Lectures on
Sieve Methods

By

H.E. Richert

Tata Institute of Fundamental Research
Bombay
1976

Lectures on
Sieve Methods

By

H.E. Richert

Notes By

S. Srinivasan

Tata Institute of Fundamental Research
Bombay
1976

Printed In India
By
Anil D. Ved At Prabhat Printers. Bombay 400 004
And Published By
The Tata Institute of Fundamental Research

Preface

THESE LECTURES were given during a seven-week course at the Tata Institute of Fundamental Research. The aim was to provide an introduction to modern sieve methods, i.e. to various forms of both the large sieve (part I) and the small sieve (part II), as well as their interconnections and applications. Being a were of the fact that such a goal cannot be reached in such a short time. I have tried to compromise between an introduction and a survey. The difficult task of deciding what to omit I have tried to overcome in most cases by presenting the simplest approach in details and a sketch of the more sophisticated results if their proof would have required too much time. Nevertheless I have decided to include a chapter an the history of the large sieve upto Bombieri's first paper, because I believe that a student coming to a new branch of mathematics can learn much more from the historical development in that then is generally expected. The final chapter contains a proof of Chen's Theorem, because I consider it the most beautiful example of the interaction between various sieve methods and other powerful tools of analytic number theory.

I am indebted to my colleagues at the Tata Institute for their generous hospitality, particularly to K.G. Ramanathan and to K. Ramachandra for many interesting discussions.

The notes have been prepared by S.Srinivasan. His critical ability has been of great value to me, and I wish to thank him his meticulous handling of the manuscript.

H.-E. Richert

Bombay, April 1976

Introduction

SIEVE METHODS, beyond that of Eratosthenes and of Legendre, can be considered to have started with the works Brun (small sieve) and of Linnik (large sieve). In first part of these lectures we confine ourselves to an introduction to the large sieve and a survey of its applications. Under Chapter 0 we give a historical introduction to the theory of the large sieve pertaining to the works, upto the first paper of Bombieri (1965), covering a period of twenty-five years.

Regarding the relative powers of elementary sieve methods and the analytical methods one usually considers that the latter should be more powerful. Further it has generally held that large sieve is than the small sieve (also that Selberg's sieve always supersedes Brun's sieve). But history has shown that such views are not totally correct.

As for the first point we elucidate the connection between the elementary large sieve method and the analytical methods in number theory by recalling briefly some of those basic methods used in number theory.

1) In multiplicative number theory one has the important problem of finding asymptotic formulae for the sums $\sum_{n\leq x} a_n$ as $x \to \infty$, where $a_n s$ are values of some number-theoretic functions. Introducing the function $F(s) = \sum_{n=1}^{\infty} a_n n^{-s}. s \in \mathbb{C}$, the connection with analytical methods is brought about through the following formula, due to Dirichlet:

$$\sum_{n\in x} a_n = \frac{1}{2\pi i} \int\limits_{Re\ s=c} F(s)\frac{x^s}{s} ds (x > 1,\ \text{non-integral }).$$

where c is a real number exceeding the abscissa of convergence (finite in practice) of $F(s)$. (Hence and in what follows integration along straight lines are always in the direction of non-decreasing real and imaginary parts.) To clarify the use of this formula let us take $F(s) = \zeta^2(s)$, which corresponds to the problems of finding the asymptotic formula (with error term) for $\sum_{n \leq x} d(n)$. In this case one can take for c above any value greater than 1. One can show now that the major contribution to the above integral comes from the part $| Im\ s| \leq T$, provided T is suitably large in relation to x (we also choose c sufficiently close to 1). Next shifting the line of integration to the left one has that the remaining (major) part is, by Cauchy's theorem.

$$= 2\pi iR + (\int\limits_{\substack{Re\ s=s \\ | Im\ s| \leq T}} + \int\limits_{\substack{s \leq Re\ s \geq c \\ Im\ s = T}} - \int\limits_{\substack{s \leq Re\ s \geq c \\ Im\ s = T}})\zeta^2(s)\frac{x^s}{s}ds,$$

where $(0 \neq)\zeta < 1$ and R denotes the sum of residues of the integer and at its poles within the rectangle bounded by the lines $| Im\ s| = T, Re\ s = \zeta$ and $Re\ s = c$. It turns out that the estimate for the first integral above dominates those of the other two and is itself 'negligible' provided T is not too large in relation to x; in otherwords, if T is appropriately chosen, then R is the main term of the asymptotic formula for $\sum_{n \leq x} d(n), x \to \infty$. Thus, in this case, we are led to the following

$$\sum_{n \in x} d(n) = x \log x + c_0 x + 0(x^\theta), x \to \infty,$$

with some constants c_o and $\theta, 0 < \theta < 1$. (Clearly, the restriction that x is not an integer can easily be dropped here.)

2) As to additive number theory we again consider a simple case only. Let γ be an (infinite) set of non-negative integers and let a_n be the characteristic function of γ; i.e.,

$$a_n = \begin{cases} 1 \text{ if } n \in \gamma, \\ 0 \text{ if } n \notin \gamma. \end{cases}$$

Introducing the function

$$f(z) = \sum_{n=0}^{\infty} a_n z^n$$

(so that the power-series has radius of convergence = 1) the formula

$$\sum_{\substack{n_1\in\gamma, n_2\in\gamma\\ n_1+n_2=N}} 1 = \frac{1}{2\pi i} \int\limits_{|z|=r<1} f^2(z) z^{-N-1} dz, N \in N$$

provides the analytical connection with the problem of finding (the asymptotic formula for) the number of representations of Nas a sum of two numbers of γ. For instance, when γ is the set of primes this corresponds to the well-known Goldbach problem. Then it turns out that the major contribution to the above (integral) comes out of the points z with arguments 'close' to fractions with 'small' denominators, while r approaches 1 as $N \to \infty$, and the set of such points z constitute the 'major arcs' and the remaining parts are termed 'minor arcs'. Thus again we see that to get information about our problem one has to move close to the singularities (on the unit circle) of our function.

The functions introduced in 1) and 2) above are particular instances of general Dirichlet series

$$G(s) = \sum_{n=0}^{\infty} a_n e^{-\lambda_n s} \quad a_n \in \mathbb{C}, s \in \mathbb{C},$$

where $\lambda_1 < \lambda_2 < \lambda_3 < \cdots \to \infty$. Still the above problems have this essential difference: Under 1) we encountered isolated singularities and on the other hand, regarding 2) one knows from gap theorems that, for examples, for the function $\sum_p z^p$ the unit circle is the natural boundary. However, both the cases illustrate the principle of the analytical methods in that the singularities of the associated function are the sources of arithmetical information regarding the concerned problem, in the sense that heavier the singularity more is its contribution to the main term.

The idea of the second of the methods sketched above, the 'Hardy - Littlewood method', goes back to Hardy and Ramanujan. It has been later developed in a series of papers by Hardy and Littlewood. (A variant of this method, introduce by I.M Vinogradov, which uses finite sums instead of series, allows integration over the unit circle.) I this method, the aforementioned principle is reflected in that the contribution, of the major arcs (i.e., neighbourhoods of heavier singularities), the so-called 'singular series' of the problem, determines the main term.

We are now in a position to indicate as to how the 'elementary' large sieve method can be regarded as being analogous to the corresponding analytical approach In its basic form the large sieve relates the mean-square contribution from the mid-points of the major arcs (the sources of arithmetical information) of the size of the associated function with the mean-square integral, similar to the singular series (above) being related to the integral over the unit circle. Thus the method links up an arithmetical information with the gross mean-square (= number of elements in γ, if a_n is the characteristic function of γ). This is to suggest that this elementary' sieve method can be considered analogous to analytical methods.

Notation

In general all the notation employed in these lecture are either standard or are given explicitly at the place of their first occurrence. So we limit our seives here to a description of the former type followed by the ones of the other kind (along with the place of their first occurrence in parenthesis, for the convenience of reference). A reference $(A.B)$ to a part in these lectures stands for "'formula $B'$of 'chapter A'".

The letter p (with or without affixes) denotes invariably a prime number. An 'almost prime' P_r (cf. (12.8)), for a given integer $r \leq 1$, is a natural number with not more than r prime factors (counted with multiplicity). The greatest common divisor of two integers m and n is denoted by (m, n). For an integer n, the divisor function $d(n)$ $(cf.p.ii)$ denotes the number of (positive) divisors of n. We also use von Mangoldt's function $\wedge(n)$ $(cf.(1.75))$ defined as log p or 0 according as n is a power of (some prime) p or not. As usual, for a real $x, [x]$ denotes the greatest integer not exceeding x. Euler's constant γ (cf. (8.33)) is $(\lim_{x\to\infty}(\sum\limits_{1\leq n\leq x} n^{-1} - \log x))$.

The order notation $O, o \ll, \gg$ have their customary meaning and the dependence of the implied constant on some auxiliary parameter (s) is (when essential) given explicitly. (The notation $\asymp$, meaning $\gg\ll$, occurs only once in connection with ((0.51).)

The symbols $\mathbb{N}, \mathbb{Z}, \mathbb{R}, \mathbb{C}$ denote respectively the sets of natural numbers, integers, real numbers, complex numbers endowed with their natural (basic) structures. Regarding intervals (of reals or integers) the convention of using the brackets), (to indicate the excluded end-point (s) and], [to indicated the included end point (s) is adopted. Also we

$|A|$ to denote the cardinality of a finite set (or sequence)$\mathcal{A}$).

Finally, the notation explicitly introduced in the lectures: We have $\|x\|$ $(cf.(2.10)), e(u)(cf.(0.52))$. The number-theoretic functions $\nu(n)$ $(cf.(1.9))$. $\mu(n)(cf.(1.8)).q(n)(cf.(1.19)), c_q(n)(cf.(1.29)), \varphi(n)$ $(cf.(1.33))$ and $r(n)(cf.(6.71))$ occur more than once. Also for the summation conventions $(\sum_{\ell=1}^{q}, \sum_{\ell \bmod q})$ and $\sum^{*}_{\chi \bmod q}$ see respectively (1.29) and (3.6). The following conditions are of repeated use in the second part of these lectures:

$$(\Omega_1)(cf.(9.16_a) \text{ or } (9.16_b)); (\Omega_2(k.L))(cf.(11.3)); (\Omega_2(k))(cf.(11.4));$$
$$(R)(cf.(11.9)); (\Omega_\circ)(cf.(11.11)); (Q)(cf.(11.23)); (R(1,\alpha))(cf.(11.62)).$$

Contents

SIEVE METHODS

Chapter 0

History of the Large Sieve

LET γ BE a set of $|\gamma|$ integers contained in an interval of length N: **1**

$$\gamma \subset (M, M+N), M \in \mathbb{Z}, N \in \mathbb{N}, S := |\gamma|. \tag{0.1}$$

Setting

$$S(q, \ell) := \sum_{\substack{n \in \gamma \\ n \equiv \ell \bmod q}} 1, \tag{0.2}$$

we have

$$\sum_{\ell=1}^{q} S(q, \ell) = S, \tag{0.3}$$

or that

$$\sum_{\ell=1}^{q} (S(q, \ell) - \frac{S}{q}) = 0. \tag{0.4}$$

Confining, for the moment, our attention to primes $q = p$ only, (0.4) tells us that the quantity

$$D(p) := \sum_{\ell=1}^{p} (S(p, \ell) - \frac{S}{p})^2 \tag{0.5}$$

measures how uniformly the set γ is distributed among the residue classes $\bmod p$. Such an information. is of great importance in various problems.

A uniform and non-trivial bound of the form

$$(D) \qquad \sum_{p \le Q} pD(p) \le K(N, Q, S) \tag{0.6}$$

by uniform we mean here that while K may depend on N, Q and S, it should be independent of the particular structure of the set γ-makes it possible to draw the following general conclusions. (Note that actually
2 the supposition $M = 0$ here involves no loss of generality.)

(A) Let N, Q and S be given. If every set γ (cf. (0.1)) is so uniformly distributed over the residue classes $\bmod p$ as expressed by (D), then for most of the p's $D(p)$ must be small. This remark tells us since $D(p)$ is bound to be large if many residue classes $\bmod p$ do not contain any element of γ, the statement that $S(p, \ell) = 0$ for 'many' ℓ's $\bmod p$ can be true for only 'few' ('exceptional') primes $p \le Q$.

We can express this remark in a quantitative form. Let $\omega(p)$ be an (integer-valued) function satisfying

$$0 < \omega(p) < p, \tag{0.7}$$

and now we ask for the number of $p's \le Q$, for which *at least* $\omega(p)$ residue classes $\bmod p$ do not contain any element of our set γ. Let us denote the set of these 'exceptional' primes by $\mathfrak{p}$, and set

$$\min_{p \in \mathfrak{p}} \frac{\omega(p)}{p} = \delta = \delta(\mathfrak{p}). \tag{0.8}$$

One has then for each $p \in \mathfrak{p}$

$$pD(p) \ge p\omega(p)\frac{S^2}{p^2} \ge \delta S^2, \tag{0.9}$$

and so

$$\sum_{p \in \mathfrak{p}} pD(p) \ge \delta S^2 |\mathfrak{p}|. \tag{0.10}$$

Trivially
$$\sum_{p\in\mathfrak{p}} pD(p) \leq \sum_{p\leq Q} pD(p), \tag{0.11}$$
and therefore (0.6) and (0.10) give
$$(A) \qquad |\mathfrak{p}| \leq \frac{K(N,Q,S)}{\delta S^2} \tag{0.12}$$

For the remaining primes p i.e., $p \notin \mathfrak{p}$ and $p \leq Q$, less than $\omega(p)$ residue classes $\bmod p$ are devoid of numbers of γ. Consequently, for these primes p each of at least $p-\omega(p)$ residue classes $\bmod p$ contains atleast one element of γ **3**

(B) The preceding result may also be considered as a sieve problem. In order to see this. let us start (cf. (0.1)) with the set of numbers
$$M+1,\ldots,M+N, \quad M\in\mathbb{Z}, N\in\mathbb{N}. \tag{0.13}$$

Now for certain primes $p \leq Q$, $p \in \mathfrak{p}$ say, strike out of numbers (0.13) all those numbers which are situated in any of certain $\omega(p)$, where $\omega(p)$ satisfies (0.7), of the p residue classes $\bmod p$. Let the remaining set of numbers be our set γ. We obtain (0.12) again under the present situation, since we have used for its proof only that at least $\omega(p)$ residue classes $\bmod\ p$ (for each $p \in \mathfrak{p}$) contain no element of γ. Next resolving (0.12) with respect to S - as shall be seen later to be possible we get an upper bound for $|\gamma|$ with respect to our set γ above. It is the type
$$(B) \qquad (|\gamma| =)S \leq K_1(N,Q,|\mathfrak{p}|,\delta). \tag{0.14}$$

(C) Our remark at the beginning of A had been that a non-trivial estimate of the type (D) implies $D(p)$ is small for most of the $p's$ under consideration. If we view this as the statement that 'most often' $\frac{S}{p}$ is a good approximation to $S(p,\ell)$ then its quantitative version leads to a more precise formulation than that under A and consequently than that under B.. In fact, we obtain a result of the type of Čebysěv's inequality.

To this end we introduce a function $c(p)$ satisfying

$$c(p) \geq 1 \tag{0.15}$$

and put

$$\max_{p \leq Q} c(p) = c = c_Q. \tag{0.16}$$

4 Now we ask for the existence of the inequality

$$|S(p, \ell) - \frac{S}{P}| < \frac{S}{pc(p)} \tag{0.17}$$

More precisely, we ask for the number of primes $p \leq Q$ for which the inequality (0.17) does not hold for at least $\omega(p)$ (cf. (0.7)) residue classes $\ell \bmod p$. Let us denote again by $\mathfrak{p}$ the set of such exceptional primes. Then, in the notation of (0.8),

$$pD(p) \geq \omega(p) \frac{S^2}{p^2 c^2(p)} \geq \frac{\delta}{c^2} S^2, \quad \forall p \in \mathfrak{p}. \tag{0.18}$$

Consequently, in view of (0.11). (0.6) gives for the number of exceptional primes p the upper bound

$$(C) \qquad |\mathfrak{p}| \leq \frac{c^2}{\delta S^2} K(N, Q, S). \tag{0.19}$$

This result may be phrased as follows: (0.6) implies that for every set γ (of (0.1)) one has for all $p \leq Q$. save for atmost

$$\frac{c^2}{\delta S^2} K(N, Q, S)$$

primes, and for all residue classes $\ell \bmod p$ with the exception of less then $\omega(p)$ of them

$$S(p, \ell) = \frac{S}{P} + \theta \frac{S}{pc(p)}, \quad \text{where} \quad |\theta| < 1. \tag{0.20}$$

Choosing $c(p) = 1$, we see that each exceptional prime of A, is also an exceptional prime of C. and conversely. Therefore C. is a generalization of A. and hence also of B.

Although the uniformity of (0.6) with respect to γ may be consid-
5 ered a defect (because it includes all 'bad' sets) it has the advantage of drawing all the conclusions (of which we have given three general examples in **A**, **B**. and **C**.) valid for all sets , including those which are otherwise not readily accessible. The quality of the $K's$ known for (0.6) allows us to obtain results which are not available by the use of other methods.

(0.6) is of interest for $Q \leq N$ only. *A* trivial estimate is obtained in the following way. By (0.5) and (0.3) we have

$$D(p) = \sum_{\ell=1}^{p} S^2(p, \ell) - \frac{S^2}{p}, \tag{0.21}$$

and trivially

$$S(p, \ell) \leq \sum_{\substack{M<n\leq M+N \\ n\equiv\ell \bmod p}} 1 \leq \frac{N}{p} + 1 \leq 2\frac{N}{p} \quad (p \leq N). \tag{0.22}$$

Using this and (0.3) again in (0.21) we find that

$$pD(p) \leq 2N \sum_{\ell=1}^{p} S(p, \ell) = 2NS, \tag{0.23}$$

and so

$$(D_0) \qquad \sum_{p\leq Q} pD(p) \leq 2NQS \quad (Q \leq N). \tag{0.24}$$

As we have seen above (A),(B) and (C) are more or less different versions of general types of results based on (D). Linnik [1] was the first to consider such problems. With respect to *B*. he made the remark that when striking out an (absolutely) *bounded* number of residue classes mod p the sieve method of Brun (or of Selberg) is applicable. However, this is no longer true if, as is permissible under **B.**, $\omega(p)$ is, for example, an increasing function of p. For this very reason, Linnik named his method of treating A. and B.

"The Large Sieve".

Linnik [1] proved that 6

$$(A_1) \qquad |\mathfrak{p}| \leq 20\pi \frac{N}{\delta^2 S} \quad \text{for} \quad Q = \sqrt{N}, \tag{0.25}$$

and consequently

$$(B_1) \qquad S \leq 20\pi \frac{N}{\delta^2 |\mathfrak{p}|}, \quad \text{for} \quad Q = \sqrt{N} \tag{0.26}$$

As an example he takes γ to be the set of primes $\leq N$, i.e., $M = 0$ and $S = \pi(N)$, and $\omega(p) = p^{3/4}$. Then $\delta \geq N^{-1/8}$ and (A_1) yields the non-trivial estimate:

$$|\mathfrak{p}| \leq 20\pi \frac{N}{N^{-1/4}\pi(N)} \leq 80 N^{1/4} \log N, \quad \text{for} \quad N \geq N_0 \tag{0.27}$$

Although the large sieve first occurred in the forms (A_1) and (B_1), I would prefer, in particular, in view of later developments, to call (D) *the* large sieve and rather consider results of type (A), (B),(C) and others as applications of the theory of the large sieve.

Following Linnik it was Rényi's ([1], [2]) merit to generalize the large sieve method in several respects. Simultaneously he noticed the fundamental importance of (D), and also treated the more precise version (C) for the first time.

Generalizing (0.2), for arbitrary complex numbers a_n we set

$$\tilde{S}(q, \ell) := \sum_{\substack{M < n \leq M+N \\ n \equiv \ell \bmod q}} a_n \tag{0.28}$$

and

$$\tilde{S} := \tilde{S}(1, 1) = \sum_{M < n \leq M+N} a_n. \tag{0.29}$$

(By taking for a_n the characteristic function of γ in (0.28) we get (0.2).)

7 Let $\mathfrak{p}$ be an arbitrary set of primes $p \leq Q$. Rényi's paper [2] has implicitly the following (explicit) generalization of (D) (with $Q < \sqrt{N}$):

$$(\tilde{D_1}) \qquad \begin{cases} \sum\limits_{p \in f} pq \sum\limits_{\ell=1}^{pq} |\tilde{S}(pq, \ell) - \frac{\tilde{S}(q,\ell)}{P}|^2 \leq \frac{1}{2\varepsilon} \sum\limits_{M<n\leq M+N} a_n^2 \\ + \frac{4\pi^2 \epsilon^2 N^4 m^2}{3} |\mathfrak{p}| \text{ for } Q < \sqrt{\frac{N}{q}}, \end{cases} \tag{0.30}$$

where $a_n's$ are ≥ 0, $m = \max\limits_{M<n\leq M+N} a_n$, q is a squarefree number not divisible by any $p \in \mathfrak{p}$, and $0 < \epsilon < \dfrac{1}{2N}$.

The use of a set $\mathfrak{p}$ of primes here is particularly suitable for applications of the type (A), (B) and (C), because $\mathfrak{p}$ can serve as the of exceptional primes, and the factor $|\mathfrak{p}|$ instead of Q improves the estimate. From this generalization to composite moduli and arbitrary coefficients he derived ([1], (Lemma 1)) a (C)-type result about

$$|\tilde{S}(pq,\ell) - \frac{\tilde{S}(q,\ell)}{p}| < \frac{\tilde{S}}{pqc(pq)}. \qquad (0.31)$$

From here he succeeded ([1], (Lemma 2)) in making the large sieve applicable also in the estimation of certain averages of character sums, i.e., sums of the form

$$\sum_{M<n\leq M+N} \chi(n)a_n. \qquad (0.32)$$

Turning back to the case $q = 1$, $a_n = \begin{cases} 1 \text{ for } n \in \gamma, \\ 0 \text{ for } n \notin \gamma, \end{cases}$ i,e. to (0.17), then Rényi's C-result corresponding to (0.19) is

$$(C_1) \qquad |\mathfrak{p}| \leq \frac{3\pi N^2 c^3}{2\delta^{3/2} S^2} \quad \text{for} \quad Q < \sqrt{N}. \qquad (0.33)$$

By using a different method of proof, Rényi ([7], (18)) next proved

$$(D_1) \qquad \sum_{p<(N/12)^{1/3}} pD(p) \leq 2NS \qquad (0.34)$$

and then applied this to C. and A. obtaining ([7], (Theorem 3)) **8**

$$(C_2) \qquad |\mathfrak{p}| \leq \frac{2Nc^2}{\delta S} \quad \text{for} \quad Q < (\frac{N}{12})^{1/3} \qquad (0.35)$$

and ([7], (Corollary 1))

$$(A_2) \qquad |\mathfrak{p}| \leq \frac{2N}{\delta S} \quad \text{for} \quad Q < (\frac{N}{12})^{1/3}. \qquad (0.36)$$

Finally Rényi improved his method and considered the large sieve in its (D)-version as a special statistical statement ([4], [8], [10], [12]) and he showed ([4])

$$(D_2) \qquad \sum_{p \le \frac{1}{2}N^{1/3}} pD(p) \le 9NS \tag{0.37}$$

and from this ([4], (Theorem 3))

$$(C_3) \qquad |\mathfrak{p}| \le \frac{9Nc^2}{\delta S} \quad \text{for} \quad Q < \frac{1}{2}N^{1/3} \tag{0.38}$$

The last results above, though stronger than Linnik's, are still valid only for a smaller range for the primes, i,e., for smaller values of Q. In [9] (cf. Halberstam and Roth [1] (Ch. IV, Theorem 6′)) he prepared the ground for the extension

$$(D_3) \qquad \sum_{p \le Q} pD(p) < (N + Q^3)S \quad \text{for} \quad Q < \sqrt{N}. \tag{0.39}$$

Barban has been the first to prove a (D)-result by using Linnik's original method. He showed ([8], (Theorem 1))

$$(D_4) \qquad \sum_{p \le Q} pD(p) \le 2\pi \frac{NQ}{K_0} S + \frac{K_o^2}{Q} S^2 \ \text{ for } \ 1 < K_0 < \min(Q, \frac{N}{Q}), \tag{0.40}$$

and for $(\tilde{D}_1)$ in the case $q = 1$ ([10], (1.3); for $q > 1$ cf. [10], (Theorem 3.1))

$$(\widetilde{D}_2) \qquad \begin{cases} \sum_{p \in \mathfrak{p}} p \sum_{\ell=1}^{p} |\tilde{S}(p, \ell) - \frac{\tilde{S}}{p}|^2 \le \frac{1}{\varepsilon} \sum\limits_{M < n \le M+N} a_n^2 + \\ + \frac{4\pi^2 \epsilon^2 N^2 \tilde{S}^2}{3} |\mathfrak{p}| \quad \text{for} \quad Q < N, \end{cases} \tag{0.41}$$

9 subject to the conditions $0 < \epsilon < \min(\frac{1}{2\pi N}, \frac{1}{2Q^2}), a_N \ge 0$. The particular case $a_n = \begin{cases} 1 \text{ for } n \in \gamma, \\ 0 \text{ for } n \notin \gamma, \end{cases}$ (i.e., $\tilde{S} = S$) of this result contains (D_4) and from it Barban ([10]. (Theorem 1.1)) derived for A.

$$(A_3) \qquad |\mathfrak{p}| \le 20 \frac{N}{\delta^{3/2} S} \quad \text{for} \quad Q = \sqrt{N}. \tag{0.42}$$

The importance of (D_4) lies in the fact that there is no longer any restriction on Q, apart from the natural one, namely $Q < N$, implied by the K_0-condition.

There is another generalization to a weighted form of (D) due to Halberstam and Roth [1] (Ch. IV, Theorem 5). They proved that for arbitrary weights δ_p, satisfying

$$0 < \delta_p < \frac{1}{2p^Q} \tag{0.43}$$

one has

$$(D*) \qquad \sum_{p \le Q} p\delta_p D(p) \ll S + N^4 \sum_{p \le Q} \delta_p^3. \tag{0.44}$$

A very important progress was made by Roth [2], who succeeded in proving

$$(D_5) \qquad \sum_{p \in \mathfrak{p}} pD(p) \ll (N + Q^2 \log K_0)S + S^2 |\mathfrak{p}| K_0^{-2} \quad \text{for} \quad K_0 \ge 2. \tag{0.45}$$

where again $\mathfrak{p}$ denotes an arbitrary set of primes $p \le Q$, and there is no restriction on Q. This result includes in particular ([2], (9))

$$(D_6) \qquad \sum_{p \le \sqrt{\frac{N}{\log N}}} pD(p) < NS, \tag{0.46}$$

as well as ([2], (Corollary 1)) **10**

$$(D_7) \qquad \sum_{p \le Q} pD(p) \ll S\,Q^2 \log Q \quad \text{for} \quad Q \ge \sqrt{\frac{N}{\log N}}. \tag{0.47}$$

With respect to (C) Roth ([2], (Corollary 2)) derived, also noticing that it is more appropriate - as can be seen from (0.18) to introduce

$$\beta := \frac{1}{4} \min_{p \le Q} \frac{\omega(p)}{pc^2(p)} \tag{0.48}$$

instead of δ and c separately, from (D_5)

$$(C_4) \qquad |\mathfrak{p}| \ll \frac{N + Q^2 \log \frac{1}{\beta}}{\beta S} \quad \text{for} \quad Q < N. \tag{0.49}$$

It is not too simple a matter to make a comparison of the various basic (D)-results of the large sieve. However, the main features are the following ones.The very effective estimate (for $K(N, Q, S)$ of (0.6))

$$\ll NS \tag{0.50}$$

which was known by Renyi for Q upto $N^{1/3}$ (cf. (0.34), (0.37), (0.39)) has been extended by Roth (cf. (0.46)) upto $Q = \sqrt{\frac{N}{logN}}$. For values of Q beyond $\sqrt{N}$, (0.40) and (0.47) still yield non-trivial estimates (compare (0.24)) upto the vicinity of N. Apart from the factor $\log Q$ and the $\ll$-constant, (0.47) is in most cases the better estimate. Only for $Q \asymp N^2 S)^{1/5}$ does (0.40) yield

$$\ll Q^2 S \tag{0.51}$$

which is (0.47) without the factor $\log Q$. On the other hand, the same result is implied by (0.45) if moreover $S \ll Q \log Q$.

11 As far as the methods of proof for the large sieve are concerned there are different ways of approach.

Recalling our notation introduced in the beginning of this chapter (i.e., considering, for simplicity, the case $a_n = \begin{cases} 1 \text{ for } n \in \gamma, \\ 0 \text{ for } n \notin \gamma, \end{cases}$ only) the first method, used in the aforementioned basic papers of Linnik [1] and of Rényi [2]. is based on a treatment of the exponential sum

$$T(x) : \sum_{n \in \gamma} e(nx) \cdot x \in \mathbb{R}, \quad e(u) := e^{2\pi i u} \tag{0.52}$$

Here the essential use is made of the Farey dissection from the method of Hardy and Littlewood and of Parseval's formula. In fact, the close connection with $T(x)$ stems from the identity (cf. (7.1))

$$\sum_{\ell=1}^{p-1} |T(\frac{\ell}{p})|^2 = p \sum_{\ell=1}^{p} (S(p, \ell) - \frac{S}{P})^2 = pD(p), \tag{0.53}$$

the second equality being (0.5). For the proof of the first equality we need only note that the left-hand side expression

$$= -S^2 + \sum_{\ell=1}^{p} |T(\frac{\ell}{p})|^2 = -S^2 + \sum_{\ell=1}^{p} \sum_{n_1 \in \gamma} \sum_{n_2 \in \gamma} e(\frac{\ell}{p}(n_1 - n_2))$$

$$= -S^2 + \sum_{n_1 \in \gamma} \sum_{n_2 \in \gamma} (\sum_{\ell}^{p} e(\frac{\ell}{p}(n_1 - n_2)))$$

and further, by (1.22). the inner sum in the last expression $= p$ or $= 0$ according as $(n_1 - n_2)$ is $\equiv 0$ or $\not\equiv 0 \bmod p$, so that our expression

$$= -S^2 + \sum_{\ell=1}^{p} S^2(p, \ell) = p(\sum_{\ell=1}^{p} (S(p, \ell) - \frac{S}{P})^2)$$

by (0.3). Hence the form (D), which we have considered so far to be the basis of the large sieve, amounts to asking for an upper bound for

$$\sum_{p \leq Q} \sum_{\ell=1}^{p-1} |T(\frac{\ell}{p})|^2. \tag{0.54}$$

Another method of proof that should be mentioned here takes a more general point of view and may serve to simplify the understanding of the large sieve method. It relates the problems with certain results in an inner product space. This has been already developed by Rényi in his early papers ([4], [7]), where he refers to Boas [1] and also to Bellman [1] for their extensions of classical results to 'quasi-orthogonal' functions. Roth's work [2] has used king of combination of both methods. **12**

Further methods and details of proofs with respect to the theory of the large sieve will not be given here, but rather would be mentioned in appropriate chapters, in part icular under chapter 2. However, since we will not have the opportunity to use the second method mentioned above, we shall present here a basic result due to A. Selberg (cf. Bombieri [4]). The proof is very elegant and the result seems to me to be most suitable for giving an idea of this method, which consists then in choosing appropriate functions in an application of (0.55).

Theorem 0.1. *Let $f, \varphi_1, \ldots, \varphi_R$ be elements of an inner product space over $\mathbb{C}$. Then*

$$\sum_{r=1}^{R} \frac{|(f, \varphi_r)|^2}{\sum_{s=1}^{R} |(\varphi_r, \varphi_S)|} \leq ||f||^2. \tag{0.55}$$

Proof. For any complex numbers c_r, $1 \leq r \leq R$, we have by the Bessel's inequality argument,

$$\begin{cases} ||f||^2 - 2Re \sum\limits_{r=1}^{R} c_r \overline{(f, \varphi_r)} + \sum\limits_{r,s=1}^{R} c_r \overline{c}_S (\varphi_r, \varphi_S) = \\ = ||f - \sum\limits_{r=1}^{R} c_r \varphi_r||^2 \geq 0. \end{cases} \tag{0.56}$$

Using here

$$\begin{cases} \sum\limits_{r,s=1}^{R} c_r \overline{c}_S (\varphi_r, \varphi_s) \leq \sum\limits_{r,s=1}^{R} (\frac{1}{2}|c_r|^2 + \frac{1}{2}|c_S|^2)|(\varphi_r, \varphi_S)| = \\ = \sum\limits_{r=1}^{R} |c_r|^2 \sum\limits_{s=1}^{R} |(\varphi_r, \varphi_s)| \end{cases} \tag{0.57}$$

13 and then choosing

$$c_r = \frac{(f, \varphi_r)}{\sum\limits_{s=1}^{R} |(\varphi_r, \varphi_s)|} \tag{0.58}$$

the proof is completed. □

From Theorem 0.1 or variants of it the main tools for this (second) method can be derived (cf. Montgomery [5] (pp. 4-8) and Huxley [7] (pp. 29-30)). A simple example is that as an immediate consequence of (0.55) we obtain, under the same assumptions. 'Bellman's inequality' (Bombieri [3])

$$\sum_{r=1}^{R} |(f, \varphi_r)|^2 \leq ||f||^2 \max_{1 \leq r \leq R} \sum_{s=1}^{R} |(\varphi_r, \varphi_s)|. \tag{0.59}$$

(0.55) as well as (0.59) generalize Bessel's inequality, or (for $R = 1$) Schwarz's inequality, in an inner product space to which they reduce when $\{\varphi_r\}$ happen to be orthonormal.

The importance of this method for the modern development of the large sieve has been noticed by Bombieri, Gallagher and *A*. Selberg (cf. Bombieri [3], [4]).

The further development in the theory of the large sieve has shown that with the work of Roth one had already come close to best possible results. The next decisive step in this direction was made in an important paper by Bombieri [1] (also independently in a paper by A.I. Vinogradov [1]. Apart from the important deductions he made from his result and other details (not to mention here), the main features of this progress in theoretical respect were (i) the extension of his ((D)-type) result from an estimate of (0.54) to an estimate (of the larger gum) where the summation is extended over all natural numbers $q \leq Q$ instead of over only primes $p \leq Q$ (cf. (2.3)). (ii) keeping thereby not only the quality of Roth's result but even removing a log-factor, and (iii) obtaining an explicit $\ll$-constant. which is of considerable importance in certain applications. **14**

After having recalled some arithmetical results in Chapter 1 we shall take up this modern version of the large sieve in Chapter 2. Applications to character sums are possible, as has been already mentioned in (0.32). and will be treated in Chapter 3. Important further applications of the large sieve, not mentioned so far, to Dirichlet series which were first noticed by Davenport (cf. Montgomery [2]) are dealt with in Chapter 4. This theme is continued in Chapter 5 where certain 'hybrid' forms of the large sieve for applications to Dirichlet series also occur. Chapter 6 is devoted to a survey on special applications of the large sieve to Dirichlet series and also to some problems of number theory: In Chapter 7 we shall turn to the large sieve in its arithmetical form (B). Lastly, in Chapter 8 we give an application of the large sieve to a special problem, Viz. Brun-Titchmarsh theorem. of number theory.

So far very little has been said about the applications of the large sieve. As can be seen from the description of the contents of the following chapters given above there is a great variety of applications. Many of these are of a 'statistical' nature, in the sense that they are concerning certain averages. Many results of number theory are known to be consequences of certain unproved hypotheses. However, in many cases we are able to apply the aforementioned statistical statements to obtain, strikingly, the same results as those which one gets by assuming certain still unproved hypotheses.

Apart from a first application in the construction of a non-basic essential component (Linnik [2]). Linnik [3] showed the power of his new
15 method, via (A), in a result about the least quadratic non-residue mod p. Rényi's first application of the large sieve yielded a surprising result in the direction of Goldbach's conjecture. In fact, Rényi has been the first to prove that every sufficiently large even number can be written as a sum of a prime and a number consisting of a *bounded* number of prime factors.

We shall not mention other applications in this historical introduction but rather defer them to later (appropriate) chapters. However, in keeping with the title of this chapter, following the Notes for this chapter we add a list of references to works, upto the first paper of Bombieri, in chronological order. This list also includes papers, upto this point, which deal only with the applications of the large sieve.

There are also generalizations of the large sieve in various directions. Some of these papers are given in a second list of references following the above one.

NOTES

In order not to make this purely historical introduction too long, we have selected only some significant results that seem to be illuminating for our way of presenting the subject. For further information the reader is referred to the surveys in Barban [10], Halberstam and Roth [1]. Davenport [1], Roth [4] Montgomery [5]. Huxley [7] and Bombieri [5], [6].

(0.25), (0.26): By using $|e^{ix} - 1| \leq |x|$ instead of Linnik's estimate $|e^{ix} - 1| \leq e|x|$ the constant 20 can be replaced by 4.

(0.33): Rényi's paper ([2], (Lemma 1)) contains only the special case $\omega(p) = p^{8/9}$, $c = p^{1/9}$. However, as he has pointed out ([7], (Theorem 2)) his method also applies to the general case.

16 (0.39): Note that by (D_0), (0.39) holds also for $Q \geq \sqrt{N}$.

(0.40): His method gives actually a factor $\frac{1}{3}$ for the second term, and also the condition on K_0 may be slightly relaxed.

(0.39), (0.40): In general (D_3) yields the better result for Q in the vicinity of $N^{1/3}$ whereas (D_4) becomes superior when Q tends to $\sqrt{N}$. Also (D_4) has been the first result to yield non-trivial estimates upto the vicinity of N.

(0.42): Actually, the more appropriate choice of $\epsilon = \frac{\sqrt{s}}{2\pi N}$ yields the factor 3π instead of 20, and a simple refinement of the proof gives even $\frac{3}{2}\pi$. A consequence of this remark is, when applied to C.,

$$(C_5) \qquad |\mathfrak{p}| \leq \frac{3}{2}\frac{\pi N c^3}{\delta^{3/2} S} \text{ for } Q = \sqrt{N}, \tag{0.60}$$

a result which is always better than Rényi'es $((C_i))$.

(0.44): Following Barban's method of proof for (D_4) one notices that in (D^*) the factor N^4 can be replaced by $N^2 S^2$.

(0.56): Here we have used that

$$\begin{cases} \|f\|^2 = (f, f), (f, g+h) = (f, g) + (f, h), \\ (cf, g) = c(f, g), (g, f) = \overline{(f, g)} \end{cases} \tag{0.61}$$

and (hence also)

$$(f, cg) = \bar{c}(f, g). \tag{0.62}$$

For an application of Theorem 0.1 see the notes following Chapter 2.

History of the large sieve. References in chronological order.

Linnik [1], Boas [1], Linnik [2], [3], Bellman [1], Rényi [1], [2], [3], [4], [5], [6], [7], [8], Bateman, Chowla and Erdös [1], Kubiliyus [1], Rényi [9], Wang [1], Rényi [10], [11], Stepanov [1], Hua [1], Rényi **17**
[12], Barban [1], Linnik [4], Barban [2], Erdös [1], Barban [3], Gelfond and Linnik [1], Pan [1], Erdös [2], Pan [2], Wang [2], Levin [1], Barban [4], [5], [6], Pan [3], Rieger [3], Pan [4], Barban [7], Roth [1], Levin [2], Barban [8], [9], M. and S. Uchiyama [1], Barban [10], Wang,

Hsieh and Yu [1], Halberstam and Roth [1], Roth [2], Buchstab [1], A.I. Vinogradov [1], Bombieri [1],

Extensions of the large sieve

Andruhaev [1], Fogels [1], Goldfeld [3], Hlawka [1], [2], Huxley [1], [3], [4], Johnsen [1], Rieger [1], [2], Samandarov [1], Schaal [1], Wilson [1].

Chapter 1

Arithmetical Aids

IN THIS chapter we shall collect for the reader's convenience some of the results, which are use in later chapters, from elementary number theory. 18

1 Multiplicative functions

By

$$\mathfrak{m} \tag{1.1}$$

we denote the set of functions (defined on $\mathbb{N}$)

$$f \not\equiv 0 \tag{1.2}$$

that satisfy

$$f(nq) = f(n)f(q) \ \forall n,\ q \in \mathbb{N} \tag{1.3}$$

whenever

$$(n, q) = 1. \tag{1.4}$$

Since we have excluded the trivial f by (1.2), there is a q such that $f(q) \neq 0$ and so (1.3) implies for $n = 1$

$$f \in \mathfrak{m} \Longrightarrow f(1) = 1. \tag{1.5}$$

Obviously

$$f(n) = n^z \in \mathfrak{m} \quad \text{for every} \quad z \in \mathbb{C}, \tag{1.6}$$

so, in particular, for $z = 0$

$$f(n) = 1 \in \mathfrak{m}, \tag{1.7}$$

and for these functions the restriction (1.4) is not even necessary. A non-trivial example is provided by the Mobius function, defined by

$$\mu(n) \begin{cases} 0 & \text{if } n \text{ is not squarefree, i.e., there is a prime } p : p|^2 n, \\ (-1)^{\nu(n)} & \text{if } n \text{ is squarefree} \end{cases} \tag{1.8}$$

19 where, as usual.

$$\nu(n) = \sum_{p|n} 1 \quad (\nu(1) = 0). \tag{1.9}$$

For ν, (1.3) is easily checked, because, if atleast one of the numbers n and q is not squarefree we have zero on both sides, and if both n and q are squarefree we get (1.3) subject to (1.4) by using

$$\nu(nq) = \nu(n) + \nu(q) \quad \text{for} \quad (n, q) = 1; \tag{1.10}$$

hence

$$\mu \in \mathfrak{m} \tag{1.11}$$

A simple way of obtaining new functions is by multiplying together (1.3) for any two such functions, so that

$$f_1, f_2 \in \mathfrak{m} \Rightarrow f_1 f_2 \in \mathfrak{m}. \tag{1.12}$$

A less trivial result is

$$f_1, f_2 \in \mathfrak{m} \Rightarrow g(n) := \sum_{d|n} f_1(d) f_2\left(\frac{n}{d}\right) \in \mathfrak{m}. \tag{1.13}$$

For a proof we first note that if $d|nq$ there is, in view of (1.4), a unique factorization $d = th$ with $t|n$, $h|q$. Therefore, by considering (1.3), for both functions, under the condition (1.4), we obtain

$$\begin{cases} g(nq) = \sum\limits_{t|n} \sum\limits_{h|q} f_1(th) f_2\left(\frac{n}{t}\frac{q}{h}\right) = \\ = \sum\limits_{t|n} f_1(t) f_2\left(\frac{n}{t}\right) \sum\limits_{h|q} f_1(h) f_2\left(\frac{q}{h}\right) = g(n) g(q). \end{cases} \tag{1.14}$$

For a $f \in \mathfrak{m}$ it follows by repeated application of (1.3), subject to (1.4), that 20

$$f \in \mathfrak{m} \Rightarrow f(n) = \prod_{p|n} f(p^{a_p}) \quad \text{where} \quad n = \prod_{p|n} p^{a_p}. \tag{1.15}$$

so that these functions need only be known at prime-powers. In particular, for squarefree number one has

$$f \in \mathfrak{m} \Rightarrow f(q) = \prod_{p|q} f(p) \text{ if } \mu(q) \neq 0. \tag{1.16}$$

A good illustration of this property of functions in $\mathfrak{m}$ is given as follows. Suppose that $f_1, f_2 \in \mathfrak{m}$, and let q be a squarefree number. Then $g \in \mathfrak{m}$, where g is defined by (1.13) abd from (1.16) we get

$$\begin{cases} g(q) = \prod\limits_{p|q} g(p) = \prod\limits_{p|q}(\sum\limits_{d|p} f_1(d) f_2(\frac{p}{d})) \\ \qquad = \prod\limits_{p|q}(f_1(p) f_2(1) + f_1(1) f_2(p)) \text{ for } \mu(q) \neq 0. \end{cases} \tag{1.17}$$

Therefore, in view of (1.5), we have proved that

$$\begin{cases} f_1, f_2 \in \mathfrak{m} \Rightarrow \sum\limits_{d|q} f_1(d) + f_2(\frac{q}{d}) = \\ \qquad\qquad = \prod\limits_{p|q}(f_1(p) f_2(p)). \text{ if } \mu(q) \neq 0. \end{cases} \tag{1.18}$$

For an application we note that by (1.11) and (1.7) one can take μ and 1 for f_1 and f_2 respectively. Then, denoting by $q(n)$, the 'kernel' of n, i.e., the largest squarefree divisor of n, (1.18) yields

$$\sum_{d|n} \mu(d) = \sum_{d|q(n)} \mu(d) = \prod_{p|n} (\mu(p) + 1), \tag{1.19}$$

which gives the well-known formula

$$\sum_{d|n} \mu(d) = \begin{cases} 1 & \text{for } n = 1 \\ 0 & \text{for } n > 1. \end{cases} \tag{1.20}$$

2 Ramanujan's function

Recalling our notation (cf. (0.52)) 21

$$e(u) := e^{2\pi i u} \tag{1.21}$$

we see, on considering the partial sums of geometric series, that

$$\sum_{\ell=1}^{q} e(n\frac{\ell}{q}) = \begin{cases} q \text{ for } q/n, \\ \\ 0 \text{ for } q \nmid n \end{cases} \quad \forall q \in \mathbb{N}, n \in \mathbb{Z}. \tag{1.22}$$

The continuous analogue of (1.22) is

$$\int_0^1 e(nx)dx = \begin{cases} 1 \text{ for } n = 0, \\ 1 \text{ for } n \neq 0, \end{cases} \quad n \in \mathbb{Z}. \tag{1.23}$$

Since $e(\frac{m}{q})$ has period q, (1.22) may also be written as

$$\sum_{\ell \bmod q} e(n\frac{\ell}{q}) = \begin{cases} q \text{ for } q|n, \\ 0 \text{ for } q \nmid n, \end{cases} \quad \forall q \in \mathbb{N}, n \in \mathbb{Z}. \tag{1.24}$$

where ℓ runs through a complete system of residues modulo q. The corresponding result for (1.23) is

$$\int_\alpha^{\alpha+1} e(nx)dx = \begin{cases} 1 \text{ for } n = 0, \\ 0 \text{ for } n \neq 0, \end{cases} \quad \forall n \in \mathbb{Z}, \alpha \in \mathbb{R}. \tag{1.25}$$

For any $a_\ell' s \in \mathbb{C}$

$$\sum_{n=1}^{q} \Big| \sum_{\ell=1}^{q} a_\ell e(\frac{n\ell}{q}) \Big|^2 = \sum_{\ell_1,\ell_2=1}^{q} a_{l_1} \overline{a}_{l_2} \sum_{n=1}^{q} e((\ell_1 - \ell_2)\frac{n}{q}), \tag{1.26}$$

and because of $|\ell_1 - \ell_2| < q$ in the innermost sum, (1.22) gives

$$\frac{1}{q} \sum_{\ell=1}^{q} | \sum_{\ell=1}^{q} a_\ell e(\frac{n\ell}{q}) \Big|^2 = \sum_{\ell=1}^{q} |a_\ell|^s, \ \forall a_\ell \in \mathbb{C}. \tag{1.27}$$

Similarly, by (1.23), assuming that $\sum_n |a_n|^2 < \infty$ (so, in particular, for any finite range for m), we have

$$\int_0^1 |\sum_n e(nx)|^2 dx = \sum_{n_1,n_2} a_{n_1}\overline{a}_{n_2} \int_0^1 e((n_1-n_2)x)dx = \sum_n |a_n|^2. \quad (1.28)$$

In view of (1.25), it is obvious that (1.28) remains true if the integral is extend over any interval of length 1. **22**

Ramanujan's sum is defined by

$$\begin{cases} c_q(n) : \sum_{\ell=1}^{q} e(n\frac{\ell}{q}) = \sum_{\ell \bmod q}' e(n\frac{\ell}{q}), \ \forall q \in \mathbb{N}, n \in \mathbb{Z} \\ \text{and where } \sum_{\ell=1}^{\prime q} : \sum_{\substack{\ell=1 \\ (\ell,q)=1}}^{q} \end{cases} \quad (1.29)$$

and $\sum\limits_{\ell \bmod q}'$ means summation as ℓ runs through a reduced system of residues modulo q. In order to compute $c_q(n)$ we use (1.20). Noting that always

$$\sum_{d|q} f(d) = \sum_{d|q} f(\frac{q}{d}) \quad (1.30)$$

and writing $\ell = \ell_1 . \dfrac{q}{d}$ in the last step below we find. by (1.20),

$$\begin{cases} c_q(n) = \sum_{\ell=1}^{q} e(n\frac{\ell}{q}) \sum_{\substack{d|q \\ d|\ell}} \mu(d) = \sum_{d|q} \mu(d) \sum_{\substack{\ell=1 \\ \ell \equiv 0 \bmod d}}^{\ell} e(n\frac{\ell}{q}) = \\ \quad = \sum_{d|q} \mu(\frac{q}{d}) \sum_{\substack{\ell=1 \\ \ell=0 \bmod \frac{q}{d}}}^{q} e(n\frac{\ell}{q}) = \sum_{d|p} \mu(\frac{q}{d}) \sum_{\ell_1=1}^{d} e(n\frac{\ell_1}{d}), \end{cases} \quad (1.31)$$

so that, by (1.22), we have

$$c_q(n) = \sum_{\substack{d|q \\ d|n}} \mu(\frac{q}{d})d \quad \forall q \in \mathbb{N}, \ n \in \mathbb{Z} \quad (1.32)$$

We note the following two special cases. For $n = 0$, we see from (1.29) that Ramanujan's sum becomes Euler's function $\varphi(q)$, and (1.32) leads via (1.30) to the well-known formulae

$$\begin{cases} \varphi(q) := \sum_{\ell=1}^{q} 1 = c_q(0) = \sum_{d|q} \mu(\frac{q}{d})d = \sum_{d|q} \mu(d)\frac{q}{d} = \\ \qquad = q \sum_{d|q(q)} \frac{\mu(d)}{d} = q \prod_{p|q(q)} (-\frac{1}{p} + 1) = q \prod_{p|q} (1 - \frac{1}{p}), \end{cases} \tag{1.33}$$

23 where we have used also (1.11), (1.12), (1.6) for $z = 0$ and $z = 0$ and $z = -1$, and (1.18). Next, for $(n, q) = 1$, the right-hand side of (1.32) reduces to $\mu(q)$, so that via (1.29) we also obtain

$$\mu(q) = c_q(n) = \sum_{\ell=1}^{q}{}' \, e(n\frac{\ell}{q}) \ \text{ if } \ (n, q) = 1, q \in \mathbb{N}, n \in \mathbb{Z}. \tag{1.34}$$

3 Dirichlet's characters and Gaussian Sums

For each $q \in \mathbb{N}$ we define the arithmetic functions, names 'characters modulo q',

$$\chi(m)(\in \mathbb{C}), \quad \forall m \in \mathbb{Z}. \tag{1.35}$$

For an elementary introduction of these functions one requires following four properties (1.36) through (1.39):

$$\chi(1) = 1 \tag{1.36}$$

$$\chi(mn) = \chi(m)\chi(n) \quad \forall m, n \in \mathbb{Z}, \tag{1.37}$$

i.e., $\chi \in \mathfrak{m}$ without the restriction $(m, n) = 1$,

$$\chi(n) = \chi(\ell) \ \text{ for } \ n = \ell \bmod q \tag{1.38}$$

and

$$\chi(n) = 0 \ \text{ for } \ (n, q) > 1. \tag{1.39}$$

The relations (1.38) and (1.36) imply

$$\chi(m) = 1 \quad \text{for} \quad m \equiv 1 \bmod q. \tag{1.40}$$

By (1.39) any character vanishes for all n which are not coprime to q. On the other hand, for $(n, q) = 1$, we have by Euler's theorem $n^{\varphi(q)} \equiv 1 \bmod q$, so that (1.37) and (1.40) give

$$(\chi(n))^{\varphi(q)} = 1 \quad \text{for} \quad (n, q) = 1, \tag{1.41}$$

i,e., for $(n, q) = 1, \chi(n)$ is a $\varphi(q)$-th root of unity; in particular.

$$|\chi(n)| = 1 \quad \text{for} \quad (n, q) = 1. \tag{1.42}$$

24

It is obvious that the function

$$\chi_0(n) = \begin{cases} 1 \text{ if } (n, q) = 1, \\ 0 \text{ if } (n, q) > 1, \end{cases} \tag{1.43}$$

is a character modulo q; it is called the *principal* character $\bmod q$. A simple formula is

$$\sum_{\ell=1}^{q} \chi(\ell) = \sum_{\ell=1}^{q}{}' \chi(\ell) = \begin{cases} \varphi(q) \text{ for } \chi = \chi_0, \\ o \text{ for } \chi \neq \chi_0 \end{cases} \tag{1.44}$$

The first statement is immediate from (1.39), and for $\chi = \chi_0$ the second statement follows by (1.43). Next for $\chi = \chi_0$ there must be a number m such that

$$(m, q) = 1, \chi(m) \neq 1. \tag{1.45}$$

So when ℓ runs through a reduced residue system $\bmod q$ then the same does $m\ell$ too. Hence, in view of (1.38) and (1.37), we have

$$\chi(m) \sum_{\ell \bmod q}{}' \chi(\ell) = \sum_{\ell \bmod q}{}' \chi(\ell m) = \sum_{\ell \bmod q}{}' \chi(\ell), \tag{1.46}$$

which implies the second statement of (1.44) for $\chi \neq \chi_0$ because of $\chi(m) \neq 1$.

It can be shown that for any given $q \in \mathbb{N}$ there are exactly $\varphi(p)$ distinct functions fulfilling (1.36) through (1.39), i.e., there are $\varphi(q)$ characters $\bmod q$, a fact that can be stated in the form

$$\sum_{\chi \bmod q} 1 = \varphi(q). \tag{1.47}$$

It is easily checked that with $\chi(n)$ also $\overline{\chi}(n)$ is a character $\bmod q$. Also, with χ_1, χ_2 the function χ_1, χ_2 is a character $\bmod q$. The second remark leads us to the following statement. If for certain characters $\chi', \chi'', \chi_1 \bmod q$, $\chi'(n)\chi_1(n) = \chi''(n)\chi_1(n)$ holds for all n, then in view
25 of (1.42) and (1.39) one has $\chi' = \chi''$. Hence, if χ_1 is a fixed character $\bmod q$ and χ runs through all characters $\bmod q$, then $\chi_1\chi$ also runs through all characters $\bmod q$, a fact which can be expressed through

$$\chi_1(m) \sum_{\chi \bmod q} \chi(m) = \sum_{\chi \bmod q} (\chi_1\chi)(m) \sum_{\chi \bmod q} \chi(m), \quad \forall m \in \mathbb{Z}. \tag{1.48}$$

This leads us to the following counterpart of (1.44):

$$\sum_{\chi \bmod q} \chi(m) = \begin{cases} \varphi(q) & \text{for } m \equiv 1 \bmod q, \\ 0 & \text{for } m \not\equiv 1 \bmod q. \end{cases} \tag{1.49}$$

For, when $m \equiv 1 \bmod q$ this assertion follows from (1.40) and (1.47) and for $(m, q) > 1$ it is trivially true in view of (1.39). In this remaining case

$$m \neq 1 \bmod q, (m, q) = 1, \tag{1.50}$$

we need the result that, for any m subject to (1.50), there is a character $\chi_1 \bmod q$ such that $\chi_1(m) \neq 1$, and then the result follows from (1.48).

Next we note the following equivalence:

$$\overline{\chi}(n)\chi_1(n) = \chi_0(n), \forall n \Leftrightarrow \chi_1(n) = \chi(n). \quad \forall n. \tag{1.51}$$

This is clear for $(n, q) > 1$. For $(n, q) = 1$ we multiply on the left by $\chi(n)$ and use (1.42) and (1.43) (the latter implies always that $\chi\chi_0 = \chi$), and this step can be reversed because $\chi(n) \neq 0$. Since $\overline{\chi}\chi_1$ is a character, it may be used in (1.44), and because of (1.37) and (1.51) we obtain

$$\sum_{\ell=1}^{q}{}' \, \overline{\chi}(\ell)\chi_1(\ell) = \begin{cases} \varphi(q) & \text{for } \chi = \chi_1, \\ 0 & \text{for } \chi \neq \chi_1. \end{cases} \tag{1.52}$$

Finally we prove that

$$\sum_{\chi \bmod q} \overline{\chi}(\ell)\chi(n) = \begin{cases} \varphi(q) & \text{for } n \equiv \ell \bmod q, \\ 0 & \text{for } n \not\equiv \ell \bmod q \end{cases} \text{ if } (\ell, q) = 1. \tag{1.53}$$

26 In view of (1.39) we may assume that $(n, q) = 1$. Since $(\ell, q) = 1$, we can determine ℓ' such that

$$(\ell', q) = 1, \ell\ell' \equiv 1 \bmod q, \tag{1.54}$$

and hence by (1.40) and (1.37) one has $\chi(\ell)\chi(\ell') = 1$, which yields after multiplication by $(\overline{\chi}(\ell))$, in view of (1.42), that $\overline{\chi}(\ell) = \chi(\ell')$, so that, by using (1.37) again follows

$$\sum_{\chi \bmod q} \overline{\chi}(\ell)\chi(n) = \sum_{\chi \bmod q}{}' \chi(\ell' n). \tag{1.55}$$

On the otherhand, by the definition of ℓ', we have

$$\ell' n \equiv 1 \bmod q \Leftrightarrow n \equiv \ell \bmod q. \tag{1.56}$$

Hence (1.53) follows from (1.49) with $m = \ell' n$, using (1.55) and (1.56).

From (1.53) we now derive the following analogue of (1.27):

$$\frac{1}{\varphi(q)} \sum_{\chi \bmod q} |\sum_{\ell=1}^{q} \overline{\chi}(\ell) u_\ell|^2 = \sum_{\ell=1}^{q}{}' |u_\ell|^2, \quad \forall u_\ell \in \mathbb{C}. \tag{1.57}$$

Keeping (1.39) in mind, we see that the left-hand side equals

$$\frac{1}{\varphi(q)} \sum_{\chi \bmod q} \sum_{\ell,n=1}^{\prime q} \overline{\chi}(\ell)\chi(n) u_\ell \overline{u}_n = \sum_{\ell,n=1}^{\prime q} u_\ell \overline{u}_n (\frac{1}{\varphi(q)} \sum_{\chi \bmod q} \overline{\chi}(\ell)\chi(n)). \tag{1.58}$$

so that (1.57) follows form (1.53).

If $d|q$ and χ_1 is a character $\bmod d$, then

$$\chi(m) := \begin{cases} \chi_1(m) & \text{for } (m, q) = 1, \\ 0 & \text{for } (m, q) > 1, \end{cases} \tag{1.59}$$

is a character $\bmod q$, and we say that $\chi_1 \bmod d$ induces the character $\chi \bmod q$. If $\chi \bmod q$ is not induced by any character $\chi_1 \bmod d$ for any

$d < q$, χ is then called a *primitive* character $\bmod q$. The smallest f, $f|q$, such that a χ^* mod f induces $\chi \bmod q$ is called *conductor*of χ. 27

For any character $\chi \bmod q$ the 'Gaussian sum' is defined by

$$\tau(\chi) := \sum_{\ell=1}^{q} \chi(\ell) e(\frac{\ell}{q}) = \sum_{\ell \bmod q} \chi(\ell) e(\frac{\ell}{q}), \tag{1.60}$$

since both $\chi(\ell)$ and $e(\frac{\ell}{q})$ are of period q. By (1.39), ℓ runs actually through a reduced system of residues $\bmod q$, and so does $n\ell$ also if $(n, q) = 1$. Therefore it follows, by using (1.60) for $\overline{\chi}$, (1.37) and (1.42), that

$$\begin{cases} \tau(\overline{\chi})\chi(n) & = \chi(n) \sum_{\ell \bmod q} \overline{\chi}(n\ell) e(n\frac{\ell}{q}) = \\ & = \sum_{\ell \bmod q} c\overline{hi}(\ell) e(n\frac{\ell}{q}), \text{ for any } \chi \bmod q \text{ and } (n, q) = 1. \end{cases} \tag{1.61}$$

It requires a little more effort to prove that for primitive characters χ (1.61) holds even without the restriction $(n, q) = 1$, i.e,,

$$\tau(\overline{\chi})\chi(n) = \sum_{\ell \bmod q} \overline{\chi}(\ell) e(n\frac{\ell}{q}), \quad \text{for primitive } \chi \bmod q, n \in \mathbb{Z}. \tag{1.62}$$

If we take $a_\ell = \overline{\chi}(\ell)$ in (1.27), it follows from (1.62) that for any primitive character $\chi \bmod q$

$$\frac{1}{q} \sum_{n=1}^{q} |\tau(\overline{\chi})\chi(n)|^2 = \varphi(q) \tag{1.63}$$

or

$$|\tau(\chi)^2| = q \quad \text{for primitive } \chi \bmod q. \tag{1.64}$$

If χ is not a primitive character, let it be induced by

$$\chi^* \bmod f, q = rf, \tag{1.65}$$

where f is the conductor of χ. It can be shown that

$$\tau(\chi) = 0 \quad \text{if } (r, f) > 1 \tag{1.66}$$

and that

$$|\tau(\chi)|^2 = \mu^2(r)|\tau(\chi^*)|^2 = \mu^2(r)f, \quad \text{if } (r, f) = 1, \tag{1.67}$$

by (1.64), since χ^* is a primitive character $\mod f$. Collecting together the results (1.64), (1.66) and (1.67). we have, for the Gaussian sums, the following 28

Lemma 1.1. *If f is the conductor of $\chi \bmod q$, then*

$$q = rf \tag{1.68}$$

and

$$|\tau(\chi)^2| = \begin{cases} \mu^2(r)f & \textit{for } (r, f) = 1, \\ 0 & \textit{otherwise}. \end{cases} \tag{1.69}$$

We close this chapter with another application which demonstrates the usefulness of our characters.

Lemma 1.2. *For any a_n's in $\mathbb{C}$ and any for character $\chi \bmod q$ put*

$$S(x; q, \ell) := \sum_{\substack{n \le x \\ n \equiv \ell \bmod q}} a_n, S(x, \chi) := \sum_{n \le x} a_n \chi(n). \tag{1.70}$$

Then

$$\sum_{\ell=1}^{q}{}' |S(x; q, \ell) - \frac{S(x, \chi_0)}{\varphi(q)}|^2 = \frac{1}{\varphi(q)} \sum_{\chi \ne \chi_0} |S(x, \overline{\chi})|^2 \tag{1.71}$$

Proof. First we note that, by (1.53) (with χ being replaced by χ_1 followed by taking the complex conjugate) and (1.52), we have

$$\begin{cases} \sum\limits_{\ell=1}^{q} \overline{x}(\ell) S(x; q.\ell) & = \sum\limits_{\ell=1}^{q} \overline{x}(\ell) \sum\limits_{n \le x} a_n \frac{1}{\varphi(q)} \sum\limits_{x_1 \bmod q} X_1(\ell)\overline{x}_1(n) = \\ & = \sum\limits_{n \le x} a_n \sum\limits_{x_1 \bmod q} \overline{x}_1(n) \frac{1}{\varphi(q)} \sum\limits_{\ell=1}^{q}{}' \overline{x}(\ell) x_1(\ell) = \\ & = S(x, \overline{x}). \end{cases} \tag{1.72}$$

□

Next, taking $u_\ell = S(x; q, \ell) - \frac{S(x, \chi_0)}{\varphi(q)}$ in (1.57), it follows that

$$\sum_{\ell=1}^{q}{}' \left| S(x; q.\ell) - \frac{S(x, \chi_0)}{\varphi(q)} \right|^2 = \frac{1}{\varphi(q)} \sum_{\chi \bmod q} \left| \sum_{\ell=1}^{q} \bar{\chi}(\ell)(S(x; q, \ell) - \frac{S(x, \chi_0)}{\varphi(q)}) \right|^2. \tag{1.73}$$

29 According to (1.72) and (1.44). we have

$$\sum_{\ell=1}^{q} \bar{\chi}(\ell)(S(x; q, \ell) - \frac{S(x, \chi_0)}{\varphi(q)}) = \begin{cases} 0 & \text{for } \chi = \chi_0, \\ S(x, \bar{\chi}) & \text{for } \chi \neq \chi_0, \end{cases} \tag{1.74}$$

and using this in (1.73) we obtain (1.71).

For the theory of prime numbers the important special case

$$\psi(x; q, \ell) := \sum_{\substack{n=x \\ n \equiv \ell \bmod q}} \Lambda(n), \psi(x, \chi) : \sum_{n \leq x} \Lambda(n)\chi(n) \tag{1.75}$$

yields, on nothing that $\Lambda(n)$ is a real-valued function, the identity

$$\sum_{\ell}^{q}{}' (\psi(x; q, \ell) - \frac{\psi(x, \chi_0)}{\varphi(q)})^2 = \frac{1}{\varphi(q)} \sum_{\chi \neq \chi_0} |\psi(x, \chi)|^2. \tag{1.76}$$

This formula was the starting point in the proof of the mean-value theorem of Davenport and Halberstam [2] (cf. (6.34)).

NOTES

(1.27): As another example, take in (1.27)

$$a_\ell = \begin{cases} 1 & \text{for } (\ell, q) = 1, \\ 0 & \text{for } (\ell, q) > 1. \end{cases}$$

Then, in view of (1.29),

$$\sum_{n=1}^{q} c_q^2(n) = q\varphi(q). \tag{1.77}$$

3: For the proofs omitted in this section we refer, for instance, to Davenport [1] (Chapter 9) and Huxley [7] (Chapter 3).

Chapter 2

The Large Sieve

WE START by recalling the notation introduction in Chapter 0 (cf. (0.1) 30
and (0.52)). Let γ be a set consisting of S integers from an interval of length N:

$$\gamma \subset (M, M+N], M \in \mathbb{Z}, N \in \mathbb{N}, S := |\gamma|. \tag{2.1}$$

and let

$$T(x) := \sum_{n \in \gamma} e(nx). \tag{2.2}$$

Then, as has been indicated in Chapter 0, the large sieve in its basic form is concerned with the estimation, of the expression

$$\sum_{q \leq Q} \sum_{\ell=1}^{q}{}' \left| T(\frac{\ell}{q}) \right|^2 \tag{2.3}$$

in terms of N, Q and S, of the form

$$\sum_{q \leq Q} \sum_{\ell=1}^{q}{}' \left| T(\frac{\ell}{q} \right|^2 \leq C(N, Q)S. \tag{2.4}$$

The simplest approach to a bound of the type in (2.4) is now due to Gallagher [1]. Gallagher's starting point is the following lemma which occurs in the earlier work of Hardy and Littlewood, and of Sobolev (for several variables):

Lemma 2.1. *Let u and $\delta(>0)$ be real numbers, and let $f(x)$ (complex-valued) be continuous on $\left[u-\frac{\delta}{2}, u+\frac{\delta}{2}\right]$ with a continuous derivative in $(u-\frac{\delta}{2}, u+\frac{\delta}{2})$. Then*

$$|f(u)|^2 \leq \int\limits_{u-\frac{\delta}{2}}^{u+\frac{\delta}{2}} |f(x)f'(x)|dx + \delta^{-1} \int\limits_{u-\frac{\delta}{2}}^{u+\frac{\delta}{2}} |f(x)|^2 dx. \tag{2.5}$$

Proof. Put

$$F(x) = f^2(x). \tag{2.6}$$

Then, by partial integration, we find the identity

$$\begin{cases} F(u) & = \delta^{-1} \int\limits_{u-\frac{\delta}{2}}^{u} (x(u-\frac{\delta}{2}))F'(x)dx + \delta^{-1} \int\limits_{u}^{u+\frac{\delta}{2}} (x-(u+\frac{\delta}{2}))F'(x)dx + \\ & +\delta^{-1} \int\limits_{u-\frac{\delta}{2}}^{u+\frac{\delta}{2}} F(x)dx. \end{cases} \tag{2.7}$$

31 Hence

$$|F(u)| \leq \frac{1}{2} \int\limits_{u-\frac{\delta}{2}}^{u+\frac{\delta}{2}} |F'(x)|dx + \delta^{-1} \int\limits_{u-\frac{\delta}{2}}^{u+\frac{\delta}{2}} |F(x)|dx, \tag{2.8}$$

which gives (2.5) on using (2.6).

Gallagher's use of the above lemma enables one to obtain a bound of the type in (2.4) when the expression on the left there is extended in two directions; namely, when the terms of the sum in (2.2) have weights (arbitrary complex numbers) a_n and the set of (fractional) points $\frac{\ell}{q}$ in (2.4) are replaced by a finite set of real numbers which are distinct modulo 1. □

Theorem 2.1. *For any complex numbers a_n, $M < n \leq M + N$, set*

$$U(x) := \sum_{M<n\leq M+N} a_n e(nx). \tag{2.9}$$

Let $x_1, \ldots, x_R$ be real numbers which are distinct $\bmod 1$ and put

$$\delta := \min_{\substack{r,s \\ r\neq s}} \|x_r - x_s\|, \text{ if } R \geq 2, \|x\| := \min_{k\in\mathbb{Z}} |x-k|; \delta := \infty, \text{ if } R = 1. \tag{2.10}$$

Then

$$\sum_{r=1}^{R} |U(x_r)|^2 \leq (\pi N + \delta^{-1}) \sum_{M<n\leq M+N} |a_n|^2. \tag{2.11}$$

We shall present the proof of this theorem in all its detail. For some proofs of large sieve inequalities, in particular, for the proof of Theorem 2.1 which is based on Lemma 2.1 (cf. also the proof of (2.95)), it is a advantageous (cf. (2.30)) to consider a shifted interval (usually, an **32**
interval symmetric about the point zero) instead of the interval $(M + M + N]$. Then the general case is easily derived by reversing the shifting procedure. The idea of the last step is contained in the following

Lemma 2.2. *Let $N \in \mathbb{N}$, let $x_1, \ldots, x_R (R \geq 2)$ be real numbers which are district $\bmod 1$ and put*

$$\delta := \min_{\substack{r,s \\ r\neq s}} \|x_r - x_s\|. \tag{2.12}$$

Set

$$V(x) := \sum_{-\frac{N}{2}<m\leq\frac{N}{2}} b_m e(mx), b_m \in \mathbb{C}. \tag{2.13}$$

Then the inequality

$$\sum_{r=1}^{R} |V(x_r)|^2 \leq \Delta(N, \delta) \sum_{-\frac{N}{2}<m\leq\frac{N}{2}} |b_m|^2, \quad \forall b_m \in \mathbb{C} \tag{2.14}$$

with some (positive) function $\Delta(N, \delta)$ depending only on N and δ implies that for any $M \in \mathbb{Z}$,

$$\sum_{r=1}^{R} |U(x_r)|^2 \leq \Delta(N, \delta) \sum_{M<n\leq M+N} |a_n|^2 \ \forall a_n \in \mathbb{C} \tag{2.15}$$

where

$$U(x) := \sum_{M<n\leq M+N} a_n e(nx). \tag{2.16}$$

Proof of lemma 2.2. If $U(x)$ is defined by (2.16), let us consider

$$\begin{cases} V(x) : & = e(-(M+\left[\frac{N+1}{2}\right])x)U(x) = \sum\limits_{M<n\leq M+N} a_n e((n-M-\left[\frac{N+1}{2}\right])x) = \\ & = \sum\limits_{-\left[\frac{N+1}{2}\right]<m\leq -\left[\frac{N+1}{2}\right]+N} b_m e(mx) = \sum\limits_{-\frac{N}{2}<m\leq\frac{N}{2}} b_m e(mx), \end{cases} \tag{2.17}$$

33 where

$$b_m = a_{m+M+\left[\frac{N+1}{2}\right]}, -\frac{N}{2} < m \leq \frac{N}{2}. \tag{2.18}$$

Now $|U(x)| = |V(x)|$ for all real x from which we easily see that (2.15) is an immediate consequence of (2.14).

Now we are in a position to prove Theorem 2.1.

Proof of Theorem 2.1. To start with we dispose off the case $R = 1$. This is easily done in view of (2.10) and Cauchy's inequality:

$$|U(x_1)|^2 \leq N \sum_{M<n\leq M+N} |a_n|^2. \tag{2.19}$$

So we assume that

$$R \geq 2 \tag{2.20}$$

and further, because of Lemma 2.2, we shall consider V instead of U. Since $V(x)$ is of period 1 we can also suppose that

$$0 \leq x_1 < x_2 < \ldots < x_R < 1. \tag{2.21}$$

By the pigeon-hole principle follows easily that

$$\delta \leq \frac{1}{R}. \tag{2.22}$$

Now, from Lemma 2.1 with $f = V$ and $u = x_r$, one has

$$|V(x_r)|^2 \leq \int\limits_{I_r} |V(x)V'(x)|dx + \delta^{-1} \int\limits_{I_r} |V(x)|^2 dx, \tag{2.23}$$

where

$$I_r := [x_r - \frac{\delta}{2}, x_r + \frac{\delta}{2}]. \tag{2.24}$$

By (2.12), we have for $r \neq s$

$$\delta \leq \|x_r - x_s\| \leq |x_r - x_s|, \tag{2.25}$$

so that our intervals I_r do not overlap and their total length, i.e., length of 34

$$\bigcup_{r=1}^{R} I_r, \tag{2.26}$$

equals

$$\delta R \leq 1, \tag{2.27}$$

on recalling (2.22). Summing now over r in (2.23) (and using that both $V(x)$ and $V'(x)$ have period 1) we can replace the integration on the right over (2.26) by $\int\limits_0^1$. Thus we get, by employing Schwarz's inequality.

$$\sum_{r=1}^{R} |V(x_r)|^2 \leq (\int\limits_0^1 |V(x)|^2 dx)^{1/2} (\int\limits_0^1 |V'(x)|^2 dx)^{\frac{1}{2}} + \delta^{-1} \int\limits_0^1 |V(x)|^2 dx. \tag{2.28}$$

Now, it follows from (1.28) that

$$\int\limits_0^1 |V(x)|^2 dx = \sum_{-\frac{N}{2} < m \leq \frac{N}{2}} |b_m|^2 \tag{2.29}$$

and also

$$\int\limits_0^1 |V'(x)|^2 dx = \sum_{-\frac{N}{2} < m \leq \frac{N}{2}} |2 b_m \pi_m|^2 \leq N^2 \pi^2 \sum_{-\frac{N}{2} < m \leq \frac{N}{2}} |b_m|^2. \tag{2.30}$$

Hence (2.28) yields

$$\sum_{r=1}^{R} |V(x_r)|^2 \leq (\pi N + \delta^{-1}) \sum_{-\frac{N}{2} < m \leq \frac{N}{2}} |b_m|^2. \tag{2.31}$$

This is (2.14) with

$$\Delta(N, \delta) = \pi N + \delta^{-1}. \tag{2.32}$$

Therefore, by lemma 2.2, (2.15) with (2.32) gives (2.11) thereby completing the proof of Theorem 2.1.

Now we discuss the result of Theorem 2.1. Due to the presence of
35 the factor δ^{-1} the efficiency of (2.11), as of other large sieve inequalities that we shall consider, depends on the information as to how 'well-spaced' (in the sense of (2.10) the points x_r are.

The simplest case is, for any $(2 \leq) R \in \mathbb{N}$. with

$$x_r = \frac{r}{R}, \quad 1 \leq r \leq R \tag{2.33}$$

so that

$$\delta = \frac{1}{R}. \tag{2.34}$$

Now Theorem 2.1 gives

$$\sum_{r=1}^{R} \left| U\left(\frac{r}{R}\right) \right|^2 \leq (\pi N + R) \sum_{M < n \leq M+N} |a_n|^2 \tag{2.35}$$

In the more interesting case

$$x_r = \frac{\ell}{q}, \ 1 \leq \ell \leq q \leq Q, \ (\ell, q) = 1, \tag{2.36}$$

the set of Farey fractions of ordér Q, the points are not quite so well-spaced. We find, on assuming that $Q \geq 2$, for any two distinct Farey fractions in (2.36)

$$\left\|\frac{\ell}{q} - \frac{\ell'}{q'}\right\| = \left\|\frac{\ell q' - q\ell'}{qq'}\right\| \geq \frac{1}{qq'} \geq \frac{1}{Q^2}, \tag{2.37}$$

i.e.

$$\delta \geq \frac{1}{Q^2}. \tag{2.38}$$

Hence we get from Theorem 2.1 and (2.19) the following

Theorem 2.2. *Let $M \in \mathbb{Z}$, $N \in \mathbb{N}$ and let $a_n (M < n \leq M+N)$ be arbitrary complex numbers. Then*

$$\sum_{q \leq Q} \sum_{\ell=1}^{q}{}' |U(\frac{\ell}{q})|^2 \leq (\pi N + Q^2) \sum_{M<n\leq M+N} |a_n|^2 \ \forall Q \in \mathbb{N}, \tag{2.39}$$

where $U(x)$ is defined through (2.9).

Regarding the quality of the preceding results we make the following remarks. Under the assumptions of Theorem 2.1, recall the estimation (2.15): **36**

$$\sum_{r=1}^{R} |U(x_r)|^2 \leq \Delta(N, \delta) \sum_{M<n\leq M+N} |a_n|^2. \tag{2.40}$$

In some application the bound in (2.11) is quite satisfactory. However, if we are inserted in better estimates we need apply more effective tools in the proof.

In order to see some necessary conditions for having a general result of the form (2.40) first note that if $x_1 = 0$ and $a_n = 1$ for $M < n \leq M+N$ then the left hand side is at least $|U(x_1)|^2 = N^2$, so that

$$\Delta(N, \delta) \geq N. \tag{2.41}$$

Next, since δ is invariant under a translation of the set $x_1, \ldots, x_R$ by any given $x \in \mathbb{R}$, (2.40) would also simply that

$$\sum_{r=1}^{R} |U(x_r + x)|^2 \leq \Delta(N, \delta) \sum_{M<n\leq M+N} |a_n|^2 \ \text{ for every } \ x \in \mathbb{R}. \tag{2.42}$$

Integrating this with respect to x over an interval of length 1 and using (1.28), we see that

$$R \sum_{M<n\leq M+N} |a_n|^2 = \sum_{r=1}^{R} \int_0^1 |U(x_r + x)|^2 dx \leq \Delta(N, \delta) \sum_{M<n\leq M+N} |a_n|^2. \tag{2.43}$$

Therefore from our example (2.33) of equally-spaced points follows that

$$\Delta(N, \delta) \geq \delta^{-1}. \tag{2.44}$$

Furthermore, Bombieri and Davenport [3] have given examples from which one gets

$$\Delta(N, \delta) \geq N + \delta^{-1} - 1. \tag{2.45}$$

These remarks can be considered as negative ones.

In the positive direction the following result, due to Montgomery and Vaughan [2], leaves only a small gap when compared with (2.45).
37 Since N is usually large in applications this difference in minor.

Theorem 2.3. *Under the hypotheses of Theorem 2.1, we have*

$$\sum_{r=1}^{R} |U(x_r)|^2 \leq (N + \delta^{-1}) \sum_{M<n\leq M+N} |a_n|^2. \tag{2.46}$$

(In what follows we assume that $R \geq 2$ (cf. (2.20)).)

Their proof uses the principle of duality due to Hellinger-Toeplitz, namely that for an $R\times N$ matrix (c_{rn}) with complex entries and a constant A

$$\sum_{n=1}^{N} |\sum_{r=1}^{R} c_{rn} v_r|^2 \leq A \sum_{r=1}^{R} |v_r|^2, \ \forall v_r \in \mathbb{C} \tag{2.47}$$

implies that (and is implied by)

$$\sum_{r=1}^{R} |\sum_{n=1}^{N} c_{rn} w_n|^2 \leq A \sum_{n=1}^{N} |w_r|^2, \forall w_n \in \mathbb{C} \tag{2.48}$$

and the following extension of Schur's result regarding Hilbert's inequality (Montgomery and Vaughan [1]):

Lemma 2.3. *Under the hypotheses of Theorem 2.1, we have*

$$|\sum_{r=1}^{R} \sum_{\substack{s=1 \\ s\neq r}}^{R} u_r \bar{u}_s cosec(\pi(x_r - x_s))| \leq \delta^{-1} \sum_{r=1}^{R} |u_r|^2, \forall u_r \in \mathbb{C}. \tag{2.49}$$

Proof of Lemma 2.3. First we can impose the normalization condition

$$\sum_{r=1}^{R} |u_r|^2 = 1. \tag{2.50}$$

Also we can assume, since the double-sum in (2.49) is a skew-hermitian form, that the (u) which makes the left-hand side there maximum satisfies

$$\sum_{\substack{r=1\\ r\neq s}} u_r \text{ cosec } (\pi(x_r - x_s)) = i\lambda u_s, \quad 1 \leq s \leq R \tag{2.51}$$

with some (real) λ. Thus it suffices to show (under (2.50) and (2.51)) that

$$|\lambda| \leq \delta^{-1} \tag{2.52}$$

38

Further, we can assume that all the x_r's lie in the interval (0,1] (without any loss of generality).

We have, by (2.15) and (2.50),

$$\begin{cases} |\lambda|^2 & = \sum\limits_{s=1}^{R} |\sum\limits_{\substack{r=1\\ r\neq s}}^{R} ur \text{ cosec } (\pi(x_r - x_s))|^2 = \\ & = \sum\limits_{s=1}^{R} \sum\limits_{\substack{r=1\\ r\neq s}}^{R} \sum\limits_{\substack{t=1\\ t\neq s}}^{R} u_r \overline{u_t} \text{ cosec } (\pi_r - x_s)) \text{ cosec } (x(x_t - x_s)) = \\ & = \sum_1 + \sum_2, \end{cases} \tag{2.53}$$

where

$$\begin{cases} \sum_1 & = \sum\limits_{s=1}^{R} | \sum\limits_{\substack{r=1\\ r\neq s}}^{R} |u_r|^2 \; cosec^2(\pi(x_r - x_s)) \\ \text{and} & \\ \sum_2 & = \sum\limits_{\substack{r=1\\ r\neq t}}^{R} \sum\limits_{t=1}^{R} u_r \overline{u_t} \sum\limits_{\substack{s=1\\ s\neq r\\ s\neq t}}^{R} \text{ cosec } (\pi(x_r - x_s)) \text{ cosec } (\pi(x_t - x_s)). \end{cases} \tag{2.54}$$

Using the identity (if $R \geq 3$-for $R = 2$ the inequality (2.63) below is trivial, because, then $\sum_2 = 0$)

$$\begin{cases} \operatorname{cosec}\ \theta_1 \operatorname{cosec}\theta_2 & = \operatorname{cosec}(\theta_1 - \theta_2)(\cot\theta_2 - \cot\theta_1), if \theta_1\theta_2(\theta_1 - \theta_2) \neq 0. \\ & -\pi < \theta_1, \theta_2, \theta_1 - \theta_2 < \pi, \end{cases}$$

we see that

$$\sum_2 = \sum_3 - \sum_4, \tag{2.55}$$

where

$$\begin{cases} \sum_3 = \sum_{\substack{r=1 \\ r\neq t}}^R \sum_{t=1}^R u_r\bar{u}_t\ cosec(\pi(x_r - x_t)) \sum_{\substack{s=1 \\ s\neq r, s\neq t}}^R cot(\pi(x_t - x_s)) \\ \text{and} \\ \sum_4 = \sum_{\substack{r=1 \\ r\neq t}}^R \sum_{t=1}^R u_r\bar{u}_t\ cosec(\pi(x_r - x_t)) \sum_{\substack{s=1 \\ s\neq r, s\neq t}}^R cot(\pi(x_t - x_s)). \end{cases} \tag{2.56}$$

39 Denoting

$$b_r := \sum_{\substack{s=1 \\ s\neq r}}^R cot(\pi(x_r - x_s)), \quad 1 \leq r \leq R, \tag{2.57}$$

we will have

$$\sum_3 = \sum_{\substack{r=1 \\ r\neq t}}^R \sum_{t=1}^R u_r\bar{u}_t\ cosec(\pi(x_r - x_t))b_t - \sum_{31} \tag{2.58}$$

with

$$\sum_{31} \sum_{\substack{r=1 \\ r\neq t}}^R \sum_{t=1}^R u_r\bar{u}_t \operatorname{cosec}(\pi(x_r - x_t)) \cot(\pi(x_t - x_r)) = \text{ Re } \sum_{31}. \tag{2.59}$$

Now, from (2.58) and (2.15), we obtain

$$\sum_3 + \sum_{31} = \sum_{r=1}^R \bar{u}_t b_t i\lambda u_t = i\lambda \sum_{t=1}^R b_t |u_t|^2. \tag{2.60}$$

Treating $\sum_4$ of (2.56) similarly, we also get

$$\sum_3 + \sum_{31} = \sum_4 + \sum_{41}, \sum_{41} = -\sum_{31}, \tag{2.61}$$

so that, from (2.55) and (2.59). follows

$$\sum_2 = -2 \text{ Re } \sum_{31}. \tag{2.62}$$

Hence

$$|\sum_2| \leq 2|\sum_{31}| \leq \sum_{\substack{r=1\\ r\neq t}}^{R} \sum_{t=1}^{R} (|u_r|^2| + |u_t|^2)| \operatorname{cosec}^2(\pi(x_r - x_t)) \cos)\pi(x_r - x_t))|, \tag{2.63}$$

which yields in view of symmetry, by (2.53) and (2.54),

$$|\lambda|^2 \leq \sum_{\substack{r=1\\ r\neq t}}^{R} \sum_{t=1}^{R} (|U_r|^2 \operatorname{cosec}^2(\pi(x_r - x_t))(1 + 2|\cos(\pi(x_r - x_t))|). \tag{2.64}$$

Observing here that $\sin^2(\pi\theta) = \sin^2(\pi\|\theta\|)$ and $|\cos(\pi\theta)| = |\cos(\pi\|\theta\|)|$ (for any real θ) and employing the inequality

$$\operatorname{cosec}^2 \theta(1 + 2\cos\theta) \leq 3\theta^{-2} \text{ for } 0 < \theta \leq \frac{\pi}{2}, \tag{2.65}$$

we obtain further **40**

$$|\lambda|^2 \leq 3\pi^{-2} \sum_{r=1}^{R} |u_r|^2 \sum_{\substack{t=1\\ t\neq r}}^{R} \|x_r - x_t\|^{-2}. \tag{2.66}$$

Since the value of the inner sum here is unaltered when x's are translated by integers, we can arrange, for any given x_r. translates of all x_t's $(t \neq r)$ to lie in the interval $(x_r - \frac{1}{2}, x_r + \frac{1}{2})$. Then the inner sum in (2.66) is easily majorized, because of (2.10), by $2 \sum_{j=1}^{\infty} (\delta j)^{-2}$. So (on recalling (2.50)) holds

$$|\lambda|^2 \leq 6\pi^{-2}\delta^{-2} \sum_{j=1}^{\infty} j^{-2} = \delta^{-2}. \tag{2.67}$$

This is (2.52) and thereby Lemma 2.3 is completely proved.

Now we are in a position to deduce Theorem 2.3 from (2.47) to (2.49).

Proof of Theorem 2.3. Taking

$$c_{rn} = e((M+n)x_r), \qquad w_n = a_n, \tag{2.68}$$

it suffices for a proof of (2.46), in view of (2.47) and (2.48), to show that

$$\sum_{n=1}^{N} |\sum_{r=1}^{R} c_{rn} v_r|^2 (N+\delta^{-1}) \sum_{r=1}^{R} |v_r|^2, \ \forall v_r \in \mathbb{C}. \tag{2.69}$$

Expanding the left-hand side here we obtain from the diagonal terms the contribution

$$N \sum_{r=1}^{R} |v_r|^2. \tag{2.70}$$

The remaining part amounts to

$$\sum_{\substack{r=1 \\ r\neq s}}^{R} \sum_{s=1}^{R} v_r \overline{v}_S \sum_{M<n\leq M+N} e(n(x_r - x_S)), \tag{2.71}$$

and the inner sum here is

$$\frac{1}{2} i\{e((M+\frac{1}{2})(x_r - x_S)) - e((M+N+\frac{1}{2})(x_r - x_S))\} \operatorname{cosec}(\pi(x_r - x_S)). \tag{2.72}$$

41

We apply (2.49) twice with the choices

$$u_r = v_r e((M+\frac{1}{2})x_r) \text{ and } u_r = V_r e((M+N+\frac{1}{2})x_r). \tag{2.73}$$

Then, because of the factor $\frac{1}{2}$ in (2.72), we obtain that the contribution (2.71) is

$$\leq \delta^{-1} \sum_{r=1}^{R} |v_r|^2. \tag{2.74}$$

and this in combination with (2.70) yields (2.69).

Montgomery and Vaughan [2] have also obtained a more sophisticated form of the large sieve, which has turned out to be extremely powerful in arithmetical applications. The weights attached here enable one to take care of the irregular spacing of Farey fraction (cf. Theorems 2.5 and 2.6 below).

Theorem 2.4. *Under the assumption of Theorem 2.1, put*

$$\delta_r := \min_{\substack{s \\ s \neq r}} \|x_r - x_s\|. \tag{2.75}$$

Then

$$\sum_{r=1}^{R} (N + \frac{3}{2}\delta_r^{-1})^{-1} |U(x_r)|^2 \leq \sum_{M < n \leq M+N} |a_n|^2. \tag{2.76}$$

The proof is very similar to that of Theorem 2.3. The essential change is the following version, involving $\delta_r' s$, of Lemma 2.3 (Montgomery and Vaughan [1]):

Lemma 2.4. *Under the hypothesis and notation of Theorem 2.4, there holds*

$$| \sum_{\substack{r=1 \\ r \neq s}}^{R} \sum_{s=1}^{R} u_r \bar{u}_S \text{ cosec } (\pi(x_r - x_S))| \leq \frac{3}{2} \sum_{r=1}^{R} |u_r|^2 \delta_r^{-1}, \ \forall u_r \in \mathbb{C}. \tag{2.77}$$

Here we shall only conclude Theorem 2.4 using (2.47), (2.48) and (2.77).

Proof of Theorem 2.4. Now we put, instead of (2.68), 42

$$c_{rn} = (N + \frac{3}{2}\delta_r^{-1})^{-\frac{1}{2}} e((M+n)x_r), w_n = a_n. \tag{2.78}$$

Then the diagonal terms, from the expression on the left of (2.47), contribute

$$N \sum_{r=1}^{R} (N + \frac{3}{2}\delta_r^{-1})^{-1} |v_r|^2, \tag{2.79}$$

and the remaining part is

$$\sum_{\substack{r=1\\ r\neq s}}^{R}\sum_{s=1}^{R} v_r(N+\frac{3}{2}\delta_r^{-1})^{-\frac{1}{2}}\overline{v_s}(N+\frac{3}{2}\delta_s^{-1})^{-\frac{1}{2}}\sum_{M<n\leq M+N} e(n(x_r-x_s)). \tag{2.80}$$

Again the inner sum is given by (2.72). Now we apply (2.77) twice with the choices

$$u_r = v_r(N+\frac{3}{2}\delta_r^{-1})^{-\frac{1}{2}}e((M+\frac{1}{2})x_r) \text{ and } u_r = v_r(N+\frac{3}{2}\delta_r^{-1})^{-\frac{1}{2}}e((M+N+\frac{1}{2})x_r). \tag{2.81}$$

Then, because of the factor $\frac{1}{2}$ in (2.72), we obtain that the contribution (2.80) is

$$\leq \frac{3}{2}\sum_{r=1}^{R}|v_r|^2(N+\frac{3}{2}\delta_r^{-1})^{-1}\delta_r^{-1}. \tag{2.82}$$

Now, this together with (2.79) proves (2.47) with $A = 1$ for our choice (2.78) of c_{rn}. Therefore (2.48) with the above choice (2.78) of w_n yields (2.76), thereby proving Theorem 2.4.

Let us now specialize again, We assume that $Q \geq 2$, because the theorems that follow are trivially true for $Q = 1$ as before (cf. (2.19)). Take

$$x_r = \frac{\ell}{q}, 1 \leq \ell \leq q \leq Q, (\ell, q) = 1, \tag{2.83}$$

so that we have (cf. (2.37)), for any two distinct Farey fractions of (2.83).

$$\left\|\frac{\ell}{q}-\frac{\ell'}{q'}\right\| \geq \frac{1}{qq'} \geq \frac{1}{qQ} \geq \frac{1}{Q^2} \tag{2.84}$$

which shows that in Theorem 2.3 and Theorem 2.4 the quantities Q^{-2}
43 and $q^{-1}Q^{-2}$ are permissible lower bounds for δ and δ_r, respectively. Therefore we obtain from these theorems

Theorem 2.5. *For any complex number a_n, put*

$$U(x) := \sum_{M<n\leq M+N} a_n e(nx). \tag{2.85}$$

Then

$$\sum_{q \le Q} \sum_{\ell=1}^{q}{}' \left| U\left(\frac{\ell}{q}\right) \right|^2 \le (N + Q^2) \sum_{M<n\le M+N} |a_n|^2 \qquad (2.86)$$

and

$$\sum_{q \le Q} (N + \frac{3}{2} qQ)^{-1} \sum_{\ell=1}^{q}{}' |U(\frac{\ell}{q})|^2 \le \sum_{M<n\le M+N} |a_n|^2. \qquad (2.87)$$

Finally, returning to the beginning of the this chapter, i.e.,

$$a_n = \begin{cases} 1 & \text{if } n \in \gamma, \\ 0 & \text{if } n \notin \gamma, \end{cases} \qquad (2.88)$$

where $U(x) \equiv T(x)$, we note that Theorem 2.5 contains, with respect to (2.4), the following

Theorem 2.6. *Let γ be a set of S integers from an interval $(M, M + N]$ and put*

$$T(x) := \sum_{n \in \gamma} e(nx). \qquad (2.89)$$

Then

$$\sum_{q \le Q} \sum_{l=1}^{q}{}' \left| T\left(\frac{\ell}{q}\right) \right|^2 \le (N + Q^2) S \qquad (2.90)$$

and

$$\sum_{q \le Q} (N + \frac{3}{2} qQ)^{-1} \sum_{\ell=1}^{q}{}' \left| T\left(\frac{\ell}{q}\right) \right|^2 \le S. \qquad (2.91)$$

NOTES

The (explicit) qualitative version of (2.39) occurs for the first time in Bombieri [1] with the factor $7 \max(N, Q^2)$. His method may be considered to be a refinement of Linnik's, **44**

An improvement of Bombieri's factor as well as the extension from Farey fractions to well-spaced points, i.e., (2.40), is due to Davenport and Halberstam [1]. Their substantial improvement of the method is based on a convolution of $U(x)$ with a suitable auxiliary function, an idea introduced by Roth [2]. Subsequent developments and discussions along these lines were given by Bombieri and Davenport [2], [3] (and by Liu [1]). The paper of Bombieri and Davenport [3] contains also various investigations under different assumptions about the relative orders of N and δ. In particular, they have proved the only result that still supersedes, under certain conditions, the Theorem 2.3; namely

$$\Delta(N, \delta) \le \delta^{-1} + 270N^3\delta^2 \quad \text{if} \quad N\delta \le \frac{1}{4}, \tag{2.92}$$

which is superior to (2.46) if $(N\delta)^2 < \frac{1}{270}$.

The best general result known to date, Theorem 2.3, is due to Montgomery and Vaughan [2]. They have pointed out (cf. Montgomery and Vaughan [1]) that it is possible to replace the factor δ^{-1} in (2.49), and so also in (2.46), by $\delta^{-1}-c$ for some $c > 0$. Indeed, from the remarks subsequent to (2.66) we see that the inner sum in (2.66) can be more precisely estimated as follows: Let J_1 and J_2 denote the number of translates of $x_t's$ which lie in the intervals $(x_r - \frac{1}{2}, x_r)$ and $(x_r, x_r + \frac{1}{2}]$ respectively, so that $J_1 + J_2 = R - 1$. Then the above-mentioned inner sum is at most (cf. (2.10))

$$\begin{cases} \delta(\sum_{j=1}^{J_1} \frac{1}{j^2} + \sum_{j=1}^{J_2} \frac{1}{J^2}) \le \delta^{-2}(\frac{\pi^{-2}}{3} - \int_{J_1+1}^{\infty} \frac{du}{U^2} - \int\limits_{J_2+1}^{\infty} \frac{du}{u^2}) = \\ = \delta^{-2}(\frac{\pi^2}{3} - \frac{(R+1)}{(J_1+1)(J_2+1)}) \le \delta^{-2}(\frac{\pi^2}{3} - \frac{4}{R+1}) \le \frac{\pi^2}{3}\delta^{-2} - \frac{8}{3}\delta^{-1} \end{cases} \tag{2.93}$$

45 on using $R \ge 2$, which implies in view of (2.22) that $(R + 1) \le \frac{3}{2}\delta^{-1}$. Now, on combining (2.66), (2.93) and (2.50) we obtain

$$|\lambda|^2 \le \delta^{-2} - \frac{8}{\pi^2}\delta^{-1} \le (\delta^{-1} - \frac{4}{\pi^2}). \tag{2.94}$$

Thus one can take, for instance, $c = \frac{4}{\pi^2}$ in the above remark. However, according to (2.45), it is not possible to obtain these general estimates with any $c > 1$. The essential tools for the proofs of the Theorem

2.3 and 2.4 were developed in Montgomery and Vaughan [1] (cf. Montgomery [6]). Earlier similar approaches had been discussed in the work of Elliott [7], Mathews [1], [2], [3], and Kobayashi [1].

An intermediate result of Bombieri ([4] and p. 17 of [6], namely

$$\Delta(N, \delta^{-1}) \leq N + 2\delta^{-1}, \tag{2.95}$$

is based on Theorem 0.1. Under the assumption of Theorem 2.1, considering the sum $\sum_{-N}^{N} a_n e(nx)$ (cf. Lemma 2.2) instead of $U(x)$, he takes the Hilbert space ℓ^2 of sequences $\alpha = \{\alpha_n\}$ with $(\alpha, \beta) := \sum_{-\infty}^{\infty} \alpha_n \overline{\beta}_n$, $\|\alpha\|^2 = \sum_{-\infty}^{\infty} |\alpha_n|^2$. Choosing

$$L \in \mathbb{N}, f = \begin{cases} a_n \text{ if } |n| \leq N, \\ 0 \text{ if } |n| > N, \varphi_r = \begin{cases} e(-nx_r) & \text{if } |n| \leq N \\ (\frac{N+L-|n|}{L})^{1/2} e(-nx_r) & \text{if } N < |n| \leq N + L. \\ 0 & \text{if } |n| > N + L, \end{cases} \end{cases} \tag{2.96}$$

one gets

$$\|f\|^2 = \sum_{-N}^{N} |a_n|^2, (f, \varphi_r) = \sum_{-N}^{N} a_n e(nx_r), \ r = 1, \ldots, R. \tag{2.97}$$

Bombieri proves that for each r, $1 \leq r \leq R$

$$\sum_{S=1}^{R} |(\varphi_r, \varphi_s)| \leq 2N + L + \frac{\pi^2}{12}.\frac{1}{L\delta^2}, \tag{2.98}$$

Now (2.95) follows from Theorem 0.1 or from (0.59), by taking **46**

$$L = \left[\frac{1}{\delta}\right]. \tag{2.99}$$

(2.10) : or $\|X\| := |x - [x + \frac{1}{2}]|$.

(2.41), (2.44) : cf. Montgomery and Vaughan [2].

(2.47), (2.48) : cf. Hardy, Littlewood and Polya, *Cambridge* Theorem 288.

(2.51) : cf. Mirsky, L., *An Introduction to Linear Algebra*(Oxford), p. 388 Theorem 12.6.5.

(2.65) : We rewrite (2.65) in the more convenient form:

$$3\sin^2\theta \geq \theta^2(1 + 2\cos\theta) \quad \text{for} \quad 0 \leq \theta \leq \frac{\pi}{2} \tag{2.100}$$

Now, from the series expansion of $\sin\theta$ and $\cos\theta$, we have

$$\sin\theta \geq \theta - \frac{1}{6}\theta^3, \cos\theta \leq 1 - \frac{1}{2}\theta^2 + \frac{1}{24}\theta^4 \text{ for } 0 \leq \theta \leq \frac{\pi}{2}, \tag{2.101}$$

since $(m!)^{-1}\theta^m \geq ((m+2)!)^{-1}\theta^{m+2}$, if $m \geq 1$ and $\theta^2 \leq 6$. This gives (2.100) on verifying

$$3(1 - \frac{\theta^2}{6})^2 = 1 + 2(1 - \frac{\theta^2}{2} + \frac{\theta^4}{24}). \tag{2.102}$$

(2.76): In connection with this very effective form of the large sieve we have in Montgomery and Vaughan [2] the remark that it may be that the constant $\frac{3}{2}$ in (2.77), and consequently also in (2.76), can be replaced by 1. Actually, their work contains the constant $\frac{\sqrt{12+\sqrt{78}}}{\pi}$(= $1.45282\cdots$) (and some slight improvements) instead of the aforementioned constant $\frac{3}{2}$ (cf. notes of Chapter 8 under (8.36)). For a previous result of this type, see Montgomery [5] (Therefore 4.1). A first weighted form of this sieve occurs in Davenport and Halberstam [1] (cf. also Davenport [1] Liu [1] and Montgomery [5]. (Theorem 2.4).

47 By comparing the results of this chapter with those of the previous forms (0.54) gives in Chapter 0, it is natural to ask the following question (Erdös [3]) Consider (in the notation of Ch. 0)

$$\sum_{p\leq Q}\sum_{l=1}^{p-1}|T(\frac{l}{p})|^2 \leq C_1(N,Q)S, \tag{2.103}$$

where now only a 'negligible' proposition of terms remain on the left-hand side. Then are there better results for $C_1(N,Q)$ than those for

$C(N,Q)$, (cf. (2.4)) or, more specifically, can one expect a gain of a factor $\log Q$ here? Erdös [3] (cf. Erdös and Rényi [1]) proved that if $Q \leq \sqrt{N}$ this is true for *almost all* sets γ. Wolke [2] (cf. Wolke [1]) has proved a slightly weaker estimate which holds also for (more general) $U(x)$, (cf. (2.9)), and for *all* sets γ, but under the severe condition $N \leq Q(\log Q)^{\delta}$ for some $\delta > 0$. On the other hand Erdös [3] (cf. Erdös and Rényi [1]) has shown that $C_1(N,Q)$ is of the same order of magnitude as $C(N,Q)$, if Q is of a higher order than $\sqrt{N \log N}$. For further literature in this connection, see Elliott [5], [7]. The result is generally speaking, that (except under special circumstances) we cannot have a better estimate in (2.103) than that for (2.4).

Another attempt, at sharpening the large sieve, is due to Burgess ([1]). He proved that for any set $\mathscr{Q} \cup \mathbb{N}$

$$\sum_{\substack{q\leq Q\\ q\in\mathscr{Q}}} \sum_{\ell=1}^{q}{}' |U(\frac{\ell}{q})| \ll (Q|\mathscr{Q}|)(N + (Q|\mathscr{Q}|))(\sum_{M<n\leq M+N} |a_n|^2), \tag{2.104}$$

from which saving is made when $\mathscr{Q}$ is a sparse set. (Note also that here U occurs to the first power on the left-hand side.)

In the other direction one may ask for general lower bounds, for instance in (2.103). In view of (0.53) such results would mean that 'general sequences cannot be too well-distributed in almost all arithmetic progressions'. The first result in this context is given by Roth [1], **48**
For further literature concerning this question, we refer to Roth [3], [5], Choi [1], Montgomery [5] (Chapter 5) and Huxley [5]. Wolke [10] (cf. Wolke [9]) has stated the explicit lower bound (in the notation (1.29))

$$\sum_{\substack{q\leq Q\\ q\in\mathscr{Q}}} \sum_{\ell=1}^{q}{}' |U(\frac{\ell}{q})|^2 \geq \sum_{\substack{q\leq Q\\ q\in\mathscr{Q}}} \frac{1}{\varphi(q)} \Big| \sum_{M<n\leq M+N} a_n c_q(n) \Big|^2 \text{ for any set } \mathscr{Q} \subset \mathbb{N}. \tag{2.105}$$

However, (2.105) should rather be considered as a reduction of the expression on the right-hand side to the large sieve (cf. (3.28)).

Chapter 3

The Large Sieve for Character Sums

THERE IS another version of the large sieve which concerns with the averaging of character sums (cf. (3.2) below). In this chapter we give three such results which are readily obtained from Theorem 2.5. We prove first **49**

Theorem 3.1. *Let $Q \in \mathbb{N}$. For any character $\chi \bmod q$ and for any complex numbers a_n, satisfying*

$$a_n = 0, \text{ unless } (n, q) = 1 \forall q \leq Q, \tag{3.1}$$

write

$$X(\chi) := \sum_{M<n\leq M+N} a_n \chi(n). \tag{3.2}$$

Then, we have

$$\sum_{q\leq Q} \frac{1}{\varphi(q)} \sum_{\chi \bmod q} |\tau(\chi)|^2 |X(\chi)|^2 \leq (N + Q^2) \sum_{M<n\leq M+N} |a_n|^2, \tag{3.3}$$

$$\sum_{q\leq Q} (N + \frac{3}{2} qQ)^{-1} \frac{1}{\varphi(q)} \sum_{\chi \bmod q} |\tau(\chi)|^2 |X(\chi)|^2 \leq \sum_{M<n\leq M+N} |a_n|^2, \tag{3.4}$$

and

$$\sum_{q\le Q} \log \frac{Q}{q} \sideset{}{^*}\sum_{\chi \bmod q} |X(\chi)|^2 \le (N+Q^2) \sum_{M<n\le M+N} |a_n|^2, \tag{3.5}$$

where $\tau(\chi)$ is defined by (1.60) and

$$\sideset{}{^*}\sum_{\chi \bmod q} := \sum_{\substack{\chi \bmod q \\ \chi \text{ primitive}}} \tag{3.6}$$

Remark. The condition (3.1) is not usually a severe restriction, since in applications either this fulfilled or the extra terms arising in the other case are separately estimated to be small.

Proof. First of all, for $(n,q)=1$, it follows from (1.61)

$$\varphi(\bar{\chi}) a_n \chi(n) = \sum_{\ell=1}^{q} \bar{\chi}(\ell) a_n e(n\frac{\ell}{q}), \tag{3.7}$$

50 which holds, by (3.1), also for $(n,q)>1$. Now (3.7) gives

$$\tau(\bar{\chi}) X(\chi) = \sum_{\ell=1}^{q} \bar{\chi}(\ell) U(\frac{\ell}{q}), \tag{3.8}$$

where $U(x)$ is defined through (2.9). Multiplying each side of (3.8) by its complex conjugate, summing over all character $\chi \bmod q$ and using (1.57) with $u_\ell = (\frac{\ell}{q})$. we get the identity

$$\frac{1}{\varphi(q)} \sum_{\chi \bmod q} |\tau(\chi)|^2 |X(\chi)|^2 = \sideset{}{'}\sum_{\ell=1}^{q} |U(\frac{\ell}{q})|^2, \tag{3.9}$$

(cf. Bombieri and Davenport [2]) and further use of this in Theorem 2.5 yields (3.3) and (3.4).

Next, for a character $\chi \bmod q$ let f be its conductor, and let χ be induced by χ^* mod f. Then by the assumption (3.1) (cf. (1.59)),

$$X(\chi) = X(\chi^*), \tag{3.10}$$

and using Lemma 1.1 as well as

$$\frac{f}{\varphi(f)}(\sum_{\substack{r\le x\\(r,f)=1}})\frac{\mu^2(r)}{\varphi(r)}) > \log x \quad \text{for} \quad x > 0, \tag{3.11}$$

we obtain, in our notation (3.6),

$$\begin{cases}\sum_{q\le Q}\frac{1}{\varphi(q)}\sum_{\chi \bmod q}|\tau(\chi)|^2|X(\chi)|^2 = \sum_{f\le Q}\frac{f}{\varphi(f)}\sum_{\substack{r\le Q/f\\(r,f)=1}}\frac{\mu^2(r)}{\varphi(r)}\sum^*_{\chi \bmod f}\\ |X(\chi)|^2 \ge \sum_{f\le Q}\log\frac{Q}{f}\sum^*_{\chi \bmod f}|X(\chi)|^2.\end{cases} \tag{3.12}$$

Thus (3.5) is a consequence of (3.3).

Regarding the quality of the results in Theorem 3.1 we note that, in view of the identity (3.9), the estimates (3.3) and (3.4) are capable of improvements only along with sharpening of Theorem 2.5. On the other hand, the statement (3.5) leaves a gap (even through the inequality (3.11) is capable of an asymptotic formulation. □

In the case when only primitive characters occur in both (3.3) and (3.4) our condition (3.1) can be removed to prove the next **51**

Theorem 3.2. *For any character χ mod f, $r \in \mathbb{N}$ and for arbitrary complex numbers a_n, set*

$$X_r(\chi) := \sum_{M<N\le M+N} a_n\chi(n)c_r(n) \tag{3.13}$$

where $c_r(n)$ is given by (1.29)*. Then, we have*

$$\sum_{\substack{rf\le Q\\(r,f)=1}}\frac{f}{\varphi rf}\sum^*_{\chi\, mod\, f}|X_r(\chi)|^2 \le (N+Q^2)\sum_{M<n\le M+N}|a_n|^2, \tag{3.14}$$

and

$$\sum_{\substack{rf\le Q\\(r,f)=1}}(\frac{N}{f}+\frac{3}{2}rQ)^{-1}\frac{1}{\varphi(rf)}\sum^*_{\chi\, mod\, f}|X_r(\chi)|^2 \le \sum_{M<n\le M+N}|a_n|^2. \tag{3.15}$$

Remark. Observe that, under the condition (3.1), we have $c_r(n)$ appearing in the non-zero terms of (3.13) as $\mu(r)$ and so, (3.14) also leads to (3.5) in view of (3.12).

Proof. We have for any $q \in \mathbb{N}$, by (1.57),

$$\frac{1}{\varphi(q)} \sum_{\chi \bmod q} \Big| \sum_{\ell=1}^{q} \overline{\chi}(\ell) U(\frac{\ell}{q}) \Big|^2 = \sum_{\ell=1}^{q}{}' \, |U(\frac{\ell}{q})|^2. \tag{3.16}$$

Now if $\chi \bmod q$ is induced by $\chi^* \bmod f$ (f conductor of χ). we have, on using (1.68) and (1.59),

$$q = rf, \chi(\ell) = \chi^*(\ell) \text{ for } (\ell, q) = 1. \tag{3.17}$$

Therefore, summing (3.16) over $q \le Q$ gives

$$\sum_{q \le Q} \sum_{\ell=1}^{q}{}' \, |U(\frac{\ell}{q})|^2 \ge \sum_{\substack{rf \le Q \\ (r,f)=1}} \frac{1}{\varphi(rf)} \sideset{}{^*}\sum_{\chi \bmod f} | \sum_{\ell=1}^{q}{}' \, \overline{\chi}(\ell) U(\frac{\ell}{q})|^2. \tag{3.18}$$

For any primitive character $\chi \bmod f$, $q = rf$, $(r, f) = 1$ and any ℓ, $(\ell, q) = 1$, on writing

$$\ell = \lambda r + \mu f, (\lambda, f) = 1, (\mu, r) = 1, \tag{3.19}$$

52 we will have

$$\begin{cases} \sum_{\ell=1}^{q}{}' \, \overline{\chi}(\ell) U(\frac{\ell}{q}) = \sum_{M<n\le M+N} a_n \sum_{\lambda=1}^{f}{}' \, \overline{\chi}(\lambda r) \sum_{\mu=1}^{r}{}' \, e(n(\frac{\lambda}{f} + \frac{\mu}{r})) = \\ = \overline{\chi}(r) \sum_{M<n\le M+N} a_n \sum_{\lambda=1}^{f}{}' \, \overline{\chi}(\lambda) e(n\frac{\lambda}{f}) \sum_{\mu=1}^{f}{}' \, e(n\frac{\mu}{r}) = \overline{\chi}(r)\tau(\overline{\chi}) X_r(\chi), \end{cases} \tag{3.20}$$

because of (1.62) and (1.29). Since $(r, f) = 1$, we have (by (1.42)) that $|\overline{\chi}(r)| = 1$ and further by Lemma 1.1 that $|\tau(\overline{\chi})|^2 = f$. Thus (3.20) and (3.18), on using (2.86), prove the part (3.14). The proof of (3.15) is the same in the before summing over q in (3.16) we need multiply by the factor $(N + \frac{3}{2}qQ)^{-1}$ and at the and employ (2.87) instead of (2.86).

Since obviously

$$c_1(n) = 1, \tag{3.21}$$

we obtain, by retaining only the parts with $r = 1$ in the expressions occurring on the left-hand sides of Theorem 3.2, as a particular case □

Theorem 3.3. *For any character $\chi \bmod q$ and for any complex numbers a_n, define*

$$X(\chi) := \sum_{M<n\leq M+N} a_n\chi(n). \tag{3.22}$$

Then, we have

$$\sum_{q\leq Q} \frac{q}{\varphi(q)} \sideset{}{^*}\sum_{x \bmod q} |X(\chi)|^2 \leq (N + Q^2) \sum_{M<n\leq M+N} |a_n|^2 \tag{3.23}$$

and

$$\sum_{q\leq Q} (\frac{N}{q} + \frac{3}{2}Q)^{-1} \frac{1}{\varphi}(q) \sideset{}{^*}\sum_{\chi \bmod q} |X(\chi)|^2 \leq \sum_{M<n\leq M+N} |a_n|^2. \tag{3.24}$$

NOTES 53

The version of the large sieve discussed in this chapter occurs for the first time in Bombieri [1] (see, however, Rényi [2]; cf (0.32)). Simplifications of the proof and improvement of the quality of the result were made by Davenport and Halberstam [1], who also obtained there the first weighted form. However, as has been mentioned earlier and can also be seen from (3.9) and (3.18), the results of this chapter are more or less direct consequences of those in Chapter 2.

A first result with conditions (3.1) was given by Bombieri and Davenport [2] (and for (3.5), see Bombieri [6] (Théorème 8)) where they also prove a generalization of (3.3) which, via Theorem 2.5 under the assumptions of Theorem 3.1, becomes

$$\sum_{\substack{q\leq Q\\(q,k)=1}} \frac{1}{\varphi(q)} \sum_{\chi \bmod q} |\tau(\chi)|^2 | \sum_{\substack{M<n\leq M+N\\ n\equiv\ell (\bmod k)}} a_n\chi(n)|^2 \leq (\frac{N}{k}+1+Q^2) \sum_{M<n\leq M+N} |a_n|^2 \tag{3.25}$$

where $k \in \mathbb{N}$ and $(\ell, k) = 1$.

(3.11): By Chapter 1, **1.** We have that, for $x > 0$,

$$\begin{cases} \frac{f}{\varphi(f)} \sum_{\substack{r \le x \\ (r,f)=1}} \frac{\mu^2(r)}{\varphi(r)} = \prod_{p|f}(1 - \frac{1}{p})^{-1} \sum_{\substack{r \le x \\ (r,f)=1}} \frac{\mu^2(r)}{r} \prod_{p|r}(1 - \frac{1}{p})^{-1} = \\ = \prod_{p|f}(1 + \sum_{\nu=1}^{\infty} p^{-\nu}) \sum_{\substack{r \le x \\ (r,f)=1}} \frac{\mu^2(r)}{r} \prod_{p|r}(1 + \sum_{\nu=1}^{\infty} p^{-\nu}) \ge \sum_{n \le x} \frac{1}{n} > \log x, \end{cases} \tag{3.26}$$

on using, for $x \ge 1$, if $N \le x < N + 1$,

$$\sum_{n \le x} \frac{1}{n} \ge \sum_{n \le N} \int_n^{n+1} \frac{dt}{t} = \log(N + 1) > \log x \tag{3.27}$$

(cf. van Lint and Richert [1]).

54 Clearly, analogous to the derivation of (3.5) from (3.3) one can obtain a corresponding result from (3.4). In the special case $a_n = \Lambda(n)$, this has been done by Montgomery and Vaughan [2] (cf. (6.29)).

The extension to sums involving Ramanujan's sum $c_r(n)$, namely, Theorem 3.2 which contains Theorem 3.3, is due to A. Selberg ([6], cf. Bombieri [6] (Théorème 7A)).

We obtained (3.23) from (3.14) by keeping only the part with $r = 1$. On the otherhand, by taking the part corresponding to $f = 1$, we get

$$\sum_{q \le Q} \frac{1}{\varphi(q)} \Big| \sum_{M < n \le M+N} a_n c_q(n) \Big|^2 \le \sum_{q \le Q} \sum_{\ell - 1}^{q}{}' \, \Big| U\Big(\frac{\ell}{q}\Big) \Big|^2 \tag{3.28}$$

which should be compared with Wolke's result ((2.105)).

The main importance of Selberg's generalization (3.14) is due to its application in proving density theorems for Dirichlet's L-functions (cf. Chapter 6, 2.). There the strongest known results (cf. Montgomery [8], Motohashi [11] and Jutila [12] are based on the following strinking identity again due to A. Selberg:

$$\begin{cases} L(s, \chi) M(s, \chi, \psi_r) = \sum_{n=1}^{\infty} \frac{\chi(n)}{n^s} \psi_r(n) \sum_{d|n} \xi_d, (\chi := \chi \bmod f) \\ \text{for any set of } \xi_d \in \mathbb{C}, \xi_d = 0(1). \end{cases} \tag{3.29}$$

where

$$M(s,\chi,\psi_r) := \sum_{n=1}^{\infty} \frac{\chi(n)\xi_n\psi_r(n)}{n^s} \prod_{p|\frac{r}{(r,n)}} (1 - \frac{\chi(p)}{p^{s-1}}), \tag{3.30}$$

$$\psi_r(n) := \mu((r,n))\varphi((r,n)), \tag{3.31}$$

and

$$\mu(r) \neq 0. \tag{3.32}$$

Actually, (3.14) is employed in the weaker from

$$\sum_{\substack{rf\leq Q\\(r,f)=1}} \frac{\mu^2(r)f}{\varphi(rf)} \sum_{\chi \bmod f}^{*} \Big|\sum_{M<n\leq M+N} a_n\chi(n)\psi_r(n)\Big|^2 \leq (N+Q^2) \sum_{M<n\leq M+N} \Big|a_n\Big|^2. \tag{3.33}$$

55

In fact, the expression on the left-hand side of (3.33) is the part of the sum in (3.14) corresponding to squarefree r's, since (under (3.32)) one has one has readily, by (1.32),

$$c_r(n) = \sum_{d|(r,n)} \mu(\frac{r}{(r,n)}\frac{(r,n)}{d})d = \mu(\frac{r}{(r,n)})\varphi((r,n)) = \mu(r)\psi_r(n)\ (\mu(r)\neq 0). \tag{3.34}$$

Motohashi [14] has in turn shown that (3.33) can be generalized to give the following estimate: Let $\chi_j \bmod f_j,\ f_j \leq F(j = 1,\dots,J)$, be distinct primitive characters. Then

$$\sum_{r\leq R} \sum_{\substack{j\leq J\\(f_j,r)=1}} \frac{\mu^2(r)f_j}{\varphi(rf_j)} |\sum_{M<n\leq M+N} a_n\chi_j(n)\psi_r(n)|^2 \leq (N + O(JFR^2\log(FR))) \sum_{M<n\leq M+n} |a_n|^2 \tag{3.35}$$

which, for sufficiently small J, improves upon (3.33).

Slightly more general forms than those of Theorem 3.1 and 3.3 (namely, without the restriction (3.1) or removing the limitation to summing over only primitive characters) are possible but at the expense of the quality of the estimates (cf. Bombieri [1] and Davenport and Halberstam [1]).

For an estimate of averages involving real characters only, see jutila [6].

Finally, there are result concerning averages of characters sums, which are useful when combined with large sieve estimates in some applications. They can be obtained without employing results of Chapter 2 (cf. Montgomery [5] (Theorems 6.2 and 6.3)). We mention, as an example, (Montgomery [5] (Theorem 6.2)):

$$\sum_{\chi \bmod q} |X(\chi)|^2 \leq \varphi(q)(1 + \left[\frac{N-1}{q}\right]) \sum_{\substack{M<n\leq M+N \\ (n,q)=1}} |a_n|^2. \tag{3.36}$$

56 For a proof, split $X(\chi)$ into $1+\left[\frac{N-1}{q}\right]$ parts of length q (introducing additions a_n's = 0, if necessary). For each part $X_1(\chi)$, say, it results from (1.57), with obvious appropriate choices for u_ℓ, that

$$\sum_{\chi \bmod q} |X_i(\chi)|^2 = \varphi(q) \sum_{\substack{n\in\mathscr{I}_i \\ (n,q)=1}} |a_n|^2, \tag{3.37}$$

where $\mathscr{I}_i$ denotes the range of n in $X_i(\chi)$. Now (3.36) following on using Minkowski's and Cauchy's inequalities. Likewise, but with a more complicated yet still elementary, reasoning (Montgomery [5] (Theorem 6.3)) one obtains

$$\sum_{\chi \bmod q} |X(\chi^*)|^2 \leq q(1 + \left[\frac{N-1}{q}\right]) \sum_{\substack{M<n\leq M+N \\ (n,q)=1}} |a_n|^2 \tag{3.38}$$

where for each $\chi \bmod q$, χ^* denotes the primitive character which induces χ.

Chapter 4

The Large Sieve for Dirichlet Polynomials and Dirichlet Series

STILL ANOTHER from of the large sieve,as has been noted Davenport, can be obtained with respect to Dirichlet polynomials **57**

$$\sum_{n=1}^{N} a_n n^{-s}, s = \sigma + \text{ it } (\sigma, t \text{ reals }), a_n \in \mathbb{C}, \tag{4.1}$$

as an application of Lemma 2.1.

Theorem 4.1. *Let T_0, $T(>0)$ and t_r be real numbers satisfying*

$$T_0 = t_0 < t_1 < \cdots < t_R < t_{R+1} = T_0 + T, \tag{4.2}$$

and put

$$\delta := \min_{0 \le r \le R} t_{r+1} - t_r \tag{4.3}$$

Then, we have

$$\sum_{r=1}^{R} \left| \sum_{n=1}^{N} a_n n^{-it_r} \right|^2 \le (T + 4N \log N)(\log N + \delta^{-1}) \sum_{n=1}^{N} |a_n|^2, \forall a_n \in \mathbb{C}. \tag{4.4}$$

Proof. By (4.2) and (4.3), we have $(R+1)\delta \leq T$ and so we can suppose that $N \geq 2$. Also, by (4.3), the intervals $\left[t_r - \frac{\delta}{2}, t_r + \frac{\delta}{2}, \right]$, $1 \leq r \leq R$, do not overlap. Therefore, taking in Lemma 2.1

$$f(u) = \sum_{n=1}^{N} a_n n^{-iu}, u = t_r (1 \leq r \leq R) \tag{4.5}$$

and proceeding as in (2.28), we obtain

$$\sum_{r=1}^{R} |f(t_r)|^2 \leq (\int_{T_0}^{T_0+T} |f(t)|^2 dt)^{\frac{1}{2}} \int_{T_0}^{T_0+T} |f'(t)|^2 dt)^{\frac{1}{2}} + \delta^{-1} \int_{T_0}^{T_0+T} |f(t)|^2 dt. \tag{4.6}$$

Now

$$\int_{T_0}^{T_0+T} |f(t)|^2 dt = T \sum_{N=1}^{N} |a_n|^2 + \sum_{\substack{m,n=1 \\ m \neq n}} a_m \bar{a}_n \int_{T_0}^{T_0+T} (\frac{n}{m})^{it} dt. \tag{4.7}$$

58 The second sum, on noting that

$$\left|\log \frac{n}{m}\right| \geq \frac{|n-m|}{\max(n,m)} \tag{4.8}$$

and using the arithmetic-geometric inequality, may be estimated by

$$2 \sum_{\substack{m,n=1 \\ m \neq n}}^{N} \frac{|a_m {}^a n|}{|\log \frac{n}{m}|} \leq N \sum_{\substack{m,n=1 \\ m \neq n}}^{N} \frac{1}{|n-m|} (|a_m|^2 + |a_n|^2). \tag{4.9}$$

Due to symmetry in m and n (of these expressions), it is enough to consider the factor of $|a_n|^2$ on the right-hand side. This equals

$$\sum_{1 \leq m < n} \frac{1}{n-m} + \sum_{n < m \leq N} \frac{1}{n-m} = \sum_{1 \leq k < n} \frac{1}{k} + \sum_{1 \leq k \leq N-n} \frac{1}{k} \leq 2 \log N. \tag{4.10}$$

Therefore, the expression (4.9) can further be estimated by

$$4N \log N \sum_{n=1}^{N} |a_n|^2. \tag{4.11}$$

Using this in (4.7), we get

$$\int\limits_{T_0}^{T_0+T} |f(t)|^2 dt \leq (T + 4N \log N) \sum_{n=1}^{N} |a_n|^2. \tag{4.12}$$

Clearly, for f' instead of f (4.12) holds with an extra factor of $\log^2 N$ on the right-hand side; in otherwords, we have

$$\int\limits_{T_0}^{T_0+T} |f(t)|^2 dt \leq (T + 4N \log N) \sum_{n=1}^{N} |a_n|^2. \tag{4.13}$$

□

Now (4.4) follows from (4.6), (4.12) and (4.13).

For applications in the direction of the classical mean-value theorems for the Dirichlet series, the powerful tools of Montgomery and Vaughan [1], namely (2.49) and (2.77), can be remodelled to the following Lemma.

Lemma 4.1. *Let $\lambda_1, \ldots, \lambda_R$ be distinct real number and set* **59**

$$\Delta : \min_{r \neq s} |\lambda_r - \lambda_s| \tag{4.14}$$

and

$$\Delta_r : \min_{\substack{s \\ s \neq r}} |\lambda_r - \lambda_s|. \tag{4.15}$$

Then

$$\Bigg| \sum_{r=1}^{R} \sum_{\substack{s=1 \\ r \neq s}}^{R} \frac{u_r \bar{u}_s}{\lambda_r - \lambda_s} \Bigg| \leq \pi \Delta^{-1} \sum_{r=1}^{R} |u_r|^2, \forall u_r \in \mathbb{C}. \tag{4.16}$$

and

$$\Bigg| \sum_{r=1}^{R} \sum_{\substack{s=1 \\ r \neq s}}^{R} \frac{u_r \bar{u}_s}{\lambda_r - \lambda_s} \Bigg| \leq \frac{3}{2} \pi \sum_{r=1}^{R} |u_r|^2, \Delta_r^{-1},\ \forall u_r \in \mathbb{C}. \tag{4.17}$$

Proof. Let $\epsilon > 0$ denote a small number to be suitably restricted below. Put

$$x_r = \epsilon\lambda_r (1 \le r \le R). \tag{4.18}$$

Then, we have in the notation of (2.10) and (4.14), for all sufficiently small ϵ,

$$\delta = \min_{r \neq s} \|\epsilon(\lambda_r - \lambda_s)\| = \epsilon \min_{r \neq s} |\epsilon(\lambda_r - \lambda_s)| = \epsilon\Delta, \tag{4.19}$$

and, similarly in the notation of (2.75) and (4.15).

$$\delta_r = \epsilon\Delta_r \quad (1 \le r \le R). \tag{4.20}$$

Multiplying both sides of (2.49) and (2.77) by $\pi\epsilon$ and further using (4.18), (4.19) and (4.20), we obtain Lemma 4.1 as a consequence of

$$\lim_{\epsilon \to +0} \frac{\pi\epsilon}{\sin(\pi(x_r - x_s))} = \frac{1}{(\lambda_r - \lambda_s)}. \tag{4.21}$$

□

We obtain almost immediately from Lemma 4.1 the following

Theorem 4.2. *Under the assumptions and notation of Lemma 4.1, we have*

$$\int_{-T}^{T} \Big| \sum_{r=1}^{R} a_r e^{i\lambda_r t} \Big|^2 dt = 2(T + \theta_1 \pi \Delta^{-1}) \sum_{r=1}^{R} |a_r|^2, \quad \forall a_r \in \mathbb{C}, \tag{4.22}$$

60 *and*

$$\int_{-T}^{T} \Big| \sum_{r=1}^{R} a_r e^{i\lambda_r t} \Big|^2 dt = \sum_{r=1}^{R} |a_r|^2 \; (2T + 3\theta_2 \pi \Delta_r^{-1}), \quad \forall a_r \in \mathbb{C}, \tag{4.23}$$

where

$$|\theta_j| \le 1, \; j = 1, 2. \tag{4.24}$$

Proof. The integral of the theorem is

$$\sum_{r,s=1}^{R} a_r \overline{a}_s \int_{-T}^{T} e^{i(\lambda_r - \lambda_s)t}{}_{dt=2T} \sum_{r=1}^{R} |a_r|^2 + \sum_{r=1}^{r} \sum_{\substack{s=1 \\ r \neq s}}^{R} a_r \overline{a}_s \frac{e^{i(\lambda_r \lambda_s)T} - e^{-i(\lambda_r - \lambda_s)T}}{i(\lambda_r - \lambda_s)}. \tag{4.25}$$

Now application of (4.16) and (4.17) with the choices

$$u_r = a_r e^{\pm i\lambda_r T} \tag{4.26}$$

to the double sum in (4.25) yield (4.22) and (4.23) respectively.

The most interesting use of Theorem 4.2 is when applied, in its from (4.23), to Dirichlet series via Dirichlet polynomials. Taking

$$\lambda_r = -\log r, \tag{4.27}$$

we find, by (4.8),

$$\Delta_r^{-1} \leq r + 1, \tag{4.28}$$

and so get, by (4.23),

$$\int_{-T}^{T} \left| \sum_{r=1}^{R} a_r r^{-it} \right|^2 dt = \sum_{r=1}^{R} |a_r|^2 (2T + 30_3 \pi (r+1)), \ |\theta_3| \leq 1, \quad \forall R \in \mathbb{N}. \tag{4.29}$$

Further, if we impose the condition

$$\sum_{r=1}^{\infty} r |a_r|^2 < \infty, \tag{4.30}$$

we can conclude from (4.29) that the Dirichlet polynomials

$$\sum_{r=1}^{R} a_r r^{-it} \tag{4.31}$$

converge in the mean to the Dirichlet series **61**

$$\sum_{r=1}^{\infty} a_r r^{-it} \in L_2(-T, T). \tag{4.32}$$

□

Thus we derive from (4.23) the following important result.

Theorem 4.3. *For $a_n \in \mathbb{C}$, suppose that*

$$\sum_{r=1}^{\infty} n|a_n|^2 < \infty. \tag{4.33}$$

Then

$$\int_{-T}^{T} \left| \sum_{n=1}^{\infty} a_n n^{-it} \right|^2 dt = \sum_{n=1}^{\infty} |a_n|^2 \; (2T + 3\theta\pi(n+1)), \tag{4.34}$$

where

$$|\theta| \leq 1. \tag{4.35}$$

NOTES

Theorem 4.1, due to Davenport, was published by Montgomery [2].

Lemma 4.1, Theorems 4.2 and 4.3 are due to Montgomery and Vaughan [1] (Theorem 2 and Corollaries 2 and 3).

(4.1): Clearly, it is no restriction to have the result of this chapter with $\sigma = 0$. Theorems 4.2 and 4.3: Obviously, it is possible to obtain with the help of Gallagher's Lemma 2.1 (cf. (4.6)) results, corresponding to Theorems 4.2 and 4.3, in with the integral is replaced by a sum over a set of well-spaced points.

(4.32): For the reasoning leading to (4.32), cf, Titchmarsh, E.C., *The Theory of Functions* (Oxford), pp. 386–387.

For a discussion of $\rho(N, T)$ in the general inequality

$$\int_{-T}^{T} | \sum_{n=1}^{N} a_n n^{-it} |^2 \; dt \leq \rho(N, T) \sum_{n=1}^{N} | \; a_n \; |^2, \tag{4.36}$$

see Elliott [7].

Chapter 5

The Hybrid Sieve

FOR SOME very important applications to number theory, via Dirichlet series, estimations of averages of the type 62

$$\sum_{q \leq Q} (\cdots) \sum_{X \bmod q}^{*} \int_{-T}^{T}, \tag{5.1}$$

which is a combined version of the forms of the large sieve considered in Chapters 3 and 4, are of much use. Prior to the innovation of an ingenious method of Halász [1] there was no method of dealing with this question without carrying out at least one of the operations $\sum^*$ or $\int$ in a trivial fashion. Methods, for the purpose of this hybrid sieve, were developed independently by Montgomery [2], combining ideas of Halász with the large sieve, and by Jutila [1], who used a method of Rodosskij with large sieve. Subsequently, a common basis for both of these was provided by Gallagher [4] through the introduction of new technical devices (see Lemmas 5.1 and 5.2 below).

We start with

Lemma 5.1. *Let*

$$D(t) : \sum_{\nu} c(\nu) e(\nu t), \tag{5.2}$$

where ν runs through a countable set of real numbers and the coefficients

$c(\nu)(\in \mathbb{C})$ *are subjected to the condition*

$$\sum_{\nu} | c(\nu) |< \infty. \tag{5.3}$$

Let δ and T be positive real numbers satisfying

$$\delta T \leq \frac{1}{2\pi}. \tag{5.4}$$

Then, for some absolute constant c_0, holds

$$\int_{-T}^{T} |D(t)|^2 dt \leq c_0 \int_{-\infty}^{\infty} |C_\delta(y)|^2 dy, \tag{5.5}$$

63 where

$$C_\delta(y) : \delta^{-1} \sum_{|y-\nu|<\frac{\delta}{2}} c(\nu). \tag{5.6}$$

Proof. For a proof of (5.5) we use two results from the theory of Fourier transforms. Introduce

$$F_\delta(y) : \begin{cases} \delta^{-1} & if \mid y \mid < \delta/2, \\ 0 & otherwise, \end{cases} (y \in \mathbb{R}) \tag{5.7}$$

so that

$$C_\delta(y) = \sum_{\nu} c(\nu) F_\delta(y - \nu). \tag{5.8}$$

In view of (5.3), $C_\delta(y)$ is a bounded integrable function and hence it belongs to $L_2(-\infty, \infty)$. Therefore, by Plancherel's theorem, C_δ has a Fourier transform $\hat{C}_\delta$ and further, by Parseval's formula, we have

$$\int_{-\infty}^{\infty} | C_\delta(y) |^2 dy = \int_{-\infty}^{\infty} | \hat{C}_\delta(t) |^2 dt. \tag{5.9}$$

Now one has

$$\begin{cases} \hat{C}_\delta(t) & = \int\limits_{-\infty}^{\infty} C_\delta(y) e(yt) dy = \sum_\nu c(\nu) \int\limits_{-\infty}^{\infty} F_\delta(y-\nu) e(yt) dy = \\ & = \sum_\gamma c(\nu) e(\nu t) \int\limits_{-\infty}^{\infty} F_\delta(x) e(xt) dx = D(t) \hat{F}_\delta(t) \end{cases} \tag{5.10}$$

say, on using (5.8) and (5.2). Also, by (5.7),

$$\hat{F}_\delta(t) = \delta^{-1} \int_{-\delta/2}^{\delta/2} e(xt)dx = \frac{\sin(\pi\delta t)}{\pi\delta t}. \tag{5.11}$$

Thus (5.9) yields

$$\int_{-\infty}^{\infty} \mid C_\delta(y) \mid^2 dy = \int_{-\infty}^{\infty} \mid D(t)\hat{F}_\delta(t) \mid^2 dt \le \int_{-T}^{T} \mid D(t)\hat{F}_\delta(t) \mid^2 dt. \tag{5.12}$$

because of (5.10). For $|t| \le T$, we use (5.11) to note (cf. (5.4)) **64**

$$|\hat{F}_\delta(t)| \ge \frac{\sin(\pi\delta T)}{\pi\delta T} \ge \frac{1}{\sqrt{c_0}}, \tag{5.13}$$

say, so that (5.5) follows from (5.12). □

Following Gallagher we use Lemma 5.1 to prove (Gallagher [4], Theorem 1).

Lemma 5.2. *For $a_n \in \mathbb{C}$, let*

$$\sum_{n=1}^{\infty} |a_n < \infty. \tag{5.14}$$

Then, for $T \ge 1$,

$$\int_{-T}^{T} \Big| \sum_{n=1}^{\infty} a_n n^{-it} \Big|^2 dt \le c_1 T^2 \int_0^{\infty} \Big| \sum_{x<n<xe^{1/T}} a_n \Big|^2 \frac{dx}{x} \tag{5.15}$$

holds with some absolute constant c_1.

Proof. In Lemma 5.1 we choose

$$\nu = -\frac{1}{2\pi} \log n, \; c(\nu) = a_n, \; n \in \mathbb{N} \tag{5.16}$$

and note that (5.3) is satisfied because of (5.14). Further, we put

$$\delta = \frac{1}{2\pi T}, y = \frac{1}{2\pi}(\log x + \frac{1}{2T}) \quad (x > 0), \tag{5.17}$$

so that (5.4) is fulfilled and the condition of summation in (5.6) reads

$$-\frac{\delta}{2} < y - \nu < \frac{\delta}{2} \Leftrightarrow \log x < \log n < \frac{1}{T} + \log x. \tag{5.18}$$

Therefore (5.5) yields (with the above choice)

$$\int\limits_{-T}^{T} \Big|\sum_{n=1}^{\infty} a_n n^{-it}\Big|^2 dt \leq c_0 \int\limits_{0}^{\infty} 2\pi T^2 \Big|\sum_{x<n<xe^{1/T}} a_n\Big|^2 \frac{dx}{x}. \tag{5.19}$$

This completes the proof of the lemma. □

Now we shall give two applications of Lemma 5.2 for the averages of the type (5.1).

65 **Theorem 5.1.** *For $a_n \in \mathbb{C}$, let*

$$\sum_{n=1}^{\infty} |a_n| < \infty. \tag{5.20}$$

Then, for $T \geq 1$,

$$\sum_{q \leq Q} \frac{q}{\varphi(q)} \sideset{}{^*}\sum_{\chi \bmod q} \int\limits_{-T}^{T} \Big|\sum_{n=1}^{\infty} a_n \chi(n) n^{-it}\Big|^2 dt \leq 2c_1 \sum_{n=1}^{\infty} (TQ^2 + n)|a_n|^2 \forall Q \in \mathbb{N} \tag{5.21}$$

holds with the constant c_1 of Lemma 5.2.

Proof. We use Lemma 5.2 with $a_n\chi(n)$ instead of a_n and then apply

$$\sum_{q \leq Q} \frac{q}{\varphi(q)} \sideset{}{^*}\sum_{\chi \bmod q} \tag{5.22}$$

to (the resulting) (5.15). Then, we have for the squared expression on the righthand side of (5.15), in the notation of Theorem 3.3, $M = [x]$ and $N \leq x(e^{1/T} - 1) + 1$ so that using (3.23) it follows that

$$\sum_{q \leq Q} \frac{q}{\varphi(q)} \sum_{\chi \bmod q}^{*} \int_{-T}^{T} \Big| \sum_{n=1}^{\infty} a_n \chi(n) n^{-it} \Big|^2 dt$$

$$\leq c_1 T^2 \int_0^{\infty} (Q^2 + 1 + x(e^{\frac{1}{T}} - 1)x(\sum_{x<n<xe^{1/T}} |a_n|^2) \frac{dx}{x}. \tag{5.23}$$

Herein, the factor of $|a - n|^2$ is

$$\Big\{ c_1 T^2 (Q^2 + 1) \int_{ne^{-1/T}}^{n} \frac{dx}{x} + c_1 T^2 (e^{\frac{1}{T}} - 1) \int_{ne^{-1/T}}^{n} dx$$

$$= c_1 T (Q^2 + 1) + c_1 T^2 (e^{\frac{1}{T}} - 1)(1 - e^{-\frac{1}{T}}) n \leq 2c_1 (TQ^2 + n), \tag{5.24}$$

where we have employed the estimate

$$T^2 (e^{1/T} - 1)(1 - e^{-1/T}) \leq 2 \text{ for } T \geq 1. \tag{5.25}$$

□

Now, putting together (5.23) and (5.24) we obtain (5.21).

Theorem 5.2. *Let $Q \in \mathbb{N}$. For $a_n \in \mathbb{C}$, let*

$$\sum_{n=1}^{\infty} |a_n| < \infty. \tag{5.26}$$

and suppose that

$$a_n = 0 \ \textit{unless} \ (n, q) = 1 \ \textit{for all} \ q \leq Q. \tag{5.27}$$

Then, for $T \geq 1$, 66

$$\sum_{q \leq Q} \log \frac{Q}{q} \sum_{\chi \bmod q}^{*} \int_{-T}^{T} | \sum_{n=1}^{\infty} a_n x(n) n^{-it} |^2 dt \leq 2c_1 \sum_{n=1}^{\infty} (TQ^2 + n) |a_n|^2 \tag{5.28}$$

holds with the constant c_1 of Lemma 5.2.

Proof. We proceed as in the proof of Theorem 5.1 but applying

$$\sum_{q\leq Q}\log\frac{Q}{q}\sum_{\bmod q}^{*} \tag{5.29}$$

instead of (5.22). Now the condition (5.27) permits us to use (3.5) of Theorem 3.1 to obtain the same bound as in (5.23) for the left-hand side of (5.28). So the proof is again completed by (5.24). □

NOTES

It is possible to put Halász's method in an abstract form (cf. Montgomery [5] (Lemma 1.7) and Huxley [7] (p. 115)). Gallagher and Bombieri ([3]) have observed that Bellman's inequality (0.59) contains both the large sieve and the idea of Halász.

All the results of this chapter are due to Gallagher [4]. Clearly, estimates corresponding to the other results of Chapter 3 can be derived in the same way. For general results conforming to the theme of this chapter and also for more sophisticated forms of the hybrid sieve, see Montgomery [5]. Huxley [7], Gallagher [4], Forti and Viola [1] (cf. Bombieri [5] and [6] (§5)), Huxley [9], [11], Jutila [6] and Huxley [12].

(5.4): This condition can be relaxed to $\delta T \leq 1-\epsilon$ for any $\epsilon > 0$ and then the constant c_0 of (5.5) depends, as can be seen from the proof (cf. (5.13)), on ϵ.

(5.5): Since there is not need in applications we do not aim at obtaining the best possible values (for instance, by a different choice of δ)
67 for the constants c_0 and c_1 in this chapter. However, just for a complete form of the proof we obtain some permissible values for these constants. From (5.13) and (5.4) we see that one can take

$$c_0 = (2\sin\frac{1}{2})^{-2} < 1.1. \tag{5.30}$$

Lemma 5.3. *As Gallagher [4] has stated, if the a_n's are irregular, Lemma 5.2 is more precise than Theorem 5.1, since in Lemma 5.2 the coefficients are first smoothed and then squared.*

(5.14): The condition (occurring in Theorems 5.1 and 5.2)

$$\sum_{n=1}^{\infty} |a_n| < \infty \tag{5.31}$$

of this chapter stems from Lemma 5.1. Further, our condition of Theorem 4.3, namely

$$\sum_{n=1}^{\infty} n|a_n|^2 \infty, \tag{5.32}$$

need also be satisfied; for, otherwise, the theorems of this chapter hold trivially true. Also, it may be noted from the examples

$$a_n = \frac{1}{n\log(n+1)} \quad \text{and} \quad a_n = \begin{cases} \frac{1}{m^2} & \text{if } n = 2^m, m \in \mathbb{N}, \\ 0 & \text{otherwise}, \end{cases} \tag{5.33}$$

that the conditions (5.31) and (5.32) are independent.

(5.15): From (5.19) and (5.30) we see that

$$c_1 = 2\pi c_0 < 7. \tag{5.34}$$

(5.25): For a proof of (5.25) if suffices to verify, by differentiation and the use of $c^{1/x} > 1 + \frac{1}{x} (x > 0)$, that the function $x(1 - e^{-1/x})$ is decreasing in $x > 0$ so that we have

$$x^2(e^{1/x} - 1)(1 - e^{-1/x}) = (x(1 - e^{-1/x}))^2 e^{1/x} \leq (e - 1)(1 - e^{-1}) \text{ for } x \geq 1, \tag{5.35}$$

which gives the upper bound in (5.25) because

$$(e - 1)^2 < 2e. \tag{5.36}$$

Chapter 6

Applications of the Large Sieve

THE SIGNIFICANCE of the large sieve is due to its usefulness towards the solution of important problems of number theory. For this purpose the large sieve is employed in two ways; namely, in proving results which have number-theoretic consequences of depth and on the other side, for direct applications to number theory. We shall defer the discussion of the (latter) arithmetical version of the large sieve to the next chapter and confine ourselves here to a brief survey of the (aforementioned) indirect applications. **68**

1. Moments of the Dirichlet's L-series

In this section we mention the applications to the moments of the Dirichlet's L-series. For the historical introduction to this topic we refer to Montgomery [5] (Chapter 10).

Gallagher [1] has shown that the large sieve (in its version of Chapter 3) can be used to prove

$$\sum_{q\leq Q}\ \sideset{}{^*}\sum_{\chi \bmod q} |L(\frac{1}{2}+it,\chi)|^4 \ll Q^2T^2\log^4(QT).|t|\leq T, T\geq 2. \qquad (6.1)$$

Montgomery ([5] (Lemma 10.5)) reduced the problem of estimating the mean fourth power of $L(s,\chi^*)$, using the work of Lavrik on approximate functional equations for the Dirichlet's L-series, to an application of the result (3.36) and (3.37), and obtained ([5] (Theorem 10.1))

$$\sideset{}{^*}\sum_{\chi \bmod q} \int_{-T}^{T} |L(\sigma + it, \chi|^4 dt \ll \varphi(q) T \log^4(qT) \text{ for } |\sigma - \frac{1}{2}| \ll \log^{-1}(qT), T \geq 2, \tag{6.2}$$

which may be considered as the average-version of the generalized Lindelof hypothesis:

$$L(\frac{1}{2} + it, \chi) \underset{\varepsilon}{\ll} (q|t|)^{\varepsilon}, \quad |t| \geq 1. \tag{6.3}$$

69 Then he proceeds to derive easily ([5] (Corollary 10.2))

$$\sideset{}{^*}\sum_{\chi \bmod q} \int_{-T}^{T} |L(\frac{1}{2} + it, \chi) L'(\frac{1}{2} + it, \chi)|^2 dt \ll \varphi(q) T \log^6(qT), T \geq 2. \tag{6.4}$$

2. Density theorems.

The next important applications of the large sieve (employed in its hybrid version of Chapter 5 along with many other ingenious ideas) concern to the 'statistical density theorems' for the zeros of the Dirichlet's L-series and, in particular, of the Riemann zeta function. For the history of this subject, see Montgomery [5] (Chapter 12).

We recall the following standard notation required for the description of the results of this section. As usual,

$$N(\sigma, T, \chi) \quad (\sigma \leq 1) \tag{6.5}$$

denotes the number of zeros

$$\rho = \beta + i\gamma \tag{6.6}$$

of the function

$$L(S, \chi) \tag{6.7}$$

in the rectangle

$$\sigma \leq \beta \leq 1, |\gamma| \leq T. \tag{6.8}$$

Particularly for the Riemann zeta function, i.e., $q = 1$, we use

$$N(\sigma, T) \tag{6.9}$$

instead of (6.5).

Regarding the average over χ of (6.5), for a fixed q, the best known results at present are the following. We have

$$\sum_{\chi \bmod q} N(\sigma, T, \chi) \ll \begin{cases} (qT)^{\frac{3}{2-\sigma}(1-\sigma} \log^9(qT), \frac{1}{2} \leq \sigma \leq \frac{3}{4} \\ \quad (\text{ Montgomery [5](Theorem 12.1)}), \\ (qT)^{\frac{3}{3\sigma-1}(1-\sigma)+\varepsilon}, \frac{3}{4} \leq \sigma \leq \frac{4}{5}(\text{ Huxley [12]}), \quad (\varepsilon > 0). \\ (qT)^{(2+\varepsilon)(1-\sigma)}, \frac{4}{5} \leq \sigma \leq 1(\text{ Jutila [12]}) \end{cases} \tag{6.10}$$

(Here and in what follows, the $\ll$ - constant is understood to depend on 70
ϵ whenever the bound contains ϵ.) It is easily deduced from (6.10) that

$$\sum_{\chi \bmod q} N(\sigma, T, \chi) \ll (qT)^{(\frac{12}{5}+\varepsilon)(1-\sigma)}, \quad \frac{1}{2} \leq \sigma \leq 1, \quad (\varepsilon > 0), \tag{6.11}$$

holds uniformly in σ, $q \geq 1$ and $T \geq 2$. And, for the average over q we have

$$\sum_{q \leq Q} \sideset{}{^*}\sum_{\chi \bmod q} N(\sigma, T, \chi) \ll \begin{cases} (Q^3T)^{\frac{3}{2-\sigma}(1-\sigma)} \log^9(qT), \frac{1}{2} \leq \sigma \leq \frac{3}{4} \\ \quad (\text{ Montgomery [5](Theorem 12.2)}) \\ (Q^2T)^{\frac{3}{3\sigma-1}(1-\sigma)+\varepsilon}, \frac{3}{4} \leq \sigma \leq \frac{4}{5}(\text{ Huxley [12]}), (\varepsilon > 0). \\ (Q^2T)^{2+\varepsilon)(1-\sigma)}, \frac{4}{5} \leq \sigma \leq 1(\text{ Jutila [12]}), \end{cases} \tag{6.12}$$

from which one gets the estimate analogous to (6.11)

$$\sum_{q \leq Q} \sideset{}{^*}\sum_{\chi \bmod q} N(\sigma, T, \chi) \ll (Q^2T)^{(\frac{12}{5}+\varepsilon)(1-\sigma)}, \frac{1}{2} \leq \sigma \leq 1, (\varepsilon > 0) \tag{6.13}$$

valid uniformly in σ, $Q \geq 1$ and $T \geq 2$.

The function $N(\sigma, T)$ has been investigated more extensively. Estimates of the form

$$N(\sigma, T) \ll T^{2(1-\sigma)+\epsilon} \quad (\epsilon > 0). \tag{6.14}$$

valid in $\sigma > \alpha$ (for some α) uniformly, are called 'density hypothesis'. It is known that the Lindelöf hypothesis:

$$\zeta(\frac{1}{2} + it) \ll (1 + |t|)^{\epsilon} \quad (\epsilon > 0) \tag{6.15}$$

implies (6.14), with $\alpha = \dfrac{1}{2}$. For some applications proved under the
71 assumption of (6.15) it suffices instead to have a result of the type (6.14). Now the density hypothesis is known for

$$\sigma \geq \frac{11}{14} = 0.78571\ldots \tag{6.16}$$

and this result is due to Jutila [9], [10] (cf. [12]). In $\sigma \geq \frac{61}{74}$ still better estimates are available; namely,

$$N(\sigma, T) \leq T^{\lambda(\sigma)(1-\sigma)+\epsilon} \quad (\epsilon > 0) \tag{6.17}$$

with

$$\lambda(\sigma) = \begin{cases} \frac{48}{37(2\sigma-1}, \frac{61}{74} \leq \sigma \leq \frac{37}{42}(\text{ Huxley [11]}), \\ \frac{3}{2\sigma}, \frac{37}{42} \leq \sigma \leq \frac{37+\sqrt{73}}{48}(\text{ Huxley [11]}), \\ \frac{4(3\sigma-2)}{3(4\sigma-3)(2\sigma-1)}, \frac{37+\sqrt{73}}{48} \leq \sigma \leq 1 \\ \qquad (\text{Montgomery [5](Corollary 12.4)}). \end{cases} \tag{6.18}$$

Close to the line $\sigma = 1$, we have (Halász and Turán [1]) even

$$N(\sigma, T) \underset{\delta}{\ll} T^{(1-\sigma)^{3/2}} \log^3(\frac{1}{1-\sigma}), \sigma \geq 1 - \delta \quad (\delta > 0) \tag{6.19}$$

and (Montgomery [5] (Corollary 12.5))

$$N(\sigma, T) \ll T^{167(1-\sigma)^{3/2}} \log^{17} T, \sigma \geq \frac{1}{2}, T \geq 2. \tag{6.20}$$

On the other side, in the vicinity of $\sigma = \frac{1}{2}$ and for $(\frac{1}{2} \leq)\sigma \leq \frac{3}{4}$ still the best known density estimates are due to A. Selberg and Ingham (cf.

Montgomery [5] (Chapter 12)). For σ between $\dfrac{3}{4}$ and Jutila's bound (6.16), the best known estimate (6.17) with

$$\lambda(\sigma) = 3(\min_{k\in\mathbb{N}} \max \frac{1}{(3\sigma-1)+\frac{2}{k}(1-\sigma)}, \frac{1}{2k(4\sigma-3)+3(1-\sigma)}), \frac{3}{4} < \sigma \leq \frac{11}{14} \tag{6.21}$$

is also due to Jutila [10].

The connection between the order of $\zeta(s)$ and the density estimates has already been indicated (cf. (6.14) and (6.15)). Indeed, some of the aforementioned bounds for $N(\sigma, T)$ can be improved slightly by using **72**
better estimates for $\zeta(s)$. For general results in this context we refer to Bombieri [3] (and for $L(s,\chi)$ to Forti and Viola [1]).

3. Mean-value Theorems of the Bombieri type.

In this section and the next we mention the applications, involving number-theoretic functions, which have important consequences in (proper) number theory. As regards the notation almost all are standard and so we repeat only one of these, namely, the Hurwitz's zeta function,defined through

$$\zeta(s,w) = \sum_{n=0}^{\infty}(n+w)^{-s}, s = \sigma + it \quad (\sigma > 1), 0 < w \leq 1, \tag{6.22}$$

and analytic continuation.

Now, one of the most important applications of the large sieve has been to what we shall call as Bombieri's prime number theorem (Bombieri [1]): For any given number $U(>0)$ there exists a value $C = C(U)$ such that

$$\sum_{q\leq x^{\frac{1}{2}}\log^{-C}x} \max_{2\leq y\leq x} \max_{(\ell,q)=1}\left|\chi(y;q,\ell) - \frac{li\, y}{\varphi(q)}\right| \ll_U \frac{x}{(\log x)^U}, \tag{6.23}$$

or equivalently

$$\sum_{q\leq x^{\frac{1}{2}}\log^{-C}x} \max_{2\leq y\leq x} \max_{(\ell,q)=1}|\psi(y;q,\ell) - \frac{y}{\varphi}| \ll_U \frac{x}{(\log x)^U}. \tag{6.24}$$

A result of this kind can be derived, either via estimates of the type (6.12) or directly, from the large sieve. Bombieri proved (6.24) (via (6.12)-type result) with the value $C = 3U + 23$, and the best known result now is with

$$C = U + \frac{7}{2}, \tag{6.25}$$

due to Vaughan [6], who obtained this by a refinement of Gallagher's [2] method (-a direct application of the large sieve-) for a proof of (6.24).
73 Jutila [1] has proved a corresponding result for short intervals which states that

$$\sum_{q\le x^{\beta}} \max_{z\le x^{\theta}} \max_{(\ell,q)=1} |\psi(x+z;q,\ell)-\psi(x;q,\ell)-\frac{z}{\varphi(q)}| \ll_{U,\varepsilon} \frac{x^{\theta}}{(\log x)^{U}}, 0<\theta<1. \tag{6.26}$$

where

$$\beta = \beta(\theta,\epsilon) = \frac{4c\theta + 2\theta - 1 - 4c}{6+4c} - \epsilon \tag{6.27}$$

in which c denotes a constant satisfying, for the functionl in (6.22),

$$\zeta(\frac{1}{2} + it, w) \ll_{\delta} (1+|t|)^{c+\delta} \text{ for every } \delta > 0. \tag{6.28}$$

For the various (more or less equivalent) forms of (6.23) and (6.24) we refer the reader to Elliott and Halberstam [2] and Montgomery [5] (Chapter 15). As an example, we mention an interesting remark of Montgomery and Vaughan [2] in connection with the following result derived from (3.4)

$$\sum_{q\le \frac{N^{\frac{1}{2}}}{200}} \log(\frac{N^{\frac{1}{2}}}{q}) \sideset{}{^*}\sum_{\chi \bmod q} |\psi(N,\chi)|^2 < N^2 \log N, \quad N > N_0, \tag{6.29}$$

where (cf. (1.75))

$$\psi(N,\chi) := \sum_{n\le N} \Lambda(n)\chi(n). \tag{6.30}$$

Now the term corresponding to $q = 1$ in (6.29) contributes already $\frac{1}{2}N^2 \log N + o(N^2)$ to the sum. Further, if $L(s,\chi_1)$ has a Siegel - zero,

$q_1 = N^\delta$, then $|\psi(N,\chi_1)| > (1-\delta)N$ and consequently the contribution from the term for χ_1 in (6.29) is atleast $(\frac{1}{2} - 2\delta)N^2 \log N$ so that

$$\sum_{N^\delta < q \leq \frac{N^{\frac{1}{2}}}{200}} \sideset{}{^*}\sum_{\chi \bmod q} |\psi(N,\chi)|^2 \ll \delta N^2 \log N, \tag{6.31}$$

which seems rather unlikely to be true for sufficiently small δ.

There are results, analogous to (6.23), concerning the average order of remainder terms with respect to other number-theoretic functions. From these results also one has been able to obtain results which could 74 only be proved earlier either by the use of the complicated Linnik's dispersion method or only under the assumption of the generalized Riemann hypothesis.

Such analogues are now available for the functions $d_k(n)$ (A.I. Vinogradov [1], Motohashi [1], [7]), $r(n)$ (Motohashi [2], Siebert and Wolke [1]) and for certain powers of these functions. Certain other special functions have also been investigated (Siebert and Wolke [1], see (6.33) below, and Wolke [7]) and interestingly we have now general results of the type (6.23), based on the large sieve, due to Wolke (cf. Wolke [5], [6], Siebert and Wolke [1]): Under certain conditions (stemming from the work of Wirsing) for a multiplicative number-theoretic function $f(n)$, one has

$$\sum_{q \leq x^{\frac{1}{2}} \log^{-C} x} \max_{y \leq x} \max_{(\ell,q)=1} \Big| \sum_{\substack{n \leq y \\ n \equiv \ell \bmod q}} f(n) - \frac{1}{\varphi(n)} \sum_{\substack{n \leq y \\ (n,q)=1}} f(n) \Big| \ll_U \frac{x}{(\log x)^U}, U > 0, C = C(U). \tag{6.32}$$

As an example, we mention the following consequence of (6.32) (Siebert and Wolke [1]):

$$\sum_{q \leq x^{\frac{1}{2}} \log^{-C} x} \max_{y \leq x} \max_{\ell} \Big| \sum_{\substack{n \leq y \\ n \equiv \ell \bmod q}} \mu(n) \Big| \ell_U \frac{x}{(\log x)^U}, U > 0, C = C(U). \tag{6.33}$$

Orr [1] (cf. [2]) has derived such a result for the number of squarefree integers (i.e., $\mu^2(n)$) in an arithmetic progression in an elementary way. For another elementary derivation of such results from a different type of mean-value theorems we refer to the next section.

4. Mean-value Theorems of the Barban-Danport-Halberstam type.

There is another type of mean-value theorem corresponding to (6.24) which deals with the mean-square instead and has a considerably much
75 wider range of validity (for q):

$$\sum_{q\leq x\log} -(U+1)_x \sum_{\ell=1}^{q}{}'(\psi(x;q,\ell)-\frac{x}{\varphi(q)})^2 \underset{U}{\ll} \frac{x^2}{(\log x)^U} \quad \text{for } U>0. \tag{6.34}$$

Such a result was found (in a slightly weaker form) first by Barban ([6] (Theorem 1), cf. [10] (Theorem 3.2)) and was later rediscovered by Davenport and Halberstam [2]. The improved form (6.34) is due to Gallagher [1]. The proof is based on the Siegel-Walfisz theorem and the identity (Davenport and Halberstam [2]) (1.76),

$$\sum_{\ell=1}^{q}{}'(\psi(x;q,\ell)-\frac{x}{\varphi(q)})^2 = \frac{1}{\varphi(q)}\sum_{\chi\neq\chi_0}|\psi(x,\chi)|^2, \tag{6.35}$$

and an application of the large sieve in its form of Chapter 3.

Montgomery ([4], cf. [5] (Chapter 17)) discovered a proof of (6.34) independent of the large sieve and also succeeded in obtaining the asymptotic formulae:

$$\sum_{q\leq Q}\sum_{\ell=1}^{q}{}'(\psi(x;q,\ell)-\frac{x}{\varphi(q)})^2 = \begin{cases} Qx\log Q + o(Qx + x^2\log^{-U}x), \\ \qquad \text{for } Q\leq x, (\text{ for any fixed } U\leq 0) \\ Qx\log x - \frac{\zeta(2)\zeta(3)}{\zeta(6)}x^2\log\frac{Q}{x} - \\ Qx + Ax^2 + o(Qx\log^{-U}x), \\ \text{for } Q>x. \end{cases} \tag{6.36}$$

The part $Q\leq x$ of (6.36) is an improved version of Montgomery's first result (Croft [1]).

Montgomery's method of proof of (6.36) is based on a deep theorem of Lavrik [1], [2], about the average order of the error-term in the generalized twin-prime problem, which may be stated as

$$\sum_{q<\frac{x}{2}}(\sum_{2q<n\leq x}\Lambda(n)\Lambda(n-2q) - 2(x-2q)\prod_{p>2}(1-\frac{1}{(p-1)^2})\prod_{2<p|q}\frac{p-1}{p-2})^2 \leq_U \frac{x^3}{\log_x^U} \tag{6.37}$$

and which in turn depends on the method of I.M. Vinogradov for the estimation of exponential sums.

Regarding the first part ($Q \leq x$) of (6.36), Hooley [5] has shown that one can replace the error term by 76

$$AQx + O(Q^{5/4}x^{3/4} + x^2 \log^{-U} x) \tag{6.38}$$

as an application of the (simpler) large sieve method only.

Hooley [8] has also proved, on the basis of the large sieve method (of Chapter 3), the general result of the Barban-Davenport-Halberstam type: Let γ be a set of positive integers and suppose that there holds, for all $U > 0$ and all integers q, ℓ, (for some function $g(q, b)$)

$$S(x; q, l) := \sum_{\substack{n \in \gamma \\ n \leq x \\ n \equiv \ell \bmod q}} 1 = g(q, (l, q))x + O_U(x \log^{-U} x) \text{ as } x \to \infty. \tag{6.39}$$

Then

$$\sum_{q \leq Q} \sum_{l=1}^{q} (S(x; q, l)) - g(q, (l, q)x)^2 = O(Qx) + O(x^2 \log^{-U} x)$$
$$\text{for } 1 \leq Q \leq x \text{ (for any fixed U).} \tag{6.40}$$

Combining this with the method of his paper [5] he derived that

$$\sum_{q \leq Q} \sum_{l=1}^{q} (\sum_{\substack{n \leq x \\ n = l \bmod q}})\mu(n))^2 = \frac{6}{\pi^2} Qx + O(x^2 \log^{-U} x) \text{ for } 1 \leq x, (\text{ for any fixed } U). \tag{6.41}$$

An interesting connection between the mean-value theorems of the squared expression and those of the preceding section was noted by Barban [6]. He observed that from a mean-value theorem of the type (6.40) it is possible to derive in an elementary manner a Bombieri-type result for a related function (depending on the parameter x). Here we shall present a proof in the simplest case: Let us put

$$\psi_2(x; q, l) := \sum_{\substack{n_1 \leq x \\ n_1 n_2 \equiv \ell \bmod q}} \sum_{n_2 \leq x} \Lambda(n_1)\Lambda(n_2), \tag{6.42}$$

$$E_2(x;q,l) := \psi_2(x;q,l) - \frac{x}{\phi(q)}, \tag{6.43}$$

$$E_1(x;q,l) := \phi(x;q,l) - \frac{x}{\phi(q)}, \tag{6.44}$$

77 and assume that

$$(\ell, q) = 1. \tag{6.45}$$

We have

$$\psi_2(x;q,l) = \sum_{h=1}^{q}{}' \sum_{\substack{n_1 \le \sqrt{x}\\ n_1 \equiv h \bmod q}} \Lambda(n_1) \sum_{\substack{n_2 \le \sqrt{x}\\ n_1 \equiv h^{-1} l \bmod q}} \Lambda(n_2)$$

$$= \sum_{h=1}^{q}{}' \psi(\sqrt{x};q,h)\psi(\sqrt{x};q,h^{-1}l) \tag{6.46}$$

(here h^{-1} represent the residue class $\bmod q$ for which $hh^{-1} \equiv 1 \bmod q$) and

$$\Big|\sum_{b=1}^{q}{}'(\psi(\sqrt{x};q,b) - \frac{\sqrt{x}}{\phi(q)}\Big| \le \Big|\psi(\sqrt{x}) - \sqrt{x}| + \log\ q. \tag{6.47}$$

Subtracting $\dfrac{x}{\phi(q)}$ from both the the sides of (6.46) and using (6.47) we obtain, by Cauchy's inequality,

$$\begin{cases} E_2(x;q,l) & \le \sum_{h=1}^{q}{}' E_1(\sqrt{x},q,h)E_1(\sqrt{x},q,h^{-1}l) \\ & \quad +2\frac{\sqrt{x}}{\phi(q)}\left\{|\psi(\sqrt{x}) - \sqrt{x}| + \log q\right\} \le \\ & \le \sum_{b=1}^{q}{}' E_1^2(\sqrt{x};q,b) + 2\frac{\sqrt{x}}{\phi(q)}\left\{|\psi(\sqrt{x}) - \sqrt{x}| + \log\ q\right\} \end{cases} \tag{6.48}$$

uniformly in ℓ subject to (6.45). Now summation over q and the use of (6.34) along with an application of the prime number theorem in the form

$$(\psi(\sqrt{x}) - \sqrt{x}) << \frac{\sqrt{x}}{(\log x)^{U+1}}, \tag{6.49}$$

give

$$\sum_{q\leq\sqrt{x}\log^{-(U+1)}x}\max_{(l,q)=1}\Bigg|\sum_{\substack{n_1\leq\sqrt{x}\\ n_1n_2\equiv}}\sum_{\substack{n_2\leq\sqrt{x}\\ l \bmod q}}\Lambda(n_1)\Lambda(n_2)-\frac{x}{\phi(q)}\Bigg|\ll_U\frac{x}{(\log x)^U}. \tag{6.50}$$

Thus we have shown that this Bombieri-type result is an elementary consequences of (6.34).

5. Some number-theoretic applications.

In this last section we merely record a few applications of the preceding result to pure number theory.

First of all we have the important consequences about the difference between consecutive primes that 78

$$p_{n+1}-p_n<p_n^{\delta+\xi}\text{ for } n\geq n_0(\epsilon),\ \epsilon>0, \tag{6.51}$$

with

$$\delta=1-\frac{1}{\lambda} \tag{6.52}$$

whenever there is a result (6.17) with uniform λ valid in $\frac{1}{2}\leq\sigma\leq 1$ (cf., for the history of this question, Montgomery [5] (Chapter 14)). Therefore the density hypothesis (6.14) with $\alpha=\frac{1}{2}$ would give (6.51) with $\delta=\frac{1}{2}$. Montgomery [3] proved (6.51) with $\delta=3/5$ and Huxley [8] (cf. [7] (p. 119)) succeeded in getting the best known value

$$\delta=\frac{7}{12}. \tag{6.53}$$

(Now, in view of our initial statement, we can get (6.53) from either of (6.11) or (6.13) also.)

An analogous application of (6.10) is known with respect to the least prime $p_1(q,\ell)$ in the arithmetic progression $\{\ell,\ell+q,\ell+2q,\ldots\}, (\ell,q)=1, 0\leq\ell<q$. This stated (Iwaniec [4])

$$p_1(q,\ell)\ll_{\epsilon,q(q)} q^{\lambda+\epsilon}, \epsilon>0, \tag{6.54}$$

so that one has from Jutila's result (6.11) that

$$P_1(q,l) \ll_{\epsilon,q(q)} q^{\frac{12}{5}+\epsilon}, \epsilon > 0. \tag{6.55}$$

(However, observe that the $\ll$-constant depend on the kernel of q (cf. (1.19)). Here, with respect to Linnik's famous theorem, the best known exponent in (6.54) with $\ll$-constant independent of $q(q)$ is 550 proved unconditionally by Jutila.)

79 Regarding the analogues, for other arithmetical functions, of the Bombieri theorem as well as of Barban-Davenport-Halberstam theorem mentioned earlier, we have their applications in the proof of various delicate asymptotic formulae, involving such function as the divisor function $d_k(n)$, $r(n)$, $\mu(n)$ etc., as also general function fulfilling certain conditions, their powers and some mixed forms of these functions, and further, with the argument running through certain polynomial sequences. For these problems we refer to Barban [6], [10], Elliot and Halbertam [1], Hooley [1], [6], Huxley and Iwaniec [1]. Indlekofer [1], [2], Iwaniec [2], [5], Katai [1], Linnik [4], Motohashi [1], [2], [4], [7], Orr [1], [2], Proter [2]. [3], Rodriquez [1], Siebert and Wolke [1], Vaughan [4], [5], A.I. Vinograndov [1], and Wolke [3], [5], [6].

NOTES

1.: Montgomery and Vaugham [1] have shown that the proof of the classical formula

$$\int\limits_0^T |\zeta(\frac{1}{2}+it)|^2 dt = T\log T + O(T) \tag{6.56}$$

can be greatly simplified by an application of Theorem 4.3. Ramachandra [6], [7] has extended this to various other moments with reports to $\zeta(s)$ and L-functions. Such results are important for the results mentioned in the second section of this chapter (as has already been indicated in the context of (6.15)).

(6.1), (6.2): For similar results involving $|L(\frac{1}{2}+it,\chi)|^2$ see Gallagher [8]. More general forms, with an averaging over certain well-spaced

t-sets of points, can be obtained by an additional use of (6.2) (see Montgomery [5]) (Theorem 10.3 and Corollary 10.4), and Huxley **80**
[7] (pp. 97,108)). Such results can be found in Ramachandra [3] for $\zeta(s)$ in $\frac{1}{2} \leq \sigma < 1$, for $L(s,\chi)$ in Ramachandra [7] and Jutila [10], and with χ mod p for $L(s,\chi)$ in Elliott [9].

(6.2): A simple proof of (6.2) has been given by Ramachandra [7]. For a result of this type with real characters only see Jutila [7].

(6.4): For a mean-value theorem for $|L'(\frac{1}{2}+it,\chi)|^2$ which can be derived in the same way, see Vaughan [6].

2.: (6.5): Gallagher [8] has proved that one has

$$N(\sigma,T,\chi) \ll T^{3(1-\sigma)} \log^C T, \text{ for } q \leq T \tag{6.57}$$

and

$$\sum_{\chi \bmod q} (N(\sigma,T+1,\chi) - N(\sigma,T,\chi)) \ll q^{3(1-\sigma)} \log^C T, \text{ for } T \leq q \tag{6.58}$$

with some constant C.

For results about the size of $L(1,\chi), \chi$ mod p, we refer to the papers of Bareman, Chowla and Erdös [1], Barban [8], [10]. Elliott [1]. Joshi [1] and for any $L(s,\chi)$ to Elliott [8].

(6.10), (6.12): For estimates of these averages, under certain restrictions on q and Q, respectively, see Ramachandra [4]. For earlier results, in particular, regarding the 'generalized density hypothesis' (namely, the estimates (6.10), (6.12) with the exponent $2(1-\sigma)+\epsilon$ instead), see Balasubramanian and Ramachandra [1], Huxley [9], Huxley and Jutila [1], and Jutila [5], [9], [10], [11].

(6.12): For a result, with the summation restriction to real characters only, which has better estimates in some cases, see Jutila [4], [6]. A. Selberg [6], cf. Montgomery [8]) had proved (using his **81** results (3.14) and (3.33)) that

$$\sum_{q\leq Q}\sideset{}{^*}\sum_{\mathcal{X} \bmod q} N(\sigma, T, \mathcal{X}) \ll_\epsilon (Q^5T^3)^{(1+\epsilon)(1-\sigma)}, \frac{1}{2} \leq \sigma \leq 1(\epsilon > 0). \tag{6.59}$$

In this connection we have also the following deep result due to Bombieri ([6] (Théorèm 14)): If there is a 'Siegel-zero' β_1 (relative to $T \geq 2$ and a certain constant $c_0 > 0$) then holds

$$\sum_{q\leq T}\sideset{}{^*}\sum_{\chi \bmod q} N(\sigma, T\chi) \ll ((1-\beta_1)\log T)T^{c(1-\sigma)} \tag{6.60}$$

with some absolute and effective c and $\ll$-constants, where on the left-hand side the exceptional zeros are not included. As an application of this Bombieri [6] derives the well-known theorem of Siegel: One has

$$1-\beta_1 \geq c(\epsilon)T^{-\epsilon}, T \geq 2(\epsilon > o), \tag{6.61}$$

with (as in all other known proofs) an ineffective $c(\epsilon) > o$. We briefly sketch this deduction. Introducing θ as the supremum of the real parts of the zeros of all the Dirichlet's L-functions, we see easily that one can suppose (for the purpose of (6.61)) $\theta = 1$. Now, taking T to satisfy

$$T \geq \max(q_0, |\gamma_0|, \exp(\frac{c_0}{1-\beta_0})), \beta_0 > 1-\epsilon. \tag{6.62}$$

where q_0 is the (least) modules of the L-function which has a zero $\rho_\circ = \beta_\circ + i\gamma_\circ$ (with $\beta_\circ > 1-\epsilon$), we see that $\rho_\circ$ is not an exceptional zero (relative to T and c_0) and also that the left-left-hand side (6.60) with $\sigma = 1-\epsilon$ is ≥ 1. Hence we have

$$(1-\beta_1) \gg ((\log T)T^{c\epsilon})^{-1} \tag{6.63}$$

which gives (6.61) for T subject to (6.62) and so for all $T \geq 2$.

(6.64): Halász and Turán [1] have shown that (6.15) gives (even)

$$N(\sigma, T) \underset{\xi\delta}{\ll} T^\epsilon, \text{ for } \sigma \geq \frac{3}{4} + \delta, (\text{ with } \epsilon > o). \tag{6.64}$$

Also they have proved in [2] that the generalized Lindel of hypothesis,

$$L(s,\chi) \underset{\epsilon\ T}{\ll} q^{\epsilon}, \text{ uniformly for } \sigma \geq \frac{1}{2}, |t| \leq T, \tag{6.65}$$

implies the uniformly estimates, for (6.12).

$$\underset{\epsilon\ T}{\ll} Q^{4(1-\sigma)+\epsilon} \text{ for } \frac{1}{2} \leq \sigma \leq 1 \text{ and } \underset{\epsilon\ T,\delta}{\ll} Q^{\epsilon} \text{ for } \sigma \geq \frac{3}{4} + \delta(\epsilon > o, \delta > o). \tag{6.66}$$

82

(6.16): In this connection, for earlier results, see Montgomery [3], Huxley [6], Ramachandra [8], Forti and Viola [2], Huxley [9], [11], [12], and Jutila [8].

(6.17): Ramachandra [9] has proved for $\sigma \geq \frac{1}{2}$ that

$$N(\sigma, T + T^{5/12}) - N(\sigma, T) \underset{\epsilon}{\ll} T^{\frac{5}{9-6\sigma}(1-\sigma+\epsilon)}, (\epsilon > 0) \tag{6.67}$$

and more generally, Balasubramanian [1] has shown that, uniformly in $\sigma \geq \frac{1}{2}$

$$N(\sigma, T + H) - N(\sigma, T) \ll H^{\frac{4}{3-2\sigma}(1-\sigma)} \log^{100} H \text{ for } T^{\frac{27}{82}} \geq H \geq T. \tag{6.68}$$

(6.21): This type of exponents in (6.10) and (6.12) occurs also in Ramachandra [8], and Balasubramanian and Ramachandran [1].

3.: (6.23), (6.24): The first result of this type is due to Rényi [2]. Replacing the sum on the left by

$$\sum_{q \leq x^{\alpha}} . \tag{6.69}$$

Barban [2] improved Rényi's result (i,e., with a certain small $\alpha > 0$) to $\alpha = \frac{1}{6} - \epsilon$ (for every $\epsilon > 0$) and succeeded then ([4], [5], [6]) in extending it to all $\alpha < 3/8$. Since the value of the result increases considerably with α, Bombieri's theorem has led to most important applications. A.I. Vinogradov [1] proved the same theorem (through in a slightly weaker form, namely, with $\alpha = \frac{1}{2} - \epsilon$ for every $\epsilon >$ 0) simultaneously. A result of this type with $\alpha > \frac{1}{2}$ would have important consequences (cf. Buchstad [2], Helberstam, Jurkat and Richert [1]) and it has been conjectured by Elliott and Halberstam [2] (cf. Elliott [6]) that every value of $\alpha < 1$ is admissible.

(6.25): Gallagher [2] uses in his proof of (6.24) the decomposition
83 $\frac{L'}{L} = (1 - LG)^2 \frac{L'}{L} + 2L'G - LL'G^2$ where G is a partial sum of the Dirichlet's series for $\frac{1}{L}$, instead of which Vaughan [6], to obtain (6.25), uses the more efficient splitting $\frac{L'}{L} = (\frac{L'}{L} + F)(1 - LG) +$ $(L' + LF)G - F$ with F being the partial sum corresponding to $\frac{L'}{L}$.

(6.26): For a different proof of (a slightly weaker form of) (6.26) see Motohashi [5]; more recent results which in certain ranges (for θ) improves (6.26) are due to Huxley and Iwaniec [1] (according to Motgomery (oral communication), S. Ricci, a student of his, has obtained a similar result).

(6.28): The best exponents known here,

$$c \leq \frac{173}{1067} < \frac{1}{6}, \tag{6.70}$$

is due to Kolesnik [1].

(6.29): For similar results, see Bombieri and Davenport [2], Montgomery [5] (Chapter 15), and Vaughan [6].

(6.32): In this context, we mentioned that the function $r(n)$ which appears in this and the last section of this Chapter and for which also (6.32) holds, denotes the general

$$r(n) := \sum_{d|n} \chi(d), \chi \bmod q, n \in N : \tag{6.71}$$

in particular, when χ is the non-principle character to the modulus 4, $r(n)$ denotes the number of representation of n as a sum of two squares (apart from a factor of 4).

4.: (6.36): For stronger and more general results, and also for those which can be obtained on the assumption of the generalized Riemann hypothesis, see Hooley [5], [7], [9]. The corresponding problem for

$$\sum_{\substack{n\geq x\\ n\equiv l \bmod q}} \mu^2(n) \tag{6.72}$$

has been investigated extensively, both with the help of the large sieve as well as by elementary methods only (Orr [1] (cf. Orr [2]). Warlimont [1], [2], and Croft [1]). Motohashi [6] has used the idea of Montgomery's proof, namely without using the large sieve, to obtain an asymptotic formula with respect to 84

$$\sum_{\substack{n\leq x\\ n\equiv l \bmod q}} d(n) \tag{6.73}$$

(6.50): Barban ([10] (Theorem 3.3)) has extended his result also to the more general sums

$$\sum_{\substack{p_i\leq x\\ i=1,\ldots,k\\ p_1\ldots p_k\equiv l \bmod q}} \tag{6.74}$$

where

$$0 < \alpha_1 \leq \ldots \leq \alpha_k, \alpha_1 + \cdots + \alpha_k = 1, k \geq 2, \tag{6.75}$$

of which (6.50) corresponds to the case

$$k = 2; \alpha_1 = \alpha_2 = \frac{1}{2}. \tag{6.76}$$

Orr [1], [2] has obtained results, with sums of the type (6.74), where primes are replaced by squarefree numbers, and also with mixed products instead.

For results of others types concerning short interval we merely mention two recent papers: Gallagher [5], and Ramachandra [10].

5.: (6.51): The history of the more general question of the number of
85 primes in short interval

$$[x, x + x^{\delta}]$$

starts with Hohesisel, who was the first to prove a result with a $\delta < 1$, actually with any $\delta > 1 - (33000)^{-1}$. Another famous question, concerning small differences of primes, is to find an estimate of the form

$$\underline{\lim}_{n\to\infty} \frac{p_{n+1} - p_n}{\log p_n} \leq \theta, \qquad (p_n : n^{\text{th}} \text{ prime}). \tag{6.77}$$

It is conjectured that this $\underline{\lim}$ is zero, an obvious consequences of the twin-prime conjecture and it has been known that, under the assumption of the generalized Riemann hypothesis, one would have (6.77) with $\theta = \dfrac{1}{2}$. However, Bombieri and Debenport [1] obtained (6.77) unconditionally with a $\theta < \dfrac{1}{2}$, in fact, with $\theta = \dfrac{1}{8}(2 + \sqrt{3}) = 0.46650\ldots$, where they used among other ideas the large sieve. Now we have (6.77), due to Huxley [10], with

$$\theta = \frac{4 + \pi}{16} = 0.44634\ldots \tag{6.78}$$

For earlier results generalizations and allied questions see, for instance, Huxley [2], Uchiyama [1], Moreno [1], and Wolke [11].

(6.54): For $q = p^r$, see Gallagher [5]: for previous result and further references see Jutila [5].

(6.55): The corresponding problem for general sequences, in particular for squarefree numbers, was dealt with by Wolke [1], [2] as an application of (2.103) (see also Warlimont [2]).

We have more application for the result of Section 2., as also of those in the notes to that section. First of all there is the question of the least character non-residues for which we refer to Montgomery [5] (Chapter 13). and Elliott [4]. For estimations of character sums, see Montgomery
86 [5] (Chapter 13), Gallagher [4] [8], and Jutila [6] (as also Jutila [2]). Regarding primitive roots and Artin's conjecture we mention the papers of Gallagher [1]. Burgess and Elliott [1], Goldfeld [1], Elliott [1], [2], Vaughan [1], and for some related questions, Gallagher [1], [3], and

Goldfeld [2]. Finally, for the applications to the problem of the largest prime factor of numbers we refer to the papers of Goldfeld [2], Hooley [2], [4], Jutila [3], Motohashi [3], [11], and Ramachandra [1], [2].

Chapter 7

The 'Sieve Form' of the Large Sieve

AS HAS been mentioned in the beginning of the preceding chapter, we **87**
consider now the (direct) general application of the large sieve in its arithmetical form, in the sense of *B.* of Chapter 0. Let us recall the situation described there. Consider the set of integers in an interval $(M, M+N]$. For each prime p (be-longing to a certain set $\mathfrak{p}$) we drop from our set all such numbers as which fall in any one of certain $\omega(p)$ of the residue classes mod p, and denote the set of remaining integers by γ. Our object is to obtain an upper bound for $S := |\gamma|$.

If $\omega(p)$ is (absolutely) bounded the sieve methods of Brun and Selberg yield satisfactory result. On the other hand, these method fail if, for instance, $\omega(p)$ is an increasing function of p. This was the reason, as has already been started in our remark preceding (0.26), that Linnik called his method the large sieve. In this context one might ask whether it is not possible to have a version of the large sieve which shall include the Selberg sieve, say when $\omega(p)$ is bounded.

This problem has not yet been solved. However as we shall see, it is possible to adapt the large sieve for this purpose provided we confine ourselves to the 'linear sieve' and aim merely for the simplest Selberg upper bound.

In connection with his problem, we recall that in the original form

(of Linnik's method) there was the defect, of having the summation restricted to primes only (cf. (0.54)), which existed upto the first paper of Gallagher [1]. Bombieri was the first to notice (cf, Bombieri and Davenport [2]) that one can solve this problem if the dimension k (cf. (11.3)) of the sieve problem in question equals one (i.e., $\omega(p) = 1$ on
88 the average). In the general case. Montgomery [1] obtained this result first, by discovering the identity (in our notation of Chapter 0)

$$q \sum_{h=1}^{q} \Big| \sum_{d|q} \frac{\mu(d)}{d} S\left(\frac{q}{d}, h\right) \Big|^2 = {\sum_{l=1}^{q}}' \Big| T\left(\frac{l}{q}\right) \Big|^2 \ \forall q \in \mathbb{N}, \tag{7.1}$$

and thereby extending (0.53) to composite numbers.

However, there is a simpler way of dealing with this problem which does not make use of (7.1) but starts instead with the well-known formula (1.34). This was first found for $\kappa = 1$ only (cf. Richert [2]). Later Huxley [5] (see also Montgomery [5] (Chapter 3), and Huxley [7] (Chapter 8) succeeded in extending this method to the general case (cf. (7.24)). In both the cases the question is reduced to an application of the large sieve in its version of Chapter 2. Therefore, the best known solution so far, for our sieve problem, is due to Montgomery and Vaughan [2] (regarding (7.6), see also Gallagher [6]), who derived this from their strong result given in Chapter 2.

Before stating the main result of this chapter we sketch now the aforementioned simpler approach in the case $\kappa = 1$: Let $\mathscr{Q}$ denote the set of all natural numbers composed of only primes p in our set $\mathfrak{p}$ and let us drop (as in Selberg sieve) from the set of integers in $(M, M+N]$ those divisible by some $p \in \mathfrak{p}$ to obtain our $\gamma(\omega(p) = 1)$. (All other notation are as usual.) By (1.34). we have (since $(n, q) = 1$)

$$\mu(q) = {\sum_{\ell=1}^{q}}' e\left(n\frac{\ell}{q}\right) \quad \forall n \in \gamma, \ \forall q \in \mathscr{Q}. \tag{7.2}$$

Summing this over all $n \in \gamma$ and interchanging summation we obtain (on squaring both sides)

$$\mu^2(q) S^2 = \Big| {\sum_{l=1}^{q}}' T\left(\frac{l}{q}\right) \Big|^2, \tag{7.3}$$

which gives, by Cauchy's inequality and summing (after a division bu $\varphi(q)$) overall $q \in \mathscr{Q}$, $q \leq Q$, 89

$$(\sum_{\substack{q \leq Q \\ q \in \mathscr{Q}}} \frac{\mu^2(q)}{\varphi(q)}) S^2 \leq \sum_{\substack{q \leq Q \\ q \in \mathscr{Q}}} \sum_{\ell=1}^{q}{}' |T(\frac{\ell}{q})|^2. \tag{7.4}$$

Now an application of (2.90) gives an upper bound of the desired from for S:

$$S \leq \frac{N + Q^2}{(\sum_{\substack{q \leq Q \\ q \in \mathscr{Q}}} \frac{\mu^2(q)}{\varphi(q)})}. \tag{7.5}$$

Now we come to the main result of this chapter. We state it as

Theorem 7.1. *Let $\gamma \subset (M, M+N]$ be a set of S integers. For each prime p let us denote by $\omega(p)$ the number of residue classes which do not have any number from γ. Then, for any $z > 0$, we have*

$$S \leq \frac{N + z^2}{L(z)}, \tag{7.6}$$

where

$$L(z) = \sum_{q \leq z} \mu^2(q) \prod_{p|q} \frac{\omega(p)}{p - \omega(p)}, \tag{7.7}$$

and also

$$S \leq \frac{1}{L^*(z)}, \tag{7.8}$$

with

$$L^*(z) = \sum_{q \leq z} (N + \frac{3}{2} qz)^{-1} \mu^2(q) \prod_{p|q} \frac{\omega(p)}{p - \omega(p)}. \tag{7.9}$$

Remarks. Observe here that the inequalities (7.6), (7.8), and also those of Theorem 7.2, do not deteriorate if z is replaced by its integral part (with an obvious interpretation if $0 < z < 1$. Therefore, we set Q=[z] and it suffices now to prove our results for $Q \geq 1$. Further, these results remain true if $\omega(p) = p$ for prime p (since then $S = 0$). Hence we assume throughout that 90

$$\omega(p) < p \text{ for all } p \tag{7.10}$$

holds.

Proof. Our first objective is to prove (7.24), an inequality which trivially if $\omega(p) = 0$ for some $p|q$ or if $\mu(q) = 0$ and also for $q = 1$. So we can impose the conditions that

$$1 < q(\leq Q) \text{ is a squarefree number,} \tag{7.11}$$

as well as (cf. (7.10))

$$0 < \omega(p) < p \quad \forall p|q. \tag{7.12}$$

Now, for each $p|q$ we have (for certain $h = h(p)$)

$$n \not\equiv h_i \bmod p, i = 1, \ldots, \omega(p), \quad \forall n \in \gamma, \tag{7.13}$$

which restrictions are equivalent to, by Chinese remainder theorem and (7.11), (with certain $f = f(q)$)

$$(n - f_j, q) = 1, j = 1, \ldots, \omega(q), \forall n \in \gamma \tag{7.14}$$

with

$$\omega(q) = \prod_{p|q} \omega(p), \tag{7.15}$$

so that $\omega(p)$ become a multiplicative function ($\not\equiv 0$, because of (7.12)) on setting

$$\omega(1) = 1; \tag{7.16}$$

i.e., we also have

$$\omega \in m. \tag{7.17}$$

Next, in view of (7.14), we have, by (1.34),

$$\mu(q) = \sideset{}{'}\sum_{l=1}^{q} e((n - f_j)\frac{\ell}{q}) \quad \forall n \in \gamma \text{ and } j = 1, \ldots \omega(q). \tag{7.18}$$

Here summing over all $n \in \gamma$ and also over all j followed by squaring
91 both the sides results in, by Cauchy's inequality,

$$\begin{cases} S^2\mu^2(q)\omega^2(q) & = (\sum_{\ell=1}^{q}{}' \sum_{n\in\gamma} e(n\frac{\ell}{q}) \sum_{j=1}^{\omega(q)} e(-f_j\frac{\ell}{q}))^2 \le \\ & \le (\sum_{\ell=1}^{q}{}' |T(\frac{\ell}{q})|^2)(\sum_{\ell=1}^{q}{}' \sum_{j,j'=1}^{\omega(q)} e((f_{j'}-f_j)\frac{\ell}{q})), \end{cases} \tag{7.19}$$

where we have employed as before (cf. (7.3)) our usual definition (cf. (2.89))

$$T(x) := \sum_{n\in\gamma} e(nx). \tag{7.20}$$

Denoting the second factor in the last expression of (7.19) by $\sum_\circ(q)$ and taking the summation inside, we see that one has on using (1.29) and (1.32)

$$\begin{cases} \sum_\circ(q) & = \sum_{j,j'=1}^{\omega(q)} c_q(f_{j'}-f_j) = \sum_{j,j'=1}^{\omega(q)} \sum_{\substack{d|q \\ d|f_{j'}-f_j}} d\mu(\frac{q}{d}) = \\ & = \sum_{d|q} d\mu(\frac{q}{d}) \sum_{b=1}^{d} (\sum_{\substack{j=1 \\ f_j\equiv b \bmod d}}^{\omega(q)} 1)^2. \end{cases} \tag{7.21}$$

Note that here the $b's$ for which the corresponding inner sum is not empty are precisely those $\omega(d)$ forbidden residue classes $\bmod d$, because of (7.14). Further, for each such b the inner sum counts the same number of f_j's, namely $\omega(q/d)$. Hence, from (7.21).

$$\sum_\circ(q) = \sum_{d|q} d\mu(\frac{q}{d})\omega(d)\omega^2(\frac{q}{d}). \tag{7.22}$$

Now taking $f_1(n) = n\omega(n)$, $f_2(n) = \mu(n)\omega^2(n)$ in (1.18) it follows, because of (7.17), (1.11) and (1.12), that

$$\sum_\circ(q) = \prod_{p|q}(p\omega(p) - \omega^2(p)) = \prod_{p|q}\{\omega(p)(p-\omega(p))\}. \tag{7.23}$$

Using (7.23) in (7.19) and noting (7.15) we obtain

$$S^2\mu^2(q)\prod_{p|q}\frac{\omega(p)}{p-\omega(p)} \le \sum_{l=1}^{q}{}' |T(\frac{l}{q})|^2. \tag{7.24}$$

This is the basic inequality providing the connection between our sieve problem for γ and large sieve method. In view of the remarks at the beginning of the proof, (7.24) is valid for all $q \epsilon \mathbb{N}$. 92

Finally, summing (7.24) over all $q \leq Q$ it follows, by (7.7) and (2.90). that

$$S^2 L(Q) \leq \sum_{q \leq Q} \sum_{\ell=1}^{q}{}' \left| T(\frac{l}{q}) \right|^2 \leq (N + Q^2) \sum_{n = \in \gamma} 1 = (N + Q^2) S. \quad (7.25)$$

This proves (7.6). on recalling our earlier remark preceding (7.10). Further, (7.8) is proved in the same manner by multiplying (7.24) by $(N + \frac{3}{2} qQ)^{-1}$ before summation and then using (2.91). Thus Theorem 7.1 is completely proved. □

The following seemingly more general result is easily derived from Theorem 7.1:

Theorem 7.2. *Under the assumptions of Theorem 7.1, let a_n be arbitrary complex numbers satisfying*

$$a_n = 0 \quad \forall n \notin \gamma. \quad (7.26)$$

Then, for any $z > 0$, we have

$$\Big| \sum_{M < n \leq M+N} a_n \Big|^2 \leq \frac{N + z^2}{L(z)} \sum_{M < n \leq M+N} |a_n|^2 \quad (7.27)$$

and

$$\Big| \sum_{M < n \leq M+N} a_n \Big|^2 \leq \frac{1}{L^*(z)} \sum_{M < n \leq M+N} |a_n|^2. \quad (7.28)$$

Proof. It follows from (7.26), that by Cauchy's inequality the left-hand side of (7.27) (and so also of (7.28)) is

$$\leq (\sum_{n \in \gamma} 1)(\sum_{M < n \leq M+N} |a_n|^2) = S \sum_{M < n \leq M+N} |a_n|^2, \quad (7.29)$$

from which our results are readily obtained from Theorem 7.1. □

93 **Remark.** One might like to consider Theorem 7.2 as a weighted from of Theorem 7.1; but the restriction (7.26) and the relation (7.29) show that Theorem 7.1 is never weaker than Theorem 7.2.

NOTES

(7.1): Montgomery's [1] proof of (7.1) proceeds in the following way: Set

$$\tilde{T}(q,h) := \sum_{\ell=1}^{q} T(\frac{\ell}{q})e(-h\frac{\ell}{q}) \tag{7.30}$$

so that, by (1.27), one has

$$\sum_{h=1}^{q} |\tilde{T}(q,h)|^2 = q \sum_{\ell=1}^{q}{}' |T(\frac{\ell}{q})|^2. \tag{7.31}$$

Now, using (1.29), (1.32) and (0.2),

$$\tilde{T}(q,h) = \sum_{n\epsilon\gamma} \sum_{h=1}^{q} e((n-h)\frac{\ell}{q}) = \sum_{n\epsilon\gamma} \sum_{\substack{d|q \\ d|n-h}} d\mu(\frac{q}{d}) = \sum_{d|q} d\mu(\frac{q}{d}) S(d,h), \tag{7.32}$$

from which (7.1) is derived by means of (1.30) and (7.31).

For identities of this type see (1.27), Montgomery [1], Huxley [7] (Chapter 18), and Sokolovskij [1].

Theorem 7.1: This result contains the (B)-version of the large sieve, as well as, in the cases mentioned in the introduction of this chapter, the 'small' sieves. On the other hand, as per an observation made by Kobayashi [1], one can also derive Theorem 7.1, and consequently also Theorem 7.2 and Theorem 8.1, in these cases from the Selberg sieve with the additional tool of Theorem 2.6 (see Halberstam and Richert [1] (pp. 125–126)).

In the case of $\omega(p)$ being close to p (at least on the average). the 'larger sieve' of Gallagher [3] is more effective. This sieve also includes prime-power moduli.

(7.6): Johnsen [1] has generalized Montgomery's [1] first result of this kind to include non-squarefree numbers also by reducing the question to an inequality of the type of (7.24), so that the improved version (2.90) leads to the following: 94

'For each p remove all but $g(p)$ residue classes $\bmod p$. In each of the remaining classes, remove all but $g(p^2)$ different residue classes $\bmod p^2$. and so on. Then the number of $n \leq N$ which remain is at most $(N+z^2)/\tilde{L}(z)$, for every $z > 0$, where

$$\tilde{L}(z) := \sum_{q\leq z} \prod_{p^\nu \| q} (\frac{p^\nu}{h(p^\nu)} - \frac{p^{\nu-1}}{h(p^{\nu-1})}) \tag{7.33}$$

with $h(p^\nu) = g(p)g(p^2)\dots g(p^\nu)$ being the number of residue classes $\bmod p^\nu$ remaining at the ν^{th} stage'.

A simpler proof of this result has been given by Gallagher [7] (cf. Gallagher [6]).

(7.22): For the remark that precedes (7.22) note that otherwise, by (7.15) and (7.11), there will be one forbidden $b \bmod d$ with more than $\dfrac{\omega(q)}{\omega(d)} = \omega(\dfrac{q}{d})$ of distinct $f_j's \equiv b \bmod d$ and hence for any $n(\in \gamma)$ and one such f_j we would have $(n - f_{j'}\dfrac{q}{d}) > 1$ contrary to (7.14).

(7.24): For a variant of the proof of (7.24) see Montgomery [5] (Chapter 3), and Bombieri [6] (p. 21).

Chapter 8

The Brun-Titchmarsh Theorem

IN THIS chapter we deal with an important application, of the arithmetical sieve result of the previous chapter, in prime number theory. One of the prominent problems of number theory is the study of distribution of primes in arithmetic progressions; i.e., to investigate 95

$$\pi(x;k,\ell) := \sum_{\substack{p\le x \\ p\equiv \ell \bmod k}} 1, \tag{8.1}$$

in particular, to obtain estimates valid uniformly in (ranges of) k (relative to x).

In this direction there is the famous Siegel-Walfisz theorem that, for any $C > 0$ and $U > 0$,

$$\pi(x;k,\ell) = \frac{li\ x}{\varphi(k)} + O_{U,C}(x \log^{-U} x) \ \text{ uniformly in } \ k \le \log^C x \tag{8.2}$$

(with an ineffective O-constant due to the possible existence of a Siegel-zero), which was one of the main tools in I.M. Vinogradov's proof of the solubility of the famous equation

$$2N + 1 = p_1 + p_2 + p_3 \quad \text{for} \quad N \ge N_0. \tag{8.3}$$

Titchmarsh used the generalized Riemann hypothesis to tackle the divisor problem (since then named after him)

$$\sum_{p \le x} d(p-1) \sim cx \quad \text{as} \quad x \to \infty \quad (\text{certain } c > 0) \tag{8.4}$$

where he also employed (cf. (8.21)) the estimate

$$\pi(x; k, \ell) \ll_\alpha \frac{x}{\alpha(k) \log x} \quad \text{for} \quad k \le x^\alpha \quad (\text{fixed } \alpha, 0 < \alpha < 1), \tag{8.5}$$

which he obtained by Brun'sieve.

96 This problem provides a good example to illustrate our earlier remark that Bombieri's prime number theorem can replace the generalized Riemann hypothesis on average, since (6.23) states that $\dfrac{li\, x}{\varphi(k)}$ is the leading term for almost all $k \le \dfrac{x^{\frac{1}{2}}}{\log^C x}$.

Now we are in a position to outline the *unconditional* proof of (8.4) employing (6.23) and (8.5). To start with have

$$\sum_{p \le x} d(p-1) = \sum_{p \le x} \sum_{d | p-1} 1 = \sum_{d \le x} \sum_{\substack{p \le x \\ p \equiv 1 \bmod d}} 1 = \sum_{d \le x} \pi(x; d, 1) \tag{8.6}$$

(in the notation of (8.1)). Let $U > 0$ denote a number to be suitably restricted and let $C = C(U)(> 1)$ be a value for which (6.23) holds. Then splitting the range of d in the last sum of (8.6) into three parts

$$d \le x^{\frac{1}{2}} \log^{-C} x, \quad x^{\frac{1}{2}} \log^{-C} x < d \le x^{\frac{1}{2}} \log^C x, x^{\frac{1}{2}} \log^C x < d \le x \tag{8.7}$$

and denoting the corresponding partial sums there by

$$\sum_1, \sum_2, \sum_3 \tag{8.8}$$

we see that, in view of (8.5), one has

$$\sum_2 \ll \frac{x}{\log x} \sum{}'' \frac{1}{\varphi(d)} \ll \frac{x \log\log x}{\log x}, \tag{8.9}$$

where $''$ denote the restriction of d to the second range in (8.7). Here we have used the fact (cf. Estermann [1]) that

$$\sum_{d \leq y} \frac{1}{\varphi(d)} = c \log y + o(1) \text{ as } y \to \infty, c = \prod_p (1 + \frac{1}{p(p-1)}). \quad (8.10)$$

Next consider (in a notation similar to the one above)

$$\sum_3 = \sum{}'' \sum_{\substack{p \leq x \\ p \equiv 1 \bmod d}} 1 = \sum_{p \leq x} \sum_{d | p-1}{}''' 1. \quad (8.11)$$

Treating $\sum_1$ similarly we see that (after a simple rearrangement) since $C > 1$,

$$\sum_1 - \sum_3 \ll \sum_{p \leq x} \sum_{d | p-1}{}'' 1 + \sum_{p \leq x \log^{-2C} x} d(p-1) \leq \sum_2 + x \log^{-1} x \quad (8.12)$$

from which one obtains, with the help of (8.9) and (8.6), **97**

$$\sum_{p \leq x} d(p-1) = 2 \sum_1 + O(\frac{x \log \log x}{\log x}). \quad (8.13)$$

Now, by our choice of C and (6.23), we have

$$\sum_1 = \sum{}' (\pi(x; d, 1) - \frac{li\, x}{\varphi(d)}) + li\, x \sum{}' \frac{1}{\varphi(d)} + li\, x \sum{}' \frac{1}{\varphi(d)} + 0(x \log^{-U} x). \quad (8.14)$$

Using here (8.10) again we get, in view of (8.13).

$$\sum_{p \leq x} d(p-1) = 2\, li\, x (c \log(x^{\frac{1}{2}} \log^{-C} x)) + O(\frac{x \log \log x}{\log x}) \quad (8.15)$$

on taking $U = 2$ say. Thus, after inserting the value of c from (8.10), we have the asymptotic formula (8.4) in a more precise form:

$$\sum_{p \leq x} d(p-1) = x \prod_p (1 + \frac{1}{p(p-1)} + O(\frac{x \log \log x}{\log x}). \quad (8.16)$$

For number-theoretic purposes the 'Brun-Titchmarsh Theorem' (8.5) is very valuable being the only known result valid in such a wide range for k.

In the direction of (8.5) we shall prove the following result of Montgomery and Vaughan [2] as an application of Theorem 7.1 ((7.8)):

Theorem 8.1. *For any positive real numbers x and y, and for any ℓ, k from $\mathbb{N}$ with*

$$(\ell, k) = 1, \tag{8.17}$$

there is an absolute constant c_0 such that

$$\pi(x+y;k,\ell) - \pi(x;k,\ell) < \frac{2y}{\varphi(k)(\log(\frac{y}{k}) + \frac{13}{15})}, \tag{8.18}$$

provided

$$\frac{y}{k} > c_0. \tag{8.19}$$

98 **Remark.** Under the assumptions of Theorem 8.1 we have, in particular, the estimate (8.18) without the term $\frac{13}{15}$. In this context, by adding some numerical computation, Montgomery and Vaughan have also shown that c_0 can be taken equal to 1, i.e,. we have the neat result

$$\pi(x+y;k,\ell) - \pi(x;k,\ell) < \frac{2y}{\varphi(k)\log(\frac{y}{k})}, 1 \leq \ell \leq k < y, (\ell, k) = 1, \forall x > 0; \tag{8.20}$$

so, in particular, choosing $x = 0$ (and replacing y by x)

$$\pi(x;k.\ell) < \frac{2y}{\varphi(k)\log(\frac{y}{k})}, 1 \leq l \leq k < x, (\ell, k) = 1. \tag{8.21}$$

Proof. Let $z > 0$ denote a number to be suitably chosen later. Consider the set

$$\gamma := \{m : x < mk + \ell \leq x + y, ((mk + \ell), \prod_{\substack{p \leq z \\ p \nmid k}}^{\infty} p) = 1\}. \tag{8.22}$$

□

In the notation of Theorem 7.1 we have

$$M\left[\frac{x-l}{k}\right], N = \left[\frac{x+y-\ell}{k}\right] - \left[\frac{x-l}{k}\right], \tag{8.23}$$

and so

$$\frac{y}{k} - 1 < N < \frac{y}{k} + 1. \tag{8.24}$$

Note that γ contains all those prime numbers counted in $\pi(x+y; k,\ell) - \pi(x;k,\ell)$ which are $> z$, and also

$$\omega(p) \geq \quad \forall p \leq z, p \nmid k. \tag{8.25}$$

Hence, by (7.8),

$$\pi(x+y;k,\ell) - \pi(x;k,\ell) \leq S + z \leq \frac{1}{L^*(z)} + \leq \frac{N}{M_k(z)} + z, \tag{8.26}$$

where, by (8.25),

$$M_k(z) : \sum_{\substack{q\leq z\\(q,k)=1}} (1 + \frac{3}{2}q\frac{z}{N})^{-1}\frac{\mu^2(q)}{\varphi(q)}. \tag{8.27}$$

We now choose **99**

$$z = (\frac{2}{3}N)^{\frac{1}{2}} \tag{8.28}$$

so that

$$\frac{2}{3}\frac{z}{N} = z^{-1} \tag{8.29}$$

which makes

$$M_k(z) = \sum_{\substack{q\leq z\\(q,k)=1}} (1 + qz^{-1})^{-1}\frac{\mu^2(q)}{\varphi(q)}. \tag{8.30}$$

The equality part of (3.26) and the observation that $(1+qz^{-1})^{-1}$ decreases as q increases, enable us to uphold

$$M_k(z) \geq \frac{\varphi(k)}{k}\sum_{q\leq z}(1+qz^{-1})^{-1}\frac{\mu^2(q)}{\varphi(q)} = \frac{\varphi(k)}{k}M_1(z). \tag{8.31}$$

Now we need, instead of the estimate in (3.26) which yielded only the leading term, the more precise formula due to D.R. Ward [1], namely that

$$\sum_{q \leq w} \frac{\mu^2(q)}{\varphi(q)} \log w + c_1 + o(1) \text{ as } w \to \infty, \tag{8.32}$$

where

$$c_1 = \gamma + \sum_p \frac{\log p}{p(p-1)} = \lim_{u \to \infty} (\log u - \sum_{p \leq u} \frac{\log p}{p}) = 1.33258 \cdots \tag{8.33}$$

(cf. Rosser and Schoenfeld [1]). Using (8.32) for partial summation we get as $N \to \infty$

$$M_1(z) = \log z + c_1 - \log 2 + \circ(1) = \frac{1}{2} \log N + c_1 - \log 2 - \frac{1}{2} \log \frac{3}{2} + \circ(1) \tag{8.34}$$

because of (8.28), and consequently

$$\frac{N-1}{N} M_1(z) = \frac{1}{2} \log(N+1) + c_2 + o(1) \text{ as } N \to \infty \tag{8.35}$$

with, in view of (8.33),

$$c_2 = c_1 - \log 2 - \frac{1}{2} \log \frac{3}{2} = c_1 - \frac{1}{2} \log 6 > \frac{13}{30}. \tag{8.36}$$

100

We obtain, on using (8.32), (8.35) and (8.28),

$$\frac{N-1}{N} M_k(z) \geq \frac{\varphi(k)}{2k} (\log(N+1) + 2c_2 + o(1)), z \leq (N-1)^{\frac{1}{2}} \text{ as } N \to \infty. \tag{8.37}$$

Thus, on choosing a sufficiently large c_0, it follows from an use of (8.37) in (8.26)

$$\begin{aligned} \pi(x+y; k, l) - \pi(x; k, l) &\leq \frac{2k(N-1)}{\varphi(k)(\log(N+1) + 2c_2 + \circ(1))} + (N-1)^{\frac{1}{2}} \\ &< \frac{2y}{\varphi(k)(\log(\frac{y}{k}) + \frac{13}{15})} \end{aligned} \tag{8.38}$$

because of (8.24) and (8.19). This completes the proof of the theorem.

NOTES

For the history of Brun-Titchmarsh theorem we refer to Halberstam and Richert [1] (Chapter 3).

(8.12): The estimations of (8.12) are obtained as follows. We have

$$\sum_{1} - \sum_{3} = \sum_{p\leq x}\left(\sum_{d|p-1}{}' 1 - \sum_{d|p-1}{}''' 1\right). \tag{8.39}$$

For any fixed $p \leq x$, the difference of the sums inside is precisely the number of $d's (\leq x^{\frac{1}{2}} \log^{-C} x)$ dividing p-1 and with $\dfrac{p-1}{d} \leq x^{\frac{1}{2}} \log^{C} x$. Now splitting such d's into parts according as $\dfrac{p-1}{d} \leq x^{\frac{1}{2}} \log^{-C} x$ or not and noting further further that the second part is empty if $p > \log^{-2C} x$ we obtain the first majorization. The remaining part employs only the simple $\sum\limits_{m\leq y} d(m) = O(y \log y)$.

(8.16): This unconditional result is due to Rodriquez [1] and Halberstam [1] (cf. Halberstam and Richert [1] (Theorem 3.9). More general results of this type have already been mentioned in the notes of Chapter 6.

Theorem 8.1: The estimate (8.20) demonstrates the power of the weighted sieve of Montgomery and Vaughan (cf. under Theorem 7.1 in the notes for Chapter 7. However, an estimate of the type (8.18) can also be derived without the use of obayashi's results: cf. Halberstam **101** and Richert [1] (pp. 124-125)). On the other hand, it is easily checked that if we use, at the beginning of the proof, the estimate (7.6) instead of (7.8) we cannot obtain (8.20), even subject to the condition (8.19), without an extra term on the right-hand side.

(8.21): A further improvement of the factor 2 in (8.21) to a constant $c < 2$ would have important consequences concerning the Siegel-zeros of Dirichlet's L-functions, as has been first pointed out by Rodosskij (cf. for example, Bombieri and Davenport [2]).

(8.36): It is easily checked to have, instead of (8.36), that

$$c_2 > 1, \tag{8.40}$$

so that one has neater (8.18) with 1 in place of $\frac{13}{15}$, it would be sufficient to improve the constant $\frac{3}{2}$ (at least in (2.91)) (cf. notes for Chapter 2, under (2.76)) to a constant Δ satisfying

$$\Delta\frac{1}{4}e^{2c_1-1} = 1.32163\cdots \tag{8.41}$$

On the other hand (see the above remark) we point out that Selberg's sieve permits one to replace 13/15 in (8.18) by any constant C (with a $c_0 = c_0(C)$ in (8.19)).

Recently, starting from the works of Hooley and Motohashi (Hooley [2], [6], and Motohashi [8], [9], [11]) there has been a remarkable progress with respect to the Brun-Titchmarsh theorem. Some of these results are concerning averages and certain others are valid only for some ranges of k. As an example, we mention one of the most recent results of the latter type (Goldfeld [4]): For every (sufficiently small) $\epsilon > 0$ holds, with a certain $c > 0$,

$$\pi(x;k,l) \leq (1+\epsilon)\frac{x}{\varphi(k)\log(\frac{x}{\sqrt{k^3}}}).if\, x^{\frac{2}{5}-c\epsilon} \leq k \leq x^{\frac{1}{2}} \tag{8.42}$$

Chapter 9

Selberg's Sieve

NOW WE turn to the small sieves. The most elegant version of a small sieve is due to Selberg. In this chapter we present its simplest version with a view to clarifying the main ideas involved. Also, later in the next chapter we importance in the proof of the remarkable theorem of Chen. **102**

This sieve method can be considered as concerning the question of finding bounding for

$$S(\mathscr{A}, \mathfrak{p}, z). \tag{9.1}$$

the number of elements in a (finite) sequence, depending on several parameters,

$$\mathscr{A} := \{a : \cdots\}, a \in \mathbb{Z}. \tag{9.2}$$

of (not necessarily distinct and not necessarily positive) integers, which are not divisible by any prime number $< z$,

$$z \geq (z \in \mathbb{R}), \tag{9.3}$$

belonging to a set of primes

$$\mathfrak{p} := \{p : \ldots\}. \tag{9.4}$$

Introducing

$$P(z) := \prod_{\substack{p<z \\ p\in\mathfrak{p}}} p, \tag{9.5}$$

we can restate this question as that of estimating

$$S(\mathscr{A}, \mathfrak{p}, z) = |\{a : a \in \mathscr{A}, (a, P(z)) = 1\}| \tag{9.6}$$

103 The required estimates would be naturally dependent on the various parameters describing $\mathscr{A}, \mathfrak{p}$, and also on z. However, we would be interested in bounds, for (9.6), which do not involve the special features of these defining arguments ($\mathscr{A}$ and $\mathfrak{p}$). To make this remark clearer, we introduce the following notation.

Let

$$\mathscr{A}_d := \{a : a \in \mathscr{A}, a \equiv 0 \bmod d\} \quad for \quad d \in \mathbb{N}. \tag{9.7}$$

First we choose a convenient approximation X to $|\mathscr{A}|$, requiring

$$X > 1. \tag{9.8}$$

we write for the remainder

$$R_1 := |\mathscr{A}| - X. \tag{9.9}$$

Next, for each prime $p \in \mathfrak{p}$, we choose $\omega(p)(\in \mathbb{R})$ so that $\dfrac{\omega(p)}{p}X$ is close to $|\mathscr{A}_p|$, and set

$$R_p : |\mathscr{A}_p| - \frac{\omega(p)}{p}X, \forall p \in \mathfrak{p}. \tag{9.10}$$

Further, denoting by

$$\overline{\mathfrak{p}} \tag{9.11}$$

the complement of $\mathfrak{p}$ with respect to the set of all primes, we also put

$$\omega(p) = \qquad \forall p \in \overline{\mathfrak{p}}. \tag{9.12}$$

(This is consistent with our interest being only with regard to the distribution of numbers of $\mathscr{A}$ in the residue class 0 modulo primes from $\mathfrak{p}$.) If we now define

$$\omega(d) := \prod_{p|d} \omega(p) \qquad \forall d \in \mathbb{N} \text{ with } \mu(d) \neq 0, \omega(1) := 1 \tag{9.13}$$

(and $\omega(d) = 0$ if $\mu(d) = 0$), we see that

$$\omega \in \mathscr{M} \tag{9.14}$$

104 and also that the definition

$$R_d := |\mathscr{A}| - \frac{\omega(d)}{d} X \ \forall d \in \mathbb{N} \text{ with } \mu(d) \neq 0, \tag{9.15}$$

is consistent with (9.9) and (9.10). (Note that ω may depend on both $\mathscr{A}$ and $\mathfrak{p}$. Now we can elaborate a little on our remarks made subsequent to (9.6). The estimates for (9.6) are allowed to depend on X, ω (and consequently on R), but not on the particular structures of $\mathscr{A}$ and $\mathfrak{p}$ (apart from those which yields information towards the most appropriate choices for X and ω as introduced above).

For the purposes of the method we also require to fulfill the condition

$$(\Omega_1) \qquad 0 \leq \frac{\omega(p)}{p} \leq 1 - \frac{1}{A_1} \text{ with some constant } A_1 \geq 1 \tag{9.16$_a$}$$

or equivalently

$$(\Omega_1) \qquad 1 \leq \frac{1}{1 - \frac{\omega(p)}{p}} \leq A_1 \text{ with some constant } A_1 \geq 1. \tag{9.16$_b$}$$

We introduce further

$$g(d) := \prod_{p|d} \frac{\omega(p)}{p - \omega(p)} \ \forall d \in \mathbb{N} \quad \text{with} \quad \mu(d) \neq 0, \tag{9.17}$$

which is well-defined (because of (Ω_1)). The product $g(d)$ reminds us of the function occurring in Theorem 7.1, but the advantage here is that $\omega(p)$ need no longer be integer-valued. On the other hand, when compared with Theorem 7.2, the condition (Ω_1) already prevents us from dealing with 'too large a sieve'

By (9.12) we see that

$$g(d) = 0 \quad \text{if} \quad (d, \overline{\mathfrak{p}}) \neq 1. \tag{9.18}$$

and also (from (9.13)) that

$$g(d) = 0 \Longleftrightarrow \omega(d) = 0. \tag{9.19}$$

(Here and in what follows $(d, \overline{\mathfrak{p}}) = 1$ means that no $p \in \overline{\mathfrak{p}}$ divides d.)

105 Finally, we put

$$W(z) := \prod_{p<z}(1 - \frac{\omega(p)}{p}). \tag{9.20}$$

$$G(z) := \sum_{d<z} \mu^2(d) g(d). \tag{9.21}$$

and more generally

$$G_k(x) := \sum_{\substack{d<x \\ (d,k)=1}} \mu^2(d) g(d), \quad 0 < x \in \mathbb{R}, k \in \mathbb{N}. \tag{9.22}$$

In view of (9.6) we could start with the identity (cf. (1.20))

$$S(\mathscr{A}, \mathfrak{p}, z) = \sum_{a \in \mathscr{A}} \sum_{d \mid (a, P(z))} \mu(d); \tag{9.23}$$

in fact, this is the sieve formula of Eratosthenes-Legendre. Selberg's sieve, for obtaining an upper bound for $S(\mathscr{A}, \mathfrak{p}, z)$, consists in the introduction of arbitrary real numbers λ_d with the only condition

$$\lambda_1 = 1. \tag{9.24}$$

which already implies that

$$S(\mathscr{A}, \mathfrak{p}, z) \leq \sum_{a \in \mathscr{A}} (\sum_{\substack{d \mid x \\ d \mid P(z)}} \lambda_d)^2 = \sum_{\substack{d_\nu \mid P(z) x \\ \nu=1,2}} \lambda_{d_1} \lambda_{d_2} \sum_{\substack{a \in \mathscr{A} \\ a = 0 \bmod [d_1, d_2]}} \tag{9.25}$$

Hence, by (9.7) and (9.15),

$$\begin{cases} S(\mathscr{A}, \mathfrak{p}, z) \leq X \sum\limits_{\substack{d_\nu \mid P(z) \\ \nu=1,2}} \lambda_{d_1} \lambda_{d_2} \frac{\omega([d_1, d_2])}{[d_1, d_2]} + \sum\limits_{\substack{d_\nu \mid P(z) \\ \nu=1,2 [d_1, d_2]}} |\lambda_{d_1} \lambda_{d_2} R_{[d_1, d_2]}| = X \sum\limits_1 + \sum\limits_2 \\ \end{cases} \tag{9.26}$$

say. With a view to keep $\sum_2$ small one takes in this method

$$\lambda_d = 0 \quad \text{for} \quad d \geq z \tag{9.27}$$

and then the remaining $\lambda'_d s(2 \leq d < zd|P(z))$ are chosen so as to minimize $\sum_1$

This leads to the choice **106**

$$\lambda_d = \mu(d) \prod_{p|d} \frac{p}{p-\omega(p)} \frac{G_d(\frac{z}{d})}{G(z)}. \tag{9.28}$$

Note here that (9.28) includes both (9.24) and (9.27). Next, it can be shown by the argument of (3.26) that

$$|\lambda_d| \leq 1. \tag{9.29}$$

Now, with the choice (9.28), one obtains

$$\sum_1 = \frac{1}{G(z)} \tag{9.30}$$

and further (9.27) and (9.29) give

$$\sum_2 \leq \sum_{\substack{d_\nu < z \\ d_\nu | P(z) \\ \nu = 1,2}} |R_{[d_1, d_2]}|. \tag{9.31}$$

Here the numbers $d = [d_1, d_2]$ are $< z^2$ and divide $P(z)$. Since d is square free, the number of terms with the same d is atmost

$$|\{d_1, d_2 : [d_1, d_2] = d\} = 3^{\nu(d)} \tag{9.32}$$

From (9.26), (9.30) and (9.31) we now obtain (in view of (9.32))

Theorem 9.1. $(\Omega_1)^1$*: We have, in the above notation* [1].

$$\begin{cases} S(\mathscr{A}, \mathfrak{p}, z) \leq \frac{X}{G(z)} + \sum\limits_{\substack{d_\nu z \\ d_\nu | P(z) \\ \nu = 1,2}} |R_{[d_1,d_2]}| \leq \frac{X}{G(z)} \\ + \sum\limits_{\substack{d < z^2 \\ d | P(z)}} 3^{\nu(d)} |R_d| \leq \frac{X}{G(z)} + \sum\limits_{\substack{d \leq z^2 \\ (d,, \bar{\mathfrak{p}}) = 1}} \mu^2(d) 3^{\nu(d)} |R_d|. \end{cases} \tag{9.33}$$

---

[1]By this notation, which is also employed in a similar way later, we mean that the subsequent statement is valid subject to the conditions in parentheses.

107 Now we give two important special cases of Theorem 9.1.

Theorem 9.2. *Suppose that*

$$\omega(d) = 1 \text{ and } |R_d| \le 1, \text{ if } \mu(d) \ne 0 \text{ and } (d, \overline{\mathfrak{p}}) = 1. \tag{9.34}$$

Then

$$S(\mathscr{A}, \mathfrak{p}, z) \le \frac{X}{(\prod\limits_{\substack{p<z\\ p\in\mathfrak{p}}}(1-\frac{1}{p}))\log z} + z^2. \tag{9.35}$$

Proof. From our assumption on ω in (9.34) it follows that the condition (Ω_1) is fulfilled and further

$$G(z) = \sum_{\substack{d<z\\ (d,k)=1}} \frac{\mu^2(d)}{\varphi(d)}, \tag{9.36}$$

where

$$k = \prod_{\substack{p<z\\ p\notin\mathfrak{p}}} p. \tag{9.37}$$

Therefore, by (3.26), we have

$$G(z) \ge \frac{\varphi(k)}{k}\log z = \prod_{\substack{p<z\\ p\notin\mathfrak{p}}}(1-\frac{1}{p})\log z, \tag{9.38}$$

so that first inequality of (9.33) yields (9.35) (since $|R_d| \le 1$ by (9.34)).

Let us set

$$\mathfrak{p}_K = \{p : p \nmid K\},\ K \in \mathbb{Z} \tag{9.39}$$

□

Theorem 9.3. *Let* $K(\ne 0)$ *be an even integer and suppose that*

$$\omega(p) = \frac{p}{p-1} \text{ for } p \in \mathfrak{p}_K. \tag{9.40}$$

Then, we have

$$S(\mathscr{A}, \mathfrak{p}_K, z) \le \mathfrak{S}(K)\frac{X}{\log z}(1 + O(\frac{1}{\log z})) + \prod_{\substack{p<z^2\\ (d,K)=1}} \mu^2(d)3^{\nu d}|R_d|, \tag{9.41}$$

where

$$\mathfrak{S}(K) = 2\prod_{p>2}(1-\frac{1}{(p-1)^2})\prod_{2<p|K}\frac{p-1}{p-2} \tag{9.42}$$

Proof. We use the last bound given in (9.33). Clearly, we need only show that

$$\frac{1}{G(z)} \leq 2\prod_{p>2}(1-\frac{1}{(p-1)^2})\prod_{2<p|K}\frac{p-1}{p-2}\frac{1}{\log z}(1+0(\frac{1}{\log z})), \tag{9.43}$$

since (Ω_1) is satisfied with $A_1 = 2$ in view of $2|K$. we note that **108**

$$g(p) = \frac{1}{p-2} = \frac{1}{p-1}(1+\frac{1}{p-2}) = \frac{1}{\varphi(p)}(1+g(p)) \text{ for } p \nmid K, \tag{9.44}$$

and so (cf. (9.17))

$$g(d) = \frac{1}{\varphi(p)}\sum_{l|d}\mu^2(l)g(l) \text{ if } (d,K)=1. \tag{9.45}$$

Then, by (3.26).

$$G(z) = \sum_{\substack{\ell<z\\(\ell,K)=1}}\frac{\mu^2(l)g(l)}{\varphi(\ell)}\sum_{\substack{m<\frac{z}{\ell}\\(m,lK)=1}}\frac{\mu^2(m)}{\varphi(m)} \geq \prod_{p|K}(1-\frac{1}{p})\prod_{\substack{\ell=1\\(\ell,K)=1}}^{\infty}\frac{\mu^2(\ell)g(\ell)}{(l)}\log(\frac{z}{\ell}) \tag{9.46}$$

on observing that for $\ell > z$ the lower bound log $(\frac{z}{l})$, for the (empty) inner sum, is negative, Further, by (9.44) and (1.18).

$$\sum_{\substack{l=1\\(l,K)=1}}^{\infty}\frac{\mu^2(\ell)g(\ell)}{(\ell)} = \prod_{p\nmid K}(1+\frac{1}{p(p-2)}) \tag{9.47}$$

and

$$\begin{cases}\prod_{\substack{\ell=1\\(\ell,K)=1}}^{\infty}\frac{\mu^2(l)g(l)}{(\ell)}\log\ell = \prod_{\substack{\ell=1\\(\ell,K)=1}}^{\infty}\frac{\mu^2(l)g(\ell)}{(=\ell)}\sum_{p|\ell}\log p\\ = \sum_{p\nmid K}\frac{\log p}{p(p-2)}\frac{1}{1+(p(p-2))^{-1}}\prod_{p'\nmid K}(1+\frac{1}{p'(p'-2)}).\end{cases} \tag{9.48}$$

Using (9.47) and (9.48) in (9.46) we obtain

$$G(z) \geq \prod_{p|K}(1-\frac{1}{p})\prod_{p\nmid K}(1+\frac{1}{p(p-2)})\{\log z - \sum_{p}\frac{\log p}{(p-1)^2}\}, \tag{9.49}$$

which upholds (9.43). Thus the theorem is completely proved. □

NOTES

Selberg's sieve occurs for the first time in Selberg [1] (cf. Selberg [3], [4], [5], [6]).

For the content of this chapter we refer the reader to Halberstam and
109 Richert [1] (Chapter 3).

(9.29): The details leading to (9.29) are the following. In view of (9.28) consider (for only squarefree d's)

$$\begin{cases} \prod_{p|d}\frac{p}{p-\omega(p)}G_d(\frac{z}{d}) = (\prod_{p|d}(1+g(p))(\sum_{\substack{d_1<z/d\\(d_1,d)=1}}\mu^2(d_1)g(d_1)) \\ = (\sum_{d_2|d}\mu^2(d_2)g(d_2))((\sum_{\substack{d_1<z/d\\(d_1,d)=1}}\mu^2(d_1)g(d_1)). \end{cases} \tag{9.50}$$

Now multiplying out the last expression and comparing with $G(z)$ we obtain (9.29)

9.1: cf. Halberstam and Richert [1] (Theorem 3.2)

The observation due to Kobayashi, which we have mentioned earlier (cf. notes for Chapters 7 and 8), consists in noticing

$$\begin{cases} G(z)\sum_{\substack{d|F(n)\\d|P(z)}}\lambda_d = \sum_{q<z}\sideset{}{'}\sum_{l=1}^{q} b_{q,l}e(-n\frac{l}{q}) \\ \qquad b_{q,l} = \frac{1}{q}\prod_{p|q}(1-\frac{\rho(p)}{p})^{-1}\sum_{\substack{h=1\\(F(h),q)=1}}^{q} e(\frac{lh}{q}). \end{cases} \tag{9.51}$$

and

$$\mu^2(q)g(q) = \sideset{}{'}\sum_{l=1}^{q}|b_{q,l}|^2. \tag{9.52}$$

and using the duality principle (cf. (2.47)–(2.48)). Here F denotes an integer-valued polynomial and $\rho(p)$ is the number of solutions of $F(n) \equiv 0 \bmod p$. Actually, Kobayashi [1] proves the following dual form of the large sieve (in our notation)

$$\sum_{M<n\leq M+N} |\sum_{r=1}^{R} a_r e(-nx_r)|^2 = (N + O(\delta^{-1}))|\sum_{r=1}^{R} a_r|^2, \forall a_r \in \mathbb{C}, \quad (9.53)$$

using the upper bound form of the large sieve as well as a smoothing technique. of Bombieri [4], for a lower bound. (The O-constant in (9.53) is absolute.) From this he derives Selberg's sieve (cf. Mathews **110**
[3]), from which on using (2.90) one obtains, instead of (9.33),

$$S(\mathscr{A}, \mathfrak{p}, z) \leq \frac{X + z^2}{G(z)}. \quad (9.54)$$

(cf. Halberstam and Richert [1] (pp. 125-126)) a result that should be compared with (7.6). It is also possible similarly to get the stronger form (7.8) by defining λ_d's in Selberg's sieve in a different way (cf. Halberstam and Richert [1] (p. 126)).

(9.33): For the second inequality in (9.33) we have used that from $d_\nu|P(z)$, $\nu = 1, 2$, one has $d := [d_1, d_2]|P(z)$ and also for each such d

$$|\{d_1, d_2 : d_\nu < d_\nu|P(z), [d_1, d_2] = d\}| \leq |\{d_1, d_2 : [d_1, d_2] = d\} = 3^{\nu(d)} \quad (9.55)$$

9.2: cf. Halberstam and Richert [1] (Theorem 3.3)

9.3: cf. Halberstam and Richert [1] (Theorem 3.10)

Chapter 10

Some Applications of the Small Sieve in the Case $\omega(p) = \dfrac{p}{p-1}$

AS ALREADY mentioned at the beginning of Chapter 9 we present here some applications of the results obtained there in the chapter for our purposes. Theorem 9.3, which is not contained in the arithmetical form of the large sieve (cf. the remark made subsequent to (9.17)), is of particular interest. **111**

We have the following two interesting (cf. Notes) applications, which we shall quote without any details of proof.

Theorem 10.1. *We have, as $N \to \infty$,*

$$|\{p : p \leq N, p + h = p'\} \leq 4\mathfrak{S}(h)\frac{N}{\log^2 N}\{1 + O(\frac{\log\log N}{\log N})\},$$
$$\forall h \in \mathbb{Z}, h \neq 0,\ \ h \equiv 0\, mod\, 2. \tag{10.1}$$

uniformly in h, and also

$$|\{p : p \leq N, N - p = p'\} \leq 4\mathfrak{S}(h)\frac{N}{\log^2 N}\{1 + O(\frac{\log\log N}{\log N})\},$$

$$for\ N \equiv 0\, mod\, 2. \tag{10.2}$$

where $\mathfrak{S}$ is defined through (9.42).

More generally one also has

Theorem 10.2. *Let $A > 0$ and let a, b, k, ℓ be integers satisfying*

$$ab \neq 0, (a, b) = 1, ab \equiv 0\, mod\, 2. \tag{10.3}$$

and

$$(k, \ell) = 1, 1 \leq k \leq \log^A x (\mathbb{R} \ni x \geq x_0). \tag{10.4}$$

Then we have, uniformly in a, b, k and ℓ, as $x \to \infty$

$$|\{p : p \leq x, p \equiv \ell \bmod k, ap+b = p'\}| \leq 4\mathfrak{S}(abk)\frac{x}{\varphi(k)\log^2 x}\{1+O_A(\frac{\log\log x}{\log x})\}. \tag{10.5}$$

A much more delicate application is the next theorem which shall be
112 used in Chapter 13. Though some of the majorizations in the proof of Theorem 10.3 are crude, many others involve rather delicate considerations. Before coming to the formulation of this theorem we shall obtain some useful auxiliary results.

Lemma 10.1. *We have*

$$\sum_{q \leq x} \frac{\mu^2(q)}{q} h^{\nu(q)} \leq_h (\ log\ x + 1)^h \ for\ x \geq 1, h \in \mathbb{N}. \tag{10.6}$$

Proof. Consider

$$\sum_{a \leq x} \frac{\mu^2(q)}{q} h^{\nu(q)} \leq \prod_{p \leq x} (1 + \frac{h}{p}) \leq \prod_{p \leq x} (1 + \frac{1}{p})^h. \tag{10.7}$$

Now (10.6) is apparent in view of the well-known formula due to Mertens,

$$\prod_{p \leq x} (1 - \frac{1}{p}) = \frac{e^{-\gamma}}{\log x} (1 + 0(\frac{1}{\log x})). \tag{10.8}$$

on noting that $(1 + \frac{1}{p}) \leq (1 - \frac{1}{q})^{-1}$. □

Lemma 10.2. *Let $A > 0$ and let $h, k \in \mathbb{N}$. Let $K \le \log^A x$ for sufficiently large x. Set*

$$E(x,d) := \max_{(\ell,d)=1} |\pi(x;d,\ell) - \frac{li\ x}{\varphi(d)}|. \tag{10.9}$$

Then for any $U_1 (> 0)$ there exists a value $C_1 = C_1(U_1, h, A)$ such that

$$\sum_{d < \frac{\sqrt{x}}{k \log^{C_1} x}} \mu^2(d) h^{\nu(d)} E(x, dk) \ll_{U_1,h,A} \frac{x}{\varphi(k) \log^{U_1} x}. \tag{10.10}$$

Proof. By the rough estimate (cf. (10.9))

$$E(x, d_1) \ll \frac{x}{d_1} \quad \text{for} \quad d_1 \le x \tag{10.11}$$

and an application of the Cauchy's inequality followed by extensions of ranges for variables in the resulting summations, we see that the expression on the left-hand side in (10.10) is

$$\ll (\frac{x}{k})^{\frac{1}{2}} (\sum_{d \le x} \frac{\mu^2(d) h^{2\nu(d)}}{d})^{\frac{1}{2}} (\sum_{d < \frac{\sqrt{x}}{\log^{C_1} x}} E(x,d))^{\frac{1}{2}}. \tag{10.12}$$

113

In view of the bound given by (10.6) for the first sum here, we make the choice of C_1, for given U_1, by means of (6.23) such that the second sum is bounded by $x(\log x)^{-(2U_1+h^2+A)}$. Then, continuing the estimation in (10.12). we obtain further

$$\ll_{U_1,h,A} (\frac{x}{k})^{\frac{1}{2}} (\log x)^{\frac{1}{2}h^2} (x(\log x)^{(2U_1+h^2+A})^{\frac{1}{2}} \ll_{U_1,h,A} x(\log x)^{U_1 - \frac{1}{2}A} (\varphi(k))^{-\frac{1}{2}}. \tag{10.13}$$

which proves (10.10) because $\varphi(k) \le k \le (\log x)^A$. □

Now we are in a position to prove the main theorem of this chapter.

Theorem 10.3. *Let v be a real number satisfying*

$$v > 3. \tag{10.14}$$

Let $h(\neq 0)$ satisfy, being determined with respect to sufficiently large x.

$$h \in \mathbb{Z}, h \equiv 0 \bmod 2, \quad \text{and either} \quad h = [x] \quad \text{or} \quad 0 < |h| \leq x^{1/3}. \tag{10.15}$$

Then, as $x \to \infty$, holds

$$\begin{cases} C_v(x,h) := |\{|h - p_1p_2p_3| : |h - p_1p_2p_3| = p, x^{1/v} \\ \leq p_1 < x^{1/3} \leq p_2 < p_3, p_1p_2p_3 \leq x\}| \leq \\ \leq 4c(v)\mathfrak{S}(h)\frac{x}{\log^2 x}\left\{1 + O_v(\frac{\log\log x}{\log x})\right\}. \end{cases} \tag{10.16}$$

where $\mathfrak{S}$ is defined by (9.42) and

$$c(v) := \int_{1.v}^{1/3} \frac{\log(2-3\alpha)}{\alpha(1-\alpha)} d\alpha \tag{10.17}$$

(Note that the O-constant in (10.16) depends at most on v.)

Proof. Let us consider (the finite sequence)

$$\mathscr{B} := \{b : b = pd, d \in \mathscr{D}, p \leq \frac{x}{d}\}. \tag{10.18}$$

114 where

$$\mathscr{D} := \left\{d : d = p_1p_2, x^{1/v} \leq p_1 < x^{1/3} \leq p_2 \leq \sqrt{\frac{x}{p_1}}\right\}. \tag{10.19}$$

We note that each $d \in \mathscr{D}$ (has a unique representation as p_1p_2 and) satisfies

$$x^{\frac{1}{3}} < x^{\frac{1}{3}+\frac{1}{v}} \leq d = p_1p_2 < \sqrt{p_1x} < x^{\frac{2}{3}}. \tag{10.20}$$

and so

$$|\mathscr{D}| < x^{\frac{3}{2}}. \tag{10.21}$$

With a view to determine a suitable approximation to $|\mathscr{B}|$, we use the formula (taking $\rho = \frac{1}{v}$ and $\rho = \frac{1}{3}$)

$$\sum_{x^\rho \leq p < y} \frac{1}{p} = \log(\frac{\log y}{\rho\log x}) + O(\frac{1}{\rho\log x}) \text{ for } y \geq x^\rho, \tag{10.22}$$

for Stieltjes integration to obtain (with $\rho = \dfrac{1}{3}$)

$$\sum_{x^{\frac{1}{3}} \le p_2 < \sqrt{\frac{x}{p_1}}} \frac{1}{p_2 \log \frac{x/p_1}{p_2}} = \int\limits_{\sqrt{x^{1/3}}}^{x/p_1} \frac{d\eta}{\eta \log \eta \log \frac{x/p_1}{\eta}} + \circ(\frac{1}{\log^2 x}). \qquad (10.23)$$

Further multiplying (10.23) by $\dfrac{1}{p_1}$ and summing over $x^{1/v} \le p_1 < x^{1/3}$ we get, by use of (10.22) (with $\rho = 1/v$),

$$\begin{cases} \sum\limits_{d\in\mathscr{D}} \frac{1}{d \log \frac{x}{d}} & = \int\limits_{x^{1/v}}^{x^{1/3}} \frac{d\xi}{\xi \log \xi} \int\limits_{x^{1/3}}^{(\frac{x}{\xi})^{1/2}} \frac{d\eta}{\eta \log \eta \log(\frac{x}{\xi\eta})} + O_v(\frac{1}{\log^2 x}) = \\ & = \frac{1}{\log x} \int\limits_{1/v}^{1/3} \frac{d\alpha}{\alpha} \int\limits_{1/3}^{(1-\alpha)/2} \frac{d\beta}{\beta(1-\alpha-\beta)} + O_v(\frac{1}{\log^2 x}) = \\ & = \frac{1}{\log x} \int\limits_{\frac{1}{v}}^{\frac{1}{3}} \frac{\log(2-3\alpha)}{\alpha(1-\alpha)} d\alpha + O_v(\frac{1}{\log^2 x}) = \frac{c(v)}{\log x} + O_v(\frac{1}{\log^2 x}), \end{cases} \qquad (10.24)$$

where we have put

$$\xi = x^{\alpha}, \eta = x^{\beta}. \qquad (10.25)$$

and have also used the notation (10.17). Thus since every $d \in \mathscr{D}$ is $< x^{2/3}$ (cf. (10.20)), by (10.18) and the prime-number theorem (in a weak from) one has 115

$$|\mathscr{B}| = \sum_{d\in\mathscr{D}} \sum_{p \le \frac{x}{d}} 1 = \sum_{d\in\mathscr{D}} \frac{x}{d \log(\frac{x}{d})} (1 + O(\frac{1}{\log x})) = \frac{c(v)x}{\log x} (1 + O_v(\frac{1}{\log x})), \qquad (10.26)$$

on using (10.24). (Note that the formula (10.26) counts the numbers in $\mathscr{B}$ according to the multiplicity of their occurrence.) □

Towards the estimate (10.17), we naturally consider

$$S(\mathscr{A}, \mathfrak{p}_h, z), \qquad (10.27)$$

where

$$\mathscr{A} := \left\{ |h - p_1 p_2 p_3| : x^{1/v} \le p_1 < x^{1/3} \le p_2 < p_3, p_1, p_2, p_3 \le x \right\}, \tag{10.28}$$

$$\mathfrak{p}_h := \{p : p \nmid h\}, \tag{10.29}$$

and (we make the choice)

$$z^2 = x^{\frac{1}{2}} \ell^{-18} \tag{10.30}$$

with the abbreviation

$$\ell : L = \log x. \tag{10.31}$$

Instead of $\mathscr{A}$ it is more convenient (cf. (10.60)) to work with

$$\mathscr{A}^* := \{|h - p_1 p_2 p_3| : x^{1/v} \le p_1 < x^{1/3} \le p_2 < \sqrt{\frac{x}{p_1}}, p_1 p_2 p_3 \le x\}. \tag{10.32}$$

Now the primes $p_i (i = 1, 2, 3)$ occurring in $\mathscr{A}$ satisfy $p_1 p_2^2 \le x$ and so $p_2 < \sqrt{\frac{x}{p_1}}$, which shows that $\mathscr{A}$ is contained in $\mathscr{A}^*$ (even with regard to multiplicity of numbers in it). Therefore

$$S(\mathscr{A}, \mathfrak{p}_h, z) \le S(\mathscr{A}^*, \mathfrak{p}_h, z). \tag{10.33}$$

Note also that all the elements of $\mathscr{A}^*$ have the same type of repre-
116 sentation from among

$$h - p_1 p_2 p_3, p_1 p_2 p_3 - h \text{ or } |h| + p_1 p_2 p_3 \tag{10.34}$$

according as $h = [x]$, $|h| \le x^{1/3}$ with $h > 0$ or $h < 0$ respectively (cf. (10.15)) and hence, in particular, that the multiplicity of a number in $\mathscr{A}^*$ is exactly the multiplicity of the corresponding $p_1 p_2 p_3$ in $\mathscr{B}$.

Next, we prepare for an application of Theorem 9.3 with respect $\mathscr{A}^*$, $\mathfrak{p}_h$ and z. Comparing the definitions of $\mathscr{A}^*$ and $\mathscr{B}$ (through $\mathscr{D}$) we see that, for $(q, h) = 1$ with $\mu(q) \ne 0$, by (1.53) (in the notation of Chapter 9) one has

$$\begin{cases} |\mathscr{A}^*_q| & = \sum\limits_{\substack{b \in \mathscr{B} \\ b \equiv h \bmod q}} 1 = \frac{1}{\varphi(q)} \sum\limits_{\chi \bmod q} \bar{\chi}(h) \sum\limits_{b \in \mathscr{B}} \chi(b) = \\ & = \frac{1}{\varphi(q)} \sum\limits_{b \in \mathscr{B}} \chi_0(b) + \frac{1}{\varphi(q)} \sum\limits_{\substack{\chi \bmod q \\ \chi \ne \chi_0}} \bar{\chi}(h) \sum\limits_{b \in \varphi} \chi(b), \end{cases} \tag{10.35}$$

in view of the remark involving (10.34). From here we get

$$|R_q| = \left||A_q^*| - \frac{1}{\varphi(q)}|\mathscr{B}|\right| \leq \frac{1}{\varphi(q)} \sum_{\substack{\chi \bmod q \\ \chi \neq \chi_0}} |\sum_{b \in \mathscr{B}} \chi(b)| + \frac{1}{\varphi(q)} \sum_{\substack{b \in \mathscr{B} \\ (b.q)>1}} 1 =: R_q^*, \tag{10.36}$$

say. In particular

$$|\mathscr{A}^*| = |\mathscr{A}_1^*| = |\mathscr{B}|, \tag{10.37}$$

which is clear otherwise also. Therefore we make the choices

$$\omega(p) = \frac{p}{p-1} \text{for} p \in \mathfrak{p}_h \tag{10.38}$$

and, by (10.26),

$$X = |\mathscr{B}| = c(v)\frac{x}{\log x}(1 + O_v(\frac{1}{\log x})). \tag{10.39}$$

Now, $S(\mathscr{A}, \mathfrak{p}_h, z)$ counts all numbers of (the set in) (10.16) which exceed z and so (because of (10.33)) applying Theorem 9.3 with $K = h$ for $\mathscr{A}^*$ it follows, by (9.41), (10.39) and (10.30). **117**

$$C_v(x, h) \leq 4c(v)\mathscr{B}(h)\frac{x}{\log^2 x}\{1 + O_v(\frac{\log\log x}{\log x})\} + \sum_{\circ}, \tag{10.40}$$

where (cf. (10.36))

$$\sum_{\circ} := z + \sum_{\substack{q<z^2 \\ (q,h)=1}} \mu^2(q)3^{\nu(q)}|R_q|. \tag{10.41}$$

The rest of the proof concerns with an estimation of $\sum_0$ which shows this contribution to (10.40) as being of the nature of an error-term. The sum in (10.41) is, by Cauchy's inequality and (10.36),

$$\leq (\sum_{\substack{q<z^2 \\ (q,h)=1}} \mu^2(q)9^{\nu(q)}|R_q|)^{\frac{1}{2}}(\sum_{\substack{q<z^2 \\ (q,h)=1}} \mu^2(q)R_q^*)^{\frac{1}{2}}. \tag{10.42}$$

Now trivially, from (10.35), (10.26), we have (cf. (10.15), cf. also (10.11))

$$|R_q| \ll_\nu \left(\frac{x + x^{1/3}}{q} + \frac{x}{\varphi(q)\log x}\right) \ll_\nu \frac{x}{q} \tag{10.43}$$

so that the first sum in (10.42) is, by Lemma 10.1,

$$= O_\nu(x \log^9 x). \tag{10.44}$$

(Observe that the first term of the middle expression in (10.43) is obtained by using the fact that the multiplicity of any member of $\mathscr{A}^*$ is absolutely bounded (cf. (10.34)).) Next we deal with the simpler part, of the second sum in (10.42), arising from the second term defining R_q^*. We have for $q < x$ with q squarefree,

$$\begin{cases} \sum\limits_{\substack{b \in \mathscr{B} \\ (b,q)>1}} 1 \leq \sum\limits_{\substack{d \in \mathscr{D} \\ p \mid q}} 1 + \sum\limits_{\substack{x^{1/\nu} \leq p_1, p_2 < x^{\frac{1}{2}} \\ p_1 \mid q}} \sum\limits_{p < \frac{x}{p_1 p_2}} 1 \leq \nu(q)(|\mathscr{D}| + x^{1-\frac{1}{\nu}} \sum\limits_{x^{1/\nu} \leq p_2 < x^{\frac{1}{2}}} \frac{1}{p_2}) \\ \ll_\nu x^{1-1/\nu} \log x, \end{cases} \tag{10.45}$$

on using (10.21), (10.22), (10.15) and the trivial estimate

$$\nu(q) \leq \frac{\log q}{\log 2} \ \forall q \in \mathbb{N}. \tag{10.46}$$

118

Hence (10.45) leads to the estimate, for the part of the sum in (10.42) which in under consideration via (8.32)

$$\ll_\nu x^{1-\frac{1}{\nu}} \log x \sum_{d < z^2} \frac{\mu^2(q)}{\varphi(q)} \ll_\nu x^{1-\frac{1}{\nu}} \log^2 x. \tag{10.47}$$

Now collecting together the bounds (10.42), (10.44) and (10.47) we see that because of the choice (10.30).

$$\sum_0^2 \ll_\nu z^2 + x \log^9 x (x^{1-\frac{1}{\nu}} \log^2 x + \sum_1) \ll_\nu x^{2-\frac{1}{\nu}} \log^{11} x + x \log^9 x. \sum_1, \tag{10.48}$$

where (cf. (10.36))

$$\sum_i := \sum_{q<z^2} \frac{\mu^2(q)}{\varphi(q)} \sum_{\substack{\chi \bmod q \\ \chi \neq \chi_0}} |\sum_{b\in\mathscr{B}} \chi(b)|. \tag{10.49}$$

Observe that h has no longer a part to play in the sequel.

Transition to primitive characters in (10.49) yields (cf. (1.59)), on writing $q = rf$ (f: conductor $\chi \bmod q$)

$$\begin{cases} \sum_1 = \sum_{r<z^2} \frac{\mu^2(r)}{\varphi(r)} \sum_{\substack{f<\frac{z^2}{r} \\ (r,f)=1}} \frac{\mu^2(f)}{\varphi(f)} \sideset{}{^*}\sum_{\substack{\chi \bmod f \\ \chi\neq\chi_0}} | \sum_{\substack{b\in\mathscr{B} \\ (b,r)=1}} \chi(b)| \leq \\ \sum_{r<z^2} \frac{\mu^2(r)}{\varphi(r)} \sum_{f<z^2} \frac{\mu^2(f)}{\varphi(f)} \sideset{}{^*}\sum_{\substack{\chi \bmod f \\ \chi\neq\chi_0}} | \sum_{\substack{b\in\mathscr{B} \\ (b,r)=1}} \chi(b)| = \sum_{r<z^2} \frac{\mu^2(r)}{\varphi(r)} \sum_2(r), \end{cases} \tag{10.50}$$

where (on replacing f by q)

$$\sum_2(r) := \sum_{q<z^2} \frac{\mu^2(q)}{\varphi(q)} \sum_{\substack{\chi \bmod q \\ \chi\neq\chi_0}} | \sum_{\substack{b\in\mathscr{B} \\ (b,r)=1}} \chi(b)|. \tag{10.51}$$

(In the remaining part we also use the abbreviation (10.31) wherever convenient.)

By the Siegel-Walfisz theorem (cf. (8.2)) we have for any character $\chi \neq \chi_0 \bmod q$, in view of (1.44) and (10.46), **119**

$$\Big|\sum_{\substack{p<y \\ p\nmid r}} \chi(p)\Big| \leq \frac{\log r}{\log 2} + \Big|\sum_{\ell=1}^{q} \chi(\ell)\pi(y;q,\ell)\Big| \ll \frac{\log r}{\log 2} + y.\varphi(q)\ell^{-3g}$$

$$\text{uniformly for } q \ll \log^g y \text{ as } y \to \infty, \tag{10.52}$$

so that (with $g = 17$)

$$\begin{cases} \sum_{q\leq\ell^{17}} \frac{\mu^2(q)}{\varphi(q)} \sum_{\substack{\chi \bmod q \\ \chi\neq\chi_0}} \Big| \sum_{\substack{b\in\mathscr{B} \\ (b,r)=1}} \chi(b) \Big| = \sum_{q\leq\ell^{17}} \frac{\mu^2(q)}{\varphi(q)} \sideset{}{^*}\sum_{\substack{\chi \bmod q \\ \chi\neq\chi_0}} \Big| \sum_{\substack{d\in\mathscr{D} \\ (d,r)=1}} \chi(d) \sum_{\substack{p\leq x/d_d \\ p\nmid r}} \chi(p)| \leq \\ \ll \sum_{q\leq\ell^{17}} \sum_{d\in\mathscr{D}} \frac{x}{d}\varphi(q)\ell^{-51} \ll_v x\ell^{-16}. \end{cases} \tag{10.53}$$

Hence

$$\sum_2(r) \ll_v \sum_3(r) + x\ell^{-16}, \tag{10.54}$$

where

$$\sum_3(r) := \sum_{\ell^{17}<q<z^2} \frac{\mu^2(q)}{\varphi(q)} \sideset{}{^*}\sum_{\chi \bmod q} \Bigg| \sum_{\substack{b\in\mathscr{B}\\(b,r)=1}} \chi(b)\Bigg|. \tag{10.55}$$

For an estimation of (10.55) we use contour integration and the hybrid sieve. To this and we put

$$T = x^3, \tag{10.56}$$

as well as

$$a = 1 + \ell^{-1}. \tag{10.57}$$

Further let

$$\ell^{17} < w \leq z^2 \tag{10.58}$$

and also note that the *supposition*

$$x = [x] + \frac{1}{2} \tag{10.59}$$

involves no loss of generality. We also introduce the Dirichlet series

$$P = P_r(s,\chi) := \sum_{\substack{p\leq w^2\\p\nmid r}} \frac{\chi(p)}{p^s}, Q = Q_r(s,\chi) := \sum_{\substack{p>w^2\\p\nmid r}} \frac{\chi(p)}{p^s},$$

$$D = D_r(s,\chi) := \sum_{\substack{d\in\mathscr{D}\\(d,r)=1}} \frac{\chi(d)}{d^s} (Res > 1). \tag{10.60}$$

120

Now, by (10.19) and a well-known formula (using (10.59))

$$\sum_{\substack{b\in\mathscr{B}\\(b,r)=1}} \chi(d) = \frac{1}{2\pi i} \int_{a-iT}^{a+iT} (P+Q)D(s)\frac{x^s}{s} ds + O(\frac{x\log x}{T}). \tag{10.61}$$

Splitting the integral here into two parts, corresponding to $(PD)(s)$ and $(QD)(s)$ respectively, and shifting the line of integration of the former only to $\sigma = \dfrac{1}{2}$ we obtain, on using (10.57).

$$\begin{cases} \sum\limits_{\substack{b \in \mathscr{B} \\ (b,r)=1}} \chi(b) \ll x^{\frac{1}{2}} \int\limits_{-T}^{T} |(PD)(\tfrac{1}{2} + it, \chi)| \frac{dt}{1+|t|} + \frac{x}{T} \int\limits_{a}^{1/2} |(PD)(\sigma \pm iT, \chi)| d\sigma + \\ +x \int\limits_{-T}^{T} |(QD)(a + it, \chi)| \frac{dt}{1+|t|} + \frac{x \log x}{T}. \end{cases} \tag{10.62}$$

For the second term on the right-hand side the crude estimate

$$|(PD)(\sigma \pm it, \chi)| \leq \sum_{p \leq w^2} \sum_{d \in \mathscr{D}} 1 \leq w^2 |\mathscr{D}| \leq x^2 \text{for} \frac{1}{2} \leq \sigma \leq a, \tag{10.63}$$

obtained from (10.61), (10.59), (10.31) and (10.22), suffices.

Towards an estimation of the remaining terms we introduce a notation (for convenience of description). For an arbitrary function $f(s, \chi)$, $s = \sigma + it$ we set (with respect to r, w, T and z as before)

$$M(\sigma, f) := \sum_{\substack{q < w \\ (q,r)=1}} \frac{q}{\varphi(q)} \sideset{}{^*}\sum_{\chi \bmod q} \int_{-T}^{T} |f(\sigma + it, \chi)| \frac{dt}{1 + |t|}, \tag{10.64}$$

and observe that, by Cauchy-Schwarz inequality, for any two o(such) functions I_1 and f_2 one has

$$M(\sigma, f_1 f_2) \leq (M(\sigma, f_1^2))^{\frac{1}{2}} (M(\sigma, f_2^2))^{-\frac{1}{2}}. \tag{10.65}$$

We would also need the estimate (valid with an absolute $\ll$-constant)

$$M(\sigma, f^2) \ll \sum_{n=1}^{\infty} (w^2 \log x + n) |a_n|^2 n^{-2\sigma}, \text{for} f = f(s, \chi) = \sum_{n=1}^{\infty} \frac{a_n}{n^s} \chi(n) \text{ if } \sum_{n=1}^{\infty} \frac{|a_n|}{n^\sigma} < \infty, \tag{10.66}$$

which is obtained by partial integration in Theorem 5.1 and using (10.56). **121**

By (10.62), (10.64), (10.55) and (10.56) there holds

$$\sum_{\ell^{17}<q<w} \frac{q}{\varphi(q)} \sideset{}{^*}\sum_{\chi \bmod q} |\sum_{\substack{b\in\mathscr{B} \\ (b,r)=1}} \chi(b)| \ll (xM(\frac{1}{2}, P^2)M(\frac{1}{2}, D^2))^{\frac{1}{2}}$$
$$+x(M(a, Q^2)M(a, D^2))^{\frac{1}{2}} + w^2. \tag{10.67}$$

Now we apply (10.66) to the four M's occurring in (10.67) to obtain in succession (with absolute $\ll$-constants)

$$M(\frac{1}{2}, P^2) \ll \sum_{p\le w^2} (w^2 \log x + p)p^{-1} \ll w^2 \log^2 x, \tag{10.68}$$

because of (10.60), (10.58) and (10.30)),

$$M(\frac{1}{2}, D^2) \ll \sum_{d\in\mathscr{D}} (w^2 \log x + d)d^{-1} \ll w^2 \log^2 x + x^{\frac{2}{3}}, \tag{10.69}$$

using in addition (10.21) further (cf. (10.57))

$$M(a, Q^2) \ll \sum_{p>w^2} (w^2 \log x + p)p^{-2a} \log x, \tag{10.70}$$

and lastly

$$M(a, D^2) \ll \sum_{d\in\mathscr{D}} (w^2 \log x + d)d^{-2a} \ll w^2(\log^2 x)x^{-\frac{1}{3}} + \log x. \tag{10.71}$$

Using the four estimates (10.68)–(10.71) in (10.67) we arrive (cf. (10.58), (10.30))

$$\begin{cases} \sum\limits_{\ell^{17}<q<w} \frac{q}{\varphi(q)} \sideset{}{^*}\sum\limits_{\chi \bmod q} |\sum\limits_{\substack{b\in\mathscr{B} \\ (b,r)=1}} \chi(b)| \ll (xw^2 \log^2 x(w^2 \log^2 x + x^{\frac{2}{3}}))^{\frac{1}{2}} \\ +x(\log^2 x(w^2 x^{-\frac{1}{3}}))^{\frac{1}{2}} + w^2 \ll x^{\frac{1}{2}} w \log x(w \log +x + x^{\frac{1}{3}}) \\ +x \log x(wx^{-\frac{1}{6}} + 1) + w^2 \ll x \log x + wx^5 6 \log x + w^2 x^{\frac{1}{2}} \log^2 x, \end{cases} \tag{10.72}$$

valid uniformly in r. This gives, by partial summation and (10.30),

$$\sum_{\ell^{17}<q<z^2} \frac{1}{\varphi(q)} \sideset{}{^*}\sum_{\chi \bmod q} \Big| \sum_{\substack{b\in\mathscr{B}\\(b,r)=1}} \chi(b) \Big| \ll x\ell^{-16} + x^{\frac{5}{6}}\ell^2 \ll x\ell^{-16} \qquad (10.73)$$

122 uniformly in r.

Note that the left-hand side expression in (10.73) majorizes $\sum_3(r)$ of (10.55). Therefore, it follows from (10.54) that

$$\sum_2 (r) \ll_v x\ell^{-16}, \qquad (10.74)$$

which when used in (10.50) yields (by means of (8.32)) the estimate

$$\sum_1 \ll_v x\ell^{-15}. \qquad (10.75)$$

Thus we obtain, on using (10.75) in (10.48),

$$\sum_0 \ll_v x(\log x)^{-3}, \qquad (10.76)$$

from which in view of (10.40), follows the estimate in (10.16). This completes the proof if Theorem 10.3.

NOTES

10.1: The results of this theorem were proved first under the assumption of the generalized Riemann hypothesis by Wang [2], and the unconditional proof of these results (with an application to have (6.77) with a $\theta < \frac{1}{2}$) is due to Bombieri and Davenport [1]. The main terms of (10.1) and (10.2) are 4 times the conjectured asymptotic formulae (without error terms), for the respective problems, of Hardy and Littlehood. In this context it is significant to mention that Montgomery has pointed out (in correspondence) that a decrease of the factor 4 here any constant $c < 4$ would have the same consequences regarding the Siegel-zeros as has been remarked in connection with the Brun-Titchmarsh theorem (cf. Notes of Chapter 8, under (8.21)).

Theorem 10.2: See Klimov [1], [2], and Halberstam and Richert [1] (Theorem 3.12). 123

(10.8): The formula (10.8) though not needed in this full force for Lemma 10.1, is of much use in later chapters.

Lemma 10.2: cf. Halberstam and Richert [1] (Lemma 3.5)

Theorem 10.3: This result is the most essential part of Chen's proof of his famous theorem with respect to Goldach's conjecture. For $v = 10$ it occurs in Chen [1], Halberstam and Richert [1] (Chapter 11, with some simplifications due to P.M. Ross) (with the weights $\wedge(n)$), and in Ross [1]. We shall use Theorem 10.3 with $v = 8$ in Chapter 13 and thereby obtaining the advantage of dealing with elementary functions in connection with the lower bound estimation via Selberg's sieve. (cf. Notes of Chapter 13, preceding (13.27).)

(10.17): For later use we obtain an estimate for $c(v)$. First, note that the function $f(x) = x^2 - 2x\log x$ is increasing for $x \geq 1$ (since the derivative $f'(x) = 2(x - 1 - \log x) \geq 0$ for $x \geq 1$). So $f(x) \geq f(1)$, for $x \geq 1$, which means that $x^2 - 2x\log x \geq 1$; i.e.,

$$\log x \leq \frac{(x-1)(x+1)}{2x} \text{ for } x \geq 1. \tag{10.77}$$

(However, observe also that (10.77) follows from

$$\int\limits_a^b F(y)dy \leq \frac{(b-a)}{2}(F(b) + F(a)) \text{ for any convex function } F, b \geq a, \tag{10.77$'$}$$

on taking $F(y) = y^{-1}$ and $a = 1$, $b = x(\geq 1)$.) Hence (10.77), with $x = 2 - 3\alpha$, gives

$$\frac{\log(2-3\alpha)}{\alpha(1-\alpha)} \leq \frac{1-3\alpha}{2}\frac{3}{\alpha(2-3\alpha)} = \frac{3}{4}(\frac{1}{\alpha} - \frac{3}{2-3\alpha}) \text{ for } 0 < \alpha \leq \frac{1}{3}, \tag{10.78}$$

from which, we obtain

$$c(v) = \int\limits_{\frac{1}{v}}^{\frac{1}{3}} \frac{\log(2-3\alpha)}{\alpha(1-\alpha)}d\alpha \leq \frac{3}{4}\log(\alpha(2-3\alpha))\Big|_{\alpha=\frac{1}{v}}^{\alpha=\frac{1}{3}} = \frac{3}{4}\log(\frac{v^2}{3(2v-3)}). \tag{10.79}$$

(10.61): Titchmarsh, *E.C. The theory of the Riemann zeta function* (Oxford), **124**

Lemma 3.12.

The effective way of treating the remainder terms by an explicit use of analytical methods. as done here subsequent to (10.61). is of recent origin. It occurs in the papers Barban and Vehow [1] (see Motohashi [10]) Hooley [2], [5]. Huxley [5], Chen [1]. Motohashi [8]. Halberstam and Richert [1]. Goldfeld [4], and Ross [1]. Excepting the first of these paper all the others employ this method for the purpose of some applications only. In that paper Barban and Vehov [1] sketch a proof (- a rigorous proof was given by Motohashi [10]-) of the following surprisingly uniform result:

If $x > z$, $\log z \gg \log z_1$, $z_1 \geq z$ and

$$\lambda_d := \begin{cases} \mu(d) & if d \leq z \\ \mu(d)\frac{\log(\frac{z_i}{d})}{\log(\frac{z_i}{d})} & if z < d \leq z_1. \end{cases} \tag{10.80}$$

then we have

$$\sum_{1 \leq n \leq x} (\sum_{\substack{d|n \\ d \geq z_1}} \lambda_d)^2 \frac{x}{\log(\frac{z_1}{z})}. \tag{10.81}$$

Chapter 11

A Generalized form of Selberg's Sieve

IN CHAPTER 9 we discussed a simple version of Selberg's sieve along 125
with two particular cases, corresponding to the choices

$$\omega(p) = 1 \text{ and } \omega(p) = \frac{P}{p-1}(\text{for} p \in \mathfrak{p}), \tag{11.1}$$

and applied the latter, in the next chapter, to obtain the important Theorem 10.3 for the purposes of Chapter 13. In this chapter we continue the theme of the small sieve of Selberg with a view to enunciating it in its best form (in a certain sense), with is useful also in the proofs of the results of Chapter 13. However, our account of this aspect of the Selberg's sieve here will be sketchy (with relevant references being included in the Notes).

At the out we recall that Selberg's Theorem 9.1 was proved subject to only the condition (Ω_1) (cf. $((9.16_a))$ a) which restriction can also stated in the form

$$0 \leq \omega(p) \leq \delta p \quad \text{for some} \quad \delta < 1. \tag{11.2}$$

(Though (11.2) does not make Selberg's sieve 'too large a sieve', nevertheless it leaves the sieve a 'large' one still.) However, to effectively deal with the function $G(z)$ and the remainder terms of Theorem 9.1 one needs information concerning the average order of magnitude of ω (and also of R_d's). In this connection the condition

$$(\Omega_2(k,L)) - L \leq \sum_{w\leq p<z} \frac{\omega(p)\log p}{p} - k\log(\frac{z}{w}) \leq A_2 \text{for} 2 \leq w \leq z, \quad (11.3)$$

which tells us that ω is 'on the average' equal to κ is useful (when R_d's are 'small' atleast on the average) for obtaining both the upper and lower
126 estimates for $S(\mathscr{A}, \mathfrak{p}, z)$. Then the O-constants in these estimates are allowed to depend on $A_1(\text{from}\Omega_1)$) and k (and hence, comparing with (11.2). the sieve is 'small') A_2 (from $(\Omega_2(k, L))$), (and some constants implicit in the restrictions on R_d's (cf. (11.62))) but not on any other parameter involved (in particular, independent of L, which is of the nature of an error term, in practice thereby requiring a separate consideration).

The constant κ in (11.3) is called the 'dimension' of the sieve problem.

As we shall see later the equation of obtaining lower bounds for our sifting function $S(\mathscr{A}, \mathfrak{p}, z)$ can be linked up, in a significant way with the problem of finding good upper estimates for it. Accordingly, we now deal with the latter problem. First, we mention that when one combines for this purpose, Theorem 9.1 with $\Omega_2(\kappa, L)$, or instead even with the one - sided restriction

$$(\Omega_2(\kappa)) \quad \sum_{w\leq p<z} \frac{\omega(p)\log p}{p} \leq \kappa\log(\frac{z}{w}) + A_2 \quad \text{if} \quad 2 \leq w \leq z, \quad (11.4)$$

the results obtained are quite satisfactory. In fact, an elementary reasoning gives, under the condition $(\Omega_2(k))$ (in addition to (Ω_1)), the estimate

$$\frac{1}{G(z)} \ll W(z) \quad (11.5)$$

for the function $G(z)$ of Theorem 9.1 in terms if $W(z)$ defined by (9.20), and with a little more effort, on using $(\Omega_2(\kappa, L))$ in place of $\Omega_2(\kappa)$ here, one obtains

$$\frac{1}{G(z)} = W(z)e^{\gamma\kappa}\Gamma(\kappa+1)\{1 + O(\frac{L}{logz})\}. \quad (11.6)$$

Thus we get, by means of (the last inequality of) Theorem 9.1 the following two theorems (respectively):

Theorem 11.1. (Ω_1), $(\Omega_2(\kappa))$:

$$S(\mathscr{A}, \mathfrak{p}, z) \ll XW(z) + \sum_{\substack{d \le z^2 \\ (d,\mathfrak{p})=1}} \mu^2(d) 3^{\nu(d)} |R_d| \tag{11.7}$$

and 127

Theorem 11.2. (Ω_1), $(\Omega_2(\kappa, L))$:

$$S(\mathscr{A}, \mathfrak{p}, z) \le XW(z) e^{\gamma\kappa} \Gamma(k+1)\{1 + O(\frac{L}{\log z})\} + \sum_{\substack{d < z^2 \\ (d,\overline{\mathfrak{p}})=1}} \mu^2(d) 3^{\nu(d)} |R_d|. \tag{11.8}$$

Here, again, we point out the O-constants in these two theorems depend atmost on A_1, A_2 and κ inherent in (Ω_1) and $\Omega(\kappa, L)$ (or $\Omega_2(\kappa)$).

Next, in accordance with the remark involving (11.3), we consider the question of the magnitude of R_d's There are many cases in which one has the following information

$$(R) \qquad |R_d| \le \omega(d) \quad \text{if} \quad \mu(d) \ne 0 \text{ and } (d.\overline{\mathfrak{p}}) = 1. \tag{11.9}$$

In such a situation one readily obtains, from (the second inequality of) Theorem 9.1 and (11.5).

Theorem 11.3. (Ω_1), $(\Omega_2(\kappa))$, (R)*: For any* $A > 0$

$$S(\mathscr{A}, \mathfrak{p}, z) \ll X \prod_{p<z} (1 - \frac{\omega(p)}{p}) \text{ if } z \le X^A, \tag{11.10}$$

where the $\ll$*-constant depends almost on* A*,* A_1*,* A_2 *and* κ*.*

In the literature, usually, the phrase 'by Brun's sieve…' refers to the statement (11.10). Here notice that Theorem 11.3 is a available the more convenient condition, instead of $(\Omega_2(\kappa))$,

$$(\Omega_0) \qquad \omega(p) \le A_0 \tag{11.11}$$

(since (Ω_0)) implies $(\Omega_2(\kappa))$). Similar to Theorem 11.3, on using (11.6) with Theorem 9.1 (cf. (10.8)), one has the explicit

Theorem 11.4. (Ω_1), $(\Omega_2(\kappa, L))$, (R)*:*

$$S(\mathscr{A}, \mathfrak{p}, z) \leq \Gamma(\kappa+1)\prod_p\{(1-\frac{\omega(p)}{p})(1-\frac{1}{p})^{-k}\}\frac{X}{\log^{k_z}}$$
$$\{1+O(\frac{\log\log 3z+L}{\log z})\} \ \ if \ \ z \leq X^{\frac{1}{2}}, \tag{11.12}$$

128 *where the infinite product converges and the O-constant depends atmost on A_1, A_2 and κ.*

On the other hand, in absence of the information (R), as happens with more delicate problems, one has to seek atleast an average result about R_d's (cf. the remark containing (11.3)). It is at this stage that the Bombieri-type results (cf. Chapter 6, **3**) are effective. It is easily seen from the last sums in Theorem 11.1 and 11.2 that the size of z^2 (which is about

$$\frac{\sqrt{X}}{\log^C X} \tag{11.13}$$

usually) is very important, since a smaller choice of z increases the leading term view of the factor $W(z)$. For making this remark a little more explicit notice that under (R) one could $z^2 \leq X$ (cf. (11.12)), whereas the use of Bombieri's theorem allows us, for example, upto the bound (11.13), so that

$$z^2 \leq \frac{\sqrt{X}}{\log^C X} \tag{11.14}$$

which worsens the leading term in (11.8) by a factor of 2^κ (in view of the fact that $W_{(z)}$ behaves like $c(\kappa)\log_z^{-\kappa}$ under $\Omega_2(\kappa, L)$). However, if we ask for a bound to the primes represented by an (irreducible) integer-valued polynomial F then we can sieve the sequence $\{F(p)\}$. Then (leaving minor details apart) the dimension of the problem would be 1 and one has to use Bombieri's theorem instead of (R), thereby (apparently) lose a factor 2. Instead one can also sieve (as was necessary before the availability of Bombieri's theorem) the sequence $\{nF(n)\}$. In that event, the dimension becomes 2 and consequently one loses a factor 4 instead of 2. This example provides an instance of how Bombieri's theorem permits linearizing a problem (-i.e., reduce the dimension by one-) and thus save a factor of 2 in the upper estimate.

Now returning to our problem of obtaining good upper estimates, with a view to achieve lower bounds (which are far more important), for $S(\mathscr{A}, \mathfrak{p}, z)$ we find that it is helpful (cf. (11.32)) to generalize the method of Chapter 9 by the introduction of a new parameter (which is possible because of the dual role of z there expressed through $d|P(z)$ and $d < z$) in the following way. **129**

Again we start with (9.26) which holds true under the single condition (9.24),

$$\lambda_1 = 1, \tag{11.15}$$

and then require (9.27) with respect to some arbitrary

$$\xi > 1, \tag{11.16}$$

instead of the z inherent in $S(\mathscr{A}, \mathfrak{p}, z)$, i.e.,

$$\lambda_d = 0 \quad \text{for} \quad d \geq \xi. \tag{11.17}$$

(Note that (11.17) is consistent with (11.15) because of (11.16).) Proceeding as in Chapter 9 one is now led to the choices

$$\lambda_d := \mu(d) \prod_{p|d} \frac{p}{p - \omega(p)} . \frac{G_d(\frac{\xi}{d}, z)}{G(\xi, z)}, \tag{11.18}$$

where

$$G_k(x, z) := \sum_{\substack{d<x \\ d|P(z) \\ (d,k)=1}} g(d), \ G(x, z) := G_1(x, z) \text{ for } 0 < x \in \mathbb{R} \tag{11.19}$$

are generalizations of the functions in (9.22) (and (9.21)). Also note here that (11.18) includes both (11.15) and (11.17). Now again, as before, one has

$$|\lambda_d| \leq 1. \tag{11.20}$$

Further, with the choice (11.18), we get (9.26) with

$$\sum_1 = \frac{1}{G(\xi, z)} \tag{11.21}$$

and also, corresponding to (9.31), **130**

$$\sum_2 \leq \sum_{\substack{d_\nu \xi \\ d_\nu | P(z) \\ \nu=1,2}} \left| R_{d_1, d_2} \right| \leq \sum_{\substack{d<\xi^2 \\ d|P(z)}} 3^{\nu(d)}{}_{|R_d|}, \tag{11.22}$$

because of (11.20), (11.17) and (9.32). Now we see that ξ has, so to speak, taken over from z the role of controlling the order of magnitude of the remainder sum (cf. the remark preceding the statement of (11.15)).

Also we would formulate (cf. again (11.32)) Theorem 9.1 generalized further so as include all $\mathscr{A}_q$'s for $q's$ restricted by

$$(Q) \qquad \mu(q) \neq 0, (q, P(z)) = 1, (q, \overline{\mathfrak{p}}) = 1. \tag{11.23}$$

Here we stress that $\mathscr{A}_q$'s are related to $\mathscr{A}$ through the approximations required by (9.15) (and (9.9)) and consequently this step is not merely a change of notation. The condition (Q) ensures, for this step, that the only changes required in the previous generalization of Theorem 9.1 are the replacements of

$$\sum_1 \quad \text{by} \quad \frac{\omega(q)}{q} \sum_1 \tag{11.24}$$

and of

$$R'_d s \quad \text{by} \quad R'_{qd} s. \tag{11.25}$$

Thus from (9.26), (11.21), (11.24), (11.22) and (11.25), in view (11.16), the required generalization of Theorem 9.1 follows; namely, one has

Theorem 11.5. (Ω_1), (Q)*: For any real number* $\xi > 1$,

$$S(\mathscr{A}_q, \mathfrak{p}, z) \leq \frac{\omega(q)}{q} \frac{X}{G(\xi, z)} + \sum_{\substack{d<\xi^2 \\ d|P(z)}} 3^{\nu(d)} |R_{qu}|. \tag{11.26}$$

Actually, for $q = 1$ and $\xi = z$, (11.26) *is the second inequality of Theorem 9.1 since*

$$\mathscr{A}_1 = \mathscr{A} \tag{11.27}$$

and 131

$$G(z,z) = G(z). \tag{11.28}$$

Here, by imposing the condition $(\Omega_2(\kappa))$ alone we cannot expect to get very useful results. However,as a starting point for the method of obtaining a lower bound for our sifting function, we use $(\Omega_2(\kappa))$ only to derive a simple estimate for $\frac{1}{G(\xi,z)}$ so that it follows from Theorem 11.5.

Theorem 11.6. (Ω_1), $(\Omega_2(\kappa))$, (Q)*: For*

$$\tau := \frac{\log \xi^2}{\log z} \geq 2, \tag{11.29}$$

there holds

$$S(\mathscr{A}_q, \mathfrak{p}, z) \leq \frac{\omega(q)}{q} XW(z)\{1 + O(\exp\{-\frac{\tau}{2}(\log \frac{\tau}{2} + 2)\})\} + \sum_{\substack{d<\xi^2 \\ d|P(z)}} 3^{\nu(d)}|R_{qd}|. \tag{11.30}$$

Now we are in a position to point out how by means of a certain very effective combinatorial argument (in the from of identities concerning our function $S(\mathscr{A}_q, \mathfrak{p}, z)$ and $W(z)$), which is really a result about arrangements from mathematical logic,one can obtain some lower bounds and also improved estimates for our sifting functions from the estimates of Theorem 11.6.

Buchstab was the first to notice the fruitful utility of this combinatorial result for the purposes of sieve methods. We state this result as

Lemma 11.1. (Q)*: If*

$$2 \leq z_1 \leq z, \tag{11.31}$$

then we have

$$S(\mathscr{A}_q, \mathfrak{p}, z) = S(\mathscr{A}_q, \mathfrak{p}, z_1) - \sum_{\substack{z_1 \leq p < z \\ p \in \mathfrak{p}}} S(\mathscr{A}_{qp}, \mathfrak{p}, p) \tag{11.32}$$

as well as

$$W(z) = W(z_1) - \sum_{z_1 \le p < z} \frac{\omega(p)}{p} W(p) \tag{11.33}$$

132

Proof. Let $p_1, p_2 \ldots$ be all the primes belonging of $\mathfrak{p}$, written in their natural order, which are greater than or equal to z_1. If $z \le p_1$ we have, by (9.15) and (9.12)

$$P(z) = P(z_1) \text{ and } \omega(p) = 0 \text{ for } z_1 \le p < z. \tag{11.34}$$

This gives, by (9.6) and (9.20),

$$S(\mathscr{A}_q, \mathfrak{p}, z) = S(\mathscr{A}_q, \mathfrak{p}, z_1) \text{ and } W(z) = W(z_1), \text{ if } z \le p_1. \tag{11.35}$$

Thus Lemma 11.1 is trivially true in this case ($z_1 \le z \le p_1$).

Now suppose that $p_1 < z$, so that defining the integer N by $p_N < z \le p_{N+1}$ we have $N \ge 1$. Then, for each integer $\nu = 1, \ldots, N$, (9.6) yields

$$\begin{cases} S(\mathscr{A}_q, \mathfrak{p}_{\nu+1}) - S(\mathscr{A}_q, \mathfrak{p}_\nu) = \\ \quad = |\{a : a \in \mathscr{A}_q, a \equiv 0 \bmod p_\nu, (a, P(p_\gamma)) = 1\}| = \\ \quad = -S(\mathscr{A}_{qp_\gamma}, \mathfrak{p}, p_\nu), \end{cases} \tag{11.36}$$

and, by (9.20), we have

$$W(p_{\nu+1}) - W(p_\nu) = -\frac{\omega(p_\nu)}{p_\nu} W(p_\nu). \tag{11.37}$$

Summing up the identities (11.36) and (11.37) over $\nu = 1, \ldots, N$ and observing that $S(\mathscr{A}_q, \mathfrak{p}, z) = S(\mathscr{A}_q, \mathfrak{p}_{N+1})$ and $W(z) = W(p_{N+1})$ we obtain (11.32) and (11.33). This completes the proof of the lemma.

To see as to how (11.32) links up the problem of obtaining a lower bound for sifting functions to that of having good upper estimates, suppose that one has a lower bound for (the larger) $S(\mathscr{A}_q, \mathfrak{p}, z_1)$. Then upper bounds for $S(\mathscr{A}_{qp}, \mathfrak{p}, p)$'s enable us to obtain a lower bound for (the
133 smaller) $S(\mathscr{A}_q, \mathfrak{p}, z)$. (We also have a similar remark the problem of obtaining upper bounds by means of (11.32).) However, the significant part of (11.32) is its iterative aspect consisting of using (11.32) to (some

of) the S-functions with (the respective) p's in place of z, thereby obtaining more terms of both signs which in turn (after 'many' iterations) would provide a more effective scheme for the above procedure. Actually , it is again this iterative process because of which (11.32) is stated with a (general) q (rather than with $q = 1$).

Since, to start with, we do not have general lower bound for $S(\mathscr{A}_q, \mathfrak{p}, z_1)$ one can only make the choice

$$z_1 = 2 \tag{11.38}$$

so that, by (9.15),

$$S(\mathscr{A}_q, \mathfrak{p}, z_1) = |\mathscr{A}_q| = \frac{\omega(q)}{q} X + R_q. \tag{11.39}$$

Then, on multiplying (11.33) by $\dfrac{\omega(q)}{q}X$ and subtracting the result from (11.32) we are led to deal with the remainders R_{qd}'s only. Now, in order not to accumulate too many terms from the upper estimate for the sum in (11.32), we use Theorem 11.5 for $S(\mathscr{A}_{qp}, \mathfrak{p}, p)$ with ξ^2 replaced ξ_0^2/p (and also with p instead of z, which ensures (Q) for qp in place of q). Thus we arrive at a first step lower bound corresponding to the upper bound of Theorem 11.6, and so we state this result in the of an asymptotic equality: □

Theorem 11.7. (Ω_1), $(\Omega_2(\kappa))$, (Q)*: For*

$$\tau := \frac{\log \xi^2}{\log z} \geq 2. \tag{11.40}$$

there holds

$$S(\mathscr{A}_q, \mathfrak{p}, z) = \frac{\omega(q)}{q} X W(z)\{1 + O(\exp\{-\frac{\tau}{2}(\log \frac{\tau}{2} + 2)\})\} + \theta \sum_{\substack{d<\xi^2 \\ d|P(z)}} 3^{\nu(d)} |R_{qd}| \cdot |\theta| \leq 1. \tag{11.41}$$

It should be mentioned here that, with $q = 1$ and a suitable choice for ξ.

One easily obtains from Theorem 11.7, under the condition (R), the **134**
so-called 'Fundamental Lemma' (cf. Kubiliyus [2], Lemma 1.4)

Theorem 11.8. (Ω_1), ($\Omega_2(k)$), (R)*: For*

$$u := \frac{\log X}{\log z} \geq 1. \tag{11.42}$$

there hold

$$S(\mathscr{A}, \mathfrak{p}, z) = XW(z)\{1 + O(e^{-\frac{1}{2}u \log u}) + O(e^{-\sqrt{\log X}})\} \tag{11.43}$$

and

$$S(\mathscr{A}, \mathfrak{p}, z) = XW(z)\{1 + O(e^{-\frac{1}{2}u}\}. \tag{11.44}$$

Observe that the preceding results (i.e., Theorem 11.6, 11.7 and 11.8) are significant only if u (or τ) is large (which means that z should be small in comparison with X (or ξ)). The reason for this limitation is, apart from the weak condition ($\Omega_2(\kappa)$), mainly due to the fact that Lemma 11.1 has been used with a 'trivial' choice for z_1 in deriving the lower estimate. However, using the stronger $\Omega_2(\kappa, L)$) instead of ($\Omega_2(\kappa)$) we obtain a more precise information about $G(\xi, z)$. And still more important is that Theorem 11.7 enables us to employ Lemma 11.1 with a 'non trivial' choice for z_1 (cf. (11.52)).

From this point onwards we shall confine ourselves to the case of dimension

$$\kappa = 1, \tag{11.45}$$

and we owe some explanation for doing so. First of all, the upper estimates for $G(\xi, z)$ under ($\Omega(\kappa, L)$) in the case of general dimension κ (and then so all for $S(\mathscr{A}, \mathfrak{p}, z)$ via Theorem 11.5) become quite complicated and further, they do not further, they do not yield satisfactory results (apart from the particular cases of $\kappa = 1$ and $\kappa = \dfrac{1}{2}$), when applied for obtaining lower bounds by means of Lemma 11.1. Also, when one is interested solely in finding upper estimates (in the most interesting ques-
135 tions) the generalized Theorem 11.5 has no advantage over (the simple) Theorem 9.1. Finally, we have that most of the prominent problems in prime number theory which can be attacked by Selberg's sieve method are dimension 1.

Now on imposing ($\Omega_2(1, L)$) one obtains

Lemma 11.2. (Ω_1), $(\Omega_2(1, L))$*: Let*

$$\tau := \frac{\log \xi^2}{\log z} > 0. \tag{11.46}$$

Then holds

$$\frac{1}{G(\xi, z)} = W(z)\{F_0(\tau) + O(\frac{L}{\log z}(\tau^3 + \tau^{-2}))\}, \tag{11.47}$$

where $F_0(\tau)$ is defined by

$$F_0(\tau) = \frac{2e^\gamma}{\tau} \text{ for } o < \tau \leq 2, \tag{11.48}$$

and by the differential-difference equation

$$(\frac{1}{\tau F_0(\tau)})' = -\frac{1}{\tau^2 F_0(\tau - 2)} \text{for } \tau > 2. \tag{11.49}$$

If we apply lemma 11.2 in Theorem 11.5 we obtain, corresponding to Theorem 11.6, the following

Theorem 11.9. (Ω_1), $(\Omega_2(1, L))$, (Q)*: For any real number* $\xi > 1$ *and*

$$\tau := \frac{\log \xi^2}{\log z} (> 0), \tag{11.50}$$

there holds

$$S(\mathscr{A}_q, \mathfrak{p}, z) \leq \frac{\omega(q)}{q} XW(z)\{F_0(\tau) + O(\frac{L}{\log z}(\tau^3 + \tau^{-2}))\} + \sum_{\substack{d<\xi^2 \\ d|P(z)}} 3^\nu(d)_{|R_{qd}|}, \tag{11.51}$$

where $F_0(\tau)$ *is defined by* (11.48) *and* (11.49).

We can now repeat that above Buchstab-procedure (cf. description preceding Theorem 11.7) to obtain a general non-trivial lower bound for **136**
our sifting functions. To this end we choose in (11.32)

$$z_1 = \exp\{\frac{\log \xi}{\log \log \xi}\} \tag{11.52}$$

(and also assume ξ to be sufficiently large; indeed, we can even suppose that $\xi \geq z$ since the result below (Theorem 11.10) is otherwise trivial because of (11.59)). This choice enables us to employ Theorem 11.7 for a lower estimate of $S(\mathscr{A}_q, \mathfrak{p}, z_1)$ and for the remaining part of the right-hand side in (11.32) we can use now Theorem 11.9 instead, but again with

$$\xi^2 \text{ replaced by } \frac{\xi^2}{p} \text{ and } z \text{ replaced by } p \tag{11.53}$$

for the same reasons as before. However, here we encounter an additional difficulty due to the presence of factors $F_0(\frac{\log(\xi^2/p)}{p})$ (instead of 1 in the previous case) stemming from our use of Theorem 11.9 with (11.53). This difficulty is overcome by deriving from (11.33) the following

Lemma 11.3. (Ω_1), $(\Omega_2)1, L))$: *Suppose that*

$$2 \leq z_1 \leq z \leq \xi, \tag{11.54}$$

and let $\psi(t)$ be a non-negative, monotonic and continuous function for $t \geq 1$. Further, define

$$M : \max_{z_1 w \leq < z} \psi(\frac{\log(\xi^2/w)}{\log w}), \tag{11.55}$$

Then holds

$$\sum_{z_1 \leq p < z} \frac{\omega(p)}{p} W(p) \psi(\frac{\log(\xi^2/p)}{\log p})$$
$$= W(z) \frac{\log z}{\log \xi p} \int_{\frac{\log \xi^2}{\log z}}^{\frac{\log \xi^2}{\log z_1}} \psi(t-1) dt + O(\frac{LMW(z)\log z}{\log^2 z_1}). \tag{11.56}$$

Thus using Lemma 11.3 (instead of (11.33)) along with (11.32) one obtain, by the above procedure, without any more difficulty the required

137 **Theorem 11.10.** $(\Omega_1), (\Omega_2(1, L)), (Q)$*: For any real number* $\xi > 1$ *and*

$$\tau := \frac{\log\log\xi^2}{\log z}(> 0). \tag{11.57}$$

there holds

$$S\,)\mathscr{A}_q.\mathfrak{p}, z) \geq \frac{\omega(q)}{q} XW(z)\Big\{ f_0(\tau) + O(\frac{L(\log\log 3\xi)^5}{\log\xi})\Big\} - \sum_{\substack{d<\xi^2\\ d|P(z)}} 3^{\nu(d)}|R_{qd}|, \tag{11.58}$$

where $f_0(u)$ *is defined by*

$$f_0(\tau) = 0 \text{ for } 0 < \tau \leq \nu_0 = 2.06\cdots, \tag{11.59}$$

and by, with the function F_0 *of Lemma 11.2,*

$$(\tau f_0(\tau))' = F_0(\tau - 1) \text{ for } \tau \geq \nu_0. \tag{11.60}$$

Now, before starting the iteration of the Buchstab-procedure (cf. the remarks made subsequent to the proof to Lemma 11.1), we make the result of Theorem 11.10 explicit (for $q = 1$) by imposing the following condition (about R_d's):

Suppose that there are constants

$$0 < \alpha \leqslant 1, \quad A_3(\geqslant 1), \quad A_4(\geqslant 1) \tag{11.61}$$

such that

$$(R(1,\alpha)) \qquad \sum_{\substack{d<\frac{X^\alpha}{\log^{A_3} X}\\ (d,\mathfrak{p})=1}} \mu^2(d)3^{\nu(d)}|R_d| \leqslant A_4 \frac{X}{\log^2 X} \quad \text{for } X \geqslant 2. \tag{11.62}$$

Then one gets from Theorem 11.10 in the case $q = 1$, on making the choice

$$\xi^2 = \frac{X^\alpha}{\log^{A_3} X} \tag{11.63}$$

and noting that $d|P(z)$ implies that $\mu(d) \neq 0$ and $(d, \overline{\mathfrak{p}}) = 1$,

Theorem 11.11. (Ω_1), $(\Omega_2(1,L))$, $(R(1,\alpha))$*:*

$$S(\mathscr{A},\mathfrak{p},z) \geqslant XW(z)\Big\{ f_0(\alpha\frac{\log X}{\log z}) + O(\frac{L(\log\log 3X)^5}{\log X})\Big\}, \tag{11.64}$$

where $f_0(\tau)$ *is defined by* (11.59) *and* (11.60). 138

Returning now to the remarkable iterative aspect of the Buchstab-procedure we see that on using Theorem 11.9, with an appropriate choice of z_1, one has an upper bound for the first term on the right- hand side in (11.32) and also that using Theorem (11.10), with the replacements mentioned in (11.53). in combination with Lemma 11.3 (for f_0 in place of ψ) there follows a lower bound for the sum in (11.32). Thus we arrive at another form of Theorem 11.9, where $F_0(\tau$ is replaced by some other (similar) function $F_1(\tau)$ and this in turn also yields another form of Theorem 11.10, where $f_0(\tau)$ is replaced by an $f_1(\tau)$ (related to $F_1(\tau)$).

Continuing this procedure we are led to results of the type Theorem 11.9 and 11.10, with (at the τ th step) the following pair of functions (instead of $F_0(\tau) f_0(\tau)$ respectively)

$$F_\mu(\tau),\ f_\mu(\tau), \quad \mu = 0,1,2,\cdots, \tag{11.65}$$

where (analogous to the first step as indicated in (11.88))

$$f_\mu(\tau) := \begin{cases} f_0(\tau) = 0 \text{ for } \tau \leq \nu_\mu, \\ 1 - \frac{1}{\tau}\int\limits_{\tau}^{\infty}(F_\mu(t-1) \,\mathrm{dt} \text{ for } \tau \geq \nu_\mu, \end{cases} \tag{11.66}$$

with the number ν_μ defined by

$$\frac{1}{\nu_\mu}\int_{\nu_\mu}^{\infty}(F_\mu(t-1)-1)dt = 1 \tag{11.67}$$

(cf. the remark following (11.88)) and similarly

$$F_{\mu+1}(\tau) := \begin{cases} F_0(\tau) \quad \text{for } \tau \leq \nu'_\mu, \\ 1 - \dfrac{1}{\tau}\int_{\tau}^{\infty}(f_\mu(t-1)-1) \,\mathrm{dt} \text{ for } \tau \geq \nu'_\mu \end{cases} \tag{11.68}$$

139 with suitably chosen numbers ν'_μ for $\mu = 0, 1, 2, \ldots$

The power of this procedure is demonstrated by the surprising fact that, at each step, the quality of the respective forms of the Theorems 11.9 and 11.10 improves. Further, the sequence of numbers $\{\nu_\mu\}$ converges to 2 from above, as also does $\{\nu'_\mu\}$, and the pair of functions $\{F_\mu, f_\mu\}$ converges to a pair of limit functions $\{F, f\}$ converges to a pair of limit functions $\{F, f\}$, as $\mu \to \infty$:

$$\lim_{\mu\to\infty} \nu_\mu = 2 = \lim_{\mu\to\infty} \nu'_\mu, \lim_{\mu\to\infty} F_\mu(\tau), = F(\tau), \lim_{\mu\to\infty} f_\mu = f(\tau). \tag{11.69}$$

Now, from (11.69), (11.66), (11.48) and (11.59), we find that

$$F(\tau) = \frac{2e^\gamma}{\tau}, f(\tau) = 0 \text{ for } 0 < \tau \leq 2 \tag{11.70}$$

and

$$(\tau F(\tau))' = f(\tau - 1) \text{ and } (\tau f(\tau))' = F(\tau - 1), \text{ for } \tau \geq 2, \tag{11.71}$$

which gives on integrating from 2 to u

$$uF(u) - 2e^\gamma = \int_1^{u-1} f(t)dt \text{ and } uf(u) = \int_1^{u-1} F(t)dt, \text{ for } u \geq 2. \tag{11.72}$$

Then we obtain from (11.72), in particular, (cf. (11.70))

$$F(u) = \frac{2\epsilon^\gamma}{u} \text{ for } 0 < u \leq 3 \tag{11.73}$$

and (so) further

$$f(u) = \frac{2e^\gamma}{u} \log(u - 1) \text{ for } 2 \leq u \leq 4. \tag{11.74}$$

Also, we have

$$\begin{gathered} F(u) \geq 0, f(u) \geq 0, F(u) \downarrow, f(u) \uparrow \text{ (for } u > 0) \\ \text{and } \lim_{u\to\infty} F(u) = 1 = \lim_{u\to\infty} f(u) \end{gathered} \tag{11.75}$$

and

$$0 < F(u_1) - F(u_2) \leq F(\delta)\frac{u_2 - u_1}{u_1},$$
$$0 \leq f(u_2) - f(u_1) \leq 2e^{\gamma}\frac{u_2 - u_1}{u_1}, \text{ for } 0 < \delta \leq u_1 < u_2. \tag{11.76}$$

140 Actually, when one knows these results (about F, f) there is no need to consider the sequence of pairs of functions $\{F_\mu, f_\mu\}$ (and so also the numbers ν_μ, ν'_μ). One can instead iterate (11.32) as well as Lemma 11.3 (with $\psi = F$ and $\psi = f$) and apply the Buchstab-procedure (as was done, for instance, to obtain Theorem 11.10) only once. More precisely, in the iterated version of (11.32) among the various functions S occurring with different signs, one can apply Theorem 11.7 for those which are within its domain (of applicability). For those of the remaining $S's$ which are to be estimated from above one can use Theorem 11.9 while the trivial lower estimate $S \geq 0$ in used for the rest. It remains only to prove that this process converges. However, this can be done if the dimension κ satisfies

$$\kappa < \kappa_0, \tag{11.77}$$

where κ_0 is some constant greater than 1. Therefore we can succeed in our case (11.45).

Now we can state the final result of the Buchstab-procedure obtained in the manner mentioned in our previous remark.

Theorem 11.12. (Ω_1), $(\Omega_2(1, L))$, (Q)*: For*

$$\xi \geq z, \tag{11.78}$$

there hold

$$S(\mathscr{A}_q, \mathfrak{p}, z) \leq \frac{\omega(q)}{q} XW(z)\left\{F(\frac{\log \xi^2}{\log z}) + O(\frac{I,}{(\log \xi)^{1/14}})\right\} + \sum_{\substack{d<\xi^2 \\ d|P(z)}} 3^{\nu(d)}|R_{qd}|, \tag{11.79}$$

and

$$S(\mathscr{A}_q, \mathfrak{p}, z) \geq \frac{\omega(q)}{q} XW(z)\left\{f(\frac{\log \xi^2}{\log z}) + O(\frac{I,}{(\log \xi)^{1/14}})\right\} + \sum_{\substack{d<\xi^2 \\ d|P(z)}} 3^{\nu(d)}|R_{qd}|, \tag{11.80}$$

where the functions F, f are defined by (11.70) *and* (11.71)*, and the O-constants depend atmost on A_1 and A_2.* **141**

This theorem is also true if

$$1 < \xi < z \text{ but } \ll \xi^{\lambda} \tag{11.81}$$

with a positive constant λ, in which case the O-constant in (11.79) depends also on λ.

Similar to Theorem 11.11 we can obtain, in the case $q = 1$, with the same choice as (11.63) the following final result from Theorem 11.12.

Theorem 11.13. (Ω_1), $(\Omega_2(1, L))$, $(R(1, \alpha))$*: For*

$$z \leq X, \tag{11.82}$$

we have

$$S(\mathscr{A}, \mathfrak{p}, z) \leq XW(z)\left\{F(\alpha\frac{\log X}{\log z}) + O(\frac{L}{(\log X)^{1/14}})\right\}, \tag{11.83}$$

and

$$S(\mathscr{A}, \mathfrak{p}, z) \geq XW(z)\left\{f(\alpha\frac{\log X}{\log z}) + O(\frac{L}{(\log X)^{1/14}})\right\}, \tag{11.84}$$

where the functions F, f are defined by (11.70) *and* (11.71)*, and the O-constants depend atmost on A_i, $i = 1, 2, 3, 4$, and α.*

Although the functions F, f are invariant under the Buchstab - procedure, in view of the fact that

$$f(u) < F(u) \quad \forall u > 0 \tag{11.85}$$

and the procedure described preceding (11.77) it is natural to have some doubt as to whether the qualities of our final results (viz. Theorems 11.12 and 11.13) cannot further be improved. However, it can be shown that, for the sets

$$\mathscr{A} = \mathscr{B}_\nu := \left\{n : |1 \leq n \leq x, \Omega(n) \equiv |\nu \bmod 2\right\}, \nu = 1, 2 \text{ and } \mathfrak{p} = \mathfrak{p}_1 \tag{11.86}$$

(cf. (9.39)). the relations (11.83) and (11.84) hold with equality signs (upto the leading term) for $\nu = 1$ and $\nu = 2$ respectively and for all 142
values of $u : \dfrac{\log X}{\log z} > 0$ in both cases); (Actually, here we take $X = \dfrac{x}{2}$, $\omega(p) = 1$ and slightly modify $R(1, \alpha)$, with $\alpha = 1$, via Theorem 11.12). It is in this sense that the final form, as stated in Theorem 11.13, of the (proper Selberg sieve is best possible.

NOTES

They survey of the Selberg's sieve given in this chapter follows the approach of Halberstam and Richert [1] and we refer to (Chapters 4, 5, 6, 7, 8, of) this book for the content of this chapter as well as for the applications of these results. So all the references below (unless otherwise explicitly stated) are referred to by 'l,c.'.

(11.5)**:** cf. Lemma 4.1

(11.6)**:** cf. Lemma 5.4

Theorem 11.1: cf. l.c Theorem 4.1.

Theorem 11.2: cf. l.c. Theorem 5.2 (see also under (11.12) below.)

Theorem 11.3: cf. l.c. Theorem 2.2. Actually, Theorem 11.3 holds for $z \geq X^A$ also when the restriction in the product, $p < z$, is replaced by $p < X$.

Theorem 11.4; cf. l.c. Theorem 5.1.

(11.12): Actually, in the O-term of (11.12) (and also of Theorem 11.2 L can be replaced by min $(L, \log z)$ so that the Theorem 11.4 (and respectively Theorem 11.2) would include Theorem 11.3 (and respectively Theorem 11.1). This is, however, not surprising because of the fact that the condition $(\Omega_2(\kappa, L))$ includes $(\Omega_2(\kappa))$.

(11.13): For more details concerning the description here, in connection with

(11.13): see l.c Chapter 5, **5.**

Theorem 11.5: cf. l.c. Theorem 6.1.

143 Theorem 11.6: cf. l.c. Theorem 6.2.

Lemma 11.1: cf. l.c. Lemma 7.1.

Theorem 11.7: cf. l.c Theorem 7.1.

Theorem 11.8: cf. l.c. Theorem 7.2. Also cf. l.c. Theorem 2.5 for a slightly stronger result derived from Brun's sieve.

Lemma 11.2: cf. l.c. Lemma 6.1 and also l.c. (4.18) on p. 201.

(11.48), (11.49): For the function F_0 defined here, it can be proved that

$$F_0(\tau) \geq 0, F_0(\tau) \downarrow (\text{ for } \tau > 0) \text{ and } \lim_{\tau\to\infty} F_0(\tau) = 1. \tag{11.87}$$

For more details about $F_0(\tau)$, cf. l.c. Chapter 6, **3** (where this function occurs as $\frac{1}{\sigma_1(\tau)}$).

Theorem 11.9: cf. l.c. Theorem 6.3.

Lemma 11.3: cf. l.c. Theorem 7.2.

Theorem (11.10): cf. l.c. Theorem 7.3.

(11.59), (11.60): To form an idea of the introduction of this function $f_0(\tau)$ it is useful to observe that, by the procedure leading to Theorem 11.10 here, the contributions to the leading term (apart from the factor $\frac{\omega(q)}{q} XW(z)$) are, in view of Lemma 11.3 here (with $\psi(t) = F_0(t)$) and Theorem 11.9 here,

$$1 - \frac{1}{\tau}\int_\tau^\infty (F_0(t-1) - 1)\,\mathrm{dt} =: f_0(\tau). \tag{11.88}$$

This gives (11.60) and the choice of (11.59) made since $f_0(\tau)$ is negative if $\tau \leq \nu_0$ (by (11.88)), where ν_0 is defined by

$$\frac{1}{\nu_0}\int_{\nu_0}^\infty (F_0(t-1) - 1)dt = 1, \tag{11.89}$$

while one has always, on the other hand, the trivial $S(\mathscr{A}_q, \mathfrak{p}, z) \geq 0$, Now (11.89) yields

$$\nu_0 = 2.06.$$

Further, one can show, corresponding to (11.87) above **144**

$$f_0(\tau) \geq 0, f_0(\tau) \uparrow (\text{ for } \tau > 0) \text{ and } \lim_{\tau\to\infty} f_0(\tau) = 1. \tag{11.90}$$

(11.62): Clearly, this condition has been modelled such that Bombieri-type theorems are directly applicable.

Theorem 11.11. cf. l.c. Theorem 7.4.

(11.70),...(11.76): Jurkat and Richert [1] (cf. l.c. (Chapter 8, **2**)).

Theorem 11.12: cf. l.c. Theorem 8.3.

Theorem 11.13: Jurkat and Richert [1], Halberstam, Jurkat and Richert [1] (cf. l.c. (Theorem 8.4)). Selberg [5] has pointed out that (according to an unpublished paper of J.B. Rosser) this result can also be derived from Brun's sieve. This has been proved, even with a better error term (namely, with $\frac{1}{14}$ replaced by 1) by Iwaniec [1]. Iwaniec [1] has also applied this improved version of Theorem 11.13 to sharpen the bounds for the Legendre-Jacobsthal function $C_0(r)$, the maximum length a block of consecutive integers each of which is divisible by at least one of the first r primes. His result is

$$C_0(r) \ll r^2 \log^2 r,$$

whereas (our) Theorem 11.13 leads only to the estimate

$$C_0(r) \ll r^2 \exp\{(\log r)^{13/14}\}.$$

In the opposite direction we have, by Rankin [1].

$$C_0(r) > e^{\gamma-\epsilon} \frac{r \log^2 r \log\log\log r}{(\log\log r)^2}$$

(cf. l.c. (p. 239)). (For the case $\kappa = \frac{1}{2}$ see Iwaniec [2], [6], and for $\kappa < \frac{1}{2}$ see Iwaniec [7]. However, for κ exceeding some constant $\kappa_1 (> 1)$ Selberg's sieve seems to be always superior than Brun's sieve (for instance, in sifting values of reducible polynomials).

145 (11.86): This fact (about (11.86) was established by Selberg [3] (for $0 < u \le 2$), and he added the remark that the sieve method "cannot distinguish between numbers with an odd or an even number of prime factors".

Chapter 12

Weighted Sieves

AS ALREADY mentioned in the previous chapter our object is to use the 'final' results there for the proofs of the results of next chapter. For this purpose we need further improve to the quality of these results,inspite of the concluding remarks of Chapter 11, and this is achieved by considering a weighted sieve, which actually consists of a combination of various sifting functions (so that the counter-examples (11.86) collapse). **146**

Before commencing the introduction of weighted sieves for the purpose mentioned above, we briefly point out as to how the final result of Chapter 11, in particular (11.84), already yields a result in the direction of Theorem 13.2. For this, let us consider

$$\mathscr{A} := \left\{p + 2 : p \leq x\right\}, \mathfrak{p} := \mathfrak{p}_2 \tag{12.1}$$

(cf. (9.39)). Here we can take

$$X = li\ x, \omega(p) = \frac{p}{p-1} \text{ for } p \in \mathfrak{p}(i.e., \text{ for } p \geq 2) \tag{12.2}$$

and then the conditions (Ω_1), $(\Omega_2(1, L))$ are verified easily for some fixed *constants* A_1, A_2 and $L(\geq 1)$. Further, Bombieri's theorem (Lemma 10.2) enables us to fulfill $(R(1, \alpha))$ with

$$\alpha = \frac{1}{2} \tag{12.3}$$

(for some suitably chosen absolute constants A_3 and A_4). Also we find, by (10.8),

$$\begin{cases} W(z) = \prod_{2<p<z}(1-\frac{1}{p-1}) & = 2\prod_{p<z}(1-\frac{1}{p})\Pi_{2<p<z}(\frac{p(p-2)}{(p-1)^2)}) = \\ & = 2e^{-\gamma}\prod_{p>2}(1-\frac{1}{(p-1)^2})\frac{1}{\log z}(1+O(\frac{1}{\log z})). \end{cases} \tag{12.4}$$

147 Therefore, taking

$$z = x^{\frac{1}{u}}, u > 4; u = 4.2 \tag{12.5}$$

say, we obtain from (11.84) and (11.74) that, for some positive constant c_0,

$$S(\mathscr{A}, \mathfrak{p}, z) \geq c_O \frac{x}{\log^2 x} \quad \text{for} \quad x \geq x_O. \tag{12.6}$$

Now, note that the numbers counted on the left-hand side here have no prime divisor $< z$ and so for each of these numbers we have

$$x + 2 \geq p + 2 = p_1 \cdots p_r \geq x^{r/u} \tag{12.7}$$

which shows tht necessarily $r \leq u$ or by (12.5) that $r \leq 4$. Thus, letting $x \to \infty$, we have that for infinitely many primes p holds

$$p + 2 = P_4 \tag{12.8}$$

At this point it is worthwhile to notice that Theorem 11.11 (instead of (11.84)) would have also led to (12.6), though with a smaller value for c_0, and so also to (12.8). We can express this remark by saying that (12.8) follows from Selberg's upper bound sieve combined with the (one-step) Buchstab-procedure on using Bombieri's Prime number theorem (cf. also the remark involving (13.21), with respect to Theorem 12.2 below).

Now we turn to the weighted sieves. The first weighted sieve, suitable for our purpose also, was introduced by Kuhn (in connection with Brun's sieve). Since then there have been other more sophisticated sieves invented. However, we shall need only a special form of the simplest of these, namely Kuhn's sieve (cf. Theorem 12.1).

Theorem 12.1. $(R(1,\alpha))$: *Let h be an even integer (determined with respect to x) satisfying*

$$0 < |h| \leq x \tag{12.9}$$

and suspect that (associated with a sequence $\mathscr{A}$) we have **148**

$$X = \mathrm{li}\, x, \omega(p) = \frac{p}{p-1} \text{ for } p \in \mathfrak{p}_h. \tag{12.10}$$

Let u and v be two real numbers (independent of x) such that

$$\frac{1}{\alpha} < u < v. \tag{12.11}$$

Define

$$W(\mathscr{A}; \mathfrak{p}_h, v, u, \lambda) := \sum_{\substack{a\in\mathscr{A} \\ (a,P(x^{1/V}))=1}} \{1 - \frac{1}{\lambda} \sum_{\substack{x^{\frac{1}{v}} \leq p < x^{\frac{1}{u}} \\ p|a \\ p\in\mathfrak{p}_h}} 1, \tag{12.12}$$

$$2 \leq \lambda \in \mathbb{R}. \tag{12.13}$$

Then

$$W(\mathscr{A} : \mathfrak{p}_h, v, u, \lambda) \geq e^{-\gamma}\mathbb{C}(h)\frac{x}{\log^2 x}v$$
$$\left\{ f(\alpha v) - \frac{1}{\lambda}\int_u^v)F(v(\alpha - \frac{1}{t}))\frac{dt}{t} + O(\frac{1}{(\log x)^{1/15}})\right\} \tag{12.14}$$

where $\mathbb{C}$ is defined by (9.42) and the O-constant depends atmost on u, v, A_3, A_4 and α.

Proof. We may assume that

$$x \geq X_0(u, v, A_3, A_4, \alpha). \tag{12.15}$$

Let us set

$$z := x^{\frac{1}{v}},\ y := x^{1/u}. \tag{12.16}$$

Then, by (12.12),

$$W(\mathscr{A} : \mathfrak{p}_h, v, u, \lambda) = S(\mathscr{A}, \mathfrak{p}_h, z)$$

$$-\frac{1}{\lambda}\sum_{\substack{z\le p<y\\ p\in\mathfrak{p}_h}} S(\mathscr{A}_p,\mathfrak{p}_h,z) = S(\mathscr{A},\mathfrak{p}_h,z)-\frac{1}{\lambda}\sum_1, \tag{12.17}$$

say. Since $2|h$, by (12.10), the condition (Ω_1) is satisfied with an *abso-*
149 *lute* constant A_1. Also $(\Omega_2(1,L))$ is fulfilled with some *absolute* constant A_2 and, by (12.9) with

$$L \le O(1)+\sum_{p|h}\frac{\log p}{p-1} \ll \log 3|h| \ll \log\log x. \tag{12.18}$$

Hence we get, by (11.84) (since $v>1$ by (12.11)),

$$S(\mathscr{A},\mathfrak{p}_h,Z) \ge XW(z)\{f(\alpha v(\frac{\log \operatorname{li} x}{\log x}))+O(\frac{1}{(\log x)^{1/15}})\}. \tag{12.19}$$

We apply Theorem 11.12 for the estimation of $\sum_1$, and for this we define (in terms of the constants from $(R(1,\alpha))$)

$$\xi^2=\frac{x^\alpha}{\log^{A_3+\alpha}X}. \tag{12.20}$$

Note that for each p in the range of $\sum_1$ we have

$$\frac{\xi^2}{p}\ge\frac{x^{\alpha-\frac{1}{u}}}{\log^{A_3+\alpha}X} \tag{12.21}$$

Therefore applying (11.79), for each term in $\sum_1$, with ξ^2/p in place of ξ^2, $\mathfrak{p}=\mathfrak{p}_h$ (and $q=p$ so that (Q) is satisfied because of $p\ge z$ and $p\in\mathfrak{p}_h$) we obtain

$$\sum_1 \le \sum_{z\le p<Y}\frac{\omega(P)}{P}XW(z)\left\{F(\frac{\log(\xi^2/p)}{\log z})+O(\frac{1}{\log^{1/15}})\right\}+\sum_{\substack{d<\xi^2\\ (d,\mathfrak{p})=1}}\mu^2(d)3^{\gamma(d)}|R_d|. \tag{12.22}$$

in view of (12.21), (12.11), (11.81) and (12.18). Estimating the last sum in (12.22) by means of $(R(1,\alpha))$ (cf. (11.62)), (12.20) and (12.10) one gets

$$\sum_1 \le XW(z)\left\{\sum_{z\le p<Y}\frac{1}{p}F(\frac{\log(\xi^2/p)}{\log z})+O(\frac{1}{(\log x)^{1/15}})\right\}+O(\frac{x}{\log^3 x}). \tag{12.23}$$

after some simple considerations involving (12.10), (11.73), (11.75), (11.75), (11.75), (12.16) and (10.22).

Now, from (12.10), (10.8) and (12.16), it follows that (cf, (12.4))

$$W(z) = \prod_{\substack{2<p<z \\ p\nmid h}} (1 - \frac{1}{p-1}) = \frac{e^{-\gamma}}{\log z} \mathfrak{S}(h) + O(\frac{\log\log x}{\log^2 x}) \tag{12.24}$$

where $\mathfrak{S}$ is defined by (9.42) and (so) satisfies, (cf. (12.9)), **150**

$$\mathfrak{S}(h) = O(\log\log 3|h|) = O(\log\log x). \tag{12.25}$$

Using (12.24) in both (12.19) and (12.23) we are led to (by (12.17) and (12.13))

$$W(\mathscr{A} : \mathfrak{p}_h, v, u, \lambda) \geq \frac{x}{\log^2 x} e^{-\gamma} \mathbb{C}(h) v$$
$$\left\{ f(\alpha v) - \frac{1}{\lambda} \sum_{Z \leq p < Y} \frac{1}{p} F(\frac{\log(x^{\alpha/P})}{\log x} v) + O(\frac{1}{(\log x)^{\frac{1}{15}}}) \right\} \tag{12.26}$$

by means of (11.76) and (12.20). It remains only to deal with the sum in (12.26) and for this we proceed as in (10.24) obtaining thereby

$$\sum_{z \leq p < y} \frac{1}{p} F(v(\frac{\log(\frac{x^\alpha}{p})}{\log x}) = \int\limits_z^y F(v \frac{\log(\frac{x^\alpha}{w})}{\log x}) \frac{dw}{w \log w} + O(\frac{1}{\log z})$$
$$= \int_u^v F(v(\alpha - \frac{1}{t})) \frac{dt}{t} + O(\frac{1}{\log x}). \tag{12.27}$$

Using this in (12.26) yields (12.14) and so the proof is completed. □

Regarding the terms containing the function f and F it turns out that for a suitable range of v one can have instead an expression involving only elementary functions (cf. (11.73)) and (11.74)). More precisely, we have

Lemma 12.1. *Let*

$$\frac{1}{\alpha} < u < v, \frac{2}{\alpha} \leq v \leq \frac{4}{\alpha}(\textit{ for } 0 < \alpha \leq 1). \tag{12.28}$$

Then (for $2 \leq \lambda \in \mathbb{R}$*)*

$$f(\alpha v) - \frac{1}{\lambda}\int_u^v F(v(\alpha - \frac{1}{t}))\frac{dt}{t} = \frac{2e^{-\gamma}}{\alpha v}\left\{\log(\alpha v - 1) - \frac{1}{\lambda}\log\frac{v - \frac{1}{\alpha}}{u - \frac{1}{\alpha}}\right\}. \tag{12.29}$$

Proof. Note that the arguments of the functions f and F in (12.29) satisfy, by (12.28),

$$2 \leq \alpha v \leq 4 \text{ and } 0 < v(\alpha - \frac{1}{u}) \leq v(\alpha - \frac{1}{v}) = \alpha v - 1 \leq 3. \tag{12.30}$$

Hence, by (11.73) and (11.74) one has that the left-hand side of (12.29) equals

$$\frac{2e^{\gamma}}{\alpha v}\left\{\log(\alpha v - 1) - \frac{1}{\lambda}\int\limits_u^v \frac{dt}{(t - \frac{1}{\alpha})}\right\}. \tag{12.31}$$

151 This proves (12.29). □

For our use in the next chapter it suffices to have the specialization of Theorem 12.1 to the sequences

$$\mathscr{A} := \{|p + h| : p \leq x\}, 2|h, 0 < |h| \leq x, \tag{12.32}$$

where h is determined with respect to sufficiently large x, and with u and v restricted by (12.28). Here note that for $\mathscr{A}$ one has (12.10) and $R_d = O(E(x, d))$ so that lemma 10.2 (with $k = 1$) fulfills $(R(1, \alpha))$ for

$$\alpha = \frac{1}{2} \tag{12.33}$$

Hence we have, by Theorem 12.1 and Lemma 12.1 the required

Theorem 12.2. *Let h denote an even integer (determined with respect to x) satisfying*

$$0 < |h| \leq x. \tag{12.34}$$

Let u and v be two real numbers (independent of x) subject to

$$2 < u < v, 4 \leq v \leq 8 \tag{12.35}$$

and let

$$2 \leq \lambda \in \mathbb{R}. \tag{12.36}$$

Then we have

$$\sum_{\substack{p \leq x \\ (p+h, \prod p')=1 \\ p' < x^{1/v} \\ p' \nmid h}} \left\{1 - \frac{1}{\lambda} \sum_{\substack{x^{1/v} \leq p < x^{1/u} \\ p' | p+h \\ p' \nmid h}} \right\} \geq \frac{4x}{\log^2 x} \mathscr{S}(h)$$
$$\left\{\log(\frac{v}{2} - 1) - \frac{1}{\lambda} \log \frac{v-2}{u-2} + 0(\frac{1}{(\log x)^{\frac{1}{15}}})\right\}. \tag{12.37}$$

where $\mathfrak{S}$ is defined by (9.42) and the O-constant depends atmost on u and v.

NOTES 152

(12.3): Here and in the sequel one recognizes the effect of Bombieri's theorem (and possible improvements, for instance, like the Elliott and Halberstam [2] conjecture mentioned in the notes of Chapter 6, **3**) when used along with Selberg's sieve.

Among the various weighted sieves introduced successfully (for applications) we mention first Kuhn [1], [2], [3]. Here the basic idea consists in forming the W-function ((12.12)). Next we have Selberg's weights of the form

$$\sum_{a \in \mathscr{A}} \left\{1 - \frac{1}{\lambda} D(a)\right\} (\sum_{d|a} \lambda_d)^2 \tag{12.38}$$

(with λ_d's given by (9.28)). By taking $\mathscr{A} = \{n(n+2) : n \leq x, 2 \nmid n\}$ and $D(a) = d(n) + d(n+2)$ here, Selberg [2] (cf. Selberg [4], [5]) succeeded in proving that

$$n(n+2) = P_5 \tag{12.39}$$

holds for infinitely many integers n. This method, in the case where $D(a)$ has (apart from a term to take care of the 'small' prime divisors of a) the form of the inner sum in Kuhn's W (cf. (12.12)) has been published by Miech [1], [2], and Porter [1], Next, Ankeny and Onishi [1] have replaced Kuhn's constant weight $\frac{1}{\lambda}$, attached to the inner sum in (12.12), by a logarithmic weight which is more effective and also has a smoothing effect on the prime divisors in that sum and this weight has been generalized by Richert [1]. Both the Kuhn weight and the logarithmic weight can also be used simultaneously in (12.38) (cf. Halberstam and Richert [1] (Theorem 10.8)). In the first case, a generalization and refinement of Selberg's second method (mentioned above in connection with (12.39)) can be found in Bombieri [6] (§ 8). Bombieri [6] (§ 9) (cf. [8]) has used this method, which is both elegant and comparatively simple (through some what weaker than the other methods described
153 above), for the problem $p+2 = P_4$. (For a more general result which can be obtained by this method, are Halberstam and Richart [1] (Theorem 10.9).) Buchstab [2] has generalised Kuhn's idea of constant weights by splitting up the inner sum in (12.12) into many parts and attaching different constant weights to each part. This method is highly effective but is, on the other hand, very complicated. Roughly speaking, it may be described as splitting the inner sum into two parts and for one portion attaching constant weights which approximate to the logarithmic weight (thereby achieving a smoothing) while in the other portion the attached weights approximate to a smoothing in the opposite direction.

Chen's ingenious idea, for which we refer to Chapter 13, leading to an improvement in respect of some very prominent problems in additive prime number theory does not need any of the more sophisticated weighted sieves described above, but just the very special Kuhn's sieve (in the form stated in Theorem 12.2).

However, contrary to the other sieves methods mentioned above, this method cannot be applied to a great variety of sieves problems. It can be directly applied to the problems of the type

$$N = p + P_2 . p + h = P_2, \quad ap + b = P_2 \tag{12.40}$$

(cf. Theorem 13.1, 13.2; the last one (cf. Theorem 10.2) according to Halberstam (oral communication)). The problem of attacking other related problems by this method has not yet been tackled (cf. Notes of Chapter 13), and in this context the logarithmic weight procedure still gives the best results known to date. In case of dimension $\kappa = 1$ we mention (Richert [1]):

Let $F(n)$ be an irreducible polynomial of degree $g(\geq 1)$ with integer coefficients. Let $\rho(p)$ denote the number of solutions of the congruence

$$F(m) \equiv 0 \bmod p. \tag{12.41}$$

and suppose that 154

$$\rho(p) < p \text{ for all } p. \tag{12.42}$$

Then, we have

$$|\{n : 1 \leq n \leq x, F(n) = P_{g+1}\}| \geq \frac{2}{3} \prod_p \frac{(1 - \frac{\zeta(p)}{p})}{(1 - \frac{1}{p})} \frac{x}{\log x} \text{ for } x \geq x_\circ(F) \tag{12.43}$$

and if further

$$\rho(p) < p - 1 \text{ for } p \nmid F(0) \text{ and } p \leq g + 1, \tag{12.44}$$

then (excluding the case $F(n) = \pm n$) we also have

$$|\{p : p \leq x, F(p) = P_{2g+1}\}| \geq \frac{4}{3} \prod_{p \nmid F(0)} \frac{(1 - \frac{\rho(p)}{p-1})}{(1 - \frac{1}{p})}$$
$$\prod_{p \mid F(0)} \frac{(1 - \frac{\rho(p)-1}{p-1})}{(1 - \frac{1}{p})} \frac{x}{\log^2 x} \quad \text{for} \quad x \geq x_0(F). \tag{12.45}$$

Hence, in particular (if $\lim\limits_{x\to\infty}\dfrac{x_0(F)}{x} = 0$), there are infinitely many natural numbers n such that

$$F(n) = P_{g+1}. \tag{12.46}$$

and also infinitely many primes p such that

$$F(p) = P_{2g+1}. \tag{12.47}$$

(Note that for (12.47) there is no need to exclude the case $F(n) = \pm n$ as was done for (12.45).)

The corresponding problems for polynomials in two variables $F(m,n)$ canceled more successfully. In this connection we refer, for instance, to the papers of Greaves [1], [2], Iwaniec [2], [3], [5], [6], and Huxley and Iwaniec [1].

Regarding the question of almost primes in arithmetic progressions, i.e.,

$$P_r \equiv l \bmod k, (l,k) = 1, \tag{12.48}$$

155

Motohashi has proved, by averaging the logarithmic weights of Richert [1], that there is a

$$P_2 \leq k^{1.1}, \tag{12.49}$$

and that there is a

$$P_3 \leq k \log^{70} k \tag{12.50}$$

for almost all k and also the corresponding results valid for almost all $\ell \bmod k (k \to \infty)$ (Motohashi [12], [13], [15]). Without any exceptions we have only the existence of

$$P_2 \ll k^{2.2}. \tag{12.51}$$

$$P_3 \ll k^{11/7} \tag{12.52}$$

and

$$P_r \ll k^{1+\frac{1}{r-\frac{9}{7}}} \quad \text{for} \quad r > 2. \tag{12.53}$$

in (12.48). (Actually, here the exponents can be replaced by slightly smaller ones.) (Richert [1]. cf. Halberstam and Richert [1] (Theorem

9.6)). These results be compared with the corresponding ones for primes (cf. the remark following (6.55)).

For more details and various other applications of weighted sieves we refer to Halberstam and Richert [1] (Chapter 9 and 10).

As to the literature pertaining to the small sieves in general we refer to the extensive list of references given in Halberstam and Richert [1]. We take this opportunity to add the following list of papers, which are neither included there nor are mentioned in these lectures so far:

Bombieri [7], Buchstab [3], Elliott [3], Hall [1], Hooley [3], [10]. **156**
Meijer [1], [2], Ramachandra [5], Scourfield [1], Wolke [4], [8].

Lastly, we take up now again (cf. Notes of Chapters 7, 8 and 9) the question of comparison of the large sieves and the small sieves. The importance of the large sieves for applications in analytic number theory should be clear from the exposition in the Chapters 2 through 6, and it can hardly be overestimated. Also in arithmetical questions the large sieves turns out to be powerful when applied, in an auxiliary capacity, along with 'small' sieves, not only via the Bombieri-type results of Chapter 6, **3** (cf. our condition $(R(1, \alpha))$, (11.62)), and under (12.3) in these notes above) but also in a wider sense as is reflected in the proof of Chen's theorem (cf. Theorem 10.3 and the beginning of notes for Chapter 13). It is only in its arithmetical version (cf.Chapter 7) that the large sieve, even in its most powerful form (namely, the weighted Montgomery-Vaughan sieve) suffers from some deficiencies when compared with Selberg's sieve (in particular, when compared with the weighted form of the latter). Without repeating our remarks, made in the Notes of Chapter 7, 8 and 9, we mention only the following facts: The large sieve (for example, Theorem 7.1) can be used (as it is) to obtain only upper bounds, while Buchstab's method (cf. Lemma 11.1) provides (at least in principle)a corresponding lower bound, however, for $\dfrac{\log N}{\log z} > 2$. Theorem 11.13 (for $X = N$) gives a better estimate than does Theorem 7.1. Therefore, a suitably iterated form of (7.9) should be investigated. With respect to the function $\omega(p)$ the large sieve has the decisive advantage over the Selberg's sieve in that Theorem 7.1 imposes no restriction on the order of magnitude of $\omega(p)$, so in particular $\omega(p)$ need not be bounded on the 'average' as required by our condition (Ω_2) **157** of the small sieve. However, the large sieve cannot deal with the very important case $\omega(p) = \dfrac{p}{p-1}$, for example with the problem

$$F(p) = P_r. \tag{12.54}$$

Here, the defect steams from the fact that the large sieves basically requires that n has to run through a sequence of *consecutive* integers. Thus the large sieve, while applicable to the problem

$$F(n) = P_r. \tag{12.55}$$

(leading to a trivial estimate when applied directly to sieve the sequence $\{F(p)\}$ requires that one has to sieve the sequence $\{nF(n)\}$ for the problem (12.54), but then (cf. our remarks following (11.14)) the constant in the upper estimate is worsened by a factor 2.

Chapter 13

On Goldbach's Conjecture and Prime-Twins

NOW WE have prepared the ground for providing the main results of (the second part of) these lectures. These results represent the best approximation (in case sense) to the two most prominent problems in the additive theory of prime numbers (namely, those mentioned in the title of this chapter). The first proof of these results is due to Chen [1] (for simplifications see Ross [1]. cf. Malberstan and Richert [1] (Chapter 11)). Here we shall give a proof with further simplifications, in respect of (numerical) calculations and more specifically, involving only elementary functions. Actually this is accomplished at the expense of the quality of the constants ($c_\circ$) occurring in these results, which do not however affect the qualitative statement of these results. For the proof we use the special form (as given in Theorem 12.2) of Kuhn's sieve and, instead of the earlier practices of translating this theorem into the language of additive number theory, follow the idea of Chen by subtracting an additional term. Then, as one would expect, reformulate the problem for an application of Theorem 10.3, thereby completing the proof. **158**

We start with (cf. (12.37), (10.16))

$$G(N) := \sum_{\substack{p<N-1 \\ (N-p,\prod\limits_{p'<N^{1/8}} p')=1 \\ \mu(N-p)\neq 0}} \{1 - \frac{1}{2} \sum_{\substack{N^{1/8}\leq p'<N^{1/3} \\ p'|N-p}} -\frac{1}{2} \sum_{\substack{p_1p_2p_3\leq N \\ N^{1/8}\leq p_1<N^{1/3}\leq p_2p_3 \\ N-p=p_1p_2p_3}} . \quad (13.1)$$

As we shall see later, $G(N) > 0$ implies (in the direction of Goldbach's conjecture) that

$$N = p + P_2 \quad (13.2)$$

159 is soluble. One can even obtain a lower bound for $G(N)$ (when N is even and large) in terms of N, by means of Theorems 12.2 and 10.3. Indeed, with suitable choices for the parameters in these theorems, namely

$$X = N, 2|N \in \mathbb{N}, v = 8 \quad (13.3)$$

in both of them and further

$$h = -N, u = 3, \lambda = 2, \quad (13.4)$$

in Theorem 12.2, while in Theorem 10.3

$$h = N, \quad (13.5)$$

one obtains an estimate for an expression similar to $G(N)$. Actually, then one would have

$$\begin{cases} \sum\limits_{\substack{p\leq N \\ (N-p,\prod\limits_{\substack{p'<N^{1/8} \\ p'\nmid N}} p')=1}} \{1 - \frac{1}{2} \sum\limits_{\substack{N^{1/8}\leq p'<N^{1/3} \\ p'|N-p \\ p'\nmid N}} 1 - \frac{1}{2} \sum\limits_{\substack{p_1p_2p_3\leq N \\ N^{1/8}\leq p_1<N^{1/3}\leq p_2<p_3 \\ N-p=p_1p_2p_3}} \geq \\ \geq 4\mathfrak{S}(N)\frac{N}{\log^2 N}\{\log 3 - \frac{1}{2}\log 6 - \frac{1}{2}c(8) + 0(\log N)^{-\frac{1}{15}}\}. \end{cases} \quad (13.6)$$

Now a comparison of the left-hand side here with $G(N)$ shows that we need the estimate

$$1 + \sum_{p'|N} 1 + \sum_{N^{1/8}\leq p'\leq N^{\frac{1}{2}}} \sum_{\substack{p\leq N \\ N-p=0 \bmod p'^2}} 1 \ll \log N + \sum_{N^{1/8}\leq p} \frac{N}{p'^2} \ll N^{7/8}. \quad (13.7)$$

which provides a bound for the (difference in) contributions to these expressions arising from (the possible $p = N - 1$,) the numbers $N - p$ satisfying either of the conditions

$$(N - p, N) > 1 \text{ and } \mu(N - p) = 0 \tag{13.8}$$

Hence we have $G(N)$ also the same lower estimate given by the right-hand side of (13.6). Since, by (10.79) with $v = 8$, 160

$$4\left\{\log\ 3 - \frac{1}{2}\log 6 - \frac{1}{2}c(8)\right\} = 2\log\frac{3}{2} - 2c(8) \geq 2\log \frac{3}{2} - \frac{3}{2}\log\frac{64}{39}(> \frac{1}{2}\log\frac{3^7}{2.10^3} =; c_0) > 0, \tag{13.9}$$

we see that one has (for instance) the lower bound

$$G(N) \geq c_0 \mathfrak{S}(N)\frac{N}{\log^2 N} \quad \text{for} \quad N \geq N_0, \tag{13.10}$$

where N_0 is some absolute constant.

Next, we elaborate on the remark above pertaining to (13.12). To being with observe that $G(N)$ does not exceed the part, ($G^*(N)$ *say*, of its defining sum comprising only of all the *positive* (i,e., > 0) terms. And also note that any term of (13.1) with its second inner sum at least 1 makes the first sum, accompanying it, ≥ 1 so that such a term is ≤ 0. Thus we see that the terms occurring in $G^*(N)$ have their second inner sum empty and the first one is atmost 1. In otherwords, one (since $\mu(N - p) \neq 0$)

$$\begin{cases} G(N) \leq G^*(N) \leq \frac{1}{2}|\{p \leq N : N - p = p', \\ N^{\frac{1}{8}} \leq p' < N^{\frac{1}{3}}\} \cup \{p \leq N : N - p = p'p_2, \\ N^{1/8} \leq p' < N^{\frac{1}{3}} \leq p_2\} + |\{p \leq N : N - p = p_1, \\ p_1 \geq N^{\frac{1}{3}}\} \cup \{p \leq N : N - p = p_1 p_2, N^{\frac{1}{3}} \leq p_1 < p_2\}| \end{cases} \tag{13.11}$$

and also

$$G(N) \leq |\{p \leq N : N - p = P_2\}|. \tag{13.12}$$

From (13.10), (13.12) we obtain the desired

Theorem 13.1. *There is an absolute constant N_0 such that for all even numbers $N \geq N_0$, we have*

$$|\{p \leq N : N - p = P_2\} \leq c_0 \mathbb{C}(N) \frac{N}{\log^2 N}, \quad (c_0 > 0), \tag{13.13}$$

where c_0 may be taken constant defined in (13.9)*; in particular, there is*
161 *always a solution of the equation*

$$N = p + P_2 \quad \text{if} \quad 2|N \quad \text{and} \quad N \geq N_0.$$

The proof of the corresponding result regarding the generalized prime-twins proceeds analogously. Now we choose again (cf. (13.3))

$$x, 2|h, 0 < |h| \leq x^{\frac{1}{3}}, v = 8 \tag{13.14}$$

in both the Theorems 12.2, 10.3 and further

$$u = 3, v = 8, \lambda = 2 \tag{13.15}$$

in Theorem 12.2. Also we note that the sum $\sum\limits_{p \leq x}$ on the left-hand side of (12.37) can be replaced by $\sum\limits_{p+h \leq x}$ (cf. (13.14)) apart from a negligible error of the order $O(x^{\frac{1}{3}})$. Lastly, by (13.14), we observe that the p's counted in (10.16) satisfy (irrespective of the sign of h) $p + h \leq x$ and also $p + h = p_1 p_2 p_3$.

Now, by the choices (13.14), (13.15) for the parameters, it follows from Theorems 12.2 and 10.3 (corresponding to (13.6))

$$\begin{cases} \sum\limits_{\substack{p+h \leq x \\ (p+h, \prod\limits_{\substack{p' < x^{1/8} \\ p' \nmid h}} p') = 1}} \left\{ 1 - \frac{1}{2} \sum\limits_{\substack{x^{1/8} \leq p' < x^{1/3} \\ p' | p+h \\ p' \nmid h}} 1 - \frac{1}{2} \sum\limits_{\substack{p_1 p_2 p_3 \leq x \\ x^{1/8} \leq p_1 < x^{1/3} \leq p_2 < p_3 \\ p+h = p_1 p_2 p_3}} 1 \right\} \\ \geq 4 \mathfrak{S}(h) \frac{x}{\log^2 x} \left\{ \frac{1}{2} \log \frac{3}{2} - \frac{1}{2} c(8) + O((\log x)^{-\frac{1}{15}}) \right\}. \end{cases} \tag{13.16}$$

From here, on using the argument of (13.7) and (13.9), one obtains (for the analogue of $G(N)$ the lower bound

$$\begin{cases} T_h(x) = \sum\limits_{\substack{p+h\le x \\ (p+h, \prod\limits_{p'<x^{1/8}} p')=1 \\ \mu(p+h)\neq 0}} \left\{1 - \frac{1}{2} \sum\limits_{\substack{x^{1/8}\le p'<x^{1/3} \\ p'|p+h}} 1 - \frac{1}{2} \sum\limits_{\substack{p_1p_2p_3\le x \\ x^{1/8}\le p_1<x^{1/3}\le p_2<p_3 \\ p+h=p_1p_2p_3}} 1\right\} \\ \ge c_0\mathfrak{S}(h)\frac{x}{\log^2 x} \text{ for } x\ge x_0. \end{cases} \tag{13.17}$$

Again the reasoning leading to (13.11), and so also to (13.12), is applicable with $p + h$ in place of $N - p$. Thus one arrives at the final **162**

Theorem 13.2. *There is an absolute constant x_0 such that for any even number h (determined with respect to x) satisfying*

$$0 < |h| \le x^{\frac{1}{3}}, \tag{13.18}$$

we have

$$|\{p + h \le x : p + h = P_2\}| c_0\mathfrak{S}(h)\frac{x}{\log^2 x} for x \ge x_0, (c_0 > 0). \tag{13.19}$$

where c_0 (again) may be taken as the constant defined in (13.9)*; is particular, for any non-zero even number h, there are infinitely many primes p such that*

$$p + h = p_2. \tag{13.20}$$

NOTES

Chen's theorem affords a beautiful instance of the effective use of various powerful tools of numbers theorey. As an inspection of its proof (and those of Theorems 12.2 and 10.3) discloses we have employed Kuhn's sieve, Selberg's sieve (several times), Bombieri's prime number theorem, Siegel-Walfisz theorem, contour integration, the hybrid form of the large sieve and Chen's new idea described below.

Chen's idea, which has already been briefly indicated at the beginning of this Chapter, can be described as follows. (We shall confine ourselves, for this purpose, to Theorem 13.1) One first sifts the sequence $N - p$ so that the remaining numbers satisfy.

$$N - p = P_3 \tag{13.21}$$

(and then the numbers of these remaining ones is estimated from below. cf. Theorem 12.2). Now to remove from the rest those which are of the
163 form $p_1 p_2 p_3$, one subtracts (from the preceding lower bound) another sieve estimate (from above) for the numbers of solutions of

$$N - p_1 p_2 p_3 = p \tag{13.22}$$

(cf. Theorem 10.3). If, as has been shown to be the case for the problems under consideration, the lower bound exceeds the upper estimate (for (13.22)), it follows then that (since the surviving members of $\{N - p\}$ are all now $P_2's$) there must be solutions of the equation

$$N = p + P_2. \tag{13.23}$$

This procedure of inverting the equation (13.21) to (13.22) has turned out to be much more fruitful than any further known improvement of the sieve method one has started with. Other successful applications of this last step have been given by Indlekofer [2], Huxley and Iwaniec [1].

In this context we recall the remarks preceding (12.40). To briefly expand on that statement, we mention that to attempt an improvement upon (12.46) (or (12.47)) by Chen's method one would require (cf. (13.21) and (13.22)), limiting ourselves now to the simplest case, for

$$n^2 + 1 = P_2 \tag{13.24}$$

a satisfactory upper bound for the number of solutions of

$$p_1 p_2 p_3 - 1 = n^2 \tag{13.25}$$

with p_i's restricted by some conditions (like those in Theorem 10.3). Surprisingly we do not any method of obtaining a satisfactory estimate for the number of square in a sequence under appropriate conditions, and specifically not even in this case. An explanation would clearly be
164 that the sequence of squares, even though more regularly distributed, is much thinner than the sequence of primes. This also indicated that the corresponding problem with respect to (12.46) (or (12.47) is much complicated.

Theorems 13.1 and 13.2: It is possible to state Theorem 13.1 in more precise form with respect to the prime factors of P_2, as can be seen from a combination of (13.10) and (13.11). A similar remarks applies to Theorem 13.2 also.

Regarding the constant c_0 (cf. (13.13), (13.19)) in these theorems we note that our value of

$$c_0 \geq 0.0446 \tag{13.26}$$

(cf. (13.9)) is rather small. One of the reasons for this is the choice $v = 8$, which was made to enable us to deal (cf. (11.73), (11.74)) with elementary functions only (and also to simplify numerical calculations), as has been mentioned in the introductions to this chapter, a convenience not available under the better choice (of all earlier proofs) of $v = 10$. Of course, we can replace (13.26) by (cf. (13.9))

$$c_0 \geq 2\log\frac{3}{2} - \frac{3}{2}\log\frac{64}{39} \geq 0.0679, \tag{13.27}$$

but a better constant can be obtained by taking $v = 10$. With this later choice, Chen [1] obtained

$$c_0 \geq 0.3354 \tag{13.28}$$

and any numerical integration one can even get

$$c_0 \geq 0.3445 \tag{13.29}$$

(see Halberstam and Richert [1] (p. 338)). It returns out that $v = 11$ is close to the optimal choice, by considering the non-elementary functions F, f in a wider range, and then one is led to

$$c_0 \geq 0.3716. \tag{13.30}$$

165

In all these cases u is kept fixed to be 3, a value convenient in the arithmetical interpretation of the estimates of the weighted sieve. The constant c_0 can further be slightly improved by taking $u > 3$ but then

Chen's procedure becomes more complicated. As to the constant c_0 under consideration it is worthwhile to compare it with 1, in view of the conjecture of Hardy and Littlewood mentioned (under Theorem 10.1) in the Notes for Chapter 10. This suggests that our present constant c_0 should be capable of much further improvement. In this context, we add that one has theorems (corresponding to Theorems 13.1 and 13.2) with P_2's in (13.13) and (13.19) replaced by P_3 for a better (corresponding) constant

$$c_0 \geq \frac{13}{6} \tag{13.31}$$

(cf. Helberstam and Richert [1] (Theorem 9.2)). Comparing the methods of proof it is considered to be likely that the constant c_0 in Chen's theorem can be further improved by using the logarithmic weights instead of Kuhn's weights.

Continuing on with related questions we mention now a few results concerning Goldbach numbers (namely, those even numbers N which can be written as sums of two primes)

$$N = p + p' \ (2|N). \tag{13.32}$$

Ramachandra [4] has derived from an estimate of (6.11)-type, actually from (the uniform)

$$\sum_{\chi} N(\sigma, T, \chi) \ll (q^2 T)^{g(1-\sigma)} \log^{14}(qT) \text{ for } \frac{1}{2} \leq \sigma \leq 1, \tag{13.33}$$

that the numbers of Goldbach numbers in the interval

$$x \leq N \leq x + x^{\lambda} \tag{13.34}$$

166 has the asymptotic formula

$$\frac{1}{2}x^{\lambda} + O_{\lambda,A}(x^{\lambda} \log^{-A} x) \ \text{ as } \ x \to \infty, \ \ \textit{if} \ (1 \geq)\lambda > 1 - \frac{1}{g}, \tag{13.35}$$

and has also deduced from (13.35), by combining with a result of Montgomery [3], that (for $x \geq x_0$) there is always a prime p in the interval (13.34) such that both $p + 1$ and $p - 1$ are Goldbach numbers.

Many results have been proved (in various forms) in order to show that ‘almost all’ (with respect to the error-term corresponding to that of (13.15) even numbers are Goldbach numbers. The best result known here, upto this time, is due to Montgomery an Vaughan [3] and states that the number of even integers $N \leq x$, which are not Goldbach numbers is

$$\ll x^{1-\delta}, \text{ for some } \delta > 0. \tag{13.36}$$

(For a previous result, see Vaughan [2].)

Further, turning to problems allied to Theorem 13.2, we have (cf. (12.40)) the observation of Vaughan [3] which yields, when combined with Chen’s method the following result: Either the equation

$$2p + 1 = p' \tag{13.37}$$

has infinitely many solutions or

$$2p + 1 = p_1 p_2 \tag{13.38}$$

has infinitely many solutions, in which extend event one , in particular, infinitely many solutions of the equation

$$d(n + 1) = d(n). \tag{13.39}$$

This statement concerning (13.39) is a conjecture of Erdös and Mirsky [1] (cf. Helberstam and Richert [1] (p. 338)). **167**

Lastly we mention an application, due to Jutila [1], to a question allied to Theorem 13.2. He deduced from his result (6.26), by combining it with a theorem of Levin [2] (reference to Richert [1] would permit to replace ‘8’ by ‘7’ in the following statement), that *for every integer* $r \geq 8$ there exists a numbers $\theta(r)$ satisfying (with c an in (6.28))

$$\left(c^* := \frac{1 + 4c}{2 + 4c} \leq\right) \theta(r) < 1 \tag{13.40}$$

(and $\theta(r)$ decreasing to C^* for increasing r), such that

$$x < p < x + x^{\theta(r)}, p + 2 = P_r, \text{ for all } x \geq x_0, \tag{13.41}$$

is soluble. Also he started that a similar result can be derived for (an almost-) Goldbach problems, which may be interpreted as that the equation

$$N = p + P_r, N \geq N_0, 2|N, r \geq 8, \tag{13.42}$$

has a solution in two 'almost equal' (-in a sense which is stronger for larger r-) numbers p, P_r.

Bibliography

168

[1] ANDRUHAEV, H.M.

1. The addition problem for prime and near-prime numbers in algebraic number fields. (Russian) *DOkal. Akad. Nauk SSSR 159* (1964), 1207-1209 = *Soviet Math. Dokl. 5* (1964), 1666-1668, MR *30*, 1116.

[2] ANKENY, N.C. and ONISHI, H.

1. The general sieve. *Acta Arith*. *10* (1964/65), 31-62. MR *29*, 4740.

[3] BALASUBRAMANIAN, R.

1. An improvement of a theorem of Titchmarsh on the mean square of $|\zeta(\frac{1}{2}+it)|$. *Proc. London Math. Soc*. (to appear).

[4] BALASUBRAMANIAN, R. and RAMACHANDRA, K.

1. Two remarks on a result of Ramachandra. *J. Indian Math. soc*. (N.S.) *38* (1974), 395-397.

[5] BARBAN, M.B.

1. Arithmetic functions on 'rare' set. (Russian) *Doki. Akad. Nauk* Uz SSR *1961*, no. 8, 10-12.
2. New applications of the 'great sieve' of Yu. V. Linnik. (Russian) *Akad. Nauk* Uzbek. SSR *Trudy Inst. Mat*. No. *22* (1961), 1-20. MR *30*, 1990.

3. Linnik's "great sieve" and a limit theorem for the class number of ideals of an imaginary quadratic field. (Russian) *Izv. Akad. Nauk* SSSR *Ser. Mat. 26* (1962), 573-580. MR *27,* 1426.

4. The density of zeros of Dirichlet *L*-series and the problems of the addition of primes and almost primes. (Russian) *Dokl. Akad. Nauk* Uz SSR *1963,* no.1, 9-10.

5. The "density" of the zeros of Dirichlet *L*-series and the problems of the sum of primes and "near primes" (Russian) *Mat. Sbornik* (N.S.) *61 (103)* (1963), 418-425. MR *30*, 1992.

6. Analogues of the divisor problem of Titchmarsh. (Russian. English summary) *Vestnik Leningrad. Univ. Ser. Mat. Meh. Astronom*. *18* (1963), no. 4, 5-13. MR *28*, 57.

7. Multiplicative functions of $\sum_R$ - equidistributed sequences. (Russian. Uzbek summary) *Izv, Akad. Nauk* Uz SSR Ser. Fiz. - Mat. Nauk *1964* no. 6, 13-19. MR *31*, 1239.

8. On a theorem of P. Bateman, S. Chowla and P. Erdös. (Russian. English summary) *Magyar Tud. Akad. Mat, Kutató Int . Közl*. *9* (1964), 429-435. MR *33*, 109.

169 9. A remark on the author's paper "New applications of the 'large sieve' of Yu. V. Linnik". (Russian) *Theory Probability Math. Static*. (Russian), pp. 130-133. *Izdat. "Nauka" Uzbek. SSR, Tashkent, 1964*, MR *33*. 4037.

10. The "large sieve" method and its application to number theorey. (Russian) *Uspehi Mat. Nauk 21* (1966), no. 1 (127), 51-102 = *Russian Math, Surveys 21* (1966), no. 1, 49-103. MR *33*, 7320.

[6] BARBAN, M.B. and VEHOV, P.P.

1. An external problem. (Russian) *Trudy Moskov. Mat. Obšč*. *18* (1968), 83-90 = *Trans. Moscow Math. Soc*. *18* (1968), 91-99. MR *38*, 4410.

[7] BATEMAN, P.T., CHOWLA, S. and ERDÖS, P.

1. Remarks on the size of $L(1,\chi)$. *Publ. Math. Debrecen 1* (1950), 165-182. MR *12*, 244; *13*, 1138.

[8] BELLMAN, R.

1. Almost orthogonal series. *Bull. Amer. Math. Soc*. (2) *50* (1944), 517-519. MR *6*, p. 48.

[9] BOAS, R.P.

1. A general moment problem. *Amer. J. Math*. *63* (1941), 361-370. MR *2*, p. 281.

[10] BOMBIERI, E.

1. On the large sieve. *Mathematika 12* (1965), 201-225. MR *33*, 5590.
2. On a theorem of van Lint and Richert. *Symposia Mathematica* Vol. *IV* (INDAM, Rome, 1968/69), 175-180. MR *43*, 4791.
3. Density theorems for the zeta function. *Proc. Sympos. Pure Math. 20* (1969). 352-358. MR *47*, 4945.
4. A Note on the large sieve. *Acta Arith*. *18* (1971), 401-404. MR *44*, 3982.
5. On large sieve inequalities and their applications. *Trudy Mat. Inst. Steklov*. *132* (1973), 251-256.
6. Le grand crible dans la theorie analytique des nombres. *Soc. Math. France, Asterisque* No. *18*. 1974.
7. On twin almost primes. *Acta Arith*. *28* (1975), 177-193.
8. Corrigendum to my paper “On twin almost primes” and an addendum on Selberg’s sieve. *Acta Arith*. *28* (1976), 457-461.

[11] BOMBIERJ, E. and DAVENPORT, H.

1. Small differences between prime numbers. *Proc. Roy. Soc*. Ser. A *293* (1966), 1-18. MR *33*, 7314.

2. *On the large sieve method*. Number theory and Analysis (Papers in Honor of Edmund Landau) pp. 9-22. New York, 1969. MR *41*, 5327. 170

3. Some inequalities involving trigonometrical polynomials. *Ann. Scuola Norm. Sup. Pisa* (3) *23* (1969), 223-241. MR *40*, 2636.

[12] BUCHSTAB, A.A.

1. New results in the investigation of the Goldbach-Euler problem and the problem of prime pairs. (Russian) *Dokl. Akod. Nauk* SSSR *162* (1965).735-738 = *Soviet Math. Dokl. 6* (1965), 729-732. MR *31*, 2226.

2. Combinatorial strengthening of the sieve method of Eratosthenes. (Russian) *Uspehi Mat. Nauk 22* (1967), no. 3 (135), 199-226 = *Russain Math. Surveys 22* (1967), no. 3, 205-233, MR *36*, 1413.

3. A simplefied modification of the combinatorial sieve. (Russian) *Moskov. Gos. Ped. Inst, Učen Zap.* 375 (1971), 187-194. MR *48*, 11019.

[13] BURGESS, D.A.

1. The average of the least primitive root modulo p^2. *Acta Arith. 18* (1971), 263-271. MR *45*, 212.

[14] BURGESS, D.A and ELLIOTT, P.D.T.A.

1. The average of the least primitive root. *Mathematika 15* (1968), 39-50. MR *38*, 5736.

[15] CHEN, J.

1. On the representation of a large even integer as the sum of a prime and the product of atmost two primes. *Sci. Sinica 16* (1973), 157-176.

[16] CHOI, S.L.G.

1. On a theorem of Roth. *Math. Ann. 179* (1969), 319-328. MR *39*, 1421.

[17] CROFT, M.J.

1. Square-free numbers in a arithmetic progressions. *Proc. London Math. Soc*. (3) *30* (1975). 143-159.

[18] DAVENPORT, H.

1. *Multiplicative numbers theorey*. Chicago , III., 1967, vii + 189 pp. MR *36*, 117.

[19] DAVENPORT, H. and HALBERSTAM, H.

1. The values of a trigionometrical polynomial at well spaced points. *Mathematika 13* (1966), 91-96. MR *33*, 5592; *Corrigendum and Addendum: ibid. 14* (1967), 229-232. MR *36*, 2569.
2. Primes in arithmetic progressions. *Michigan Math. J. 13* (1966), 485-489, MR *34*, 156; *Corrigendum: ibid. 15* (1968), 505. MR *38*, 2099.

[20] ELLIOTT, P.D.T.A. **171**

1. On the size of $L(1,\chi)$. *J. Reine Angew. Math. 236* (1969), 26-36. MR *40*, 2619.
2. The distribution of primitive roots. *Canad. J. Math. 21* (1969), 822-841. MR *40*, 104.
3. A restricted mean value theorem. *J. London Math. Soc.* (2) *1* (1969), 447-460 MR *40*, 111.
4. On mean value of $f(p)$. *Proc. London Math. Soc.* (3) *21* (1970), 28-96. MR *42*, 1783.
5. The Turan-Kubilius inequality, and a limitation theorem for the large sieve. *Amet. J. Math. 92* (1970), 293-300. MR *41*, 8360.

6. Some applications of a theorem of Raikov to number theory. *J. Number Theory 2* (1970), 22-55. MR *41*, 3422.
7. One inequalities of large sieve type. *Acta Arith. 18* (1971), 405-422. MR *44*, 3983.
8. On the distribution of arg $L(s,\chi)$ in the half-plane $\sigma > \frac{1}{2}$. *Acta Arith. 20* (1972), 155-169. MR *48*, 2090.
9. On the distribution of the values of Dirichlet L-series in the half-plane $\sigma > \frac{1}{2}$. *Indag. Math. 33* (1971), 222-234. MR *45*, 194.

[21] ELLIOTT, P.D.T.A. and HALBERSTAM, H.

1. Some applications of Bombieri's theorem. *Mathematika 13* (1966), 196-203. MR *34*, 5788.
2. A conjecture in prime number theory. *Symposia Mathematica* Vol. *IV* (INDAM, Rome, 1968/69), 59-72. MR *43*, 1943.

[22] ERDÖS, P.

1. Remarks on number theorey. I. (Hungarian. Russian and English summaries) *Mat. Lapok 12* (1961), 10-17. MR *26*, 2410.
2. Some remarks on number theory. III. (Hungarian. Russian and English summaries) *Mat. Lapok 13* (1962), 28-38. MR *26*, 2412.
3. Remarks on number theorey, V. (Hungarian. English summary) *Mat. Lapok 17* (1966), 135-155. MR *36*, 133.

[23] ERDÖS, P. and MIRSKY, L.

1. The distribution of values of the divisor function $d(n)$. *Proc. London Math. Soc.* (3) *2* (1952), 257-271. MR *14*, p. 249.

[24] ERDÖS, P. RÉNYI, A.

1. Some remarks on the large sieve of Yu. V. Linnik. *Ann. Univ. Sci. Budapest. Eötvös Sect. Math. 11* (1968). 3-13. MR 39, 2718.

[25] ESTERMANN, T. **172**

1. On an asymptotic formula due to Titchmarsh. *J. London Math. Soc. 6* (1931), 250-251.

[26] FOGELS, E.K.

1. Analogue of Brun-Titchmarsh theorem. (Russian. Latvian summary) *Latvijas PSR Zinātnu Akad. Fiz. Mat. Inst. Raksti 2* (1950), 46-58. MR *13*, p. 725.

[27] FORTI. M. and VIOLA, C.

1. On large sieve type estimated for the Direchlet series operator. *Proc. Sympos. Pure Math. 24* (1973), 31-49.
2. Density estimates for the zeros of *L*-functions. *Acta Arith. 23* (1973), 379-391. MR *51*, 3081.

[28] GALLAGHER, P.X.

1. The large sieve. *Mathematika 14* (1967), 14-20. MR *35*, 5411.
2. Bombieri's mean value theorem. *Mathematika 15* (1968), 1-6. MR *38*, 5724.
3. A larger sieve. *Acta Arith. 18* (1971), 77-81. MR *45*, 214.
4. A large sieve density estimate near $\sigma = 1$. *Invent. Math. 11* (1970), 329-339. MR *43*, 4775.
5. Primes in progressions to prime-power modulus. *Invent. Math. 16* (1972), 191-201. MR *46*, 3462.
6. Sieving by prime powers. *Proc. of the Number Theory Conference* (Univ. Colorado, Boulder, Colo., 1972) 95-99. MR *50*, 4457.
7. Sieving by prime powers. *Acta Arith. 24* (1973/74), 491-497. MR *49*, 2613.
8. Local mean value and density estimates for Dirichlet *L*-functions. *Indag. Math. 37* (1975), 259-264. MR *51*, 5517.

[29] GELFOND, A.O. and LINNIK, Yu. V.

1. *Elementary methods in analytic number theory*. Moscow, 1962, 272 pp. = Chicago, 1965, xii + 242 pp. MR *32*, 5575 = Oxford - New York - Toronto, Ont. 1966, xi + 232 pp. MR *34*, 1252.

[30] GOLDFELD, M.

1. Artin's conjecture on the average. *Mathematika 15* (1968), 223-226. MR *39*, 2711.
2. On the number of primes p for which $p+a$ has a large prime factor. *Mathematika 16* (1969), 23-27. MR *39*, 5493.

173

3. A large sieve for a class of non-abelian L-functions. *Israel J. Math. 14* (1973), 39-49. MR *47*, 8489.
4. A further improvement of the Brun-Titchmarsh theorem. *J. London Math. Soc*. (2) *11* (1975), 434-444.

[31] GREAVES, G.

1. Large prime factors of binary forms. *J. Number Theory 3* (1971), 35-59. MR *42*, 5909.
2. An application of a theorem of Barban-Davenport and Halberstam. *Bull. London Math. Soc. 6* (1974), 1-9.

[32] HALÁSZ, G.

1. Über die Mittelwerte multiplilkativer zahlentheoretischer Funktionen. *Acta Math. Acad. Sci. Hungar. 19* (1968), 365-403. MR *37*, 6254.

[33] HALÁSZ, G. and TURÁN, P.

1. On the distribution of roots of Riemann zeta and allied functions. I. *J. Number Theory 1* (1969). 121-137. MR *38*, 4422.
2. On the distribution of roots of Riemann zeta and allied functions. II. *Acta Math. Acad. Sci. Hungar. 21* (1970), 403-419. MR *42*, 3035.

[34] HALBERSTAM, H.

1. Footnote to the Titchmarsh-Linnik divisor problem. *Proc. Amer. Math. Soc. 18* (1967), 187-188. MR *34*, 4221.

[35] HALBERSTAM, H., JURKAT, W. and RICHERT, H.-E.

1. Un nouveau résultat de la méthode du crible. *C.R. Acad. Sci. Paris* Sér. A-B *264* (1967), A920-A923. MR *36*, 6374.

[36] HALBERSTAM, H. and RICHERT, H.-E.

1. *Sieve methods*. London-New York - San Francisco, 1974, xiv + 364 pp.

[37] HALBERSTAM, H. and ROTH, K.F.

1. *Sequences*. Vol. I. Oxford, 1966. xx + 291 pp. MR *35*, 1565.

[38] HALL, R.R.

1. Halving an estimate obtained from Selberg's upper bound method. *Acta Arith. 25* (1973/74), 347-351. MR *49*, 4949.

[39] HLAWKA, E.

1. Bemerkungen zum grossen Sieb von Linnik. *Öterreich. Akad. Wiss. Math. - Natur. Kl.* S.-B. II. *178*(1970), 13-18. MR *42*, 224.
2. Zum grossen Sieb von Linnik. *Acta Arith. 27*, (1975), 89-100. MR *51*, 3091.

[40] HOOLEY, C. **174**

1. An asymptotic formula in the theory of numbers. *Proc. London Math. Soc.* (3) *7*(1957), 396-413. MR *19*, p. 839.
2. On the Brun-Titchmarsh theorem. *J. Reine Angew. Math. 255* (1972), 60-79. MR *46*, 3463.
3. On the intervals between numbers that are sums of two squares. III. *J. Reine Angew. Math. 267* (1974), 207-218.

4. On the greatest prime factor of $p+a$. *Mathematika 20* (1973), 135-143. MR *50*, 7057.

5. On the Barban-Davenport-Halberstam theorem. I. *J. Reine Angew. Math. 274/275* (1975), 206-223.

6. On the Brun-Titchmarsh theorem. II. *Proc. London Math. Soc.* (3) *30* (1975), 114-128. MR *51*, 5531.

7. On the Barban-Davenpoet-Halberstam theorem. II. *J. London Math. Soc.* (2) *9* (1975), 625-636.

8. On the Barban-Davenport-Halberstam theorem. III. *J. London Math. Soc*. (2) *10* (1975), 249-256.

9. On the Barban-Davenport-Halberstam theorem. IV. *J. London Math. Soc*. (2) *11* (1975), 399-407.

10. *Applications of sieve methods to the theory of numbers*. Cambridge, 1976, 122 pp.

[41] HUA, L.-K.

1. *Die Abschatzung von Exponentialsummen und ihre Anwendung in der Zahlentheorie*. Enzykl. math. Wiss., I. 2, Heft 13, Teil I, Leipzig, 1959. 123 pp. MR *24*, A94.

[42] HUXLEY, M.N.

1. The large sieve inequality for algebraic number fields. *Mathematika 15* (1968), 178-187. MR *38*. 5737.

2. On the differences of primes in arithmetical progressions. *Acta Arith. 15* (1968/69), 367-392. MR *39*, 5494.

3. The large sieve inequality for algebraic number fields. II. Means of moments of Hecke zeta-functions. *Proc. London Math. Soc*. (3) *21* (1970), 108-128, MR *42*, 5944.

4. The large sieve inequality for algebraic number fields. III. Zero-density results. *J. London Math. Soc*. (2) *3* (1971), 233-240. MR *43*, 1944.

175 5. Irregularity in sifted sequences. *J. Number Theory 4* (1972), 437-454. MR *47*, 180.

6. On the difference between consecutive primes. *Inven. Math. 15* (1972), 164-170. MR *45*, 1856.
7. *The distribution of prime numbers*. Oxford, 1972, x + 128 pp.
8. The difference between consecutive primes. *Proc. Sympos. Pure Math. 24* (1973), 141-145. MR *50*, 9816.
9. Large values of Dirichlet polynomials. *Acta Arith. 24* (1973), 329-346. MR *48*, 8406.
10. Small differences between consecutive primes. *Mathematika 20* (1973), 229-223. MR *50*, 4509.
11. Large values of Dirichlet polynomials. II. *Acta Arith. 27* (1975), 159-169. MR *50*, 12934.
12. Large values of Dirichlet polynomials. III. *Acta Arith. 26* (1975), 435-444.

[43] HUXLEY, M.N. and IWANIEC, H.

1. Bombieri's theorem in short intervals. *Mathematika 22* (1975), 188-194.

[44] HUXLEY, M.N. AND JUTILA, M.

1. Large values of Dirichlet polynomials. IV. *Acta Arith.* (to appear).

[45] INDLEKFER, K.-H.

1. Eine asymptotische Formel in der Zahlentheoric. *Arch. Math. 23* (1972), 619-624. MR *47*, 6629.
2. Scharfe untere Absch$\bar{\text{a}}$tzung f$\bar{\text{u}}$r die Anzahlfunktion der *B*-Zwillinge. *Acta Arith. 26* (1974/75), 207-212. MR *50*, 7064.

[46] IWANIEC, H.

1. On the error term in the linear sieve. *Acta Arith. 19* (1971), 1-30. MR *45*, 5104.

2. Primes of the type $\varphi(x, y) + A$ where φ is a quadratic form. *Acta Arith. 21* (1972), 203-234. MR *46*, 3466.

3. Primes represented by quadratic polynomials in two variables. *Acta Arith. 24* (1973/74), 435-459. MR *49*, 7210.

4. On zeraos of Dirichlet's L series. *Invent. Math. 23* (1974), 97-104. MR *49*, 8947.

5. The generalized Hardy-Littlewood's problem involving a quadratic polynomial with coprime discriminants. *Acta Arith. 27* (1975), 421-46. MR *51*, 3.

6. The half dimensional sieve. *Acta Arith. 29* (1976), 69-95.

7. The sieve of Eratosthenes-Legendre. (to appear).

176 [47] JOHNSEN, J.

1. On the large sieve method in $GF[q, x]$. *Mathematika 18* (1971), 172-184. MR *46*, 1761.

[48] JOSHI, P.T.

1. The size of $L(1, \chi)$ for real nonprincipal residue characters χ with prime modulus. *J. Number Theory 2* (1970), 58-73. MR *40*, 4215.

[49] JURKAT, W.B. and RICHERT, H.-E.

1. An improvement of Selberg's sieve method. I. *Acta Arith. 11* (1965), 217-240. MR *34*, 2540.

[50] JUTILA, M.

1. A statistical density theorem for L-functions with applications. *Acta Arith. 16* (1969/70), 207-216. MR *40*, 5557.

2. On character sums and class numbers. *J. Number Theory 5* (1973), 203-214. MR *49*, 230.

3. On numbers with a large prime factor. *J. Indian Math. Soc.* (N.S.) *37* (1973), 43-53. MR *50*, 12936.

4. A density estimate for *L*-functions with a real character. *Ann. Acad. Sci. Fenn*. Ser. AI No. 508 (1972), 10 pp. MR *46*, 1725.

5. On a density theorem of H.L. Montgomery for *L*-functions. *Ann. Acad. Sci. Fenn*, Ser. AI No. 520 (1972), 13 pp. MR *48*, 6023.

6. On mean values of Dirichlet polynomials with real characters. *Acta Arith. 27* (1975), 191-198.

7. On mean values of *L*-functions and short character sums with real characters. *Acta Arith. 26* (1975), 405-410.

8. *On the large values of Dirichlet polynomials*. Proceedings of the Colloquium on Number Theory (Debrecen, 1974).

9. *Zero-density estimates for L-functions*. Zahlentheorie (Tagung, Forachungzinst., Oberwolfach, 1975) (to appear).

10. Zero-density estimates for *L*-functions. *Acta Arith.* (to appear).

11. Zero-density estimates for *L*-functions. II *Acta Arith.* (to appear).

12. On Linnik's density theorem. (to appear).

[51] KATAI, I.

1. On an application of the large sieve: Shifted prime numbers, which have no prime divisors from a given arithmetical progression. *Acta Math. Acad. Sci. Hungar. 21* (1970), 151-173. MR *41*, 8373.

[52] KLIMOV, N.I. **177**

1. Upper estimates of some number theoretical functions. (Russian) *Dokl. Akad. Nauk* SSSR *111* (1956), 16-18. MR *19*, p. 251.

2. Combination of elementary and analytic methods in the theory of numbers. (Russian) *Uspehi Mat. Nauk 13* (1958), no. 3(81), 145-164.

[53] KOBAYASHI, I.

1. A note on the Selberg sieve and the large sieve. *Proc. Japan Acad. 49* (1973), 1-5. MR *48*, 3903.

[54] KOLESNIK, G.A.

1. An estimate for certain trigonometric sums. (Russian) *Acta Arith. 25* (1973/74), 7-30. MR *48*, 11002.

[55] KUBILIUS, I.P.

1. Probabilistic methods in the theory of numbers. (Russian) *Uspehi Mat. Nauk* (N.S.) *11* (1956), no. 2(68), 31-66. MR *18*, p. 17 = *Amer. Math. Soc. Transl*. (2) *19* (1962), 47-85. MR *24*, A 1266.
2. *Probabilistic methods in the theory of numbers*. Vilna, 1959. 164 pp. MR *23*, A 134; 2nd ed. Vilna, 1962, 221 pp. MR *26*, 3691 = Providence, R.I., 1964, xvii + 182 pp. MR *28*, 3956.

[56] KUHN, P.

1. Zur Viggo Brun'schen Siebmethode. I. *Norske Vid. Selsk. Forh. Trondhjem 14* (1941). no. 39, 145-148. MR *8*, p. 503.
2. *Neue Abschatzungen auf Grund der Viggo Brunschen Siebmethode*. 12. Skand. Mat. Kongr., Lund, 1953, 160-168. MR *16*, p. 676.
3. *Über die Primteiler eines Polynoms*. Proc. Internat. Congress Math., Amsterdam, 1954, *2*, 35-37.

[57] LAVRIK, A.F.

1. On the twin prime hypothesis of the theory of primes by the method of I.M. Vinogradov. (Russian) *Dokl. Akad. Nauk* SSSR *132* (1960), 1013-1015 = *Soviet Math. Dokl. 1* (1960), 700-702. MR *28*, 1183.

2. An approximate functional equation for the Dirichlet *L*-function. (Russian) *Trudy Moscov. Mat. Obšč. 18* (1968), 91-104 = *Trans. Moscow Math. Soc. 18* (1968), 101-115. MR *38*, 4424.

[58] LEVIN, B.V. **178**

1. Distribution of "near primes" in polynomial sequence. (Russian) *Mat. Sbornik* (N.S.) *61(103)* (1963), 389-407. MR *30*, 1991.
2. A one-dimensional sieve. (Russian) *Acta Arith. 10* (1964/65), 387-397. MR *31*, 4774.

[59] LINNIK, Yu. V.

1. "The large sieve". *C.R.(Doklady) Acad. Sci. URSS* (N.S.) *30* (1941), 292-294 (290-292). MR *2*, p. 349.
2. On Erdös's theorem on the addition of numerical sequences. (English. Russian summary) *Rec. Math. (Mat.Sbornik)* N.S. *10 (52)* (1942). 67-78. MR *4*, p. 131.
3. A remark on the least quadratic non-residue. *C.R. (Doklady) Acad. Sci. URSS* (N.S.) *36* (1942), 119-120. (131-132). MR *4*, p. 189.
4. *The dispersion method in binary additive problems*. Leningrad, 1961, 208 pp. MR *25*,3920 = Providence, R.I., 1963, x + 186 pp. MR *29*, 5804.

[60] VAN LINT, J.H. AND RICHERT, H.-E.

1. On primes in arithmetic progressions. *Acta Arith. 11* (1965), 209-216. MR *32*, 5613.

[61] LIU, M.-C.

1. On a result of Davenport and Halberstam. *J. Number Theory 1* (1969), 385-389. MR *40*, 2637.

[62] MATHEWS, K.R.

1. On an inequality of Davenport and Halberstam. *J. London Math. Soc*. (2) *4* (1972), 638-642. MR *46*, 1726.
2. On a bilinear form associated with the large sieve. *J. London Math. Soc*. (2) *5* (1972), 567-570. MR *47*, 6631.
3. Hermitian forms and the large and small sieves. *J. Number Theory 5* (1973), 16-23.

[63] MEIJER, H.G.

1. Sets of primes with intermediate density. *Math. Scand. 34* (1974), 37-43. MR *50*. 7082.

179 2. A problem on prime numbers. *Nieuw Arch. Wisk.* (3) *22* (1974), 125-127. MR *50*, 12937.

[64] MIECH,R.J

1. Almost primes generated by a polynomial. *Acta Arith. 10* (1964/65), 9-30. MR *29*, 1174.
2. A uniform result on almost primes. *Acta Arith. 11* (1966), 371-391. MR *34*, 1287.

[65] MONTGOMERY, H.L.

1. A note on the large sieve. *J. London Math. Soc. 43* (1968), 93-98. MR *37*, 184.
2. Mean and large values of Dirichlet polynomials. *Invent. Math 8* (1969), 334-345. MR *42*, 3029.
3. Zeros of *L*-functions. *Invent. Math. 8* (1969), 346-354. MR *40*, 2620.
4. Primes in arithmetic progressions. *Michigan Math. J. 17* (1970), 33-39. MR *41*, 1660.
5. *Topics in multiplicative number theory*. Lecture Notes in Math. *227* (1971), ix + 178 pp., Berlin-New York. MR *49*, 2616.

6. Hilbert's inequality and the large sieve. *Proc. of the Number Theory Conference* (Univ. Colorado, Boulder, Colo., 1972), 156-161. MR *50*, 4457.
7. The pair correlation of zeros of the zeta function. *Proc. Sympos. Pure Math. 24* (1973), 181-193. MR *49*, 2590.
8. Linnik's theorem. (unpublished).

[66] MONTGOMERY, H.L. and VAUGHAN, R.C.

1. Hilbert's inequality. *J. London Math. Soc.* (2) *8* (1974), 73-82. MR *49*, 2544.
2. The large sieve. *Mathematika 20* (1973), 119-134.
3. The exceptional set in Goldbach's problem. *Acta Arith. 27* (1975), 353-370.

[67] MORENO, C.J.

1. The average size of gaps between primes. *Mathematika 21* (1974), 96-100.

[68] MOTOHASHI, Y. **180**

1. An asymptotic formula in the theory of numbers. *Acta Arith. 16* (1969/70), 255-264. MR *42*, 1786.
2. On the distribution of prime numbers which are of the form $x^2 + y^2 + 1$. *Acta Arith. 16* (1969/70), 351-363. MR *44*, 5284.
3. A note on the least prime in an arithmetic progression with a prime difference. *Acta Arith. 17* (1970), 283-285. MR *42*, 3030.
4. On the distribution of prime numbers which are of the form "$x^2 + y^2 + 1$". II. *Acta Math. Acad. Sci. Hungar. 22* (1971), 207-210. MR *44*, 5285.
5. On the mean value theorem for the remainder term in the prime number theorem for short arithmetic progressions. *Proc. Japan Acad. 47* (1971), 653-657. MR *46*, 3464.

6. On the distribution of the divisor function in arithmetic progressions. *Acta Arith. 22* (1973), 175-199. MR *49*, 4952.
7. On the representations of an integer as a sum of two squares and a product of four factors. *J. Math. Soc. Japan 25* (1973), 475-505. MR *48*, 3899.
8. On some improvements of the Brun-Titchmarsh theorem. *J. Math. Soc. Japan 26* (1974), 306-323. MR *49*, 2617.
9. On some improvements of the Brun-Titchmarsh theorem. II. (Japanese) *Sūrikaiseki Kenkyūjo Kōkyūroku 193* (1973), 97-109.
10. A study of the paper of M.B. Barban and P.P. Vehov entitled "On an extremal problem". (Japanese) *Sūrikaiseki Kenkyūjo Kōlkyūroku 222* (1974), 9-50.
11. On some improvements of the Brun-Titchmarsh theorem. III. *J. Math. Soc. Japan 27* (1975), 444-453.
12. On almost primes in arithmetic progressions. *J. Math. Soc. Japan*. (to appear).
13. On almost primes in arithmetic progressions. II. *Proc. Japan Acad. 51* (1975), 545-547.
14. On the zero-density theorem of Linnik. (to appear).
15. *On almost primes in arithmetic progressions*. Zahlentheorie (Tagung, Math. Forschungzinst., Oberwolfach, 1975). (to appear).

[69] ORR, R.C.

1. *Remainder estimates for squarefree integers in arithmetic progressions*. Dissertation, Syracuse University, 1969.

181 2. Remainder estimates for squarefree integers in arithmetic progression. *J. Number Theory 3* (1971), 474-497. MR *44*, 3973.

[70] PAN, C.

1. On the representations of an even integer as the sum of a prime and an almost prime. (Chinese) *Acta Math. Sinica 12* (1962), 95-106= *Chinese Math. -Acta 3* (1963), 101-112. MR *29*, 4727. = *Sci. Sinica 11* (1962), 873-888. (Russian) MR *27*, 1427.

2. On the representation of even numbers as the sum of the prime and a product of not more than 4 primes. (Chinese) *Acta Sci. Natur. Univ. Shangtung 1962*, no. 2, 40-62 = *Sci. Sinica 12* (1963), 455-473. (Russian) MR *28*, 73.

3. A note on the large sieve method and its applications. (Chinese) *Acta Math. Sinica 13* (1963), 262-268 = *Chinese Math. - Acta. 4* (1963), 283-290. MR *29*, 82.

4. A new application of the Yu. V. Linnik large sieve method. (Chinese) *Acta Math. Sinica 14* (1964), 597-608 = *Chinese Math. - Acta 5* (1964), 642-652. MR *30*, 3871 = *Sci. Sinica 13* (1964), 1045-1053. (Russian) MR *30*, 3872.

[71] PORTER, J.W.

1. *Some divisor problems in the theory of numbers*. Ph.D. thesis, Nottingham 1973. (to appear in Acta Arith.).

2. The generalized Titchmarsh-Linnik divisor problem. *Proc. London Math. Soc*. (3) *24* (1972), 15-26. MR *45*, 5098.

3. The mean value of multiplicative functions on the set of numbers of the form $p + a$. *J. London Math. Soc*. (2) *10* (1975), 447-456.

[72] RAMACHANDRA, K.

1. A note on numbers with a large prime factor. *J. London Math. Soc*. (2) *1* (1969), 303-306. MR *40*, 118.

2. A note on numbers with a large prime factor. II. *J. Indian Math. Soc*. (N.S.) *34* (1970), 39-48 (1971). MR *45*, 8616.

3. On a discrete mean value theorem for $\zeta(s)$. *J. Indian Math. Soc*. (N.S.) *36* (1972), 307-316 (1973). MR *50*, 2095.

4. On the number of Goldbach numbers in small intervals. *J. Indian Math. Soc*. (N.S) *37* (1973), 157-170. MR *50*, 12941.

5. Largest prime factor of the product of *k*-consecutive integers. (Russian summary) *Trudy Mat. Inst. Steklov*. *132* (1973), 77-81, 264. MR *49*, 2600.

182 6. Application of a theorem of Montgomery and Vaughan to the zeta-function. *J. London Math. Soc*. (2) *10* (1975), 482-486.

7. A simple proof of the mean fourth power estimate for $\zeta(\frac{1}{2}+it)$ and $L(\frac{1}{2}+it\chi)$. *Ann. Scuola Norm. Sup. Pisa* (4) *1* (1974), 81-97.

8. Some new density estimates for the zeros of the Riemann zeta-function. *Ann. Acad. Sci. Fenn.* Ser. AI *1* (1975), 177-182.

9. On the frequency of Titchmarsh's phenomenon for $\zeta(s)$. II. *Acta Math. Acad. Sci. Hungar*. (to appear).

10. Some problems of analytic number theory. *Acta Arith*. (to appear).

[73] RANKIN, R.A.

1. The difference between consecutive prime numbers. V. *Proc. Edinburgh Math. Soc.* (2) *13* (1962/63). 331-332. MR *28*, 3978.

[74] RÉNTI,A.

1. On the representation of an even number as the sum of a single prime and single almost-prime number. (Russian) *Dokl. Akak. Nauk* SSSR (N.S) *56* (1947), 455-458. MR *9*, p. 136.

2. On the representation of an even number as the sum of a single almost-prime number. (Russian) *Izv. Akad. Nauk* SSSR. Ser. Mat. *12* (1948), 57-78. MR *9*, p. 413 = *Amer. Math. Soc. Transl*. (2) *19* (1962), 299-231. MR *24*, A 1264.

3. Generalization of the "large sieve" of Yu. V. Linnik. *Math. Centrum Amsterdam 1948,* 5 pp.

4. Un noveau théorème concernant les functions indépendantes et ses applications à la théorie des numbers. *J. Math. Pures Appl*. (9) *28* (1949), 137-149, MR *11*, p. 161.

5. On a theorem of the theory of probability and its application in number theory. (Russian. Czech summary) *Časopis Pěst. Mat. Eys*. *74* (1949), 167-175 (1950). MR *12*, p. 590.

6. *Probability methods in number theory*. Publ. Math. Collectae Budapest *1* (1949), no. 21, 9 pp. MR *12*, p. 161.

7. On the large sieve of Yu. V. Linnik. *Compositio Math*. *8* (1951), 68-75. MR *11* (1950). p. 581.

8. Sur un théorèm général de probabilité. *Ann. Inst. Fourier 1* (1949), 43-52 (1950), MR *14*, p. 886.

9. Probability methods in number theory. (Chinese) *Advancement in Math. 4* (1958), 465-510. MR *20*, 4535.

10. On the probabilistic generalization of the large sieve of Linnik. (Hungarian and Russian summaries) *Magyar Tud. Akad. Mat. Kutató Int. Közl*. *3* (1958), 199-206. MR *22*, B 1937. **183**

11. Probabilistic methods in number theory. *Proc. Int. Congr. Math. 1958*, 529-539. MR *22*, 9478.

12. New version of the probabilistic generalization of the large sieve. (Russian summary) *Acta Math. Acad. Sci. Hungar*. *10* (1959), 217-226. MR *22*, 1938.

[75] RICHERT, H.-E.

1. Selberg's sieve with weights. *Mathematika 16* (1969), 1-22. MR *40*, 119.

2. Selberg's sieve with weights. *Proc. Sympos. Pure Math. 20* (1971), 287-310. MR *47*, 3286, 6632.

[76] RIEGER, G.J.

1. Zum Sieb von Linnik. *Arch. Math. 11* (1960), 14-22. MR *27*, 3595.

2. Das grosse Sieb von Linnik für algebraische Zahlen. *Arch. Math. 12* (1961), 184-187. MR *24*, A 3152.

3. Über die Folge der Zahlen der Gestalt $p_1 + p_2$. *Arch. Math. 15* (1964), 33-41. MR *28*, 3023.

[77] RODRIQUEZ, G.

1. Sul problema dei divisori di Titchmarsh. (English summary) *Boll. Un. Mat. Ital*. (3) *20* (1965), 358-366. MR *33*, 5574.

[78] ROSS, P.M.

1. On Chen's theorem that each large even number has the form $p_1 + p_2$ or $p_1 + p_2p_3$. *J. London Math. Soc*. (2) *10* (1975), 500-506.

[79] ROSSER, J.B. and SCHOENFELD, L.

1. Approximate formulas for some functions of prime numbers. *Illinois J. Math*. *6* (1962), 64-94. MR *25*, 1139.

[80] ROTH. K.F

1. Remark concerining integer sequences. *Acta Arith. 9* (1964). 257-260. MR *29*, 5806.

2. On the large sieves of Linnik and Rényi *Mathematika 12* (1965), 1-9. MR *33*, 5589.

184 3. Irregularities of sequences relative to arithmetic progressions *Math. Ann. 169* (1967), 1-25. MR *34*, 4235.

4. *The large sieve*. Inaugural Lecture, 23 january 1968, Imp. Coll. London *1968*, 9 pp. MR *38*, 5740.

5. Irregulararities of sequences relative to arithmatic progressions. III. *J. Number Theory 2* (1970), 125-142. MR *41*, 6180.

[81] SAMANDAROV, A.G.

1. The large sieve the algebraic fields. (Russain) *Mat. Zametki 2* (1967), 673-680. MR *36*, 6379.

[82] SCHAAL, W.

1. On the large sieve method in algebraic number fields. *J. Number Theory 2* (1970), 249-270. MR *42*. 7826.

[83] SCOURFIELD, E.J.

1. An asymoptomic formula for the property $(n, f(n)) = 1$ for a class of multiplicative funcitons. *Acta Arith. 29*, no. 4. (to appear).

[84] SELBERG, A,.

1. On an elementary method in the theory of primes. *Norske Vid. Selsk. Forh., Trondhjem 19* (1947), no. 18, 64-67. MR *9*, p. 271.
2. *Twin prime problem*. Manuscript 16 pp. + Appendix 7 pp. (unpublished)
3. On elementary method in prime number-theory and their limitations *11. Skand. Mat. Kongr., Trondheim 1949*, 13-22. MR *14*, p. 726.
4. The general sieve method and its place in prime number theory. *Proc. Internat. Congress Math.*, Cambridge, Mass *1* (1950), 286-292. MR *13*, p. 438.
5. Sieve methods. *Proc, Sympos, Pure Math. 20* (1971), 311-351. MR *47*, 3286.
6. Remarks on sieves. *Proc. of the Number Theory Conference* (Univ. Colorado, Boulder, Colo., 1972) 205-216. MR *50*. 4457.

[85] SIEBERT, H. and WOLKE, D.

1. Über einige Analoga zum Bombierischen Primazahlsatz. *Math. Z. 122* (1971), 327-341.

[86] SOKOLOVSKIJ, A.V. **185**

1. On the large sieve. (Russian) *Acta Arith. 25* (1974), 301-306.

[87] STEPANOV, B.V.

1. On the mean value of the k^{th} power of the number of classes for an imaginary quadratic field. (Russain) *Dokl. Akad. Nauk* SSSR *124* (1959). 984-986. MR *21*, 4948.

[88] TURÁN, P.

1. Über die Primzahlen der arithmetischen Progression. *Acta Sci. Math.* (Szaged) *8* (1937), 226-235,

[89] UCHIYAMA, S.

1. On the difference between consecutive prime numbers. *Acta Arith. 27* (1975), 153-157. MR *51*, 3085.

[90] UCHIYAMA, M. AND UCHIYAMA, S.

1. On the representation of large even integers as sums of a prime and an almost prime. *Proc. Japan Acad. 40* (1964), 150-154. MR *29*, 2234.

[91] VAUGHAN, R.C.

1. Some applications of Montogomery's sieve, *J. Number Theory 5* (1973), 64-69. MR *49*, 7222.
2. On Goldbach's problem. *Acta Arith. 22* (1972), 21-48. MR *48*, 6045.
3. A remark on the divisor function $d(n)$. *Glasgow Math. J. 14* (1973), 54-55. MR *47*, 169.
4. On the number of solutions of the equation $p = a + n_1 \dots n_k$ with $a < p \leq x$ *J. London Math. Soc*. (2) *6* (1972), 43-45. MR *46*, 8994 a.
5. On the number of solutions of the equation $p + n_1 \dots n_k = N$ *J. London Math. Soc.* (2) *6* (1973), 326-328. MR *46*, 8994 b.

6. Mean value theorems in prime number theory. *J. London Math. Soc.* (2) *10* (1975), 153-162.

[92] VINOGRADOV, A.I.

1. The density hypothesis for Dirichelet *L*-series. (Russain) *lzv. Akad. Nauk* SSSR Ser. Mat. *29* (1965), 903-934. MR *33*, 5579; *Corrigendum: ibid 30* (1966), 719-720. MR *33*, 2607 = *Amer. Math. Soc. Transl.* (2) *82* (1969) 9-46.

[93] WANG, Y. **186**

1. A note on some properties of the arithmatical functions $\phi(n)$, $\sigma(n)$ and $d(n)$. (Chinese. English summary) *Acta Math. Sinica 8* (1958), $1-11$ = *Chinese Math. - Acta 8* (1966), 585-598 = *Amer. Math. Soc. Transl.* (2) *37* (1964), 143-156. MR *20*, 4533.
2. On the representation of large integer as a sum of a prime and an almost prime. *Sci. Sinica 11* (1962), 1033-1054. MR *27*, 1424.

[94] WANG, Y., HSIEH , S. and Yu. K.

1. Remarks on the difference of consecutive primes. *Sci. Sinica 14* (1965), 786-788. MR *32*, 5617.

[95] WARD, D.R.

1. Some series involving Euler's function *J. London Math. Soc. 2* (1927), 210-214.

[96] WARLIMONT, R.

1. On squarefree members in arithmatic progressions. *Monatsh. Math. 73* (1969), 433-448. MR *41*, 3430.
2. Über die kleinsten quadratfreien Zahlen in arithmetischen Progression mit primen Differenzen. *J. Reine Angew. Math. 253* (1972), 19-23. MR *46*, 1728.

[97] WILSON, R.J.

1. The large sieve in algebraic number fields. *Mathematika 16* (1969), 189-204. MR *41*, 8374.

[98] WOLKE, D.

1. Farrey-Br̄uche mit primem Nenner und das grosse Sieb. *Math. Z. 114* (1970), 145-158. MR *41*, 5228.
2. On the large sieve with primes. *Acta Math. Ascad. Sci. Hungar. 22* (1971), 239-247. MR *45*, 215.
3. Multiplikative Funktionen auf schnell wachsenden Folgen. *J. Reine Angew. Math. 251* (1971), 54-67. MR *44*, 6629.
4. Polynominal values with small prime divisors. *Acta Arith. 19* (1971), 327-333. MR *45*, 1871.
5. Über das summatorische Verhalten zalentheoretischer Funktionen. *Math. Ann. 194* (1971), 147-166. MR *46*, 3465.

187

6. Über die mittlere verteilung der Werte zahlentheoretischer Funktionen auf Restklassen. I.*Math. Ann. 202* (1973), 1-25. MR *48*, 6030.
7. Über die mittlere Verteilung der Werte zahlentheoretischer Funktionen auf Restklassem. II. *Math. Ann. 204* (1973), 145-153. MR *50*, 12940.
8. Eine Bemerkung zum Sieb des Erastothnes. *Monatsh. Math. 78* (1974), 264-266.
9. Das Selbegasche Sieb für zahletheoretische Funktionen, I. *Arch. Math. 24* (1973), 632-639. MR *49*, 2618.
10. A lower bound for the large sieve inequality. *Bull. London Math. Soc. 6* (1974), 315-318. MR *50*, 7065.
11. Grosse Differenzen zwischen aufeinanderfolgenden Primizahlen, *Math. Ann. 218* (1975), 269-271.