# The Distribution of Prime Numbers

Angel Kumchev

Spring 2005

# Preface

These notes cover the material for the graduate course with the same title that I taught at the University of Texas at Austin during the Spring 2005 semester. They draw heavily on *The Distribution of Prime Numbers* by M. Huxley and *Multiplicative Number Theory* by H. Davenport (as revised by H. L. Montgomery). I also acknowledge the use of notes by Jeff Vaaler and numerous discussions with him that helped improve the exposition.

# Notation

For functions $f$ and $g$ with $g \geq 0$, we write $f(x) = O(g(x))$ or $f(x) \ll g(x)$ when there is a constant $c$ such that $|f(x)| \leq cg(x)$; when $f$ and $g$ are both non-negative, we may also write $f \gg g$ instead of $g \ll f$. We write $f(x) \sim g(x)$ when $\lim f(x)/g(x) = 1$ as $x$ tends to some limit to be specified at each occurrence. We use $c, c_0, c_1, c_2, \ldots$ to denote implicit constants; these and the constants implied by $O$- and $\ll$-symbols are presumed absolute, unless stated otherwise. Throughout these notes (and much of number theory outside them) the letter $p$, with or without subscripts or superscripts, is reserved for prime numbers. We also use $s = \sigma + \mathrm{i}t$ to denote a complex variable.

For the most part, we use the standard notations for common number-theoretic functions. These are usually defined at their first appearance, but for convenience we also list them here:

| | |
|---|---|
| $\|\|\theta\|\|$ | the distance from the real number $\theta$ to the nearest integer; |
| $[\theta]$ | the integral part of the real number $\theta$; |
| $\{\theta\}$ | the fractional part of the real number $\theta$; |
| $e(z)$ | $e^{2\pi \mathrm{i} z}$; |
| $\operatorname{Log} z$ | the principal branch of the complex logarithm ($\operatorname{Log} x = \ln x$ when $x > 0$); |
| $d(n)$ | the number of positive divisors of $n$; |
| $\phi(n)$ | Euler's totient function: the number of reduced residue classes modulo $n$; |
| $\mu(n)$ | the Möbius function (see (1.1)); |
| $\Lambda(n)$ | von Mangoldt's function (see (1.3)); |
| $\pi(x)$ | the number of primes $p \leq x$; |
| $\pi(x; q, a)$ | the number of primes $p \leq x$, with $p \equiv a \pmod q$; |
| $\theta(x)$ | Chebyshev's function (see (1.4)); |
| $\psi(x)$ | the sum of the values of $\Lambda(n)$ over $n \leq x$; |
| $\psi(x; q, a)$ | the sum of the values of $\Lambda(n)$ over $n \leq x$, with $n \equiv a \pmod q$; |
| $\psi(x, \chi)$ | the sum of the values of $\Lambda(n)\chi(n)$ over $n \leq x$ (see (3.50)); |
| $\tau(\chi, a)$ | the Gaussian sum (see (3.7)); |
| $N(\sigma, T)$ | the number of zeros $\rho = \beta + \mathrm{i}\gamma$ of $\zeta(s)$ with $\sigma \leq \beta \leq 1$ and $\|\gamma\| \leq T$; |

# Contents

# Chapter 0

# Historical background

## 0.1 Early history

The first result on the distribution of primes is Euclid's theorem (*circa* 300 B.C.) on the infinitude of the primes. In 1737 Euler went a step further and proved that, in fact, the series of the reciprocals of the primes diverges. In the opposite direction, Euler observed that the rate of divergence of this series is much slower than the rate of divergence of the harmonic series:

> "*The sum of the series of the reciprocals of the prime numbers,*
>
> $$\frac{1}{2}+\frac{1}{3}+\frac{1}{5}+\frac{1}{7}+\frac{1}{11}+\frac{1}{13}+\cdots,$$
>
> *is infinitely large, but it is infinitely many times less than the sum of the harmonic series,*
>
> $$1+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\frac{1}{5}+\cdots.$$
>
> *Furthemore, the sum of the former series is like the logarithm of the sum of the latter series.*"

This statement appears to be the earliest attempt to quantify the frequency of the primes among the positive integers.

Consider the prime counting function

$$\pi(x)=\sum_{p\le x}1.$$

In 1798 Legendre conjectured that $\pi(x)$ satisfies the asymptotic relation

$$\lim_{x\to\infty}\frac{\pi(x)}{x/(\log x)}=1; \tag{0.1}$$

this is the *prime number theorem* (PNT). Years later, Gauss wrote that back in his adolescent years he had observed that the logarithmic integral

$$\operatorname{Li} x=\int_2^x\frac{dt}{\log t}$$

seemed to provide a very good approximation to $\pi(x)$. This, of course, is consistent with (0.1), as can be seen from the formula

$$\operatorname{Li} x = \frac{x}{\log x} + \frac{1!x}{(\log x)^2} + \cdots + \frac{k!x}{(\log x)^{k+1}} + O\left(\frac{x}{(\log x)^{k+2}}\right). \tag{0.2}$$

The first theoretical evidence in support of the PNT was given by Chebyshev in the 1850s. He proved that:

- (0.1) predicts correctly the order of magnitude of $\pi(x)$, that is, there exist absolute constants $c_2 > c_1 > 0$ such that
$$\frac{c_1 x}{\log x} \le \pi(x) \le \frac{c_2 x}{\log x}. \tag{0.3}$$
Chebyshev showed that for sufficiently large $x$ one may take $c_1 = 0.9212$ and $c_2 = 1.1056$. In his honor, bounds for $\pi(x)$ of this type are now known as *Chebyshev's estimates*.

- If the limit on the left side of (0.1) exists, then it must be equal to 1.

Chebyshev used the methods that he developed for the proof of (0.3) to establish *Bertrand's postulate:* the interval $(n, 2n]$ contains a prime number for all integers $n \ge 1$. Furthermore, in 1874 Mertens used Chebyshev's estimates (0.3) to show that

$$\sum_{p \le x} \frac{1}{p} = \log\log x + B + O\big((\log x)^{-1}\big), \tag{0.4}$$

$B$ being an absolute constant. This provided the first rigorous proof of Euler's observation that "the sum of the [series of the reciprocals of the primes] is like the logarithm of the sum of the [harmonic series]." We sketch the proofs of (0.3), (0.4), and some related results in §1.2.

## 0.2 The Riemann $\zeta$-function and the prime number theorem

The *Riemann zeta-function* is defined in the half-plane $\operatorname{Re}(s) > 1$ as

$$\zeta(s) = \sum_{n=1}^{\infty} n^{-s} = \prod_{p} \left(1 - p^{-s}\right)^{-1}. \tag{0.5}$$

The identity between the infinite series and the infinite product on the right (which runs over all primes) is an analytic expression of the fundamental theorem of arithmetic and was discovered by Euler in 1737 (in the same paper as his proof of the infinitude of the primes), at least in the case when $s$ is real. The first to consider $\zeta(s)$ as a function of a complex variable was Riemann. In 1859 he published his seminal paper [47] (his only paper on number theory), in which he observed that $\zeta(s)$ is holomorphic in the half-plane $\operatorname{Re}(s) > 1$ and that it can be continued analytically to a meromorphic function, whose only singularity is a simple pole at $s = 1$. Riemann was interested in $\zeta(s)$, because Euler's identity (0.5) provides a connection between the analytic properties of $\zeta(s)$

and the PNT. It is not difficult to deduce from (0.5) that $\zeta(s)$ does not vanish in the half-plane $\mathrm{Re}(s) > 1$. Riemann proved that $\zeta(s)$ satisfies the *functional equation*

$$\pi^{-s/2}\Gamma\left(\frac{s}{2}\right)\zeta(s) = \pi^{(s-1)/2}\Gamma\left(\frac{1-s}{2}\right)\zeta(1-s),$$

from which it is easy to deduce that the only zeros of $\zeta(s)$ in the half-plane $\mathrm{Re}(s) < 0$ are the negative even integers; these are the *trivial zeros* of $\zeta(s)$. Besides the trivial zeros, $\zeta(s)$ has infinitely many zeros in the strip $0 \le \mathrm{Re}(s) \le 1$: the *non-trivial zeros* of $\zeta(s)$. Riemann proposed several conjectures about the non-trivial zeros of $\zeta(s)$:

C1. If

$$N(T) = \#\big\{\rho \in \mathbb{C} : \zeta(\rho) = 0,\ 0 \le \mathrm{Re}(\rho) \le 1,\ 0 < \mathrm{Im}(\rho) \le T\big\},$$

then

$$N(T) = \frac{T}{2\pi}\log\left(\frac{T}{2\pi e}\right) + O(\log T).$$

C2. The entire function

$$\xi(s) = \frac{1}{2}s(s-1)\pi^{-s/2}\Gamma\left(\frac{s}{2}\right)\zeta(s)$$

has a product representation

$$\xi(s) = e^{A+Bs}\prod_{\rho}\left(1-\frac{s}{\rho}\right)e^{-s/\rho},$$

the product being over all non-trivial zeros of $\zeta(s)$.

C3. If $x > 1$, there is an explicit formula that represents $\pi(x)$ as a series over the non-trivial zeros of $\zeta(s)$.

C4. **Riemann Hypothesis (RH).** All zeros of $\zeta(s)$ with $0 \le \mathrm{Re}(s) \le 1$ lie on the line $\mathrm{Re}(s) = \frac{1}{2}$.

By the end of the 19th century, conjectures C1–C3 were proved: C1 and C3 were established by von Mangoldt, and C2 is a consequence of the general theory of entire functions of finite order developed by Hadamard. In particular, the Riemann–Mangoldt explicit formula for $\pi(x)$ demonstrated that the PNT follows from the nonvanishing of $\zeta(s)$ on the line $\mathrm{Re}(s) = 1$. Thus, when in 1896 Hadamard and de la Vallée Poussin proved (independently) that $\zeta(1+it) \neq 0$ for all real $t$, the PNT was finally proved. In contrast, the Riemann Hypothesis is still an open problem that has been selected by the Clay Mathematics Institute as one of the seven Millennium Problems. We remark that under RH, the Riemann–Mangoldt formula implies the asymptotic formula

$$\pi(x) = \mathrm{Li}\, x + O\big(x^{1/2}\log x\big), \tag{0.6}$$

which is essentially best possible.

The last observation has motivated the investigations of the error term in the PNT. In 1899 de la Vallée Poussin refined the original proof that $\zeta(1+it) \neq 0$ and showed that, in fact, $\zeta(\sigma+it)$ does not vanish in the region

$$\sigma \geq 1 - \frac{c}{\log(|t|+10)}, \tag{0.7}$$

for some absolute constant $c > 0$. This suffices to establish the following quantitative version of the PNT, which will be the main subject of Chapter 2 of these notes.

**Theorem 1.** *There exists an absolute constant $c > 0$ such that*

$$\pi(x) = \operatorname{Li} x + O\big(x \exp\big(-c\sqrt{\log x}\big)\big).$$

Further improvements on the error term in the PNT have been quite limited. In 1922 Littlewood proved that

$$\pi(x) - \operatorname{Li} x \ll x \exp\big(-c\sqrt{\log x \log\log x}\big), \tag{0.8}$$

while the best result to date was obtained by Korobov [38] and I. M. Vinogradov [58] in 1958:

$$\pi(x) - \operatorname{Li} x \ll x \exp\big(-c(\log x)^{3/5}(\log\log x)^{-1/5}\big). \tag{0.9}$$

Both (0.8) and (0.9) are consequences of repsective improvements on the estimate of the zerofree region (0.7). Unfortunately, it is known that the approach employed in these works can never yield a bound of the form $\pi(x) - \operatorname{Li} x \ll x^{\theta}$, with a fixed $\theta < 1$.

## 0.3 Primes in arithmetic progressions

In a couple of memoirs published in 1837 and 1840, Dirichlet proved that if $a$ and $q$ are natural numbers with $(a,q) = 1$, then the arithmetic progression $a, a+q, a+2q, \ldots$ contains infinitely many primes. By refining Dirichlet's argument, Mertens established the asymptotic formula

$$\sum_{\substack{p \leq x \\ p \equiv a \pmod q}} \frac{1}{p} \sim \frac{1}{\phi(q)} \sum_{p \leq x} \frac{1}{p} \qquad \text{as } x \to \infty, \tag{0.10}$$

where $\phi(q)$ is Euler's totient function. Fix $q$ and consider the various reduced residue classes modulo $q$. Since all but finitely many primes lie in residue classes $a \bmod q$ with $(a,q) = 1$, (0.10) suggests that the primes are uniformly distributed among the reduced residue classes to a given modulus $q$. Thus, one may expect that if $(a,q) = 1$, then

$$\pi(x; q, a) = \sum_{\substack{p \leq x \\ p \equiv a \pmod q}} 1 \sim \frac{\operatorname{Li} x}{\phi(q)} \qquad \text{as } x \to \infty. \tag{0.11}$$

This is the *prime number theorem for arithmetic progressions*. We may approach this statement in two different ways. First, we may fix $a$ and $q$ and ask whether (0.11) holds (allowing the

convergence to depend on $q$ and $a$). Posed in this form, the problem is a minor generalization of the PNT. In fact, shortly after proving Theorem 1, de la Vallée Poussin showed that

$$\pi(x;q,a) = \frac{\operatorname{Li} x}{\phi(q)} + O\big(x\exp\big(-c\sqrt{\log x}\big)\big),$$

where $c = c(q,a) > 0$ and the $O$-implied constant depends on $q$ and $a$. The problem becomes much more difficult if we want an estimate that is explicit in $q$ and uniform in $a$. The first result of this kind was obtained by Page [43], who proved that there exists a (small) positive number $\delta$ such that

$$\pi(x;q,a) = \frac{\operatorname{Li} x}{\phi(q)} + O\big(x\exp\big(-(\log x)^{\delta}\big)\big),$$

whenever $1 \le q \le (\log x)^{2-\delta}$ and $(a,q) = 1$. In 1935 Siegel [49] proved the following result, which we will establish in Chapter 3.

**Theorem 2.** *For any fixed $A > 0$, there exists a constant $c = c(A) > 0$ such that*

$$\pi(x;q,a) = \frac{\operatorname{Li} x}{\phi(q)} + O\big(x\exp\big(-c\sqrt{\log x}\big)\big),$$

*whenever $1 \le q \le (\log x)^A$ and $(a,q) = 1$.*

**Remark.** While this result is clearly sharper than Page's, it does have one significant drawback: it is ineffective, that is, given a particular value of $A$, the proof does not allow the constant $c(A)$ or the $O$-implied constant to be computed.

The proofs of the above results rely on the analytic properties of a class of generalizations of the Riemann zeta-function known as *Dirichlet L-functions*. For each positive integer $q$ there are $\phi(q)$ functions $\chi : \mathbb{Z} \to \mathbb{C}$ called *Dirichlet characters modulo q* (we will define these in Chapter 3). Given a character $\chi$ modulo $q$, we define the respective Dirichlet $L$-function by

$$L(s,\chi) = \sum_{n=1}^{\infty} \chi(n) n^{-s} \qquad (\operatorname{Re}(s) > 1).$$

Similarly to $\zeta(s)$, $L(s,\chi)$ is holomorphic in the half-plane $\operatorname{Re}(s) > 1$ and can be continued analytically to a meromorphic function that has at most one pole, which (if present) must be a simple pole at $s = 1$. Furthermore, just as $\zeta(s)$, the continued $L(s,\chi)$ has infinitely many zeros in the strip $0 \le \operatorname{Re}(s) \le 1$, and the horizontal distribution of those zeros has important implications on the distribution of primes in arithmetic progressions. For example, the results of de la Vallée Poussin, Page, and Siegel mentioned above were proved by showing that no $L$-function has zeros "close" to the line $\operatorname{Re}(s) = 1$. We also have the following generalization of the Riemann Hypothesis:

> **Generalized Riemann Hypothesis (GRH).** Let $L(s,\chi)$ be a Dirichlet $L$-function. Then all zeros of $L(s,\chi)$ with $0 \le \operatorname{Re}(s) \le 1$ lie on the line $\operatorname{Re}(s) = \frac{1}{2}$.

Assuming GRH, we can deduce easily that

$$\pi(x;q,a) = \frac{\operatorname{Li} x}{\phi(q)} + O\big(x^{1/2}\log x\big), \tag{0.12}$$

whenever $(a,q)=1$. This estimate establishes (0.11) for $1 \le q \le x^{\theta}$, $\theta < \frac{1}{2}$.

In many applications one only needs approximations like (0.12) in some average sense over the moduli $q$. During the 1950s and 1960s several authors obtained such results. In particular, the following quantity was studied extensively:

$$E(x,Q) = \sum_{q\le Q} \max_{(a,q)=1} \max_{y\le x} \left|\pi(y;q,a) - \frac{\operatorname{Li} y}{\phi(q)}\right|.$$

The trivial bound for this sum is $E(x,Q) \ll x$, whereas (0.12) implies

$$E(x,Q) \ll Qx^{1/2}\log x. \tag{0.13}$$

In 1965 Bombieri [5] and A. I. Vinogradov [55] proved (independently) the following result.

**Theorem 3.** *For any fixed $A > 0$, there exists a constant $B = B(A) > 0$ such that*

$$E(x,Q) \ll x(\log x)^{-A},$$

*whenever $Q \le x^{1/2}(\log x)^{-B}$.*

We observe that this result provides a nontrivial estimate for $E(x,Q)$ under essentially the same restrictions on $Q$ as GRH. Because of this fact, the Bombieri–Vinogradov theorem has seen numerous applications in which it has been used as a *de facto* replacement for GRH. In Chapter 5 we will give a proof of Theorem 3 with $B = A + 4$.

It should be noted that unlike the error term in (0.6), the error term in (0.12) is not necessarily best possible. In fact, there is some evidence in support of the bold conjecture that

$$\pi(x;q,a) = \frac{\operatorname{Li} x}{\phi(q)} + O_{\epsilon}\big((x/q)^{1/2+\epsilon}\big)$$

for any fixed $\epsilon > 0$. In Chapter 5 we will establish the so-called *Barban–Davenport–Halberstam theorem*, which asserts that this bound holds in the mean-square over all arithmetic progressions with differences $q \le x^{1-\epsilon}$. We should also mention that during the mid 1980s Bombieri, Friedlander, and Iwaniec [6, 7, 8] obtained several variants of the Bombieri–Vinogradov theorem, in which the value of $Q$ can exceed $x^{1/2}$. However, since their methods go beyond the reach of these notes, we will only state one of their results (see [7]).

**Theorem 4.** *Let $a \ne 0$ and $x \ge y \ge 3$. Then*

$$\sum_{\substack{q\le \sqrt{xy}\\(q,a)=1}} \left|\pi(x;q,a) - \frac{\operatorname{Li} x}{\phi(q)}\right| \ll (\operatorname{Li} x)\left(\frac{\log y}{\log x}\right)^2 (\log\log x)^c.$$

*Here $c$ is an absolute constant and the $\ll$-implied constant depends only on $a$.*

## 0.4 Primes in short intervals

It is an old problem in the theory of prime numbers to prove that for any integer $n \geq 1$, the interval $(n^2, (n+1)^2]$ contains a prime number. This problem leads quickly to the more general question of estimating the differences between consecutive primes. Cramér was the first to study this question systematically. Let $p_n$ denote the $n$th prime number. In 1920 Cramér [12] proved that under RH

$$p_{n+1} - p_n \ll p_n^{1/2} \log p_n.$$

Cramér also proposed a probabilistic model of the prime numbers that leads to very precise (and very bold) predictions of the asymptotic properties of the primes. In particular, he conjectured [13] that

$$\limsup_{n\to\infty} \frac{p_{n+1} - p_n}{(\log p_n)^2} = 1. \tag{0.14}$$

Nontrivial upper bounds for $p_{n+1} - p_n$ can be derived from the quantitative versions of the PNT stated above, but the ensuing results are rather poor, because the known bounds for the error term in the PNT are barely smaller than the main term. However, in 1930 Hoheisel [26] obtained a much sharper result. He proved (unconditionaly) the asymptotic formula

$$\pi(x + x^\theta) - \pi(x) \sim x^\theta (\log x)^{-1} \qquad \text{as } x \to \infty, \tag{0.15}$$

with $\theta = 1 - (33000)^{-1}$. Subsequently several authors made further contributions that produced the following improvements on Hoheisel's result:

| | |
|---|---|
| Heilbronn [25] (1933) | $\theta = 0.996$ |
| Chudakov [11] (1936) | $\theta > 3/4 \quad = 0.750$ |
| Ingham [30] (1937) | $\theta > 5/8 \quad = 0.625$ |
| Montgomery [41] (1971) | $\theta > 3/5 \quad = 0.600$ |
| Huxley [27] (1972) | $\theta > 7/12 = 0.583\ldots$ |
| Heath-Brown [23] (1988) | $\theta = 7/12 = 0.583\ldots$ |

We will see the proof of Huxley's result in Chapter 5 of these notes. Furthermore, since the late 1970s, several mathematicians have shown that even shorter intervals must contain primes (without establishing an asymptotic formula for the number of primes in such intervals). Such results usually take the form

$$\pi(x + x^\theta) - \pi(x) \gg x^\theta (\log x)^{-1} \qquad \text{for } x \geq x_0(\theta). \tag{0.16}$$

The following list displays the progress in that direction over the last 30 years:

| | |
|---|---|
| Iwaniec and Jutila [34] (1979) | $\theta = 13/23 = 0.565\ldots$ |
| Heath-Brown and Iwaniec [24] (1979) | $\theta > 11/20 = 0.550$ |
| Iwaniec and Pintz [35] (1984) | $\theta = 23/42 = 0.547\ldots$ |
| Lou and Yao [40] (1992) | $\theta = 6/11 \quad = 0.545\ldots$ |
| Baker and Harman [1] (1996) | $\theta = 0.535$ |
| Baker, Harman, and Pintz [2] (2001) | $\theta = 0.525$ |

Selberg [48] considered the distribution of primes in "almost all short intervals." Let $h(x)$ be an increasing function of $x$. We say that *almost all* intervals $(x, x+h(x)]$ contain primes if the measure of the set of $x \in (1, X]$ for which the interval $(x, x+h(x)]$ contains no prime is $o(X)$. Selberg proved that if $h(x)$ grows faster than $(\log x)^2$ as $x \to \infty$, the Riemann Hypothesis implies that almost all intervals $(x, x+h(x)]$ contain a prime number (and also that the asymptotic formula (0.15) holds for each inexceptional interval). Further, Selberg showed unconditionally that if $\theta > 1/4$, then almost all intervals $(x, x+x^\theta]$ contain a prime number. The latter result has been the subject of a long series of successive improvements, similar to the improvements on Hoheisel's result described above. In particular, the best result to date obtained in 1996 by Jia [36] extends the range for $\theta$ in Selberg's result to $\theta > 1/20$.

In the opposite direction, Erdös [16] showed in 1935 that

$$p_{n+1} - p_n \geq c \log p_n \log\log p_n (\log\log\log p_n)^{-2} \tag{0.17}$$

infinitely often. In 1938 Rankin [46] showed that one can replace the right side of (0.17) by

$$(1/3 + o(1)) \log p_n \log\log p_n \log\log\log\log p_n (\log\log\log p_n)^{-2},$$

but subsequent attempts at further improvements have not been very successful: the best result to date (see Pintz [44]) replaces the constant $1/3$ in Rankin's bound by $2e^\gamma$, where $\gamma$ is Euler's constant. In fact, the problem appears to be so notoriously difficult that Erdös—who was known for offerring monetary prizes for solutions of problems he was intrigued by—announced that he would pay $10,000 to anyone who proved that the constant $1/3$ in Rankin's result can be taken arbitrarily large!

# Chapter 1

# Introduction: basic estimates

The purpose of this chapter is to introduce some of the basic techniques and functions appearing in the later chapters. The results are mostly elementary and the reader may be familiar with some (and possibly all) of them.

## 1.1 Multiplicative functions

A function $f : \mathbb{N} \to \mathbb{C}$ is said to be *multiplicative* if it is not identically zero and

$$f(mn) = f(m)f(n) \qquad \text{whenever } \gcd(m, n) = 1.$$

If $f$ satisfies the stronger condition that $f(mn) = f(m)f(n)$ for all pairs $m, n$, it is said to be *completely (or totally) multiplicative*.

Some functions, such as $f(n) = n^s$ ($s \in \mathbb{C}$), are obviously multiplicative. Others are defined so that they are. One such function is the *Möbius function*

$$\mu(n) = \begin{cases} 1 & \text{if } n = 1, \\ (-1)^r & \text{if } n \text{ is the product of } r \text{ distinct primes,} \\ 0 & \text{if } n \text{ is divisible by the square of a prime.} \end{cases} \tag{1.1}$$

The following lemma provides an easy way to deduce the multiplicativity of a large class of arithmetic functions. We leave its proof as an exercise.

**Lemma 1.1.** *Suppose that $f$ and $g$ are multiplicative functions. Then the arithmetic function $f * g$, defined by*

$$(f * g)(n) = \sum_{d|n} f(d)g(n/d),$$

*is also multiplicative.*

The next lemma contains the most important property of the Möbius function.

**Lemma 1.2.** $\displaystyle\sum_{d|n}\mu(d) = \begin{cases} 1 & \text{if } n = 1, \\ 0 & \text{otherwise.} \end{cases}$

*Proof.* It suffices to consider the case when $n$ is squarefree. Suppose that $n = p_1 p_2 \cdots p_r$, where $p_1, p_2, \ldots, p_r$ are distinct primes, and write $m = p_1 p_2 \cdots p_{r-1}$. Then

$$\sum_{d|n}\mu(d) = \sum_{d|m}\mu(d) + \sum_{d|m}\mu(dp_r) = \sum_{d|m}\mu(d) + \sum_{d|m}(-\mu(d)) = 0,$$

where the second to last step uses the multiplicativity of $\mu$. ∎

**Corollary 1.3.** *(Möbius inversion formula) Suppose that $f : \mathbb{N} \to \mathbb{C}$ is an arithmetic function and define*

$$F(n) = \sum_{d|n} f(d).$$

*Then $f = F * \mu$, that is,*

$$f(n) = \sum_{d|n} F(d)\mu(n/d).$$

*In particular, if $F$ is multiplicative, so is $f$.*

*Proof.* We have

$$\sum_{d|n} F(d)\mu(n/d) = \sum_{d|n}\sum_{k|d} f(k)\mu(n/d) = \sum_{k|n}\sum_{\substack{d|n \\ k|d}} f(k)\mu(n/d) = \sum_{k|n}\sum_{m|(n/k)} f(k)\mu(n/km).$$

By Lemma 1.2, the sum over $m$ vanishes unless $k = n$, so the first claim follows. The second claim is a consequence of the first, Lemma 1.1, and the multiplicativity of $\mu$. ∎

## 1.2 Partial summation

We now discuss a simple trick that is put to a great use in analytic number theory.

**Lemma 1.4 (Abel).** *Suppose that $a_n$ are complex numbers and $f(x)$ is continuously differentiable on $[\alpha, \beta]$. Then*

$$\sum_{\alpha < n \le \beta} a_n f(n) = A(\beta) f(\beta) - \int_\alpha^\beta A(x) f'(x)\, dx,$$

*where $A(x) = \sum_{\alpha < n \le x} a_n$.*

*Proof.* Using Stieltjes integration by parts, we have

$$\sum_{\alpha < n \le \beta} a_n f(n) = \int_{\alpha^+}^{\beta^+} f(x)\, dA(x) = f(x)A(x)\Big|_\alpha^\beta - \int_\alpha^\beta A(x)\, df(x),$$

and the desired result follows. ∎

**Corollary 1.5.** *There is a constant $c_1$ such that*

$$\sum_{n\le x}\frac{1}{n}=\log x+c_1+O\left(x^{-1}\right).$$

**Corollary 1.6.** $\displaystyle\sum_{n\le x}\log n = x\log x - x + O(\log x)$.

**Remark.** The constant $c_1$ is known as *Euler's constant* and usually is denoted by $\gamma$:

$$\gamma=\lim_{x\to\infty}\left(\sum_{n\le x}\frac{1}{n}-\log x\right)\approx 0.5772\ldots. \tag{1.2}$$

Next, we define three arithmetic functions that play an important role in prime number theory. These are *von Mangoldt's function*

$$\Lambda(n)=\begin{cases}\log p & \text{if } n \text{ is a power of a prime } p,\\ 0 & \text{otherwise,}\end{cases} \tag{1.3}$$

and the functions

$$\theta(x)=\sum_{p\le x}\log p \qquad \text{and} \qquad \psi(x)=\sum_{n\le x}\Lambda(n), \tag{1.4}$$

which were introduced first by Chebyshev. Our first result about these functions is a Chebyshev-type bound for $\psi(x)$.

**Theorem 1.7.** *Suppose that* $0<c_2<\log 2$ *and* $c_3>\log 4$. *Then for sufficiently large* $x$,

$$c_2x\le\psi(x)\le c_3x.$$

*Proof.* Define

$$T(x)=\sum_{n\le x}\log n.$$

Taking logarithms in the prime factorization of $n$, we see that

$$\log n=\sum_{d\mid n}\Lambda(d),$$

so we can rewrite $T(x)$ as

$$T(x)=\sum_{n\le x}\sum_{d\mid n}\Lambda(d)=\sum_{d\le x}\Lambda(d)\sum_{\substack{n\le x\\ d\mid n}}1=\sum_{d\le x}\Lambda(d)\left[\frac{x}{d}\right].$$

Thus,

$$T(x)-2T(x/2)=\sum_{d\le x}\Lambda(d)\left(\left[\frac{x}{d}\right]-2\left[\frac{x}{2d}\right]\right).$$

We now note that

$$0 \le \left[\frac{x}{d}\right] - 2\left[\frac{x}{2d}\right] \le 1$$

and

$$\left[\frac{x}{d}\right] - 2\left[\frac{x}{2d}\right] = 1 \qquad \text{for } x/2 < d \le x.$$

Hence,

$$\psi(x) - \psi(x/2) \le T(x) - 2T(x/2) \le \psi(x).$$

On the other hand, by Corollary 1.6,

$$T(x) - 2T(x/2) = x\log 2 + O(\log x).$$

We deduce that

$$\psi(x) \ge x\log 2 + O(\log x)$$

and

$$\begin{aligned}
\psi(x) &\le \psi(x/2) + x\log 2 + O(\log x)\\
&\le \psi(x/4) + \left(1+\tfrac{1}{2}\right)x\log 2 + O(\log x)\\
&\le \psi(x/8) + \left(1+\tfrac{1}{2}+\tfrac{1}{4}\right)x\log 2 + O(\log x)\\
&\vdots\\
&\le \psi(x/2^r) + \left(1+\tfrac{1}{2}+\tfrac{1}{4}+\cdots\right)x\log 2 + O(r\log x)\\
&\le x\log 4 + O\big((\log x)^2\big),
\end{aligned}$$

on choosing $r$ so that $2^r \le x < 2^{r+1}$. ∎

**Lemma 1.8.** $\displaystyle\sum_{p\le x}\frac{\log p}{p} = \log x + O(1).$

**Theorem 1.9 (Mertens).** *There is an absolute constant $B$ such that*

$$\sum_{p\le x}\frac{1}{p} = \log\log x + B + O\big((\log x)^{-1}\big). \tag{1.5}$$

*Proof.* Define the function

$$R(x) = \sum_{2<p\le x}\frac{\log p}{p} - \log x$$

and the sequence

$$a_n = \begin{cases} (\log n)/n & \text{if } n \text{ is a prime number,}\\ 0 & \text{otherwise.}\end{cases}$$

Then, by Lemma 1.4,

$$\begin{aligned}\sum_{p\le x}\frac{1}{p} &= \sum_{2<n\le x}\frac{a_n}{\log n}+\frac{1}{2}\\ &= \frac{1}{\log x}\sum_{2<n\le x}\frac{\log p}{p}+\int_2^x\Bigg(\sum_{2<n\le y}\frac{\log p}{p}\Bigg)\frac{dy}{y(\log y)^2}+\frac{1}{2}\\ &= \frac{1}{\log x}\big(\log x+R(x)\big)+\int_2^x\frac{\log y+R(y)}{y(\log y)^2}\,dy+\frac{1}{2}\\ &= \log\log x+\frac{3}{2}-\log\log 2+\int_2^x\frac{R(y)}{y(\log y)^2}\,dy+\frac{R(x)}{\log x}.\end{aligned}$$

Using Lemma 1.8 to bound $R(y)$, we obtain

$$\int_2^x\frac{R(y)}{y(\log y)^2}\,dy+\frac{R(x)}{\log x}=\int_2^\infty\frac{R(y)}{y(\log y)^2}\,dy+O\big((\log x)^{-1}\big),$$

and the desired conclusion follows with

$$B=\frac{3}{2}-\log\log 2+\int_2^\infty\frac{R(y)}{y(\log y)^2}\,dy.$$

■

The final result of this section quantifies the relation between the error term in the PNT and the difference $\psi(x)-x$.

**Theorem 1.10.** *Suppose that $f$ is an integrable function such that $x^{1/2}\ll f(x)\ll x$ and*

$$\int_2^x\frac{f(t)}{t}\,dt\ll f(x)\log x.$$

*Then*

$$\psi(x)-x\ll f(x)\qquad\Leftrightarrow\qquad\pi(x)-\operatorname{Li}x\ll f(x)(\log x)^{-1}.$$

*Proof.* Theorem 1.7 implies that

$$\theta(x)=\psi(x)+O\big(x^{1/2}\big),$$

whence

$$\theta(x)=x+O(f(x)).$$

Let $a_n$ be the sequence

$$a_n=\begin{cases}\log n & \text{if } n \text{ is a prime number,}\\ 0 & \text{otherwise.}\end{cases}$$

As in the proof of Theorem 1.9,

$$\begin{aligned}\pi(x) &= \sum_{2<n\le x} \frac{a_n}{\log n} + 1 \\ &= \frac{\theta(x)}{\log x} + \int_2^x \frac{\theta(y)\,dy}{y(\log y)^2} + O(1) \\ &= \frac{x}{\log x} + \int_2^x \frac{dy}{(\log y)^2} + O\Big(\frac{f(x)}{\log x}\Big) + O\Big(\int_2^x \frac{f(y)\,dy}{y(\log y)^2}\Big) \\ &= \operatorname{Li} x + O\big(f(x)(\log x)^{-1}\big),\end{aligned}$$

by (0.2) and the properties of $f$. This proves the direct implication, the converse is left as an exercise. ∎

## 1.3 Dirichlet series

A *Dirichlet series* is an infinite series of the form

$$\sum_{n=1}^{\infty} a_n n^{-s}, \tag{1.6}$$

where $a_n$ are complex numbers and $s = \sigma + \mathrm{i}t$ is a complex variable. The following lemma shows that if a Dirichlet series converges at any finite complex number $s_0 = \sigma_0 + \mathrm{i}t_0$, then it converges to a holomorphic function in the half-plane $\operatorname{Re}(s) > \sigma_0$.

**Lemma 1.11.** *Suppose that $s_0 = \sigma_0 + \mathrm{i}t_0$ and the series*

$$\sum_{n=1}^{\infty} a_n n^{-s_0}$$

*converges. Then the Dirichlet series* (1.6) *converges uniformly on the compact subsets of the half-plane $\operatorname{Re}(s) > \sigma_0$ and the sum-function*

$$f(s) = \sum_{n=1}^{\infty} a_n n^{-s}$$

*is holomorphic in that half-plane.*

*Proof.* It suffices to show that (1.6) converges uniformly in the regions

$$\big\{ s \in \mathbb{C} \ : \ \operatorname{Re}(s) \ge \sigma_0 + \delta,\ |\operatorname{Im}(s) - t_0| \le T \big\}$$

where $\delta, T > 0$. By Lemma 1.4 with $a_n = a_n n^{-s_0}$ and $f(n) = n^{-(s-s_0)}$,

$$\sum_{\alpha<n\le\beta} a_n n^{-s} \ll_{\delta,T} \max_{\alpha<x\le\beta} \Big| \sum_{\alpha<n\le x} a_n n^{-s_0} \Big|, \tag{1.7}$$

so the uniform convergence of (1.6) follows from the convergence of $\sum_n a_n n^{-s_0}$ and Cauchy's criterion. ∎

The number

$$\inf\left\{\operatorname{Re}(s) : (1.6) \text{ converges}\right\}$$

is called the *abscissa of convergence* of the Dirichlet series (1.6). Here, of course, we allow the possibility that the infimum could be $\pm\infty$. The abscissa of convergence of the Dirichlet series $\sum_n |a_n| n^{-s}$ is called the *abscissa of absolute convergence* of (1.6). The two abscissas are related by the following inequality.

**Lemma 1.12.** *Suppose that $\sigma_c$ and $\sigma_a$ are the abscissa of convergence and the abscissa of absolute convergence of the Dirichlet series* (1.6)*. Then $\sigma_c \le \sigma_a \le \sigma_c + 1$.*

Dirichlet series are an important class of generating functions in number theory. In the remainder of this section, we discuss their properties related to their use in number theory. We first consider the relation between the sum-function of a Dirichlet series,

$$f(s) = \sum_{n=1}^{\infty} a_n n^{-s},$$

and the running sums of its coefficient sequence,

$$A(x) = \sum_{n \le x} a_n.$$

The passage from $A(x)$ to $f(s)$ is easy (at least, when $\operatorname{Re}(s)$ is sufficiently large):

$$f(s) = \sum_{n=1}^{\infty} a_n \int_n^{\infty} s x^{-s-1}\, dx = \int_1^{\infty} \left( \sum_{n \le x} a_n \right) s x^{-s-1}\, dx = \int_1^{\infty} A(x) s x^{-s-1}\, dx.$$

The inverse relation requires a little bit more work.

**Lemma 1.13 (Perron's formula).** *Suppose that $\alpha > 0$. Then*

$$\frac{1}{2\pi \mathrm{i}} \int_{\alpha - \mathrm{i}T}^{\alpha + \mathrm{i}T} \frac{u^s}{s}\, ds = \begin{cases} 1 + O\left(u^{\alpha}(T|\log u|)^{-1}\right) & \text{if } u > 1, \\ \frac{1}{2} + O\left(\alpha T^{-1}\right) & \text{if } u = 1, \\ O\left(u^{\alpha}(T|\log u|)^{-1}\right) & \text{if } 0 < u < 1. \end{cases}$$

*Proof.* This is an exercise in contour integration. ∎

**Corollary 1.14.** *Let $f(s)$ be the sum-function of the Dirichlet series* (1.6)*. Suppose that $x \notin \mathbb{Z}$ and $\alpha > \sigma_a$, where $\sigma_a$ is the abscissa of absolute convergence of* (1.6)*. Then*

$$\sum_{n \le x} a_n = \frac{1}{2\pi \mathrm{i}} \int_{\alpha - \mathrm{i}T}^{\alpha + \mathrm{i}T} f(s) x^s s^{-1}\, ds + O\left( \frac{x^{\alpha}}{T} \sum_{n=1}^{\infty} \frac{|a_n| n^{-\alpha}}{|\log(x/n)|} \right).$$

**Corollary 1.15.** *Suppose that $a_1, a_2, a_3, \ldots$ and $b_1, b_2, b_3, \ldots$ are sequences of complex numbers and the holomorphic functions $f(s)$ and $g(s)$ are defined in the half-plane* $\mathrm{Re}(s) > \sigma_0$ *by*

$$f(s) = \sum_{n=1}^{\infty} a_n n^{-s} \qquad \text{and} \qquad g(s) = \sum_{n=1}^{\infty} b_n n^{-s}.$$

*If $f(s) = g(s)$ whenever* $\mathrm{Re}(s) > \sigma_0$*, then $a_n = b_n$ for all $n = 1, 2, 3, \ldots$.*

*Proof.* We apply Corollary 1.14 with $x \notin \mathbb{Z}$, $\alpha = \sigma_0 + 2$ (this ensures the absolute convergence of the series on the line $\mathrm{Re}(s) = \alpha$), and $T = \infty$. We get

$$\sum_{n \le x} a_n = \frac{1}{2\pi \mathrm{i}} \int_{\alpha - \mathrm{i}\infty}^{\alpha + \mathrm{i}\infty} f(s) x^s s^{-1} \, ds = \frac{1}{2\pi \mathrm{i}} \int_{\alpha - \mathrm{i}\infty}^{\alpha + \mathrm{i}\infty} g(s) x^s s^{-1} \, ds = \sum_{n \le x} b_n.$$

Since this holds for all non-integer $x > 1$, it follows that $a_n = b_n$ for all $n = 1, 2, 3, \ldots$. ∎

The next two lemmas and their corollaries illustrate why Dirichlet series are convenient generating functions in multiplicative number theory.

**Lemma 1.16.** *Suppose that $f(n)$ is a multiplicative function. Then the identity*

$$\sum_{n=1}^{\infty} f(n) n^{-s} = \prod_{p} \left(1 + f(p)p^{-s} + f(p^2)p^{-2s} + \cdots\right)$$

*holds whenever the series on the left converges absolutely.*

*Proof.* The absolute convergence of the series $\sum_n f(n)n^{-s}$ implies the absolute convergence of the series $\sum_m f(p^m)p^{-ms}$ for all primes $p$. Let $x \ge 2$ and $r = \pi(x)$. Then

$$\begin{aligned}
\prod_{p \le x} \left(1 + f(p)p^{-s} + f(p^2)p^{-2s} + \cdots\right) &= \sum_{m_1=0}^{\infty} \cdots \sum_{m_r=0}^{\infty} f(p_1^{m_1}) \cdots f(p_r^{m_r})(p_1^{m_1} \cdots p_r^{m_r})^{-s} \\
&= \sum_{m_1=0}^{\infty} \cdots \sum_{m_r=0}^{\infty} f(p_1^{m_1} \cdots p_r^{m_r})(p_1^{m_1} \cdots p_r^{m_r})^{-s} \\
&= \sum_{\substack{n=1 \\ p \mid n \Rightarrow p \le x}}^{\infty} f(n) n^{-s},
\end{aligned}$$

where we have used the multiplicativity of $f$. Noting that the last sum contains, in particular, all the terms with $n \le x$, we conclude that

$$\left| \sum_{n \le x} f(n)n^{-s} - \prod_{p \le x} \left(1 + f(p)p^{-s} + f(p^2)p^{-2s} + \cdots\right) \right| \le \sum_{n > x} |f(n)| n^{-\sigma},$$

which establishes the desired identity. ∎

**Corollary 1.17.** *Suppose that $f(n)$ is a completely multiplicative function. Then the identity*

$$\sum_{n=1}^{\infty} f(n)n^{-s} = \prod_{p} \left(1 - f(p)p^{-s}\right)^{-1}$$

*holds whenever the series on the left converges absolutely.*

**Lemma 1.18.** *Suppose that $a_1, a_2, a_3, \ldots$ and $b_1, b_2, b_3, \ldots$ are sequences of complex numbers such that the Dirichlet series*

$$f(s) = \sum_{n=1}^{\infty} a_n n^{-s} \qquad \text{and} \qquad g(s) = \sum_{n=1}^{\infty} b_n n^{-s}$$

*converge absolutely in the half-plane $\mathrm{Re}(s) > \sigma_0$. Then the Dirichlet series*

$$h(s) = \sum_{n=1}^{\infty} c_n n^{-s}, \qquad c_n = \sum_{uv=n} a_u b_v,$$

*is also absolutely convergent in $\mathrm{Re}(s) > \sigma_0$ and $h(s) = f(s)g(s)$.*

*Proof.* Suppose first that $\sigma > \sigma_0$. Then

$$\begin{aligned}
\sum_{n\le x} |c_n| n^{-\sigma} &= \sum_{n\le x} \left| \sum_{uv=n} a_u b_v \right| n^{-\sigma} \\
&\le \sum_{n\le x} \left( \sum_{uv=n} |a_u b_v| \right) n^{-\sigma} = \sum_{uv\le x} |a_u b_v| (uv)^{-\sigma} \\
&\le \left( \sum_{u\le x} |a_u| u^{-\sigma} \right) \left( \sum_{v\le x} |b_v| v^{-\sigma} \right) \\
&\le \left( \sum_{u=1}^{\infty} |a_u| u^{-\sigma} \right) \left( \sum_{v=1}^{\infty} |b_v| v^{-\sigma} \right),
\end{aligned}$$

which proves the absolute convergence of $\sum_n c_n n^{-s}$ for $\mathrm{Re}(s) = \sigma$. In particular, we have that

$$\sum_{n>x} \left( \sum_{uv=n} |a_u b_v| \right) n^{-\sigma} \to 0 \qquad \text{as } x \to \infty,$$

so the second part of the lemma follows from the inequality

$$\left| \sum_{n\le x} c_n n^{-s} - \left( \sum_{u\le x} a_u u^{-s} \right) \left( \sum_{v\le x} b_v v^{-s} \right) \right| \le \sum_{n>x} \left( \sum_{uv=n} |a_u b_v| \right) n^{-\sigma}. \tag{1.8}$$

■

In the next series of corollaries $\zeta(s)$ is the Riemann zeta-function.

**Corollary 1.19.** *Suppose that* $\mathrm{Re}(s) > 1$. *Then*

$$\sum_{n=1}^{\infty} \mu(n) n^{-s} = \frac{1}{\zeta(s)}.$$

*Proof.* By Lemmas 1.2 and 1.18,

$$\left( \sum_{n=1}^{\infty} \mu(n) n^{-s} \right) \left( \sum_{n=1}^{\infty} n^{-s} \right) = 1.$$

■

**Corollary 1.20.** *Suppose that* $\mathrm{Re}(s) > 1$. *Then*

$$\sum_{n=1}^{\infty} \Lambda(n) n^{-s} = -\frac{\zeta'(s)}{\zeta(s)}.$$

## 1.4 Divisor functions

In this section we collect several standard estimates for the number of divisors function $d(n)$ and its averages. First of all, we note that $d(n)$ is multiplicative (by Lemma 1.1) and satisfies $d(p^k) = k+1$. These two observations lead (after some work) to the following upper bound for $d(n)$.

**Lemma 1.21.** *For any* $\epsilon > 0$, $d(n) \ll_\epsilon n^\epsilon$.

The bound in Lemma 1.21 is not tight, but it is also not too far from the best possible general bound (see Exercise 21). On the other hand, the next lemma shows that for most values of $n$, $d(n)$ is significantly smaller: its average value is $\log n$.

**Theorem 1.22 (Dirichlet).** *Suppose that* $x \geq 2$. *Then*

$$\sum_{n \leq x} d(n) = x \log x + (2\gamma - 1)x + O\left(x^{1/2}\right), \tag{1.9}$$

*where* $\gamma$ *is Euler's constant.*

*Proof.* Let $D(x)$ denote the left side of (1.9). We have

$$D(x) = \sum_{n \leq x} \sum_{uv=n} 1 = \sum_{uv \leq x} 1 = \sum_{\substack{uv \leq x \\ u \leq \sqrt{x}}} 1 + \sum_{\substack{uv \leq x \\ v \leq \sqrt{x}}} 1 - \sum_{u,v \leq \sqrt{x}} 1 = D_1(x) + D_2(x) - D_3(x), \qquad \text{say.}$$

Thus, (1.9) follows from the estimates

$$\begin{aligned} D_1(x) = D_2(x) &= \sum_{u \leq \sqrt{x}} \left[ \frac{x}{u} \right] = \sum_{u \leq \sqrt{x}} \frac{x}{u} + O\left(x^{1/2}\right) \\ &= x \log \sqrt{x} + \gamma x + O\left(x^{1/2}\right) \qquad \text{(by Corollary 1.5);} \\ D_3(x) &= \left[ \sqrt{x} \right]^2 = x + O\left(x^{1/2}\right). \end{aligned}$$

■

**Remark.** The estimation of the error term in (1.9) is a famous problem in analytic number theory. It is not too difficult to show that

$$\Delta(x) = \sum_{n\le x} d(n) - x\log x - (2\gamma - 1)x \ll x^{1/3}\log x.$$

Attempts to improve further on this and other similar bounds have stimulated the development of the theory of exponential sums (see Graham and Kolesnik [17] and Huxley [28]). The best result to date was obtained recently by Huxley [29]:

$$\Delta(x) \ll_\epsilon x^{131/416+\epsilon},$$

where $131/416 = 0.3149\ldots$. It is conjectured that

$$\Delta(x) \ll_\epsilon x^{1/4+\epsilon},$$

which if proven would be essentially best possible, as it is an old result of Hardy [20] that the bound $\Delta(x) \ll x^{1/4}$ does not hold for all $x$.

Often one needs upper bounds for higher moments of $d(n)$. The following theorem provides such an estimate.

**Theorem 1.23.** *Suppose that $x \ge 1$ and $k \in \mathbb{N}$. Then*

$$\sum_{n\le x}(d(n))^k \ll_k x(\log x)^{2^k-1} + 1. \tag{1.10}$$

*Proof.* By induction on $k$. The case $k = 1$ follows from Theorem 1.22. Now suppose that (1.10) holds for some $k \ge 1$. Then

$$\sum_{n\le x}(d(n))^{k+1} = \sum_{uv\le x}(d(uv))^k \le \sum_{uv\le x}(d(u)d(v))^k,$$

where the last step uses that $d(mn) \le d(m)d(n)$. Hence, by the inductive hypothesis,

$$\begin{aligned}\sum_{uv\le x}(d(u)d(v))^k &\le \sum_{u\le x}(d(u))^k \sum_{v\le x/u}(d(v))^k \\ &\ll_k x(\log x)^{2^k-1}\sum_{u\le x}\frac{(d(u))^k}{u} + \sum_{u\le x}(d(u))^k \ll_k x(\log x)^{2^{k+1}-1},\end{aligned}$$

on using the bound

$$\sum_{u\le x}\frac{(d(u))^k}{u} \ll_k (\log x)^{2^k},$$

which follows from the inductive hypothesis by partial summation. ∎

## Exercises

1. Prove (0.2).
2. Prove that Euler's function $\phi(n)$ is multiplicative.
3. Prove Lemma 1.1.
4. Prove Corollary 1.5. [HINT: The value of $c_1$ is $1 - \int_1^\infty \{x\}x^{-2}\,dx$.]
5. Prove Corollary 1.6.
6. Prove Lemma 1.8. [HINT: First show that the given sum equals $x^{-1}T(x)+O(1)$, where $T(x)$ is the sum appearing in the proof of Theorem 1.7.]
7. Prove that
$$\prod_{p\le x}\left(1-\frac{1}{p}\right) = \frac{C}{\log x}\left(1+O\left(\frac{1}{\log x}\right)\right),$$
where $C$ is an absolute constant. (It can be shown that, in fact, $C = e^{-\gamma}$, where $\gamma$ is Euler's constant.)
8. Let $B$ be the constant appearing in Theorem 1.9 and $C$ be the constant appearing in the last problem. Prove that
$$B + \log C = \sum_p \left(\frac{1}{p} + \log\left(1-\frac{1}{p}\right)\right).$$
9. Prove that under the hypotheses of Theorem 1.10,
$$\int_2^x \frac{f(y)\,dy}{y(\log y)^2} \ll \frac{f(x)}{\log x}.$$
10. Prove the converse part of Theorem 1.10.
11. Modify the proof of Theorem 1.10 to show that the PNT is equivalent to the statement that $\psi(x) \sim x$ as $x \to \infty$.
12. Verify (1.7).
13. Suppose that in Lemma 1.11 the assumption that the series $\sum_n a_n n^{-s_0}$ converges is weakened to the assertion that the partial sums
$$\sum_{n\le N} a_n n^{-s_0} \qquad (N = 1, 2, 3, \dots)$$
are bounded. Prove that the conclusion of the lemma stays true.
14. Prove Lemma 1.12.
15. Prove Lemma 1.13.
16. Prove (1.8).
17. Prove Corollary 1.20.
18. Prove the following identities:

(a) $\displaystyle\sum_{n=1}^{\infty} d(n)n^{-s} = \zeta^2(s)$ whenever $\mathrm{Re}(s) > 1$;

(b) $\displaystyle\sum_{n=1}^{\infty} |\mu(n)|n^{-s} = \zeta(s)/\zeta(2s)$ whenever $\mathrm{Re}(s) > 1$;

(c) $\displaystyle\sum_{n=1}^{\infty} \phi(n)n^{-s} = \zeta(s-1)/\zeta(s)$ whenever $\mathrm{Re}(s) > 2$.

19. Define a multiplicative function $f : \mathbb{N} \to \mathbb{C}$ by

$$f(p^k) = \binom{k - 1/2}{k} = (-1)^k \binom{-1/2}{k},$$

where the generalized binomial coefficient $\binom{s}{k}$, $s \in \mathbb{C}$, is the coefficient of $z^k$ in the Maclaurin expansion of $(1+z)^s$:

$$\binom{s}{k} = \frac{s(s-1)\cdots(s-k+1)}{k!}.$$

(a) Prove that the Dirichlet series $F(s) = \sum_n f(n)n^{-s}$ converges absolutely and uniformly on the compact subsets of the half-plane $\mathrm{Re}(s) > 1$.

(b) Prove that $F(s)^2 = \zeta(s)$ whenever $\mathrm{Re}(s) > 1$.

20. Prove that $d(n) \leq \sqrt{3n}$ for all $n \in \mathbb{N}$.

21. (a) Prove that there exists an absolute constant $c_1 > 0$ such that

$$d(n) \ll \exp\left(\frac{c_1 \log n}{\log\log n}\right).$$

(b) Let $n = p_1 p_2 \cdots p_k$, where $p_k$ denotes the $k$th prime. Prove that there exists an absolute constant $c_2 > 0$ such that

$$d(n) \gg \exp\left(\frac{c_2 \log n}{\log\log n}\right).$$

# Chapter 2

# The prime number theorem

In this chapter we develop the basic theory of the Riemann zeta-function to the level needed for the proof of the PNT in the form of Theorem 1. However, for technical reasons, instead of Theorem 1 we will establish the following result, which is equivalent to it (by Theorem 1.10).

**Theorem 2.1.** *There exists an absolute constant $c > 0$ such that for $x \geq 2$*

$$\psi(x) = x + O\big(x \exp\big(-c\sqrt{\log x}\big)\big).$$

## 2.1 Definition of $\zeta(s)$. The functional equation

We start by providing a rigorous definition of the zeta-function. As we said in the Introduction, we initially define $\zeta(s)$ for $\mathrm{Re}(s) > 1$ by

$$\zeta(s) = \sum_{n=1}^{\infty} n^{-s}. \tag{2.1}$$

Since the series converges uniformly on compact subsets of the half-plane $\mathrm{Re}(s) > 1$, it follows that $\zeta(s)$ is holomorphic in this half-plane. Moreover, since the convergence is absolute, Corollary 1.17 applies to $\zeta(s)$. In this way, we get the Euler product representation of $\zeta(s)$:

$$\zeta(s) = \prod_{p} \left(1 - \frac{1}{p^s}\right)^{-1}. \tag{2.2}$$

In particular, we deduce from (2.2) that $\zeta(s)$ is does not vanish in the half-plane $\mathrm{Re}(s) > 1$.

Our next goal is to extend the definition of $\zeta(s)$ to the whole complex plane $\mathbb{C}$. There are numerous ways to do this. We will follow the original approach of Riemann, which yields one of the most elegant and illuminating treatments of the analytic continuation of $\zeta(s)$ even today. Alternative proofs can be found in most monographs on the theory of the zeta-function (for example, Titchmarsh [50] gives seven such proofs). However, before we are in position to present Riemann's argument, we need to build up our knowledge about two classical functions.

### 2.1.1 The theta-series

If $\mathrm{Re}(z) > 0$ and $\alpha \in \mathbb{C}$, we define the theta-function $\vartheta(z;\alpha)$ by

$$\vartheta(z;\alpha) = \sum_{n=-\infty}^{\infty} \exp\left(-\pi z(n+\alpha)^2\right). \tag{2.3}$$

The property of $\vartheta(z;\alpha)$ that we are interested in is the following transformation formula.

**Lemma 2.2.** *Let $x > 0$ and $0 \le \alpha < 1$ be real. Then*

$$\vartheta(x;\alpha) = x^{-1/2}\exp\left(-\pi\alpha^2 x\right)\vartheta\left(x^{-1};-\mathrm{i}\alpha x\right). \tag{2.4}$$

*Proof.* Suppressing the dependence on $x$, we write $f(\alpha) = \vartheta(x;\alpha)$. Since the series

$$\sum_{n=-\infty}^{\infty} \exp\left(-\pi x(n+\alpha)^2\right) \qquad \text{and} \qquad \sum_{n=-\infty}^{\infty} n\exp\left(-\pi x(n+\alpha)^2\right)$$

converge uniformly in $\alpha$, $f(\alpha)$ is continuously differentiable. It is also clear that $f(\alpha)$ is 1-periodic. Hence, $f(\alpha)$ equals its Fourier series:

$$f(\alpha) = \sum_{n=-\infty}^{\infty} \hat{f}_n e(n\alpha).$$

Here $\hat{f}_n$ is the $n$th Fourier coefficient of $f(\alpha)$,

$$\hat{f}_n = \int_0^1 f(t)e(-nt)\,dt.$$

Thus, it suffices to show that

$$\int_0^1 f(t)e(-nt)\,dt = x^{-1/2}\exp\left(-\pi n^2 x^{-1}\right).$$

By the absolute convergence of the theta-series, we can integrate it term-by-term, whence

$$\begin{aligned}
\int_0^1 f(t)e(-nt)\,dt &= \sum_{m\in\mathbb{Z}} \int_0^1 \exp\left(-\pi x(m+t)^2\right) e(-nt)\,dt \\
&= \sum_{m\in\mathbb{Z}} \int_m^{m+1} \exp\left(-\pi x t^2\right) e(-n(t-m))\,dt \\
&= \sum_{m\in\mathbb{Z}} \int_m^{m+1} \exp\left(-\pi x t^2\right) e(-nt)\,dt \\
&= \int_{\mathbb{R}} \exp\left(-\pi x t^2\right) e(-nt)\,dt \\
&= x^{-1/2} \int_{\mathbb{R}} \exp\left(-\pi u^2\right) e\left(-nx^{-1/2}u\right) du \\
&= x^{-1/2}\exp\left(-\pi n^2 x^{-1}\right),
\end{aligned}$$

where the last step uses that the function $\exp(-\pi u^2)$ equals its Fourier transform. ∎

While the above proof can be generalized to all complex $\alpha$ and $z$ with $\mathrm{Re}(z) > 0$, it is easier to extend Lemma 2.2 to those cases by means of the identity theorem for analytic functions. This yields the following result.

**Corollary 2.3.** *Let* $\mathrm{Re}(z) > 0$ *and* $\alpha \in \mathbb{C}$. *Then*

$$\vartheta(z;\alpha) = z^{-1/2}\exp\big(-\pi\alpha^2 z\big)\vartheta\big(z^{-1};-\mathrm{i}\alpha z\big). \tag{2.5}$$

*Here* $z^{-1/2} = \exp\big(-\frac{1}{2}\operatorname{Log} z\big)$ *denotes the principal branch of the function* $z^{-1/2}$.

### 2.1.2 The gamma-function

The other special function that will figure prominently in our analysis is *Euler's gamma-function* $\Gamma(s)$. We define $\Gamma(s)$ by

$$\frac{1}{\Gamma(s)} = se^{\gamma s}\prod_{n=1}^{\infty}\left(1+\frac{s}{n}\right)e^{-s/n}, \tag{2.6}$$

where $\gamma$ is Euler's constant (recall (1.2)). As the infinite product is uniformly convergent on the compact subsets of $\mathbb{C} - \{0,-1,-2,\dots\}$, (2.6) defines $\Gamma(s)$ as a meromorphic function on $\mathbb{C}$ with simple poles at 0 and at the negative integers and with no zeros. We now state several important properties of $\Gamma(s)$ that will be needed later.

**Lemma 2.4.** *Suppose that* $s \neq 0,-1,-2,\dots$. *Then*

$$\Gamma(s+1) = s\Gamma(s). \tag{2.7}$$

*Proof.* Define

$$\varepsilon(x) = \sum_{k\le x}\frac{1}{k} - \log x - \gamma, \tag{2.8}$$

so that $\varepsilon(x) \to 0$ as $x \to \infty$ (in fact, $\varepsilon(x) = O(x^{-1})$). By (2.6),

$$\begin{aligned}\frac{\Gamma(s+1)}{\Gamma(s)} &= \frac{se^{\gamma s}}{(s+1)e^{\gamma(s+1)}}\lim_{n\to\infty}\prod_{k=1}^{n}\frac{\big(1+s/k\big)e^{-s/k}}{\big(1+(s+1)/k\big)e^{-(s+1)/k}}\\ &= \frac{s}{e^{\gamma}(s+1)}\lim_{n\to\infty}\frac{(s+1)(s+2)\cdots(s+n)}{(s+2)(s+3)\cdots(s+n+1)}\exp\left(1+\frac{1}{2}+\cdots+\frac{1}{n}\right)\\ &= se^{-\gamma}\lim_{n\to\infty}\frac{\exp\{\log n+\gamma+\varepsilon(n)\}}{n+s+1} = s\lim_{n\to\infty}\frac{ne^{\varepsilon(n)}}{n+s+1} = s.\end{aligned}$$

■

Similar arguments can be used to establish other relations between the function values of $\Gamma(s)$. In particular, we have the following two formulas:

$$\Gamma(s)\Gamma(1-s) = \frac{\pi}{\sin\pi s}, \qquad \Gamma(s)\Gamma(s+1/2) = \pi^{1/2}2^{1-2s}\Gamma(2s). \tag{2.9}$$

**Lemma 2.5 (Euler's integral formula).** *Suppose that* $\mathrm{Re}(s) > 0$. *Then*

$$\Gamma(s) = \int_0^\infty e^{-t} t^{s-1}\, dt. \tag{2.10}$$

*Proof.* We first show that

$$\Gamma(s) = \lim_{n\to\infty} \frac{n! \cdot n^s}{s(s+1)(s+2)\cdots(s+n)}.$$

By (2.6),

$$\begin{aligned}
\Gamma(s) &= s^{-1} e^{-\gamma s} \lim_{n\to\infty} \prod_{k=1}^{n} \left(1 + \frac{s}{k}\right)^{-1} e^{s/k} \\
&= s^{-1} e^{-\gamma s} \lim_{n\to\infty} \frac{1\cdot 2\cdots n}{(s+1)(s+2)\cdots(s+n)} \cdot \exp\left(s + \frac{s}{2} + \cdots + \frac{s}{n}\right) \\
&= \lim_{n\to\infty} \frac{n! \exp\{s(\log n + \varepsilon(n))\}}{s(s+1)(s+2)\cdots(s+n)} = \lim_{n\to\infty} \frac{n! \cdot n^s}{s(s+1)(s+2)\cdots(s+n)},
\end{aligned}$$

where $\varepsilon(x)$ is the function defined in (2.8). We now observe that

$$\int_0^n \left(1 - \frac{t}{n}\right)^n t^{s-1}\, dt = \frac{n! \cdot n^s}{s(s+1)(s+2)\cdots(s+n)}.$$

Indeed, when $s > 0$ the above integral converges and we have

$$\begin{aligned}
\int_0^n \left(1 - \frac{t}{n}\right)^n t^{s-1}\, dt &= n^s \int_0^1 (1-u)^n u^{s-1}\, du \\
&= n^s \frac{n}{s} \int_0^1 (1-u)^{n-1} u^s\, du \\
&= n^s \frac{n(n-1)}{s(s+1)} \int_0^1 (1-u)^{n-2} u^{s+1}\, du \\
&\;\;\vdots \\
&= n^s \frac{n(n-1)\cdots 1}{s(s+1)\cdots(s+n-1)} \int_0^1 u^{s+n-1}\, du \\
&= \frac{n! \cdot n^s}{s(s+1)(s+2)\cdots(s+n)}.
\end{aligned}$$

Thus, it suffices to prove that

$$\lim_{n\to\infty} \int_0^n \left(1 - \frac{t}{n}\right)^n t^{s-1}\, dt = \int_0^\infty e^{-t} t^{s-1}\, dt.$$

To this end, we consider the functions

$$f_n(t) = \begin{cases} (1 - t/n)^n t^{s-1} & \text{if } 0 \le t \le n, \\ 0 & \text{if } t > n. \end{cases}$$

Each of these functions is in $L^1[0,\infty)$ and satisfies the inequality

$$|f_n(t)| \le e^{-t}t^{\sigma-1},$$

where $\sigma = \mathrm{Re}(s)$. The last inequality is easily verified by taking logarithms and noting that

$$n\log\left(1-\frac{t}{n}\right) = -t - \frac{t^2}{2n} - \frac{t^3}{3n^2} - \cdots < -t.$$

Furthermore,

$$\lim_{n\to\infty} f_n(t) = t^{s-1}\lim_{n\to\infty}\left(1-\frac{t}{n}\right)^n = e^{-t}t^{s-1}.$$

Since the function $e^{-t}t^{\sigma-1}$ is in $L^1[0,\infty)$, the dominated convergence theorem yields

$$\lim_{n\to\infty}\int_0^\infty f_n(t)\,dt = \int_0^\infty \lim_{n\to\infty} f_n(t)\,dt = \int_0^\infty e^{-t}t^{s-1}\,dt,$$

which completes the proof of the lemma. ∎

We conclude our discussion of $\Gamma(s)$ with Stirling's formula, which provides an asymptotic expansion for $\log\Gamma(s)$ when $|s|\to\infty$.

**Lemma 2.6 (Stirling's formula).** *Suppose that* $|\arg s| < \pi$. *Then*

$$\log\Gamma(s) = (s-1/2)\log s - s + \log\sqrt{2\pi} + \int_0^\infty \frac{\Psi(u)}{u+s}\,du.$$

*Here* $\log s$ *denotes the principal branch of the logarithm and* $\Psi(u) = \{u\} - 1/2$.

**Corollary 2.7.** *Suppose that* $0 < \delta < \pi$ *and* $|\arg s| < \pi - \delta$. *Then*

$$\log\Gamma(s) = (s-1/2)\log s - s + \log\sqrt{2\pi} + O\big(|s|^{-1}\big)$$

*and*

$$\frac{\Gamma'(s)}{\Gamma(s)} = \log s + O\big(|s|^{-1}\big),$$

*the implied constants depending at most on* $\delta$.

**Corollary 2.8.** *Suppose that* $\alpha \le \sigma \le \beta$ *and* $|t| \ge 1$. *Then*

$$|\Gamma(\sigma+it)| = \sqrt{2\pi}|t|^{\sigma-1/2}\exp(-\pi|t|/2)\big\{1+O\big(|t|^{-1}\big)\big\},$$

*the implied constant depending at most on* $\alpha$ *and* $\beta$.

### 2.1.3 The functional equation

We are now in position to obtain the analytic continuation of $\zeta(s)$.

**Proposition 2.9.** *Suppose that* $\mathrm{Re}(s) > 1$. *Then*

$$\pi^{-s/2}\Gamma\left(\frac{s}{2}\right)\zeta(s) = \frac{1}{s(s-1)} + \frac{1}{2}\int_1^\infty \left(x^{s/2-1} + x^{-s/2-1/2}\right)(\vartheta(x;0) - 1)\,dx, \tag{2.11}$$

*where* $\vartheta(x;0)$ *is defined by* (2.3).

*Proof.* From (2.10), we have

$$\Gamma\left(\frac{s}{2}\right) = \int_0^\infty e^{-y}y^{s/2-1}\,dy = \pi^{s/2}n^s\int_0^\infty e^{-\pi x n^2}x^{s/2-1}\,dx.$$

Summing this identity over $n$, we get

$$\pi^{-s/2}\Gamma\left(\frac{s}{2}\right)\zeta(s) = \sum_{n=1}^\infty \int_0^\infty e^{-\pi x n^2}x^{s/2-1}\,dx.$$

After interchanging the order of summation and integration, the right side becomes

$$\int_0^\infty \sum_{n=1}^\infty e^{-\pi x n^2}x^{s/2-1}\,dx = \frac{1}{2}\int_0^\infty (\vartheta(x;0) - 1)x^{s/2-1}\,dx.$$

Next we write

$$\int_0^1 (\vartheta(x;0) - 1)x^{s/2-1}\,dx = \int_1^\infty \left(\vartheta(t^{-1};0) - 1\right)t^{-s/2-1}\,dt$$

and use (2.4) with $\alpha = 0$ to put the last integral in the form

$$\int_1^\infty \left(t^{1/2}\vartheta(t;0) - 1\right)t^{-s/2-1}\,dt = \int_1^\infty (\vartheta(t;0) - 1)t^{-s/2-1/2}\,dt + \frac{2}{s(s-1)}.$$

Clearly this completes the proof of (2.11). ∎

**Theorem 2.10.** *The function* $\zeta(s)$ *can be continued to a meromorphic function on* $\mathbb{C}$, *whose only singularity is a simple pole at* $s = 1$ *with residue* 1. *Furthermore, the* $\zeta(s)$ *satisfies the functional equation*

$$\pi^{-s/2}\Gamma\left(\frac{s}{2}\right)\zeta(s) = \pi^{(s-1)/2}\Gamma\left(\frac{1-s}{2}\right)\zeta(1-s). \tag{2.12}$$

*Proof.* Proposition 2.9 was obtained under the assumption $\mathrm{Re}(s) > 1$, but since for $x > 1$

$$\vartheta(x;0) - 1 \ll e^{-\pi x},$$

the integral on the right side of (2.11) is an entire function. Therefore, the function

$$\xi(s) = \frac{s(s-1)}{2}\int_1^\infty (\vartheta(x;0) - 1)\left(x^{s/2-1} + x^{(1-s)/2-1}\right)dx + 1 \tag{2.13}$$

is entire. Thus, the right side of the equation

$$\zeta(s) = \frac{\pi^{s/2}}{s(s-1)}\Gamma\left(\frac{s}{2}\right)^{-1}\xi(s) \tag{2.14}$$

(which is equivalent to (2.11) when $\mathrm{Re}(s) > 1$) is a meromorphic function whose only singularity is a simple pole at $s = 1$ with residue

$$\frac{\pi^{1/2}\xi(1)}{\Gamma(\frac{1}{2})} = 1.$$

This establishes the first part of the theorem.

The functional equation (2.12) follows from (2.14) and the invariance of $\xi(s)$ under the transformation $s \mapsto 1 - s$. ∎

## 2.2 The zeros of $\zeta(s)$

Since we want to work with the logarithmic derivative $\zeta'(s)/\zeta(s)$, we need to understand the zeros and the poles of the zeta-function. Theorem 2.10 provides the necessary information about the single pole of $\zeta(s)$. In this section, we concentrate on the zeros. So far we know that there are no zeros in the half-plane $\mathrm{Re}(s) > 1$. Since $1/\Gamma(s)$ is entire with simple zeros at the nonpositive integers, it follows from the functional equation (2.12) that the only zeros of $\zeta(s)$ in the half-plane $\mathrm{Re}(s) < 0$ are simple zeros at the negative even integers and that $\zeta(0) \neq 0$. The remaining part of the complex plane—the vertical strip $0 \leq \mathrm{Re}(s) \leq 1$— is a twilight zone which may, and indeed does, contain more zeros of the zeta-function. These come from the factor $\xi(s)$ on the right of (2.14). In this section we study $\xi(s)$ as an entire function of order 1.

In general, an entire function $f$ with $f(0) \neq 0$ is said to be of *finite order* if there is a number $\eta > 0$ such that

$$M_f(r) = \max\left\{|f(z)| : |z| \leq r\right\} \ll_{f,\eta} \exp(r^\eta).$$

When it is finite, the infimum of all such $\eta > 0$ is called the *order* of $f$. Entire functions of finite order enter our discussion because of the following result.

**Lemma 2.11.** *The function $\xi(s)$ is an entire function of order 1. Furthermore,*

$$\limsup_{|s|\to\infty} \frac{\log|\xi(s)|}{|s|\log|s|} = \frac{1}{2}. \tag{2.15}$$

*Proof.* Since $\xi(s) = \xi(1-s)$, it suffices to bound $|\xi(s)|$ in the half-plane $\mathrm{Re}(s) \geq 1/2$. There, we can estimate $\xi(s)$ by means of (2.14), Stirling's formula, and elementary upper bounds for $\zeta(s)$. When $\mathrm{Re}(s) > 1$, a variant of Corollary 1.5 yields

$$\zeta(s) = \frac{s}{s-1} - s\int_1^\infty \{x\}x^{-s-1}\,dx, \tag{2.16}$$

where $\{x\}$ denotes the fractional part of $x$. However, since the last integral represents a holomorphic function in $\mathrm{Re}(s) > 0$, the identity holds in this larger domain. In particular, we have

$$(s-1)\zeta(s) \ll |s|^2 \qquad \text{whenever } \mathrm{Re}(s) \geq 1/2.$$

Moreover, since $\log|\Gamma(s)| \leq |\log\Gamma(s)|$, Stirling's formula yields

$$\log|\Gamma(s)| \leq |s|\log|s| + O(|s|) \qquad \text{whenever } \mathrm{Re}(s) \geq 1/2.$$

Combining the last two estimates and (2.14), we obtain

$$\log|\xi(s)| \leq \tfrac{1}{2}|s|\log|s| + O(|s|) \qquad \text{whenever } \mathrm{Re}(s) \geq 1/2,$$

which establishes the first claim of the lemma. For the second claim, we note that

$$\log|\Gamma(|s|)| = \log\Gamma(|s|) = |s|\log|s| + O(|s|),$$

whence

$$\log\xi(|s|) = \tfrac{1}{2}|s|\log|s| + O(|s|) \qquad \text{as } |s| \to \infty.$$

■

We now take a short detour into the theory of entire functions of finite order.

**Lemma 2.12.** *Suppose that $f(s)$ is holomorphic in the closed disk $|s| \leq R$. Let $0 < r < R$ and let $a_1, \dots, a_n$ be the zeros of $f(s)$ lying inside the disk $|s| \leq r$ (listed according to multiplicities). Then*

$$\frac{|f(0)|R^n}{|a_1 a_2 \cdots a_n|} \leq \max_{|s|=R} |f(s)|.$$

*Proof.* Consider the function

$$F(s) = f(s)\prod_{k=1}^{n} \frac{R^2 - s\bar{a}_k}{R(s-a_k)}.$$

This function is holomorphic in $|s| \leq R$ and $|F(s)| = |f(s)|$ on the circle $|s| = R$. Thus, by the maximum modulus principle,

$$\frac{|f(0)|R^n}{|a_1 a_2 \cdots a_n|} = |F(0)| \leq \max_{|s|=R} |F(s)| = \max_{|s|=R} |f(s)|.$$

■

**Lemma 2.13.** *Suppose that $f(s)$ is an entire function of order $\eta$, and let $N(R)$ denote the number of zeros of $f(s)$ inside the disk $|s| < R$, counted according to their multiplicities. Then for any $R \geq 1$ and $\epsilon > 0$,*

$$N(R) \ll_{f,\eta,\epsilon} R^{\eta+\epsilon}.$$

*Proof.* Without loss of generality, we may assume that $f(0) = 1$. Let $a_1, \dots, a_n$ be the zeros of $f(s)$ inside the disk $|s| < R$, listed according to multiplicities, and choose an $r > 0$ so that

$$\max\left\{|a_k| : 1 \le k \le n\right\} \le r < R.$$

We apply Lemma 2.12 with $r$ and $2R$ and obtain

$$2^{N(R)} \le \prod_{|a_k|<R} \left(\frac{2R}{|a_k|}\right) \le \max_{|s|=2R} |f(s)| \ll_{f,\epsilon} \exp((2R)^{\eta+\epsilon}).$$

The desired conclusion now follows by taking logarithms. ∎

**Lemma 2.14.** *Let $R > 0$ and suppose that $f(s)$ is holomorphic in the disk $|s| \le R$. Then*

$$\left|f^{(n)}(0)\right| \le 2n!R^{-n} \max_{|s|=R} \mathrm{Re}(f(s) - f(0)).$$

*Proof.* It suffices to consider the case $f(0) = 0$. We write

$$\frac{f^{(n)}(0)}{n!} = r_n e^{i\phi_n}, \qquad s = Re^{i\theta}.$$

Then

$$\mathrm{Re}\ f(Re^{i\theta}) = \sum_{n=1}^{\infty} r_n R^n \cos(n\theta + \phi_n).$$

Since the last series is absolutely convergent, we can integrate it term-by-term to obtain the identity

$$\int_0^{2\pi} \left(1 + \cos(n\theta + \phi_n)\right) \mathrm{Re}\ f(Re^{i\theta})\, d\theta = \pi r_n R^n.$$

Hence,

$$\pi r_n R^n \le M \int_0^{2\pi} \left(1 + \cos(n\theta + \phi_n)\right) d\theta = 2\pi M,$$

where $M = \max\{\mathrm{Re}\ f(s) : |s| = R\}$. ∎

**Corollary 2.15.** *Suppose that $f(s)$ is an entire function such that* $\mathrm{Re}\ f(s) = o(|s|^n)$ *as $|s| \to \infty$. Then $f(s)$ is a polynomial of degree at most $n - 1$.*

**Theorem 2.16.** *Suppose that $f(s)$ is an entire function of order* 1 *with $f(0) \ne 0$, and let $a_1, a_2, a_3, \dots$ denote the zeros of $f(s)$ listed according to their multiplicities and arranged so that*

$$0 < |a_1| \le |a_2| \le \cdots \le |a_n| \le \cdots.$$

*Then $f(s)$ can be written as*

$$f(s) = e^{A+Bs} \prod_{n=1}^{\infty} \left(1 - \frac{s}{a_n}\right) e^{s/a_n},$$

*where $A$ and $B$ are constants.*

*Proof.* From Lemma 2.13 we deduce by partial summation that $\sum_n |a_n|^{-2}$ converges. Hence, the product

$$P(s) = \prod_{n=1}^{\infty} \left(1 - \frac{s}{a_n}\right) e^{s/a_n}$$

converges uniformly on compact sets. Thus, $P(s)$ is an entire function having the same zeros as $f(s)$ and $F(s) = f(s)/P(s)$ is an entire function with no zeros. Therefore, $F(s)$ has a holomorphic logarithm, that is, there is an entire function $g(s)$ such that $F(s) = e^{g(s)}$. We now prove that

$$g(s) = A + Bs \tag{2.17}$$

for some constants $A, B$.

In view of Corollary 1.14, (2.17) follows from the estimate $\operatorname{Re} g(s) = o(|s|^2)$ as $|s| \to \infty$. Since $\operatorname{Re} g(s)$ is a harmonic function, it suffices to establish this on a sequence of circles $|s| = R$ with $R \to \infty$. We choose the radii so that

$$\big|R - |a_n|\big| \gg R^{-1.1} \tag{2.18}$$

for all zeros $a_n$ of $f(s)$; this is possible because of Lemma 2.13. With such a choice of $R$, we have

$$-\log\left|(1 - s/a_n)e^{s/a_n}\right| \ll \begin{cases} R|a_n|^{-1} + \log R & \text{if } |a_n| \le 2R, \\ R^2|a_n|^{-2} & \text{if } |a_n| > 2R, \end{cases}$$

whenever $|s| = R$. Hence, by Lemma 2.13 and partial summation (see Exercise 6),

$$-\log|P(s)| \ll \log R \sum_{|a_n|\le 2R} 1 + R \sum_{|a_n|\le 2R} |a_n|^{-1} + R^2 \sum_{|a_n|>2R} |a_n|^{-2} \ll_\epsilon R^{1+\epsilon}$$

for any $0 < \epsilon < 1$. Therefore,

$$\operatorname{Re} g(s) = \log\left|e^{g(s)}\right| = \log|f(s)| - \log|P(s)| \ll_\epsilon R^{1+\epsilon}$$

whenever $|s| = R$ and $R$ is chosen so that (2.18) holds. This establishes (2.17) and completes the proof of the theorem. ∎

**Theorem 2.17.** *The function $\xi(s)$ has infinitely many zeros in the strip $0 \le \operatorname{Re}(s) \le 1$ and no zeros outside that strip. It can be written as*

$$\xi(s) = e^{Bs} \prod_{\rho} \left(1 - \frac{s}{\rho}\right) e^{s/\rho}, \tag{2.19}$$

*where $\rho$ runs through the zeros of $\xi(s)$ counted according to their multiplicities and $B$ is a constant.*

*Proof.* Because of Lemma 2.11, we can apply Theorem 2.16 to $\xi(s)$. Upon noting that $\xi(0) = 1$ (recall (2.13)), this proves (2.19). The infinitude of the zeros of $\xi(s)$ is also a consequence of Lemma 2.11. Indeed, if $\xi(s)$ had only a finite number of zeros, (2.19) would imply the estimate

$$\log|\xi(s)| \ll |s|,$$

which contradicts (2.15). ∎

**Lemma 2.18.** *There exists a constant $B_1$ such that*

$$\frac{\zeta'(s)}{\zeta(s)} = \frac{-1}{s-1} + B_1 + \sum_{n=1}^{\infty}\left(\frac{1}{s+2n} - \frac{1}{2n}\right) + \sum_{\rho}\left(\frac{1}{s-\rho} + \frac{1}{\rho}\right), \tag{2.20}$$

*where $\rho$ runs through the zeros of $\xi(s)$ counted according to their multiplicities.*

*Proof.* By logarithmic differentiation of (2.14), (2.6), and (2.19), we get

$$\begin{aligned}
\frac{\zeta'(s)}{\zeta(s)} &= \frac{-1}{s-1} - \frac{1}{s} + \log\sqrt{\pi} - \frac{\Gamma'(s/2)}{2\Gamma(s/2)} + \frac{\xi'(s)}{\xi(s)} \\
-\frac{\Gamma'(s)}{\Gamma(s)} &= \frac{1}{s} + \gamma + \sum_{n=1}^{\infty}\left(\frac{1}{s+n} - \frac{1}{n}\right) \\
\frac{\xi'(s)}{\xi(s)} &= B + \sum_{\rho}\left(\frac{1}{s-\rho} + \frac{1}{\rho}\right).
\end{aligned}$$

Combining these three formulas, we obtain (2.20) with $B_1 = \log\sqrt{\pi} + \frac{1}{2}\gamma + B$, where $\gamma$ is Euler's constant and $B$ is the constant appearing in (2.19). ∎

**Theorem 2.19.** *Suppose that $s = \sigma + \mathrm{i}t$, $-1 \le \sigma \le 2$. Then*

$$\frac{\zeta'(s)}{\zeta(s)} = \frac{-1}{s-1} + \sum_{|\operatorname{Im}\rho - t| \le 1} \frac{1}{s-\rho} + O(\log(|t|+2)). \tag{2.21}$$

*Proof.* We write $\tau = |t| + 2$. Under the hypotheses of the theorem, we have

$$\sum_{n=1}^{\infty}\left|\frac{1}{s+2n} - \frac{1}{2n}\right| \le \sum_{n\le\tau}\frac{3}{2n} + \sum_{n>\tau}\frac{|s|}{n^2} \ll \log\tau.$$

Substituting this into (2.20), we obtain

$$\frac{\zeta'(s)}{\zeta(s)} = \frac{-1}{s-1} + \sum_{\rho}\left(\frac{1}{s-\rho} + \frac{1}{\rho}\right) + O(\log\tau). \tag{2.22}$$

We now apply (2.22) to $s = 2 + \mathrm{i}t$. Logarithmic differentiation of the Euler product (2.2) yields

$$\left|\frac{\zeta'(2+\mathrm{i}t)}{\zeta(2+\mathrm{i}t)}\right| = \left|\sum_{p}\frac{\log p}{p^{2+\mathrm{i}t}-1}\right| \le \sum_{p}\frac{\log p}{p^2-1},$$

so (2.22) with $s = 2 + \mathrm{i}t$ gives

$$\operatorname{Re}\sum_{\rho}\left(\frac{1}{2+\mathrm{i}t-\rho} + \frac{1}{\rho}\right) \ll \log\tau. \tag{2.23}$$

Writing a typical zero of $\xi(s)$ in the form $\rho = \beta + \mathrm{i}\gamma$ and noting that $0 \le \beta \le 1$, we now find

$$\operatorname{Re}\frac{1}{2+\mathrm{i}t-\rho} = \frac{2-\beta}{(2-\beta)^2+(t-\gamma)^2} \gg \frac{1}{1+(t-\gamma)^2}, \qquad \operatorname{Re}\frac{1}{\rho} = \frac{\beta}{\beta^2+\gamma^2} \ge 0.$$

These inequalities and (2.23) give

$$\sum_{\rho} \frac{1}{1+(t-\gamma)^2} \ll \log\tau. \tag{2.24}$$

To prove (2.21) we subtract from (2.22) the respective formula for $s = 2 + \mathrm{i}t$ and obtain

$$\frac{\zeta'(s)}{\zeta(s)} = \frac{-1}{s-1} + \sum_{\rho}\left(\frac{1}{s-\rho} - \frac{1}{2+\mathrm{i}t-\rho}\right) + O(\log\tau). \tag{2.25}$$

By (2.24),

$$\sum_{|\operatorname{Im}\rho - t|>1} \left|\frac{1}{s-\rho} - \frac{1}{2+\mathrm{i}t-\rho}\right| \le \sum_{|\operatorname{Im}\rho-t|>1} \frac{3}{(t-\gamma)^2} \le \sum_{\rho} \frac{6}{1+(t-\gamma)^2} \ll \log\tau$$

and

$$\sum_{|\operatorname{Im}\rho - t|\le 1} \left|\frac{1}{2+\mathrm{i}t-\rho}\right| \le \sum_{|\operatorname{Im}\rho-t|\le 1} 1 \le \sum_{\rho} \frac{2}{1+(t-\gamma)^2} \ll \log\tau.$$

Hence, (2.21) follows from (2.25). ∎

We conclude this section by recording a direct consequence of inequality (2.24).

**Corollary 2.20.** *Suppose that $T \ge 2$. The number of the zeros of $\zeta(s)$ in the region*

$$0 \le \operatorname{Re}(s) \le 1, \qquad T \le |\operatorname{Im}(s)| \le T+1$$

*is $O(\log T)$.*

## 2.3 The zerofree region

**Theorem 2.21 (de la Vallée Poussin).** *There exists an absolute constant $c_1 > 0$ such that $\zeta(s)$ has no zero $\rho = \beta + \mathrm{i}\gamma$ with*

$$\beta \ge 1 - \frac{c_1}{\log(|\gamma|+2)}. \tag{2.26}$$

*Proof.* Suppose that $s = \sigma + \mathrm{i}t$ with $\sigma > 1$. Taking real parts in Corollary 1.20, we obtain

$$-\operatorname{Re}\left(\frac{\zeta'(s)}{\zeta(s)}\right) = \sum_{n=1}^{\infty} \Lambda(n) n^{-\sigma}\cos(t\log n).$$

Since for real $\theta$,

$$3 + 4\cos\theta + \cos 2\theta = 2(1 + \cos\theta)^2 \geq 0, \tag{2.27}$$

we have

$$-3\frac{\zeta'(\sigma)}{\zeta(\sigma)} - 4\,\mathrm{Re}\left(\frac{\zeta'(\sigma + \mathrm{i}t)}{\zeta(\sigma + \mathrm{i}t)}\right) - \mathrm{Re}\left(\frac{\zeta'(\sigma + 2\mathrm{i}t)}{\zeta(\sigma + 2\mathrm{i}t)}\right) \geq 0. \tag{2.28}$$

We now consider a particular zero $\rho = \beta_0 + \mathrm{i}\gamma_0$ and write inequalities for the terms on the left side of (2.28) when $t = \gamma_0$. Since $\zeta(s)$ has a pole of residue 1 at $s = 1$, we have

$$-\frac{\zeta'(\sigma)}{\zeta(\sigma)} = \frac{1}{\sigma - 1} + O(1). \tag{2.29}$$

In view of Exercise 10, we have $|\gamma_0| \geq c_2 > 0$. Thus, we obtain from (2.21) that

$$\begin{aligned} -\mathrm{Re}\left(\frac{\zeta'(\sigma + \mathrm{i}\gamma_0)}{\zeta(\sigma + \mathrm{i}\gamma_0)}\right) &\leq -\mathrm{Re}\sum_{|\gamma - \gamma_0| \leq 1} \frac{1}{(\sigma - \beta) + \mathrm{i}(\gamma_0 - \gamma)} + c_3 \log(|\gamma_0| + 2) \\ &\leq \frac{-1}{(\sigma - \beta_0)} + c_3 \log(|\gamma_0| + 2), \end{aligned} \tag{2.30}$$

and similarly,

$$-\mathrm{Re}\left(\frac{\zeta'(\sigma + 2\mathrm{i}\gamma_0)}{\zeta(\sigma + 2\mathrm{i}\gamma_0)}\right) \leq c_4 \log(|\gamma_0| + 2). \tag{2.31}$$

Inserting (2.29)–(2.31) into (2.28), we deduce that for $\sigma$ close to 1,

$$4(\sigma - \beta_0)^{-1} - 3(\sigma - 1)^{-1} \leq c_5 \log(|\gamma_0| + 2).$$

Choosing

$$\sigma = 1 + \frac{1}{2c_5 \log(|\gamma_0| + 2)},$$

we obtain

$$\beta_0 \leq 1 - \frac{1}{14c_5 \log(|\gamma_0| + 2)},$$

which establishes (2.26) when $|\gamma_0| \geq c_2$. We dispense with the last condition by noting that there are only $O(1)$ zeros with $|\gamma_0| \leq c_2$ and that none of them can be too close to the pole at $s = 1$. ∎

As we mentioned in the Introduction, there are more precise estimates for the zerofree region of $\zeta(s)$ than that in Theorem 2.21. While their proofs are too technical to present here in full detail, we will describe briefly the main idea. In the proof of Theorem 2.21, we derived bounds for $\zeta'(s)/\zeta(s)$ from Theorem 2.19, which in turn relied on estimates for $\Gamma'(s)/\Gamma(s)$. The more sophisticated approach towards the zerofree region bounds $\zeta'(s)/\zeta(s)$ using estimates for the sum

$$\sum_{n \leq N} n^{\mathrm{i}t} = \sum_{n \leq N} \exp(\mathrm{i}t \log n).$$

In fact, both Littlewood's and the Vinogradov–Korobov improvements on Theorem 1 (recall (0.8) and (0.9)) stem from improved estimates for this exponential sum. Those interested in learning more about those results should consult the specialized monographs on the theory of the zeta-function, e.g., Ivić [31], Karatsuba and Voronin [37], or Titchmarsh [50].

## 2.4 Proof of the prime number theorem

Our proof of Theorem 2.1 combines Theorem 2.21 and the following result, known as the *Riemann–Mangoldt explicit formula* for $\psi(x)$. An alternative proof is sketched in Exercise 17.

**Theorem 2.22.** *Suppose that* $2 \le T \le x$ *and* $\{x\} = 1/2$. *Then*

$$\psi(x) = x - \sum_{|\operatorname{Im}\rho|\le T} \frac{x^\rho}{\rho} + O\left(\frac{x(\log x)^2}{T}\right), \tag{2.32}$$

*where the summation is over the nontrivial zeros of* $\zeta(s)$ *with* $|\operatorname{Im}\rho| \le T$.

*Proof.* We apply Corollary 1.14 with $f(s) = -\zeta'(s)/\zeta(s)$ and $\alpha = 1 + (\log x)^{-1}$. In view of Corollary 1.20, this gives

$$\psi(x) = \frac{1}{2\pi\mathrm{i}} \int_{\alpha-\mathrm{i}T}^{\alpha+\mathrm{i}T} \left(-\frac{\zeta'(s)}{\zeta(s)}\right) \frac{x^s}{s}\, ds + O\left(\frac{x}{T} \sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^\alpha |\log(x/n)|}\right).$$

The error term is easily seen to be

$$\ll \frac{x\log x}{T}\left(\sum_{n\le x/2} \frac{1}{n} + \sum_{x/2<n\le 2x} \frac{1}{|x-n|} + \sum_{n>2x} \frac{1}{n^\alpha}\right) \ll \frac{x(\log x)^2}{T}.$$

Hence,

$$\psi(x) = \frac{1}{2\pi\mathrm{i}} \int_{\alpha-\mathrm{i}T}^{\alpha+\mathrm{i}T} \left(-\frac{\zeta'(s)}{\zeta(s)}\right) \frac{x^s}{s}\, ds + O\left(\frac{x(\log x)^2}{T}\right). \tag{2.33}$$

Observe that, by Corollary 2.20, we can always choose a number $T$ of a certain size so that

$$\big|T - |\gamma|\big| \gg (\log T)^{-1} \tag{2.34}$$

for all zeros $\rho = \beta + \mathrm{i}\gamma$ of $\zeta(s)$. With such a value of $T$ we move the path of integration to the contour $C$ on Fig. 2.1. The contribution from the poles of the integrand lying between the two contours is

$$x - \sum_{|\rho|\le T} \frac{x^\rho}{\rho} - \frac{\zeta'(0)}{\zeta(0)},$$

so it remains to show that the integral along $C$ is negligible. To this end, we note that by Theorem 2.19 and Corollary 2.20,

$$\left|\frac{\zeta'(s)}{\zeta(s)}\right| \ll (\log T)^2$$

whenever $s \in C$. Hence,

$$\begin{aligned}\left|\int_C \left(-\frac{\zeta'(s)}{\zeta(s)}\right) \frac{x^s}{s}\, ds\right| &\ll (\log T)^2 \left(\frac{1}{T}\int_{-1/2}^{\alpha} x^u\, du + \int_{-T}^{T} \frac{x^{-1/2}\, dy}{1+|y|}\right) \\ &\ll xT^{-1}(\log T)^2 + x^{-1/2}(\log T)^3 \ll xT^{-1}(\log x)^2.\end{aligned}$$

Once (2.32) has been established for $T$ subject to (2.34), removing that constraint by means of Corollary 2.20 is straightforward. ∎

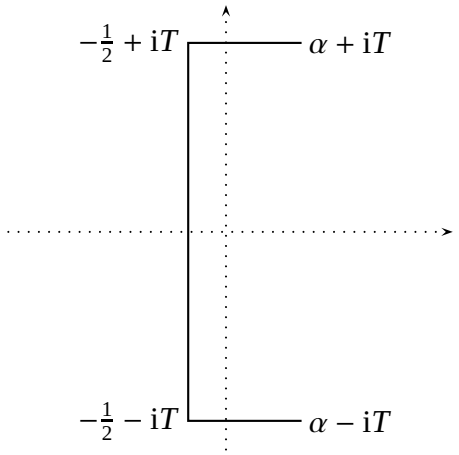


Figure 2.1:

*Proof of Theorem 2.1.* Suppose that $T \geq 10$ and set $\delta = c_1(\log T)^{-1}$, where $c_1 > 0$ is the constant from Theorem 2.21. By Corollary 2.20 and Theorem 2.21, the sum over the zeros on the right of (2.32) is

$$\ll x^{1-\delta} \sum_{r \leq 2\log T} 2^{-r} \sum_{|\operatorname{Im}\rho| \leq 2^r} 1 \ll x^{1-\delta}(\log T)^2.$$

Thus, the result follows from (2.32) on choosing

$$\log T = (\log x)^{1/2} + O(1).$$

■

## Exercises

1. Prove identities (2.9).
2. Prove Stirling's formula.
3. Suppose that $\operatorname{Re}(s) > 0$ and $N \in \mathbb{N}$. Prove that
$$\zeta(s) = \sum_{n \leq N} n^{-s} + \frac{N^{1-s}}{1-s} - s\int_N^\infty \{u\}u^{-s-1}\,du.$$
4. Prove Corollary 2.15.
5. Prove that the convergence of the series $\sum_n |a_n|^{-2}$ implies the uniform convergence on compact sets of the product
$$\prod_{n=1}^{\infty}\left(1-\frac{s}{a_n}\right)e^{s/a_n}.$$

6. Suppose that $f(s)$ is an entire function of order 1 with $f(0) \neq 0$ and $a_1, a_2, a_3, \dots$ are its zeros, labeled as in Theorem 2.16. Prove that:

   (a) $\displaystyle\sum_{|a_n| \le R} |a_n|^{\alpha} \ll_{\alpha,\epsilon} R^{\alpha+1+\epsilon} + 1$ for any $\epsilon > 0$;

   (b) $\displaystyle\sum_{|a_n| > R} |a_n|^{\alpha} \ll_{\alpha,\epsilon} R^{\alpha+1+\epsilon}$ for any $\alpha < -1$ and $0 < \epsilon < -\alpha - 1$.

7. Prove that $\zeta(0) = -1/2$ and $\zeta(-1) = -1/12$.

8. Prove that the Laurent expansion of $\zeta'(s)/\zeta(s)$ about $s = 1$ is

$$\frac{\zeta'(s)}{\zeta(s)} = \frac{-1}{s-1} + \gamma + \cdots,$$

   where $\gamma$ is Euler's constant. [Hint: Use (2.16).]

9. Prove that the value of $B$ in (2.19) is $B = \frac{1}{2}\log 4\pi - \frac{1}{2}\gamma - 1$, where $\gamma$ is Euler's constant.

10. Prove that any zero $\rho$ of the function $\xi(s)$ satisfies

$$|\operatorname{Im}\rho| \ge \sqrt{-B^{-1} - 1} > 6.503695\ldots,$$

    where $B$ is the constant appearing in (2.19) (and in the last problem).

11. The purpose of this problem is to show that the value of the constant $B$ in Theorem 1.9 is

$$B = \gamma - \sum_{p} \left( \frac{1}{p} + \log\left(1 - \frac{1}{p}\right)\right), \tag{$*$}$$

    where $\gamma$ is Euler's constant. Note that combining this identity and the result of Exercise 1.8, we find that the constant $C$ in Exercise 1.7 equals $e^{-\gamma}$.

    (a) Let $S(x)$ denote the sum on the left side (1.5) and define $f(s) = \sum_p p^{-s}$. Prove that if $\operatorname{Re}(s) > 1$, then

$$f(s) = (s-1)\int_1^{\infty} S(x) x^{-s}\, dx.$$

    (b) Suppose that $\sigma > 1$. Combining Theorem 1.9 and part (a), prove that

$$f(\sigma) = -\log(\sigma - 1) - \gamma + B + O(-(\sigma-1)\log(\sigma-1)).$$

    (c) Suppose that $\operatorname{Re}(s) > 1$. Prove that $f(s) = \log\zeta(s) + g(s)$, where $g(s)$ is holomorphic in the half-plane $\operatorname{Re}(s) > 1/2$.

    (d) Derive $(*)$ from parts (b) and (c).

12. Suppose that $T \ge 10$ and $N(T)$ denotes the number of zeros $\rho$ of $\xi(s)$ with $0 < \operatorname{Im}\rho \le T$. Prove that

$$N(T) = \frac{T}{2\pi}\log\left(\frac{T}{2\pi e}\right) + O(\log T).$$

    [Hint: Apply the argument principle to $\xi(s)$ and the rectangle with vertices $-1 \pm iT$, $2 \pm iT$. Use the results in §2.2 to estimate the contribution from the horizontal lines. Use the functional equation to replace the integral over the line $\operatorname{Re}(s) = -1$ by an integral over the line $\operatorname{Re}(s) = 2$. Finally, use that on the line $\operatorname{Re}(s) = 2$, $\zeta'(s)/\zeta(s)$ has a Dirichlet series representation.]

13. Suppose that $\rho_1, \rho_2, \rho_3, \dots$ are the zeros of $\xi(s)$ in the upper half-plane, listed according to multiplicities and arranged so that $0 < \gamma_1 \le \gamma_2 \le \gamma_3 \le \cdots$, where $\gamma_n = \operatorname{Im} \rho_n$. Prove that

$$\lim_{n\to\infty} \frac{\gamma_n \log n}{2\pi n} = 1.$$

14. Prove Theorem 2.22 for an arbitrary $x \ge 2$. [HINT: Use Lemma 2.18.]

15. Define $\psi_0(x) = \frac{1}{2}\{\psi(x^+) + \psi(x^-)\}$. Prove that for $x > 1$,

$$\psi_0(x) = x - \sum_\rho \frac{x^\rho}{\rho} - \frac{\zeta'(0)}{\zeta(0)} - \frac{1}{2}\log\left(1 - x^{-2}\right).$$

Here $\sum_\rho = \lim_{T\to\infty} \sum_{|\rho|\le T}$.

16. Define $\psi_1(x) = \sum_{n\le x}(x-n)\Lambda(n)$. Prove that for $x > 1$,

$$\psi_1(x) = \frac{x^2}{2} - \sum_\rho \frac{x^{\rho+1}}{\rho(\rho+1)} - x\frac{\zeta'(0)}{\zeta(0)} + \frac{\zeta'(-1)}{\zeta(-1)} - \sum_{k=1}^{\infty} \frac{x^{1-2k}}{2k(2k-1)}.$$

17. The purpose of this exercise is to deduce the PNT directly from (2.33) instead from the explicit formula (2.32). Starting with (2.33) move the integration to a polygonal contour $C$ that is similar to the contour displayed on Fig. 2.1 but has vertices at $\alpha \pm iT$ and $\eta \pm iT$, where $\eta = 1 - \frac{1}{2}c_1(\log T)^{-1}$. Estimate the integral over $C$ to obtain the PNT.

# Chapter 3

# Prime numbers in arithmetic progressions

In this chapter we study Dirichlet characters and $L$-functions and prove Theorem 2.

## 3.1 Characters

### 3.1.1 Characters of finite abelian groups

Let $G$ be a finite abelian group of order $m$, written multiplicatively. A group homomorphism $\chi : G \to \mathbb{C}^\times$ is called a *character* of $G$, that is,

$$\chi(xy) = \chi(x)\chi(y) \qquad \text{for all } x, y \in G.$$

In particular, $\chi(e) = 1$ and (by Lagrange's theorem on finite groups) $\chi(x)^m = \chi(x^m) = \chi(e) = 1$. That is, $\chi(x)$ is an $m$th root of unity.

The characters of $G$ form a group $\hat{G}$ under pointwise multiplication:

$$(\chi_1\chi_2)(x) = \chi_1(x)\chi_2(x) \qquad \text{for all } x \in G.$$

The identity element of $\hat{G}$ is the trivial character

$$\chi_0(x) = 1 \qquad \text{for all } x \in G,$$

and the inverse of $\chi$ is its complex-conjugate character $\overline{\chi}$.

**Theorem 3.1.** $\hat{G} \cong G$.

*Proof.* Suppose first that $G$ is cyclic, $G = \langle g \rangle$. Writing a generic element $x$ of $G$ as $x = g^y$, we find that every character $\chi \in \hat{G}$ must be of the form

$$\chi_a(x) = \chi_a(g^y) = e(ay/m) \qquad (a \in \mathbb{Z}),$$

that is, $\hat{G}$ is a cyclic group of order $m$ generated by $\chi_1$.

Now, suppose that $G$ is an arbitrary finite abelian group. By the structural theorem for finite abelian groups, $G$ can be written as the direct product of cyclic groups, $G = C_1 \times C_2 \times \cdots \times C_k$. Given an $x = x_1 x_2 \cdots x_k$, $x_j \in C_j$, we define a character $\chi \in \hat{G}$ by

$$\chi(y) = \chi_1(y_1)\chi_2(y_2)\cdots\chi_k(y_k) \qquad \text{for } y = y_1 y_2 \cdots y_k \in G,\ y_j \in C_j;$$

here $\chi_j$ is the character in $\hat{C}_j$ corresponding to $x_j$ under the above isomorphism. Since the map $x \mapsto \chi$ is an isomorphism of abelian groups, the result follows. ∎

**Corollary 3.2.** *Suppose that $G$ is a finite abelian group and $x$ is an element of $G$ other than the identity. Then there is a character $\chi \in \hat{G}$ such that $\chi(x) \neq 1$.*

*Proof.* This is a consequence of the proof of the theorem. As in that proof, we write $G$ as the direct product of cyclic groups, $G = C_1 \times C_2 \times \cdots \times C_k$. Then $x = x_1 x_2 \cdots x_k$, $x_j \in C_j$, and some $x_j$ is not the identity. Without loss of generality, we may assume that $x_1 \neq e$. Let $g$ be the generator of $C_1$. The character $\chi$ corresponding to $ge \cdots e$ under the isomorphism from the proof of the theorem has the desired property. ∎

**Lemma 3.3.** *Let $G$ be a finite abelian group and denote by $e$ and $\chi_0$ the identity element and the trivial character of $G$. Then the following two orthogonality relations hold:*

$$\sum_{x \in G} \chi(x) = \begin{cases} |G| & \text{if } \chi = \chi_0, \\ 0 & \text{otherwise}; \end{cases} \tag{3.1}$$

*and*

$$\sum_{\chi \in \hat{G}} \chi(x) = \begin{cases} |G| & \text{if } x = e, \\ 0 & \text{otherwise}. \end{cases} \tag{3.2}$$

*Proof.* Suppose that $\chi \neq \chi_0$. Then for some $x_0 \in G$, $\chi(x_0) \neq 1$. We now observe that

$$\chi(x_0) \sum_{x \in G} \chi(x) = \sum_{x \in G} \chi(x_0 x) = \sum_{y \in x_0 G} \chi(y) = \sum_{y \in G} \chi(y).$$

Since $\chi(x_0) \neq 1$, the sum on the right must be equal to 0. This establishes (3.1) when $\chi$ is nontrivial; the alternative case is straightforward.

Now suppose that $x \neq e$. By Corollary 3.2, there is a character $\chi_0 \in \hat{G}$ such that $\chi_0(x) \neq 1$. But

$$\chi_0(x) \sum_{\chi \in \hat{G}} \chi(x) = \sum_{\chi \in \hat{G}} \chi_0 \chi(x) = \sum_{\psi \in \chi_0 \hat{G}} \psi(x) = \sum_{\psi \in \hat{G}} \psi(x),$$

and since $\chi_0(x) \neq 1$, the sum on the right must be equal to 0. Again, the remaining case is straightforward. ∎

### 3.1.2 Dirichlet characters

Let $q \geq 1$ be an integer. Then $\mathbb{Z}/q\mathbb{Z}$ is a commutative ring. Let $G_q = (\mathbb{Z}/q\mathbb{Z})^\times$ denote the multiplicative group of $\mathbb{Z}/q\mathbb{Z}$. (Recall that a residue class in $\mathbb{Z}/q\mathbb{Z}$ is invertible if and only if it is relatively prime to the modulus $q$.) Then $G_q$ is an abelian group of order $\phi(q)$, where $\phi(q)$ is Euler's function. Of course, a character of $G_q$ is a homomorphism $\chi : G_q \to \mathbb{C}^\times$. It will be convenient to extend the domain of each character to all elements of the ring $\mathbb{Z}/q\mathbb{Z}$ by setting

$$\chi(n) = 0 \qquad \text{if } \gcd(n, q) > 1.$$

This extended function will be called a *Dirichlet character modulo q*, or simply a *Dirichlet character*. We will often regard each Dirichlet character modulo $q$ as a $q$-periodic function $\chi : \mathbb{Z} \to \mathbb{C}$. Although Dirichlet characters are not group homomorphisms, they are still completely multiplicative:

$$\chi(mn) = \chi(m)\chi(n) \qquad \text{for all } m, n \in \mathbb{Z}. \tag{3.3}$$

We refer to the extension of the trivial group character $\chi_0$ as the *principal character modulo q*; we will denote the principal character by $\chi_0$. The orthogonality relations in Lemma 3.3 yield the following orthogonality relations among Dirichlet characters.

**Lemma 3.4.** *Suppose that $q \geq 1$. If $\chi$ is a Dirichlet character modulo $q$, then*

$$\sum_{n=1}^{q} \chi(n) = \begin{cases} \phi(q) & \text{if } \chi = \chi_0, \\ 0 & \text{otherwise.} \end{cases} \tag{3.4}$$

*Furthermore,*

$$\sum_{\chi \bmod q} \chi(n) = \begin{cases} \phi(q) & \text{if } n \equiv 1 \pmod{q}, \\ 0 & \text{otherwise,} \end{cases} \tag{3.5}$$

*where the sum on the right side is over all Dirichlet characters modulo $q$.*

Let $\chi$ be a non-principal Dirichlet character modulo $q$, let $q_1$ be a proper divisor of $q$, and let $\chi_1$ be a non-principal character modulo $q_1$ such that

$$\chi(n) = \chi_1(n)\chi_0(n) \qquad \text{for all } n \in \mathbb{Z}, \tag{3.6}$$

where $\chi_0$ is the principal character modulo $q$. Then we say that $\chi_1$ *induces* $\chi$. If $\chi$ is a non-principal Dirichlet character modulo $q$ and there exists a character $\chi_1$ as in (3.6), then $\chi$ is called *imprimitive*; otherwise, $\chi$ is called *primitive*. Note that principal characters are neither primitive, nor imprimitive. If $\chi$ is an imprimitive Dirichlet character modulo $q$, we define[1] its *conductor* to be the least modulus $q^*$ such that there exists a (necessarily primitive) character $\chi^*$ modulo $q^*$ which induces $\chi$. If $\chi$ is primitive, we define its conductor to be equal to the modulus $q$, and if $\chi$ is principal, we define the conductor to be equal to 1.

[1] This definition requires some justification; see Exercise 2.

### 3.1.3 Gaussian sums

We now introduce the *Gaussian sum*. If $\chi$ is a Dirichlet character modulo $q$ and $a$ is an integer, we define

$$\tau(\chi, a) = \sum_{m \bmod q} \chi(m) e(am/q), \tag{3.7}$$

where the summation is over any complete system of residues modulo $q$.

**Lemma 3.5.** *Let $\chi$ be a Dirichlet character modulo $q$ and suppose that either* $\gcd(a, q) = 1$ *or $\chi$ is primitive. Then*

$$\tau(\chi, a) = \overline{\chi}(a)\tau(\chi, 1). \tag{3.8}$$

*Proof.* First, suppose that $(a, q) = 1$. Then

$$\tau(\chi, a) = \overline{\chi}(a) \sum_{m \bmod q} \chi(am) e(am/q) = \overline{\chi}(a) \sum_{n \bmod q} \chi(n) e(n/q) = \overline{\chi}(a)\tau(\chi, 1).$$

Here, we used that if $m$ runs through a complete system of residues modulo $q$, then so does $am$.

Now, suppose that $\chi$ is primitive and $(a, q) = k > 1$. We write $a = ka_1$, $q = kq_1$, and note that there exists an integer $b$ such that

$$(b, q) = 1, \qquad b \equiv 1 \pmod{q_1}, \qquad \chi(b) \neq 1.$$

Then

$$\chi(b)\tau(\chi, a) = \sum_{m \bmod q} \chi(bm) e(a_1 m/q_1) = \sum_{m \bmod q} \chi(bm) e(a_1 bm/q_1) = \tau(\chi, a).$$

Since $\chi(b) \neq 1$, it follows that $\tau(\chi, a) = 0$, which establishes the second claim of the lemma. ∎

Identity (3.8) is useful for transforming exponential sums into character sums and *vice versa*. However, for such applications it is crucial to be sure that $\tau(\chi, 1)$ is nonzero. The next lemma determines exactly the characters for which this is the case.

**Lemma 3.6.** *Let $\chi$ be a Dirichlet character modulo $q$ induced by a primitive character $\chi^*$ modulo $q^*$. Then*

$$\tau(\chi, 1) = \mu\left(\frac{q}{q^*}\right) \chi^*\left(\frac{q}{q^*}\right) \tau(\chi^*, 1). \tag{3.9}$$

*Moreover, if $\chi$ is primitive, then $|\tau(\chi, 1)| = \sqrt{q}$.*

*Proof.* Assume first that $\chi$ is primitive. Summing (3.8) over all $a$ modulo $q$, we get

$$\begin{aligned} |\tau(\chi, 1)|^2 \sum_{a \bmod q} |\chi(a)|^2 &= \sum_{a \bmod q} |\tau(\chi, a)|^2 \\ &= \sum_{a \bmod q} \sum_{m \bmod q} \chi(m) e(am/q) \sum_{n \bmod q} \overline{\chi}(n) e(-an/q) \\ &= \sum_{m \bmod q} \sum_{n \bmod q} \chi(m)\overline{\chi}(n) \sum_{a \bmod q} e(a(m-n)/q). \end{aligned} \tag{3.10}$$

The innermost sum on the right side of (3.10) is $q$ or 0 according as $a \equiv b \pmod q$ or not. Hence,

$$|\tau(\chi,1)|^2 \sum_{a \bmod q} |\chi(a)|^2 = q \sum_{m \bmod q} |\chi(m)|^2.$$

This proves the second claim of the lemma.

We now turn to (3.9). Using Lemma 1.2, we can write the principal character $\chi_0$ modulo $q$ as

$$\chi_0(n) = \sum_{d|(n,q)} \mu(d).$$

Thus,

$$\begin{aligned}\tau(\chi,1) &= \sum_{m \bmod q} \chi^*(m)\chi_0(m)e(m/q) = \sum_{m \bmod q} \chi^*(m)e(m/q) \sum_{d|(m,q)} \mu(d)\\ &= \sum_{d|q} \mu(d) \sum_{n \bmod q/d} \chi^*(nd)e(nd/q)\\ &= \sum_{d|q} \mu(d)\chi^*(d) \sum_{n \bmod q/d} \chi^*(n)e(nd/q).\end{aligned}$$

Note that the terms with $(d, q^*) > 1$ do not contribute to the last sum. Thus, we may restrict the summation over $d$ to the divisors of $q_0 = q/q^*$:

$$\tau(\chi,1) = \sum_{d|q_0} \mu(d)\chi^*(d) \sum_{n \bmod q/d} \chi^*(n)e(nd/q).$$

We now write the summation variable $n$ modulo $q/d$ as $q^*v + u$, where $u$ runs over a complete system of residues modulo $q^*$ and $v$ runs over a complete system of residues modulo $q/dq^* = q_0/d$. We get

$$\begin{aligned}\tau(\chi,1) &= \sum_{d|q_0} \mu(d)\chi^*(d) \sum_{u \bmod q^*} \sum_{v \bmod q_0/d} \chi^*(q^*v+u)e((q^*v+u)d/q)\\ &= \sum_{d|q_0} \mu(d)\chi^*(d) \sum_{u \bmod q^*} \chi^*(u)e(ud/q) \sum_{v \bmod q_0/d} e(vd/q_0).\end{aligned}$$

Since the innermost sum vanishes when $q_0/d > 1$, the result follows. ∎

### 3.1.4 The Pólya–Vinogradov theorem

Suppose that $\chi$ is a non-principal character modulo $q$. From the orthogonality relation (3.4),

$$\left| \sum_{M<n\le M+N} \chi(n) \right| \le \phi(q)/2$$

for all $M, N \ge 1$. In this section, we will improve on this trivial bound. The next result was obtained independently in 1918 by Pólya [45] and I. M. Vinogradov [56] and is known as the *Pólya–Vinogradov inequality*.

**Theorem 3.7.** *Suppose that $M, N$ are positive integers and $\chi$ is a non-principal character modulo $q$. Then*

$$\left| \sum_{M<n\le M+N} \chi(n) \right| \le \sqrt{3q}\log q. \tag{3.11}$$

*Proof.* First, suppose that $\chi$ is primitive. Then, by (3.8),

$$\tau(\bar{\chi},1) \sum_{M<n\le M+N} \chi(n) = \sum_{M<n\le M+N} \sum_{m \bmod q} \bar{\chi}(m) e(nm/q) = \sum_{m \bmod q} \bar{\chi}(m) \sum_{M<n\le M+N} e(mn/q).$$

Hence, we deduce from Lemma 3.6 that

$$\left| \sum_{M<n\le M+N} \chi(n) \right| \le q^{-1/2} \sum_{m=1}^{q-1} \left| \sum_{M<n\le M+N} e(mn/q) \right|.$$

On noting that the modulus of the inner sum is $|\sin(\pi mN/q)/\sin(\pi m/q)|$, we get the inequality

$$\left| \sum_{M<n\le M+N} \chi(n) \right| \le q^{1/2} \sum_{m=1}^{q-1} \csc(\pi m/q).$$

We now use apply the inequality $\csc(\pi x) \le (2x)^{-1}$ for $0 < x \le 1/2$. When $q = 2k$, we obtain

$$\begin{aligned} \sum_{m=1}^{q-1} \csc(\pi m/q) &\le q \sum_{m=1}^{k-1} m^{-1} + 1 \le q \sum_{m=1}^{k-1} \log\left(\frac{2m+1}{2m-1}\right) + 1 \\ &= q\log(q-1) + 1 \le q\log q; \end{aligned}$$

and when $q = 2k+1$,

$$\sum_{m=1}^{q-1} \csc(\pi m/q) \le q \sum_{m=1}^{k} m^{-1} \le q \sum_{m=1}^{k} \log\left(\frac{2m+1}{2m-1}\right) = q\log q.$$

This establishes (3.11) for primitive characters.

On the other hand, if $\chi$ is induced by a primitive character $\chi^*$ modulo $r$, $r < q$, we have

$$\sum_{M<n\le M+N} \chi(n) = \sum_{M<n\le M+N} \chi^*(n) \sum_{d|(q,n)} \mu(d) = \sum_{d|q} \mu(d)\chi^*(d) \sum_{M/d<m\le (M+N)/d} \chi^*(m).$$

Since $\chi^*$ is primitive, the sum over $m$ is bounded above by $r^{1/2}\log r$. Hence,

$$\left| \sum_{M<n\le M+N} \chi(n) \right| \le r^{1/2}\log r \sum_{d|q} |\chi^*(d)| \le d(q/r) r^{1/2}\log r,$$

where the last step uses that the terms with $(d,r) > 1$ do not contribute to the sum over $d$. The desired result now follows from the elementary bound $d(n) \le \sqrt{3n}$. ∎

## 3.2 Dirichlet $L$-functions

If $\chi$ is a Dirichlet character modulo $q$, we define the *Dirichlet L-function* $L(s,\chi)$ by

$$L(s,\chi) = \sum_{n=1}^{\infty} \chi(n) n^{-s} \qquad (\mathrm{Re}(s) > 1). \tag{3.12}$$

Since $|\chi(n)| \leq 1$, this series converges absolutely and uniformly on the compact subsets of the half-plane $\mathrm{Re}(s) > 1$. Furthermore, when $\chi$ is a non-principal character, the series in (3.12) converges uniformly (but not absolutely) on the compact subsets of $\mathrm{Re}(s) > 0$:

$$L(s,\chi) = \int_{1/2}^{\infty} S(x) s x^{-s-1}\, dx, \qquad S(x) = \sum_{n \leq x} \chi(n). \tag{3.13}$$

Thus, $L(s,\chi)$ is holomorphic in the half-plane $\mathrm{Re}(s) > 1$, and for non-principal $\chi$ even in $\mathrm{Re}(s) > 0$. By Lemma 1.16, every $L$-function has an Euler product:

$$L(s,\chi) = \prod_{p} \left(1 - \chi(p) p^{-s}\right)^{-1} \qquad (\mathrm{Re}(s) > 1). \tag{3.14}$$

In particular, (3.14) implies that $L(s,\chi) \neq 0$ when $\mathrm{Re}(s) > 1$.

Next, we want to obtain an analytic continuation of $L(s,\chi)$ to a meromorphic function on $\mathbb{C}$. We observe that it suffices to consider the case when $\chi$ is primitive. Indeed, if $\chi$ is an imprimitive character modulo $q$ induced by a primitive character $\chi^*$ modulo $q^*$, then (3.14) yields

$$L(s,\chi) = \prod_{p} \left(1 - \chi(p) p^{-s}\right)^{-1} = \prod_{p \nmid q} \left(1 - \chi^*(p) p^{-s}\right)^{-1} = L(s,\chi^*) \prod_{p \mid q} \left(1 - \chi^*(p) p^{-s}\right).$$

Therefore, the analytic continuation of $L(s,\chi)$ is a straightforward consequence from the analytic continuation of $L(s,\chi^*)$ and the holomorphy of the finite product on the right. Similarly, if $\chi_0$ is the principal character modulo $q$, we have

$$L(s,\chi_0) = \zeta(s) \prod_{p \mid q} \left(1 - p^{-s}\right),$$

so $L(s,\chi_0)$ is holomorphic in $\mathbb{C} - \{1\}$ and has a simple pole at $s = 1$ with residue

$$\mathrm{Res}\left(L(s,\chi_0); 1\right) = \prod_{p \mid q} \left(1 - p^{-1}\right) = \frac{\phi(q)}{q}.$$

We now turn toward primitive characters.

**Lemma 3.8.** *Suppose that $\chi$ is a primitive Dirichlet character modulo $q$, $a \in \{0,1\}$, and define*

$$\theta_a(x;\chi) = \sum_{n=-\infty}^{\infty} n^a \chi(n) \exp\left(-\pi x n^2/q\right).$$

*Then for all $x > 0$,*

$$\tau(\overline{\chi}, 1) \theta_a(x^{-1};\chi) = (\mathrm{i}x)^a (qx)^{1/2} \theta_a(x;\overline{\chi}). \tag{3.15}$$

*Proof.* Suppose that $a = 0$. Because $\chi$ is primitive, we can use Lemma 3.5 to obtain

$$\begin{aligned}\tau(\overline{\chi},1)\theta_a(x^{-1};\chi) &= \sum_{n=-\infty}^{\infty} \tau(\overline{\chi},n)\exp\big(-\pi n^2/qx\big)\\ &= \sum_{k \bmod q} \overline{\chi}(k) \sum_{n=-\infty}^{\infty} \exp\big(-\pi n^2/qx\big)e(kn/q).\end{aligned}$$

Introducing the theta-series $\vartheta(z;\alpha)$ defined in (2.3), we can write this identity as

$$\tau(\overline{\chi},1)\theta_a(x^{-1};\chi) = \sum_{k \bmod q} \overline{\chi}(k)\exp\big(\pi k^2 x/q\big)\vartheta\big((qx)^{-1};-\mathrm{i}kx\big).$$

We now appeal to Lemma 2.2 and get

$$\begin{aligned}\tau(\overline{\chi},1)\theta_a(x^{-1};\chi) &= (qx)^{1/2} \sum_{k \bmod q} \overline{\chi}(k)\vartheta(qx;k/q)\\ &= (qx)^{1/2} \sum_{k \bmod q} \overline{\chi}(k) \sum_{n=-\infty}^{\infty} \exp\big(-\pi x(nq+k)^2/q\big)\\ &= (qx)^{1/2} \sum_{k \bmod q} \overline{\chi}(k) \sum_{m\equiv k \,(\mathrm{mod}\, q)} \exp\big(-\pi x m^2/q\big) = (qx)^{1/2}\theta_a(x;\overline{\chi}).\end{aligned}$$

The proof for $a = 1$ is similar, except that instead of Lemma 2.2 it uses the identity

$$\sum_{n=-\infty}^{\infty} (n+\alpha)e^{-\pi x(n+\alpha)^2} = -\mathrm{i}x^{-3/2} \sum_{n=-\infty}^{\infty} n\exp\left(-\pi n^2 x^{-1} + 2\pi\mathrm{i}\alpha n\right),$$

which follows from (2.4) via term-by-term differentiation with respect to $\alpha$. ∎

**Theorem 3.9.** *Suppose that $\chi$ is a primitive Dirichlet character modulo $q$ and choose $a \in \{0,1\}$ so that $\chi(-1) = (-1)^a$. Then the Dirichlet L-function $L(s,\chi)$ can be extended to an entire function satisfying the functional equation*

$$\left(\frac{q}{\pi}\right)^{(s+a)/2} \Gamma\left(\frac{s+a}{2}\right) L(s,\chi) = \frac{\mathrm{i}^a q^{1/2}}{\tau(\chi,1)} \left(\frac{q}{\pi}\right)^{(1-s+a)/2} \Gamma\left(\frac{1-s+a}{2}\right) L(1-s,\overline{\chi}). \tag{3.16}$$

*Proof.* As in the proof of Theorem 2.10, we start with a change of variables in the integral representation for the gamma-function:

$$\Gamma\left(\frac{s+a}{2}\right) = m^s \left(\frac{\pi}{q}\right)^{(s+a)/2} \int_0^\infty m^a \exp\big(-\pi x m^2/q\big) x^{(s+a)/2-1}\,dx.$$

Multiplying this identity by $(q/\pi)^{(s+a)/2}\chi(m)m^{-s}$ and then summing over $m$, we get

$$\begin{aligned}\left(\frac{q}{\pi}\right)^{(s+a)/2} \Gamma\left(\frac{s+a}{2}\right) L(s,\chi) &= \sum_{m=1}^{\infty} \chi(m) \int_0^\infty m^a \exp\big(-\pi x m^2/q\big) x^{(s+a)/2-1}\,dx\\ &= \frac{1}{2}\int_0^\infty \theta_a(x,\chi) x^{(s+a)/2-1}\,dx,\end{aligned} \tag{3.17}$$

since $m^a\chi(m)$ is an even function. By a change of variables and an appeal to Lemma 3.8,

$$\begin{aligned}\int_0^1 \theta_a(x,\chi)x^{(s+a)/2-1}\,dx &= \int_1^\infty \theta_a(t^{-1},\chi)t^{-(s+a)/2-1}\,dt \\ &= \frac{\mathrm{i}^a q^{1/2}}{\tau(\overline{\chi},1)}\int_1^\infty \theta_a(t,\overline{\chi})t^{-(s-a)/2-1/2}\,dt. \end{aligned} \tag{3.18}$$

Combining (3.17) and (3.18), we find that the left side of (3.16) equals

$$\frac{1}{2}\int_1^\infty \theta_a(x,\chi)x^{(s+a)/2-1}\,dx + \frac{\mathrm{i}^a q^{1/2}}{2\tau(\overline{\chi},1)}\int_1^\infty \theta_a(x,\overline{\chi})x^{-(s-a)/2-1/2}\,dx. \tag{3.19}$$

Since $\theta_a(x,\chi)$ decays exponentially for $x \to \infty$, this expression represents an entire function and thus provides an analytic continuation of $L(s,\chi)$ to $\mathbb{C}$. Furthermore, the substitution $s \mapsto 1-s$ transforms (3.19) into

$$\frac{1}{2}\int_1^\infty \theta_a(x,\chi)x^{-(s-a)/2-1/2}\,dx + \frac{\mathrm{i}^a q^{1/2}}{2\tau(\overline{\chi},1)}\int_1^\infty \theta_a(x,\overline{\chi})x^{(s+a)/2-1}\,dx. \tag{3.20}$$

Noting that when $\chi$ is primitive

$$\tau(\chi,1)\tau(\overline{\chi},1) = \chi(-1)q = (-1)^a q, \tag{3.21}$$

we see that (3.20) is equal to

$$\frac{\tau(\overline{\chi},1)}{\mathrm{i}^a q^{1/2}}\left\{\frac{1}{2}\int_1^\infty \theta_a(x,\overline{\chi})x^{(s+a)/2-1}\,dx + \frac{\mathrm{i}^a q^{1/2}}{2\tau(\chi,1)}\int_1^\infty \theta_a(x,\chi)x^{-(s-a)/2-1/2}\,dx\right\}.$$

This establishes the functional equation (3.16). ∎

## 3.3 The zeros of $L(s,\chi)$

Again, we want to use the logarithmic derivative $L'(s,\chi)/L(s,\chi)$, so we need first to study the zeros and the poles of the Dirichlet $L$-functions. As we already mentioned in the previous section, $L(s,\chi)$ is entire, unless $\chi$ is principal, in which case $L(s,\chi)$ has a single singularity—a simple pole at $s = 1$. We already know (from (3.14)) that $L(s,\chi)$ is non-zero in the half-plane $\mathrm{Re}(s) > 1$. Furthermore, using the functional equation (3.16), we can show that the only zeros of $L(s,\chi)$ in the half-plane $\mathrm{Re}(s) < 0$ are simple zeros at the even or odd integers, depending on the sign of $\chi(-1)$. Also, if $\chi$ is a non-principal character with $\chi(-1) = 1$, we see that $s = 0$ must be a zero. To study the zeros of $L(s,\chi)$ in the strip $0 \le \mathrm{Re}(s) \le 1$, we introduce the entire function

$$\xi(s,\chi) = \left(\frac{q}{\pi}\right)^{(s+a)/2}\Gamma\left(\frac{s+a}{2}\right)L(s,\chi),$$

which will play the same role the function $\xi(s)$ defined by (2.13) played in the study of the zeros of the zeta-function.

**Theorem 3.10.** *If $\chi$ is a Dirichlet character modulo $q$, then $L(1,\chi) \neq 0$.*

*Proof.* For principal characters the result is trivial (the $L$-function has a pole at $s = 1$), so we may assume that $\chi$ is non-principal. The case when $\chi$ is complex is easy. Suppose that $L(1,\chi) = 0$ for a complex character $\chi$. Then $\overline{\chi}$ is another character modulo $q$ with

$$L(1,\overline{\chi}) = \overline{L(1,\chi)} = 0.$$

Therefore, the product

$$f(s) = \prod_{\chi \bmod q} L(s,\chi)$$

represents an entire function, which vanishes at $s = 1$. On the other hand, when $\sigma > 1$, we have

$$\sum_{\chi \bmod q} \log L(\sigma,\chi) = \sum_{\chi \bmod q} \sum_{p} \sum_{m=1}^{\infty} \frac{\chi(p^m)}{mp^{m\sigma}} = \sum_{\substack{p \\ p^m \equiv 1 \pmod q}} \sum_{m=1}^{\infty} \frac{\phi(q)}{mp^{m\sigma}} \geq 0.$$

Hence, $f(\sigma) \geq 1$ for $\sigma > 1$, which is inconsistent with $f(1) = 0$. Thus, our assumption must be false.

To prove the lemma for a non-principal real character, we consider the function

$$f(s) = \frac{L(s,\chi)L(s,\chi_0)}{L(2s,\chi_0)},$$

where $\chi_0$ is the principal character modulo $q$. If $L(1,\chi) = 0$, this function is holomorphic in the half-plane $\mathrm{Re}(s) > 1/2$ and vanishes at $s = 1/2$ (since the denominator has a pole and the numerator is entire). On the other hand, when $\mathrm{Re}(s) > 1$, (3.14) yields

$$f(s) = \prod_{\chi(p)=1} \left( \frac{p^s + 1}{p^s - 1} \right) = \sum_{n=1}^{\infty} a_n n^{-s}.$$

Note that the coefficients $a_n$ are nonnegative. We now look at the Taylor expansion of $f(s)$ in $|s - 2| < 3/2$. We have

$$f(s) = \sum_{m=0}^{\infty} \frac{f^{(m)}(2)}{m!}(s-2)^m,$$

where from the Dirichlet series representation,

$$f^{(m)}(2) = (-1)^m \sum_{n=1}^{\infty} a_n (\log n)^m n^{-2} = (-1)^m b_m, \qquad \text{say.}$$

Thus,

$$f(s) = f(2) + \sum_{m=1}^{\infty} \frac{b_m}{m!}(2-s)^m,$$

with non-negative coefficients $b_m$. Letting $s \to 1/2$, we find that all the coefficients on the right are zero (since $f(1/2) = 0$), and in particular, that $f(2) = 0$. This, however, contradicts the Euler product representation of $f(s)$. ∎

We now commence our investigation of the zeros of $\xi(s,\chi)$. The following two results are analogues of Lemma 2.11 and Theorem 2.17.

**Lemma 3.11.** *Suppose that $\chi$ is a primitive character modulo $q$. There is an absolute constant $c_1 > 0$ such that*

$$|\xi(s,\chi)| \ll e^{c_1|s|\log(q|s|)}. \tag{3.22}$$

*Proof.* Because $\xi(s,\chi)$ satisfies the functional equation

$$\xi(s,\chi) = w(\chi)\xi(1-s,\overline{\chi}), \qquad |w(\chi)| = 1,$$

it suffices to prove (3.22) when $\mathrm{Re}(s) \geq 1/2$. For such $s$, the desired bound follows from the definition of $\xi(s,\chi)$, Stirling's formula, and the estimate

$$|L(s,\chi)| \ll q|s|.$$

■

**Theorem 3.12.** *Suppose that $\chi$ is a primitive character modulo $q$. Then the function $\xi(s,\chi)$ has infinitely many zeros in the strip $0 \leq \mathrm{Re}(s) \leq 1$ and can be written as*

$$\xi(s,\chi) = e^{A+Bs}\prod_{\rho}\left(1-\frac{s}{\rho}\right)e^{s/\rho}.$$

*Here $A = A(\chi)$ and $B = B(\chi)$ are constants depending only on the character $\chi$ and the product is over the zeros of $\xi(s,\chi)$ listed according to their multiplicities.*

**Corollary 3.13.** *Suppose that $\chi$ is a primitive character modulo $q$ and $a \in \{0,1\}$ is such that $\chi(-1) = (-1)^a$. Then*

$$\frac{L'(s,\chi)}{L(s,\chi)} = B(\chi) - \frac{1}{2}\log(q/\pi) - \frac{1}{2}\frac{\Gamma'((s+a)/2)}{\Gamma((s+a)/2)} + \sum_{\rho}\left(\frac{1}{s-\rho}+\frac{1}{\rho}\right). \tag{3.23}$$

*Here $B(\chi)$ is the constant appearing in Theorem 3.12.*

*Proof.* This follows from Theorem 3.12 by logarithmic differentiation. ■

**Corollary 3.14.** *The constant $B(\chi)$ satisfies*

$$\mathrm{Re}\, B(\chi) = -\,\mathrm{Re}\sum_{\rho}\frac{1}{\rho}. \tag{3.24}$$

*Proof.* We first observe that, by the functional equation of $\xi(s,\chi)$, if $\rho$ is a zero of $\xi(s,\chi)$, then so is $1-\overline{\rho}$, while $\overline{\rho}$ and $1-\rho$ are zeros of $\xi(s,\overline{\chi})$. From Theorem 3.12,

$$\frac{\xi'(s,\chi)}{\xi(s,\chi)} = B(\chi) + \sum_{\rho}\left(\frac{1}{s-\rho}+\frac{1}{\rho}\right), \tag{3.25}$$

and from the functional equation,

$$\frac{\xi'(s,\chi)}{\xi(s,\chi)} = -\frac{\xi'(1-s,\overline{\chi})}{\xi(1-s,\overline{\chi})}.$$

Substituting $s = 0$, we get

$$B(\chi) = \frac{\xi'(0,\chi)}{\xi(0,\chi)} = -\frac{\xi'(1,\overline{\chi})}{\xi(1,\overline{\chi})} = -B(\overline{\chi}) - \sum_{\rho}\left(\frac{1}{1-\overline{\rho}} + \frac{1}{\overline{\rho}}\right). \tag{3.26}$$

Note that since $\xi(\overline{s},\overline{\chi}) = \overline{\xi(s,\chi)}$, (3.25) implies $B(\overline{\chi}) = \overline{B(\chi)}$. Thus, by (3.25) and our starting remark,

$$2\,\mathrm{Re}\,B(\chi) = B(\chi) + B(\overline{\chi}) = -\sum_{\rho}\left(\frac{1}{\rho} + \frac{1}{\overline{\rho}}\right) = -2\,\mathrm{Re}\sum_{\rho}\frac{1}{\rho}.$$

■

**Theorem 3.15.** *Suppose that $\chi$ is a primitive character modulo $q$ and $s = \sigma + \mathrm{i}t$, $-1/2 \le \sigma \le 2$. Then*

$$\frac{L'(s,\chi)}{L(s,\chi)} = \frac{-1}{s+a} + \sum_{|\mathrm{Im}\,\rho - t| \le 1}\frac{1}{s-\rho} + O(\log q(|t|+2)), \tag{3.27}$$

*where $a \in \{0,1\}$ is such that $\chi(-1) = (-1)^a$.*

*Proof.* We write $\tau = q(|t|+2)$. The starting point is (3.23). The term involving the gamma-function can be estimated as

$$\frac{1}{s+a} + O(\log\tau).$$

Thus, (3.23) may be rewritten as

$$\frac{L'(s,\chi)}{L(s,\chi)} = B(\chi) - \frac{1}{s+a} + \sum_{\rho}\left(\frac{1}{s-\rho} + \frac{1}{\rho}\right) + O(\log\tau). \tag{3.28}$$

We view this approximate equation as an analogue of (2.22) and want to deduce from it an analogue of (2.21). We let $s = 2 + \mathrm{i}t$ and take real parts. Then the left side of (3.28) is bounded, so using (3.24), we obtain

$$\mathrm{Re}\sum_{\rho}\frac{1}{2+\mathrm{i}t-\rho} \ll \log\tau.$$

Since $\rho = \beta + \mathrm{i}\gamma$, $0 \le \beta \le 1$, we have

$$\mathrm{Re}\,\frac{1}{2+\mathrm{i}t-\rho} = \frac{2-\beta}{(2-\beta)^2 + (t-\gamma)^2} \gg \frac{1}{1+(t-\gamma)^2},$$

whence

$$\sum_{\rho}\frac{1}{1+(t-\gamma)^2} \ll \log\tau. \tag{3.29}$$

Subtracting from (3.28) the corresponding equation with $s = 2 + it$, we deduce that

$$\frac{L'(s,\chi)}{L(s,\chi)} = \frac{-1}{s+a} + \sum_{\rho}\left(\frac{1}{s-\rho} - \frac{1}{2+it-\rho}\right) + O(\log\tau).$$

When $|\gamma - t| > 1$,

$$\left|\frac{1}{s-\rho} - \frac{1}{2+it-\rho}\right| \ll \frac{1}{(t-\gamma)^2}.$$

Hence, in view of (3.29),

$$\frac{L'(s,\chi)}{L(s,\chi)} = \frac{-1}{s+a} + \sum_{|\gamma-t|\le 1}\left(\frac{1}{s-\rho} - \frac{1}{2+it-\rho}\right) + O(\log\tau).$$

Using (3.29) once more, we see that the terms $(2 + it - \rho)^{-1}$ are also superfluous, and so (3.27) follows from the last equation. ∎

From (3.29), we obtain the following result.

**Corollary 3.16.** *Suppose that $T \ge 2$. The number of zeros of $L(s,\chi)$ in the region*

$$0 \le \operatorname{Re} s \le 1, \qquad T \le |\operatorname{Im} s| \le T+1$$

*is $O(\log qT)$.*

**Theorem 3.17.** *Suppose that $\chi$ is a primitive character modulo $q$. There exists an absolute constant $c_2 > 0$ such that at most one zero $\rho = \beta + i\gamma$ of the function $L(s,\chi)$ does not satisfy*

$$\beta \le 1 - \frac{c_2}{\log q(|\gamma|+2)}.$$

*If such a zero does exist, the character $\chi$ must be real and the zero itself must be simple and real.*

*Proof.* As in the proof of Theorem 2.21, we deduce from (2.27) that

$$-3\frac{L'(\sigma,\chi_0)}{L(\sigma,\chi_0)} - 4\operatorname{Re}\frac{L'(\sigma+it,\chi)}{L(\sigma+it,\chi)} - \operatorname{Re}\frac{L'(\sigma+2it,\chi^2)}{L(\sigma+2it,\chi^2)} \ge 0. \tag{3.30}$$

Here, $\sigma > 1$ and $\chi_0$ is the principal character modulo $q$. Let $\rho = \beta_0 + i\gamma_0$ be a particular zero of $L(s,\chi)$ and write $\tau = q(|\gamma_0| + 2)$. We now estimate the left side of (3.30) when $t = \gamma_0$. From Theorem 3.15,

$$-\operatorname{Re}\frac{L'(\sigma+i\gamma_0,\chi)}{L(\sigma+i\gamma_0,\chi)} \le -\operatorname{Re}\sum_{|\operatorname{Im}\rho-\gamma_0|\le 1}\frac{1}{\sigma+i\gamma_0-\rho} + O(\log\tau) \le \frac{-1}{\sigma-\beta_0} + O(\log\tau). \tag{3.31}$$

To estimate the term involving $\chi^2$, we observe that if $\chi^2$ is induced by a character $\chi_1$ modulo $q_1$, then

$$\frac{L'(s,\chi^2)}{L(s,\chi^2)} - \frac{L'(s,\chi_1)}{L(s,\chi_1)} \ll \sum_{(m,q)>1}\Lambda(m)m^{-\sigma} \ll \sum_{p|q}\log p\left(p^{-\sigma} + p^{-2\sigma} + \cdots\right) \ll \log q. \tag{3.32}$$

Moreover, a similar relation holds for $L'(s,\chi_0)/L(s,\chi_0)$ and $\zeta'(s)/\zeta(s)$. In particular, combining that relation with (2.29), we obtain

$$-\frac{L'(\sigma,\chi_0)}{L(\sigma,\chi_0)} \le \frac{1}{\sigma-1} + O(\log q). \tag{3.33}$$

When $\chi^2$ is non-principal, (3.27) and (3.32) give

$$-\operatorname{Re}\frac{L'(\sigma+2i\gamma_0,\chi^2)}{L(\sigma+2i\gamma_0,\chi^2)} \ll \log\tau. \tag{3.34}$$

Combining (3.30), (3.31), (3.33), and (3.34), we deduce that

$$4(\sigma-\beta_0)^{-1} \le 3(\sigma-1)^{-1} + c_3\log\tau. \tag{3.35}$$

On choosing $\sigma = 1+(2c_3\log\tau)^{-1}$, this establishes the theorem in the case of complex characters.

We now turn to real characters $\chi$ (so that $\chi^2=\chi_0$). In this case, we replace (3.34) by

$$-\operatorname{Re}\frac{L'(\sigma+2i\gamma_0,\chi_0)}{L(\sigma+2i\gamma_0,\chi_0)} \le \frac{\sigma-1}{(\sigma-1)^2+4\gamma_0^2} + O(\log\tau), \tag{3.36}$$

the extra term accounting for the pole at $s=1$. Accordingly, (3.35) becomes

$$\frac{4}{\sigma-\beta_0} \le \frac{3}{\sigma-1} + \frac{\sigma-1}{(\sigma-1)^2+4\gamma_0^2} + c_4\log\tau.$$

Thus, if $|\gamma_0| \ge \delta(\log\tau)^{-1}$, the desired conclusion follows on choosing

$$\sigma = 1+c_5(\log\tau)^{-1}, \qquad 0<c_5<\min\left((4c_4)^{-1}, 4c_4\delta^2\right).$$

Finally, suppose that $|\gamma_0| \le \delta(\log q)^{-1}$. For $\sigma>1$, Theorem 3.15 yields

$$-\frac{L'(\sigma,\chi)}{L(\sigma,\chi)} \le \sum_{|\operatorname{Im}\rho|\le 1}\frac{-1}{\sigma-\rho} + c_6\log q. \tag{3.37}$$

(Note that for a real character $\chi$, $\rho$ and $\bar\rho$ are both zeros of $L(s,\chi)$, so the sum on the right is real.) On the other hand,

$$-\frac{L'(\sigma,\chi)}{L(\sigma,\chi)} = \sum_{n=1}^{\infty}\Lambda(n)\chi(n)n^{-\sigma} \ge -\sum_{n=1}^{\infty}\Lambda(n)n^{-\sigma} = \frac{\zeta'(\sigma)}{\zeta(\sigma)}. \tag{3.38}$$

Combining (2.29), (3.37), and (3.38), we obtain

$$\sum_{|\operatorname{Im}\rho|\le 1}\frac{1}{\sigma-\rho} \le \frac{1}{\sigma-1} + c_7\log q. \tag{3.39}$$

We now assume, as we may, that $\delta < (10c_7)^{-1}$ and set $\sigma = 1 + 2\delta(\log q)^{-1}$. If $\rho = \beta_0 + i\gamma_0$ is a complex zero or a double real zero, we estimate the left side of (3.39) from below by the contribution from $\rho$ and $\overline{\rho}$ or by the doubled contribution from $\rho$. We get

$$\frac{2(\sigma-\beta_0)}{(\sigma-\beta_0)^2+\gamma_0^2} \le 0.6\delta^{-1}\log q. \tag{3.40}$$

Also, by our choices,

$$\frac{\sigma-\beta_0}{(\sigma-\beta_0)^2+\gamma_0^2} \ge \frac{\sigma-\beta_0}{(\sigma-\beta_0)^2+\delta^2(\log q)^{-2}} \ge \frac{0.8}{\sigma-\beta_0}.$$

Combining this inequality with (3.40), we obtain

$$1.6(\sigma-\beta_0)^{-1} \le 0.6\delta^{-1}\log q \qquad \Rightarrow \qquad \beta_0 < 1 - 0.5\delta(\log q)^{-1}.$$

Finally, if $\beta_0$ is a real zero and $\beta_1$ is another real zero, we replace (3.40) by

$$(\sigma-\beta_0)^{-1} + (\sigma-\beta_1)^{-1} \le 0.6\delta^{-1}\log q,$$

whence

$$\min(\beta_0,\beta_1) \le 1 - \delta(\log q)^{-1}.$$

■

## 3.4 The exceptional zero

The possible real zero appearing in Theorem 3.17 is known as an *exceptional zero*, a *Siegel zero*, or a *Siegel–Landau zero*. The purpose of this section is to show that such zeros cannot lie too close to 1. First, we prove that if two $L$-functions both have exceptional zeros, one of them must have a modulus that is much larger than the modulus of the other.

**Theorem 3.18 (Landau).** *Let $\chi_1$ and $\chi_2$ be distinct primitive real characters modulo $q_1$ and $q_2$, respectively. Suppose that $\beta_1$ and $\beta_2$ are real numbers such that*

$$L(\beta_1,\chi_1) = L(\beta_2,\chi_2) = 0.$$

*There exists an absolute constant $c_8 > 0$ such that*

$$\min(\beta_1,\beta_2) \le 1 - c_8(\log q_1q_2)^{-1}.$$

*Proof.* Using the inequality

$$(1+\chi_1(m))(1+\chi_2(m)) \ge 0,$$

we find that

$$-\frac{\zeta'(\sigma)}{\zeta(\sigma)} - \frac{L'(\sigma,\chi_1)}{L(\sigma,\chi_1)} - \frac{L'(\sigma,\chi_2)}{L(\sigma,\chi_2)} - \frac{L'(\sigma,\chi_1\chi_2)}{L(\sigma,\chi_1\chi_2)} \ge 0. \tag{3.41}$$

Since $\chi_1$ and $\chi_2$ are distinct, $\chi_1\chi_2$ is non-principal and we can deduce the theorem from (3.41) by referring to (2.29), (3.31), and an obvious analogue of (3.34). ■

**Corollary 3.19.** *Let $Q > 1$. There is an absolute constant $c_9 > 0$ such that no L-function $L(s,\chi)$ modulo $q$, $q \le Q$, has a zero $\rho = \beta + i\gamma$ with*

$$\beta < 1 - c_9(\log Q(|\gamma| + 2))^{-1}, \tag{3.42}$$

*except possibly at a point $\beta_0$ on the real axis, where $L(s,\chi)$ may have a simple zero. Furthermore, any character $\chi$ modulo $q$, $q \le Q$, for which this zero does occur is real and induced by the same primitive real character.*

So far, we know from Theorem 3.10 that the exceptional zero (if it exists) is less than 1, but we have no quantitative form of this result. By a classical result of Dirichlet—the analytic class number formula (see Davenport [14, §6, (15)]), for a primitive quadratic character $\chi$ modulo $q$, we have

$$L(1,\chi) = C(\chi)h(q)q^{-1/2},$$

where $C(\chi) \ge 1$ and $h(q)$ is a positive integer (the number of "classes" of binary quadratic forms of discriminant $q$). Using the trivial observation that $h(q) \ge 1$, we conclude that, in fact, we may strenghten Theorem 3.10 to

$$L(1,\chi) \gg q^{-1/2},$$

which in turn leads to the following bound for the exceptional zero $\beta_0$:

$$\beta_0 \le 1 - c_{10}q^{-1/2}(\log q)^{-2}. \tag{3.43}$$

The following remarkable result of Siegel's provides a much stronger bound on $\beta_0$.

**Theorem 3.20 (Siegel).** *Let $\epsilon > 0$ be fixed. There is a constant $c_0(\epsilon) > 0$ such that if $\chi$ is a primitive real character modulo $q$, then $L(s,\chi) \ne 0$ in the region*

$$|\operatorname{Im}(s)| \le 1, \qquad \operatorname{Re}(s) \ge 1 - c_0(\epsilon)q^{-\epsilon}. \tag{3.44}$$

*Proof.* We consider primitive real characters $\chi_1$ and $\chi_2$ with moduli $q_1$ and $q_2$, respectively, and introduce the function

$$F(s) = \zeta(s)L(s,\chi_1)L(s,\chi_2)L(s,\chi_1\chi_2). \tag{3.45}$$

Since $\chi_1\chi_2$ is a non-principal (though not necessarily primitive) character modulo $q_1q_2$, $F(s)$ is holomorphic everywhere except at $s = 1$, where it has a simple pole with residue

$$\lambda = L(1,\chi_1)L(1,\chi_2)L(1,\chi_1\chi_2).$$

We now proceed to show that

$$F(\sigma) > 1/2 - c_{11}\lambda(q_1q_2)^{8(1-\sigma)}(1-\sigma)^{-1} \qquad \text{for } 7/8 < \sigma < 1. \tag{3.46}$$

When $\operatorname{Re}(s) > 1$, we can write $F(s)$ as a Dirichlet series

$$F(s) = \sum_{n=1}^{\infty} a_n n^{-s}.$$

It follows from the Euler product representations of the factors in (3.45) that $a_1 = 1$ and that $a_n \geq 0$ for all $n = 1, 2, \ldots$. When $|s-2| < 1$, $F(s)$ has also a Taylor expansion

$$F(s) = \sum_{n=0}^{\infty} b_n (2-s)^n,$$

where

$$b_n = (-1)^n F^{(n)}(2)/n! = \sum_{m=1}^{\infty} a_m (\log m)^n m^{-2} \geq 0.$$

Hence,

$$G(s) = F(s) - \lambda(s-1)^{-1} = \sum_{n=0}^{\infty} (b_n - \lambda)(2-s)^n, \tag{3.47}$$

and this representation is, in fact, valid in the larger disk $|s-2| \leq 3/2$, because $G(s)$ is an entire function. We now estimate the coefficients of the series in (3.47). On the circle $C : |s-2| = 3/2$, we have

$$|\zeta(s)| \ll 1, \qquad |s-1|^{-1} \ll 1, \qquad \text{and} \qquad |L(s,\chi)| \ll q$$

for any non-principal character $\chi$ modulo $q$. Thus,

$$G(s) \ll (q_1 q_2)^2 \qquad (s \in C).$$

We now use Cauchy's formula for the Taylor coefficients. Integrating along $C$, we find that

$$|b_n - \lambda| = \left| \frac{1}{2\pi \mathrm{i}} \int_C \frac{G(s)}{(s-2)^{n+1}}\, ds \right| \ll \left(\frac{2}{3}\right)^n (q_1 q_2)^2.$$

When $N > 1$ and $7/8 \leq \sigma \leq 1$, this gives

$$\sum_{n=N}^{\infty} |b_n - \lambda|(2-\sigma)^n \ll (q_1 q_2)^2 (3/4)^N \ll (q_1 q_2)^2 e^{-N/4}.$$

Hence,

$$\begin{aligned} F(\sigma) - \lambda(\sigma-1)^{-1} &\geq \sum_{n=0}^{N-1} (b_n - \lambda)(2-\sigma)^n - c_{12}(q_1 q_2)^2 e^{-N/4} \\ &\geq 1 - \lambda \sum_{n=0}^{N-1} (2-\sigma)^n - c_{12}(q_1 q_2)^2 e^{-N/4}, \end{aligned}$$

upon noting that $b_0 = F(2) \geq 1$. Thus, choosing $N$ so that

$$c_{12}(q_1 q_2)^2 e^{-N/4} < 1/2 \leq c_{12}(q_1 q_2)^2 e^{-(N-1)/4},$$

we obtain

$$F(\sigma) \geq 1/2 - \lambda(2-\sigma)^N (1-\sigma)^{-1},$$

and (3.46) follows from the inequality

$$(2-\sigma)^N \le \exp\big(N(1-\sigma)\big) \le c_{13}(q_1q_2)^{8(1-\sigma)}.$$

We may assume that $L(s,\chi)$ has a real zero $\beta_0 \ge 1-(\log q)^{-1}$ and that there is some primitive character $\chi_2$ such that $L(s,\chi_2)$ has a real zero $\beta_2$ with

$$1-\epsilon/10 \le \beta_2 < 1.$$

We then consider $F(s)$ with $\chi_1 = \chi$ and $\chi_2$ being this special character. Since $F(\beta_2) = 0$, (3.46) yields

$$\lambda q^{8(1-\beta_2)} \gg_\epsilon 1; \tag{3.48}$$

here the constant depends on the choice of $\chi_2$, and hence, on $\epsilon$. By the bounds in Exercise 12,

$$\lambda \ll_\epsilon L(1,\chi)\log q. \tag{3.49}$$

Furthermore, Lagrange's mean-value theorem and another appeal to Exercise 12 give

$$L(1,\chi) = L(1,\chi) - L(\beta_0,\chi) = (1-\beta_0)L'(\sigma,\chi) \ll (1-\beta_0)(\log q)^2,$$

for some $\sigma \in (\beta_0, 1)$. Combining this estimate, (3.48) and (3.49), we obtain

$$1 \ll_\epsilon (1-\beta_0)q^{8(1-\beta_2)}(\log q)^3 \ll_\epsilon (1-\beta_0)q^\epsilon,$$

and (3.44) follows. ∎

**Remark.** Theorem 3.20 is *ineffective*. That is, given $\epsilon > 0$, the proof does not allow us to calculate the constant $c_0(\epsilon)$. Indeed, in the above proof, we essentially used a possible counterexample to a strong conjecture[2] (i.e., $\beta_2$) to show that any possible counterexample to a weaker conjecture does not fail that conjecture too miserably. In particular, in order to compute $c_0(\epsilon)$, we must exhibit a particular character $\chi_2$ as in the proof of Siegel's theorem. Of course, if GRH is true—as is the popular belief—neither $\beta_0$ nor $\beta_2$ exist and we will never find an actual character $\chi_2$ that we can use to calculate $c_0(\epsilon)$.

## 3.5 The prime number theorem for arithmetic progressions

For a Dirichlet character $\chi$, define

$$\psi(x,\chi) = \sum_{n\le x}\Lambda(n)\chi(n). \tag{3.50}$$

The next theorem is an analogue of Theorem 2.22.

[2]Namely, that all real zeros $\beta$ of $L$-functions with real characters satisfy $\beta \le 1-\epsilon/10$.

**Theorem 3.21.** *Suppose that $\chi$ is a non-principal character modulo $q$ and $2 \le T \le x$, where $\{x\} = 1/2$. Then*

$$\psi(x,\chi) = -\sum_{|\operatorname{Im}\rho|\le T} \frac{x^\rho}{\rho} + O\left(\left(xT^{-1} + q^{1/2}\right)(\log qx)^2\right), \tag{3.51}$$

*where the summation is over the nontrivial zeros $\rho$ of $L(s,\chi)$ with $|\operatorname{Im}\rho| \le T$.*

*Proof.* First, suppose that $\chi$ is primitive. As in the proof of Theorem 2.22, we set $\alpha = 1+(\log qx)^{-1}$ and apply Corollary 1.14 with $f(s) = -L'(s,\chi)/L(s,\chi)$. Then an argument similar to that leading to (2.33) yields

$$\psi(x,\chi) = \frac{1}{2\pi\mathrm{i}} \int_{\alpha-\mathrm{i}T}^{\alpha+\mathrm{i}T} \left(-\frac{L'(s,\chi)}{L(s,\chi)}\right) \frac{x^s}{s}\, ds + O\left(\frac{x(\log qx)^2}{T}\right). \tag{3.52}$$

Because of Corollary 3.16, we may assume that $T$ is chosen so that

$$\left|\,|\operatorname{Im}\rho| - T\right| \gg (\log qT)^{-1} \qquad \text{whenever } L(\rho,\chi) = 0. \tag{3.53}$$

It then follows from Theorem 3.15 and Corollary 3.16 that the inequality

$$\left|\frac{L'(s,\chi)}{L(s,\chi)}\right| \ll (\log qT)^2$$

holds on the contour $C$ shown on Fig. 2.1. Thus,

$$\int_C \left(-\frac{L'(s,\chi)}{L(s,\chi)}\right) \frac{x^s}{s}\, ds \ll \frac{x(\log qx)^2}{T}, \tag{3.54}$$

the details being similar to those in the proof of the respective bound in the proof of Theorem 2.22. Combining (3.52) and (3.54), we find that

$$\psi(x,\chi) = \Sigma + O\left(\frac{x(\log qx)^2}{T}\right),$$

where $\Sigma$ is the sum of the residues of the function

$$\left(-\frac{L'(s,\chi)}{L(s,\chi)}\right) \frac{x^s}{s}$$

at its poles lying between $C$ and the vertical line $\operatorname{Re}(s) = \alpha$. This function has simple poles at the zeros of $L(s,\chi)$ in the critical strip and a simple or double pole at $s = 0$ (according as $L(0,\chi) \ne 0$ or $L(0,\chi) = 0$). Hence,

$$\psi(x,\chi) = -\sum_{|\operatorname{Im}\rho|\le T} \frac{x^\rho}{\rho} + C(\chi) + O\left(\frac{x(\log qx)^2}{T}\right), \tag{3.55}$$

where $C(\chi)$ is the residue at 0. It remains to estimate $C(\chi)$.

Let $b(\chi)$ be the constant term in the Laurent expansion of $L'(s,\chi)/L(s,\chi)$ near $s$. Then

$$C(\chi) = \begin{cases} b(\chi) & \text{if } L(0,\chi) \neq 0, \\ b(\chi) + \log x & \text{if } L(0,\chi) = 0. \end{cases}$$

From (3.27),

$$b(\chi) = \sum_{|\rho|\leq 1} \frac{-1}{\rho} + O(\log q).$$

By Theorem 3.17, the last sum contains $O(\log q)$ terms and each of them, except for possibly one, is $O(\log q)$. The exceptional term occurs only when $\chi$ is a real character such that $L(s,\chi)$ has an exceptional zero; furthermore, by (3.43), the exceptional term is $\ll q^{1/2}(\log q)^2$. Altogether, we have

$$C(\chi) \ll q^{1/2}(\log qx)^2.$$

The desired result follows from this estimate and (3.55).

Finally, we remove the restriction to primitive characters. Suppose that $\chi$ is a character modulo $q$ induced by a primitive character $\chi^*$ modulo $r$, $1 < r < q$. Then

$$\psi(x,\chi) - \psi(x,\chi^*) \leq \sum_{\substack{n\leq x \\ (n,q)>1}} \Lambda(n) \leq \sum_{p|q} (\log p) \sum_{\substack{k: \\ p^k \leq x}} 1 \ll (\log qx)^2. \tag{3.56}$$

Hence,

$$\psi(x,\chi) = -\sum_{|\operatorname{Im}\rho|\leq T} \frac{x^\rho}{\rho} + O\left(\left(xT^{-1} + r^{1/2}\right)(\log qx)^2\right),$$

where the summation is over the nontrivial zeros of $L(s,\chi^*)$. Thus, (3.51) follows on noting that $L(s,\chi)$ and $L(s,\chi^*)$ have the same nontrivial zeros. ∎

**Theorem 3.22.** *Suppose that $x \geq 2$, $q \geq 1$, and $(a,q) = 1$. There is an absolute constant $c_{14} > 0$ such that*

$$\psi(x;q,a) = \frac{x}{\phi(q)} - \delta_q \frac{\chi_1(a)}{\phi(q)} \frac{x^{\beta_1}}{\beta_1} + O\left(x \exp\left(-c_{14}\sqrt{\log x}\right)\right), \tag{3.57}$$

*where $\delta_q = 1$ if there is a real character $\chi_1$ modulo $q$ such that $L(s,\chi_1)$ has an exceptional zero $\beta_1$ and $\delta_q = 0$ otherwise.*

*Proof.* We assume, as we may, that $c_{14} < 1/2$. It suffices to consider the case

$$1 \leq q \leq \exp\left(\sqrt{\log x}\right),$$

for otherwise (3.57) is trivial. By (3.5),

$$\psi(x;q,a) = \frac{1}{\phi(q)} \sum_{\chi \bmod q} \overline{\chi}(a)\psi(x,\chi). \tag{3.58}$$

Suppose that $\chi$ is a nonprincipal character modulo $q$ and set

$$T = \exp\big(\sqrt{\log x}\big), \qquad \delta(T) = c_9(2\log T)^{-1}.$$

By Corollary 3.19,

$$\operatorname{Re}\rho \le 1 - c_9(\log qT)^{-1} \le 1 - \delta(T)$$

for all zeros of $L(s,\chi)$, except possibly for the exceptional zero $\beta_1$. Thus, (3.51) yields

$$\psi(x,\chi) = -\delta(\chi)\frac{x^{\beta_1}}{\beta_1} + O\bigg(x^{1-\delta(T)}\sum_{|\operatorname{Im}\rho|\le T}\frac{1}{|\rho|}\bigg) + O\big(xT^{-1}(\log x)^2\big),$$

where $\delta(\chi)$ is 1 or 0 according as $\chi = \chi_1$ or not. Using Corollary 3.16 to bound the last sum, we deduce that

$$\sum_{\chi \bmod q}{}' \bar{\chi}(a)\psi(x,\chi) = -\delta_q\chi_1(a)\frac{x^{\beta_1}}{\beta_1} + O\big(\phi(q)x\exp\big(-c_{15}\sqrt{\log x}\big)\big), \tag{3.59}$$

where the summation on the right side is restricted to the non-principal characters modulo $q$. Furthermore, by a variant of (3.56) and Theorem 2.1,

$$\psi(x,\chi_0) = \psi(x) + O\big((\log x)^2\big) = x + \big(x\exp\big(-c_{16}\sqrt{\log x}\big)\big) \tag{3.60}$$

Clearly, the desired result follows from (3.58)–(3.60). ∎

Note that the constant $c_{14}$ is effective, but Theorem 3.22 itself is not, since in general we do not know whether the exceptional zero exists.

*Proof of Theorem 2.* We now deduce Theorem 2 from Theorem 3.22. We use Siegel's theorem with $\epsilon = (2A)^{-1}$. It gives

$$\beta_1 \le 1 - c_0(A)q^{-1/(2A)} \le 1 - c_0(A)(\log x)^{-1/2}, \tag{3.61}$$

whence

$$x^{\beta_1}/\beta_1 \ll x\exp\big(-c_1(A)\sqrt{\log x}\big).$$

Therefore, even if there is an exceptional zero, the corresponding term on the right side of (3.57) is superfluous and

$$\psi(x;q,a) = \frac{x}{\phi(q)} + O\big(x\exp\big(-c_2(A)\sqrt{\log x}\big)\big). \tag{3.62}$$

Theorem 2 now follows by partial summation. ∎

**Remark.** Observe that the constants $c_0(A), c_1(A), \dots$ above depend on the constant $c_0(\epsilon)$ in Theorem 3.20 and are therefore ineffective.

## Exercises

1. Suppose that $q \geq 1$ is an integer and $f : \mathbb{Z} \to \mathbb{C}$ is a nontrivial, $q$-periodic, completely multiplicative function such that $f(n) = 0$ whenever $\gcd(n, q) > 1$. Prove that $f$ is a Dirichlet character modulo $q$.

2. Suppose that the character $\chi$ modulo $q$ is induced by the characters $\chi_1$ modulo $q_1$ and $\chi_2$ modulo $q_2$. Write $q_3 = \gcd(q_1, q_2)$. Prove that $\chi$ is induced also by a character $\chi_3$ modulo $q_3$.

3. Suppose that $\gcd(q_1, q_2) = 1$, $\chi_1$ is a character modulo $q_1$, and $\chi_2$ is a character modulo $q_2$. Prove that the character $\chi_1\chi_2$ modulo $q_1q_2$ is primitive if and only if both $\chi_1$ and $\chi_2$ are primitive.

4. (a) Suppose that $p$ is an odd prime. Prove that there is no primitive real character modulo $p^\alpha$, $\alpha \geq 2$, and that the only primitive real character modulo $p$ is the Legendre symbol

$$\left(\frac{n}{p}\right) = \begin{cases} +1 & \text{if } n \equiv \square \pmod{p}, \\ -1 & \text{if } n \not\equiv \square \pmod{p}. \end{cases}$$

   (b) Prove that there is no primitive real character modulo $2^\alpha$, $\alpha \geq 4$.

   (c) Prove that the only primitive real character modulo 4 is the character $\chi_4$ given by

$$\chi_4(n) = \begin{cases} +1 & \text{if } n \equiv 1 \pmod 4, \\ -1 & \text{if } n \equiv 3 \pmod 4. \end{cases}$$

   (d) Prove that the only primitive real characters modulo 8 are $\chi_8$ and $\chi_4\chi_8$, where $\chi_8$ is given by

$$\chi_8(n) = \begin{cases} +1 & \text{if } n \equiv \pm 1 \pmod 8, \\ -1 & \text{if } n \equiv \pm 3 \pmod 8. \end{cases}$$

5. Suppose that $m$ is a positive integer and define the exponential sum

$$G(m) = \sum_{n=1}^{m} e\big(n^2/m\big).$$

   (a) Define the function $f : \mathbb{R} \to \mathbb{C}$ by

$$f(x) = \frac{f_0(x-0) + f_0(x+0)}{2}, \qquad f_0(x) = \begin{cases} e\big(x^2/m\big) & \text{if } 0 \leq x \leq m, \\ 0 & \text{otherwise.} \end{cases}$$

   We can apply to $f$ the Poisson summation formula (see Zygmund [59, eq. (II.13.4)]):

$$\sum_{n\in\mathbb{Z}} f(n) = \sum_{n\in\mathbb{Z}} \hat{f}(n), \qquad \hat{f}(t) = \int_{\mathbb{R}} f(x)e(-xt)\,dx.$$

   Use this to prove that

$$G(m) = m \sum_{n\in\mathbb{Z}} e\big(-mn^2/4\big) \int_{-n/2}^{-n/2+1} e\big(my^2\big)\,dy.$$

   (b) Using the result of (a), show that

$$G(m) = C\big(1 + \mathrm{i}^{-m}\big)\sqrt{m}, \qquad \text{where } C = \int_{-\infty}^{\infty} e\big(t^2\big)\,dt.$$

(c) Deduce that

$$G(m) = \frac{1 + \mathrm{i}^{-m}}{1 + \mathrm{i}^{-1}} \sqrt{m}.$$

(d) Deduce the formula for the Fresnel integrals:

$$\int_0^\infty \cos\left(x^2\right) dx = \int_0^\infty \sin\left(x^2\right) dx = \sqrt{\frac{\pi}{2}}.$$

6. The purpose of this problem is to establish the law of quadratic reciprocity:

$$\left(\frac{p}{q}\right)\left(\frac{q}{p}\right) = (-1)^{(p-1)(q-1)/4} \tag{$*$}$$

for all pairs of distinct odd primes $p, q$.

(a) Let $G(m)$ be the exponential sum defined in the last problem. Prove that

$$G(pq) = \left(\frac{p}{q}\right)\left(\frac{q}{p}\right) G(p)G(q).$$

(b) Deduce $(*)$ from part (a) and the explicit formula for $G(m)$.

7. Suppose that $\chi$ is a character modulo $q$ induced by a primitive character $\chi^*$ modulo $q^*$. Suppose also that $a$ is an integer and write $q_1 = q/(a,q)$, $a_1 = a/(a,q)$. Prove that:

(a) If $q^* \nmid q_1$, then $\tau(\chi, a) = 0$.

(b) If $q^* \mid q_1$, then

$$\tau(\chi, a) = \mu\left(\frac{q_1}{q^*}\right)\chi^*\left(\frac{q_1}{q^*}\right)\frac{\phi(q)}{\phi(q_1)}\tau(\chi^*, a_1).$$

8. Suppose that $\gcd(q_1, q_2) = 1$, $\chi_1$ is a character modulo $q_1$, and $\chi_2$ is a character modulo $q_2$. Prove that

$$\tau(\chi_1\chi_2, 1) = \chi_1(q_2)\chi_2(q_1)\tau(\chi_1, 1)\tau(\chi_2, 1).$$

9. (a) Suppose that $p$ is an odd prime and $\chi$ is the primitive real character modulo $p$ (i.e., $\chi$ is the Legendre symbol $(\cdot/p)$). Prove that $\tau(\chi, 1) = G(p)$, where $G(m)$ is the exponential sum defined in Problem 5.

(b) Suppose that $q$ is an odd squarefree integer and $\chi$ is the primitive real character modulo $q$. Prove that

$$\tau(\chi, 1) = \varepsilon_q \sqrt{q}, \qquad \varepsilon_q = \begin{cases} 1 & \text{if } q \equiv 1 \pmod 4, \\ \mathrm{i} & \text{if } q \equiv 3 \pmod 4. \end{cases}$$

10. Verify (3.21).

11. If $\chi_1$ and $\chi_2$ are two characters modulo $q$, the *Jacobi sum* is defined by

$$J(\chi_1, \chi_2) = \sum_{n \bmod q} \chi_1(n)\chi_2(1-n).$$

(a) Prove that when $\chi_1\chi_2$ is primitive, $\tau(\chi_1, 1)\tau(\chi_2, 1) = J(\chi_1, \chi_2)\tau(\chi_1\chi_2, 1)$.

(b) Prove that when $\chi$ is primitive, $J(\chi, \overline{\chi}) = \chi(-1)\mu(q)$.

12. Suppose that $\chi$ is a non-principal character modulo $q$, $k \geq 0$, and $\sigma \geq 1 - (\log q)^{-1}$. Prove that:

(a) $\displaystyle\sum_{n=1}^{q} \chi(n)(\log n)^k n^{-\sigma} \ll (\log q)^{k+1}$.

(b) $\displaystyle\sum_{n=q+1}^{\infty} \chi(n)(\log n)^k n^{-\sigma} \ll (\log q)^k$.

(c) $L^{(k)}(\sigma,\chi) \ll (\log q)^{k+1}$.

13. The purpose of this exercise is to establish *Dirichlet's theorem on primes in arithmetic progressions:* if $a$ and $q$ are integers with $\gcd(a,q) = 1$, then the arithmetic progression $a \mod q$ contains infinitely many prime numbers.

    (a) Suppose that $\gcd(a,q) = 1$ and $\mathrm{Re}(s) > 1$. Observe that

    $$\sum_{p \equiv a \pmod q} p^{-s} = \frac{1}{\phi(q)} \sum_{\chi \bmod q} \bar{\chi}(a) \sum_p \chi(p) p^{-s}.$$

    (b) Suppose that $\mathrm{Re}(s) > 1$ and $\chi$ is a Dirichlet character. Show that

    $$\sum_p \chi(p) p^{-s} = \log L(s,\chi) + f(s,\chi),$$

    where $f(s,\chi)$ is holomorphic in $\mathrm{Re}(s) > 1/2$.

    (c) Suppose that $\chi$ is non-principal. Then $L(s,\chi)$ is holomorphic in $\mathrm{Re}(s) > 0$. Together with Theorem 3.10, this establishes that $\log L(s,\chi)$ is holomorphic near $s = 1$. Combine this observation with the results of (a) and (b) to conclude that when $\mathrm{Re}(s) > 1$,

    $$\sum_{p \equiv a \pmod q} p^{-s} = \frac{1}{\phi(q)} \log \zeta(s) + g(s),$$

    where $g(s)$ is holomorphic near $s = 1$.

    (d) Prove that $\displaystyle\sum_{p \equiv a \pmod q} p^{-1}$ diverges. This establishes Dirichlet's theorem.

# Chapter 4

# The large sieve

In modern analytic number theory we use the term "large sieve" to describe any among several analytic lemmas, none of which is really a "sieve" (in the usual sense attached to that word in number theory). Suppose that we have a sequence $\mathcal{A} = (a_n)$ of some arithmetic interest and that we want to study its distribution in a certain sense. A typical example is the case where $\mathcal{A}$ is an integer sequence and we want to understand its distribution in arithmetic progressions or in short intervals. Often such problems can be reduced to the estimation of generating functions

$$\sum_{a\in\mathcal{A}} X(a), \tag{4.1}$$

where the function $X$ belongs to a suitably chosen class $\mathcal{X}$ of "harmonics"; nontrivial bounds for the sums (4.1) then lead to information about the distribution of the sequence $\mathcal{A}$. For example, in the proof of Theorem 2 we used this approach to establish that the primes are uniformly distributed among the reduced residue classes modulo $q$. In that case, the "harmonics" were the Dirichlet characters modulo $q$ and the generating functions (4.1) were the sums $\psi(x,\chi)$, with $\chi$ non-principal.

In a large-sieve inequality, we seek estimates for mean-square averages over $\mathcal{X}$ of general linear forms in the "harmonics" $X \in \mathcal{X}$. That is, we want to bound

$$\sum_{X\in\mathcal{X}} \left| \sum_{n\le N} a_n X(n) \right|^2,$$

for any choice of the coefficients $a_n$. In this chapter, we prove several estimates of the form

$$\sum_{X\in\mathcal{X}} \left| \sum_{n\le N} a_n X(n) \right|^2 \le C(\mathcal{X}, N) \sum_{n\le N} |a_n|^2,$$

where the harmonics are additive characters $e(\alpha m)$, Dirichlet characters $\chi(m)$, or powers $m^{-it}$. In the next chapter, we will see several applications of these results to the distribution of primes in progressions and in intervals.

## 4.1 Two results from analysis

We say that an entire function $f(z)$ is of *exponential type* $\tau$, if

$$|f(z)| \ll_\epsilon e^{(\tau+\epsilon)|z|} \qquad \text{for all fixed } \epsilon > 0.$$

The entire functions of exponential type $\tau$ whose restrictions to the real line are in $L^2(\mathbb{R})$ are characterized by the Paley–Wiener theorem (see Boas [4, §6.8] or Zygmund [59, §XVI.7]): an entire function $f(z)$ has these two properties if and only if its Fourier transform on the real line,

$$\hat{f}(t) = \int_{-\infty}^{\infty} f(x)e(-xt)\,dx,$$

is supported in the interval $|t| \le \tau/(2\pi)$.

Consider the entire function

$$B(z) = \frac{\sin^2 \pi z}{\pi^2}\left\{ \sum_{n=-\infty}^{\infty} \frac{\operatorname{sgn}(n)}{(z-n)^2} + z^{-2} + 2z^{-1} \right\},$$

where for a real number $x$,

$$\operatorname{sgn}(x) = \begin{cases} +1 & \text{if } x > 0, \\ 0 & \text{if } x = 0, \\ -1 & \text{if } x < 0. \end{cases}$$

This function, discovered by Beurling in 1930 and then rediscovered by Selberg in 1974, has the following three important properties:

($B_1$) $\operatorname{sgn}(x) \le B(x)$ for all $x \in \mathbb{R}$;

($B_2$) the Fourier transform of the function $B(x) - \operatorname{sgn}(x)$ is a continuous function, supported in $[-1, 1]$;

($B_3$) $\int_{-\infty}^{\infty} \big(B(x) - \operatorname{sgn}(x)\big)\,dx = 1$.

Furthermore, the function $B(z)$ is the unique entire function that satisfies ($B_1$) and ($B_2$) and minimizes the integral appearing in ($B_3$). The proofs of these facts can be found in Graham and Kolesnik [17, Appendix] or in Vaaler [51] (see also Exercise 1 after the chapter). We can use the function $B(z)$ to establish the following result.

**Lemma 4.1.** *Suppose that $\alpha, \beta, \delta$ are real numbers such that $\alpha < \beta$ and $\delta > 0$. There exists an entire function $F(z) = F(z; \alpha, \beta, \delta)$ such that:*

*(i)* $F(x) \ge \mathbb{1}(x; \alpha, \beta)$, *where $\mathbb{1}(x; \alpha, \beta)$ is the characteristic function of the interval $[\alpha, \beta]$;*

*(ii) its Fourier transform $\hat{F}(t)$ is supported in $[-\delta, \delta]$;*

*(iii)* $\hat{F}(0) = \beta - \alpha + \delta^{-1}$.

*Proof.* The function

$$F(z) = \tfrac{1}{2}\big(B(\delta(\beta - z)) + B(\delta(z - \alpha))\big)$$

has all the desired properties. ∎

**Lemma 4.2.** *Suppose that $F(z)$ is an entire function of exponential type $\tau$. Suppose furhter that $F \in L^2(\mathbb{R})$ and $F(x) \geq 0$ for all real $x$. Then there exists an entire function $f(z)$ of exponential type $\tau/2$ such that $F(x) = |f(x)|^2$ when $x$ is real.*

*Proof.* This follows from the main result in Boas [4, §7.5]. Hypothesis (7.5.2) in [4] follows from our hypothesis that $F \in L^2(\mathbb{R})$ and the discussion in [4, §8.1]. ∎

**Corollary 4.3.** *Suppose that $\alpha, \beta, \delta$ are real numbers such that $\alpha < \beta$ and $\delta > 0$. There exists an entire function $f(z) = f(z; \alpha, \beta, \delta)$ such that:*

*(i) $|f(x)|^2 = F(x)$ for all $x \in \mathbb{R}$, where $F(z) = F(z; \alpha, \beta, \delta)$ is the function from Lemma 4.1;*

*(ii) the Fourier transform $\hat{f}(t) = \int_{-\infty}^{\infty} f(x)e(-xt)\,dx$ is supported in $[-\delta/2, \delta/2]$.*

*Proof.* The function $F(z)$ is of exponential type $\delta$. Since by construction $F \in L^p(\mathbb{R})$ for all $p \geq 1$, we can apply Lemma 4.2 to $F(z)$. The resulting function $f(z)$ has property (i), and therefore, is of exponential type $\delta/2$ and square integrable. Hence, an appeal to the Paley–Wiener theorem proves that $\hat{f}$ is supported in $[-\delta/2, \delta/2]$. ∎

## 4.2 Large-sieve inequalities

Suppose that $\xi_1 < \xi_2 < \cdots < \xi_R$ are real numbers such that

$$T_0 + \tfrac{1}{2}\delta \leq \xi_r \leq T_0 + T - \tfrac{1}{2}\delta \tag{4.2}$$

and

$$|\xi_r - \xi_s| \geq \delta > 0 \qquad \text{whenever } r \neq s. \tag{4.3}$$

Further, suppose that $\nu_1 < \nu_2 < \cdots < \nu_K$ are real numbers such that

$$M \leq \nu_k \leq M + N$$

and

$$|\nu_k - \nu_l| \geq \Delta > 0 \qquad \text{whenever } k \neq l. \tag{4.4}$$

**Lemma 4.4.** *Suppose that $\xi_1, \xi_2, \ldots, \xi_R$ and $\nu_1, \nu_2, \ldots, \nu_K$ are as above and define*

$$S(\alpha) = \sum_{k=1}^{K} a_k e(\alpha\nu_k),$$

*where $a_1, a_2, \ldots, a_K$ are complex numbers. Then*

$$\sum_{r=1}^{R} |S(\xi_r)|^2 \leq \left(N + \delta^{-1}\right)\left(T + \Delta^{-1}\right)\sum_{k=1}^{K} |a_k|^2. \tag{4.5}$$

*Proof.* Assuming the notation of Lemma 4.1 and Corollary 4.3, we introduce the functions

$$G(z) = F(z; M, M+N, \delta), \quad g(z) = g(z; M, M+N, \delta), \quad H(z) = F(z; T_0, T_0+T, \Delta).$$

and define the sum

$$S^*(\alpha) = \sum_{k=1}^{K} a_k g(\nu_k)^{-1} e(\alpha \nu_k).$$

By Fourier inversion,

$$S(\alpha) = \int_{-\infty}^{\infty} \hat{g}(u) S^*(u+\alpha)\, du.$$

Recalling that $\hat{g}(u)$ is supported in $[-\delta/2, \delta/2]$, we obtain

$$|S(\xi_r)|^2 \le \left( \int_{-\infty}^{\infty} |\hat{g}(u)|^2\, du \right) \left( \int_{-\delta/2}^{\delta/2} |S^*(u+\xi_r)|^2\, du \right).$$

Since, by Plancherel's theorem,

$$\int_{-\infty}^{\infty} |\hat{g}(u)|^2\, du = \int_{-\infty}^{\infty} |g(u)|^2\, du = \int_{-\infty}^{\infty} G(u)\, du = \hat{G}(0),$$

it follows that

$$|S(\xi_r)|^2 \le \hat{G}(0) \int_{\xi_r-\delta/2}^{\xi_r+\delta/2} |S^*(x)|^2\, dx.$$

Hence, by (4.2) and (4.3),

$$\sum_{r=1}^{R} |S(\xi_r)|^2 \le \hat{G}(0) \int_{T_0}^{T_0+T} |S^*(x)|^2\, dx. \tag{4.6}$$

Next, we bound the integral on the right side of (4.6). Let $b_k = a_k g(\nu_k)^{-1}$. We have

$$\int_{T_0}^{T_0+T} |S^*(x)|^2\, dx \le \int_{-\infty}^{\infty} H(x) |S^*(x)|^2\, dx = \sum_{k=1}^{K} \sum_{l=1}^{K} b_k \bar{b}_l \hat{H}(\nu_l - \nu_k).$$

By (4.4), $\hat{H}(\nu_l - \nu_k)$ vanishes unless $k = l$. Thus,

$$\int_{T_0}^{T_0+T} |S^*(x)|^2\, dx \le \hat{H}(0) \sum_{k=1}^{K} |b_k|^2 = \hat{H}(0) \sum_{k=1}^{K} |a_k|^2 G(\nu_k)^{-1} \le \hat{H}(0) \sum_{k=1}^{K} |a_k|^2.$$

Combining this inequality and (4.6), we get

$$\sum_{r=1}^{R} |S(\xi_r)|^2 \le \hat{G}(0)\hat{H}(0) \sum_{k=1}^{K} |a_k|^2,$$

so the desired conclusion follows from the identities

$$\hat{G}(0) = N + \delta^{-1}, \qquad \hat{H}(0) = T + \Delta^{-1}.$$

■

**Corollary 4.5.** *Assume the notation of Lemma 4.4. Suppose that $T \le 1$ and all $\nu_k$'s are integers. Then*

$$\sum_{r=1}^{R} |S(\xi_r)|^2 \le \left(N + \delta^{-1}\right) \sum_{k=1}^{K} |a_k|^2.$$

*Proof.* Under the present hypotheses, we can estimate the right side of (4.6) as

$$\int_{T_0}^{T_0+T} |S^*(x)|^2\, dx \le \int_0^1 |S^*(x)|^2\, dx = \sum_{k=1}^{K} |b_k|^2 \le \sum_{k=1}^{K} |a_k|^2,$$

where we have used Parseval's identity. ∎

**Corollary 4.6.** *Suppose that $N, Q$ are positive integers and $a_M, \dots, a_{M+N}$ are complex numbers. Then*

$$\sum_{q\le Q} \sum_{\substack{1\le b\le q\\(b,q)=1}} \left| \sum_{n=M}^{M+N} a_n e(bn/q) \right|^2 \le \left(N + Q^2\right) \sum_{n=M}^{M+N} |a_n|^2. \tag{4.7}$$

*Proof.* When $Q = 1$, (4.7) follows from Cauchy's inequality, so we may assume that $Q \ge 2$. Then, we can view the double sum over $q$ and $b$ as a sum over all reduced fractions $b/q$, $q \le Q$, such that

$$(b/q) \in \left[Q^{-1}, 1\right] \subset \left[2Q^{-2}, 1\right].$$

If $b'/q'$ and $b''/q''$ are two such fractions, we have

$$\left| \frac{b'}{q'} - \frac{b''}{q''} \right| = \frac{|b'q'' - b''q'|}{q'q''} \ge \frac{1}{Q^2},$$

unless $b' = b''$ and $q' = q''$. Thus, (4.7) follows from Corollary 4.5 with $\delta = Q^{-2}$. ∎

We now turn to averages of character sums.

**Lemma 4.7.** *Suppose that $q, M, N$ are positive integers and $a_{M+1}, \dots, a_{M+N}$ are complex numbers. Then*

$$\sum_{\chi \bmod q} \left| \sum_{n=M+1}^{M+N} a_n \chi(n) \right|^2 \le \left(N + \phi(q)\right) \sum_{n=M+1}^{M+N} |a_n|^2. \tag{4.8}$$

*Proof.* The sum on the left of (4.8) equals

$$\sum_{m,n}^{*} a_m \bar{a}_n \sum_{\chi \bmod q} \chi(m)\bar{\chi}(n) = \sum_{m,n}^{*} a_m \bar{a}_n \sum_{\chi \bmod q} \chi(m\bar{n}),$$

where $\sum_{m,n}^{*}$ denotes a summation restricted to integers relatively prime to $q$ and $\bar{m}$ is the multiplicative inverse of $m$ modulo $q$: $m\bar{m} \equiv 1 \pmod q$. By (3.5), the latter sum is

$$\le \phi(q) \sum_{m\equiv n \,(\mathrm{mod}\, q)} |a_m a_n| \le \phi(q) \sum_{m\equiv n \,(\mathrm{mod}\, q)} \tfrac{1}{2}\left(|a_m|^2 + |a_n|^2\right) \le \phi(q)\left(Nq^{-1} + 1\right) \sum_{m} |a_m|^2,$$

and (4.8) follows. ∎

If we want to sum over $q$ as well, we need to restrict our attention to primitive characters or to impose some restrictions on the coefficients $a_n$. The next lemma provides such a result.

**Lemma 4.8.** *Suppose that $Q, M, N$ are positive integers and $a_{M+1}, \dots, a_{M+N}$ are complex numbers. Then*

$$\sum_{q\le Q} \frac{q}{\phi(q)} \sideset{}{^*}\sum_{\chi \bmod q} \left| \sum_{n=M+1}^{M+N} a_n \chi(n) \right|^2 \le \left(N + Q^2\right) \sum_{n=M+1}^{M+N} |a_n|^2.$$

*Proof.* When $\chi$ is primitive, Lemmas 3.5 and 3.6 yield

$$q\left| \sum_{n=M+1}^{M+N} a_n \chi(n) \right|^2 = \left| \sum_{n=M+1}^{M+N} a_n \tau(\bar{\chi}, n) \right|^2 = \left| \sum_{\substack{1\le b\le q \\ (b,q)=1}} \bar{\chi}(b) S(b/q) \right|^2,$$

where

$$S(\alpha) = \sum_{n=1}^{N} \tilde{a}_n e(\alpha n), \qquad |\tilde{a}_n| = |a_n|.$$

Hence,

$$\frac{q}{\phi(q)} \sideset{}{^*}\sum_{\chi \bmod q} \left| \sum_{n=M+1}^{M+N} a_n \chi(n) \right|^2 \le \frac{1}{\phi(q)} \sum_{\chi \bmod q} \left| \sum_{\substack{1\le b\le q \\ (b,q)=1}} \bar{\chi}(b) S(b/q) \right|^2$$

$$= \frac{1}{\phi(q)} \sum_{\substack{1\le b_1, b_2\le q \\ (b_1 b_2, q)=1}} S(b_1/q)\overline{S(b_2/q)} \sum_{\chi \bmod q} \chi(b_2 \bar{b}_1) = \sum_{\substack{1\le b\le q \\ (b,q)=1}} |S(b/q)|^2;$$

here $\bar{b}_1$ denotes the multiplicative inverse of $b_1$ modulo $q$. Thus, the lemma follows from (4.7). ∎

Next, we consider averages of Dirichlet polynomials of the form

$$D(s) = \sum_{n=1}^{N} a_n n^{-s}. \tag{4.9}$$

**Lemma 4.9.** *Suppose that $\delta > 0$ and $t_1 < t_2 < \cdots < t_R$ are real numbers such that*

$$T_0 + \tfrac{1}{2}\delta \le t_r \le T_0 + T - \tfrac{1}{2}\delta$$

*and*

$$|t_{r_1} - t_{r_2}| \ge \delta > 0 \qquad \textit{whenever } r_1 \ne r_2.$$

*Further, suppose that $a_1, \dots, a_N$ are complex numbers and $D(s)$ is defined by* (4.9)*. Then*

$$\sum_{r=1}^{R} |D(\mathrm{i}t_r)|^2 \le \left(\delta^{-1} + \tfrac{1}{2\pi}\log N\right)\left(T + 2\pi N\right) \sum_{n=1}^{N} |a_n|^2. \tag{4.10}$$

*Proof.* We apply Lemma 4.4 with $\xi_r = t_r$ and $\nu_n = -\frac{1}{2\pi}\log n$. Then

$$-\tfrac{1}{2\pi}\log N \le \nu_n \le 0,$$

and for $1 \le m < n \le N$,

$$2\pi|\nu_m - \nu_n| = \log(n/m) \ge \begin{cases} \log 2 & \text{if } n \ge 2m, \\ N^{-1} & \text{if } n < 2m. \end{cases}$$

We note that the estimate in the case $m < n < 2m$ uses the inequalities

$$\log\Big(\frac{n}{m}\Big) \ge \frac{n-m}{m}\Big(1 - \frac{n-m}{2m}\Big) \ge \frac{1}{N-1}\Big(1 - \frac{1}{2N-2}\Big) \ge \frac{1}{N}.$$

Thus, in the application of Lemma 4.4, $N = \frac{1}{2\pi}\log N$ and $\Delta = (2\pi N)^{-1}$ and (4.10) is an immediate consequence of (4.5). ∎

**Lemma 4.10.** *Suppose that $a_1, \dots, a_N$ are complex numbers and $D(s)$ is defined by* (4.9)*. Then*

$$\int_{T_0}^{T_0+T} |D(\mathrm{i}t)|^2\, dt \le \big(T + 2\pi N\big) \sum_{n=1}^{N} |a_n|^2.$$

*Proof.* We argue similarly to the second part of the proof of Lemma 4.4. Let $\Delta = (2\pi N)^{-1}$ and $G(z) = F(z; T_0, T, \Delta)$, where $F(z)$ is the function from Lemma 4.1. Then

$$\int_{T_0}^{T_0+T} |D(\mathrm{i}t)|^2\, dt \le \int_{-\infty}^{\infty} G(t)|D(\mathrm{i}t)|^2\, dt = \sum_{m=1}^{N}\sum_{n=1}^{N} a_m\overline{a}_n \hat{G}\big(\tfrac{1}{2\pi}\log(n/m)\big).$$

Recalling from the proof of the previous lemma that $\frac{1}{2\pi}|\log(n/m)| \ge \Delta$ unless $m = n$, we get

$$\int_{T_0}^{T_0+T} |D(\mathrm{i}t)|^2\, dt \le \hat{G}(0)\sum_{n=1}^{N}|a_n|^2 = \big(T + 2\pi N\big)\sum_{n=1}^{N}|a_n|^2.$$

∎

## 4.3 Dirichlet polynomials with characters: a hybrid sieve

So far we have obtained large-sieve results in the form of inequalities with two terms on the right side, one of which corresponds to the maximum size of the summands and the other to the mean square of the function times the number of terms. In applications to the distribution of primes, we sometimes consider Dirichlet polynomials of the form

$$D(s,\chi) = \sum_{n=1}^{N} a_n\chi(n)n^{-s}. \tag{4.11}$$

We can sieve $D(s,\chi)$ over $\chi$ or over $s$; we can also sieve over both $\chi$ and $s$ and obtain a hybrid result. Oftentimes, such hybrid results are superior to either of the estimates resulting from sieving over one of $\chi$ or $s$ and then summing (or integrating) over the other.

**Lemma 4.11.** *Suppose that $q, Q, N$ are positive integers, $T_0$ and $T$ are real, and $D(s,\chi)$ is defined by* (4.11)*. Then*

$$\sum_{\chi \bmod q} \int_{T_0}^{T_0+T} |D(\mathrm{i}t,\chi)|^2\,dt \ll \big(N + \phi(q)T\big) \sum_{n=1}^{N} |a_n|^2 \tag{4.12}$$

*and*

$$\sum_{q\le Q} \frac{q}{\phi(q)} \sideset{}{^*}\sum_{\chi \bmod q} \int_{T_0}^{T_0+T} |D(\mathrm{i}t,\chi)|^2\,dt \ll \big(N + Q^2T\big) \sum_{n=1}^{N} |a_n|^2. \tag{4.13}$$

*Proof.* Let $G(z) = F(z; T_0, T, T^{-1})$ and $g(z) = f(z; T_0, T, T^{-1})$, where $F$ and $f$ are the functions from Lemma 4.1 and Corollary 4.3, respectively. Then

$$\int_{T_0}^{T_0+T} |D(\mathrm{i}t,\chi)|^2\,dt \le \int_{-\infty}^{\infty} G(t)|D(\mathrm{i}t,\chi)|^2\,dt = \int_{-\infty}^{\infty} |g(t)D(\mathrm{i}t,\chi)|^2\,dt. \tag{4.14}$$

Note that

$$g(-t)D(-\mathrm{i}t,\chi) = \hat{S}_\chi(t), \qquad S_\chi(x) = \sum_{n=1}^{N} a_n\chi(n)\hat{g}\big(x + \tfrac{1}{2\pi}\log n\big).$$

Applying Plancherel's theorem to the right side of (4.14), we find that

$$\int_{T_0}^{T_0+T} |D(\mathrm{i}t,\chi)|^2\,dt \le \int_{-\infty}^{\infty} |S_\chi(x)|^2\,dx.$$

Hence,

$$\sum_{\chi \bmod q} \int_{T_0}^{T_0+T} |D(\mathrm{i}t,\chi)|^2\,dt \le \int_{-\infty}^{\infty} \sum_{\chi \bmod q} |S_\chi(x)|^2\,dx. \tag{4.15}$$

Since $\hat{g}(x)$ is supported in the interval $|x| \le (2T)^{-1}$, the summation in $S_\chi(x)$ is supported in an interval of length

$$\le e^{-2\pi x}\big(e^{\pi/T} - e^{-\pi/T}\big) \ll T^{-1}e^{-2\pi x} = H(x), \qquad \text{say.}$$

Thus, an appeal to Lemma 4.7 gives

$$\sum_{\chi \bmod q} |S(x,\chi)|^2 \ll \big(H(x) + \phi(q)\big) \sum_{n=1}^{N} \big|a_n\hat{g}\big(x + \tfrac{1}{2\pi}\log n\big)\big|^2.$$

Inserting this bound into the right side of (4.15), we obtain

$$\begin{aligned}\sum_{\chi \bmod q} \int_{T_0}^{T_0+T} |D(\mathrm{i}t,\chi)|^2\,dt &\ll \int_{-\infty}^{\infty} \big(H(x) + \phi(q)\big) \sum_{n=1}^{N} \big|a_n\hat{g}\big(x + \tfrac{1}{2\pi}\log n\big)\big|^2\,dx \\ &\ll \sum_{n=1}^{N} |a_n|^2 \int_{-\infty}^{\infty} \big(H(x) + \phi(q)\big)\big|\hat{g}\big(x + \tfrac{1}{2\pi}\log n\big)\big|^2\,dx.\end{aligned}$$

We now observe that the last integral is supported in the interval $\left|x + \frac{1}{2\pi}\log n\right| \le (2T)^{-1}$. For such values of $x$, we have $H(x) \ll nT^{-1}$, whence

$$\sum_{\chi \bmod q} \int_{T_0}^{T_0+T} |D(\mathrm{i}t,\chi)|^2\,dt \ll \sum_{n=1}^{N} \big(nT^{-1} + \phi(q)\big)|a_n|^2 \int_{-\infty}^{\infty} \left|\hat{g}\big(x + \tfrac{1}{2\pi}\log n\big)\right|^2 dx. \tag{4.16}$$

Finally, we observe that the integral on the right side of (4.16) equals

$$\int_{-\infty}^{\infty} |\hat{g}(x)|^2\,dx = \int_{-\infty}^{\infty} |g(t)|^2\,dt = \int_{-\infty}^{\infty} G(t)\,dt = 2T,$$

and so (4.12) follows from (4.16). The proof of (4.13) is similar, using Lemma 4.8 instead of Lemma 4.7. ∎

An important alternative form of the hybrid large sieve is the following. Consider a collection of Dirichlet characters, such as the collection of all characters modulo $q$ or the collection of all primitive characters to moduli $q \le Q$. Suppose that for each character $\chi$ in that collection we have selected real numbers $t_1(\chi) < t_2(\chi) < \cdots < t_R(\chi)$, $R = R(\chi)$, such that

$$T_0 + \delta/2 \le t_r(\chi) \le T_0 + T - \delta/2$$

and

$$|t_r(\chi) - t_s(\chi)| \ge \delta > 0 \qquad \text{whenever } r \ne s.$$

We will refer to such a collection $\mathfrak{S}$ of pairs $(\chi, t_r(\chi))$ of characters and real numbers as a *$\delta$-spaced set of points*. When the given collection of characters is that of all characters modulo $q$ or of all primitive characters to moduli $q \le Q$, we further define

$$|\mathfrak{S}| = \phi(q)T \qquad \text{or} \qquad |\mathfrak{S}| = Q^2T,$$

respectively.

**Lemma 4.12.** *Suppose that $D(s,\chi)$ is defined by* (4.11) *and $\mathfrak{S}$ is a $\delta$-spaced set of points $(\chi, t_r(\chi))$ of one of the two special kinds described above. Then*

$$\sum_{(\chi,t_r(\chi))\in\mathfrak{S}} |D(\mathrm{i}t_r(\chi),\chi)|^2 \ll \big(\delta^{-1} + \log N\big)\big(N + |\mathfrak{S}|\big)\sum_{n=1}^{N}|a_n|^2.$$

*Proof.* Fix a character $\chi$ and consider the respective numbers $t_1(\chi), \dots, t_R(\chi)$. Following the proof of Lemma 4.4 (with $\nu_n = \log n$, $M = 0$, $N = \log N$, and $\Delta = N^{-1}$) up to (4.6), we get

$$\sum_{r=1}^{R(\chi)} |D(\mathrm{i}t_r(\chi),\chi)|^2 \le \big(\delta^{-1} + \log N\big)\int_{T_0}^{T_0+T} |D^*(\mathrm{i}t,\chi)|^2\,dt, \tag{4.17}$$

where

$$D^*(s,\chi) = \sum_{n=1}^{N} b_n\chi(n)n^{-s},$$

with coefficients satisfying

$$\sum_{n=1}^{N} |b_n|^2 \le \sum_{n=1}^{N} |a_n|^2. \tag{4.18}$$

Summing (4.17) over all characters appearing in $\mathfrak{S}$, we obtain

$$\sum_{(\chi,t_r(\chi))\in\mathfrak{S}} |D(\mathrm{i}t_r(\chi),\chi)|^2 \le \big(\delta^{-1} + \log N\big) \sum_{\chi} \int_{T_0}^{T_0+T} |D^*(\mathrm{i}t,\chi)|^2\, dt. \tag{4.19}$$

The desired result follows from (4.18), (4.19), and Lemma 4.11. ∎

## Exercises

1. For $z \in \mathbb{C}$, define

$$H(z) = \frac{\sin^2 \pi z}{\pi^2}\Big\{ \sum_{n=-\infty}^{\infty} \frac{\mathrm{sgn}(n)}{(z-n)^2} + 2z^{-1} \Big\}, \qquad K(z) = \Big(\frac{\sin \pi z}{\pi z}\Big)^2.$$

The Beurling–Selberg function $B(z)$ from §4.1 is then $B(z) = H(z) + K(z)$.

(a) Let $\psi(x) = \{x\} - 1/2$ and $\psi_1(x) = \int_0^x \psi(u)\, du$. Show that when $x > 0$,

$$\sum_{n=1}^{\infty} \frac{1}{(x+n)^2} = \frac{1}{x} - 2\int_0^{\infty} \frac{\{u\}\, du}{(u+x)^3} = \frac{1}{x} - \frac{1}{2x^2} + 6\int_0^{\infty} \frac{\psi_1(u)\, du}{(u+x)^4}.$$

(b) Prove that $|H(x)| \le 1$ for all $x \in \mathbb{R}$.

(c) Prove that $|H(x) - \mathrm{sgn}(x)| \le K(x)$ for all $x \in \mathbb{R}$.

(d) Prove that $B(x)$ satisfies property $(\mathrm{B}_1)$ in §4.1.

(e) Prove that $B(z)$ has exponential type $2\pi$.

(f) By part (c), $B(x) - \mathrm{sgn}(x)$ is integrable, and so

$$\int_{-\infty}^{\infty} \big(B(x) - \mathrm{sgn}(x)\big)\, dx = \lim_{N\to\infty} \int_{-N}^{N} \big(B(x) - \mathrm{sgn}(x)\big)\, dx.$$

Use this to show that

$$\int_{-\infty}^{\infty} \big(B(x) - \mathrm{sgn}(x)\big)\, dx = \int_{-\infty}^{\infty} K(x)\, dx.$$

(g) Show that

$$K(z) = \int_{-1}^{1} (1-|t|)e(tz)\, dt \qquad \text{and} \qquad \hat{K}(t) = \max(1-|t|, 0).$$

Combine the latter identity and the result of part (f) to prove that $B(x)$ satisfies property $(\mathrm{B}_3)$ in §4.1.

**Remark**. Note that we stopped just short of establishing property $(\mathrm{B}_2)$ in §4.1: if $B(x)$ belonged to $L^2(\mathbb{R})$, we would be able to deduce $(\mathrm{B}_2)$ from (e) above and the Paley–Wiener theorem, but of course, $B(x)$ does not belong to $L^p(\mathbb{R})$ for any $p < \infty$. On the other hand, the function $F(z)$ constructed in Lemma 4.1 does belong to $L^2(\mathbb{R})$, so the above properties of $B(z)$ suffice to prove that that function has all the desired properties.

2. Suppose that $T > 0$ and $\nu_1, \nu_2, \dots, \nu_K$ are real numbers, and define

$$S(t) = \sum_{k=1}^{K} a_k e(\nu_k t),$$

where $a_1, a_2, \dots, a_K$ are complex numbers. Prove that

$$\int_{-T}^{T} |S(t)|^2 \, dt \le (\pi T)^2 \int_{-\infty}^{\infty} \left| \sum_{|\nu_k - x| \le (2T)^{-1}} a_k \right|^2 dx.$$

[HINT: Let $\delta = (2T)^{-1}$. Start by arguing similarly to the proof of (4.15), but choose $g(t) = (\sin \pi\delta t)/(\pi\delta t)$ and $G(t) = g(t)^2$. Then $G(t) \ge 4\pi^{-2}$ for all $t \in [-T, T]$ and $\hat{g}(x)$ is the characteristic function of $[-\delta/2, \delta/2]$, normalized in $L^1(\mathbb{R})$.]

3. Consider the *Ramanujan sum*

$$c_q(n) = \sum_{\substack{1 \le b \le q \\ (b,q)=1}} e(bn/q).$$

   (a) Prove that $c_q(n)$ is multiplicative as a function of $q$, that is, $c_{q_1 q_2}(n) = c_{q_1}(n) c_{q_2}(n)$ whenever $(q_1, q_2) = 1$.

   (b) Prove that

$$c_q(n) = \phi(q) \mu\left(\frac{q}{(q,n)}\right) \phi\left(\frac{q}{(q,n)}\right)^{-1}.$$

4. (a) Let $\mathcal{N}_x$ denote the set of integers not divisible by primes $p > x$. Prove that

$$\sum_{n \le x} \mu(n)^2 \phi(n)^{-1} = \sum_{n \in \mathcal{N}_x} n^{-1} \ge \log x.$$

   (b) Suppose that $q$ is a positive integer. Prove that

$$\sum_{\substack{n \le x \\ (n,q)=1}} \mu^2(n) \phi(n)^{-1} \ge \frac{\phi(q)}{q} \sum_{n \le x} \mu(n)^2 \phi(n)^{-1}.$$

5. Suppose that $\mathcal{N}$ is a set of positive integers contained in $[M, M+N]$. For each $q \le Q$, define

$$\mathcal{R}_q = \big\{ h \in \mathbb{Z} \,:\, 1 \le h \le q,\ (n-h, q) = 1 \text{ for all } n \in \mathcal{N} \big\}, \qquad \omega(q) = |\mathcal{R}_q|.$$

The purpose of this exercise is to prove that

$$|\mathcal{N}| \le \big(N + Q^2\big) \left\{ \sum_{q \le Q} \mu(q)^2 \prod_{p \mid q} \left( \frac{\omega(p)}{p - \omega(p)} \right) \right\}^{-1}. \tag{$*$}$$

   (a) Prove that $\omega(q)$ is multiplicative.

   (b) Define

$$S(\alpha) = \sum_{n \in \mathcal{N}} e(\alpha n).$$

   Use the result of Exercise 3 to show that

$$\sum_{h \in \mathcal{R}_q} \sum_{\substack{1 \le b \le q \\ (b,q)=1}} S(b/q) e(-bh/q) = \mu(q) \omega(q) |\mathfrak{N}|.$$

(c) Suppose that $q$ is squarefree. Prove that

$$\sum_{\substack{1\le b\le q\\(b,q)=1}}\left|\sum_{h\in\mathcal{R}_q}e(-bh/q)\right|^2=\omega(q)^2\sum_{d|q}\frac{d\mu(q/d)}{\omega(d)}=\omega(q)\prod_{p|q}\big(p-\omega(p)\big).$$

(d) Show that

$$|\mathcal{N}|^2\sum_{q\le Q}\mu(q)^2\prod_{p|q}\left(\frac{\omega(p)}{p-\omega(p)}\right)\le\sum_{q\le Q}\sum_{\substack{1\le b\le q\\(b,q)=1}}|S(b/q)|^2.$$

Use this inequality and Corollary 4.6 to establish $(*)$.

6. Suppose that $M, N, Q$ are positive integers with $Q \le M$ and let $\mathcal{N}$ be the set of primes in $[M+1, M+N]$. Apply the result of the previous exercise to show that

$$\pi(M+N)-\pi(M)\le\big(N+Q^2\big)\left\{\sum_{q\le Q}\mu^2(q)\phi(q)^{-1}\right\}^{-1}.$$

Deduce that

$$\pi(M+N)-\pi(M)\le\big(N+Q^2\big)(\log Q)^{-1}.$$

Upon choosing $Q=N^{1/2}(\log N)^{-1/2}$, this provides a Chebyshev-type upper bound for primes in short intervals:

$$\pi(M+N)-\pi(M)\le\frac{N}{\log N}\left\{2+O\left(\frac{\log\log N}{\log N}\right)\right\},$$

whenever $N > 2$ and $M \ge N^{1/2}$.

7. Suppose that $M, N, q$ are positive integers and $a$ is an integer with $(a, q) = 1$. Generalize the result of the previous exercise to prove that

$$\pi(M+N;q,a)-\pi(M;q,a)\le\frac{N}{\phi(q)\log(N/q)}\left\{2+O\left(\frac{\log\log(N/q)}{\log(N/q)}\right)\right\},$$

whenever $N \ge 3q$ and $M \ge (N/q)^{1/2}$. This result is one of the many versions of the *Brun–Titchmarsh inequality*. The sharpest result in this direction was obtained by Montgomery and Vaughan [42]:

$$\pi(M+N;q,a)-\pi(M;q,a)\le\frac{2N}{\phi(q)\log(N/q)},$$

whenever $N > q$.

# Chapter 5

# Applications of the large sieve

## 5.1 Sums over primes and double sums

Suppose that the function $f : \mathbb{N} \to \mathbb{C}$ is such that, when $x$ is large, the sums

$$\sum_{n \le x} f(n)$$

exhibit certain cancellation, and that we want to show that the same is true for the sums

$$\sum_{p \le x} f(p). \tag{5.1}$$

The first general method for obtaining such results was developed by I. M. Vinogradov in the late 1930s. His starting point is the sieve of Eratosthenes. Let $P(z)$ denote the product of all primes $p \le z$ and write $P_x = P(x^{1/2})$. Then

$$\sum_{\substack{n \le x \\ (n, P_x)=1}} f(n) = f(1) + \sum_{x^{1/2} < p \le x} f(p), \tag{5.2}$$

since the only numbers $n \le x$ that are not divisible by any prime $\le x^{1/2}$ are 1 and the primes in $(x^{1/2}, x]$. Using the properties of the Möbius function, we can write the sum on the left side of (5.2) as

$$\sum_{\substack{n \le x \\ (n, P_x)=1}} f(n) = \sum_{n \le x} f(n) \sum_{d \mid (n, P_x)} \mu(d) = \sum_{d \mid P_x} \mu(d) \sum_{m \le x/d} f(md).$$

The crux of Vinogradov's method is a clever (and complicated) combinatorial argument that decomposes the latter sum into several subsums of two major types:

- *type I sums:* double sums of the form

$$\sum_{m \le M} \sum_{n \le x/m} a_m f(mn), \tag{5.3}$$

  where $M$ is not too large and the coefficients $a_m$ are small on average, but otherwise arbitrary;

- *type II sums:* double sums of the form

$$\sum_{m\le M}\sum_{n\le N} a_m b_n f(mn), \tag{5.4}$$

  where $M$ and $N$ are neither too small nor too large and the coefficients $a_m, b_n$ are small on average, but otherwise arbitrary.

This reduces the estimation of (5.1) to the estimation of type I and type II double sums.

In 1977 Vaughan [53] found an alternative way for decomposing sums over primes into double sums that is much more straightforward than Vinogradov's. His result is as follows.

**Lemma 5.1 (Vaughan).** *Suppose that* $2 \le U, V < X$. *Then*

$$\sum_{U<n\le X} \Lambda(n) f(n) = \Sigma_1 - \Sigma_2 - \Sigma_3, \tag{5.5}$$

*where*

$$\Sigma_1 = \sum_{m\le V}\sum_{U<mk\le X} \mu(m)(\log k) f(mk), \qquad \Sigma_2 = \sum_{m\le UV}\sum_{U<mk\le X} a_m f(mk),$$

*and*

$$\Sigma_3 = \sum_{\substack{m>U\\ mk\le X}}\sum_{k>V} \Lambda(m) b_k f(mk),$$

*with coefficients* $|a_m| \le \log m$ *and* $|b_k| \le d(k)$.

*Proof.* Our main tool is the identity

$$-\frac{\zeta'(s)}{\zeta(s)} = L(s) - M(s)\zeta'(s) - L(s)M(s)\zeta(s) + \left(-\frac{\zeta'(s)}{\zeta(s)} - L(s)\right)\big(1 - M(s)\zeta(s)\big), \tag{5.6}$$

in which we choose $L(s)$ and $M(s)$ to be the Dirichlet polynomials

$$L(s) = \sum_{n\le U} \Lambda(n) n^{-s} \qquad \text{and} \qquad M(s) = \sum_{n\le V} \mu(n) n^{-s}.$$

Suppose that $n > U$. Comparing the coefficients of $n^{-s}$ in the Dirichlet series representations of the left and right sides of (5.6) we obtain the following identity for $\Lambda(n)$:

$$\Lambda(n) = -\sum_{\substack{mk=n\\ m\le U}} \mu(m)(-\log k) - \sum_{\substack{uvk=n\\ u\le U, v\le V}} \Lambda(u)\mu(v) + \sum_{\substack{mk=n\\ m>U, k>V}} \Lambda(m)\left(-\sum_{\substack{uv=k\\ u\le V}} \mu(u)\right).$$

Multiplying both sides of by $f(n)$ and summing over $U < n \le X$, we obtain (5.5) with

$$a_m = \sum_{\substack{uv=m\\ u\le U, v\le V}} \Lambda(u)\mu(v), \qquad b_k = \sum_{\substack{uv=k\\ u\le V<uv}} \mu(u).$$

Clearly, $|a_m| \le \sum_{u|m} \Lambda(u) = \log m$ and $|b_k| \le d(k)$. ∎

Heath-Brown [22] proposed a different decomposition for von Mangoldt's function, which provides more flexibility than Lemma 5.1 and sometimes leads to superior results. Like Vaughan's, Heath-Brown's identity arises from an identity for $\zeta'(s)/\zeta(s)$. In this case, the underlying identity is

$$\frac{\zeta'(s)}{\zeta(s)} = \sum_{j=1}^{k} (-1)^{j-1} \binom{k}{j} \zeta(s)^{j-1} \zeta'(s) M(s)^j + \zeta(s)^{-1} (1 - \zeta(s)M(s))^k \zeta'(s), \tag{5.7}$$

here $k \geq 1$ is an integer and $M(s)$ is the Dirichlet polynomial

$$M(s) = \sum_{n \leq X} \mu(n) n^{-s}.$$

Suppose that $n$ is an integer with $n \leq X^k$ and consider the coefficients of $n^{-s}$ on both sides of (5.7). The coefficient of $n^{-s}$ on the left side of (5.7) is $-\Lambda(n)$ and the last term on the right side of (5.7) does not contribute to the coefficient of $n^{-s}$. We thus find that

$$\Lambda(n) = \sum_{j=1}^{k} \binom{k}{j} (-1)^j \sum_{\substack{m_1 \cdots m_{2j} = n \\ m_1, \dots, m_j \leq X}} (\log m_1) \mu(m_{j+1}) \cdots \mu(m_{2j}), \tag{5.8}$$

whenever $n \leq X^k$. We will come back to Heath-Brown's identity when we discuss the distribution of primes in short intervals later in this chapter.

## 5.2 The Bombieri–Vinogradov theorem

In this section we will use Vaughan's identity, the large sieve for character sums in the form of Lemma 4.8, and the Pólya–Vinogradov theorem to establish the following result equivalent to Theorem 3.

**Theorem 5.2.** *Suppose that $2 \leq Q \leq x$. Then, for any fixed $A > 0$,*

$$\sum_{q \leq Q} \max_{(a,q)=1} \max_{y \leq x} \left| \psi(y; q, a) - \frac{y}{\phi(q)} \right| \ll x(\log x)^{-A} + Qx^{1/2} (\log x)^5. \tag{5.9}$$

### 5.2.1 Preparations

Define

$$\delta_\chi = \begin{cases} 1 & \text{if } \chi \text{ is principal,} \\ 0 & \text{otherwise.} \end{cases}$$

By (3.58),

$$\psi(y; q, a) - \frac{y}{\phi(q)} = \frac{1}{\phi(q)} \sum_{\chi \bmod q} \overline{\chi}(a) \big( \psi(y, \chi) - \delta_\chi y \big),$$

whence

$$\max_{(a,q)=1} \Big|\psi(y;q,a) - \frac{y}{\phi(q)}\Big| \le \frac{1}{\phi(q)} \sum_{\chi \bmod q} \big|\psi(y,\chi) - \delta_\chi y\big|.$$

Writing $\Sigma(x,Q)$ for the left side of (5.9), we find that

$$\Sigma(x,Q) \le \sum_{q\le Q} \frac{1}{\phi(q)} \sum_{\chi \bmod q} \max_{y\le x} \big|\psi(y,\chi) - \delta_\chi y\big| = \Sigma_0 + \Sigma_1, \quad \text{say,} \tag{5.10}$$

where $\Sigma_0$ denotes the contribution from the principal characters and $\Sigma_1$ denotes the contribution from all the other characters. By (3.60) (that inequality holds for all $q \le x$) and the elementary bound (see Exercise 1)

$$\sum_{n\le z} \phi(mn)^{-1} \ll \phi(m)^{-1} \log z, \tag{5.11}$$

we have

$$\Sigma_0 \ll x \exp\big(-c_1\sqrt{\log x}\big) \sum_{q\le Q} \frac{1}{\phi(q)} \ll x(\log x)^{-A}. \tag{5.12}$$

As usual, for a non-principal character $\chi$ modulo $q$, we denote by $\chi^*$ the primitive character inducing $\chi$. By (3.56),

$$\Sigma_1 \ll \sum_{q\le Q} \frac{1}{\phi(q)} \sum_{\chi \bmod q} \max_{y\le x} |\psi(y,\chi^*)| + Q(\log x)^2.$$

Rearranging the sum over the characters as to combine the contributions of all characters to moduli $q = rq_1 \le Q$ induced by the same primitive character $\chi$ modulo $r$, we deduce that

$$\begin{aligned}\Sigma_1 &\ll \sum_{r\le Q} \sideset{}{^*}\sum_{\chi \bmod r} \max_{y\le x} |\psi(y,\chi)| \sum_{q_1\le Q/r} \frac{1}{\phi(rq_1)} + Q(\log x)^2 \\ &\ll (\log x) \sum_{r\le Q} \frac{1}{\phi(r)} \sideset{}{^*}\sum_{\chi \bmod r} \max_{y\le x} |\psi(y,\chi)| + Q(\log x)^2,\end{aligned} \tag{5.13}$$

where we have used (5.11) again. We can estimate the contribution from the "small" moduli $r$ using the Siegel–Walfisz theorem. Indeed, by (3.60) and (3.61), we have

$$\max_{y\le x} |\psi(y,\chi)| \ll x \exp\big(-c(A)\sqrt{\log x}\big)$$

for all primitive characters to moduli $r \le (\log x)^{A+5} = Q_0$, say. Thus,

$$\sum_{r\le Q_0} \frac{1}{\phi(r)} \sideset{}{^*}\sum_{\chi \bmod r} \max_{y\le x} |\psi(y,\chi)| \ll xQ_0 \exp\big(-c(A)\sqrt{\log x}\big) \ll x(\log x)^{-A-1}.$$

Combining this inequality and (5.13), we obtain

$$\Sigma_1 \ll (\log x)\Sigma_2 + x(\log x)^{-A} + Q(\log x)^2, \tag{5.14}$$

where

$$\Sigma_2 = \sum_{Q_0<r\le Q} \frac{1}{\phi(r)} \sideset{}{^*}\sum_{\chi \bmod r} \max_{y\le x} |\psi(y,\chi)|.$$

### 5.2.2 Application of Vaughan's identity

By (5.5) with $f(n) = \chi(n)$, $X = y$, and $U = V \le x^{1/2}$ (we will specify our choice of $U$ later),

$$\psi(y,\chi) = S_1(y,\chi) - S_2(y,\chi) - S_3(y,\chi) + \psi(U,\chi),$$

where $S_j(y,\chi)$ denotes the sum $\Sigma_j$ on the right side of (5.5). Hence,

$$\Sigma_2 \ll \Sigma_3 + \Sigma_4 + \Sigma_5 + QU, \tag{5.15}$$

where

$$\Sigma_j = \sum_{Q_0 < r \le Q} \frac{1}{\phi(r)} \sideset{}{^*}\sum_{\chi \bmod r} \max_{y \le x} |S_{j-2}(y,\chi)| \qquad (j = 3,4,5).$$

We can estimate $\Sigma_3$ right away. By partial summation,

$$S_1(y,\chi) \le \sum_{m \le U} \Bigg| \sum_{U < mk \le y} (\log k)\chi(k) \Bigg| \ll (\log y) \sum_{m \le U} \Bigg| \sum_{U < mk \le z} \chi(k) \Bigg|$$

for some $z$ with $U < z \le y$. Thus, by the Pólya–Vinogradov inequality,

$$\max_{y \le x} |S_1(y,\chi)| \ll r^{1/2} U (\log x)^2.$$

We conclude that

$$\Sigma_3 \ll Q^{3/2} U (\log x)^2. \tag{5.16}$$

Next we estimate $\Sigma_5$ using the large sieve. We then split $\Sigma_4$ into two subsums: one similar to $\Sigma_3$ and one similar to $\Sigma_5$.

### 5.2.3 Estimation of $\Sigma_5$

Suppose that $a_1, \dots, a_M$ and $b_1, \dots, b_K$ are complex numbers. Then, by Cauchy's inequality and Lemma 4.8,

$$\begin{aligned}
&\sum_{r \le R} \frac{r}{\phi(r)} \sideset{}{^*}\sum_{\chi \bmod r} \Bigg| \sum_{m=1}^{M} \sum_{k=1}^{K} a_m b_k \chi(mk) \Bigg| \\
&\ll \Bigg\{ \sum_{r \le R} \frac{r}{\phi(r)} \sideset{}{^*}\sum_{\chi \bmod r} \Bigg| \sum_{m=1}^{M} a_m \chi(m) \Bigg|^2 \Bigg\}^{1/2} \Bigg\{ \sum_{r \le R} \frac{r}{\phi(r)} \sideset{}{^*}\sum_{\chi \bmod r} \Bigg| \sum_{k=1}^{K} b_k \chi(k) \Bigg|^2 \Bigg\}^{1/2} \\
&\ll \big(M + R^2\big)^{1/2} \big(K + R^2\big)^{1/2} \Bigg( \sum_{m=1}^{M} |a_m|^2 \Bigg)^{1/2} \Bigg( \sum_{k=1}^{K} |b_k|^2 \Bigg)^{1/2}.
\end{aligned} \tag{5.17}$$

We would like to apply this bound to $\Sigma_5$, but before we can do that we must deal with the summation condition $mk \le y$ appearing in the definition of $S_3(y,\chi)$.

We start by splitting the interval $U < m \le xU^{-1}$ into $O(\log x)$ subintervals $M < m \le M_1$ such that $M_1 \le 2M$. Then, for some choice of $M, M_1$, we have

$$S_3(y,\chi) \ll (\log x)\Big|\sum_{M<m\le M_1}\sum_{U<k\le y/m}\Lambda(m)b_k\chi(mk)\Big|. \tag{5.18}$$

Next, we use Perron's formula (Lemma 1.13) with $\alpha = (\log x)^{-1}$, $T = x^2$, and $u = y/(mk)$. We get

$$\sum_{M<m\le M_1}\sum_{U<k\le y/m}\Lambda(m)b_k\chi(mk) = \frac{1}{2\pi}\int_{-T}^{T} S(\chi,t)\frac{y^{\alpha+\mathrm{i}t}}{\alpha+\mathrm{i}t}\,dt + O(\Delta), \tag{5.19}$$

where

$$S(\chi,t) = \sum_{M<m\le M_1}\sum_{U<k\le xM^{-1}}\Lambda(m)b_k\chi(mk)(mk)^{-\alpha-\mathrm{i}t}, \qquad \Delta = \frac{y^\alpha}{T}\sum_{M<m\le M_1}\sum_{k\le yM^{-1}}\frac{\Lambda(m)d(k)}{|\log(y/mk)|}.$$

If we assume, as we may, that $\|y\| = \frac{1}{2}$, we have $|\log(y/mk)| \ge y^{-1}$. Hence, by the PNT and Theorem 1.22,

$$\Delta \ll x^{-1}\sum_{M<m\le M_1}\sum_{k\le yM^{-1}}\Lambda(m)d(k) \ll \log x.$$

Note also that

$$\int_{-T}^{T}|\alpha+\mathrm{i}t|^{-1}\,dt \ll \log x.$$

Substituting these bounds into (5.19), we find that, for some $|t_0| \le T$,

$$\sum_{M<m\le M_1}\sum_{U<k\le y/m}\Lambda(m)b_k\chi(mk) \ll \big(|S(\chi,t_0)|+1\big)\log x.$$

Since $S(\chi,t_0)$ is independent of $y$, combining this inequality and (5.18), we get

$$\max_{y\le x}|S_3(y,\chi)| \ll (\log x)^2\big(|S(\chi,t_0)|+1\big).$$

Thus,

$$\Sigma_5 \ll (\log x)^2\sum_{Q_0<r\le Q}\frac{1}{\phi(r)}\sideset{}{^*}\sum_{\chi \bmod r}|S(\chi,t_0)| + Q(\log x)^2. \tag{5.20}$$

We now observe that

$$\sum_{M<m\le M_1}\Lambda(m)^2m^{-2\alpha} \ll M\log x \quad\text{and}\quad \sum_{U<k\le xM^{-1}}d(k)^2k^{-2\alpha} \ll xM^{-1}(\log x)^3;$$

the former bound follows from the PNT and the latter from Theorem 1.23. Hence, (5.17) yields

$$\sum_{r\le R}\frac{r}{\phi(r)}\sideset{}{^*}\sum_{\chi \bmod r}|S(\chi,t_0)| \ll (\log x)^2\big(x + xRU^{-1/2} + x^{1/2}R^2\big).$$

From this inequality and (5.20), we derive

$$\Sigma_5 \ll (\log x)^4\big(xQ_0^{-1} + xU^{-1/2}(\log x) + x^{1/2}Q\big). \tag{5.21}$$

### 5.2.4 Completion of the proof

Suppose that $U = V \le x^{1/3}$. Then we can write $S_2(y,\chi)$ as

$$S_2(y,\chi) = S_2'(y,\chi) + S_2''(y,\chi),$$

where $S_2'$ is the portion of $S_2$ where $m \le U$ and $S_2''$ is the portion where $U < m \le UV \le xU^{-1}$. We have

$$|S_2'(y,\chi)| \le (\log x) \sum_{m \le U} \Big| \sum_{U < mk \le y} \chi(k) \Big|,$$

so the contribution of $S_2'$ to $\Sigma_4$ can be bounded similarly to $\Sigma_3$. Moreover, we can estimate the contribution of $S_2''$ to $\Sigma_4$ similarly to $\Sigma_5$, and the resulting bound is slightly sharper, because in this case we apply (5.17) with coefficients $a_m$ and $b_k$ subject to $|a_m| \le \log m$ and $|b_k| \le 1$. Altogether, we conclude that

$$\Sigma_4 \ll (\log x)^4 \big( Q^{3/2}U + xQ_0^{-1} + xU^{-1/2} + x^{1/2}Q \big). \tag{5.22}$$

Combining (5.10), (5.12), (5.14)–(5.16), (5.21), and (5.22), we get

$$\Sigma(x,Q) \ll (\log x)^5 \big( Q^{3/2}U + xQ_0^{-1} + xU^{-1/2}(\log x) + x^{1/2}Q \big),$$

where $U \le x^{1/3}$ is a parameter at our disposal. Any choice of $U$ subject to

$$(\log x)^{2A+12} \le U \le (x/Q)^{1/2}$$

then yields (5.9). This proves the theorem when $Q \le x(\log x)^{-4A-24}$; in the alternative case, (5.9) is worse than the trivial bound for $\Sigma(x,Q)$. ∎

## 5.3 The Barban–Davenport–Halberstam theorem

Using the large sieve and reductions such as those leading to (5.10), we can also establish the following result.

**Theorem 5.3.** *Suppose that $2 \le Q \le x$. Then, for any fixed $A > 0$,*

$$\sum_{q \le Q} \sum_{\substack{1 \le a \le q \\ (a,q)=1}} \Big| \psi(x;q,a) - \frac{x}{\phi(q)} \Big|^2 \ll x^2(\log x)^{-A} + Qx \log x. \tag{5.23}$$

The first results of this form were obtained by Barban [3] and Davenport and Halberstam [15]; hence, the name of the theorem. Note that by (5.23), the error term in the prime number theorem for arithmetic progressions is $O\big((x/q)^{1/2+\epsilon}\big)$, at least on average over all progressions to moduli $q \le Q$. Except when the modulus $q$ is very small, so strong a bound for an individual progression does not follow even from GRH! Furthermore, (5.23) appears to be (essentially) the best result within the reach of present methods. Indeed, a substantial improvement of the first term on the right side of (5.23) would yield a subsequent improvement on Theorem 2 (see Exercise 3). While

such an improvement would represent a significant achievement in the field, it seems more likely that it would occur independent of the above problem than as a consequence to it. As to the second term in the bound (5.23), Montgomery [41, Ch. 17] has shown that when $Q \geq x(\log x)^{-A}$,

$$\sum_{q\leq Q}\sum_{\substack{1\leq a\leq q\\(a,q)=1}}\left|\psi(x;q,a)-\frac{x}{\phi(q)}\right|^2 \sim Qx\log x \qquad \text{as } x\to\infty.$$

Thus, the second term on the right side of (5.23) is needed when $Q$ is large.

## 5.4 The three primes theorem

Our goal in this section is to establish the following result from additive prime number theory.

**Theorem 5 (I. M. Vinogradov).** *For a positive integer $n$, define*

$$R(n)=\sum_{p_1+p_2+p_3=n}(\log p_1)(\log p_2)(\log p_3),$$

*where the summation is over all representations of $n$ as the sum of three primes. Then, for any given $A>0$,*

$$R(n)=\tfrac{1}{2}n^2\mathfrak{S}(n)+O\big(n^2(\log n)^{-A}\big), \tag{5.24}$$

*where*

$$\mathfrak{S}(n)=\prod_{p\mid n}\big(1-(p-1)^{-2}\big)\prod_{p\nmid n}\big(1+(p-1)^{-3}\big). \tag{5.25}$$

*In particular, every sufficiently large odd integer is the sum of three primes.*

This theorem was first proved in 1923 by Hardy and Littlewood [21] under the assumption of GRH. In 1937 Vinogradov [57] applied his method for estimating sums over primes to the exponential sum $f(\alpha)$ below to give an unconditional proof of the three primes theorem.

### 5.4.1 The Hardy–Littlewood circle method

Using the orthogonality relation

$$\int_0^1 e(\alpha m)\,d\alpha=\begin{cases}1 & \text{if } m=0,\\ 0 & \text{if } m\neq 0,\end{cases} \tag{5.26}$$

we can express $R(n)$ as a Fourier integral. Indeed, by (5.26),

$$\begin{aligned}R(n)&=\sum_{p_1,p_2,p_3\leq n}(\log p_1)(\log p_2)(\log p_3)\int_0^1 e\big(\alpha(p_1+p_2+p_3-n)\big)\,d\alpha\\&=\int_0^1\left(\sum_{p\leq n}(\log p)e(\alpha p)\right)^3 e(-\alpha n)\,d\alpha.\end{aligned} \tag{5.27}$$

This identity is the starting point of the application of the circle method: we will use it to derive the asymptotic formula for $R(n)$ from estimates for the exponential sum

$$f(\alpha) = \sum_{p \le n} (\log p) e(\alpha p). \tag{5.28}$$

Our analysis is motivated by two observations:

- when $\alpha$ is near a rational number $a/q$ with a small denominator, $f(\alpha)$ should be large and should have certain asymptotic behavior, suggested by the behavior of $f(a/q)$;
- otherwise, the numbers $e(\alpha p)$, $p \le n$, should be approximately uniformly distributed on the unit circle, and hence, $f(\alpha)$ should be "small".

Let $B = B(A)$ be a positive number to be chosen later and set

$$P = (\log n)^B. \tag{5.29}$$

If $a$ and $q$ are integers, we define the *major arc*[1]

$$\mathfrak{M}(q,a) = \big[a/q - P/(qn), a/q + P/(qn)\big]. \tag{5.30}$$

The integration in (5.27) can be taken over any interval of unit length, and in particular, over $\big[Pn^{-1}, 1 + Pn^{-1}\big]$. We partition this interval into two subsets:

$$\mathfrak{M} = \bigcup_{q \le P} \bigcup_{\substack{1 \le a \le q \\ (a,q)=1}} \mathfrak{M}(q,a) \qquad \text{and} \qquad \mathfrak{m} = \big[Pn^{-1}, 1 + Pn^{-1}\big] \setminus \mathfrak{M}, \tag{5.31}$$

called respectively the *set of major arcs* and the *set of minor arcs*. Then, from (5.27) and (5.31),

$$R(n) = \left( \int_{\mathfrak{M}} + \int_{\mathfrak{m}} \right) f(\alpha)^3 e(-\alpha n)\, d\alpha. \tag{5.32}$$

In §5.4.2, we use the Siegel–Walfisz theorem to prove that

$$\int_{\mathfrak{M}} f(\alpha)^3 e(-\alpha n)\, d\alpha = \tfrac{1}{2} n^2 \mathfrak{S}(n) + O\big(n^2 P^{-1}\big) \tag{5.33}$$

for any choice of $P$. Then, in §5.4.3, we show that

$$\int_{\mathfrak{m}} f(\alpha)^3 e(-\alpha n)\, d\alpha \ll n^2 (\log n)^{-A} \tag{5.34}$$

for $B \ge 3A + 18$. Clearly, the asymptotic formula (5.24) follows from (5.32)–(5.34).

[1] This term may seem a little peculiar, considering that $\mathfrak{M}(q,a)$ is in fact an interval. The explanation is that, in the original version of the circle method, Hardy and Littlewood used power series and Cauchy's integral formula instead of exponential sums and (5.26) (see Vaughan [54, §1.2]). In that setting, the role of $\mathfrak{M}(q,a)$ is played by a small circular arc near the root of unity $e(a/q)$; hence, the terminology.

### 5.4.2 The major arcs

It is easy to see that the major arcs are comprised of mutually disjoint intervals $\mathfrak{M}(q,a)$. Thus,

$$\int_{\mathfrak{M}} f(\alpha)^3 e(-\alpha n)\,d\alpha = \sum_{q\le P}\sum_{\substack{1\le a\le q\\(a,q)=1}} \int_{\mathfrak{M}(q,0)} f(a/q+\beta)^3 e\big(-(a/q+\beta)n\big)\,d\beta. \tag{5.35}$$

We now proceed to approximate $f(a/q+\beta)$. We will prove the following result.

**Lemma 5.4.** *Suppose that $B>0$, $1\le a\le q\le(\log n)^B$, $(a,q)=1$, $|\beta|\le n^{-1}(\log n)^B$. Suppose also that $f(\alpha)$ is defined by* (5.28) *and define*

$$v(\beta)=\int_0^n e(\beta x)\,dx.$$

*Then*

$$f(a/q+\beta)=\mu(q)\phi(q)^{-1}v(\beta)+O\big(n(\log n)^{-B}\big).$$

*Proof.* We split the summation in $f(\alpha)$ according to the residue of $p$ modulo $q$. We get

$$\begin{aligned} f(a/q+\beta) &= \sum_{1\le h\le q}\sum_{\substack{p\le n\\ p\equiv h \pmod q}} (\log p)e\big((a/q+\beta)p\big)\\ &= \sum_{1\le h\le q} e(ah/q)\sum_{\substack{p\le n\\ p\equiv h \pmod q}} (\log p)e(\beta p)\\ &= \sum_{\substack{1\le h\le q\\(h,q)=1}} e(ah/q)\sum_{\substack{p\le n\\ p\equiv h \pmod q}} (\log p)e(\beta p)+O(q). \end{aligned} \tag{5.36}$$

When $(h,q)=1$, we have

$$\begin{aligned} \sum_{\substack{p\le n\\ p\equiv h \pmod q}} (\log p)e(\beta p) &= \sum_{\substack{m\le n\\ m\equiv h \pmod q}} \Lambda(m)e(\beta m)+O\big(\sqrt{n}\big)\\ &= \int_0^n e(\beta x)\,d\psi(x;q,h)+O\big(\sqrt{n}\big). \end{aligned} \tag{5.37}$$

By the Siegel–Walfisz theorem in the form of (3.62),

$$\Delta(x;q,h)=\psi(x;q,a)-x/\phi(q)\ll n(\log n)^{-3B},$$

for all $x\le n$. Hence,

$$\begin{aligned} \int_0^n e(\beta x)\,d\Delta(x;q,h) &\ll |\Delta(n;q,h)|+1+|\beta|\int_2^n |\Delta(x;q,h)|\,dx\\ &\ll n(\log n)^{-3B}+|\beta|\int_0^n n(\log n)^{-3B}\,dx \ll n(\log n)^{-2B}. \end{aligned}$$

Combining this estimate and (5.37), we get

$$\sum_{\substack{p\le n\\ p\equiv h \pmod q}} (\log p)e(\beta p) = \frac{1}{\phi(q)}\int_0^n e(\beta x)\,dx + O\big(n(\log n)^{-2B}\big). \tag{5.38}$$

Since (see Exercise 4.3)

$$c_q(a) = \sum_{\substack{1\le h\le q\\ (h,q)=1}} e(ah/q) = \mu(q),$$

the desired conclusion follows from (5.36) and (5.38). ∎

By Lemma 5.4 with $3B$ in place of $B$,

$$f(a/q+\beta)^3 = \mu(q)\phi(q)^{-3}v(\beta)^3 + O\big(n^3P^{-3}\big).$$

Since the measure of $\mathfrak{M}$ is $O\big(P^2n^{-1}\big)$, inserting this approximation into the right side of (5.35), we obtain

$$\int_{\mathfrak{M}} f(\alpha)^3e(-\alpha n)\,d\alpha = \sum_{q\le P}\frac{\mu(q)c_q(-n)}{\phi(q)^3}\int_{\mathfrak{M}(q,0)} v(\beta)^3e(-\beta n)\,d\beta + O\left(n^2P^{-1}\right). \tag{5.39}$$

At this point, we extend the integration over $\beta$ to the whole real line. Since

$$v(\beta) \ll n(1+n|\beta|)^{-1}, \tag{5.40}$$

the error we incur from doing this is

$$\ll \sum_{q\le P}\phi(q)^{-2}\int_{P/(qn)}^{\infty}\frac{n^3\,d\beta}{(1+n\beta)^3} \ll n^2P^{-2}\sum_{q\le P}\frac{q^2}{\phi(q)^2} \ll n^2P^{-1}.$$

The last step uses the result of Exercise 5, which at the same time implies that

$$\sum_{q\le P}\mu(q)c_q(-n)\phi(q)^{-3} \ll \sum_{q\le P}\phi(q)^{-2} \ll 1.$$

Hence, we deduce from (5.39) that

$$\int_{\mathfrak{M}} f(\alpha)^3e(-\alpha n)\,d\alpha = \mathfrak{S}(n,P)J(n) + O\big(n^2P^{-1}\big),$$

where

$$\mathfrak{S}(n,X) = \sum_{q\le X}\mu(q)c_q(-n)\phi(q)^{-3}, \qquad J(n) = \int_{-\infty}^{\infty} v(\beta)^3e(-\beta n)\,d\beta. \tag{5.41}$$

By Fourier's inversion formula, $J(n) = \frac{1}{2}n^2$, and by Exercise 5,

$$\mathfrak{S}(n,P) - \mathfrak{S}(n,\infty) \ll \sum_{q>P}\phi(q)^{-2} \ll P^{-1}.$$

It follows that

$$\int_{\mathfrak{M}} f(\alpha)^3 e(-\alpha n)\, d\alpha = \tfrac{1}{2} n^2 \mathfrak{S}(n,\infty) + O\big(n^2 P^{-1}\big).$$

Finally, we remark that $\mathfrak{S}(n,\infty)$ equals the product $\mathfrak{S}(n)$ defined in (5.25). Indeed, this follows from Lemma 1.16, on noting that the function $g(q) = \mu(q) c_q(-n)\phi(q)^{-3}$ is multiplicative and

$$g(p^u) = \begin{cases} (p-1)^{-3} & \text{if } u = 1 \text{ and } p \nmid n, \\ -(p-1)^{-2} & \text{if } u = 1 \text{ and } p \mid n, \\ 0 & \text{if } u \geq 2. \end{cases}$$

Thus, (5.33) is established.

### 5.4.3 The minor arcs

We now turn to (5.34). The modulus of the left side does not exceed

$$\int_{\mathfrak{m}} |f(\alpha)|^3\, d\alpha \leq \Big( \sup_{\mathfrak{m}} |f(\alpha)| \Big) \int_0^1 |f(\alpha)|^2\, d\alpha. \tag{5.42}$$

By Parseval's identity and the PNT,

$$\int_0^1 |f(\alpha)|^2\, d\alpha = \sum_{p \leq n} (\log p)^2 \ll n \log n.$$

Thus, (5.34) will follow from (5.42), if we show that

$$\sup_{\mathfrak{m}} |f(\alpha)| \ll n (\log n)^{-A-1}. \tag{5.43}$$

We note that the trivial estimate for $f(\alpha)$ is

$$f(\alpha) \ll \sum_{p \leq n} (\log p) \ll n,$$

so our goal is to save a power of $\log n$ over the trivial estimate for $f(\alpha)$. We can do this using the following lemma, which provides such a saving under the assumption that $\alpha$ can be approximated by a reduced fraction whose denominator $q$ is "neither too small, nor too large."

**Lemma 5.5.** *Suppose that $\alpha, \delta$ are real and $a, q$ are integers satisfying*

$$1 \leq q \leq n, \quad (a,q) = 1, \quad |\alpha - a/q| \leq \delta.$$

*Then*

$$f(\alpha) \ll (\log n)^5 (1 + \delta n) \big( n q^{-1/2} + n^{5/6} + n^{2/3} q^{1/3} \big).$$

This estimate for $f(\alpha)$ is not quite the best known, but it has the advantage that it can be deduced quickly from the work in §5.2. The sharpest known bound for $f(\alpha)$ is

$$f(\alpha) \ll (\log n)^4 \big(nq^{-1/2} + n^{4/5} + n^{1/2}q^{1/2}\big); \tag{5.44}$$

this holds under the standard assumption that $|\alpha - a/q| \le q^{-2}$. For the proof of (5.44) see Vaughan [54, Theorem 3.1] or Exercise 7 after the chapter. It is also possible to apply more carefully the ideas used in the proof of Lemma 5.5 to prove, again when $|\alpha - a/q| \le q^{-2}$, that

$$f(\alpha) \ll (\log n)^4 \big(nq^{-1/2} + n^{7/8}q^{-1/8} + n^{3/4}q^{1/8} + n^{1/2}q^{1/2}\big).$$

For the proof of this result, see Vaughan [52]; that paper is also where identity (5.6) first appeared and contains a proof of the (so far) strongest version of the Bombieri–Vinogradov theorem.

*Proof.* We will derive the lemma from the inequality

$$\frac{\sqrt{q}}{\phi(q)} \sum_{\chi \bmod q} |\psi(x,\chi)| \ll (\log qx)^5 \big(xq^{-1/2} + x^{5/6} + x^{2/3}q^{1/3}\big). \tag{5.45}$$

First, we note that the contribution from the principal character modulo $q$ is

$$\le xq^{1/2}\phi(q)^{-1} \ll xq^{-1/2} \log\log q,$$

by Exercise 5(a). We estimate the average over the non-principal characters similarly to the sum $\Sigma_2$ in the proof of Theorem 5.2. In place of (5.15), we have

$$\frac{\sqrt{q}}{\phi(q)} \sum_{\substack{\chi \bmod q \\ \chi \ne \chi_0}} |\psi(x,\chi)| \ll q^{-1/2} (\log qx) \big(\Sigma_3' + \Sigma_4' + \Sigma_5' + qU\big),$$

where $\Sigma_j'$ is similar to $\Sigma_j$ and $1 \le U \le x^{1/3}$ is a parameter at our disposal. We can estimate each $\Sigma_j'$ analogously to the respective $\Sigma_j$, the only difference being that instead of (5.17), we appeal to the inequality

$$\sum_{\chi \bmod q} \left| \sum_{m=1}^{M} \sum_{k=1}^{K} a_m b_k \chi(mk) \right| \ll \big(M+q\big)^{1/2} \big(K+q\big)^{1/2} \left( \sum_{m=1}^{M} |a_m|^2 \right)^{1/2} \left( \sum_{k=1}^{K} |b_k|^2 \right)^{1/2}.$$

Altogether, we obtain

$$\frac{\sqrt{q}}{\phi(q)} \sum_{\chi \bmod q} |\psi(x,\chi)| \ll (\log qx)^5 \big(xq^{-1/2} + xU^{-1/2} + x^{1/2}q^{1/2} + qU\big).$$

We now choose

$$U = \min\big(x^{1/3}, (x/q)^{2/3}\big),$$

so that

$$xU^{-1/2} \ll x^{5/6} + x^{2/3}q^{1/3} \qquad \text{and} \qquad qU \ll x^{2/3}q^{1/3}.$$

Since $x^{1/2}q^{1/2} \le x^{2/3}q^{1/3}$, (5.45) follows.

We now turn to the estimation of $f(\alpha)$. We have

$$f(\alpha) = \sum_{m\le n} \Lambda(m)e(\alpha m) + O\big(n^{1/2}\big). \tag{5.46}$$

By partial summation,

$$\sum_{m\le n} \Lambda(m)e(\alpha m) \ll \big(1 + n|\alpha - a/q|\big)\left|\sum_{m\le x} \Lambda(m)e(am/q)\right|, \tag{5.47}$$

for some $x \le n$. Similarly to (3.56) and (3.58), we obtain

$$\begin{aligned}\sum_{m\le x} \Lambda(m)e(am/q) &= \sum_{\substack{1\le h\le q\\(h,q)=1}} e(ah/q)\psi(x;q,h) + O\big((\log qx)^2\big)\\ &= \frac{1}{\phi(q)} \sum_{\chi \bmod q} \sum_{\substack{1\le h\le q\\(h,q)=1}} e(ah/q)\bar{\chi}(h)\psi(x,\chi) + O\big((\log qx)^2\big)\\ &= \frac{1}{\phi(q)} \sum_{\chi \bmod q} \tau(\bar{\chi},a)\psi(x,\chi) + O\big((\log qx)^2\big).\end{aligned}$$

Since $(a,q) = 1$, it follows from Lemmas 3.5 and 3.6 that $|\tau(\bar{\chi},a)| \le \sqrt{q}$. Hence,

$$\sum_{m\le x} \Lambda(m)e(am/q) \ll \frac{\sqrt{q}}{\phi(q)} \sum_{\chi \bmod q} |\psi(x,\chi)| + (\log qx)^2, \tag{5.48}$$

and the desired conclusion follows from (5.45)–(5.48). ∎

Before we can derive (5.43) from Lemma 5.5, we need to state a simple lemma known as *Dirichlet's theorem on Diophantine approximation*.

**Lemma 5.6 (Dirichlet).** *Let $\alpha$ and $Q$ be real and $Q \ge 1$. There exist integers $a$ and $q$ such that*

$$1 \le q \le Q, \qquad (a,q) = 1, \qquad |q\alpha - a| < Q^{-1}.$$

*Proof.* See Vaughan [54, Lemma 2.1] or Exercises 8 and 9. ∎

*Proof of* (5.43). Suppose that $\alpha \in \mathfrak{m}$. By Lemma 5.6 with $Q = nP^{-1}$, there are integers $a$ and $q$ such that

$$1 \le q \le nP^{-1}, \qquad (a,q) = 1, \qquad |q\alpha - a| < Pn^{-1}.$$

Since $\alpha \notin \mathfrak{M}$, it is not possible to have $q \le P$, and so,

$$P \le q \le nP^{-1}, \qquad (a,q) = 1, \qquad |\alpha - a/q| < n^{-1}.$$

We now apply Lemma 5.5 (with $\delta = n^{-1}$) and obtain

$$\begin{aligned} f(\alpha) &\ll (\log n)^5 \big(nq^{-1/2} + n^{5/6} + n^{2/3}q^{1/3}\big) \\ &\ll (\log n)^5 \big(nP^{-1/2} + n^{5/6} + nP^{-1/3}\big) \ll n(\log n)^{5-B/3}. \end{aligned}$$

Thus, (5.43) follows on choosing $B \ge 3A + 18$. ∎

## 5.5 Primes in short intervals

In this section we discuss the following result mentioned in the Introduction.

**Theorem 6 (Huxley).** *Let $\epsilon > 0$ be fixed. Then for $x \ge x_0(\epsilon)$ and $x^{7/12+\epsilon} \le y \le x$,*

$$\psi(x) - \psi(x-y) = y + O\big(y(\log x)^{-1}\big). \tag{5.49}$$

We deduce Huxley's theorem from the following two results.

**Theorem 5.7 (Korobov; I. M. Vinogradov).** *There is an absolute constant $c_1 > 0$ such that*

$$\beta \ge 1 - c_1(\log(|\gamma|+3))^{-2/3}(\log\log(|\gamma|+3))^{-1/3}.$$

**Theorem 5.8 (Huxley).** *Given $0 \le \sigma \le 1$ and $T \ge 2$, define*

$$N(\sigma, T) = \big\{\rho = \beta + \mathrm{i}\gamma \ : \ \zeta(\rho) = 0,\ \sigma \le \beta \le 1,\ |\gamma| \le T\big\}.$$

*There is an absolute constant $c_2 > 0$ such that*

$$N(\sigma, T) \ll T^{2.4(1-\sigma)}(\log T)^{c_2}.$$

Theorem 5.7 is the Vinogradov–Korobov zerofree region underlying the modern error term (0.9) in the PNT. The reader will find its proof in Ivić [31], Karatsuba and Voronin [37], or Titchmarsh [50]. Theorem 5.8, whose proof forms the bulk of this section, is an example of a *zero-density theorem*. By virtue of Corollary 2.20, we have the trivial bound

$$N(\sigma, T) \le N(0, T) \ll T(\log T). \tag{5.50}$$

A zero-density theorem is an inequality of the form

$$N(\sigma, T) \ll T^{A(\sigma)(1-\sigma)}(\log T)^{c(\sigma)}, \tag{5.51}$$

where $A(\sigma)$ and $c(\sigma)$ are such that (5.51) represents an improvement over (5.50) for $\sigma$ in some subinterval of $1/2 \le \sigma \le 1$ (when $\sigma \le 1/2$, (5.50) is best possible). Results in which $A(\sigma)$ is a constant are of particular interest for applications. Theorem 5.8 above was established by Huxley [27] and is the sharpest known result of this type. It should be compared with the conjectural bound

$$N(\sigma, T) \ll T^{2(1-\sigma)}(\log T)^{c_3} \qquad (1/2 \le \sigma \le 1), \tag{5.52}$$

which is known as the Density Hypothesis and in many situations can be used as a substitute for RH.

### 5.5.1 Proof of Theorem 6

By Theorem 2.22 with $T = x^{5/12-\epsilon/2}$,

$$\psi(x) - \psi(x-y) = y - \sum_{|\operatorname{Im}\rho|\le T} \frac{x^\rho - (x-y)^\rho}{\rho} + O\big(x^{7/12+\epsilon/2}(\log x)^2\big). \tag{5.53}$$

On writing $\rho = \beta + \mathrm{i}\gamma$, we have

$$\left|\frac{x^\rho - (x-y)^\rho}{\rho}\right| = \left|\int_{x-y}^{x} u^{\rho-1}\,du\right| \le yx^{\beta-1},$$

whence

$$\sum_{|\operatorname{Im}\rho|\le T} \frac{x^\rho - (x-y)^\rho}{\rho} \ll y\sum_{|\gamma|\le T} x^{\beta-1}. \tag{5.54}$$

Clearly, (5.49) follows from (5.53), (5.54), and the inequality

$$\sum_{|\gamma|\le T} x^{\beta-1} \ll (\log x)^{-1}, \tag{5.55}$$

which we now proceed to prove.

Let

$$\delta(T) = c_1(\log T)^{-2/3}(\log\log T)^{-1/3},$$

where $c_1$ is the constant from Theorem 5.7. By Theorem 5.7 and partial integration,

$$\begin{aligned}\sum_{|\operatorname{Im}\rho|\le T} x^{\beta-1} &= -\int_0^{1-\delta(T)} x^{\sigma-1}\,dN(\sigma,T)\\ &= x^{-1}N(0,T) + (\log x)\int_0^{1-\delta(T)} N(\sigma,T)x^{\sigma-1}\,d\sigma.\end{aligned}$$

We use (5.50) to bound $N(0,T)$ and Theorem 5.8 to bound $N(\sigma,T)$ under the sign of the integral. We find that

$$\begin{aligned}\sum_{|\operatorname{Im}\rho|\le T} x^{\beta-1} &\ll x^{-1}T\log T + (\log x)^{c_2+1}\int_0^{1-\delta(T)} \big(x^{-1}T^{2.4}\big)^{1-\sigma}\,d\sigma\\ &\ll x^{-1/2} + (\log x)^{c_2+1}x^{-\epsilon\delta(T)} \ll (\log x)^{-1},\end{aligned}$$

on noting that

$$x^{-\epsilon\delta(T)} \ll \exp\big(-\epsilon c_4(\log x)^{1/4}\big) \ll (\log x)^{-c_2-2}.$$

This establishes (5.55) and completes the proof of the theorem. ∎

### 5.5.2 Huxley's density theorem: zero detection

Without loss of generality we may assume that $7/12 \le \sigma \le 1$ and $T \ge T_0$. We start from the integral transform

$$e^{-x} = \frac{1}{2\pi i}\int_{2-i\infty}^{2+i\infty} \Gamma(s)x^{-s}\,ds \qquad (x>0), \tag{5.56}$$

which follows from (2.10) by Mellin inversion. We introduce parameters $X$ and $Y$, which we will specify later. We define

$$M_X(s) = \sum_{m\le X} \mu(m)m^{-s}$$

and observe that, by Lemma 1.2,

$$F_X(s) = \zeta(s)M_X(s) = \sum_{n=1}^{\infty} a_n n^{-s} = 1 + \sum_{n>X} a_n n^{-s}, \qquad a_n = \sum_{\substack{k\mid n\\ k\le X}} \mu(k) \ll d(n).$$

Applying (5.56) with $x = n/Y$ and summing the resulting identities over $n$, we get

$$e^{-1/Y} + \sum_{n>X} a_n n^{-w} e^{-n/Y} = \frac{1}{2\pi i}\int_{2-i\infty}^{2+i\infty} F_X(w+s)\Gamma(s)Y^s\,ds \qquad (\mathrm{Re}(w)>0). \tag{5.57}$$

Suppose that $\rho = \beta + i\gamma$ is a zero of $\zeta(s)$ counted by $N(\sigma, T)$. We move the integral in (5.57) to the line $\mathrm{Re}\, s = \frac{1}{2} - \beta$. The only singularities of the integrand in the strip $\frac{1}{2} - \beta < \mathrm{Re}\, s < 2$ are the pole of $\zeta(s+w)$ at $s = 1-w$ and the pole of $\Gamma(s)$ at $s = 0$. Furthermore, when $w = \rho$, $\zeta(\rho + s)$ has a zero at $s = 0$ that cancels the pole of the gamma-function. Hence,

$$e^{-1/Y} + \sum_{n>X} a_n n^{-\rho} e^{-n/Y} = M_X(1)\Gamma(1-\rho)Y^{1-\rho} + \frac{1}{2\pi i}\int_{1/2-\beta-i\infty}^{1/2-\beta+i\infty} F_X(\rho+s)\Gamma(s)Y^s\,ds. \tag{5.58}$$

In order to simplify this identity, we now suppose that

$$T^{0.01} \le X \le T^{10} \qquad \text{and} \qquad T^{0.01} \le Y \le T^{10}. \tag{5.59}$$

The terms with $n \ge Y(\log T)^2$ contribute $o(1)$ to the left side of (5.59). Also, by Corollary 2.8, the first term on the right of (5.59) is $o(1)$ unless $|\gamma| \le (\log T)^2$. Therefore, apart from the zeros counted by $N\big(\sigma, (\log T)^2\big)$, all zeros counted by $N(\sigma, T)$ fall in one of the following classes:

- *Class I:* zeros $\rho$ with
$$\left| \sum_{X<n\le Y(\log T)^2} a_n n^{-\rho} e^{-n/Y} \right| > 1/3; \tag{5.60}$$
- *Class II:* zeros $\rho$ with
$$\left| \int_{1/2-\beta-i\infty}^{1/2-\beta+i\infty} F_X(s+\rho)\Gamma(s)Y^s\,ds \right| > 1/3. \tag{5.61}$$

We subdivide the interval $X < n \le Y(\log T)^2$ into $O(\log T)$ subintervals $N < n \le N_1$, $N_1 \le 2N$, and note that the number of zeros of class I does not exceed $R_1(\log T)$, where $R_1$ is the number of zeros $\rho$ with

$$\left| \sum_{N<n\le N_1} a_n n^{-\rho} e^{-n/Y} \right| \gg (\log T)^{-1}. \tag{5.62}$$

Furthermore, by Corollary 2.20,

$$R_1 \ll |\mathcal{R}_1|(\log T),$$

where $\mathcal{R}_1$ is a set of zeros of $\zeta(s)$ that satisfy (5.62) and

$$|\operatorname{Im}\rho_1 - \operatorname{Im}\rho_2| \ge 1 \qquad \text{whenever } \rho_1 \ne \rho_2,\ \rho_i \in \mathcal{R}_1. \tag{5.63}$$

Similarly, the number of class II zeros is bounded by $|\mathcal{R}_2|(\log T)$, where $\mathcal{R}_2$ is a set of zeros of $\zeta(s)$ that satisfy (5.61) and (5.63). We conclude that

$$\begin{aligned} N(\sigma, T) &\ll N\big(\sigma, (\log T)^2\big) + |\mathcal{R}_1|(\log T)^2 + |\mathcal{R}_2|(\log T) \\ &\ll (|\mathcal{R}_1| + |\mathcal{R}_2| + \log T)(\log T)^2. \end{aligned} \tag{5.64}$$

### 5.5.3 Huxley's density theorem: $1/2 \le \sigma \le 3/4$

We first bound $|\mathcal{R}_1|$. We write $b_n = a_n e^{-n/Y}(\log n)$ and denote by $\mathcal{S}_1$ the set of imaginary parts $\gamma$ of zeros $\rho \in \mathcal{R}_1$. Then

$$\begin{aligned} |\mathcal{R}_1| &\ll (\log T)^2 \sum_{\rho\in\mathcal{R}_1} \left| \sum_{N<n\le N_1} b_n \int_\beta^\infty n^{-u-i\gamma}\, du \right|^2 \\ &\ll (\log T)^2 \sum_{\gamma\in\mathcal{S}_1} \left\{ \int_\sigma^\infty \left| \sum_{N<n\le N_1} b_n n^{-u-i\gamma} \right| du \right\}^2 \\ &\ll (\log T)^2 N^{-\sigma} \sum_{\gamma\in\mathcal{S}_1} \int_\sigma^\infty N^u \left| \sum_{N<n\le N_1} b_n n^{-u-i\gamma} \right|^2 du \\ &\ll (\log T)^2 N^{-\sigma} \int_\sigma^\infty N^u \sum_{\gamma\in\mathcal{S}_1} \left| \sum_{N<n\le N_1} b_n n^{-u-i\gamma} \right|^2 du. \end{aligned}$$

An appeal to Lemma 4.12 with $\delta = 1$ now gives

$$\begin{aligned} |\mathcal{R}_1| &\ll N^{-\sigma}(N+T)(\log T)^2 \int_\sigma^\infty N^u \sum_{N<n\le 2N} |b_n|^2 n^{-2u}\, du \\ &\ll N^{-2\sigma}(N+T)(\log T) \sum_{N<n\le 2N} d(n)^2 (\log n)^2 \ll \big(N^{2-2\sigma} + TN^{1-2\sigma}\big)(\log T)^6, \end{aligned}$$

by Theorem 1.23 with $k = 2$. Recalling that $X \le N \le Y(\log T)^2$, we deduce that

$$|\mathcal{R}_1| \ll \big(Y^{2-2\sigma} + TX^{1-2\sigma}\big)(\log T)^8. \tag{5.65}$$

We now turn to class II zeros. By (5.61),

$$|\mathcal{R}_2| \ll \sum_{\rho\in\mathcal{R}_2} Y^{2(1-2\beta)/3}\left\{\int_{-\infty}^{\infty}\left|F_X(1/2+\mathrm{i}(t+\gamma))\Gamma(1/2-\beta+\mathrm{i}t)\right|dt\right\}^{4/3}. \tag{5.66}$$

Since $7/12 \le \beta \le 1$, Lemmas 2.4 and 2.5 and Corollary 2.8 yield

$$\Gamma(1/2-\beta+\mathrm{i}t) \ll e^{-|t|}. \tag{5.67}$$

Thus, by (5.66) and Hölder's inequality,

$$\begin{aligned}|\mathcal{R}_2| &\ll Y^{2(1-2\sigma)/3}\sum_{\gamma\in\mathcal{S}_2}\left\{\int_{-\infty}^{\infty}\left|F_X(1/2+\mathrm{i}(t+\gamma))\right|e^{-|t|}\,dt\right\}^{4/3}\\ &\ll Y^{2(1-2\sigma)/3}\sum_{\gamma\in\mathcal{S}_2}\int_{-\infty}^{\infty}\left|F_X(1/2+\mathrm{i}(t+\gamma))\right|^{4/3}e^{-|t|}\,dt\\ &\ll Y^{2(1-2\sigma)/3}\Sigma_1^{1/3}\Sigma_2^{2/3},\end{aligned} \tag{5.68}$$

where $\mathcal{S}_2$ denotes the set of imaginary parts of the zeros in $\mathcal{R}_2$,

$$\begin{aligned}\Sigma_1 &= \sum_{\gamma\in\mathcal{S}_2}\int_{-\infty}^{\infty}\left|\zeta(1/2+\mathrm{i}(t+\gamma))\right|^4 e^{-|t|}\,dt,\\ \Sigma_2 &= \sum_{\gamma\in\mathcal{S}_2}\int_{-\infty}^{\infty}\left|M_X(1/2+\mathrm{i}(t+\gamma))\right|^2 e^{-|t|}\,dt.\end{aligned}$$

We have

$$\begin{aligned}\Sigma_2 &= \sum_{\gamma\in\mathcal{S}_2}\sum_{m=-\infty}^{\infty}\int_{m-1/2}^{m+1/2}\left|M_X(1/2+\mathrm{i}(t+\gamma))\right|^2 e^{-|t|}\,dt\\ &\ll \sum_{m=-\infty}^{\infty} e^{-|m|}\sum_{\gamma\in\mathcal{S}_2}\int_{m-1/2}^{m+1/2}\left|M_X(1/2+\mathrm{i}(t+\gamma))\right|^2 dt\\ &\ll \sum_{m=-\infty}^{\infty} e^{-|m|}\int_{m-T-1}^{m+T+1}\left|M_X(1/2+\mathrm{i}u)\right|^2 du,\end{aligned} \tag{5.69}$$

the last inequality being a consequence of (5.63). Since Lemma 4.10 yields

$$\int_{m-T-1}^{m+T+1}\left|M_X(1/2+\mathrm{i}u)\right|^2 du \ll (X+T)\sum_{n\le X} n^{-1} \ll (X+T)(\log T),$$

we conclude that

$$\Sigma_2 \ll (X+T)(\log T)\sum_{m=-\infty}^{\infty} e^{-|m|} \ll (X+T)(\log T). \tag{5.70}$$

For the estimation of $\Sigma_1$, we use the following result.

**Theorem 5.9.** *Suppose that $T \geq 2$. Then*

$$\int_0^T \left|\zeta\left(\tfrac{1}{2} + \mathrm{i}t\right)\right|^4 dt \ll T(\log T)^4. \tag{5.71}$$

*Proof.* See Ivić [31, Ch. 5], Montgomery [41, Ch. 10], or Titchmarsh [50, §7.5 and §7.6]. ∎

By a variant of (5.69),

$$\Sigma_1 \ll \sum_{m=-\infty}^{\infty} e^{-|m|} \int_{-U(m)}^{U(m)} \left|\zeta\left(1/2 + \mathrm{i}u\right)\right|^4 du,$$

where $U(m) = |m| + T + 1$. Hence, we derive from (5.71) that

$$\Sigma_1 \ll \sum_{m=-\infty}^{\infty} e^{-|m|}(T + |m|)(\log(T + |m|))^4 \ll T(\log T)^4. \tag{5.72}$$

Combining (5.68), (5.70), and (5.72), we obtain

$$|\mathcal{R}_2| \ll Y^{2(1-2\sigma)/3}\left(X^{2/3}T^{1/3} + T\right)(\log T)^2. \tag{5.73}$$

We now choose $X = T$, which is consistent with (5.59). Then, by (5.64), (5.65), and (5.73),

$$N(\sigma, T) \ll \left(Y^{2-2\sigma} + T^{2-2\sigma} + Y^{2(1-2\sigma)/3}T + 1\right)(\log T)^{10}.$$

Finally, we put $Y = T^{3/(4-2\sigma)}$ (note that $Y \geq T$) and obtain

$$N(\sigma, T) \ll T^{3(1-\sigma)/(2-\sigma)}(\log T)^{10}.$$

In particular, we have

$$N(\sigma, T) \ll T^{2.4(1-\sigma)}(\log T)^{10} \qquad \text{whenever } 1/2 \leq \sigma \leq 3/4. \tag{5.74}$$

### 5.5.4 The Halász–Montgomery method

We consider the Dirichlet polynomial

$$D(s) = \sum_{N < n \leq 2N} a_n n^{-s},$$

where $a_n$ are complex numbers. Suppose that $s_1, \dots, s_R$, $s_r = \sigma_r + \mathrm{i}t_r$, are complex numbers such that

$$T_0 \leq t_1 < t_2 < \cdots < t_R \leq T_0 + T, \qquad t_{r+1} - t_r \geq 1, \qquad \alpha \leq \sigma_r \leq 1, \tag{5.75}$$

and

$$|D(s_r)| \geq V \qquad \text{for all } r = 1, \dots, R. \tag{5.76}$$

We choose complex numbers $b_1, \dots, b_R$ such that $|b_r| = 1$ and $|D(s_r)| = b_r D(s_r)$. Then

$$\begin{aligned}
\sum_{1\le r\le R} |D(s_r)| &= \sum_{1\le r\le R} b_r \sum_{N<n\le 2N} a_n n^{-s_r} = \sum_{N<n\le 2N} a_n \sum_{1\le r\le R} b_r n^{-s_r} \\
&\le \Big\{ \sum_{N<n\le 2N} \Big| \sum_{1\le r\le R} b_r n^{-s_r} \Big|^2 \Big\}^{1/2} \Big\{ \sum_{N<n\le 2N} |a_n|^2 \Big\}^{1/2} \\
&= \Big\{ \sum_{N<n\le 2N} \sum_{1\le r,q\le R} b_r \bar{b}_q n^{-s_r-\bar{s}_q} \Big\}^{1/2} \Big\{ \sum_{N<n\le 2N} |a_n|^2 \Big\}^{1/2} \\
&\le \Big\{ \sum_{1\le r,q\le R} |b_r \bar{b}_q| \Big| \sum_{N<n\le 2N} n^{-s_r-\bar{s}_q} \Big| \Big\}^{1/2} \Big\{ \sum_{N<n\le 2N} |a_n|^2 \Big\}^{1/2} \\
&= \Big\{ \sum_{1\le r,q\le R} \Big| \sum_{N<n\le 2N} n^{-\sigma_r-\sigma_q+\mathrm{i}(t_q-t_r)} \Big| \Big\}^{1/2} \Big\{ \sum_{N<n\le 2N} |a_n|^2 \Big\}^{1/2}. \qquad (5.77)
\end{aligned}$$

By partial summation, for some $N < M \le 2N$,

$$\sum_{N<n\le 2N} n^{-\sigma_r-\sigma_q+\mathrm{i}(t_q-t_r)} \ll N^{-\sigma_r-\sigma_q} \Big| \sum_{N<n\le M} n^{\mathrm{i}(t_q-t_r)} \Big| \ll N^{-2\alpha} \Big| \sum_{N<n\le M} n^{\mathrm{i}(t_q-t_r)} \Big|. \qquad (5.78)$$

We deal with the sum over $n$ by means of the following exponential sum estimate.

**Lemma 5.10.** *Suppose that $N \ge 2$, $X > 0$, and $f : [N, 2N] \to \mathbb{R}$ has two continuous derivatives that satisfy the conditions*

$$X \ll |f'(x)| \ll X \qquad \text{and} \qquad XN^{-1} \ll |f''(x)| \ll XN^{-1}$$

*for all $x \in [N, 2N]$. Then, for any interval $I \subseteq [N, 2N]$,*

$$\sum_{n\in I} e(f(n)) \ll (XN)^{1/2} + X^{-1}.$$

When $r \ne q$, we apply the lemma with $f(x) = \frac{1}{2\pi}(t_q - t_r)\log x$. We have $X = |t_q - t_r|N^{-1}$, so

$$\sum_{N<n\le M} n^{\mathrm{i}(t_q-t_r)} = \sum_{N<n\le M} e(f(n)) \ll |t_q - t_r|^{1/2} + N|t_q - t_r|^{-1}.$$

Because of (5.75), we can put this inequality in the form

$$\sum_{N<n\le M} n^{\mathrm{i}(t_q-t_r)} \ll |t_q - t_r|^{1/2} + N(|t_q - t_r| + 1)^{-1}, \qquad (5.79)$$

in which it is valid even when $r = q$. Inserting (5.79) into the right side of (5.78), we find that

$$\sum_{N<n\le 2N} n^{-\sigma_r-\sigma_q+\mathrm{i}(t_q-t_r)} \ll N^{-2\alpha} T^{1/2} + N^{1-2\alpha}(|t_q - t_r| + 1)^{-1}.$$

Thus,

$$\sum_{1\le r,q\le R}\Big|\sum_{N<n\le 2N} n^{-\sigma_r-\sigma_q+\mathrm{i}(t_q-t_r)}\Big| \ll N^{1-2\alpha}\Sigma_3 + N^{-2\alpha}R^2T^{1/2}, \tag{5.80}$$

where

$$\Sigma_3 = \sum_{1\le r,q\le R}(|t_r-t_q|+1)^{-1} \le \sum_{1\le r\le R}\sum_{|m|\le T}(|m|+1)^{-1} \ll R(\log T). \tag{5.81}$$

Combining (5.75)–(5.77), (5.80), and (5.81), we deduce that

$$R^2V^2 \ll \big(RN(\log T) + R^2T^{1/2}\big)N^{-2\alpha}G, \qquad G = \sum_{N<n\le 2N}|a_n|^2. \tag{5.82}$$

Let $T_1 = c_4N^{4\alpha}V^4G^{-2}$ for some sufficiently small constant $c_4 > 0$. When $T \le T_1$, (5.82) yields

$$R \ll GN^{1-2\alpha}V^{-2}(\log T).$$

When $T \ge T_1$, we first partition the interval $[T_0, T_0+T]$ into $O(T/T_1+1)$ subintervals of length at most $T_1$ and then bound the number of $t_r$'s in each subinterval using (5.82). We find that

$$R \ll \big(T/T_1+1\big)GN^{1-2\alpha}V^{-2}(\log T) \ll \big(GN^{1-2\alpha}V^{-2} + G^3TN^{1-6\alpha}V^{-6}\big)(\log T). \tag{5.83}$$

### 5.5.5 Huxley's density theorem: $3/4 \le \sigma \le 1$

In §5.5.3, we estimated $|\mathcal{R}_1|$ and $|\mathcal{R}_2|$ using the large sieve and Theorem 5.9. In this section we derive alternative bounds for $|\mathcal{R}_1|$ and $|\mathcal{R}_2|$ using the Halász–Montgomery large value method. To bound $|\mathcal{R}_1|$ we apply the Halász–Montgomery method to the Dirichlet polynomial on the right side of (5.62). By (5.83) with $V = (\log T)^{-1}$ and $\alpha = \sigma$,

$$|\mathcal{R}_1| \ll \big(GN^{1-2\sigma} + G^3TN^{1-6\sigma}\big)(\log T)^{c_5},$$

where

$$G \le \sum_{N<n\le N_1}|a_n|^2 \ll \sum_{n\le 2N} d(n)^2 \ll N(\log N)^3.$$

Hence, when $2/3 \le \sigma \le 1$,

$$|\mathcal{R}_1| \ll \big(Y^{2-2\sigma} + TX^{4-6\sigma}\big)(\log T)^{c_6}. \tag{5.84}$$

We now proceed with the estimation of $|\mathcal{R}_2|$. By (5.61) and (5.67), a class II zero $\rho = \beta + \mathrm{i}\gamma$ satisfies the inequality

$$\int_{-10\log T}^{10\log T}\big|F_X(1/2+\mathrm{i}(t+\gamma))\big|\,dt \gg Y^{\sigma-1/2}.$$

Let $U$ be a parameter to be chosen later. We partition $\mathcal{R}_2$ into two subsets: the set $\mathcal{R}_{2,1}$ of zeros $\rho \in \mathcal{R}_2$ such that

$$\left|M_X(1/2 + \mathrm{i}(t+\gamma))\right| \le U^{-1}Y^{\sigma-1/2}$$

for all $t$ with $|t| \le 10\log T$; and the set $\mathcal{R}_{2,2}$ of the remaining zeros. For zeros $\rho \in \mathcal{R}_{2,1}$, we have

$$U \ll \int_{-10\log T}^{10\log T} \left|\zeta(1/2 + \mathrm{i}(t+\gamma))\right| dt,$$

whence

$$U^4 \ll \left\{ \int_{-10\log T}^{10\log T} \left|\zeta(1/2 + \mathrm{i}(t+\gamma))\right| dt \right\}^4 \ll (\log T)^3 \int_{-10\log T}^{10\log T} |\zeta(1/2 + \mathrm{i}(t+\gamma))|^4\, dt.$$

Thus,

$$\begin{aligned}
|\mathcal{R}_{2,1}| &\ll U^{-4}(\log T)^3 \sum_{\rho\in\mathcal{R}_{2,2}} \int_{-10\log T}^{10\log T} |\zeta(1/2 + \mathrm{i}(t+\gamma))|^4 dt \\
&\ll U^{-4}(\log T)^3 \int_{-2T}^{2T} |\zeta(1/2 + \mathrm{i}u)|^4 \Bigg\{ \sum_{\substack{\rho\in\mathcal{R}_{2,1} \\ |\gamma-u|\le 10\log T}} 1 \Bigg\} du \\
&\ll U^{-4}(\log T)^5 \int_{-2T}^{2T} |\zeta(1/2 + \mathrm{i}u)|^4\, du.
\end{aligned}$$

Using Theorem 5.9, we obtain

$$|\mathcal{R}_{2,1}| \ll TU^{-4}(\log T)^9. \tag{5.85}$$

If $\rho$ is a zero in $\mathcal{R}_{2,2}$, there exists a real number $t_\gamma$, $|t_\gamma - \gamma| \le 10\log T$, such that

$$\left|M_X(1/2 + \mathrm{i}t_\gamma)\right| \ge U^{-1}Y^{\sigma-1/2}.$$

We partition the interval $[1, X]$ into $O(\log X)$ subintervals $[N, N_1]$, $N_1 < 2N$, some of which must satisfy

$$\left| \sum_{N<n\le N_1} \mu(n) n^{-1/2-\mathrm{i}t_\gamma} \right| \gg U^{-1}Y^{\sigma-1/2}(\log X)^{-1}. \tag{5.86}$$

Let $\mathcal{S}(N)$ denote the subset of $\mathcal{R}_{2,1}$ containing those zeros $\rho$ for which (5.86) holds. Then

$$|\mathcal{R}_{2,2}| \ll |\mathcal{S}(N)|(\log T) \tag{5.87}$$

for some $N$, $1 \le N \le X$. By (5.83) with $\alpha = 1/2$ and $V = U^{-1}Y^{\sigma-1/2}(\log T)^{-1}$,

$$|\mathcal{S}(N)| \ll \left(NV^{-2} + NTV^{-6}\right)\log T. \tag{5.88}$$

From (5.85), (5.87), and (5.88),

$$|\mathcal{R}_2| \ll \left(TU^{-4} + XY^{1-2\sigma}U^2 + TXY^{3-6\sigma}U^6\right)(\log T)^{c_7}. \tag{5.89}$$

Upon choosing

$$U = \left(X^{-1}Y^{-3(1-2\sigma)}\right)^{1/10},$$

(5.89) becomes

$$|\mathcal{R}_2| \ll \left(TX^{2/5}Y^{6(1-2\sigma)/5} + X^{4/5}Y^{2(1-2\sigma)/5}\right)(\log T)^{c_7}.$$

Combining this inequality with (5.64) and (5.84), we find that

$$N(\sigma, T) \ll \left(Y^{2-2\sigma} + TX^{4-6\sigma} + TX^{2/5}Y^{6(1-2\sigma)/5} + X^{4/5}Y^{2(1-2\sigma)/5}\right)(\log T)^{c_8}. \tag{5.90}$$

Under the assumption that

$$X^2Y^{4(2\sigma-1)} \le T^5,$$

the third term on the right side of (5.90) dominates the fourth, and so

$$N(\sigma, T) \ll \left(Y^{2-2\sigma} + TX^{4-6\sigma} + TX^{2/5}Y^{6(1-2\sigma)/5}\right)(\log T)^{c_8}.$$

Setting

$$X = Y^{(2\sigma-1)/(5\sigma-3)},$$

we derive

$$N(\sigma, T) \ll \left(Y^{2-2\sigma} + TY^{(4-6\sigma)(2\sigma-1)/(5\sigma-3)}\right)(\log T)^{c_8}, \tag{5.91}$$

provided that

$$Y \le T^{\frac{1}{2}(5\sigma-3)/(2\sigma-1)^2}. \tag{5.92}$$

Finally, we take

$$Y = T^{\frac{1}{2}(5\sigma-3)/(\sigma^2+\sigma-1)},$$

which satisfies (5.92) for $2/3 \le \sigma \le 1$ and turns the bound (5.91) into

$$N(\sigma, T) \ll T^{(5\sigma-3)(1-\sigma)/(\sigma^2+\sigma-1)}(\log T)^{c_8}.$$

In particular,

$$N(\sigma, T) \ll T^{2.4(1-\sigma)}(\log T)^{c_8} \qquad \text{whenever } 3/4 \le \sigma \le 1.$$

Together with (5.74), this completes the proof of Huxley's theorem. ∎

## 5.6 Primes in almost all short intervals

In this section, we use Huxley's density theorem (Theorem 5.8) to prove the following result.

**Theorem 7.** *Let $\mathcal{E}(X, \delta)$ denote the set of real numbers $x \in [X, 2X]$ such that*

$$\left|\psi(x) - \psi(x - \delta x) - \delta x\right| \ge \delta x(\log x)^{-1}. \tag{5.93}$$

*Suppose that $\epsilon > 0$ and $A > 0$ are fixed, $X \ge X_0(\epsilon, A)$, and $X^{-5/6+\epsilon} \le \delta \le 1$. Then*

$$|\mathcal{E}(X, \delta)| \ll_A X(\log X)^{-A},$$

*the left side representing the Lebesgue measure of $\mathcal{E}(X, \delta)$.*

*Proof.* We have

$$|\mathcal{E}(X,\delta)| \ll (\delta X)^{-2}(\log X)^2 \int_X^{2X} \left|\psi(x) - \psi(x-\delta x) - \delta x\right|^2 dx,$$

so it suffices to show that

$$\int_X^{2X} \left|\psi(x) - \psi(x-\delta x) - \delta x\right|^2 dx \ll \delta^2 X^3 (\log X)^{-A-2}. \tag{5.94}$$

By Theorem 2.22 with $T = X^{5/6-\epsilon/2}$,

$$\begin{aligned}\psi(x) - \psi(x-\delta x) - \delta x &= \sum_{|\operatorname{Im}\rho|\le T} \frac{x^\rho - (x-\delta x)^\rho}{\rho} + O\big(X^{1/6+\epsilon/2}(\log X)^2\big) \\ &= \sum_{|\operatorname{Im}\rho|\le T} x^\rho \omega(\rho) + O\big(X^{1/6+2\epsilon/3}\big), \qquad \omega(\rho) = \int_{1-\delta}^1 u^{\rho-1}\,du.\end{aligned}$$

Hence,

$$\int_X^{2X} \left|\psi(x) - \psi(x-\delta x) - \delta x\right|^2 dx \ll \int_X^{2X} \Big|\sum_{|\operatorname{Im}\rho|\le T} x^\rho \omega(\rho)\Big|^2 dx + \delta^2 X^{3-\epsilon/2}. \tag{5.95}$$

Upon noting that $|\omega(\rho)| \le \delta$, we obtain

$$\begin{aligned}\int_X^{2X} \Big|\sum_{|\operatorname{Im}\rho|\le T} x^\rho \omega(\rho)\Big|^2 dx &= \sum_{|\operatorname{Im}\rho_1|\le T} \sum_{|\operatorname{Im}\rho_2|\le T} \omega(\rho_1)\overline{\omega(\rho_2)} \int_X^{2X} x^{\rho_1+\bar{\rho}_2}\,dx \\ &\ll \delta^2 \sum_{|\operatorname{Im}\rho_1|\le T} \sum_{|\operatorname{Im}\rho_2|\le T} \Big|\int_X^{2X} x^{\rho_1+\bar{\rho}_2}\,dx\Big|. \end{aligned}\tag{5.96}$$

We now appeal to the inequality

$$\int_X^{2X} x^{\beta_1+\beta_2+\mathrm{i}(\gamma_1-\gamma_2)}\,dx \ll \frac{X^{\beta_1+\beta_2+1}}{|\gamma_1-\gamma_2|+1},$$

which follows by partial integration. Using this to bound the right side of (5.96), we get

$$\begin{aligned}\int_X^{2X} \Big|\sum_{|\operatorname{Im}\rho|\le T} x^\rho \omega(\rho)\Big|^2 dx &\ll \delta^2 \sum_{|\gamma_1|\le T} \sum_{|\gamma_2|\le T} \frac{X^{\beta_1+\beta_2+1}}{|\gamma_1-\gamma_2|+1} \\ &\ll \delta^2 \sum_{|\gamma_1|\le T} \sum_{|\gamma_2|\le T} \frac{X^{2\beta_1+1}}{|\gamma_1-\gamma_2|+1} \ll \delta^2 (\log T) \sum_{|\gamma|\le T} X^{2\beta+1}.\end{aligned}\tag{5.97}$$

We now write $\theta(T) = (\log T)^{-3/4}$. By Theorems 5.7 and 5.8 and (5.50),

$$\begin{aligned}\sum_{|\gamma|\le T} X^{2\beta+1} &= -\int_0^{1-\theta(T)} X^{2\sigma+1}\, dN(\sigma,T)\\ &\ll XN(0,T) + (\log X)\int_0^{1-\theta(T)} X^{2\sigma+1}N(\sigma,T)\,d\sigma\\ &\ll XT(\log T) + (\log X)^{c_2+1}\int_0^{1-\theta(T)} X^{2\sigma+1}T^{2.4(1-\sigma)}\,d\sigma\\ &\ll X^{3-\epsilon\theta(T)}(\log X)^{c_2+1} \ll X^3(\log X)^{-A-3}. \end{aligned} \tag{5.98}$$

The desired bound (5.94) follows from (5.95), (5.97), and (5.98). ∎

## 5.7 The linear sieve

This section is a (very) brief introduction to sieve methods, without proofs and in the special case of a "linear sieve".

### 5.7.1 The fundamental problem of sieve theory

Let $\mathcal{A}$ be a finite integer sequence. We will be concerned with the existence of elements of $\mathcal{A}$ that are primes or, more generally, *almost primes* $P_r$, that is, integers having at most $r$ prime divisors, counted according to multiplicity. We consider a set of prime numbers $\mathcal{P}$ and a real parameter $z \ge 2$ and define the *sifting function*

$$S(\mathcal{A},\mathcal{P},z) = \#\big\{a\in\mathcal{A} \;:\; (a,P(z))=1\big\}, \qquad P(z)=\prod_{\substack{p<z\\ p\in\mathcal{P}}} p. \tag{5.99}$$

In applications, the set $\mathcal{P}$ is usually taken to be the set of possible prime divisors of the elements of $\mathcal{A}$, so the sifting function (5.99) counts the elements of $\mathcal{A}$ free of prime divisors $p < z$.

For our first attempt at bounding $S(\mathcal{A},\mathcal{P},z)$, we recall Lemma 1.2. It yields

$$S(\mathcal{A},\mathcal{P},z) = \sum_{a\in\mathcal{A}}\sum_{d\mid(a,P(z))}\mu(d) = \sum_{d\mid P(z)}\mu(d)|\mathcal{A}_d|, \tag{5.100}$$

where

$$|\mathcal{A}_d| = \#\big\{a\in\mathcal{A} \;:\; a\equiv 0 \pmod d\big\}.$$

To this end, we suppose that there exist a (large) quantity $X$ and a multiplicative function $\omega(d)$ such that $|\mathcal{A}_d|$ can be approximated by $X\omega(d)/d$, and we write $r(\mathcal{A},d)$ for the remainder in this approximation:

$$|\mathcal{A}_d| = X\frac{\omega(d)}{d} + r(\mathcal{A},d). \tag{5.101}$$

We expect $r(\mathcal{A}, d)$ to be 'small', at least in some average sense over $d$. Substituting (5.101) into the right side of (5.100), we find that

$$S(\mathcal{A}, \mathcal{P}, z) = XV(z) + R(\mathcal{A}, z), \tag{5.102}$$

where

$$V(z) = \sum_{d|P(z)} \mu(d)\frac{\omega(d)}{d}, \qquad R(\mathcal{A}, z) = \sum_{d|P(z)} \mu(d) r(\mathcal{A}, d). \tag{5.103}$$

We would like to believe that, under 'ideal circumstances', (5.103) is an asymptotic formula for the sifting function $S(\mathcal{A}, \mathcal{P}, z)$, $XV(z)$ being the main term and $R(\mathcal{A}, z)$ the error term. However, such expectations turn out to be unrealistic (see Exercise 10). Therefore, we need to adjust our strategy.

Let $D > 0$ be a parameter to be chosen later in terms of $X$. Suppose that $\Lambda^+(d)$ and $\Lambda^-(d)$ are real-valued functions supported on the squarefree integers $d$ such that

$$|\Lambda^{\pm}(d)| \le 1 \qquad \text{and} \qquad \Lambda^{\pm}(d) = 0 \quad \text{for } d \ge D. \tag{5.104}$$

Furthermore, suppose that

$$\sum_{d|n} \Lambda^-(d) \le \sum_{d|n} \mu(d) \le \sum_{d|n} \Lambda^+(d) \qquad \text{for all } n \in \mathcal{A}. \tag{5.105}$$

Using (5.100), (5.101), and the left inequality in (5.105), we obtain

$$\begin{aligned} S(\mathcal{A}, \mathcal{P}, z) &\ge \sum_{a\in\mathcal{A}} \sum_{d|(a,P(z))} \Lambda^-(d) = \sum_{d|P(z)} \Lambda^-(d)|\mathcal{A}_d| \\ &= \sum_{d|P(z)} \Lambda^-(d) \left( X\frac{\omega(d)}{d} + r(\mathcal{A}, d) \right) \ge X\mathcal{M}^- - \mathcal{R}, \end{aligned}$$

where

$$\mathcal{M}^{\pm} = \sum_{d|P(z)} \Lambda^{\pm}(d)\frac{\omega(d)}{d}, \qquad \mathcal{R} = \sum_{\substack{d|P(z)\\ d<D}} |r(\mathcal{A}, d)|. \tag{5.106}$$

In a similar fashion, we can use the right inequality in (5.105) to estimate the sifting function from above. That is, we have

$$X\mathcal{M}^- - \mathcal{R} \le S(\mathcal{A}, \mathcal{P}, z) \le X\mathcal{M}^+ + \mathcal{R}. \tag{5.107}$$

We are now in a position to overcome the difficulty caused by the "error term" in (5.100). The sum $\mathcal{R}$ is similar to the error term $R(\mathcal{A}, z)$ defined in (5.101), but unlike $R(\mathcal{A}, z)$ we can use the parameter $D$ to control the number of terms in $\mathcal{R}$. Thus, our general strategy will be to construct functions $\Lambda^{\pm}(d)$ which satisfy (5.104) and (5.105) and for which the sums $\mathcal{M}^{\pm}$ are of the same order as the sum $V(z)$ defined in (5.102). There are various constructions of such functions $\Lambda^{\pm}(d)$. We will simply state one of the modern sieves in a form suitable for application in §5.8.

### 5.7.2 The Rosser–Iwaniec sieve

The basic form of this sieve method appeared for the first time in an unpublished manuscript by Rosser, and the full-fledged version was developed independently by Iwaniec [32, 33]. Suppose that the multiplicative function $\omega(d)$ in (5.101) satisfies the condition

$$\prod_{w_1\le p<w_2}\left(1-\frac{\omega(p)}{p}\right)^{-1}\le\left(\frac{\log w_2}{\log w_1}\right)^{\kappa}\left(1+\frac{K}{\log w_1}\right)\quad(2\le w_1<w_2),\tag{5.108}$$

where $\kappa>0$ is an absolute constant known as the *sieve dimension* and $K>0$ is independent of $w_1$ and $w_2$. This inequality is usually interpreted as an average bound for the values taken by $\omega(p)$ when $p$ is prime, since it is consistent with the inequality $\omega(p)\le\kappa$. In the applications we are interested in, (5.108) holds with $\kappa=1$, so we will state the Rosser–Iwaniec sieve in this special case: this is the so-called the *linear sieve*.

Suppose that $\omega(p)$ satisfies (5.108) with $\kappa=1$ and that

$$0<\omega(p)<p\quad\text{when } p\in\mathfrak{P}\qquad\text{and}\qquad\omega(p)=0\quad\text{when } p\notin\mathfrak{P}.\tag{5.109}$$

We put $\Lambda^{\pm}(1)=1$ and $\Lambda^{\pm}(d)=0$ if $d$ is not squarefree. If $d>1$ is squarefree and has prime decomposition $d=p_1\cdots p_r$, $p_1>p_2>\cdots>p_r$, we define

$$\Lambda^{+}(d)=\begin{cases}(-1)^r & \text{if } p_1\cdots p_{2l}p_{2l+1}^3<D \text{ whenever } 0\le l\le(r-1)/2,\\ 0 & \text{otherwise,}\end{cases}\tag{5.110}$$

$$\Lambda^{-}(d)=\begin{cases}(-1)^r & \text{if } p_1\cdots p_{2l-1}p_{2l}^3<D \text{ whenever } 1\le l\le r/2,\\ 0 & \text{otherwise.}\end{cases}\tag{5.111}$$

It can be shown (see Greaves' book [18] or Iwaniec's original paper [33]) that these two functions satisfy conditions (5.104) and (5.105). Furthermore, if the quantities $\mathcal{M}^{\pm}$ are defined by (5.106) with $\Lambda^{\pm}(d)$ given by (5.110) and (5.111), we have

$$V(z)\le\mathcal{M}^{+}\le V(z)\left(F(s)+O\big(e^{-s}(\log D)^{-1/3}\big)\right)\qquad\text{for } s\ge1,\tag{5.112}$$

$$V(z)\ge\mathcal{M}^{-}\ge V(z)\left(f(s)+O\big(e^{-s}(\log D)^{-1/3}\big)\right)\qquad\text{for } s\ge2,\tag{5.113}$$

where $s=\log D/\log z$ and the functions $f(s)$ and $F(s)$ are the continuous solutions of the following system of differential delay equations:

$$\begin{aligned}f(s)&=0 && \text{if } 0<s\le2,\\ F(s)&=2e^{\gamma}s^{-1} && \text{if } 0<s\le3,\\ (sf(s))'&=F(s-1) && \text{if } s>2,\\ (sF(s))'&=f(s-1) && \text{if } s>3.\end{aligned}$$

Here $\gamma$ is Euler's constant. The analysis of this system reveals that the function $F(s)$ is strictly decreasing for $s>0$, that the function $f(s)$ is strictly increasing for $s>2$, and that

$$0<f(s)<1<F(s)\qquad\text{for } s>2.\tag{5.114}$$

Furthermore, both functions are very close to 1 for large $s$: they satisfy

$$F(s),\ \ f(s) = 1 + O(s^{-s}) \qquad \text{as } s \to \infty. \tag{5.115}$$

Substituting (5.112) and (5.113) into (5.107), we obtain

$$S(\mathcal{A}, \mathcal{P}, z) \le XV(z)\left(F(s) + O\big((\log D)^{-1/3}\big)\right) + \mathcal{R} \qquad \text{for } s \ge 1, \tag{5.116}$$

$$S(\mathcal{A}, \mathcal{P}, z) \ge XV(z)\left(f(s) + O\big((\log D)^{-1/3}\big)\right) - \mathcal{R} \qquad \text{for } s \ge 2. \tag{5.117}$$

### 5.7.3 Two applications

**Example 5.7.1.** Suppose that $2 \le y \le x$, where $x$ is a (large) real number. We choose $\mathcal{A}$ to be the sequence of integers $n \in (x - y, x]$ and $\mathcal{P}$ to be the set of all primes. Then

$$|\mathcal{A}_d| = \sum_{x-y<md\le x} 1 = \left[\frac{x}{d}\right] - \left[\frac{x-y}{d}\right] = \frac{y}{d} - \left\{\frac{x}{d}\right\} + \left\{\frac{x-y}{d}\right\},$$

so (5.101) holds with

$$X = y, \qquad \omega(d) = 1, \qquad \text{and} \qquad r(\mathcal{A}, d) = -\left\{\frac{x}{d}\right\} + \left\{\frac{x-y}{d}\right\},$$

and in (5.117), one has

$$XV(z) = y\prod_{p<z}\left(1 - p^{-1}\right) \gg y(\log z)^{-1} \qquad \text{and} \qquad \mathcal{R} \ll D.$$

Hence, combining (5.114) and (5.117), we obtain

$$S(\mathcal{A}, \mathcal{P}, z) \gg y(\log x)^{-1}, \tag{5.118}$$

provided that

$$D \le y^{1-\epsilon} \qquad \text{and} \qquad z \le D^{1/2-\epsilon}$$

for some fixed $\epsilon > 0$.

Choosing $y = x^{\theta}$, $D = y^{1-\epsilon}$, and $z = D^{1/2-\epsilon}$, we find that there are $\gg y(\log x)^{-1}$ integers $n \in (x - x^{\theta}, x]$ that have no prime divisor smaller than $x^{\theta/2-2\epsilon}$. Since the numbers in question do not exceed $x$, each of the elements of $\mathcal{A}$ counted on the left side of (5.118) has at most $2/\theta$ prime divisors. In particular, we are able to conclude that:

*For sufficiently large $x$, the interval $(x - x^{1/2}, x]$ contains a $P_4$-number.*

Note that in this case we just miss to show the existence of $P_3$-numbers in $(x - x^{1/2}, x]$. If we increase the length of the intervals just slightly, we obtain:

*For $\delta > 0$ and $x \ge x_0(\delta)$, the interval $(x - x^{1/2+\delta}, x]$ contains a $P_3$-number.*

■

**Example 5.7.2.** Suppose that $n$ is a (large) even integer and set

$$\mathcal{A} = \{n - p \ : \ 2 < p < n\} \qquad \text{and} \qquad \mathcal{P} = \{p \ : \ p \nmid n\}.$$

Then (5.101) is expected to hold with

$$X = \operatorname{Li} n \qquad \text{and} \qquad \omega(d) = \begin{cases} d/\phi(d) & \text{if } (d,n) = 1, \\ 0 & \text{otherwise.} \end{cases}$$

The main term in (5.117) is

$$XV(z) = (\operatorname{Li} n) \prod_{\substack{p<z \\ p\nmid n}} \left(1 - (p-1)^{-1}\right) \gg (\operatorname{Li} n)(\log z)^{-1}, \tag{5.119}$$

and the error term $\mathcal{R}$ is bounded by

$$D + \sum_{d\le D} \max_{(a,d)=1} \left| \pi(n; d, a) - \frac{\operatorname{Li} n}{\phi(d)} \right|.$$

In particular, when $D \le n^{1/2-\epsilon}$, the Bombieri–Vinogradov theorem yields

$$\mathcal{R} \ll n(\log n)^{-3}. \tag{5.120}$$

We now choose $D = n^{0.49}$ and $z = n^{2/9}$, so that we have

$$s = \frac{\log D}{\log z} > 2.2.$$

Combining (5.114), (5.117), (5.119), and (5.120), we find that

$$S(\mathcal{A}, \mathcal{P}, z) \gg n(\log n)^{-2}. \tag{5.121}$$

That is, there are $\gg n(\log n)^{-2}$ elements of $\mathcal{A}$ that have no prime divisors smaller than $n^{2/9}$. Since the numbers in $\mathcal{A}$ do not exceed $n$, the elements of $\mathcal{A}$ counted on the left side of (5.121) have at most four prime divisors each, that is, the left side of (5.121) counts solutions of $n - p = P_4$. We conclude that:

> *Every sufficiently large even integer n can be represented as the sum of a prime and a $P_4$-number.*

■

The results of both examples can be strengthened significantly. Chen [9, 10] has proved the following two theorems.

**Theorem 8 (Chen).** *For sufficiently large x, the interval $(x - x^{1/2}, x]$ always contains a $P_2$-number.*

**Theorem 9 (Chen).** *Every sufficiently large even integer $n$ can be represented as the sum of a prime and a $P_2$-number.*

Obviously, Theorem 9 is the best possible result of its kind, short of a proof of the binary Goldbach conjecture. Its proof is too involved to include in these lectures, but those interested can find all the details in Halberstam and Richert [19, Ch. 11]. We present the proof of Theorem 8 in §5.8. However, unlike Theorem 9, Theorem 8 can and has been improved on (several times). The best result to date is due to Liu [39] and states that all intervals of the form $(x - x^{0.436}, x]$, $x \ge x_0$, contain a $P_2$-number.

### 5.7.4 The bilinear form of the error term in the linear sieve

Assume that $\mathcal{A}$ is an integer sequence such that (5.101) holds with a function $\omega$ subject to (5.108) with $\kappa = 1$ and

$$\sum_{w_1 \le p < w_2} \sum_{\nu \ge 2} \frac{\omega(p^\nu)}{p^\nu} \le \frac{L}{\log 3w_1} \qquad (2 \le \omega_1 < \omega_2),$$

where $L > 0$ is independent of $w_1, w_2$. Inspired by Chen's original proof of Theorem 8, Iwaniec [32] obtained the following more flexible form of the linear Rosser–Iwaniec sieve.

**Theorem 5.11 (Iwaniec).** *Suppose that $0 < \epsilon < 1/3$, $M, N \ge 2$, $D = MN$. Then*

$$S(\mathcal{A}, \mathcal{P}, z) \le XV(z)\,(F(s) + E(\epsilon, D, K, L)) + \mathcal{R}^+(M, N), \tag{5.122}$$

$$S(\mathcal{A}, \mathcal{P}, z) \ge XV(z)\,(f(s) - E(\epsilon, D, K, L)) - \mathcal{R}^+(M, N), \tag{5.123}$$

*where $s = \log D/\log z$, $E(\epsilon, D, K, L) \ll \epsilon + \epsilon^{-8} e^{K+L} (\log D)^{-1/3}$, and*

$$\mathcal{R}^\pm(M, N) = \sum_{j \le J} \sum_{\substack{m \le M \\ m \mid P(z)}} \sum_{\substack{n \le N \\ n \mid P(z)}} a^\pm_{m,j} b^\pm_{n,j} r(\mathcal{A}, mn), \qquad J = \exp\left(8\epsilon^{-3}\right).$$

*The coefficients $a^\pm_{m,j}, b^\pm_{n,j}$ depend at most on $\epsilon, M, N$ (but not on $\mathcal{A}$) and satisfy $|a^\pm_{m,j}| \le 1$, $|b^\pm_{n,j}| \le 1$.*

The importance of this result is that it allows us to replace the error term $\mathcal{R}$ defined by (5.106) with a bounded number of sums of the form

$$\sum_{m \le M} \sum_{n \le N} a_m b_n r(\mathcal{A}, mn),$$

where $|a_m| \le 1$, $|b_n| \le 1$. In many applications, one can exploit the arithmetic properties of the sequence $\mathcal{A}$ to estimate such double sums more effectively. To illustrate this, we return to Example 5.7.1. Suppose that $y = x^\theta$, $2/5 < \theta < 3/5$. We will show (see Lemma 5.13 below) that in this situation one can obtain a satisfactory bound for $\mathcal{R}^\pm(M, N)$ under the hypotheses

$$M \le x^{\theta - 6\epsilon}, \qquad MN^2 \le x^{(5\theta-1)/2 - 10\epsilon}.$$

Therefore, upon choosing $0 < \epsilon < \epsilon_0(\delta)$, $M = x^{\theta-6\epsilon}$, and $N = x^{(3\theta-1)/4-2\epsilon}$ in Theorem 5.11, we can replace the parameter $D = x^{\theta-\epsilon}$ in Example 5.7.1 by $D = MN = x^{(7\theta-1)/4-8\epsilon}$ to obtain:

*For $\delta > 0$ and $x \ge x_0(\delta)$, the interval $(x - x^{3/7+\delta}, x]$ contains a $P_3$-number.*

## 5.8 Almost primes in short intervals

In this section we establish Theorem 8. That result and its subsequent improvements rest on two new ideas: more sophisticated sieve machinery and the Fourier analysis of the remainders $r(\mathcal{A}, d)$.

### 5.8.1 The remainders $r(\mathcal{A}, d)$

In Example 5.7.1 we estimated the error term in (5.117) trivially. In this section, we use Fourier analytic techniques to exploit the oscillation of the remainders $r(\mathcal{A}, mn)$ in (5.122) and (5.123).

**Lemma 5.12.** *Suppose that $\epsilon > 0$, $x^{2\epsilon} \le y \le x$, $2 \le M < M_1 \le 2M$, $2 \le N < N_1 \le 2N$, $y \le MN \le x$, and $a_m, b_n$ are complex numbers with $|a_m| \le 1$, $|b_n| \le 1$. We define*

$$r(x, y; d) = \left[\frac{x}{d}\right] - \left[\frac{x-y}{d}\right] - \frac{y}{d}.$$

*There exist a real number $X \in [x/2, 2x]$ and (complex) coefficients $a_m^*, b_n^*$, with $|a_m^*| \le 1$, $|b_n^*| \le 1$, such that*

$$\sum_{M<m\le M_1} \sum_{N<n\le N_1} a_m b_n r(x, y; mn) \ll y(MN)^{-1} \left| \sum_{M<m\le M_1} \sum_{N<n\le N_1} \sum_{1\le h\le H} a_m^* b_n^* e\left(\frac{Xh}{mn}\right) \right| + yx^{-\epsilon},$$

*with $H = MNy^{-1}x^{3\epsilon}$.*

*Proof.* Let $f$ be a $C^\infty$-function, supported in $[x - y - yx^{-2\epsilon}, x + yx^{-2\epsilon}]$ and such that

$$f(u) = 1 \quad (x - y \le u \le x) \qquad \text{and} \qquad f^{(j)}(u) \ll (yx^{-2\epsilon})^{-j} \quad (j \ge 0). \tag{5.124}$$

(See Exercise 12 for one possible construction of such a function.) Then, by Lemma 1.21,

$$\begin{aligned} &\sum_{M<m\le M_1} \sum_{N<n\le N_1} \sum_{x-y<kmn\le x} a_m b_n - \sum_{M<m\le M_1} \sum_{N<n\le N_1} \sum_k a_m b_n f(kmn) \\ &\ll \sum_{x-y-yx^{-2\epsilon}<u\le x-y} d(u)^2 + \sum_{x<u\le x+yx^{-2\epsilon}} d(u)^2 \ll yx^{-\epsilon}, \end{aligned}$$

that is,

$$\sum_{M<m\le M_1} \sum_{N<n\le N_1} a_m b_n \left( \left[\frac{x}{mn}\right] - \left[\frac{x-y}{mn}\right] \right) = \sum_{M<m\le M_1} \sum_{N<n\le N_1} \sum_k a_m b_n f(kmn) + O(yx^{-\epsilon}). \tag{5.125}$$

Let $g_r(u) = f(ur)$. Applying the Poisson summation formula (see Zygmund [59, eq. (II.13.4)]) to the sum over $k$, we obtain

$$\sum_k f(kmn) = \sum_k g_{mn}(k) = \sum_h \hat{g}_{mn}(h) = (mn)^{-1} \sum_h \hat{f}\left(\frac{h}{mn}\right),$$

where $\hat{f}$ is the Fourier transform of $f$. Hence, we can rewrite (5.125) as

$$\sum_{M<m\le M_1}\sum_{N<n\le N_1} a_m b_n \left(\left[\frac{x}{mn}\right]-\left[\frac{x-y}{mn}\right]\right) = \sum_{M<m\le M_1}\sum_{N<n\le N_1}\sum_h \frac{a_m b_n}{mn}\hat{f}\left(\frac{h}{mn}\right)+O(yx^{-\epsilon}). \quad (5.126)$$

The contribution to the right side from the terms with $h=0$ is

$$\hat{f}(0)\sum_{M<m\le M_1}\sum_{N<n\le N_1}\frac{a_m b_n}{mn} = y\sum_{M<m\le M_1}\sum_{N<n\le N_1}\frac{a_m b_n}{mn}+O(yx^{-\epsilon}),$$

so it follows from (5.126) that

$$\sum_{M<m\le M_1}\sum_{N<n\le N_1} a_m b_n r(x,y;mn) = \sum_{M<m\le M_1}\sum_{N<n\le N_1}\sum_{h\neq 0}\frac{a_m b_n}{mn}\hat{f}\left(\frac{h}{mn}\right)+O(yx^{-\epsilon}). \quad (5.127)$$

We now proceed to estimate the tails of the series over $h$. Choose an integer $r\ge 3+\epsilon^{-1}$. By (5.124) and $r$-fold partial integration,

$$\hat{f}(t) = (-2\pi i t)^{-r}\int_{-\infty}^{\infty} f^{(r)}(u)e(-ut)\,du \ll y(yx^{-2\epsilon}|t|)^{-r} \qquad (r\ge 0).$$

Thus, the contribution to the right side of (5.127) from terms with $|h|>H$ is

$$\ll y\sum_{|h|>H}\left(\frac{yx^{-2\epsilon}|h|}{MN}\right)^{-r} \ll yH\left(\frac{yx^{-2\epsilon}H}{MN}\right)^{-r} \ll MNx^{(3-r)\epsilon}\ll 1.$$

Therefore, (5.127) yields

$$\sum_{M<m\le M_1}\sum_{N<n\le N_1} a_m b_n r(x,y;mn) \ll \int_{-\infty}^{\infty}\left|\sum_{M<m\le M_1}\sum_{N<n\le N_1}\sum_{0<|h|\le H}\frac{a_m b_n}{mn}e\left(\frac{-uh}{mn}\right)\right||f(u)|\,du + yx^{-\epsilon}.$$

Recalling that $f$ is supported on a subset of $[x/2,2x]$ of measure $O(y)$, we conclude that

$$\sum_{M<m\le M_1}\sum_{N<n\le N_1} a_m b_n r(x,y;mn) \ll y(MN)^{-1}\left|\sum_{M<m\le M_1}\sum_{N<n\le N_1}\sum_{1\le h\le H} a_m^* b_n^* e\left(\frac{Xh}{mn}\right)\right| + yx^{-\epsilon},$$

where $|a_m^*|\le 1$, $|b_n^*|\le 1$, and $x/2\le X\le 2x$. ∎

**Lemma 5.13.** *Suppose that $\epsilon>0$, $x^{2\epsilon}\le y\le x$, $2\le M<M_1\le 2M$, $2\le N<N_1\le 2N$, $y\le MN\le x$, and $a_m, b_n$ are complex numbers with $|a_m|\le 1$, $|b_n|\le 1$. Also, suppose that*

$$M\le yx^{-6\epsilon}, \qquad MN\le y^{1/2}x^{1/2-3\epsilon}, \qquad MN^2\le y^{5/2}x^{-1/2-10\epsilon}. \quad (5.128)$$

*Then*

$$\sum_{M<m\le M_1}\sum_{N<n\le N_1} a_m b_n r(x,y;mn)\ll yx^{-\epsilon},$$

*where $r(x,y;mn)$ is the function defined in Lemma 5.12.*

*Proof.* Let $H = MNy^{-1}x^{3\epsilon}$. By Lemma 5.12, it suffices to show that

$$S(H,M,N) = \sum_{M<m\le M_1}\sum_{N<n\le N_1}\sum_{1\le h\le H} a_m^* b_n^* e\left(\frac{Xh}{mn}\right) \ll MNx^{-\epsilon}. \tag{5.129}$$

Here $a_m^*$, $b_n^*$, and $X$ are as in Lemma 5.12. By Cauchy's inequality,

$$\begin{aligned}|S(H,M,N)|^2 &\ll M\sum_{M<m\le 2M}\left|\sum_{N<n\le N_1}\sum_{1\le h\le H} b_n^* e\left(\frac{Xh}{mn}\right)\right|^2\\ &\ll M\sum_{N<n_1,n_2\le 2N}\sum_{1\le h_1,h_2\le H}\left|\sum_{M<m\le 2M} e\left(\frac{X}{m}\left(\frac{h_1}{n_1}-\frac{h_2}{n_2}\right)\right)\right|.\end{aligned} \tag{5.130}$$

We now group the quadruples $(h_1,h_2,n_1,n_2)$ according to the value of the determinant $\Delta = h_1n_2 - h_2n_1$. When $\Delta = 0$, we bound the sum over $m$ in (5.130) trivially by $M$. When $\Delta \neq 0$, we appeal to Lemma 5.10 with $f(m) = \Delta X(n_1n_2m)^{-1}$. We get

$$\sum_{M<m\le 2M} e\left(\frac{\Delta X}{mn_1n_2}\right) \ll \left(\frac{|\Delta|X}{MN^2}\right)^{1/2} + \frac{M^2N^2}{|\Delta|X}.$$

Writing $\delta(k)$ for the number of quadruples with $\Delta = k$, we conclude that

$$|S(H,M,N)|^2 \ll \delta(0)M^2 + \sum_{0<|k|\le 2HN}\delta(k)\left(\left(\frac{|k|X}{MN^2}\right)^{1/2} + \frac{M^2N^2}{|k|X}\right). \tag{5.131}$$

For $|k| \le 2HN$, we have

$$\delta(k) \le \sum_{N<n\le 2N}\sum_{1\le h\le H} d(hn+|k|) \ll (HN)^{1+\epsilon/2},$$

so (5.131) yields

$$\begin{aligned}(HN)^{-\epsilon}|S(H,M,N)|^2 &\ll M^2NH + X^{1/2}M^{1/2}N^{3/2}H^{5/2} + X^{-1}M^3N^3H\\ &\ll M^2N^2\left(My^{-1}x^{3\epsilon} + MN^2y^{-5/2}x^{1/2+7.5\epsilon} + M^2N^2y^{-1}x^{-1+3\epsilon}\right).\end{aligned} \tag{5.132}$$

Since $(HN)^\epsilon \le x^{\epsilon/2}$, (5.129) follows from (5.132) and the hypotheses (5.128). ∎

## 5.8.2 Proof of Theorem 8

Let $y = x^{1/2}$, $z = x^\delta$, $\mathfrak{P}$ be the set of all primes, $\mathcal{A}$ the sequence of integers $n \in (x-y, x]$. We write $S(\mathcal{A}, w)$ for $S(\mathcal{A}, \mathfrak{P}, w)$,

$$P(w) = \prod_{p<w} p, \qquad V(w) = \prod_{p<w}\left(1-p^{-1}\right).$$

Our starting point is the sum

$$\Sigma(\alpha,\beta,\delta)=\sum_{\substack{x-y<n\le x\\(n,P(z))=1}}\left(1-\sum_{p\mid n}\left(\alpha-\beta\frac{\log p}{\log x}\right)\right),$$

where $\alpha,\beta,\delta$ are positive absolute constants to be chosen later. On the one hand, we have

$$\Sigma(\alpha,\beta,\delta)\le\sum_{\substack{x-y<n\le x\\(n,P(z))=1}}\left(\beta+1-\alpha\sum_{p\mid n}1\right).$$

Hence, if we assume that

$$3\alpha\ge\beta+1, \tag{5.133}$$

only $P_2$-numbers $n$ will contribute positive terms to $\Sigma(\alpha,\beta,\delta)$. In particular, the theorem will follow if we show that

$$\Sigma(\alpha,\beta,\delta)>0$$

for some $\alpha,\beta,\delta$ satisfying (5.133). On the other hand,

$$\begin{aligned}\Sigma(\alpha,\beta,\delta)&=S(\mathcal{A},z)-\sum_{p\ge z}\left(\alpha-\beta\frac{\log p}{\log x}\right)S(\mathcal{A}_p,z)\\&\ge S(\mathcal{A},z)-\sum_{z\le p\le x^{\alpha/\beta}}\left(\alpha-\beta\frac{\log p}{\log x}\right)S(\mathcal{A}_p,z).\end{aligned} \tag{5.134}$$

First, we proceed to obtain a lower bound for $S(\mathcal{A},z)$. Put $\epsilon_0=10^{-6}$. By Lemma 5.13 with $M\le x^{1/2-6\epsilon_0}$ and $N\le x^{1/8-2\epsilon_0}$, we have

$$\sum_{M<m\le M_1}\sum_{N<n\le N_1}a_mb_nr(\mathcal{A},mn)\ll yx^{-\epsilon_0} \tag{5.135}$$

for any choice of $M_1\le 2M$, $N_1\le 2N$, $|a_m|\le 1$, $|b_n|\le 1$. We now appeal to Theorem 5.11 with $M=x^{1/2-6\epsilon_0}$, $N=x^{1/8-2\epsilon_0}$, $X=y$, $z=x^{\delta}$. It yields

$$S(\mathcal{A},z)\ge yV(z)\left(f\left(\delta^{-1}(5/8-8\epsilon_0)\right)-c_9\epsilon-O_\epsilon\left((\log x)^{-1/3}\right)\right).$$

Note that we have used (5.135) to estimate the remainder $\mathcal{R}^-(M,N)$ in (5.123). Choosing $\epsilon=(2000c_9)^{-1}$, we deduce that, for $x\to\infty$,

$$S(\mathcal{A},z)\ge yV(z)\left(f\left(\delta^{-1}(5/8-8\epsilon_0)\right)-0.001\right). \tag{5.136}$$

Next, we turn to the sum on the right side of (5.134). To this end we require that

$$\alpha/\beta\le 1/2-7\epsilon_0. \tag{5.137}$$

Using a dyadic argument, we split the interval $[z, x^{\alpha/\beta}]$ into $O(\log x)$ subintervals $[x^u, Cx^u]$, $C \le 2$. For $p \in [x^u, Cx^u]$, we apply Theorem 5.11 with $M = M_u = x^{1/2-u-6\epsilon_0}$, $N = x^{1/8-2\epsilon_0}$, $X = X_p = yp^{-1}$, $z = x^\delta$:

$$S(\mathcal{A}_p, z) \le X_p V(z) \left( F\left(\frac{5/8 - u - 8\epsilon_0}{\delta}\right) + c_9\epsilon + O_\epsilon\big((\log x)^{-1/3}\big) \right) + \mathcal{R}_p(M_u, N), \tag{5.138}$$

where

$$\mathcal{R}_p(M_u, N) = \sum_{j \le J} \sum_{\substack{m \le M_u \\ m \mid P(z)}} \sum_{\substack{n \le N \\ n \mid P(z)}} a_{m,j} b_{n,j} r(\mathcal{A}_p, mn).$$

Here, the coefficients $a_{m,j}, b_{n,j}$ independent of $p$ and satisfy $|a_{m,j}| \le 1$, $|b_{n,j}| \le 1$. When $m$ and $n$ are divisors of $P(z)$ and $p \ge z$, we have $r(\mathcal{A}_p, mn) = r(\mathcal{A}, mnp)$. Hence, on writing $k = pm$, we get

$$\sum_{x^u < p \le Cx^u} \left(\alpha - \beta\frac{\log p}{\log x}\right) \mathcal{R}_p(M_u, N) \ll_\epsilon \left| \sum_{k \le K} \sum_{n \le N} c_k b_n r(\mathcal{A}, kn) \right|,$$

where $K = 2x^{1/2-6\epsilon_0}$, $|c_k| \le 1$, $|b_n| \le 1$. Thus, by (5.135),

$$\sum_{x^u < p \le Cx^u} \left(\alpha - \beta\frac{\log p}{\log x}\right) \mathcal{R}_p(M_u, N) \ll_\epsilon yx^{-\epsilon_0/2}. \tag{5.139}$$

Furthermore, when $x^u \le p \le 2x^u$, we have

$$F\left(\frac{5/8 - u - 8\epsilon_0}{\delta}\right) - F\left(\frac{5/8 - 8\epsilon_0}{\delta} - \frac{\log p}{\log z}\right) \ll (\log x)^{-1}. \tag{5.140}$$

Combining (5.138)–(5.140), we find that

$$\sum_{z \le p \le x^{\alpha/\beta}} \left(\alpha - \beta\frac{\log p}{\log x}\right) S(\mathcal{A}_p, z) \le yV(z)\big(\sigma_1 + \sigma_2 c_9\epsilon + O_\epsilon\big(\sigma_2(\log x)^{-1/3}\big)\big),$$

where

$$\sigma_1 = \sum_{x^\delta \le p \le x^{\alpha/\beta}} \frac{1}{p} \left(\alpha - \beta\frac{\log p}{\log x}\right) F\left(\frac{5/8 - 8\epsilon_0}{\delta} - \frac{\log p}{\log z}\right),$$

$$\sigma_2 = \sum_{x^\delta \le p \le x^{\alpha/\beta}} \frac{1}{p} \left(\alpha - \beta\frac{\log p}{\log x}\right).$$

By Theorem 1.9, $\sigma_2 \ll 1$, so choosing $\epsilon$ sufficiently small, we conclude that when $x \to \infty$,

$$\sum_{z \le p \le x^{\alpha/\beta}} \left(\alpha - \beta\frac{\log p}{\log x}\right) S(\mathcal{A}_p, z) \le yV(z)(\sigma_1 + 0.001). \tag{5.141}$$

Finally, we evaluate $\sigma_1$. Let $g(u) = x^{-u}(\alpha - \beta u)F\left(\delta^{-1}(5/8 - 8\epsilon_0 - u)\right)$ and note that

$$g(u) \ll x^{-u} \qquad \text{and} \qquad g'(u) \ll x^{-u}(\log x).$$

Using Stieltjes integration by parts and the PNT, we obtain

$$\begin{aligned}
\sigma_1 &= \int_\delta^{\alpha/\beta} g(u)\, d\big(\pi(x^u) - \pi(x^\delta)\big) \\
&= -\int_\delta^{\alpha/\beta} \big(\pi(x^u) - \pi(x^\delta)\big) g'(u)\, du \\
&= -\int_\delta^{\alpha/\beta} \int_{x^\delta}^{x^u} \frac{dt}{\log t}\, dg(u) + O\big((\log x)^{-1}\big) \\
&= \int_\delta^{\alpha/\beta} \frac{g(u)x^u(\log x)}{\log(x^u)}\, du + O\big((\log x)^{-1}\big) \\
&= \int_\delta^{\alpha/\beta} u^{-1}(\alpha - \beta u)F\left(\delta^{-1}(5/8 - 8\epsilon_0 - u)\right)\, du + O\big((\log x)^{-1}\big).
\end{aligned}$$

From the last calculation, (5.134), (5.136), and (5.141), we deduce that

$$\Sigma(\alpha, \beta, \delta) \geq yV(z)(\sigma_3 - 0.003),$$

where

$$\sigma_3 = f\left(\delta^{-1}(5/8 - 8\epsilon_0)\right) - \int_\delta^{\alpha/\beta} u^{-1}(\alpha - \beta u)F\left(\delta^{-1}(5/8 - 8\epsilon_0 - u)\right)\, du.$$

Hence, it remains to choose $\alpha, \beta, \delta$ satisfying (5.133) and (5.137) and such that $\sigma_3 > 0.003$. In order to simplify the calculations, we choose

$$\delta = 5/32 - 2\epsilon_0, \qquad \alpha = 4/3, \qquad \beta = 3,$$

although a slightly better choice would have been

$$\delta = 5/32 - 2\epsilon_0, \qquad \alpha = (\beta + 1)/3, \qquad \beta = 2.1. \tag{5.142}$$

Then (5.133) and (5.137) hold and

$$\begin{aligned}
\sigma_3 &= f(4) - \frac{1}{3}\int_\delta^{4/9} u^{-1}(4 - 9u)F\left(4 - \delta^{-1}u\right)\, du \\
&= 2e^\gamma\left(\frac{\ln 3}{4} - \frac{1}{3}\int_\delta^{4/9} \frac{4 - 9u}{u(4 - \delta^{-1}u)}\, du\right) \\
&= 2e^\gamma\left(\frac{\ln 3}{4} - \frac{1}{3}\ln\left(\frac{3}{9\delta - 1}\right) + 3\delta\ln\left(\frac{6.75\delta}{9\delta - 1}\right)\right) \geq 0.11e^\gamma > 0.11.
\end{aligned}$$

This completes the proof of Chen's theorem. ∎

**Remark.** It is not difficult to obtain *some* improvement on Theorem 8. Indeed, after changing the choice of $y$ to $y = x^\theta$, we alter our choices so that

$$\delta = (7\theta - 1)/16 - 2\epsilon_0, \qquad \alpha/\beta \le \theta - 7\epsilon_0. \tag{5.143}$$

Choosing first $\alpha = (\beta + 1)/3$ and then $\beta$ so that the second inequality in (5.143) is nearly exact (cf. (5.142)), we derive a lower bound

$$\Sigma(\alpha, \beta, \delta) \ge x^\theta V(z)(\sigma_4(\theta) - 0.003),$$

where $\sigma_4(\theta)$ is a function of $\theta$ similar to $\sigma_3$ above. Finally, we try to choose $\theta$ so that $\sigma_4(\theta) > 0.003$. Following this strategy, one easily finds that one can replace the interval $(x - x^{1/2}, x]$ in Theorem 8 by $(x - x^{0.46}, x]$. As we mentioned at the end of §5.7.3, further improvements arise from the use of sharper exponential sum estimates and/or more sophisticated versions of the sum $\Sigma(\alpha, \beta, \delta)$ above. Such matters, however, go beyond the scope of these lectures. ■

## Exercises

1. Prove (5.11).
2. Prove Theorem 5.3.
3. (a) Suppose that $f(x)(\log x)^{-1} \to \infty$ and

$$\sum_{q \le Q} \max_{(a,q)=1} \left| \psi(x; q, a) - \frac{x}{\phi(q)} \right| \ll x f(x)^{-1}.$$

   Show that the asymptotic formula

$$\psi(x; q, a) = \frac{x}{\phi(q)}\big(1 + o(1)\big) \qquad \text{as } x \to \infty \tag{$*$}$$

   holds for all arithmetic progressions $a \bmod q$, with $1 \le q \le \min\big(Q, f(x)(\log x)^{-1}\big)$ and $(a, q) = 1$. In particular, a version of the Bombieri–Vinogradov theorem with $x \exp\big(-2(\log x)^\delta\big)$, $\delta > 0$, in place of the term $x(\log x)^{-A}$ on the right side of (5.9) would establish $(*)$ for all arithmetic progressions with moduli $q \le \exp\big((\log x)^\delta\big)$, thus yielding an improvement on the Siegel–Walfisz theorem.

   (b) Obtain a variant of the result of part (a) relating to the Barban–Davenport–Halberstam theorem.
4. Prove (5.40).
5. Prove that:

   (a) $\phi(n) \gg n(\log\log n)^{-1}$ for all $n \ge 10$;

   (b) $\displaystyle\sum_{n \le x} \frac{n^2}{\phi(n)^2} \ll x$;

   (c) $\displaystyle\sum_{n > x} \phi(n)^{-2} \ll x^{-1}$.
6. Let $J(n)$ be defined by (5.41). Prove that $J(n) = \frac{1}{2}n^2$.

7. The point of this exercise is to establish (5.44).

   (a) Use Lemma 5.1 with $U = V$ to decompose $f(\alpha)$ into type I sums of the form (5.3) with $f(n) = e(\alpha n)$ and $M \ll U^2$ and type II sums of the form (5.4) with the same $f(n)$ and $U \ll M, N \ll xU^{-1}$.

   (b) Let $\Sigma_1$ be any of the type I sums arising from the decomposition in part (a). Prove that

$$\Sigma_1 \ll (\log n) \sum_{m \le U^2} \min\big(n/m, \|\alpha m\|^{-1}\big).$$

   (c) Let $\Sigma_2$ be any of the type II sums arising from the decomposition in part (a). Prove that

$$|\Sigma_2|^2 \ll M(\log n)^7 \sum_{u \le nM^{-1}} \sum_{v \le nM^{-1}} \min\big(M, \|\alpha(u-v)\|^{-1}\big),$$

   for some $M$ with $U \le M \le xU^{-1}$.

   (d) Suppose that $M, N, \alpha$ are real numbers with $M, N \ge 1$, and that $|\alpha - a/q| \le q^{-2}$ with $(a, q) = 1$. Then

$$\sum_{m \le M} \min\big(MNm^{-1}, \|\alpha m\|^{-1}\big) \ll \big(MNq^{-1} + M + q\big)(\log 2MNq).$$

   [Hint: See Vaughan [54, Lemma 2.2].]

   (e) Suppose that $|\alpha - a/q| \le q^{-2}$. Using the results of parts (b)–(d), show that

$$\Sigma_1 \ll (\log n)^2\big(nq^{-1} + U^2 + q\big)$$

   and

$$\Sigma_2 \ll (\log n)^4\big(nq^{-1/2} + nU^{-1/2} + n^{1/2}U^{1/2} + n^{1/2}q^{1/2}\big).$$

   Noting that $U = n^{2/5}$ is the choice that optimizes these bounds, deduce (5.44).

8. The purpose of this exercise is to establish Lemma 5.6. Let $N = [Q]$ and consider the numbers

$$0, \{\alpha\}, \{2\alpha\}, \dots, \{N\alpha\}, 1. \tag{$*$}$$

   Show that some interval $\big[(k-1)(N+1)^{-1}, k(N+1)^{-1}\big]$, $1 \le k \le N+1$, contains at least two of the numbers $(*)$. From this, deduce Dirichlet's theorem.

9. The purpose of this exercise is to give an alternative proof of Lemma 5.6.

   (a) Show that Dirichlet's theorem is equivalent to the inequality

$$\delta_\alpha = \min\big\{\|n\alpha\| \ : \ 1 \le n \le N\big\} \le (N+1)^{-1}. \tag{$*$}$$

   (b) Suppose that $0 < \delta \le \frac{1}{2}$ and define the 1-periodic function

$$f_\delta(x) = \max(\delta - \|x\|, 0).$$

   Prove that the $n$th Fourier coefficient of $f_\delta$ is given by

$$\hat{f}_\delta(n) = \begin{cases} \delta^2 & \text{if } m = 0, \\ (\sin \pi\delta m)^2/(\pi m)^2 & \text{if } m \ne 0. \end{cases}$$

(c) Suppose that $\delta_\alpha > 0$ and set $\delta = \delta_\alpha$. Observe that

$$\delta = \sum_{|n|\le N} \left(1 - \frac{|n|}{N+1}\right) f_\delta(n\alpha).$$

Deduce that

$$\delta = \sum_{m=-\infty}^{\infty} \hat{f}_\delta(m) K_N(m\alpha),$$

where $K_N(x)$ is the Fejér kernel

$$K_N(x) = \sum_{|n|\le N} \left(1 - \frac{|n|}{N+1}\right) e(nx) = \frac{1}{N+1}\left(\frac{\sin \pi(N+1)x}{\sin \pi x}\right)^2.$$

(d) Use the result of part (c) to prove (∗).

10. Let $\mathcal{A}$ be the set of integers $n \le X$, $\mathfrak{P}$ the set of all primes, $z = X^{1/2}$. Observe that with these choices, (5.102) takes the form

$$S(\mathcal{A}, \mathfrak{P}, z) = X \prod_{p<z} \left(1 - p^{-1}\right) + R(\mathcal{A}, z).$$

Hence, under the hypothesis

$$R(\mathcal{A}, z) = o\big(X(\log X)^{-1}\big) \qquad \text{as } X \to \infty, \tag{∗}$$

one obtains

$$\pi(X) \sim X \prod_{p<z} \left(1 - p^{-1}\right) \sim \frac{e^{-\gamma} X}{\log z} = \frac{2e^{-\gamma} X}{\log X},$$

which contradicts the PNT. Therefore, (∗) must be false.

11. Fill the details of Example 5.7.2.

12. The purpose of this exercise is to construct a $C^\infty$-function $f$ with the properties required in the proof of Lemma 5.12.

(a) Define the function

$$g(x) = \begin{cases} \exp\left((x-1)^{-1} - x^{-1}\right) & \text{if } 0 < x < 1, \\ 0 & \text{otherwise.} \end{cases}$$

Show that $g \in C^\infty(\mathbb{R})$.

(b) Let $G(x) = \int_{-\infty}^{x} g(t)\,dt$, where $g$ is the function from part (a). Show that the function $h(x) = G(x)/G(1)$ is a non-decreasing $C^\infty$-function such that $h(x) = 0$ when $x \le 0$ and $h(x) = 1$ when $x \ge 1$.

(c) Suppose that $\alpha < \beta$ and $\delta > 0$. Let $h(x)$ be the function from part (b) and define

$$f(x) = h((x-\alpha)/\delta + 1) - h((x-\beta)/\delta).$$

Then $f$ is a $C^\infty$-function, supported in $[\alpha - \delta, \beta + \delta]$ and such that

$$f(x) = 1 \quad (\alpha \le x \le \beta) \qquad \text{and} \qquad f^{(j)}(x) \ll \delta^{-j} \quad (j \ge 0).$$

13. Prove (5.140).

14. In the remark at the end of §5.8, we sketched the proof of the following result:

For sufficiently large $x$, the interval $(x - x^{0.46}, x]$ always contains a $P_2$-number.

Fill the details of the proof.

# Bibliography


[1] R. C. Baker and G. Harman, *The difference between consecutive primes*, Proc. London Math. Soc. (3) **72** (1996), 261–280.

[2] R. C. Baker, G. Harman, and J. Pintz, *The difference between consecutive primes II*, Proc. London Math. Soc. (3) **83** (2001), 532–562.

[3] M. B. Barban, *The "large sieve" method and its applications to number theory*, Uspehi Mat. Nauk **21** (1966), 51–102, in Russian.

[4] R. P. Boas, *Entire Functions*, Academic Press, 1954.

[5] E. Bombieri, *On the large sieve*, Mathematika **12** (1965), 201–225.

[6] E. Bombieri, J. B. Friedlander, and H. Iwaniec, *Primes in arithmetic progressions to large moduli*, Acta Math. **156** (1986), 203–251.

[7] ———, *Primes in arithmetic progressions to large moduli II*, Math. Ann. **277** (1987), 361–393.

[8] ———, *Primes in arithmetic progressions to large moduli III*, J. Amer. Math. Soc. **2** (1989), 215–224.

[9] J. R. Chen, *On the representation of large even integer as the sum of a prime and the product of at most two primes*, Sci. Sinica **16** (1973), 157–175.

[10] ———, *On the distribution of almost primes in an interval*, Sci. Sinica **18** (1975), 611–627.

[11] N. G. Chudakov, *On zeros of Dirichlet's L-functions*, Mat. Sb. (N.S.) **1** (1936), 591–602.

[12] H. Cramér, *Some theorems concerning prime numbers*, Arkiv Mat. Astronom. Fysik **15** (1920), 1–32.

[13] ———, *On the order of magnitude of the difference between consecutive primes*, Acta Arith. **2** (1937), 23–46.

[14] H. Davenport, *Multiplicative Number Theory*, 3rd ed., Springer-Verlag, 2000.

[15] H. Davenport and H. Halberstam, *Primes in arithmetic progressions*, Michigan Math. J. **13** (1966), 485–489.

[16] P. Erdös, *On the difference of consecutive primes*, Quart. J. Math. Oxford **6** (1935), 124–128.

[17] S. W. Graham and G. A. Kolesnik, *Van der Corput's Method of Exponential Sums*, Cambridge University Press, 1991.

[18] G. Greaves, *Sieves in Number Theory*, Springer-Verlag, 2001.

[19] H. Halberstam and H.-E. Richert, *Sieve Methods*, Academic Press, 1974.

[20] G. H. Hardy, *On Dirichlet's divisor problem*, Proc. London Math. Soc. (2) **15** (1916), 1–25.

[21] G. H. Hardy and J. E. Littlewood, *Some problems of 'Partitio Numerorum' III: On the expression of a number as a sum of primes*, Acta Math. **44** (1923), 1–70.

[22] D. R. Heath-Brown, *Prime numbers in short intervals and a generalized Vaughan identity*, Canad. J. Math. **34** (1982), 1365–1377.

[23] ———, *The number of primes in a short interval*, J. Reine Angew. Math. **389** (1988), 22–63.

[24] D. R. Heath-Brown and H. Iwaniec, *On the difference between consecutive primes*, Invent. Math. **55** (1979), 49–69.

[25] H. Heilbronn, *Über den Primzahlsatz von Herrn Hoheisel*, Math. Z. **36** (1933), 394–423.

[26] G. Hoheisel, *Primzahlprobleme in der Analysis*, Sitz. Preuss. Akad. Wiss. **2** (1930), 1–13.

[27] M. N. Huxley, *On the difference between consecutive primes*, Invent. Math. **15** (1972), 164–170.

[28] ———, *Area, Lattice Points, and Exponential Sums*, Oxford University Press, 1996.

[29] ———, *Exponential sums and lattice points III*, Proc. London Math. Soc. (3) **87** (2003), 591–609.

[30] A. E. Ingham, *On the difference between consecutive primes*, Quart. J. Math. Oxford **8** (1937), 255–266.

[31] A. Ivić, *The Riemann Zeta-Function*, Wiley, 1985.

[32] H. Iwaniec, *A new form of the error term in the linear sieve*, Acta Arith. **37** (1980), 307–320.

[33] ———, *Rosser's sieve*, Acta Arith. **36** (1980), 171–202.

[34] H. Iwaniec and M. Jutila, *Primes in short intervals*, Ark. Mat. **17** (1979), 167–176.

[35] H. Iwaniec and J. Pintz, *Primes in short intervals*, Monatsh. Math. **98** (1984), 115–143.

[36] C. H. Jia, *Almost all short intervals containing prime numbers*, Acta Arith. **76** (1996), 21–84.

[37] A. A. Karatsuba and S. M. Voronin, *The Riemann Zeta-Function*, Walter de Gruyter, 1992.

[38] N. M. Korobov, *Estimates of trigonometric sums and their applications*, Uspehi Mat. Nauk **13** (1958), 185–192, in Russian.

[39] H. Q. Liu, *Almost primes in short intervals*, J. Number Theory **57** (1996), 303–322.

[40] S. T. Lou and Q. Yao, *A Chebychev's type of prime number theorem in a short interval II*, Hardy–Ramanujan J. **15** (1992), 1–33.

[41] H. L. Montgomery, *Topics in Multiplicative Number Theory*, Springer-Verlag, 1971.

[42] H. L. Montgomery and R. C. Vaughan, *The large sieve*, Mathematika **20** (1973), 119–134.

[43] A. Page, *On the number of primes in an arithmetic progression*, Proc. London Math. Soc. (2) **39** (1935), 116–141.

[44] J. Pintz, *Very large gaps between consecutive primes*, J. Number Theory **63** (1997), 286–301.

[45] G. Pólya, *Über die Verteilung der quadratischen Reste und Nichtreste*, Nachr. Akad. Wiss. Göttingen Math.-Phys. (1918), 21–29.

[46] R. A. Rankin, *The difference between consecutive prime numbers*, J. London Math. Soc. **13** (1938), 242–247.

[47] G. F. B. Riemann, *Über die Anzahl der Primzahlen unter einen gegebenen Grösse*, Monatsber. Berliner Akad. (1859), 671–680.

[48] A. Selberg, *On the normal density of primes in small intervals, and the difference between consecutive primes*, Arch. Math. Naturvid. **47** (1943), 87–105.

[49] C. L. Siegel, *Über die Classenzahl quadratischer Körper*, Acta Arith. **1** (1935), 83–86.

[50] E. C. Titchmarsh, *The Theory of the Riemann Zeta-Function*, 2nd ed., Oxford University Press, 1986.

[51] J. D. Vaaler, *Some extremal functions in Fourier analysis*, Bull. Amer. Math. Soc. (N.S.) **12** (1985), 183–216.

[52] R. C. Vaughan, *Mean value theorems in prime number theory*, J. London Math. Soc. (2) **10** (1975), 153–162.

[53] ———, *Sommes trigonométrique sur les nombres premiers*, C. R. Acad. Sci. Paris Sér. A **258** (1977), 981–983.

[54] ———, *The Hardy–Littlewood Method*, 2nd ed., Cambridge University Press, 1997.

[55] A. I. Vinogradov, *The density hypothesis for Dirichlet L-series*, Izv. Akad. Nauk SSSR Ser. Mat. **29** (1965), 903–934, in Russian.

[56] I. M. Vinogradov, *Sur la distribution des résidus et non résidus de puissances*, Perm Univ. Fiz.-Mat. Zh. **1** (1918), 18–24.

[57] ———, *Representation of an odd number as the sum of three primes*, Dokl. Akad. Nauk SSSR **15** (1937), 291–294, in Russian.

[58] ———, *A new estimate for* $\zeta(1 + it)$, Izv. Akad. Nauk SSSR Ser. Mat. **22** (1958), 161–164, in Russian.

[59] A. Zygmund, *Trigonometric Series*, 3rd ed., Cambridge University Press, 2002.