Lectures on
Sieve Methods and Prime Number Theory

By

Y. Motohashi

Tata Institute of Fundamental Research
Bombay
1983

Lectures on
Sieve Methods and Prime Number Theory

By

Y. Motohashi

Published for the

Tata Institute of Fundamental Research

Springer-Verlag

Berlin Heidelberg New York

1983

Author

Y. Motohashi
College of Science and Technology
Nihon University
Surugadai, Tokyo 101
Japan

ISBN 3-540-12281-8 Springer-Verlag, Berlin.
Heidelberg. New York
ISBN 0-387-12281-8 Springer-Verlag, New York.
Heidelberg. Berlin

Printed by M. N. Joshi at The Book Centre Limited,
Sion East, Bombay 400 022 and published by H. Goetze.
Springer-Verlag, Heidelberg, West Germany

Printed in India

To

My wife Kazuko and my daughter Haruko

Preface

In the last years we have witnessed penetrations of sieve methods into the core of analytic number theory - the theory of the distribution of prime numbers. The aim of these lectures which I delivered at the Tata Institute of Fundamental Research during a two-month course early 1981 was to introduce my hearers to the most fascinating aspects of the fruitful unifications of sieve methods and analytical means which made possible such deep developments in prime number theory.

I am much indebted to Professor K. Ramachandra and Dr. S. Srinivasan for their generous hospitality. I can still remember quite vividly many interesting discussions we made on the Institute beach aglow with the magnificent setting sun.

The whole manuscript was read by Dr. Srinivasan with utmost care, and I wish to thank him sincerely for his help.

Chiba, JAPAN — Yoichi Motohashi
October, 1983

Acknowledgement

In preparing these lectures I could freely refer to important unpublished works which my friend Dr. Iwaniec kindly put at my disposal. I thank him for his great generosity. I am deeply grateful to him and to Professors Halberstam, Jutila, Richert and Wolke for their encouragement which sustained me through the last decase.

Y. Motohashi

NOTATION

Most of the notations and conventions employed in these lectures are standard, but the following briefing may help the readers. The letter p with or without suffix stands for a prime number. For an integer d, $\omega(d)$ and $\tau_k(d)$ denote the number of different prime factors of d and the number of ways of expressing d as a product of k factors, respectively; $u|d^\infty$ implies that u divides a power of d. φ and μ are the Eiler and the möbius functions, respectively. For two integers d_1 and $d_2, (d_1, d_2)$ and $[d_1, d_2]$ are the greatest common divisor and the least comomon multiple of them, respectively. We use usual notation from that set theory; in particular, if A is a finite set, $|A|$ is its cardinality.

Most of Dirichlet characters are denoted typically by χ, and $\sum^*$ is as usual a sum over primitive characters.

If the letter s stands for a complex variable which will be clear from the context, we use the convention: $Re(s) = \sigma$ and $Im(s) = t$. The letters ϵ and c denote a suficiently small positive constant and a certain positive constant, respectively, whose value may differ at each occurrence.

The constants implied b the $0-$, $o-$ and $\ll$ symols are always absolute apart from their possible dependednce on ϵ which is also effective, i.e. once the value of ϵ is fixed the value of those constants is explicity computable.

Contents

Part I

Topics in Sieve Methods

Chapter 1

The Λ^2-Sieve

ONE OF THE primary purposes in these lectures is to appreciate the 1
power as well as the sharpness of Selberg's fundamental sieve idea —- the Λ^2 - sieve – by employing it as a principal tool in the investigation of the zeta- and L-functions. Our actual applications of his idea will, however, be made not in its original form but rather in its hybridized version with the large sieve of Linnik; indeed, there is a sort of duality relation between these two fundamental sieve methods because of which they admit of a fruitful unification.

In the present chapter, we shall first study this aspect of the Λ^2-sieve to some extent of generality and then, by specializing main results, prepare basic aids for the applications to be made in PART II.

1.1 Selberg's Sieve for Intervals

To begin with, we shall give a formulation of Selberg's fundamental idea:

Let Ω be a map of $\{p^\alpha\}$ the set of all prime- powers into the family of all subsets of $\mathbb{Z}$, and for an arbitray sequence A of integers let us consider

$$A_\Omega = \{a\epsilon A; a \not\epsilon \Omega(P^\alpha) \text{ for all } p^\alpha (\alpha > 0)\}$$

which may be called the resulatant of sifting A by Ω. We extend the 2 domain of Ω to $\mathbb{N}$ by putting

$$\Omega(d) = \bigcap_{P^\alpha \| d} \Omega(P^\alpha), \Omega(1) = \mathbb{Z},$$

and denote by δ_d and $\tilde{\omega}$ the charcteristic functions of the sets $\Omega(d)$ and $\mathbb{Z}_{\Omega'}$ respectively. Then Selberg's idea may be formulated as follows.

Theorem 1. *Let λ be an arbitrary real-valued function with a compact support and satisfying $\lambda(1) = 1$. Then*

$$\tilde{\omega} \leq \left(\sum_d \lambda(d)\delta_d \right)^2 .$$

PROOF is immediate.

In partticular, we have, for any finite sequence A of integers,

$$| A_\Omega | \leq \sum_{a \epsilon A} \left(\sum_d \lambda(d)\delta_d(a) \right)^2 . \tag{1.1.1}$$

Naturally, it is desirable to have the minimum value taken by this quadratic form of λ under the side condition $\lambda(1) = 1$; but obviously that would be intractable without imposing certain reasonable conditions on Ω, A and λ. Hence we shall introduce the following specialization of them in order to illustrate the process leading to the determination of a quasi-optimal λ, and thus a satisfactory upper bound for $|A_\Omega|$.

We assume that A is in an interval of length $N \epsilon \mathbb{N}$, i.e., there is an $M \epsilon \mathbb{Z}$ such that $A \subseteq [M, M+N)$ and that Ω is defined locally by congruence relations, i.e.,

$$\Omega(p^{\alpha)} = \left\{ \begin{array}{l} \quad n \quad (\bmod\ p^\alpha) \text{ belongs to a given } \\ n; \\ \quad \text{set if residues} \quad (\bmod\ p^\alpha) \end{array} \right\} \tag{1.1.2}$$

3

Also, we assume, for the sake of simplicity, that for each prime p there is an $\alpha_p \geq 1$ such that

$$\Omega(p^\alpha) = \text{ empty for all } \alpha \geq \alpha_p. \tag{1.1.3}$$

Further, we restrict λ by requiring that its support be contained in the interval $[0, Q]$, $Q > 1$ being a parameter.

On these assumptions, we have, by (1.1.1),

$$|A_\Omega| \leq \sum_{M \leq n < M+N} \left(\sum_{d < Q} \lambda(d)\delta_d(n) \right)^2 . \tag{1.1.4}$$

However, to avoid the complexity arising from the possible inter-relation among $\Omega(p^\alpha), \alpha \geq 1$, it is expendient to transform Ω into $\tilde{\Omega}$ which is defined by

$$\tilde{\Omega}(p^\alpha) = \Omega(p^\alpha) - \bigcup_{j=1}^{\alpha-1} \Omega(p^j)$$

so that $\tilde{\Omega}(p^\alpha)$, $\alpha \geq 1$, are independent of each other, i.e.,

$$\tilde{\Omega}(p^\alpha) \cap \tilde{\Omega}(p^\beta) = \text{ empty if } \alpha\beta(\alpha - \beta) \neq 0. \tag{1.1.5}$$

This does not cause any change in our present sieve situation, for we have evidently $A_\Omega = A_{\tilde{\Omega}}$. Thus we shall consider, instead of (1.1.4), the expression

$$|A_\Omega| \leq \sum_{M \leq n < M+N} \left(\sum_{d < Q} \lambda(d)\tilde{\delta}_d(n) \right)^2 , \tag{1.1.6}$$

where $\tilde{\delta}_d$ is the characteristic function of the set **4**

$$\tilde{\Omega}(d) = \bigcap_{p^\alpha \| d} \tilde{\Omega}(p^\alpha),$$

i.e.,

$$\tilde{\delta}_d = \delta_d \prod_{p^\alpha \| d} \prod_{j=1}^{\alpha-1} (1 - \delta_p^j).$$

Now we have to estimate the right side of (1.1.6). The conventional way of doing this is to expand out the lamda-squares, change the order of summations, and single out the main-term while estimating drastically the error thus caused; by an obvious reason, this does not work well in our present situation. Thus we need to devise an alternative argument. To this end, we observe that since we have defined Ω by the congruence condition (1.1.2), the characteristic function $\tilde{\delta}_d$ can be expressed as

$$\begin{aligned}\tilde{\delta}_d(n) &= \frac{1}{d}\sum_{\ell=1}^{d}\sum_{k=1}^{d}\exp\left(2\pi i\frac{k}{d}(n-\ell)\right)\tilde{\delta}_d\ell\\ &= \frac{1}{d}\sum_{q|d}\sum_{\substack{h=1\\(h,q)=1}}^{q}\exp\left(2\pi i\frac{h}{q}n\right)\sum_{\ell=1}^{d}\exp\left(-2\pi i\frac{h}{q}\ell\right)\tilde{\delta}_d(\ell).\end{aligned}\tag{1.1.7}$$

Insertion of this into the right side of (1.1.6) gives

$$|A_\Omega| \le \sum_{M\le n<M+N}|\sum_{q<Q}\sum_{\substack{h=1\\(h,q)=1}}^{q} b\left(\frac{h}{q}\right)\exp\left(2\pi i\frac{h}{q}n\right)|^2,\tag{1.1.8}$$

where

$$b\left(\frac{h}{q}\right) = \sum_{\substack{d\equiv 0 \pmod q\\ d<Q}}\frac{\lambda(d)}{d}\sum_{\ell=1}^{d}\exp\left(-2\pi i\frac{h}{q}\ell\right)\tilde{\delta}_d(\ell).$$

5 Thus we have got an expression fairly familiar in the theory of the large sieve, and we recall the fundamental

Lemma 1. *Let $\{x_r\}$ be a set of points in the unit interval which are spaced by $\delta > 0$. Then, for any $M\epsilon\mathbb{Z}, N\epsilon\mathbb{N}$ and complex numbers $\{b_r\}$, we have*

$$\sum_{M\le n<M+N}|\sum_r b_r\exp(2\pi i n x_r)|^2 \le (N-1+\delta^{-1})\sum_r |b_r|^2.$$

Applied to the right side of (1.1.8), this yields readily

$$\begin{aligned}|A_\Omega| &\le (N-1+Q^2)\sum_{q<Q}\sum_{\substack{h=1\\(h,q)=1}}^{q}|b\left(\frac{h}{q}\right)|^2\\ &= (N-1+Q^2)\sum_{d_1,d_2<Q}\lambda(d_1)\lambda(d_2)f(d_1,d_2),\end{aligned}\tag{1.1.9}$$

say; here we have

$$f(d_1,d_2) = \frac{1}{[d_1,d_2]} \sum_{\substack{\ell_1=1\\ \ell_1\equiv\ell_2}}^{d_1} \sum_{\substack{\ell_2=1\\ (\bmod\ (d_1,d_2))}}^{d_2} \tilde{\delta}_{d_1}(\ell_1)\tilde{\delta}_{d_2}(\ell_2).$$

Taking into account the multilicative property of $\tilde{\delta}_d$, this may be written as

$$f(d_1,d_2) = \frac{1}{[d_1,d_2]} \prod_{\substack{p^\alpha\|d_1\\ p^\beta\|d_2}} \left\{ \sum_{\substack{\ell_1=1\\ \ell_1\equiv\ell_2}}^{p^\alpha} \sum_{\substack{\ell_2=1\\ (\bmod\ p^{\min(\alpha,\beta)})}}^{p^\beta} \tilde{\delta}_{p^\alpha}(\ell_1)\tilde{\delta}_{p^\beta}(\ell_2) \right\}.$$

If $\alpha\beta(\alpha-\beta)\neq 0$ then this double sum is zero, for we have (1.1.5); on the other hand, if, either $\alpha=\beta$ or $\beta=0$, then it is equal to the number of 6
residue classes $\pmod{p^\alpha}$ defining the set $\tilde{\Omega}(p^\alpha)$, which we shall denote by $\|\tilde{\Omega}(p^\alpha)\|$, in what follows. Hence we have

(1.1.10)
$$f(d_1,d_2) = \prod_{\substack{p^\alpha\|d_1\\ p^\beta\|d_2}} f(p^\alpha,p^\beta),$$

where

(1.1.11)
$$f(p^\alpha,p^\beta) = f(p^\beta,p^\alpha) = \begin{cases} 0 & \text{if } \alpha\beta(\alpha-\beta)\neq 0,\\ \|\tilde{\Omega}(p^\alpha)\|p^{-\alpha} & \text{if either } \alpha=\beta \text{ or } \beta=0. \end{cases}$$

Now, let us proceed to the computation of the minimum value taken by the quadratic form

$$I = \sum_{d_1,d_2<Q} \lambda(d_1)\lambda(d_2) f(d_1,d_2)$$

on the side condition $\lambda(1)=1$. To this end, we need to have *a* diagonalization of the infinite matrix

$$F = (f(d_1,d_2))(d_1,d_2\epsilon\mathbb{N}).$$

Formula (1.1.10) implies that F can be expressed as the infinite Kronecker product:

$$F = \bigotimes_{p} F_{p'} \tag{1.1.12}$$

where

$$F_p = (f(p^{\alpha}, p^{\beta}))(\alpha, \beta \leq \alpha_p \quad \text{(cf. (1.1.13)))}.$$

This should be taken for a symbolic interpretation of the multiplicative property of the function f; so we may neglect the question of the order of multiplication.

7 Thus, it suffices to consider a diagonalization of F_P.For this sake we use the familiar algorithm of Gauss, and get

$$F_P = T_P D_P T_P^t, \tag{1.1.13}$$

where D_P is diagonal, and T_P is lower triangular with all diagonal entries being equal to 1. To see the precise form of D_P and T_P, we consider the quadratic form $K(x_0, x_1, \ldots, x_r)r = \alpha_P$, for which F_P is the coefficient matrix; (1.1.11) gives

$$k = x_{0^2} + 2x_0(f_1x_1 + f_2x_2 + \cdots + f_rx_r) + f_1x^{2_1} + f_2x^{2_2} + \cdots + f_rx^{2'_r}$$

where $f_j = \|\tilde{\Omega}(p^j)\| p^- j$.

We have

$$\begin{aligned} K &= (x_0 + f_1x_1 + \cdots + f_rx_r)^2 + f_1(1-f_1)\left(x_1 - \frac{1}{1-f_1}(f_2x_2 + \cdots + f_rx_r)\right)^2 \\ &\quad - \frac{1}{1-f_1}(f_2x_2 + \ldots + f_rx_r)^2 + f_2x_2^2 + \cdots + f_rx_r^2 \\ &= (x_0 + f_1x_1 + \cdots + f_rx_r)^2 + f_1(1-f_1)\left(x_1 - \frac{1}{1-f_1}(f_2x_2 + \cdots + f_rx_r)\right)^2 \\ &\quad + f_2\frac{1-f_1-f_2}{1-f_1}\left(x_2 - \frac{1}{1-f_1-f_2}(f_3x_3 + \ldots + f_rX_r)\right)^2 \\ &\quad - \frac{1}{1-f_1-f_2}(f_3x_3 + \ldots + f_rx_r)^2 + f_3x_3^2 + \ldots f_rx_r^2; \end{aligned}$$

thus inductiveIy we find

(1.1.14)
$$K = y_0^2 + f_1(1-f_1)y_1^2 + f_2\frac{1-f_1-f_2}{1-f_1}y_2^2 + \cdots + f_r\frac{1-f_1-f_2-\cdots-f_r}{1-f_1-f_2-\cdots-f_{r-1}}y_r^2$$

where

$$y_0 = x_0 + f_1x_1 + \cdots + f_rx_r \tag{1.1.15}$$

and,for $1 \leq j \leq r$, **8**

$$y_j = x_j - \frac{1}{1-f_1-f_2-\cdots-f_j}(f_{j+1}x_{j+1} + \cdots + f_rx_r). \tag{1.1.16}$$

It should be remarked here that in the above transformation of K we have assumed that for $1 \leq j \leq$

$$\theta(p^j) = 1 - f_1 - f_2 - \cdots - f_j \tag{1.1.17}$$

does not vanish. This causes no loss of generality. For, $p^j(1-\theta(p^j))$ is obviously the number of residue of classes $\pmod{p^j}$ defining the set $\Omega(p) \cup \Omega(P^2) \cup \cdots \cup \Omega(p^j)$, and if $\theta(p^j) = 0$, then this sum coincides with $\mathbb{Z}$, that is $|A_\Omega| = 0$.

Using the nitation (1.1.17), we may put the transformation (1.1.15) – (1.1.16) in the matrix from (1.1.13) with

$$D_p = \begin{pmatrix} g(1)_{g(p)} & & \bigcirc \\ & \ddots & \\ \bigcirc & & g(p^{\alpha_p}), \end{pmatrix}$$

$$T_p = (t(p^\alpha, mp^\beta))(0 \leq \alpha, \beta \leq \alpha_p),$$

where

$$g(1) = 1, g(p^\alpha) = (\theta(p^{\alpha-1}) - \theta(p^\alpha))\theta(p^\alpha)\theta(p^{\alpha-1})^{-1} \tag{1.1.18}$$

and

$$t(p^\alpha, p^\beta) = \begin{cases} 1 & \text{if} \alpha = \beta, \\ \theta(p^{\alpha-1}) - \theta(p^\alpha) & \text{if} \alpha > 0, \beta = 0, \\ (\theta(p^\alpha) - \theta(p^{\alpha-1}))\theta(p^\alpha)^{-1} & \text{if} \alpha > \beta > 0, \\ 0 & \text{if} \alpha < \beta. \end{cases} \tag{1.1.19}$$

9

In particular, we have

$$f(p^\alpha, p^\beta) = \sum_{r=0}^{\min(\alpha,\beta)} g(p^r)t(p^\alpha, p^r)t(p^\beta, p^r).$$

Thus, in view of (1.1.10), we obtain

$$f(d_1, d_2) = \sum_{u|(d_1,d_2)} g(u)t(d_1, u)t(d_2, u), \tag{1.1.20}$$

in which we have put

$$g(u) = \prod_{p^\alpha \| u} g(p^\alpha) \tag{1.1.21}$$

and

$$t(d, u) = \prod_{\substack{p^\alpha \| d \\ p^\beta \| u}} t(p^\alpha, p^\beta).$$

The formula (1.1.20) provides the quadratic form I with the diagonalized form:

$$\begin{aligned} I &= \sum_{u<Q} g(u)\Big(\sum_{\substack{d<Q \\ d\equiv 0 \pmod u}} t(d, u)\lambda(d)\Big)^2 \\ &= \sum_{u<Q} g(u)\xi_u^2, \end{aligned}$$

say.

10 To proceed further, we need to express the side condition $\lambda(1) = 1$ in terms of ξ_u. To this end, we compute the inverse matrix of T_p; this may be performed easily with the aid of (1.1.15) and (1.1.16). We have

$$T_p^{-1} = (t^*(p^\alpha, p^\beta))(0 \le \alpha, \beta \le \alpha_p)$$

with

$$t^*(p^\alpha, p^\beta) = \begin{cases} 1 & \text{if } \alpha = \beta, \\ (\theta(p^\alpha) - \theta(p^{\alpha-1}))\theta(p^{\alpha-1})^{-1} & \text{if } \alpha > 0, \beta = 0, \\ -(\theta(p^\alpha) - \theta(p^{\alpha-1}))\theta(p^{\alpha-1}) & \text{if } \alpha > \beta > 0, \\ 0 & \text{if } \alpha < \beta. \end{cases} \tag{1.1.22}$$

Then, putting

$$t^*(d,u) = \prod_{\substack{p^\alpha \| d \\ p^\beta \| u}} t^*(p^\alpha, p^\beta),$$

we have

$$\sum_{\substack{u \mid d_1. \\ u \equiv 0 \pmod{d_2}}} t(d_1,u)t^*(u,d_2) = \delta_{d_1,d_2} \text{(Kronecker's delta)} \tag{1.1.23}$$

as well as

$$\sum_{\substack{u \mid d_1 \\ u \equiv 0 \pmod{d_2}}} t^*(d_1,u)t(u,d_2) = \delta_{d_1,d_2}. \tag{1.1.24}$$

In view of the definition of ξ_u, (1.1.23) implies **11**

$$\lambda(d) = \sum_{\substack{u<Q \\ u \equiv 0 \pmod{d}}} t^*(u,d)\xi_u.$$

Specifically, we have transformed the side condition $\lambda(1) = 1$ into

$$1 = \sum_{u<Q} t^*(u,1)\xi_u.$$

Thus we have

$$I = \sum_{d<Q} g(d)\left(\xi_d - \frac{t^*(d,1)}{g(d)}D\right)^2 + D$$

where

$$\begin{aligned} D &= \left\{\sum_{d<Q} \frac{t^*(d,1)^2}{g(d)}\right\}^{-1} \\ &= \left\{\sum_{d<Q} \prod_{p^\alpha \| d} \left(\frac{1}{\theta(p^\alpha)} - \frac{1}{\theta(p^{\alpha-1})}\right)\right\}^{-1}. \end{aligned}$$

Hence we find that

$$\min_{\lambda(1)=1} I = D,$$

and this is attained at

$$\lambda(d) = D \sum_{\substack{u<Q \\ u\equiv 0 \pmod d}} \frac{t^*(u,1)t^*(u,d)}{g(u)}. \tag{1.1.25}$$

Summing up the above discussion, we have now established

Theorem 2 (SELBERG'S SIEVE FOR INTERVALS)**.** *Let A be a se-*
12 *quence of integers in an interval of length $N\epsilon\mathbb{N}$, and Ω be defined by the*
conqruence relation (1.1.2)*. Then have, for any $Q > 1$,*

$$|A_\Omega| \leq (N - 1 + Q^2)\left\{\sum_{d<Q}\prod_{p^\alpha\|d}\left(\frac{1}{\theta(p^\alpha)} - \frac{1}{\theta(p^{\alpha-1})}\right)\right\}^{-1}$$

where θ is defined by (1.1.17)*.*

Remark. By the exclusion-inclusion principle, we can show easily that

$$\theta(p^\alpha) = 1 + \sum_{r=1}^{\alpha}(-1)^r \sum_{1\leq j_1<j_2<\cdots<j_r\leq\alpha} \|\Omega(p^{j_1}) \cap \cdots \cap \Omega(p^{-j_r})\| p^{-j_r}$$

where $\|\Omega(p^{j_1}) \cap \cdots \cap \Omega(p^{j_r})\|$, $j_1 < j_2 < \cdots < j_r$, denote the number of residue classes $\pmod{p^{j_r}}$ defining the set $\Omega(p^{j_1}) \cap \cdots \cap \Omega(p^{j_r})$.

1.2 The Hybrid Dual Sieve for Intervals

Next, we shall show a hybridization of THEOREM 2 with the multiplicative large sieve inequality, and by doing so, we shall stress that the occurrence of the additive large sieve inequality (LEMMA 1) in our discussion on the Selberg sieve for intervals is by no means accidental; in fact, as already mentioned in the introduction to this chapter, behind this phenomenon is an important relation between Selberg's and Linnik's sieve methods which may be termed a duality.

In the present section, we shall retain the notation and conventions introduced in the above; in particular, Ω is defined by the congruence relation (1.1.2).

First, we make an observation on the nature of the optimal λ which **13**
has been obtained at (1.1.25). It gives

$$\sum_{u<Q} \lambda(u)\tilde{\delta}_u(n) = D \sum_{d<Q} \frac{t^*(d,1)}{g(d)} \Psi_d(n,\Omega), \tag{1.2.1}$$

where

$$\Psi_d(n,\Omega) = \sum_{u|d} t^*(d,u)\tilde{\delta}_u(n). \tag{1.2.2}$$

Recȧlling the definitions of $\tilde{\delta}$ and t^*, this may be written as

$$\psi_d(n,\Omega) = \prod_{p^\alpha\|d} \Theta(p^{\alpha-1})^{-1}\left\{\triangle_{p^{\alpha-1}}(n)\Theta(p^\alpha) - \triangle_{p^\alpha}(n)\Theta(p^{\alpha-1})\right\} \tag{1.2.3}$$

where

$$\triangle_p^\alpha = \prod_{j-1}^{\alpha}(1-\delta_p{}^j).$$

And we have actually proved in the preceding section the inequality:

$$\sum_{M\le n<M+N}\left\{\sum_{d<Q}\frac{t^*(d,1)}{g(d)}\Psi_d(n,\Omega)\right\}^2 \le (N-1+Q^2)\sum_{d<Q}\left(\frac{t^*(d,1)}{g(d)}\right)^2 g(d).$$

This relation raises the anticipation that the norm of the matrix

$$(\Psi_d(n,\Omega)g(d)^{-\frac{1}{2}})(d<Q, M\le n<M+N)$$

may not exceed $(N-1+Q^2)^{\frac{1}{2}}$, i.e., for any complex numbers $\{b_d\}$

$$\begin{aligned}\sum_{M\le n<M+N}\Big|\sum_{d<Q}\frac{\psi_d(n,\Omega)}{\sqrt{g(d)}}b_d\Big|^2 \\ \le (N-1+Q^2)\sum_{d<Q}|b_d|^2.\end{aligned} \tag{1.2.4}$$

In order to press the matter further, we quote the well-known duality **14**
principle:

Lemma 2. *Let (c_{ij}) be a matrix, and E be such that for any complex numbers $\{x_i\}$*

$$\sum_j |\sum_i c_{ij}x_i|^2 \leq E\sum_i |x_i|^2.$$

Then we have, for any complex numbers $\{y_j\}$,

$$\sum_i |\sum_j c_{ij}y_j|^2 \leq E\sum_j |y_j|^2,$$

and vice-versa.

Thus, if (1.2.4) is true, then its dual

$$\begin{aligned}\sum_{d<Q}\frac{1}{g(d)}|\sum_{M\leq n<M+N}\psi_d(n,\Omega)a_n|^2 \\ \leq (N-1+Q^2)\sum_{M\leq n<M+N}|a_n|^2\end{aligned} \tag{1.2.5}$$

with arbitrary complex numbers $\{a_n\}$ will also be true. We should note the strong similarity of this to the multiplicative large sieve inequality.

Now we shall develop a descussion to confirm that (1.2.5), and indeed a much more general result actually hold. To begin with, we introduce the expression

$$D = \sum_{\substack{qr<Q\\(q,r)=1\\(qr,k)=1}}\ \sum^{*}_{X \pmod q}\frac{q}{\varphi(q)g(r)}|\sum_{\substack{M\leq n<M+N\\ n\equiv\ell \pmod k}}\chi(n)\Psi_r(n,\Omega)a_n|^2,$$

15 where $k, \ell, Q, N\epsilon\mathbb{N}, M\epsilon\mathbb{Z}$ are arbitary. But the direct estimation of D is tedious if not difficult; so we consider, instead, the dual

$$D^* = \sum_{\substack{M\leq n<M+N\\ n\equiv\ell \pmod k}}|\sum_{\substack{qr<Q\\(q,r)+1\\(qr,k)=1}}\ \sum^{*}_{\chi \pmod q}\left(\frac{q}{\varphi(q)g(r)}\right)^{\frac{1}{2}}X(n)\Psi_r(n,\Omega)b(r,X)|^2,$$

where $\{b(r,\chi)\}$ are arbitrary complex numbers. Inserting the expression (1.2.2) for $\psi_r(n,\Omega)$, we have

$$D^* = \sum_{\substack{M\leq n<M+N\\ n\equiv\ell \pmod k}} |\sum_{\substack{qd<Q\\ (q,d)=1\\ (qd,k)=1}} \ \sum^{*}_{X \pmod q} \left(\frac{q}{\varphi(q)}\right)^{\frac{1}{2}} X(n)\tilde{\delta}_d(n)S(d,X)|^2,$$

where

$$(1.2.6) \qquad S(d,X) = \sum_{\substack{u<Q/q\\ u\equiv 0 \pmod d\\ (u,qk)=1}} b(u,X)t^*(u,d)g(u)^{-}\frac{1}{2}, (X \pmod q)).$$

Recalling (1.1.7), we transfrom D^* further into

$$(1.2.7) \quad D^* = \sum_{\substack{M\leq n<M+N\\ n\equiv\ell \pmod k}} |\sum^{**}_{u,h,q,X} \left(\frac{q}{\varphi(q)}\right)^{\frac{1}{2}} \chi(n)\exp\left(2\pi i\frac{h}{u}n\right) y(u,h,X)|^2$$

where

(1.2.8)

$$y(r,\chi) = \sum_{\substack{d<Q/q\\ d\equiv 0 \pmod u\\ (d,qk)=1}} \frac{s(d,X)}{d}\sum_{\ell}^{d} = 1\exp\left(-2\pi i\frac{h}{u}\ell\right)\tilde{\delta}_d(\ell),\chi \mod q.$$

and $\sum^{**}$ denotes the sum over u, h, q, χ satisfying the conditions: $uq <$ **16**
$Q, (u,q) = (uq,k) = 1;\ 1 \leq h \leq u, (h,u) = 1; \chi$ primitive $\pmod q$.

Then, regarding the right side of (1.2.7) as an Hermitian form of the variables $y(u,h,\chi)$, we its dual:

$$(1.2.9) \qquad \sum^{**}_{u,h,q,\chi} \frac{q}{\varphi(q)} |\sum_{\substack{M\leq n<M+N\\ n\equiv\ell \pmod k}} \chi(n)\exp\left(2\pi i\frac{h}{u}n\right)c_n|^2,$$

where $\{c_n\}$ are arbitrary complex numbers. as usual, we express $\chi(n)$ as a linear combination of additive characters via Gauss sum, and by the orthogonality of characters, we infer that (1.2.9) is not larger than

$$\sum_{\substack{up<Q\\(u,p)=1\\(uq,k)=1}} \sum_{\substack{h=1\\(u,h)=1}}^{u} \sum_{\substack{a=1\\(a,q)=1}}^{q} \Big| \sum_{\substack{M\leq n<M+N\\n\equiv\ell \pmod{k}}} \exp\left(2\pi i n\left(\frac{h}{u}+\frac{a}{q}\right)\right) c_n\Big|^2$$

$$\leq \sum_{f<Q} \sum_{\substack{s=1\\(s,f)=1}}^{f} \Big| \sum_{\frac{M-\ell}{k}\leq M<\frac{M+N-\ell}{k}} \exp\left(2\pi i \frac{s}{f} m\right) c'_m\Big|^2,$$

where $c'_m = c_{km+\ell}$. Thus, by the dual of LEMMA 1, we see that (1.2.9) is not larger than

$$\left(\frac{N}{K}+Q^2\right) \sum_{\substack{M\leq n<M+N\\n\equiv\ell \pmod{k}}} |c_n|^2,$$

whence we obtain, by LEMMA 2,

$$D^* \leq \left(\frac{N}{K}+Q^2\right) \sum_{u,h,q,\chi}^{**} |y(u,h,\chi)|^2.$$

17

Recalling (1.2.8), we can compute the last sum just as (1.1.9), getting

$$D^* \leq \left(\frac{N}{K}+Q^2\right) \sum_{\substack{d_1q<Q\\d_2q<Q\\(d_1d_2,q)=1\\(d_1d_2q,k)=1}} \sum_{\chi}^{*} \pmod{q} s(d_1,\chi)\overline{s(d_2,\chi)} f(d_1,d_2).$$

Thus, by virtue of (1.1.20), we have

$$D^* \leq \left(\frac{N}{K}+Q^2\right) \sum_{\substack{dq<Q\\(d,q)=1\\(dq,k)=1}} \sum_{\chi \pmod{q}}^{*} g(d) \Big| \sum_{\substack{u<Q/q\\u=0 \pmod{d}\\(u,qk)=1}} t(u,d) S(u,\chi)\Big|^2.$$

But by (1.1.24) and (1.2.6), the last sum over u is equal to $b(d,\chi)$ $g(d)^{-\frac{1}{2}}$, whence

$$D^* \leq \left(\frac{N}{K}+Q^2\right) \sum_{\substack{dq<Q\\(d,q)=1\\(dq,k)=1}} \sum_{\chi \pmod{q}}^{*} |b(d,\chi)|^2.$$

Therefore, appealing to LEMMA 2 once more, we obtain

Theorem 3 (The Hybrid Dual Sieve for Intervals)**.** *Let* Ω *be defined by the corgruence relation* (1.1.2)*,* $\psi_d(n,\Omega)$ *by* (1.2.2)*, and* $g(d)$ *by* (1.1.18) *and* (1.1.21)*. Then we have, for arbitrary* k, ℓ, Q, N, $\epsilon\mathbb{N}$, $M\epsilon\mathbb{Z}$ *and complex* $\{a_n\}$*,*

$$\sum_{\substack{qr<Q\\(q,r)=1\\(qr,k)=1}} \sum_{\chi \pmod q}^{*} \frac{q}{\varphi(q)g(r)} \Big| \sum_{\substack{M\le n<M+N\\ n\equiv \ell \pmod k}} \chi(n)\psi_r(n,\Omega)a_n\Big|^2 \le \left(\frac{N}{K}+Q^2\right) \sum_{\substack{M\le n<M+N\\ n\equiv \ell \pmod k}} |a_n|^2.$$

The sieve-effect of this remarkably uniform result is embodied in **18**

Corollary to Theorem 3. Let Ω be defined by (1.1.2), and $\Theta(p^\alpha)$ by (1.1.17). Let $\{a_n\}$ be an arbitrary sequence of complex numbers satisfying $a_n = 0$ whenever there is a $p^\alpha(\alpha > 0)$ such that $n\epsilon\Omega(p^\alpha)$. Then we have, for arbitrary $k, \ell, Q, N\epsilon\mathbb{N}$ and $M\epsilon\mathbb{Z}$,

$$\sum_{\substack{qr<Q\\(q,r)=1\\(qr,k)=1}} \sum_{\chi \pmod q}^{*} \frac{q}{\varphi(q)} \prod_{p^\alpha \| r} \left(\frac{1}{\Theta(p^\alpha)} - \frac{1}{\Theta(p^{\alpha-1})}\right) \Big| \sum_{\substack{M\le n<M+N\\ n\equiv \ell \pmod k}} \chi(n)a_n\Big|^2 \le \left(\frac{N}{K}+Q^2\right) \sum_{\substack{M\le n<M+N\\ n\equiv \ell \pmod k}} |a_n|^2.$$

To deduce this from THEOREM 3, we need only to note that ψ_r $(n,\Omega) = t^*(r,1)$ if $a_n \neq 0$.

Specializing THEOREM 3 and the Corollary to it, we can deduce various important inequalities known at present in the theory of the large sieve; for instance, THEOREM 2 is contained in the corollary. Also, a special attention should be paid for the case arising from the simplest choice of Ω : $\Omega(p^\alpha)$ is empty for $\alpha \ge 2$ and $n\epsilon\Omega(p)$ is equivalent to $p|n$. For this Ω, we have $g(r) = \mu^2(r)\varphi(r)r^{-2}$ and $\psi_r(n,\Omega) =$ **19**

$u(r)u((r,n))\varphi((r,n))r^{-1}$. thus THEOREM 3 gives, for arbitrary k, l, Q, $N\epsilon\mathbb{N}$, $M\epsilon\mathbb{Z}$ and complex numbers $\{a_n\}$,

$$\sum_{\substack{qr<Q\\(q,r)=1\\(qr,k)=1}}\ \sum_{\chi\pmod q}^{*}\frac{u^2(r)q}{\varphi(qr)}|\sum_{\substack{m\leq n<M+N\\n\equiv\ell\pmod k}}\chi(n)\Psi_r(n)a_n|^2 \leq \left(\frac{N}{K}+Q^2\right)\sum_{\substack{M\leq n<M+N\\n\equiv\ell\pmod k}}|a_n|^2, \tag{1.2.10}$$

where

$$\psi_r(n)=\mu((r,n))\varphi((r,n)). \tag{1.2.11}$$

Also, recalling the well-known estimate

$$\sum_{\substack{r<R\\(r,f)=1}}\frac{\mu^2(r)}{\varphi(r)}\geq\frac{\varphi(f)}{f}\log R \tag{1.2.12}$$

for arbitrary $f,R\epsilon\mathbb{N}$, we see readily that the Corollary to THEOREM 3, or rather (1.2.10), gives rise to the assertion that if $a_n=0$ whenever n has a prime factor less than Q, then we have, for arbitrary k, ℓ, Q, $N\epsilon\mathbb{N}$ and $M\epsilon\mathbb{Z}$,

$$\sum_{\substack{q<Q\\(q,k)=1}}\log\frac{Q}{q}\sum_{\chi\pmod q}^{*}|\sum_{\substack{M\leq n<M+N\\n\equiv\ell\pmod k}}\chi(n)a_n|^2 \leq\frac{1}{\varphi(k)}(N+kQ^2)\sum_{\substack{M\leq n<M+N\\n\equiv\ell\pmod k}}|a_n|^2. \tag{1.2.13}$$

20 Specifically, we get, for $M\geq Q$,

$$\sum_{\substack{q<Q\\(q,k)=1}}\log\frac{Q}{q}\sum_{\chi\pmod q}^{*}|\sum_{\substack{M\leq p<M+N\\p\equiv\ell\pmod k}}\chi(p)|^2 \leq\frac{1}{\varphi(k)}(N+KQ^2)(\pi(M+N;k,\ell)-\pi(M;k,\ell)), \tag{1.2.14}$$

which is a refinement of the Brun-Titchmarsh theorem:

$$\pi(M+N;k,\ell)-\pi(M;k,\ell)\leq\frac{(2+0(1))N}{\varphi(k)\log\frac{N}{K}} \tag{1.2.15}$$

as N/k tends to infinity.

1.3 An Auxiliary Result Relating to the Λ^2-Sieve

In the above, we have seen that the optimal lamda-weight (1.1.25) has an important arithmetical property which makes it possible to unite Selberg's and Linnik's sieve methods. In the present section, digressing somewhat from the main theme of this chapter, we shall take up a subject related to the asymptotic behaviour of the optimal λ which the simplest choice of Ω mentioned above; this will also have important applications in PART II.

Thus, let Ω be such that $\Omega(p^\alpha)$ is empty for all $\alpha\geq 2$, and $n\epsilon\Omega(p)$ is equivalent to P/N. Then we have the simplest case of the Selberg sieve: the number of integers $\leq N$ which are free of prime factors less than z is bounded by

$$\sum_{1\leq n\leq N}\left(\sum_{\substack{d|n\\ d<z}}\lambda(d)\right)^2\ (\lambda(1)=1); \tag{1.3.1}$$

here, for our convenience, we use $z>1$ instead of Q. (1.1.25) gives the 21
optimal weight

$$\lambda(d)=\mu(d)\frac{d}{\varphi(d)}\left\{\sum_{\substack{r<z/d\\ (r,d)=1}}\frac{\mu^2(r)}{\varphi(r)}\right\}\left\{\sum_{r<z}\frac{\mu^2(r)}{\varphi(r)}\right\}-1 \tag{1.3.2}$$

which gives

$$\sum_{1\leq n\leq N}\left(\sum_{\substack{d|n\\ d<z}}\lambda(d)\right)^2\leq(N-l+z^2)/\log z.. \tag{1.3.3}$$

On the other hand, if we fix d and let z tend to infinity then (1.3.2) becomes the asymptotic relation:

$$\lambda(d) = (l + O(l))\mu(d)\frac{\log z/d}{\log z}.$$

Now the new weight

$$\tilde{\lambda}(d) = \begin{cases} \mu(d)\frac{\log z/d}{\log z} if d < z, \\ 0 if d \geq z \end{cases}$$

has a striking property: we have, for any $N \geq z$,

$$\sum_{1 \leq n \leq N} \left(\sum_{d|n} \tilde{\lambda}(d) \right)^2 \ll \frac{N}{\log z.} \tag{1.3.4}$$

The significance of this result lies in that, apart from a constant mul-
22 tiplier to the main-term, the error-term corresponding to z^2 of (1.3.3) does not appaer at all. Because of this uniformity, (1.3.4) has some important applications especially to the theory of the zeta-and L-functions. In our later discussion on these functions, however, we shall not require (1.3.4) in its full force, but rather the following consequence of it:

$$\sum_{n=1}^{\infty} \left(\sum_{d|n} \tilde{\lambda}(d) \right)^2 n^{-\omega} = O(l), \tag{1.3.5}$$

provided $\omega \geq 1 + c(\log z)^{-1}$. And for some special problems on L-functions, it is more desrable to heve a similar result in which the factor $\tau_k(n)$ occur in a sum corresponding to the left side of (1.3.5). For this sake, it would be expedient to consider the Selberg sieve problem:

$$\sum_{m \leq N} \tau_k(n) \left(\sum_{\substack{d|n \\ d<z}} \lambda(d) \right)^2 \quad (\lambda(l) = l).$$

The standard argument shows that the quasi-optimal λ is such that, for each fixed d,

$$\lambda(d) = (l + O(l))\mu(d)\left(\frac{\log z/d}{\log z}\right)^k$$

as z tends to infinity.

Thus we are led to the problem of estimating

$$J = \sum_{n=1}^{\infty} \tau_k(n)(\sum_{d|n} \xi(d))^2 n^{-\eta}, \tag{1.3.6}$$

where $\eta = 1 + c(\log z)^{-1}$ and

$$\xi(d) = \begin{cases} \mu(d)(\log\ z/d)^k & \text{if } d < z, \\ 0 & \text{if } d \geq z. \end{cases}$$

23

Expanding out the squares and changing the order of summation, we see that

$$\begin{aligned} J &= \zeta(\eta)^k \sum_{d_1,d2<z} \xi(d_1)\xi(d_2) \prod_{p|d_1d_2} (1-(1-p^{-n})^k) \\ &= \zeta(\eta)^k E, \end{aligned} \tag{1.3.7}$$

say. To diagonalize E, we employ a well-known device of Selberg, getting

$$E = \sum_{d<z} \mu^2(d) \prod_{p|d} (1-p^{-n})^k (1-(1-p^{-n})^k |R_d(z/d)^2, \tag{1.3.8}$$

where

$$R_d(x) = \sum_{\substack{u<x \\ (u,d)=1}} \mu(u)(\log \frac{x}{u})^k \prod_{p|u} (1-(1-p^{-n})^k).$$

There is an elementary argument to estimate $R_d(x)$ which relies on the elementary prime number theorem with remainder term (cf. § 4.1).

But, for the sake of simplicity, we take here an alternative way, an analytic one. We note first that

$$R_d(x) = \frac{k!}{2\pi i}\int_{2-i\infty}^{2+i\infty} \zeta(s+\eta)^{-k} P_d(s)\frac{x^s}{s^{k+1}}ds,$$

where $p_d(s) = \prod_{p|d}(l - \frac{l}{p^{s+n}})^{-k} \prod_{p\nmid}\left(l - \frac{l}{p^{s+n}}\right)^{-k}\left(l - \frac{l}{p^s}\left(l - \left(l - \frac{l}{p^\eta}\right)^k\right)\right)$,

24 which converges absolutely for Re $(s) > -c$. Then we quote an elementary estimate of $\zeta^{-1}(s)$ (cf. §4.1): in the region

$$Re(s) > 1 - c(\log(|t|+2))^{-9}$$

we have

$$\zeta^{-1}(s) \ll (\log(|t|+2))^7.$$

Thus shifiting the line of integration to the left appropriately, we get

$$R_d(x) = k!\mathrm{Res}_{s=0}\left\{\zeta(s+\eta)^{-k}P_d(s)x^s s^{-k-1}\right\} + 0\left(\prod_{p|d}\left(1+\frac{1}{\sqrt{p}}\right)k\right).$$

After some elementary estimations of derivatives of $\zeta^{-k}(s+\eta)$ and $P_d(s)$ at $s = 0$, we obtian

$$R_d(x) << \prod_{p|d}\left(1+\frac{1}{\sqrt{p}}\right)^k \sum_{j=0}^{k} ((\eta-1)\log x)^j.$$

Inserting this into (1.3.8) we see, via (1.3.7), that

$$J \ll (\log z)^{2k}. \tag{1.3.9}$$

After these preparations, we can show

Theorem 4. *Let $z > 1$ and $\vartheta > 0$, and let us put*

$$\Lambda_d^{(k)} = \frac{1}{k!}(\vartheta\log z)^{-k}\sum_{j=0}^{k}(-1)^{k-j}\tbinom{k}{j}\lambda_d^{(j,k)}, \tag{1.3.10}$$

where

$$\lambda_d^{(j,k)} = \begin{cases} \mu(d)\left(\log \frac{z^{1+j\vartheta}}{d}\right)^k & \textit{if } d < z^{1+j\vartheta}, \\ 0 & \textit{otherwise} . \end{cases}$$ 25

Then we have

$$\Lambda_d^{(k)} = \mu(d) \textit{ for } d < z. \tag{1.3.11}$$

Also

$$\sum_{n=1}^{\infty} \tau_k(n) \left(\sum_{d|n} \Lambda_d^{(k)} \right)^2 n^{-\omega} = O(l), \tag{1.3.12}$$

provided

$$\omega \geq 1 + c(\log z)^{-1}.$$

In fact, the second statement follows immediately from (1.3.9). As for the forst we note that for $d < z$

$$\sum_{j=0}^{k} (-1)^{k-j} \binom{k}{j} \lambda_d^{(j,k)}$$
$$= \mu(d) \sum_{\ell=0}^{k} (-1)^{k-\ell} \binom{k}{\ell} (\log z)^{\ell} (\log d)^{k-\ell} \sum_{j=0}^{k} (-1)^{k-j} \binom{k}{j} (1 + j\vartheta)^{\ell}.$$

But the last sum over j is equal to $\vartheta^k k!$ if $\ell = k$, and to 0 if $\ell < k$, whence we have (1.3.11).

1.4 The Hybrid Dual Sieve for Multiplicative Functions

In the first two sections, we were concerned with problems of sifting integers in an interval to each of which the simplest weight i.e. 1 is 26
attached, and we had a very powerful tool: the additive large sieve inequality. We are now going to investigate a similar problem on the assumption that weights not necessarily equal to 1 are given to the elements to be sifted. Then we have no longer such useful an aid as

LEMMA 1, but can appeal only to the conventional way of manipulating the Λ^2-sieve.

Let us denote by f the weight function, and consider

$$K = \sum_{n<N} f(n) \left(\sum_{\substack{d|n \\ d<R}} \lambda(d) \right)^2 .$$

Here $R > 1$ is a parameter. We look for the optimal λ which makes K as small as possible on the side condition $\lambda(1) = 1$. We may discuss this problem on some fairly general assumption on the average property of the sequence $\{f(n)\}$. But, since we have particular applications in mind which will be made in PART II, we shall confine ourselves to those f which satisfy the following practical conditions:

(C_1) f is a non-negative multiplicative function such that

$$f(n) = O(n^{\epsilon})$$

for all $n\epsilon\mathbb{N}$.

(C_2) There exist $A > 0$ and $\alpha \geq 1$ such that for all prime p we have

$$F_p - 1 \geq Ap^{-\alpha},$$

where
$$F_p = \sum_{m=0}^{\infty} f(p^m)p^{-m}.$$

27

(C_3) There exist $\beta \geq 0$, $0, 0 < \gamma < 1, \mathcal{F} > 0$, $D \geq 1$ such that

$$\sum_{n<y} \chi(n) f(n) = E(\chi)\mathcal{F} K(q)y + O(Dq^{\beta}y^{\gamma}),$$

where $\chi \pmod q$, $K(q) = \prod_{p|q} F_p^{-1}$; the constant implied by the O-symbol is absolute.

Now let us estimate K on these assumptions. As usual, we may restrict λ by

$$|\lambda(d)| \leq |\mu(d)|; \tag{1.4.1}$$

in fact, this will be confirmed later for the optimal λ. Expanding out the lamda-squares and changing the order of summation, we have

$$K = \sum_{d_1, d_2 < R} \lambda(d_1)\lambda(d_2) \sum_{n < N/d} f(dn), \tag{1.4.2}$$

where $d = [d_1, d_2]$ is square-free. Introducing the convolution inverse f_1 of f, the last factor is expressed as

$$f(dn) = \mu(d) \sum_{\substack{u|n \\ u|d^\infty}} f\left(\frac{n}{n}\right) f_1(du).$$

This and (C_3) with the trivial character give

$$\sum_{n<y} f(dn) = \mu(d) \sum_{\substack{u|d^\infty \\ u<y}} f_1(du) \left\{\mathcal{F}\frac{y}{u} + o\left(D\left(\frac{y}{u}\right)^\gamma\right)\right\}.$$

We note that we have 28

$$\begin{aligned}
\sum_{\substack{u|d^\infty \\ u<y}} \frac{f_1(du)}{u} &= \sum_{u|d^\infty} \frac{f_1(du)}{u} \sum_{\substack{u|d^\infty \\ u\geq y}} \frac{f_1(du)}{u} \\
&= \mu(d) d \prod_{p|d} (1 - F_p^{-1}) + o\left\{y^{\gamma-1} \sum_{u|d^\infty} \frac{|f_1(du)|}{u^\gamma}\right\}.
\end{aligned}$$

Hence

$$\sum_{n<y} f(dn) = \mathcal{F} yd \prod_{p|d} (1 - F_p^{-1}) + 0\left\{(D + \mathcal{F}) y^\gamma \sum_{u|d^\infty} |f_1(du)| u^{-\gamma}\right\}.$$

Inserting this into (1.4.2) and reacalling (1.4.1) we have

$$K = \mathcal{F}N \sum_{d_1,d_2<R} \lambda(d_1)\lambda(d_2) \prod_{p|d_1d_2} (1-F_p^{-1}) + O\left\{(D+\mathcal{F})N^{\gamma}R^{2(l-\gamma)+\epsilon}\right\}, \tag{1.4.3}$$

where we have used the fact that (C_1) implies $f_1(n) = O(n^{\epsilon})$ for all n.

Then, by a routine argument, we can conclude that the optimal λ is given by

$$\lambda(d) = \mu(d)\frac{G_d(R/d)}{G_1(R)} \prod_{p|d} F_p, \tag{1.4.4}$$

where

$$G_d(x) = \sum_{\substack{r<\chi \\ (r,d)=1}} \mu^2(r)/g(r) \tag{1.4.5}$$

29 with

$$g(\Gamma) = \prod_{p|\Gamma} (F_p - 1)^{-1}. \tag{1.4.6}$$

And this choice of λ gives

$$\sum_{d_1,d_2<R} \lambda(d_1)\lambda(d_2) \prod_{p|d_1d_2} (1-F_p^{-1}) = G_1(R)^{-1}. \tag{1.4.7}$$

Also, we have

$$G_d(R/d) \leq G_1(R) \prod_{p|d} F_p^{-1} \tag{1.4.8}$$

which implies, in particular, (1.4.1) for the λ defined by (1.4.4). Further, we should note that we have

$$G_d(R) \geq K(d)G_1(R). \tag{1.4.9}$$

Now, let us observe that the optimal λ defined by (1.4.4) yields the relation

$$\sum_{\substack{d|n \\ d<R}} \lambda(d) = G_1(R)^{-1} \sum_{r<R} \frac{\mu^2(r)}{g(r)} \Phi_r(n),$$

where

$$\Phi_R(n) = \mu((r,n))\, g((r,n)). \tag{1.4.10}$$

This should be compared with (1.2.1); we may expect that for this Φ_r there will be an analogue of THEOREM 3. The object of the present section is to show that this is indeed the case.

To this end, we shall consider the estimation of the expression **30**

$$J = \sum_{\substack{q<Q\\ r<R\\ (q,r)=1}} \frac{\mu^2(r)}{K(q)g(r)} \sideset{}{^*}\sum_{\chi \pmod q} \Big| \sum_{M\le n<M+N} \chi(n)\Phi_r(n) f(n)^{\frac{1}{2}} a_n\Big|^2$$

with $\{a_n\}$ being arbitrary complex numbers; we assume $(C_1), (C_2), (C_3)$ naturally, and also

$$N = O(M).$$

But, as before, it is advantageous to estimate, instead, the dual form

$$J* = \sum_{M\le n<M+N} f(n) \Big| \sum_{\substack{q<Q\\ r<R\\ (q,r)=1}} \left(\frac{\mu^2(r)}{k(q)g(r)}\right)^{1/2} \Phi_r(n) \sideset{}{^*}\sum_{\chi \pmod q} \chi(n) b(r,\chi)\Big|^2,$$

where $\{b(r,\chi)\}$ are arbitary complex numbers. Expanding out the squares and changing the order of summation we have

$$J* = \sum_{\substack{q,q'<Q\\ r,r'<r\\ (q,r)=(q',r')=1}} \left\{\frac{\mu^2(r)\mu^2(r')}{K(q)K(q')g(r)g(r')}\right\}^{\frac{1}{2}}$$

(1.4.11)

$$\times \sideset{}{^*}\sum_{\substack{\chi(\bmod q\\ \chi' \pmod{q'})}} \{S(M+N,\chi\overline{\chi'}) - S(M,\chi\overline{\chi'}; r, r'\}\, b(r,\chi)\overline{b(r',\chi')}$$

where

$$S(y,\chi; r, r') = \sum_{n<y} \chi(n)\Phi_{r'}(n) f(n).$$

In order to estimate the last sum, we consider first the function **31**

$$\sum_{n=1}^{\infty}\chi(n)\Phi_r(n)f(n)n^{-s}, \tag{1.4.12}$$

which converges absolutely for $Re(s) > 1$. Recalling that Φ_r is multiplicative, and r, r' are square-free, this can be decomposed as

$$\begin{aligned}\left\{\sum_{(n,rr')=1}\chi(n)f(n)n^{-s}\right\}&\left\{\sum_{n|(\frac{[r,r']}{(r,r')})^{\infty}}\chi(n)\Phi_r(n)\Phi_{r'}(n)f(n)n^{-s}\right\}\\ &\times\left\{\sum_{(n,rr')^{\infty}}\chi(n)\Phi_r(n)\Phi_{r'}(n)f(n)n^{-s}\right\}\\ &= P_1P_2P_3,\end{aligned}$$

say. Introducing the functions

$$\begin{aligned}F(s,\chi) &= \prod_p F_p(s,\chi),\\ F_p(s,\chi) &= \sum_{m=0}^{\infty}\chi(p^m)f(p^m)p^{-ms},\end{aligned} \tag{1.4.13}$$

we have

$$p_1 = F(s,\chi)\prod_{p|rr'}(F_p(s,\chi)^{-1},$$

if $Re(s)$ is sufficiently large. Also (1.4.10) implies

$$P_2 = \prod_{p|\frac{[r,r']}{(r,r')}}(1-(F_p-1)^{-1}(F_p(s,\chi)-1))$$

and

$$P_3 = \prod_{p|(r,r')}(1+(F_p-1)^{-2}(F_p(s,\chi)-1)).$$

32 Thus, we see that (1.4.12) is equal to

$$F(s,\chi)A_{r,r'}(s,\chi),$$

where

$$A_{r,r'}(s,\chi), = P_2 P_3 \prod_{p|rr'} F_p(s,\chi)^{-1}$$

(1.4.14)

$$= \sum_{n|(rr')^\infty} \chi(n) f_2(n) n^{-1},$$

say; here $Re(s)$ is only to be positive. In particular, we have

$$\Phi_r(n)\Phi_{r'}(n) f(n) = \sum_{\substack{d|n \\ d|(rr')^\infty}} f\left(\frac{n}{d}\right) f_2(d).$$

This and the condition (C_3) give, for $\chi \pmod q$,

$$S(y,\chi;r,r') = \sum_{\substack{d<y \\ d|(rr')^\infty}} \chi(d) f_2(d) \left\{ E(\chi)\mathcal{F}K(q)\frac{y}{d} + 0(Dq^\beta \left(\frac{y}{d}\right)^\gamma \right\},$$

whence

$$S(y,\chi;r,r') = \mathcal{F}E(\chi)K(q)A_{r,r'}(1,\chi)y + 0\left\{(D+\mathcal{F})q^\beta y^\gamma \sum_{d|(rr')^\infty} |f_2(d)| d^{-\gamma}\right\}.$$

By the definition (1.4.14) of f_2, the last sum over d is equal to

$$\sum_{P|\left(\frac{[r,r']}{(r,r')}\right)^\infty} \left\{1 + (F_p - 1)^{-1} \sum_{m=1}^{\infty} \frac{f(p^m)}{p^{my}}\right\} \prod_{P|(r,r')}$$

$$\left\{1 + (F_p - 1)^{-2} \sum_{m=1}^{\infty} \frac{f(p^m)}{p^{my}}\right\} \times \prod_{p|rr'} \left\{\sum_{m=0}^{\infty} \frac{|f_1(p^m)}{p^{m\gamma}}\right\},$$

and, by (C_1) and (C_2), this is **33**

$$0((rr')^{\alpha+\epsilon}[r,r']^{-\gamma}).$$

We should remark also that if χ is principal $\pmod q$ and $(rr',q) = 1$ them, by (1.4.14), we have

$$A_{r,r'}(1,\chi) = g(r)\delta_{r,r'} \quad \text{(Kronecker's delta)}.$$

Inserting these into (1.4.11), and recalling that $N = O(M)$ we obtain, after some elementary estimations,

$$J* = \left\{\mathcal{F}N + 0(D + \mathcal{F})M_Q^{2(1+\beta)+\epsilon_R 2\alpha+\epsilon}\right\} \sum_{\substack{q<Q \\ r<R \\ (q,r)=1}} \sum_{\chi \pmod q}^{*} |b(r,\chi)|^2.$$

Hence, returning to J via the duality principle (LEMMA 2), we get the following hybridization of the Selberg sieve for multiplicative functions and the multiplicative large sieve inequality:

Theorem 5 (THE HYBRID DUAL SIEVE FOR MULTIPOICATIVE FUNCTIONS)**.** *On the assumptions* $(C_1), (C_2), (C_3)$ *we have, for any* $N = O(M)$ *and arbitrary complex numbers* $\{a_n\}$*,*

$$\sum_{\substack{q<Q \\ r<R \\ (q,r)=1}} \frac{\mu^2(r)}{K(q)g(r)} \sum_{\chi \pmod q}^{*} | \sum_{M\leq n<M+N} \chi(n)\Phi_r(n)f(n)^{\frac{1}{2}}a_n|2$$

$$\leq \{\mathcal{F}N + 0(Y_f(M; Q, R))\} \sum_{M\leq n<M+N} |a_n|^2,$$

34 *where* $g(r)$ *and* $\Phi_r(n)$ *are defined by* (1.4.6) *and* (1.4.10)*, respectively, and*

$$Y_f(M; Q, R) = (\mathcal{F} + D)M^{\gamma}Q^{2(1+\beta)+\epsilon_R 2\alpha+\epsilon}.$$

Next we turn to the basic lemmas which will be utilised in PART II when we make important applications of THEOREM 5 TO Dirichlet's L-functions.

First we qote the fundamental

Lemma 3. *We have, for* $T \geq 1$*,*

$$\int\limits_{-T}^{T} | \sum_{n=1}^{\infty} a_n n^{it}|^2 dt << T^2 \int_0^{\infty} | \sum_{y\leq n<ye^{1/T}} a_n|^2 \frac{dy}{y},$$

provided the right side converges.

The combination of THEOREM 5 and LEMMA 3 yields immediately

Lemma 4. *We have, for $T \geq 1$,*

$$\sum_{\substack{q<Q\\ r<R\\ (q,r)=1}} \frac{\mu^2(r)}{K(q)g(r)} \sum_{\chi \pmod q}^{*} \int_{-T}^{T} |\sum_{n=1}^{\infty} \chi(n)\Phi_r(n)f(n)a_n n^{it}|^2 dt$$
$$\ll \sum_{n=1}^{\infty} (\mathcal{F}n + TY_f(n : Q, R))f(n)|a_n|^2,$$

provided the right side converges.

In our applications of THEOREM 5, an important rôle will be played by the multiplicative property of Φ_r, which is embodied in

Lemma 5. *Let r be square free and let $b\ \xi_d = o(|\mu(d)|d^{\epsilon})$. Then we have,* 35
for s with sufficiently large real part,

$$\sum_{n=1}^{\infty} \chi(n)\Phi_r(n)f(n)\left(\sum_{d|n} \xi_d\right) n^{-s} = F(s,\chi)M_r(s,\chi;\xi), \tag{1.4.15}$$

where

$$M_r(s,\chi;\xi) = g(r)\sum_{d=1}^{\infty} \xi_d \mu((r,d)) \prod_{p|d}(1 - f_p)(s,\chi)^{-1}) \prod_{p\nmid d}(f_p(s,X)^{-1} f_p - 1), \tag{1.4.16}$$

$F(s,\chi)$and$F_p(s,\chi)$ being defined by (1.4.13)

To show this, we note first that, for square-free r, we have

$$\Phi_r(dn) = \Phi_r(d)\Phi_u(n), u = r/(r,d).$$

Thus, the left side of (1.4.15) is equal to

$$\sum_{d=1}^{\infty} \chi(d)\xi_d \Phi_r(d)d^{-s} \sum_{n=1}^{\infty} \chi(n)\Phi_u(n)f(dn)_n^{-s}. \tag{1.4.17}$$

Because of the multiplicativity of Φ_u, this inner-sum may be written as

$$\left\{\sum_{(n,d)=1} \chi(n)\Phi_u(n) f(n)n^{-s}\right\}\left\{\sum_{n|d^\infty} \chi(n)\Phi_u(n) f(dn)n^{-s}\right\}.$$

But $n|d^\infty$ implies $\Phi_u(n) = 1$. Hence this product is equal to

$$\begin{aligned} & d^{s-}\chi^{(d)}\prod_{p\nmid d}(1+\Phi_u(p)(F_p(s,X)-1))\prod_{p|d}(F_p(s,\chi)-1)\\ &= d^{s-}\chi^{(d)} f(s,\chi)\prod_{\substack{p\nmid d\\ p|r}}(F_p-1)^{-1}\prod_{p|d}(1-F_p(s,\chi)^{-1})\prod_{\substack{p\nmid d\\ p|r}}(F_p(s,\chi)^{-1}F_p-1); \end{aligned}$$

here we have used the fact that d can be assumed to be square free. Inserting this into (1.4.17) and noticing that

$$\Phi_r(d)\prod_{\substack{p\nmid d\\ p|r}}(F_p-1)^{-1} = \mu((r,d))g(r)$$

36 we obtain the assertion of the lemma.

We now introduce THEOREM 4 into our discussion: but, for this sake, we have to replace the condition (C_1) on f by the stronger $(C_1')f$ is a non-negative multiplicative function such that there exists a k satisfying

$$f(n) = o(\tau_k(n))$$

for all n.

Then we have

Lemma 6. *On the conditions* (C_1'), (C_2) *and* (C_3)

$$\sum_{r\le z} 1 + k\zeta\frac{\mu^2(r)}{g(r)}M_r\left(1,\chi_0;\Lambda^{(k)}\right)^2 = O((\mathscr{F}\log z)^{-1})$$

for any $z > (D+\mathscr{F})^\epsilon$, *where* χ_0 *is the trivial character, and the functions* $\Lambda^{(k)}$ *and* M_r *are defined by* (1.3.10) *and* (1.4.16), *respectively.*

To prove this, we note first that

$$M_r(1,\chi_0;\Lambda^{(k)}) = \mu(r)g(r) \sum_{\substack{d<z \\ d\equiv 0 \pmod r}} 1 + k\zeta^{\Lambda} d^{(k)} \prod_{p|d}(1-F_p^{-1}).$$

Hence, denoting by H the sum to be estimated, we have

$$\begin{aligned} H &= \sum_{r<z} 1 + k\zeta\mu^2(r)g(r) \left\{ \sum_{\substack{d<z \\ d\equiv 0 \pmod r}} \Lambda_d^{(k)} \prod_{p|d}(1-F_p^{-1}) \right\}^2 \\ &= \sum_{d_1,d_2<z} 1 + k\zeta\Lambda_{d_1}^{(k)}\Lambda_{d_1}^{(k)} \prod_{p|d_1d_2}(1-F_p^{-1}). \end{aligned}$$

37

Thus, just as (1.4.3), we have

$$\sum_{n<n} f(n) \left(\sum_{d|n} \Lambda_d^{(k)} \right)^2 = \mathscr{F}HN + 0\left\{(D+\mathscr{F})N^{\gamma}{}_{z}{}^{2(1+k\zeta)(1_{-\gamma})+\epsilon}\right\}$$

whence, by partial summation, we have for $\omega > 1$ and $b > 0$,

$$\begin{aligned} &\sum_{n>z} bf(n) \left(\sum_{d|n} \Lambda_d^{(k)} \right)^2 n^{-\omega} \\ &\quad = (\omega-1)^{-1}\mathscr{F}Hz^{b(1-\omega)} + 0\left\{(D+\mathscr{F})z^{b(\gamma-\omega)+2(1+k\zeta)(1_{-\gamma})+\epsilon}\right\}. \end{aligned} \tag{1.4.18}$$

If we set $\omega = 1 + (\log z)^{-1}$ and take b sufficiently large then this is equal to

$$e^{-b}H\mathscr{F}\log z + 0(z^{-\epsilon}).$$

But, by virtue of (C_1') and THEOREM 4, the left side of (1.4.18) is bounded, whence the assertion of the lemma.

NOTES (I)

The origin of Selberg's Λ^2 -sieve can be found in his deep investigations [68] [71] (see also [70]) on the distribution of zeros of the Riemann zeta-function in the vicinity of the cirtical line. Refining the ideas of Bohr, Landau and Carleson, Selberg was led to the problem of **38**
making the following quadratic for of λ

$$\int_{-T}^{T} |\zeta(\frac{1}{2}+\mathrm{it}) \sum_{d<z} \lambda(d) d^{-\frac{1}{2}-\mathrm{it}} - 1|^2 dt$$

as small as possible on the side conditon $\lambda(1) = 1$, where z is to be taken suitably in connection with the sufficiently large parameter T. Applying certain mean-value theorems for $\zeta(s)$, he could reduce the problem to the one of determining the minimum value of the quadratic from

$$\sum_{d_1,d_2<z} \frac{\lambda(d_1)\lambda(d_2)}{[d_1,d_2]} (\lambda(1) = 1),$$

which corresponds just to (1.3.1) The sieve-effect of the argument with which Selberg solved this extremal problem was explicitly formulated on a general setting in his later papers [72] [73] [74]. It is noteworthy that the Λ^2-sieve was created in the course of deeper studies of the analytical behaviour of the Riemann zeta function, and that, as we shall see in PART II, our account of his theory has also important applications to $\zeta(s)$ and $L(s,\chi)$; this seems to agree appreciably with Selberg's opinion expressed in the last lines of [73] [74].

We formulated Selberg's idea in a generalized from as THEOREM 1, for we have hope that one may find applications of it to the problems with Ω not necessarily defined by the congruence condition (1.1.2), on which, however, all applications know at present are made.

39 One may want to see how well the right side of (1.1.4) approximates to the left side. For this,we refer to NOTES (II) where we shall give an explicit representation of the difference between the two sides, revealing the mechanisum behind the device of Selberg which at first may look somewhat *ad* hoc.

It is a remarkable coincidence that two fundamental sieve ideas, Selberg's and Linnik's were created almost simultaneously, and this fact becomes more interesting when we know that between them is a duality relation as we have shown in the second section.

LEMMA 1 is the latest version of Linnik's large sieve, and is due to Selberg. To prove this, Selberg employed a delicately chosen function in conjuction with his inequility of Bessel's type [[48] Lemma 1.8]. P. Cohen has shown, however, that LEMMA 1 is an immediate consequence of an inequality of Montgomery and Vaughan [51] in which occurs the factor $N + \delta^{-1}$ instead of $N - 1 + \delta^{-1}$. For the details see the expository article [49] of Montgomery.

THEOREM 2 is due to Selberg [77]. This remarkable result implies as its speical cases the large sieves of Montgomery [[48], p. 25], Jhonsen [35] and Gallagher [16]. Our proof of THEOREM 2 has a difference from Selberg's in that we have appealed to LEMMA 1, an argument which was employed formerly by Motohashi [[54], II] in his alternative proof of Montgomery's large sieve. We should point out the possibility **40**
of generalizing THEOREM 2 into the directions indicated by Salerno-Viola [67] and Gallagher [17].

The duality relation between Selberg's Λ^2-sieve and Linnik's large sieve was observed by not a few people simultaneously in published and unpublished forms. THEOREM 3 which is due to Motohashi [[54], III] summarises the former discussions on this matter each of which was made on some special assumptions on Ω It shows that $\{\psi_r(n, \Omega) g(r)^{-\frac{1}{2}}\}$ behaves just like $\{\chi(n); \chi$ primitive$\}$, i.e., they share the property which may be called quasi-orthogonality. This was first observed by Selberg [76] when he obtained (1.2.10), and called $\{\psi_r(n)\}$ pseudo-charecters; but the relation betwen ψ_r and the Selberg sieve was remarked explicity by Motohashi [[58], p. 166].

(1.2.13) and (1.2.14) are due to Bombieri and Devenport [8] (see also Bombieri [4]), which, apart from the fundamental work [45] of Linnik, was the first instance that the sieve effect of the large sieve was clearly perceived (1.2.13) has had a deep application to the theory of L-functions, as Gallagher showed in his important work [15]. The same can be said about (1.2.10), as we shall show in § 5.2.

The Brun-Titchmarsh theorem (1.2.15) is introduced here only for the sake of illustrating the generality of THEOREM 3; a further discus- **41** sion on this basic sieve result will be given in §4.3.

We have seen that a drasitc specialization of THEOREM 3 yields important results knows already. In the proop of THEOREM 3,we have used, however, nothing deeper than the additive large sieve inequality and the duality principle, both of which are, in fact, of very elementary character. Thus one may expect that, on some special conditions, more sophisticated tools will produce improvements upon THEOREM 3. In the case of (1.2.10) this was confimed by Motohashi [[54], the first note], but the general case seems to be a difficult problem. Relating to this question we should note that these might be a possibility to improve, in some sense, upon LEMMA 1 for the Farey sequence in place of general well-spaced sequence $\{x_r\}$.

(1.3.4) is due to Barban ([2]), and (1.3.5) to Selberg [69]. Graham [19] gave an elegant proof of (1.3.4), and even succeeded in replacing it by an asymptotic relation. THEOREM 4 is due to Motohashi [58]; Jutila [41] obtained an analogue of Graham's result for the weights $\{\Lambda_d^{(k)}\}$.

THEOREM 5 is due to Motohashi [58]. In deriving the important artimitic function Φ_r from f, we used the standard argument of manipulating the Λ^2-sieve, for, as already mentioned,we do not have anything analogous to LEMMA 1 in the situation of the fourth section. Hence it is desrable to have an additive large sieve inequality which admits the
42 weight f. But to this end, we would have to find first a sort additive characters derived from f which substitute for $\exp(2\pi iX)$ of LEMMA 1.

LEMMA 3 is the famous inequality of Gallagher [15]. LEMMA 5 is essentially due to Selberg ([50]) who showed it for ψ_r; this will be a key lemma in our application of THEOREM 5 to Dirichlet's L-functions.

Chapter 2

Elements of the Combinatorial Sieve

WE NOW TURN to another topic in sieve methods; in the present and 43
the next chapters, we shall develop a detailed study of some important aspects of the combinatotial sieve method, which is essentially a system of devices of introducing effective truncations into the exact -seive of Eratosthenes, and, as contrasted with the Λ^2-sieve, the most notable feature of which lies in that it leads simultaneously to upper and lower sieve bounds on some fairly general conditions.

Partly because of their independent interest, we shall discuss in the present chapter the basic combinatorial or logical identities, and then in the next chapter exhibit their power in the particular application to the linear sieve situation. We shall try to explain the motivation behind those combinatorial identities, for, it seems that the combinatorial sieve methods lacks the straight forwardness which characterises the Λ^2-sieve and makes easiers to understand it.

2.1 Rosser's Identity

To begin with, we repeat the conventions introduced in §1.1 in a much simplified from

We suppose that to each prime p is assigned a set $\Omega(p)$ of residue 44 classes $\pmod p$. For a square-free d, we denote by $\Omega(d)$ the set of residue classes $\pmod d$ arising (in the way of Chinese Remainder Theorem) from those of $\Omega(p)$, $p|d$. We shall write, in the sequel, $n\epsilon\Omega(d)$ instead of $n \pmod d \epsilon\Omega(d)$, and naturally, we have $n\epsilon\Omega(1)$ for all integer n.

Next let $z \geq 2$ be a parameter, and put

$$\triangle(n,z) = \prod_{\substack{p<z \\ n\epsilon\Omega(p)}} P$$

and, as usual,

$$p(z) = \prod_{p<z} p.$$

Let A be a finitie sequence of integers, and put, for a square-free d,

$$A_d = \{a\epsilon A; a\epsilon\Omega(d)\}.$$

Further, let Θ be an arbitrary function defined on $\mathbb{N}$, and put

$$S(A,z;\Theta) = \sum_{a\epsilon A} \Theta(\triangle(a,z)).$$

Then the sieve problem on which we are going to discuss is to find a good(in one or another sense) estimate of $S(A,z;\Theta)$ in terms of $|A_d|$ under suitable condition on the nature of A, Ω, and Θ. To solve this
45 problem in a very special but highly important case, i.e. the linear sieve situation which will be defined in the next chapter, we shall employ the combinatorial sieve method; the whole theory of it is built on the very simple

Lemma 7 (THE BUCHYSTAB IDENTITY)**.** *We have*

$$S(A,z;\Theta) = \Theta(1)|A| \sum_{p<z} S(A_p, p; \Theta_p), \tag{2.1.1}$$

where Θ *is defined by*

$$\theta_p(n) = \Theta(n) - \Theta(pn).$$

To show this, let

$$\triangle(a,z) = p_1 p_2 \ldots p_r, \; p_1 < p_2 < \cdots < p_r < z.$$

Then we have

$$\begin{aligned}\Theta(1) - \Theta(p_1 p_2 \cdots p_r) &= \sum_{j=0}^{r-1} (\Theta(p_1 p_2 \cdots p_j)) - (\Theta(p_1 p_2 \cdots p_{j+1})) \\ &= \sum_{j=0}^{r-1} \Theta_{p_{j+1}}(p_1 p_2 \cdots p_j) \\ &= \sum_{j=0}^{r-1} \Theta_{p_{j+1}}(\triangle(a, p_{j+1})),\end{aligned}$$

which amounts to

$$\Theta(\triangle(a,z)) = \Theta(1) - \sum_{\substack{p<z \\ a\epsilon\Omega(p)}} \Theta_p(\triangle(a,p)),$$

and this is apparently equivalent to the assertion of the lemma.

The Buchstab identity obviously admits of iteration. And to state **46**
the result of the infinite iteration in a compact from we introduce the function Θ_d defined by

$$\Theta_d(n) = \sum_{r|d} \mu(r)\Theta(rn).$$

Then LEMMA 7 yields readily

$$S(A,z;\Theta) = \sum_{d|p(z)} \mu(d)\Theta_d(1)|A_d| \tag{2.1.2}$$

which is a little generalized version of the exact-sieve of Eratosthenes.

As is commonly remarked in sieve literature, (2.1.2) is a useless identity, for, it involves too many terms to be handled with. Thus, if we want to keep the number of terms within a manageable size, we have to discard certain summands on the right side; the cost of doing so is to give up the exact identity. Since the process of casting away some summands in (2.1.2) is equivalent to attaching the weight 1 to some specially chosen divisors of $P(z)$ and the weight 0 to all other, we are naturally led to the problem to find a weighted version of (2.1.2).

To formulate the answer to this problem, we introduce an arbitrary function ρ defined $\mathbb{N}$ and satisfying

$$\rho(1) = 1,$$

and put

$$\sigma(d) = \rho(\frac{d}{p(d)}) - \rho(d), \sigma(1) = 0.$$

47 Here and throughout the sequel, the symbol

$$p(d)$$

stands for the least prime factor of $d > 1$.

Then we have the fundamental

Theorem 6.

$$S(A, z; \Theta) = \sum_{d|p(z)} \mu(d)\rho(d)\Theta_d(1)|A_d| + \sum_{d|p(z)} \mu(d)\sigma(d)S(A_d, p(d); \Theta_d).$$

Proof. is quite simple. Inserting into the right side the expression

$$S(A_d, p(d); \Theta_d) = \sum_{\ell|P(p(d))} \mu(\ell)\Theta_{d\ell}(1)|A_{d\ell}|,$$

which is a particular case of (2.1.2) we immediately recover the left side. □

But the following alternative argument seems to be more instructive, if tedious. We introduce an arbitrary function λ defined on $\mathbb{N}$ and satisfying $\lambda(1) = 1$, and as a first step we modify (2.1.1) trivially as

$$(2.1.3) \quad S(A, z; \Theta) = \Theta(1)|A| - \sum_{p<z} \lambda(p)S(A_p, p; \Theta_p)$$

$$-\sum_{p<z}(1-\lambda(p))S(A_p,p;\Theta_p).$$

Similarly, we have

$$S(A_p,p;\Theta_p)=\Theta_p(1)|A_p|-\sum_{p'<p}\lambda(pp')S(A_{pp'},p';\Theta_{pp'})$$
$$-\sum_{p'<p}(1-\lambda(pp'))S(A_{pp'},p';\Theta_{pp}). \tag{2.1.4}$$

Inserting this into the first sum over p on the right side of (2.1.3) we **48**
get

$$\begin{aligned}S(A.z;\Theta)&=\Theta(1)|A|-\sum_{p<z}\Theta_p(1)\lambda(p)|A_p|\\&+\sum_{p'<p<z}\lambda(p)\lambda(pp')S(A_{pp'},p';\Theta_{pp'})-\sum_{p<z}(l-\lambda(p))S(A_p,p;\Theta_p)\\&+\sum_{p'<p<z}\lambda(p)(1\lambda(pp'))S(A_{pp'},p';\Theta_{pp'}).\end{aligned}$$

This is the case $k=2$ of the identity

$$\begin{aligned}S(A,z;\Theta=&\sum_{\substack{d|p(z)\\ \omega(d)<k}}\mu(d)\Theta_d(1)\tilde{\rho}(d)|A_d|\\&+(-1)^k\sum_{\substack{d|p(z)\\ \omega(d)=k}}\tilde{\rho}(d)S(A_d,p(d);\Theta_d)\\&+\sum_{\substack{d|p(z)\\ \omega(d)\leq k}}\mu(d)\tilde{\sigma}(d)S(A_d,p(d);\Theta_d),\end{aligned}\tag{2.1.5}$$

where $\tilde{\rho}$ and $\tilde{\sigma}$ are defined by

$$\tilde{\rho}(d)=\lambda(p_1)\lambda(p_1p_2)\cdots\lambda(p_1p_2\cdots p_r),\tilde{\rho}(1)=1$$

and

$$\tilde{\sigma}(d)=\tilde{\rho}\left(\frac{d}{p(d)}\right)-\tilde{\rho}(d),\tilde{\sigma}(1)=0$$

if $d = p_1 p_2 \cdots p_r, p_1 > p_2 > \cdots > p_r$. We may establish (2.1.5) by the induction on k: we need only to replace $S(A_d, p(d); \Theta_d)$ of the second sum on the right of (2.1.5) by the expression

$$S(A_d, p(d); \Theta_d) = \Theta_d(1)|A_d| - \sum_{p<p(d)} \lambda(dp) S\left(A_{dp}, p; \Theta_{dp}\right) - \sum_{p<p(d)} (1 - \lambda(dp)) S\left(A_{dp}, p; \Theta_{dp}\right),$$

49 which is a special case of (2.1.3), getting the formula (2.1.5) with $k+1$ in place of k. We then take k in (2.1.5) sufficiently large ($> \pi(z)$, say), and obtain

(2.1.6)
$$S(A, z; \Theta) = \sum_{d|P(z)} \mu(d)\Theta_d(1)\tilde{\rho}(d)|A_d| + \sum_{d|P(z)} \mu(d)\tilde{\sigma}(d) S(A_d, p(d); \Theta_d).$$

This is equivalent to the assertion of THEOREM 6, because, as is easily seen, we can always find a λ such that $\tilde{\rho} = \rho$.

Compared with the first, the second proof has an advantage in that the procedure of truncation-iteration of the Buchstab identity is clearly exibited in it. Moreover, it will turn out that the formulation (2.1.6) of THEOREM 6 is more convenient for our later purpose.

We now restict ourselves to the case where Θ is the unit measure placed at 1 so that $S(A, z; \Theta)$ is equal to

$$S(A, z) = |\{a \in A; a \notin \Omega(p) \text{ for all } p < z\}|$$

and

$$\Theta_d(1) = 1$$

for all d. Then we have, by (2.1.6),

(2.1.7) $$S(A, z) = \sum_{d|P(z)} \mu(d)\tilde{\rho}(d)|A_d| + \sum_{d|P(z)} \mu(d)\tilde{\sigma}(d) S(A_d, p(d)).$$

50 Further, let us set

(2.1.8) $$0 \leq \lambda(d) \leq 1$$

so that

$$0 \le \tilde{\rho}(d) \le 1, 0 \le \tilde{\sigma}(d) \le 1. \tag{2.1.9}$$

And let us try to express 'good' upper and lower bounds of $S(A, z)$ in terms of $|A_d|$ via the formula (2.1.7). This means, in other words, that we have to discard the terms $S(A_d, p(d))$ on the right side of (2.1.7); this should be done, of course, in the manner to keep at a minimum the loss caused by doing so. In general, we can assume, however, nothing more than the trivial information

$$S(A_d, p(d)) \ge 0.$$

This implies, in particular, for $\nu = 0$ and 1,

$$(-1)^\nu\{S(A, z) - \sum_{d|P(z)} \mu(d)\tilde{\rho}(d)|A_d|\} \le \sum_{\substack{d|P(z)\\ \omega(d)\equiv\nu \pmod 2}} \tilde{\sigma}(d)S(A_d, p(d)),$$

since we have (2.1.9). Here the equality holds if we set $\tilde{\sigma}(d) = 0$ for all $d|P(z)$ such that $\omega(d) \equiv \nu + 1 \pmod 2$. The simplest way to attain this is to set

$$\lambda(d) = 1 \text{ it } = 1\omega(d) \equiv +1 \pmod 2, \tag{2.1.10}$$

which we shall impose on λ henceforth; we write $\tilde{\rho}_\nu, \tilde{\sigma}_\nu$ for $\tilde{\rho}, \tilde{\sigma}$ with λ satisfying this condition. Then we have

$$S(A, z) = \sum_{d|P(z)} \mu(d)\tilde{\rho}_\nu(d)|A_d| + (-1)^\nu \sum_{d|P(z)} \tilde{\sigma}_\nu(d)S(A_d, p(d)). \tag{2.1.11}$$

In paticular, we have **51**

$$(-1)^\nu\{S(A, z) - \sum_{d|P(z)} \mu(d)\tilde{\rho}_\nu(d)|A_d|\} \ge 0;$$

this means that we have neglected all $S(A_d, p(d))$ on the right side of (2.1.11), and thus a certain inaccuray is brought in.

Now, we note trivial but crucial fact that $S(A, z)$ is a decreasing function of the parameter z. Thus the negligence of $S(A_d, p(d))$ with $p(d)$

which is 'small' for A_d causes most likely a relatively 'large' loss, to avoid this we should better set $\tilde{\sigma}_\nu = 0$ for such d. One of the most fruitful devices to make explicit the 'smallness' of $p(d)$ for A_d is to introduce two parameters y and $\beta > 1$, and to define $p(d)$ to be 'small' for A_d if $p(d) < (y/d)^{1/\beta}$. The simplest way to realize this in terms of λ is to set

$$\lambda(d) = \begin{cases} 1 & \text{if } \omega(d) \equiv \nu \pmod 2,\ p(d)^\beta d < y. \\ 0 & \text{if } \omega(d) \equiv \nu \pmod 2,\ p(d)^\beta d \geq y. \end{cases}$$

besides (2.1.10). Then $\tilde{\rho}_\nu$ and $\tilde{\sigma}_\nu$ are the characteristic functions $\rho_\nu(d) = \rho_\nu(d; y, \beta)$ and $\sigma_\nu(d) = \sigma_\nu(d; y, \beta)$ of the sets

$$(2.1.12) \quad D_1^\nu(y,\beta) = \left\{ d; \begin{array}{l} d = p_1 p_2 \cdots p_r, p_1 > p_2 > \cdots > p_r; \\ p_{2k+\nu}^{\beta+1} p_{2k+\nu-1} \ldots p_1 < y \text{ for } 1 \leq 2k+\nu \leq r \end{array} \right\}$$

and

(2.1.13)
$$D_1^\nu(y,\beta) = \left\{ d; \begin{array}{l} d = p_1 p_2 \cdots p_r, p_1 > p_2 > \cdots > p_r, r \equiv \nu \pmod 2 \\ \rho_\nu(p_1 p_2 \cdots p_{r-1} = 1, p_r^{\beta+1} p_{r-1} \cdots p_1 \geq y \end{array} \right\},$$

52 respectively.

In this way, we are led to

Lemma 8 (ROSSER'S IDENTITY). *Let ρ_ν and σ_ν be as above. Then we have*

$$(2.1.14) \quad S(A,z) = \sum_{d|P(z)} \mu(d)\rho_\nu(d)|A_d| + (-1)^\nu \sum_{d|P(z)} \sigma_\nu(d) S(A_d, p(d)).$$

We should note here that this is a logical dentity, so the choice of the parmaeters y and β is at our disposal. Since the larger y and the smaller β give the wider $D_1^{(\nu)}(y,\beta)$, the support of ρ_ν, as can be seen from (2.1.12) if is desirable to take y and β as large and small as possible, respectively. Under a fairly general condition to be specified in the next chapter, we shall show how to determine the smallest, i.e. the optimal value of β, and also a very penetrating device which allows us take y unexpectedly large in some practically important situations.

2.2 The Fundamental Lemma

In this section, we shall first show an important application of Rosser's formula which is also a basis preparation for the next chapter. We shall then turn to a tentative explanation of Rosser's motivation behind his 53
formula which was introduction rather abruptly in the above.

First of all, we have to make precise the information on $|A_d|$. We assume that there exists a non-negative multiplication function δ and a parameter X such that

$$\delta(p) < p \quad \text{for all} \quad p,$$

and

$$R_d = |A_d| - \frac{\delta(d)}{d} X \tag{2.2.1}$$

is small, in one or another sense, for d, $d|P(z)$, lying in a certain range. Then we introduce the notation

$$V(z) = \prod_{p<z} \left(1 - \frac{\delta(p)}{p}\right).$$

We note that we have an analogue of (2.1.1) for $V(z)$:

$$V(z) = 1 - \sum_{p<z} \frac{\delta(p)}{p} V(p).$$

And this is utilization, in much the same way as in the same way as in the proof of (2.1.14), to prove the identity

$$V(z) = \sum_{d|P(z)} \mu(d)\rho_\nu(d)\frac{\delta(d)}{d} + (-1)^\nu \sum_{d|P(z)} \sigma_\nu(d)\frac{\delta(d)}{d} V(p(d)). \tag{2.2.2}$$

We shall need also an information on the size of the elements of $D_1^{(\nu)}(y, \beta)$.

Lemma 9. *If $z \leq y^{1/2}$ and $\rho_\nu(d) = 1$, then we have*

$$\log d < \left(1 - \frac{1}{2}\left(\frac{\beta - 1}{\beta + 1}\right)^{\omega(d)/2}\right) \log y.$$

To show this, we may restrict ourselves in the case $\nu = 1$, $\omega(d) = 2r$, for others can be treated quite similarly. Thus let $\rho_1(d) = 1$, $d = p_1 p_2 \cdots p_{2r}$, $y^{1/2} \geq z > p_1 > \cdots > p_{2r-1} > p_{2r}$. By (2.1.12), we have, for $0 \leq j \leq r-1$, **54**

$$p_{2j+2} < p_{2j+1} < \left(\frac{y}{p_1 p_2 \cdots p_{2j}}\right)^{\frac{1}{\beta+1}}.$$

This implies

$$\log\left(\frac{y}{p_1 p_2 \ldots p_{2j+2}}\right) > \left(1 - \frac{2}{\beta+1}\right)\log\left(\frac{y}{p_1 p_2 \ldots p_{2j}}\right),$$

whence inductively we get

$$\log\frac{y}{d} > \left(\frac{\beta-1}{\beta+1}\right)^r \log y$$

which gives the assertion of the lemma for our present case.

We can now prove the very important

Theorem 7 (THE FUNDAMENTAL LEMMA)**.** *Let δ be such that, uniformly for any $2 \leq u \leq v$,*

$$\prod_{u \leq p < v}\left(1 - \frac{\delta(p)}{p}\right)^{-1} \leq C\left(\frac{\log v}{\log u}\right)^k \tag{2.2.3}$$

with certain positive constants C and k. Also, let $z = y^{1/s}$, $s \geq 2$. Then there are two sequences $\left\{\xi_d^{(\nu)}\right\}$ $(\nu = 0, 1)$ depending only on y and k such that

(i) $$\xi_1^{(\nu)} = 1; |\xi_d^{(\nu)}| \leq 1; \xi_d^{(\nu)} = 0 \text{ for } d \geq y,$$

55 *and uniformly for q, $(q, P(z)) = 1$,*

$$(-1)^\nu\left\{S(A_q, z) - XV(z)\frac{\delta(q)}{q}\left(1 + O\left(\exp\left(-\frac{s}{2}\log s\right)\right)\right)\right\}$$

(ii) $$\geq (-1)^{\nu} \sum_{\substack{d|P(z)\\ d<y}} \xi_d^{(\nu)} R_{dq},$$

where the constant involved in the O-symbol depends in C and k of (2.2.3) *at most.*

To prove this, we put

$$\xi_d^{(\nu)} = \mu(d)\rho_\nu(d; y, \beta)$$

with a sufficiently large β. Then (i) is immediate. As for (ii) we apply LEMMA 8 to A_q, and modify the dentity by (2.2.1), getting

$$(-1)^{\nu} \left\{ S(A_q, z) - \frac{\delta(q)}{q} X U_\nu(y, z) \right\} \geq (-1)^{\nu} \sum_{\substack{d|P(z)\\ d<y}} \xi_d^{(\nu)} R_{dq},$$

where

$$U_\nu(y, z) = \sum_{d|P(z)} \frac{\delta(d)}{d} \mu(d)\rho_\nu(d).$$

Then (2.2.2) gives

$$U_\nu(y, z) = V(z) + (-1)^{\nu-1} \sum_{r=1}^{\infty} \sum_{\substack{d|P(z)\\ \omega(d)=2r+\nu}} \frac{\delta(d)}{d} \sigma_\nu(d) V(p(d))$$

(2.2.4) $$= V(z) + (-1)^{\nu-1} \tilde{U}_\nu(y, z),$$

say.

We note that if $\sigma_\nu(d) = 1$ then $\rho_\nu(d/p(d)) = 1$, and thus by LEMMA 9 for we have, $\omega(d) = 2r + \nu$,

$$\log \frac{d}{p(d)} < \left(1 - c\left(\frac{\beta-1}{\beta+1}\right)^r\right) \log y$$

or $$\log p(d) > \frac{c}{\beta}\left(\frac{\beta-1}{\beta+1}\right)^r \log y,$$

since $\sigma_\nu(d) = 1$ implies $p(d)^\beta d \geq y$. On the other hand, the last in- **56** equality implies also $z^{\omega(d)+\beta} \geq y$, so $\beta + \omega(d) \geq s$, for, we have $z = y^{1/s}$. Thus, on the right side of (2.2.4), we have $r > (s - \beta - 1)/2$.

Collecting these observations, we have

$$\widetilde{U}_\nu(y,z) \ll \sum_{r>1/2(S-\beta-1)} \frac{1}{(2r+\nu)!} V\left(y^{\frac{c}{\beta}\left(\frac{\beta-1}{\beta+1}\right)^r}\right) \left\{ \sum_{y^{\frac{c}{\beta}\left(\frac{\beta-1}{\beta+1}\right)^r} \leq p < z} \frac{\delta(p)}{p} \right\}^{2r+\nu}$$

Then, noticing that (2.2.3) gives

$$\sum_{y^{\frac{c}{\beta}(\frac{\beta-1}{\beta+1})^r} \leq p < z} \frac{\delta(p)}{p} \leq kr \log\left(\frac{\beta+1}{\beta-1}\right) + k \log\left(\frac{\beta}{s} c\right),$$

we have

$$\tilde{U}_\nu(y,z) \ll V(z) \left(\frac{\beta}{s}\right)^k \sum_{r>1/2(s-\beta-1)} \frac{1}{(2r+\nu)!} \left\{ k(\frac{\beta+1}{\beta-1})^{\frac{k}{2}} \left(r \log\left(\frac{\beta+1}{\beta-1}\right) + \log\left(\frac{\beta}{s} c\right)\right)\right\}^{2r+\nu} \tag{2.2.5}$$

We now suppose that s is large, and we put $\beta = s/3$. Then we have

$$\begin{aligned} \tilde{U}_\nu(y,z) &\ll V(z) \sum_{r>s/3} \left(\frac{ck}{s}\right)^{2r} \\ &\ll V(z) \exp\left(-\frac{s}{2} \log s\right). \end{aligned}$$

57 If s is not large enough then we take β so large that the right side of (2.2.5) converges. This ends the proof of the theorem.

Now, let us digress briefly from rigorous discussion and explain the motivation of Rosser's device introduced in the proceding section; this may help one to see that Rosser's seemingly complicated identity is a sort of logical conclusion when we try to seek for optimal sieve procedures.

One may have the impression that the introduction of the parameter y and β is abrupt and arbitrary though the idea to eliminate $S(A_d.p(d))$

with $p(d)$ 'small' for A_d from the identity (2.1.11) is quite natural. But this is actually related closely to the concept of the sieving limit, which may be roughly formulated as follows.

In many practical problems, the information on the size of $|A_d|$ is given in the form

$$\sum_{\substack{d<y\\ d|P(z)}} |R_d| = o(XV(z)) \tag{2.2.6}$$

uniformly for $z \leq y$, and δ is almost constant at primes, but, for the sake of simplicity, we assume here that

$$\prod_{u\leq p<v}\left(1-\frac{\delta(p)}{p}\right)^{-1} \leq \left(\frac{\log v}{\log u}\right)^{k} (2 \leq u < v) \tag{2.2.7}$$

where k is a positive constant.

Returning to the identity (2.1.11), because of (2.2.6) we may restrict **58**
λ by the condition

$$\lambda(d) = 0 \text{ for } d \geq y, \tag{2.2.8}$$

in addition to (2.1.10), without loss of much generality; then we have, by (2.2.1),

$$S(A,z) \geq XV(z)\{U_\delta(y,z;\tilde{\rho}_0) - o(1)\}$$

where

$$V(z)U_\delta(y,z;\tilde{\rho}_0) = \sum_{\substack{d|P(z)\\ d<y}} \mu(d)\frac{\delta(d)}{d}\tilde{\rho}_0(d).$$

But, since $S(A,z)$ is non-nagetive, we have more precisely

$$S(A,z) \geq XV(z)\{T_\delta(y,z;\tilde{\rho}_0) - o(1)\},$$

where

$$T_\delta(y,z;\tilde{\rho}_0) = \max(0, U_\delta(y,z;\tilde{\rho}_0));$$

the identity

$$V(z) = \sum_{d|P(z)} \mu(d)\tilde{\rho}_0(d)\frac{\delta(d)}{d} + \sum_{d|P(z)} \tilde{\rho}_0(d)\frac{\delta(d)}{d}V(p(d))$$

implies

$$1 \geq T_\delta(y, z; \tilde{\rho}_0) \geq 0.$$

Next we put

$$\varphi_k(y, z) = \inf_\delta \sup_\lambda T_\delta(y, z; \tilde{\rho}_0),$$

where δ satisfies (2.2.7), and λ (2.1.8), (2.1.10) and (2.2.8).

59 Obviously we have

$$S(A, z) \geq XV(z)\{\varphi_K(y, z) - 0(1)\}.$$

Our interest lies, naturally, in such a choice of y and z that $\varphi_K(y, z) > 0$; so we consider the quantity

$$\alpha_K(y) = \inf\{s; \varphi_K(y, y^{1/s}) > 0\}.$$

And we point out the important fact that $\alpha_K(y)$ remains bounded as y tends to infinity. This can be seen easily from THEOREM 7[1]. Thus we may consider, further, the quantity

$$\beta(k) = \limsup_{y \to \infty} \alpha_k(y).$$

If $s > \beta(k)$, then we have the possibility of

$$S(A, y^{1/s}) > 0$$

for a sufficiently large y, but, otherwise, we can say nothing definite about the lower bound for $S(A, y^{1/s})$ than that it is non - negative. This is the reason that $\beta(k)$ is called the *sieving limit.*

Now, if we want to keep at minimum the loss caused by discarding certain terms $S(A_d, p(d))$ on the right side of (2.1.11) we should, of course, put $\tilde{\sigma}_\nu(d) = 0$ for all d such that there is the possibility of the existence of atleast one A with

60 $S(A_d, p(d)) > 0$. But, if A is such that all $A_d(d|p(z), d < y)$ satisfy

[1]One may say that this is a tautology, for our proof of THOEOREM 7 depends on Rosser's sieve idea. To avoid such a confusion, we remark that we could have proved THEOREM 7 by Brun's (cf. [21, Chap. 2])

the analogue of (2.2.6), i.e.,

$$\sum_{\substack{\ell<y/d \\ \ell|p(p(d))}} |R_{d\ell}| = 0(\frac{\delta(d)}{d}XV(p(d))),$$

then we have the possiblity of $S(A_d, p(d)) > 0$ for $p(d) < (y/d)^{\frac{1}{\beta(k)}}$ provided y/d is sufficiently large. And this observation leads us immediately to Rosser's device.

We should deep it in our mind, however, that although Rosser's weights ρ_ν may simulate well the extremal (or optimal) sieving procedure, there is no reason to believe that they give actually the optimal estimate of $S(A, z)$ generally. In fact, it is known that, for the sieve problem with $k > 1$, the Rosser weights do not yield optimal results. But, very fortunately, for the linear sieve problems (i.e., $k = 1$) which contain most of important classical problem Rosser's method can indeed produce optimal results as we shall show in detail in the next chapter.

2.3 A Smoothed Version of Rosser's Identity

Returing to the main theme of this chapter, we shall give an important modification of the fundametal identity (2.1.7): we shall inject a smoothing device into it. This will play a vital role in the investigation of the error term in the linear sieve which will be developed in §3.4.

To this effect, we take up an interval $[z_1, z), 2 \le z_1 < z$, and dissect it **61**
into smaller ones which we shall denote generally by I with or without suffix; so, we have

$$[z_1, z) = \bigcup I(\text{ disjoint }).$$

Next, let K with or without suffix stand for the set - theoretic direct product of a sequence of Is', and $\omega(K)$ for the number of constituent I's. If $K = I_1 I_2 \cdots I_r$ then $I < K$ means that $(I) < \min(I_j)$ where (I) is the right end point of I; also, $d \in K$ implies that $d = p_1 p_2 \cdots p_r$ with $P_j \in I_j$. Here we have to introduce the convention that $1 \in K$ for empty K. Note that we do not reject non- squarefree d; this convention will have effect in the formula of LEMMA 10 below.

Theorem 8. *Let λ be an arbitrary function defined on the set of all K and satisfying $\lambda(K) = 1$ for empty K.*

Put $\phi(K) = \lambda(I_1)\lambda(I_1I_2)\cdots\lambda(I_1I_2\cdots I_r)$; $\phi(K) = 1$ if K is empty, and $\psi(K) = \phi(I_1I_2\cdots I_{r-1}) - \phi(I_1I_2\cdots I_r)$; $\psi(K) = 0$ if K is empty, where

$$K = I_1I_2\cdots I_r, I_1 > I_2 > \cdots > I_r.$$

62 Then we have

$$\begin{aligned} S(A,z) = {} & \sum_K (-1)^{\omega(K)}\phi(K)\sum_{d\in K} S(A_d, z_1) \\ & + \sum_{I<K}(-1)^{\omega(K)}\phi(KI)\sum_{\substack{p'<p\\ p',p\in I\\ d\in K}} S(A_{dpp'}, p') \\ & + \sum_K (-1)^{\omega(K)}\psi(K)\sum_{d\in K} S(A_d, p(d)). \end{aligned}$$

To prove this, we first modify modify the Bushstab identity trivially as

$$\begin{aligned} S(A,z) = S(A,z_1) & - \sum_I \lambda(I)\sum_{p\in I} S(A_p, p) \\ & - \sum_I (I - \lambda(I))\sum_{p\in I} S(A_p, p). \end{aligned} \tag{2.3.1}$$

Also, for each $p \in I$, we have

$$\begin{aligned} S(A_p, p) & = S(A_p, z_1) - \sum_{\substack{p'<p\\ p'\in I}} S(A_{pp}, p') - \sum_{I'<I}\sum_{p'\in I'}{}' S(A_{pp}, p') \\ & = S(A_p, z_1) - \sum_{\substack{p'<p\\ p'\in I}} S(A_{pp}, p') - \sum_{I'<I}\lambda(II')\sum_{p'\in I'}{}' S(A_{pp}, p') \\ & - \sum_{I'<I}(1 - \lambda(II'))\sum_{p'\in I'} S(A_{pp}, p'). \end{aligned}$$

Inserting this into the first double sum on the right side of (2.3.1), we get

$$S(A,z) = S(A,z_1) - \sum_I \lambda(I)\sum_{p\in I} S(A,z_1) + \sum_I \lambda(I)\sum_{\substack{p'<p\\ p',p\in I}} S(A_{pp'}, p')$$

$$- \sum_{I}(1-\lambda(I))\sum_{p\in I} S(A_p, p) + \sum_{I_2<I_1} \lambda(I_1)(1-\lambda(I_1 I_2))$$
$$\sum_{\substack{p_1\in I_1\\ p_2\in I_2}} S(A_{P_1 p_2}, p_2) + \sum_{I_2<I_1} \lambda(I_1)\lambda(I_1 I_2) \sum_{\substack{p_1\in I_1\\ p_2\in I_2}} S(A_{p_1 p_2}, p_2).$$

63

This is obviously the case $r = 2$ of the identity

$$\begin{aligned} S(A,z) = &\sum_{\omega(K)<r} (-1)^{\omega(K)}\phi(K)\sum_{d\in K} S(A_d, z_1)\\ &+ \sum_{\substack{I<K\\ \omega(K)<r-1}} (-1)^{\omega(K)}\phi(KI) \sum_{\substack{p'<p\\ p'p\in I\\ d\in K_2}} S(A_{dpp'}, p')\\ &+ \sum_{\omega(K)\leq r} (-1)^{\omega(K)}\psi(K)\sum_{d\in K} S(A_d, p(d))\\ &+ (-1)^r \sum_{\omega(K)=r} \phi(K)\sum_{d\in K} S(A_d, p(d)). \end{aligned} \tag{2.3.2}$$

We may establish this by induction on r; we need only to insert in the last double sum the expression

$$\begin{aligned} S(A_d, p(d)) = S(A_d, z_1) &- \sum_{\substack{p<p(d)\\ pd\in K}} S(A_{dp}, p)\\ &- \sum_{I<K}\lambda(KI)\sum_{p\in I} S(A_{dp}, p) - \sum_{I<K}(1-\lambda(KI))\sum_{p\in I} S(A_{dp}, p). \end{aligned}$$

Having obtained (2.3.2), we take r sufficiently large and conclude the proof of the theorem. 64

Now, let us introduce two parameters y and $\beta \geq 1$, and imitate Rosser's device. We set $\lambda = \lambda_\nu$ the characteristic function of the set

$$\left\{ K = I_1 I_2 \cdots I_r; I_1 > I_2 > \cdots > I_r, \quad \begin{array}{c} r \equiv \nu + 1 \pmod 2\\ \text{or}\\ r \equiv \nu \pmod 2 \text{ and } (I_r)^{\beta+1}(I_{r-1})\cdots(I_1) < y \end{array} \right\} \tag{2.3.3}$$

And let Θ_ν and Δ_ν stand for ϕ and ψ with this choice of λ, respectively. Then Θ_ν and Δ_ν are the characteristic functions of the sets

(2.3.4)
$$\left\{K = I_1 I_2 \cdots I_r; I_1 > I_2 > \cdots > I_r, \begin{array}{r} (I_{2k+\nu})^{\beta+1}(I_{2k+\nu-1})\cdots(I_1) < y \\ \text{for all k with } \leq 2k+\nu \leq r \end{array}\right\}$$

and
$$\left\{K = I_1 I_2 \cdots I_r; \begin{array}{r} I_1 > I_2 > \cdots > I_r, r \equiv \nu \pmod 2 \\ \Theta_\nu(I_1 I_2 \cdots I_{r-1}) = 1 \\ (I_r)^{\beta+1}(I_{r-1}\cdots(I_1) \geq y) \end{array}\right\},$$

respectively. Then THEOREM 8 gives the following smoothed version of LEMMA 8.

Lemma 9.
$$\begin{aligned} S(A,z) &= \sum_K (-1)^{\omega(K)}\Theta_\nu(K)\sum_{d\in K} S(A_d, z_1) \\ &+ \sum_{I<K}(-1)^{\omega(K)}\Theta_\nu(KI)\sum_{\substack{p'<p \\ p',p\in I \\ d\in K}} S(A_d, z_1) \\ &+ (-1)^\nu \sum_K \Delta_\nu(K)\sum_{d\in K} S(A_{dpp'}, p'). \end{aligned}$$

65 Further, in this we replace $S(A_{dpp,}p')$ by larger $S(A_{dpp'}, z_1)$ and discard the condition $p' < p$, getting

Lemma 10.
$$\begin{aligned} (-1)^\nu S(A,z) \geq (-1)^\nu &\sum_K (-1)^{\omega(K)}\Theta_\nu(K)\sum_{d\in K} S(A_d, z_1) \\ &- \sum_{\substack{I<K \\ \omega(K)\equiv\nu+1 \pmod 2}} \Theta_\nu(KI)\sum_{\substack{p,p'\in I \\ d\in K}} S(A_{dpp'}, z_1). \end{aligned}$$

NOTES (II)

In our definition of a sieve problem, we have introduced the weight θ, but this has nothing to do with later development of our discussion.

However, THEOREM 6 will probably serve for the future developments of the theory of 'weighted' combinatorial sieve methods which has been initiated by Greaves; in fact THEOREM 6 is generalized version of his identity [[20], (2.8)].

THEOREM 6 in its conventional form can be found in Halberstam and Richert [[21], p. 39], which seems to originate in Levis's work [44] on Brun's sieve. On the other hand, the identity (2.1.6) with the simplest choice of θ occors in Iwaniec [30]. 66

We may call THEOTEM 6 *the fundamental theorem in sieve methods*, for various specializations of ρ give all sieve method known at present, except for the local sieve of Selberg. Especially, with the aid of THEOREM 6, we can reveal the mechanism lying behind the $\wedge^2$-sieve([1]).

We set

$$\rho(d) = \mu(d) \sum_{[d_1,d_2]=d} \lambda_{d_1}\lambda_{d_2}(\lambda_1 = 1).$$

Then, after some rearrangement, we get

$$\begin{aligned} S(A,z) &= \sum_{a\in A}\Big(\sum_{\substack{d|P(z)\\ a\in\Omega(d)}} \lambda_d\Big)^2 \\ &- \sum_{p<z} \sum_{\substack{a\in A_p\\ a\notin\cup\Omega(q)\\ q<p.}} \Big(\sum_{\substack{a\in\Omega(h)\\ h|p(p^+,z)}} (\lambda_h + \lambda_{hp})\Big)^2 \end{aligned}$$

where p, q are primes, p^+ is the prime which succeeds p, and $P(p^+, z) = P(z)/P(p^+)$. This remarkable identity is due to Halberstam. [1]

We should remark also that, via a special case of THEOREM 6, Fouvry and Iwaniec [13] obtained a stricking result pertaining to Bombieri's mean prime number theorem. 67

In the proof of THEOREM 7, we followed the argument of Friedlander and Iwaniec [14] which is quick and elegant compared with the one via Brun's sieve; here, already we can have the glimpse of the power of Rosser's idea.

[1]By the courtesy of Professor Halberstam
[1]By the courtesy of Professor Halberstam

In explaining Rosser's idea, we had to appeal to a rough image of the concept of the sieving - limit. We stress that our definition of the seieving limit applies to some restricted class of sieve procedures only; for a more general treatment of the matter, see Selberg [75].

The idea of introducing a smoothing device into Rosser's sieve method is an outstanding contribution of Iwaniec [31] to the theory of sieve methods. This will result in a highly flexible error - term in the linear sieve, as we shall see in the next chapter.

We are not in a position to speculate how Iwaniec was led to his novel idea; readers are referred to his own account [33].

The argument developed in §2.3 is due to Motoghashi [[60], II], which is a refinement of Iwaniec's

Chapter 3

The Linear Sieve

THE OBJECT OF this chapter is to develop a detailed account of the **68**
fundamental result of Rosser and Iwaniec on the linear sieve. Rosser's theory determines the optimal mainterm in the upper and lower bounds for linear sieve problems, and Iwaniec's theory enhances its power greatly by introducing into it a highly flexible error - term.

We shall first study the nature of expected optimal upper and lower bounds for linear sieve problems by employing Rosser's sieving procedure described in the preceding chapter. This will lead us to a difference - differential equation, and solving it, we shall find the most suitable choice of the parameter β, which will, in turn, be fed back to a rigorous argument to prove Rosser's linear sieve. And an example due to Selberg will be used to confirm that Rosser's result is indeed optimal. Then we will focus our attention on the error - term in Rosser's linear sieve; we shall inject into our discussion the smoothing device introduced in the last section of the preceding chapter, and obtain Iwaniec's bilinear form for the error - term.

Throughout this chapter, we shall retain the notation and convention introduced in the preceding chapter.

3.1 A Difference-Differential Equation

First of all, we have to give a precise notion of the dimension of a sieve problem. 69

We require that δ which is introduced at (2.2.1) be not wild locally, and constant on average. Namely, we assume that there is a constant $A_1 > 0$ such that for all prime p

$$0 \leq \frac{\delta(p)}{P} \leq 1 - \frac{1}{A_1}, \tag{3.1.1}$$

and that there are constants $k, A_2 > 0$ and a positive parameter L which is not too large such that for any $2 \leq u < v$

$$-L \leq \sum_{u \leq p < v} \frac{\delta(p)}{P} \log\ p - k\ \log \frac{v}{u} \leq A_2.$$

Then $S(A, z)$ is called a k - *dimensional sieve problem*. And, in the present chapter, we are concerned with the case $k = 1$ exclusively. Thus we assume throughout the sequel the conditions (3.1.1) and

$$-L \leq \sum_{u \leq p < v} \frac{\delta(p)}{P} \log p - \log \frac{v}{u} \leq A_2. \tag{3.1.2}$$

for any $2 \leq u < v$.

It is known that (3.1.1) and (3.1.2) imply

$$\prod_{u \leq p < v} \left(1 - \frac{\delta(p)}{P}\right)^{-1} \leq \frac{\log v}{\log u} \left(1 + o\left(\frac{1}{\log u}\right)\right) \tag{3.1.3}$$

70 as well as

$$\prod_{u \leq p < v} \left(1 - \frac{\delta(p)}{p}\right)^{-1} = \frac{\log v}{\log u} \left(1 + 0\left(\frac{1}{\log u}\right)\right) \tag{3.1.4}$$

for any $2 \leq u < v$, where the implied constants depend on A_1 and A_2 at most. In particular, (3.1.3) allows us to use THEOREM 7. Also, in our argument, we shall make multiple use of the basic

Lemma 11. *We assume* (3.1.1) *and* (3.1.2). *Let* $\psi(t)$ *be a non-nequative, monotone and continuous fuction for* $t \geq \alpha > 0$. *Then we have, for any* $2 \leq u < v \leq x^{1/1+\alpha}$,

$$\sum_{u \leq p < v} \frac{\delta(p)}{P} V(p) \psi\left(\frac{\log \frac{x}{p}}{\log p}\right) = V(v) \frac{\log v}{\log x} \int\limits_{\frac{\log x}{\log v}}^{\frac{\log x}{\log u}} \psi(t-1) dt + 0(LMV(v) \frac{\log v}{\log^2 u})$$

where

$$M = \max_{u \leq \xi \leq v} \psi\left(\frac{\log \frac{x}{\xi}}{\log \xi}\right).$$

After these initial remarks, we now start the investigation leading to the determination of the optimal β in the Rosser weights ρ_ν under the basic assumptions (3.1.1) and (3.1.2).

To simplify the convergence problem which we shall encounter later, we introduce here another parameter z_1 such that

$$z_1 \leq \exp\left(\frac{\log y}{(\log \log y)^2}\right), \tag{3.1.5}$$

where y is the parameter which occurs in the definition of ρ_ν. Further, we put 71

$$z = y^{1/s},$$

and assume that

$$0 < s < \frac{\log y}{\log z_1} \tag{3.1.6}$$

so that $z_1 < z$. Then we apply Rosser's identity (LEMMA 8) to the sequence

$$\{a \in A; a \notin \Omega(p) \text{ for all } p < z_1\},$$

getting

$$S(A, z) = \sum_{d \mid P(z_1, z)} \mu(d)_{\rho_\nu}(d) S(A_d, z_1) + (-1)^\nu \sum_{d \mid P(z_1, z)} \sigma_\nu(d) S(A_d, \rho(d)),$$

and thus

$$(3.1.7) \qquad (-1)^{\nu}\left\{S(A,z) - \sum_{d|P(z_1,z)} \mu(d)_{\rho_\nu}(d) S(A_d, z_1)\right\} \geq 0,$$

where $P(z_1,z) = P(z)/P(z_1)$. Similarly, we have

(3.1.8)

$$V(z) = V(z_1) \sum_{d|P(z_1,z)} \mu(d)_{\rho_\nu} \frac{\delta(d)}{d} + (-1)^{\nu} \sum_{d|P(z_1,z)} \sigma_\nu(d) \frac{\delta(d)}{d} V(p(d)).$$

To each $S(A_d, z_1)$ of (3.1.7) we apply THEOREM 7, and get

$$(-1)^{\nu-1}\left\{S(A_d,z_1) - XV(z_1)\frac{\delta(d)}{d}\left(1 + O\left(\exp\left(-\frac{h}{2}\log h\right)\right)\right)\right\} \leq \sum_{\substack{r|P(z_1)\\ r<z_1^h}} |R_{dr}|,$$

72 where h is at our disposal. Insertion of this into (3.1.7) gives

$$(3.1.9) \qquad \begin{aligned} &(-1)^{\nu-1}\{S(A,z) - XV(z)K_\nu(y,z;\delta)\} \\ &\leq \sum_{\substack{r|P(z_1)\\ r<z_1^h}} |R_d| + o\left\{\exp\left(-\frac{h}{2}\log h\right) XV(z_1) \sum_{d|P(z_1,z)} \frac{\delta(d)\rho_\nu(d)}{d}\right\}, \end{aligned}$$

where

$$(3.1.10) \qquad V(z)K_\nu(y,z;\delta) = V(z_1) \sum_{d|P(z_1,z)} \mu(d)_{\rho_\nu}(d) \frac{\delta(d)}{d}.$$

Since

$$\sum_{d|P(z_1,z)} \frac{\delta(d)\rho_\nu(d)}{d} \leq V(z_1)/V(z)$$

and also we have (3.1.3), the O-term of (3.1.9) is

$$O\left\{\exp\left(-\frac{h}{2}\log h\right) XV(z)(\log z)^2\right\}.$$

We now set

$$h = \log\log y,$$

and collecting above estimates, we get

$$(3.1.11) \quad (-1)^{\nu-1}\{S(A,z) - XV(z)K_\nu(y,z;\delta)\} \\ \leq \sum_{\substack{d|P(z)\\ d<y_0}} |R_d| + XV(z)(\log y)^{-10_s-2}$$

where $y_0 = y\exp(\log z_1 \log\log y)$. 73

Next, we put

$$H_\nu(y,z;\delta) = \max(0, K_\nu(y,z;\delta));$$

obiviously, we have

$$H_1(y,z;\delta) = K_1(y,z;\delta).$$

We should note also that (3.1.8) implies

$$(3.1.12) \qquad H_1(y,z;\delta) \geq 1 \geq H_0(y,z;\delta) \geq 0.$$

Hence our problem is now transformed into the asymtotic evaluation of $H(y,y^{1/s};\delta)$ in terms of s, i.e., we will seek for the continuous function $\phi_\nu(s)$ such that

$$(3.1.13) \qquad \lim_{y\to\infty} H_\nu(y,y^{1/s};\delta) = \phi_\nu(s),$$

if it ever exists. Note that we are going to find a $\phi_\nu(s)$ not depending on δ apart the basic conditions (3.1.1) and (3.1.2).

If (3.1.13) holds, and if we assume (2.2.6) with $y = y_0$, then we would have

$$(3.1.14) \quad XV(y^{1/s})(\phi_0(s) - o(1)) \leq S(A,y^{1/s}) \leq XV(y^{1/s})(\phi_1(s) + o(1)).$$

The direct proof of (3.1.13) seems to be quite difficult if not impossible. Thus we make a round-about, by assuming first the existence of the limit $\phi_\nu(s)$; on this assumption, we investigate its nature, and then feed the obtained information back to the actual proof of the asymptotic formula for $H_\nu(y,y^{1/s};\delta)$. The optimal choice of β will emerge out of this process.

Thus let us assume more preciesely that (3.1.13) holds uniformly for all bounded s. Then 74

(3.1.15) $$\phi_\nu \text{ is monotone },$$

since it is clear from (3.1.8) that $K_\nu(y, y^{1/s}; \delta)$ is monotone with respect to s for each fixed y. Also, because of (3.1.12), we have

(3.1.16) $$\phi_1(s) \geq 1 \geq \phi_0(s) \geq 0.$$

Now, since we have (3.1.14), the obervation made in §2.2 on the sieving limit suggests that if we want to let Rosser's sieving procedure simulate well the optimal one which is supported to exist, we should confine ourselves to the most critical case

(3.1.17) $$\beta = \inf\{s; \phi_0(s) > 0\}.$$

This shall we assume henceforth, and will turn out to be decisive. (3.1.10) gives

(3.1.18)
$$V(y^{1/s})K_1(y, y^{1/s}; \delta) = V(z_1) - \sum_{z_1 \leq p < \min(y^{1/s}, y^{\frac{1}{\beta+1}})} \frac{\delta(p)}{p} V(p) K_0(\frac{y}{p}, p; \delta),$$

since $\rho_1(p) = 1$ implies $p < y^{\frac{1}{\beta + 1}}$. So we have, for $s \leq \beta + 1$,

$$V(y^{1/s})K_1(y, y^{1/s}; \delta) = V(y^{\frac{1}{\beta+1}})K_1(y, y^{\frac{1}{\beta+1}}; \delta),$$

75 that is, for $s \leq \beta + 1$,

$$\begin{aligned} s\phi_1(s) &= (\beta + 1)_{\phi_1}(\beta + 1) \\ &= D, \end{aligned} \qquad (3.1.19)$$

say. On the other hand, if $\beta + 1 + \varepsilon \leq s$, then we have $(\log y/p)/\log p \geq \beta + \varepsilon$ in (3.1.18), and by the assumption (3.1.17), we have

$$K_0\left(\frac{y}{p}, p; \delta\right) = H_0\left(\frac{y}{p}, p; \delta\right)$$

for sufficeintly large y. Thus, for $v > u \geq \beta + 1 + \varepsilon$ we have, by (3.1.18),

$$V(y^{1/v})H_1(y, y^{1/v}; \delta) - V(y^{1/u})H_1(y, y^{1/u}; \delta) = \sum_{y^{\frac{1}{v}} \leq p < y^{\frac{1}{u}}} \frac{\delta(p)}{p} V(p) H_0\left(\frac{y}{p}, p; \delta\right).$$

But, by our present assumption, the last sum is equal to

$$(1 + o(1))_{y^{\frac{v}{1}}} \sum_{\leq p < y^1 u} \frac{\delta(p)}{p} V(p) \phi_0 \left(\frac{\log \frac{y}{p}}{\log p}\right),$$

which, since ϕ_0 is monotone and bounded (cf. (3.1.15) and (3.1.16)), can be expressed, with the aid of LEMMA 11, as

$$(1 + 0(1)) \frac{1}{u} V(y^{\frac{1}{u}}) \int_u^v \phi_0(t - 1) dt,$$

provided

(3.1.20) $$L = 0(\log y)$$

76

We shall assume, in the sequel, this harmless condition on L. Thus we get

$$u \frac{V(y^{1/v})}{V(y^{1/u})} H_1(y, y^{\frac{1}{v}}; \delta) - u H_1(y, y^{\frac{1}{v}}; \delta) = (1 + o(1)) \int_u^v \phi_0(t - 1) dt,$$

and by (3.1.4)

$$v\phi_1(v) - u\phi_1(u) = \int_u^v \phi_0(t - 1) dt,$$

for $\beta + 1 + \varepsilon \leq a < v$ with any fixed $\varepsilon > 0$. But, because of the continuity, we see that this holds for $\beta + 1 + \leq < v$; namely, we have

(3.1.21) $$(s\phi_1(s))' = \phi_0(s - 1) \text{ for } \beta + 1 \leq s.$$

Simirlarly, we have, for $v \geq u \geq \beta + \varepsilon$,

$$V(y^{\frac{1}{v}})H_0(y, y^{\frac{1}{v}}) - V(y^{\frac{1}{V}})H_0(y, y^{\frac{1}{v}}; \delta) = \sum_{y^{\frac{1}{v}} \leq p < y^{\frac{1}{u}}} \frac{\delta(p)}{p} V(p) H_1\left(\frac{y}{P}, P; \delta\right),$$

provided y is sufficiently large. Thus, as much the same way as above, we have

$$v\phi_0(v) - u\phi_0(u) = \int_u^v \phi_1(t-1)dt$$

for $\max(1, \beta) < u < v$; here the condition $1 < u$ is needed bacause of
77 (3.1.19). And, we have $\beta \leq 1$, the last equation contradicts the boundedness of ϕ_0. Hence we may assume hereafter that

(3.1.22) $$\beta > 1.$$

Then we have

(3.1.23) $$(s\phi_0(s))' = \phi_1(s-1) \text{ for } \beta \leq s,$$

which is of course supplemented by

(3.1.24) $$\phi_0(s) = 0 \text{ for } s \leq \beta.$$

Collecting (3.1.17), (3.1.19), (3.1.21)-(3.1.24), we are now led to the investigation of the difference-differential equation [1]

(3.1.25) $$(s\phi_\nu(s))' = \phi_{\nu+1}(s-1) \text{ for } \beta \leq s$$

on the boundary condition

(3.1.26) $$s\phi_1(s) = D, \phi_0(s) = 0 \text{ for } s \leq \beta,$$

(3.1.27) $$\phi_0(s) \leq 1 \leq \phi_1(s) \text{ for all } s > 0,$$

where $\beta > 1$ and $D > 0$ are to be determined so that

(3.1.28) $$\phi_0(s) > 0 \text{ for } s > \beta$$

and the asymptotic formula (3.1.13) holds on the condition (3.1.20).

---

[1]In the sequel, we shall use the convention: $\phi_j \equiv \phi_\nu$ if $j \equiv \nu \pmod 2$.

3.2 The Optimal Value of β

In this section, we shall now show a detailed solution of the last problem. This requires a little lengthy discussion, and we start with the following 78
two important observations on the nature of the expected solution ϕ_v.

Lemma 12. *If ϕ_v satisfies* (3.1.25)-(3.1.27)*, then ϕ_1 and ϕ_0 are strictly decreasing and increasing, respectively. In particular,* (3.1.28) *is redundant.*

To prove this, let u_o be the least root of $\phi_1'(u) = 0$, if exists. By (3.1.25) and (3.1.26) $u_0 > \beta + 1$. But we have, by (3.1.25) and (3.1.27),

$$\begin{aligned}
0 &= u_0\phi_1'(u_0) = \phi_0(u_0 - 1) - \phi_1(u_0) \leq \phi_0(u_0 - 1) - \phi_0(u_0) \\
&= \phi_0'(u') = \frac{1}{u'}(\phi_0(u')) - \phi_1(u' - 1)) \leq \frac{1}{u'}(\phi_1(u') - \phi_1(u' - 1)) \\
&= \frac{1}{u'}\phi_1'(u''),
\end{aligned}$$

where $u_0 - 1 < u' < u_0, u' - 1 < u'' < u'$. However, we have $\phi_1'(u'') < 0$ because of the definition of u_0. Hence $\phi_1(u)$ is strictly decreasing. And so we have, for $u \geq \beta$,

$$u\phi_0'(u) = \phi_1(u - 1) - \phi_0(u) \geq \phi_1(u - 1) - \phi_1(u) > 0$$

whence $\phi_0(u)$ is strictly incresing for $u \geq \beta$.

Lemma 13. *We assume* (3.1.1) *and* (3.1.2)*. Let ϕ_v be a solution of* (3.1.25)-(3.1.27)*. Then we have, for $2 \leq u \leq v \leq y^{1/\beta}$,*

$$\begin{aligned}
V(v)\phi_v\left(\frac{\log y}{\log v}\right) = V(u) \sum_{d|P(u,v)} \mu(d)\rho_v(d)\frac{\delta(d)}{d}\phi_{v+\omega(d)}\left(\frac{\log \frac{y}{d}}{\log u}\right) \\
+ 0\left(LV(v)\frac{\log^2 v}{\log^3 u}\right),
\end{aligned}$$

where $\rho_v(d) = \rho_v(d; y, \beta)$. 79

To show this, we note firstly that the previous lemma allows us to appeal to LEMMA 11, and we have

$$\sum_{u\le p<v}\frac{\delta(p)}{p}\phi_{\nu+1}\left(\frac{\log\frac{y}{p}}{\log p}\right)=V(v)\frac{\log v}{\log y}\int\limits_{\frac{\log y}{\log v}}^{\frac{\log y}{\log u}}\phi_{\nu+1}(t-1)dt+0\left(V(v)L\frac{\log v}{\log^2 u}\right),$$

since $\phi_{\nu+1}\left(\frac{\log\frac{y}{\xi}}{\log\xi}\right)$ is bounded if $\xi\le y^{1/\beta}$; here, we should observe also that we have $\beta>1$. But this integral is, by (3.1.4) and (3.1.25),

$$\begin{aligned}&V(v)\frac{\log v}{\log y}\left\{\frac{\log y}{\log u}\phi_\nu\left(\frac{\log y}{\log u}\right)-\frac{\log y}{\log v}\phi_\nu\left(\frac{\log y}{\log v}\right)\right\}\\&=V(u)\phi_\nu\left(\frac{\log y}{\log u}\right)-V(v)\phi_\nu\left(\frac{\log y}{\log v}\right)+o\left(LV(v)\frac{\log v}{\log^2 u}\right).\end{aligned}$$

Thus, noting that (3.1.26) implies $\phi_0\left(\frac{\log\frac{y}{p}}{\log P}\right)=0$ if $p^{\beta+1}\ge y$, we get

$$\begin{aligned}V(v)\phi_\nu\left(\frac{\log y}{\log v}\right)&=V(u)\phi_\nu\left(\frac{\log y}{\log u}\right)\\&\quad-\sum_{u\le p<v}\frac{\delta(p)}{p}\rho_\nu(p)\phi_{\nu+1}\left(\frac{\log\frac{y}{p}}{\log p}\right)+O\left(LV(v)\frac{\log v}{\log^2 u}\right).\end{aligned}$$

80 This is obviously the case $r=1$ of the formula

$$\begin{aligned}V(v)\phi_\nu\left(\frac{\log y}{\log v}\right)&=V(u)\sum_{\substack{d\mid P(u,v)\\ \omega(d)<r}}\mu(d)\rho_\nu(d)\frac{\delta(d)}{d}\phi_{\nu+\omega(d)}\left(\frac{\log\frac{y}{d}}{\log u}\right)\\&\quad+(-1)^r\sum_{\substack{d\mid P(u,v)\\ \omega(d)=r}}\rho_\nu(d)\frac{\delta(d)}{d}V(p(d))\phi_{\nu+r}\left(\frac{\log\frac{y}{d}}{\log p(d)}\right)\\&\quad+O\left(\frac{L}{\log^2 u}\right)\left\{V(v)\log v+\sum_{\substack{\omega(d)<r\\ d\mid P(u,v)}}\frac{\delta(d)}{d}V(p(d))\log p(d)\right\}\end{aligned}\tag{3.2.1}$$

We may establish this by induction on r: If $v+\omega(d) \equiv 0(mod 2)$ then $\rho_v(d) = 1$ implies $p(d)^\beta < y/d$, so

$$\phi_{v+\omega(d)+1}\left(\frac{\log \frac{y}{d\xi}}{\log \xi}\right) = o(1)$$

for $\xi < p(d)$. If $v+\omega(d) \equiv 1 \pmod 2$, then the same holds obviously. Hence we have, as before,

$$V(p(d))\phi_{v+\omega(d)}\left(\frac{\log \frac{y}{d}}{\log p(d)}\right) = V(u)\phi_{v+\omega(d)}\left(\frac{\log \frac{y}{d}}{\log u}\right)$$
$$- \sum_{u \le p < p(d)} \frac{\delta(p)}{p} V(p)\phi_{v+\omega(d)+1}\left(\frac{\log \frac{y}{pd}}{\log p}\right) + O\left(LV(p(d))\frac{\log p(d)}{\log^2 u}\right),$$

where the left side is the one appearing in the second sum of (3.2.1). Inserting this into (3.2.1) and eliminating the terms with those pd such that

$$\phi_{v+\omega(d)+1}\left(\frac{\log \frac{y}{pd}}{\log p}\right) = 0,$$

we readily ontain (3.2.1) for $r+1$ in place of r. To conclude the proof of **81**
the lemma, we need only to take r sufficiently large in (3.2.1) and note that the error-term is, by (3.1.3),

$$0\left\{\frac{LV(v)\log v}{\log^2 u}\sum_{j=0}^{\infty}\frac{1}{j!}\left(\sum_{u \le p < v}\frac{\delta(p)}{p}\right)^j\right\} = O\left(LV(v)\frac{\log^2 v}{\log^3 u}\right).$$

After these preperations, we can now proceed to the determination of β and D, and so ϕ_v.

Put

$$G(u) = \phi_1(u) + \phi_0(u).$$

Then by (3.1.25) we have, for $\beta \le u$,

$$(uG(u))' = G(u-1)$$

which gives

$$|G'(u)| = \frac{1}{u}|G(u-1) - G(u)| \leq \frac{1}{u}\max_{u-1\leq t\leq u}|G'(t)|.$$

Thus we have

$$G'(u) = O\left(r(u+1)^{-1}\right)$$

which implies obviously that there exists a constant A such that

$$G(u) = A + O(r(u)^{-1}) \tag{3.2.2}$$

for $u \geq \beta$.

On the other hand, if we put

$$g(u) = \phi_1(u) - \phi_0(u)$$

82 then $g(u) \geq 0$ by (3.1.27), and we have, for $u \geq \beta$,

$$\begin{aligned}\frac{d}{du}\int_{u-1}^{u}\xi g(\xi)d\xi &= ug(u) - (u-1)g(u-1)\\ &= ug(u) + (u-1)(ug(u))'\\ &= (u(u-1)g(u))',\end{aligned}$$

since $(ug(u))' = -g(u-1)$ is implied by (3.1.25). Hence we have

$$\int_{u-1}^{u}\xi g(\xi)d\xi = u(u-1)g(u) + C \tag{3.2.3}$$

for $u \geq \beta$; setting $u = \beta$ and recalling (3.1.26), we, have

$$C = (2-\beta)D.$$

Then, by the monotonicity of $ug(u)$, one may deduce from (3.2.3) the asymptotic formula

$$g(u) = \frac{(\beta-2)D}{u^2}\left(1 + o\left(\frac{1}{u}\right)\right) + O(r(u)^{-1}). \tag{3.2.4}$$

But we have $g(u) \geq 0$; so, we get, in particular,

$$\beta \geq 2. \tag{3.2.5}$$

From (3.1.27) and (3.2.4), we infer that $A = 2$ in (3.2.2), and thus

$$\phi_\nu(u) = 1 + (-1)^{\nu-1}\frac{(\beta-2)D}{u^2}\left(1 + o\left(\frac{1}{u}\right)\right) + O\left(r(u)^{-1}\right).$$

This shows clearly that $\beta = 2$ is likely the most favourable choice in the sense that then $\phi_\nu(u)$ would converage to 1 very strongly as u tends to infinity. The same can also be inferred from the combination of (3.1.10) and LEMMA 13, for it gives

$$V(y^{1/s})|\phi_\nu(s) - K_\nu(y, y^{1/s}; \delta)|$$

(3.2.6)

$$\leq V(z_1)\sum_{d|P(z_1,y^{1/s})}\rho_\nu(d)\frac{\delta(d)}{d}|1-\phi_{\nu+\omega(d)}\left(\frac{\log y/d}{\log z_1}\right)|+o\left(LV(y^{1/s})\frac{\log^2 y}{\log^3 z_1}\right),$$

provided $s \geq \beta$; this could be small only when $\beta = 2$. **83**

Therefore, we now put

$$\beta = 2; \tag{3.2.7}$$

this would be the optimal value of β, for as we have mentioned already β was to be taken as small as possible, and we have (3.2.5).

Next, we shall determine the value of D on the condition (3.2.7). To this end, we consider the Laplace transform of $G(u)$:

$$A(\tau) = \int_2^\infty e^{-\tau u}G(u)du.$$

(3.1.25) and (3.1.26) with $\beta = 2$ give

$$(\tau(A(\tau))' = A(\tau)(1 - e^{-\tau}) - D(e^{-2\tau} + \int_2^3 e^{-\tau u}(u-1)^{-1}(du).$$

Solving this differential equation on the boundary condition $A(\infty) = O$, we get

$$\tau A(\tau) = D\int\limits_{\tau}^{\infty}\left(e^{-2t} + \int\limits_{2}^{\infty}\frac{e^{-tu}}{u-1}du\right)\exp$$
$$\left(-\int\limits_{1}^{\infty}\frac{e^{-u}}{u}du - \int\limits_{t}^{1}\frac{e^{-u}-1}{u}du\right)dt \times \exp\left(\int\limits_{1}^{\infty}\frac{e^{-t}}{t}dt + \int\limits_{\tau}^{1}\frac{e^{-t}}{t}dt\right).$$

84 Then observing that

(3.2.8) $$\lim_{\tau\to+O}\tau A(\tau) = A = 2,$$

and that

$$\int\limits_{0}^{1}\frac{1-e^{-t}}{t}dt - \int\limits_{1}^{\infty}\frac{e^{-t}}{t}dt = \gamma \text{ (Euler's contant)},$$

we have

$$2e^{\gamma} = D\{h(2) + \int\limits_{2}^{3}\frac{h(u)}{u-1}du\}$$

where

$$h(u) = \int\limits_{o}^{\infty}\exp\left(-tu - \int\limits_{1}^{\infty}\frac{e^{-\xi}}{\xi}d\xi - \int\limits_{t}^{1}\frac{e^{-\xi}-1}{\xi}d\xi\right)dt.$$

But it is easy to check

(3.2.9) $$uh'(u) = -h(u+1)(u > O);$$

this implies that

$$\int\limits_{2}^{3}\frac{h(u)}{u-1}du = h(1) - h(2),$$

whence

$$D = 2e^{\gamma}/h(1).$$

On the other hane, (3.2.9) implies also that

$$\begin{aligned}
h(1) &= -\lim_{u\to+O} uh'(u)\\
&= \lim_{u\to+O} u\int\limits_{o}^{\infty} \exp\left(-tu - \int\limits_{1}^{\infty}\frac{e^{-\xi}}{\xi}d\xi - \int\limits_{t}^{1}\frac{e^{-\xi}-1}{\xi}d\xi + \log t\right)dt\\
&= \lim_{u\to+O} u\int\limits_{o}^{\infty} \exp\left(-tu - \int\limits_{t}^{\infty}\frac{e^{-\xi}}{\xi}d\xi\right)dt,
\end{aligned}$$

and this limit is equal to 1, whence we obtain **85**

$$D = 2e^{\gamma}.$$

Collecting the above discussions, we see that (3.1.25) and (3.1.26) have now the new form:

$$(s\phi_\nu(s))' = \phi_{\nu+1}(s-1) \text{ for } s \geq 2,$$

$$s\phi_1(s) = 2e^{\gamma}, \phi_0(s) = 0 \text{ for } 0 < s \leq 2. \tag{3.2.10}$$

And, in the sequal, we let ϕ_1 and ϕ_0 stand for the functions defined by this equation; in fact, it is clear that (3.2.10) defines two continuous functions inductively starting from the range $0 < s \leq 2$.

Then apply the above argument to the equation (3.2.1).

We now have $C = 0$ in (3.2.3), whence we have

$$\phi_1(s) > \phi_o(s). \tag{3.2.11}$$

On the other hand, this time we have $D = 2e^{\gamma}ab$ initio, and through the analysis of the Laplace transform of $G = \phi_o + \phi_1$, we get $A = 2$ again. Hence, by (3.2.2) and (3.2.3) with $\beta = 2$, we obtain

$$\phi_\nu(s) = 1 + 0(\Gamma(s)^{-1})(s \geq 2). \tag{3.2.12}$$

Finally, we note that by LEMMA 12 ϕ_1 and ϕ_O are strictly decreasing and increasing, respectively, as s increases in the range $s \geq 2$. **86**

3.3 Rosser's Linear Sieve

In this section, we shall demonstrate that the asymptotic relation (3.1.13) actually holds for the function ϕ_ν defined by the equation (3.2.10), and thus establish the fundamental result of Rosser on the linear sieve.

According to (3.2.6) and (3.2.12), it suffices to consider the estimation of the sum

$$\sum_{d|P(z_1,y^{1/s})} \frac{\rho_\nu(d)}{d}\delta(d)\exp\left(-\frac{\log y/d}{\log z_1}\right), \tag{3.3.1}$$

where we should stress that we have $s \geq 2$; this is due to the fact that we have already fixed the value of β to be 2, and thus, in the sequel, we shall work on those ρ_ν with $\beta = 2$.

To this end, we shall prove first the crucial

Lemma 14. *Assuming* (3.1.1) *and* (3.1.2) *we have, for any* $2 \leq u \leq v \leq x^{1/2}$,

$$\sum_{\substack{u\leq p_2<p_1<v\\ p_2^3p_1<x}} \frac{\delta(p_1p_2)}{p_1p_2}V(p_2)\exp\left(-\frac{\log x/p_1p_2}{\log p_2}\right)$$
$$\leq \eta V(v)\exp\left(-\frac{\log x}{\log v}\right)\left\{1+0\left(L\frac{\log v}{\log^2 u}\right)\right\}^2$$

where

$$\eta = \frac{1}{2}\left(\frac{1}{3}+\log 3\right) < 1.$$

87 To show this, we divide the sum to be estimated into two parts $\sum_1$ and $\sum_{II}$ according to $u \leq p_1 < \min(v, x^{1/4})$ and to $\max(u, x^{1/4}) \leq p_1 < v$, respectively. First, we consider the case where

$$v \leq x^{1/4}. \tag{3.3.2}$$

Then we have $\sum_{II} = 0$, and bu LEMMA 11,

$$\sum_I = \sum_{u\leq p_1<v}\ \sum_{u\leq p_2<p_1}$$

$$= \sum_{u \le p_1 < v} \frac{\delta(p_1)}{p_1} \left\{ \frac{V(p_1)\log p_1}{\log \frac{x}{p_1}} \int\limits_{\frac{\log \frac{x}{p_1}}{\log v}}^{\frac{\log \frac{x}{p_1}}{\log u}} e^{1-t} dt \right.$$

$$\left. + 0\left(\frac{LV(p_1)\log p_1}{\log^2 u} \exp\left(-\frac{\log \frac{x}{p_1}}{\log p_1} \right)\right)\right\}$$

$$\le \frac{e}{3}\left(1 + 0\left(L\frac{\log v}{\log^2 u}\right)\right) \sum_{u \le p < v} \frac{\delta(p)}{p} V(p) \exp\left(-\frac{\log \frac{x}{p}}{\log p} \right),$$

since, according to (3.3.2), we have

$$\log p_1 / \log \frac{x}{p_1} \le \frac{1}{3}.$$

Thus, again bu LEMMA 11, we have

$$\sum_1 \le \frac{e}{3}\left(1 + 0\left(L\frac{\log v}{\log^2 u}\right)\right)$$

$$\times \left\{ \frac{V(v)\log v}{\log x} \int\limits_{\frac{\log x}{\log v}}^{\frac{\log x}{\log u}} e^{1-t} dt + 0\left(\frac{LV(v)\log v}{\log^2 u} \exp\left(-\frac{\log x}{\log v}\right)\right)\right\}$$

$$\le \frac{e^2}{12}\left(1 + 0\left(L\frac{\log v}{\log^2 u}\right)\right)^2 V(v)\exp\left(-\frac{\log x}{\log v}\right).$$

Next, we consider the case where **88**

$$x^{1/4} \le v \le x^{1/2}. \tag{3.3.3}$$

As before, we have

$$\sum_I = \sum_{u \le p_1 < x^{1/4}} \sum_{u \le p_2 < p_1}$$

$$\leq \frac{e}{3}\left(1+0\left(L\frac{\log v}{\log^2 u}\right)\right)\sum_{u\leq p<x^{1/4}}\frac{\delta(p)}{p}V(p)\exp\left(-\frac{\log x/p}{\log p}\right)$$

$$=\frac{e}{3}\left(1+0\left(L\frac{\log v}{\log^2 u}\right)\right)\left\{\frac{V(v)\log v}{\log x}\int_4^{\frac{\log x}{\log u}} e^{1-t}dt+0\left(\frac{LV(v)\log v}{\log^2 u}\right)\right\}$$

$$\leq \frac{e^{-2}}{2}\left(1+0\left(L\frac{\log v}{\log^2 u}\right)\right)^2 V(v)\frac{\log v}{\log x}.$$

On the other hand, we have

$$\sum_{II}\leq \sum_{x^{1/4}\leq p_1<v}\ \sum_{u\leq p_2<(\frac{x}{p_1})^{1/3}}$$

$$=\sum_{x^{1/4}\leq p<v}\frac{\delta(p)}{p}\left\{\frac{V(p)\log p}{\log\frac{x}{p}}\int_3^{\frac{\log\frac{x}{p}}{\log u}} e^{1-t}dt+0\left(\frac{LV(p)\log p}{\log^2 u}\right)\right\}$$

$$\leq e^{-2}\left(1+0\left(L\frac{\log v}{\log^2 u}\right)\right)\sum_{x^{1/4}\leq p_1<v}\frac{\delta(p)}{p}\frac{\log p}{\log\frac{x}{p}}V(p)$$

$$=e^{-2}\left(1+0\left(L\frac{\log v}{\log^2 u}\right)\right)\left\{\frac{V(v)\log v}{\log x}\int_{\frac{\log x}{\log v}}^{4}\frac{dt}{t-1}+0\left(L\frac{V(v)\log v}{\log^2 u}\right)\right\}$$

$$=e^{-2}\left(1+0\left(L\frac{\log v}{\log^2 u}\right)\right)^2 V(v)\frac{\log v}{\log x}\log\left(\frac{3}{\frac{\log x}{\log v}-1}\right).$$

89 Hence we get

$$\sum_{I}+\sum_{II}\leq \Delta\left(\frac{\log x}{\log v}\right)V(v)e^{-\frac{\log x}{\log v}}\left(1+0\left(L\frac{\log v}{\log^2 u}\right)\right)^2,$$

where

$$\Delta(\xi)=\left(\frac{1}{3}+\log\frac{3}{\xi-1}\right)\frac{e\xi-2}{\xi},$$

and (3.3.3) is equivalent to $2 \le \xi \le 4$.

Now we have

$$\frac{d}{d\xi}\Delta(\xi) = \frac{\xi-1}{\xi^2}e^{\xi-2}\left\{\frac{-\xi}{(\xi-1)^2} + \frac{1}{3} + \log\frac{3}{\xi-1}\right\}$$
$$= \frac{\xi-1}{\xi^2}e^{\xi-2}\Delta_0(\xi),$$

say. In the interval $2 \le \xi \le 4$, $\Delta_0(\xi)$ attains its maximum at $\xi = 3$ and

$$\Delta_0(3) = \log\frac{3}{2} - \frac{5}{12} < 0.$$

Thus

$$\max_{2\le\xi\le 4}\Delta(\xi) = \Delta(2) = \frac{1}{2}\left(\frac{1}{3} + \log 3\right) > \frac{e^2}{12},$$

Which gives rise to the assertion of the lemma.

We now proceed to the proof of (3.1.13). For this sake, we replace (3.1.20) by the stricter, but still harmless, condition

$$L = 0\left(\frac{\log y}{(\log\log y)^5}\right), \tag{3.3.4}$$

and we set in the above discussion

$$z_1 = \exp\left(\frac{\log y}{(\log\log y)^2}\right) \tag{3.3.5}$$

so that (3.1.5) is satisfied. **90**

We divide the sum (3.3.1) into two parts $\sum_1$ and $\sum_2$ according to $\omega(d) < 2B$ and $\omega(d) \ge 2B$, respectively: here B is to satisfy

$$3^B = \frac{1}{2}\log\log y. \tag{3.3.6}$$

Then LEMMA 9 (with $\beta = 2$) implies that in $\sum_1$ we have

$$\frac{\log y/d}{\log z_1} > \frac{3^{-B}}{2}\frac{\log y}{\log z_1} = \log\log y\,(\rho_\nu(d) = 1),$$

whence we have, uniformly for $s \geq 2$,

$$V(z_1)\sum_1 << V(y^{1/s})\frac{(\log\log y)^4}{\log y} \tag{3.3.7}$$

because of (3.3.5).

To estimate $\sum_2$, we note first that if $\rho_\nu(d) = 1$ and $p(d) \geq w$ then

$$V(w)\exp\left(-\frac{\log y/d}{\log w}\right) << V(p(d))\exp\left(-\frac{\log y/d}{\log p(d)}\right); \tag{3.3.8}$$

this follows from (3.1.3) and the fact that $\rho_\nu(d) = 1$ implies $p(d)d < y$. Thus, for instance, we have

$$\begin{aligned} V(z_1)\sum_{\substack{\omega(d)=2r+1\\ d|P(z_1,y^{1/s})}}\frac{\rho_0(d)\delta(d)}{d}\exp\left(-\frac{\log y/d}{\log z_1}\right) \\ << \sum_{\substack{z_1\leq p<y^{1/s}\\ \ell|P(p,y^{1/s})\\ \omega(\ell)=2r}}\frac{\rho_0(l)\delta(p)\delta(\ell)}{p^\ell}V(p)\exp\left(-\frac{\log y/p^\ell}{\log p}\right), \end{aligned} \tag{3.3.9}$$

91 for $\rho_0(p\ell) = \rho_0(\ell)$ in this sum. Using (3.3.8) once more, we see that the last sum is

$$\begin{aligned} &<< \sum_{z_1\leq p<y^{1/s}}\frac{\delta(p)}{p}\sum_{\substack{\ell|P(p,y^{1/s}\\ \omega(\ell)=2r}}\frac{\rho_0(\ell)\delta(\ell)}{\ell}V(p(\ell))\exp\left(-\frac{\log y/\ell}{\log p(\ell)}\right) \\ &<< \log\log\log y\sum_{\substack{\ell|P(p,y^{1/s}\\ \omega(\ell)=2r}}\frac{\rho_0(\ell)\delta(\ell)}{\ell}V(p(\ell))\exp\left(-\frac{\log y/\ell}{\log p(\ell)}\right), \end{aligned}$$

since we have

$$\sum_{z_1\leq p<y^{1/s}}\frac{\delta(p)}{p} \leq \log\prod_{z_1\leq p<y^{1/s}}\left(1-\frac{\delta(p)}{p}\right)-1$$

$$<< \log\log\log y$$

because of (3.1.3) and (3.3.5). But the last sum over ℓ is equal to

$$\sum_{\substack{k|P(z_1,y^{1/s})\\ \omega(k)=2(r-1)}} \frac{\delta(k)\rho_0(k)}{k} \sum_{\substack{p_2^3 p_1<y/k\\ z_1\leq p_2<p_1<p(k)}} \frac{\delta(p_1p_2)}{p_1p_2}\exp\left(-\frac{\log\frac{y}{p_1p_2k}}{\log p_2}\right).$$

To the inner-sum we can apply LEMMA 14, since $\rho_0(k) = 1$ and $\omega(k) \equiv 0 \pmod 2$ imply $p(k) < (y/k)^{1/2}$; thus the last sum is

$$\leq \eta\left(1+0\left(L\frac{\log y}{\log^2 z_1}\right)\right)^2 \sum_{\substack{k|p(z_1,y^{1/s})\\ \omega(k)=2(r-1)}} \frac{\delta(k)\rho_0(k)}{k} V(p(k))\exp\left(-\frac{\log y/k}{\log p(k)}\right).$$

Hence, by induction on r, we see that the left side of (3.3.9) is, uniformly for $s \geq 2$,

$$<< \left\{\eta\left(1+0\left(L\frac{\log y}{\log^2_{z_1}}\right)\right)^2\right\}^r V(y^{1/s})\log\log\log y.$$

92

In much the same way, we can show, more generally, that

$$V(z_1)\sum_{\substack{\omega(d)=j\\ d|P(z_1,y^{1/s})}} \frac{\rho_\nu(d)\delta(d)}{d}\exp\left(-\frac{\log y/d}{\log z_1}\right)$$

$$<< \left\{\eta\left(1+0\left(L\frac{\log y}{\log^2_{z_1}}\right)\right)^2\right\}^{j/2} V(y^{1/s})\log\log\log y.$$

uniformly for $s \geq 2$ and for all $j \geq 1$.

Hence by (3.3.4)-(3.3.6), we have, for any fixed $\eta' > \eta$,

$$\sum_2 << V(y^{1/s})(\log\log y)^{\frac{\log\eta'}{\log 3}}.$$

By this and (3.3.7), we see that (3.3.1) is

$$0\left(\frac{V(y^{1/s})}{V(z_1)}(\log\log y)^{-3/10}\right),$$

provided (3.3.4) holds and $s \leq 2$.

Therefore, by (3.2.6), we obtain

$$K_\nu(y, y^{1/s}; \delta) = \phi_\nu(s) + O(\log\log y)^{-3/10})$$

uniformly for all bounded $s \geq 2$ on the assumptions (3.1.1), (3.1.2) and (3.3.4), whence we have indeed proved (3.1.13).

Summing up the above discussions, we have established

93 **Theorem 9** (ROSSER'S LINEAR SIEVE)**.** *Provided* (3.1.1) *and* (3.1.2) *with* $L = O\left(\dfrac{\log y}{(\log\log y)^5}\right)$ *we have, uniformly for all* $s \geq 2$*,*

$$(-1)^{\nu-1}\left\{S(A, y^{1/s}) - XV(y^{1/s})(\phi_\nu(s) + O((\log\log y)^{-3/10}))\right\} \leq \sum_{d<y} |R_d|,$$

where ϕ_ν *is defined by* (3.2.10).

Remark. According to (3.1.11), the sum over d on the right side should have been extended up to y_0. But this blemish can easily be removes by taking into account the basic properties of $\phi_\nu(s)$. We should note also that we have actually established this theorem on the assumption (3.1.6), but if it is violated, then the theorem follows from the fundamental lemma (THEOREM 7).

Now it remains to show that Rosser's linear sieve is an optimal result in the sense that it is impossible to improve upon the main-term under the prescribed general conditions.

To this end, we introduce the sequence $A^{(\nu)}(x) = \{n < x$; the total number of prime factors of n is$\}$. congruent to ν (mod 2) where x is to tend to infinity. We have, for any $d < x$,

$$|A_d^{(\nu)}(x)| = \frac{x}{2d} + o\left(\frac{x}{d}\exp\left(-c\left(\log\frac{x}{d}\right)^{1/2}\right)\right).$$

Thus we have

$$X = x/2, \delta \equiv 1,$$

94 and we put

$$y = x \exp\left(-(\log x)^{1/2}\right).$$

Then

$$\sum_{d<y} |R_d^{(\nu)}(x)| = O\left(x \exp\left(-c(\log x)^{1/4}\right)\right). \tag{3.3.10}$$

Also we have, for $s \leq 2$,

$$S(A^{(1)}(x), x^{1/s}) = \pi(x) + O(x^{1/s}),$$
$$S(A^{(0)}(x), x^{1/s}) = 0;$$

that is, we have, for $1 < s \leq 2$,

$$S(A^{(\nu)}(x), x^{1/s}) = XV(x^{1/s})\left(\phi_\nu(s) + O\left(\frac{1}{\log x}\right)\right), \tag{3.3.11}$$

where we have used Mertens' theorem:

$$V(w) = \prod_{p<w}\left(1 - \frac{1}{p}\right) = \frac{e^{-\gamma}}{\log w}\left(1 + O\left(\frac{1}{\log w}\right)\right).$$

On the other hand, the Buchstab identity gives, for $s < t$,

$$S(A^{(\nu)}(x), x^{1/s}) = S(A^{(\nu)}(x), x^{1/t}) - \sum_{x^{1/t} \leq p < x^{1/s}} S\left(A^{(\nu+1)}\left(\frac{x}{p}\right), p\right).$$

Thus, by LEMMA 11 and (3.2.10), we can inducively confirm that (3.3.11) holds for all bounded $s \geq 2$.

But the relation (3.3.11) yields readily **95**

$$S(A^{(\nu)}(x), y^{1/s}) = XV(y^{1/s})(\phi_\nu(s) + 0((\log\ x)^{-1/2}))$$

for all bounded $s \geq 2$. Recalling (3.3.10), this means that the main-term $XV(y^{1/s})\phi_\nu(s)$ of THEOREM 9 is asymptotically attained by the sequence $A^{(\nu)}(x)$.

It may be worth remarking that for $A^{(\nu)}(x)$ Rosser's formula (2.1.14) with $\rho_\nu(d) = \rho_\nu(d; x, 2)$ and $\sigma_\nu(d) = \sigma_\nu(d; x, 2)$ takes the critical form

$$S(A^{(\nu)}(x), z) = \sum_{d|P(z)} \mu(d)\rho_\nu(d)|A_d^{(\nu)}(x)|;$$

namely, the second sum of (2.1.14) does not appear at all. Thus for $A^{(\nu)}(x)$, Rosser's truncation-iteration procedure of the Buchstab identity causes no essential loss.

3.4 Iwaniec's Linear Sieve

Having determined the main-term in the linear sieve, we can now focus our attention onto the error-term, which has been left in a crude form in THEOREM 9: We shall inject the smoothing device developed in § 2.3 into the argument leading to THEOREM 9.

As before, we assume always (3.1.1) and (3.1.2).

We begin our discussion by showing a smoothed version of LEMMA
96 13 (with $\beta = 2$ and ϕ_ν defined by (3.2.10)). But to this end, we have to specify the mode of the dissection of the interval $[z_1, z)$ which was introduced in § 2.3. We put

$$z = z_1 z_2^J$$

where J is a large integer, and z_1, z_2 are large parameters to be determined later in terms of z. And we define I to be one of the intervals

$$[z_1 z_2^{j-1}, z_1 z_2^j)(1 \leq j \leq J). \tag{3.4.1}$$

Further, in view of the result of § 3.2, we set

$$\beta = 2$$

in the definitions of Θ_ν and Δ_ν (cf. (2.3.4)).

Then we shall show

Lemma 15. *We assume* (3.1.1) *and* (3.1.2). *Let ϕ_ν be defined by* (3.2.10), *and let $z \leq y^{1/2}$. Then we have*

$$V(z)\phi_\nu\left(\frac{\log y}{\log z}\right) = V(z_1)\sum_K (-1)^{\omega(K)}\Theta_\nu(K)\sum_{d\epsilon K}\frac{\delta(d)}{d}\phi_{\nu+\omega(d)}\left(\frac{\log \frac{y}{d}}{\log z_1}\right)$$
$$+ 0\left\{V(z)\frac{\log^2 z}{\log^3 z_1}\left(L + \log z_2 \log\left(\frac{\log z}{\log z_1}\right)\right)\right\}.$$

We shall prove first that, for

$$d\epsilon K, \Theta_\nu(K) = 1, \tag{3.4.2}$$

we have

$$V(p(d))\phi_{\nu+\omega(d)}\left(\frac{\log\frac{y}{d}}{\log p(d)}\right) = V(z_1)\phi_{\nu+\omega(d)}\left(\frac{\log\frac{y}{d}}{\log z_1}\right)$$

$$-\sum_{I<K}\lambda_\nu(KI)\sum_{p\epsilon I}\frac{\delta(p)}{p}V(p)\phi_{\nu+\omega(d)+1}\left(\frac{\log\frac{y}{dp}}{\log p}\right)$$

$$+\,0\left\{\frac{V(p(d))\log p(d)}{\log^2 z_1}(L+\omega(d)\log z_2)\right\}, \tag{3.4.3}$$

where λ_ν is defined at (2.3.3). In fact, since (3.4.2) implies 97

$$\phi_{\nu+\omega(d)+1}\left(\frac{\log\frac{y}{d\xi}}{\log\xi}\right) = 0(1)$$

for $\xi < p(d)$, we have, just as in the proof of LEMMA 13,

$$V(p(d))\phi_{\nu+\omega(d)}\left(\frac{\log\frac{y}{d}}{\log p(d)}\right) = V(z_1)\phi_{\nu+\omega(d)}\left(\frac{\log\frac{y}{d}}{\log z_1}\right)$$

$$-\sum_{z_1\le p<p(d)}\frac{\delta(p)}{p}V(p)\phi_{\nu+\omega(d)+1}\left(\frac{\log\frac{y}{d}}{\log p}\right) + O\left(L\frac{V(p(d))\log p(d)}{\log^2 z_1}\right).$$

If $\nu+\omega(d)\equiv 0 \pmod 2$, then the last sum over p is

(3.4.4)

$$\sum_{I<K}\lambda_\nu(KI)\sum_{p\epsilon I}\frac{\delta(p)}{p}V(p)\phi_{\nu+\omega(d)+1}\left(\frac{\log\frac{y}{dp}}{\log P}\right) + o\left\{\sum_{p<p(d)pd\epsilon K}\frac{\delta(p)}{P}V(p)\right\}$$

because $\lambda_\nu(KI)=1$, on our present assumption and (3.4.2). This error-term is, by (3.1.3),

$$0\left(V(p(d))\log p(d)\frac{\log z_2}{\log^2 z_1}\right). \tag{3.4.5}$$

On the other hand, if $\nu+\omega(d)\equiv 1 \pmod 2$, then the sum over p in 98
question is represented as

$$\sum_{\substack{I<K\\(I)^3(K)<y}}\sum_{p\epsilon I} + \sum_{\substack{I<K\\(I)^3(K)\ge y}}\sum_{p\epsilon I} + \sum_{\substack{p<p(d)\\pd\epsilon K}} \tag{3.4.6}$$

where $(K) = (I_1)(I_2)\dots(I_r)$ if $K = I_1 I_2 \dots I_r$. The first double sum can be put in the form (3.4.4) without the error-term, and the last sum over p has obviously the upper bound (3.4.5). It remains to estimate the middle sum; it is equal to

$$\sum_{\substack{I<K\\ (I)^3(K)\geq y}} \sum_{\substack{p\epsilon I\\ p^3<y/d\\ p<p(d)}} \frac{\delta(p)}{p} V(p)\phi_0\left(\frac{\log\frac{y}{dp}}{\log p}\right). \tag{3.4.7}$$

Here we have, by the mean value theorem,

$$\begin{aligned}\phi_0\left(\frac{\log\frac{y}{dp}}{\log p}\right) &= \phi_0\left(\frac{\log\frac{y}{dp}}{\log p}\right) - \phi_0(2)\\ &\ll \log\left(\frac{y}{dp^3}\right)/\log p,\end{aligned}$$

which is

$$\ll \omega(d)\log z_2/\log z_1,$$

for $p\epsilon I, d\epsilon K, (I)^3(K)\geq y$ imply

$$p^3 d z_2^{\omega(d)+3} \geq (I)^3(K)\geq y.$$

99 Thus (3.4.7) is less than a constant multiple of

$$\begin{aligned}&\omega(d)\frac{\log z_2}{\log z_1}\sum_{z_1\leq p<p(d)}\frac{\delta(p)}{p}V(p)\\ &\ll \omega(d)\frac{\log z_2}{\log^2 z_1}V(p(d))\log p(d).\end{aligned}$$

Collecting these observations, we obtain (3.4.3).

In much the same way, we get

$$\begin{aligned}V(z)\phi_\nu\left(\frac{\log y}{\log z}\right) &= V(z_1)\phi_\nu\left(\frac{\log y}{\log z_1}\right)\\ &\quad - \sum_I \lambda_\nu(I)\sum_{p\epsilon I}\frac{\delta(p)}{p}\phi_{\nu+1}\left(\frac{\log\frac{y}{p}}{\log p}\right)V(p)\end{aligned}$$

$$+ O\left\{\frac{V(z)\log z}{\log^2 z_1}(L + \log z_2)\right\}.$$

Then the formula (3.4.3) allows us to iterate the last one, and after the infinite iteration we get the formula of the lemma, apart from the error-term which is

$$O\left\{\frac{V(z)\log z}{\log^2 z_1}\sum_{r=0}^{\infty}\frac{L + r\log z_2}{r!}\left(\sum_{z_1 \le p < z}\frac{\delta(p)}{p}\right)^r\right\}$$
$$= o\left\{V(z)\frac{\log^2_z}{\log^3 z_1}\left(L + \log z_2 \log\left(\frac{\log z}{\log z_1}\right)\right)\right\},$$

whence the assertion of the lemma.

We are now at the stage to combine LEMMA 10 with LEMMA 15.

For this sake, we introduce very mild restrictions on δ and L. We **100**
assume that, for any $3 \le u < v$,

$$\sum_{u \le p < v}\frac{\delta(p^2)}{p^2} = 0((\log\log u)^{-1}), \tag{3.4.8}$$

and that

$$L = 0\left(\frac{\log z}{\log\log z}\right). \tag{3.4.9}$$

Further, we set in (3.4.1)

$$z_1 = z^{\tau^2}, z_2 = z^{\tau^9}, \tau = (\log\log z)^{-\frac{1}{10}}; \tag{3.4.10}$$

thus, in particular,

$$J \le (\log\log z)^{9/10}. \tag{3.4.11}$$

Also, we assume, in the sequel, that

$$y \ge z^2. \tag{3.4.12}$$

Now by the fundamental lemma (THEOREM 7) we have, for a certain sequence $\{\xi_f^{(\nu)}\}$ which is independent of d,

$$(-1)^\nu \left\{S(A_d, z_1) - \frac{\delta(d)}{d} XV(z_1)\left(1 + O\left(\exp\left(-\frac{H}{2}\log H\right)\right)\right)\right\} \geq (-1)^\nu \sum_{\substack{f<z_1^H \\ f|P(z_1)}} \xi_f^{(\nu)} R_{df} \tag{3.4.13}$$

where $d|P(z_1, z)$, and H is at our disposal. We set

$$H = \tau^{-1} = (\log\log z)^{1/10}.$$

101 On the other hand, modifying the inequality of LEMMA 10 (with $\beta = 2$), we have

$$\begin{aligned}
(-1)^\nu &\left\{S(A,z) - XV(z_1)\sum_K \Theta_\nu(K)(-1)^{\omega(K)} \sum_{d\epsilon K} \frac{\delta(d)}{d}\right\} \\
&\geq \sum_K \Theta_\nu(K)(-1)^{\nu+\omega(K)} \sum_{d\epsilon K}\left(S\left(A_d, z_1\right) - \frac{\delta(d)}{d} XV(z_1)\right) \\
&- \sum_{\substack{I<K \\ \omega(K)\equiv \nu+1 \pmod 2}} \Theta_\nu(KI) \sum_{\substack{d\epsilon K \\ p,p\epsilon I}} \left(S(A_{dpp'}, z_1) - \frac{\delta(dpp')}{dpp'} XV(z_1)\right) \\
&- \sum_{\substack{I<K \\ \omega(K)\equiv \nu+1 \pmod 2}} \Theta_\nu(KI) \sum_{\substack{d\epsilon K \\ p,p\epsilon I}} \frac{\delta(dpp')}{dpp'} XV(z_1).
\end{aligned}$$

Insertion of (3.4.13) into this yields

$$\begin{aligned}
(-1)^\nu &\left\{S(A,z) - XV(z_i)\sum_K \Theta_\nu(K)(-1)^{\omega(K)} \sum_{d\epsilon K} \frac{\delta(d)}{d}\right\} \\
&\geq \sum_K \Theta_\nu(K)(-1)^{\nu+\omega(K)} \sum_{\substack{d\epsilon K \\ f<z_1^H \\ f|P(z_1)}} \xi f^{\nu+\omega(K))} R_d f
\end{aligned} \tag{3.4.14}$$

$$- \sum_{\substack{I<K \\ \omega(K)\equiv \nu+1 \pmod 2}} \Theta_\nu(KI) \sum_{\substack{d\epsilon K \\ P,P'\epsilon I \\ f<z_1^H \\ f|P(z_1)}} \xi_f^{(1)} R_{dpp'f}$$

$$- o\left(\exp\left(-\frac{H}{2}\log H\right)\right) XV(z_1)\sum_K \Theta_\nu(K)\sum_{d\epsilon K}\frac{\delta(d)}{d}$$

$$- o(XV(z_1))\sum_{I<K}\Theta_\nu(KI)\sum_{\substack{d\epsilon K \\ p,p'\epsilon I}}\frac{\delta(dpp')}{dpp'}.$$

102

By (3.1.3), (3.4.8) and (3.4.10), the last 0− terms are easily estimated to be

(3.4.15) $$0(XV(z)\tau^6).$$

Also, by virtue of LEMMA 15 the sum over K, d on the left side of (3.4.14) is equal to

$$XV(z)\left\{\phi_\nu\left(\frac{\log y}{\log z}\right) + 0(\tau^2)\right\}$$

(3.4.16) $$+0\left\{XV(z_1)\sum_{d|P(z)}\frac{\rho_\nu(d)\delta(d)}{d}| - \phi_{\nu+\omega(d)}\left(\frac{\log\frac{y}{d}}{\log z_1}\right)|\right\},$$

because $\Theta_\nu(K) = 1$, $d\epsilon K$ imply readily $\rho_\nu(d) = 1$, where $\rho_\nu(d) = \rho_\nu(d; y, 2)$ is defined at (2.1.12). We have to estimate this error-term on our present suppositions. As before, it is sufficient to consider sum

$$V(z_1)\sum_{d|P(z)}\frac{\rho_\nu(d)}{d}\delta(d)\exp\left(-\frac{\log\frac{y}{d}}{\log z_1}\right),$$

which is divided into two parts as

(3.4.17) $$V(z_1)\sum_{\omega(d)<2B'} + V(z_1)\sum_{\omega(d)\geq 2B'},$$

where B' is to satisfy

$$3^{B'} = \tau^{-1} = (\log\log z)^{\frac{1}{10}}.$$

103

We have, by LEMMA 9,

$$V(z_1)\sum_{\omega(d)<2B'} \leq V(z_1)\sum_{d|p(z)}\frac{\rho_\nu(d)\delta(d)}{d}\exp\left(-\frac{3^{-B'}\log y}{2\log z_1}\right)$$

$$= 0\left(V(z)\tau^{-4}\exp\left(-\frac{1}{\tau}\right)\right). \tag{3.4.18}$$

On the other hand, LEMMA 14 gives, as before,

$$V(z_1)\sum_{\omega(d)\geq 2B'} \ll V(z)\left\{\eta\left(1+0\left(L\frac{\log z}{\log^2 z_1}\right)\right)^2\right\}^{B'}$$

$$= V(z)\left\{\eta(1+0(\tau^6))\right\}^{B'}$$

$$\leq V(z)(\log\log z)^{-\frac{1}{50}}, \tag{3.4.19}$$

in which we have used (3.4.9).

Collecting (3.4.14) - (3.4.19), we obtain

$$(-1)^\nu\left\{S(A,z)-XV(z)\left(\phi_\nu\left(\frac{\log y}{\log z}\right)+0((\log\log z))^{-\frac{1}{50}}\right)\right\}$$

$$\geq \sum_K \Theta_\nu(K)(-1)^{\nu+\omega(K)}\sum_{\substack{d\epsilon K\\ f<z^\tau\\ f|P(z_1)}}\xi^{\nu+\omega(K))_f}R_{df}$$

$$-\sum_{\substack{I<K\\ \omega(K)\equiv\nu+1 \pmod 2}}\Theta_\nu(KI)\sum_{\substack{d\epsilon K;p,p'\epsilon I\\ f<z^\tau, f|p(z_1)}}\xi_f^{(1)}R_{dpp'f}. \tag{3.4.20}$$

104 In order to transform further these double sums, we make here a crucial observation.

Lemma 16. *Let $y = MN \geq z^2$ with arbitrary $M, N \geq 1$. Then $\Theta_\nu(K) = 1$ implies that there exists a decomposition $K = K_1K_2$ such that $(K_1) < m, (K_2) < N$. Also, if $\Theta_\nu(KI) = 1, I < K$ and $\omega(K) \equiv \nu + 1 \pmod 2$, then we have a decomposition $K = K_1K_2$ as above, and moreover, at least one of the following three cases occurs:*

$$\{(K_1)(I) < M, (K_2)(I) < N\}, \left\{(K_1)(I)^2 < M, (K_2) < N\right\},$$
$$\left\{(K_1) < M, (K_2)(I)^2 < N\right\}.$$

To show this, let $K = I_1 I_2 \ldots I_r$, $I_1 > I_2 \ldots I_r$. We have $(I_I) < z \leq \sqrt{y} \leq \max(M, N)$; so $(I_1) < M$ or $(I_1) < N$. Let us assume that we have already the decomposition $I_1 I_2 \ldots I_j = K_1^{(j)} K_2^{(j)}$ such that $(K_1^{(j)}) < M$ and $(K_2^{(i)}) < N$. Since $\Theta_\nu(K) = 1$ gives obviously $(I_{j+1})^2 (I_j) \ldots (I_1) < y$ for any $j \leq r-1$, we have either $(K_1^{(j)})(I_{j+1}) < M$ or $(K_1^{(j)})(I_{j+1}) < N$; for, otherwise, we would have $(I_{j+1})^2 (K_1^{(j)})(K_2^{(j)}) \geq MN = y$, a contradiction. Thus we get, inductively, the first assertion of the lemma. As for the second assertion, we note that the stated condition on K, I implies $(K)(I)^3 < y$, which readily yields the claim.

We now return to (3.4.20), and we assume that

$$y = MN \geq z^2; M, N \geq 1.$$

We have, by the lemma just proved, **105**

$$\sum_K \Theta_\nu(K) \Big| \sum_{\substack{d \epsilon K \\ f < z^\tau \\ f | P(z_1)}} \xi_f^{(\nu+\omega(K))} R_{df} \Big| = \sum_K \Theta_\nu(K) \Big| \sum_{\substack{d_1 \epsilon K_1 \\ d_2 \epsilon K_2 \\ f < z^\tau \\ f | P(z_1)}} \xi_f^{(\nu+\omega(K))} R_{d f_1 d_2} \Big|,$$

where $K = K_1 K_2$, $(K_1) < M$, $(K_2) < N$. But the last absolute value is obviously not larger than the expression

$$\sup_{\alpha,\beta} \Big| \sum_{\substack{m < Mz^\tau \\ n < N}} \alpha_m \beta_n R_{mn} \Big| \tag{3.4.21}$$

where $\alpha = \{\alpha_m\}$, $\beta = \{\beta_n\}$ are variable vectors such that $|\alpha_m| \leq 1, |\beta_n \leq 1|$. On the other hand, by the second assertion of LEMMA 16, the second sum on the right hand side of (3.4.20) can be written as

$$\sum_{\substack{I < K \\ \omega(K) \equiv \nu+1 \pmod 2}} \Theta_\nu(KI) \sum_{\substack{d_1 \epsilon K_1 \\ d_2 \epsilon K_2 \\ p, p' \epsilon I \\ f < z^\tau \\ f | P(z_1)}} \xi_f^{(1)} R_{d_1 d_2 p p' f'}$$

where $K = K_1K_2$ and one of the three cases listed in the lemma occurs. Let us assume, for example, that we have $(K_1)(I)^2 < M, (K_2) < N$. Then **106**
we have $fd_1pp' < Mz^\tau$ and $d_2 < N$; thus, again the inner sum is, in absolute value, not larger than the expression (3.4.21). Other cases can be treated in just the same way. Finally, we observe that the number of admissible K, I in the formula (3.4.20) does not exceed 2^{J+2} which is less than $\log z$, because of (3.4.11).

Therefore we have now established

Theorem 10 (IWANIEC'S LINEAR SIEVE)**.** *We assume* (3.1.1), (3.1.2) *and* (3.4.8) *with* $L = 0\left(\frac{\log z}{\log\log z}\right)$. *Then we have, for any* $MN \geq z^2$,

$$(-1)^{\nu-1}\left\{S(A,z) - \left(\phi_\nu\left(\frac{\log MN}{\log z}\right) + O\left((\log\log z)^{-\frac{1}{50}}\right)\right)XV(z)\right\}$$
$$< \log z \sup_{\alpha,\beta} |\sum_{\substack{m<M \\ n<N}} \alpha_m\beta_n R_{mn}|,$$

where ϕ_ν *is defined by* (3.2.10), *and* $\alpha = \{\alpha_m\}, \beta = \{\beta_n\}$ *are variable vectors such that* $|\alpha_m| \leq 1, |\beta_n| \leq 1$.

Remark 3.4.1. In the above, we have actually proved this inequality with Mz^τ in place of M on the right side, but this blemish can easily be removed in much the same way as in the remark to THEOREM 9.

NOTES (III)

THEOREM 9 was first proved by Rosser, but his work has never
107 been published. Being anticipated some ten years by Rosser's work, but independently, Jurkat and Richert [36] proved essentially the same result as THEOREM 9 completely; in their remarkable proof Selberg's sieve was used as an aid to start the truncation-iteration procedure of Buchstab's identity which is quite similar to that of Rosser. Rosser's argument is briefly sretched in Selberg's expository paper [75], and also Iwaniec [28] worked out the full detail of this fundamental sieve idea.

Our proof of THEOREM 9 follows the argument of Motohashi [[60], I] who combined some of important ideas of Jurkat and Richert with those of Rosser. This fact is embodied in LEMMA 13 and LEMMA 14; especially, LEMMA 13 shows well how natural Rosser's idea is. The analysis of the difference-differential equation (3.1.25) which is developed is § 3.2 is conducted partly by employing the ideas of de Brui jn [9]; the use of Laplace transform is also indicated by Selberg [75]. Also we note that (3.1.3), (3.1.4) and LEMMA 11, LEMMA 12 are quoted from Halberstam and Richert [[12], see p.53, p.144, p.214 and p.227, respectively].

Our argument may be generalized, at least in principle, so as to include the K-dimensional sieve problems with $K \neq 1$, but then we should have t overcome anew the difficulty pertaining to the convergence problem arising from the infinite iteration procedure; in our case, this was **108**
solved in LEMMA 14. For the general case, see Iwaniec's work [30] to which we owe much.

The observation that Rosser's linear sieve is optimal is due to Selberg [75] (cf. also [73]), and our example is quoted from there; a related subject was studied by Bombieri [7] (cf. also Friedlamder and Iwanied [14]) in a more general setting.

THEOREM 10 is due to Iwaniec [31]. This far-reaching improvement of Rosser's linear sieve was a major event in the theory of sieve methods; it allows us to combine very effectively the linear sieve with various powerful analytical means, e.g. hybrid mean vale theorems for Dirichlet polynomials. Some of the deep consequences of such applications to fundamental problems in analytic number theory are surveyed in Iwaniec's own expository paper [33]; later in PART II we shall give an important application to the theory of the distribution of prime numbers.

One should note how nicely Iwaniec exploited the particular form of Rosser's weights ρ_ν. Prior to Iwaniec's discovery Motohashi [52] did the same for Selberg's Λ^2-sieve.

The argument of 3.4 is due to Motohashi [[60], II] which is a straightforward refinement of the one developed in the preceding section; LEMMA 16 is a refined version of Iwaniec's decisive observation.

ADDENDUM ([2]). After studying the first draft of the present chap- **109**
ter. Professor Halberstam kindly showed us the following penetrating

[2]By the courtesy of Professor Halberstam

observation.

From (3.1.10) and LEMMA 13 (with $\rho = 2$ and ϕ_ν defined by (3.2.10)) we get, for $2 \le z_1 < z \le \sqrt{y}$,

$$\text{(i)}\quad (-1)^\nu V(z)\left\{K_\nu(y,z;\delta) - \left(\phi_{-\nu}\left(\frac{\log y}{\log z}\right) + O\left(L\frac{\log^2 z}{\log^3 z_1}\right)\right)\right\}$$
$$= V(z_1)\sum_{d|P(z_1.z)}(-1)^\nu\mu(d)\rho_\nu(d)\frac{\delta(d)}{d}\left(1-\phi_{\nu+\omega(d)}\left(\frac{\log\frac{y}{d}}{\log z_1}\right)\right).$$

But all summands on the right side are non-negative, and hence

$$\text{(ii)}\qquad (-1)^\nu K_\nu(y,z;\delta) \ge (-1)^\nu\left(\phi_\nu\left(\frac{\log y}{\log z}\right) + O\left(L\frac{\log^2 z}{\log^3 z_1}\right)\right),$$

which is essentially equivalent to the assertion of THEOREM 9.

Namely, we can demonstrate Rosser's linear sieve without proving painstakingly the convergence lemma (LEMMA 14); just the same can be said about the corresponding part of our proof of Iwaniec's linear sieve. This is a remarkable observation, for, it may be applied equally well to the higher dimensional sieve situation and provide Rosser's sieve (in the sense of [30]) with a more accessible proof.

Our LEMMA 14 is thus to be regarded as a means to ensure that what was disregarded in deducing (ii) from (i) is, in fact, negligible.

Part II

Topics in Prime Number Theory

Chapter 4

Zero-Free Regions for $\zeta(s)$ and $L(s,\chi)$

WE NOW TURN to the applications of the results obtained in the pre- 110
ceding chapters to some basic problems in the theory of the distribution of prime numbers. As the first of such applicatons we shall show in this chapter that, to some extent, the sieve method can take the place which has long been occupied solely by the complex variable method in the investigations of the fundamental properties of $\zeta(s)$ and $L(s,\chi)$. More precisely, we shall demonstrate that, by employing, instead, the Selberg sieve for multiplicative functions, the classical function-theoretical convexity argument can be dispensed with in deducing Vinogradov's zero-free region and Page-Landau-Siegel-Linnik's theorem from the relevant elementary estimates of the zeta-and L-funtions.

We shall also dwell on the Brun-Titchamarsh theorem; this is included here, because of its relation with the exceptional zeros of L-functions.

4.1 Vinogradov's Zero-Free Region for $\zeta(s)$

In order to extract the informations on the distribution of prime numbers from the Euler product for $\zeta(s)$ which connects prime numbers with
a fairly smooth analytical expression, we need to extend the zero-free 111

region of $\zeta(s)$ as far as possible to the left of the line $\sigma = 1$. Probably the simplest way to get an effective zero-free region of this sort is the one due to Landau; he deduced, for $t > 2$,

$$\frac{\zeta'}{\zeta}(s) = O((\log t)^7) \text{ for } \sigma > 1 - c(\log t)^{-9} \tag{4.1.1}$$

from de la Vallee Poussin's inequality

$$\zeta^3(\sigma)|\zeta(\sigma + it)^4\zeta(\sigma + 2it)| > 1(\sigma > 1). \tag{4.1.2}$$

For the sake of a later purpose, we stress that (4.1.1) is an elementary result in the sense that in deriving it we do not need to appeal to the complex variable method.

Although (4.1.2) already yields a relatively good estimate of the error-term in the prime number theorem, to get finer results we have to seek for a wider zero-free region. And a general theorem of Landau which is a consequence of Hadamard-Borel-Caratheodory's convexity theorem, thus much involve in the complex variable method, converts our problem to the one of estimating $\zeta(s)$ in the vicinity of $\sigma = 1$. Further, the elementary formula

$$\zeta(s) = \sum_{n<N} n^{-s} + \frac{N^{1-s}}{s-1} + O(N^{-\sigma}) \tag{4.1.3}$$

112 which holds uniformly for $\sigma > 0$, $|t| < N$ reduces it to that of the sum

$$\sum_{n<N} n^{it}.$$

And for our purpose, it is desirable to have an estimate which is particularly effective for those N much smaller than $|t|$. In this context, the following *purely elementary* result of Vinogradov is the hitherto best one:

Lemma 17. *For $N < ct,\ t > 2$, we have*

$$\sum_{n<N} n^{it} << N \exp\left(-c\frac{\log^3 N}{\log^2 t}\right).$$

This yields, via the general theorem of Landau mentioned above,

$$\frac{\zeta'}{\zeta}(s) = 0\left((\log t)^{2/3}(\log\log t)^{1/3}\right) \quad \text{for } \sigma > 1 - \frac{c}{(\log t)^{2/3}(\log\log t)^{1/3}} (t \geq 3), \tag{4.1.4}$$

which is the deepest zero-free region for $\zeta(s)$ known at present, and has had profound influence on diverse problems involving prime numbers.

Since its discovery, it has long been maintained that Vinogradov's zero-free region (4.1.4) represents one of the most important analytical properties of $\zeta(s)$, partly because only Hadamard's global theory of integral functions and the convexity principle of the Borél-Carathéodory type have been able to derive it from the result stated in LEMMA 17. **113**
We can, however, break away from this prevalent notion in the theory of the zeta-function, for, as we shall show below, there exists an elementary argument with which we can deduce a result of the same depth as (4.1.4) from LEMMA 17.

Our proof of the last assertion depends largely on a special instance of the Selberg sieve for multiplicative functions as well as an auxiliary result ((4.1.6) below) from the theory of elementary proofs of the prime number theorem with remainder term.

We begin our discussion by making the second point explicit. We shall require a special case of THEOREM 4, and as we have remarked in § 1.3 the necessary upper bound for $R_d(x)$ (cf. (1.3.8)) can also be obtained in an elementary manner. Here we prove this fact for the case $k = 1$ only, since this is sufficient for our present purpose.

We see from (1.3.8) (with $k = 1$) that, retaining the notations of § 1.3, it suffices to deal with the sums

$$\sum_{\substack{u<x \\ (u,d)=1}} \mu(u)u^{-\eta}, \quad \sum_{\substack{u<x \\ (u,d)=1}} \mu(u)u^{-\eta}\log u, \tag{4.1.5}$$

where $x < z$ can be assumed to be sufficiently large. A routine argument transforms the first sum into the expression:

$$\sum_{\substack{uf<x\\ f|d^\infty\\ f<\sqrt{x}}} \mu(u)(uf)^{-\eta} + 0\left(\sum_{\substack{uf<x\\ f|d^\infty\\ f\geq\sqrt{x}}} (uf)^{-\eta}\right)$$

$$= \sum_{\substack{f|d^\infty\\ f<\sqrt{x}}} f^{-\eta}\left\{\zeta(\eta)^{-1} - \sum_{u>x/f} \mu(u)u^{-\eta}\right\} + 0\left(x^{-\frac{1}{10}} \prod_{p|d}\left(1+\frac{1}{\sqrt{p}}\right)\right).$$

114 Then we appeal to the elementary estimate

$$\sum_{n<y} \mu(n)n^{-1} << (\log y)^{-1}(y\geq 2). \tag{4.1.6}$$

This implies

$$\sum_{u>x/f} \mu(u)u^{-\eta} << \left(\frac{x}{f}\right)^{1-\eta} (\log x)^{-1},$$

whence the first sum in (4.1.5) is

$$<< (\log x)^{-1} \prod_{p|d}\left(1+\frac{1}{\sqrt{p}}\right).$$

In much the same way, we can show that the second sum in (4.1.5) is

$$<< \prod_{p|d}\left(1+\frac{1}{\sqrt{p}}\right).$$

Inserting these into (1.3.8), $k = 1$, we immediately obtain an elementary account of THEOREM 4 for the case $k = 1$.

Now, to make explicit the first point, i.e. the sieve aspect of our argument we have to make a rather lengthy preparation; the complexity is caused mainly by our elementary treatment of various estimates.

115 First, we introduce two parameters δ and B such that

$$(\log t)^{-10} \leq \delta \leq (\log t)^{-2/3}, (\log\log g)^{1/3} \geq B > 0 \tag{4.1.7}$$

and we shall assume always in the sequel that B and t are sufficiently large. We shall use the notations

$$\begin{aligned} Y(t) &= (\log t)^{2/3}(\log\log t)^{1/3}, \\ Q(t) &= \exp(Y(t))(\log\log t)^{2/3}, \\ E(t) &= (\log t)^{c}|\zeta(1-\delta+it)| + (\log t)^{-cB^3}; \end{aligned}$$

the last one allows us to use the convention:

$$0\{(\log t)^{c}E(t)\} = 0\{E(t)\}.$$

Further, we introduce the multiplicative function

$$f(n) = |\sigma(n,-\delta-it)|^2,$$

where $\sigma(n,a)$ is the sum of the a-th powers of divisors of n; here, we should note that as a special case of an identity due to Ramanujan we have, for $\sigma > 1$,

$$\sum_{n=1}^{\infty} f(n)n^{-s} = \zeta(s)\zeta(s+2\delta)\zeta(s+\delta-it)\zeta(2(s+\delta))^{-1}. \tag{4.1.8}$$

Afterwards, we shall apply the Selberg sieve to f, and for this sake, we prove first the following lemma; the argument employed in the proof is the one common in the problems pertaining to sums of divisor functios, and so we may be brief.

Lemma 18. *Let* **116**

$$H(x) = \sum_{n<x} f(n)n^{-1+2\delta}$$

and

$$\mathscr{F} = \zeta(1+2\delta)|\zeta(1+\delta+it)|^2\zeta(2(1+\delta))^{-1}.$$

Then we have

$$H(x) = (2\delta)^{-1}\mathscr{F}\times^{2\delta}(1+O(E(t)))$$

provided

$$Q(t) \geq \times \geq \exp(BY(t)).$$

To show this, we note that (4.1.8) implies

$$f(n) = \sum_{d^2 d_1 k_2 d_3 d_4 = n} \mu(d) d^{-2\delta} d_2^{-2\delta} d_3^{-\delta+it} d_4^{-\delta-it}$$

so

$$H(x) = \sum_{d<x^{\frac{1}{4}}} \mu(d) d^{-2(1-\delta)} H_1(xd^{-2}) + O(x^{-1/5}), \tag{4.1.9}$$

where

$$H_1(y) = \sum_{d_1 d_2 d_3 d_4 < y} d_1^{-1} d_2^{-1+2\delta} d_3^{-1+\delta+it} d_4^{-1+\delta-it}.$$

We decompose $H_1(y)$ into three parts as follows:

$$\begin{aligned} H_1(y) = &\sum_{uv \le \sqrt{y}} u^{-1} v^{-1+2\delta} K\left(1-\delta it,\ 1-\delta-\text{it};\ \frac{y}{uv}\right) \\ &+ \sum_{uv \le \sqrt{y}} u^{-1+\delta+\text{it}} v^{-1+\delta-it} K\left(1-2\delta, 1; \frac{y}{uv}\right) \\ &- K(1-\delta+it,\ 1-\delta-it;\ \sqrt{y}) K(1-2\delta, 1;\ \sqrt{y}), \end{aligned} \tag{4.1.10}$$

117 where

$$K(s, w; y) = \sum_{mn \le y} m^{-s} n^{-w}.$$

Similarly, we have

$$\begin{aligned} K(s, w; y) = &\sum_{m \le \sqrt{y}} m^{-s} U\left(w, \frac{y}{m}\right) + \sum_{n \le \sqrt{y}} n^{-w} U\left(s, \frac{y}{n}\right) \\ &- U(w,\ \sqrt{y}) U(s,\ \sqrt{y}). \end{aligned} \tag{4.1.11}$$

where

$$U(s, y) = \sum_{n \le y} n^{-s}.$$

Hence the problem is reduced to an asymptotic evaluation of $U(s,y)$ at the points $s = 1, 1-2\delta, 1-\delta \pm$ it. The first two cases give no difficulty; we have

$$U(1, y) = \log y + \gamma + O\left(\frac{1}{y}\right),$$

$$U(1-2\delta, y) = \frac{1}{2\delta} y^{2\delta} + \zeta(1-2\delta) + O(y^{-1+2\delta}), \tag{4.1.12}$$

where γ is the Euler constant. As for the points $s = 1-\delta \pm$ it we require LEMMA 17. We set $N = t$ in (4.1.3), getting

$$U(\sigma + it, y) = \zeta(\sigma + it) - \sum_{y<n\leq t} n^{-\sigma - it} + O(t^{-\sigma}),$$

but the lemma implies

$$\sum_{y<n\leq t} n^{-\sigma-it} << \max_{y\leq \times \leq t} \times^{1-\sigma} \exp\left(-c\frac{(\log x)^3}{(\log t)^2}\right);$$

thus we get **118**

$$U(\sigma + \text{ it }, y) = \zeta(\sigma + \text{ it }) + o((\log t)^{-cB^3}), \tag{4.1.13}$$

provided

$$\sigma \geq 1-\delta, Q(t) \geq y \geq \exp(BY(t)).$$

In particular, we have, for $\sigma \geq 1-\delta$,

$$\zeta(\sigma + \text{it}) = o(\log^c t). \tag{4.1.14}$$

Inserting (4.1.12) into (4.1.11) (but with $s = 1-2\delta, w = 1$) we get

$$\begin{aligned} K(1-2\delta, 1; y) = {} & \frac{1}{2}\log y\, U(1-2\delta, \sqrt{y}) + \left(\frac{1}{2}\log y + \gamma\right)\zeta(1-2\delta) \\ & + \frac{y^{2\delta}}{2\delta} U(1+2\delta, \sqrt{y}) + U'(1-2\delta, \sqrt{y}) + 0(y^{-\frac{1}{2}}\log^c t), \end{aligned}$$

where $U'(s, y) = \frac{d}{ds} U(s, y)$. But we have, for $y \leq Q(t)$,

$$U'(1-2\delta, \sqrt{y}) = \zeta'(1-2\delta) - \frac{y^{\delta}}{4\delta}\log y + \frac{y^{\delta}}{4\delta^2} + o(y^{-\frac{1}{2}}\log^c t).$$

Hence we have, for $y \leq Q(t)$,

$K(1-2\delta, 1; y)$

$$= \frac{1}{2\delta} y^{2\delta}\zeta(1+2\delta) + (\gamma + \log y)\zeta(1-2\delta) + \zeta'(1-2\delta) + o(y^{-\frac{1}{2}}(\log t)^{c}).$$

On the other hand, noting (4.1.14), we see that (4.1.13) gives

$$K(\sigma + \mathrm{it}; \sigma - \mathrm{it}; y) = |\zeta(\sigma + \mathrm{it}|^2 + o((\log t)^{-cB^3})$$

119 on the same condition as that for (4.1.13). Inserting these into (4.1.10), we obtain

$$\begin{aligned} H_1(y) &= (2\delta)^{-1}\zeta(1+2\delta)|\zeta(1+\delta+\mathrm{it}|^2 y^{2\delta} \\ &\quad - \zeta(1-2\delta) \sum_{mn \le \sqrt{y}} m^{-1+\delta+\mathrm{it}} n^{-1+\delta-\mathrm{it}} \log mn + 0(E(t)) \end{aligned}$$

for $Q(t) \ge y \ge \exp(BY(t))$. But this sum over m, n admits the similar decomposition as (4.1.11), and again by virtue of LEMMA 17 we can readily estimate it to be $0(E(t))$ Hence we have

$$H_1(y) = (2\delta)^{-1}\zeta(1+2\delta)|\zeta(1+\delta+it)|^2 y^{2\delta} + 0(E(t))$$

for $Q(t) \ge y \ge \exp(BY(t))$ Then by (4.1.19), we obtain the assertion of the lemma.

Also we shall need

Lemma 19. *Let*

$$I_D(x) = \sum_{\substack{n<x \\ (n,D)=1}} f(n),$$

and

$$F_p = \sum_{m=0}^{\infty} f(p^m)p^{-m}.$$

Then we have

$$I_D(x) = \mathscr{F} \times \prod_{p|D} F_p^{-1}(1 + 0(E(t)))$$

provided

$$\log \times >> \log D, Q(t) \ge \times \ge \exp(BY(t)).$$

120 This corresponds to the condition C_3) of § 1.4. The proof is quite similar to that of the preceding lemma, so we omit it. We should remark, however, that at a point in the proof, we require the following observation:

$$F_p = \left(1 - \frac{1}{p}\right)^{-1}\left(1 - \frac{1}{p^{1+2\delta}}\right)^{-1} | - \frac{1}{p^{1+\delta+\mathrm{it}}}|^{-2}\left(1 - \frac{1}{p^{2(1+\delta)}}\right)$$
$$> \left(1 - \frac{1}{p^{2(1+\delta)}}\right)^{-1},$$

whence

$$F_p - 1 > p^{-2(1+\delta)}.$$

This corresponds precisely to (C_2) of § 1.4.

Now we consider the Selberg sieve for the multiplicative function f:

$$\sum_{n \le N} f(n)\left(\sum_{\substack{d|n \\ d<R}} \Theta_d\right)^2 \quad (\Theta_1 = 1).$$

By LEMMA 19 and the general theory developed in § 1.4, we see that the optimal choice of Θ_d is given by

$$\Theta_d = \mu(d)\frac{G_d(R/d)}{G_1(R)} \prod_{p/d} F_p, \tag{4.1.15}$$

where

$$G_d(y) = \sum_{\substack{r \le y \\ (r,d)=1}} \mu^2(r) \prod_{p|r} (F_p - 1).$$

And this yields 121

$$\sum_{n \le N} f(n)\left(\sum_{d|n} \Theta_d\right)^2 = \mathscr{F} G_1(R)^{-1} N(1 + o(E(t))) \tag{4.1.16}$$

provided $Q(t) \ge N \ge R^{40} \ge \exp(BY(t))$.

On the other hand, the sieve-effect of (4.1.15) is embodied in the assertion that

$$G_1(R) \leq (2\delta)^{-1}\mathscr{F}(1 + o(E(t))) \tag{4.1.17}$$

provided $Q(t) \geq R \geq \exp(BY(t))$; this follows immediately from LEMMA 18, if we note that

$$G_1(R) \geq R^{-2\delta} \sum_{n<R} f(n) n^{-1+2\delta}.$$

Having these preparations at our hands, we can now proceed to our elementary proof of Vinogradov's zero-free region.

So, let us assume that $\zeta(s)$ takes a small value at $s = 1 - \delta +$ it, or, more precisely, the ineqality

$$|\zeta(1 - \delta + \text{ it } | \geq (\log t)^{-A} \tag{4.1.18}$$

holds for a δ satisfying (4.1.7); the value of A is to be fixed later but, for a while, let us take it for a large parameter.

Using $\Lambda_d^{(1)}$ and Θ_d which are defined in THEOREM 4 and at (4.1.15), respectively, we put

$$\omega_d = \sum_{[d_1,d_2]=d} \Theta_{d_1} \Lambda_{d_2}^{(1)},$$

so that, for all n,

$$\sum_{d|n} \omega_d = \left(\sum_{d|n} \Theta_d\right)\left(\sum_{d|n} \Lambda_d^{(1)}\right); \tag{4.1.19}$$

122 in particular, we have

$$\sum_{d|n} \omega_d = \begin{cases} 1 \text{ if } n = 1 \\ 0 \text{ if } n \leq z. \end{cases} \tag{4.1.20}$$

Further, we set

$$\vartheta = 1, z = \exp(40AY(t)), R = \exp(AY(t)), x = \exp(100AY(t))$$

where ϑ, z occur in the definition of $\Lambda_d^{(1)}$.

Then we consider the sum

$$\begin{aligned} Z &= \sum_{n<x} \sigma(n, -\delta - \mathrm{it}) \left(\sum_{d|n} \omega_d \right) n^{-1+\delta-\mathrm{it}} \\ &= \sum_{d<z^2R} \omega_d d^{-1+\delta-\mathrm{it}} \sum_{n<\times/d} \sigma(dn, -\delta - \mathrm{it}) n^{-1+\delta-\mathrm{it}}. \end{aligned}$$

This inner-sum can be readily estimated to be

$$0_A \left\{ (\log t)^c |\zeta(1 - \delta + \ \mathrm{it}\ )| + (\log t)^{-cA^3} \right\}$$

by appealing to LEMMA 17. Thus, if A is sufficiently large, the assumption (4.1.18) implies $Z = o(1)$, whence recalling (4.1.20), we have

$$\frac{1}{2} < \sum_{z \leq n < \times} f(n)^{\frac{1}{2}} |\sum_{d|n} \omega_d| n^{-1+\delta}.$$

Hence, by (4.1.19) and Schwarz's inequality, we get **123**

$$1 <<_A \times^{2\delta} \sum_{z \leq n < \times} f(n) \left(\sum_{d|n} \Theta_d \right)^2 n^{-1} \sum_{n=1}^{\infty} \left(\sum_{d|n} \Lambda_d^{(1)} \right)^2 n^{-\xi},$$

where $\xi = 1 + Y(t)^{-1}$. But, by virtue of THEOREM 4, $k = 1$, with its elementary account given above, the last infinite sum is $0_A(1)$. Then (4.1.16) gives

$$1 <<_A \mathscr{F} \times^{2\delta} Y(t) G_1(R)^{-1} \{1 + (\log t)^c |\zeta(1 - \delta + \mathrm{it}\ )|\}.$$

Hence noting (4.1.17), we infer that for an appropriately chosen A the assumption (4.1.18) implies

$$1 << \delta \times^{2\delta} Y(t),$$

which is apparently equivalent to

$$Y(t)^{-1} << \delta.$$

This and (4.1.1) give rise to the assertion:

Theorem 11. *The estimate*

$$\frac{\zeta'}{\zeta}(\sigma+it) = o((\log t)^c) \textit{for } \sigma > 1 - \frac{c}{(\log t)^{2/3}(\log\log t)^{1/3}} (t > 3)$$

is obtainable without using the theory of functions.

4.2 The Deuring-Heilbronn Phenomenon

Now let us apply similar considerations to Dirichlet's L-functions.

124 The important point in the study of the distribution of zeros of $L(s,\chi)$ is that it is required to have results which hold uniformly for varying χ. This raises difficult problems, and the incompleteness of our knowledge on $L(s,\chi)$ is reflected in the fact that the following statement, the theorem of Page, Landau and Siegel, is the best zero-free region for $L(s,\chi)$ known at present.

Let us denote by $Z(T)$ the set of all zeros of all $L(s,\chi)$ for primitive χ (mod q), $q \leq T$, which are in the region $|t| \leq T, 0 < \sigma < 1$. Then, except for at most one element β_1 of $Z(T)$, we have, for all $\rho \epsilon Z(T)$,

$$Re\rho < 1 - \frac{c_0}{\log T} \tag{4.2.1}$$

where $c_0 > 0$ is effectively computable. This exceptional zero $\beta_1 = \beta_1(T)$ which may also be called T-exceptional, if exists, is real and simple, and comes from $L(s,\chi_1)$ for a unique real primitive character χ_1. Further, for any fixed $\epsilon > 0$, there exists a $c(\epsilon) > 0$ such that

$$\beta < 1 - c(\epsilon)T^{-\epsilon}. \tag{4.2.2}$$

(4.2.1) is due to Page and Landau, and (4.2.2) to Siegel. The non-existence of such exceptional zeros has never been proved, and indeed this seems to be one of the most difficult problems in analytic number
125 theory. It can be shown, however, that if β_1 ever exists, then a strange phenomenon occurs among other elements of $Z(T)$. This was discovered by Deuring and Heilbronn in their effort to determine the asymptotic behaviour of class numbers of imaginary quadratic fields. Afterwards, Linnik succeeded in obtaining a quantitative version of their finding, which he called the Deuring-Heilbronn phenomenon, and runs as follows.

There exists an effectively computable constant $c_1 > 0$ such that for all $\rho \epsilon Z(T), \rho \neq \beta_1$, we have

$$Re\rho < 1 - \frac{c_1}{\log T}\log\left(\frac{c_0 e}{(1-\beta_1)\log T}\right). \tag{4.2.3}$$

It seems worth remarking that this implies (4.2.2). We note that, by an obvious reason, we may assume that there is a zero $\beta_\epsilon + i_{\gamma_\epsilon}$ of an $L(s, \chi_\epsilon), \chi_\epsilon(\mod q_\epsilon)$, such that $\beta_\epsilon > 1 - \epsilon$. Then let us take T so large that $q_\epsilon < T$, $|\gamma_\epsilon| < T, \beta_\epsilon < 1 - c_0(\log T)^{-1}$. This means that we may put $1 - \epsilon$ on the left side of (4.2.3), and we get (4.2.2).

It should be stressed that Linnik's result has the important feature that a sieve estimate, i.e. the Brun-Titchmarsh theorem, played a crucial rôle in its proof. In this context, perhaps it may not be surprising that we can show the following statement by means of Selberg's sieve method.

Theorem 12. *Page - Landau-Siegel's theorem and the Deuring-Heilbronn phenomenon can be proved without appealing to the theory of* **126** *functions.*

The proof is quite similar to that of THEOREM 11, save for the point that we have to be careful in obtaining an elementary lower bound of $L(s, \chi)$ for real χ in the vicinity of $s = 1$. Thus, to avoid unnecessary repetition, we shall show only the main steps of our argument.

We observe first that modifying the reasoning employed in the proof of (4.1.1) and using the well-known elementary result

$$L(1, \chi) > cq^{-\frac{1}{2}}(\log q)^{-1} \tag{4.2.4}$$

for real χ (mod q) it can be shown easily that, for all $\rho \epsilon Z(T)$, Re $\rho < 1 - T^{-2}$, provided T is sufficiently large as we assume hereafter.

Now let $1 - \delta + i\tau \epsilon Z(T)$ be a zero of $L(s, \psi)$; we may assume of course that $T^{-2} \leq \delta \leq 1/4$, say. We put

$$h(n) = |\sum_{d|n} \psi(d) d^{-\delta - i\tau}|^2.$$

Then we have

$$\sum_{n=1}^{\infty} h(n) n^{-s}$$

$$= \zeta(s)L(s+2\delta, \psi_0)L(s+\delta+i\tau, \psi)L(s+\delta-i\tau, \overline{\psi})L(2(s+\delta), \psi_0)^{-1},$$

127 where $\sigma > 1$ and $\psi_0 = \psi\overline{\psi}$. And we consider the Selberg sieve for the multiplicative function h:

$$\sum_{n<N} h(n) \left(\sum_{\substack{d|n \\ d<R}} \Theta'_d \right)^2 \qquad (\Theta'_1 = 1).$$

The optimal value of Θ'_d can be found by the argument of § 1.4, and we can infer that it yields the estimate

$$<<_B \delta N \tag{4.2.5}$$

for the last sum, if $N \geq R^4 \geq T^B$ with sufficiently large B. This is proved in much the same way as in the case of $\zeta(s)$; in fact, we need only (4.1.3) and its analogue for $L(s, \chi)$.

Then, as before, we define ω'_d by

$$\sum_{d|n} \omega'_d = \left(\sum_{d|n} \Theta'_d \right) \left(\sum_{d|n} \Lambda^{(1)}_d \right).$$

And we set $\vartheta = 1, z = T^{4A}, R = T^A, \times = T^{100A}$ with a large constant A. After some elementary estimations we have, for any non-principal χ (mod q), $q \leq T$,

$$\sum_{z \leq n < \times} \chi(n) \left(\sum_{d|n} \psi(d) d^{-\delta-\mathrm{it}} \right) \left(\sum_{\ell|n} \omega'_\ell \right) n^{-s} \tag{4.2.6}$$
$$= -1 + K(s, \chi)M(s, \chi) + o(T^{-cA}),$$

provided

$$Re(s) \geq 3/4, |Im(s)| \leq T. \tag{4.2.7}$$

128 Here

$$K(s, \chi) = L(s, \chi)L(s+\delta+i\tau), \chi\psi)$$

and

$$M(s,\chi) = \sum_{d<z^2R} \omega'_d \chi(d) d^{-s} \prod_{p|d} \left(1 + \frac{\psi(p)}{p^{\delta+i\tau}} - \frac{\chi\psi(p)}{p^{s+\delta+i\tau}}\right).$$

Now let $\rho = \beta + i\gamma$ be a zero of $K(s,\chi)$ in the region (4.2.7). Then setting $s = \rho$ in (4.2.6), we get

$$\sum_{z\leq n<\times} h(n)^{\frac{1}{2}} |\sum_{\ell|n} \omega'_\ell| n^{-\beta} \geq \frac{1}{2}.$$

Thus, by virtue of (4.2.5) and THEOREM 4, $k = 1$, we obtain

$$1 << \delta T^{20A(1-\beta)} \log T. \tag{4.2.8}$$

We now observe that either if ψ is complex, if ψ is real, non-principal and $\tau \neq 0$, or if ψ is real, non-principal, $\tau = 0$ and $1-\delta$ is a multiple zero of $L(s,\psi)$, then we have $K(1-\delta+i\tau,\psi) = 0$. Namely, in these cases, we may put $\rho = 1 - \delta + i\tau$ in (4.2.8), getting

$$\delta > \frac{c}{\log T}.$$

But if ψ is trivial we have already proved this, in fact much more, in the previous section. Thus, in the remaining case ψ is real, non-principal, and $1 - \delta$ is a simple zero of $L(s,\psi)$. Here we may assume obviously that $\delta \leq c'(\log T)^{-1}$ with a certain small constant $c' > 0$. Then (4.2.8) implies that all elements of $Z(T)$ except for $1-\delta$ are in the region 129

$$\sigma < 1 - \frac{c''}{\log T}$$

with an effectively computable $c'' > 0$. This proves (4.2.1). Finally, if $1-\delta$ is the T-exceptional zero then (4.2.8) implies the Deuring-Heilbronn phenomenon (4.2.3). This ends the proof of the theorem.

4.3 The Brun-Titchmarsh Theorem

The undesirable possibility of the existence of exceptional zeros causes much trouble in most applications of Page-Landau's theorem; thus many attempts to eliminate this defect in the theory of L-functions have been made from various directions. Among them is a sieve-theoretical one which, despite not much prospect of its success, seems to be worth describing explicitly because of its simplicity as well as the completeness of the hypothetical assertion deduced by it.

This idea rests on the plausibility of the estimate

$$\pi(\times; k, \ell) < (2-\eta)\frac{\times}{\varphi(k)\log(\times/k)}(k < \times^{\xi}) \tag{4.3.1}$$

with some effective constants $\eta, \xi > 0$. From this, we can deduce the non-existence of exceptional zeros. The proof is quite simple. In fact, let us assume that $L(s,\chi)$, real $_\chi$ (mod k), has a real zero $1-\delta$. Then we put

$$b(n) = \sum_{d|n} \chi^{(d)} d^{-\delta}$$

130 which is positive and multiplicative. We apply Selberg's sieve to $b(n)$:

$$I(N) = \sum_{n<N} b(n)\left(\sum_{\substack{d|n\\ d<r}} \lambda_d\right)^2 \quad (\lambda_1 = 1).$$

By an argument similar to (in fact, much simler than) that of the preceding section, we can infer that, for $N \geq R^4 \geq k^c$, the optimal chioce of λ_d gives $I(N) \ll \delta N$. On the other hand, we have

$$\begin{aligned} I(N) &> \sum_{R\leq p<N} b(p) \\ &\geq \pi(N) - \pi(R) - \sum_{\substack{p<N\\ \chi(p)=-1}} 1. \end{aligned}$$

Thus the hypothetical estimate (4.3.1) implies

$$\begin{aligned} I(N) &> (1-o(1))\frac{N}{\log N} - \frac{\varphi(k)}{2}(2-\eta)\frac{N}{\varphi(k)\log N/k} \\ &> \frac{\eta}{3}\frac{N}{\log N} \end{aligned}$$

provided $\log N \gg_{\xi,\eta} \log k$, whence $\delta \gg_{\xi,\eta} (\log k)^{-1}$ or the non-existence of the exceptional zeros as claimed.

But we have (1.2.15) which is close to (4.3.1). Because of this fact, considerable efforts have been spent to improve upon (1.2.15), and they are closely connected with the development of the sieve method itself. **131**
And the purpose of this section is to see how far the modern account of the linear sieve takes us on this matter.

Precisely speaking, we are going to improve upon (1.2.14) a generalised version of the Brun-Titchmarsh theorem. For this sake, we shall first show briefly a special instance of the hybridization of Iwaniec's linear sieve and the multiplicative large sieve.

Thus, let $(k, \ell) = 1$ and put

$$S(x, z, \chi) = \sum_{\substack{r \equiv \ell \pmod k \\ (r, P(z))=1 \\ r<x}} \chi(r) a_r,$$

where $\{a_r\}$ are arbitrary complex number, and as usual $P(z)$ is the product of all primes less then $z > 2$. And we consider the estimation of the expression

$$\sum_{\chi \epsilon \square} |S(x, z, \chi)|^2,$$

where

$$\square = \{X;\ \text{primitive} \pmod q, q < Q, (q, k) = 1\}.$$

But, by the duality principle (LEMMA 2), it suffices to deal with

$$D(x, z) = \sum_{\substack{r \equiv \ell \pmod k \\ (r, P(z))=1 \\ r<x}} |\sum_{\chi \epsilon \square} \chi(r) b_\chi|^2,$$

where $\{b_\chi\}$ are arbitrary complex numbers. The argument leading to **132**
LEMMA 10 gives

$$D(x, z) \leq \sum_K (-1)^{\omega(K)} \Theta_1(K) \sum_{d \epsilon K} D_d(x, z_1)$$
$$+ \sum_{\substack{I<K \\ \omega(K)\equiv 0 \pmod 2}} \Theta_1(KI) \sum_{\substack{p, p' \epsilon I \\ d \epsilon K}} D'_{dpp}(x, z_1),$$

where $2 < z_1 < z$ and

$$D_d(x, z_1) = \sum_{\substack{r\equiv\ell \pmod k\\ r\equiv\ell \pmod d\\ (r,p(z_1))=1\\ r<x}} |\sum_{\chi\in\square}\chi(r)b_\chi|^2;$$

the mode of the dissection of the interval $[z_1, z)$ is the one given at (3.4.10), and of course $\beta = 2$ in the defnition of Θ_1. We then follow closely the reasoning of § 3.4 up to (3.4.20); by using the notation

$$R_d(x,\chi) = \begin{cases} 0 & \text{if } (d,k) > 1\\ \sum\limits_{\substack{n\equiv\ell \pmod k\\ n\equiv 0 \pmod d\\ n<x}} \chi(n) - E(X)\chi(d)\frac{\varphi(q)}{dqk}x & \text{if } (d,k) = 1, \end{cases}$$

where χ is to mod q,we may express the result as

$$D(x,z) \leq \frac{xe^{-\gamma}}{\varphi(k)\log z}\left(\phi_1\left(\frac{\log y}{\log z}\right) + o(1)\right)\sum_{\chi\epsilon\square}|b_X|^2$$

$$+ \sum_{\chi,\psi\epsilon\square} b_\chi\bar{b}_\psi \left\{\sum_K \Theta_1(K)(-1)^{\omega(k)}\sum_{\substack{d\epsilon K\\ f<z^\tau}}\xi_f^{(1+\omega(K))}{}_{Rdf}(x,\chi\bar{\psi})\right.$$

$$\left. + \sum_{\substack{I<k\\ \omega(k)\equiv 0 \pmod 2}} \Theta_1(KI)\sum_{\substack{d\epsilon K\\ p,p'\epsilon I\\ f<z^\tau\\ f|P(z_1)}}\xi_f^{(1)}R_{dpp'f}^{(x,\chi\bar{\psi})}\right\}$$

133 provided $\log kQz \ll \log x \log \ll \log z$, where γ is the Euler constant, and the conventions are just the same as those in (3.4.20).

Next, to this we apply the smoothing device:

$$D(x,z) \leq \frac{1}{\eta}\int_X^{xe^\eta} D(w,z)\frac{dw}{w} \text{ for any } \eta > 0,$$

and then appeal to LEMMA 16. We get

$$D(x,z) \le \left\{ \frac{xe^{-\gamma}}{\varphi(k)\log z}\left(\phi_1\left(\frac{\log MN}{\log z}\right) + o(1)\right) + o((\log x)^2 E)\right\} \sum_{\chi\epsilon\square} |b_\chi|^2,$$

where $MN \ge z^2$, and

$$E = \sup_{w\le 2x} \sup_{\alpha,\beta} \sup_{\psi\epsilon\square} \sum_{\chi\epsilon\square} |\sum_{\substack{m<M\\ n<N\\ (mn,k)=1}} \alpha_m\beta_n R^{(1)}_{mn}(w,\chi\bar{\psi})|;$$

$$R^{(1)}_d(w,\chi) = \sum_{\substack{r\equiv\ell \pmod k\\ r\equiv 0 \pmod d\\ r<w}} \chi(r)\log\frac{w}{r} - E(\chi)\chi(d)\frac{\varphi(q)}{dqk} w,\chi \pmod q.$$

134

Here $\alpha = \{\alpha_m\}, \beta = \{\beta_m\}$ are, as before, variable vectors such that $|\alpha_m| \le 1, |\beta_n| \le 1$.

Hence, by LEMMA 2, we obtain

(4.3.2)
$$\sum_{\chi\epsilon\square} |S(x,z,\chi)|^2 \le \left\{\frac{xe^{-\gamma}}{\varphi(k)\log z}\phi_1\left(\frac{\log MN}{\log z}\right) + o((\log x)^2 E)\right\} \sum_{\substack{r\equiv\ell \pmod k\\ (r,p(z))=1\\ r<x}} |a_r|^2$$

provided $\log kQz \ll \log x \ll \log z$, and $MN \ge z^2$.

Now we proceed to the setimation of E. For this sake, we quote the following basic aids.

Lemma 20. *For any $\psi\epsilon\square$ and $T \ge 1$, we have*

(i) $$\sum_{\chi\epsilon\square} \sum_{\xi \pmod k} \int_{-T}^{T} |L\frac{1}{2} + it, \chi\bar{\psi}\xi|^4 dt \ll (kQ^2T)^{1+\epsilon},$$

and also, for any $H \ge 1$ and t,

(ii) $$\sum_{\chi\epsilon\square} \sum_{\xi \pmod k} |\sum_{h<H} \chi\bar{\psi}\xi(h) h^{-\frac{1}{2}-it}|^4 \ll ((|t|+1)kQ^2)^{i+\epsilon},$$

Lemma 21. *Let a_n be arbitrary complex numbers, and let $G = \sum_{n<N} |a_n|^2$. Then we have, for any $V > 0$,*

$$\left|\left\{(\chi,\xi);\chi\epsilon\Box,\xi \pmod k \text{ such that } |\sum_{n<N} a_n\chi\xi(n)| > V\right\}\right| \ll GNV^{-2} + G^3NV^{-6}(kQ^2)^{1-\epsilon}.$$

135 **Lemma 22.** *Let χ be non-principal $\pmod q$. Then we have*

$$\sum_{n<L}\chi(n)n^{it} \ll (|t|+1)L^{I-\epsilon},$$

provided $L > q^{3/8+\eta}$ with $\eta = \eta(\epsilon) > 0$.

To estimate E, it is sufficient to treat

$$E_\psi(A,B) = \sum_{\chi\epsilon\Box}|\sum_{\substack{A<m\leq 2A\\ B<n\leq 2B\\ (mn,k)=1}} \alpha_m\beta_n R^{(1)}_{mn}(w,\chi\bar{\psi})|,$$

where $\psi\epsilon\Box$; A, B are independent variables, and $\log ABkQ \ll \log x, w \ll x$, as we shall assume below. To this we shall apply two methods.

The *first method* rests on the expression

$$\sum_{\substack{A<m\leq 2A\\ B<n\leq 2B\\ (mn,k)=1}} \alpha_m\beta_n R^{(1)}_{mn}(w,\chi\bar{\psi}) = \frac{1}{2\pi i\varphi(k)} \sum_{\xi \pmod k} \bar{\xi}(\ell) \int\limits_{1/2-i\infty}^{1/2+i\infty} L(s,\chi\bar{\psi}\bar{\xi})A(s,\bar{\psi}\bar{\xi})B(s,\bar{\psi}\bar{\xi})\frac{w^s}{s^2}ds, \tag{4.3.3}$$

where

$$A(s,\chi) = \sum_{A<m\leq 2A}\chi(m)\alpha_m{m}^{-s}, B(s,\chi) = \sum_{B<n\leq 2B}\chi(n)\beta_n n^{-s}.$$

Thus using Hölder's inequality, we get

$$E_\psi(A,B) \ll \frac{x^{1/2+\epsilon}}{k}\left\{\int\sum|L(s,\chi\bar{\psi}\xi)|^4\frac{|ds|}{|s|^2}\right\}^{1/4}$$

$$\times \left\{ \int \sum |M(s,\chi\bar{\psi}\xi)|^4 \frac{|ds|}{|s|^2} \right\}^{1/4} \left\{ \int \sum |N(s,\chi\bar{\psi}\xi)|^2 \frac{|ds|}{|s|^2} \right\}^{1/2},$$

where integerals are along the line $\sigma = \dfrac{1}{2}$, and sums are over $\chi\epsilon\square$, $\xi \pmod k$. Then the multiplicative large sieve inequality and (i) of LEMMA 20 give **136**

$$E_\psi(A,B) \ll x^{1/2+\epsilon} \left(\frac{Q}{\sqrt{k}} \right)^{1/2} \left(A^2 + kQ^2 \right)^{1/4} (B + kQ^2)^{1/2}.$$

Hence, if we set

$$M = \left(\frac{x}{Q\sqrt{k}}\right)^{\frac{1}{3}-\eta}, N = M^2 \geq kQ^2$$

with a small fixed $\eta > 0$, we have

$$E \ll x^{1-\epsilon}/k.$$

Inserting this into (3.4.2), we obtain

$$\sum_{\chi\epsilon\square} |S\left(x, D^{\frac{1}{3}}, \chi\right)|^2 \leq \frac{(2+o(1))x}{\varphi(k)\log D} \sum_{\substack{r\equiv\ell \pmod k \\ (r,p(D^{1/3}))=1 \\ r<x}} |a_r|^2 \tag{4.3.4}$$

for

$$D = \left(\frac{x}{Q\sqrt{k}}\right)^{1-\epsilon}, kQ^2 \leq x^{\frac{1}{2}-\epsilon}.$$

The *second method* is more involved, and rests on the observation that, apart from a negligible error, the left side of (4.3.3) is equal to **137**

$$\frac{1}{2\pi i\varphi(k)} \sum_{\xi \pmod k} \bar{\xi}(\ell) \int_{1/2-\mathrm{i}x^c}^{1/2+\mathrm{i}x^c} H(s,\chi\bar{\psi}\xi)A(s,\chi\bar{\psi}\xi)B(s,\chi\bar{\psi}\xi)\frac{w^s}{s^2}ds,$$

where c is sufficiently large, and

$$H(s,\chi) = \sum_{h<H} \chi(h)h^{-2}, H = \frac{w}{AB}.$$

Now, for each pair of ψ and s with $Re(s) = 1/2$, let $K_{\psi,s}(U, V, W)$ be the number of charecters $\chi\xi, \chi\epsilon\Box, \xi$ (mod k), satisfying simultaneously

$$U < |H(s, \chi\bar{\psi}\bar{\xi})| \leq 2U, V < |A(s, \chi\bar{\psi}\bar{\xi})| \leq 2U, W < |B(s, \chi\bar{\psi}\bar{\xi})| \leq 2W.$$

Here, by an obious reason, we may assume that

$$|\log U|, |\log V|, |\log W| \ll \log x. \tag{4.3.5}$$

By virtue of (ii) of LEMMA 20 and LEMMA 21, we find readily

$$K_{\psi,s}(U, V, W) < x^{\epsilon} F$$

with

$$F = \min\left\{\frac{A + kQ^2}{V^2}, \frac{B + kQ^2}{W^2}, \frac{kQ^2|s|}{U^4}, \frac{A}{V^2} + \frac{kQ^2 A}{V^6}, \frac{B}{W^2} + \frac{kQ^2 B}{W^6}, \frac{H^2}{U^4} + \frac{kQ^2 H^2}{U^{12}}\right\};$$

in particular, we have

$$E_{\psi}(A, B) \ll \frac{x^{1/2+\epsilon}}{k} \int\limits_{1/2-ix^c}^{1/2+ix^c} \sup_{U,V,W} UVWF \frac{|ds|}{|s|^2},$$

where U, V, W are to satisfy (4.3.5).

138 Then it suffices to show

$$UVWF \leq |s| x^{1/2-\eta} \tag{4.3.6}$$

for some $\eta > 0$. Actually, we shall prove this on the assumption

$$A, B < \left(\frac{x}{(kQ^2)^{3/8}}\right)^{1/2-\delta}, kQ^2 < x^{\frac{9}{20}-\delta} \tag{4.3.7}$$

for some $\delta = \delta(\eta) > 0$. For this sake, we shall consider four cases separately: (i) $F \ll AV^{-2}$, BW^{-2}; (ii) $F \gg AV^{-2}$, BW^{-2}; (iii) $F \ll$

AV^{-2}, $F \ll BW^{-2}$; (iv) $F \gg AV^{-2}$, $F \ll BW^{-2}$. But, because of the symmetry, we do not need to treat the case (iv).

Now if (i) holds, then we have

$$UVWF \ll UVW \min\left(\frac{A}{V^2}, \frac{B}{W^2}\right) \ll U(AB)^{1/2}.$$

But, by virtue of (4.3.7) we may appeal to LEMMA 22 which gives $U \ll |s|L^{1/2}{}_X{}^{-\eta}$, whence we have (4.3.6). Before treating the case (ii), we remark that for any $\alpha_i > 0$ we have

$$\min\{\alpha_1, \alpha_2, \ldots, \alpha_r\} \leq \alpha_1^{e_1}\alpha_2^{e_2}\ldots\alpha_r^{e_r}$$

with any $e_i \geq 0$, $\sum\limits_{i=1}^{r} e_i = 1$. Thus, in case (ii), we have

$$\begin{aligned}
F &\ll \min\left\{\frac{kQ^2}{V^2}, \frac{kQ^2}{W^2}, \frac{AkQ^2}{V^6}, \frac{BkQ^2}{W^6}, \frac{H^2}{U^4}, \frac{kQ^2|s|}{U^4}\right\} \\
&\quad + \min\left\{\frac{kQ^2}{V^2}, \frac{kQ^2}{W^2}, \frac{AkQ^2}{V^6}, \frac{BkQ^2}{W^6}, \frac{H^2kQ^2}{U^{12}}, \frac{kQ^2|s|}{U^4}\right\} \\
&\ll |s|\Bigg[\left(\frac{kQ^2}{V^2}\right)^{\alpha}\left(\frac{kQ^2}{W^2}\right)^{\alpha}\left(\frac{AkQ^2}{V^6}\right)^{\beta}\left(\frac{BkQ^2}{W^6}\right)^{\beta}\left\{\min\left(\frac{H^2}{U^4}, \frac{kQ^2}{U^4}\right)\right\}^{\gamma} \\
&\quad + \min\Bigg\{\left(\frac{kQ^2}{V^2}\right)^{\alpha}\left(\frac{kQ^2}{W^2}\right)^{\alpha}\left(\frac{AkQ^2}{V^6}\right)^{\beta}\left(\frac{BkQ^2}{W^6}\right)^{\beta}\left(\frac{kQ^2}{U^4}\right)^{\gamma} \\
&\qquad \left(\frac{kQ^2}{V^2}\right)^{\alpha'}\left(\frac{kQ^2}{W^2}\right)^{\alpha'}\left(\frac{AkQ^2}{V^6}\right)^{\beta'}\left(\frac{BkQ^2}{W^6}\right)^{\beta'}\left(\frac{H^2kQ^2}{U^{12}}\right)\gamma'\Bigg\}\Bigg].
\end{aligned}$$

Here $\alpha, \beta, \gamma, \alpha', \beta', \gamma'$ are to be chosen in such a way that $UVWF$ is 139
bounded by a quantity not depending on U, V, W, and also $2\alpha + 2\beta, +\gamma = 1, 2\alpha' + 2\beta' + \gamma' = 1$. We should put, obviously, $\gamma = \frac{1}{4}, \gamma' = \frac{1}{12}$, and $2\alpha + 6\beta = 1, 2\alpha' + 6\beta' = 1$. So we find $\alpha = \frac{5}{16}, \beta = \frac{1}{16}, \alpha' = \frac{7}{16}, \beta' = \frac{1}{48}$. Inserting these into the last expression, we get

$$F \ll |s|(UVW)^{-1}kQ^2(AB)^{\frac{1}{16}}\left\{\min(1, H^{\frac{1}{2}}(kQ^2)^{-\frac{1}{4}}) + \min(1, H^{\frac{1}{6}}(AB)^{-\frac{1}{24}})\right\}$$

$$F \ll |s|(UVW)^{-1}kQ^2(AB)^{\frac{1}{16}}\left\{\left(H^{\frac{1}{2}}(kQ^2)^{-\frac{1}{4}}\right)^{\frac{1}{8}} + \left(H^{\frac{1}{6}}(AB)^{-\frac{1}{24}}\right)^{\frac{3}{10}}\right\}$$

$$\ll |s|(UVW)^{-1}\left\{x^{\frac{1}{16}}(kQ^2)^{\frac{31}{32}} + x^{\frac{1}{20}}kQ^2\right\}.$$

And this is, by (4.3.7),

$$\ll |s|(UVW)^{-1}x^{1/2-\eta},$$

140 whence we get (4.3.6) again. In much the same way, we can show that if (iii) holds, then

$$F \ll |s|(UVW)^{-1}\left\{x^{\frac{1}{8}}(kQ^2)^{\frac{7}{16}}N^{\frac{3}{8}} + x^{\frac{1}{12}}(kQ^2)^{\frac{1}{2}}N^{\frac{5}{12}}\right\},$$

which is, by (4.3.7),

$$\ll |s|(UVW)^{-1}\left(x^{\frac{5}{16}}(kQ^2)^{\frac{47}{128}} + x^{\frac{7}{24}}(kQ^2)^{\frac{27}{64}}\right)$$
$$\ll |s|(UVW)^{-1}x^{0.49}.$$

This amply implies (4.3.6).

This, we infer that we have

$$\sum_{X\epsilon\square}|S(x, D_1^{\frac{1}{3}}, \chi)|^2 \leq \frac{(2+o(1))x}{\varphi(k)\log D_1}\sum_{\substack{r\equiv\ell \pmod k\\ (r,P(D_1^{1/3}))=1\\ r<x}}|a_r|^2 \tag{4.3.8}$$

if

$$D_1 = \left(\frac{x}{(kQ^2)^{3/8}}\right)^{1-\epsilon}, kQ^2 < x^{\frac{9}{20}-\epsilon}.$$

Now we specialize the sequence $\{a_r\}$ by setting $a_r = 1$ if r is a prime, and $= 0$ otherwise. Introducing this into (4.3.4) and (4.3.8), and estimating trivially the contribution of primes less than D or D_1, we obtain

Theorem 13. *We have*

$$\sum_{\substack{q<Q\\(q,k)=1}}\ \sum_{\chi \pmod q}|\sum_{\substack{p\equiv\ell \pmod k\\ p<x}}\chi(p)|^2$$

$$\leq \begin{cases} \frac{(2+0(1))x}{\varphi(k)\log\frac{x}{{kQ^2}^{3/8}}}\pi(x;k,\ell) & \text{if } kQ^2 < x^{\frac{9}{20}-\epsilon}, \\ \frac{(2+0(1))x}{\varphi(k)\log\frac{x}{\sqrt{kQ}}}\pi(x;k,\ell) & \text{if } kQ^2 < x^{\frac{1}{2}-\epsilon} \end{cases}$$

141

Specifically,

$$\pi(x;k,\ell) \leq \begin{cases} \frac{(2+0(1))x}{\varphi(k)\log\frac{x}{k^{3/8}}} & \text{if } k < x^{\frac{9}{20}-\epsilon}, \\ \frac{(2+o(1))x}{\varphi(k)\log\frac{x}{\sqrt{k}}} & \text{if } k < x^{\frac{1}{2}-\epsilon}. \end{cases} \tag{4.3.9}$$

These are genuine improvement upon (1.2.14) *and* (1.2.15) *but far from* (4.3.1).

NOTES (IV)

A clear-cut proof of Vinogradov's result states in LEMMA 17 can be found in the text book [42].

Montgomery [[48], Chap 11] has proved, by analytical means, that if there is a zero of $\zeta(s)$ very near the line $\sigma = 1$ there are other (in fact many) zeros nearby; actually, he obtained a quantitative account of this phenomenon, from which the zero-free region of Vinogradov follows. On this matter see also Ramachandra [64].

For a proof of (4.1.6), see Wirsing [82], where much more than what 142
is required here is proved.

THEOREM 11 is due to Motohashi [61]. Apart from Selberg's sieve and THEOREM 4 an important ingredinent in his argument is Ramanujan's identity. It seems that Ingham [27] is the first who exploited the information about the zeros of $\zeta(s)$ from Ramanujan's indentitly. Balasubramanian and Ramachandra [1] discussed Ingham's idea in detail. In this context, it may be worth remarking that de la Vallee Poussin's inequality (4.1.2) corresponds to the following identity of the Ramanujan type: for Re(s)> 1,

$$\sum_{n=1}^{\infty} |\sigma(n,iu)|^4 n^{-s} = \zeta(s)^6\zeta(s+iu)^4\zeta(s-iu)^4\zeta(s+2iu)\zeta(s-2iu)G(s,u),$$

where u is real, and $G(s,u)$ is regular for $\mathrm{Re}(s) > 1/2$. From this, we can extract interesting information on the relation between the size of $\zeta(1+it)$ and the existence of zeros in the vicinity of the line $\sigma = 1$.

Linnik's proof [[47], II] of the Deuring-Heilbronn phenomenon (4.2.3) is formidable. Considerable simplifications were made by Knapowski [43]; in his argument, the power sum method of Turan is vital. Further simplifications were given by Montgomery [50]; he employed a special version due to himself of the power sum method. Later Jutila [40] and Motohashi [[55], I] worked out a conceptually much simpler
143 proof of (4.2.3) via the Selberg sieve. The argument developed in § 4.2 is quoted from Motohashi [57].

For an elementary proof of (4.2.4), see e.g. Gel'fond [18]. The elementary treatment of the basic theory of L-functions originates in Linnik's work [46]. Pintz made an extensive study on this matter.

Brun-Titchmarsh's theorem states, in its original form, that there exists an absolute constant $C > 0$, such that, for any $k < x$, we have

$$\pi(x; k, l) < C\frac{x}{\varphi(k)\log\frac{x}{k}}.$$

This was obtained as a particular application due to Titchmarsh of the really revolutionary idea of Brun, and shows clearly the advantage of his elementary method over analytical methods, for, the latter has never been able to yield such effective and uniform an estimate of $\pi(x; k, \ell)$ as this. For a detailed history of the Brun-Titchmarch theorem, see the relevant part of the text book [21].

The assertion that (4.3.1) implies the non-existence of exceptional zeros is due to Motohashi [59]; formerly, it was known only that, if (4.3.1) holds, then $L(s,\chi), \chi \pmod k$, does not vanish in the interval

$$1 - \frac{c}{\log k \log\log k} < s < 1.$$

144 Siebert [78] has extended the matter so that any effective improvements on the main-term in the limear seive applied to arithmetic progressions would yield a result similar to our assertion. It should be noted also that, as can be seen easily from our argument, in order to obtain our result it suffices to have (4.3.1) for all but $0(\varphi(k))$ residue classes $\pmod k$. Such statistical study of the Brun-Titchmarsh theorem was initiated by Hooley [24].

In this context, it is quite intersting that the estimate of the sort

$$\pi(x+h) - \pi(x) << \frac{h}{\log h}$$

which is closely related to the Brun-Titchmarsh theorem can yield an effective zero-free region for $\zeta(s)$. This was observed by balasubramanina and Ramachandra in the paper quoted above; in fact, they obtained

$$|\zeta(1+it) >> (\log(|t|+2))^{-3}$$

by using the above result on $\pi(x)$, and moreover, this was achieved without using de la Vallee Poussin's inequality unlike all other methods.

THEOREM 13 is due to Motohashi [[54]. IV], which is a large sieve extension of the results of Iwaniec [32] who showed (4.3.9); the estimation of E is done by following the relevant part of Iwaniec's work.

(i) of LEMMA 20 can be proved by employing an idea of Ramachandra [63] (see also [6]), and (ii) is an easy consequence of (i). LEMMA 21 is a simplified version of the large value theorem of Huxley [26]; for its quick and elegant proof, see Jutila [39], LEMMA 22 is due to Burgess [10]. **145**

In the paper quoted above, Iwaniec proved also the estimate:

$$\pi(x;k,\ell) \leq \frac{(2+0(1))}{\varphi(k)\log(x^{3/2}/k^{7/4})} x \text{ for } k \leq x^{2/3}.$$

This remarkable result was obtained by an ingenious combination of his linear sieve and a special instance of the dispersion method of Linnik. These results of Iwaniec are substantial improvements upon those due to Motohasi [52] who using the Selberg sieve proved (4.3.9) but, for smaller values of k.

Chapter 5

Zero-Density Theorems

THE MAIN OBJECT of the present chapter is to prove in detail a well-known zero density estimate for the Riemann zeta-function which will be required in the next chapter. The argument has nothing to do with the sieve methods developed in PART I, but the result will be combined with the linear sieve to produce a deep consequence on the difference between consective primes. **146**

We shall prove also a zero-density estimate of the Linnik type for Dirichlet's L-functions; there the hybrid dual sieve for intervals will play an impotant rôle, and we shall have a nice instance of a fruitful unification of sieve methods and analytic methods.

5.1 A Zero-Density Estimate for $\zeta(s)$

As usual, we denote by $N(\alpha, T)$ the number of zeros $\rho = \beta + i\gamma$ of $\zeta(s)$ satisfying $\alpha \le \beta \le 1, |\gamma| \le T$. It is expedient to consider the estimate of $N(\alpha, T) - N(\alpha, T/2)$ instead of $N(\alpha, T)$; so we assume henceforth that ρ satisfies

(5.1.1) $$\alpha \le \beta \le 1, \frac{T}{2} \le |\gamma| \le T,$$

and that T is sufficiently large.

We divide our discussion into three parts according to the value of α **147**

Case 1. $0 \leq \alpha \leq 3/4$.

For a while, we assume further that

$$\frac{1}{2} + (\log T)^{-1} \leq \alpha \leq \frac{3}{4}. \tag{5.1.2}$$

Let x, y be two parameters such that $2 \leq x \leq y, \log xy = 0(\log T)$, and put

$$M(s) = \sum_{n<x} \mu(n) n^{-s}.$$

Then considering the Mellin integral

$$\frac{1}{2\pi i} \int_{2-i\infty}^{2+i\infty} \zeta(\rho + w) M(\rho + w) r(w) y^w dw$$

we get

$$e^{-1/y} + \sum_{n \geq x} a(n) n^{-\rho} e^{-n/y} = M(1) y^{1-\rho} r(1-\rho)$$
$$+ \frac{1}{2\pi} \int_{-\infty}^{\infty} \zeta\left(\frac{1}{2} + i(\gamma + u)\right) M\left(\frac{1}{2} + i(\gamma + u)\right) y^{\frac{1}{2}-\beta+iu} r\left(\frac{1}{2} - \beta + iu\right) du,$$

where we should observe (5.1.2), and $a(n) = \sum_{\substack{d|n \\ d<x}} \mu(d)$. On the left side, we may truncte the sum at $y(\log T)^2$ with a negligible error; on the right side, the first term is negligible while the integral may be truncted at $u = \pm(\log T)^2$. Thus we have

$$1 << |\sum_{x \leq n \leq y(\log T^2)} A^{A(n) n^{-\rho e - n/y}}| \tag{5.1.3}$$
$$+ y^{\frac{1}{2}-\alpha} \log T \int_{-(\log T)^2}^{(\log T)^2} |\zeta\left(\frac{1}{2} + i(\gamma + u)\right) M\left(\frac{1}{2} + i(\gamma + u)\right)| du.$$

Next, from each each horizontal strip

$$2n + \nu \leq t < 2n + \nu + 1 (\nu = 0, ; n = 0, \pm 1, \pm 2, \ldots),$$

we pick up a zero of $\zeta(S)$ satisfying (5.1.1), and let $\mathscr{R}_\nu$ be the resulting set of zeros. Then we have

$$N(\alpha, T) - N\left(\alpha, \frac{T}{2}\right) \ll (|\mathscr{R}_0| + |\mathscr{R}_1|) \log T \tag{5.1.4}$$

since, as is well-known, $N(0, u+1) - N(0, u) \ll \log(u+2)$. Hence it suffices to estimate $|\mathscr{R}_\nu|$. By (5.1.3) and Hölder's inequality, we have

(5.1.5)

$$|\mathscr{R}_\nu| \ll |\mathscr{R}_\nu|^{\frac{1}{2}} \left\{ \sum_{\rho \epsilon \mathscr{R}_\nu} | \sum_{x \leq n \leq y (\log T^2)} a(n) n^\rho e^{-n/y} |^2 \right\}^{\frac{1}{2}} + y^{\frac{1}{2} - \alpha} \log T |\mathscr{R}_\nu|^{\frac{1}{4}}$$
$$\int\limits_{-(\log T)^2}^{(\log T)^2} \left\{ \sum_{\rho \epsilon \mathscr{R}_\nu} |\zeta \left(\frac{1}{2} + i(\gamma + u) \right)^4 | \right\}^{\frac{1}{4}} \left\{ \sum_{\rho \epsilon \mathscr{R}_\nu} |M \left(\frac{1}{2} + i(\gamma + u) \right)|^2 \right\}^{\frac{1}{2}} du.$$

To proceed further, we require discrete mean-value theorems for the Riemann zeta-function and Dirichlet polynomials:

Lemma 23. (i) *Let $\{t_r\}$ be a set of real numbers such that $|t_r| \leq T$ and $|t_r - t_{r'}| \geq \delta > 0$ for $r \neq r'$. Then we have*

$$\sum_r |\zeta \left(\frac{1}{2} + it_r \right)|^4 << (\delta^{-1} + \log T) T (\log T)^4.$$

(ii) *Let $\{s_r\}$ be set of complex numbers such that $Re(s_r) \geq 0, |Im(s_r)| \leq T$, and $|Im(s_r) - Im(s'_r)| \geq \Delta > 0$ for $r \neq r'$. Then we have, for arbitrary complex numbers $\{a_m\}$,* **149**

$$\sum_r | \sum_{M < m \leq 2M} a_m m^{-s_r} |^2 \ll (\delta^{-1} + \log M)(T + M) \sum_{M < m \leq 2M} |a_m|^2.$$

(iii) *Let* $\{s_r\}$ *be as above. Then we have also*

$$\sum_r |\sum_{M<m\leq 2M} a_m m^{-s_r}|^2 << (\Delta^{-1})(M + |\{s_r\}|T^{\frac{1}{2}} \log T) \sum_{M<m\leq 2M} |a_m|^2.$$

Applying (i) and (ii) of this lemma to (5.1.5), we get readily

$$|\mathscr{R}_\nu| << |\mathscr{R}_\nu|^{\frac{1}{2}} \left(y^{2(1-\alpha)} + Tx^{1-2\alpha}\right)^{\frac{1}{2}} \log^c T$$
$$+ |\mathscr{R}_\nu|^{\frac{1}{4}} y^{\frac{1}{2}-\alpha} T^{\frac{1}{4}} (x+T)^{\frac{1}{2}} \log^c T.$$

Namely, we have

$$|\mathscr{R}_\nu| << \left(y^{2(1-\alpha)} + Tx^{1-2\alpha} + T^{\frac{2}{3}} y^{\frac{2}{4}(1-2\alpha)}\right) \log^c T;$$

in this, we set

$$x = T, y = T^{\frac{2}{2(2-\alpha)}},$$

getting

$$|\mathscr{R}_\nu| \ll T^{\frac{3}{2-\alpha}(1-\alpha)} \log^c T.$$

150 This and (5.1.4) yield

$$N(\alpha, T) \ll \frac{3}{T^{2-\alpha}}(1-\alpha) \log^c T \tag{5.1.6}$$

in which we may now neglect (5.1.2) by an obvious reason.

Case 2. $5/6 \leq \alpha \leq 1$.

We proceed as before, up to (5.1.3): But this time, we set there

$$x = T^{\frac{6\alpha-5}{12(3\alpha-1)}} + \eta, y = T^{\frac{3}{4(3\alpha-1)}+2}\eta \tag{5.1.7}$$

with a small fixed $\eta > 0$. Then recalling the well-known estimate $\zeta(\frac{1}{2} + it) << (|t|+1)^{1/6} \log(|t|+2)$, we see that the second term on the right side (5.1.3) is $(T^{-\eta/10})$. Hence we have

$$1 << |\sum_{x\leq n\leq y(\log T)^2} a(n) n^{-\rho} e^{-n/y}|.$$

From this, we can infer that there exist an N, $x \leq N \leq y(\log T)^2$ and a subset $\mathscr{R}_\nu^{(1)}$ of $\mathscr{R}_\nu$ such that

$$|\mathscr{R}_\nu| <<_\eta |\mathscr{R}_\nu^{(1)}| \log T \tag{5.1.8}$$

and for all $\rho \epsilon \mathscr{R}_\nu^{(1)}$

$$(\log T)^{-1} <<_\eta |\sum_{N \leq n \leq 2N} a(n) n^{-\rho} e^{-n/y}|. \tag{5.1.9}$$

To proceed further, we require a large-value theorem of Dirichalet polynomials.

Lemma 24. *Let $\{s_r\}$ satisfy the condition given in (ii) of the previous lemm. And let us assume that there is a $V > 0$, such that for all r* 151

$$V < |\sum_{M < m \leq 2M} a_m m^{-s_r}.$$

Then we have

$$|\{s_r\}| << (1 + \delta^{-1})^3 M(V^{-2} G + V^{-6} T G^3 \log^2 T)$$

where

$$G = |\sum_{M < m \leq 2M} |a_m|^2.$$

Before applying this to our situation, we choose the integer k such that for the N of (5.1.9)

$$N^{k-1} < T^{\frac{1}{3\alpha - 1}} \leq N^k; \tag{5.1.10}$$

obviously, we have $2 \leq k <<_\eta 1$. Then we raise the both sides of (5.1.9) to $2k$-th power, and use LEMMA 24, getting

$$|\mathscr{R}_\nu^{(1)}| <<_\eta \left(N^{2k(1-\alpha)} + (T N^{2k(2-3\alpha)}\right) \log^{c(\eta)} T.$$

Now if $k = 2$, then we have, by (5.1.7) and (5.1.10),

$$T^{\frac{2}{3\alpha - 1}} \leq N^4 < y^4 (\log T)^8 \leq T^{\frac{3}{3\alpha - 1} + 9\eta},$$

and if $k \geq 3$, then (5.1.10) implies

$$T^{\frac{2}{3\alpha-1}} \leq N^{2k} < T^{\frac{2}{3\alpha-1}\left(1+\frac{1}{k-1}\right)} \leq T^{\frac{3}{3\alpha-1}}.$$

Hence we get 152

$$\mathscr{R}_\nu^{(1)} <<_\eta T^{\left(\frac{3}{3\alpha-1}+9\eta\right)(1-\alpha)} \log^{c(\eta)} T.$$

Thus, by (5.1.4) and (5.1.8), we obtain

$$N(\alpha, T) << T^{\left(\frac{3}{3\alpha-1}+\epsilon\right)(1-\alpha)} \log^{c(\epsilon)} T. \tag{5.1.11}$$

Case 3. $3/4 \leq \alpha \leq 5/6$.

In this case, we require a zero-detecting method different from the above. We note first that, by an elementary consideration, we can confine ourselves to those zeros $\rho = \beta + i\gamma$ of $\zeta(s)$ that satisfy (5.1.1) and

$$\zeta(s) \neq 0 \text{ for } \alpha + \eta^4 \leq \sigma \leq 1, |t - \gamma| < \log^2 T, \tag{5.1.12}$$

where η is a small positive parametr to be fixed later. We pick up such a zero which lies in one of the horizontal strips $2n + \nu \leq t < 2n + \nu + 1(\nu = 0, 1; n = 0, \pm 1, \pm 2, \ldots)$, and let $\tilde{\mathscr{R}}_\nu$ be the obtained set of zeros. Then we have

$$N(\alpha, T) - N(\alpha, T/2) <<_\eta (|\tilde{\mathscr{R}}_0| + |\tilde{\mathscr{R}}_1|) \log^{c(\eta)} T.$$

Also, we remark that (5.1.12) implies

$$\zeta(s)^{-1} = 0_\eta(T^{\eta^5}) \tag{5.1.13}$$

in the region $\alpha + 2\eta^4 \leq \sigma, |t - \gamma| < \dfrac{1}{2}(\log T)^2$.

Then we put

$$N(s) = \sum_{n=1}^{\infty} \frac{\mu(n)}{n^s} \exp\left(-\left(\frac{n}{x}\right)^2\right)$$

153 where x is to be chosen later, and for a while, we assume only $2 \leq x \leq T$.

We have, for $Re(s) \leq 3/2$,

$$N(s) = \frac{1}{2\pi i}\int_{2-i\infty}^{2+i\infty} \frac{1}{\zeta(W)} r\left(\frac{w-s}{2}\right) x^{w-s} dw.$$

In this formula, let us confine s to the region $\sigma \leq \alpha + \eta^4, |t-\gamma| \leq \frac{1}{8}(\log T)^2$, and then shift the line of integration to the broken line $\ni w = u + iv : u = 2, |v-\gamma| \geq \frac{1}{4}(\log T)^2; \alpha + 2\eta^4 \leq u \leq 2, v = \gamma \pm \frac{1}{4}(\log T)^2$; $u = \alpha + 2\eta^4, |v-\gamma| \leq \frac{1}{4}(\log T)^2$. Then, by (5.1.13), we get

(5.1.14) $$N(s) = 0_\eta\left(x^{\alpha+2\eta^4-\sigma} T^{\eta 5}\right).$$

Next, we consider the function

$$N(s)\zeta(s) = \sum_{n=1}^{\infty} b(n) n^{-s} (\sigma > 1)$$

where $$b(n) = \sum_{d|n} \mu(d) \exp\left(-\left(\frac{d}{x}\right)^2\right).$$

As before, we consider the expression

(5.1.15) $$\sum_{n=1}^{\infty} b(n) n^{-\rho} e^{-n/y} = \frac{1}{2\pi i}\int_{2-i\infty}^{2+i\infty} \zeta(s) N(s) r(s-\rho) y^{s-\rho} ds$$

where ρ satisfies (5.1.12). Again, we shift the line of integration to the broken line $\ni s : \sigma = 2, |t-\gamma| \geq \dfrac{1}{8}(\log T)^2; \alpha + \eta^4 - \eta^2 \leq \sigma \leq 2, t = \gamma \pm \dfrac{1}{8}(\log T)^2; \sigma = \alpha + \eta^4 - \eta^2, |t-\gamma| \leq \dfrac{1}{8}(\log T)^2$.

Then using (5.1.14), we see that the last integral is

$$0_\eta\left(y^{\eta^4-\eta^2} x^{\eta^4+\eta^2} T^{\eta^5} C_p(T)\right),$$

where $$C_p(T) = \max_{|t-\gamma| \leq \frac{1}{8}(\log T)^2} |\zeta(\alpha + \eta^4 - \eta^2 + it)|.$$

But by a convexity argument we can show that

$$C_p(T) \ll_\eta T^{\frac{\eta^2}{2}+\eta^5}, \tag{5.1.16}$$

which implies that the right side of (5.1.15) is $0_\eta(1)$ if we set

$$y = T^{\frac{1}{2}+\eta}, xT^{\eta^4}. \tag{5.1.17}$$

Thus we put these values of x, y in (5.1.15); we can truncate the sum on the left side at $y(\log T)^2$, and noting that $b(1) = 1 + 0(x^{-2})$ and $b(n) = 0((n/x)^2)$ for $1 < n \leq x$, we can infer that

$$1 \ll_\eta |\sum_{x(\log T)^{-2} \leq n \leq y(\log T)^2} b(n) n^{-\rho_e - n/y}|$$

for all ρ which satisfy (5.1.12).

Then, as in CASE 2, we have an N, $x(\log T)^{-2} \leq N \leq y(\log T)^2$, for which there exists a subset $\tilde{\mathscr{R}_\nu}(1)$ of $\tilde{\mathscr{R}}_\nu$ such that $|\tilde{\mathscr{R}}_\nu(1)| \gg |\tilde{\mathscr{R}}_\nu|$ $(\log T)^{-1}$ and for all $\rho \epsilon \tilde{\mathscr{R}}_\nu(1)$

$$(\log T)^{-1} \ll_\eta |\sum_{N < n \leq 2N} b(n) n^{-\rho_e - n/y}|.$$

155 We raise the both sides of this to $2k - th$ power so that $N^{k-1} < T^{\frac{1}{3\alpha-1}} \leq N^k$. Obviously, we have $2 \leq k \ll_\eta 1$. If $k \geq 3$, we can argue as before, and get (5.1.11) for our present value of α. But, if $k = 2$, we have

$$T^{\frac{1}{3\alpha-1}} \leq N^2 \leq T^{1+2\eta} (\log T)^4$$

because of (5.1.17), and using LEMMA 24 we obtain

$$\begin{aligned} |\tilde{\mathscr{R}}_\nu(1)| &\ll_\eta (\log T)^c \left(T^{2(1+2\eta)(1-\alpha)} + T^{\frac{3}{3\alpha-1}(1-\alpha)}\right) \\ &\ll_\eta (\log T)^c T^{\left(\frac{3}{3\alpha-1}+4_\eta\right)(1-\alpha)} \end{aligned}$$

since $3/(3\alpha - 1) \geq 2$ for $\alpha \geq 5/6$. Thus we get (5.1.11) in CASE 3 as well.

Finally, taking into account the zero-free region of Vinogradov, we may summarize the above discussion as

Theorem 14. *For $0 \leq \alpha \leq 1$, we have*

$$N(\alpha, T) \ll T^{(\phi(\alpha)+\epsilon)(1-\alpha)}$$

where

$$\phi(\alpha) = \begin{cases} \frac{3}{3\alpha-1} & \text{if } 3/4 \leq \alpha \leq 1, \\ \frac{3}{2-\alpha} & \text{if } 0 \leq \alpha \leq 3/4. \end{cases}$$

The zero-density result which we shall require in the next chapter is not this theorem but rather the following consequence of it.

Lemma 25. *Let $\{a_n\}$ be arbitrary complex numbers such that $|a_n| \ll 1$, and put* 156

$$K(s) = \sum_{K<n\leq 2K} a_n n^{-s}.$$

Then we have, for $0 \leq \alpha \leq 1$,

$$\sum_{\substack{\rho \\ |\gamma|\leq T \\ \beta\geq\alpha}} |K(\rho)| \ll \begin{cases} \left(T^{\frac{6}{5}+\epsilon} K\right)^{1-\alpha} & \text{if } T \leq K \leq T^c, \\ \left(\frac{T^{\frac{16}{5}+\epsilon}}{K}\right)^{1-\alpha} & \text{if } T^{\frac{4}{5}} \leq K \leq T \end{cases}$$

where $\rho = \beta + i\gamma$ is a complex zero of $\zeta(s)$.

To prove this, let us choose a ρ among those in the rectangle $2n+\nu \leq t < 2n+\nu+1, \alpha \leq \sigma \leq 1, (\nu = 0, 1)$, for which $|K(\sigma)|$ is the greatest, and $\mathscr{R}'_\nu$ be the obtained set of zeros. Then we have obviously

$$\sum_{\substack{\rho \\ |\gamma|\leq T \\ \beta\geq\alpha}} |K(\rho)| \ll \log T \left\{ \sum_{\rho\epsilon\mathscr{R}'_0} + \sum_{\rho\epsilon\mathscr{R}'_1} \right\} |K(\rho)|.$$

We have, by Schwarz's inequality,

$$\sum_{\rho\epsilon\mathscr{R}'_\nu} |K(\rho)| \leq N(\alpha, T)^{\frac{1}{2}} \left\{ \sum_{\rho\epsilon\mathscr{R}'_\nu} |K(\rho)|^2 \right\}^{\frac{1}{2}}.$$

If $K \geq T$, then, by (ii) of LEMMA 23, we get

$$\sum_{\rho \epsilon \mathscr{R}'_{\nu}}, |K(\rho)|^2 \ll K^{2(1-\alpha)} \log^c T;$$

on the other hand, the last theorem implies

$$N(\alpha, T) \ll T^{\left(\frac{12}{5}+\epsilon\right)(1-\alpha)}, \tag{5.1.18}$$

157 whence the first assertion of the lemma. As for the second we consider three cases separately. Firstly, if $0 \leq \alpha \leq 3/4$, then, by the last theorem,

$$\begin{aligned}
\sum_{\rho \epsilon \mathscr{R}'_{\nu}} |K(\rho)| &\ll \left(TN(\alpha, T)K^{1-2\alpha}\right)^{\frac{1}{2}} \log^c T \\
&\ll K^{\alpha-1}\left(T^{\frac{3-4\alpha}{2(1-\alpha)}} T^{\frac{1}{T^{2(1-\alpha)}}+\frac{3}{2(2-\alpha)}+\epsilon}\right)^{(1-\alpha)} \log^c T \\
&\ll K^{\alpha-1}\left(T^{\frac{3-4\alpha}{2(1-\alpha)}+\frac{1}{T^{2(1-\alpha)}}+\frac{3}{2(2-\alpha)}+\epsilon}\right)^{(1-\alpha)} \log^c T \\
&\ll \left(T^{\frac{16}{5}+\epsilon}/K\right)^{1-\alpha} \log^c T,
\end{aligned}$$

since

$$2 + \frac{3}{2(2-\alpha)} \leq 16/5$$

if $0 \leq \alpha \leq 3/4$. Secondly, if $3/4 \leq \alpha \leq 11/12$, then again, by the last theorem,

$$\begin{aligned}
\sum_{\rho \epsilon \mathscr{R}'_{\nu}} |K(\rho)| &\ll K^{\alpha-1}\left(K^{\frac{3-4\alpha}{2(1-\alpha)}} T^{\frac{1}{2(1-\alpha)}+\frac{3}{2(3-\alpha)}+\epsilon}\right)^{(1-\alpha)} \log^c T \\
&\ll K^{\alpha-1}\left(T^{\frac{6-8\alpha}{5(1-\alpha)}+\frac{1}{2(1-\alpha)}+\frac{3}{2(3\alpha-1)}+\epsilon}\right)^{(1-\alpha)} \log^c T.
\end{aligned}$$

But, for $\alpha \leq 11/12$, we have

$$\frac{6-8\alpha}{5(1-\alpha)} + \frac{1}{2(1-\alpha)} + \frac{3}{2(3\alpha-1)} = \frac{8}{5} + \frac{1}{10(1-\alpha)} + \frac{3}{2(3\alpha-1)} \leq 16/5.$$

Finally, if $11/12 \leq \alpha \leq 1$, then we appeal to (iii) of LEMMA 23,
158 getting

$$\sum_{\rho \epsilon \mathscr{R}'_v} |K(\rho)| \ll (N(\alpha, T)(K + N(\alpha, T)T^{\frac{1}{2}})K^{1-2\alpha})^{\frac{1}{2}} \log^c T.$$

But, for $\alpha \geq 11/12$, we have by the last theorem

$$T^{\frac{1}{2}} N(\alpha, T) \ll T^{\frac{9}{14}+\epsilon} < K.$$

Hence the sum in question is

$$\ll \left(KT^{\frac{3}{2(3\alpha-1)}+\epsilon}\right)^{1-\alpha} \log^c T,$$

and noting that for $\alpha \geq 11/12, K \leq T$, we have

$$KT^{\frac{3}{2(3\alpha-1)}} \leq T^{\frac{16}{5}} K^{-1},$$

which ends the proof of the second assertion of the lemma.

Here we should note that in the statement of the last lemma we have neglected log-factors, because of Vinogradov's zero-free region.

5.2 A Zero-Density Estimate of the Linnik Type

Most of the estimates of $N(\alpha, T)$ can be extended to those of $N(\alpha, T\chi)$ the number of zeros of $L(s, \chi)$ in the rectangle $\alpha \leq \sigma \leq 1, |t| \leq T$. But they are of limited value, because the theory of Dirichlet L-functions is greatly hampered by the lack of a zero-free region comparable to that of Vinogradov for the Riceman zeta-function. For some important problems in prime number theory, however, this deficiency can be circumscribed by the combination of the Deuring- Helibronn phenomenon **159**
(4.2.3) and the zero-density estimate of the Linnik type:

(5.2.1) $$\sum_{\chi \pmod q} N(\alpha, T, \chi) \ll (qT)^{c(1-\alpha)} (0 \leq \alpha \leq 1),$$

which is especially strong near the line $\sigma = 1$.

We have seen already that the Selberg sieve for multiplicative functions is capable to yield a very simple proof of the Deuring-Heilbronn. In this section, we shall apply again a similar idea to L-functions, and show that the same holds for once-difficult zero-density estimates of the Linnik type.

Our main tool is inequality (1.2.10), $k = 1$, or more precisely, the following one more step hybridized version of it.

Lemma 26. *Let $S(\chi)$ be a set of complex numbers such that for any $s, s' \epsilon S(\chi)$ we have $Re(s) \geq 0, |Im(s)| \leq T$, and $|Im(s) - Im(s')| \geq \delta > 0$ if $s \neq s'$. Then we have, for any complex numbers $\{a_n\}$,*

$$\sum_{\substack{rq<Q\\(r,q)=1}} \frac{\mu^2(r)q}{\varphi(rq)} \sideset{}{^*}\sum_{\chi \pmod q} \sum_{s\epsilon S(\chi)} |\sum_{n\leq N} a_n \chi(n)\psi_r(n) n^{-s}|^2$$
$$\ll (\delta^{-1} + \log N) \sum_{n\leq N} (n + Q^2 T)|a_n|^2 \left(1 + \log\left(\frac{\log^2 N}{\log 2n}\right)\right).$$

160 *Here, ψ_r is, of course, the one the defined at* (1.2.11)*. Here-after, we shall take T for a sufficiently large variable, and for the sake of simiplicity, we assume, up to* (5.2.7)*, that all Dirichlet characters are non-principal and have conductors less than T.*

Now let us denote, by $\bar{N}(\alpha, T, \chi)$, the number of zeros of $L(s, \chi)$ in the rectangle $\alpha \leq \sigma \leq 1, |t| \leq T$, save for the T-exceptional zero (cf. § 4.2). We note first that because of Page-Landau's theorem and a reason to be disclosed later, we may assume that

$$1 - \eta \leq \alpha \leq 1 - \frac{c}{\log T} \tag{5.2.2}$$

with a fixed small $\eta > 0$. Then, for each $\nu = 0, 1$ we pick up a zero of $L(s, \chi)$ which lies in the above rectangle and also in one of the horizontal strips

$$\frac{2n + \nu}{\log T} \leq t < \frac{2n + \nu + 1}{\log T} (n = 0, \pm 1, \pm 2, \ldots),$$

and denote by $Z_\nu(\chi)$ the resulting set of zeros of $L(s, \chi)$. Here, we should quote the zero-density lemma: the number of zeros of $L(s, \chi)$ contained in the disk $|s - (1 + iu)| \leq 1 - \alpha$ is $0((1 - \alpha)\log T)$, if $-T \leq u \leq T$, and α satisfies (5.2.2).

Thus we have

$$\bar{N}(\alpha, T, \chi) \ll (1-\alpha)\log T(|Z_0(\chi)| + |Z_1(\chi)|).$$

Next, let us recall the formula (1.4.15); there, we set $f \equiv 1$ and $\xi = \Lambda^{(1)}$ (cf. THEOREM 4). Then LEMMA 5 implies that, for each square-free r, the function **161**

$$M_r(s,\chi;\Lambda^{(1)}) = \frac{1}{\varphi(r)}\sum_{d=1}^{\infty}\Lambda_d^{(1)}\mu((r,d))\frac{\chi(d)}{d^s}\prod_{\substack{p\nmid d\\ p|r}}\left(\left(1-\frac{\chi(p)}{p^s}\right)\left(1-\frac{1}{p}\right)^{-1}-1\right) \tag{5.2.3}$$

satisfies

$$L(s,\chi)M_r(s,\chi;\Lambda^{(1)}) = \sum_{n=1}^{\infty}\chi(n)\psi_r(n)\left(\sum_{d|n}\Lambda_d^{(1)}\right)n^{-s}. \tag{5.2.4}$$

We note also that (5.2.3) gives, for $0 \le \sigma \le 1$,

$$M_r(s,\chi;\Lambda^{(1)}) \ll z^{(1+\vartheta)(1-\sigma)+\epsilon}r^{-\sigma-1+\epsilon} \tag{5.2.5}$$

where ϑ, z are the parameters appearing in the definition of $\Lambda^{(1)}$.

Then we consider the Mellin integral

$$\frac{1}{2\pi i}\int_{2-i\infty}^{2+i\infty}L(s,\chi)M_r(s,\chi;\Lambda^{(1)})r(s-\rho)x^{s-\rho}ds$$

where $\rho\epsilon Z_\nu(\chi)$; in this, we set

$$\vartheta = \eta^2, r \le R = T^{\eta^2}, z = T^3R^2, x = T^{\frac{7}{2}+\eta} \tag{5.2.6}$$

with the same η as that of (5.2.2). Shifting the line of integration $\sigma = (\log T)^{-1}$, and noting (5.2.2) and (5.2.5), we see that this integral is $0(T^{-\eta/2})$. On the other hand, we have (1.3.11) and (5.2.4).

Hence, after some simple consideration, we get **162**

$$1 \ll |\sum_{z \le n \le x(\log T)^2} \chi(n)\psi_r(n)\left(\sum_{d|n} \Lambda_d^{(1)}\right) n^{\sigma_e - n/x}|$$

for any $\sigma \epsilon Z_\nu(\chi)$ and square - free $r \le R$.

Now, noting (1.2.12), we see that gives

$$\log R \sum_{1<q<T} \sideset{}{^*}\sum_{\chi \pmod q} |Z_\nu(\chi)| \ll \sum_{\substack{r \le R \\ 1<q<T \\ (r,q)=1}} \frac{\mu^2(r)q}{\varphi(rq)}$$

$$\sideset{}{^*}\sum_{\chi \pmod q} \sum_{\rho \epsilon Z_\nu(\chi)} | \sum_{z \le n \le x(\log T)}^{2} \chi(n)\psi_r(n)\left(\sum_{d|n} \Lambda_d^{(1)}\right) n^{-\rho} e^{-n/x}|^2.$$

Then, appealing to LEMMA 26, we infer that the right side is

$$\ll \log T \sum_{z \le n \le x(\log T)^2} \left(\sum_{d|n} \Lambda_d^{(1)}\right)^2 n^{1-\alpha}$$

$$\ll (x \log^2 T)^{2(1-\alpha)} \log T \sum_{n=1}^{\infty} \left(\sum_{d|n} \Lambda_d^{(1)}\right)^2 n^{-1-(\log T)^{-}1}$$

$$\ll (x \log^2 T)^{2(1-\alpha)} \log T;$$

the last line is due to THEOREM 4. Hence we have

$$\sum_{1<q<T} \sideset{}{^*}\sum_{\chi \pmod q} |Z_\nu(\chi)| \ll_\eta T^{(7+2\eta)(1-\alpha)},$$

which implies

$$\sum_{1<q<T} \sideset{}{^*}\sum_{\chi \pmod q} \bar{N}(\alpha, T, \chi) \ll_\eta T^{(7+3_\eta)(1-\alpha)}. \tag{5.2.7}$$

But, as is well-knowm, we have, for $0 \le \alpha \le 1$,

$$\sum_{q<Q} \sideset{}{^*}\sum_{\chi \pmod q} N(\alpha, T, \chi) \ll (Q^2 T)^{\frac{5}{2}(1-\alpha)} \log^c QT. \tag{5.2.8}$$

163

Thus, taking into account Vinogradov's zero-free region for $\zeta(s)$, we obtain, from (5.2.7) and (5.2.8),

Theorem 15. *We have, for* $0 \leq \alpha \leq 1$ *and* $T \geq 1$,

$$\sum_{q<T} \sideset{}{^*}\sum_{\chi \pmod q} N(\alpha, T, \chi) \ll T^{8(1-\alpha)}.$$

NOTES (V)

The zero-density method originates in the discovery made by Bohr and Landau [3] of evidence which supports statistically Riemann's hypothesis. But the actual emergence of the zero-density method as an indispensable tool for the study of the distribution of primes started, when Hoheisel [23] found that the estimate of the type

$$N(\alpha, T) \ll T^{\lambda(1-\alpha)} \log^c T (0 \leq \alpha \leq 1)$$

yields a result on the difference between consecutive primes which had never been obtained without assuming the Riemann hypothesis for $\zeta(s)$ or sometimes similar to it. Namely, Hoheisel found a way to avoid the Riemann hyothesis for the investigaton of the distribution of primes. Afterwards, the discovery of the zero-free for $\zeta(s)$ of Vinogradov's type 164
made it clear that the smaller λ in the above estimate of $N(\alpha, T)$ would yield the better results on prime numbers.

In this context, Huxley's result [25]:

$$N(\alpha, T) \ll \begin{cases} T^{\frac{3}{2-\alpha}(1-\alpha)} \log^c T & \text{if } 0 \leq \alpha \leq 3/4 \\ T^{\frac{3}{3\alpha-1}(1-\alpha)} \log^c T & \text{if } 3/4 \leq \alpha \leq 1 \end{cases}$$

is so far the best among various estimates of $N(\alpha, T)$, for it gives the smallest λ, i.e., $12/5$, ever obtained.

In Huxley's proof of this, a difficult estimate of $\zeta\left(\frac{1}{2} + \text{it}\ \right)$ due to Haneke was employed. The reason that we develped, for CASE 3, a zero-detecting method of Bombieri [5] is that we wanted to dispense with Haneke's result. This caused a slight decrease in the quality of the obtained estimate, THEOREM 14, of $N(\alpha, T)$ if compared with Huxley's, but, for the applications which we have in mind, this will give no difference.

(5.1.13) can be proved in just the same way as in Titchmarsh [[79], p. 77]; (5.1.16) can be proved similarly, but we need also the functional equation for $\zeta(s)$.

It should be remarked that, in § 5.1, we used twice the device of
165 raising a Dirichlet polynomial to a high power so as to let it take a form suitable for the application of LEMMA 24 - the large-value theorem of Huxley [25]; this nice idea is due to Jutila [37].

LEMMA 25 is due to Iwaniec and Jutila [34]; the weighted version of the zero-density estimates was first considered and used by Jutila [38].

What Hoheisel did for $\zeta(s)$ Linnik did for L-functions; namely, he found the way to avoid the extended Riemann hypothesis in the investigation of some important problems concerning primes in arithmetic progressions. This possibility was first realized in his famous work [[47], I] in which he proved a result similar to (5.2.1). But it should be stressed that it is Fogel's [12] who actually obtained the estimate of the type (5.2.1). In Fogel's argument, Turan's idea [80] is vital, and this is the same in Gallagher's important work [15] where an estimate similar to that of THEOREM 15 was first proved.

In Linnik's Turán's and Fogels' works, a sieve result, i.e. the Brun-Titchmarsh theorem occupies an important place; the same can be said about Gallagher's quoted above, for he used Bombieri-Davenport's theorem (1.2.13) which is apparently a large sieve extension of the Brun-Tichmarsh theorem. This sieve aspect of the theory is now made more explicit in our proof of THEOREM 15, for, as we have shown in § 1.2,
166 the pseudo-character ψ_r is directly connected with the Selberg siever for intervals.

LEMMA 26 can readily be proved by combining (1.2.10), $k = 1$, with the argument of Montgomery [[48], Chap. 7 and 8].

One should note that in our proof of THEOREM 15 Selberg's observation (5.2.4) is vital (cf. Montgomery [50]. For a more refinned treatment of the matter related to THEOREM 15, see Motohashi [53] and Jutila [40].

For the history of the zero-density method, we refer to Mongomery [48] and Richert [66].

Chapter 6

Prims in Short Intervals and Short Arithmetic Progressions

IN THIS FINAL chapter, we shall demonstrate that the sieve method 167
can actually detect prime numbers in some very difficult and important situations, if it is correctly combined with analytical means.

In the first section, we shall inuect Iwaniec's linear sieve into the study of primes in short intervals and prove a remarkable result which, in spite of much efforts, has never been attained by the sole use of analytical means. On the other hand, in the second section, we shall empoly Selberg's sieve to prove a deep result pertaining to the least prime in an arithmetic progression, and illustrate the versatility of this fundamental sieve method.

6.1 Existence of Primes in Short Intervals

As is well-Known, THEOREM 14 or rather (5.1.18) yields

$$\pi(x) - \pi(x - x^{\theta}) = (1 + 0(1))\frac{x^{\theta}}{\log x} \tag{6.1.1}$$

Whenever $\theta > 7/12$. This implies of course

$$p_{n+1} - p_n \ll p_n^{\frac{7}{12}+\epsilon}, \tag{6.1.2}$$

p_n being the n-th prime.

168 Our aim is to show that if we geve up an asymptotic estimate but ask for a positive lower bound for $\pi(x) - \pi(x - x^\theta)$, then the value of θ can be taken less then $7/12$, so that (6.1.2) can be improved.

Adopting the notations introduced in the second chapter, we may write

$$\pi(x) - \pi(x-h) = S\left(A, x^{\frac{1}{2}}\right),$$

where

$$A = \{n; x - h \le n < x\}, x^\theta (\theta < 1),$$

and Ω is the simplest one: $\Omega(p) \ni n$ implies $p|n$. Thus Buchstab's formula (2.1.1) gives, for any $2 \le z < x^{\frac{1}{2}}$,

$$\pi(x) - \pi(x-h) = S(A, z) - \sum_{z \le q < x^{\frac{1}{2}}} S(A_q, q). \qquad (^1)$$

The remarkabale fact in this identity is that we can compute asymptotically the sum

$$\sum_{Q \le q < 2Q} S(A_q, q) \tag{6.1.3}$$

for some θ which is definitely smaller than $7/12$, if Q is in a certain range.

169 To show this, we set

$$x^{\frac{1}{3}} < Q \le x^{\frac{1}{2}}, \theta > \frac{1}{2}, \tag{6.1.4}$$

and we assume hereafter that x is sufficiently large. Then (6.1.3) is obviously equal to

$$\sum_{Q \le q < 2Q} \left(\pi\left(\frac{x}{q}\right) - \pi\left(\frac{x-h}{q}\right)\right).$$

[1] In this section, the letter q stands for prime numbers.

But it is easy to see that this is also equal to

$$\frac{1}{\log\frac{x}{Q}}\sum_{Q\le q<2Q}\left(\psi\left(\frac{x}{q}\right)-\psi\left(\frac{x-h}{q}\right)\right)+o\left(\frac{h}{(\log x)^3}\right) \tag{6.1.5}$$

where ψ is usual Chebyshev function. Replacing ψ by its explicit formula we get readily

$$\sum_{Q\le q<2Q}\left(\psi\left(\frac{x}{q}\right)-\psi\left(\frac{x-h}{q}\right)\right)$$
$$= hU(1)-\sum_{\substack{|\gamma|^{\rho}<T\\ \beta>0}}U(\rho)\frac{x^{\rho}-(x-h)^{\rho}}{\rho}+0\left(\frac{x(\log X)^3}{T}\right).$$

Here $\rho=\beta+i\gamma$ is a cmoplex zero of $\zeta(s)$, and

$$U(s)=\sum_{Q\le q<2Q}q^{-s};$$

also

$$T = x^{1-\theta+\eta}, \tag{6.1.6}$$

where η is a small positive constant. This sum over ρ is

$$\ll h\int_0^{1-(\log T)^{-\frac{3}{4}}} x^{\alpha-1}\sum_{\substack{|\gamma|<T\\ \beta>\alpha}}|U(\rho)|d\alpha,$$

because of Vinogradov's zero-free region. Thus appealing to LEMMA 25 we can infer that this is $o(h(\log x)^{-10})$ either, if $Q\ge T$ and $T^{6/5+\epsilon}Q\le x^{1-\eta}$ or, if $T^{4/5}\le Q\le T$ and $T^{16/5+\epsilon}<Qx^{1-\eta}$. Namely, if **170**

$$x^{\frac{11-16\theta}{5}+4\eta}\le Q\le x^{\frac{6\theta-1}{5}-2\eta} \tag{6.1.7}$$

and

$$\frac{6}{11}+2\eta\le\theta\le\frac{7}{12}, \tag{6.1.8}$$

then we have

$$\sum_{Q\le q<2Q}\left(\psi\left(\frac{x}{q}\right)-\psi\left(\frac{x-h}{q}\right)\right)=hU(1)(1+o((\log x)^{-5})). \tag{6.1.9}$$

We should note here that (6.1.7) and (6.1.8) imply (6.1.4).

Now we set

$$z=x^{\frac{11-16\theta}{5}+4\eta}, \tag{6.1.10}$$

and put

$$Z=x^{\frac{6\theta-1}{5}-2\eta}. \tag{6.1.11}$$

Then, by (6.1.5), (6.1.9) and by partial summation, we get

$$\sum_{z\le q<Z}\left(\pi\left(\frac{x}{q}\right)-\pi\left(\frac{x-h}{q}\right)\right)=(C_2(\theta)+o(\eta))\frac{h}{\log x}$$

171 where θ satisfies (6.1.8) and

$$C_2(\theta)=\log\left(\frac{(6\theta-1)(8\theta-3)}{3(1-\theta)(11-6\theta)}\right).$$

Thus we have

$$\pi(x)-\pi(x-h)=S(A,z)-(C_2(\theta)+o(\eta))(h/\log x)-\sum_{Z\le q<x^{\frac{1}{2}}}S(A_q,q), \tag{6.1.12}$$

provioded (6.1.8) holds.

Next, we appeal to THEOREM 10; we set there A,Ω as above, $\delta\equiv 1$, $X=h$, and

$$R_d=[x/d]-[(x-h)/d]-(h/d).$$

Obviously, all conditions required there are amply satisfied. Hence we have, for any $M,N\ge 1$ such that $MN\ge z^2$,

$$S(A,z)\ge\frac{e^{-\gamma}h}{\log z}\left(\phi_0\left(\frac{\log MN}{\log z}\right)-o(1)\right)-\log z\sup_{\alpha,\beta}\Big|\sum_{\substack{m<M\\ n<N}}\alpha_m\beta_nR_{mn}\Big| \tag{6.1.13}$$

where γ is the Euler constant, and $|\alpha_m|\le 1, |\beta_n|\le 1$.

Now we introduce the crucial

Lemma 27. *Let Z be as in* (6.1.11), *and let*

$$\frac{11}{20} + 2\eta \leq \theta \leq \frac{7}{12}. \tag{6.1.14}$$

Then we have 172

$$|\sum_{m,n<Zx^{\epsilon}} a_m b_n R_{mn}| \ll hx^{-c\eta^3}$$

for any $\{a_m\}, \{b_n\}$ *such that* $|a_m|, |b_n| < x^{\epsilon}$.

Before giving the proof, let us see the implication of this for our problem: We may set $M = N = Z$ in (6.1.13) provided θ satisfies (6.1.14), which we shall assume henceforth. Then, we note that (6.1.10), (6.1.11) imply $2 < 2\log Z/\log z < 4$, and that $\phi_0(u) = \frac{2e^{\gamma}}{u}\log(u-1)$ for $2 \leq u \leq 4$ because of (3.2.10). Thus we have

$$S(A,z) \geq (1 - o(\eta))C_1(\theta)\frac{h}{\log x},$$

where

$$C_1(\theta) = \frac{5}{6\theta - 1}\log\left(\frac{28\theta - 13}{11 - 16\theta}\right).$$

This and (6.1.12) give

$$\pi(x) - \pi(x-h) \geq (C_1(\theta) - C_2(\theta) - o(\eta))\frac{h}{\log x} - \sum_{z \leq q < x^{1/2}} S(A_q, q), \tag{6.1.15}$$

provided (6.1.11) and (6.1.14) hold.

Now let us proceed to the proof of LEMMA 27. Obviously, it suffices to consider the estimate of

$$E(A,B) = \sum_{\substack{A \leq m < 2A \\ B \leq n < 2B}} a_m b_n R_{mn}$$

under the assumption

$$AB \geq hx^{-\eta};\, A, B \leq Zx^{\epsilon}. \tag{6.1.16}$$

173

We put

$$A(s) = \sum_{A \le m < 2A} a_m m^{-s}, B(s) = \sum_{B \le n < 2B} b_n n^{-s}, L(s) = \sum_{\frac{L}{8} \le \ell < L} \ell^{-s},$$

where

$$L = \frac{x}{AB} \ge x^{3_\eta}, \tag{6.1.17}$$

because of (6.1.11), (6.1.14) and (6.1.16). Then, by Perron's inversion formula, we get

$$E(A, B) = \frac{1}{2\pi i} \int_{a-iT}^{a+iT} A(s)B(s)L(s)\frac{x^s - (x-h)^s}{s} ds - h(A(1)B(1) + 0(x^{-\eta/2})),$$

where $a = 1 + (\log x)^{-1}$, and T is as (6.1.6). We divide this integral into two parts, according to $|t| < \sqrt{L}$ and $\sqrt{L} \le |t| \le T$. And we observe the following: If $|t| < \sqrt{L}$, then, by (4.1.3), we have, for $\sigma = \alpha$,

$$L(s) = \frac{L^{1-s} - (L/8)^{1-s}}{1-s} + 0(L^{-1}),$$

and also

$$\frac{x^s - (x-h)^s}{s} = hx^{s-1} + o(|s|h^2 x^{-1}),$$

since $|6|hx^{-1} < x^{-\epsilon}$, because of (6.1.16). Inserting these into

$$\int\limits_{a-i\sqrt{L}}^{a+i\sqrt{L}} A(s)B(s)L(s)\frac{x^s - (x-h)^s}{s} ds,$$

174 we see readily that this is equal to

$$\frac{h}{2\pi i} \int\limits_{a-i\sqrt{L}}^{a+i\sqrt{L}} \frac{L^{1-s} - (L/8)^{1-s}}{1-s} A(s)B(s)x^{s-1} ds + o(hx^{-\eta/2})$$
$$= hA(1)B(1) + o(hx^{-\eta/2}).$$

Hence we get

$$E(A,B) = \frac{1}{2\pi i}\left\{\int_{a-iT}^{a-i\sqrt{L}} \int_{a+i\sqrt{L}}^{a+iT}\right\} A(s)B(s)L(s)\frac{x^s-(x-h)^s}{s}ds + o(hx^{-\eta/2}).$$

Then it is apparent that there exists a set $\{t_r\}$ such that $\sqrt{L} \leq |t_r| \leq T, |t_r - t_{r'}| \geq 1$ if $r \neq r'$, and

$$E(A,B) \ll h\sum_r |A(a+it_r)B(a+it_r)L(a+it_r)| + hx^{-\eta/2}$$

Now let $S(U,V,W)$ be the nmber of t_r such that $\nu < |A(a+it_r)| \leq 2\nu, W < |B(a+it_r)| \leq 2W, U < |L(a+it_r)| \leq 2U$ hold simultaneously. Here, we can, of course, assume that

$$|\log U|, |\log V|, \log W| \ll \log x.$$

Thus

$$E(A,B) \ll h(\log x)^3 \max_{U,V,W} UVWS(U,V,W) + hx^{-\eta/2}. \tag{6.1.18}$$

To estimate $S(U,V,W)$, let us see first an implication of $U < |L(a+it_r)| \leq 2U$. For this sake, let $\sqrt{L} \leq T_1 \leq T$, and let S_1 be the set of all t_r such that $T_1 \leq t_r < 2T_1, U < |L(a+it_r)| \leq 2U$. We note that, for any $T_i \leq t < 2T_1$,

$$L(a+it) = \frac{1}{2\pi i}\int_{\epsilon-iT_1/2}^{\epsilon+iT_1/2} \zeta(w +_a +it)\frac{L^w-(L/8)^w}{w}dw + 0\left(\frac{\log x}{T_1}\right).$$

175

Thus shifting the line of integration to $Re(w) = \frac{1}{2} - a, |Im(w)| \leq T_1/2$, we get

$$|L(a+it)| \ll L^{-1/2}\int_{-T_1/2}^{T_1/2} |\zeta\left(\frac{1}{2} + i(u+t)\right)|\frac{du}{1+|u+t|} + L^{-1/2}\log x.$$

Raising both sides to $4-th$ power and using Hölder's inequality, we have

$$|L(a+it)|^4 \ll L^{-2}\log^3 x \int\limits_{-T_1/2}^{T_1/2} \left|\zeta\left(\frac{1}{2}+i(u+t)\right)\right|^4 \frac{du}{1+|u+t|} + L^{-2}\log^4 x,$$

and thus

$$|S_1|U^4 \ll L^{-2}\log^3 x \int_{-T_1/2}^{T_1/2} \sum_{t_r \epsilon S_1} \left|\zeta\left(\frac{1}{2}+i(u+t_r)\right)\right|^4 \frac{du}{1+|ut_r|} + T_1 L^{-2}\log^4 x.$$

Hence, by (i) of LEMMA 23, we obtain

$$|S_1| \ll L^{-2}U^{-4}T_1 \log^c x.$$

This implies obviously

$$S(U,V,W) \ll L^{-2} \cup^{-4} T \log^c x.$$

Other estimates of $S(U,V,W)$ can be obtained by (ii) of LEMMA 23 and LEMMA 24, and we find readily

$$S(U,V,W) \ll x^{\epsilon} F_1 \tag{6.1.19}$$

with

$$F_1 = \min\left\{\frac{1}{V^2}+\frac{T}{V^2A}, \frac{1}{W^2}+\frac{T}{W^2B}, \frac{1}{V^2}+\frac{T}{V^6A^2}, \frac{1}{W^2}+ \frac{T}{W^6B^2}, \frac{T}{U^4L^2}, \frac{1}{U^4}+\frac{T}{U^{12}L^4}\right\}.$$

176

Then we consider the following four cases separately: (i) $F_1 \ll V^{-2}$, W^{-2}, (ii) $F_1 \gg V^{-2}$, W^{-2}, (iii) $F_1 \ll V^{-2}$, $F_1 \gg W^{-2}$, (iv) $F_1 \gg V^{-2}$, $F_1 \ll W^{-2}$. If (i) holds, we have $F_1 \ll (UVW)^{-1}U$. But LEMMA 17 yields

$$U \ll \exp\left(-c\frac{(\log L)^3}{(\log T)^2}\right) \ll x^{-cn^3},$$

because of (6.1.17). Hence in case (i), we have

$$UVWF_1 \ll x^{-cn^3}. \tag{6.1.20}$$

On the other hand, in the cases (ii) - (iV) we can argue just as in the estimation of F treated in § 4.3, and we get readily

$$F_1 \ll (UVW)^{-1}\left\{x^{-\frac{7}{16}}T^{\frac{31}{32}} + x^{-\frac{9}{20}}T\right\},$$
$$F_1 \ll (UVW)^{-1}\left\{x^{-\frac{3}{8}}T^{\frac{7}{16}}B^{\frac{3}{8}} + x^{-\frac{5}{12}}T^{\frac{1}{2}}B^{\frac{5}{12}}\right\},$$
$$F_1 \ll (UVW)^{-1}\left\{x^{-\frac{3}{8}}T^{\frac{7}{16}}A^{\frac{3}{8}} + x^{-\frac{5}{12}}T^{\frac{1}{2}}A^{\frac{5}{12}}\right\},$$

respectively. And by virtue of (6.1.14) and (6.1.16) these are all $0(x^{-\eta/4})$. Thus by this and (6.1.18) - (6.1.20), we obtain the assertion of the lemma.

Now, returning to (6.1.15), we have to seek for a good upper bound for

$$\sum_{Z \le q < x^{1/2}} S(A_q, q)$$

so that the left side of (6.1.15) is positive for a θ in the range (6.1.14). **177**

For this sake, we consider the sums

$$\sum_{Q \le q < 2Q} S(A_q, q),\ Z \le Q < x^{1/2}.$$

Obviously, this is not greater than

$$\frac{1}{\log Q}\sum_{Q \le q < 2Q} S(A_q, (Z^2/Q)^{\frac{1}{3}}) \log q.$$

Then to each summand we apply (3.4.20) with $\nu = 1$ and

$$y = Z^2/Q; \tag{6.1.21}$$

in particular, the function $\Theta_\nu(K)$ is defined by (2.3.4) with $z = (Z^2/Q)^{1/3}$ and $y = Z^2/Q$, and, of course, independent of q. Thus we have, on noting $\phi_1(s) = 2e^\gamma/s$ for $s \le 3$,

$$\sum_{Q \le q < 2Q} S(A_q, (Z^2/Q)^{1/3}) \log q$$

$$\leq (1+o(1))\frac{2h}{\log(Z^2/Q)}\sum_{Q\leq q<2Q}\frac{\log q}{q}$$
$$+\sum_K (-1)^{\omega(K)}\Theta_1(K)\sum_{\substack{d\epsilon K\\ f<z^{\tau}, f|P(z_1)\\ Q\leq q<2Q}} \xi_f^{(1+\omega(K))} R_{dfq}\log q$$
$$+\sum_{\substack{I<K\\ \omega(K)\equiv 0 \pmod 2}} \Theta_1(KI)\sum_{\substack{d\epsilon K\\ p,p'\epsilon I\\ f<z^{\tau}, f|P(z_1)\\ Q\leq q<2Q}} \xi_f^{(1)} R_{dpp'fq}\log q, \tag{6.1.22}$$

178 in which τ, z_1, and the mode of dissection of the interval (z_1, z) are as in § 3.4. Here we should note that, more precisely, we should have written $(R_q)_{df}$ and $(R_q)_{dpp'f'}$ instead of R_{qdf} and $R_{qdpp'f}$ respectively, but our present choice of q allows us to put the foumula as above.

Now let us estimate

$$E = \sum_{\substack{d\epsilon K\\ f<Z^{\tau}, f|P(Z_1)\\ Q\leq q<2Q}} \xi_f^{(1+\omega(K))} R_{dfq}\ \log\ q, \Theta_1\ (K) = 1.$$

It will turn out that this can be reduced to an application of LEMMA 27. To this end, we transform the factor $\log q$, q being a prime, into a sum of certain arithmetic functions. Let us put, for $\sigma > 1$,

$$\sum_{n=1}^{\infty} \lambda^{(1)}(n)n^{-s} = \sum_{q>U} q^{-s}\ \log\ q,$$
$$\sum_{n=1}^{\infty} \lambda^{(2)}(n)n^{-s} = -\sum_{n>U} \Big(\sum_{\substack{r|n\\ r\leq U}} \mu(r)\Big)\, n^{-s},$$
$$\sum_{n=1}^{\infty} \lambda^{(3)}(n)n^{-s} = -\zeta'(s)M(s) + (G(s)+N(s))(1-\zeta(s)M(s)) \tag{6.1.23}$$

with

$$M(s) = \sum_{n\leq U} \mu(n)n^{-s},$$

$$N(s) = \sum_{q \leq U} q^{-s} \ \log \ q,$$

$$G(s) = \sum_{q} \frac{\log \ q}{q^s(q^s - 1}).$$

179

Then it is easy to see that, for any $U \geq 1$,

$$\sum_{q} q^{-s} \ \log \ q \ = \sum_{n=1}^{\infty} (\lambda^{(1)} * \lambda^{(2)}(n) n^{-s} + \sum_{n=1}^{\infty} \lambda^{(3)}(n) n^{-s},$$

i.e.,

$$(\lambda^{(1)} * \lambda^{(2)})(n) + \lambda^{(3)}(n) = \begin{cases} \log n & \text{if } n \text{ is a prime,} \\ 0 & \text{otherwise .} \end{cases}$$

We shall use this identity with

(6.1.24) $$U = Q/Z.$$

Hence E is divided into two parts:

$$\begin{aligned} E = & \sum_{\substack{d \epsilon K \\ f < z^{\tau, f|P(z_1)} \\ Q \leq n < 2Q}} (\lambda^{(1)} * \lambda^{(2)}(n) \, \xi_f^{(1+\omega(K))} R_{dfn} \\ & + \sum_{\substack{d \epsilon K \\ f < Z^{\tau, f|P(z_1)} \\ Q \leq n < 2Q}} \lambda(3)(n) \, \xi_f^{(1+\omega(K))} R_{dfn} \\ = & E_1 + E_2, \end{aligned}$$

say, We estimate E_2 first. For this sake, we put

$$V(s) = \sum_{Q \leq n < 2Q} \lambda^{(3)}(n) \, n^{-s}$$

and 180

$$W(s) = \sum_{\substack{d \epsilon K \\ f < z^{\tau}, f|P(z_1)}} (df)^{-s}.$$

Then Perron's inversion formula gives

$$E_2 = \frac{1}{2\pi i}\int\limits_{1/2-iT'}^{1/2+iT'} \zeta(s)V(s)W(s)\,\frac{x^s-(x-h)^s}{s}\,ds + 0(hx^{-\eta/2}). \tag{6.1.25}$$

Here T' is such that

$$\int\limits_{1/2}^{2} |\zeta(\sigma+iT')|d\sigma \ll \log T, T \le T' < 2T,$$

T being as (6.1.6), which is an easy consequence of the mean value estimate for $|\zeta(\frac{1}{2}+\text{ it })|^2$, and guarantees us that, in deriving (6.1.25), the shift of the line of integration causes only a negligible error. Also, noting that $Q > T$ and (6.1.23), we have, again by Perron's inversion formula

$$V(s) = \frac{1}{2\pi i}\int\limits_{\sigma_1+iQ}^{\sigma_1-iQ}$$

$$\{-\zeta'(s+w)M(s+w)+(G(s+w)+N(s+w))(1-\zeta(s+w)M(s+w))\}$$
$$\frac{(2Q)^w-Q^w}{w}\,dw + o\left(\frac{Q^{1/2}}{|S|}(\log x)^2\right) + 0(x^\epsilon)$$

where $s = \frac{1}{2}+\text{ it }, |t| \le T'$, and $\sigma_1 = (\log\times)^{-1}$; the first 0-term being due to the pole at $w = 1-s$ of the integrand. Inserting this into (6.1.25), we have

$$E_2 = -\frac{1}{4\pi^2}\int\limits_{\sigma_1+iQ}^{\sigma_1-iQ}\frac{(2Q)^w-Q^w}{w}\int\limits_{1/2-iT'}^{1/2+iT'}\zeta(s)W(s)\times$$

$$\{-\zeta'(s+w)M(s+w)+(G(s+w)+N(s+w))(1-\zeta(s+w)M(s+w))\}$$

$$\frac{x^s-(x-h)^s}{s}\,\text{dsdw} + 0\left\{hx^{\epsilon-1/2}\int\limits_{1/2-iT'}^{1/2+iT'}\left(\frac{Q^{1/2}}{|S|}+1\right)|\zeta(s)W(s)|\,|ds|\right\}.$$

181

Using the mean value estimates for $|\zeta(\frac{1}{2}+\text{ it }|^2$ and $|W(\frac{1}{2}+\text{ it })|^2$ we can easily show that this 0-term is $0(hx^{-\eta/2})$. On the other hand, for the inner integral of the first term we have, on noting (6.1.21), (6.1.24) and $G(s+w) \ll \log x$,

$$\ll hx^{-1/2}\log\times\left\{Z+\left(\int\limits_{-T}^{T}|\,\zeta\left(\frac{1}{2}+\text{ it }\right)|^4 dt\right)^{1/4}\left(\int\limits_{-T}^{T}|\zeta'\left(\frac{1}{2}+\text{ it }+w\right)|^4\right.\right.$$

$$\left.+|\zeta\left(\frac{1}{2}+\text{ it }+w\right)|^4\right)dt)^{\frac{1}{4}}\times\left(\int\limits_{T}^{-T}|W\left(\frac{1}{2}+\text{ it }\right)M\left(\frac{1}{2}+\text{ it }+w\right)|^2\right.$$

$$\left.\left.\left(1+N\left(\frac{1}{2}+\text{ it }+w\right)|^2\right)dt\right)^{1/2}\right\}$$

$$\ll\ hx^{-1/2}\left(Z+T^{1/4}\ Q^{1/4}\right)\ \log^c x$$

$$\ll\ hx^{-\eta},$$

in which we have used the estimate

$$\int\limits_{T}^{-T}\left(|\zeta'\left(\frac{1}{2}+\text{ it }+w\right)|^4+|\zeta\left(\frac{1}{2}+\text{ it }+w\right)|^4\right)dt\ \ll\ Q\log^c x.$$

Thus $E_2 = 0(hx^{-n/2})$, and we have

$$E = E_1 + 0(hx^{-\eta/2}).$$

Now let us estimate E_1. Noting that both $\lambda^{(1)}(n)\lambda^{(2)}(n)$ vanish for 182
$n \leq U = Q/Z$, we infer that

$$E_1 \ll (\log\ x)^2\ \sup_{G,L}\ |\sum_{\substack{G\leq g\leq 2G\\ L\leq \ell<2L\\ d\epsilon K\\ f<z^\tau, f|P(z_1)}}\lambda^{(1)}(g)\lambda^{(2)}(\ell)\xi_f^{(1+\omega(K))}\,R_{dfg\ell}\,|$$

where

$$\frac{Q}{Z} < G, L \leq 2Z, \ Q \leq GL < 2Q.$$

We then recall that we have $\Theta_1(K) = 1$. Hence by virtue of LEMMA 16 there exists a decomposition $K = K_1 K_2$ such that $(K_1) \leq c_1 Z/G$, $(K_2) \leq c_2 Z/L$ since we have $y = z^3 = Z^2/Q$. Here $c_1 c_2 = GL/Q$. Thus we see immediately that we can again appeal to LEMMA 27, and we obtain

$$E_1 \ll hx^{-c\eta^3},$$

whence

$$E \ll hx^{-c\eta^3}.$$

Obviously, we can apply the same argument to the inner sums of the third term on the right side of (6.1.22). Thus (6.1.22) gives

$$\sum_{Q \leq q < 2Q} S(A_q, (Z^2/Q)^{\frac{1}{3}}) \leq (1 + 0(1)) \frac{2h}{\log(Z^2/Q)} \sum_{Q \leq q < 2Q} \frac{\log q}{q}$$

183 if $Z \leq Q < x^{1/2}$ and Θ satisfies (6.1.14). Then, by partial summation, we get

$$\sum_{Z \leq q < x^{1/2}} S(A_q, q) \leq (2 + 0(1))h \sum_{Z \leq q < x^{1/2}} \frac{1}{q \log(Z^2/q)}$$
$$= (C_3(\Theta) + o(n)) \frac{h}{\log x},$$

where

$$C_3(\Theta) = \frac{5}{6\Theta - 1} \log\left(\frac{5}{3(8\Theta - 3)}\right).$$

This and (6.1.15) give rise to

$$\pi(x) - \pi(x - x^{\Theta}) \geq (H(\Theta) - cn) \frac{x^{\Theta}}{\log x}$$

with

$$H(\Theta) = C_1(\Theta) - C_2(\Theta) - C_3(\Theta)$$

$$= \frac{5}{6\Theta - 1} \log\left(\frac{3(28\Theta - 13)(8\Theta - 3)}{5(11 - 16\Theta)}\right) - \log\left(\frac{(6\Theta - 1)(8\Theta - 3)}{3(1 - \Theta)(11 - 16\Theta)}\right)$$

provided

$$\frac{11}{20} + 2\eta \le \Theta \le \frac{7}{12}.$$

In particular, we have $H(0.558) > 0$. Therefore we have established

Theorem 16.

$$p_{n+1} - p_n \ll p_n^{0.56}.$$

6.2 Existence of Primes in Short Arithmetic Progressions

Now we turn to the problem of finding primes in short arithmetic progressions. The result in our mind is the celebrated theorem of Linnik: **184**

There exists an effectively computable constant $\mathscr{L}$ such that the least prime in any arithmetic progression (mod q) does not exceed $q^{\mathscr{L}}$.

By a combination of a dualized form of the Selberg sieve and analytical means, we shall prove a fairly generalized version of this important result.

We begin by making explicit the notion of the exceptional character which occurs in our discussion. Thus, let Q be a sufficiently large parameter, and let $1 - \delta$ be the Q-exceptional zero (cf. § 4.2), if exists, which comes from the L-function for χ_1 a unique real primitive character (mod q_1), $q_1 < Q$. Then refining (4.2.1), we have the following assertion: there exists an effective constant k, $0 < k \le 1$, such that for all primitive χ (mod q), $q < Q$,

$$\frac{L'}{L}(s,\chi) + o(\log Q) = \begin{cases} 0 & \text{if } \chi \ne \chi_0, \chi_1 \\ (s - 1 + \delta)^{-1} & \text{if } \chi = \chi_1, \\ -(s - q)^{-1} & \text{if } \chi = \chi_0 \end{cases} \tag{6.2.1}$$

in the region

$$\sigma \ge 1 - k(\log Q)^{-1}, |t| \le Q^{10},$$

where χ_0 is the trivial character, and 0-constant is effective. 185

If, in the above, we have

$$0 < \delta \leq \frac{k}{2 \log Q} \tag{6.2.2}$$

then we call χ_1 the Q-exceptional character; hereafter, we shall assume always that χ_1 stands for the Q-exceptional character, and $1 - \delta$ is the zero of $L(s, \chi_1)$ satisfying (6.2.2).

Then Linnik's theorem is apparently contained in

Theorem 17. *If χ_1 exists, then we put $\triangle = \delta \log Q$, and otherwise $\triangle = 1$. Also, we put*

$$\tilde{\psi}(x, \chi) = \begin{cases} \sum\limits_{n<x} \chi^{(n)}\, \Lambda(n) & \text{if } \chi \neq \chi_0, \chi_1, \\ \sum\limits_{n<x} \chi_1(n)\Lambda(n) + \frac{x^{1-\delta}}{1-\delta} & \text{if } \chi = \chi_1, \\ \sum\limits_{n<x} \Lambda(n) - x & \text{if } \chi = \chi_0, \end{cases}$$

where Λ is the von Mangoldt function. Then, there exist effectively computable positive constants a_0, a_1 and a_2 such that

$$\sum_{q<Q} \sum_{\chi} \sideset{}{^*}\sum_{(\text{mod } q)} |\tilde{\psi}(x, \chi) - \tilde{\psi}(x - h, \chi)| \leq a_1 \triangle\, h \exp\left(-a_2 \frac{\log x}{\log Q}\right),$$

186 *provided*

$$Q^{a_0} < \frac{x}{Q} < h < x, \log x \leq (\log Q)^2.$$

We may prove this by employing the Deuring-Heilbronn phenomenon, a zero-density estimate of the Linnik type (THEOREM 15) as well as the explicit formula for $\tilde{\psi}(x, \chi)$. But we shall exhibit below that there is a more direct and conceptually simpler way to achieve this.

First we introuduce the multiplicative function B defined by

$$B(n) = \begin{cases} 1 & \text{if } \chi_1 \text{ does not exist,} \\ \sum\limits_{d|n} \chi_1(d) d^{-\delta} & \text{if } \chi_1 \text{ exists .} \end{cases}$$

And, throughout the sequel, we shall use the results and the notations of § 1.4 by setting $f = B$ always: It is quite easy to see that B satisfies the conditions $(C_1), (C_2)$ and (C_3) introduced there with

$$\alpha = 2 + \epsilon, \beta = \frac{1}{2} + \epsilon, \gamma = \frac{1}{2} + \epsilon, D = q_1^{1/4+\epsilon},$$

$$\mathscr{F} = \begin{cases} 1 & \text{if } \chi_1 \text{ does not exist ,} \\ L(1+\delta, \chi_1) & \text{if } \chi_1 \text{ exists .} \end{cases}$$

Among these, the fact $\alpha = 2 + \epsilon$ is obvious, if χ_1 does not exist, and otherwise, it is a consequence of

$$F_p = \left(1 - \frac{1}{p}\right)^{-1} \left(1 - \frac{\chi_1(p)}{p^{1+\delta}}\right)^{-1} > \left(1 - \frac{1}{p^2(1+\delta)}\right)^{-1}. \tag{6.2.3}$$

187

Hence we have in THEOREM 5

$$Y_B(M; Q, R) \ll (Q^4 R^4 M^{1/2})^{1+\epsilon}. \tag{6.2.4}$$

We should observe also that B satisfies (C_1') of § 1.4 with $k = 2$.

Further, we remark that we have, on setting $f = B$ in (1.4.5),

$$G_1(R) \gg \vartriangle^{-1} \mathscr{F} \log Q \tag{6.2.5}$$

provided

$$R \geq Q^{1/2+\epsilon}. \tag{6.2.6}$$

If χ_1 does not exist, then this is implies by (1.2.12). On the other hand, if χ_1 exists we argue as follows. We note that

$$G_1(R) \geq R^{-2\delta} \sum_{r<R} \frac{\mu^2(r)}{g(r)} r^{2\delta},$$

and we have, in our present case,

$$\sum_{r=1}^{\infty} \frac{\mu^2(r)}{g(r)} r^{2\delta - s} = \zeta(s + 1 - 2\delta) L(1 + s - \delta, \chi_1) A(s),$$

where, as is easily seen, $A(2\delta) = 1$, and $A(s)$ is bounded for $\sigma \geq -3/4$. Thus, using Perron's inversion formula and an elementary estimate for $L(s, \chi_1)$, we get

$$\sum_{r<R} \frac{\mu^2(r)}{g(r)} r^{2\delta} = \frac{R^{2\delta}}{2\delta} L(1+\delta, \chi_1) + 0(R^{-1/2+\epsilon} q_1^{1/4+\epsilon}). \tag{6.2.7}$$

188 Taking R appropriately in this, we get, in particular,

$$L(1+\delta, \chi_1) \gg \delta, \tag{6.2.8}$$

since the left side of (6.2.7) is not less than 1, whence (6.2.5).

Also we shall need a lower bound of δ, and this is supplied by (4.2.4), which yields

$$\delta \gg Q^{-1/2} (\log Q)^{-4}. \tag{6.2.9}$$

Having these preparations in our hands, we may enter into the actual proof of the theorem. We observe first that (6.2.1) and (6.2.2) imply

$$\frac{L'}{L}(s, \chi) = o(\log Q) \tag{6.2.10}$$

for all primitive $\chi \pmod q$, $q \leq Q$, and for s on the segment

$$|t| \leq Q^{10}, \sigma = \sigma_0 = \begin{cases} 1 - \frac{k}{4 \log Q} & \text{if } \chi_1 \text{ does not exist ,} \\ 1 - \frac{k}{\log Q} & \text{if } \chi_1 \text{ exists .} \end{cases}$$

Thus, specifically, we have

$$\tilde{\psi}(x, \chi) = \frac{1}{2\pi i} \int\limits_{\sigma_0+iT}^{\sigma_0-iT} \frac{L'}{L}(s, \chi) \frac{x^s}{s} ds + o(xQ^{-9}), \tag{6.2.11}$$

189 where $T = Q^{10}$ and $\log x \leq (\log Q)^2$, as we shall henceforth assume.

Next, we put

$$V_r(s, \chi) = F(s, \chi) M_r(s, \chi; \Lambda^{(2)}) - 1,$$

where, for $\sigma > 1$,

$$F(s,\chi) = \sum_{n=1}^{\infty} B(n)\chi(n)n^{-s},$$

and $\Lambda^{(2)}$ and $M_r(s,\chi;\Lambda^{(2)})$ are defined in THEOREM 4 and LEMMA 5 (with $f = B$) respectively. In particular, we have, for $\sigma > 1$,

$$F(s,\chi)M_r(s,\chi;\Lambda^{(2)}) = \sum_{n=1}^{\infty}\chi(n)\Phi_r(n)B(n)\left(\sum_{d|n}\Lambda_d^{(2)}\right)n^{-s}, \tag{6.2.12}$$

where Φ_r is defined by (1.4.10) with $f = B$.

Then (6.2.11) is transformed into

$$\tilde{\psi}(x,\chi) = -\frac{1}{2\pi i}\int_{\sigma_0-iT}^{\sigma_0+iT}\frac{L'}{L}(s,\chi)V_r(s,\chi)^2\frac{x^s}{s}ds$$

$$+\frac{1}{2\pi i}\int_{\sigma_0-iT}^{\sigma_0+iT} W_r(s,\chi)\frac{x^s}{s}ds + o(xQ^{-9}), \tag{6.2.13}$$

where

$$W_r(s,\chi) = (V_r(s,\chi) - 1)M_r(s,\chi;\Lambda^{(2)})L'(s,\chi)H(s,\chi)$$

with

$$H(s,\chi) = \begin{cases} 1 & \text{if } \chi_1 \text{ does not exist,} \\ L(s+\delta,\chi\chi_1) & \text{if } \chi_1 \text{ exists ,} \end{cases}$$

190

We should note here that we have, for $0 \le \sigma \le 1$,

$$M_r(s,x;\Lambda^{(2)}) \ll z^{(1+z\vartheta)(1-\sigma)+\epsilon}g(r)r^{-\sigma+\epsilon},$$

where z and ϑ occur in the definition of $\Lambda^{(2)}$. Also, we have $g(r) \ll r^{2+\epsilon}$, because of (1.4.6) and (6.2.3).

Now let us set in the above

$$r \le R = Q, z = Q^{40}, \vartheta = \epsilon \tag{6.2.14}$$

Then using a simple estimete for $L(s,\chi)$, we have

$$V_r(s,\chi) \ll Q^{55}, W_r(s,\chi) \ll Q^{110} \tag{6.2.15}$$

for $\sigma(\log Q)^{-2}, |t| \leq T = Q^{10}, \chi \pmod q$, $q < Q$. Hence shifting the line of integration to $Re(s) = (\log Q)^{-2}, |t| \leq T$ in the second integral of (6.2.13), we get

$$\tilde{\psi}(x,\chi) = -\frac{1}{2\pi i}\int\limits_{\sigma_0-iT}^{\sigma_0+iT} \frac{L'}{L}(s,\chi)V_r(s,\chi)^2\frac{x^2}{s}ds + o(xQ^{-9}) \text{ if } x \geq Q^{120}.$$

Thus, recalling (6.2.10), we have

$$\begin{aligned}&|\tilde{\psi}(x,\chi) - \tilde{\psi}(x-h,\chi)|\\ &\ll h\exp(-\frac{k\log\times}{4\log Q}\log Q\int\limits_{-T}^{T}|V_r(\sigma_0+it,\chi)|^2dt + xQ^{-9}.\end{aligned}$$

191

We multiply both sides by $\mu^2(r)(k(q)g(r))^{-1}$, and sum first over $r < R, (r,q) = 1$, and next over primitive $\chi \pmod q), q < Q$, getting

$$\begin{aligned}&\sum_{q<Q} K(q)^{-1}G_p(R)\sum_{X}\sideset{}{^*}\sum_{(\text{mod } q)}|\tilde{\psi}(x,\chi) - \tilde{\psi}(x-h,\chi)|\\ &\ll \Psi h\exp\left(-\frac{k\log\times}{4\log Q}\right)\log Q + xQ^{-6},\end{aligned}$$

where

$$\Psi = \sum_{\substack{r<R\\ q<Q\\ (r,q)=1}}\frac{\mu^2(r)}{g(r)k(q)}\sum_{\chi}\sideset{}{^*}\sum_{(\text{mod } q)}\int\limits_{-T}^{T}|V_r(\sigma_0+it,\chi)|^2dt.$$

Then we observe (1.4.9), (6.2.5) and (6.2.8), and we get

$$\sum_{q<Q}\sum_{\chi}\sideset{}{^*}\sum_{\text{mod } q}|\tilde{Psi}(x,\chi) - \tilde{\Psi}(x-h,\chi)|$$

(6.2.16) $$\ll \Psi\mathscr{F}^{-1}\Delta h \exp\left(-\frac{k\log x}{4\log Q}\right) + xQ^{-6}.$$

Thus it suffices to show

(6.2.17) $$\psi \ll \mathscr{F}$$

For this sake, we consider the Mellin integral

$$X_r^{(1)}(s,\chi) = \frac{1}{2\pi i}\int\limits_{2-i\infty}^{2+i\infty} V_r(s+w,\chi)\Gamma(w)Z^w dw$$

where $Z = Q^{150}, \sigma = \sigma_0, |t| \le T$. Because of (1.3.11) and (6.2.12), this is equal to **192**

$$\sum_{n\ge z}\chi(n)B(n)\phi_r(n)\left(\sum_{d|n}\Lambda_d^{(2)}n^{-s}e^{-n/Z}\right)$$

On other hand, we have, shifting the line of integration to $Re(w) = -\sigma_0$,

$$\begin{aligned} X_r^{(1)}(s,\chi) &= V_r(s,\chi) + E(\chi)\mathscr{F}K(q)M_r(1,\chi;\Lambda^{(2)})r(1-s)Z^{1-s} + o(Q^{-29}) \\ &= V_r(s,\chi) + \chi_r^{(2)}(s,\chi) + o(Q^{-29}) \end{aligned}$$

say. Hence we have

$$\psi \ll \psi_1 + \psi_2 + Q^{-30}$$

where ψ_j is obtained by replacing $V_r(\sigma_0+it,\chi)$ in the definition of ψ by $X_r^{(j)}(\sigma_0+it,\chi)$, $j = 1, 2$. Then appealing to LEMMA 4 and noting (6.2.4), we get

$$\psi_1 \ll \sum_{n\ge z}(\mathscr{F}n + Q^{19}n^{1/2+\epsilon})B(n)\left(\sum_{d|n}\Lambda_d^{(2)}\right)^2 n^{-2\sigma_0}e^{-2n/Z}.$$

Thus, by (6.2.8), (6.2.9) and (6.2.14), we have

$$\psi_1 \ll \mathscr{F}\sum_{n\ge z}\tau_2(n)\left(\sum_{d|n}\Lambda_d^{(2)}\right)^2 n^{1-2\sigma_0}e^{-2n/Z}$$

since $B(n) \leq \tau_2(n)$. This sum may be truncated at $n = Z^2$, and then the exponent $1-2\sigma_0$ of n can be decreased to $-1-(\log Q)^{-1}$. Then by viutue of THEOREM 4, we get

$$\Psi_1 \ll \mathscr{F}.$$

193 As for Ψ_2, we remark that

$$\int_{\sigma_0-iT}^{\sigma_0+iT} |r(1-s)|^2 |ds| \ll \log Q.$$

Hence

$$\Psi_2 \ll \mathscr{F}^2 \log Q \sum_{r<R} \mu^2(r) g(r) |M_r(1, \chi_0; \Lambda^{(2)})|^2;$$

by this and LEMMA 6, we have

$$\Psi_2 \ll \mathscr{F}.$$

Thus we have proved (6.2.17). Then the assertain of the theorem follows immediately from (6.2.8) and (6.2.16).

NOTES (VI)

The extraordinary argument developed in § 6.1 is taken from Iwaniec-Jutila [34] and its subsequent improvement [22] due to Heath-Brown and Iwaniec. The achievement of Iwaniec and Jutila was a real breakthrough that had come after long efforts of searching for new methods which could oversome the difficulty in improving upon (6.1.1)-the prime number theorem of Huxley [25]. One should observe that most of the best results and the sharpest tools in today's analytic number theorey are mobilized in their argument.

194 We have obtained the exponent 0.56 as stated in THEOREM 16, but Heath-Brown and Iwnaniec have indeed obtained the exponent $0.55+\epsilon$, i.e.,

$$p_{n+1} - p_n \ll p_n^{\theta}$$

whenever $\Theta > 0.55$. We indicate here how to achieve this.

In our argument, we esimated $S(A, z)$ from below by appealing to THEOREM 10. But as a matter of fact we started our proof of THEOREM 10 at the inequality stated in LEMMA 10. This means that we cast away the thrid sum on the right side of the identity stated in LEMMA 9. Namely, we have actually proved in the above

$$S(A, z) - \sum_{K} \Delta_0(K) \sum_{d \epsilon K} S(A_d, p(d)) \geq (C_1(\Theta) - o(\eta))\frac{h}{\log x},$$

where conventions are as in §6.1. Now, as is apparent by the definition of Δ_0, the part corresponding to those K with $\omega(K) = 2$ continuous essentially

$$I_1 = \sum_{(Z^2/p)^{1/3} \leq q < p < z} S(A_{pq}, q).$$

On the other hand, for the sum over q in (6.1.12), we have, by Buchstab's indentity,

$$\sum_{Z \leq q < x^{1/2}} S(A_q, q) = \sum_{Z \leq q < x^{1/2}} S(A_q, (Z^2/q)^{1/3}) - \sum_{\substack{Z \leq q < x^{1/2} \\ (Z^2/q)^{1/3} \leq p < q}} S(A_{qp}, p).$$

And we have obtained the upper bound $(C_3(\Theta) + 0(\eta))h(\log x)^{-1}$, for 195
the first sum on the right side. Collecting these observations, we see that we have, for $11/20 < \Theta < 7/12$,

$$\pi(x) - \pi(x - x^{\Theta}) \geq (H(\Theta) - o(1))\frac{x^{\Theta}}{\log x} + I_1 + I_2,$$

where I_2 is the sum over p, q in the last indetity. Now Heath-Brown and Iwaniec have able to give good lower bounds for I_1 and I_2 by means of weighted zero-density estimates similar to LEMMA 25, so that they could conclude the right side of the last inequality is positive at $\Theta = 0.55 + \epsilon$ even though $H(0.55) < 0$.

In reducing the estimate of E to LEMMA 27, we used a variant of Vaughan's idea [81]. Our argument there should be compared with the corresponding part of Heath - Brown and Iwaniec [22].

THEOREM 17 is the prime number theorem of Gallagher [15]. The argument developed in §6.2 is due to Motohashi [58]. It should be stressed that our proof be compared with Linnik's formidable works [47]. The simplification is definitely due to the injection of sieve in to the theory.

Gallagher's important work [15] contains two novel ideas; one is embodied in LEMMA 3 and the other is his effective use of Bomberi-
196 Daveport's extension (1.2.13) of the Burn- Titchmarsh theorem, as has been already mentiond in NOTES (V). These ideas were combained with Turan's power-sum method to produce a zero-density estimate similar to THEOREM 15. Then using the Deuring- Heilbronn phenomenon, Gallaher obtained THEOREM 17. In this context, it should be stressed that we have dispensed with zero-density esimates of Linnik's type the Deuring Heilbronn phenomenon and the power sum method altogeher.

It is also quite remarkable that, in Linnik's works, the sieve aspect of the theory was almost implict, but the succeding simplifications pushed it qradually to the surface and in our proof of THEOREM 17, Selberg's sieve method governs the whole affair.

For a proof of (6.2.1), see Prachar [[62], Kap. IV].

We did not take care for the numerical percison of various constants, which is important in the actual computation of the Linnik constant $\mathscr{L}$. On this matter, see Jutila [40], Graham [19] and Chen [11]; in the last work, it is claimed that $\mathscr{L} \leq 17$, for sufficiently large modulus.

It seems worth remarking that our argument of §6.2 yields also

$$\sum_{q<T} \sideset{}{^*}\sum_{\chi \pmod q} \bar{N}(\alpha, T, \chi) \ll \Delta T^{50(1-\alpha)},$$

197 where $\bar{N}(\alpha, T, \chi)$ is defined in §5.2. This should be compared with Bombieri [[6], Théorèm 14]. For the proof, see Motohashi [[55], II].

Bibliography

[1] R. Balasubrmanian and K. Ramachandra: The place of an indentity of Ramanujan in prime number theory. Proc. Indian Acad. Sci., 83 (A) (1976). 156-165. **198**

[2] M.B Barban and P.P Vehov: An extermal problem. Trans Mosow Math. Soc., 18 (1968), 91-99.

[3] H. Bohr and E. Landu: Beitrage zur Theorie der Riemannschen Zetafunktion. Math. Ann., 74(1913), 3-30.

[4] E. Bomboeri: On a theorem of van Lint and Richert. Symposia Mathematic, 4 (1968/69), 175-180.

[5] E Bombieri: Density theorems for the zeta-function. Proc Sympos. Pure Math., 20 (1971), 352-358.

[6] E. Bombieri: Le grand crible dans la theorie analytiqe des nombers. Soc. Math. France, Asterisque No. 18, 1974.

[7] E. Bombieri: The asympotic sieve. Rend. Accad. Naz. XL(5)1/2 (1975/76), 243-269.

[8] E. Bombieri and H.Davenport: On the large sieve method. Number Theory and Analysis (Papers in honour of Edmund Landau), New York 1969, pp.9-22.

[9] N.G. de Bruijn: On the number of uncancelled elements in the sieve of Eratoshenes. Indag. Math., 12 (1950), 247-256.

[10] D.A. Burgess: On character sums and L-series II. Proc London 199
Math. Soc., (3)13(1963),524-536.

[11] J. Chen: On the least prime in an arithmetical progression and theorems concerning the zero of Dirichelt's L-functions II. Sci. Sinica, 22 (1979), 3-31.

[12] E. Fogels: On the zeros of L-functions. Acta Arith., 11 (1965), 67-96.

[13] E. Fouvry and H. Iwaniec: On Bombieri- Vinogradov's type theorems (to appear).

[14] J.B. Friedlander and H. Iwaniec: On Bombieri's asympototic sieve. Ann. Scu. Norm. Sup. Pisa, (4)4(1978), 719-756.

[15] P.X. Gallager: A large sieve density esimate near $\sigma = 1$. Invent. Math., 11 (1970), 329-339.

[16] P.X. Gallahger: Sieving by prime powers. Acta Arith., 24 (1973/74), 491-497.

[17] P.X. Gallagher : The large sieve and probabilistic Galois theory. Proc. Sympos. Pure Math., 24(1973), 91-101.

[18] A.O Gel'fond.: On an elementary approach to some problems from the field of distribution of prime numbers. Vest. Moskov. Univ. Ser. Fiz.Mat. Estest. Nauk., 8 (1953), 21-26.

200 [19] S. Graham : Applications of sieve methods Ph.D. Dissertation, Univ. Of Michigan, 1977.

[20] G. Greaves : A weighted sieve of Brun's type (to appear).

[21] H. Halberstam and H.-E. Richert: Sieve methods. Acad, Press, New York 1974.

[22] D.R Heath- Brown and H. Iwaniec: On the difference between consecutive primes. Invent. Math., 55(1979), 49-69.

[23] G. Hoheisel: Primzahproblem in der Analysis. Stiz. Preuss. Akad. Wiss.., 33 (1930), 3-11.

[24] C. Hooley: On the Brun-Titchmarsh theorem. J. Reine Angew. Math., 255 (1972). 60-79.

[25] M.N. Huxley: On the diffenence between consecutive primes. Invent. Math., 15 (1972), 164-170.

[26] M.N. Huxley: Large values of Dirichlet polynominals III. Acta Arith., 26(1975), 435-444.

[27] A.E. Ingham: Note on Riemann's ζ function and Dirichlet's L-functions. J. London Math. Soc., 5(1930), 107-112.

[28] H. Iwaniec: On the error eerm in the linear sieve. Acta Arith., 19 (1971), 1-30.

[29] H. Iwaniec: The sieve of Eratosthenes- legendre. Ann. Suc. Norm. Sup. Pisa, (4)4(1977), 257-268.

[30] H. Iwaniec: Rosser's sieve. (to appear).

[31] H. Iwaniec: A new form of the error-therm in the linear sieve. (to **201**
appear)

[32] H. Iwaniec: On the Brun - Titchmarsh theorem. (to appear)

[33] H. Iwaniec: Sieve methods. A talk deliverd at IMU Cibgressm Helsinki 1978. (to appear)

[34] H. Iwaniec and M. jutila: Primes in short intervals. Arkiv for Mat., 17 (1979), 167-176.

[35] J. Johnsen: On the large sieve method in GF $[q, x]$. Mathematika, 18 (1971), 172-184.

[36] W.B. Jurkat and H. -E. Richert: An improvement of Selberg's sieve method I. Acta Arith., 11 (1965), 217-240.

[37] M. Jutila: On a density theorem of H.L. Montgomery for *L*-functions. Ann. Acad. Sci. Fenn. Ser. AI., No. 520 (1972), 13 pp.

[38] M. Jutila: On numbers with a large prime factor. J. Indian Math. Soc., (N.S)., 37 (1973), 43-53.

[39] M. Jutila: Zero-density estimates for *L*-functions. Acta Arith, 32 (1977), 55-62.

[40] M. Jutila: On Linnik's constant. Math. Scand., 41 (1977), 45 -62.

[41] M. Jutila: On a problem of Barban and Vehov. (to appear)

[42] A.A. Karacuba: Foundations of analytic number theory. Nauka, Moscow 1975. (Russian)

202 [43] S. Knapowski: On Linnik's theorem concerning exceptional L-zeros. Publ. Math. Debrecen, 9 (1962), 168 - 178.

[44] B.V. LEVIN: Comparison of A. Selberg's and V. Brun's sieves. Uspehi Mat. Nauk, 20 (1965), 214-220. (Russian)

[45] Yu.V. Linnik: The large sieve. C.R. Acad. Sci. URSS (N.S.), 30 (1941), 292-294.

[46] YU.V. Linnik: An elementary proof of Siegel's theorem based on a method of Vinogradov. Mat. Sb., 12 (1943), 225-230. (Russian)

[47] Yu.V. Linnik: On the least prime in an arithmetic progression I. The basic theorem. Rec. Math. (Sbornik), 15 (57) (1944), 139 - 178; II. The Deuring -Geibronn phenomenon. ibid, 347-368.

[48] H.L. Montgomery: Topics in multiplicative number theory Lecture Notes in Math., 227 (1971), 178 pp.

[49] H.L. Montgomery: The analytic principle of the large sieve. Bull. Amer. Math. Soc., 84 (1978), 547 -567.

[50] H.L. Montgomery: Linnik's theorem. (unpulished)

[51] H.L. Montgomery and R.C. Vaughan: The large sieve. Mathematika, 20 (1973), 119-134.

[52] Y. Motohashi: On some improvements of the Brun -Titchmarsh theorem. J. Soc. Japan, 26 (1974), 306- 323.

[53] Y. Motohashi: On a density theorem of Linnik. Proc. Japan Acad., 51 (1975), 815-817.

[54] Y. Motohashi: A note on the large sieve. Proc. Japan Acad., 53 203
(1977), 17-19; II. ibid, 122-124 ; III. ibid, 55 (1979), 92-94; IV. ibid, 56 (1980), 288-290.

[55] Y. Motohashi: on the Deuring-Heilbronn phenomenon I. Proc. Japan Acad., 53 (1977), 1-2; II. ibid, 25-27.

[56] Y. Motohashi: On Vinogradov's zero-free region for the Riemann zeta-function. Proc. Japan Acad., 54 (1978), 300-302.;

[57] Y. Motohashi: On the zero-free region of Dirichlet's L- functions. Proc. Japan Acad., 54 (1978), 332-334.

[58] Y. Motohashi: Primes in arithemtic progrssions. Invent. Math., 44 (1978), 163-178.

[59] Y. Motohashi: A note on Siegel's zeros. Proc. Japan Acad., 55 (1979), 190-192.

[60] Y. Motohashi: On the linear sieve I. Proc. Japan Acad., 56 (1980), 285-287; II. ibid, (to appear).

[61] Y. Motohashi: An elementary proof of Vinogradov's zerofree region for the Riemann zeta-function. (to appear)

[62] K. Prachar: Primzahlverteilung. Springer, Berlin 1957.

[63] K. Ramachandra: A simple proof of the mean fourth power estimate for $\zeta\left(\frac{1}{2}+\mathrm{it}\right)$ and $L\left(\frac{1}{2}+\mathrm{it},\chi\right)$. Ann. Scu. Norm. Sup. Pisa, (4) 1 (1974), 81-97.

[64] K. Ramachandra: On the zeros of the Riemann zeta-function and 204
L-series. Acta Arith., 34 (1978), 211-218.

[65] H.-E. Richert: Selberg's sieve with weights. Proc. Sympos. Pure Math., 20 (1971), 287 -310.

[66] H.-E. Richert: Lectures on sieve methods. Tata Inst. Fund. Research, Bombay 1976.

[67] S. Salerno and C. Viola: Sieving by almost primes. J. London Math. Soc., (2) 14 (1976), 221-234.

[68] A. Selberg: On the zeros of Riemann's zeta-function. SKr. Norske Vid. Akad. Oslo (1942), No.1.

[69] A. Selberg: Contribution to the theory of Dirichlet's *L*-functions. Skr. Norske Vid. Akad. Oslo (1946), No.3.

[70] A. Selberg: The zeta-function and the Riemann hypothesis. 10.Skand. Math. Kongr., (1946), 187-200.

[71] A. Selberg: Contribution to the theory of the Riemann zetafunction. Arch. for Math. og Naturv. B, 48 (1946), No.5.

[72] A. Selberg: On an elementary method in the theory of primes. Norske Vid. Selsk. Forth., Trondhjem 19 (1947), 64-67.

[73] A. Selberg: On elementary methods in prime number theory and thier limitations. 11. Skand. Math. Kongr., Trondhjeim (1949), 13-22.

[74] A. Selberg: The general sieve-method and its place in prime number theory. Proc. IMU Congress, Cambridge, Mass., 1 (1950), 286-292.

205 [75] A. Selberg: Sieve methods. Proc. Sympos. Pure Math., 20 (1971), 311-351.

[76] A. Selberg: Remarks on sieves. Proc. 1972 Number Theory Conf., Boulder (1972), 205-216.

[77] A. Selberg: Remarks on multiplicative functions. Lecture Notes in Math., 626 (1977), 232 -241.

[78] H. Soebert: Sieve methods and Siegel's zeros. (to appear)

[79] E.c. Titchmarhs: The zeta-function of Riemann. Cambridge Univ. Press, 1930.

[80] P. Tur*án*: On a density theorem of Yu. V. Linnik. Magyar Tud. Akad. Mat. Kutato Int. Ser. 2, 6 (1961), 165 - 179.

[81] R.C. Vaughan: On the estimation of trigonometrical sums over primes, and related questions. Report No. 9, Inst. Mittag-Leffler, 1977.

[82] E. Wirsing: ELementare Beweise des Primzahlsatzes mit Restglied II. J. Reine Angew. Math., 214/215 (1964), 1-18.