b_0} \ll \frac{1}{r^2\sqrt{q}},$$

$$\frac{U}{A} \ll \frac{Nb_0^{-k+1}}{Nb_1^{-1}} \leqq \frac{b_1}{b_0} \ll \frac{1}{r^2\sqrt{q}}.$$

Hence the error in question is

$$\ll \frac{WA}{Ur^2\sqrt{q}}.$$

After making this change, the part of L_k corresponding to an interval (14) of values of p is

$$M = \sum_{U<p\leqq U+W} \sum_{\frac{N-A}{U}<v\leqq\frac{N}{U}} e_q(apv).$$

Let $\xi(l)$ denote the number of primes p in the last sum for which $p \equiv l \pmod q$, and let $\eta(t)$ denote the number of values of v in the last sum for which $v \equiv t \pmod q$. These numbers are 0 unless $(l, q) = 1$ and $(t, q) = 1$, and in the latter case they can be estimated by Lemma 4. Since $W \geqq b$ and $A/U \gg (Nb_1^{-1})/(Nb_0^{-k+1}) > b$, and since the upper bound for q in Lemma 4 is amply satisfied, we have

$$\xi(l) \ll \frac{W(rq)^{2\varepsilon}}{rq}, \quad \eta(t) \ll \frac{A(rq)^{2\varepsilon}}{Urq}.$$

Now

$$M = \sum_{l=0}^{q-1} \sum_{t=0}^{q-1} \xi(l)\eta(t)e_q(alt);$$

and on applying Lemma 10a of Chapter I we obtain

$$M \ll \{q(\Sigma \mid \xi(l) \mid^2)(\Sigma \mid \eta(t)\mid^2)\}^{\frac{1}{2}}$$
$$\ll \left(q^3 \frac{W^2(rq)^{4\varepsilon}}{r^2q^2} \frac{A^2(rq)^{4\varepsilon}}{U^2r^2q^2}\right)^{\frac{1}{2}}$$
$$= \frac{WA(rq)^{4\varepsilon}}{Ur^2\sqrt{q}}.$$

It follows that the contribution of an interval of the form (14) to L_k satisfies the estimate just given. The contribution of an interval of the form (13) is therefore

$$\ll \frac{WA(rq)^{4\varepsilon}}{Ur^2\sqrt{q}} \frac{Y}{b} \ll \frac{A(rq)^{4\varepsilon}}{r^2\sqrt{q}},$$

since $W \ll b$ and $U \gg Y$. Finally, on adding $\ll r$ such contributions to give L_k, we obtain the required result (12).

THEOREM 2b. *Suppose that*

$$\tau = N/H, \text{ where } 1 < H \leqq \exp(r^\varepsilon).$$

Suppose that

$$\alpha = \frac{a}{q} + z, \text{ where } (a, q) = 1, 0 < q \leqq \exp(r^{\frac{1}{2}\varepsilon}), \ |z| \leqq \frac{1}{q\tau}.$$

Let

$$S = \sum_{p \leqq N} e(\alpha p).$$

Then

$$S \ll \frac{N(rq)^{5\varepsilon}}{r\sqrt{q}}.$$

Proof. We split up the interval $0 < p \leqq N$ into $[\exp r^{\frac{1}{2}}]$ intervals of length $A = N[\exp r^{\frac{1}{2}}]^{-1}$. Denote one such interval by $N_1 - A < p \leqq N_1$. Putting $r_1 = \log N_1$, we have $r_1 \leqq r$ and

$$r_1 \geqq \log A \geqq r - \sqrt{r}.$$

Also

$$N_1 \exp(-\sqrt{r}) < A < N_1.$$

The contribution of an interval $N_1 - A < p \leqq N_1$ to the sum S is

$$\sum_{N_1 - A < p \leqq N_1} e\left(\frac{a}{q} p + zp\right),$$

and

$$e(zp) = e(zN_1) + O\left(\frac{A}{q\tau}\right).$$

To the sum

$$\sum_{N_1 - A < p \leqq N_1} e_q(ap)$$

we can apply Theorem 2*a*, the conditions of which are satisfied since

$$q \leqq \exp(r^{\frac{1}{2}\varepsilon}) < \exp(r_1^{\varepsilon})$$

and

$$A \geqq N_1 \exp(-r^{\frac{1}{2}}) > N_1 \exp(-r_1^{1-2\varepsilon}).$$

It follows that

$$\sum_{N_1 - A < p \leqq N_1} e\left(\frac{a}{q} p + zp\right) \ll \frac{A(r_1 q)^{5\varepsilon}}{r_1\sqrt{q}} + \frac{A^2}{q\tau} \ll \frac{A(rq)^{5\varepsilon}}{r\sqrt{q}},$$

since $A\tau^{-1} \ll N \exp(-r^{\frac{1}{2}}) N^{-1} H \ll r^{-1}$. Finally,

$$S \ll \exp(r^{\frac{1}{2}}) \frac{A(rq)^{5\varepsilon}}{r\sqrt{q}} \ll \frac{N(rq)^{5\varepsilon}}{r\sqrt{q}}.$$

LEMMA 5. *Suppose*

$$0 < c \leqq \tfrac{1}{6},\ 0 < \sigma \leqq \tfrac{1}{3}.$$

Let d run through all distinct positive integers not exceeding N which are products of distinct primes not exceeding N^{σ}. Then these numbers d can be distributed into at most

$$D = \exp\frac{(\log r)^2}{\log(1 + c)}$$

disjoint sets, with the following properties.

(i) *To each set there corresponds a positive number φ such that all the numbers in the set satisfy*

$$\varphi \leqq d \leqq \varphi^{1+c}.$$

(ii) *Suppose that, for a particular set,* $\varphi > N^\gamma$, *where the number* γ *(which may vary from set to set) satisfies*

$$0 < \gamma \leqq 1 - \sigma.$$

Then there exist for that set two increasing sequences of positive integers x *and* y *such that all the integers* x *lie in an interval*

$$\varphi_0 \leqq x \leqq \varphi_0^{1+c}, \text{ where } N^\gamma < \varphi_0 \leqq N^{\gamma+\sigma},$$

and such that the products xy, *with*

$$(x, y) = 1, \; xy \leqq N,$$

comprise precisely the numbers d *of the set in question, each repeated the same number of times.*

Proof. Define a function $\varphi(t)$ for integers $t \geqq 0$ by

$$\varphi(0) = 1, \; \varphi(1) = 2, \; \varphi(t+1) = (\varphi(t))^{1+c} \text{ for } t \geqq 1.$$

Let τ be the greatest integer for which $\varphi(\tau) \leqq N^\sigma$; then

$$(1+c)^{\tau-1} \log 2 \leqq \sigma \log N = \sigma r.$$

Since $\sigma \leqq \frac{1}{3}$ and $\log 8 > (1+\frac{1}{6})^2 \geqq (1+c)^2$, this implies

$$\tau + 1 < \frac{\log r}{\log (1+c)}.$$

Let $R = [r]$. Then each number d under consideration has at most R prime factors, since any product of $R+1$ distinct primes is

$$\geqq 2 \,.\, 3 \ldots (R+2) > e^{R+1} > N$$

for sufficiently large N. We can represent each d as

$$d = p_1 p_2 \ldots p_R,$$

where each p_j is either a prime or 1, and where

$$N^\sigma \geqq p_1 \geqq p_2 \geqq \ldots \geqq p_R \geqq 1$$

(with $p_j > p_{j+1}$ unless $p_j = 1$).

For each $j = 1, \ldots, R$ we define $t_j = t_j(d)$ to be the unique integer t_j for which

$$\varphi(t_j) \leqq p_j < \varphi(t_j + 1). \tag{15}$$

Thus $t_j = 0$ if $p_j = 1$ and $t_j = 1$ if $p_j = 2$. Then each number d defines a sequence $t_1, \ldots, t_R$ of integers satisfying

$$\tau \geqq t_1 \geqq t_2 \geqq \ldots \geqq t_R \geqq 0. \tag{16}$$

It should be observed that the number $d = 1$, which arises as the empty product of primes, corresponds to the sequence $0, \ldots, 0$.

We collect together all numbers d which give rise to the same sequence $t_1, \ldots, t_R$, and thereby distribute all the numbers d into sets. It remains to be proved that these sets have the properties stated in the enunciation.

A sequence $t_1, \ldots, t_R$ satisfying (16) is uniquely determined if we know how often each of the numbers $\tau, \ldots, 0$ occurs in it. The number of times each of these numbers occurs is at most R, so the number of distinct sequences $t_1, \ldots, t_R$ is

$$\leqq R^{\tau+1} \leqq r^{\tau+1} < D,$$

by the inequality for τ proved earlier.

For each sequence we define φ by

$$\varphi = \varphi(t_1) \ldots \varphi(t_R).$$

Then for all the numbers d which correspond to this sequence, we have $d \geqq \varphi$ by (15). Also, if $t_j > 0$ and $t_{j+1} = 0$, we have

$$d = p_1 \ldots p_j \leqq \varphi(t_1 + 1) \ldots \varphi(t_j + 1) = (\varphi(t_1) \ldots \varphi(t_j))^{1+c} = \varphi^{1+c},$$

and this conclusion holds also (with equality) when all the t_j are 0, in which case $\varphi = 1$. Hence (i) holds.

Now consider a sequence for which

$$\varphi = \varphi(t_1) \ldots \varphi(t_R) > N^\gamma.$$

Define an integer s, where $1 \leqq s \leqq R$, by

$$\varphi(t_1) \ldots \varphi(t_{s-1}) \leqq N^\gamma < \varphi(t_1) \ldots \varphi(t_s).$$

Putting

$$\varphi_0 = \varphi(t_1) \ldots \varphi(t_s),$$

we have

$$\varphi_0 \leqq p_1 \ldots p_s < \varphi_0^{1+c},$$

and

$$N^\gamma < \varphi_0 \leqq \varphi(t_1) \ldots \varphi(t_{s-1})N^\sigma \leqq N^{\gamma+\sigma}.$$

Now let x run through all distinct products $p_1 \ldots p_s$ for which (15) holds for $j = 1, \ldots, s$ and for which

$$p_1 > p_2 > \ldots > p_s,$$

and note that $p_s > 1$ since $\varphi(t_s) > 1$. Let y run through all distinct products $p_{s+1} \ldots p_R$ for which (15) holds for $j=s+1, \ldots, R$ and for which

$$p_{s+1} > \ldots > p_R,$$

where it is understood that the inequality $p_j > p_{j+1}$ is to be replaced by $p_j = p_{j+1}$ if $p_j = 1$. The inequalities for x and φ_0 are satisfied. Each product xy, where $(x, y) = 1$ and $xy \leqq N$, is a number d of the set corresponding to the sequence $t_1, \ldots, t_R$, and it remains to consider how often each d is representable as xy.

If, in the particular sequence $t_1, \ldots, t_R$ under consideration, we have $t_s > t_{s+1}$, the choice of x and y for given d is unique; for then x is the product of all those prime factors of d which are $\geqq \varphi(t_s)$ and y is the product of the others. Now suppose $\varphi(t_s) = \varphi(t_{s+1})$, and let

$$t_{s_1}, \ldots, t_s, \; t_{s+1}, \ldots, t_{s_2}$$

be all the t_j equal to t_s. Then the intervals for the prime factors $p_{s_1}, \ldots, p_s$ of x and the prime factors $p_{s+1}, \ldots, p_{s_2}$ of y all coincide. Any d which belongs to the set under consideration has exactly $s_2 - s_1 + 1$ distinct prime factors in this interval. Of these, any $s - s_1 + 1$ can be assigned to x and the remaining ones to y. Hence the number of representations of d as xy is

$$\binom{s_2 - s_1 + 1}{s - s_1 + 1},$$

and this is the same for all numbers d in the particular set.

LEMMA 6. *Let x run through one increasing sequence of positive integers and y run through another. Let u run through all products of c_1 factors, one from each of c_1 increasing sequences of positive* *integers, and let v run through all products of c_2 factors, one from each of c_2 increasing sequences of positive integers. Suppose that $U \geqq 1$, $X \geqq 1$, and*

$$N^{\frac{1}{4}} \ll UX \ll N^{\frac{3}{4}}. \tag{17}$$

Suppose that

$$U < U' \ll U,\ X < X' \ll X.$$

Let τ satisfy

$$N^{\frac{1}{2}} \leqq \tau \leqq N \exp(-r^{\varepsilon_0}).$$

Let

$$\alpha = \frac{a}{q} + \frac{\theta}{q\tau},\ \textit{where } (a, q) = 1 \textit{ and } \exp(r^{\varepsilon_0}) \leqq q \leqq \tau.$$

Let

$$\Delta = (q^{-1} + qN^{-1})^{\frac{1}{2}},$$

and let K be a positive integer satisfying

$$K \ll \Delta^{-2}.$$

Let

$$S = \sum_{k=1}^{K} \left| \sum_u \sum_x \sum_y \sum_v e(\alpha kuxyv) \right|,$$

where the summation is extended over

$$U < u \leqq U',\ X < x \leqq X',\ uxyv \leqq N,\ (x, y) = 1.$$

Then

$$S \ll KN \left(\Delta^{1-\varepsilon} + \left(\frac{1}{UX} + \frac{UX}{N} \right)^{\frac{1}{2}-\varepsilon} \right).$$

Proof. We recall first an elementary identity which is applicable to sums which are subject to the condition $(x, y) = 1$. If $\mathscr{R}$ is any set of pairs x, y of integers, and $f(x, y)$ any function, the identity is

$$\sum_{\substack{x,y \text{ in } \mathscr{R} \\ (x,y)=1}} f(x, y) = \sum_{d \geqq 1} \mu(d) \sum_{dx',\, dy' \text{ in } \mathscr{R}} f(dx',\, dy').$$

The proof is immediate, since the coefficient of $f(x, y)$ on the right is $\Sigma\mu(d)$ extended over all d which divide both x and y.

Applying this principle to the sum S, we obtain

$$S \leqq \sum_{k=1}^{K} \sum_{d \geqq 1} \Big| \sum_{u} \sum_{x'} \sum_{y'} \sum_{v} e(\alpha k d^2 u x' y' v) \Big|,$$

where x' and y' run through the sequences obtained by dividing by d those terms of the x and y sequences which are divisible by d, and where the summation is extended over

$$(18) \qquad U < u \leqq U', \ X < dx' \leqq X', \ d^2 u x' y' v \leqq N.$$

We consider first the terms with $d > \Delta^{-1}$ in the sum. The number of representations of an integer t as $ux'y'v$ does not exceed $\tau_{c_1+c_2+2}(t)$, and consequently the number of solutions of the inequality $ux'y'v \leqq \lambda$, where $\lambda \leqq N$, is $\ll \lambda r^{c_3}$ by Lemma 17 of Chapter I, where c_3 depends only on c_1 and c_2. The part of the multiple sum which corresponds to $d > \Delta^{-1}$ is therefore

$$\ll K r^{c_3} \sum_{d > \Delta^{-1}} \frac{N}{d^2} \ll K N r^{c_3} \Delta \ll K N \Delta^{1-\varepsilon},$$

since $\Delta^2 \ll \exp(-r^{\varepsilon_0})$. Thus this part satisfies the desired inequality.

It now suffices to consider the sum S' defined by

$$(19) \qquad S' = \sum_{k=1}^{K} \sum_{d \leqq \Delta^{-1}} S_{d,k},$$

where

$$S_{d,k} = \sum_{u} \sum_{x'} \Big| \sum_{y'} \sum_{v} e(\alpha k d^2 u x' y' v) \Big|,$$

the conditions of summation being those stated in (18). It will be convenient also to write

$$(20) \qquad S' = \sum_{d \leqq \Delta^{-1}} S_d, \ \text{where } S_d = \sum_{k=1}^{K} S_{d,k}.$$

The number of representations of an integer z as ux' does not exceed $\tau_{c_1+1}(z)$. Hence

$$S_{d,k} \leqq \sum_{\frac{UX}{d} < z \leqq \frac{U'X'}{d}} \tau_{c_1+1}(z) \Big| \sum_{\substack{y' \\ d^2 z y' v \leqq N}} \sum_{v} e(\alpha k d^2 z y' v) \Big|,$$

where z runs through all integers of the interval indicated, and y', v have the same meaning as before. Applying Cauchy's inequality, and using Lemma 17 of Chapter I, we obtain

$$(S_{d,k})^2 \ll \frac{UX}{d} r^{c_4} \sum_{\frac{UX}{d} < z \leqq \frac{U'X'}{d}} \left| \sum_{y'} \sum_{\substack{v \\ d^2zy'v \leqq N}} e(\alpha k d^2 z y' v) \right|^2,$$

where c_4 depends only on c_1. We write the square of the sum as a double sum over y', v, y_1', v_1, and interchange the order of summation. For specified y', v, y_1', v_1, the variable z runs through all integers of a certain interval of length $\ll UX/d$. Hence

$$(S_{d,k})^2$$
$$\ll \frac{UX}{d} r^{c_4} \sum\sum_{y'v \leqq N(dUX)^{-1}} \sum\sum_{y_1'v_1 \leqq N(dUX)^{-1}} \min\left(\frac{UX}{d}, \frac{1}{\| \alpha d^2 k (y'v - y_1'v_1) \|}\right).$$

Let $\psi(t)$ denote the number of representations of an integer t as $y'v$. Then the number of representations of an integer s as $y'v - y_1'v_1$ is

$$\sum_t \psi(t)\psi(t+s),$$

and here t runs through the integers of an interval of length $\ll N(dUX)^{-1}$. Noting that $\psi(t) \leqq \tau_{c_2+1}(t)$, and using Lemma 17 of Chapter I, we have

$$\sum_t \psi(t)\psi(t+s) \leqq \left(\sum_t \psi^2(t)\right)^{\frac{1}{2}} \left(\sum_t \psi^2(t+s)\right)^{\frac{1}{2}} \ll \frac{N}{dUX} r^{c_5},$$

where c_5 depends only on c_2. Using this in the inequality for $(S_{d,k})^2$, we obtain

$$(S_{d,k})^2 \ll \frac{N}{d^2} r^{c_6} \sum_{0 \leqq s \leqq N(dUX)^{-1}} \min\left(\frac{UX}{d}, \frac{1}{\| \alpha d^2 k s \|}\right).$$

Consequently, applying Cauchy's inequality in the definition of S_d in (20), we have

$$(21) \quad (S_d)^2 \ll \frac{KN}{d^2} r^{c_6} \sum_{k=1}^{K} \sum_{0 \leqq s \leqq N(dUX)^{-1}} \min\left(\frac{UX}{d}, \frac{1}{\| \alpha d^2 k s \|}\right).$$

We now consider three cases, depending on the magnitude of q in relation to N. In all three cases the estimate will be derived from Lemma 8a of Chapter I, but the details vary.

Case 1. *Suppose that* $N^{\frac{1}{10}} \leqq q \leqq N^{\frac{9}{10}}$. We split up the interval of summation for d in (20) into $\ll r$ subintervals of the form

$$D < d \leqq D', \text{ where } D < D' \leqq 2D.$$

For all values of d, k, s under consideration, we have $d^2ks \leqq 2DKN(UX)^{-1}$. The number of representations of a positive integer t as d^2ks is $\ll N^\varepsilon$. Hence (21) implies

$$\sum_{D<d\leqq D'} (S_d)^2 \ll \frac{KN^{1+\varepsilon}}{D^2} \sum_{0<t\leqq\frac{2DKN}{UX}} \min\left(\frac{UX}{D}, \frac{1}{\|\alpha t\|}\right) + \frac{K^2N^{1+\varepsilon}}{D} \frac{UX}{D},$$

the last term being an estimate for the contribution of the terms in (21) with $s = 0$. We split the interval of summation for t into at most

$$\frac{2DKN}{UXq} + 1$$

intervals of length $\leqq q$. For the sum over any subinterval, Lemma 8a of Chapter I (with $\lambda = 1$ and U replaced by UX/D, which is $\gg 1$) gives the estimate $UX/D + q\log q$. Hence

$$\sum_{D<d\leqq D'} (S_d)^2 \ll \frac{KN^{1+\varepsilon}}{D^2} \left(\frac{DKN}{UXq} + 1\right)\left(\frac{UX}{D} + q\right) q^\varepsilon$$

$$\ll \frac{K^2N^{2+2\varepsilon}}{D} \left(\frac{1}{q} + \frac{q}{N} + \frac{1}{UX} + \frac{UX}{N}\right).$$

Thus

$$\left(\sum_{D<d\leqq D'} S_d\right)^2 \ll K^2N^{2+2\varepsilon}\left(\Delta^2 + \frac{1}{UX} + \frac{UX}{N}\right).$$

Summing over the subintervals for d, which are $\ll r$ in number, we obtain from (20) the same estimate for $(S')^2$, apart from a factor r^2 which can be absorbed in N^ε. This gives the result enunciated, since N^ε can be transferred as required, in view of the fact that q and $q^{-1}N$ are greater than positive powers of N by the hypothesis of the present case.

Case 2. *Suppose that* $q < N^{\frac{1}{10}}$. Let $\delta = (d^2, q)$ and put $d^2 = \delta d_1$, $q = \delta q_1$. Let $\varkappa = (k, q_1)$ and put $k = \varkappa k_1$, $q_1 = \varkappa q_2$. Since $d \leqq \Delta^{-1}$, we have

$$\alpha d^2 k = \frac{ad_1 k}{q_1} + \frac{\theta d_1 k}{q_1 \tau} = \frac{ad_1 k_1}{q_2} + \frac{\theta d_1 k_1}{q_2 \tau}.$$

Here $(ad_1 k_1, q_2) = 1$, and

$$\frac{d_1 k_1}{\tau} \leqq \frac{d^2 k}{\tau} \leqq \frac{\Delta^{-4}}{\sqrt{N}} < \frac{q^2}{\sqrt{N}} < \frac{1}{q} \leqq \frac{1}{q_2}.$$

For a particular value of k, the sum

$$\sum_{0 \leqq s \leqq N(dUX)^{-1}} \min\left(\frac{UX}{d}, \frac{1}{\|\alpha d^2 ks\|}\right) \tag{22}$$

on the right of (21) can be estimated by Lemma 8a of Chapter I, with $\lambda = 1$ and a/q replaced by ad_1k_1/q_2. Taking intervals of length $\leqq q_2$ for s, we obtain for the above sum the estimate

$$\left(\frac{N}{dUXq_2} + 1\right)\left(\frac{UX}{d} + q_2 \log q_2\right)$$
$$\ll q^{\varepsilon}\left(\frac{N}{d^2 q_2} + q_2 + \frac{N}{dUX} + \frac{UX}{d}\right).$$

Using this in (21), we obtain

$$(S_d)^2 \ll \frac{KN}{d^2} q^{2\varepsilon} \sum_{k=1}^{K} \left(\frac{N}{d^2 q_2} + q_2 + \frac{N}{dUX} + \frac{UX}{d}\right).$$

We recall that $q_1 = q_2 \varkappa$, where $\varkappa = (q_1, k)$. The values of k which correspond to a particular $\varkappa$ are multiples of $\varkappa$, and their number is at most $K/\varkappa$. Hence

$$(S_d)^2 \ll \frac{KN}{d^2} q^{2\varepsilon} \sum_{\varkappa | q_1} \left(\frac{N}{d^2 q_2} + q_2 + \frac{N}{dUX} + \frac{UX}{d}\right) \frac{K}{\varkappa}$$

$$\ll \frac{K^2 N}{d^2} q^{2\varepsilon} \sum_{\varkappa | q_1} \left(\frac{N}{d^2 q_1} + q_1 + \frac{N}{dUX} + \frac{UX}{d}\right)$$

$$\ll \frac{K^2 N}{d^2} q^{3\varepsilon} \left(\frac{N}{d^2 q_1} + q_1 + \frac{N}{UX} + UX\right)$$

$$\ll \frac{K^2 N}{d^2} q^{3\varepsilon} \left(\frac{N}{q} + q + \frac{N}{UX} + UX\right).$$

Using this in (20), we obtain

$$S' \ll KNq^{2\varepsilon} \left(\frac{1}{q} + \frac{q}{N} + \frac{1}{UX} + \frac{UX}{N}\right)^{\frac{1}{2}} \sum_{d \leqq \Delta^{-1}} \frac{1}{d}.$$

This gives the inequality enunciated (with a different ε).

Case 3. *Suppose that* $q > N^{\frac{9}{10}}$. In this case we must have $\tau > N^{\frac{9}{10}}$ also. Let $\delta = (d^2k, q)$ and put $d^2k = d_1\delta$, $q = q_1\delta$. Then

$$\alpha d^2 k = \frac{a d_1}{q_1} + \frac{\theta d_1}{q_1 \tau},$$

and

$$\frac{d_1}{q_1 \tau} = \frac{d^2 k}{q\tau} \leqq \frac{d^2 k}{q^2} \leqq \frac{d^2 k}{q_1{}^2}.$$

Hence we can apply Lemma 8a of Chapter I to the sum (22), with $\lambda = d^2k$. We obtain for it the estimate

$$\left(\frac{N}{dUXq_1} + 1\right)\left(d^2k \frac{UX}{d} + q_1 \log q_1\right).$$

Here the first term in the first bracket can be omitted, since

$$\frac{N}{dUX} \ll \frac{N^{\frac{3}{4}}}{d} \leqq \frac{N^{\frac{3}{4}} dk}{\delta} \ll \frac{N^{\frac{3}{4}} \Delta^{-3}}{\delta} < \frac{N^{\frac{9}{10}}}{\delta} < q_1.$$

Hence, on noting that $\log q_1 \leqq r < \Delta^{-\varepsilon}$, we obtain on substitution in (21)

$$(S_d)^2 \ll \frac{KN}{d^2} \Delta^{-\varepsilon} \sum_{k=1}^{K} (dkUX + q_1),$$

$$(S_d)^2 \ll \frac{K^2N}{d^2} \Delta^{-\varepsilon}(dKUX + q). \tag{23}$$

We now use the hypothesis that $UX \ll N^{\frac{3}{4}}$. This implies

$$dKUX \ll \Delta^{-3}N^{\frac{3}{4}} \ll \left(\frac{N}{q}\right)^{\frac{3}{2}} N^{\frac{3}{4}} \ll q.$$

Hence

$$(S_d)^2 \ll \frac{N^2K^2\Delta^{-\varepsilon}}{d^2} \left(\frac{q}{N}\right).$$

Finally, by (20),

$$S' \ll NK\Delta^{-\varepsilon} \left(\frac{q}{N}\right)^{\frac{1}{2}} \sum_{d \leq \Delta^{-1}} \frac{1}{d} \ll NK\Delta^{1-2\varepsilon},$$

whence the result.

LEMMA 7. *Suppose that in Lemma 6 the sequences for x and y reduce to the single number 1. Let the other hypotheses be as before, except that* (17) *is replaced by the weaker hypothesis*

$$N^{\frac{1}{4}} \ll U \ll N\Delta^4.$$

Then

$$S \ll KN\Delta^{-\varepsilon} \left(\Delta + \left(\frac{1}{U}\right)^{\frac{1}{2}-\varepsilon}\right).$$

Proof. In this case, since $(x, y) = 1$ necessarily, the summation over d in the proof of Lemma 6 disappears and we can put $d = 1$ throughout.

In Cases 1 and 2 of the proof of Lemma 6, the hypothesis $UX \ll N^{\frac{3}{4}}$ was used only to transfer powers involving ε. In Case 1 we used the hypothesis when replacing N^{ε} by $\left(\frac{1}{UX} + \frac{UX}{N}\right)^{-\varepsilon}$. In Case 2 we used it when replacing $\Delta^{-\varepsilon}$ by the same expression. These replacements are now no longer needed, since the result asserted in the present lemma allows a factor $\Delta^{-\varepsilon}$.

We therefore need only consider the proof of Case 3 above. Putting $d = 1$ and $X = 1$ in (23), we obtain

$$S_1 \ll KN^{\frac{1}{2}}\Delta^{-\varepsilon}(KU + q)^{\frac{1}{2}},$$

and this estimate is valid also for S' since there is now no summation over d in (20). Hence

$$S' \ll KN\Delta^{-\varepsilon}\left(\frac{KU}{N} + \frac{q}{N}\right)^{\frac{1}{2}}$$
$$\ll KN\Delta^{-\varepsilon}\,(\Delta^{-2}\Delta^{4} + \Delta^{2})^{\frac{1}{2}},$$

since $U/N \ll \Delta^4$ and $q/N < \Delta^2$. This gives the estimate stated.

THEOREM 3. *Let* τ *satisfy*

$$N^{\frac{1}{2}} \leqq \tau \leqq N \exp\,(-r^{\varepsilon_0}).$$

Suppose that

$$\alpha = \frac{a}{q} + \frac{\theta}{q\tau},\ \textit{where}\ (a, q) = 1\ \textit{and}\ \exp\,(r^{\varepsilon_0}) \leqq q \leqq \tau.$$

Let

$$\Delta = (q^{-1} + qN^{-1})^{\frac{1}{2}}.$$

Let

$$S = \sum_{k=1}^{K}\left|\sum_{p \leqq N} e(\alpha kp)\right|,$$

where

$$K \ll \Delta^{-2}.$$

Then

$$S \ll KN(\Delta^{1-\varepsilon} + N^{-\frac{1}{5}+\varepsilon}).$$

Proof. Let P denote the product of all primes $p \leqq N^{\frac{1}{5}}$, and let Q denote the product of all primes p satisfying $N^{\frac{1}{5}} < p \leqq N$. For $j = 1, 2, 3, 4$, let D_j run through all distinct products of j distinct primes. Define S_j by

$$S_j = S_j(k) = \sum_{D_j | Q} e(\alpha k D_j).$$

Then obviously

$$S = \sum_{k=1}^{K} | S_1(k) | + O(KN^{\frac{1}{5}}),$$

so it suffices to prove the result for the sum on the right here.

For $s = 1, 2, 3, 4$ we define a sum W_s by

$$W_s = W_s(k) = \sum_{\substack{y_1 | Q \\ y_1 \ldots y_s \leq N}} \ldots \sum_{y_s | Q} e(\alpha k y_1 \ldots y_s).$$

Each y is either 1 or a product of distinct primes all greater than $N^{\frac{1}{5}}$. The contribution to this sum of all terms for which $y_1 \ldots y_s$ is divisible by the square of a prime is

$$\ll \sum_{p > N^{\frac{1}{5}}} \frac{N^{1+\varepsilon}}{p^2} \ll N^{\frac{4}{5}+\varepsilon}.$$

Excluding these terms, each product $y_1 \ldots y_s$ is a product of at most 4 distinct primes, and so is a divisor of Q. Any particular product of j distinct primes, where $j = 1, \ldots, 4$ will occur s^j times as a value of $y_1 \ldots y_s$, since each of the primes can occur in any of $y_1, \ldots, y_s$. Hence

$$sS_1 + s^2 S_2 + \ldots + s^4 S_4 = W_s + O(N^{\frac{4}{5}+\varepsilon}).$$

Thus the sums $S_1, \ldots, S_4$ are related to the sums $W_1, \ldots, W_4$ by the four linear approximative equations

$$\begin{aligned}
S_1 + S_2 + S_3 + S_4 &= W_1 + O(N^{\frac{4}{5}+\varepsilon}),\\
2S_1 + 4S_2 + 8S_3 + 16S_4 &= W_2 + O(N^{\frac{4}{5}+\varepsilon}),\\
3S_1 + 9S_2 + 27S_3 + 81S_4 &= W_3 + O(N^{\frac{4}{5}+\varepsilon}),\\
4S_1 + 16S_2 + 64S_3 + 256S_4 &= W_4 + O(N^{\frac{4}{5}+\varepsilon}).
\end{aligned}$$

It is now plain that in order to prove the desired result it suffices to prove that

$$\sum_{k=1}^{K} (| W_1 | + \ldots + | W_4 |) \ll KN(\Delta^{1-\varepsilon} + N^{-\frac{1}{5}+\varepsilon}).$$

We shall carry out the proof for $\sum_k |W_4|$, and it will be plain that the same method (with rather simpler details) applies to the sums of the other W's.

We have

$$W_4 = \sum_{\substack{y_1|Q \\ y_1 \dots y_4 \leqq N}} \dots \sum_{y_4|Q} e(\alpha k y_1 \dots y_4),$$

and the first step is to express this in another form. We replace the condition $y_1|Q$ by the condition $(y_1, P) = 1$, and so on. The only effect of this is to admit additional terms for which some y is divisible by the square of a prime greater than $N^{\frac{1}{5}}$. The sum of such terms is $O(N^{\frac{4}{5}+\varepsilon})$, as already proved. Hence it suffices to consider W_4', given by

$$W_4' = \sum_{\substack{(y_1, P)=1 \\ y_1 \dots y_4 \leqq N}} \dots \sum_{(y_4, P)=1} e(\alpha k y_1 \dots y_4).$$

In accordance with Lemma 2, we can express this alternatively as

$$W_4' = \sum_{d_1|P} \mu(d_1) \sum_{\substack{m_1 \\ d_1 m_1 \dots d_4 m_4 \leqq N}} \dots \sum_{d_4|P} \mu(d_4) \sum_{m_4} e(\alpha k d_1 m_1 \dots d_4 m_4),$$

where the d's and m's take all positive integral values subject to the restrictions indicated.

We now apply Lemma 5 to each of the four variables d_1, d_2, d_3, d_4. The values of d_1 can be distributed among at most D disjoint sets which satisfy (i) and (ii) of Lemma 5, and similarly for d_2, d_3, d_4. We also split up the interval $0 < m_1 \leqq N$ into $\ll r$ subintervals, each of the form

$$M_1 < m \leqq M_1', \text{ where } M_1 < M_1' \leqq 2M_1,$$

and similarly for m_2, m_3, m_4.

Now consider the contribution to $\sum_k |W_4'|$ corresponding to four sets of values for d_1, d_2, d_3, d_4 and to four subintervals for m_1, m_2, m_3, m_4. This is of the form

$$T = \sum_{k=1}^{K} \left| \sum_{\substack{d_1 \\ d_1 \dots m_4 \leqq N}} \dots \sum_{m_4} e(\alpha k d_1 \dots m_4) \right|$$

where d_1, d_2, d_3, d_4 run through certain integers in the intervals

$$\varphi_1 \leqq d_1 \leq \varphi_1^{1+c}, \ldots, \varphi_4 \leqq d_4 < \varphi_4^{1+c},$$

and m_1, m_2, m_3, m_4 run through all integers in the intervals

$$M_1 < m_1 \leqq M_1', \ldots, M_4 < m_4 \leqq M_4'.$$

The number of sums such as T which we have to consider is

$$\ll r^4 D^4 \ll r^4 \exp\left((\log r)^3\right) \ll \Delta^{-\varepsilon},$$

and consequently it suffices to prove the inequality of the enunciation for each sum T. We can restrict ourselves to cases in which

$$M_1 \ldots M_4 \; \varphi_1 \ldots \varphi_4 > N^{\frac{4}{5}}, \tag{24}$$

for in other cases we have the trivial estimate

$$T \ll KN^{\frac{4}{5}(1+c)},$$

where c can be taken arbitrarily small, and this gives the desired result.

Case 1. *Suppose that* $M_1 M_2 M_3 M_4 \leqq N^{\frac{2}{5}}$. Let t be the least integer for which

$$M_1 M_2 M_3 M_4 \varphi_1 \ldots \varphi_t > N^{\frac{2}{5}};$$

plainly $1 \leqq t \leqq 4$. Define γ by

$$M_1 M_2 M_3 M_4 \varphi_1 \ldots \varphi_{t-1} N^\gamma = N^{\frac{2}{5}}, \tag{25}$$

so that $\varphi_t > N^\gamma$. Let $\sigma = \frac{1}{5}$; then $0 < \gamma \leqq 1 - \sigma$.

The set of values of d_t under consideration satisfies the condition $\varphi_t > N^\gamma$ of part (ii) of Lemma 5. Hence there exist two increasing sequences of positive integers x and y, with

$$N^\gamma < x \leqq N^{(\gamma+\frac{1}{5})(1+c)}, \tag{26}$$

such that the values of xy with $(x, y) = 1$ and $xy \leqq N$ give all the numbers d_t, each with the same multiplicity.

We put $u = m_1 m_2 m_3 m_4 d_1 \ldots d_{t-1}$, and we divide the values of u into $\ll r$ sets, each of which is contained in an interval

$$U < u \leqq U', \text{ where } U < U' \leqq 2U. \tag{27}$$

We also divide the values of x into $\ll r$ sets, each of which is contained in an interval

$$(28) \qquad X < x \leqq X', \text{ where } X < X' \leqq 2X.$$

If T_1 is the part of the sum T which corresponds to two such intervals of u and x, we have

$$sT_1 = \sum_{k=1}^{K} \left| \sum_u \sum_x \sum_y \sum_v e(\alpha kuxyv) \right|,$$

where u runs through integers of the interval (27) which are of the form $m_1m_2m_3m_4d_1 \ldots d_{t-1}$, and x runs through those numbers of its particular sequence lying in the interval (28), and where y runs through the sequence already mentioned, and where v runs through integers of the form $d_{t+1} \ldots d_4$ (or $v = 1$ if this product is empty). Also s is a positive integer, namely the multiplicity of the representations of the numbers d_t in the form xy.

The above multiple sum is of the type estimated in Lemma 6. All the hypotheses of that lemma are satisfied (c_1 and c_2 being at most 7), except possibly those relating to the magnitude of UX. As regards this, we have

$$UX \gg (M_1M_2M_3M_4\varphi_1 \ldots \varphi_{t-1})N^{\gamma} = N^{\frac{2}{5}}$$

by (25) and (26), and

$$\begin{aligned} UX &\ll M_1M_2M_3M_4(\varphi_1 \ldots \varphi_{t-1})^{1+c}N^{(\gamma+\frac{1}{5})(1+c)} \\ &\ll N^{\frac{2}{5}(1+c)+\frac{1}{5}(1+c)} = N^{\frac{3}{5}(1+c)}. \end{aligned}$$

Thus, if c is sufficiently small, all the hypotheses of Lemma 6 are satisfied, and it follows that

$$T_1 \ll KN(\Delta^{1-\varepsilon} + N^{-\frac{1}{5}+\varepsilon}).$$

Since T_1 was any one of $\ll r^2$ parts of T, the desired result follows.

Case 2. Suppose that $M_1M_2M_3M_4 > N^{\frac{2}{5}}$ *and that some product of M's lies between* $N^{\frac{2}{5}}$ *and* $N^{\frac{3}{5}}$. (The product of M's may be a single M or a product of two or three distinct M's or may be $M_1M_2M_3M_4$.) By permuting the M's we can suppose that

$$N^{\frac{2}{5}} \leqq M_1 \ldots M_t \leqq N^{\frac{3}{5}}$$

for some t with $1 \leqq t \leqq 4$.

We introduce variables u and v by writing

$$u = m_1 \ldots m_t, \quad v = m_{t+1} \ldots m_4 d_1 d_2 d_3 d_4.$$

Then u is restricted to an interval of the form

$$U < u \leqq U', \text{ where } U < U' \leqq 16U \text{ and } N^{\frac{2}{5}} \leqq U \leqq N^{\frac{3}{5}}.$$

We apply Lemma 6 with the sequences for x and y consisting each of the single number 1, so that in particular we can replace X by 1 in the conclusion. It follows from Lemma 6 that

$$T_1 \ll KN\left(\Delta^{1-\varepsilon} + \left(\frac{1}{U} + \frac{U}{N}\right)^{\frac{1}{2}-\varepsilon}\right)$$

$$\ll KN(\Delta^{1-\varepsilon} + N^{-\frac{1}{5}+\varepsilon}),$$

whence the result, as before.

Case 3. Suppose that $M_1M_2M_3M_4 > N^{\frac{2}{5}}$ *and that no product of M's lies between* $N^{\frac{2}{5}}$ *and* $N^{\frac{3}{5}}$. In particular, therefore,

$$M_1M_2M_3M_4 > N^{\frac{3}{5}}.$$

We can suppose without loss of generality that $M_1 \leqq M_2 \leqq M_3 \leqq M_4$.

We prove first that $M_4 > N^{\frac{1}{4}}$. In the contrary case, we should have $M_3 \leqq M_4 \leqq N^{\frac{1}{4}}$, whence $M_3M_4 \leqq N^{\frac{1}{2}}$ and therefore, by the hypothesis of this case, $M_3M_4 < N^{\frac{2}{5}}$. This implies $M_2 \leqq M_3 \leqq N^{\frac{1}{5}}$. Hence $M_2M_3M_4 \leqq N^{\frac{3}{5}}$, and so, by the hypothesis, $M_2M_3M_4 < N^{\frac{2}{5}}$. But now $M_1M_2M_3M_4 < N^{\frac{3}{5}}$, contrary to the inequality stated above.

We consider two cases. Suppose first that

$$M_4 \leqq N\Delta^4 \text{ and either } q < N^{\frac{1}{10}} \text{ or } q > N^{\frac{9}{10}}.$$

We apply the special form of Lemma 6 given in Lemma 7, taking

$$u = m_4, \ v = m_1 m_2 m_3 d_1 d_2 d_3 d_4.$$

We can take $U = M_4$, since then $N^{\frac{1}{4}} < U \leqq N\Delta^4$. The result is

$$\begin{aligned} T_1 &\ll KN\Delta^{-\varepsilon}(\Delta + (M_4)^{-\frac{1}{2}+\varepsilon}) \\ &\ll KN\Delta^{1-3\varepsilon}, \end{aligned}$$

since $M_4 > N^{\frac{1}{4}} > \Delta^{-2}$ because of the hypothesis on q. Thus the result follows.

Suppose next that

$$\text{either } M_4 > N\Delta^4 \text{ or } N^{\frac{1}{10}} \leqq q \leqq N^{\frac{9}{10}}.$$

In the latter of these alternative cases, $N \ll \Delta^{-20}$. In the former, every value of

$$d_1 d_2 d_3 d_4 m_1 m_2 m_3$$

is $\leqq N/M_4 < \Delta^{-4}$, and also $k \leqq K \ll \Delta^{-2}$. Hence, in either case, the number of representations of any integer z as $kd_1d_2d_3d_4m_1m_2m_3$ is $\ll \Delta^{-\varepsilon}$. Thus

$$T_1 \ll \Delta^{-\varepsilon} \sum_z \Big| \sum_{m_4} e(\alpha z m_4) \Big|,$$

where $0 < z \leqq KN/M_4$ and m_4 runs through all integers of an interval of length $\ll KN/z$. It follows that

$$T_1 \ll \Delta^{-\varepsilon} \sum_z \min\left(\frac{KN}{z}, \frac{1}{\|\alpha z\|}\right).$$

Applying Lemma 8b of Chapter I, with $W = KN$ and $W_0 = KN/M_4$, we obtain

$$\begin{aligned} T_1 &\ll \Delta^{-\varepsilon}\left(\frac{KN}{M_4} + q + \frac{KN}{q}\right) \log (KN) \\ &\ll KN\Delta^{-2\varepsilon}\left(\frac{1}{M_4} + \frac{q}{N} + \frac{1}{q}\right) \\ &\ll KN\Delta^{-2\varepsilon}(N^{-\frac{1}{4}} + q^{-1} + qN^{-1}) \\ &\ll KN(N^{-\frac{1}{5}+\varepsilon} + \Delta^{2-2\varepsilon}), \end{aligned}$$

whence the result.

NOTES ON CHAPTER IX

The proofs in this chapter have been given in greater detail than in the original. It was found necessary to modify the final stages of the proof of Theorem 3, and in this connection a special form of Lemma 6, given in Lemma 7, has been introduced.

Theorems 1 and 2b give estimates for the sum

$$\sum_{p \leqq N} e(\alpha p)$$

which are used in the solution of Goldbach's Problem for three primes in Chapter X. The estimate of Theorem 1 is significant only if $q > r^{11}$ (and of course $q < Nr^{-11}$). The estimate of Theorem 2b, on the other hand, is significant if q is greater than any fixed positive power of r, however small, provided again that q is not too large. As q is small, it is natural that the proof of Theorem 2a, on which Theorem 2b depends, should make use of the regularity of distribution of the primes among the residue classes (mod q).

It is possible to obtain the result of Chapter X without using Theorems 2a and 2b by basing the proof on Siegel's class-number theorem instead of on Page's work: see the Notes on Chapter X.

Though we do not propose to comment in detail on the contents of Chapter IX, we would point out that Lemma 1, which is fundamental for the work of the chapter, gives an estimate for a certain sum of the general form $\sum\sum e(\alpha uv)$ mentioned in the Note on Vinogradov's Method. Many of the later complications are due to the need for arranging that in applications of this lemma the number $U^{-1} + UN^{-1}$ shall be reasonably small.

CHAPTER X

Goldbach's Problem

In the present chapter I give a solution of Goldbach's problem concerning the representability of every sufficiently large odd number N as the sum of three primes, and I establish an asymptotic formula for the number of representations.

The method used here enables one also to solve more general additive problems involving primes, for example the question of the representability of large numbers N in the form

$$N = p_1^n + \ldots + p_s^n$$

(Waring's problem for primes). But I do not consider these more general questions here.

For the solution of Goldbach's problem I express the number of representations by an integral, similar to the one first used by Hardy and Littlewood in the same connection. As in Chapters IV and VII, I divide the interval of integration into basic and supplementary intervals. For the contribution of the basic intervals, a general method (that of Estermann, based on Page's results on the distribution of primes in arithmetical progressions) had been worked out by English scholars not long before the appearance of my work on Goldbach's problem in 1937. This method applies both to Goldbach's problem and to more general additive problems with primes, and it is based on the modern theory of L-series. In the present chapter I treat the contribution of the basic intervals by using a simplified form of Page's result (Lemma 1) in conjunction with Theorem 2a of Chapter IX, which was based on Brun's method. In estimating the contribution of the supplementary intervals I use solely my own method, in the form of Theorems 1 and 2b of Chapter IX.

Notation in this chapter. The letter p will always denote a prime.

N will be an arbitrarily large positive integer, and we put $r = \log N$.

LEMMA I (*Page*). *Let* $\pi(N; q, l)$ *denote the number of primes* p *satisfying* $p \leqq N$ *and* $p \equiv l \pmod q$, *where* $(l, q) = 1$. *Let* ε_0, c, c_1 *be fixed positive numbers, and suppose that*

$$0 < q \leqq r^{c_1}.$$

Then, if q *is not exceptional, we have*

$$\pi(N; q, l) = \frac{1}{\varphi(q)} \int_2^N \frac{dx}{\log x} + O\left(\frac{Nr^{-c}}{r\varphi(q)}\right).$$

The exceptional values of q, *if any, are multiples of some one number* q_0 *which satisfies*

$$q_0 > r^{2-\varepsilon_0}.$$

Proof. This result is a deduction from those of A. Page, *Proc. London Math. Soc.* (2), 39 (1935), 116—141. The proof depends on the theory of L-series developed by Dirichlet, Riemann, Hadamard, de la Vallée Poussin, Hardy and Littlewood, Landau and others; and cannot be given here.

Page's Theorem 1 asserts that

$$\pi(N; q, l) = \frac{1}{\varphi(q)} \int_2^N \frac{dx}{\log x} + O(N \exp(-Cr^{\frac{1}{2}})) + O\left(\frac{N^{\sigma_1}}{r\varphi(q)}\right),$$

where $\sigma_1 = \sigma_1(q)$ is the greatest real zero possessed by any L-function to the modulus q, and C is a positive absolute constant. As regards the first error term here, we have

$$N \exp(-Cr^{\frac{1}{2}}) \ll Nr^{-c-c_1-1} \ll \frac{Nr^{-c}}{r\varphi(q)},$$

since $\varphi(q) \leqq q \leqq r^{c_1}$. As regards the second error term, this also will be of the form stated in the present lemma if $N^{1-\sigma_1} \geqq r^c$, that is, if

$$\sigma_1 \leqq 1 - \frac{c \log r}{r}. \tag{1}$$

Page's Lemma 9 asserts that the more precise inequality $\sigma_1 < 1 - C_1/(\log r)$ is true for all but certain exceptional values of q, these exceptional values being multiples of some one number q_0. Also $\sigma_1(q) = \sigma_1(q_0)$. Now Page's Theorem 2 asserts that

$$\sigma_1(q_0) \leqq 1 - \frac{C_2}{q_0^{\frac{1}{2}} \log^2 (q_0 + 1)},$$

and since (1) is supposed not to hold for σ_1 we obtain

$$c q_0^{\frac{1}{2}} \log^2 (q_0 + 1) > C_2 r/(\log r).$$

This implies the lower bound for q_0 stated in the present lemma, if N is sufficiently large.

LEMMA 2. *Let* $\tau = N r^{-c}$, *where* $c \geqq 4$, *and let*

$$R = \int_{-1/\tau}^{1/\tau} (J(z))^3 e(-zN) dz,$$

where

$$J(z) = \int_2^N \frac{e(zx)}{\log x} dx.$$

Then

$$R = \frac{N^2}{2r^3} + O\left(\frac{N^2}{r^4}\right).$$

Proof. Let

$$I(z) = \int_2^N \frac{e(zx)}{r} dx.$$

Then

$$| J(z) - I(z) | \leqq \int_2^N \left(\frac{1}{\log x} - \frac{1}{r}\right) dx \ll N r^{-2}.$$

Also, for $| z | \leqq 1/\tau$, both $J(z)$ and $I(z)$ are majorized by

$$r^{-1} \min (N, | z |^{-1}),$$

by Lemma 14b of Chapter I; the condition $| z | < N^{-\frac{1}{2}}$ being amply satisfied if $| z | \leqq 1/\tau$. Hence

$$\int_{-1/\tau}^{1/\tau} \left| (J(z))^3 - (I(z))^3 \right| dz \ll \int_0^\infty N r^{-2}(r^{-1} \min (N, z^{-1}))^2 dz$$
$$\ll N^2 r^{-4}.$$

Further, since $I(z) \ll r^{-1} | z |^{-1}$ always, we have

$$\int_{1/\tau}^{\frac{1}{2}} | I(z) |^3 \, dz \ll r^{-3} \int_{1/\tau}^{\infty} z^{-3} dz \ll N^2 r^{-4}.$$

These two results show that if

$$R_0 = \int_{-\frac{1}{2}}^{\frac{1}{2}} (I(z))^3 e(-zN) dz,$$

then

$$R - R_0 \ll N^2 r^{-4}.$$

Next, putting

$$S(z) = r^{-1} \sum_{x=3}^{N} e(zx),$$

we have $I(z) - S(z) \ll r^{-1}$ for $| z | \leqq \frac{1}{2}$, as a special case of Lemma 13 of Chapter I (or directly). Hence

$$\int_{-\frac{1}{2}}^{\frac{1}{2}} \left| (I(z))^3 - (S(z))^3 \right| dz \ll \int_0^{\frac{1}{2}} r^{-1}(r^{-1} \min (N, z^{-1}))^2 dz \ll N r^{-3}.$$

Thus, if

$$R_1 = \int_{-\frac{1}{2}}^{\frac{1}{2}} (S(z))^3 e(-zN) dz,$$

then

$$R - R_1 \ll N^2 r^{-4}.$$

Now, by the definition of $S(z)$,

$$r^3 R_1 = \int_{-\frac{1}{2}}^{\frac{1}{2}} \sum_{x_1=3}^{N} \sum_{x_2=3}^{N} \sum_{x_3=3}^{N} e(z(x_1 + x_2 + x_3 - N)) dz,$$

and this equals the number of representations of N as $x_1 + x_2 + x_3$ where x_1, x_2, x_3 are integers $\geqq 3$. Plainly this number is $\frac{1}{2}N^2 + O(N)$, whence

$$R_1 = r^{-3}(\tfrac{1}{2}N^2 + O(N)),$$

and the result follows.

THEOREM. *The number $I(N)$ of representations of an odd positive integer N as $p_1 + p_2 + p_3$ can be expressed by the formula*

$$I(N) = \frac{N^2}{2r^3}\mathfrak{S}(N) + O\left(\frac{N^2}{r^{\frac{7}{2}-\varepsilon}}\right),$$

where $\varepsilon > 0$ and

$$\mathfrak{S}(N) = \prod_p \left(1 + \frac{1}{(p-1)^3}\right) \prod_{p|N} \left(1 - \frac{1}{p^2 - 3p + 3}\right),$$

the first product being extended over all primes. Moreover, the product $\mathfrak{S}(N)$ satisfies

$$\mathfrak{S}(N) > \frac{6}{\pi^2}$$

for all N.

COROLLARY (*Goldbach's theorem*). *There exists a number c_0 such that every odd integer $N \geqq c_0$ can be represented as the sum of three primes.*

Proof. Define $S(\alpha)$ by

$$S(\alpha) = \sum_{p \leq N} e(\alpha p).$$

By Lemma 4 of Chapter I, we have

$$I(N) = \int_{-1/\tau}^{1-1/\tau} (S(\alpha))^3 e(-\alpha N) d\alpha.$$

We divide the interval of integration into basic and supplementary intervals. Let

$$\tau = Nr^{-14}.$$

We define the basic intervals to consist of all α of the form

$$(2) \quad \alpha = \frac{a}{q} + z, \text{ where } (a, q) = 1, \ |z| \leqq 1/\tau, \ 0 < q \leqq r^3;$$

and the intervals which remain after the removal of these from the interval of integration will be the supplementary intervals. Plainly no two basic intervals overlap, since (with an obvious notation)

$$\left| \frac{a_1}{q_1} - \frac{a_2}{q_2} \right| \geqq \frac{1}{q_1 q_2} \geqq r^{-6} > \frac{2}{\tau} \geqq | z_1 | + | z_2 |.$$

By Lemma 7 of Chapter I, any α of a supplementary interval can be represented by

$$(3) \quad \alpha = \frac{a}{q} + z, \text{ where } (a, q) = 1, \ | z | \leqq (\tau q)^{-1}, \ r^3 < q \leqq \tau.$$

We write $I(N)$ as

$$I(N) = I_1(N) + I_2(N),$$

where $I_1(N)$ denotes the contribution of the basic intervals to the integral for $I(N)$, and $I_2(N)$ denotes the contribution of the supplementary intervals.

The proof of the theorem falls into five stages.

1. *The estimation of* $I_2(N)$. Let α be in a supplementary interval, and therefore of the form (3). If $q > r^{14}$ we apply Theorem I of Chapter IX, the hypotheses of which are satisfied since $| z | \leqq (\tau q)^{-1} \leqq q^{-2}$. We obtain

$$S(\alpha) \ll N r^{\frac{9}{2}} \left((q^{-1} + qN^{-1})^{\frac{1}{2}} + \exp\left(- \tfrac{1}{2} r^{\frac{1}{3}}\right) \right) \ll N r^{-\frac{5}{2}}.$$

If $r^3 < q \leqq r^{14}$ we apply Theorem 2b of Chapter IX, the hypotheses of which are satisfied, and obtain

$$S(\alpha) \ll \frac{N (rq)^{5\varepsilon}}{r\sqrt{q}} \ll N r^{-\frac{5}{2} + \varepsilon_1}.$$

Hence

$$I_2(N) \ll N r^{-\frac{5}{2} + \varepsilon_1} \int_0^1 | S(\alpha) |^2 \, d\alpha.$$

The integral on the right represents the number of solutions of $p = p' \leqq N$, and is therefore $\ll N r^{-1}$. Hence

$$I_2(N) \ll N^2 r^{-\frac{7}{2} + \varepsilon_1}.$$

2. *Basic intervals corresponding to a non-exceptional value of* q. Putting $c_1 = 3$ and $c = 48$ in Lemma 1, we have

$$\text{(4)} \qquad \pi(N; q, l) = \frac{1}{\varphi(q)} \int_2^N \frac{dx}{\log x} + O\left(\frac{Nr^{-49}}{\varphi(q)}\right)$$

for $q \leqq r^3$, provided that q is not exceptional. We shall apply this result with N replaced by various other numbers. Strictly speaking, the concept of "exceptional" and "non-exceptional" values of q is one which is relative to N. But in the applications of (4) which are made later, N will be replaced by numbers whose logarithms are always asymptotically equal to r, and it is apparent from the proof of Lemma 1 that the same definition of "exceptional q" and of q_0 will serve for all such numbers.

Let $I_{a,q}$ denote the contribution to $I_1(N)$ made by an individual basic interval corresponding to a fraction a/q for which q is not exceptional. Let

$$D = [r^{31}], \ A = N/D.$$

We split the sum $S(\alpha)$ into $D - 1$ sums of the form

$$S(\alpha, N_1) = \sum_{N_1 - A < p \leq N_1} e(\alpha p),$$

where N_1 takes the values

$$\text{(5)} \qquad N_1 = N - sA \text{ for } s = 0, 1, \ldots, D - 2.$$

There remains a sum over $p \leqq N - (D - 1)A = A$, and this is $\ll Nr^{-31}$. Hence

$$\text{(6)} \qquad S(\alpha) = \sum_{N_1} S(\alpha, N_1) + O(Nr^{-31}),$$

where N_1 takes the values (5). The lower limit of any sum is at least A, the logarithm of which is asymptotically equal to r.

In each term of the sum $S(\alpha, N_1)$, we have

$$e(zp) = e(zN_1) + O(|z| A).$$

The number of terms in the sum is

$$\int_{N_1 - A}^{N_1} \frac{dx}{\log x} + O(Nr^{-49}) = O(Ar^{-1}),$$

using the special case $q = 1$ of (4). Hence, since

$$(|z| \cdot A)Ar^{-1} \ll \tau^{-1}A^2r^{-1} \ll Ar^{-18},$$

we have

$$S(\alpha, N_1) = e(zN_1) \sum_{N_1 - A < p \leqq N_1} e_q(ap) + O(Ar^{-18}). \tag{7}$$

The number of terms in the sum on the right satisfying $p \equiv l \pmod{q}$, where $(l, q) = 1$, is

$$\frac{1}{\varphi(q)} \int_{N_1 - A}^{N_1} \frac{dx}{\log x} + O\left(\frac{Ar^{-18}}{\varphi(q)}\right),$$

by Lemma 1 in the form (4). The number of terms for which p divides q is $\ll q \ll r^3 \ll Ar^{-18}$. Hence

$$\begin{aligned} S(\alpha, N_1) &= \frac{1}{\varphi(q)} \sum_{\substack{l=0 \\ (l,q)=1}}^{q-1} e_q(al)e(zN_1) \int_{N_1-A}^{N_1} \frac{dx}{\log x} + O(Ar^{-18}) \\ &= \frac{\mu(q)}{\varphi(q)} \int_{N_1-A}^{N_1} \frac{e(zN_1)}{\log x} dx + O(Ar^{-18}) \\ &= \frac{\mu(q)}{\varphi(q)} \int_{N_1-A}^{N_1} \frac{e(zx)}{\log x} dx + O(Ar^{-18}), \end{aligned}$$

since $e(zN_1) - e(zx) \ll |z| A$ in the interval of integration, and $|z| A \cdot Ar^{-1} \ll Ar^{-18}$ as verified above.

Adding the various sums $S(\alpha, N_1)$, we obtain from (6)

$$S(\alpha) = \frac{\mu(q)}{\varphi(q)} J(z) + O(Nr^{-18}).$$

The contribution of the particular basic interval in question is given by

$$I_{a,q} = \int_{-1/\tau}^{1/\tau} \left(S\left(\frac{a}{q} + z\right)\right)^3 e\left(-\left(\frac{a}{q} + z\right)N\right) dz.$$

Now $J(z) \ll r^{-1} \min(N, |z|^{-1})$ in the interval of integration, by Lemma 14b of Chapter I. Hence

$$\int_{-1/\tau}^{1/\tau} \left| \left(S\left(\frac{a}{q}+z\right)\right)^3 - \left(\frac{\mu(q)}{\varphi(q)} J(z)\right)^3 \right| dz$$
$$\ll \frac{N\gamma^{-18}}{\varphi^2(q)} \int_0^{-1/\tau} (\gamma^{-1} \min (N, z^{-1}))^2 dz \ll \frac{N^2\gamma^{-20}}{\varphi^2(q)}.$$

Hence

$$I_{a,q} = \frac{\mu(q)}{(\varphi(q))^3} e_q(-Na) \int_{-1/\tau}^{1/\tau} (J(z))^3 e(-zN) dz + O\left(\frac{N^2\gamma^{-20}}{(\varphi(q))^2}\right).$$

Summing for a over $0 \leqq a < q$, $(a, q) = 1$, we obtain

$$\sum_a I_{a,q} = G(q)R + O\left(\frac{N^2\gamma^{-20}}{\varphi(q)}\right),$$

where

$$G(q) = \frac{\mu(q)}{(\varphi(q))^3} \sum_a e_q(-Na)$$

and

$$R = \int_{-1/\tau}^{1/\tau} (J(z))^3 e(-zN) dz.$$

By Lemma 2 together with the obvious inequality $|G(q)| \leqq (\varphi(q))^{-2}$, it follows that

$$\sum_a I_{a,q} = \frac{N^2}{2\gamma^3} G(q) + O\left(\frac{N^2}{\gamma^4 (\varphi(q))^2}\right).$$

3. *Basic intervals corresponding to an exceptional value of q.* Again we split up $S(\alpha)$ into $D - 1$ sums $S(\alpha, N_1)$ as in (6), where N_1 takes the values (5); and again we express $S(\alpha, N_1)$ in the form (7).

The contribution $I_{a,q}$ of an individual basic interval is

$$I_{a,q} = \int_{-1/\tau}^{1/\tau} \left(S\left(\frac{a}{q}+z\right)\right)^3 e\left(-\left(\frac{a}{q}+z\right)N\right) dz$$
$$= \sum_{N_1} \sum_{N_2} \sum_{N_3} I_{a,q}(N_1, N_2, N_3) + O(N^2\gamma^{-31}),$$

where N_1, N_2, N_3 assume the values (5) and where

$I_{a,q}(N_1, N_2, N_3) =$

$$\int_{-1/\tau}^{1/\tau} S\left(\frac{a}{q} + z, N_1\right) S\left(\frac{a}{q} + z, N_2\right) S\left(\frac{a}{q} + z, N_3\right) e\left(-\left(\frac{a}{q} + z\right)N\right) dz,$$

and where the error term $O(N^2\gamma^{-31})$ arises from the error term in (6). If we replace each S here by the main term on the right of (7), the error introduced is

$$\ll \tau^{-1}(A\gamma^{-1})^2 A\gamma^{-18} \ll A^3 N^{-1}\gamma^{-6}.$$

Hence

$$I_{a,q}(N_1, N_2, N_3) = S'(N_1)S'(N_2)S'(N_3)e_q(-Na)W + O(A^3N^{-1}\gamma^{-6}),$$

where

$$S'(N_1) = \sum_{N_1 - A < p \leq N} e_q(ap)$$

and

$$W = \int_{-1/\tau}^{1/\tau} e(z(N_1 + N_2 + N_3 - N))dz.$$

To obtain an estimate for the sums S' we appeal to Theorem 2a of Chapter IX. This shows that

$$S'(N_1) \ll \frac{A\gamma^{\varepsilon_2}}{r\sqrt{q}}.$$

To estimate W we observe that if $N_1 = N - s_1 A$, etc., as in (5), then

$$W \ll \min\left(\tau^{-1}, |A(2D - s_1 - s_2 - s_3)|^{-1}\right).$$

Substituting these estimates in the expression for $I_{a,q}(N_1, N_2, N_3)$ and then summing over s_1, s_2, s_3, we obtain

$$I_{a,q} \ll \left(\frac{A\gamma^{\varepsilon_2}}{r\sqrt{q}}\right)^3 \sum_{s_1=0}^{D-1} \sum_{s_2=0}^{D-1} \sum_{s_3=0}^{D-1} \min\left(\tau^{-1}, |A(2D - s_1 - s_2 - s_3)|^{-1}\right)$$
$$+ N^3N^{-1}\gamma^{-6} + N^2\gamma^{-31}.$$

For a given integer h, the number of solutions of the equation $2D - s_1 - s_2 - s_3 = h$ will be $\ll D^2$. Hence

$$\sum_{s_1=0}^{D-1}\sum_{s_2=0}^{D-1}\sum_{s_3=0}^{D-1}\min\left(\tau^{-1},\ |A(2D-s_1-s_2-s_3)|^{-1}\right)$$
$$\ll D^2\tau^{-1}+D^2\sum_{h=1}^{2D}(hA)^{-1}\ll D^2A^{-1}\gamma^{\varepsilon_3}.$$

Thus

$$I_{a,q}\ll\frac{N^2\gamma^{\varepsilon_4}}{\gamma^3q^{\frac{3}{2}}}+N^2\gamma^{-6}\ll\frac{N^2\gamma^{\varepsilon_4}}{\gamma^3q^{\frac{3}{2}}}.$$

Summing over a, for a particular q, we obtain

$$\sum_a I_{a,q}\ll\frac{N^2\gamma^{\varepsilon_4}}{\gamma^3q^{\frac{1}{2}}}.$$

It will be convenient to express this in a form comparable with that of the corresponding result obtained in the previous stage. Since $N^2\gamma^{-3}G(q)\ll N^2\gamma^{-3}(\varphi(q))^{-2}$, we can write the result as

$$\sum_a I_{a,q}=\frac{N^2}{2\gamma^3}G(q)+O\left(\frac{N^2\gamma^{\varepsilon_4}}{\gamma^3q^{\frac{1}{2}}}\right).$$

4. *The formula for $I(N)$.* By the results of the last two stages, on recalling that the exceptional values of q considered in the last stage are multiples of q_0, we have

$$I_1(N)-\frac{N^2}{2\gamma^3}\sum_{q\leqq\gamma^3}G(q)$$
$$\ll\sum_{q\leqq\gamma^3}\frac{N^2}{\gamma^4(\varphi(q))^2}+\sum_{s\leqq\gamma^3/q_0}\frac{N^2\gamma^{\varepsilon_4}}{\gamma^3(q_0s)^{\frac{1}{2}}}$$
$$\ll\frac{N^2}{\gamma^4}+\frac{N^2\gamma^{\varepsilon_4}}{\gamma^3q_0^{\frac{1}{2}}}\left(\frac{\gamma^3}{q_0}\right)^{\frac{1}{2}}\ll\frac{N^2\gamma^{\varepsilon_5}}{\gamma^{\frac{7}{2}}},$$

since $q_0\gg\gamma^{2-\varepsilon_0}$. Also

$$\frac{N^2}{2\gamma^3}\sum_{q>\gamma^3}G(q)\ll\frac{N^2}{\gamma^3}\sum_{q>\gamma^3}(\varphi(q))^{-2}\ll\frac{N^2}{\gamma^5}.$$

Thus, taking into account the estimate for $I_2(N)$ obtained in the first stage, we have

$$I(N)=\frac{N^2}{2\gamma^3}\sum_{q=1}^{\infty}G(q)+O\left(\frac{N^2\gamma^{\varepsilon}}{\gamma^{\frac{7}{2}}}\right).$$

5. *The investigation of* $\mathfrak{S}(N)$. We first prove that if $(q_1, q_2) = 1$ then

$$G(q_1)G(q_2) = G(q_1 q_2).$$

Since $\mu(q)$ and $\varphi(q)$ are well known to be multiplicative, it suffices to prove that the sum

$$\sum_{\substack{a=0 \\ (a,q)=1}}^{q-1} e_q(-Na)$$

is a multiplicative function of q. This is immediate, since if a_1 runs through a reduced set of residues (mod q_1) and a_2 runs through a reduced set of residues (mod q_2), then $a_1 q_2 + a_2 q_1$ runs through a reduced set of residues (mod $q_1 q_2$).

The series

$$\mathfrak{S}(N) = \sum_{q=1}^{\infty} G(q)$$

is absolutely convergent, since $G(q) \ll (\varphi(q))^{-2}$. Putting

$$\xi_p = 1 + G(p) + G(p^2) + \dots$$

we have

$$\mathfrak{S}(N) = \prod_p \xi_p.$$

Now $G(p^s) = 0$ if $s > 1$, since then $\mu(p^s) = 0$. Also

$$G(p) = -\frac{1}{(p-1)^3} \sum_{a=1}^{p-1} e_p(-Na) = \begin{cases} (p-1)^{-3} & \text{if } N \text{ is not divisible by } p, \\ -(p-1)^{-2} & \text{if } N \text{ is divisible by } p. \end{cases}$$

Hence

$$\mathfrak{S}(N) = \prod_{p \nmid N} \left(1 + \frac{1}{(p-1)^3}\right) \prod_{p \mid N} \left(1 - \frac{1}{(p-1)^2}\right). \tag{8}$$

This is equivalent to the formula stated in the enunciation, since

$$\frac{1-(p-1)^{-2}}{1+(p-1)^{-3}} = 1 - \frac{1}{p^2 - 3p + 3}.$$

As regards the magnitude of $\mathfrak{S}(N)$, we have, by (8),

$$\mathfrak{S}(N) > \prod_{p \mid N} \left(1 - \frac{1}{(p-1)^2}\right)$$
$$> \prod_{p} \left(1 - \frac{1}{p^2}\right) = \frac{6}{\pi^2},$$

since N is odd and consequently in the former product we have $p - 1 \geqq p_1$, where p_1 is the greatest prime less than p.

This completes the proof of the theorem.

NOTES ON CHAPTER X

This chapter has been expanded, and details of arguments supplied, but otherwise there is no significant change. The deduction of Lemma 1 from the results explicity stated by Page has been inserted, and attention has been drawn to the relative nature of the concept of "exceptional q".

It will be noted that in the present chapter the lengths of the basic intervals are defined by $|z| \leqq \tau^{-1}$ instead of by $|z| \leqq q^{-1}\tau^{-1}$ as in the chapters on Waring's Problem. The reason why this simplification is possible lies in the small order of magnitude of q in the present definition.

An alternative way of obtaining the results of this chapter is to base the treatment on a theorem of Siegel instead of on the work of Page. Siegel's theorem implies that

$$\pi(N; q, l) = \frac{1}{\varphi(q)} \int_2^N \frac{dx}{\log x} + O(N \exp(-c (\log N)^{\frac{1}{2}})),$$

for some positive absolute constant c. This is a stronger result than that of Page, and there are no exceptional values of q. It enables one to approximate to $\sum e(\alpha p)$ when α is near to a rational number a/q with $q \leqq r^{c_1}$ for any fixed c_1. If c_1 is chosen fairly large, Theorem I of Chapter IX provides a sufficiently good estimate when α is not of the above form. For a treatment of Goldbach's Problem on these lines, see T. Estermann, *Introduction to modern prime number theory* (Cambridge, 1952). Estermann's Theorem **56** is a special case of Theorem I of Chapter IX.

A detailed account of recent researches on Waring's Problem for primes, using Vinogradov's method, is given in L. K. Hua's monograph *Additive theory of prime numbers*, Travaux de l'Institut mathématique Stekloff, XXII (Moscow and Leningrad, 1947; Russian with English summary).

CHAPTER XI

The Distribution of the Fractional Parts of the Values of the Function αp

In this chapter, the results of Chapter IX are applied to the solution of the problem of the distribution of the fractional parts of the values of the function αp when p runs through primes not exceeding N.

The method used here makes it possible to solve the general problem of the distribution of the fractional parts of the values of a function $f(p)$ when $f(p)$ is a polynomial of degree higher than the first, and also when $f(p)$ is a function which approximates well in a certain sense to a polynomial. But we shall not consider these questions here.

Notation in this chapter. The letter p will always denote primes. N will be an arbitrarily large positive integer, and $\log N = r$. We denote by $\pi(N)$ the number of primes not exceeding N.

THEOREM. *Let τ satisfy*

$$N^{\frac{1}{2}} \leqq \tau \leqq N \exp(-r^{\varepsilon_0}).$$

Let α be real, and suppose that

$$\alpha = \frac{a}{q} + \frac{\theta}{q\tau}, \text{ where } (a, q) = 1 \text{ and } \exp(r^{\varepsilon_0}) \leqq q \leqq \tau.$$

Let $H(N)$ denote the number of primes p satisfying

$$p \leqq N, \qquad \{\alpha p\} \leqq \beta,$$

where $0 < \beta < 1$. Then

$$H(N) = \beta\pi(N) + O(N\gamma),$$

where

$$\gamma = (q^{-1} + qN^{-1})^{\frac{1}{2}-\varepsilon} + N^{-\frac{1}{5}+\varepsilon}.$$

Proof. Let

$$S_m = \sum_{p \leqq N} e(\alpha m p),$$

where m is a positive integer. The hypotheses on τ, α, q are the same as in Theorem 3 of Chapter IX.

Let ε_1 be any fixed positive number, and let

$$\Delta_1 = (q^{-1} + qN^{-1})^{\frac{1}{2}-\varepsilon_1} + N^{-\frac{1}{5}+\varepsilon_1},$$
$$K_1 = [\Delta_1^{-2}].$$

Then taking ε_1 in place of ε in Theorem 3 of Chapter IX, and noting that $\Delta_1 > \Delta$, it follows that if

$$U_K = \sum_{m=1}^{K} | S_m |,$$

then

$$U_K \ll K N \Delta_1$$

for any positive integer $K \leqq K_1$.

The proof of the present theorem, starting from this inequality, is essentially the same as the proof of the theorem in Chapter VIII. We can suppose that $\Delta_1 < \frac{1}{4}$, since N is large. For any real A, B with $0 \leqq B - A \leqq 1 - 2\Delta_1$, we apply Lemma 12 of Chapter I with

$$r = 1,\ \alpha = A - \tfrac{1}{2}\Delta_1,\ \beta = B + \tfrac{1}{2}\Delta_1.$$

We obtain a function $\psi(z)$ with period 1, whose value always lies between 0 and 1 and is

$$\begin{aligned} &1 \text{ if } A \leqq z \leqq B \pmod 1, \\ &0 \text{ if } B + \Delta_1 \leqq z \leqq 1 + A - \Delta_1 \pmod 1). \end{aligned}$$

This function has the Fourier series expansion

$$\psi(z) = (B - A + \Delta_1) + \sum_{m=1}^{\infty} (a_m \cos 2\pi m z + b_m \sin 2\pi m z),$$

where $a_m \ll h_m$, $b_m \ll h_m$, if h_m is defined by $h_m = m^{-1}$ if $m \leqq \Delta_1^{-1}$, $h_m = \Delta_1^{-1} m^{-2}$ if $m > \Delta_1^{-1}$. It follows that

$$\sum_{p \leq N} \psi(\alpha p) = (B - A + \Delta_1)\pi(N) + \sum_{m=1}^{\infty} (a_m S_m' + b_m S_m''),$$

where S_m' and S_m'' are defined by $S_m = S_m' + iS_m''$.

We have

$$\sum_{m=1}^{\infty} (a_m S_m' + b_m S_m'') \ll \sum_{m=1}^{\infty} h_m \mid S_m \mid$$

$$= \sum_{m \leq K_1} h_m \mid S_m \mid + \sum_{m > K_1} h_m \mid S_m \mid.$$

We apply partial summation to the first sum. Noting that $0 \leqq h_m - h_{m+1} \ll m^{-2}$ for all m, we obtain

$$\begin{aligned} \sum_{m \leq K_1} h_m \mid S_m \mid &= \sum_{m \leq K_1} h_m(U_m - U_{m-1}) \\ &= \sum_{m \leq K_1 - 1} (h_m - h_{m+1})U_m + h_{K_1} U_{K_1} \\ &\ll \sum_{m \leq K_1 - 1} m^{-2} m N \Delta_1 + (\Delta_1^{-1} K_1^{-2})(K_1 N \Delta_1) \\ &\ll N\Delta_1 \log K_1 + N K_1^{-1} \\ &\ll N\Delta_1 \log (\Delta_1^{-1}) + N\Delta_1^2 \\ &\ll N\gamma, \end{aligned}$$

if we take ε in the definition of γ to be $2\varepsilon_1$. Since $\mid S_m \mid \leqq N$ trivially, we have also

$$\sum_{m > K_1} h_m \mid S_m \mid \ll \sum_{m > K_1} \Delta_1^{-1} m^{-2} N \ll N\Delta_1 \ll N\gamma.$$

Hence

$$\sum_{p \leq N} \psi(\alpha p) = (B - A)\pi(N) + O(N\gamma).$$

The rest of the proof is completed in the same way as in Chapter VIII, and the conclusion follows.

NOTES ON CHAPTER XI

If α is an irrational number whose partial quotients are bounded, we can choose q to lie between two constant multiples of $\sqrt{N}$, and the conclusion of the Theorem is that

$$H(N) = \beta\pi(N) + O(N^{-\frac{1}{8}+\varepsilon}).$$

The error term is remarkably good, and is of course much better than any error term that is known for the asymptotic expression of $\pi(N)$ itself as a function of N.

The Theorem has one imperfection; as it stands it does not establish that *if α is any fixed irrational number then the fractional parts of αp are uniformly distributed in the interval* (0, 1). For if N is an arbitrarily large integer, and we approximate to α by a/q in the usual way, with $\tau = N \exp(-r^{\varepsilon_0})$, then q may not satisfy the hypothesis $q \geq \exp(r^{\varepsilon_0})$. Indeed, this will certainly happen if α is an irrational number whose partial quotients increase with great rapidity.

The result stated above is in fact true, and it is easy to complete the proof by dealing separately with the case when $q < \exp(r^{\varepsilon_0})$. Theorem 2b of Chapter IX covers this case if $q \geq r^{11\varepsilon}$, and if $q < r^{11\varepsilon}$ the result can easily be deduced from what is known concerning the distribution of primes in arithmetical progressions.

The result can be expressed in another form, as was suggested by Professor Heilbronn (in conversation): *if $\alpha_0 > 1$ and α_0 is not an integer, the numbers $[n\alpha_0]$, for $n = 1, 2, \ldots$, include infinitely many primes.* For the integers m given by $m = [n\alpha_0]$ are precisely those integers m for which

$$1 - \frac{1}{\alpha_0} < \left\{\frac{1}{\alpha_0} m\right\} \leq 1,$$

and if α_0 is irrational the above result with $\alpha = 1/\alpha_0$ implies that there are infinitely many prime values of m with this property. If α_0 is rational, but not an integer, the numbers of the form $[n\alpha_0]$ include at least one arithmetical progression $ax + b$ with $(a, b) = 1$, and the result follows from Dirichlet's theorem.