SIEVE METHODS

ANDREW M. ODLYZKO

A senior thesis

CALIFORNIA INSTITUTE OF TECHNOLOGY
PASADENA, CALIFORNIA

1971

CONTENTS

	Preface	i
1.	The Sieves of Eratosthenes and Brun	1
	Notes	15
2.	Selberg's Sieve	17
	Notes	27
3.	Applications of Selberg's Sieve	29
	Notes	46
4.	Primes in Arithmetic Progressions	48
	Notes	61
5.	Gallagher's Sieve	62
	Notes	68
6.	Trigonometric Polynomial Inequalities	69
	Notes	94
7.	The Large Sieve	96
	Notes	110
8.	Bombieri's Theorem	111
	Notes	128
	Bibliography	129

PREFACE

The main goal of this work is to give an introductory account of sieve methods that would be understandable with only a slight knowledge of analytic number theory. These notes are based to a large extent on lectures on sieve methods given by Professor Van Lint and the author in a number theory seminar during the 1970-71 academic year, but rather extensive changes have been made in both the content and the presentation.

Several developments related to the subject of these notes are not discussed in them at all. One such is Rényi's probabilistic version of the large sieve, for which the reader is referred to Rényi [4]-[9]. Another is Vinogradov's method of trigonometric sums which was used to prove the famous theorem that any sufficiently large odd integer is representable as a sum of at most three primes. This method is discussed in I. M. Vinogradov [1]. Neither of the two methods mentioned above would fit in very well in this work, however. A much more serious omission, due to lack of time, is that of Selberg's lower bound method. Although much more complicated and in many ways less satisfactory than the upper bound method, it gives the best results known in many cases. We might mention here that using this method it has been shown that there are infinitely many primes p such that $p + 2$ is a product of at most three primes, and that every sufficiently large even integer can be represented as a sum of a prime and a product of at most three primes. While these results do not prove the twin prime conjecture (which states that $p + 2$ is a prime for infinitely many primes p) nor Goldbach's conjecture

(that every even integer ≥ 4 is a sum of two primes), they are still significant achievements. A new unified proof of both of the above results is given in Richert [1]. A good introduction to Selberg's sieve is provided in Halberstam and Roth [1; Chapter 4]. A much more comprehensive and more up-to-date presentation is given in Richert [3].

All references for results quoted in any chapter are given in the notes at the end of that chapter. These notes also contain some general bibliographic information and often some additional facts. The bibliography contains practically all the publications on sieve methods of which the author is aware, and an attempt has been made to supply Mathematical Reviews references for as many as possible.

I am greatly indebted to Professor Gallagher for permission to use the material of Chapter 5, which has not yet been published. I would also like to thank Professor Van Lint for lecturing on the material of Chapters 3 and 4, and for help on many problems. I am very grateful to the Mathematics Department of the California Institute of Technology for a 1970 Summer Research and Independent Study Fellowship, which made this work possible. Most of all I would like to thank Professor Apostol for his advice, encouragement, and guidance. His help in the writing and editing of these notes has been invaluable. Finally, I would like to express my appreciation to Mrs. Lorayne Decker for her patience and skill in typing the manuscript.

Andrew Odlyzko
Pasadena, California
June 1971

1

THE SIEVES OF ERATOSTHENES AND BRUN

The name "sieve method" comes from the sieve of Eratosthenes, an algorithm for finding all primes. It utilizes the fact that a natural number is prime if and only if it is not divisible by any prime smaller than itself. To find all the primes $\leq x$, one writes down the natural numbers 2, 3, 4, ..., [x] in this order. Since 2 is the first prime, it is left untouched, but every proper multiple of it (that is, every second number starting with 4) is crossed out, since it is composite. The next number in the sequence is 3, and it has not been crossed out yet. Hence it is not divisible by any prime smaller than itself, and so it is prime. Therefore 3 is left as it is, but every proper multiple of 3, being composite, is crossed out. The next number, 4, has already been crossed out, and therefore must be composite. The next one, 5, has not been crossed out and hence must be a prime. It is left alone but all its proper multiples are crossed out. Since if an integer $\leq x$ is composite, at least one of its prime factors has to be $\leq \sqrt{x}$, it is sufficient to continue this process only up to $[\sqrt{x}]$. The numbers which have not been crossed out are exactly the primes $\leq x$. Thus this procedure "sieves out" all the composite numbers.

Let $\pi(y)$ denote the number of primes $\leq y$. The sieve of Eratosthenes can be used to obtain an exact formula for $\pi(x) - \pi(\sqrt{x})$, the number of primes between $\sqrt{x}$ and x. Let

$$s'(n) = \begin{cases} 1 & \text{if } n = 1, \\ 1 & \text{if } n \text{ is prime and } \sqrt{x} < n \leq x, \\ 0 & \text{otherwise.} \end{cases}$$

Then

$$\pi(x) - \pi(\sqrt{x}) = \sum_{2 \leq n \leq x} s'(n).$$

Now let

$$\Pi = \prod_{p \leq \sqrt{x}} p.$$

We observe that the sieve of Eratosthenes removes all those natural numbers $\leq x$ which are not relatively prime to Π, except for the primes $\leq \sqrt{x}$. Since the Möbius function satisfies the relation

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

we have

$$s'(n) = \sum_{d|(n,\Pi)} \mu(d) \tag{1.1}$$

for each $n \leq x$, and hence

$$\pi(x) - \pi(\sqrt{x}) + 1 = \sum_{n \leq x} s'(n) = \sum_{n \leq x} \sum_{d|(n,\Pi)} \mu(d) = \sum_{d|\Pi} \mu(d)[\tfrac{x}{d}]$$

$$= x \sum_{d|\Pi} \frac{\mu(d)}{d} + \sum_{d|\Pi} \mu(d)\{[\tfrac{x}{d}] - \tfrac{x}{d}\}.$$

But

$$\sum_{d|\Pi} \frac{\mu(d)}{d} = \prod_{p|\Pi} (1 - \frac{1}{p}) = \prod_{p \le \sqrt{x}} (1 - \frac{1}{p})$$

and so we obtain the exact formula

$$(1.2) \qquad \pi(x) - \pi(\sqrt{x}) + 1 = x \prod_{p \le \sqrt{x}} (1 - \frac{1}{p}) + \sum_{d|\Pi} \mu(d)\{[\frac{x}{d}] - \frac{x}{d}\}.$$

Each term in $\sum_{d|\Pi} \mu(d)\{[\frac{x}{d}] - \frac{x}{d}\}$ is ≤ 1 in absolute value, but there are $2^{\pi(\sqrt{x})}$ terms, a number much larger than x. Still, one might hope that the first term on the right side of (1.2) yields the correct order of magnitude of $\pi(x) - \pi(\sqrt{x})$ and that the second term is of a smaller order because of cancellation due to alterations in the sign of $\mu(d)$ and the small size of the factors $[\frac{x}{d}] - \frac{x}{d}$ (since for most divisors d of Π, x/d is much smaller than 1). However, a theorem of Mertens states that

$$(1.3) \qquad \prod_{p \le y} (1 - \frac{1}{p}) \sim \frac{e^{-\gamma}}{\log y} \quad \text{as } y \to \infty,$$

where γ is Euler's constant, so the first term on the right of (1.2) is asymptotic to $2e^{-\gamma}x/\log x$. But, by the prime number theorem, the left side of (1.2) is asymptotic to $x/\log x$, so that the second term on the right is asymptotic to $(1 - 2e^{-\gamma})x/\log x$. Hence each term on the right of (1.2) is asymptotic to a constant times $\frac{x}{\log x}$, so that we cannot expect to obtain useful estimates for $\pi(x) - \pi(\sqrt{x})$ from this identity.

Nevertheless, the same underlying ideas can be used to obtain estimates for the number of integers $\leq x$ which are not divisible by any prime $p \leq z$, provided z is much smaller than $\sqrt{x}$. (In the foregoing discussion we considered the case $z = \sqrt{x}$.) This time we let

$$\Pi = \prod_{p \leq z} p$$

and define $s'(n)$ by (1.1). Then the number $S(x,z)$ of integers $\leq x$ not divisible by any prime $\leq z$ is given by

$$S(x,z) = \sum_{n \leq x} s'(n) = \sum_{n \leq x} \sum_{d \mid (n,\Pi)} \mu(d) = \sum_{d \mid \Pi} \mu(d)[\tfrac{x}{d}]$$

$$= x \prod_{p \leq z} (1 - \tfrac{1}{p}) + \sum_{d \mid \Pi} \mu(d)\{[\tfrac{x}{d}] - \tfrac{x}{d}\}.$$

Since each term in the sum on the right has absolute value ≤ 1, the sum itself is $O(2^{\pi(z)})$, and therefore

$$S(x,z) = x \prod_{p \leq z} (1 - \tfrac{1}{p}) + O(2^{\pi(z)}).$$

By Mertens' theorem the first term is asymptotic to $e^{-\gamma} x/\log z$ as $x \to \infty$, provided $z \to \infty$ also. Now if $z = \log x$ we have

$$2^{\pi(z)} \leq 2^{2z/\log z} = 4^{\frac{\log x}{\log z}} = e^{\frac{\log x \log 4}{\log z}} = x^{\frac{\log 4}{\log z}} = x^{o(1)},$$

so the second term is much smaller than the first term. Thus, if $z \to \infty$ as $x \to \infty$ but $z \leq \log x$ we have

$$S(x,z) \sim \frac{e^{-\gamma} x}{\log z} \quad \text{as } x \to \infty.$$

As is shown below, the above method can be generalized to give information about the number of integers $\leq x$ which are not divisible by any prime $\leq z$, when these integers belong to sequences other than the sequences of natural numbers. Unfortunately the requirement that z be very small in comparison with x limits the usefulness of such generalizations.

Modern sieve methods originated with Viggo Brun around 1920. He used a new sieve to obtain several important number-theoretic results, notably an estimate of the density of twin primes. In the next few pages we will develop a very simple form of Brun's sieve which will enable us to prove a theorem on the density of twin primes (Theorem 1.4). While weaker than the best results obtainable with Brun's sieve, it will be sufficient to prove the celebrated theorem of Brun that the sum of the reciprocals of the twin primes converges (Theorem 1.5). We will first state Theorem 1.4 and then use it to prove Theorem 1.5. The rest of the chapter will then be devoted to the proof of Theorem 1.4.

(1.4) THEOREM. Let T denote the set of twin primes (that is, primes p such that either p - 2 or p + 2 is also a prime), and let $T(x) = \sum_{\substack{p \leq x \\ p \in T}} 1$. Then

$$T(x) \ll x\left(\frac{\log \log x}{\log x}\right)^2 \quad \text{for } x \geq 3,$$

where $\ll$ is the Vinogradov symbol*).

*) $F(x) \ll G(x)$ is equivalent to $F(x) = O(G(x))$; that is, both imply that there is a $c > 0$ such that $|F(x)| \leq c\, G(x)$ for x in the range indicated.

So far we have only utilized concepts from the sieve of Eratosthenes. Since good estimates for $B(x,d)$ can be obtained quite easily (see (1.12) below), we could write $S(x,z)$ as the sum of a main term for which an asymptotic estimate exists and a remainder. Unfortunately, just as before, the remainder would be a sum over all divisors d of Π, and therefore we would need to take z of a much smaller order of magnitude than x to ensure that the main term is the dominant one.

The main idea of Brun's sieve is that in order to get an effective upper bound for $S(x,z)$ when z is fairly large compared with x one should sum in (1.6) over only a relatively small subset of divisors of Π, where this subset is chosen so that the resulting sum is greater than or equal to $S(x,z)$ (and to get a lower estimate one should sum over a subset that makes that sum less than or equal to $S(x,z)$). In this chapter we will obtain an upper bound for $S(x,z)$ when $z = x^{\frac{1}{20 \log\log x}}$. To do this, let us take an even natural number $m \leq r$ (which will be specified more exactly later) and define

$$(1.7) \qquad s(n) = \sum_{\substack{d \mid (n(n+2),\Pi) \\ \nu(d) \leq m}} \mu(d)$$

where $\nu(k)$ is the number of distinct prime divisors of k.

(1.8) **LEMMA.** For all integers n, $s'(n) \leq s(n)$.

Proof: Let $t = (n(n+2), \Pi)$. Then we have

$$s'(n) = \sum_{d|t} \mu(d), \quad s(n) = \sum_{\substack{d|t \\ \nu(d) \le m}} \mu(d).$$

If $t = 1$, then $s'(n) = s(n) = 1$. If $t > 1$, then $s'(n) = 0$. Therefore in order to prove $s'(n) \le s(n)$, we only need to show $s(n) \ge 0$ when $t > 1$.

Let us assume $t > 1$. Since t divides Π, all divisors of t are squarefree and hence

$$s(n) = \sum_{\substack{d|t \\ \nu(d) \le m \\ \nu(d) \text{ even}}} 1 - \sum_{\substack{d|t \\ \nu(d) \le m \\ \nu(d) \text{ odd}}} 1 .$$

Fix a prime p which divides t. Suppose $\delta | t$, $\nu(\delta) \le m$, and $\nu(\delta)$ is odd (so that $1 \le \nu(\delta) \le m - 1$ because m is even).

Let

$$\delta' = \begin{cases} p\delta & \text{if } p \nmid \delta , \\ \delta/p & \text{if } p \mid \delta . \end{cases}$$

Then

$$\nu(\delta') = \begin{cases} \nu(\delta) + 1 & \text{if } p \nmid \delta , \\ \nu(\delta) - 1 & \text{if } p \mid \delta . \end{cases}$$

Thus δ' is a divisor of t with $\nu(\delta')$ even and $0 \le \nu(\delta') \le m$. Since the correspondence between δ and δ' is one-to-one, there are at least as many d with $d|t$, $\nu(d) \le m$ for which $\nu(d)$ is even as there are those for which $\nu(d)$ is odd. Therefore $s(n) \ge 0$, and the proof of the lemma is complete.

Applying Lemma 1.8, we obtain

$$S(x,z) \le \sum_{\substack{n=1 \\ n \text{ odd}}}^{x} s(n) = \sum_{\substack{n=1 \\ n \text{ odd}}}^{x} \sum_{\substack{d \mid (n(n+2),\Pi) \\ \nu(d) \le m}} \mu(d) \tag{1.9}$$

$$= \sum_{\substack{d \mid \Pi \\ \nu(d) \le m}} \mu(d) B(x,d).$$

In order to simplify further work we will denote by $\rho^{(f)}$ all $d \mid \Pi$ with $\nu(d) = f$ (with the convention $\rho^{(0)} = 1$). $\sum_{\rho^{(f)}}$ will be understood to be the sum over all $d \mid \Pi$ with $\nu(d) = f$. Then (1.9) becomes

$$S(x,z) \le \sum_{f=0}^{m} (-1)^f \sum_{\rho^{(f)}} B(x,\rho^{(f)}). \tag{1.10}$$

Next let us compute $B(x,\rho^{(f)}) = B(x,p_{i_1} \cdots p_{i_f})$ = number of odd $n \le x$ such that $n(n+2) \equiv 0 \pmod{(p_{i_1} \cdots p_{i_f})}$. Since for odd n we have $(n,n+2) = 1$, each p_{i_j} divides exactly one of n, n + 2. Thus the congruence $n(n+2) \equiv 0 \pmod{(p_{i_1} \cdots p_{i_f})}$ is equivalent to the congruences

$$\begin{aligned} n &\equiv 0 \pmod{\Big(\prod_{p \in P_1} p\Big)} \\ n+2 &\equiv 0 \pmod{\Big(\prod_{p \in P_2} p\Big)} \end{aligned} \tag{1.11}$$

where P_1 and P_2 are any two sets such that

$$P_1 \cup P_2 = \{p_{i_1}, \ldots, p_{i_f}\} \text{ and } P_1 \cap P_2 = \emptyset.$$

By the Chinese Remainder Theorem, (1.11) has exactly one solution modulo $\rho^{(f)}$ for fixed P_1 and P_2. Therefore for fixed P_1 and P_2 there will be $\frac{x}{2\rho^{(f)}} + \theta$ odd $n \leq x$ satisfying (1.11), where $|\theta| \leq 1$. Since there are 2^f possible choices for the sets P_1 and P_2,

$$B(x,\rho^{(f)}) = 2^f \frac{x}{2\rho^{(f)}} + 2^f\theta, \quad |\theta| \leq 1. \tag{1.12}$$

Combining (1.12) with (1.10) leads to

$$S(x,z) \leq \frac{x}{2} \sum_{f=0}^{m} (-1)^f \sum_{\rho^{(f)}} \frac{2^f}{\rho^{(f)}} + \sum_{f=0}^{m} \binom{r}{f} 2^f. \tag{1.13}$$

Now

$$\sum_{f=0}^{m} \binom{r}{f} 2^f \leq \sum_{f=0}^{m} 2^f r^f < (2r)^{m+1}, \tag{1.14}$$

and

$$\sum_{f=0}^{m} (-1)^f \sum_{\rho^{(f)}} \frac{2^f}{\rho^{(f)}} = \sum_{f=0}^{r} (-1)^f \sum_{\rho^{(f)}} \frac{2^f}{\rho^{(f)}} + \sum_{f=m+1}^{r} (-1)^{f-1} \sum_{\rho^{(f)}} \frac{2^f}{\rho^{(f)}} \tag{1.15}$$

$$= \prod_{3 \leq p \leq z} \left(1 - \frac{2}{p}\right) + \sum_{f=m+1}^{r} (-1)^{f-1} s_f$$

where s_f is the f-th elementary symmetric function of $\frac{2}{3}, \frac{2}{5}, \ldots, \frac{2}{p_{r+1}}$ *) .

---

*) The k-th elementary symmetric function of $a_1,\ldots,a_t$ is defined as $\Sigma\, a_{i_1} \cdots a_{i_k}$ where the sum is over the $\binom{t}{k}$ possible choices of $i_1,\ldots,i_k$ from $1,2,\ldots,t$; thus, for example, $s_1(x,y,z) = x+y+z$, $s_2(x,y,z) = xy + xz + yz$, and $s_3(x,y,z) = xyz$.

We thus find from (1.12) - (1.15) that

$$(1.16) \qquad S(x,z) \leq (2r)^{m+1} + \frac{x}{2} \prod_{3\leq p\leq z} (1 - \frac{2}{p}) + \frac{x}{2} \sum_{f=m+1}^{r} (-1)^{f-1} s_f .$$

We next observe that

$$(1.17) \qquad s_1 \cdot s_f \geq (f+1)s_{f+1}, \quad f = 1,2,\ldots$$

since any product of (f+1) terms can be written in (f+1) ways as a product of a single factor and a product of f factors. Therefore $s_2 \leq \frac{s_1^2}{2}$, $s_3 \leq \frac{s_1^3}{6}$, and in general (by induction)

$$s_f \leq \frac{s_1^f}{f!} .$$

Relation (1.17) also shows that $s_f \geq s_{f+1}$ if $s_1 \leq f + 1$. Let us now choose $m + 1 \geq s_1 = \sum_{3\leq p\leq z} \frac{2}{p}$ (the only restriction on m so far was that it be even). Then $\sum_{f=m+1}^{r} (-1)^{f-1} s_f$ becomes an alternating sum with terms decreasing in absolute magnitude, so that it is $\leq s_{m+1} \leq \frac{s_1^{m+1}}{(m+1)!} \leq \left(\frac{es_1}{m+1}\right)^{m+1}$ (in the last inequality we used the fact that $n! > (n/e)^n$). In addition, since by a theorem of Mertens

$$(1.18) \qquad s_1 = \sum_{3\leq p\leq z} \frac{2}{p} = 2 \log\log z + O(1),$$

we can choose m so that $e^2 s_1 < m + 1 < 9s_1$ for sufficiently large z (this guarantees $m < r$), and then

$$\left(\frac{es_1}{m+1}\right)^{m+1} \leq \left(\frac{1}{e}\right)^{m+1} \leq e^{-e^2 s_1} \leq e^{-s_1}.$$

Also, $1 - y \leq e^{-y}$ for all real y, and thus

$$\prod_{3 \leq p \leq z} \left(1 - \frac{2}{p}\right) \leq e^{-s_1}.$$

Since $r = \pi(z) - 1$, we can use $2r \leq z$ in (1.16). Then the inequalities above lead to

$$S(x,z) \leq z^{9s_1} + xe^{-s_1}.$$

Because of (1.18) there is a constant C such that for $z \geq 3$,

$$2 \log\log z - C < s_1 < 2 \log\log z + C$$

and therefore for sufficiently large x,z

$$S(x,z) \leq z^{(18 \log\log z + 9C)} + \frac{xe^C}{(\log z)^2}.$$

We now take $z = x^{\frac{1}{20 \log\log x}}$. Then for sufficiently large x

$$\log z = \frac{\log x}{20 \log\log x}, \quad \log\log z \leq \log\log x,$$

and hence

$$S\left(x, x^{\frac{1}{20 \log\log x}}\right) \leq x^{9/10 + o(1)} + 400\, e^C x \left(\frac{\log\log x}{\log x}\right)^2.$$

Therefore for all $x \geq 3$, say,

$$S\left(x, x^{\frac{1}{20 \log\log x}}\right) \ll x\left(\frac{\log\log x}{\log x}\right)^2.$$

To complete the proof of Theorem 1.8 we observe that

$$T(x) \leq 2z + 2S(x,z) \ll x\left(\frac{\log\log x}{\log x}\right)^2 .$$

The method used in proving Theorem 1.4 can easily be generalized. We could, for example, investigate the density of primes p for which $p + 2$ and $p + 6$ are also primes (it has been conjectured that there are infinitely many such prime triplets). We would find that there are $\ll x\left(\frac{\log\log x}{\log x}\right)^3$ of them below x. However, in Chapter 3 we will obtain more accurate and more general results by using Selberg's sieve, so we will not deal further with this subject here.

So far we have dealt only with Brun's upper bound method. However, the lower bound method is analogous (actually one of the advantages of Brun's over Selberg's sieve is that the two methods are almost identical in case of Brun's sieve). The most important difference is that in (1.7) we would take m odd in order to obtain a lower bound for $S(x,z)$. After making a few obvious modifications in the proof of Theorem 1.4 we would find that $S\left(x,x^{\frac{1}{20\log\log x}}\right) > cx\left(\frac{\log\log x}{\log x}\right)^2$ for some positive constant c. Unfortunately, this result does not allow us to obtain a lower bound for $T(x)$.

The best versions of Brun's sieve give estimates for $S(x,x^{\alpha})$ where $\alpha < 1/2$ is a constant. There a set of primes is chosen:

$$p_t < p_{t-1} < \cdots < p_1 .$$

In the upper bound method s(n) is defined as in (1.8) except that summation is over those d which have at most 2i of their prime divisors greater than or equal to p_i (in the lower bound method: at most 2i - 1). By choosing those t primes appropriately it is possible to show that

$$T(x) \ll x(\log x)^{-2}.$$

We will obtain this result using Selberg's sieve.

Notes on Chapter 1.

Proofs of Mertens' theorems (Eqs. (1.3) and (1.18)) may be found in Hardy and Wright [1; Chapter 22].

The exact formula (1.2) not only fails to give a good asymptotic estimate for $\pi(x)$, but it is also not very useful in actually calculating $\pi(x)$ for a specific x. However, Meissel [1] has found another exact formula for $\pi(x)$ which leads to an effective (even though laborious) procedure for calculating this function. Uspensky and Heaslet [1; Chapter 5] also give a presentation of this method.

Several mathematicians have investigated the number $\Phi(x,y)$ of positive integers $\leq x$ and free of prime factors $\leq y$. Buchstab [1] proved that for a fixed $u \geq 2$,

$$\lim_{y \to \infty} \Phi(y^u,y)y^{-u}\log y = \omega(u),$$

where

$$\frac{d}{du}\{u\omega(u)\} = \omega(u-1)$$

for $u \geq 2$. Further results were later obtained by De Bruijn [1]-[3] and Ramaswami [1], [2]. Their methods, however, were analytic.

Our version of Brun's sieve largely follows the presentation of Landau [2; Part 2, Chapter 2], Rademacher [2; Chapter 15], and Gelfond and Linnik [1; Chapter 5]. The main difference is our explicit use of the sieving function s(u), a concept borrowed from Selberg's sieve.

The literature on Brun's sieve is rather extensive. Practically all the papers listed in the Bibliography that were published before 1942

deal with this subject. Moreover, there is a presentation of Brun's sieve in Gelfond and Linnik [1; Chapter 5]. Although it had seemed for a while that Selberg's sieve had superseded Brun's, Miech [2] recently used the latter to prove important results on the almost-prime values assumed by a polynomial.

2

SELBERG'S SIEVE

Let us suppose that $A = \{a_\nu\}$ is a sequence defined by $a_\nu = h(\nu)$, $\nu = 1,2,\ldots$ where h is an integer-valued polynomial. Let P be a finite set of primes ($\leq z$ for some z). Many important number theoretic problems reduce to the problem of estimating the number $S(A,P,n)$ of a_ν, $\nu=1,\ldots, n$ which are not divisible by any prime $p \in P$. For example, such estimates in the case $h(x) = x(x+2)$ give information about the twin primes. In this chapter we prove a result (Theorem 2.17) which will enable us to obtain effective upper bounds for $S(A,P,n)$ in a wide variety of cases.

Selberg's sieve can be formulated so as to apply to general sequences of integers. However, the sequences generated by polynomials are the most important ones for which effective estimates can be obtained, and so we will deal only with them.

Let $N(d)$ denote the number of solutions of

$$h(\nu) \equiv 0 \pmod{d}, \quad 1 \leq \nu \leq d. \tag{2.1}$$

Then by a property of congruences N is a multiplicative function. Those primes p for which $N(p) = 0$ do not contribute to the "sieving out" process and so we may assume that none of them belong to P. In addition, if $N(p) = p$ for some prime $p \in P$, then $S(A,P,n) = 0$. Therefore we will require $1 \leq N(p) < p$ for all $p \in P$. For convenience in later work we define

$$f(d) = \frac{d}{N(d)}. \tag{2.2}$$

The function f is multiplicative and $1 < f(p) \leq p$ for all $p \in P$.

The ν between 1 and n are divided into $[\frac{n}{d}]$ complete residue classes modulo d plus at most d - 1 additional ν. There are exactly N(d) integers ν in every complete residue class such that $a_\nu \equiv 0 \pmod{d}$ and there are at most N(d) such ν among those left at the end. Hence

$$(2.3) \qquad \sum_{\substack{\nu \le n \\ d \mid a_\nu}} 1 = [\tfrac{n}{d}]N(d) + \theta \cdot N(d) = \frac{n}{f(d)} + R(d), \quad (0 \le \theta \le 1)$$

where

$$(2.4) \qquad |R(d)| \le \frac{d}{f(d)} .$$

Let us now define $\Pi = \prod_{p \in P} p$ and

$$s'(a) = \sum_{d \mid (a,\Pi)} \mu(d) = \begin{cases} 1 & \text{if } (a,\Pi) = 1, \\ 0 & \text{if } (a,\Pi) > 1. \end{cases}$$

Then

$$(2.5) \qquad S(A,P,n) = \sum_{\nu \le n} s'(a_\nu) = \sum_{\nu \le n} \sum_{d \mid (a_\nu, \Pi)} \mu(d)$$

$$= \sum_{d \mid \Pi} \mu(d) \sum_{\substack{\nu \le n \\ d \mid a_\nu}} 1 = n \sum_{d \mid \Pi} \frac{\mu(d)}{f(d)} + \sum_{d \mid \Pi} \mu(d) R(d).$$

Unfortunately this relation suffers from the disadvantage that the second term, the remainder term, is a sum over all the divisors of Π, and so is very difficult to estimate. The main problem is to circumvent this obstacle.

First we observe that if $s(a)$ is a function defined by

$$s(a) = \sum_{d|(a,\Pi)} \lambda(d) \tag{2.6}$$

where $\lambda(d)$ is any function defined for all $d|\Pi$, then (just as in (2.5))

$$\sum_{\nu\leq n} s(a_\nu) = n \sum_{d|\Pi} \frac{\lambda(d)}{f(d)} + \sum_{d|\Pi} \lambda(d)R(d)$$

where the left side depends on λ. If λ is chosen so that $s(a_\nu) \leq s'(a_\nu)$ for all ν, then

$$S(A,P,n) \leq n \sum_{d|\Pi} \frac{\lambda(d)}{f(d)} + \sum_{d|\Pi} \lambda(d)R(d). \tag{2.7}$$

The essential part of both Brun's and Selberg's sieves is the choice of λ so that $S(A,P,n)$ can be effectively estimated. The difficulty lies in the need to make the remainder smaller than the main term, while at the same time minimizing the latter (since we would like to minimize the right side of (2.7)). The task of making the remainder term fairly small is accomplished in both sieve methods by defining λ to be zero outside a relatively small subset D^* of divisors of Π. In Brun's sieve, λ is chosen equal to μ on D^*, and D^* is chosen (in a rather complicated way) to ensure that $s(a) \geq s'(a)$. In Selberg's sieve, on the other hand, the possibility of choosing λ different from μ is utilized, while D^* is chosen in a rather simple fashion. The resulting estimates are, as it turns out, more effective than those obtainable with Brun's sieve and are easier to find.

(It should be pointed out, however, that the transition from Brun's to Selberg's sieve was not as obvious as it might appear from this chapter. Our whole presentation utilizes many ideas introduced by Selberg in connection with his sieve, while Brun's sieve was originally formulated quite differently.)

The set of all functions defined by (2.6) is too wide to work with. However, it turns out that very good sieving functions can be selected from a subset of it that is particularly well-behaved.

Suppose we define

$$s(a) = \sum_{d|(a,\Pi)} \Lambda(d) \tag{2.8}$$

where Λ is any function defined on all the divisors of Π. Then

$$s^2(a) = \Big(\sum_{d|(a,\Pi)} \Lambda(d) \Big)^2 = \sum_{d|(a,\Pi)} \lambda(d) \tag{2.9}$$

where

$$\lambda(d) = \sum_{\substack{d_1,d_2|\Pi \\ d=[d_1,d_2]}} \Lambda(d_1)\Lambda(d_2). \tag{2.10}$$

Thus if we denote by T the set of all functions s defined by (2.6), then $s^2 \in T$ whenever $s \in T$. Selberg's decisive observation was that we should look for our sieving function among the functions s^2, where $s \in T$. One great advantage of this choice, as we shall see, is that it leads to a

quadratic form which enables us, for every subset D of divisors of Π, to find the minimum value of $\sum_{d|\Pi} \frac{\lambda(d)}{f(d)}$ over all functions Λ which are zero outside D. The main reason for this is the great freedom of choice as to Λ. To make $s^2(a) \geq s'(a)$ it is sufficient to ensure that $s^2(a) \geq 1$ when $(a,\Pi) = 1$ and that $s^2(a) \geq 0$ when $(a,\Pi) > 1$. The first condition is easily satisfied by specifying $\Lambda(1) = 1$. But the second condition is satisfied trivially no matter what real-valued function we choose for Λ. Thus (2.7) will hold subject to the single restriction $\Lambda(1) = 1$. This is of great practical importance, since for a general function s defined by (2.6) it might be very difficult to prove that it satisfies $s(a) \geq s'(a)$. As a result we can, when looking for the minimum of $\sum_{d|\Pi} \frac{\lambda(d)}{f(d)}$, treat $\Lambda(d)$ as a free variable whenever $d > 1$.

Before we prove the main theorem (Theorem 2.13), we will derive a few preliminary results. Let us define

$$g(k) = \sum_{d|k} \mu(d) f(k/d) \tag{2.11}$$

so that

$$f(k) = \sum_{d|k} g(d).$$

If $k|\Pi$ then $(d, \frac{k}{d}) = 1$ for each divisor d of k and hence $f(k/d) = f(k)/f(d)$ since f is multiplicative. Therefore if $k|\Pi$ we have

$$g(k) = f(k) \sum_{d|k} \mu(d)/f(d) = f(k) \prod_{p|k} \left(1 - \frac{1}{f(p)}\right).$$

In particular, we see that $g(k) > 0$. Also, since f is multiplicative,

$$f((d_1,d_2))\cdot f([d_1,d_2]) = f(d_1)\cdot f(d_2)$$

and therefore

$$\frac{1}{f([d_1,d_2])} = \frac{1}{f(d_1)\cdot f(d_2)} \sum_{\ell \mid (d_1,d_2)} g(\ell).$$

Next, let D be a finite, divisor-closed set (i.e., if $d \in D$ then all divisors of d belong to D) and let

$$F(k) = \sum_{\substack{d\in D \\ k\mid d}} G(d),$$

where G is arbitrary. Then we have the inversion formula

$$\sum_{\substack{t\in D \\ d\mid t}} \mu(\tfrac{t}{d})F(t) = G(d) \tag{2.12}$$

because

$$\sum_{\substack{t\in D \\ d\mid t}} \mu(\tfrac{t}{d})F(t) = \sum_{\substack{t\in D \\ d\mid t}} \mu(\tfrac{t}{d}) \sum_{\substack{\ell\in D \\ d\mid \ell}} G(\ell) = \sum_{\substack{\ell\in D \\ t\mid \ell}} G(\ell) \sum_{\substack{d\mid t \\ t\mid \ell}} \mu(\tfrac{t}{d}) = \sum_{\substack{\ell\in D \\ d\mid \ell}} G(\ell) \sum_{\delta\mid \frac{\ell}{d}} \mu(\delta) = G(d).$$

(2.13) **THEOREM.** Let D be a divisor-closed set of divisors of Π, and let $D^* = \{d;\ d\mid\Pi,\ d = [d_1,d_2]$, where $d_1,d_2 \in D\}$. Assume that $\Lambda(1) = 1$ and that $\Lambda(d) = 0$ if $d \notin D$, and let

$$\lambda(d) = \sum_{\substack{d_1,d_2 \in D \\ [d_1,d_2]=d}} \Lambda(d_1)\Lambda(d_2) \quad \text{if} \quad d \in D^*,$$

and define $\lambda(d) = 0$ if $d \notin D^*$. Let

$$H(\Lambda) = \sum_{d \in D^*} \frac{\lambda(d)}{f(d)} .$$

Then

$$H(\Lambda) \geq \frac{1}{Q} , \tag{2.14}$$

where

$$Q = \sum_{d \in D} \frac{1}{g(d)} , \tag{2.15}$$

and this lower bound is attained when

$$\Lambda(d) = \frac{\mu(d)f(d)}{Q} \sum_{\substack{t \in D \\ d|t}} \frac{1}{g(t)} . \tag{2.16}$$

Proof: We have

$$H(\Lambda) = \sum_{d_1,d_2 \in D} \frac{\Lambda(d_1)\Lambda(d_2)}{f([d_1,d_2])} = \sum_{d_1,d_2 \in D} \frac{\Lambda(d_1)\Lambda(d_2)}{f(d_1)f(d_2)} \sum_{t|(d_1,d_2)} g(t)$$

$$= \sum_{t \in D} g(t) \sum_{\substack{d_1,d_2 \in D \\ t|d_1, t|d_2}} \frac{\Lambda(d_1)\Lambda(d_2)}{f(d_1)f(d_2)} = \sum_{t \in D} g(t) \Big(\sum_{\substack{d \in D \\ t|d}} \frac{\Lambda(d)}{f(d)} \Big)^2 .$$

Now we write

$$y(t) = \sum_{\substack{d\in D \\ t|d}} \frac{\Lambda(d)}{f(d)} .$$

By (2.12) this means

$$\Lambda(d) = f(d) \sum_{\substack{t\in D \\ d|t}} \mu(t/d)y(t).$$

Since f is multiplicative, $f(1) = 1$, and by taking $d = 1$ we find

$$1 = \sum_{t\in D} \mu(t)y(t).$$

This leads to

$$H(\Lambda) = \sum_{t\in D} g(t)y^2(t) = \sum_{t\in D} g(t)y^2(t) - \frac{2}{Q} \sum_{t\in D} \mu(t)y(t) + \frac{1}{Q^2} \sum_{t\in D} \frac{\mu^2(t)}{g(t)} + \frac{1}{Q}$$

$$= \sum_{t\in D} \frac{1}{g(t)} \left\{g(t)y(t) - \mu(t) \frac{1}{Q}\right\}^2 + \frac{1}{Q} .$$

Therefore $\min H(\Lambda) = 1/Q$ and this minimum is attained if and only if

$$y(t) = \frac{\mu(t)}{Qg(t)} .$$

But in view of (2.12) that is equivalent to

$$\Lambda(d) = \frac{\Lambda(d)}{Q} \sum_{\substack{t\in D \\ d|t}} \frac{\mu(t/d)\mu(t)}{g(t)} = \frac{f(d)\mu(d)}{Q} \sum_{\substack{t\in D \\ d|t}} \frac{1}{g(t)} .$$

(2.17) THEOREM. If Q and Λ are those of Theorem 2.13, then

$$S(A,P,n) \leq \frac{n}{Q} + \sum_{d_1,d_2 \in D} |\Lambda(d_1)\Lambda(d_2)R([d_1,d_2])| . \tag{2.18}$$

For $D = \{d: d|\Pi, d \leq z\}$ this leads to

$$S(A,P,n) \leq \frac{n}{Q} + z^2 \prod_{p \in P} \left(1 - \frac{1}{f(p)}\right)^{-2} . \tag{2.19}$$

Proof: Relation (2.18) follows immediately from (2.7) and (2.14). We thus only need to prove that

$$R = \sum_{d_1,d_2 \in D} |\Lambda(d_1)\Lambda(d_2)R([d_1,d_2])| \leq z^2 \prod_{p \in P} \left(1 - \frac{1}{f(p)}\right)^{-2} . \tag{2.20}$$

By Theorem 2.13

$$|\Lambda(d)| = \frac{f(d)}{Q} \sum_{\substack{t \in D \\ d|t}} \frac{1}{g(t)} \leq \frac{f(d)}{g(d)} \frac{1}{Q} \sum_{\delta \in D} \frac{1}{g(\delta)} = \frac{f(d)}{g(d)}$$

since $g(t) > 0$ and g is multiplicative. Also, by (2.4) and the fact that $f(k) \leq k$

$$|R([d_1,d_2])| \leq \frac{[d_1,d_2]}{f([d_1,d_2])} = \frac{d_1 d_2}{(d_1,d_2)} \cdot \frac{f((d_1,d_2))}{f(d_1)\cdot f(d_2)} \leq \frac{d_1 \cdot d_2}{f(d_1)\cdot f(d_2)} .$$

Hence

$$R \leq \sum_{d_1,d_2 \in D} \frac{f(d_1)}{g(d_1)} \cdot \frac{f(d_2)}{g(d_2)} \cdot \frac{d_1}{f(d_1)} \cdot \frac{d_2}{f(d_2)} \leq \left(\sum_{d \in D} \frac{d}{g(d)}\right)^2 \tag{2.21}$$

$$\leq z^2 \left(\sum_{d \in D} \frac{1}{g(d)}\right)^2 = z^2 Q^2 .$$

Now

$$Q = \sum_{d \in D} \frac{1}{g(d)} \leq \prod_{p \in P} (1 + \frac{1}{g(p)}), \tag{2.22}$$

and

$$1 + \frac{1}{g(p)} = 1 + \frac{1}{f(p)-1} = \frac{f(p)}{f(p)-1} = (1 - \frac{1}{f(p)})^{-1}.$$

Thus

$$R \leq z^2 \prod_{p \in P} (1 - \frac{1}{f(p)})^{-2}.$$

This finishes the proof.

In the proof of Theorem 2.17 we could have combined (2.18) with (2.21) to obtain

$$S(A,P,n) \leq \frac{n}{Q} + z^2 Q^2.$$

The reason we replaced the Q in the second term above by a simpler expression (simpler than the expression defining Q, that is) is that the size of z^2Q^2 is determined mostly by z, and we do not lose much by using the estimate (2.22) for Q. The size of Q is critically important, however, for the main term. The main difficulty in the applications we will be discussing in the next chapter will be in finding a good lower bound for Q.

Notes on Chapter 2.

An obvious way to generalize the results of this chapter is to drop the restriction that the a_v be generated by a polynomial. In fact, the only place where we used this property of the a_v was in proving (2.3). If we were to start with some general sequence $\{a_v\}$ and defined, say,

$$R(d) = \sum_{\substack{v \le n \\ d \mid a_v}} 1 - \frac{n}{f(d)}$$

for some multiplication function f satisfying $1 < f(p) \le p$ for all $p \in P$, then Theorem 2.13 and the inequality (2.18) would still be valid (with the R(d) defined as above). Whether effective use could be made of these results would then depend on whether the function f can be chosen so as to make the second term on the right side of (2.18) small, which would be equivalent to making R(d) small on the average. This can, in fact, be done for many sequences. Perhaps the most important cases are those of values of polynomials at primes. Going back to the example of twin primes that was treated in Chapter 1, we could, instead of taking $a_v = (2n-1)(2n+1)$, consider $a_v = p_v + 2$, where p_v is the v-th prime, with $v = 1, \ldots, n = \pi(x)$ for some x. Then we would have

$$\sum_{\substack{v \le n \\ d \mid a_v}} 1 = |\{p;\ p \le x \text{ and } p + 2 \equiv 0 \pmod{d}\}| = \pi(x,d,-2).$$

It would then be natural to write for d odd

$$\sum_{\substack{v \le n \\ d \mid a_v}} 1 = \frac{li(x)}{\varphi(d)} + \{\pi(x,d,-2) - \frac{li(x)}{\varphi(d)}\}.$$

(Notice that we are not literally following the suggestion at the beginning of this note.) In place of (2.7) we would then obtain (with $2 \notin P$)

$$S(A,P,\pi(x)) \leq \mathrm{li}(x) \sum_{d|\Pi} \frac{\lambda(d)}{\varphi(d)} + \sum_{d|\Pi} \lambda(d)\{\pi(x,d,-2) - \frac{\mathrm{li}(x)}{\varphi(d)}\}.$$

We could then apply Theorem 2.13 and obtain an inequality analogous to (2.18); namely

$$S(A,P,\pi(x)) \leq \frac{\mathrm{li}(x)}{Q} + \sum_{d_1,d_2 \in D} |\Lambda(d_1)\Lambda(d_2)\{\pi(x,d,-2) - \frac{\mathrm{li}(x)}{\varphi(d)}\}|$$

where the Q and Λ are defined in Theorem 2.13. It is here that the large sieve becomes very useful. Through results such as Bombieri's theorem (Chapter 8), which says that the terms $\pi(x,d,-2) - \frac{\mathrm{li}(x)}{\varphi(d)}$ are small on the average, it enables us to conclude that the second term above is small. For further discussion the reader is referred to Richert [3].

3
APPLICATIONS OF SELBERG'S SIEVE

This chapter discusses applications of Selberg's sieve. Our notation is the same as in Chapter 2; that is, we will work with a sequence $A = \{a_\nu\}$ given by $a_\nu = h(\nu)$, where h is an integer-valued polynomial. $N(d)$ will denote the number of solutions of $h(\nu) \equiv 0 \pmod{d}$ for $\nu = 1, \ldots, d$, P will be a set of primes such that $p \leq z$ and $0 < N(p) < p$ for all $p \in P$, $\Pi = \prod_{p \in P} p$, $f(d) = \frac{d}{N(d)}$ for $d \mid \Pi$. We seek estimates of $S = S(A,P,n)$, the number of elements of $\{a_1, \ldots, a_n\}$ which are not divisible by any $p \in P$. Theorem 2.17 then states that

$$S \leq \frac{n}{Q} + z^2 \prod_{p \in P} \left(1 - \frac{1}{f(p)}\right)^{-2}, \tag{3.1}$$

where $Q = \sum_{\substack{d \mid \Pi \\ d \leq z}} \frac{1}{g(d)}$ and $g(d) = \sum_{\ell \mid d} \mu(\ell) f(d/\ell)$. In all applications we follow the same basic procedure; namely, choose the polynomial h, find f, estimate Q from below, choose z so as to minimize the right side of (3.1), and complete the estimate.

Our first application is to primes in arithmetic progressions. We consider the polynomial $h(\nu) = \ell + k\nu$, where ℓ and k are relatively prime integers. Then

$$N(p) = \begin{cases} 1 & \text{if } p \nmid k, \\ 0 & \text{if } p \mid k. \end{cases}$$

We define $P = \{p \leq z;\ p \nmid k\}$, so that $f(p) = \frac{p}{N(p)} = p$ for $p \in P$. Hence for $d \mid \Pi$ we obtain

$$g(d) = \sum_{\ell|d} \mu(\ell) f(d/\ell) = \sum_{\ell|d} \mu(\ell)\,\frac{d}{\ell} = d \prod_{p|d} (1 - \frac{1}{p}) = \varphi(d).$$

Since the greatest squarefree divisor of every natural number $d \leq z$ which is relatively prime to k divides Π, we find

$$(3.2) \qquad Q = \sum_{\substack{d|\Pi \\ d\leq z}} \frac{1}{\varphi(d)} = \sum_{\substack{d|\Pi \\ d\leq z}} \prod_{p|d} (\frac{1}{p} + \frac{1}{p^2} + \cdots) > \sum_{\substack{d\leq z \\ (d,k)=1}} \frac{1}{d}\,.$$

To estimate the last sum, we use the following result:

(3.3) **LEMMA.** For $y \geq 1$ and any positive integer K we have

$$\sum_{\substack{m\leq y \\ (m,K)=1}} \frac{1}{m} \geq \frac{\varphi(K)}{K} \log y.$$

Proof: We have

$$\Big(\prod_{p|K} (1 - \frac{1}{p})^{-1}\Big) \sum_{\substack{m\leq y \\ (m,K)=1}} \frac{1}{m} = \Big(\prod_{p|K} (1 + \frac{1}{p} + \frac{1}{p^2} + \cdots)\Big) \sum_{\substack{m\leq y \\ (m,K)=1}} \frac{1}{m} \geq \sum_{m\leq y} \frac{1}{m} > \log y.$$

Hence

$$\sum_{\substack{m\leq y \\ (m,K)=1}} \frac{1}{m} \geq \log y \cdot \prod_{p|K} (1 - \frac{1}{p}) = \frac{\varphi(K)}{K} \log y.$$

Using the above estimate we now deduce from (3.2) that

$$Q \geq \frac{\varphi(k)}{k} \log z.$$

Also,

$$z^2 \prod_{p\in P}\left(1-\frac{1}{f(p)}\right)^{-2} = z^2 \prod_{p\in P}\left(1-\frac{1}{p}\right)^{-2} \le z^2 \prod_{p\le z}\left(1-\frac{1}{p}\right)^{-2} \ll z^2\log^2 z\ .$$

Therefore (3.1) now yields

$$S \ll \frac{k}{\varphi(k)}\ \frac{n}{\log z} + z^2\log^2 z \quad \text{for} \quad z \ge 2, \tag{3.4}$$

when the constant implied by the $\ll$ notation is absolute; that is, it is independent of k, ℓ, n, and z. To minimize the right side of (3.4) we now choose $z = n^{1/2}/\log^2 n$. Then for n sufficiently large we will have $z \ge 2$ and $\log n \ll \log z$. Therefore we finally obtain

$$S \ll \frac{k}{\varphi(k)}\ \frac{n}{\log n} \tag{3.5}$$

for n sufficiently large. We now use this result to prove a Brun-Titchmarsh type estimate on primes in arithmetic progressions.

Let $\pi(x,k,\ell)$ be the number of primes $p \le x$ such that $p \equiv \ell \pmod k$.

(3.6) **THEOREM.** If x and y are real numbers, k and ℓ integers satisfying $1 \le k < y \le x$, $(k,\ell) = 1$, then

$$\pi(x,k,\ell) - \pi(x-y,k,\ell) \ll \frac{y}{\varphi(k)\log(y/k)}$$

where the implied constant is absolute.

Proof: Let m be the largest integer such that $m \le x - y$, $m \equiv \ell \pmod k$, and let $n = [\frac{x-m}{k}]$, so that $n \le y/k + 1$. Then the integers a such that

$x - y < a \leq x$ and $a \equiv \ell \pmod{k}$ are precisely $m + k$, $m + 2k$, ..., $m + nk$. Since $(m,k) = (\ell,k) = 1$, we find from (3.5) that of these

$$\ll \frac{k}{\varphi(k)} \frac{y/k}{\log y/k}$$

are not divisible by any prime $p \leq n^{1/2}/\log^2 n$, provided y/k is sufficiently large. But then

$$\pi(x,k,\ell) - \pi(x-y,k,\ell) \ll \frac{k}{\varphi(k)} \frac{y/k}{\log(y/k)} + (y/k)^{1/2}/(\log y/k)^2$$

$$\ll \frac{y}{\varphi(k)\log(y/k)} ,$$

again provided that y/k is greater than some constant. But for y/k bounded, this result is trivially true.

(3.7) COROLLARY. If x is a real number, k and ℓ integers such that $1 \leq k < x$, $(k,\ell) = 1$, then

$$\pi(x,k,\ell) \ll \frac{x}{\varphi(k)\log(x/k)} .$$

Proof: Take $x = y$ in Theorem 3.6.

It might be expected at first that the above estimates could be easily derived from the prime number theorem for arithmetic progressions, which gives an asymptotic formula for $\pi(x,k,\ell)$. In general that is not the case, however. Even under the assumption of the very powerful (and unproved) generalized Riemann hypothesis, we could only conclude that for $(k,\ell) = 1$

$$\pi(x,k,\ell) = \frac{\ell i(x)}{\varphi(k)} + O(x^{1/2} \log x). \tag{3.8}$$

If x is much larger than y^2, for example, (3.8) implies only that $\pi(x,k,\ell) - \pi(x-y,k,\ell) = O(x^{1/2} \log x)$, while if k is much larger than $\sqrt{x}$, it would only imply that $\pi(x,k,\ell) = O(x^{1/2} \log x)$. These results are much weaker than the estimates (3.6) and (3.7) (in many cases weaker even than the estimates one obtains by considering the total number of integers n, $n \equiv \ell \pmod k$, that are in the appropriate interval).

The main purpose of proving first the estimates (3.5) - (3.7) was to acquaint the reader with the methods used in applying Selberg's sieve. We will now prove a very general theorem which will include those estimates as special cases.

In our first application of Selberg's sieve we have considered one linear factor $k\nu + h$, while in the second part of Chapter 1 we have applied Brun's sieve to the product of the two linear factors ν and $\nu + 2$. We will now generalize to the case of s linear factors $a_1\nu+b_1, \ldots, a_s\nu+b_s$, where the a_i and b_i are arbitrary integers. More specifically, we will obtain upper estimates as to how often all of the factors $|a_i\nu + b_i|$ are prime at the same time, and then will apply these estimates to several important problems of number theory.

To obtain the desired estimates we will consider $h(\nu) = (a_1\nu+ b_1) \cdots (a_s\nu+ b_s)$. To avoid the case where one of the linear factors of h has a constant prime divisor we will require that $a_i \neq 0$ and $(a_i,b_i) = 1$ for $i = 1, \ldots, s$. More generally, we will require that h as a whole

should not have a constant prime divisor; that is, we will require that $N(p) < p$ for all primes p. (A non-trivial case when this condition is violated is given by $h(\nu) = \nu(\nu+2)(\nu+4)$. Here $N(3) = 3$, and therefore 3, 5, 7 is the only prime triplet of the form p, $p+2$, $p+4$.) Verifying that this condition is satisfied is quite easy, since $N(p) \leq s$ for all primes p (as will be shown in Lemma 3.9), and therefore only the primes $p \leq s$ have to be checked. Finally, we will require that $a_i b_j - a_j b_i \neq 0$ for $i \neq j$, since we would not get any additional information by allowing one of the linear factors to be a multiple of another (in view of the requirement that $(a_i, b_i) = 1$, the only way $a_i b_j - a_j b_i = 0$ can happen is if $a_i \nu + b_i = \pm(a_j \nu + b_j)$).

Our next step is to estimate $N(p)$ for a prime p.

(3.9) **LEMMA.** We have $N(p) \leq s$ for all primes p. Moreover, $N(p) < s$ if and only if $p \mid E$, where

$$E = \prod_{i=1}^{s} a_i \cdot \prod_{i<j} (a_i b_j - a_j b_i).$$

Proof: Let p be a fixed prime. Consider any linear factor $a_i \nu + b_i$. If $p \mid a_i$, then

$$a_i \nu + b_i \equiv 0 \pmod{p} \tag{3.10}$$

has no solution, since $(a_i, b_i) = 1$. If $p \nmid a_i$, then (3.10) has exactly one solution for $\nu = 1, \ldots, p$. Since $h(\nu) \equiv 0 \pmod{p}$ implies that (3.10) holds for some i, and (3.10) has at most one solution for each i, we must have $N(p) \leq s$.

Let us now investigate the conditions under which $N(p) < s$. We have already seen that this happens if $p|a_i$ for some i. Another (and the only remaining) way this can occur is if two of the linear factors are divisible simultaneously by p; that is, if there are i,j with $i \neq j$ such that

$$a_i\nu + b_i \equiv 0 \pmod{p}$$

and

$$a_j\nu + b_j \equiv 0 \pmod{p}$$

for some integer ν. However, this occurs if and only if $p|(a_ib_j - a_jb_i)$. This finishes the proof.

We define $P = \{p \leq z;\ N(p) \geq 1\}$ to be our set of sieving primes. Then $\Pi = \prod_{\substack{p\leq z \\ N(p)>0}} p$, and $S = S(A,P,n)$ is the number of positive integers $m \leq n$ such that none of $a_im + b_i$ for $i = 1, \ldots, s$ is divisible by a prime $p \leq z$. As in the previous application of Selberg's sieve, our main task will be to estimate $Q = \sum_{\substack{d|\Pi \\ d\leq z}} \frac{1}{g(d)}$ from below.

Let $P_k = \prod_{\substack{p<z \\ N(p)=k}} p$ for $k = 0, 1, \ldots, s-1$. We write each natural number m as $m = \prod_{p|m} p^{m_p}$; that is, m_p will be the exponent of p in the prime power expansion of m. We also define $m^{(i)} = \prod_{\substack{p|m \\ N(p)=i}} p^{m_p}$, for

$i = 0, \ldots, s$, so that $m = m^{(0)} m^{(1)} \cdots m^{(s)}$, where $p|m^{(i)}$ if and only if $p|m$ and $N(p) = i$ (and $m^{(i)} = 1$ if there are no primes p such that $p|m$ and $N(p) = i$).

If k is a positive integer such that $k|\Pi$ (and thus k is squarefree), then

$$g(k) = f(k) \sum_{d|k} \mu(d) \frac{1}{f(d)} = f(k) \prod_{p|k} (1 - \frac{1}{f(p)})$$

$$= \frac{k}{N(k)} \prod_{p|k} (1 - \frac{N(p)}{p}).$$

Therefore, since Π is the product of all primes $\leq z$ which do not divide P_0,

$$(3.11) \qquad Q = \sum_{\substack{d|\Pi \\ d\leq z}} \frac{1}{g(d)} = \sum_{\substack{d|\Pi \\ d\leq z}} \frac{N(d)}{d} \prod_{p|d} (1 - \frac{N(p)}{p})^{-1}$$

$$= \sum_{\substack{d|\Pi \\ d\leq z}} \prod_{p|d} (\frac{N(p)}{p} + \frac{N^2(p)}{p^2} + \cdots)$$

$$\geq \sum_{\substack{d\leq z \\ (d,P_0)=1}} \frac{1}{d} \prod_{p|d} (N(p))^{d_p}.$$

Let us now define $d_t(m)$ for positive integers t,m to be the number of representations of m as a product of t positive integers, where two such representations $k_1 \cdots k_t$ and $k_1' \cdots k_t'$ are to be regarded as

identical only if $k_i = k_i'$ for $i = 1, \ldots, t$. (Thus, for example, $2\cdot2\cdot3$ and $2\cdot3\cdot2$ would be regarded as different representations of 12.) It is clear that for a fixed t, d_t is a multiplicative function.

(3.12) **LEMMA.** $d_t(p^a) \leq t^a$ for all primes p.

Proof: The lemma is true for $t = 1$, since $d_1(p^a) = 1 = 1^a$. Assume that it is true for t. Since each product of $t + 1$ factors is a product of one factor with the product of the remaining t factors, we obtain

$$d_{t+1}(p^a) = \sum_{0 \leq b \leq a} d_t(p^b) \leq \sum_{0 \leq b \leq a} t^b \leq (t+1)^a.$$

Hence the lemma is proved by induction.

(3.13) **LEMMA.**
$$\sum_{\substack{m \leq z \\ (m, P_0)=1}} \frac{1}{m} \prod_{i=1} d_i(m^{(i)}) \geq \prod_{i=1} \sum_{\substack{k \leq z^{1/s} \\ (k, P_0 \cdots P_{i-1})=1}} \frac{1}{k}.$$

Proof: Let $\frac{1}{k_i}$ be a term appearing in $\sum_{\substack{k \leq z^{1/s} \\ (k, P_0 \cdots P_{i-1})=1}} \frac{1}{k}$.

Then $k_i = k_i^{(i)} k_i^{(i+1)} \cdots k_i^{(s)}$. From the following array

$$
\begin{array}{lllll}
k_1 = k_1^{(1)} & k_1^{(2)} & k_1^{(3)} & \cdots\cdots\cdots & k_1^{(s)} \\
k_2 = & k_2^{(2)} & k_2^{(3)} & \cdots\cdots\cdots & k_2^{(s)} \\
k_3 = & & k_3^{(3)} & \cdots\cdots\cdots & k_3^{(s)} \\
\vdots & & & & \vdots \\
\hline
m = m^{(1)} & m^{(2)} & m^{(3)} & \cdots\cdots\cdots & m^{(s)}
\end{array}
\tag{3.14}
$$

we see that if $m = k_1 k_2 \cdots k_s$, then

$$
m^{(i)} = k_1^{(i)} \; k_2^{(i)} \cdots\cdots k_i^{(i)} \quad \text{for} \quad i = 1, \ldots, s. \tag{3.15}
$$

Now $k_i \leq z^{1/s}$ for each i, and so $m \leq z$. Likewise, $(k_i, P_0) = 1$ for all i implies $(m, P_0) = 1$. Thus if we expand the right side of (3.13), each of the terms $\dfrac{1}{k_1 k_2 \cdots k_s}$ will be equal to some term $\dfrac{1}{m}$ on the left side. Moreover, it is apparent from (3.15) that a given m appearing on the left side of (3.13) cannot be represented as $k_1 \cdots k_s$ in more than $d_1(m^{(1)}) \cdots d_s(m^{(s)})$ ways. Hence the inequality (3.13) is valid.

The remainder of our work is now very easy. From (3.11) and Lemma 3.12 we obtain

$$
Q \geq \sum_{\substack{m \leq z \\ (m, P_0)=1}} \frac{1}{m} \prod_{p \mid m} (N(p))^{m_p} = \sum_{\substack{m \leq p \\ (m, P_0)=1}} \frac{1}{m} \prod_{k=1}^{s} \prod_{\substack{p \mid m \\ N(p)=k}} k^{m_p}
$$

$$
\geq \sum_{\substack{m \leq z \\ (m, P_0)=1}} \frac{1}{m} \prod_{k=1}^{s} \prod_{\substack{p \mid m \\ N(p)=k}} d_k(p^{m_p}).
$$

But d_t is a multiplicative function for each fixed t, so

$$Q \geq \sum_{\substack{m \leq z \\ (m, P_0)=1}} \frac{1}{m} \prod_{k=1}^{s} d_k(m^{(k)})$$

$$\geq \prod_{i=1}^{s} \sum_{\substack{k \leq z^{1/s} \\ (k, P_0 \cdots P_{i-1})=1}} \frac{1}{k} \geq \prod_{i=1}^{s} \left(\frac{\varphi(P_0 \cdots P_{i-1})}{P_0 \cdots P_{i-1}} \frac{\log z}{s} \right)$$

utilizing Lemmas (3.13) and (3.3). Since E and $P_0 \cdots P_{s-1}$ have the same prime factors, we find

$$Q \geq \left(\frac{\log z}{s} \right)^s \prod_{i=0}^{s-1} \prod_{p \mid P_0 \cdots P_i} (1 - \frac{1}{p}) = s^{-s} (\log z)^s \prod_{p \mid E} (1 - \frac{1}{p})^{s - N(p)} .$$

This is the required lower bound for Q. To estimate the remainder term we note that $N(p) \leq s$ for all primes p, and therefore for $s < p$

$$(1 - \frac{1}{f(p)})^{-1} \leq (1 - \frac{s}{p})^{-1} \leq (1 - \frac{1}{p})^{-s}.$$

If $p \leq s$, then $N(p) \leq p - 1$ (since $N(p) < p$ for all p), and so

$$(1 - \frac{1}{f(p)})^{-1} = (1 - \frac{N(p)}{p})^{-1} \leq p.$$

Hence

$$\prod_{p \in P} (1 - \frac{1}{f(p)})^{-2} = \prod_{\substack{p \in P \\ p \leq s}} (1 - \frac{1}{f(p)})^{-2} \cdot \prod_{\substack{p \in P \\ s < p}} (1 - \frac{1}{f(p)})^{-2}$$

$$\leq \left(\prod_{p \leq s} p \right)^2 \cdot \left(\prod_{\substack{p \in P \\ s < p}} (1 - \frac{1}{p}) \right)^{-2s}$$

$$\leq \Big(\prod_{p \leq s} p^2\Big) \cdot \Big(\prod_{p \leq z} (1 - \tfrac{1}{p})\Big)^{-2s}$$

$$\leq c'(\log z)^{2s} \quad \text{for } z \geq 2,$$

where c' is a constant depending only on s. From (3.1) we now obtain

$$S \leq \frac{s^s}{\prod_{p|E} (1 - \frac{1}{p})^{s-N(p)}} \cdot \frac{n}{(\log z)^s} + c' \cdot z^2 \cdot (\log z)^{2s} \quad \text{for} \quad z \geq 2.$$

If we now choose $z = n^{1/2} (\log n)^{-2s}$, then we find that

$$S \leq c''(s) \prod_{p|E} (1 - \frac{1}{p})^{-s+N(p)} \cdot \frac{n}{(\log n)^s} \tag{3.16}$$

for n sufficiently large (bound depending on s only) and c'' a function of s alone.

Because of its definition, S counts all those ν for which none of the factors $a_i\nu + b_i$ is divisible by any prime $\leq z$, but it does not count those ν for which each of $|a_i\nu + b_i|$ is a prime if at least one of them is $\leq z$. However, for a given linear factor $a_i\nu + b_i$, $|a_i\nu + b_i| \leq z$ has at most $2z + 1$ solutions. Hence the total number of ν such that $|a_i\nu + b_i|$ is a prime for $0 = 1, \ldots, s$ which were sieved out is $\leq s(2z + 1)$. We now obtain

(3.17) THEOREM. Let s be a positive integer, and suppose that for $i = 1, \ldots, s$, a_i and b_i are integers such that $(a_i, b_i) = 1$ and

$E = \prod_{i=1}^{s} a_i \cdot \prod_{i<j} (a_i b_j - a_j b_i) \neq 0$. Let $N(p)$ be the number of solutions of $(a_1\nu + b_1) \ldots (a_s\nu + b_s) \equiv 0 \pmod{p}$ for $\nu = 1, \ldots, p$ for each prime p, and assume $N(p) < p$ for all primes p. Then the number of positive integers $m \leq n$ such that each of $|a_i m + b_i|$ $(i = 1, \ldots, s)$ is prime is

$$\leq c(s) \frac{n}{(\log n)^s} \cdot \prod_{p \mid E} \left(1 - \frac{1}{p}\right)^{-s+N(p)} .$$

Proof: For n larger than a certain bound (depending only on s) this follows from (3.16) and the remark preceding the statement of the theorem. But for n bounded it is trivially true!

We can make a few simple deductions from this theorem.

(3.18) COROLLARY. The number $T(x)$ of twin primes $\leq x$ satisfies

$$T(x) \ll \frac{x}{(\log x)^2} \quad \text{for} \quad x \geq 2.$$

Proof: Take $h(\nu) = \nu(\nu + 2)$. Then $|E| = 2$ and $N(p) < p$ for all primes p since $N(2) = 1$. Hence the number of primes $p \leq [x]$ such that $p + 2$ is also a prime is

$$\leq 2\, c(2) \frac{[x]}{(\log [x])^2} ,$$

and this implies the corollary.

It has been conjectured that there are infinitely many primes p such that $p + 2$, $p + 6$, and $p + 8$ are also primes, but just as in the case of

twin primes nothing is known for certain. The next corollary, however, gives an upper bound for their density.

(3.19) COROLLARY. The number of primes $p \leq x$ such that $p + 2$, $p + 6$, and $p + 8$ are also primes is

$$\ll \frac{x}{(\log x)^4} \quad \text{for } x \geq 2.$$

Proof: Take $h(\nu) = \nu(\nu + 2)(\nu + 6)(\nu + 8)$. Then $|E| = 2^9 \cdot 3^2$, and $N(2) = 1$, $N(3) = 2$, proving $N(p) < p$ for all primes p.

For our last application we will consider the number of representations of an even positive integer as a sum of two primes. If Golbach's conjecture is correct, then this number is always positive (except for 2 and 4). We are going to give an upper bound for it.

(3.20) COROLLARY. Let a be a positive even integer, and let $T(a)$ be the number of primes p such that $a - p = q$ is also a prime. Then

$$T(a) \ll \frac{a}{(\log a)^2} \prod_{p \mid a} \left(1 + \frac{1}{p}\right) \quad \text{for } a \geq 2.$$

Proof: We consider $h(\nu) = (\nu + 2)(a - 2 - \nu)$ for $\nu = 1, \ldots, a-5$. $|E| = a$, $N(2) = 1$, and in general $N(p) = 1$ for $p|a$. The corollary follows easily from Theorem 3.17 for a sufficiently large, and it is trivially true for a bounded.

As an important application of the above estimates we will now prove the following important result.

(3.21) THEOREM. There is a constant c such that every integer ≥ 2 can be represented as a sum of not more than c prime numbers.

Proof: The basic idea is to prove that the integers representable as a sum of two primes, together with 1, form a sequence of positive Schnirelmann density. The theorem will then follow by virtue of well-known results on addition of sequences.

Let n' denote those integers that can be represented as a sum of two primes. Then by the Cauchy-Schwarz inequality

$$\Big(\sum_{n\leq x} T(n)\Big)^2 \leq \Big(\sum_{n\leq x} T^2(n)\Big)\Big(\sum_{n'\leq x} 1\Big) . \tag{3.22}$$

Now

$$\sum_{n\leq x} T(n) \geq \sum_{\substack{p_1+p_2\leq x \\ p_1,p_2\leq x/2}} 1 = \pi^2(\tfrac{x}{2}) \gg \frac{x^2}{(\log x)^2} \quad \text{for } x \geq 4.$$

Also, by Corollary 3.20, we have

$$\sum_{n\leq x} T(n) \ll \frac{x^2}{(\log x)^4} \sum_{n\leq x} \prod_{p\mid n} (1+\tfrac{1}{p})^2$$

$$\leq \frac{x^2}{(\log x)^4} \sum_{n\leq x} \Big(\sum_{d\mid n} \tfrac{1}{d}\Big)^2 = \frac{x^2}{(\log x)^4} \sum_{\substack{d_1,d_2 \\ [d_1,d_2]\leq x}} \frac{1}{d_1,d_2} \sum_{\substack{n\leq x \\ n\equiv 0 (\mathrm{mod}\,[d_1,d_2])}} 1$$

$$\leq \frac{x^2}{(\log x)^4} \quad x \sum_{\substack{d_1,d_2 \\ [d_1,d_2]\leq x}} \frac{1}{d_1 d_2 [d_1,d_2]} + \frac{x^2}{(\log x)^4} \sum_{\substack{d_1,d_2 \\ [d_1,d_2]\leq x}} \frac{1}{d_1 d_2} .$$

In the second sum above we use the fact that each $k \le x$ can be expressed as $[d_1,d_2]$ in at most k ways, and each such expression contributes $\frac{1}{d_1 d_2} \le \frac{1}{k}$ to the sum. Therefore that sum is $\le x$. In the first term we sum over $k = (d_1,d_2)$ to find

$$\sum_{\substack{d_1,d_2 \\ [d_1,d_2]\le x}} \frac{1}{d_1 d_2 [d_1,d_2]} = \sum_{k\le x} \sum_{\substack{d_1,d_2 \\ [d_1,d_2]\le x \\ k=(d_1,d_2)}} \frac{k}{(d_1 d_2)^2} \le \sum_{k\le x} \frac{1}{k^3} \sum_{\substack{\delta_1,\delta_2\le x \\ (\delta_1,\delta_2)=1}} \frac{1}{\delta_1^2\delta_2^2}$$

$$\le \Big(\sum_{k\le x} \frac{1}{k^3}\Big)\Big(\sum_{\delta\le x} \frac{1}{\delta^2}\Big)^2 \ll 1.$$

Therefore

$$\sum_{n\le x} T^2(n) \ll \frac{x^3}{(\log x)^4},$$

and hence we find from (3.22) that

$$\sum_{n'\le x} 1 \ge \frac{\big(\sum_{n\le x} T(n)\big)^2}{\sum_{n\le x} T^2(n)} \gg \frac{x^4/(\log x)^4}{x^3/(\log x)^4} \ge x \quad \text{for } x \ge 4.$$

This implies that the sequence consisting of 1 and the numbers representable as a sum of two primes has positive Schnirelmann density, and therefore every positive integer is the sum of $\le c'$ primes and ones, for some constant c'. To prove the theorem we have to show that given $m \ge 2$ there is a representation of it which does not involve ones. This is clearly true if $m = 2$ or 3. Suppose therefore that $m \ge 4$ and that

$$m - 2 = \sum_{i \le k} 1 + \sum_{k < i \le C} p_i \quad \text{for some } C \le c',$$

is a representation of $m - 2$ that we have proved exists. If $k = 1$ or 0, then we write

$$m = 3 + \sum_{1 < i \le c} p_i, \quad \text{or} \quad m = 2 + \sum_{i \le c} p_i ,$$

respectively, and if $k \ge 2$ then we write the k ones as a sum of $\le k/2$ twos and threes, and obtain

$$m = 2 + \sum 2 + \sum 3 + \sum_{k < i \le c} p_i .$$

In any event we have found a representation of m as a sum of $\le c' + 1$ primes, and this completes the proof.

Notes on Chapter 3.

The presentation of this chapter largely follows Prachar [1; Chapter 2], who also gives several additional applications of Theorem 3.17. One problem which is dealt with neither in Prachar's book nor in this chapter is the explicit determination of constants in the inequalities above. The next chapter, on the other hand, will be denoted entirely to obtaining the best estimates of the Brun-Titchmarsh type that are known so far, with particular attention being paid to the constants and the relative sizes of the main term and the remainder. However, it has been found in many cases that the upper bounds given by Selberg's sieves are several times larger than either the asymptotic estimates obtained through analytic methods (where these exist) or the conjectured asymptotic estimates. As an example, it has been conjectured that $T(x)$, the number of twin prime pairs $\leq x$, satisfies

$$T(x) \sim 2C \frac{x}{(\log x)^2} \quad \text{as } x \to \infty,$$

where

$$C = \prod_{p \geq 3} \left(1 - \frac{1}{(p-1)^2}\right).$$

(For heuristic arguments supporting this conjecture see either the original paper of Hardy and Littlewood [1], or Hardy and Wright [1; Chapter 22], or Golomb [1].) The best upper estimates obtained so far are asymptotic to $8C\frac{x}{(\log x)^2}$. It is not clear how far this bound can be lowered. Selberg [2], [3] has actually proved that his sieve is in general incapable of giving the exact constants.

Material on the Schnirelmann density and the addition of sequences can be found in Niven and Zuckerman [1], Halberstam and Roth [1; Chapter 1], or Gelfond and Linnik [1; Chapter 1]. It is easy to see from the proof of Theorem 3.21 that we can obtain a numerical value for c. This was actually done by Schnirelmann [1] in his original proof (in which, however, he used Brun's sieve, since Selberg's method was not known at that time), but his value was very large. After many subsequent improvements Shapiro and Warga [1] showed that Theorem 3.21 holds with c = 20. (It should be mentioned, however, that in the meantime I. M. Virogradov [1] had proved that every sufficiently large integer is a sum of at most four primes.)

4

PRIMES IN ARITHMETIC PROGRESSIONS

This chapter uses Selberg's sieve to obtain new estimates of the Brun - Titchmarsh type due to Van Lint and Richert . These are the best estimates of this type known to date. They can be stated as follows:

(4.1) THEOREM. If x and y are real numbers, k and ℓ integers such that

$$1 \leq k < y \leq x, \quad (k,\ell) = 1,$$

then

$$(4.2) \qquad \pi(x,k,\ell) - \pi(x-y,k,\ell) < \frac{2y}{\varphi(k)\log(y/k)} \left(1 + \frac{8}{\log(y/k)}\right)$$

and

$$(4.3) \qquad \pi(x,k,\ell) - \pi(x-y,k,\ell) < \frac{3y}{\varphi(k)\log(y/k)} .$$

By taking x = y in the above relations we immediately obtain

(4.4) COROLLARY. If x is a real number and k and ℓ are integers such that $1 \leq k < x$, $(k,\ell) = 1$, then

$$(4.5) \qquad \pi(x,k,\ell) < \frac{2x}{\varphi(k)\log(x/k)} \left(1 + \frac{8}{\log(x/k)}\right)$$

and

$$(4.6) \qquad \pi(x,k,\ell) < \frac{3x}{\varphi(k)\log(x/k)} .$$

These results are of the same nature as those of (3.6) and (3.7). In proving the earlier estimates, however, as well as in our other applications of Selberg's sieve, we absorbed the remainder in the main term

by appropriate choice of z. We used the $\ll$ notation quite freely in this process, without determining the implied constants. Had we been more careful, we could have obtained an estimate similar to (4.2) (for y/k sufficiently large, at least), but with $\frac{8}{\log(y/k)}$ in the remainder replaced by $\frac{c \log\log(y/k)}{\log(y/k)}$ for some constant c. The improvement in the Van Lint - Richert result comes from a very careful treatment of the remainder term in Selberg's sieve. An important feature of the proof is the repeated application of the sieve; first a sieve estimate is proved (equation (4.13)) with a remainder somewhat smaller than we had previously, and then this estimate is used to reduce the size of the remainder even further (Lemma 4.16).

Our starting point will be, in essence, equation (2.18). However, (2.18) deals with the values of the polynomial $k\nu + \ell$ for n consecutive values of ν, while we are interested in the values of that polynomial which fall in the interval (x-y,x]. Therefore we derive first a slightly modified version of (2.18) which gives a better estimate in our case. It will be quite clear, however, that we are not doing anything basically new at this stage.

Let $|T|$ denote the number of elements of a finite set T. We define

$$S = S(k,\ell,x,y,z,K) = |\{n;\ x - y < n \leq x,\ n \equiv \ell \pmod{k},\ (n,\Pi) = 1\}|$$

for $1 < y \leq x$, $z > 1$, $(k,\ell) = 1$, $k|K$, where

$$\Pi = \Pi_K(z) = \prod_{\substack{p \leq z \\ p \nmid K}} p .$$

In particular, $(d,k) = 1$ for all $d|\Pi$. Therefore if $d|\Pi$, there is exactly one integer n in each residue class modulo kd such that $n \equiv \ell \pmod{k}$ and $n \equiv 0 \pmod{d}$. Hence

$$|\{n;\ x - y < n \leq x,\ n \equiv \ell \pmod{k},\ d|n\}| = \frac{y}{kd} + \theta, \text{ where } 0 \leq \theta \leq 1.$$

Therefore, if s is a sieving function defined by (2.8), in view of (2.9) we find

$$\sum_{\substack{x-y<n\leq x \\ n\equiv\ell(\mathrm{mod}\ k)}} s^2(n) = \sum_{\substack{x-y<n\leq x \\ n\equiv\ell(\mathrm{mod}\ k)}} \sum_{d|(n,\Pi)} \lambda(d) = \sum_{d|\Pi} \lambda(d) \sum_{\substack{x-y<n\leq x \\ n\equiv\ell(\mathrm{mod}\ k) \\ d|n}} 1$$

$$\leq \frac{y}{k} \sum_{d|\Pi} \frac{\lambda(d)}{d} + \sum_{d|\Pi} |\lambda(d)|$$

$$= \frac{y}{k} \sum_{d|\Pi} \frac{\lambda(d)}{d} + \sum_{d_1,d_2|\Pi} |\Lambda(d_1)\Lambda(d_2)|.$$

We again choose Λ to be a real-valued function on the divisors of Π such that $\Lambda(1) = 1$ and $\Lambda(d) = 0$ if $d > z$. Then $s^2(n) \geq 1$ whenever $(n,\Pi) = 1$, and therefore

$$(4.7) \qquad S \leq \sum_{\substack{x-y<n\leq x \\ n\equiv\ell(\mathrm{mod}\ k)}} s^2(n) \leq \frac{y}{k} \sum_{\substack{d|\Pi \\ d\leq z}} \frac{\lambda(d)}{d} + \sum_{\substack{d_1,d_2|\Pi \\ d_1,d_2\leq z}} |\Lambda(d_1)\Lambda(d_2)|$$

$$= \frac{y}{k} \sum_{\substack{d|\Pi \\ d\leq z}} \frac{\lambda(d)}{d} + \Big(\sum_{\substack{d|\Pi \\ d\leq z}} |\Lambda(d)| \Big)^2 .$$

In order to find the minimum of the first term on the right of (4.7) we apply Theorem 2.13, where for $d|\Pi$ we take $f(d) = d$ and define $g(d)$ by (2.11), so that $g(d) = \varphi(d)$. We then find from (2.14) and (4.7) that the main term is minimized when

$$\Lambda(d) = \frac{\mu(d)d}{Q} \sum_{\substack{t|\Pi \\ t\le z \\ t\equiv 0 \pmod d}} \frac{1}{\varphi(t)} ,$$

where

$$Q = \sum_{\substack{d|\Pi \\ d\le z}} \frac{1}{\varphi(d)} .$$

For this choice of Λ we obtain the estimate

$$S \le \frac{y}{kQ} + \Big(\sum_{\substack{d|\Pi \\ d\le z}} |\Lambda(d)| \Big)^2 . \tag{4.8}$$

This is the modification of (2.18) that we have been seeking. Had we applied (2.18) directly, we would have obtained a similar estimate, but with $\frac{y}{kQ}$ in (4.8) replaced by $(y/k + 1)/Q$.

For later convenience we define

$$H_k(x) = \sum_{\substack{n\le x \\ (n,k)=1}} \mu^2(n) \frac{\sigma(n)}{\varphi(n)} , \quad Q_k(x) = \sum_{\substack{n\le x \\ (n,k)=1}} \frac{\mu^2(n)}{\varphi(n)} , \quad x > 0.$$

Then

$$Q_k(x) = \sum_{d|K} \sum_{\substack{n\le x \\ (n,k)=1 \\ (n,K)=d}} \frac{\mu^2(n)}{\varphi(n)} = \sum_{d|K} \frac{\mu^2(d)}{\varphi(d)} \sum_{\substack{m\le x/d \\ (m,k)=(m,K/d)=(m,d)=1}} \frac{\mu^2(m)}{\varphi(m)}$$

$$= \sum_{d|K} \frac{\mu^2(d)}{\varphi(d)} Q_{Kk}\left(\frac{x}{d}\right) \quad \text{for } (K,k) = 1.$$

Since $Q_{Kk}(x)$ is an increasing function of x,

$$Q_k(x) \ge \sum_{d|K} \frac{\mu^2(d)}{\varphi(d)} Q_{Kk}\left(\frac{x}{K}\right) = \frac{K}{\varphi(K)} Q_{Kk}\left(\frac{x}{K}\right) \quad \text{if } (K,k) = 1.$$

Therefore

$$(4.9) \qquad H_k(x) = \sum_{\substack{n\le x \\ (n,k)=1}} \frac{\mu^2(n)}{\varphi(n)} \sum_{d|n} d = \sum_{\substack{d\le x \\ (d,k)=1}} \sum_{\substack{m\le x/d \\ (m,k)=1 \\ (m,d)=1}} \frac{\mu^2(md)}{\varphi(md)}$$

$$= \sum_{\substack{d\le x \\ (d,k)=1}} \frac{\mu^2(d)d}{\varphi(d)} Q_{kd}\left(\frac{x}{d}\right) \le Q_k(x) \sum_{\substack{d\le x \\ (d,k)=1}} \mu^2(d).$$

Also,

$$(4.10) \qquad Q_K(x) = \sum_{\substack{n\le x \\ (n,K)=1}} \frac{\mu^2(n)}{n} \prod_{p\,|\,n} \left(1 + \frac{1}{p} + \frac{1}{p^2} + \cdots\right) \ge \sum_{\substack{n\le x \\ (n,K)=1}} \frac{1}{n} \ge \frac{\varphi(K)}{K} \log x$$

by Lemma 3.3.

We now return to the task of estimating $S = S(k,\ell,x,y,z,K)$. In our new notation

$$(4.11)\qquad Q = \sum_{\substack{d\mid\Pi\\ d\le z}} \frac{1}{\varphi(d)} = \sum_{\substack{n\le z\\ (n,K)=1}} \frac{\mu^2(n)}{\varphi(n)} = Q_K(z)$$

and for $d\mid\Pi$

$$\Lambda(d) = \frac{\mu(d)d}{Q_K(z)} \sum_{\substack{t\mid\Pi\\ t\le z\\ t\equiv 0 \pmod d}} \frac{1}{\varphi(t)} = \frac{\mu(d)d}{Q_K(z)\varphi(d)} \sum_{\substack{u\mid\Pi\\ u\le z/d\\ (u,d)=1}} \frac{1}{\varphi(u)}$$

$$= \frac{\mu(d)d}{Q_K(z)\varphi(d)} \sum_{\substack{u\le z/d\\ (u,Kd)=1}} \frac{\mu^2(u)}{\varphi(u)} = \mu(d)\,\frac{d}{\varphi(d)}\,\frac{Q_{Kd}(z/d)}{Q_K(z)}\,.$$

Therefore $\Lambda(d) = 0$ for $d > z$, and

$$(4.12)\qquad \sum_{\substack{d\mid\Pi\\ d\le z}} |\Lambda(d)| = \frac{1}{Q_K(z)} \sum_{\substack{d\le z\\ (d,K)=1}} \mu^2(d)\,\frac{d}{\varphi(d)} \sum_{\substack{m\le z/d\\ (m,Kd)=1}} \frac{\mu^2(m)}{\varphi(m)}$$

$$= \frac{1}{Q_K(z)} \sum_{\substack{n\le z\\ (n,K)=1}} \frac{\mu^2(n)}{\varphi(n)} \sum_{d\mid n} d = \frac{H_K(z)}{Q_K(z)}\,.$$

Combining (4.8), (4.11), and (4.12) we obtain

$$S = S(k,\ell,x,y,z,K) \le \frac{y}{kQ_K(z)} + \frac{H_K^2(z)}{Q_K^2(z)}\,.$$

Equation (4.9) implies that $\left(\frac{H_K(z)}{Q_K(z)}\right)^2 \le z^2$. Hence

$$(4.13)\qquad S(k,\ell,x,y,z,K) \le \frac{y}{kQ_K(z)} + z^2.$$

This already represents an improvement over the estimates of Chapter 3, when we knew only that the remainder term was $\ll z^2(\log z)^2$. It turns out, however, that the bound on the remainder can be reduced still further. We will carry out this reduction (which will use (4.13) in a very important way) in the next two lemmas.

(4.14) **LEMMA.** If $p(k)$ denotes the greatest prime divisor of k (with the convention $p(1) = 1$), then for $x \geq 10^3$ and $p(k) \leq x$ we have

$$T_k(x) = \sum_{\substack{n \leq x \\ (n,k)=1}} 1 < \frac{15}{2} \frac{\varphi(k)}{k} x.$$

Proof: Let z satisfy $1 < z < x$ and define

$$k_1 = \prod_{\substack{p \leq z \\ p \mid k}} p \quad \text{and} \quad K = \prod_{\substack{p \leq z \\ p \nmid k}} p .$$

Then

$$T_{k_1}(x) = S(1,1,x,x,z,K).$$

Therefore by (4.13) and (4.10)

$$T_{k_1}(x) \leq \frac{x}{\log z \prod\limits_{\substack{p \leq z \\ p \nmid k}} (1 - \frac{1}{p})} + z^2.$$

Now $k_1 | k$ implies that $T_k(x) \leq T_{k_1}(x)$, and thus

$$T_k(x) \leq \frac{x}{\log z \prod\limits_{\substack{p \leq z \\ p \nmid k}} (1 - \frac{1}{p})} + z^2.$$

Since $p(k) \leq x$, we have

$$\frac{k}{\varphi(k)} \frac{T_k(x)}{x} \leq \frac{1}{\prod_{p \leq x} (1 - \frac{1}{p})} \left(\frac{1}{\log z} + \frac{z^2}{x} \right) .$$

It is known that for $x \geq 1$

$$\prod_{p \leq x} (1 - \frac{1}{p})^{-1} \leq e^{\gamma} \log(2x), \tag{4.15}$$

where γ is Euler's constant, and therefore by taking $z = (2x)^{1/3}$ we find

$$\frac{k}{\varphi(k)} \frac{T_k(x)}{x} \leq e^{\gamma} \left(3 + 2 \frac{\log(2x)}{(2x)^{1/3}} \right) .$$

By considering its derivative, we find that the expression on the right decreases for $2x > e^3$, and for $x = 10^3$ it is $< \frac{15}{2}$.

We might remark here that for a fixed k, $T_k(x)$ is asymptotic to $\frac{\varphi(k)}{k} x$. The importance of Lemma 4.14 comes from the fact that it holds even when x is small, and that it gives a constant valid for all k and x satisfying the hypotheses. The restriction $p(k) \leq x$ is quite natural, since the value of $\sum_{\substack{n \leq x \\ (n,k)=1}} 1$ is independent of prime factors of k which are greater than x, while $\frac{\varphi(k)}{k} = \prod_{p \mid k} (1 - \frac{1}{p})$ is a product over all prime factors of k. Without the restriction $p(k) \leq x$ we could take a sequence of values of k such that (for a fixed x) $T_k(x)$ would be constant, while $\frac{\varphi(k)}{k}$ would tend to zero.

(4.16) **LEMMA.** For $z \geq 10^3$ and h even

$$\frac{H_h^2(z)}{Q_h^2(z)} < 24 \frac{h}{\varphi(h)} \cdot \frac{z^2}{\log^2 z} .$$

Proof: Define

$$J_h(z) = \sum_{\substack{n \leq z \\ (n,h)=1}} \mu^2(n) \frac{\sigma^2(n)}{\varphi^2(n)} .$$

Then by the Cauchy-Schwartz inequality

$$H_h^2(z) \leq T_h(z) J_h(z). \tag{4.17}$$

Since both σ and φ are multiplicative, we have for squarefree n

$$\frac{\sigma^2(n)}{\varphi^2(n)} = \prod_{p \mid n} \frac{(p+1)^2}{(p-1)^2} = \prod_{p \mid n} \left(1 + \frac{4p}{(p-1)^2}\right) = \sum_{d \mid n} \frac{4^{\nu(d)} d}{\varphi^2(d)} ,$$

where $\nu(d)$ is the number of distinct prime factor of d. Now h is even, and therefore

$$J_h(z) \leq J_2(z) = \sum_{\substack{n \leq z \\ (n,2)=1}} \mu^2(n) \sum_{d \mid n} \frac{4^{\nu(d)} d}{\varphi^2(d)} \tag{4.18}$$

$$= \sum_{\substack{d \leq z \\ (d,2)=1}} \mu^2(d) \frac{4^{\nu(d)} d}{\varphi^2(d)} \sum_{\substack{m \leq z/d \\ (m,2d)=1}} \mu^2(m)$$

$$\leq z \sum_{\substack{d \leq z \\ (d,2)=1}} \mu^2(d) \frac{4^{\nu(d)} d}{\varphi^2(d)} = z \prod_{p > 2} \left(1 + \frac{4}{(p-1)^2}\right) < \frac{16}{5} z,$$

since

$$\prod_{p} \left(1 + \frac{4}{(p-1)^2}\right) = 15.9396\ldots .$$

Combining (4.10), (4.17), (4.18), and Lemma 4.14 we obtain

$$\frac{H_h^2(z)}{Q_h^2(z)} < \frac{\frac{15}{2} \cdot \frac{\varphi(h)}{h} \cdot z \cdot \frac{16}{5} z}{\frac{\varphi^2(h)}{h^2} \log^2 z} = 24 \frac{h}{\varphi(h)} \cdot \frac{z^2}{\log^2 z} ,$$

provided $p(h) \leq z$. But both $H_h(z)$ and $Q_h(z)$ are independent of those prime factors of h which are greater than z, while $\frac{h}{\varphi(h)}$ is increased by them, so that the above inequality is true even when $p(h) > z$.

Proof of Theorem 4.1: Let

$$\Delta(x,y,k,\ell) = \pi(x,k,\ell) - \pi(x-y,k,\ell).$$

Since for odd values of k the positive integers n such that $n \equiv \ell \pmod{k}$ are alternately even, and at most one of the even terms can be prime, we will reduce the proof to the case of even k. More precisely, we define

$$h = \begin{cases} k & \text{if } k \text{ even}, \\ 2k & \text{if } k \text{ odd} . \end{cases}$$

Then for a suitable ℓ_1 we will have

$$\Delta(x,y,k,\ell) \leq \Delta(x,y,h,\ell_1) + 1.$$

Since $S(h,\ell_1,x,y,z,h)$ counts at least those primes p which satisfy $x - y < p \leq x$, $p > z$, and $p \equiv \ell_1 \pmod{h}$, we have

$$\Delta(x,y,k,\ell) \leq S(h,\ell_1,x,y,z,h) + \pi(z,h,\ell_1) + 1 \text{ for } z > 1.$$

Therefore by (4.13) we find

$$\Delta(x,y,k,\ell) \leq \frac{y}{hQ_h(z)} + \frac{H_h^2(z)}{Q_h^2(z)} + \pi(z,h,\ell_1) + 1. \tag{4.19}$$

We define

$$u = \sqrt{y/k} .$$

If we take $z = 2$ in (4.19), then

$$\Delta(x,y,k,\ell) \leq \frac{y}{h} + 2 \leq \frac{y}{k} + 2 = u^2 + 2.$$

Therefore if $W = \frac{\varphi(k)\log(y/k)}{2y}\Delta(x,y,k\ell)$, we have

$$W \leq \log u\left(\frac{1}{2} + \frac{2}{u^2}\right) < \frac{3}{2} \quad \text{for } 1 < u \leq e^{2.9} .$$

Also, since

$$\pi(z,h,\ell_1) + 1 \leq \sum_{\substack{m\leq z \\ (m,2)=1}} \mu^2(m) \leq \frac{z-1}{2} \quad \text{for } z \geq 10,$$

we find from (4.19), (4.10), and (4.9)

$$W \leq \log u \left\{\frac{1}{\log z} + \frac{z^2}{4u^2}\right\} \quad \text{for } z \geq 10.$$

Defining ν by

$$u = \frac{\nu}{\sqrt{2}} e^{\nu}$$

we choose

$$z = e^{\nu} .$$

Then

$$(4.20) \qquad W \le \frac{\mathrm{Log}(\nu/\sqrt{2}) + \nu}{\nu} \left\{1 + \frac{1}{2\nu}\right\} \quad \text{for } \nu \ge \log 10.$$

By examining the derivative of the function on the right we find that it decreases for $\nu \ge \sqrt{2e}$ ($\sqrt{2e} > \log 10$), and for $\nu = \sqrt{2e}$ it is $< \frac{3}{2}$. This proves (4.3). It remains to prove (4.2).

Since $\varphi(h) = \varphi(k)$, (4.10) implies that

$$\frac{y}{hQ_h(z)} \le \frac{y}{\varphi(k)\log z}.$$

Also,

$$\pi(z,h,\ell_1) + 1 \le \frac{z}{k} + 2 \le \frac{k}{\varphi(k)}\, z \quad \text{for } z \ge 2.$$

Combining these inequalities and Lemma 4.16 (h is even) with (4.19) we find that for $z \ge 10^3$

$$(4.21) \qquad \frac{1}{2}\log(y/k)\left\{\frac{\varphi(k)\log(y/k)}{2y}\,\Delta(x,y,k,\ell) - 1\right\}$$

$$\le \log u \left\{\frac{\log u}{\log z} - 1 + 48\,\frac{\log u}{u^2}\cdot\frac{z^2}{\log^2 z} + \frac{\log u}{u^2}\cdot z\right\}.$$

Choosing

$$\log z = \log u - 2$$

we find that the right side of (4.21) is

$$\log u \left\{\frac{2}{\log u - 2} + \frac{48}{e^4}\cdot\frac{\log u}{(\log u - 2)^2} + \frac{\log u}{e^2 u}\right\}.$$

This function is decreasing in u, and for $u = e^{10}$ it is < 4, which proves (4.2) for $u \geq e^{10}$. If $u \leq e^{8}$, then (4.2) follows directly from (4.3). If $e^{8} < u < e^{10}$, then we use (4.20). It says that, in that range, we have

$$W < 1.4 < 1 + \frac{4}{\log u} ,$$

which is the desired estimate.

Notes on Chapter 4.

The presentation of this chapter follows closely the original paper of van Lint and Richert [1]. The only significant modification was made in the first part in order to show the connection with the general theory of Selberg's sieve.

The inequality (4.15) follows immediately from the inequality (3.30) of Rosser and Schoenfeld [1] for $x \geq e^{\frac{1}{\log 2}}$. For x smaller than this bound it is easily verified directly.

A large sieve estimate very similar to the one presented above has been recently obtained by Bombieri [4]. He showed that (4.2) holds with $O(1)$ in place of the 8 in the remainder term.

5
GALLAGHER'S SIEVE

Both Brun's and Selberg's sieve methods are applicable only to a relatively small class of sequences. In particular, only a relatively small number of residue classes modulo each prime can be sieved out of a sequence of consecutive integers. However, it is often desirable to find estimates for the case in which a large number of congruence classes is sieved out modulo each prime; for example, to investigate the density of integers which are not primitive roots for any prime in a given set, one would sieve out the $\varphi(p)$ residue classes of primitive roots modulo each of those primes p. Linnik's large sieve was invented, as its name suggests, especially for dealing with such problems. In its most refined form it is applicable to all integer sequences, and it gives upper bounds similar to those of Selberg's sieve where both methods can be used. Chapters 6 and 7 will be devoted to an exposition of the large sieve. This chapter presents a new and very elegant sieve due to Gallagher. Although very elementary, it is larger than the large sieve in the sense that it gives better results than the large sieve when the number of residue classes sieved out is very large.

We consider the case in which all but $g(q)$ residue classes modulo q are removed for a given set of prime powers q. Gallagher's basic result can be stated as follows:

(5.1) THEOREM. Let S be a finite set of prime powers, and let W be a subset of $\{M+1, \ldots, M+N\}$ (M and N any integers with $N > 0$) such that for any q in S all elements of W fall into at most $g(q)$ residue classes modulo q. If $Z = |W|$ (the number of elements of W) then

$$Z \leq \frac{\sum_{q \in S} \Lambda(q) - \log N}{\sum_{q \in S} \frac{\Lambda(q)}{g(q)} - \log N} \tag{5.2}$$

provided the denominator is positive*).

Proof: We define, for each q in S,

$$Z(q,h) = |\{n;\ n \in W \text{ and } n \equiv h \pmod q\}|.$$

As we will see in Chapter 6, the function $Z(q,h)$ arises quite naturally and plays a very important role in the large sieve method. For any given q in S we have

$$Z = \sum_{h=0}^{q-1} Z(q,h)$$

and therefore by applying the Cauchy-Schwarz inequality and the fact that at most $g(q)$ of the $Z(q,h)$ are nonzero we find that

$$Z^2 = \Big(\sum_{h=0}^{q-1} Z(q,h)\Big)^2 \leq g(q) \sum_{h=0}^{q-1} (Z(q,h))^2 .$$

---

*) Mangoldt's Λ-function is defined by $\Lambda(q) = \log p$ if $q = p^{\alpha}$ for some prime p and some $\alpha > 1$, and $\Lambda(q) = 0$ otherwise. One of its basic properties is that $\sum_{d|n} \Lambda(d) = \log |n|$ whenever $n \neq 0$.

Multiplying by $\frac{\Lambda(q)}{g(q)}$ and summing over S, we get

$$Z^2 \sum_{q\in S} \frac{\Lambda(q)}{g(q)} \leq \sum_{q\in S} \Lambda(q) \sum_{h=0}^{q-1} (Z(q,h))^2 . \tag{5.3}$$

Now

$$(Z(q,h))^2 = \Big(\sum_{\substack{n \in W \\ n \equiv h (\mathrm{mod}\ q)}} 1 \Big)^2 = \sum_{\substack{n,m \in W \\ n,m \equiv h (\mathrm{mod}\ q)}} 1 .$$

Therefore (5.3) implies

$$Z^2 \sum_{q\in S} \frac{\Lambda(q)}{g(q)} \leq \sum_{q\in S} \Lambda(q) \sum_{h=0}^{q-1} \sum_{\substack{n,m \in W \\ n,m \equiv h (\mathrm{mod}\ q)}} 1 = \sum_{q\in S} \Lambda(q) \sum_{\substack{n,m \in W \\ n \equiv m (\mathrm{mod}\ q)}} 1$$

$$= \sum_{|d|\leq N} \Big(\sum_{\substack{m,n\in W \\ m-n=d}} 1 \Big) \cdot \Big(\sum_{\substack{q|d \\ q\in S}} \Lambda(q) \Big)$$

$$= \Big(\sum_{\substack{m,n\in W \\ m=n}} 1 \Big) \cdot \Big(\sum_{q\in S} \Lambda(q) \Big) + \sum_{1\leq |d| \leq N} \Big(\sum_{\substack{m,n\in W \\ m-n=d}} 1 \Big) \cdot \log|d|$$

$$\leq Z \sum_{q\in S} \Lambda(q) + (Z^2 - Z) \cdot \log N .$$

Dividing by Z (the result is trivial if Z = 0) we obtain

$$Z \cdot \Big(\sum_{q\in S} \frac{\Lambda(q)}{g(q)} - \log N \Big) \leq \sum_{q\in S} \Lambda(q) - \log N$$

which implies (5.2) if $\sum_{q\in S} \frac{\Lambda(q)}{g(q)} > \log N$.

We now derive two consequences of Theoem 5.1, both better than the corresponding large sieve estimates.

(5.4) COROLLARY. If all but at most G residue classes modulo each prime power $q \in S$ are removed from the sequence M+1, ..., M+N, then the number Z of remaining integers satisfies

$$Z \leq G \quad \text{if} \quad \sum_{q \in S} \Lambda(q) > G^2 \log N \tag{5.5}$$

and

$$Z \leq 2G-1 \quad \text{if} \quad \sum_{q \in S} \Lambda(q) > 2G \log N. \tag{5.6}$$

Proof: We can assume that G is an integer (otherwise work with [G]). We rewrite (5.2) as

$$Z \leq \frac{G \sum_{q \in S} \Lambda(q) - G \log N}{\sum_{q \in S} \Lambda(q) - G \log N} = G + \frac{G^2 \log N - G \log N}{\sum_{q \in S} \Lambda(q) - G \log N}.$$

This implies that $Z < G+1$ if $\sum_{q \in S} \Lambda(q) > G^2 \log N$ and $Z < G + G$ if $\sum_{q \in S} \Lambda(q) > 2G \log G$. Since Z and G are integers, we obtain (5.5) and (5.6), respectively.

(5.7) THEOREM. Let $\epsilon > 0$ be given. Then the number of positive integers $n \leq N$ for which $\exp_p(n) \leq N^\theta$ for all primes $p \leq N^{\theta+\epsilon}$ is $\ll N^\theta$ for $0 \leq \theta \leq 1$, with the implied constant depending only on ϵ *).

*) $\exp_p(n)$ is the smallest positive integer k such that $n^k \equiv 1 \pmod p$ if $(p,n) = 1$; otherwise $\exp_p(n) = 0$.

Proof: The result is clear for $\theta = 0$. Therefore we assume from now on that $\theta > 0$.

For each prime $p \le y$ we remove all residue classes of exponent $> x$. Since there are $\varphi(k)$ classes of exponent k for each $k|(p-1)$, as well as the zero class (which is of exponent 0), we have

$$g(p) = 1 + \sum_{\substack{k|(p-1) \\ k\le x}} \varphi(k).$$

We need to evaluate $\sum_{p\le y} \frac{\log p}{g(p)}$. By the Cauchy-Schwarz inequality we find

$$\Big(\sum_{p\le y} \frac{\log p}{g(p)}\Big) \cdot \Big(\sum_{p\le y} g(p)\log p\Big) \ge \Big(\sum_{p\le y} \log p\Big)^2 . \tag{5.8}$$

By the prime number theorem

$$\sum_{p\le y} \log p \sim y \quad \text{as} \quad y \to \infty,$$

and so the right side of (5.8) is $\gg y^2$ for $y \ge 2$, say. Also

$$\sum_{p\le y} g(p)\log p = \sum_{p\le y} \log p + \sum_{p\le y} \sum_{\substack{k|(p-1) \\ k\le x}} \varphi(k)\log p$$

$$\le \sum_{p\le y} \log p + \log y \sum_{k\le x} \varphi(k)\pi(y,k,1).$$

By the Brun-Titchmarsh estimate (3.7)

$$\pi(y,k,1) \ll \frac{y}{\varphi(k)\log y}$$

if $k^{1+\epsilon} \leq y$ and $y \geq 2$, with the implied constant depending only on ϵ. Therefore if $x \geq 2$ and $y \geq x^{1+\epsilon}$, we find

$$\sum_{p \leq y} g(p) \log p \ll y + xy \ll xy$$

and thus

$$\sum_{p \leq y} \frac{\log p}{g(p)} \ll \frac{y}{x},$$

where the applied constants again depend only on ϵ. We now take $x = N^{\theta}$, $y = N^{\theta+\epsilon}$. Theorem 5.1 then says that

$$Z \ll \frac{N^{\theta+\epsilon}}{cN^{\epsilon} - \log N}$$

where c is some positive constant, provided $cN^{\epsilon} > \log N$ (again, the constants depend only on ϵ). Therefore

$$Z \ll N^{\theta}$$

for N sufficietly large. But the result is clearly true for bounded N.

Notes on Chapter 5.

The material of this chapter is drawn from Gallagher [3]. That paper contains also a comparison of the effectiveness of Gallagher's sieve and the large sieve as well as another interesting application of the above estimates. If for every prime power q, except perhaps for the powers of a finite number of primes, there is an integer $\alpha(q)$ such that $a \equiv b^{\alpha(q)} \pmod{q}$, then Gallagher proved that $a = b^{\alpha}$ for some integer α. (A somewhat stronger result had been proved previously by Schinzel [1].)

6

TRIGONOMETRIC POLYNOMIAL INEQUALITIES

The large sieve was invented by Linnik for the purpose of investigating certain sequences obtained by sieving out a relatively large number of congruence classes modulo each prime from a given set of primes. It was discovered soon afterwards, however, that the large sieve can be applied to all integer sequences (although it gives best estimates for sequences defined by a sieving process), giving quantitive estimates as to how much they deviate from uniform distribution into congruence classes to various moduli. Moreover, the inequalities used in proving these estimates lead to important results in analytic number theory, perhaps the most important so far being Bombieri's theorem on the average of the remainder term in the prime number theorem for arithmetic progressions. Bombieri's theorem is discussed in Chapter 8, while this chapter develops the trigonometric polynomial inequalities that are the basis of the large sieve method, and Chapter 7 utilizes those estimates to study integer sequences.

The underlying idea of the large sieve is to relate the properties of an integer sequence to the behavior of a trigonometric polynomial. The two very basic (and very easy) properties of the exponential function which make this process possible are the fact that if n, m, and q are integers, then $n \equiv m \pmod q$ is equivalent to $e(n/q) = e(m/q)$ *), and the equation

$$(6.1) \qquad \sum_{a=1}^{q} e\left(a\,\frac{n}{q}\right) = \begin{cases} q & \text{if } n \equiv 0 \pmod q, \\ 0 & \text{otherwise .} \end{cases}$$

*) We use the definition $e(t) = e^{2\pi i t}$.

(Equation (6.1) follows immediately from the formula for the sum of the first q terms of a geometric progression; the common ratio in our case is $e(n/q)$, which is 1 if $n \equiv 0 \pmod q$ and unequal to 1 otherwise). As a result, if $n_1 < \cdots < n_Z$ are integers, and we define for integers q and h

$$Z(q,h) = |\{i;\ 1 \le i \le Z,\ n_i \equiv h \pmod q\}| \tag{6.2}$$

and

$$S(x) = \sum_{i=1}^{Z} e(n_i x), \tag{6.3}$$

then we have

$$S(\tfrac{a}{q}) = \sum_{h=1}^{q} Z(q,h)e(h\,\tfrac{a}{q}) \tag{6.4}$$

for all integers a. Now if the sequence $n_1, \ldots, n_Z$ is evenly distributed modulo q (that is, if $Z(q,h) = Z/q$ for all h), then by (6.1) we have $S(a/q) = 0$ whenever $a \not\equiv 0 \pmod q$. If, on the other hand, all the n_i belong to the same congruence class modulo q, then $|S(a/q)| = Z$ for all integers a. Thus the values of $|S(x)|$ at the points a/q are somehow related to how evenly the n_i are distributed modulo q*). This fact by itself does not help us, since we do not have any way of estimating $|S(a/q)|$ without reference to the sequence defining S(x). It was

---

*) Note that values of S(x) at the point a/q can be used to determine Z(q,h) explicitly; using (6.1) we obtain

$$q\,Z(q,h) = \sum_{a=1}^{q} S(a/q)e(-h\,\tfrac{a}{q}).$$

discovered by Linnik, however, that useful upper estimates can be obtained for

$$(6.5) \qquad \sum_{p \le x} \sum_{a=1}^{p-1} |S(\tfrac{a}{p})|^2$$

which are largely independent of the nature of the sequence $n_1,\dots,n_Z$. (It is this fact that is responsible for the great generality of the large sieve.) That the expression (6.5) reflects how evenly the n_i are distributed modulo p for all primes $p \le X$ can be seen best from the following identity:

(6.6) LEMMA. If $q \ge 2$ is an integer, then

$$(6.7) \qquad \sum_{a=1}^{q-1} |S(\tfrac{a}{q})|^2 = q \sum_{h=1}^{q} (Z(q,h) - \tfrac{Z}{q})^2.$$

Proof: We write out the left side, using (6.4), as

$$\sum_{a=1}^{q-1} |S(\tfrac{a}{q})|^2 = \sum_{a=1}^{q-1} \Big(\sum_{h=1}^{q} Z(q,h) e(h \tfrac{a}{q}) \Big) \cdot \Big(\sum_{k=1}^{q} Z(q,h) e(-k \tfrac{a}{q}) \Big)$$

$$= \sum_{a=1}^{q-1} \sum_{h,k=1}^{q} Z(q,h) Z(q,k) e(\tfrac{a(h-k)}{q})$$

$$= \sum_{h,k=1}^{q} Z(q,h) Z(q,k) \sum_{a=1}^{q-1} e(\tfrac{a(h-k)}{q}).$$

Now by (6.1) we have

$$\sum_{a=1}^{q-1} e\left(\frac{a(h-k)}{q}\right) = \begin{cases} q-1 & \text{if } h \equiv k \pmod{q}, \\ -1 & \text{otherwise.} \end{cases}$$

Therefore

$$\begin{aligned} \sum_{a=1}^{q-1} \left|S\left(\frac{a}{q}\right)\right|^2 &= q \sum_{h=1}^{q} (Z(q,h))^2 - \sum_{h,k=1}^{q} Z(q,h)Z(q,k) \\ &= q \sum_{h=1}^{q} (Z(q,h))^2 - \Big(\sum_{h=1}^{q} Z(q,h)\Big)^2 \\ &= q \sum_{h=1}^{q} (Z(q,h))^2 - Z^2 \\ &= q \sum_{h=1}^{q} \left(Z(q,h) - \frac{Z}{q}\right)^2. \end{aligned}$$

Using the above lemma, we can now rewrite (6.5) in the form

$$\sum_{p \le X} p \sum_{h=1}^{p} \left(Z(p,h) - \frac{Z}{p}\right)^2 .$$

The so-called "variance" $\sum_{h=1}^{p} \left(Z(p,h) - \frac{Z}{p}\right)^2$ shows how far the n_i depart from an even distribution modulo p. Our upper bounds for (6.5) will enable us to conclude that if the n_i are sufficiently dense, then they cannot be very unevenly distributed modulo many primes. This will mean that if we start with a sequence that is unevenly distributed, then it cannot contain many elements. As an example, suppose that none of the n_i fall into any of f(p) residue classes for each prime $p \le X$.

Then $Z(p,h) = 0$ for $f(p)$ values of h, and therefore

$$Z^2 = \Big(\sum_{h=1}^{p} Z(p,h)\Big)^2 \leq (p - f(p)) \cdot \sum_{h=1}^{p} (Z(p,h))^2$$

by the Cauchy-Schwarz inequality. But then

$$\sum_{h=1}^{p} (Z(p,h) - \frac{Z}{p})^2 = \sum_{h=1}^{p} (Z(p,h))^2 - \frac{2Z}{p} \sum_{h=1}^{p} Z(p,h) + \frac{Z^2}{p}$$

$$\geq \frac{Z^2}{p - f(p)} - \frac{Z^2}{p} = \frac{Z^2}{p} \frac{f(p)}{p - f(p)} ,$$

and therefore

$$\sum_{p \leq X} \sum_{a=1}^{p-1} |S(\frac{a}{p})|^2 \geq Z^2 \sum_{p \leq X} \frac{f(p)}{p - f(p)} .$$

Since we will prove that the left side is $\ll (N + X^2)Z$ for $N = n_Z - n_1 + 1$, this will give the estimate

$$Z \ll \frac{N + X^2}{\sum_{p \leq X} \frac{f(p)}{p - f(p)}} . \tag{6.8}$$

The above discussion demonstrates how sums of the form (6.5) can be used to study integer sequences. The next chapter will be devoted mainly to proving modifications of (6.7) and (6.8), while the rest of this chapter will deal mostly with estimates for sums similar to (6.5). We will work with trigonometric polynomials more general than those defined by (6.3), and we will estimate sums of values of $|S(x)|^2$ more general than (6.5). These generalizations are necessary for the analytic number theory applications of the large sieve, and they do not cause any material

difference in the proofs. It turns out that the crucial feature of (6.5) is that it is a sum of $|S(x)|^2$ over points x which are well-spaced; that is, if a/p and b/p′ are two distinct points which appear in the sum (6.5), then

$$\left|\frac{a}{p} - \frac{b}{p'}\right| = \left|\frac{ap' - bp}{pp'}\right| \geq \frac{1}{pp'} \geq \frac{1}{X^2} .$$

We will work with trigonometric polynomials of the form

$$S(x) = \sum_{n=-K}^{K} a_n \, e(nx), \tag{6.9}$$

where K is a positive integer and the a_n are any numbers, real or complex. If $x_1, \ldots, x_R$ are real numbers satisfying *)

$$\|x_r - x_s\| \geq \delta > 0 \quad \text{for } r \neq s, \tag{6.10}$$

then we will obtain several estimates of the form

$$\sum_{r=1}^{R} |S(x_r)|^2 \leq D(K,\delta) \sum_{n=-K}^{K} |a_n|^2, \tag{6.11}$$

where $D(K,\delta)$ will be functions of K and δ only. The restriction that S be of the symmetric form (6.9) is only temporary, imposed to facilitate our proofs. Estimates for trigonometric polynomials of the form

$$\sum_{n=M+1}^{M+N} a_n \, e(nx)$$

will follow easily from estimates of the form (6.11).

---

*) We denote by $\|t\|$ the distance from t to the nearest integer, so that $\|t\| = \min_n |t-n|$.

Our first estimate is a beautiful result, on the lines of Linnik's original work, due to Gallagher. Its basic idea is to relate the sum

$$\sum_{r=1}^{R} |S(x_r)|^2 \tag{6.12}$$

to the integral $\int_0^1 |S(x)|^2 dx$. Since the exponential function satisfies the relation

$$\int_0^1 e(nx)dx = \begin{cases} 1 & \text{if } n = 0, \\ 0 & \text{otherwise,} \end{cases} \tag{6.13}$$

for integers n, we have

$$\int_0^1 |S(x)|^2 \, dx = \int_0^1 S(x)\overline{S}(x)dx = \int_0^1 \sum_{n,m=-K}^{K} a_n \overline{a_m} \, e((n-m)x)dx \tag{6.14}$$

$$= \sum_{n=-K}^{K} |a_n|^2.$$

The sum (6.12) is related to the integral (6.14), and in fact, if S is kept fixed and the x_i are evenly spaced (say $x_r = r\delta$ for $r = 1,\ldots,R$ with $\delta = \frac{1}{R}$) then

$$\sum_{r=1}^{R} \delta |S(x_r)|^2 \to \int_0^1 |S(x)|^2 dx = \sum_{n=-K}^{K} |a_n|^2 \quad \text{as } R \to \infty.$$

(This shows why one might expect $\delta^{-1}\Sigma |a_n|^2$ to appear in our estimates of (6.12).) In general our situation is different, however, since the x_r and R are fixed, and therefore we cannot argue using limits. What we will use is our knowledge of $S'(x)$. Since values of $S(x)$ in a neighborhood of x_r are determined by $S(x_r)$ and values of $S'(x)$ in that

neighborhood, we will estimate (6.10) in terms of the integral (6.12) and a function of $S'(x)$, and then complete the estimate.

(6.15) THEOREM. If $S(x)$ is defined by (6.9), and $x_1, \ldots, x_R$ are real numbers satisfying (6.10), then

$$\sum_{r=1}^{R} |S(x_r)|^2 \leq (\delta^{-1} + 2\pi K) \sum_{n=-K}^{K} |a_n|^2. \tag{6.16}$$

Proof: Since for any u

$$S^2(x_r) = S^2(u) + 2\int_u^{x_r} S'(t)S(t)dt,$$

we have

$$|S(x_r)|^2 \leq |S(u)|^2 + 2\left| \int_u^{x_r} |S'(t)S(t)|dt \right|.$$

We now integrate this inequality over the interval $I_r = (x_r - \frac{\delta}{2}, x_r + \frac{\delta}{2})$ to obtain

$$\delta|S(x_r)|^2 \leq \int_{I_r} |S(u)|^2 du + 2\int_{I_r} \left| \int_u^{x_r} |S'(t)S(t)|dt \right| du.$$

Now

$$\int_{I_r} \left| \int_u^{x_r} |S'(t)S(t)|dt \right| du = \int_{x_r}^{x_r + \frac{\delta}{2}} \left(\int_{x_r}^{u} |S'(t)S(t)|dt \right) du$$

$$+ \int_{x_r - \frac{\delta}{2}}^{x_r} \left(\int_u^{x_r} |S'(t)S(t)|dt \right) du$$

$$= \int_{x_r}^{x_r + \frac{\delta}{2}} |S'(t)S(t)|(x_r + \tfrac{\delta}{2} - t)dt + \int_{x_r - \frac{\delta}{2}}^{x_r} |S'(t)S(t)|(t - x_r + \tfrac{\delta}{2})dt$$

$$\leq \frac{\delta}{2} \int_{I_r} |S'(t)S(t)|dt$$

and thus

$$(6.17) \qquad \delta|S(x_r)|^2 \le \int_{I_r} |S(u)|^2 du + \delta \int_{I_r} |S'(t)S(t)|dt.$$

Since by (6.10) the intervals I_r are disjoint modulo 1 (that is, if $r \neq s$ then no point of I_r differs by an integer from a point of I_s) the integral of any positive function over the I_r will be not larger than the integral of that function over [0,1], provided that function is periodic with period 1. Since S is periodic with period 1, we find by summing (6.17) over r that

$$\delta \sum_{r=1}^{R} |S(x_r)|^2 \le \int_0^1 |S(t)|^2 dt + \delta \int_0^1 |S'(t)S(t)|dt.$$

The first integral is $\sum_{n=-K}^{K} |a_n|^2$. The second satisfies

$$\int_0^1 |S'(t)S(t)|dt \le \Big(\int_0^1 |S(t)|^2 dt \Big)^{1/2} \Big(\int_0^1 |S'(t)|^2 dt \Big)^{1/2}$$

$$= \Big(\sum_{n=-K}^{K} |a_n|^2 \Big)^{1/2} \Big(\sum_{n=-K}^{K} |2\pi n a_n|^2 \Big)^{1/2}$$

$$\le 2\pi K \sum_{n=-K}^{K} |a_n|^2$$

since

$$S'(t) = \sum_{n=-K}^{K} 2\pi i n a_n e(nt).$$

Hence

$$\delta \sum_{r=1}^{R} |S(x_r)|^2 \leq (1 + \delta 2\pi K) \sum_{n=-K}^{K} |a_n|^2 ,$$

which completes the proof.

We will now give two other estimates which were found by Bombieri and Davenport. The basic idea here is to write $S(x)$ as the convolution of two appropriately chosen functions, so that the value of S at any given point is defined by an integral over a neighborhood of that point. The proof is much more complicated than that of Gallagher's estimate, but the results that follow from it represent an important improvement in some applications. (Note, however, that for some values of $K\delta$ the estimates below are weaker than Gallagher's.) First we prove our auxiliary result.

(6.18) **LEMMA.** For $x > 0$ we have

$$\left| \int_x^\infty \frac{\sin t}{t} \, dt \right| < \frac{1}{x} .$$

Proof: We use contour integration. The integral is

$$\frac{1}{2i} \int_x^\infty \frac{e^{it} - e^{-it}}{t} \, dt = \frac{e^{ix}}{2i} \int_0^\infty \frac{e^{it}}{x+t} \, dt - \frac{e^{-ix}}{2i} \int_0^\infty \frac{e^{-it}}{x+t} \, dt.$$

Let us take $r > 0$ and consider the integral of $\frac{e^{iz}}{x+z}$ over the path consisting of

(a) 0 to r along the x-axis,

(b) r to ri along the circle $|z| = r$,

(c) ri to 0 along the y-axis.

Since the integrand has no poles inside the path of integration, the integral is zero. Thus

$$\int_0^r \frac{e^{it}}{x+t}\,dt + i\int_r^0 \frac{e^{-t}}{x+it}\,dt + ir\int_0^{\pi/2} \frac{e^{ire^{i\theta}}}{x+re^{i\theta}}\,d\theta = 0.$$

Now $\sin\theta \geq \frac{2}{\pi}\theta$ for $0 \leq t \leq \frac{\pi}{2}$, and therefore

$$\left|ir\int_0^{\pi/2} \frac{e^{ire^{i\theta}}}{x+re^{i\theta}}\,d\theta\right| \leq r\int_0^{\pi/2} \frac{e^{-r\sin\theta}}{r}\,d\theta$$

$$\leq \int_0^{\pi/2} e^{-\frac{2}{\pi}\theta r}\,d\theta$$

$$\leq \frac{\pi}{2r}(1 - e^{-r}) \to 0 \text{ as } r \to \infty.$$

Hence

$$\int_0^\infty \frac{e^{it}}{x+t}\,dt = i\int_0^\infty \frac{e^{-t}}{x+it}\,dt.$$

Similarly we can move the integral of $\frac{e^{-it}}{x+it}$ to the negative imaginary axis to obtain

$$\int_0^\infty \frac{e^{-it}}{x+t}\,dt = i\int_0^\infty \frac{e^{it}}{x-it}\,dt.$$

Therefore

$$\left|\int_x^\infty \frac{\sin t}{t}\,dt\right| = \left|\frac{1}{2}e^{ix}\int_0^\infty \frac{e^{-t}}{x+it}\,dt + \frac{1}{2}e^{-ix}\int_0^\infty \frac{e^{-t}}{x-it}\,dt\right|$$

$$\leq \frac{1}{2}\int_0^\infty \frac{e^{-t}}{x}\,dt + \frac{1}{2}\int_0^\infty \frac{e^{-t}}{x}\,dt = \frac{1}{x}.$$

(6.19) **THEOREM.** If $S(x)$ is defined by (6.9), and $x_1, \ldots, x_R$ are real numbers satisfying (6.10), then

$$\sum_{r=1}^{R} |S(x_r)|^2 \leq 2 \max(2K, \delta^{-1}) \sum_{n=-K}^{K} |a_n|^2 \tag{6.20}$$

and

$$\sum_{r=1}^{R} |S(x_r)|^2 \leq ((2K)^{1/2} + \delta^{-1/2})^2 \sum_{n=-K}^{K} |a_n|^2. \tag{6.21}$$

Proof: Let

$$\psi(x) = \sum_{n=-\infty}^{\infty} b_n e(nx), \text{ with } b_n \text{ real, and } b_n = b_{-n},$$

be a real function of integrable square such that

$$\psi(x) = 0 \text{ when } \|x\| > \delta/2.$$

Suppose also that $b_n \neq 0$ for $|n| \leq K$. Define

$$T(x) = \sum_{n=-K}^{K} \frac{a_n}{b_n} e(nx).$$

Then

$$S(x) = \int_0^1 \psi(y) T(x-y)\,dy$$

$$= \int_{-\delta/2}^{\delta/2} \psi(y) T(x-y)\,dy$$

since

$$\int_0^1 \psi(y)T(x-y)dy = \sum_{m=-\infty}^{\infty} \sum_{n=-K}^{K} \frac{a_n}{b_n} b_m \int_0^1 e(my)e(n(x-y))dy$$

$$= \sum_{m=-\infty}^{\infty} \sum_{n=-K}^{K} \frac{a_n}{b_n} b_m e(nx) \int_0^1 e((m-n)y)dy$$

$$= \sum_{n=-K}^{K} a_n e(nx).$$

Therefore

$$|S(x)|^2 \leq \Big(\int_{-\delta/2}^{\delta/2} \psi^2(y)dy \Big)\Big(\int_{-\delta/2}^{\delta/2} |T(x-y)|^2 dy \Big)$$

$$= 2\Big(\int_0^{\delta/2} \psi^2(y)dy \Big)\Big(\int_{x-\delta/2}^{x+\delta/2} |T(z)|^2 dz \Big).$$

Using the fact that the intervals $(x_r - \delta/2,\ x_r + \delta/2)$ are disjoint modulo 1, and that T is periodic with period 1, we find (just as in the proof of Theorem 6.15) that

$$\sum_{r=1}^{R} |S(x_r)|^2 \leq 2\Big(\int_0^{\delta/2} \psi^2(y)dy \Big)\Big(\int_0^1 |T(z)|^2 dz \Big) \tag{6.22}$$

$$= 2\Big(\int_0^{\delta/2} \psi^2(y)dy \Big) \cdot \Big(\sum_{n=-K}^{K} |a_n|^2 b_n^{-2} \Big) .$$

Now

$$b_n = \int_0^1 \psi(y)e(-ny)dy = \int_0^{\delta/2} \psi(y)e(-ny)dy + \int_{-\delta/2}^{0} \psi(y)e(-ny)dy$$

$$= 2\int_0^{\delta/2} \psi(y)\cos(2\pi ny)dy.$$

We define

$$y = \frac{1}{2}\,\delta t,\quad \varphi(t) = \psi(y),\quad u = K\delta.$$

Then $\varphi(t)$ is an arbitrary function of integrable square, and

$$\int_0^{\delta/2} \psi^2(y)dy = \frac{1}{2}\,\delta\int_0^1 \varphi^2(t)dt,$$

$$b_n = \delta\int_0^1 \varphi(t)\cos(\pi st)dt,$$

where $s = n\delta$. Therefore for $|n| \leq K$,

$$|b_n| \geq \delta \min_{0\leq s\leq u} \left|\int_0^1 \varphi(t)\cos(\pi st)dt\right|$$

Therefore

$$\sum_{r=1}^{R} |S(x_r)|^2 \leq \delta^{-1}D(u) \sum_{n=-K}^{K} |a_n|^2,$$

where

$$D(u) = \sup_{0\leq s\leq n} \frac{\int_0^1 \varphi^2(t)dt}{\left|\int_0^1 \varphi(t)\cos(\pi st)dt\right|}.$$

To prove the theorem it will then suffice to show that for a suitable function $\varphi(t)$ we have

(6.23) $$D(u) \leq 2 \max(2u,1),$$

and

(6.24) $$D(u) \leq ((2u)^{1/2} + 1)^2.$$

Case 1: $u \geq \frac{3}{2} + \sqrt{2}$. Since in this range

$$4u \geq ((2u)^{1/2} + 1)^2,$$

it suffices to prove (6.24). We take $\lambda > u$ (which will be specified more exactly later) and define

$$\varphi(t) = \frac{\sin \pi\lambda t}{t} .$$

Since

$$\int_0^\infty \frac{\sin t}{t}\, dt = \frac{\pi}{2} ,$$

Lemma 6.18 implies that

$$\int_0^x \frac{\sin t}{t}\, dt = \frac{\pi}{2} - \int_x^\infty \frac{\sin t}{t}\, dt > \frac{\pi}{2} - \frac{1}{x} \quad \text{for } x > 0.$$

Therefore for $0 \leq s \leq u$ we have

$$(6.25) \qquad \int_0^1 \varphi(t)\cos(\pi st)dt = \frac{1}{2}\int_0^1 \frac{\sin \pi(\lambda+s)t + \sin \pi(\lambda-s)t}{t}\, dt$$

$$= \frac{1}{2}\int_0^{\pi(\lambda+s)} \frac{\sin t}{t}\, dt + \frac{1}{2}\int_0^{\pi(\lambda-s)} \frac{\sin t}{t}\, dt$$

$$> \left(\frac{\pi}{4} - \frac{1}{2}\,\frac{1}{\pi(\lambda+s)}\right) + \left(\frac{\pi}{4} - \frac{1}{2}\,\frac{1}{\pi(\lambda-s)}\right)$$

$$= \frac{\pi}{2} - \frac{\lambda}{\pi(\lambda^2-s^2)} \geq \frac{\pi}{2} - \frac{\lambda}{\pi(\lambda^2-u^2)} ,$$

which is positive for λ sufficiently large. Also

$$\int_0^\infty \left(\frac{\sin t}{t}\right)^2 dt = \frac{\pi}{2}$$

implies that

$$\int_0^1 \varphi^2(t)dt = \pi\lambda\int_0^{\pi\lambda} \left(\frac{\sin t}{t}\right)^2 dt < \pi\lambda\int_0^{\infty} \left(\frac{\sin t}{t}\right)^2 dt = \frac{1}{2}\pi^2\lambda.$$

Hence for this choice of $\varphi(t)$ we have

$$D(u) < 2\lambda\left(1 - \frac{2\lambda}{\pi^2(\lambda^2 - u^2)}\right)^{-2} .$$

To prove (6.24) it suffices to prove that

$$2\lambda\left(1 - \frac{2\lambda}{\pi^2(\lambda^2 - u^2)}\right)^{-2} < ((2u)^{1/2} + 1)^2$$

for an appropriate λ. But that is equivalent to showing that

$$\frac{\sqrt{2}}{(2u)^{1/2} + 1} < \lambda^{-1/2} - \frac{2\lambda^{1/2}}{\pi^2(\lambda^2 - u^2)} .$$

We put $\lambda = u(1 + \gamma)^2$, with $\gamma > 0$. Then

$$\lambda^2 - u^2 = u^2(1 + 4\gamma + 6\gamma^2 + 4\gamma^3 + \gamma^4) - u^2 > 4u^2\gamma(1 + \gamma).$$

Thus we only need to show that

$$\frac{(2u)^{1/2}}{(2u)^{1/2} + 1} < \frac{1}{1 + \gamma} - \frac{1}{2\pi^2 u\gamma} \tag{6.26}$$

for an appropriate γ. By differentiating the right side we find that it attains its maximum when

$$\gamma = \frac{1}{\pi(2u)^{1/2} - 1} .$$

With this choice of γ the last expression in (6.25) is positive and (6.26) becomes equivalent to

$$\pi^2\eta^3 < (\eta + 1)(\pi\eta - 1)^2,$$

where $\eta = \sqrt{2u}$. But the last inequality reduces to

$$0 < \pi(\pi-2)\eta^2 - (2\pi-1)\eta + 1,$$

and this one holds for $\eta > 1.5$, say. Since we only need to prove it for $u \geq \frac{3}{2} + \sqrt{2}$, which corresponds to $\eta \geq 1 + \sqrt{2}$, the theorem is true in this case.

Case 2: $0 < u \leq 1/2$. We take $\varphi(t) = \cos(\pi ut)$. The derivative of

$$\int_0^1 \varphi(t)\cos(\pi st)dt = \int_0^1 \cos(\pi ut)\cos(\pi st)dt \tag{6.27}$$

with respect to s is

$$-\pi\int_0^1 t\cos(\pi ut)\sin(\pi st)dt,$$

which is negative for $0 < s \leq u \leq 1/2$. Hence the minimum of (6.27) occurs at $s = u$, and it is

$$\int_0^1 \cos^2(\pi ut)dt = \frac{1}{2} + \frac{\sin 2\pi u}{4\pi u},$$

which is also equal to $\int_0^1 \varphi^2(t)dt$. Therefore for this φ we have

$$D(u) = 2\left\{1 + \frac{\sin 2\pi u}{2\pi u}\right\}^{-1}.$$

Since $0 < u \leq 1/2$, $\sin 2\pi u$ is non-negative, and therefore $D(u) \leq 2$, proving (6.23). To prove (6.24) we need to show that

$$2 \leq \left(1 + \frac{\sin 2\pi u}{2\pi u}\right)((2u)^{1/2} + 1)^2. \tag{6.28}$$

If $u \geq 1/4$, then $((2u)^{1/2} + 1)^2 > 2$. Now suppose that $u < 1/4$. From the Taylor series expansion of sin x we find that $\sin 2\pi u \geq 2\pi u - \frac{(2\pi u)^3}{6}$. Hence

$$\left(1 + \frac{\sin 2\pi u}{2\pi u}\right)((2u)^{1/2} + 1)^2 \geq 2\left(1 - \frac{\pi^2 u^2}{3}\right)((2u)^{1/2} + 1)^2.$$

But $u < 1/4$ implies that $\pi^2 u < 3$, so that

$$\left(1 - \frac{\pi^2 u^2}{3}\right)((2u)^{1/2} + 1)^2 > (1-u)(2u+1) > 1,$$

and therefore (6.28) is valid.

Case 3: $1/2 \leq u \leq 3/2 + \sqrt{2}$. In this range we have

$$4u \leq ((2u)^{1/2} + 1)^2,$$

so it suffices to show that $D(u) \leq 4u$ for a suitable $\varphi(t)$. We take

$$\varphi(t) = \begin{cases} \cos(\pi u t) & \text{for } 0 \leq t \leq (2u)^{-1}, \\ 0 & \text{for } (2u)^{-1} \leq t \leq 1. \end{cases}$$

Then the derivative of

$$(6.29) \qquad \int_0^1 \varphi(t)\cos(\pi s t)dt = \int_0^{(2u)^{-1}} \cos(\pi u t)\cos(\pi s t)dt$$

with respect to s is

$$-\pi\int_0^{(2u)^{-1}} t\cos(\pi u t)\sin(\pi s t)dt,$$

which is negative for $0 < s \leq u$, and therefore the minimum of (6.29) is obtained when $s = u$. Therefore

$$D(u) = \left\{\int_0^{(2u)^{-1}} \cos^2(\pi u t)dt\right\}^{-1} = 4u$$

for this choice of φ, and this completes the proof.

Theorems 6.15 and 6.19 apply only to trigonometric polynomials of the special form (6.9). However, we easily deduce from them the following estimates:

(6.30) THEOREM. Let

$$S(x) = \sum_{n=M+1}^{M+N} a_n e(nx), \tag{6.31}$$

when M and N are integers, with $N > 0$, and the a_n are any complex numbers. Suppose that $x_1, \ldots, x_R$ are real numbers satisfying

$$\|x_r - x_s\| \geq \delta > 0 \quad \text{for } r \neq s.$$

Then we have

$$\sum_{r=1}^{R} |S(x_r)|^2 \leq (\delta^{-1} + \pi N) \sum_{n=M+1}^{M+N} |a_n|^2 , \tag{6.32}$$

$$\sum_{r=1}^{R} |S(x_r)|^2 \leq 2 \max(N, \delta^{-1}) \sum_{n=M+1}^{M+N} |a_n|^2, \tag{6.33}$$

and

$$\sum_{r=1}^{R} |S(x_r)|^2 \leq (N^{1/2} + \delta^{-1/2})^2 \sum_{n=M+1}^{M+N} |a_n|^2 . \tag{6.34}$$

Proof: Let $K = [\frac{N}{2}]$, and define

$$S_*(x) = S(x)\, e(-(M+N-K)x).$$

Then

$$S_*(x) = \sum_{n=-K}^{K} c_n e(nx),$$

where $c_n = a_{n+M+N-K}$ (except when N is even and n = -K, in which case $c_{-K} = 0$). To obtain the estimates (6.32) - (6.34) we apply Theorems 6.15 and 6.19 to $S_*(x)$, noting that

$$|S_*(x)| = |S(x)|, \quad \sum_{n=-K}^{K} |c_n|^2 = \sum_{n=M+1}^{M+N} |a_n|^2, \text{ and } 2K \leq N.$$

In such applications as Bombieri's theorem any one of the estimates (6.32) - (6.34) will be sufficient; in fact, only estimates of the form $\ll (N + \delta^{-1})\Sigma|a_n|^2$ will be used. In some applications, however, such as those to primes in arithmetic progressions, it is inequality (6.34) that is most useful. The main reason is that we deal then with bounds of the form

$$\frac{D(X^2,N)}{\log X}, \tag{6.35}$$

when N is the length of the interval we are investigating, X is a variable, and where $D(X^2,N)$ is $X^2 + \pi N$ if we use (6.32), $2\max(N,X^2)$ if we use (6.33), and $(N^{1/2} + X)^2$ if we use (6.34). By taking $X = N^{1/2}/\log N$ and using (6.34) we find that the main term of (6.35) is

$$\frac{2N}{\log N}.$$

Inequalities (6.32) and (6.33) lead to similar estimates, but with 2 above replaced by 2π and 4, respectively.

We now state, without proof, a more recent result of Bombieri and Davenport, which gives improved estimates when $N\delta$ is very large or very small. It also shows that those estimates are essentially the best possible for the class of all trigonometric polynomials defined by (6.31).

(6.36) THEOREM. With the notation of Theorem (6.30) we have

I. If $N\delta > 1$, then

$$\sum_{r=1}^{R} |S(x_r)|^2 < (N + 5\delta^{-1}) \sum_{n=M+1}^{M+N} |a_n|^2 .$$

On the other hand, if c is a constant less than 1 then there exist sums $S(x)$ with δ arbitrarily small and $N\delta$ arbitrarily large for which

$$\sum_{r=1}^{R} |S(x_r)|^2 > (N + c\delta^{-1}) \sum_{n=M+1}^{M+N} |a_n|^2 .$$

II. If $N\delta \leq 1/4$, then

$$\sum_{r=1}^{R} |S(x_r)|^2 < (\delta^{-1} + 270N^3\delta^2) \sum_{n=M+1}^{M+N} |a_n|^2.$$

On the other hand there exist sums $S(x)$ with $N\delta$ arbitrarily small for which

$$\sum_{r=1}^{R} |S(x_r)|^2 > (\delta^{-1} + \tfrac{1}{12} N^3\delta^2) \sum_{n=M+1}^{M+N} |a_n|^2.$$

So far we have been considering only sums involving the exponential function. The inequalities of Theorem 6.30 can, however, be used to

give estimates for sums involving Dirichlet's characters, and it is these estimates that make the large sieve a valuable tool in analytic number theory. We now present one way of obtaining such estimates.

Suppose $S(\chi)$ is a function of the character χ defined by

$$S(\chi) = \sum_{n=M+1}^{M+N} a_n \chi(n), \tag{6.37}$$

where, as before, M and N are integers with $N > 0$, and the a_n are any complex numbers. We then define the corresponding trigonometric polynomial $S(x)$ by (6.31); that is,

$$S(x) = \sum_{n=M+1}^{M+N} a_n e(nx),$$

where M,N, and the a_n are the same as in (6.37). Then we have

(6.38) LEMMA. If $\sum_{\chi \bmod q}^{*}$ denotes summation over the primitive characters to the modulus q, then for a positive integer q we have

$$\frac{q}{\varphi(q)} \sum_{\chi \bmod q}^{*} |S(\chi)|^2 \leq \sum_{\substack{a=1 \\ (a,q)=1}}^{q} |S(\tfrac{a}{q})|^2 .$$

Proof: Let q be a fixed positive integer. If χ is a character to the modulus q, then for any integer r we define the Gauss sum

$$G(r,\chi) = \sum_{a=1}^{q} \chi(a) e(\tfrac{ra}{q}).$$

If $(r,q) = 1$, then $G(r,\chi)$ is separable; that is,

$$G(r,\chi) = \bar{\chi}(r)G(1,\chi).$$

If χ is primitive, then $G(r,\chi)$ is separable for all r (separability of $G(r,\chi)$ for all r actually characterizes primitive characters), and

$$|G(1,\chi)| = q^{1/2}.$$

Now if χ is primitive, so is $\bar{\chi}$. Hence we can write for a primitive character χ

$$G(1,\bar{\chi})\chi(n) = \sum_{a=1}^{q} e(\frac{na}{q}).$$

Multiplying both sides by a_n and summing over n, we find

$$G(1,\bar{\chi}) \sum_{n=M+1}^{M+N} a_n\chi(n) = \sum_{a=1}^{q} \bar{\chi}(a) \sum_{n=M+1}^{M+N} a_n e(\frac{na}{q}),$$

or

$$G(1,\bar{\chi})S(\chi) = \sum_{a=1}^{q} \bar{\chi}(a)S(\frac{a}{q}).$$

Taking squares of absolute values of both sides and summing over the primitive characters we obtain

$$q \sum_{\chi \bmod q}^{*} |S(\chi)|^2 = \sum_{\chi \bmod q}^{*} \left| \sum_{a=1}^{q} \bar{\chi}(a)S(\frac{a}{q}) \right|^2$$

$$\leq \sum_{\chi \bmod q} \sum_{a,b=1}^{q} \bar{\chi}(a)\chi(b)S(\frac{a}{q})\overline{S(\frac{b}{q})}$$

$$= \sum_{a,b=1}^{q} S(\tfrac{a}{q})\overline{S(\tfrac{b}{q})} \sum_{\chi \bmod q} \bar{\chi}(a)\chi(b)$$

$$= \varphi(q) \sum_{\substack{a=1 \\ (a,q)=1}}^{q} |S(\tfrac{a}{q})|^2$$

since

$$\sum_{\chi \bmod q} \bar{\chi}(n)\chi(m) = \begin{cases} \varphi(q) & \text{if } n \equiv m \pmod q \text{ and } (n,q) = 1, \\ 0 & \text{otherwise.} \end{cases}$$

This proves the lemma.

Since bounds for

$$\sum_{q \le X} \sum_{\substack{a=1 \\ (a,q)=1}}^{q} |S(\tfrac{a}{q})|^2$$

follow easily from Theorem 6.30, Lemma 6.38 gives us bounds for the sum of $\frac{q}{\varphi(q)} |S(\chi)|^2$ over all primitive characters χ to all moduli $q \le X$. In some applications where only primitive characters occur, this result is then immediately applicable. In others we will use the fact that every character is induced by a primitive character. The following lemma will be useful in such cases.

(6.39) LEMMA. If $S(x)$ is defined by (6.31) and $\lambda(x)$ is a positive decreasing continuous function on $0 < D \le x \le Q$, then

$$(6.40) \quad \sum_{D<q\le Q} \lambda(q) \sum_{\substack{a=1 \\ (a,q)=1}}^{q} |S(\tfrac{a}{q})|^2 \ll \Big(\lambda(D)(D^2+N) + \int_D^Q x\lambda(x)dx\Big) \sum_{n=M+1}^{M+N} |a_n|^2.$$

Proof: Let $Y = \sum_{n=M+1}^{M+N} |a_n|^2$, and let

$$T(x) = \sum_{q \le x} \sum_{\substack{a=1 \\ (a,q)=1}}^{q} |S(\tfrac{a}{q})|^2.$$

Since the points a/q, with $1 \le a \le q$, $(a,q) = 1$, and $q \le x$ are separated by at least $\frac{1}{x^2}$, Theorem 6.30 says that

$$T(x) \ll (N + x^2)Y.$$

We now write the left side of (6.40) as a Riemann-Stieltjes integral. It is equal to

$$\begin{aligned}
\int_D^Q \lambda(x)dT(x) &= \lambda(Q)T(Q) - \lambda(D)T(D) - \int_D^Q T(x)d\lambda(x) \\
&\ll \lambda(Q)(N+Q^2)Y - \int_D^Q (N+x^2)Y\, d\lambda(x) \\
&= \lambda(Q)(N+Q^2)Y - NY(\lambda(Q) - \lambda(D)) - Y\int_D^Q x^2 d\lambda(x) \\
&= \lambda(D)NY + \lambda(Q)Q^2Y - Y(Q^2\lambda(Q) - D^2\lambda(D) - 2\int_D^Q x\lambda(x)dx) \\
&\ll \lambda(D)NY + \lambda(D)D^2Y + Y\int_D^Q x\lambda(x)dx,
\end{aligned}$$

and this proves the lemma.

Notes on Chapter 6.

Although the large sieve was originated by Linnik [1], much of its development is due to Rényi, who was the first to use the variance expression and to apply the large sieve to estimate the remainder term in Selberg's sieve. The results listed in this chapter were obtained in the late 1960's after several important advances had been made, notably by Roth [1], Bombieri [3], and Davenport and Halberstam [1]. Theorem 6.15 (and its corollary, the estimate (6.32)), as well as Lemmas 6.38 and 6.39, are derived from Gallagher [1]. Theorem 6.19 and the inequalities (6.33) and (6.34) come from Bombieri and Davenport [2]. The recent results of Bombieri and Davenport [3], which we listed as Theorem 6.36, effectively concludes this part of the large sieve; by indicating what the best possible results are for general trigonometric polynomials they show that further advances can be obtained only for particular cases by utilizing some special properties of the polynomials under consideration.

Properties of Dirichlet characters and Gauss sums are discussed in Ayoub [1; Chapter 5] and Prachar [1; chapters 4 and 7]. That separability of $G(r,\chi)$ for all r implies that χ is primitive has been proved by Apostol [1].

One weakness of estimates such as those of Theorem 6.30 is that they are affected greatly by two of the x_i being close together. This weakness has been avoided by a result of Davenport and Halberstam [1] (also given in Davenport [1; Section 23] which states, in the notation of

Theorem 6.30, that if

$$\delta_r = \min_{i \neq r} \|x_i - x_r\|,$$

then

$$\sum_{r=1}^{R} \min(1, \tfrac{N}{2}\delta_r)|S(x_r)|^2 \ll N \sum_{n=M+1}^{M+N} |a_n|^2.$$

In the Corrigendum and Addendum to that paper they also gave inequalities for sums of $|S(\chi)|^2$ and over all characters to the modulus q for all $q \leq x$.

7

THE LARGE SIEVE

We have seen in the first part of Chapter 6 how the number of integers that remain in an interval after removing a given number of congruence classes modulo each prime $p \leq X$ can be estimated from bounds for the sum

$$\sum_{p \leq X} \sum_{a=1}^{p-1} \left| S\left(\frac{a}{p}\right) \right|^2, \tag{7.1}$$

where S is a trigonometric polynomial of the form (6.31). Now if a/p and b/p' are two points that appear in (7.1), and $a/p \neq b/p'$, then $ap' - bp$ is a non-zero integer, and hence

$$\left| \frac{a}{p} - \frac{b}{p'} \right| = \left| \frac{ap' - bp}{pp'} \right| \geq \frac{1}{pp'} \geq \frac{1}{X^2} .$$

Since the same argument holds for $|k - \frac{a}{p} + \frac{b}{p'}|$ for any integer k, we have the lower bound

$$\min_{a/p \neq b/p'} \| \frac{a}{p} - \frac{b}{p'} \| \geq \frac{1}{X^2} ,$$

which is best possible in general. Therefore we can estimate (7.1) by applying Theorem 6.30 with $\delta = X^{-2}$. But by the same argument we can take $\delta = X^{-2}$ in estimating the sum of $|S(x)|^2$ over all rational points, in their lowest terms, which lie in $(0,1]$, and for which the denominator is $\leq X$. Therefore the bound for (7.1) that we obtain from Theorem 6.30 is the same as the bound for

$$(7.2) \qquad \sum_{q \le X} \sum_{\substack{a=1 \\ (a,q)=1}}^{q} \left|S\left(\frac{a}{q}\right)\right|^2 .$$

Since the numbers of terms in (7.1) and (7.2) are asymptotic to $\frac{1}{2} X^2/\log X$ and $3X^2/\pi^2$, respectively, it is natural that in looking for improvements on (6.8) one should try to use the terms in (7.2) with q composite. We will do this by proving an identity for

$$\sum_{\substack{a=1 \\ (a,q)=1}}^{q} \left|S\left(\frac{a}{q}\right)\right|^2$$

similar to the one of Lemma 6.6 (which is not applicable to (7.2) because of the condition $(a,q) = 1$). The next step will be to use this identity, together with Theorem 6.30, to give estimates for sequences that have no elements in a given number of residue classes modulo p, for each prime $p \le X$. This whole chapter is devoted to the proof of these estimates (which were discovered by Montgomary) and a few applications.

Instead of considering simply a sequence of integers from the interval $[M+1, M+N]$, (M and N integers, $N > 0$), we will generalize to the case where every integer in that interval has attached to it a certain complex "weight" a_n. Our inequalities will then measure the distribution of these weights into various congruence classes. Inequalities for the number of elements in a given sequence will then be obtained by choosing a_n to be 1 if n is in that sequence and 0 otherwise.

We let a_n be any complex numbers defined for $n = M+1, \ldots, M+N$.

Define

$$Z(q,h) = \sum_{\substack{n=M+1 \\ n\equiv h(\mathrm{mod}\ q)}}^{M+N} a_n \tag{7.3}$$

and

$$Z = Z(1,1) = \sum_{n=M+1}^{M+N} a_n, \tag{7.4}$$

thus generalizing our previous definition (6.2). We will work with the associated trigonometric polynomial

$$S(x) = \sum_{n=M+1}^{M+N} a_n\, e(nx). \tag{7.5}$$

(7.6) **LEMMA.** If q is a positive integer, then

$$\sum_{\substack{a=1 \\ (a,q)=1}}^{q} \left|S(\tfrac{a}{q})\right|^2 = q \sum_{h=1}^{q} \left| \sum_{d|q} \frac{\mu(d)}{d}\, Z(\tfrac{q}{d},h)\right|^2. \tag{7.7}$$

Proof: Since for a an integer

$$S(\tfrac{a}{q}) = \sum_{h=1}^{q} Z(q,h)\ e(\tfrac{ah}{q}),$$

we have, by (6.1)

$$qZ(q,h) = \sum_{a=1}^{q} S(\tfrac{a}{q})e(\tfrac{-ah}{q})$$

$$= \sum_{d|q} \sum_{\substack{b=1 \\ (b,q/d)=1}}^{q/d} S(\tfrac{bd}{q})e(\tfrac{-bdh}{q}).$$

If we define

$$T(q,h) = \sum_{\substack{a=1 \\ (a,q)=1}}^{q} S(\tfrac{a}{q})e(\tfrac{-ah}{q}), \tag{7.8}$$

then

$$qZ(q,h) = \sum_{d|q} T(q/d,h). \tag{7.9}$$

Applying the Möbius inversion formula to (7.6) we obtain

$$T(q,h) = q \sum_{d|q} \frac{\mu(d)}{d} Z(q/d,h).$$

But from (7.8) we find that

$$q \sum_{h=1}^{q} \Big| \sum_{d|q} \frac{\mu(d)}{d} Z(q/d,h) \Big|^2 = \frac{1}{q} \sum_{h=1}^{q} |T(q,h)|^2$$

$$= \frac{1}{q} \sum_{h=1}^{q} \sum_{\substack{a,b=1 \\ (a,q)=(b,q)=1}}^{q} S(\tfrac{a}{q})\overline{S(\tfrac{b}{q})}\, e(\tfrac{(b-a)h}{q})$$

$$= \frac{1}{q} \sum_{\substack{a,b=1 \\ (a,q)=(b,q)=1}}^{q} S(\tfrac{a}{q})\overline{S(\tfrac{b}{q})} \sum_{h=1}^{q} e(\tfrac{(b-a)h}{q})$$

$$= \sum_{\substack{a=1 \\ (a,q)=1}}^{q} |S(\tfrac{a}{q})|^2$$

by (6.1), and this completes the proof.

We may note that for a prime p, (7.7) is the same (except for somewhat generalized a_n) as (6.7), since

$$p\sum_{h=1}^{p}\Big|\sum_{d|p}\frac{\mu(d)}{d}Z(\tfrac{p}{d},h)\Big|^2 = p\sum_{h=1}^{p}|Z(p,h) - \tfrac{Z}{p}|^2.$$

We now easily deduce

(7.10) THEOREM. If $Z(q,h)$ is defined by (7.3), then for $X \geq 1$ we have

$$\sum_{q\leq X} q\sum_{h=1}^{q}\Big|\sum_{d|q}\frac{\mu(d)}{d}Z(\tfrac{q}{d},h)\Big|^2 \leq (N^{1/2}+X)^2\sum_{n=M+1}^{M+N}|a_n|^2.$$

Proof: Let $S(x)$ be defined by (7.5). By (6.34) we have

$$\sum_{q\leq X}\sum_{\substack{a=1\\(a,q)=1}}^{q}|S(\tfrac{a}{q})|^2 \leq (N^{1/2}+X)^2\sum_{n=M+1}^{M+N}|a_n|^2.$$

But now we just substitute (7.7) in the left side above.

We deduce from Theorem 7.10

(7.11) THEOREM. Let $Z(q,h)$ and Z be defined by (7.3) and (7.4), respectively, and let $X \geq 1$. For each prime $p \leq X$ let $H(p)$ be the union of $f(p)$ distinct residue classes modulo p. Let the a_n satisfy

(7.12) $\qquad a_n = 0$ if $n \in H(p)$ for some $p \leq X$.

Then

$$|Z|^2 \leq \frac{(N^{1/2}+X)^2}{Q}\sum_{n=M+1}^{M+N}|a_n|^2,$$

where

$$Q = \sum_{q \le X} \mu^2(q) \prod_{p|q} \frac{f(p)}{p - f(p)} .$$

Proof: In view of Theorem 7.10 it will suffice to show that for each $q \le X$ we have

$$(7.13) \qquad \mu^2(q)|Z|^2 \prod_{p|q} \frac{f(p)}{p - f(p)} \le q \sum_{h=1}^{q} \left| \sum_{d|q} \frac{\mu(d)}{d} Z(q/d,h) \right|^2 .$$

This is clearly true if $\mu(q) = 0$. Therefore assume that $q \le X$ is a fixed, squarefree integer.

If $d|q$, we define

$$(7.14) \quad K(d) = \{h;\ 1 \le h \le q,\ h \in H(p) \text{ if } p|d,\ h \notin H(p) \text{ if } p|(q/d)\}.$$

The sets $K(d)$, for d going through all the divisors of q, form a partitioning of $\{1, \ldots, q\}$, since for each h we can write q uniquely as

$$q = \Big(\prod_{\substack{p|q \\ h \in H(p)}} p \Big) \cdot \Big(\prod_{\substack{p|q \\ h \notin H(p)}} p \Big).$$

Therefore we will be able to write

$$\sum_{h=1}^{q} = \sum_{d|q} \sum_{h \in K(d)} .$$

Let us fix a δ, $\delta|q$. Then by the Cauchy-Schwarz inequality

$$(7.15)\qquad \Big|\sum_{d|q}\mu(q/d)d\sum_{h\in K(\delta)} Z(d,h)\Big|^2 = \Big|\sum_{d|q}\frac{\mu(d)q}{d}\sum_{h\in K(\delta)} Z(q/d,h)\Big|^2$$

$$= \Big|\sum_{h\in K(\delta)}\sum_{d|q}\frac{\mu(d)q}{d} Z(q/d,h)\Big|^2$$

$$\le \Big\{\sum_{h\in K(\delta)} 1\Big\}\Big\{\sum_{h\in K(\delta)}\Big|\sum_{d|q}\frac{\mu(d)q}{d} Z(q/d,h)\Big|^2\Big\}.$$

Let us consider the left side above. Suppose $(\delta,d) > 1$, and choose a prime p such that $p|(\delta,d)$. Now $Z(d,h)$ is a sum of a_n with $n \equiv h \pmod d$. But $p|d$ implies that for such an a_n we have $n \equiv h \pmod p$. But $p|\delta$ and $h \in K(\delta)$ mean, in view of the definition (7.14), that $n \in H(p)$. Therefore by (7.12) $a_n = 0$ whenever $n \equiv h \pmod d$ and $h \in K(\delta)$. Hence the inner sum on the left side of (7.15) vanishes when $(\delta,d) > 1$, and so

$$(7.16)\qquad \sum_{d|q}\mu(q/d)d\sum_{h\in K(\delta)} Z(d,h) = \sum_{d|(q/\delta)}\mu(q/d)d\sum_{h\in K(\delta)} Z(d,h).$$

Let us fix d, with $d|(q/\delta)$. We have

$$\sum_{h\in K(\delta)} Z(d,h) = \sum_{\substack{k=1\\ k\notin H(p)\ \forall\ p|d}}^{d} Z(d,k)\cdot |\{h;\ h\in K(\delta),\ h\equiv k \pmod d\}|,$$

where the condition $k \notin H(p)$ for all p dividing d follows from the fact that if $k \in H(p)$ for some $p|d$, then $Z(d,k) = 0$. We can compute

$$S(\delta,d,k) = |\{h;\ h\in K(\delta),\ h\equiv k \pmod d\}|$$

for k such that $k \in H(p)$ for all primes p dividing d. Now in view of the Chinese remainder theorem $h \equiv k \pmod{d}$ is equivalent to $h \equiv k \pmod{p}$ for all primes p dividing d. Also, $h \in K(\delta)$ is equivalent to $h \in H(p)$ for all primes p dividing δ and $h \notin H(p)$ for all primes p dividing q/δ. Therefore $h \in K(\delta)$ and $h \equiv k \pmod{d}$ if and only if the following three conditions are satisfied:

a) $p|d \Rightarrow h \equiv k \pmod{p}$ and $h \notin H(p)$,

b) $p|\delta \Rightarrow h \in H(p)$,

c) $p|(q/d\delta) \Rightarrow h \notin H(p)$.

Since we are dealing with k such that $k \notin H(p)$ for all primes p dividing d, the second part of a) is satisfied whenever the first part is. We now notice that if $p|d$, then there is exactly one solution of a) modulo p. If $p|\delta$, then there are $f(p)$ solutions of b) modulo p, and if $p|(q/d\delta)$, then there are $p - f(p)$ solutions of c) modulo p. Applying the Chinese remainder theorem we find that

$$S(\delta,d,k) = |\{h;\ 1 \le h \le q,\ h \text{ satisfies a), b), and c)}\}|$$

$$= \{\prod_{p|\delta} f(p)\}\ \{\prod_{p|\frac{q}{d\delta}} (p - f(p))\}.$$

This number is independent of k, and hence

$$\sum_{h \in K(\delta)} Z(d,h) = \sum_{\substack{k=1 \\ k \notin H(p)\ \forall\ p|d}}^{d} Z(d,k) \prod_{p|\delta} f(p) \prod_{p|\frac{q}{d\delta}} (p-f(p))$$

$$= \sum_{k=1}^{d} Z(d,k)\Big(\prod_{p|\delta} f(p)\Big)\Big(\prod_{p|\frac{q}{d\delta}} (p - f(p))\Big)$$

$$= Z\Big(\prod_{p|\delta} f(p)\Big)\Big(\prod_{p|\frac{q}{d\delta}} (p - f(p))\Big).$$

From (7.16) we now obtain

$$(7.17)\quad \sum_{d|q} \mu(q/d)d \sum_{h\in K(\delta)} Z(d,h) = \sum_{d|(q/\delta)} \mu(q/d)d\, Z\Big\{\prod_{p|\delta} f(p)\Big\}\Big\{\prod_{p|\frac{q}{d\delta}} (p - f(p))\Big\}$$

$$= \mu(q)Z\Big\{\prod_{p|\delta} f(p)\Big\}\Big\{\prod_{p|(q/\delta)} (p - f(p))\Big\} \sum_{d|(q/\delta)} \mu(d)d \prod_{p|d} (p - f(p))^{-1}$$

$$= \mu(q)Z\Big\{\prod_{p|\delta} f(p)\Big\}\Big\{\prod_{p|(q/\delta)} (p - f(p))\Big\} \prod_{p|(q/\delta)} \Big(1 - \frac{p}{p - f(p)}\Big)$$

$$= \mu(\delta)Z\Big\{\prod_{p|\delta} f(p)\Big\}\Big\{\prod_{p|q/\delta} f(p)\Big\} = \mu(\delta)Z \prod_{p|q} f(p).$$

The first factor on the right side of (7.15) is

$$S(\delta,1,1) = \Big\{\prod_{p|\delta} f(p)\Big\}\Big\{\prod_{p|(q/\delta)} (p - f(p))\Big\}$$

Dividing by it and using (7.17) we now find from (7.15) that

$$|Z|^2\Big\{\prod_{p|q} f^2(p)\Big\}\Big\{\prod_{p|\delta} f(p)\Big\}^{-1}\Big\{\prod_{p|(q/\delta)} (p - f(p))\Big\}^{-1} \le \sum_{h\in K(\delta)} \Big|\sum_{d|q} \frac{\mu(d)q}{d} Z(q/d,h)\Big|^2.$$

We now sum on all $\delta|q$. The right side is just

$$\sum_{h=1}^{q} \Big|\sum_{d|q} \frac{\mu(d)q}{d} Z(q/d,h)\Big|^2$$

since the sets $K(\delta)$ form a partition of $\{1, \ldots, q\}$. The left side is

$$|Z|^2\{\prod_{p|q} f(p)\}^2 \sum_{\delta|q} \{\prod_{p|\delta} f(p)\}^{-1} \cdot \{\prod_{p|(q/\delta)} (p - f(p))\}^{-1}$$

$$= |Z|^2\{\prod_{p|q} f(p)\} \sum_{\delta|q} \{\prod_{p|(q/\delta)} f(p)\}\{\prod_{p|(q/\delta)} (p-f(p))^{-1}\}$$

$$= |Z|^2\{\prod_{p|q} f(p)\} \prod_{p|q} (1 + \frac{f(p)}{p-f(p)})$$

$$= q|Z|^2 \prod_{p|q} \frac{f(p)}{p-f(p)} .$$

This completes the proof.

(7.18) COROLLARY. Let $X \geq 1$ and let Z be the number of integers n such that $M + 1 \leq n \leq M + N$ and n does not fall into any of $f(p)$ residue classes modulo p for any prime $p \leq X$. Then

$$Z \leq \frac{(N^{1/2} + X)^2}{Q},$$

where

$$Q = \sum_{q \leq X} \mu^2(q) \prod_{p|q} \frac{f(p)}{p-f(p)} .$$

Proof: Let a_n be 1 if n is one of the given integers and 0 otherwise. Apply Theorem 7.11, noting that

$$|Z| = Z = \sum_{n=M+1}^{M+N} |a_n|^2 .$$

We first use Corollary 7.18 to obtain estimates which we proved previously using Selberg's sieve.

(7.19) THEOREM. Let k, M, and N be positive integers satisfying $k < \frac{1}{3} N$, and let ℓ be an integer such that $(k,\ell) = 1$. Then

$$(7.20) \qquad \pi(M+N,k,\ell) - \pi(M,k,\ell) < \frac{2N}{\varphi(k)\log(N/k)}\left(1 + O\left(\frac{\log\log N/k}{\log N/k}\right)\right).$$

Proof: We consider $m + rk$ for $r = 1, \ldots, n$, where n is the largest integer such that $m \leq M$ and $m \equiv \ell \pmod k$, and n is the largest integer such that $m + nk \leq M+N$.

Let $X \geq 2$. We consider those $r \leq n$ for which $m + rk$ is not divisible by any prime $p \leq X$. Let p be a fixed prime. If $p|k$, then $p\nmid(m+rk)$ for all r. If $p\nmid k$, then $p|(m+rk)$ only when $r \equiv mk^{-1} \pmod p$. Hence we may apply Corollary 7.18 with $f(p) = 1$ if $p|k$ and $f(p) = 0$ if $p|k$. We then find that Z, the number of $r \leq n$ for which $m + rk$ is not divisible by any prime $p \leq X$ satisfies

$$Z \leq \frac{(n^{1/2} + X)^2}{Q},$$

where

$$Q = \sum_{q\leq X} \mu^2(q) \prod_{p|q} \frac{f(p)}{p-f(p)} = \sum_{\substack{q\leq X \\ (q,k)=1}} \frac{\mu^2(q)}{\varphi(q)} \geq \sum_{\substack{q\leq X \\ (q,k)=1}} \frac{1}{q} \geq \frac{\varphi(k)}{k} \log X$$

by Lemma 3.3. We now choose

$$X = n^{1/2}/\log n .$$

Then

$$Z \leq \frac{k}{\varphi(k)} \frac{2n(1 + \frac{1}{\log n})^2}{\log n - 2 \log\log n} \leq \frac{2N}{\varphi(k)\log N/k} (1 + O(\frac{\log\log N/k}{\log N/k}))$$

for N/k sufficiently large (as $n \leq N/k + 1$). Since

$$\pi(M+N,k,\ell) - \pi(M,k,\ell) \leq Z + \pi(X) < Z + X,$$

we obtain (7.20), again for N/k sufficiently large. But the theorem is clearly true for N/k bounded.

Next we give a new proof of Theorem 3.17.

(7.21) THEOREM. Let s be a positive integer, and suppose that for $i = 1, \ldots, s$, a_i and b_i are integers such that $(a_i,b_i) = 1$ and $E = \prod_{i=1}^{s} a_i \cdot \prod_{i<j} (a_ib_j - a_jb_i) \neq 0$. Let $N(p)$ be the number of solutions of $(a_1v + b_1) \cdots (a_sv + b_s) \equiv 0 \pmod{p}$ for $v = 1, \ldots, p$ for each prime p, and assume $N(p) < p$ for all primes p. Then the number of positive integers $m \leq n$ such that each of $|a_im + b_i|$ $(i = 1, \ldots, s)$ is prime is

$$\leq c(s) \frac{n}{(\log n)^s} \cdot \prod_{p|E} (1 - \frac{1}{p})^{-s+N(p)}.$$

Proof: Let $X \geq 1$, and consider

$$Z = |\{m;\ m \leq n,\ p \nmid (a_im+b_i) \text{ for any } p \leq X \text{ and } i = 1, \ldots, s\}|.$$

Then the integers counted by Z do not fall into any of the $N(p)$ residue classes modulo p corresponding to solutions of

$$(a_1 r + b_1) \cdots (a_s r + b_s) \equiv 0 \pmod{p}$$

for all $p \leq X$, and therefore Corollary 7.18 says that

$$Z \leq \frac{(n^{1/2} + X)^2}{\mathcal{Q}},$$

where

$$\mathcal{Q} = \sum_{q \leq X} \mu^2(q) \prod_{p \mid q} \frac{N(p)}{p - N(p)}.$$

Let $P_0 = \prod_{\substack{p \leq X \\ N(p)=0}}$. Then

$$\mathcal{Q} = \sum_{\substack{q \leq X \\ (q, P_0)=1}} \mu^2(q) \prod_{p \mid q} \frac{N(p)}{p - N(p)} = \sum_{\substack{q \leq X \\ (q, P_0)=1}} \mu^2(q) \prod_{p \mid q} \left(\frac{N(p)}{p} + \frac{N^2(p)}{p^2} + \cdots\right)$$

$$\geq \sum_{\substack{q \leq X \\ (q, P_0)=1}} \frac{1}{q} \prod_{p \mid q} N(p),$$

which is the estimate (3.11). The rest of the proof is the same as that of Theorem 3.17.

For our final application of Corollary 7.18 we consider a problem to which Selberg's sieve is not applicable due to the large number of congruence classes sieved out. It has been conjectured by Artin that every integer except for 0, -1, and the perfect squares is a primitive root for infinitely many primes. While we cannot prove that conjecture, we can give an upper bound on the number of positive integers $\leq N$ which are not primitive roots for small primes.

(7.22) THEOREM. The number of positive integers $\leq N$ which are not primitive roots for any prime $\leq N^{1/2}$ is

$$\ll N^{1/2} \cdot \log N \cdot \log \log N \quad \text{for } N \geq 3.$$

Proof: Here we sieve out the $\varphi(p-1)$ residue classes of primitive roots for each $p \leq N^{1/2}$. By Corollary 7.18 the number of remaining integers is

$$\ll \frac{N}{Q},$$

where

$$Q = \sum_{q \leq N^{1/2}} \mu^2(q) \prod_{p \mid q} \frac{\varphi(p-1)}{p - \varphi(p-1)} \geq \sum_{p \leq N^{1/2}} \frac{\varphi(p-1)}{p}.$$

But

$$\liminf_{n \to \infty} \frac{\varphi(n) \log \log n}{n} = e^{-\gamma}, \tag{7.23}$$

where γ is Euler's constant, and therefore

$$\varphi(n) \gg \frac{n}{\log \log n} \quad \text{for} \quad n \geq 3.$$

It follows that

$$Q \gg \frac{1}{\log \log N^{1/2}} \pi(N^{1/2}) \gg \frac{N^{1/2}}{\log N \cdot \log \log N} \quad \text{for } N \geq 3,$$

and this completes the proof.

Notes on Chapter 7.

The material of this chapter is drawn primarily from Montgomery [1]. It should be mentioned, however, that Theorem 7.19 can be proved using large sieve inequalities for sums over Dirichlet characters, as was done by Bombieri and Davenport [2]. Recently Bombieri [3] has proved, using a specialized large sieve method, a result similar to the one of Van Lint and Richert (Theorem 4.1); namely, that $O(\frac{\log \log N/k}{\log N/k})$ in (7.20) can be replaced by $O(\frac{1}{\log N/k})$.

Eq. (7.23) is proved in Hardy and Wright [1; Chapter 18].

Theorem 7.22 is not the best possible. While applying an old form of the large sieve, Gallagher [1] proved that

$$\sum_{p \leq N^{1/2}} \frac{\varphi(p-1)}{p - \varphi(p-1)} \gg \frac{N^{1/2}}{\log N} .$$

Using the contribution to Q of composite integers it can be proved that Q is of an even larger order of magnitude.

8

BOMBIERI'S THEOREM

It has been known for some time that estimates for sums of Dirichlet series can be obtained from large sieve inequalities. The purpose of this chapter is to use such estimates to prove a very important theorem of Bombieri about the average of the remainder term in the prime number theorem for arithmetic progression.

Let

$$\psi(x,q,a) = \sum_{\substack{n\le x \\ n\equiv a \pmod q}} \Lambda(n),$$

where Λ is Mangoldt's Λ-function. Bombieri's result can then be stated as

(8.1) THEOREM. For each positive constant A, there is a positive constant B such that if $Q = x^{1/2}(\log x)^{-B}$, then

$$(8.2) \qquad \sum_{q\le Q} \max_{y\le x} \max_{(a,q)=1} \left|\psi(y,q,a) - \frac{y}{\varphi(q)}\right| \ll x(\log x)^{-A},$$

where the constant implied by the $\ll$ symbol depends on A.

An analogous estimate holds with $\psi(y,q,a)$ and $y/\varphi(q)$ replaced by $\pi(y,q,a)$ and $\mathrm{li}(y)/\varphi(q)$, respectively. Now the extended Riemann hypothesis, which states that all the non-trivial zeros of Dirichlet L-series $L(s,\chi)$ lie on the line $\mathrm{Re}(s) = 1/2$, implies that

$$(8.3) \qquad \psi(y,q,a) - \frac{y}{\varphi(q)} \ll y^{1/2}\log^2 y,$$

which in turn implies that Theorem 8.1 holds with $B = A + 2$. Thus Bombieri's theorem gives a result comparable to the one implied by the extended Riemann hypothesis (which has been neither proved nor disproved), and in many cases where only the average of the remainder term is important it can be used in place of the assumption that (8.3) is valid. Our proof, due to Gallagher, shows that B may be taken as $16A + 103$, and even better results are possible, but they are relatively unimportant since no way has been found for determining the constant implied by the $\ll$ notation.

Our proof of Bombieri's theorem will proceed through a series of lemmas. The basic idea is to express the left side of (8.2) in terms of L-series, and then apply the large sieve inequalities of Chapter 6 to it. Actually, however, we will first apply a smoothing device to the functions $\psi(y,q,a) - y/\varphi(q)$ and work with the resulting functions.

For functions F piecewise continuous on $[1,\infty]$, we put

$$F_0 = F, \quad \text{and} \quad F_{k+1}(x) = \int_1^x F_k(y)\,\frac{dy}{y}.$$

The main part of our proof will consist of showing that Theorem 8.1 holds with $\psi(y,q,a)$ replaced by $\psi_k(y,q,a)$ for k sufficiently large (note that if $F(x) = x$, then $F_1(x) = x - 1$, so that our smoothing device affects x to an extent that is negligible for our purposes). We will actually show this for $k = 3$. At the end of the proof we will show that such estimates imply (8.2).

First we prove an auxiliary result.

(8.4) LEMMA. For $x \geq 1$ we have

$$\sum_{n \leq x} \frac{1}{\varphi(n)} \ll 1 + \log x.$$

Proof: This is clearly true for $x \leq 3$. Assume therefore that $x \geq 3$. Since $\frac{1}{\varphi(n)}$ is a multiplicative function, we have

$$\sum_{n \leq x} \frac{1}{\varphi(n)} \leq \prod_{p \leq x} \left(1 + \frac{1}{\varphi(p)} + \frac{1}{\varphi(p^2)} + \cdots\right) = \prod_{p \leq x} \left(1 + \frac{1}{p-1} + \frac{1}{p(p-1)} + \cdots\right)$$

$$= \prod_{p \leq x} \left(1 + \frac{p}{(p-1)^2}\right) \leq e^{\sum_{p \leq x} \frac{p}{(p-1)^2}} = e^{\sum_{p \leq x} \frac{1}{p} + O(1)}$$

$$= e^{\log \log x + O(1)} \ll \log x .$$

The next lemma shows that in order to estimate

$$\sum_{q \leq Q} \max_{y \leq x} \max_{(a,q)=1} \left|\left(\psi_k(y,q,a) - \frac{y}{\varphi(q)}\right)\right|$$

it is sufficient to estimate a sum involving primitive characters to fairly large moduli. Before we prove it we should make a few remarks about the values of A and k. Our goal is to prove (8.2) with the implied constant depending only on A. Due to the use of the smoothing device, however, the A's we will be working with will be functions of the A of Theorem 8.1. In order to ensure that all constants can be chosen so as

to depend only on the original A for which the theorem is being proved, which is given, we will require at each stage that the A and k be bounded (actually we will only need $k \le 3$, and if A' is the given A of Theorem 8.1, then we will always have $A \le 8A' + 7$), so that the implied constants will be valid for all values of A and k that we will use.

From now on we write ℓ for log x. We use the usual definition

$$\psi(x,\chi) = \sum_{n \le x} \chi(n)\Lambda(n).$$

(8.5) LEMMA. For $Q \le x^{1/2}$ and for bounded k, A, and C, we have

$$\sum_{q \le Q} \max_{y \le x} \max_{(a,q)=1} |\psi_k(y,q,a) - \frac{y}{\varphi(q)}| \ll \sum_{D < q \le Q} \epsilon(q) \sideset{}{^*}\sum_{\chi \bmod q} \max_{y \le x} |\psi_k(y,\chi)| + x\ell^{-A},$$

where $D = \ell^C$, $\epsilon(q) = \frac{1+\log(Q/q)}{\varphi(q)}$, and $\sideset{}{^*}\sum_{\chi \bmod q}$ denotes summation over primitive characters to the modulos q.

Proof: We have (by induction)

$$\psi_k(x,\chi) = \frac{1}{k!} \sum_{n \le x} \chi(n)\Lambda(n)\log^k(\tfrac{x}{n}).$$

Also, since

$$\psi(x,q,a) = \frac{1}{\varphi(q)} \sum_{\chi \bmod q} \overline{\chi}(a)\psi(x,\chi),$$

we find

$$\psi_k(x,q,a) = \frac{1}{\varphi(q)} \sum_{\chi \bmod q} \overline{\chi}(a)\psi_k(x,\chi). \tag{8.6}$$

If χ is induced by the primitive character χ^*, then $\chi(n) = \chi^*(n)$ except possibly when $(n,q) > 1$. Therefore

$$|\psi_k(x,\chi) - \psi_k(x,\chi^*)| \leq \frac{1}{k!} \sum_{n\leq x} |\chi(n) - \chi^*(n)|\Lambda(n)\log^k(\tfrac{x}{n})$$

$$\leq \frac{2}{k!}\, \ell^k \sum_{\substack{n\leq x \\ (n,q)>1}} \Lambda(n) = \frac{2}{k!}\, \ell^k \sum_{p|q} \log p \sum_{\substack{a \\ p^a\leq x}} 1$$

$$\leq \frac{2}{k!}\, \ell^{k+1} \log q \ll \ell^{k+1} \log q.$$

If χ_0 denotes the principal character to the modulus q, then χ_0 is induced by the identity character and thus

$$\psi_k(x,\chi_0^*) = \frac{1}{k!} \sum_{n\leq x} \Lambda(n)\log^k(\tfrac{x}{n}) = \psi_k(x).$$

Therefore we obtain from (8.6)

$$\max_{(a,q)=1} \left|\psi_k(x,q,a) - \frac{\psi_k(x)}{\varphi(q)}\right| \ll \frac{1}{\varphi(q)} \sum_{\chi\neq\chi_0} \psi_k(x,\chi^*) + \ell^{k+1}\log q,$$

and hence

$$(8.7) \qquad \sum_{q\leq Q} \max_{y\leq x} \max_{(a,q)=1} \left|\psi_k(y,q,a) - \frac{\psi_k(y)}{\varphi(q)}\right|$$

$$\ll \sum_{q\leq Q} \frac{1}{\varphi(q)} \sum_{\chi\neq\chi_0} \max_{y\leq x} |\psi_k(y,\chi^*)| + Q\ell^{k+1}\log Q.$$

In the first term on the right of (8.7) we group together terms arising from the same primitive character. Since if χ is induced by the primitive character χ^*, then the modulus of χ^* divides the modulus of χ, we get

$$\sum_{q\le Q}\frac{1}{\varphi(q)}\sum_{\chi\neq\chi_0}\max_{y\le x}|\psi_k(y,\chi^*)| = \sum_{1<q\le Q}\sum_{\chi \bmod q}\max_{y\le x}|\psi_k(y,\chi)|\cdot\sum_{\substack{n\le Q\\ n\equiv 0(\bmod q)}}\frac{1}{\varphi(n)}$$

$$\le \sum_{1<q\le Q}\sum_{\chi \bmod q}^{*}\max_{y\le x}|\psi_k(y,\chi)|\cdot\frac{1}{\varphi(q)}\sum_{d\le Q/q}\frac{1}{\varphi(d)}$$

$$\ll \sum_{1<q\le Q}\epsilon(q)\sum_{\chi \bmod q}^{*}\max_{y\le x}|\psi_k(y,\chi)|,$$

where the last relation follows from Lemma 8.4. Hence for $Q \le x^{1/2}$ and bounded k and A,

$$(8.8)\qquad \sum_{q\le Q}\max_{y\le x}\max_{(a,q)=1}\left|\psi_k(y,q,a)-\frac{\psi_k(y)}{\varphi(q)}\right|$$

$$\ll \sum_{1<q\le Q}\epsilon(q)\sum_{\chi \bmod q}^{*}\max_{y\le x}|\psi_k(y,\chi)| + x\ell^{-A}.$$

Siegel's theorem implies that

$$(8.9)\qquad \max_{y\le x}|\psi(y,\chi)| \ll x\ell^{-E}, \text{ for } \chi\neq\chi_0 \text{ and } q\le \ell^C,$$

with arbitrarily large constants C and E. Therefore if k, A, and C are bounded, the terms with $q \le D = \ell^C$ contribute $\ll x\ell^{-A}$ to the right side

of (8.8). Similarly, the prime number theorem implies that

$$\max_{y \leq x} |\psi(y) - y| \ll x\ell^{-E}, \tag{8.10}$$

for arbitrarily large E, and therefore

$$\max_{y \leq x} |\psi_k(y) - y| \ll x\ell^{-E},$$

for arbitrarily large E, provided k is bounded. Hence for $Q \leq x^{1/2}$ and bounded k and A we obtain

$$\sum_{q \leq Q} \max_{y \leq x} \frac{|\psi_k(y)-y|}{\varphi(q)} \ll x\ell^{-A-1} \sum_{q \leq Q} \frac{1}{\varphi(q)} \ll x\ell^{-A}.$$

Combining all our estimates we obtain the statement of the lemma.

Thus our task has been reduced to that of estimating

$$\sum_{D<q\leq Q} \epsilon(q) \sideset{}{^*}\sum_{\chi \bmod q} \max_{y \leq x} |\psi_k(y,\chi)|.$$

The next lemma expresses this sum in a form suitable for application of the large sieve.

(8.11) LEMMA. Let $L = L(s,\chi)$ be a Dirichlet L-series for the character χ and let $S = S(s,\chi)$ be any function bounded and analytic in $\sigma \geq 1/2$. Define $\alpha = 1 + \ell^{-1}$ and $\beta = 1/2 + \ell^{-1}$, and denote by (α) and (β) the paths $\alpha + it$ $(-\infty < t < \infty)$ and $\beta + it$ $(-\infty < t < \infty)$, respectively. Then for $k \geq 2$ we have

$$\sum_{D<q\leq Q} \epsilon(q) \sideset{}{^*}\sum_{\chi \bmod q} \max_{y \leq x} |\psi_k(y,\chi)| \ll x\ell \int_{(\alpha)} \frac{A(s)}{|s|^{k+1}}\, dt + x^{1/2} \int_{(\beta)} \frac{B(s)}{|s|^{k+1}}\, dt,$$

where

$$A(s) = \sum_{D<q\leq Q} \epsilon(q) \sum_{\chi \bmod q}^{*} |1 - LS|^2$$

and

$$B(s) = \sum_{D<q\leq Q} \epsilon(q) \sum_{\chi \bmod q}^{*} (|L'LS^2| + |L'S|).$$

Proof: Since for $k \geq 1$ we have

$$\frac{1}{2\pi i} \int_{(\alpha)} \frac{x^s}{s^{k+1}} ds = \begin{cases} \frac{1}{k!} \log^k x & \text{if } x \geq 1, \\ 0 & \text{if } 0 < x < 1, \end{cases} \tag{8.12}$$

and the series

$$-\frac{L'}{L}(s,\chi) = \sum_{n=1}^{\infty} \frac{\chi(n)\Lambda(n)}{n} \tag{8.13}$$

converges absolutely for $\sigma > 1$, we obtain

$$\psi_k(y,\chi) = \frac{1}{2\pi i} \int_{(\alpha)} \frac{y^s}{s^{k+1}} \left(-\frac{L'}{L}(s,\chi)\right) ds.$$

Now

$$\frac{L'}{L} = \frac{L'}{L}(1 - LS)^2 + (2L'S - L'LS^2). \tag{8.14}$$

For $\chi \neq \chi_0$, L and L' are analytic in $\sigma \geq 1/2$. Moreover, $L(s) \ll |s|^{1/2}$ for $\sigma \geq 1/2$, and since by Cauchy's formula

$$L'(s) = \frac{1}{2\pi i} \int_{\gamma} \frac{L(z)}{(z-s)^2} dz, \tag{8.15}$$

where γ is a circle of radius ℓ^{-1} centered at s, we have $L'(s) \ll |s|^{1/2}$ for $\sigma \geq \beta = 1/2 + \ell^{-1}$. Therefore the second term in (8.14) is analytic and $\ll |s|$ in $\sigma \geq \beta$, and hence for $k \geq 2$ we can transfer the path of integration of the second term from (α) to (β). We find

$$\psi_k(y,\chi) = \frac{1}{2\pi i}\int_{(\alpha)} \frac{y^s}{s^{k+1}}\left(-\frac{L'}{L}\right)(1 - LS)^2 ds + \frac{1}{2\pi i}\int_{(\beta)} \frac{y^s}{s^{k+1}}(L'LS^2 - 2L'S)ds.$$

On $\sigma = \alpha$, we have for $1 \leq y \leq x$,

$$y^s = y^{1+\ell^{-1}+it} = e^{(1+\ell^{-1})\log y + it\cdot\log y} \ll e^{\log x + 1} \ll x.$$

Similarly, on $\sigma = \beta$ we have $y^s \ll x^{1/2}$ for $1 \leq y \leq x$. Also, $\psi(u) \ll u$ and (8.13) imply that on $\sigma = \alpha$ we have

$$\frac{L'}{L}(s,\chi) \ll \sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^{\alpha}} = \alpha\int_1^{\infty} \frac{\psi(u)}{u^{1+\alpha}}\,du + O(1) \ll \int_1^{\infty} \frac{du}{u^{\alpha}} = \ell.$$

Hence we find that

$$\max_{y\leq x} |\psi_k(y,\chi)| \ll x\ell\int_{(\alpha)} \frac{|1 - LS|^2}{|s|^{k+1}}\,dt + x^{1/2}\int_{(\beta)} \frac{|L'LS^2| + |L'S|}{|s|^{k+1}}\,dt.$$

To finish the proof we just sum this estimate, multiplied by $\epsilon(q)$, over all q with $D < q \leq Q$ and all **primitive** characters to the modulus q.

Application of the large sieve inequalities. We now use the inequalities of Chapter 6 to estimate A(s) on $\sigma = \alpha$ and B(s) on $\sigma = \beta$.

Let a_n be any complex numbers. Then by Lemma 6.38

$$\frac{q}{\varphi(q)} \sum_{\chi \bmod q}^{*} \left|\sum_{n=M+1}^{M+N} a_n\chi(n)\right|^2 \leq \sum_{\substack{a=1\\(a,q)=1}}^{q} \left|\sum_{n=M+1}^{M+N} a_n e\left(\frac{na}{q}\right)\right|^2, \tag{8.16}$$

and by Theorem 6.30

$$\sum_{q\le X}\sum_{\substack{a=1\\(a,q)=1}}\Big|\sum_{n=M+1}^{M+N}a_n e(\tfrac{na}{q})\Big|^2 \ll (X^2+N)\sum_{n=M+1}^{M+N}|a_n|^2.$$

Combining these two inequalities leads to

$$(8.17)\qquad \sum_{q\le X}\frac{q}{\varphi(q)}\cdot\sideset{}{^*}\sum_{\chi \bmod q}\Big|\sum_{n=M+1}^{M+N}a_n\chi(n)\Big|^2 \ll (X^2+N)\sum_{n=M+1}^{M+N}|a_n|^2.$$

Lemma 6.39 implies that for $D \le Q/\log Q$ we have

$$\sum_{D<q\le Q}\frac{\log(Q/q)+1}{q}\sum_{\substack{a=1\\(a,q)=1}}^{q}\Big|\sum_{n=M+1}^{M+N}a_n\, e(\tfrac{na}{q})\Big|^2$$

$$\ll \Big\{\frac{\log(Q/D)+1}{D}(D^2+N)+\int_D^Q(1+\log(Q/x))dx\Big\}\sum_{n=M+1}^{M+N}|a_n|^2$$

$$\ll (Q+\frac{N\log Q}{D})\sum_{n=M+1}^{M+N}|a_n|^2.$$

Combining this with (8.15) leads to

$$(8.18)\qquad \sum_{D<q\le Q}\epsilon(q)\sideset{}{^*}\sum_{\chi \bmod q}\Big|\sum_{n=M+1}^{M+N}a_n\chi(n)\Big|^2 \ll (Q+\frac{N\log Q}{D})\sum_{n=M+1}^{M+N}|a_n|^2,$$

again with the restriction $D \le Q/\log Q$.

For χ a character to the modulus q, define

$$R(x,\chi)=\sum_{n\le x}\chi(n).$$

Then by a theorem of Pólya and Vinogradov

$$R(x,\chi) \ll q^{1/2} \log q, \quad \text{for } \chi \neq \chi_0. \tag{8.19}$$

Since the Dirichlet series for $L(s,\chi)$ converges for $\sigma > 0$ whenever $\chi \neq \chi_0$, we have under these assumptions, for each integer $H \geq 1$,

$$L(s,\chi) = \sum_{n=1}^{\infty} \frac{\chi(n)}{n^s} = \sum_{n=1}^{H} \frac{\chi(n)}{n^s} + s\int_N^{\infty} \frac{R(t,\chi)}{t^{s+1}}\,dt - \frac{R(H,\chi)}{H^s} \tag{8.20}$$

$$= \sum_{n=1}^{H} \frac{\chi(n)}{n^s} + O\Big(\frac{q^{1/2}\log q}{H^{\sigma}}\,(1 + \frac{|s|}{\sigma})\Big)\,.$$

Therefore for $\sigma \geq 1$, $H \geq 1$, and $\chi \neq \chi_0$, we have

$$L(s,\chi) = \sum_{n=1}^{H} \frac{\chi(n)}{n^s} + O(\frac{|s|q^{1/2}\log q}{H}).$$

We choose

$$S(s,\chi) = \sum_{n=1}^{H} \frac{\mu(n)\chi(n)}{n^s}\,,$$

and notice that it satisfies the hypothesis of Lemma 8.11. Moreover, for $H \leq x$ and $\sigma \geq 1$ we have $S(s,\chi) \ll \log H \leq \ell$, so that for $q \leq x$,

$$1 - L(s,\chi)S(s,\chi) = \sum_{1}^{\infty} \frac{c(n)\chi(n)}{n^s} + O(\frac{|s|q^{1/2}\ell^2}{H}),$$

where $c(1) = 0$, and for $n > 1$,

$$c(n) = \sum_{\substack{d|n \\ d,(n/d)\leq H}} \mu(d).$$

In particular, $c(n) = 0$ for $n > H^2$ and $n \le H$, and $|c(n)| \le d(n)$, the number of divisors of n, for $H < n \le H^2$. We now separate the integers $H+1, \ldots, H^2$ into m sets: $\{H+1, \ldots, 2H\}$, $\{2H+1, \ldots, 4H\}$, ..., $\{2^{m-1}H+1, \ldots, 2^m H\}$, where $2^{m-1}H+1 \le H^2 \le 2^m H$, so that $m \ll \log H \le \ell$. Using the Cauchy-Schwarz inequality, we obtain

$$\Big|\sum_{n=1}^{\infty} \frac{c(n)\chi(n)}{n^s}\Big|^2 = \Big|\sum_{h=0}^{m-1}\Big(\sum_{n=2^hH+1}^{2^{h+1}H} \frac{c(n)\chi(n)}{n^s}\Big)\Big|^2 \ll \ell \sum_{h=0}^{m-1}\Big|\sum_{n=2^hH+1}^{2^{n+1}H} \frac{c(n)\chi(n)}{n^s}\Big|^2.$$

Next we apply (8.17) with $a_n = c(n)/n^s$ to each of the factors on the right side above. For each $N = 2^hH$ ($\le x^2$), with $D \le Q/\log Q$ and $Q \le x$, we find

$$\sum_{D<q\le Q} \epsilon(q) \sideset{}{^*}\sum_{\chi \bmod q} \Big|\sum_{n=N+1}^{2N} \frac{c(n)\chi(n)}{n^s}\Big|^2 \ll (Q + \tfrac{N\ell}{D}) \sum_{n=N+1}^{2N} \frac{d^2(n)}{n^2} \ll (\tfrac{Q}{N} + \tfrac{\ell}{D})\ell^3,$$

since the **inequality**

$$T(M) = \sum_{n\le M} d^2(n) \ll M \log^3 M \tag{8.21}$$

implies

$$\sum_{n=N+1}^{2N} \frac{d^2(n)}{n^2} < \frac{1}{N^2}\sum_{n=N+1}^{2N} d^2(n) \ll \frac{\log^3 N}{N}.$$

Therefore, using $m \ll \ell$,

$$A(s) \ll \sum_{h=0}^{m-1} (\frac{Q}{2^hH} + \tfrac{\ell}{D})\ell^4 + \sum_{D<q\le Q} (\log(Q/q) + 1)\frac{|s|^2 q\ell^4}{H^2}$$

$$\ll QH^{-1}\ell^4 + D^{-1}\ell^6 + |s|^2 Q^2 H^{-2}\ell^4.$$

We can choose $H = QD\ell^{-2}$ to get

$$A(s) \ll D^{-1}|s|^2\ell^6, \text{ on } \sigma = \alpha, \tag{8.22}$$

provided $\ell \le D \le Q/\log Q$.

To estimate $B(s)$ on $\sigma = \beta$, we apply (8.17) to

$$S^2(s,\chi) = \sum_{n=1}^{H^2} \frac{b(n)\chi(n)}{n^s},$$

where $|b(n)| \le d(n)$, and get for $\sigma \ge 1/2$

$$\sum_{q\le X} \frac{q}{\varphi(q)} \sum_{\chi \bmod q}^{*} |S(s,\chi)|^4 \ll (X^2+H^2)\sum_{n=1}^{H^2} \frac{d^2(n)}{n} \ll (X^2+H^2)\ell^4, \tag{8.23}$$

when the last estimate again follows from (8.21) by partial summation.

It follows from (8.20) that for each $N \ge 1$,

$$L(s,\chi) = \sum_{n=1}^{N} \frac{\chi(n)}{n^s} + O\left(\frac{|s|q^{1/2}\log q}{N^{1/2}}\right), \text{ for } \chi \ne \chi_0 \text{ and } \sigma \ge 1/2.$$

Therefore

$$\sum_{1<q\le X} \frac{q}{\varphi(q)} \sum_{\chi \bmod q}^{*} |L(s,\chi)|^4 \ll \tag{8.24}$$

$$\ll \sum_{q\le X} \frac{q}{\varphi(q)} \sum_{\chi \bmod q}^{*} \left| \sum_{n=1}^{N} \frac{\chi(n)}{n^s} \right|^4 + \frac{|s|^4X^4\log X}{N^2}.$$

We now apply (8.17) to the sum on the right. Since

$$\left(\sum_{n=1}^{N} \frac{\chi(n)}{n^s}\right)^2 = \sum_{n=1}^{N^2} \frac{a(n)\chi(n)}{n^s},$$

with $|a(n)| \le d(n)$, we find that (8.24) is

$$\ll (X^2 + N^2) \sum_{n=1}^{N^2} \frac{d^2(n)}{n} \ll (X^2 + N^2)\log^4 N.$$

We now choose $N = [X|s| + 2]$. Then the above inequalities give

$$\sum_{1<q\le X} \frac{q}{\varphi(q)} \sideset{}{^*}\sum_{\chi \bmod q} |L(s,\chi)|^4 \ll X^2|s|^2\log^4(X|s| + 2).$$

Since $\zeta(s) \ll |s|^{1/2}$ for $1/2 \le \sigma < 3/4$, say, we actually have

$$(8.25)\quad \sum_{q\le X} \frac{q}{\varphi(q)} \sideset{}{^*}\sum_{\chi \bmod q} |L(s,\chi)|^2 \ll X^2|s|^2\log^4(X|s| + 2) \text{ for } 1/2 \le \sigma < 3/4.$$

Applying Hölder's inequality to (8.15) twice gives

$$|L'(s,\chi)|^4 \ll \ell^5 \int_\gamma |L(z,\chi)|^4.$$

Therefore on $\sigma = \beta$ we have

$$(8.26)\qquad \sum_{q\le X} \frac{q}{\varphi(q)} \sideset{}{^*}\sum_{\chi \bmod q} |L'(s,\chi)|^4 \ll X^2|s|^2\log^4(X|s| + 2)\ell^4.$$

Using the Cauchy-Schwarz inequality, we now find that

$$\sum_{q\le X} \frac{q}{\varphi(q)} \sideset{}{^*}\sum_{\chi \bmod q} |L'LS^2| \le \Big(\sum_{q\le X} \frac{q}{\varphi(q)} \sideset{}{^*}\sum_{\chi \bmod q} |L'L|^2\Big)^{1/2}\Big(\sum_{q\le X} \frac{q}{\varphi(q)} \sideset{}{^*}\sum_{\chi \bmod q} |S|^4\Big)^{1/2}$$

$$\le \Big(\sum_{q\le X} \frac{q}{\varphi(q)} \sideset{}{^*}\sum_{\chi \bmod q} |L'|^4\Big)^{1/4}\Big(\sum_{q\le X} \frac{q}{\varphi(q)} \sideset{}{^*}\sum_{\chi \bmod q} |L|^4\Big)^{1/4}\Big(\sum_{q\le X} \frac{q}{\varphi(q)} \sideset{}{^*}\sum_{\chi \bmod q} |S|^4\Big)^{1/2}$$

$$\ll (X + H)\cdot X\cdot|s|\cdot\ell^3\cdot\log^2(X|s| + 2).$$

If we now choose $H = DQ\ell^{-2}$, then we easily obtain

$$B(s) \ll DQ\ell^5|s|\log^2(|s| + 2) \text{ for } \sigma = \beta. \tag{8.27}$$

Completion of the proof. Using Lemma 8.5, Lemma 8.11, (8.22), and (8.27), we find that for $k \geq 2$, $Q \leq x^{1/2}$, and bounded k and A,

$$\sum_{q\leq Q} \max_{y\leq x} \max_{(a,q)=1} |\psi_k(y,q,a) - \frac{y}{\varphi(q)}| \ll x\ell \int_{(\alpha)} \frac{D^{-1}|s|^2\ell^6}{|s|^{k+1}}\,dt + x^{1/2}\int_{(\beta)} \frac{DQ\ell^5|s|\log^2(|s|+2)}{|s|^{k+1}}\,dt + x\ell^{-A}, \tag{8.28}$$

provided $\ell \leq D \leq Q/\log Q$. For $k = 3$ the right side of (8.28) is

$$\ll xD^{-7}\ell^7 + x^{1/2}DQ\ell^5 + x\ell^{-A}.$$

Choosing $D = \ell^{A+7}$ and $Q = x^{1/2}\ell^{-(2A+12)}$, we find

$$\sum_{q\leq Q} \max_{y\leq x} \max_{(a,q)=1} |\psi_3(y,q,a) - \frac{y}{\varphi(q)}| \ll x\ell^{-A}. \tag{8.29}$$

We now deduce (8.2) from (8.29). Suppose that the following statement is true for a positive integer k:

(8.30) For every positive constant A, there is a constant $B_k(A)$ such that if

$$Q = x^{1/2}\ell^{-B_k(A)},$$

then

$$\sum_{q\leq Q} \max_{y\leq x} \max_{(a,q)=1} |\psi_k(y,q,a) - \frac{y}{\varphi(q)}| \ll x\ell^{-A},$$

with the implied constant depending only on A.

We will show that then (8.30) also holds for k - 1, with $B_{k-1}(A) = B_k(2A + 1)$.

Since $\psi_{k-1}(y,q,a)$ is an increasing function of y, we have, for $0 < \lambda \leq 1$,

$$\frac{1}{\lambda}\int_{e^{-\lambda}y}^{y} \psi_{k-1}(z,q,a)\frac{dz}{z} \leq \psi_{k-1}(y,q,a) \leq \frac{1}{\lambda}\int_{y}^{e^{\lambda}y} \psi_{k-1}(z,q,a)\,\frac{dz}{z}\,.$$

Evaluating the integrals we obtain

$$(8.31)\qquad \frac{1}{\lambda}\{\psi_k(y,q,a) - \psi_k(e^{-\lambda}y,q,a)\} \leq \psi_{k-1}(y,q,a) \leq \frac{1}{\lambda}\{\psi_k(e^{\lambda}y,q,a) - \psi_k(y,q,a)\}.$$

If we write

$$\psi_k(x,q,a) = \frac{x}{\varphi(q)} + r_k(x,q,a),$$

then (8.31) implies that

$$\frac{1}{\lambda}\{\frac{y(1-e^{-\lambda})}{\varphi(q)} - 2\max_{z\leq y}|r_k(z,q,a)|\} \leq \frac{y}{\varphi(q)} + r_{k-1}(y,q,a) \leq \frac{1}{\lambda}\{\frac{y(e^{\lambda}-1)}{\varphi(q)} + 2\max_{z\leq e^{\lambda}y}|r_k(z,q,a)|\}.$$

But $1 - e^{-\lambda} = \lambda + O(\lambda^2)$ and $e^{\lambda} - 1 = \lambda + O(\lambda^2)$, so that

$$|r_{k-1}(y,q,a)| \ll \frac{\lambda y}{\varphi(q)} + \frac{1}{\lambda}\max_{z\leq e^{\lambda}y}|r_k(z,q,a)|.$$

Therefore for $Q \leq x^{1/2}$ we have

$$\sum_{q\leq Q}\max_{y\leq x}\max_{(a,q)=1}|r_{k-1}(y,q,a)| \ll \lambda x\ell + \lambda^{-1}\sum_{q\leq Q}\max_{y\leq ex}\max_{(a,q)=1}|r_k(y,q,a)|.$$

But by (8.30) the right side is

$$\ll \lambda x\ell + \lambda^{-1}x\ell^{-A}, \quad \text{for } Q \leq x^{\frac{1}{2}}\ell^{-B_k(A)}$$

$$\ll x\ell^{-\frac{1}{2}(A-1)}, \quad \text{if } \lambda = \ell^{-\frac{1}{2}(A+1)}.$$

This proves our assertion about (8.30). Since we have proved that $B_3(A) = 2A + 12$, this gives us $B_0(A) = 16A + 103$.

Notes on Chapter 8.

The proof given above is due to Gallagher [2]. The original proof (Bombieri [3]), which is considerably more involved, relates the sum (8.2) to the density of zeros of Dirichlet L-series in the initial and avoids the use of our smoothing device. (There is also a good exposition of this method in Davenport [1; Sections 24-28].) Bombieri proved that B can be taken as 3A + 23, but this is not very important since the use of Siegel's theorem (which plays a crucial role in both Bombieri's and Gallagher's proofs) has so far prevented the determination of the constant implied by the $\ll$ sign.

Our proof presupposes a fair knowledge of analytic number theory, but practically everything used is quite standard and may be found in Prachar [1] and Ayoub [1]. The inequality (8.21), which is not known too well, is proved in Prachar [1; Chapter 1].

Many applications of Bombieri's theorem have been made. One of the most interesting is the proof by Elliott and Halberstam [1] that every sufficiently large integer n may be represented as $n = p + x^2 + y^2$; that is, as a sum of a prime and two squares. Their paper lists previous proofs of this theorem, some of which used unproved conjectures such as the extended Riemann hypothesis.

The large sieve estimates have been used in proving several other important results in analytic number theory. We might mention here the papers of Jutila [1] and Montgomery [2], [3].

BIBLIOGRAPHY

The following list contains, besides some general references, all the publications on sieve methods of which the author is aware. Relatively few of them are referred to in the text.

ANKENY, N. C.

1. Applications of the sieve, Proceedings of Symposia in Pure Mathematics, 8(1965), 113-118. MR 31, no. 141.

ANKENY, N. C., and ONISHI, M.

1. The general sieve, Acta Arith. 10(1964), 31-62. MR 29, no. 4740.

APOSTOL, T. M.

1. Euler's φ-function and separable Gauss sums, Proc. Amer. Math. Soc. 24(1970), 482-485. MR 41, no. 1661.

AYOUB, R.

1. An Introduction to the Analytic Theory of Numbers, Mathematical Surveys, no. 10, Amer. Math. Soc., Providence, R. I., 1963. MR 28 no. 3954.

BARBAN, M. B.

1. On a theorem of I. P. Kubilius, Izv. Akad. Nauk UZSSR Ser. Fiz.-Mat. Nauk 1961, no. 5, 3-9 and 1963, no. 1, 82-83. (Russian). MR 25 no.2051.

2. New applications of the "large sieve" of Ju. V. Linnik, Akad. Nauk UzSSR Trudy Inst. Mat. 1961, no. 22, 1-20. (Russian). MR 30 no. 1990. A remark on the author's paper "New applications .....," Theory Prob. Math. Statist., Tashkent, 1964, 130-133. (Russian).

3. Linnik's "large sieve" and a limit theorem for the class number of ideals of an imaginary quadratic field, Izv. Akad. Nauk SSSR Ser. Mat. 26 (1962), 573-580. (Russian). MR 27, no. 1426.

4. Analogues of the divisor problem of Titchmarsh, Vestnik Leningrad. Univ. 1963, no. 19, 5-13. (Russian).

5. The density of zeros of Dirichlet L-series and the problem of sums of prime and "almost prime" numbers, Mat. Sb. N.S. 61(103)(1963), 418-425. (Russian). MR 30, no. 1992.

BARBAN, M. B. - continued

6. On the number of divisors of "translations" of the prime number twins, Acta Math. Akad. Sci. Hungar. 15(1964), 285-288. (Russian). MR 30, no. 1102.

7. The "large sieve" method and its applications in the theory of numbers, Uspehi Mat. Nauk 21(1966), no. 1 (127), 51-102 = Russian Math. Surveys, 21(1966), 49-103. MR 33 no. 7320.

8. On Bombieri's density conjecture, Dokl. Akad. Nauk SSSR 172(1967), 999-1000 = Soviet Math. Dokl. 8(1967), 202-203. MR 35 no. 5402.

BATEMAN, P. T., and Horn, R. A.

1. Primes represented by polynomials in one variable, Proceedings of Symposia in Pure Mathematics, 8(1965), 119-132, MR 31, no. 1234.

BATEMAN, P. T., and STEMMLER, R. M.

1. Waring's problem for algebraic number fields and primes of the form $(p^r-1)/(p^d-1)$, Illinois J. Math. 6(1962), 142-156. MR 25 no. 2059.

BOMBIERI, E.

1. Maggiorazione del resto nel "Primzahlsatz" col metodo di Erdös-Selberg, Inst. Lombardo Acad. Sci. Latt. Rend. A96(1962), 343-350. MR 27, no 120.

2. Sulle formule di A. Selberg generalizzate per classi di funzioni aritmetiche e le applicazioni al problema del resto nel "Primzahlsatz", Riv. Mat. Univ. Parma (2) 3(1962), 393-440. MR 27, no. 4804.

3. On the large sieve, Mathematika 12(1965), 201-225. MR 33, no. 5590.

4. On a theorem of Van Lint and Richert, Symposia Mathematica 4(1970), 175-180.

BOMBIERI, E., and DAVENPORT, H.

1. Small differences between prime numbers, Proc. Royal Soc. A, 293(1966), 1-18. MR 33, no. 7314.

BOMBIERI, E and DAVENPORT, H. - continued

2. On the large sieve method, Number Theory and Analysis (papers in honor of E. Landau, ed. by P. Turán), Berlin-New York 1969, 9-22. MR 41, no.5327.

3. Some inequalities involving trigonometrical polynomials, Ann. Scuola Norm. Sup. Pisa. Sci. fis. mat. III, Ser. 23 (1969), 223-241. MR 40, no. 2636.

BRUIJN, N. G. DE

1. On the number of uncancelled elements in the sieve of Eratosthenes, Nederl. Akad. Wetensch. Proc. 52 (1950), 803-812 = Indagationes Math. 12 (1950), 247-256. MR 12-11.

2. The asymptotic behavior of a function occurring in the theory of primes, Indian Math. Soc. N. S. 15(1951), 25-32. MR 13-326.

3. On the number of positive integers $\leq x$ and free of prime factors $> y$, Nederl. Akad. Wetensch.Proc. 54(1951), 50-60 = Indagationes Math. 13 (1951), 2-12. MR 13-724.

BRUN, V.

1. Le crible d'Eratosthene et le theoreme de Goldbach, Norske Vid-Selsk. Skr., I M.-N. Kl., Kristiania, 3(1920), 1-36.

2. Le série $\frac{1}{3} + \frac{1}{5} + \frac{1}{7} + \frac{1}{11} + \frac{1}{13} + \frac{1}{17} + \frac{1}{19} + \frac{1}{29} + \frac{1}{31} + \cdots$ est convergente ou finie, Bulletin des Sciences Mathematiques 43(1919), 100-104, 124-128.

BUCHSTAB, A. A.

1. Asymptotische Abschätzungen einer allgemeinen zahlentheoretischen Funktion, Mat. Sb. N.S. 44(1937), 1239-1246.

2. Neue Verbesserung in der Methode des Eratosthenischen Siebes, Rec. Math. N.S. 4(1938), 375-387.

3. On an additive representation of integers, Mat. Sb. N.S. 10(52), (1942), 87-91. (Russian).

BUCHSTAB, A. A. - continued

4. On those numbers in an arithmetic progression all prime factors of which are small in magnitude, Dokl. Akad. Nauk SSSR N.S. 67(1949), 5-9. (Russian). MR 11-84, 871.

5. Sur la decomposition des nombres pairs en somme de deux composantes dont chacune est formée d'un nombre borné de facteurs premiers, Dokl. Akad. Nauk SSSR N.S. 29(1949), 544-548. MR 2-348.

6. On an asymptotic estimate of the number of numbers of an arithmetic progression which are not divisible by "relatively" small prime numbers, Mat. Sb. N.S. 28(70), (1951), 165-184. (Russian). MR 13-626.

7. New results in the Goldbach-Euler problem and the twin prime problem, Dokl. Akad. Nauk SSSR, 162(1965), 735-738 = Soviet Math. Doklady, 6(1965), 729-732. MR 31, no. 2226.

8. A combinatorial strengthening of the Eratosthenes' sieve method, Uspehi Mat. Nauk 22 (3)(135), (1967), 199-226 = Russian Math. Surveys 22(1967), 205-233. MR 36, no. 1413.

BURGESS, D. A., and ELLIOTT, P. D. T. A.

1. On the average order of the least primitive root, Mathematika 15(1968), 39-50. MR 38, no. 5736.

DAVENPORT, H.

1. Multiplicative Number Theory, Chicago 1967. MR 36, no. 117. (See also Davenport and Halberstam [2], below.)

DAVENPORT, H., and HALBERSTAM, H.

1. The values of a trigonometrical polynomial at well spaced points, Mathematika 13(1966), 91-96. MR 33, no. 5592. (See also [2] below.)

2. Corrigendum and Addendum, Mathematika 14(1967), 229-232. MR 36, no. 2569.

3. Primes in arithmatic progressions, Michigan Math. J. 13(1966), 485-489. MR 34, no. 156.
+ Corrigendum, Michigan Math. J., 15(1968), 505. MR 38, no. 2099.

ELLIOTT, P. D. T. A.

1. The distribution of primitive roots, Canad. J. Math. 21(1969), 822-841. MR 40, no. 104.

ELLIOTT, P. D. T. A., and HALBERSTRAM, H.

1. Some applications of Bombieri's theorem, Mathematika 13 (1966), 196-203. MR 34, no. 5788.

ERDÖS, P.

1. The difference of consecutive primes, Duke Math. J. 6(1940), 438-441. MR 1-292.
2. On some applications of Brun's method, Acta Sci. Math. Szeged. 13(1949), 57-63. MR 10-684.
3. Problems and results on the difference of consecutive primes, Publ. Math. Debrecen 1(1949), 33-37. MR 11-84.

ERDÖS, P., and RÉNYI, A.

1. Some problems and results on consecutive primes, Simon Stevin 27(1950), 115-125. MR 11-644.
2. Some remarks on the large sieve of Yu. V. Linnik, Ann. Univ. Sci. Budapest, Eötvös Sect. Math. 11(1968), 3-13. MR 39, no. 2718.

ESTERMANN, T.

1. Eine neue Darstellungen und neue Anwendungen der Viggo Brunschen Methode, J. Reine Angew. Math. 168(1932), 106-116.

FAINLEIB, A. S.

1. The limit theorem for the number of classes of primitive quadratic form with negative discriminant , Dokl. Akad. Nauk SSSR 184(1969), 1048-1049 = Soviet Math. Doklady 10(1969), 206-207. MR 39, no. 5474.

FLUCH, W.

1. Verwendung der Zeta-Funktion beim Sieb von Selberg, Acta Arith. 5(1959), 381-405. MR 23, no. A1614.

GALLAGHER, P. X.

1. The large sieve, Mathematika 14(1967), 14-20. MR 35, no. 5411.
2. Bombieri's mean value theorem, Mathematika 14 (1968), 1-6. MR 38, no. 5724.
3. A larger sieve, to appear in Acta Arith. vol. 18.

GELFOND, A. O., and LINNIK, Yu. V.

1. Elementary Methods in Analytic Number Theory, Chicago 1965. MR 32, no. 5575.

GOLDFELD, M.

1. On the number of primes p for which p + a has a large prime factor, Mathematika 16(1969), 23-27. MR 39, no. 5493.

GOLOMB, S. W.

1. The twin prime constant, Amer. Math. Monthly 67(1960), 767-769.

HALBERSTAM, H.

1. Footnote to the Titchmarsh-Linnik divisor problem, Proc. Amer. Math. Soc. 18(1967), 187-188. MR 34, no. 4221.

HALBERSTAM, H., JURKAT, W., and RICHERT, H. E.

1. Un nouveau résultat de la méthode du crible,Comptes Rendus A, 264(1967), 920-923. MR 36, no. 6374.

HALBERSTAM, H., and RICHERT, H. E.

1. Sieve Methods, to be published by Markham Publ. Co., Chicago.

HALBERSTAM, H., and ROTH, K. F.

1. Sequences, vol. I, Oxford, 1966. MR 35, no. 1565.

HALBERSTAM, H., and ROTKIEWICZ, A.

1. A gap theorem for pseudo primes in arithmetic progressions, Acta Arith. 13(1967/68), 395-404. MR 38, no. 1329.

HARDY, G. H., and LITTLEWOOD, J. E.

1. Some problems of partitio numerorum, Acta Math. 44(1923), 1-70.

HARDY, G. H., and WRIGHT, E. M.

1. An Introduction to the Theory of Numbers, 4th ed., Oxford, 1960.

HEILBRONN, H.

1. Über die Verteilung der Primzahlen in Polynomen, Math. Annalen, 104(1931), 794-799.

HEILBRONN, H., LANDAU, E., and SCHERK, P.

1. Alle grossen ganzen Zahlen lassen sich als Summe von höchsten 71 Primzahlen darstellen, Casopis Pro Pestorani Matematiky a Fysiky, 65 (1936), 117-140.

HUA, L. K.

1. Die Abschätzung von Exponentialsummen und ihre Anwendung in der Zahlentheorie, Enzyklopädie der Math. Wissenschaften, Leipzig 1959, vol. I2, book 13, part 1, Art. 29. MR 24, no. A94.

2. Additive Theory of Prime Numbers, Translations of Mathematical Monographs, no. 13, Providence, R. I., 1965.

HUXLEY, M. N.

1 The large sieve inequality for algebraic number fields, Mathematika 15(1968), 178-187. MR 38, no. 5737.

2. On the differences of primes in arithmetical progressions, Acta Arith. 15(1968/69), 367-392. MR 39, no. 5494.

IZUMI, M., and IZUMI, S.

1. On a theorem concerning trigonometrical polynomials, Proc. Japan Acad. 43(1967), 71-76. MR 36, no. 592.

JAMES, R. D.

1. On the sieve method of Viggo Brun, Bull. Amer. Math. Soc. 49(1943), 422-432. MR 4-265.

2. Recent progress on the Goldbach problem, Bull. Amer. Math. Soc. 55(1949), 246-260. MR 10-515, 856.

JURKAT, W., and RICHERT, H. -E.

1. An improvement of Selberg's sieve method I, Acta Arith. 11(1965), 217-240. MR 34, no. 2540.

JUTILA, M.

1. A statistical density theorem for L-functions with applications, Arta Arith. 16(1969), 207-216. MR 40, no. 5557.

KÁTAI, I.

1. A note on a sieve method, Publ. Math. Debrecen 15(1968), 69-73. MR 38, no. 5730.

KLIMOV, N. I.

1. Combination of elementary and analytic methods in the theory of numbers, Uspehi Mat. Nauk. 13:3 (81) (1958), 145-164. (Russian). MR 20, no. 3841.

2. Almost prime numbers, Uspehi Mat. Nauk. 16(3)(99)(1961), 181-188. (Russian). MR 23, no. A2398.

KUHN, P.

1. Zur Viggo Brunschen Siebmethode, Norske Vid. Selsk. Forh., Trondheim, 14, no. 39 (1941), 145-148. MR 8-503.

2. Neue Abschätzungen auf Grund der Viggo Brunschen Siebmethode, 12. Skand. Mat. Kongr., Lund 1953, 160-168. MR 16-676.

3. Über die Primteilen eines Polynomes , Proc. Inter. Math. Congress, Amsterdam (1954), 35-37.

LANDAU, E.

1. Die Goldbache Vermutung und der Schnirelmannsche Satz, Göttinger Nachrichten, 1930, 255-276.

2. Elementary Number Theory, New York, 1958.

LEVIN, B. V.

1. Lower estimates for the number of almost primes in some general sequences, Vestnik Leningrad. Univ. 15(1960), no. 7, 48-65. (Russian). MR 22, no. 7985.

2. Distribution of almost prime numbers in sequences generated by integer-valued polynomials, Dokl. Akad. Nauk Uz3SR. 7(1962), no. 11. (Russian).

3. On a class of number-theoretic problems that reduce to differential equations, Trudy Tashkent Gos. Univ. 228 (1963), 56-68. (Russian).

4. Estimates of special sums and products which are connected to sieve methods, Trudy Tashkent Gos. Univ. 228(1963), 69 - 80. (Russian).

5. Distribution of "almost prime" numbers in polynomial sequences, Mat. Sb. N.S. 61(103)(1963), 389-407. (Russian). MR 30, no 1991.

6. A one-dimensional sieve, Acta Arith. 10(1965), 387-397. (Russian). MR 31, no. 4774.

7. Comparison of A. Selberg's and V. Brun's sieves, Uspehi Mat. Nauk. 20(1965), no. 5 (125), 214-220. (Russian). MR 32, no. 7534.

8. On the least almost prime number in an arithmetic progression and the sequence $k^2x^2 + 1$, Uspehi Mat. Nauk. 20(1965), no. 4(124), 158-162. (Russian). MR 32, no. 5612.

LEVIN, B. V., and MAKSUDOV, I. G.

1. Distribution of almost prime numbers in polynomials in n variables, Izv. Akad. Nauk UzSSR Ser. Fiz.-Mat. Nauk 10(1966), no. 3, 15-23. (Russian). MR 33, no. 7316.

LEVIN, B. V., and TULJAGANOVA, M. I.

1. The sieve of A. Selberg in algebraic number fields, Litovsk. Mat. Sb. 6(1966), 59-73. (Russian). MR 34, no. 7492.

LINNIK, Yu. V.

1. The large sieve, Dokl. Akad. Nauk SSSR 30(1941), 292-294. MR 2-349.

2. A remark on the least quadratic non-residue, Dokl. Akad. Nauk SSSR, 36(1942), 119-120. MR 4-189.

LINT, J. H. VAN, and RICHERT, H. -E.

1. On primes in arithmetic progressions, Acta Arith. 11(1965), 209-216. MR 32, no. 5613.

MEISSEL, E.

1. Über die Bestimmung der Primzahlenmenge innerhalb gegebener Grenze, Math. Ann. 2(1870), 636-642.

MIECH, R. J.

1. Almost primes generated by a polynomial, Acta Arith. 10(1964), 9-30. MR 29, no. 1174.

2. Primes, polynomials, and almost primes, Acta Arith. 11(1965), 35-56. MR 31, no. 3390.

3. A uniform result on almost primes, Acta Arith. 11(1966), 371-391. MR 34, no. 1287.

MIENTKA, W. E.

1. An application of the Selberg sieve method, Indian Math. Soc. J. N.S. 25-26 (1961-62), 129-138. MR 27, no. 2490.

MONTGOMERY, H. L.

1. A note on the large sieve, J. London Math. Soc. 43(1968), 93-98. MR 37, no. 184.

2. Mean and large values of Dirichlet polynomials, Invent. Math. 8(1969), 334-345.

3. Zeros of L-functions, Invent. Math. 8(1969), 346-354. MR 40, no. 2620.

NAGEL, T.

1. Generalisation d'un théoréme de Tchebycheff, Journal de Mathématiques (8), 4(1921), 343-356.

NIVEN, I., and ZUCKERMAN, H. S.

1. An Introduction to the Theory of Numbers, New York, 1960. MR 22, no. 5605.

PAN, C.

1. On the representation of even numbers as the sum of a prime and an almost prime, Acta Math. Sinica 12(1962), 95-106 = Chinese Math.-Acta 3(1963), 101-112. MR 27, no. 1427.

2. On the representation of even numbers as the sum of a prime and a product of not more than four primes, Sci. Sinica 12(1963), 455-473. (Russian). MR 28, no. 73.

3. A note on the large sieve method and its applications, Acta Math. Sinica 13(1963), 262-268 = Chinese Math.-Acta 4(1963), 283-290. MR 29, no. 82.

4. A new application of the Ju. V. Linnik large sieve method, Acta Math. Sinica 14(1964), 597-606 = Chinese Math. - Acta 5(1964), 642-652. MR 30, no. 3871.

PRACHAR, K.

1. Primzahlverteilung , Berlin - Göttingen - Heidelberg, 1957. MR 19-393.

RADEMACHER, H.

1. Beitrage zur Viggo-Brunschen Methode in der Zahlentheorie, Abh. Hamburg 3 (1924), 12-30.

2. Lectures on Elementary Number Theory, New York-Toronto-London, 1964.

RAMASWAMI, V.

1. On the number of positive integers $\leq x$ and free of prime divisors $> x^c$, Bull. Amer. Math. Soc. 55 (1949), 1122-1127. MR 11-233.

2. On the number of positive integers $\leq x$ and free of prime factors $> x^c$ and a problem of S. S. Pillai, Duke Math. J. 16 (1949), 99-109, MR 10-597.

RANKIN, R. A.

1. The difference between consecutive prime numbers III, J. London Math. Soc. 22 (1947), 226-230. MR 9-498.

RÉNYI, A.

1. On the representation of an even number as the sum of a single prime and a single almost prime number, Dokl. Akad. Nauk SSSR N.S. 56 (1947), 455-458. (Russian). MR 9-136.

2. On the representation of an even number as the sum of a single prime and a single almost prime number, Izv. Akad. Nauk SSSR Ser. Mat. 12(1948), 57-78 = Translations Amer. Math. Soc. (2) 19 (1962), 299-321. MR 9-413.

3. On the large sieve of Ju. V. Linnik, Compositio Math. 8 (1950), 68-75. MR 11-581.

4. Un nouveau théoréme concernent les fonctions indépendantes et ses applications á la théorie des nombres, J. Math. Pures Appl. 28 (1949), 137-149. MR 11-161.

5. Sur un théoréme general de probabilite, Annals Inst. Fourier Univ. Grenoble 1 (1949), 43-52. MR 14-886.

6. On a general theorem in probability theory and its application in the theory of numbers, Casopis Pest. Mat. Fys. 74 (1949), 167-175. (Russian). MR 12-590.

RÉNYI, A.- continued

7. Probability methods in number theory, Advancement in Math. 4(1958), 465-510. (Chinese). MR 20, no. 4535.

8. On the probabilistic generalization of the large sieve of Linnik, Magyar Tud. Akad. Mit. Kutató Inst. Közl. 3(1958), 199-206. MR 22, no. 1937.

9. New version of the probabilistic generalization of the large sieve, Acta Math. Acad. Sci. Hungar. 10(1959), 217-226. MR 22, no. 1938.

RICCI, G.

1. Ricerche aritmetiche sui polynomi, Rend. Mat. Palermo 57 (1933), 433-475.

2. Sur la congettura di Goldbach e la constante di Schnirelmann, Ann. Scuola Norm. Sup. Pisa (2), 6(1937), 71-116.

3. Sull' andamento della differenza di numeri primi consecutivi, Rivista Mat. Univ. Parma 5(1954), 3-54. MR 16-675.

RICHERT, H. -E.

1. Selberg's sieve with weights, Mathematika 16(1969), 1-22. MR 40, no. 119.

2. Selberg's sieve with weights, Symposia Mathematica, vol. 4(1970), 73-80.

3. Selberg's sieve with weights,to appear in an Amer. Math. Soc. volume on the 1969 Summer Institute on Number Theory at Stony Brook.

RIEGER, G.

1. Anwendung der Siebmethode auf einige Fragen der additiven Zahlentheorie. I., J. Reine Angew. Math. 214/215 (1964), 373-385. MR 29, no. 86.

2. Anwendung der Siebmethode auf einige Fragen der additiven Zahlentheorie. II., Math. Nachr. 28 (1964/5), 207-217. MR 31, no. 131.

3. On polynomials and almost-primes, Bull. Amer. Math. Soc. 75(1969), 100-103. MR 38, no. 2104.

RODRIQUES, G.

1. Sul problema dei divisori di Titchmarsh, Bolletino Unione Math. Italiana (3), 20(1965), 358-366. MR 33, no. 5574.

ROSSER, J. B., and SCHOENFELD, L.

1. Approximate formulas for some functions of prime numbers, Illinois J. Math. 6 (1962), 64-94. MR 25, no. 1139.

ROTH, K. F.

1. On the large sieves of Linnik and Rényi, Mathematika 12 (1965), 1-9. MR 33, no. 5589.
2. The Large Sieve, Imperial College of Science and Technology, Univ. of London, 1968. MR 38, no. 5740.

SAMANDAROV, A. G.

1. The large sieve in algebraic number fields, Mat. Zametki 2 (1967), 673-680. (Russian). MR 36, no. 6379.
2. Equidistribution of prime ideals in classes H mod f, Dokl. Akad. Nauk Tadzik SSR 11(1968), no. 2, 11-13. (Russian). MR 37, no. 1341.

SCHALL, W.

1. On the large sieve method in algebraic number fields, J. of Number Theory 3(1970), 249-270.

SCHINZEL, A.

1. On the congruence $a^x \equiv b \pmod{p}$, Bull. Acad. Polon. Sciences 8(1960), 307-309. MR 23, no. A2377.

SCHNIRELMANN, L. G.

1. Über additive Eigenschaften von Zahlen, Math. Annalen 107(1933), 649-690.

SELBERG, A.

1. On an elementary method in the theory of primes, Norske Vid. Selsk. Forh. Trondheim 19, (18) (1947), 64-67. MR 9-271.
2. On elementary methods in prime number theory and their limitations, 11, Skand. Mat. Kongr., Trondheim 1949, 13-22. MR 14-726.
3. The general sieve method and its place in prime number theory, Proc. Inter. Congr., Cambridge, Mass. (1950), 286-292. MR 13-438.

SHAPIRO, H. N., and WARGA, T.

1. On the representation of large integers as sums of primes. I., Comm. Pure Appl. Math. 3 (1950), 153-176, MR 12-244.

TARTAKOVSKII, V. A.

1. La méthode du crible approximatif "électif", Dokl. Akad. Nauk SSSR, 23(1939), 127-130.
2. The sieve method in group theory, Mat. Sb. N. S. 25(67) (1949), 3-50 = Amer. Math. Soc. Translations, no. 60(1952). MR 11-493.
3. Application of the sieve method to the solution of the word problem for certain types of groups, Mat. Sb. N.S. 25(67) (1949), 251-274 = Amer. Math. Soc. Translations, no. 60 (1952). MR 13-528.

TCHUDAKOV, N. G.

1. The density of even numbers which cannot be represented as the sum of two primes, Izv. Akad. Nauk SSSR Ser. Mat. Nauk, no. 1 (1938), 25-40. (Russian).

TCHULANOVSKI, I. V.

1. Certain estimates connected with a new method of Selberg in elementary number theory, Dokl. Akad. Nauk SSSR N.S. 63(1948), 491-494. (Russian). MR 10-355.

TITCHMARSH, E. C.

1. The Theory of the Riemann Zeta-function, Oxford, 1951. MR 13-741.

TOGASHI, A., and UCHIYAMA, S.

1. On the representation of large even integers as sums of two almost primes. I., J. Fac. Sci. Hokkaido Univ., Ser. I, 18(1964), 60-68. MR 29, no. 3420.

TURÁN, P.

1. Certain function-theoretic sieve methods in the theory of numbers, Dokl. Akad. Nauk SSSR 171(1966), 1289-1292 = Soviet Math. Doklady 7(1966), 1661-1662. MR 34, no. 5781.

UCHIYAMA, M., and UCHIYAMA, S.

1. On the representation of large even integers as sums of a prime and an almost prime, Proc. Japan Akad. 40(1964), 150-154. MR 29, no. 2234.

UCHIYAMA, S.

1. A note on the sieve method of A. Selberg, J. Fac. Sci. Hokkaido Univ. Ser. I, 16(1962), 189-192. MR 27, no. 1432.

2. A further note on the sieve method of A. Selberg, J. Fac. Sci. Hokkaido Univ., Ser I, 17(1963), 79-83. MR 29, no. 1196.

3. On a theorem concerning the distribution of almost primes, J. Fac. Sci. Hokkaido Univ., Ser I, 17(1963), 152-159. MR 28, no. 2101.

4. On the distribution of almost primes in an arithmetic progression, J. Fac. Sci. Hokkaido Univ., Ser. I, 18(1964), 1-22. MR 30, no. 1108.

5. On the representation of large even integers as sums of two almost primes. II., J. Fac. Sci. Hokkaido Univ., Ser. I, 18(1964), 69-77. MR 29, no. 3421.

6. On the representation of large even integers as sums of a prime and an almost prime, II., Proc. Japan Acad. 43(1967), 567-571. MR 37, no.179.

USPENSKY, J. V., and HEASLET, M. A.

1. Elementary Number Theory, New York, 1939.

VINOGRADOV, A I.

1. On numbers with small prime divisors, Dokl. Akad. Nauk SSSR N.S. 109(1956), 683-686. (Russian). MR 19-16.

2. Application of $\zeta(s)$ to the sieve of Eratosthenes, Mat. Sb. N.S. 41(83) (1957), 49-80 (+ 415-416) = Amer. Math. Soc. Translations, ser. 2, 13(1960), 29-60. MR 20, no. 3836.

VINOGRADOV, A. I. - continued

3. The sieve method in algebraic fields. Lower bounds, Mat. Sb. N.S. 64(1964), 52-78. (Russian). MR 29, no. 1195.

4. The density conjecture for Dirichlet L-series, Izv. Akad. Nauk SSSR, Ser. Mat. 29(1965), 903-934 + 30(1966), 719-720. (Russian). MR 33, no. 2607 + 33, no. 5579.

VINOGRADOV, I. M.

1. The Method of Trigonometric Sums in the Theory of Numbers, New York, 1954. MR 15-941.

WANG, Y.

1. On the representation of a large even integer as a sum of a prime and a product of at most 4 primes, Acta Math. Sinica 6(1956), 565-582. (Chinese). MR 20,no.4530.

2. On sieve methods and some of their applications, Science Record N.S. 1 (3)(1957), 1-5. MR 19-533.

3. On sieve methods and some of the related problems, Sci. Record N.S. 1 (1957), 9-12. MR 23, no. A2400.

4. On the representation of a large even number as a sum of two almost primes, Sci. Record N.S. 1(1957), 291-295. (Chinese). MR 23, no. A874.

5. On some properties of integral valued polynomials, Advancement in Math. 3(1957), 416-423. (Chinese). MR 20, no. 4531.

6. On sieve methods and some of their applications, Acta Math. Sinica 8(1958), 413-429 = Sientia Sinica 8(1959), 357-381. MR 21, no. 4944 + 21, no. 1958.

7. A note on the least primitive root of a prime, Sci. Record N.S. 3(1959), 174-179. (Chinese). MR 24, no. A1894.

8. On sieve methods and some of their applications, Acta Math. Sinica 9(1959), 87-100 = Scientia Sinica 11(1962), 1607-1624. MR 21, no. 5615 and 26, no. 3685.

9. On the representation of a large integer as a sum of a prime and an almost prime, Acta Math. Sinica 10(1960), 168-181 = Chinese Math.-Acta 1(1962), 181-195. MR 27, no. 1424 and 5733.

WARLIMONT, R.

1. On squarefree numbers in arithmetic progressions, Monatsh. Math. 73(1969), 433-448. MR 41, no. 3430.

2. On divisor problems in connection with squarefree numbers, Monatsh. Math. 74(1970), 154-165. MR 41, no. 3431.

WILSON, R. J.

1. The large sieve in algebraic number fields, Mathematika 16(1969), 189-204.

WIRSING, E.

1. Über die Zahlen, deren Primteiler einer gegebener Menge angehören, Archiv. Math. 7(1956), 263-272. MR 18-642.

2. Das asymptotische Verhalten von Summen über multiplikative Funktionen, Math. Ann. 143(1961), 75-102. MR 24, no. A1241.

3. Elementare Beweise des Primzahlsatzes mit Restglied. I., J. Reine Angew. Math. 211(1962), 205-214. MR 27, no. 119.

4. Elementare Beweise des Primzahlsatzes mit Restglied. II., J. Reine Angew. Math. 214/215(1964), 1-18. MR 29, no. 3457.

5. Das asymptotische Verhalten von Summen über multiplikative Funktionen. II., Acta Math. Hung. 18(1967), 411-467.

WOLKE, D.

1. A note on the least prime quadratic residue (mod p), Acta Arith. 16(1969/70), 85-87. MR 39, no. 6842.

2. Farey-Brüche mit primem Nenner und das grosse Sieb, Math. Z. 114(1970), 145-158. MR 41, no. 5328.