Probabilistic Number Theory

Dr. Jörn Steuding

Dedicated to Prof. JONAS KUBILIUS
and the members of his working group at Vilnius University,
for their outstanding work in probabilistic number theory
and their kind hospitality!

This is a first introduction to *Probabilistic Number Theory*, based on a course given at the JOHANN WOLFGANG GOETHE-Universität Frankfurt in 2001. We focus ourselves to some classical results on the prime divisor counting function $\omega(n)$ which were discovered in the first half of the 20th century. Nowadays, these facts are the basics for heuristical arguments on the expected running time of algorithms in cryptography. Furthermore, this gives a first view inside the methods and problems in this modern field of research. Especially the growing interest in probabilistic algorithms, which give with a certain *probability* the right answer (e.g. probabilistic prime number tests), underlines the power and influence of doing number theory from a probability theoretical point of view.

For our studies we require only a small background in *elementary* number theory as well as in probability theory, and, for the second half additionally, the fundamentals of complex analysis; good recommendations to refresh the knowledge on these topics are [16], [14] and [21]. We will use the same standard notations as in [30], which is also the main source of this course.

I am very grateful to RASA ŠLEŽEVIČIENĖ for her interest and her several helpful comments, remarks and corrections.

JÖRN STEUDING, Frankfurt 01/30/2002.

Contents

Chapter 1

Introduction

Instead of *probabilistic number theory* one should speak about *studying arithmetic functions with probabilistic methods*. First approaches in this direction date back to

- Gauss, who used in 1791 probabilistic arguments for his speculations on the number of products consisting of exactly k distinct prime factors below a given bound; the case $k = 1$ led to the prime number theorem (see [10], vol.10, p.11) - we shall return to this question in Chapter 16;

- Cesaro, who observed in 1881 that the *probability* that two *randomly* chosen integers are coprime is $\frac{6}{\pi^2}$ (see [1]) - we will prove this result in Chapter 4.

In number theory one is interested in the value distribution of **arithmetic functions** $f : \mathbb{N} \to \mathbb{C}$ (i.e. complex-valued sequences). An arithmetic function f is said to be **additive** if

$$f(m \cdot n) = f(m) + f(n) \qquad \text{for} \qquad \gcd(m, n) = 1,$$

and f is called **multiplicative** if

$$f(m \cdot n) = f(m) \cdot f(n) \qquad \text{for} \qquad \gcd(m, n) = 1;$$

f is **completely additive**- and **completely multiplicative**, resp., when the condition of coprimality can be removed (the symbol $\gcd(m, n)$ stands, as usual, for the greatest common divisor of the integers m and n). Obviously, the values of additive or multiplicative functions are determined by the values on the prime powers, or even on the primes when the function in question is completely additive or completely multiplicative. But prime number distribution is a difficult task.

We shall give two important examples. Let the **prime divisor counting functions** $\omega(n)$ and $\Omega(n)$ of a positive integer n (with and without multiplicities, resp.)

be defined by

$$\omega(n) = \sum_{p|n} 1 \qquad \text{and} \qquad \Omega(n) = \sum_{p|n} \nu(n;p),$$

resp., where $\nu(n;p)$ is the exponent of the prime p in the unique prime factorization of n:

$$n = \prod_p p^{\nu(n;p)};$$

here and in the sequel p denotes always a prime number (we recall that $p|n$ means that the prime p divides the integer n, and when this notation occurs under a product or a sum, then the product or the summation is taken over all p which divide n). Obviously, n is a prime number if and only if $\Omega(n) = 1$. Therefore, the distribution of prime numbers is hidden in the values of $\Omega(n)$.

We note

Lemma 1.1 *$\omega(n)$ is an additive, and $\Omega(n)$ is a completely additive arithmetic function.*

Exercise 1.1 *(i) Prove the lemma above.*

(ii) Give examples of multiplicative and completely multiplicative arithmetic functions.

When we investigate arithmetic functions we should not expect *exact* formulas. Usually, the values $f(n)$ are spread too widely. For example, EULER**'s totient** $\varphi(n)$ counts the number of prime residue classes mod n:

$$\varphi(n) := \sharp\{1 \leq a \leq n \, : \, \gcd(a,n) = 1\}.$$

It was proved by SCHINZEL [26] that the values $\frac{\varphi(n+1)}{\varphi(n)}, n \in \mathbb{N}$, lie everywhere dense on the positive real axis. Further, it is easy to see that

$$\liminf_{n\to\infty} \frac{\varphi(n)}{n} = 0 \qquad \text{and} \qquad \limsup_{n\to\infty} \frac{\varphi(n)}{n} = 1. \tag{1.1}$$

Exercise 1.2 *(i) Prove the identity*

$$\varphi(n) = n \prod_{p|n} \left(1 - \frac{1}{p}\right).$$

In particular, $\varphi(n)$ is multiplicative.
(Hint: remind that $am + bn$ runs through a complete residue system mod mn *when a and b run through complete residue systems* mod n *and* mod m*, resp., if m and n are coprime; see for this and for some basics on congruences and residues [14],* § *V.)*

(ii) Prove formulae (1.1).
(Hint: make use of formula (2.3) below.)

(iii) Try to find lower and upper bounds for $\omega(n)$ and $\Omega(n)$.

In our studies on the value distribution of arithmetic functions we are restricted to *asymptotic* formulas. Hence, we need a notion to deal with error terms. We write

$$f(x) = O(g(x)) \qquad \text{and} \qquad f(x) \ll g(x),$$

resp., when there exists a positive function $g(x)$ such that

$$\limsup_{x\to\infty} \frac{|f(x)|}{g(x)}$$

exists. Then the function $f(x)$ grows not faster than $g(x)$ (up to a multiplicative constant), and, hopefully, the growth of the function $g(x)$ is easier to understand than the one of $f(x)$, as $x \to \infty$. This is not only a convenient notation due to LANDAU and VINOGRADOV, but, in the sense of developping the right language, an important contribution to mathematics as well.

We illustrate this with an easy example. What is the order of growth of the truncated (divergent) harmonic series

$$\sum_{n\le x} \frac{1}{n},$$

as $x \to \infty$? Obviously, for $n \geq 2$,

$$\frac{1}{n} < \int_{n-1}^{n} \frac{\mathrm{d}t}{t} < \frac{1}{n-1}.$$

Denote by $[x]$ the maximum over all integers $\leq x$, then, by summation over $2 \leq n \leq [x]$,

$$\sum_{n=2}^{[x]} \frac{1}{n} < \int_1^{[x]} \frac{\mathrm{d}t}{t} < \sum_{n=1}^{[x]-1} \frac{1}{n}.$$

Therefore integration yields

$$\sum_{n\le x} \frac{1}{n} = \int_1^x \frac{\mathrm{d}t}{t} + O(1) = \log x + O(1); \tag{1.2}$$

here and in the sequel log denotes always the natural logarithm, i.e. the logarithm to the base $e = \exp(1)$. We learned above an important trick which we will use in the following several times: the sum over a *sufficiently smooth* function can be considered - up to a certain error - as a Riemann sum and its integral, resp., which is hopefully calculable.

Exercise 1.3 *Prove for $x \to \infty$ that*

(i) the number of squares $n^2 \le x$ is $\sqrt{x} + O(1)$;

(ii) $\log x \ll x^{\varepsilon}$ for any $\varepsilon > 0$;

(iii) $x^m \ll \exp(x)$ for any $m > 0$;

(iv) $\sum_{n \le x} n = \frac{1}{2}x^2 + O(x)$.

We return to number theory. In 1917 Hardy and Ramanujan [13] discovered the first deep result on the prime divisor counting function, namely that for fixed $\delta \in (0, \frac{1}{2})$ and $N \ge 3$

$$\frac{1}{N}\sharp\{n \le N : |\omega(n) - \log\log n| > (\log\log n)^{\frac{1}{2}+\delta}\} \ll \frac{1}{(\log\log N)^{2\delta}}. \tag{1.3}$$

Since the right hand side above tends to zero, as $N \to \infty$, the values of $\omega(n)$ with $n \le N$ are concentrated around $\log\log n$ (the set of integers n, for which $\omega(n)$ deviates from $\log\log n$, has zero density, in the language of densities; see Chapter 2). For example, a 50-digit number has on average only about 5 distinct prime divisors!

Moreover, Hardy and Ramanujan proved with similar arguments the corresponding result for $\Omega(n)$. Unfortunately, their approach is complicated and not extendable to other functions. In 1934 Turán [31] found a new proof based on the estimate

$$\sum_{n \le N} (\omega(n) - \log\log n)^2 \ll N \log\log N, \tag{1.4}$$

and an argument similar to Čebyšev's proof of the *law of large numbers* in probability theory (which was unknown to the young Turán). His approach allows generalizations (and we will deduce the Hardy-Ramanujan result (1.3) as an immediate consequence of a much more general result which holds for a large class of additive functions, namely the Turán-Kubilius inequality; see Chapter 6). The effect of Turán's paper was epoch-making. His ideas were the starting point for the development of *probabilistic number theory* in the following years.

To give finally a first glance on the influence of probabilistic methods on number theory we mention one of its highlights, discovered by Erdös and Kac [7] in 1939, namely that $\omega(n)$ satisfies (after a certain normalization) the Gaussian error law:

$$\lim_{N\to\infty} \frac{1}{N}\sharp\left\{n \le N : \frac{\omega(n) - \log\log N}{\sqrt{\log\log N}} \le x\right\} = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} \exp\left(-\frac{\tau^2}{2}\right) \mathrm{d}\tau. \tag{1.5}$$

Therefore, the values $\omega(n)$ are asymptotically *normally distributed* with expectation $\log\log n$ and standard deviation $\sqrt{\log\log n}$ (this goes much beyond (1.3); we will prove a stronger version of (1.5) in Chapter 13).

This classical result has some important implications to cryptography. For an analysis of the expected running time of many modern primality tests and factorization tests one needs heuristical arguments on the distribution of prime numbers and so-called **smooth** numbers, i.e. numbers which have only *small* prime divisors (see [27], §11).

For a deeper and more detailed history of probabilistic number theory read the highly recommendable introductions of [6] and [18].

Chapter 2

Densities on the set of positive integers

It is no wonder that probabilistic number theory has its roots in the 1930s. Only in 1933 Kolmogorov gave the first widely accepted axiomization of probability theory.

We recall these basics. A **probability space** is a triple $(\Omega, \mathcal{B}, \mathbf{P})$ consisting of the **sure event** Ω (a non-empty set), a **σ-algebra** $\mathcal{B}$ (i.e. a system of subsets of Ω, for example, the power set of Ω), and a **probability measure P**, i.e. a function $\mathbf{P} : \mathcal{B} \to [0,1]$ satisfying

- $\mathbf{P}(\Omega) = 1$,
- $\mathbf{P}(A) \geq 0$ for all $A \in \mathcal{B}$,
- $\mathbf{P}\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mathbf{P}(A_n)$ for all pairwise disjoint $A_n \in \mathcal{B}$.

Then, $\mathbf{P}(A)$ is the **probability** of $A \in \mathcal{B}$. We say that two events $A, B \in \mathcal{B}$ are **independent** if

$$\mathbf{P}(A \cap B) = \mathbf{P}(A) \cdot \mathbf{P}(B).$$

Based on Kolmogorov's axioms one can start to define random variables, their expectations and much more to build up the powerful theory of probability (see [16] for more details).

But our aim is different. We are interested to obtain knowledge on the value distribution of arithmetic functions. The first idea is to define a probability law on the set of positive integers. However, we are restricted to be very careful as the following statement shows: by intuition we expect that the *probability*, that a *randomly* chosen integer is even, equals $\frac{1}{2}$, but:

Theorem 2.1 *There exists no probability law on* $\mathbb{N}$ *such that*

$$\mathbf{P}(a\mathbb{N}) = \frac{1}{a} \qquad (a \in \mathbb{N}), \tag{2.1}$$

where $a\mathbb{N} := \{n \in \mathbb{N} \,:\, n \equiv 0 \bmod a\}$*.*

Proof via contradiction. By the Chinese remainder theorem (see [14], §VIII.1), one has for coprime integers a, b

$$a\mathbb{N} \cap b\mathbb{N} = ab\mathbb{N}.$$

Now assume additionally that $\mathbf{P}$ is a probability measure on $\mathbb{N}$ satisfying (2.1), then

$$\mathbf{P}(a\mathbb{N} \cap b\mathbb{N}) = \mathbf{P}(ab\mathbb{N}) = \frac{1}{ab} = \mathbf{P}(a\mathbb{N}) \cdot \mathbf{P}(b\mathbb{N}).$$

Thus, the events $a\mathbb{N}$ and $b\mathbb{N}$, and their complements

$$\mathbb{N}_a := \mathbb{N} \setminus a\mathbb{N} \qquad \text{and} \qquad \mathbb{N}_b := \mathbb{N} \setminus b\mathbb{N},$$

resp., are independent. Furthermore

$$\mathbf{P}(\mathbb{N}_a \cap \mathbb{N}_b) = (1 - \mathbf{P}(a\mathbb{N}))(1 - \mathbf{P}(b\mathbb{N})) = \left(1 - \frac{1}{a}\right)\left(1 - \frac{1}{b}\right).$$

By induction, we obtain for arbitrary integers $m < x$

$$\mathbf{P}(\{m\}) \leq \mathbf{P}\left(\bigcap_{m<p\leq x} \mathbb{N}_p\right) = \prod_{m<p\leq x}\left(1 - \frac{1}{p}\right); \tag{2.2}$$

here the inequality is caused by $m \in \mathbb{N}_p$ for all $p > m$). In view to the unique prime factorization of the integers and (1.2) we get

$$\prod_{p\leq x}\left(1 + \frac{1}{p} + \frac{1}{p^2} + \ldots\right) \geq \sum_{n\leq x} \frac{1}{n} = \log x + O(1).$$

Hence, by the geometric series expansion,

$$\prod_{p\leq x}\left(1 - \frac{1}{p}\right) \leq \frac{1}{\log x + O(1)}. \tag{2.3}$$

This leads with $x \to \infty$ in formula (2.2) to $\mathbf{P}(\{m\}) = 0$, giving the contradiction. •

In spite of that we may define a probability law on $\mathbb{N}$ as follows. Assume that

$$\sum_{n=1}^{\infty} \lambda_n = 1 \qquad \text{with} \qquad 0 \leq \lambda_n \leq 1,$$

then we set for any sequence $\mathcal{A} \subset \mathbb{N}$

$$\mathbf{P}(\mathcal{A}) = \sum_{n\in\mathcal{A}} \lambda_n.$$

Obviously, this defines a probability measure. Unfortunately, the probability of a sequence depends drastically on its initial values (since for any $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that $\mathbf{P}(\{1, 2, \ldots, N\}) \geq 1 - \varepsilon$).

To construct a model which fits more to our intuition we need the notion of *density.* Introducing a divergent series

$$\sum_{n=1}^{\infty} \lambda_n = \infty \qquad \text{with} \qquad \lambda_n \geq 0,$$

we define the **density** $\mathbf{d}(\mathcal{A})$ of a sequence $\mathcal{A}$ of positive integers to be the limit (when it exists)

$$\mathbf{d}(\mathcal{A}) = \lim_{x\to\infty} \frac{\sum_{n\leq x; n\in\mathcal{A}} \lambda_n}{\sum_{n\leq x} \lambda_n}. \tag{2.4}$$

This yields not a measure on $\mathbb{N}$ (since sequences do not form a σ-algebra, and densities are not subadditive). Nevertheless, the concept of density allows us to build up a model which matches to our intuition. Putting $\lambda_n = 1$ in (2.4), we obtain the **natural density** (when it exists)

$$\mathbf{d}\mathcal{A} = \lim_{x\to\infty} \frac{1}{x} \sharp\{n \leq x \,:\, n \in \mathcal{A}\};$$

moreover, the **lower** and **upper natural density** are given by

$$\underline{\mathbf{d}}\mathcal{A} = \liminf_{x\to\infty} \frac{1}{x} \sharp\{n \leq x \,:\, n \in \mathcal{A}\} \qquad \text{and} \qquad \overline{\mathbf{d}}\mathcal{A} = \limsup_{x\to\infty} \frac{1}{x} \sharp\{n \leq x \,:\, n \in \mathcal{A}\},$$

respectively. We give some examples. Any arithmetic progression $n \equiv b \bmod a$ has the natural density

$$\lim_{x\to\infty} \frac{1}{x} \left(\left[\frac{x}{a}\right] + O(1) \right) = \frac{1}{a},$$

corresponding to our intuition.

Exercise 2.1 *Show that*

(i) the sequence $a_1 < a_2 < \ldots$ has natural density $\alpha \in [0, 1]$ if, and only if,

$$\lim_{n\to\infty} \frac{n}{a_n} = \alpha;$$

(Hint: for the implication of necessity note that $n = \sharp\{j \,:\, a_j \leq a_n\}$.)

(ii) the sequence $\mathcal{A}$ of positive integers n with leading digit 1 in the decimal expansion has no natural density, since

$$\underline{\mathbf{d}}\mathcal{A} = \frac{1}{9} < \frac{5}{9} = \overline{\mathbf{d}}\mathcal{A}.$$

We note the following important(!) connection between natural density and probability theory: if $\nu_{\mathbf{N}}$ denotes the probability law of the **uniform distribution** with weight $\frac{1}{N}$ on $\{1, 2, \ldots, N\}$, i.e.

$$\nu_{\mathbf{N}}\mathcal{A} = \sum_{n\in\mathcal{A}} \lambda_n \qquad \text{with} \qquad \lambda_n = \begin{cases} \frac{1}{N} & \text{if} \quad n \le N, \\ 0 & \text{if} \quad n > N, \end{cases}$$

then (when the limit exists)

$$\lim_{N\to\infty} \nu_{\mathbf{N}}\mathcal{A} = \lim_{N\to\infty} \frac{1}{N}\sharp\{n \le N \,:\, n \in \mathcal{A}\} = \mathbf{d}\mathcal{A}.$$

Therefore, the natural density of a sequence is the limit of its frequency in the first N positive integers, as $N \to \infty$.

Setting $\lambda_n = \frac{1}{n}$ in (2.4), we obtain the **logarithmic density**

$$\delta\mathcal{A} := \lim_{x\to\infty} \frac{1}{\log x} \sum_{\substack{n\le x\\ n\in\mathcal{A}}} \frac{1}{n};$$

the **lower** and **upper logarithmic density** are given by

$$\underline{\delta}\mathcal{A} = \liminf_{x\to\infty} \frac{1}{\log x} \sum_{\substack{n\le x\\ n\in\mathcal{A}}} \frac{1}{n} \qquad \text{and} \qquad \overline{\delta}\mathcal{A} = \limsup_{x\to\infty} \frac{1}{\log x} \sum_{\substack{n\le x\\ n\in\mathcal{A}}} \frac{1}{n},$$

respectively; note that the occuring $\log x$ comes from (1.2).

Exercise 2.2 *Construct a sequence which has no logarithmic density.*

The following theorem gives a hint for the solution of the exercise above.

Theorem 2.2 *For any sequence $\mathcal{A} \subset \mathbb{N}$,*

$$\underline{\mathbf{d}}\mathcal{A} \le \underline{\delta}\mathcal{A} \le \overline{\delta}\mathcal{A} \le \overline{\mathbf{d}}\mathcal{A}.$$

In particular, a sequence with a natural density has a logarithmic density as well, and both densities are equal.

Before we give the proof we recall a convenient technique in number theory.

Lemma 2.3 (ABEL's partial summation) *Let $\lambda_1 < \lambda_2 < \ldots$ be a divergent sequence of real numbers, define for $\alpha_n \in \mathbb{C}$ the function $A(x) = \sum_{\lambda_n \leq x} \alpha_n$, and let $f(x)$ be a complex-valued, continuous differentiable function for $x \geq \lambda_1$. Then*

$$\sum_{\lambda_n \leq x} \alpha_n f(\lambda_n) = A(x)f(x) - \int_{\lambda_1}^{x} A(u)f'(u)\, du.$$

For those who are familiar with the RIEMANN-STIELTJES integral there is nearly nothing to show. Nevertheless,

Proof. We have

$$A(x)f(x) - \sum_{\lambda_n \leq x} \alpha_n f(\lambda_n) = \sum_{\lambda_n \leq x} \alpha_n (f(x) - f(\lambda_n)) = \sum_{\lambda_n \leq x} \int_{\lambda_n}^{x} \alpha_n f'(u)\, \mathrm{d}u.$$

Since $\lambda_1 \leq \lambda_n \leq u \leq x$, changing integration and summation yields the assertion. •

Proof of Theorem 2.2. Defining $A(x) = \sum_{n \leq x, n \in \mathcal{A}} 1$, partial summation yields, for $x \geq 1$,

$$L(x) := \sum_{\substack{n \leq x \\ n \in \mathcal{A}}} \frac{1}{n} = \frac{A(x)}{x} + \int_1^x \frac{A(t)}{t^2}\, \mathrm{d}t \tag{2.5}$$

For any $\varepsilon > 0$ exists a t_0 such that, for all $t > t_0$,

$$\underline{\mathbf{d}}\mathcal{A} - \varepsilon \leq \frac{A(t)}{t} \leq \overline{\mathbf{d}}\mathcal{A} + \varepsilon.$$

Thus, for $x > t_0$,

$$(\underline{\mathbf{d}}\mathcal{A} - \varepsilon)(\log x - \log t_0) = (\underline{\mathbf{d}}\mathcal{A} - \varepsilon) \int_{t_0}^{x} \frac{\mathrm{d}t}{t} \leq \int_1^x \frac{A(t)}{t^2}\, \mathrm{d}t,$$

and

$$\int_1^x \frac{A(t)}{t^2}\, \mathrm{d}t \leq \int_1^{t_0} \frac{\mathrm{d}t}{t} + (\overline{\mathbf{d}}\mathcal{A} + \varepsilon) \int_{t_0}^{x} \frac{\mathrm{d}t}{t} = (\overline{\mathbf{d}}\mathcal{A} + \varepsilon)(\log x - \log t_0) + \log t_0.$$

In view to (2.5) we obtain

$$(\underline{\mathbf{d}}\mathcal{A} - \varepsilon)\left(1 - \frac{\log t_0}{\log x}\right) \leq \frac{L(x)}{\log x} - \frac{A(x)}{x \log x} \leq (\overline{\mathbf{d}}\mathcal{A} + \varepsilon)\left(1 - \frac{\log t_0}{\log x}\right) + \frac{\log t_0}{\log x}.$$

Taking $\liminf$ and $\limsup$, as $x \to \infty$, and sending then $\varepsilon \to 0$, the assertion of the theorem follows. •

Exercise 2.3 *Show that the existence of the logarithmic density does not imply the existence of natural density.*
(Hint: have a look on the sequence $\mathcal{A}$ in Exercise 2.1.)

Taking $\lambda_n = \lambda_n(\sigma) = n^{-\sigma}$ in (2.4), we define the **analytic density** of a sequence $\mathcal{A} \subset \mathbb{N}$ by the limit (when it exists)

$$\lim_{\sigma \to 1+} \frac{1}{\zeta(\sigma)} \sum_{n \in \mathcal{A}} \frac{1}{n^\sigma}, \tag{2.6}$$

where

$$\zeta(s) := \sum_{n=1}^{\infty} \frac{1}{n^s} = \prod_p \left(1 - \frac{1}{p^s}\right)^{-1} \tag{2.7}$$

is the famous Riemann **zeta-function**; obviously the series converges for $s > 1$ (resp., from the complex point of view, in the half plane Re $s > 1$). Note that the equality between the infinite series and the infinite product is a consequence of the *unique prime factorization* in $\mathbb{Z}$ (for more details see [30], §II.1). By partial summation it turns out that one may replace the reciprocal of $\zeta(\sigma)$ in (2.6) by the factor $\sigma - 1$. We leave this training on the use of Lemma 2.3 to the interested reader.

Exercise 2.4 *Write $s = \sigma + it$ with $i := \sqrt{-1}$ and $\sigma, t \in \mathbb{R}$. Prove for $\sigma > 0$*

$$\zeta(s) = \frac{s}{s-1} - s \int_1^\infty \frac{x - [x]}{x^{s+1}}\, dx.$$

In particular, $\zeta(s)$ has an analytic continuation to the half plane $\sigma > 0$ except for a simple pole at $s = 1$ with residue 1.
(Hint: partial summation with $\sum_{N<n\leq M} n^{-s}$; the statement about the analytic continuation requires some fundamentals from the theory of functions.)

The analytic and arithmetic properties of $\zeta(s)$ make the analytic density very useful for a plenty of applications. We note

Theorem 2.4 *A sequence $\mathcal{A}$ of positive integers has analytic density if and only if $\mathcal{A}$ has logarithmic density; in this case the two densities are equal.*

A proof can be found in [30], §III.1.

We conclude with a further density, which differs from the above given examples, but is very useful in questions concerning the addition of sequences of positive integers, defined by

$$\mathcal{A} + \mathcal{B} := \{a + b \,:\, a \in \mathcal{A}, b \in \mathcal{B}\}.$$

The SCHNIRELMANN **density** is defined by

$$\sigma(\mathcal{A}) = \inf_{n\geq 1} \frac{1}{n} \sharp\{m \leq n \,:\, m \in \mathcal{A}\}.$$

$\sigma(\mathcal{A})$ stresses the initial values in the sequence $\mathcal{A}$. For the addition of sequences one has MANN*'s inequality*

$$\sigma(\mathcal{A}+\mathcal{B}) \geq \min\{1, \sigma(\mathcal{A})\sigma(\mathcal{B})\};$$

the interested reader can find a proof of this result and its implication to problems in additive number theory (for example, WARING**'s problem** of the representation of posiitve integers as sums of k-th powers, or the famous GOLDBACH **conjecture** which asks whether each even positive integer is the sum of two primes or not) in [12], §I.2.

As we will see in the sequel, the concept of density makes it possible in our investigations on the value distribution of an arithmetic function to exclude *extremal* values, and to have a look on its *normal* behaviour.

Chapter 3

Limiting distributions of arithmetic functions

We recall from probability theory some basic notions. A **random variable** on a probability space $(\Omega, \mathcal{B}, \mathbf{P})$ is a measurable function X defined on Ω. When, for example, $\Omega = \mathbb{R}$, then the function $F(x) := \mathbf{P}(X(\omega) \in (-\infty, x])$ contains a lot of information about the random variable X and its values $X(\omega), \omega \in \Omega$. A **distribution function** is a non-decreasing, right-continuous function $\mathbf{F} : \mathbb{R} \to [0,1]$, satisfying

$$\mathbf{F}(-\infty) = 0 \qquad \text{and} \qquad \mathbf{F}(+\infty) = 1.$$

Denote by $\mathcal{D}(\mathbf{F})$ and $\mathcal{C}(\mathbf{F})$ the set of discontinuity points and continuity points of $\mathbf{F}$, respectively. Obviously, $\mathcal{D}(\mathbf{F}) \cup \mathcal{C}(\mathbf{F}) = \mathbb{R}$. Each discontinuity point z has the property $\mathbf{F}(z + \varepsilon) > \mathbf{F}(z - \varepsilon)$ for any $\varepsilon > 0$ (the converse is not true). Write $\mathcal{D}(\mathbf{F}) = \{z_k\}$, then the function

$$\mathcal{F}(z) = \sum_{z_k \le z} (\mathbf{F}(z_k) - \mathbf{F}(z_k-))$$

increases exclusively for $z = z_k$, and is constant in any closed interval free of discontinuity points z_k (it is a step-function). If $\mathcal{D}(\mathbf{F})$ is not empty, then $\mathcal{F}$ is up to a multiplicative constant a distribution function; such a distribution function is called **atomic**. Obviously, the function $\mathbf{F} - \mathcal{F}$ is continuous. A distribution function $\mathbf{F}$ is said to be **absolutely continuous** if there exists a positive, Lebesgue-integrable function h with

$$\mathbf{F}(z) = \int_{-\infty}^{z} h(t)\,\mathrm{d}t.$$

Finally, a distribution function $\mathbf{F}$ is **purely singular** if $\mathbf{F}$ is continuous with support on a subset $\mathcal{N} \subset \mathbb{R}$ with LEBESGUE measure zero, i.e.

$$\int_{\mathcal{N}} \mathrm{d}\mathbf{F}(z) = 1.$$

We note:

Theorem 3.1 (Lebesgue) *Each distribution function F has a unique representation*

$$\mathbf{F} = \alpha_1 \mathbf{F}_1 + \alpha_2 \mathbf{F}_2 + \alpha_3 \mathbf{F}_3,$$

where $\alpha_1, \alpha_2, \alpha_3$ are non-negative constants with $\alpha_1 + \alpha_2 + \alpha_3 = 1$, and where $\mathbf{F}_1$ is absolutely continuous, $\mathbf{F}_2$ is purely singular and $\mathbf{F}_3$ is atomic.

The proof follows from the observations above and the Theorem of RADON-NIKODYM; see [30], §III.2 and [16], §28.

The next important notion is *weak convergence.* We say that a sequence $\{\mathbf{F}_n\}$ of distribution functions **converges weakly** to a function $\mathbf{F}$ if

$$\lim_{n\to\infty} \mathbf{F}_n(z) = \mathbf{F}(z) \qquad \text{for all} \qquad z \in \mathcal{C}(\mathbf{F}),$$

i.e. pointwise convergence on the set of continuity points of the limit.

We give an interesting example from probability theory (without details). Let (X_j) be a sequence of independent and identically distributed random variables with expectation μ and variance $\sigma^2 \in (0, \infty)$. By the *central limit theorem* (see [16], §21), the distribution functions of the sequence of random variables

$$Y_n := \frac{1}{\sqrt{n\sigma^2}} \left(\sum_{j=1}^{n} X_j - n\mu \right)$$

converge weakly to the **standard Normal distribution**

$$\mathbf{\Phi}(x) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left(-\frac{\tau^2}{2}\right) \mathrm{d}\tau \tag{3.1}$$

(with expectation 0 and variance 1). In particular, we obtain for a sequence of independent random variables X_j with

$$\mathbf{P}(X_j = -1) = \mathbf{P}(X_j = +1) = \frac{1}{2}$$

for the **random walk** $\{Z_n\}$, given by

$$Z_0 := 0 \qquad \text{and} \qquad Z_{n+1} := Z_n + X_n \qquad (n \in \mathbb{N}),$$

that

$$\lim_{n\to\infty} \mathbf{P}\left(\frac{Z_n}{\sqrt{n}} < x\right) = \mathbf{\Phi}(x). \tag{3.2}$$

The distribution functions of the Z_n are atomic whereas their limit is absolutely continuous. Note that one can construct Brown*ian motion* as a certain limit of random walks; see [9], §VI.6.

We return to probabilistic number theory. An arithmetic function $f : \mathbb{N} \to \mathbb{C}$ may be viewed as a sequence of random variables

$$f_N = (f, \nu_{\mathbf{N}})$$

which takes the values $f(n), 1 \leq n \leq N$, with probability $\frac{1}{N}$, i.e. the uniform distribution $\nu_{\mathbf{N}}$ on the set $\{n : n \leq N\}$. The fundamental question is: *does there exist a distribution law, as $N \to \infty$?*

Therefore, we associate to an arithmetic function f for each $N \in \mathbb{N}$ the atomic distribution function

$$\mathbf{F_N}(z) := \nu_{\mathbf{N}}\{n : f(n) \leq z\} = \frac{1}{N}\sharp\{n \leq N : f(n) \leq z\}. \tag{3.3}$$

We say that f possesses a **limiting distribution function F** if the sequence $\mathbf{F_N}$, defined by (3.3), converges weakly to a limit $\mathbf{F}$, and if $\mathbf{F}$ is a distribution function. Then f is said to have a **limit law**.

An arithmetic function f is completely determined by the sequence of the associated $\mathbf{F_N}$, defined by (3.3). However, we may hope to obtain sufficiently precise knowledge on the global value distribution of f when its limiting distribution function (when it exists) can be described adequately precise.

Important for practical use is the following

Theorem 3.2 *Let f be a real-valued arithmetic function. Suppose that for any positive ε there exists a sequence $a_\varepsilon(n)$ of positive integers such that*

(i) $\lim_{\varepsilon\to 0} \limsup_{T\to\infty} \overline{\mathbf{d}}\{n : a_\varepsilon(n) > T\} = 0$,

(ii) $\lim_{\varepsilon\to 0} \overline{\mathbf{d}}\{n : |f(n) - f(a_\varepsilon(n))| > \varepsilon\} = 0$, *and*

(iii) for each $a \geq 1$ the density $\mathbf{d}\{n : a_\varepsilon(n) = a\}$ exists.

Then f has a limit law.

Before we give the proof we recall some useful notation. Related to the O-notation, we write

$$f(x) = o(g(x)),$$

when there exists a positive function $g(x)$ such that

$$\lim_{x\to\infty} \frac{|f(x)|}{g(x)} = 0.$$

In view to Exercises 1.2 and 1.3 do

Exercise 3.1 *Show for any $\varepsilon > 0$*

(i) $\log x = o(x^\varepsilon)$ *and* $x^\varepsilon = o(\exp(x))$*, as* $x \to \infty$*;*

(ii) $\frac{\varphi(n)}{n^{1+\varepsilon}} = o(1)$*, as* $n \to \infty$*.*

We return to our observations on limit laws for arithmetic functions to give the

Proof of Theorem 3.2. Let $\varepsilon = \varepsilon(\eta)$ and $T = T(\varepsilon)$ be two positive functions defined for $\eta > 0$ with

$$\lim_{\eta\to 0+} \varepsilon(\eta) = 0 \qquad \text{and} \qquad \lim_{\eta\to 0+} T(\varepsilon(\eta)) = \infty$$

such that $\overline{\mathbf{d}}\{n : a_\varepsilon(n) > T\} \le \eta$. Further, define

$$F(z,\eta) = \sum_{\substack{a\le T(\varepsilon)\\ f(a)\le z}} \mathbf{d}\{n : a_\varepsilon(n) = a\} \qquad \text{and} \qquad \mathbf{F}(z) = \limsup_{\eta\to 0} F(z,\eta).$$

With $\mathbf{F_N}$, given by (3.3), it follows in view to the conditions of the theorem that, for any $z \in \mathcal{C}(\mathbf{F})$,

$$\begin{aligned}\mathbf{F_N}(z) &\le \frac{1}{N}\sharp\{n\le N : a_\varepsilon(n)\le T(\varepsilon), f(a_\varepsilon(n))\le z+\varepsilon\}\\ &\quad +\frac{1}{N}\sharp\{n\le N : a_\varepsilon(n) > T(\varepsilon)\}\\ &\quad +\frac{1}{N}\sharp\{n\le N : |f(n)-f(a_\varepsilon(n))|>\varepsilon\}\\ &= \mathbf{F}(z+\varepsilon,\eta)+o(1),\end{aligned}$$

as $N \to \infty$; recall that the notation $o(1)$ stands for some quantity which tends with $N \to \infty$ to zero. Therefore,

$$\limsup_{N\to\infty} \mathbf{F_N}(z) \le \limsup_{\eta\to 0} F(z+\varepsilon(\eta),\eta) = \mathbf{F}(z),$$

and, analogously,

$$\liminf_{N\to\infty} \mathbf{F_N}(z) \geq \limsup_{\eta\to 0} F(z+\varepsilon(\eta),\eta) = \mathbf{F}(z);$$

here we used that $F(z,\eta)$ is non-decreasing in z, and that $z \in \mathcal{C}(\mathbf{F})$. Thus, $\mathbf{F_N}$ converges weakly to $\mathbf{F}$, and by normalization we may assume that $\mathbf{F}$ is right-continuous. Since

$$\mathbf{F}(z) = \lim_{N\to\infty} \mathbf{F_N}(z) \qquad \text{for} \qquad z \in \mathcal{C}(\mathbf{F}),$$

we have $0 \leq \mathbf{F}(z) \leq 1$. For $\varepsilon > 0$ choose $z \in \mathcal{C}(\mathbf{F})$ with $z > \max\{f(a) : a \leq T(\varepsilon)\} + \varepsilon$. Then $f(n) > z$ implies either

$$a_\varepsilon(n) > T \qquad \text{or} \qquad |f(n) - f(a_\varepsilon(n))| > \varepsilon.$$

In view to the conditions of the theorem the corresponding density $1 - \mathbf{F}(z)$ tends with $\eta \to 0+$ to zero. This gives $\mathbf{F}(+\infty) = 0$, and $\mathbf{F}(-\infty) = 0$ can be shown analogously. Thus $\mathbf{F}$ is a limiting distribution function. •

We give an application:

Theorem 3.3 *The function* $\frac{\varphi(n)}{n}$ *possesses a limiting distribution function.*

Sketch of proof. For $\varepsilon > 0$ let

$$a_\varepsilon(n) := \prod_{p|n; p\leq\varepsilon^{-2}} p^{\nu(n;p)} = n \cdot \prod_{p|n; p>\varepsilon^{-2}} p^{-\nu(n;p)}.$$

Therefore, one finds with a simple sieve-theoretical argument, for any $a \in \mathbb{N}$,

$$\begin{aligned}\sharp\left\{n \leq N \,:\, a = \prod_{p|n;\, p\leq\varepsilon^{-2}} p^{\nu(n;p)}\right\} &= \sharp\left\{n \leq N \,:\, \frac{n}{a} = \prod_{p|n;\, p>\varepsilon^{-2}} p^{\nu(n;p)}\right\} \\ &= \frac{N}{a}\left(\prod_{p\leq\varepsilon^{-2}}\left(1-\frac{1}{p}\right) + o(1)\right)\end{aligned}$$

(for details have a look on the *sieve of* ERATOSTHENES in [30], §I.4). Thus, condition (iii) of Theorem 3.2 holds. Further,

$$\left|\frac{\varphi(n)}{n} - \frac{\varphi(a_\varepsilon(n))}{a_\varepsilon(n)}\right| \leq \sum_{\substack{p|n \\ p>\varepsilon^{-2}}} \frac{1}{p},$$

which yields condition (ii). Finally,

$$\sum_{n\leq N}\log a_\varepsilon(n) \ll \log\frac{1}{\varepsilon}\cdot\sum_{n\leq x}\sum_{p|n;p\leq\varepsilon^{-2}}\nu(n;p) \ll x\left(\log\frac{1}{\varepsilon}\right)^2,$$

which implies (i). Hence, applying Theorem 3.2, yields the existence of a limiting distribution function for $\frac{\varphi(n)}{n}$. •

For more details on EULER's totient and its limit law see [17], §4.2. In Chapter 10 we will get to know a more convenient way to obtain information on the existence of a limit law and the limiting distribution itself.

Chapter 4

Expectation and variance

Now we introduce, similarly to probability theory, the **expectation** and the **variance** of an arithmetic function f with respect to the uniform distribtuion $\nu_{\mathbf{N}}$ by

$$\mathbf{E_N}(f) := \int_{-\infty}^{\infty} z \, \mathrm{d}\mathbf{F_N}(z) = \frac{1}{N} \sum_{n \le N} f(n)$$

and

$$\mathbf{V_N}(f) := \int_{-\infty}^{\infty} (z - \mathbf{E_N}(f))^2 \, \mathrm{d}\mathbf{F_N}(z) = \frac{1}{N} \sum_{n \le N} (f(n) - \mathbf{E_N}(f))^2,$$

resp., where $\mathbf{F_N}$ is defined by (3.3).

We give an example. In (1.1) we have seen that $\lim_{n\to\infty} \frac{\varphi(n)}{n}$ does not exist. Actually, if we replace $\frac{\varphi(n)}{n}$ by its expectation value $\mathbf{E_N}$, then the corresponding limit exists.

Theorem 4.1 (Mertens, 1874) *As $N \to \infty$,*

$$\sum_{n \le N} \frac{\varphi(n)}{n} = \frac{6}{\pi^2} N + O(\log N).$$

In particular,

$$\lim_{N \to \infty} \mathbf{E_N}\left(\frac{\varphi(n)}{n}\right) = \frac{6}{\pi^2} = 0.6079\ldots.$$

Moreover, we are able to give Cesaro's statement on coprime integers, mentioned in the introduction. His interpretation of the theorem above is that the *probability* that two *randomly* chosen integers are coprime equals

$$\begin{aligned}\mathbf{d}\{(a,b)\in\mathbb{N}^2 \,:\, \gcd(a,b)=1\} &= \lim_{N\to\infty}\mathbf{E_N}\left(\frac{\sharp\{a\le n \,:\, \gcd(a,n)=1\}}{\sharp\{a\le n\}}\right)\\ &= \lim_{N\to\infty}\mathbf{E_N}\left(\frac{\varphi(n)}{n}\right)=\frac{6}{\pi^2}.\end{aligned}$$

Before we give the proof of Theorem 4.1 we recall some well-known facts from number theory. The Möbius **μ-function** is defined by

$$\mu(n)=\begin{cases}(-1)^{\omega(n)} & \text{if}\quad \omega(n)=\Omega(n),\\ 0 & \text{otherwise.}\end{cases}$$

Integers n with the property $\omega(n)=\Omega(n)$ are called **squarefree**. $\mu(n)$ vanishes exactly on the complement of the squarefree numbers.

Exercise 4.1 *(i) Prove that μ is multiplicative.*

(ii) Show

$$\sum_{d|n}\mu(d)=\begin{cases}1 & \textit{if}\quad n=1,\\ 0 & \textit{else.}\end{cases}\tag{4.1}$$

(Hint: use the multiplicativity of μ.)

Proof of Theorem 4.1. Using (4.1), we find

$$\varphi(n)=\sum_{a\le n}\sum_{d|\gcd(a,n)}\mu(d)=\sum_{d|n}\mu(d)\sum_{\substack{a\le n\\ d|a}}1=\sum_{d|n}\mu(d)\frac{n}{d}.$$

This yields

$$\begin{aligned}\sum_{n\le N}\frac{\varphi(n)}{n} &= \sum_{n\le N}\sum_{d|n}\frac{\mu(d)}{d}=\sum_{d\le N}\frac{\mu(d)}{d}\left(\frac{N}{d}+O(1)\right)\\ &= N\sum_{d=1}^{\infty}\frac{\mu(d)}{d^2}+O\left(N\sum_{d>N}\frac{1}{d^2}+\sum_{d\le N}\frac{1}{d}\right).\end{aligned}\tag{4.2}$$

Again with (4.1) we get

$$\sum_{b=1}^{\infty}\frac{1}{b^2}\cdot\sum_{d=1}^{\infty}\frac{\mu(d)}{d^2}=\sum_{n=1}^{\infty}\frac{1}{n^2}\sum_{d|n}\mu(d)=1,$$

and therefore, in view to (2.7),

$$\sum_{d=1}^{\infty}\frac{\mu(d)}{d^2}=\frac{1}{\zeta(2)}.$$

It is well-known that $\zeta(2)=\frac{\pi^2}{6}$; however, we sketch in Exercise 4.2 below a simple proof of this classical result. Further, we have

$$\sum_{d>N}\frac{1}{d^2}=\int_N^{\infty}\frac{\mathrm{d}t}{t^2}+O\left(\frac{1}{N}\right)\ll\frac{1}{N},$$

as $N\to\infty$. Hence, in view to (1.2), we deduce from (4.2) the assertion. •

Exercise 4.2 (Calabi, 1993) *Show that*

$$\begin{aligned}\sum_{m=0}^{\infty}\frac{1}{(2m+1)^2} &= \sum_{m=0}^{\infty}\int_0^1\int_0^1 x^{2m}y^{2m}\,dx\,dy=\int_0^1\int_0^1\sum_{m=0}^{\infty}(xy)^{2m}\,dx\,dy\\ &= \int_0^1\int_0^1\frac{dx\,dy}{1-x^2y^2}=\frac{\pi^2}{8}\end{aligned}$$

(Hint: for the last equality use the transformation $x=\frac{\sin u}{\cos v}, y=\frac{\sin v}{\cos u}$*), and deduce* $\zeta(2)=\sum_{n=1}^{\infty}\frac{1}{n^2}=\frac{\pi^2}{6}$.

For a fixed complex number α we define the arithmetic function

$$\sigma_\alpha(n)=\sum_{d|n}d^\alpha.$$

It is easily shown that $\sigma_\alpha(n)$ is multiplicative. We write traditionally

- **divisor function**: $\tau(n)=\sigma_0(n)$;
- **sum of divisors-function**: $\sigma(n)=\sigma_1(n)$.

Exercise 4.3 *(i) Prove the identity* $\sigma_\alpha(n)=n^\alpha\sigma_{-\alpha}(n)$*;*

(ii) Show

$$\sigma_\alpha(n)=\begin{cases}\prod_{p|n}(1+\nu(n;p)) & \textit{if}\quad \alpha=0,\\ \prod_{p|n}\frac{p^{\alpha(\nu(n;p)+1)}-1}{p^\alpha-1} & \textit{otherwise};\end{cases}$$

in particular, $\sigma_\alpha(n)$ *is multiplicative.*

(iii) Prove, as $N \to \infty$*,*

$$\sum_{n \leq N} \frac{\sigma(n)}{n} = \zeta(2)N + O(\log N),$$

and deduce $\lim_{N\to\infty} \mathbf{E_N}\left(\frac{\sigma(n)}{n}\right) = \frac{\pi^2}{6}$.

(iv) What is $\lim_{N\to\infty} \mathbf{E_N}(\sigma_{-1}(n))$*?*

As we have seen above, the mean value $\frac{1}{N}\sum_{n\leq N} f(n)$ of an arithmetic function f contains interesting information on the value distribution of f. In the following chapter we will give further examples, but also draw down the limits.

Chapter 5

Average order and normal order

We say that an arithmetic function f has **average order** g if g is an arithmetic function such that

$$\lim_{N\to\infty} \frac{\sum_{n\leq N} f(n)}{\sum_{n\leq N} g(n)} = 1.$$

Obviously, the above limit can be replaced by the condition $\mathbf{E}_N(f) = \mathbf{E}_N(g)(1 + o(1))$, as $N \to \infty$.

To give a first example we consider the divisor function.

Theorem 5.1 *As $N \to \infty$,*

$$\sum_{n\leq N} \tau(n) = N \log N + O(N).$$

In particular, $\tau(n)$ has average order $\log n$.

Proof. We have

$$\sum_{n\leq N} \tau(n) = \sum_{bd\leq N} 1 = \sum_{b\leq N} \sum_{d\leq \frac{N}{b}} 1 = \sum_{b\leq N} \left(\frac{N}{b} + O(1)\right).$$

In view to (1.2) we obtain the asymptotic formula of the theorem. Further,

$$\sum_{n\leq N} \log n = \int_1^N \log u \,\mathrm{d}u + O(\log N) = N \log N + O(N),$$

which proves the statement on the average order. •

With a simple geometric idea one can improve the above result drastically.

Exercise 5.1 (Dirichlet's hyperbola method) *Prove the asymptotic formula*

$$\sum_{n\le N}\tau(n)=N\log N+(2\gamma-1)N+O(N^{\frac{1}{2}}),$$

where γ is the Euler-Mascheroni **constant**, *given by*

$$\gamma:=\lim_{N\to\infty}\left(\sum_{n=1}^{N}\frac{1}{n}-\log N\right)=1-\int_1^\infty\frac{u-[u]}{u^2}\,du=0.577\ldots.$$

(Hint: interpret the sum in question as the number of integral lattice points under the hyperbola $bd=N$ in the (b,d)-plane; the integral representation of γ follows from manipulating the defining series by partial summation.)

The situation for the prime divisor counting functions is more delicate.

Theorem 5.2 *As $N\to\infty$,*

$$\sum_{n\le N}\omega(n)=N\log\log N+O(N).$$

In particular, $\omega(n)$ has average order $\log\log n$.

For the proof we need some information on the distribution of prime numbers; the reader having a thorough knowledge of that subject can jump directly to the proof of Theorem 5.2.

Theorem 5.3 (Mertens, 1874) *As $x\to\infty$,*

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

Sketch of proof. Let $n\in\mathbb{N}$. By the formula

$$\nu(n!;p)=\sum_{k\ge1}\left[\frac{n}{p^k}\right], \tag{5.1}$$

we find

$$\log n!=\sum_{p\le n}\nu(n!;p)\log p=\sum_{p\le n}\sum_{k\ge1}\left[\frac{n}{p^k}\right]\log p=\sum_{p\le n}\left[\frac{n}{p}\right]\log p+O(n).$$

By the so-called **weak** Stirling **formula**,

$$\log n!=n\log n-n+O(\log n), \tag{5.2}$$

we obtain

$$(5.3) \qquad \sum_{p\leq n}\left[\frac{n}{p}\right]\log p = n\log n + O(n).$$

Since

$$\left[\frac{2n}{p}\right] - 2\left[\frac{n}{p}\right] = \begin{cases} 1 & \text{if} \quad n < p \leq 2n, \\ 0 & \text{if} \quad p \leq n, \end{cases}$$

we find, using formula (5.3) with n and with $2n$ instead of n,

$$\sum_{n<p\leq 2n}\log p = \sum_{p\leq 2n}\left(\left[\frac{2n}{p}\right] - 2\left[\frac{n}{p}\right]\right)\log p = 2n\log(2n) - 2\cdot n\log n + O(n) \ll n.$$

Obviously, the same estimate holds with an arbitrary real x instead of $n \in \mathbb{N}$. Furthermore,

$$(5.4) \qquad \vartheta(x) := \sum_{p\leq x}\log p = \sum_{k\geq 1}\sum_{\frac{x}{2^k}<p\leq\frac{x}{2^{k-1}}}\log p \ll x.$$

Now, removing the Gauss brackets in (5.3), gives in view to the latter estimate the assertion of the theorem. •

For the sake of completeness

Exercise 5.2 *Prove*

(i) formula (5.1);
(Hint: the p-exponent in $n!$ is $= \sum_{k\geq 1} k \sum_{\substack{m\leq n \\ \nu(m;p)=k}} 1$.)

(ii) the weak Stirling *formula (5.2).*
(Hint: express the left hand side by a sum and, up to an error term, an integral, respectively.)

As an immediate consequence of Mertens' theorem we deduce

Corollary 5.4 *As $x \to \infty$,*

$$\sum_{p\leq x}\frac{1}{p} = \log\log x + O(1).$$

In particular, the set of prime numbers has logarithmic density zero: $\delta\mathbb{P} = 0$.

Proof. According to MERTENS' theorem 5.3 let $A(x) := \sum_{p\le x} \frac{\log p}{p} = \log x + O(1)$. Then partial summation yields

$$\begin{aligned}\sum_{p\le x}\frac{1}{p} &= \sum_{p\le x}\frac{\log p}{p}\cdot\frac{1}{\log p} = \frac{A(x)}{\log x} + \int_2^x \frac{A(u)}{u(\log u)^2}\,\mathrm{d}u.\\ &= 1 + O\left(\frac{1}{\log x}\right) + \int_2^x \frac{\mathrm{d}u}{u\log u} + O\left(\int_2^x \frac{\mathrm{d}u}{u(\log u)^2}\right),\end{aligned}$$

which gives the asymptotic formula. Consequently,

$$\overline{\delta}\mathbb{P} = \limsup_{x\to\infty}\frac{1}{\log x}\sum_{p\le x}\frac{1}{p} = \lim_{x\to\infty}\frac{\log\log x}{\log x} = 0.$$

This proves the corollary. •

Now we are able to give the

Proof of Theorem 5.2. We have

$$\sum_{n\le N}\omega(n) = \sum_{n\le N}\sum_{p|n}1 = \sum_{p\le N}\left[\frac{N}{p}\right] = N\sum_{p\le N}\frac{1}{p} + O(N).$$

Application of Corollary 5.4 yields the asymptotic formula of the theorem. The statement on the normal order is an easy exercise in integration. •

Exercise 5.3 *Prove*

(i) As $N\to\infty$*,*

$$\sum_{n\le N}\Omega(n) = N\log\log N + O(N);$$

(ii) $\Omega(n)$ *has average order* $\log\log n$*.*

Arithmetic functions do not necessarily take values in the neighbourhood of their average orders. For example, a simple combinatorial argument shows that for any $n\in\mathbb{N}$

$$(5.5)\qquad 2^{\omega(n)} \le \tau(n) \le 2^{\Omega(n)}.$$

Since $\omega(n)$ and $\Omega(n)$ both have average order $\log\log n$, one might expect that $\tau(n)$ has many values of order $(\log n)^{\log 2}$ while its average order is $\log n$. It seems that the average order depends too much on *extreme* values to give deeper insights in the value distribution of an arithmetic function.

A fruitful concept in probability theory is the one of *almost sure* events. According to that we introduce now a notion which allows us to exclude extremal values from our investigations on the value distribution of arithmetic functions. We say that an arithmetic function f has **normal order** g if g is an arithmetic function such that for any positive ε the inequality

$$|f(n) - g(n)| \leq \varepsilon |g(n)|$$

holds on a set of integers $n \in \mathbb{N}$ with natural density 1; we may write equivalently

$$f(n) = (1 + o(1))g(n) \qquad \text{almost everywhere.}$$

This important notion was introduced by Hardy and Ramanujan in [13], and can be seen as the first step towards using probabilistic concepts in number theory.

In terms of distribution functions the existence of a normal order can be seen after a suitable renormalization as the convergence to a certain limit law: assuming that f, g are positive arithmetic functions, then, f has a normal order g if, and only if, the distribution functions

$$\nu_{\mathbf{N}}\{n : f(n) \leq z \cdot g(n)\} = \frac{1}{N}\sharp\{n \leq N : f(n) \leq z \cdot g(n)\}$$

converge weakly to the one-point step-function

$$\mathbf{1}_{[1,\infty]}(z) = \begin{cases} 1 & \text{if} \quad 1 \leq z, \\ 0 & \text{else.} \end{cases}$$

Therefore, normal order seems to be the right concept for studying the value distribution of arithmetic functions with probabilistic methods.

We conclude with an easy example which is related to the above observations on prime number distribution. Define the **prime counting function** by

$$\pi(x) = \sharp\{p \leq x\}.$$

Then the characteristic function on the prime numbers $\mathbf{1}_{\mathbb{P}}(n) = \pi(n) - \pi(n-1)$ has normal order 0. This follows from

Theorem 5.5 (Čebyšev, 1852) *As $x \to \infty$,*

$$\pi(x) \ll \frac{x}{\log x}.$$

In particular, the set of prime numbers has natural density zero: $\mathbf{d}\mathbb{P} = 0$.

Proof. In view to (5.4),

$$\vartheta(x) \geq \sum_{\sqrt{x}<p\leq x} \log p \geq \log\sqrt{x}(\pi(x)-\pi(\sqrt{x})),$$

and therefore

$$\pi(x) \leq \frac{2\vartheta(x)}{\log x} + \pi(\sqrt{x}) \ll \frac{x}{\log x},$$

which proves the estimate in the theorem. Consequently,

$$\overline{\mathbf{d}}\mathbb{P} = \limsup_{x\to\infty} \frac{\pi(x)}{x} \leq \lim_{x\to\infty} \frac{1}{\log x} = 0,$$

and the assertion about the natural density follows immediately. •

Actually, Čebyšev proved much more, namely that the estimate

$$0.956\ldots \leq \pi(x)\frac{\log x}{x} \leq 1.045\ldots.$$

holds for all sufficiently large x.

Exercise 5.4 *(i) Prove that, as $x \to \infty$,*

$$\pi(x) \gg \frac{x}{\log x}.$$

(Hint: consider $\sum_{\alpha x<p\leq x}\frac{\log p}{p}$ for a sufficiently small $\alpha > 0$ with regard to Mertens*' theorem.)*
Can you give explicit values for the implicit constants in the formula above as well as in the one of Theorem 5.5?

(ii) Show that, as $N \to \infty$,

$$\frac{1}{N}\sum_{n\leq N}(\Omega(n)-\omega(N)) = \sum_p \frac{1}{p(p-1)} + o(1).$$

After having been conjectured by Gauss in 1792 the celebrated *prime number theorem*,

$$\pi(x) = (1+o(1))\frac{x}{\log x}, \tag{5.6}$$

was proved independently in 1896 by Hadamard and de La Vallée-Poussin; a proof of this deep result can be found in [30], §II.4; in Chapter 16 we will give an unconvenient proof of an interesting generalization of the prime number theorem. Note that we have not used deeper knowledge on prime number distribution - i.e. Čebyšev's theorem 5.5 or even the prime number theorem - to prove the mean value results of this chapter.

Chapter 6

The TURÁN-KUBILIUS inequality

Let $\{X_j\}$ be a sequence of random variables with expectation value $\mathbf{E}X_j = \mu$ and variance $\leq M < \infty$, and let $\varepsilon > 0$. Then the *weak law of large numbers* states that

$$\mathbf{P}\left(\left|\frac{1}{n}\sum_{j=1}^{n} X_j - \mu\right| \geq \varepsilon\right) \leq \frac{M}{\varepsilon^2 n},$$

which tends with $n \to \infty$ to zero. This is a fundamental result in probability theory, justifying the *frequency concept* of probability. The weak law of lage numbers is an immediate consequence of the ČEBYŠEV *inequality*

$$\mathbf{P}(|X - \mathbf{E}X| \geq \varepsilon) \leq \frac{\sigma^2}{\varepsilon^2},$$

which holds for any random variable X with finite variance σ^2. That means, in a sense, that the best prediction for the value of a random variable is its expectation value. This idea can be extended to additive arithmetic functions.

An additive arithmetic function $f(n)$ is called **strongly additive** if $f(p^k) = f(p)$ holds for all primes p and all positive integers k. For example, $\omega(n)$ is strongly additive whereas $\Omega(n)$ is not strongly additive. If f is strongly additive, then

$$\begin{aligned} \mathbf{E_N}(f) &= \frac{1}{N}\sum_{n\leq N} f(n) = \frac{1}{N}\sum_{n\leq N}\sum_{p|n} f(p) = \frac{1}{N}\sum_{p\leq N} f(p)\left[\frac{N}{p}\right] \\ &= \sum_{p\leq N}\frac{f(p)}{p} + O\left(\frac{1}{N}\sum_{p\leq N} f(p)\right), \qquad (6.1) \end{aligned}$$

and we may expect that $f(n)$ has *many* values of order $\sum_{p\leq N}\frac{f(p)}{p}$.

However, for an analogue of the ČEBYŠEV inequality we have to define for an arithmetic function f

$$\begin{aligned}\mathcal{E}(x) &:= \mathcal{E}(x;f) := \sum_{p^k\le x}\frac{f(p^k)}{p^k}\left(1-\frac{1}{p}\right),\\ \mathcal{D}(x) &:= \mathcal{D}(x;f) := \left(\sum_{p^k\le x}\frac{|f(p^k)|^2}{p^k}\right)^{\frac{1}{2}},\end{aligned}$$

where $\mathcal{D}(x)$ is the non-negative root. These quantities can be interpreted as the expectation and the deviation of f (but may differ from the expectation $\mathbf{E_N}(f)$ and the root of the variance $\mathbf{V_N}(f)$ of our probabilistic model defined in Chapter 4).

Exercise 6.1 *(i) Let f be a strongly additive function. Show that*

$$\mathcal{E}(N;f) = \sum_{p\le N}\frac{f(p)}{p} + O\left(\sum_{p\le N}\frac{f(p)}{p^{[\frac{\log N}{\log p}]}}\right).$$

(This should be compared with (6.1).)

(ii) Show that $\log\frac{\varphi(n)}{n}$ is strongly additive. Do the limits $\lim_{N\to\infty}\mathbf{E_N}(\log\frac{\varphi(n)}{n})$ and $\lim_{N\to\infty}\mathcal{E}(N;\log\frac{\varphi(n)}{n})$ exist?

The following theorem gives an estimate for the difference of the values of $f(n), 1\le n\le x$, from its *expectational value* $\mathcal{E}(x;f)$ in terms of its *deviation* $\mathcal{D}(x;f)$.

Theorem 6.1 (TURÁN-KUBILIUS inequality, 1955) *There exists a function $\varepsilon(x)$ with $\lim_{x\to\infty}\varepsilon(x)=0$ such that the estimate*

$$\frac{1}{x}\sum_{n\le x}|f(n)-\mathcal{E}(x)|^2 \le (2+\varepsilon(x))\mathcal{D}(x)^2 \tag{6.2}$$

holds uniformly for all additive arithmetic functions f and real $x\ge 2$.

Proof. In the sequel we denote by q always a prime number. We define

$$\varepsilon(x) = \frac{4}{x}\left(\sum_{\substack{p^kq^l\le x\\ p\ne q}} p^kq^l\right)^{\frac{1}{2}} + \frac{4}{x}\left(\sum_{\substack{p^k\le x\\ q^l\le x}} p^{-k}q^l\right)^{\frac{1}{2}}. \tag{6.3}$$

By Corollary 5.4,

$$\sum_{p^k\leq x} p^{-k} = \sum_{p\leq x}\frac{1}{p} + \sum_{\substack{p^k\leq x\\ k\geq 2}}\frac{1}{p^k} = \log\log x + O(1), \tag{6.4}$$

and, by ČEBYŠEV's theorem 5.5,

$$\sum_{q^l\leq x} q^l = \sum_{q\leq x}\sum_{l\leq\frac{\log x}{\log q}} q^l \ll \sum_{q\leq x} q^{\frac{\log x}{\log q}} = x\pi(x) \ll \frac{x^2}{\log x}.$$

This yields

$$\sum_{\substack{p^k\leq x\\ q^l\leq x}} p^{-k}q^l \ll x^2\frac{\log\log x}{\log x}.$$

Further,

$$\begin{aligned}\sum_{\substack{p^kq^l\leq x\\ p\neq q}} p^kq^l &= 2\sum_{\substack{p^kq^l\leq x\\ p>q}} p^kq^l \ll \sum_{p^k\leq x} p^k \sum_{p<q\leq\frac{x}{p^k}}\sum_{l\leq\frac{\log(x/p^k)}{\log q}} q^l \ll \sum_{p^k\leq x} p^k \sum_{p<q\leq\frac{x}{p^k}}\frac{x}{p^k}\\ &\ll x\sum_{p^k\leq x}\pi\left(\frac{x}{p^k}\right),\end{aligned}$$

which is, by ČEBYŠEV's theorem and (6.4),

$$\ll \frac{x^2}{\log x}\sum_{p^k\leq x} p^{-k} \ll x^2\frac{\log\log x}{\log x}.$$

This gives in (6.3) the upper bound

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

which tends to zero as $x\to\infty$.

Without loss of generality we may assume that $x\in\mathbb{N}$.

First, assume that f is real and non-negative. Then

$$\frac{1}{x}\sum_{n\leq x}(f(n)-\mathcal{E}(x))^2 = \frac{1}{x}\sum_{n\leq x} f(n)^2 - 2\frac{\mathcal{E}(x)}{x}\sum_{n\leq x} f(n) + \mathcal{E}(x)^2. \tag{6.5}$$

We have, by the additivity of f,

$$\begin{aligned}\frac{1}{x}\sum_{n\le x} f(n)^2 &= \frac{1}{x}\sum_{n\le x}\sum_{p|n,q|n} f(p^{\nu(n;p)})f(q^{\nu(n;q)})\\ &= \frac{1}{x}\sum_{p^k\le x} f(p^k)^2 \sum_{\substack{n\le x\\ \nu(n;p)=k}} 1+\frac{1}{x}\sum_{\substack{p^kq^l\le x\\ p\neq q}} f(p^k)f(q^l) \sum_{\substack{n\le x\\ \nu(n;p)=k,\nu(n;q)=l}} 1.\end{aligned}$$

The first inner sum does not exceed $\frac{x}{p^k}$ while the second inner sum is, by the *inclusion-exclusion principle*,

$$\begin{aligned}&\sharp\{n\le x\,:\,\nu(n;p)=k,\nu(n;q)=l\}\\ &= \left[\frac{x}{p^kq^l}\right]-\left[\frac{x}{p^{k+1}q^l}\right]-\left[\frac{x}{p^kq^{l+1}}\right]+\left[\frac{x}{p^{k+1}q^{l+1}}\right]\\ &\le \frac{x}{p^kq^l}\left(1-\frac{1}{p}\right)\left(1-\frac{1}{q}\right)+2.\end{aligned}$$

Thus

$$\frac{1}{x}\sum_{n\le x} f(n)^2\le \mathcal{D}(x)^2+\mathcal{E}(x)^2+\frac{2}{x}\sum_{\substack{p^kq^l\le x\\ p\neq q}} f(p^k)f(q^l). \tag{6.6}$$

Furthermore, we find

$$\frac{1}{x}\sum_{n\le x} f(n)=\frac{1}{x}\sum_{n\le x}\sum_{p|n} f(p^{\nu(n;p)})=\frac{1}{x}\sum_{p^k\le x} f(p^k)\sum_{\substack{n\le x\\ \nu(n;p)=k}} 1.$$

The inner sum is bounded below by

$$\sharp\{n\le x\,:\,\nu(n;p)=k\}=\left[\frac{x}{p^k}\right]-\left[\frac{x}{p^{k+1}}\right]\ge\frac{x}{p^k}\left(1-\frac{1}{p}\right)-1,$$

from which we deduce that

$$\frac{1}{x}\sum_{n\le x} f(n)\ge\mathcal{E}(x)-\frac{1}{x}\sum_{p^k\le x} f(p^k).$$

This and (6.6) give in (6.5)

$$\frac{1}{x}\sum_{n\le x}(f(n)-\mathcal{E}(x))^2\le\mathcal{D}(x)^2+\frac{2}{x}\sum_{\substack{p^kq^l\le x\\ p\neq q}} f(p^k)f(q^l)+2\frac{\mathcal{E}(x)}{x}\sum_{p^k\le x} f(p^k). \tag{6.7}$$

Note that the quadratic term $\mathcal{E}(x)^2$ is cancelled. By the Cauchy-Schwarz inequality, we obtain

$$\begin{aligned}
\mathcal{E}(x) &\leq \sum_{p^k\leq x} \frac{f(p^k)}{p^{\frac{k}{2}}}\cdot p^{\frac{k}{2}} \leq \mathcal{D}(x)\left(\sum_{p^k\leq x} p^{-k}\right)^{\frac{1}{2}}, \\
\sum_{p^k\leq x} f(p^k) &= \sum_{p^k\leq x} \frac{f(p^k)}{p^{\frac{k}{2}}}\cdot p^{-\frac{k}{2}} \leq \mathcal{D}(x)\left(\sum_{p^k\leq x} p^{k}\right)^{\frac{1}{2}}, \\
\sum_{\substack{p^kq^l\leq x\\ p\neq q}} f(p^k)f(q^l) &= \sum_{\substack{p^kq^l\leq x\\ p\neq q}} \frac{f(p^k)}{p^{\frac{k}{2}}}\frac{f(q^l)}{q^{\frac{l}{2}}}\cdot p^{\frac{k}{2}}q^{\frac{l}{2}} \leq \mathcal{D}(x)^2\left(\sum_{\substack{p^kq^l\leq x\\ p\neq q}} p^kq^l\right)^{\frac{1}{2}}.
\end{aligned}$$

This gives in (6.7)

$$\frac{1}{x}\sum_{n\leq x}(f(n)-\mathcal{E}(x))^2 \leq \left(1+\frac{1}{2}\varepsilon(x)\right)\mathcal{D}(x)^2, \tag{6.8}$$

which is even stronger than the estimate (6.2) in the theorem (by a factor of 2).

Now assume that f is real-valued but takes values of both signs. Then we introduce the functions $f^{\pm}$ defined by

$$f^{\pm}(p^k) = \max\{\pm f(p^k), 0\}.$$

Obviously, $f = f^+ - f^-$. Since f^+f^- vanishes identically, we have $f^2 = (f^+)^2 + (f^-)^2$, and we obtain for $1 \leq n \leq x$

$$\begin{aligned}
\mathcal{D}(x;f)^2 &= \mathcal{D}(x;f^+)^2 + \mathcal{D}(x;f^-)^2, \\
(f(n)-\mathcal{E}(x;f))^2 &= (f^+(n)-\mathcal{E}(x;f^+)-(f^-(n)-\mathcal{E}(x;f^-))) \\
&\leq 2(f^+(n)-\mathcal{E}(x;f^+))^2 + 2(f^-(n)-\mathcal{E}(x;f^-))^2.
\end{aligned}$$

Thus, an application of the previous estimate (6.8) gives (6.2).

Finally, when f is complex-valued, then an application of the above estimate to the real part and the imaginary parts of f seperately yield (6.2). The theorem is proved. •

In 1983 Kubilius [19] showed that the constant 2 in the Turán-Kubilius inequality can be replaced by $\frac{3}{2}+o(1)$, and also that this is optimal. On the other side, the corresponding probabilistic model gives an upper estimate with the constant 1, which shows not only the similarity but also the discrepancy between probabilistic number theory and probability theory; for details see [30], §III.4.

Exercise 6.2 *Deduce from the* Turán-Kubilius *inequality, for sufficiently large x, the estimate*

$$\frac{1}{x}\sum_{n\le x}|f(n)-\mathcal{A}(x)|^2 \le 6\mathcal{D}(x)^2\ , \qquad \textit{where} \qquad \mathcal{A}(x) := \sum_{p^k\le x}\frac{f(p^k)}{p^k}.$$

(Hint: use the Cauchy-Schwarz *inequality.)*

In the following chapter we shall derive from the Turán-Kubilius inequality the celebrated Hardy-Ramanujan result (1.3) mentioned in the introduction.

Chapter 7

The theorem of Hardy-Ramanujan

The expectation of an arithmetic function f is a good candidate for a normal order of f. The Turán-Kubilius inequality gives a sufficient condition for $f(n)$ to have normal order $\mathcal{E}(n;f) \approx \mathbf{E_n}(f)$.

Theorem 7.1 *Let f be an additive arithmetic function. If*

$$\mathcal{D}(N) = o(\mathcal{E}(N)),$$

as $N \to \infty$, then $\mathcal{E}(n)$ is a normal order for $f(n)$.

Proof. Using the Cauchy-Schwarz inequality, we obtain for $\sqrt{N} < n \leq N$

$$|\mathcal{E}(N) - \mathcal{E}(n)| = \left| \sum_{n < p^k \leq N} \frac{f(p^k)}{p^k} \left(1 - \frac{1}{p}\right) \right| \ll \left(\sum_{\sqrt{N} < p \leq N} p^{-k} \sum_{p^k \leq N} \frac{|f(p^k)|^2}{p^k} \right)^{\frac{1}{2}}.$$

Using Lemma 5.4, we find

$$\sum_{\sqrt{N} < p \leq N} \frac{1}{p^k} = \sum_{\sqrt{N} < p \leq N} \frac{1}{p} + O(1) = \log\log N - \log\log\sqrt{N} + O(1) \ll 1,$$

which gives above $|\mathcal{E}(N) - \mathcal{E}(n)| \ll \mathcal{D}(N)$. Since the right hand side is under the assumption of the theorem $= o(\mathcal{E}(N))$, it follows that $\mathcal{E}(n) = \mathcal{E}(N)(1 + o(1))$ for all $n \leq N$ except at most $o(N)$. To prove the assertion of the theorem we may use the Turán-Kubilius inequality to estimate, for any $\varepsilon > 0$,

$$\nu_{\mathbf{N}}\{n \, : \, |f(n) - \mathcal{E}(N)| > \varepsilon |\mathcal{E}(N)|\} < \frac{1}{N} \sum_{n \leq N} \left| \frac{f(n) - \mathcal{E}(N)}{\varepsilon \mathcal{E}(N)} \right|^2 \ll \left| \frac{\mathcal{D}(N)}{\varepsilon \mathcal{E}(N)} \right|^2,$$

which is $= o(1)$ by assumption. The theorem is proved. •

Now we apply our results to the prime divisor counting function $\omega(n)$. In view to Theorem 5.2 and Corollary 5.4 (resp. Exercise 6.1):

$$\mathcal{E}(N;\omega) = \log\log N + O(1) \quad \text{and} \quad \mathcal{D}(N;\omega)^2 = \log\log N + O(1).$$

The TURÁN-KUBILIUS inequality yields TURÁN's estimate (1.4): since $\omega(n)$ is non-negative, we may use (6.8) to obtain

$$\frac{1}{N}\sum_{n\leq N}(\omega(n) - \log\log N)^2 \leq \log\log N + O(1).$$

Further, Theorem 7.1 gives the normal order of $\omega(n)$, and we obtain immediately the following improvement of (1.3):

Theorem 7.2 (HARDY+RAMANUJAN, 1917; TURÁN, 1934) *For any $\xi(N) \to \infty$,*

$$\nu_{\mathbf{N}}\{n\,:\,|\omega(n) - \log\log N| > \xi(N)\sqrt{\log\log N}\} \ll \xi(N)^{-2},$$

and

$$\mathbf{d}\{n\,:\,|\omega(n) - \log\log n| > \xi(N)\sqrt{\log\log N}\} = 0.$$

In particular, $\log\log n$ *is a normal order of* $\omega(n)$.

It is easy to do the same for $\Omega(n)$.

Exercise 7.1 *(i) Show that, for any $\xi(N) \to \infty$,*

$$\nu_{\mathbf{N}}\{n\,:\,|\Omega(n) - \log\log N| > \xi(N)\sqrt{\log\log N}\} \ll \xi(N)^{-2},$$

and deduce that $\Omega(n)$ has normal order $\log\log n$*;*

(ii) calculate $\mathbf{E_N}(\Omega)$ and $\mathbf{V_N}(\Omega)$, and compare these values with $\mathcal{E}(N;\Omega)$ and $\mathcal{D}(N;\Omega)^2$.

We continue our discussion on the value distribution of the divisor function started in Chapter 5. In view to (5.5) we get as an immediate consequence of the HARDY-RAMANUJAN results on $\omega(n)$ and $\Omega(n)$

Corollary 7.3 *We have*

$$\tau(n) = (\log n)^{\log 2 + o(1)} \qquad \textit{almost everywhere.}$$

In particular, $\log\tau(n)$ *has normal order* $\log 2\cdot\log\log n$.

That means that the divisor function $\tau(n)$ has a normal order different to its average order $\log n$. This is caused by some *extraordinary large* values of $\tau(n)$.

We say that an arithmetic function f has **maximal order** g if g is a positive non-decreasing arithmetic function such that

$$\limsup_{n\to\infty}\frac{f(n)}{g(n)}=1,$$

and we say that f has **minimal order** g if g is a positive non-decreasing arithmetic function such that

$$\liminf_{n\to\infty}\frac{f(n)}{g(n)}=0.$$

In (1.1) we have seen that the identity $n\mapsto n$ is both a minimal and a maximal order for EULER's φ-function.

Theorem 7.4 *A maximal order for* $\log\tau(n)$ *is* $\frac{\log 2\cdot\log n}{\log\log n}$.

Proof. By the multiplicativity of $\tau(n)$ (see Exercise 4.3),

$$\begin{aligned}\tau(n) &= \prod_{p|n}(1+\nu(n;p))\leq\prod_{\substack{p\leq x\\ p|n}}(1+\nu(n;p))\prod_{\substack{p>x\\ p|n}}2^{\nu(n;p)}\\ &\leq \left(1+\frac{\log n}{\log 2}\right)^x\left(\prod_{p|n}p^{\nu(n;p)}\right)^{\frac{\log 2}{\log x}}\\ &\leq \exp\left(x(2+\log\log n)+\frac{\log 2\cdot\log n}{\log x}\right).\end{aligned}$$

The choice $x=\frac{\log n}{(\log\log n)^3}$ yields

$$\tau(n)\leq\exp\left(\frac{\log 2\cdot\log n}{\log\log n}\left(1+O\left(\frac{\log\log\log n}{\log\log n}\right)\right)\right).$$

This shows that

$$\limsup_{n\to\infty}\log\tau(n)\cdot\frac{\log\log n}{\log 2\cdot\log n}\leq 1.$$

In order to prove that the above lim sup is also ≥ 1 we have a look on integers with *many* prime divisors. Denote by p_j the jth prime number (ordered with respect to their absolute value), and define $n_k=\prod_{j=1}^k p_j$ for $k\in\mathbb{N}$. Then $\tau(n_k)=2^k$, and

$$\log n_k=\sum_{j=1}^k\log p\leq k\log p_k.$$

Since by Exercise 5.4

$$p_k \ll \vartheta(p_k) = \sum_{j=1}^{k} \log p_j = \log n_k,$$

where the implicit constant does not depend on k, we obtain

$$\log \tau(n_k) = k \cdot \log 2 \geq \frac{\log 2 \cdot \log n_k}{\log p_k} \geq \frac{\log 2 \cdot \log n_k}{\log\log n_k}\left(1 + O\left(\frac{1}{\log\log n_k}\right)\right).$$

This shows the theorem. •

Via (5.5) Theorem 7.4 has also an effect on the prime divisor counting functions (answering one question posed in Exercise 1.2):

Exercise 7.2 *Show that*

(i) $\omega(n)$ *has maximal order* $\frac{\log n}{\log\log n}$*;*

(ii) $\Omega(n)$ *has maximal order* $\frac{\log n}{\log 2}$.

The value distribution of the divisor function is ruled by the *arcsine law.*

Theorem 7.5 (Deshouillers+Dress+Tenenbaum, 1979) *Uniformly for* $x \geq 2, 0 \leq z \leq 1$*,*

$$\frac{1}{x}\sum_{n\leq x}\frac{1}{\tau(n)}\sum_{\substack{d|n\\ d\leq n^z}} 1 = \frac{2}{\pi}\arcsin\sqrt{z} + O\left((\log x)^{-\frac{1}{2}}\right).$$

Rewriting the asymptotic formula of the theorem, we have

$$\frac{1}{x}\sum_{n\leq x}\nu_{\mathbf{n}}\{d|n\,:\, d \leq n^z\} = \frac{2}{\pi}\arcsin\sqrt{z} + O\left((\log x)^{-\frac{1}{2}}\right).$$

This shows that, on average, an integer has *many small* (resp., *many large*) divisors! This can be proved by the Selberg-Delange method, which we shall derive in Chapter 15. Nevertheless, the proof is beyond the scope of this course; the interested reader can find a detailed proof of this result in [30], §II.5, II.6.

Chapter 8

A duality principle

The TURÁN-KUBILIUS inequality has an interesting dual variant.

Theorem 8.1 (ELLIOTT, 1979) *The inequality*

$$\sum_{p^k\le N} p^k \left| \sum_{\substack{n\le N\\ k=\nu(n;p)}} x_n - \frac{1}{p^k}\left(1-\frac{1}{p}\right)\sum_{n\le N} x_n \right|^2 \le (2+o(1))N \sum_{n\le N} |x_n|^2$$

holds uniformly for all N, and complex numbers $x_n, 1\le n\le N$.

This theorem has several nice consequences as, for example, in the theory of quadratic residues. Let p be an odd prime, and assume that $a\in\mathbb{Z}$ is not divisible by p. Then we say that a is a **quadratic residue** $\bmod p$, if the congruence $X^2\equiv a \bmod p$ is soluble; otherwise, a is called **quadratic non-residue**. ELLIOTT proved for the least pair of consecutive quadratic non-residues $\bmod p$, the upper bound

$$\ll p^{\frac{1}{4}\left(1-\frac{1}{2}\exp(-10)\right)+\varepsilon},$$

where $p\ge 5$, and the implicit constant does not depend on p. For details and much more on dual versions of the TURÁN-KUBILIUS inequality, as for example their appearance in the theory of the *large sieve*, see [6], §4.

For the proof of Theorem 8.1 we will make use of

Lemma 8.2 (Duality principle) *Let (c_{nr}) be an $N\times R$ matrix with complex entries, and let C be an arbitrary positive constant. Then the following three inequalities are equivalent:*

(i) for all $x_n \in \mathbb{C}$,

$$\sum_r \left|\sum_n c_{nr} x_n\right|^2 \leq C \sum_n |x_n|^2;$$

(ii) for all $x_n, y_r \in \mathbb{C}$,

$$\left|\sum_{n,r} c_{nr} x_n y_r\right|^2 \leq C \sum_n |x_n|^2 \sum_r |y_r|^2;$$

(iii) for all $y_r \in \mathbb{C}$,

$$\sum_n \left|\sum_r c_{nr} y_r\right|^2 \leq C \sum_r |y_r|^2.$$

Proof. It suffices to show the equivalence of (i) and (ii) (since the equivalence of (ii) and (iii) follows by exchanging the indices r and n).

First, assume that (i) holds. Then, by the CAUCHY-SCHWARZ inequality,

$$\begin{aligned}\left|\sum_{n,r} c_{nr} x_n y_r\right|^2 &= \left|\sum_r y_r \sum_n c_{nr} x_n\right|^2 \leq \sum_r |y_r|^2 \sum_r \left|\sum_n c_{nr} x_n\right|^2 \\ &\leq C \sum_n |x_n|^2 \sum_r |y_r|^2.\end{aligned}$$

For the converse implication assume that (ii) holds. Let $L_r := \sum_n c_{nr} x_n$ for $r \leq R$. Then, applying (ii) with $y_r = \overline{L_r}$, yields

$$\left(\sum_r |L_r|^2\right)^2 \leq C \sum_n |x_n|^2 \sum_r |L_r|^2,$$

which implies (i). The lemma is proved. •

Proof of Theorem 8.1. Let f be an arbitrary additive function. For $n \in \mathbb{N}$ put

$$r := p^k\ , \quad y_r := \frac{f(r)}{r^{\frac{1}{2}}}\ , \qquad \text{and} \qquad c_{nr} := \begin{cases} r^{\frac{1}{2}} - r^{-\frac{1}{2}}\left(1 - \frac{1}{p}\right) & \text{if} \quad \nu(n;p) = k, \\ -r^{-\frac{1}{2}}\left(1 - \frac{1}{p}\right) & \text{otherwise.} \end{cases}$$

This gives

$$f(n) - \mathcal{E}(N) = \sum_{\substack{r|n \\ n \not\equiv 0 \bmod pr}} f(r) - \sum_{r \leq N} \frac{f(r)}{r}\left(1 - \frac{1}{p}\right) = \sum_{r \leq N} c_{nr} y_r.$$

Thus, we can rewrite the TURÁN-KUBILIUS inequality (6.2) as

$$\sum_{n\le N}\left|\sum_{r\le N} c_{nr}y_r\right|^2 \le (2+o(1))N\sum_{r\le N}|y_r|^2.$$

Since the y_r are arbitrary complex numbers, application of Lemma 8.2 shows that the inequality

$$\sum_{r\le N}\left|\sum_{n\le N} c_{nr}x_n\right|^2 \le (2+o(1))N\sum_{n\le N}|x_n|^2$$

holds for arbitrary complex numbers x_n. In view to the definition of the c_{nr} the assertion of the theorem follows. •

We conclude with an interesting interpretation of the dual form of the TURÁN-KUBILIUS inequality: since

$$\sum_{p^k\le N} p^k \approx \frac{N^2}{\log N} \qquad \text{and} \qquad \frac{1}{p^k}\left(1-\frac{1}{p}\right) \approx \nu_{\mathbf{N}}\{n\,:\,\nu(n;p)=k\},$$

we may deduce that every *sufficiently dense* sequence of integers x_n is *well distributed* among the residue classes $n\equiv 0 \bmod p^k$.

For deeper knowledge on the value distribution of arithmetic functions we have to recall some facts from the beginnings of *analytic* number theory. The reader who is familiar with these fundamentals may jump to Chapter 10.

Chapter 9

Dirichlet series and Euler products

In probability theory many information on random variables can be derived by studying their *generating* functions. The same concept applies to number theory as well (and has even its origins there).

We write $s = \sigma + it$ with $\sigma, t \in \mathbb{R}$ and $i := \sqrt{-1}$, and associate to every arithmetic function $f : \mathbb{N} \to \mathbb{C}$ its Dirichlet **series**

$$\sum_{n=1}^{\infty} \frac{f(n)}{n^s};$$

here n^s is defined by $n^s = \exp(s \cdot \log n)$. The prototype of such series is the Riemann zeta-function (2.7). First, we consider these series only as *formal* objects. With the usual addition and multiplication of series the set of Dirichlet series form a commutative ring isomorphic to the ring of arithmetic functions $\mathcal{R}$, where the multiplication is the **convolution**

$$(f * g)(n) := \sum_{d|n} f(d) g\left(\frac{n}{d}\right),$$

and where the addition is given by superposition.

Exercise 9.1 *(i) Prove the identities*

$$\sum_{n=1}^{\infty} \frac{\mu(n)}{n^s} = \frac{1}{\zeta(s)}, \quad \sum_{n=1}^{\infty} \frac{\varphi(n)}{n^s} = \frac{\zeta(s-1)}{\zeta(s)} \quad \textit{and} \quad \sum_{n=1}^{\infty} \frac{\sigma_\alpha(n)}{n^s} = \zeta(s)\zeta(s-\alpha);$$

(ii) *verify that the set of arithmetic functions $\mathcal{R}$ is a commutative ring with (multiplicative) identity*

$$\eta(n) := \begin{cases} 1 & \textit{if} \quad n = 1, \\ 0 & \textit{if} \quad n \neq 1; \end{cases}$$

(iii) *show that an arithmetic function f is a unit in the ring $\mathcal{R}$ if and only if $f(1) \neq 0$;*

(iv) *let $\varepsilon(n) := 1, n \in \mathbb{N}$, and prove for f and $F := f * \varepsilon \in \mathcal{R}$ the* MÖBIUS **inversion formula***: $f = F * \mu$.*

In the case of DIRICHLET series with multiplicative coefficients we obtain a product representation, the so-called EULER **product**.

Lemma 9.1 *Assume that $\sum_{n=1}^{\infty} |f(n)| < \infty$. If $f(n)$ is a multiplicative arithmetic function, then*

$$\sum_{n=1}^{\infty} f(n) = \prod_{p} (1 + f(p) + f(p^2) + \ldots),$$

and if f is completely multiplicative, then

$$\sum_{n=1}^{\infty} f(n) = \prod_{p} \frac{1}{1 - f(p)}.$$

The well-known formula (2.7) is here the standard example. We may extend this for $z \in \mathbb{C}, z \neq 0$, and $\sigma > 1$, which leads to

$$\zeta(s)^z = \prod_{p} \left(1 - \frac{1}{p^s}\right)^{-z} = \sum_{n=1}^{\infty} \frac{\tau_z(n)}{n^s}, \tag{9.1}$$

where $\tau_z(n)$ is the multiplicative function given by $\tau_z(1) = 1$ and

$$\tau_z\left(p^k\right) = \binom{z + k - 1}{k} := \frac{1}{k!} \prod_{j=1}^{k} (z + k - j);$$

this is an immediate consequence of the binomial series expansion in the factors of the EULER product.

In view to later applications we introduce two more EULER products. Let $z \in \mathbb{C}$ with $0 < |z| \leq 1$. Since $\omega(n)$ is additive, the arithmetic function $\frac{z^{\omega(n)}}{n^s}$ is multiplicative, and therefore a simple calculation shows

$$L(s, z, \omega) := \sum_{n=1}^{\infty} \frac{z^{\omega(n)}}{n^s} = \prod_{p} \left(1 + \sum_{k=1}^{\infty} \frac{z}{p^s}\right) = \prod_{p} \left(1 + \frac{z}{p^s - 1}\right), \tag{9.2}$$

where all series and product representations are valid in the half plane $\sigma > 1$ (we shall return to the question of convergence later on).

Exercise 9.2 *Let $z \in \mathbb{C}$ with $0 < |z| \leq 1$. Prove, for $\sigma > 1$,*

$$L(s,z,\Omega) := \sum_{n=1}^{\infty} \frac{z^{\Omega(n)}}{n^s} = \prod_p \left(1 - \frac{z}{p^s}\right)^{-1}. \tag{9.3}$$

Proof of Lemma 9.1. By the multiplicativity of $f(n)$ and the unique prime factorization of the integers,

$$\prod_{p\leq x} (1 + f(p) + f(p^2) + \ldots) = \sum_{\substack{n \\ p|n \Rightarrow p \leq x}} f(n).$$

Since

$$\left| \sum_{n=1}^{\infty} f(n) - \sum_{\substack{n \\ p|n \Longrightarrow p \leq x}} f(n) \right| \leq \sum_{n>x} |f(n)|,$$

the convergence of $\sum_{n=1}^{\infty} |f(n)|$ implies the first assertion; the second follows in view to $f(p^k) = f(p)^k$ and application of the formula for the geometric series. •

We can obtain new insights on the value distribution of an arithmetic function by studying the associated Dirichlet series as an *analytic* function. Since

$$|n^s| = |n^\sigma \exp(it \log n)| = n^\sigma,$$

Dirichlet series converge in half planes; it is possible that this half plane is empty, or that it is the whole complex plane.

Theorem 9.2 *Suppose that the series $\sum_{n=1}^{\infty} \frac{f(n)}{n^c}$ converges for some $c \in \mathbb{R}$. Then the* Dirichlet *series*

$$F(s) := \sum_{n=1}^{\infty} \frac{f(n)}{n^s}$$

converges for any $\delta > 0$ uniformly in

$$\mathcal{H}_\delta := \left\{ s \in \mathbb{C} \, : \, |\arg(s-c)| \leq \frac{\pi}{2} - \delta \right\}.$$

In particular, the function $F(s)$ is analytic in the half plane $\sigma > c$.

Proof. Let $s \in \mathcal{H}_\delta$. Partial summation shows, for $0 \leq M < N$,

$$\begin{aligned} \sum_{M<n\leq N} \frac{f(n)}{n^s} &= \sum_{M<n\leq N} \frac{f(n)}{n^c \cdot n^{s-c}} \\ &= N^{c-s} \sum_{M<n\leq N} \frac{f(n)}{n^c} + (s-c) \int_M^N \sum_{M<n\leq x} \frac{f(n)}{n^c} \frac{\mathrm{d}x}{x^{s+1-c}}. \end{aligned}$$

By the convergence of $\sum_{n=1}^{\infty}\frac{f(n)}{n^c}$, there exists for any $\varepsilon > 0$ an index M_0 such that

$$\left|\sum_{M<n\leq N}\frac{f(n)}{n^c}\right| < \varepsilon \qquad \text{for all} \qquad M \geq M_0.$$

Hence, for those M,

$$\begin{aligned}\sum_{M<n\leq N}\frac{f(n)}{n^s} &\ll \varepsilon\left(N^{c-\sigma}+|s-c|\int_M^N x^{c-\sigma-1}\,\mathrm{d}x\right)\\ &\ll \varepsilon\left(N^{c-\sigma}+\frac{|s-c|}{\sigma-c}M^{c-\sigma}\right)\ll\varepsilon\left(1+\frac{1}{\sin\delta}\right),\end{aligned}$$

since $|s-c| < (\sigma-c)\sin\delta$. This proves the uniform convergence (by fixed δ). WEIERSTRASS' theorem states that the limit $F(s)$ of the uniform convergent sequence of analytic functions $\sum_{n\leq M}\frac{f(n)}{n^s}$, as $M\to\infty$, is analytic itself (see [21], §V.1). This proves the theorem. •

Exercise 9.3 *Assume that the series $\sum_{n=1}^{\infty}\frac{f(n)}{n^s}$ converges exactly in the (non-empty) half plane $\sigma > c$. Show that the series converges absolutely for $\sigma > c+1$.*

The proof of Theorem 9.2 yields, in the region of absolute convergence,

$$\sum_{n=1}^{\infty}\frac{f(n)}{n^s} = s\int_1^{\infty}\sum_{n\leq x}f(n)\frac{\mathrm{d}x}{x^{s+1}} \tag{9.4}$$

(this should be compared with Exercise 2.4); here and in the sequel we write $\int^{\infty}$ for $\lim_{T\to\infty}\int^T$ when the limit exists. We are interested in an inversion, i.e. a formula where the transform $\sum_{n\leq x}f(n)$ is expressed by an integral over the associated DIRICHLET series $\sum_{n=1}^{\infty}\frac{f(n)}{n^s}$.

Lemma 9.3 *Let c and y be positive and real. Then*

$$\frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty}\frac{y^s}{s}\,ds = \begin{cases} 0 & \text{if} \quad 0<y<1,\\ \frac{1}{2} & \text{if} \quad y=1,\\ 1 & \text{if} \quad y>1.\end{cases}$$

Proof. First, if $y=1$, then the integral in question equals

$$\frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{\mathrm{d}t}{c+it}=\frac{1}{\pi}\lim_{T\to\infty}\int_0^T\frac{c}{c^2+t^2}\,\mathrm{d}t=\frac{1}{\pi}\lim_{T\to\infty}\arctan\frac{T}{c}=\frac{1}{2},$$

by well-known properties of the arctan-function.

Secondly, assume that $0 < y < 1$ and $r > c$. Since the integrand is analytic in $\sigma > 0$, CAUCHY's theorem implies, for $T > 0$,

$$\int_{c-iT}^{c+iT} \frac{y^s}{s}\,\mathrm{d}s = \left\{ \int_{c-iT}^{r-iT} + \int_{r-iT}^{r+iT} + \int_{r+iT}^{c+iT} \right\} \frac{y^s}{s}\,\mathrm{d}s.$$

It is easily be shown that

$$\begin{aligned} \int_{r\pm iT}^{c\pm iT} \frac{y^s}{s}\,\mathrm{d}s &\ll \frac{1}{T}\int_r^c y^\sigma\,\mathrm{d}\sigma \ll \frac{y^c}{T|\log y|}, \\ \int_{r-iT}^{r+iT} \frac{y^s}{s}\,\mathrm{d}s &\ll \frac{y^r}{r} + y^r \int_1^T \frac{\mathrm{d}t}{t} \ll y^r\left(\frac{1}{r} + \log T\right). \end{aligned}$$

Sending now r and then T to infinity, the first case follows.

Finally, if $y > 1$, then we bound the corresponding integrals over the rectangular contour with corners $c \pm iT$, $-r \pm iT$, analogously. Now the pole of the integrand at $s = 0$ with residue

$$\mathrm{Res}_{\,s=0}\frac{y^s}{s} = \lim_{s\to 0} s \cdot \frac{y^s}{s} = 1$$

gives via the calculus of residues $2\pi i$ as the value for the integral in this case. •

Exercise 9.4 *Prove*

(i) for $\alpha \in \mathbb{R}$,

$$\int_{-\infty}^{\infty} \frac{\exp(i\alpha u) - \exp(-i\alpha u)}{iu}\,du = sgn\,(\alpha)2\pi,$$

where $sgn\,(\alpha) = 0$ if $\alpha = 0$, and $= \frac{\alpha}{|\alpha|}$ otherwise;
(Hint: shift the path of integration into the right half plane by use of CAUCHY*'s theorem, and apply Lemma 9.3.)*

(ii) for $\alpha > 0$,

$$\int_{-\infty}^{\infty} \left(\frac{\sin \alpha u}{\alpha u}\right)^2 du = \frac{\pi}{\alpha}.$$

(Hint: partial summation and part (i).)

We deduce from Lemma 9.3

Theorem 9.4 (PERRON's formula) *Suppose that the* DIRICHLET *series $\sum_{n=1}^\infty \frac{f(n)}{n^s}$ converges for $\sigma = c$ absolutely. Then, for $x \notin \mathbb{Z}$,*

$$\sum_{n\leq x} f(n) = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} \sum_{n=1}^{\infty} \frac{f(n)}{n^s}\frac{x^s}{s}\,ds, \tag{9.5}$$

and, for arbitrary x,

$$\int_0^x \sum_{n\le u} f(n)\,du = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} \sum_{n=1}^{\infty} \frac{f(n)}{n^s}\frac{x^{s+1}}{s(s+1)}\,ds. \tag{9.6}$$

PERRON's formula gives a first glance on the intimate relation between arithmetic functions (number theory) and their associated DIRICHLET series (analysis).

Proof. Obviously, the integral in formula (9.5) equals

$$\int_{c-i\infty}^{c+i\infty} \sum_{n=1}^{\infty} \frac{f(n)}{n^s}\frac{x^s}{s}\,\mathrm{d}s = \sum_{n=1}^{\infty} f(n)\int_{c-i\infty}^{c+i\infty} \left(\frac{x}{n}\right)^s \frac{\mathrm{d}s}{s};$$

here interchanging integration and summation is allowed by the absolute convergence of the series. In view to Lemma 9.4 formula (9.5) follows.

In order to prove formula (9.6) we apply (9.5) with $f(n)n^w, w \ge 0$, instead of $f(n)$, and obtain

$$\sum_{n\le x} f(n)n^w = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} \sum_{n=1}^{\infty} \frac{f(n)}{n^s}\frac{x^{s+w}}{s+w}\,\mathrm{d}s.$$

Thus we get by subtraction

$$\sum_{n\le x} f(n)(x^w - n^w) = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} \sum_{n=1}^{\infty} \frac{f(n)}{n^s}\frac{wx^{s+w}}{s(s+w)}\,\mathrm{d}s.$$

Obviously, this formula holds for $x \in \mathbb{N}$ too. We set $w = 1$, and note

$$\int_0^x \sum_{n\le u} f(n)\,\mathrm{d}u = \sum_{n\le x} f(n)\int_n^x \mathrm{d}u = \sum_{n\le x} f(n)(x-n).$$

Thus we obtain (9.6), and the theorem is shown. •

As an immediate application we note, for $0 < |z| \le 1$ and $c > 1$,

$$\int_0^x \sum_{n\le u} z^{\omega(n)}\,\mathrm{d}u = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} L(s,z,\omega)\frac{x^{s+1}}{s(s+1)}\,\mathrm{d}s; \tag{9.7}$$

a similar formula holds when we replace $\omega(n)$ by $\Omega(n)$. Later we shall prove an asymptotic formula for the arithmetic expression on the left hand side by evaluating the analytic right hand side.

Chapter 10

Characteristic functions

Many information on a probability law can be derived by studying the related *characteristic function.* Let $\mathbf{F}$ be a distribution function, then its **characteristic function** is given by the FOURIER transform of the STIELTJES measure $\mathrm{d}\mathbf{F}(z)$, namely

$$\varphi_{\mathbf{F}}(\tau) := \int_{-\infty}^{\infty} \exp(i\tau z)\, \mathrm{d}\mathbf{F}(z).$$

This defines a uniformly continuous function on the real line which satisfies, for $\tau \in \mathbb{R}$,

$$|\varphi_{\mathbf{F}}(\tau)| \leq \int_{-\infty}^{\infty} \mathrm{d}\mathbf{F}(z) = 1 = \varphi_{\mathbf{F}}(0).$$

The intimate relationship between the distribution function $\mathbf{F}$ and its characteristic function $\varphi_{\mathbf{F}}$ is ruled by the following

Lemma 10.1 (Inversion formula) *Let $\mathbf{F}$ be a distribution function with charactersitic function $\varphi_{\mathbf{F}}$. Then, for $\alpha, \beta \in \mathcal{C}(\mathbf{F})$,*

$$\mathbf{F}(\beta) - \mathbf{F}(\alpha) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{\exp(-i\tau\alpha) - \exp(-i\tau\beta)}{i\tau} \varphi_{\mathbf{F}}(\tau)\, d\tau.$$

In particular, the distribution function is uniquely determined by its charcteristic function.

Note that the singularity of the integrand in the formula of the above lemma is removable.

Proof. Without loss of generality $\alpha \leq \beta$. Using FUBINI's theorem, we can rewrite the integral on the right hand side of the formula in the lemma as

$$\int_{-\infty}^{\infty} \frac{\exp(-i\tau\alpha) - \exp(-i\tau\beta)}{i\tau} \int_{-\infty}^{\infty} \exp(i\tau w)\, \mathrm{d}\mathbf{F}(w)\, \mathrm{d}\tau$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{\exp(i\tau(w-\alpha)) - \exp(i\tau(w-\beta))}{i\tau}\, \mathrm{d}\tau\, \mathrm{d}\mathbf{F}(w).$$

By Exercise 9.4 the inner integral equals

$$\begin{aligned} &\frac{1}{2}\int_{-\infty}^{\infty} \frac{\exp(iu(w-\alpha)) - \exp(-iu(w-\alpha))}{iu}\, \mathrm{d}u \\ &\qquad -\frac{1}{2}\int_{-\infty}^{\infty} \frac{\exp(iu(w-\beta)) - \exp(-iu(w-\beta))}{iu}\, \mathrm{d}u \\ &= \pi(\operatorname{sgn}(w-\alpha) - \operatorname{sgn}(w-\beta)), \end{aligned}$$

which is $= 2\pi$ if $\alpha < w < \beta$, and $= 0$ if $w < \alpha$ or $w > \beta$. This leads to the formula in the lemma. If the distribution function $\mathbf{G}$ has the same characteristic function, then the formula proved above yields $\mathbf{F}(\alpha) = \mathbf{G}(\alpha)$ for almost all α. Since $\mathbf{F}$ and $\mathbf{G}$ both are right-continuous and non-decreasing, we finally obtain $\mathbf{F} = \mathbf{G}$. The lemma is shown. •

This lemma has some powerful consequences which we will use in what follows.

Exercise 10.1 *Let $h > 0$ and let $\mathbf{F}$ be a distribution function with characteristic function $\varphi_{\mathbf{F}}$.*

(i) Prove, for $z \in \mathbb{R}$,

$$\frac{1}{h}\left\{\int_{z}^{z+h} - \int_{z-h}^{z}\right\} \mathbf{F}(t)\, dt = \frac{1}{2\pi}\int_{-\infty}^{\infty} \left(\frac{\sin\frac{\tau}{2}}{\frac{\tau}{2}}\right)^2 \exp\left(-\frac{i\tau z}{h}\right) \varphi_{\mathbf{F}}\left(\frac{\tau}{h}\right)\, d\tau;$$

(Hint: apply Lemma 10.1 to the integrals on the left hand side, and calculate their characteristic functions by partial integration.)

(ii) show that

$$\frac{1}{h}\int_{z}^{z+h} \mathbf{F}(t)\, dt \qquad \textit{and} \qquad \frac{1}{h}\int_{z-h}^{z} \mathbf{F}(t)\, dt$$

both define distribution functions.
(Hint: for all $\varepsilon > 0$ there exists an t_0 such that $F(t) \geq F(+\infty) - \varepsilon$ for all $t \geq t_0$.)

It is time to give an example. We note for the standard normal distribution (3.1):

Lemma 10.2 *The characteristic function of the standard normal distribution Φ is given by*

$$\varphi_{\mathbf{\Phi}}(\tau) = \exp\left(-\frac{\tau^2}{2}\right).$$

Proof. By definition,

$$\begin{aligned}
\varphi_{\mathbf{\Phi}}(\tau) &= \int_{\infty}^{\infty} \exp(i\tau z)\,\mathrm{d}\mathbf{\Phi}(z) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} \exp(i\tau z)\exp\left(-\frac{z^2}{2}\right)\,\mathrm{d}z \\
&= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} (\cos(\tau z) + i\sin(\tau z))\exp\left(-\frac{z^2}{2}\right)\,\mathrm{d}z \\
&= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} \cos(\tau z)\exp\left(-\frac{z^2}{2}\right)\,\mathrm{d}z,
\end{aligned}$$

since $\sin(\tau z)\exp(-\frac{z^2}{2})$ is an odd function. Differentiation on both sides with respect to τ (which is obviously allowed), and integration by parts (with $\sin(\tau z)$ and $z\exp(-\frac{z^2}{2})$), yields

$$\begin{aligned}
{\varphi_{\mathbf{\Phi}}}'(\tau) &= -\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} z\sin(\tau z)\exp\left(-\frac{z^2}{2}\right)\,\mathrm{d}z \\
&= -\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} \tau\cos(\tau z)\exp\left(-\frac{z^2}{2}\right)\,\mathrm{d}z \\
&= -\tau\varphi_{\mathbf{\Phi}}(\tau).
\end{aligned}$$

Therefore, the characteristic function $\varphi_{\mathbf{\Phi}}(\tau)$ solves the differential equation

$$\frac{y'}{y} = -\tau,$$

and hence, integration yields

$$\log|\varphi_{\mathbf{\Phi}}(\tau)| = \int \frac{\varphi_{\mathbf{\Phi}}'}{\varphi_{\mathbf{\Phi}}}(\tau)\,\mathrm{d}\tau + c' = -\int \tau\,\mathrm{d}\tau + c' = -\frac{\tau^2}{2} + c,$$

where c', c are constants. Taking the exponential gives

$$\varphi_{\mathbf{\Phi}}(\tau) = \exp\left(c - \frac{\tau^2}{2}\right).$$

In view to $\varphi_{\mathbf{\Phi}}(0) = 1$ we obtain $c = 0$, which finishes the proof. •

The proof above is not straightforward but elementary. It is much easier to find the characteristic function of the uniform distribution.

Exercise 10.2 *Prove that the characteristic function of the uniform distribution* ν *on the interval* $[0,1]$ *is given by*

$$\varphi_\nu(\tau) = \frac{\exp(i\tau) - 1}{i\tau}.$$

The following theorem links the weak convergence of a sequence of distribution functions to the pointwise convergence of the corresponding sequence of their characteristic functions.

Theorem 10.3 (Lévy's continuity theorem, 1925) *Let $\{\mathbf{F_n}\}$ be a sequence of distribution functions and $\{\varphi_{\mathbf{F_n}}\}$ the corresponding sequence of their charactersitic functions. Then $\mathbf{F_n}$ converges weakly to a distribution function $\mathbf{F}$ if and only if $\varphi_{\mathbf{F_n}}$ converges pointwise on $\mathbb{R}$ to a function φ which is continuous at 0. Additionally, in this case, φ is the characteristic function of $\mathbf{F}$, and the convergence of $\varphi_{\mathbf{F_n}}$ to $\varphi = \varphi_{\mathbf{F}}$ is uniform on any compact subset.*

The following proof is due to Cramér.

Proof. We start with the necessity. If $\mathbf{F_n}$ converges weakly to $\mathbf{F}$, then there exists for any $\varepsilon > 0$ a real number $T = T(\varepsilon)$ such that

$$\sup_{n\in\mathbb{N}}\sup_{\tau\in\mathbb{R}}\left|\int_{|z|>T}\exp(i\tau z)\,\mathrm{d}\mathbf{F_n}(z)\right| \leq \sup_{n\in\mathbb{N}}\int_{|z|>T}\mathrm{d}\mathbf{F_n}(z) \leq \varepsilon.$$

Without loss of generality we may assume that $\pm T \in \mathcal{C}(\mathbf{F})$, then

$$\int_{-T}^{T}\exp(i\tau z)\,\mathrm{d}\mathbf{F_n}(z) \to \int_{-T}^{T}\exp(i\tau z)\,\mathrm{d}\mathbf{F}(z),$$

in any finite τ-interval, as $n \to \infty$ (Stieltjes integrals behave sufficiently *smooth*). The last integral equals $\varphi_{\mathbf{F}}(\tau) + O(\varepsilon)$, which implies that $\varphi_{\mathbf{F_n}} \to \varphi_{\mathbf{F}}$ uniformly on any compact subset, as $n \to \infty$.

To prove the converse, it is sufficient to show that, if $\varphi_{\mathbf{F_n}}$ converges pointwise to a limit φ, and if φ is continuous at 0, then $\mathbf{F_n}$ converges weakly to a distribution function $\mathbf{F}$. By the above given part of the proof it will then follow that φ is the characteristic function of $\mathbf{F}$, and that the convergence $\varphi_{\mathbf{F_n}} \to \varphi_{\mathbf{F}}$, as $n \to \infty$, is uniform on compact subsets.

Let $\mathcal{Z} := \{z_1, z_2, \ldots\}$ be a dense subset of $\mathbb{R}$ consisting of continuity points of $\mathbf{F}$ and all $\mathbf{F_n}$. Since the values of $\mathbf{F_n}(z_k)$ lie in $[0,1]$, the theorem of Bolzano-Weierstrass yields the existence of a convergent subsequence $\{\mathbf{F_{n1}}(z_1)\}$, and, by a standard diagonal argument, there exists a sub-subsequence $\{\mathbf{F_{nn}}\}$ of $\{\mathbf{F_n}\}$ which converges on $\mathcal{Z}$. Using the properties of the distribution functions $\mathbf{F_n}$, one can even find a subsequence $\{\mathbf{F_{n_j}}\}$ which converges weakly to a non-decreasing right-continuous function $\mathbf{F}$. Obviously, $0 \leq \mathbf{F}(z) \leq 1$ for all $z \in \mathbb{R}$. It remains to show that $\mathbf{F}(+\infty) - \mathbf{F}(-\infty) = 1$. We have, by Exercise 10.1,

$$\frac{1}{h}\left\{\int_0^h - \int_{-h}^0\right\}\mathbf{F_{n_j}}(t)\,\mathrm{d}t = \frac{1}{2\pi}\int_{-\infty}^{\infty}\left(\frac{\sin\frac{\tau}{2}}{\frac{\tau}{2}}\right)^2\varphi_{\mathbf{F_{n_j}}}\left(\frac{\tau}{h}\right)\,\mathrm{d}\tau.$$

Sending $j \to \infty$ and applying LEBESGUE's theorem, we arrive at

$$\frac{1}{h}\left\{\int_0^h - \int_{-h}^0\right\}\mathbf{F}(t)\,\mathrm{d}t = \frac{1}{2\pi}\int_{-\infty}^{\infty}\left(\frac{\sin\frac{\tau}{2}}{\frac{\tau}{2}}\right)^2 \varphi\left(\frac{\tau}{h}\right)\,\mathrm{d}\tau;$$

here $\mathbf{F}$ is the weak limit of $\mathbf{F_{n_j}}$, and φ is the pointwise limit of $\varphi_{\mathbf{F_{n_j}}}$. By Exercise 10.1, we may interpret the left hand side as the difference of distribution functions. Hence, as $h \to \infty$,

$$\mathbf{F}(+\infty) - \mathbf{F}(-\infty) = \lim_{h\to\infty}\frac{1}{2\pi}\int_{-\infty}^{\infty}\left(\frac{\sin\frac{\tau}{2}}{\frac{\tau}{2}}\right)^2 \varphi\left(\frac{\tau}{h}\right)\,\mathrm{d}\tau.$$

Since φ is bounded and continuous at 0, we may interchange by LEBESGUE's theorem the limit with the integration and use Exercise 9.4 to obtain

$$\begin{aligned}\mathbf{F}(+\infty) - \mathbf{F}(-\infty) &= \frac{1}{2\pi}\int_{-\infty}^{\infty}\left(\frac{\sin\frac{\tau}{2}}{\frac{\tau}{2}}\right)^2 \lim_{h\to\infty}\varphi\left(\frac{\tau}{h}\right)\,\mathrm{d}\tau \\ &= \varphi(0)\cdot\frac{1}{2\pi}\int_{-\infty}^{\infty}\left(\frac{\sin\frac{\tau}{2}}{\frac{\tau}{2}}\right)^2\,\mathrm{d}\tau \\ &= \varphi(0).\end{aligned}$$

Further, $\varphi(0) = \lim_{n\to\infty}\varphi_{\mathbf{F_{n_j}}}(0)$, and $\varphi_{\mathbf{F_{n_j}}}(0) = 1$ for all $n \in \mathbb{N}$. Therefore, $\mathbf{F}(+\infty) - \mathbf{F}(-\infty) = 1$, and the weak limit $\mathbf{F}$ of the sequence $\mathbf{F_{n_j}}$ is a distribution function. Obviously, this holds also for any other weak limit $\mathbf{G}$. But since $\mathbf{G}$ has also the characteristic function φ, and distribution functions are uniquely determined by their characteristic functions, we obtain $\mathbf{F} = \mathbf{G}$. Hence any weakly convergent subsequence of $\{\mathbf{F_n}\}$ converges to the same limit $\mathbf{F}$, and hence, $\{\mathbf{F_n}\}$ itself converges weakly to $\mathbf{F}$. The theorem is shown. •

For a special class of distribution functions $\mathbf{F}$ one can find a quantative estimate for approximations of $\mathbf{F}$ in terms of the corresponding characteristic functions by the following result. Define for a real-valued function f, given on the compact set $\mathbb{R}\cup\{\pm\infty\}$,

$$\|f\|_\infty = \max_{-\infty\leq x\leq\infty}|f(x)|.$$

Then

Theorem 10.4 (BERRY-ESSEEN inequality, 1941/1945) *Let $\mathbf{F}, \mathbf{G}$ be two distribution functions with characteristic functions $\varphi_{\mathbf{F}}, \varphi_{\mathbf{G}}$. Suppose that $\mathbf{G}$ is differentiable and that $\mathbf{G}'$ is bounded on $\mathbb{R}$. Then, for all $T > 0$,*

$$\|\mathbf{F} - \mathbf{G}\|_\infty \ll \frac{\|\mathbf{G}'\|_\infty}{T} + \int_{-T}^{T}\left|\frac{\varphi_{\mathbf{F}}(\tau) - \varphi_{\mathbf{G}}(\tau)}{\tau}\right|\,d\tau,$$

where the implicit constant is absolute.

We omit the lengthy proof, which, for example, can be found in [30] or in [6], §1, but give an interesting application to a result mentioned in Chapter 3. The convergence of the scaled random walk to the normal distribution (3.2) satisfies the quantitative estimate

$$\mathbf{P}\left(\frac{Z_n}{\sqrt{n}} < x\right) = \mathbf{\Phi}(x) + O\left(n^{-\frac{1}{2}}\right),$$

as $n \to \infty$. For this and other applications we refer to [6], §1 and §3.

Chapter 11

Mean value theorems

Lévy's continuity theorem has an important consequence, namely a criterion whether an arithmetic function possesses a limit law or not.

Corollary 11.1 *Let f be a real-valued arithmetic function. Then f possesses a limit law $\mathbf{F}$ if and only if the sequence of functions*

$$\frac{1}{N}\sum_{n\leq N}\exp(i\tau f(n))$$

converges with $N\to\infty$ pointwise on $\mathbb{R}$ to a function $\varphi(\tau)$ which is continuous at 0. In this case $\varphi=\varphi_{\mathbf{F}}$ is the characteristic function of $\mathbf{F}$.

Proof. By (3.3) the characteristic function of the distribution function $\mathbf{F_N}$ of f is

$$\varphi_{\mathbf{F_N}}(\tau)=\int_{-\infty}^{\infty}\exp(i\tau z)\,\mathrm{d}\mathbf{F_N}(z)=\frac{1}{N}\sum_{n\leq N}\exp(i\tau f(n)).$$

Consequently, Lévy's continuity theorem translates the weak convergence of the distribution functions to the pointwise convergence of the corresponding characteristic functions. •

If f is an additive function, then the function $n\mapsto\exp(i\tau f(n))$ is for each fixed τ a multiplicative arithmetic function. Thus, the problem of the existence of a limit law for f is equivalent to the problem of the existence of the mean value of a certain multiplicative function. A complete solution was found by Erdös and Wintner [8]:

Theorem 11.2 (Erdös+Wintner, 1939) *A real-valued additive function $f(n)$ possesses a limit law if and only if the following three series converge simultaneously for at least one value $R > 0$:*

$$\sum_{|f(p)|>R} \frac{1}{p}\,, \qquad \sum_{|f(p)|\le R} \frac{f(p)^2}{p}\,, \qquad \sum_{|f(p)|\le R} \frac{f(p)}{p}.$$

If this is the case, then the characteristic function of the limiting distribution function $\mathbf{F}$ *is given by the convergent product*

$$\varphi_{\mathbf{F}}(n) = \prod_p \left(1 - \frac{1}{p}\right) \sum_{k=0}^{\infty} \frac{\exp(i\tau f(p^k))}{p^k}.$$

The idea of proof is based on Kolmogorov's three series theorem on sums of independent random variables; see [9], §IX.9.

In the following years the question arose when a multiplicative function of modulus ≤ 1 has a non-zero mean value. The ultimative answer was given by Halász [11], namely

Theorem 11.3 (Halász, 1968) *Let g be a multiplicative function with values in the unit disc. If there exists some $\tau \in \mathbb{R}$ such that*

$$\sum_p \frac{1 - Re\ g(p)p^{-i\tau}}{p}$$

converges, then

$$\frac{1}{x}\sum_{n\le x} g(n) = \frac{x^{i\tau}}{1+i\tau} \prod_{p\le x} \left(1 - \frac{1}{p}\right) \sum_{k=0}^{\infty} \frac{g(p^k)}{p^{k(1+i\tau)}} + o(1),$$

as $x \to \infty$. If there exists no τ with the above property, then

$$\frac{1}{x}\sum_{n\le x} g(n) = o(1).$$

Note that Wirsing [33] obtained a similar result for real-valued multiplicative functions in 1967. To indicate the power of these results note that an application to Möbius' μ-function yields

$$\sum_{n\le x} \mu(n) = o(x),$$

which is equivalent to the prime number theorem (5.6) (for the equivalence see [30], §I.3).

Unfortunately, the proofs of these mean value theorems are beyond the scope of this course, we refer the interested reader to the original papers and [30], §III.4; further mean value results can be found in [28].

A further application of characteristic functions is to find in the theory of *uniform distribution modulo* 1.

Chapter 12

Uniform distribution modulo 1

We say that a sequence of non-negative real numbers α_n is **uniformly distributed modulo** 1 if for any interval $\mathcal{I} \subset [0,1)$

$$\mathbf{d}\{n \,:\, \alpha_n - [\alpha_n] \in \mathcal{I}\} = \lambda(\mathcal{I}),$$

where $\lambda(\mathcal{I})$ is the LEBESGUE-measure of $\mathcal{I}$ (i.e. the length of $\mathcal{I}$). This means that the proportion of α_n, which fractional parts $\alpha_n - [\alpha_n]$ lie in $\mathcal{I}$, corresponds to the proportion of the interval $\mathcal{I}$ in $[0,1)$.

H. WEYL's celebrated criterion on uniform distribution [32] states

Theorem 12.1 (H. WEYL; **1916**) *A sequence of real numbers α_n is uniformly distributed* mod 1 *if, and only if, for each non-zero integer m*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n\leq N} \exp(2\pi i m\alpha_n) = 0.$$

Proof. Assume that the sequence $\{\alpha_n\}$ is uniformly distributed mod 1, then the corresponding distribution functions

$$\mathbf{F_N}(z) = \frac{1}{N} \sum_{\substack{n\leq N\\ \alpha_n - [\alpha_n]\leq z}} 1$$

converge weakly to the uniform distribution on $[0,1]$, as $N \to \infty$, and, by LEVY's continuity theorem, the corresponding characteristic functions converge pointwise to the characteristic function of the uniform distribution, i.e.

$$\varphi_{\mathbf{F_N}}(\tau) = \int_0^1 \exp(i\tau z)\,\mathrm{d}\mathbf{F_N}(z) \quad \to \quad \varphi_\nu(\tau) = \int_0^1 \exp(i\tau z)\,\mathrm{d}\mathbf{F}_\nu(z),$$

as $N \to \infty$. Setting $\tau = 2\pi m, m \neq 0$, we obtain in view to Exercise 10.2 that

$$\frac{1}{N}\sum_{n\leq N}\exp(2\pi i m\alpha_n) = \int_0^1 \exp(2\pi i m z)\,\mathrm{d}\mathbf{F}_{\mathbf{N}}(z)$$

tends with $N \to \infty$ to

$$\int_0^1 \exp(2\pi i m z)\,\mathrm{d}z = \frac{\exp(2\pi i m) - 1}{2\pi i m} = 0.$$

We give only a sketch of the argument for the converse implication. For simplicity, we may assume that $\mathbf{F}$ is absolutely continuous. With a little help from FOURIER analysis one can show that $\mathbf{F}$ has a representation

$$\mathbf{F}(z) = \int_0^1 \mathbf{F}(u)\,\mathrm{d}u + \sum_{\substack{m=-\infty \\ m\neq 0}}^{\infty} \frac{c_m - 1}{2\pi i m}\exp(2\pi i m z),$$

where

$$c_m := \int_0^1 \exp(-2\pi i m z)\,\mathrm{d}\mathbf{F}(z).$$

Since

$$\int_0^1 \exp(2\pi i m z)\,\mathrm{d}\mathbf{F}_{\mathbf{N}}(z) = \frac{1}{N}\sum_{n\leq N}\exp(2\pi i m\alpha_n)$$

tends with $N \to \infty$ to zero, it follows that $c_m = 0$ for non-zero m. This gives above

$$F(z) = \frac{1}{2} - \sum_{\substack{m=-\infty \\ m\neq 0}}^{\infty} \frac{\exp(2\pi i m z)}{2\pi i m} = z,$$

which implies the uniform distribution for $\{\alpha_m\}$. •

Exercise 12.1 *(for experts in* FOURIER *analysis) Fill the gaps in the sketch of proof of the converse implication above.*

We note a nice application to indicate the power of this criterion.

Corollary 12.2 (KRONECKER's approximation theorem; 1884) *The sequence $\{n\xi\}$ is uniformly distributed* mod 1 *if and only if ξ is irrational.*

Proof. Let ξ be irrational. By the formula for the geometric series, we have, for any non-zero integer m,

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n\leq N}\exp(2\pi i m n\xi) = \lim_{N\to\infty}\frac{1}{N}\frac{\exp(2\pi i m\xi) - \exp(2\pi i m(N+1)\xi)}{1-\exp(2\pi i m\xi)} = 0.$$

Otherwise, if $\xi = \frac{a}{b}$, then the limit is non-zero for multiples m of b. Thus, WEYL's criterion yields the assertion of the corollary. •

Moreover, we say that the sequence $\{\alpha_n\}$ lies **dense** mod 1 if for any $\varepsilon > 0$, and any $\alpha \in [0, 1)$, exists an α_n such that

$$|\alpha - (\alpha_n - [\alpha_n])| < \varepsilon.$$

Obviously, a sequence which is uniformly distributed mod 1 lies also dense mod 1. However, the converse implication is not true in general.

Exercise 12.2 *Show that the sequence* $\{\log n\}$

(i) lies dense mod 1*;*
(Hint: consider the subsequence $\{\log(2^k)\}$*.)*

(ii) is not uniformly distributed mod 1*.*
(Hint: replace the sum in Theorem 12.1 by the corresponding integral.)

An interesting open problem is whether the sequence $\{\exp(n)\}$ is uniformly distributed or not. A further application of uniform distribution modulo 1 is numerical integration. The interested reader can find more details on this and allied topics in [15].

Chapter 13

The theorem of ERDÖS-KAC

Now we are going to prove the explicit form (1.5) of the limit distribution of the prime divisor counting functions $\omega(n)$ and $\Omega(n)$. The easiest and first proof due to ERDÖS and KAC [7] is elementary but tricky and quite delicate. We will give a proof more or less following the one of RÉNYI and TURÁN [24], including a certain modification due to SELBERG [29], which enables one to obtain further knowledge concerning the speed of convergence to the normal distribution. Moreover, this method applies to other problems as well. For some interesting historical comments see [6], §12, pp.18.

Let z be a non-zero complex constant of modulus ≤ 1. We shall prove in Chapter 15 by analytic methods the asymptotic formula

$$\sum_{n\leq x} z^{\omega(n)} = \lambda(z)x(\log x)^{z-1} + O\left(x(\log x)^{Rez-2}\right), \tag{13.1}$$

where $\lambda(z)$ is an entire function with $\lambda(1) = 1$. This implies

Theorem 13.1 (ERDÖS+KAC, 1939; RÉNYI+TURÁN, 1957) *As $N \to \infty$,*

$$\nu_{\mathbf{N}}\left\{n\,:\,\frac{\omega(n)-\log\log N}{\sqrt{\log\log N}}\leq x\right\} = \mathbf{\Phi}(x) + O\left((\log\log N)^{-\frac{1}{2}}\right).$$

Proof. We consider

$$\mathbf{F_N}(x) = \nu_{\mathbf{N}}\left\{n\,:\,\frac{\omega(n)-\log\log N}{\sqrt{\log\log N}}\leq x\right\},$$

and denote by $\varphi_{\mathbf{F_N}}(\tau)$ its characteristic function, i.e.

$$\varphi_{\mathbf{F_N}}(\tau) = \int_{\infty}^{\infty}\exp(i\tau z)\,\mathrm{d}\mathbf{F_N}(z) = \frac{1}{N}\sum_{n\leq N}\exp\left(\frac{i\tau(\omega(n)-\log\log N)}{\sqrt{\log\log N}}\right).$$

By (13.1), we have, uniformly for $N \geq 2, t \in \mathbb{R}$,

$$\frac{1}{N}\sum_{n\leq N} \exp(it\omega(n)) = \lambda(\exp(it))(\log N)^{\exp(it)-1} + O((\log N)^{\cos t - 2}).$$

Putting $T := \sqrt{\log\log N}$, and $t := \frac{\tau}{T}$, then the latter formual implies for $|\tau| \leq T$

$$\begin{aligned} \varphi_{\mathbf{F_N}}(\tau) &= \frac{1}{N}\sum_{n\leq N} \exp\left(\frac{i\tau(\omega(n) - T^2)}{T}\right) \\ &= \lambda(\exp(it))\exp((\exp(it)-1)T^2 - i\tau T) + O\left(\exp(T^2(\cos t - 2))\right). \end{aligned} \tag{13.2}$$

Since $\cos t - 1 \leq -2(\frac{t}{\pi})^2$ for $|t| \leq 1$, we deduce

$$\varphi_{\mathbf{F_N}}(\tau) \ll \exp\left(-\frac{2\tau^2}{\pi^2}\right); \tag{13.3}$$

we shall use this estimate later for *large* values of τ. Since, for $|t| \leq 1$,

$$\exp(it) - 1 = it - \frac{t^2}{2} + O(|t|^3),$$

and since $\lambda(z)$ is an entire function with $\lambda(1) = 1$, we have

$$\lambda(\exp(it)) = \sum_{k=0}^{\infty} \frac{\lambda^{(k)}(1)}{k!}(\exp(it)-1)^k = 1 + O(|t|)$$

for $|t| \leq 1$. Therefore, we obtain in view to (13.2), for $|\tau| < T^{\frac{1}{3}}$,

$$\varphi_{\mathbf{F_N}}(\tau) = \exp\left(-\frac{\tau^2}{2}\right)\left(1 + O\left(\frac{|\tau| + |\tau|^3}{T}\right)\right) + O\left(\frac{1}{\log N}\right); \tag{13.4}$$

we shall use this formula later for $\frac{1}{\log N} \leq |\tau| < T^{\frac{1}{3}}$. Sending $N \to \infty$, we deduce in view to Lemma 10.2 that

$$\varphi_{\mathbf{F_N}}(\tau) \to \exp\left(-\frac{\tau^2}{2}\right) = \varphi_{\mathbf{\Phi}}(\tau),$$

i.e. the characteristic functions $\varphi_{\mathbf{F_N}}$ converge pointwise to the characteristic function of the normal distribution $\mathbf{\Phi}(x)$. Applying Levy's continuity theorem, we get

$$\lim_{N\to\infty} \nu_{\mathbf{N}}\left\{n : \frac{\omega(n) - \log\log N}{\sqrt{\log\log N}} \leq x\right\} = \mathbf{\Phi}(x);$$

this is exactly Erdös' and Kac's formula (1.5).

In order to obtain the quantitive result of the theorem we need a further estimate of $\varphi_{\mathbf{F_N}}(\tau)$ for *small* values of τ. When $|\tau| \le \frac{1}{\log N}$, then the trivial estimate $\exp(iy) = 1 + O(y), y \in \mathbb{R}$, yields in combination with the Cauchy-Schwarz inequality

$$\begin{aligned}\varphi_{\mathbf{F_N}}(\tau) &= 1 + O\left(\frac{|\tau|}{TN}\sum_{n\le N}|\omega(n) - \log\log N|\right)\\ &= 1 + O\left(\frac{|\tau|}{TN}\left(N\sum_{n\le N}|\omega(n) - \log\log N|^2\right)^{\frac{1}{2}}\right),\end{aligned}$$

which leads in view to the Hardy-Ramanujan Theorem 7.2 to

$$(13.5) \qquad \varphi_{\mathbf{F_N}}(\tau) = 1 + O(|\tau|).$$

Now we apply the Berry-Esseen inequality Theorem 10.4. In view to Lemma 10.2 we get

$$\begin{aligned}\|\mathbf{F_N} - \mathbf{\Phi}\|_\infty &\ll \frac{\|\mathbf{\Phi}'\|_\infty}{T} + \int_{-T}^{T}\left|\frac{\varphi_{\mathbf{F_N}}(\tau) - \varphi_{\mathbf{\Phi}}(\tau)}{\tau}\right| \mathrm{d}\tau\\ &\ll \frac{1}{T} + \int_{-T}^{T}\left|\varphi_{\mathbf{F_N}}(\tau) - \exp\left(-\frac{\tau^2}{2}\right)\right|\frac{\mathrm{d}\tau}{|\tau|}.\end{aligned}$$

We split the appearing integral into three parts, and estimate in view to (13.5), (13.4) and (13.3)

$$\begin{aligned}\int_{-\frac{1}{\log N}}^{\frac{1}{\log N}} &\ll \int_{-\frac{1}{\log N}}^{\frac{1}{\log N}} \mathrm{d}\tau \ll \frac{1}{\log N},\\ \int_{\pm\frac{1}{\log N}}^{\pm T^{\frac{1}{3}}} &\ll \int_{-\infty}^{\infty}\left(\frac{1+\tau^2}{T}\right)\exp\left(-\frac{\tau^2}{2}\right)\mathrm{d}\tau + \frac{1}{\log N}\int_{\frac{1}{\log N}}^{T^{\frac{1}{3}}}\frac{\mathrm{d}\tau}{\tau} \ll \frac{1}{T},\\ \int_{\pm T^{\frac{1}{3}}}^{\pm\infty} &\ll \int_{T^{\frac{1}{3}}}^{\infty}\exp\left(-\frac{2\tau^2}{\pi^2}\right)\frac{\mathrm{d}\tau}{\tau} \ll \frac{1}{T}.\end{aligned}$$

This proves the theorem. •

It can be shown that the error term in Theorem 13.1 is best possible. This follows by studying the frequencies of positive integers n with $\nu(n) = k, k \in \mathbb{N}$; for details we refer to [30], §III.4.

Exercise 13.1 *Deduce from the asymptotic formula*

$$\sum_{n\le x} z^{\Omega(n)} = \mu(z)x(\log x)^{z-1} + O\left(x(\log x)^{Rez-2}\right), \tag{13.6}$$

where $\mu(z)$ is an entire function with $\mu(1)=1$, the limit law

$$\nu_{\mathbf{N}}\left\{n\,:\,\frac{\Omega(n)-\log\log N}{\sqrt{\log\log N}}\le x\right\} = \mathbf{\Phi}(x) + O\left((\log\log N)^{-\frac{1}{2}}\right),$$

as $N\to\infty$.

It remains to show formula (13.1). The first step towards a proof was done by (9.7) in Chapter 9. In Chapter 15 we will calculate the appearing integral by moving the path of integration to the left of the line $\sigma=1$. Therefore, we need an analytic continuation of $L(s,z,\omega)$.

However, it suffices to find a zero-free region for the RIEMANN zeta-function. The EULER product representation (2.7) implies immediately the non-vanishing of $\zeta(s)$ in the half plane of absolute convergence $\sigma>1$. As we shall see in the following chapter one can extend this zero-free region to the left.

We observe that the EULER product representation (9.2) of $L(s,z,\omega)$ is similar to (9.1). Define $G(s,z)=L(s,z,\omega)\zeta(s)^{-z}$, then, for $\sigma>1$,

$$G(s,z) = \prod_p\left(1+\frac{z}{p^s-1}\right)\left(1-\frac{1}{p^s}\right)^z = \sum_{n=1}^{\infty}\frac{b_z(n)}{n^s}, \tag{13.7}$$

where $b_z = z^\omega * \tau_{-z}$ is multiplicative with

$$b_z(1)=1\ ,\quad b_z(p^k) = (-1)^k\binom{z}{k} + z\sum_{j=0}^{k-1}(-1)^j\binom{z}{j}.$$

Since $b_z(p)=0$, we have, for $\sigma>\frac{1}{2}$,

$$\log G(s,z) = \sum_p \log\left(1+\sum_{k=2}^{\infty}\frac{b_z(p^k)}{p^{ks}}\right) \ll \sum_p \frac{1}{p^{2\sigma}} \ll 1. \tag{13.8}$$

This shows that

$$L(s,z,\omega) = G(s,z)\zeta(s)^z$$

is analytically continuable to any zero-free region of $\zeta(s)$ covering the half plane of absolute convergence $\sigma>1$.

Chapter 14

A zero-free region for $\zeta(s)$

To establish a zero-free region for the Riemann zeta-function to the left of the half plane of absolute convergence of its series expansion is a rather delicate problem. In view to (9.4) (Exercise 2.4, resp.) we have, for $\sigma > 0$,

$$\zeta(s) = \sum_{n\leq N}\frac{1}{n^s} + \frac{N^{1-s}}{s-1} + s\int_N^\infty \frac{[x]-x}{x^{s+1}}\,\mathrm{d}x \tag{14.1}$$

$$= \sum_{n\leq N}\frac{1}{n^s} + \frac{N^{1-s}}{s-1} + O\left(N^{-\sigma}\left(1+\frac{|s|}{\sigma}\right)\right). \tag{14.2}$$

This gives an analytic continuation of $\zeta(s)$ to the half plane $\sigma > 0$ except for a simple pole at $s = 1$.

Lemma 14.1 *For* $|t| \geq 1, 1-\frac{1}{2}(\log(|t|+1))^{-1} \leq \sigma \leq 2$,

$$\zeta(s) \ll \log(|t|+1)\ , \qquad \textit{and} \qquad \zeta'(s) \ll (\log(|t|+1))^2.$$

Proof. Since $n^{\overline{s}} = \overline{n^s}$, it follows that

$$\zeta(\overline{s}) = \sum_{n=1}^\infty \frac{1}{n^{\overline{s}}} = \sum_{n=1}^\infty \overline{\frac{1}{n^s}} = \overline{\zeta(s)}$$

for $\sigma > 1$, and by analytic contiunation elsewhere. Therefore, it suffices to consider only $t > 1$. Let $1-(\log(t+1))^{-1} \leq \sigma \leq 3$, then formula (14.2) implies

$$\zeta(s) \ll \sum_{n\leq t+1}\frac{1}{n} + \frac{(t+1)^{1-\sigma}}{\sigma} \ll \log(t+1).$$

The estimate for $\zeta'(s)$ follows immediately from CAUCHY's formula

$$\zeta'(s) = \frac{1}{2\pi i} \oint \frac{\zeta(z)}{(z-s)^2}\,\mathrm{d}z,$$

and standard estimates of integrals. •

For $\sigma > 1$,

$$|\zeta(\sigma+it)| = \exp\left(\sum_p \sum_{k=1}^{\infty} \frac{1}{kp^{k\sigma}} \cos(kt\log p)\right).$$

Since

$$(14.3) \qquad 17 + 24\cos\alpha + 8\cos(2\alpha) = (3+4\cos\alpha)^2 \geq 0,$$

it follows that

$$(14.4) \qquad \zeta(\sigma)^{17}|\zeta(\sigma+it)|^{24}|\zeta(\sigma+2it)|^{8} \geq 1.$$

Therefore

Lemma 14.2 $\zeta(1+it) \neq 0$ *for* $t \in \mathbb{R}$.

Proof. For small $\sigma > 1$, $\zeta(\sigma) \ll \frac{1}{\sigma-1}$ by (14.1). Assuming that $\zeta(1+it)$ has a zero for $t = t_0 \neq 0$, then it would follow that

$$|\zeta(\sigma+it_0)| \leq \zeta(\sigma) \ll \sigma - 1.$$

This leads to

$$\lim_{\sigma\to 1+} \zeta(\sigma)^{17}|\zeta(\sigma+it_0)|^{24} = 0,$$

contradicting (14.4). •

It can be shown that this non-vanishing of $\zeta(1+it)$ is equivalent to the prime number theorem (5.6) (see [22], §2.3.

A simple refinement of the argument in the proof of Lemma 14.2 allows a lower estimate of $\zeta(1+it)$: for $|t| \geq 1$ and $1 < \sigma < 2$, we deduce from (14.4) and Lemma 14.1

$$\frac{1}{|\zeta(\sigma+it)|} \leq \zeta(\sigma)^{\frac{17}{24}}|\zeta(\sigma+2it)|^{\frac{1}{3}} \ll (\sigma-1)^{-\frac{17}{24}}(\log(|t|+1))^{\frac{1}{3}}.$$

Furthermore, with Lemma 14.1,

$$(14.5) \qquad \zeta(1+it) - \zeta(\sigma+it) = -\int_1^{\sigma} \zeta'(u+it)\,\mathrm{d}u \ll |\sigma-1|(\log(|t|+1))^2.$$

Hence

$$\begin{aligned}|\zeta(1+it)| &\geq |\zeta(\sigma+it)| - c_1(\sigma-1)(\log(|t|+1))^2 \\ &\geq c_2(\sigma-1)^{\frac{17}{24}}(\log(|t|+1))^{-\frac{1}{3}} - c_1(\sigma-1)(\log(|t|+1))^2,\end{aligned}$$

where c_1, c_2 are certain positive constants. Chosing a constant $B > 0$ such that $A := c_2 B^{\frac{17}{24}} - c_1 B > 0$ and putting $\sigma = 1 + B(\log(|t|+1)^{-8}$, we obtain now

$$|\zeta(1+it)| \geq \frac{A}{(\log(|t|+1))^6}. \tag{14.6}$$

This gives even an estimate on the left of the line $\sigma = 1$.

Lemma 14.3 *There exists a positive constant δ such that*

$$\zeta(s) \neq 0 \qquad \textit{for} \qquad \sigma \geq 1 - \delta \min\{1, (\log(|t|+1))^{-8}\};$$

further, under the assumption $|s-1| \geq 1$, the estimates

$$\frac{\zeta'}{\zeta}(s) \ll (\log(|t|+1))^8 \ , \qquad \log\zeta(s) \ll \log(2\log(|t|+1))$$

hold.

Here we choose that branch of logarithm $\log\zeta(s)$ which is real on the real axis; the other values are defined by analytic continuation in a standard way.

Proof. In view to Lemma 14.1 the estimate (14.5) holds for $1 - \delta(\log(|t|+1))^{-8} \leq \sigma \leq 1$. Using (14.6), it follows that

$$|\zeta(\sigma+it)| \geq \frac{A - c_1\delta}{(\log(|t|+1))^6},$$

where the right hand side is positve for sufficiently small δ. This yields the zero-free region of Lemma 14.3; the estimate of the logarithmic derivative follows from the estimate above by use of Lemma 14.1. Finally, to obtain the bound for $\log\zeta(s)$ let $s_0 = 1 + \eta + it$ with some positive parameter η. Then

$$\log\left(\frac{\zeta(s)}{\zeta(s_0)}\right) = \int_{s_0}^{s} \frac{\zeta'}{\zeta}(u)\,\mathrm{d}u \ll |s - s_0|(\log(|t|+1))^8.$$

Using (14.1),

$$|\log\zeta(s_0)| \leq \log\zeta(1+\eta) = \log\left(\frac{1}{\eta}\right) + O(1).$$

Setting $\eta = c(\log(|t|+1))^{-8}$, we obtain

$$\log\zeta(s) \ll \log\left(\frac{1}{\eta}\right) + |\sigma - 1 - \eta|(\log(|t|+1))^8 \ll \log(2\log(|t|+1)).$$

The lemma is shown. •

Exercise 14.1 *Show that (14.3) gives the best possible estimates (by the method above).*

The famous and yet unproved Riemann **hypothesis** states that all complex zeros of $\zeta(s)$ lie on the so-called **critical line** $\sigma = \frac{1}{2}$, or equivalently, the non-vanishing of $\zeta(s)$ in the half plane $\sigma > \frac{1}{2}$. It seems that this hypothetical distribution of zeros is connected with the *functional equation*

$$\pi^{-\frac{s}{2}}\Gamma\left(\frac{s}{2}\right)\zeta(s) = \pi^{-\frac{1-s}{2}}\Gamma\left(\frac{1-s}{2}\right)\zeta(1-s), \tag{14.7}$$

which implies a symmetry of the zeros of $\zeta(s)$ with repsect to the critical line; for a proof of (14.7) see [30], §II.3.

In fact, the *first* zero on the critical line (i.e. the one with minimal imaginary part in the upper half plane) is

$$\frac{1}{2} + i\,14.13472\ldots.$$

Nevertheless, we show

Lemma 14.4 $\zeta(s) \neq 0$ *for* $|s-1| < 1$.

Proof. Integration by parts in (14.1) (resp. Euler*'s summation formula*) yields

$$\zeta(s) = \frac{1}{s-1} + \frac{1}{2} + \frac{s}{12} - \frac{s(s+1)}{2}\int_1^\infty \frac{B_2(u-[u])}{u^{s+2}}\,\mathrm{d}u,$$

where

$$B_2(u) = u^2 - u + \frac{1}{6}$$

is the second Bernoulli **polynomial**; note that $|B_2(u-[u])| \leq \frac{1}{6}$. Suppose that $\varrho = \beta + i\gamma$ is a zero of $\zeta(s)$ with $|s-1| \leq 1$. By symmetry, we may assume in view to the functional equation (14.7) that $\beta \geq \frac{1}{2}$. Setting $s = \varrho$ in the formula above, an application of the mean-value theorem yields the existence of some θ with $|\theta| \leq 1$ such that

$$0 = \frac{1}{\varrho - 1} + \frac{1}{2} + \frac{\varrho}{12} - \frac{\theta\varrho(\varrho+1)}{12}\int_1^\infty \frac{\mathrm{d}u}{u^{\beta+2}}\,\mathrm{d}u.$$

Multiplying this with $1-\varrho$, we may rewrite this as

$$(14.8)\qquad 1=\frac{1-\varrho}{2}\left(1+\frac{\varrho}{6}-\frac{\theta\varrho(\varrho+1)}{6}\int_1^\infty\frac{\mathrm{d}u}{u^{\frac{5}{2}}}\,\mathrm{d}u\right).$$

The modulus of the right hand side is

$$\frac{1}{2}\left(1+\frac{|\varrho|}{6}+\frac{|\varrho(\varrho+1)|}{9}\right)<\frac{1}{2}\left(1+\frac{1}{3}+\frac{2}{3}\right)=1,$$

which gives a contradiction to (14.8). This proves the lemma. •

In order to prove formula (13.1) we state some consequences we will use later on. By Lemma 14.4, the function

$$\frac{1}{s}\left((s-1)\zeta(s)\right)^z$$

is analytic in $|s-1|<1$, and hence, has there a power series expansion

$$(14.9)\qquad \frac{1}{s}\left((s-1)\zeta(s)\right)^z=\sum_{j=0}^{\infty}\frac{\gamma_j(z)}{j!}(s-1)^j,$$

where, by Cauchy's formula,

$$(14.10)\qquad \gamma_j(z)=\frac{j!}{2\pi i}\oint_{|s-1|=r}\frac{1}{s}\left((s-1)\zeta(s)\right)^z\frac{\mathrm{d}s}{(s-1)^{j+1}}.$$

Note that the $\gamma_j(z)$ are entire functions in z, satisfying the estimate

$$\frac{\gamma_j(z)}{j!}\ll(1+\varepsilon)^j,$$

where the implicit constant depends only on z and $\varepsilon\in(0,1)$.

Exercise 14.2 *Show that* $\gamma_0(z)=1$*,*

$$\gamma_j(1)=(-1)^j\int_1^\infty\frac{(u-[u])(\log u)^{j-1}}{u^2}\,du,$$

and, in particular, $\gamma_1(1)=\gamma-1$*, where* γ *is the* Euler-Masceroni *constant. (Hint: see Exercise 5.1.)*

For more details on the fascinating topic of the location of zeros of the Riemann zeta-function and its implications to number theory see [22], §2.4+2.5, as well as the monography [23].

Chapter 15

The Selberg-Delange method

Our aim is to prove the asymptotic formula (13.1); the proof of (13.6) is left to the reader. The proof bases on the Selberg-Delange method which works for more general Dirichlet series than (9.2) and (9.3) which we have to consider. This powerful method was developped by Selberg [29]; later it was generalized by Delange [3].

Theorem 15.1 *There exist constants $c_1, c_2 > 0$ such that, uniformly for sufficiently large $x, N \geq 0, 0 < |z| \leq 1$,*

$$\sum_{n \leq x} z^{\omega(n)} = x(\log x)^{z-1}\left(\sum_{k=0}^{N} \frac{\lambda_k(z)}{(\log x)^k} + O\left(\exp(-c_1\sqrt{\log x}) + \left(\frac{c_2 N + 1}{\log x}\right)^{N+1}\right)\right)$$

with

$$\lambda_k(z) := \frac{1}{\Gamma(z-k)} \sum_{h+j=k} \frac{\gamma_j(z)}{h!j!}\left[\frac{d^h}{ds^h} L(s,z,\omega)\zeta(s)^{-z}\right]_{s=1},$$

where $\gamma_j(z)$ is defined in (14.10).

Before we are able to start with the proof of Theorem 15.1 we quote a classical integral representation for the **Gamma-function**, which is for $\text{Re } z > 0$ defined by

$$\Gamma(z) = \int_0^\infty u^{z-1} \exp(-u)\, \mathrm{d}u,$$

and by analytic continuation elsewhere except for simple poles at $z = 0, -1, -2, \ldots$; for this and other properties of the Gamma-function, which we need later on, see [21].

Lemma 15.2 (Hankel's formula) *Denote by $\mathcal{H}$ the path formed by the circle $|s| = r > 0$, excluding the point $s = -r$, together with two copies of the half line $(-\infty, -r]$ with respective arguments $\pm\pi$. Then, for any complex z,*

$$\frac{1}{\Gamma(z)} = \frac{1}{2\pi i}\int_{\mathcal{H}} s^{-z}\exp(s)\,ds. \tag{15.1}$$

If $\mathcal{H}(x)$ denotes the part of $\mathcal{H}$ which is located in the half plane $\sigma > -x$, then uniformly for $x > 1$

$$\frac{1}{2\pi i}\int_{\mathcal{H}(\S)} s^{-z}\exp(s)\,ds = \frac{1}{\Gamma(z)} + O\left((2e)^{\pi|z|}\Gamma(1+|z|)\exp\left(-\frac{x}{2}\right)\right).$$

Proof. Obviously, the integral appearing in (15.1) is absolutely and uniformly convergent for all z. Hence, it defines an entire function of z, which, by the calculus of residues, does not dependent on r. When Re $z < -1$, the integral over the circle part $|s| = r$ of $\mathcal{H}$ tends with r to zero, and the the integral over the remaining path tends to

$$\begin{aligned}&\frac{1}{2\pi i}\int_0^\infty (\exp(i\pi z) - \exp(-i\pi z))\sigma^{-z}\exp(-\sigma)\,\mathrm{d}\sigma \\ = \;&\frac{\sin\pi z}{\pi}\int_0^\infty \sigma^{-z}\exp(-\sigma)\,\mathrm{d}\sigma = \frac{\sin\pi z}{\pi}\Gamma(1-z) = \frac{1}{\Gamma(z)};\end{aligned}$$

here we used the well-known identity $\Gamma(z)\Gamma(1-z) = \frac{\pi}{\sin\pi z}$. This proves the first formula for Re $z < 1$, and for arbitrary z by analytic continuation.

Now we consider the integral over the truncated contour $\mathcal{H}(x)$. Writing $s = \varrho\exp(\pm i\pi)$, we have

$$|s^{-z}\exp(s)| \leq (\exp(\pi)\sigma)^{|z|}\exp(-\sigma).$$

Thus,

$$\begin{aligned}\left\{\int_{\mathcal{H}} - \int_{\mathcal{H}(\S)}\right\} s^{-z}\exp(s)\,\mathrm{d}s &\ll \exp(\pi|z|)\int_x^\infty \varrho^{|z|}\exp(-\varrho)\,\mathrm{d}\varrho \\ &\leq \exp\left(\pi|z| - \frac{x}{2}\right)\int_0^x \varrho^{|z|}\exp\left(-\frac{\varrho}{2}\right)\,\mathrm{d}\varrho.\end{aligned}$$

Changing the variable $\varrho = 2t$ yields the estimate of the lemma. •

Now we are able to give the

Proof of Theorem 15.1. In view to our observations on the function $G(s,z)$, defined by (13.7), in Chapter 13, it came out that $L(s,z,\omega) = G(s,z)\zeta(s)^z$ can be analytically continued to any zero-free region of $\zeta(s)$ covering the half plane $\sigma > 1$. Hence Lemma 14.3 implies that $L(s,z,\omega)$ is analytic in the region

$$\sigma \geq 1 - \delta \min\{1, (\log(|t|+1))^{-8}\}, \tag{15.2}$$

where δ is some small positive constant. Furthermore, using (13.8), $L(s,z,\omega)$ satisfies there the estimate

$$L(s,z,\omega) = G(s,z)\exp(z\log\zeta(s)) \ll (|t|+1)^{\varepsilon}.$$

Hence, setting $c := 1 + \frac{1}{\log x}$, we find

$$\int_{c\pm iT}^{c\pm i\infty} L(s,z,\omega)\frac{x^{s+1}}{s(s+1)}\,\mathrm{d}s \ll x^{1+c}\int_T^\infty t^{\varepsilon-2}\,\mathrm{d}t \ll x^2 T^{\varepsilon-1}.$$

Therefore, we can deduce from (9.7)

$$\int_0^x \sum_{n\leq u} z^{\omega(n)}\,\mathrm{d}u = \frac{1}{2\pi i}\int_{c-iT}^{c+iT} L(s,z,\omega)\frac{x^{s+1}}{s(s+1)}\,\mathrm{d}s + O\left(x^2T^{\varepsilon-1}\right). \tag{15.3}$$

Now denote by $\mathcal{C}$ the path (symmetrical with respect to the real axis) consisting of the truncated Hankel contour $\mathcal{H}(r)$ surrounding the point $s=1$ with radius $r = \frac{1}{2}(\log x)^{-1}$, linear parts joining $1-r$ to $1-\frac{1}{2}\delta$, the arcs $\mathcal{A}_\pm$

$$\sigma = \sigma(t) := 1 - \frac{\delta}{2}\min\{1, (\log(|t|+1))^{-8}\},$$

and the linear segments $[\sigma(T) \pm iT, c \pm iT]$. Here, let x be sufficiently large such that $\mathcal{C}$ is contained in the region (15.2). Applying Cauchy's theorem, we obtain

$$\int_{c-iT}^{c+iT} L(s,z,\omega)\frac{x^{s+1}}{s(s+1)}\,\mathrm{d}s = \int_{\mathcal{C}} L(s,z,\omega)\frac{x^{s+1}}{s(s+1)}\,\mathrm{d}s,$$

since the integrand is analytic in (15.2). Obviously,

$$\begin{aligned}\int_{\sigma(T)\pm iT}^{\kappa\pm iT} L(s,z,\omega)\frac{x^{s+1}}{s(s+1)}\,\mathrm{d}s &\ll x^2T^{\varepsilon-2},\\ \int_{\mathcal{A}_\pm} L(s,z,\omega)\frac{x^{s+1}}{s(s+1)}\,\mathrm{d}s &\ll x^{1+\sigma(T)}\int_0^T (1+t)^{\varepsilon-2}\,\mathrm{d}t \ll x^{1+\sigma(T)}.\end{aligned}$$

Putting $T = \exp\left(\sqrt{\frac{\delta}{2-2\varepsilon}\log x}\right)$ for sufficiently large x, it follows from (15.3) that

$$\begin{aligned}\int_0^x \sum_{n\leq u} z^{\omega(n)}\,\mathrm{d}u &= \frac{1}{2\pi i}\int_{\mathcal{H}(r)} L(s,z,\omega)\frac{x^{s+1}}{s(s+1)}\,\mathrm{d}s\\ &\quad +O\left(x^2\exp\left(-c\sqrt{\log x}\right)\right),\end{aligned} \tag{15.4}$$

where $c = \sqrt{(1-\varepsilon)\delta}$. Obviously, the integral

$$\ell(x) := \frac{1}{2\pi i}\int_{\mathcal{H}(r)} L(s,z,\omega)\frac{x^{s+1}}{s(s+1)}\,\mathrm{d}s,$$

appearing in (15.4), is an infinitely differentiable function of $x > 0$, and, in particular, we have

$$\ell'(x) := \frac{1}{2\pi i}\int_{\mathcal{H}(r)} L(s,z,\omega)\frac{x^{s}}{s}\,\mathrm{d}s\ , \qquad \ell''(x) := \frac{1}{2\pi i}\int_{\mathcal{H}(r)} L(s,z,\omega)x^{s-1}\,\mathrm{d}s. \tag{15.5}$$

By the expansion (14.9) it follows, for $s \in \mathcal{H}(r)$, that

$$L(s,z) \ll \frac{1}{|s-1|} = \frac{1}{r}.$$

Consequently, a trivial estimate gives

$$\ell''(x) \ll \log x. \tag{15.6}$$

In view to (13.8) formula (14.9) implies, for $s \in \mathcal{H}(r)$,

$$G(s,z)\frac{((s-1)\zeta(s))^z}{s} = \sum_{k=0}^{\infty} g_k(z)(s-1)^k$$

with

$$g_k(z) := \frac{1}{k!}\sum_{h+j=k}\binom{k}{j}\gamma_j(z)\left[\frac{\mathrm{d}^h}{\mathrm{d}s^h}L(s,z,\omega)\zeta(s)^{-z}\right]_{s=1} = \Gamma(z-k)\lambda_k(z),$$

where

$$g_k(z) = \frac{k!}{2\pi i}\oint G(s,u)\frac{((s-1)\zeta(s))^u}{s(u-z)^{k+1}} \ll \delta^{-k}, \tag{15.7}$$

since the integrand is analytic in $|s-1| \le \delta$. Therefore, we have on the truncated Hankel contour $\mathcal{H}(r)$

$$G(s,z)\frac{((s-1)\zeta(s))^z}{s} = \sum_{k=0}^{N} g_k(z)(s-1)^k + O\left(\left(\frac{|s-1|}{\delta}\right)^{N+1}\right).$$

Substituting this in (15.5) gives

$$\ell'(x) = \sum_{k=0}^{N} g_k(z)\,\frac{1}{2\pi i}\int_{\mathcal{H}(r)} x^s(s-1)^{k-z}\,\mathrm{d}s + O(\delta^{-N}R(x)), \tag{15.8}$$

where

$$\begin{aligned} R(x) &:= \int_{\mathcal{H}} |x^s(s-1)^{N+1-z}||\,\mathrm{d}s| \\ &\ll \int_{1-\frac{1}{2}\delta}^{1-r} (1-\sigma)^{N+1-Rez} x^\sigma \,\mathrm{d}\sigma + x^{1+r} r^{N+2-Rez}. \end{aligned}$$

Since the Gamma-function interpolates the factorials we have in view to (5.2)

$$\Gamma(n+1) = n! = \exp(n\log n - n + O(\log n)$$

for $n \in \mathbb{N}$. Therefore, putting $u = (1-\sigma)\log x$, it follows that

$$\begin{aligned} R(x) &\ll x(\log x)^{Rez-N-2}\left(\int_{\frac{1}{2}}^{\infty} u^{N+1-Rez}\exp(-u)\,\mathrm{d}u + 2^{-N}\right) \\ &\ll x(\log x)^{Rez-N-2}\Gamma(N+3) \ll x(\log x)^{Rez-1}\left(\frac{BN+1}{\log x}\right)^{N+1}; \end{aligned}$$

here, and in the sequel, B denotes some positive absolute constant, not necessarily always the same. In order to simplify formula (15.8), changing the variable by $w = (s-1)\log x$, and applying Lemma 15.2, yields

$$\begin{aligned} \frac{1}{2\pi i}\int_{\mathcal{H}(r)} x^s(s-1)^{k-z}\,\mathrm{d}s &= \frac{x(\log x)^{z-1-k}}{2\pi i}\int_{\mathcal{H}(\frac{\delta}{2}\log x)} w^{k-z}\exp(w)\,\mathrm{d}w \\ &= x(\log x)^{z-1-k}\left(\frac{1}{\Gamma(z-k)} + O\left((Bk+1)^k x^{-\frac{\delta}{4}}\right)\right). \end{aligned}$$

Therefore, we get for the main term in (15.8)

$$x(\log x)^{z-1}\left(\sum_{k=0}^{N}\frac{\lambda_k(z)}{(\log x)^k} + E_N\right),$$

where

$$\begin{aligned} E_N &\ll x^{-\frac{\delta}{4}}\sum_{k=0}^{N}|g_k(z)|\left(\frac{Bk+1}{\log x}\right)^k \ll x^{-\frac{\delta}{4}}B^N\sum_{k=0}^{N}k!\left(\frac{5}{\delta\log x}\right)^k \\ &\ll x^{-\frac{\delta}{4}}\left(\frac{B}{\log x}\right)^N\sum_{k=0}^{N}\frac{N!}{(N-k)!}\left(\frac{\log x}{B}\right)^{N-k} \\ &\ll x^{-\frac{\delta}{4}}N!\left(\frac{B}{\log x}\right)^N \ll \left(\frac{BN+1}{\log x}\right)^{N+1}; \end{aligned}$$

here we used the weak Stirling formula (5.2) and (15.7). This lengthy calculation leads in (15.8) to

$$\ell'(x) = x(\log x)^{z-1}\left(\sum_{k=0}^{N}\frac{\lambda_k(z)}{(\log x)^k} + O\left(\left(\frac{BN+1}{\log x}\right)^{N+1}\right)\right). \tag{15.9}$$

Now we are able to finish the proof. Applying formula (15.9) with $x+h$ and x, where $0 < h < \frac{x}{2}$, leads to

$$\int_x^{x+h}\sum_{n\leq u} z^{\omega(n)}\,\mathrm{d}u = \ell(x+h) - \ell(x) + O\left(x^2\exp(-B\sqrt{\log x})\right).$$

By (15.6)

$$\ell(x+h) - \ell(x) = h\ell'(x) + h^2\int_0^1 (1-u)\ell''(x+uh)\,\mathrm{d}u = h\ell'(x) + O(h^2\log x),$$

which leads to

$$\begin{aligned}\sum_{n\leq x} z^{\omega(n)} &= \frac{1}{h}\int_x^{x+h}\sum_{n\leq u} z^{\omega(n)}\,\mathrm{d}u + O\left(\frac{L}{h}\right)\\ &= \ell'(x) + O\left(\frac{x^2}{h}\exp(-B\sqrt{\log x}) + h\log x + \frac{L}{h}\right),\end{aligned}$$

where

$$L := \int_x^{x+h}\left|\sum_{n\leq x} z^{\omega(n)} - \sum_{n\leq u} z^{\omega(n)}\right|\mathrm{d}u.$$

In view to $|z^{\omega(n)}| \leq 1$, we get

$$L \leq \int_x^{x+h}\sum_{x<n\leq u} 1\,\mathrm{d}u = \int_x^{x+h}(u-x)\,\mathrm{d}u + O(h) \ll h^2.$$

Thus, choosing $h := x\exp(-B\sqrt{\log x})$, we obtain

$$\sum_{n\leq x} z^{\omega(n)}\,\mathrm{d}u = \ell'(x) + O\left(\frac{x^2}{h}\exp(-B\sqrt{\log x}) + h\log x\right).$$

Now the assertion of the theorem follows from (15.9). •

Exercise 15.1 *Prove a similar result as in Theorem 15.1 for* $\sum_{n\leq x} z^{\Omega(n)}$.

In the last chapter we give an application of the Selberg-Delange method on the frequency of integers n with $\omega(n) = k$. This returns us to Gauss' conjecture with which we started in the introduction.

Chapter 16

The prime number theorem

As a generalization of the prime counting function $\pi(x)$ we define, for $k \in \mathbb{N}$,

$$\pi_k(x) = \sharp\{n \leq x \,:\, \omega(n) = k\}.$$

Note that the influence of prime powers $p^j \leq x$ for fixed k is small. For example, if $k = 1$, then

$$\pi_1(x) = \pi(x) + \sharp\{p^j \leq x \,:\, j \geq 2\} = \pi(x) + O(x^{\frac{1}{2}}).$$

Therefore, $\pi_k(x)$ counts asymptotically the number of integers n which are the product of exactly k distinct prime numbers. Now we shall prove

Theorem 16.1 (Sathe, 1953/1954; Selberg, 1954) *We have, uniformly for sufficiently large x and $1 \leq k \leq \log\log x$,*

$$\pi_k(x) = \frac{x}{\log x}\frac{(\log\log x)^{k-1}}{(k-1)!}\left(\lambda\left(\frac{k-1}{\log\log x}\right) + O\left(\frac{k}{(\log\log x)^2}\right)\right),$$

where

$$\lambda(z) := \frac{1}{\Gamma(z+1)}\prod_p \left(1 + \frac{z}{p-1}\right)\left(1 - \frac{1}{p}\right)^z.$$

In particular, the asymptotic formula of the theorem yields

$$\pi_k(x) = (1 + o(1))\frac{x}{\log x}\frac{(\log\log x)^{k-1}}{(k-1)!} = (1 + o(1))\pi(x)\frac{(\log\log x)^{k-1}}{(k-1)!}.$$

As we mentioned in the introduction, this result was first conjectured by Gauss. The first proof was given by Landau [20]; see also [14] where it is proved as a consequence of the prime number theorem by induction on k. Our proof is based on Theorem 15.1.

Proof of Theorem 16.1. Obviously,

$$\sum_{k\geq 0}\pi_k(x)z^k=\sum_{n\leq x}z^{\omega(n)}.$$

Consequently, $\pi_k(x)$ equals, up to a small error, the coefficient of z^k in the main term of the asymptotic formula of Theorem 15.1. Therefore, we obtain

$$\pi_k(x)=\frac{x}{\log x}\frac{1}{k!}\frac{\mathrm{d}^k}{\mathrm{d}z^k}\Bigg[(\log x)^z\lambda_k(z)+(\log x)^zE(x)\Bigg]_{z=0},$$

where $E(x)\ll(\log x)^{-1}$. By CAUCHY's formula, it turns out that

$$\frac{\mathrm{d}^k}{\mathrm{d}z^k}\Bigg[(\log x)^zE(x)\Bigg]_{z=0}=\frac{1}{2\pi i}\oint_{|z|=r}\frac{(\log x)^zE(x)}{z^{k+1}}\,\mathrm{d}z\ll\frac{\log\log x}{k!}E(x)$$

by putting $r=\frac{k}{\log\log x}$. Furthermore,

$$\frac{\mathrm{d}^k}{\mathrm{d}z^k}\Bigg[(\log x)^z\lambda_k(z)\Bigg]_{z=0}=\frac{1}{2\pi i}\oint_{|z|=r}\frac{(\log x)^z\lambda_k(z)}{z^{k+1}}\,\mathrm{d}z.$$

Recall that $\lambda_k(0)=0$. Thus, using the functional equation for the Gamma-function $\Gamma(z+1)=z\Gamma(z)$, we may write $\lambda_k(z)=z\lambda(z)$. Then we can replace the integrand above by

$$\lambda(z)\sum_{j=0}^{\infty}\frac{1}{j!}(\log\log x)^jz^{j-k},$$

which gives, by the calculus of residues, the asymptotic formula of the theorem. •

The case $k=1$ yields the celebrated prime number theorem (5.6).

Corollary 16.2 (Prime number theorem) *As* $x\to\infty$,

$$\pi(x)=\frac{x}{\log x}+O\left(\frac{x}{\log x(\log\log x)^2}\right).$$

The prime number theorem allows a plenty of interesting speculations in number theory. In view to

$$\pi(x)=(1+o(1))\frac{x}{\log x}=\int_2^x\frac{\mathrm{d}u}{\log u}+\text{error},$$

we can build up a probabilistic model for primality, in which an integer n is prime with probability $\frac{1}{\log n}$. This idea dates back to CRAMÉR [2], who discussed with his model the still open conjecture that there is always a prime number in between two consecutive squares

$$n^2 < p < (n+1)^2.$$

We give an easier example. When $a, b > 1$ are integers, then

$$2^{a \cdot b} - 1 = (2^a - 1) \cdot (2^{a(b-1)} + \ldots + 2^a + 1).$$

But if the composite integer ab is replaced by a prime number, then the situation is different; for example

$$2^3 - 1 = 7, \quad 2^5 - 1 = 31, \quad 2^7 - 1 = 127, \quad 2^{13} - 1 = 8191 \ \in \mathbb{P}.$$

For a prime p the MERSENNE **number** M_p is defined by

$$M_p = 2^p - 1.$$

The LUCAS-LEHMER *algorithm*,

$$s := 4, \ \textsf{for} \ \ i \ \textsf{from} \ \ 3 \ \textsf{to} \ \ p \ \textsf{do} \ s := s^2 - 2 \bmod \ (2^p - 1),$$

returns the value $s = 0$ if and only if M_p is prime; for a proof of this deep theorem see [13], §XV.5. Iteration yields

$$s = 4 \quad \mapsto \quad 14 = 2 \cdot \mathbf{7} \quad \mapsto \quad 194 \quad \mapsto \quad 37\,634 = 2 \cdot \mathbf{31} \cdot 607,$$

which gives the first MERSENNE primes. Not all prime p lead to prime M_p; for example $M_{11} = 23 \cdot 89$. Meanwhile, 39 MERSENNE primes are known; recently CAMERON discovered by intensive computer calculations that

$$M_{13\,466\,917} = 2^{13\,466\,917} - 1$$

is prime. This largest known prime number exceeds the number of atoms in the universe, and has more than four million digits! It is an open question whether there exist infinitely many MERSENNE primes or not. In view to our probabilistic model the probability that M_p is prime equals

$$\mathbf{P}(M_p \in \mathbb{P}) \approx \frac{1}{\log M_p} \approx \frac{1}{p \log 2}.$$

Therefore, the expectation value for the number of **Mersenne** primes is

$$\mathbf{E}(\sharp M_p \in \mathbb{P}) = \sum_p \mathbf{P}(M_p \in \mathbb{P}) \approx \frac{1}{\log 2} \sum_p \frac{1}{p},$$

which diverges by Corollary 5.4. Thus we expect that there are infinitely many MERSENNE primes.

Exercise 16.1 *For a non-negative integer k the kth* FERMAT **number** *is defined by* $F_k = 2^{2^k} + 1$.

(i) Calculate the first seven FERMAT *numbers.*

(ii) Do you think that there are infinitely many or only finitely many FERMAT *primes?*

(The FERMAT *numbers are of special interest for the problem of the construction of the regular polygon of n sides; see [13], § V.8.)*

However, the probabilistic model has also limits; see [22], §3. For many problems in number theory a deeper knowledge on the prime number distribution is needed than that what is known yet. One can show that

$$\pi(x) = \int_2^x \frac{\mathrm{d}u}{\log u} + O(x^{\theta+\varepsilon}) \qquad \Longleftrightarrow \qquad \zeta(s) \neq 0 \quad \text{in} \quad \sigma > \theta$$

(for a proof see [30], §2.4). Since there are zeros of $\zeta(s)$ on the critical line $\sigma = \frac{1}{2}$, RIEMANN's hypothesis states that the prime numbers are distributed as uniformly as possible!

It is known that $\zeta(s)$ has infinitely many zeros in the strip $0 < \sigma < 1$. Many computations were done to find a counter example to the RIEMANN hypothesis, that is to find a zero in the half plane $\sigma > \frac{1}{2}$. However the *first* 1 500 000 001 zeros lie without exception on $\sigma = \frac{1}{2}$. Further, it is known that at least 40 percent have the predicted distribution; for more details we refer the interested reader to [23].

We conclude with a probabilistic interpretation of RIEMANN's hypothesis due to DENJOY [4]. If and only if the RIEMANN hypothesis is true, i.e. that $\zeta(s)$ is free of zeros in $\sigma > \frac{1}{2}$, the reciprocal

$$\frac{1}{\zeta(s)} = \prod_p \left(1 - \frac{1}{p^s}\right)$$

has an analytic continuation to the half plane $\sigma > \frac{1}{2}$. This turns out to be equivalent to the estimate

$$(16.1) \qquad \sum_{n \leq x} \mu(n) \ll x^{\frac{1}{2}+\varepsilon}.$$

Now assume that the values $\mu(n)$ behave like independent random variables X_n with

$$\mathbf{P}(X_n = +1) = \mathbf{P}(X_n = -1) = \frac{1}{2}.$$

Then

$$Z_0 := 0 \quad \text{und} \quad Z_n := \sum_{j=1}^{n} X_j$$

defines a random walk, and formula (3.2) yields

$$\mathbf{P}\left(\left|\sum_{n\leq x} X_n\right| \leq cn^{\frac{1}{2}}\right) \longrightarrow \mathbf{\Phi}(c).$$

From that point of view the validity of formula (16.1), and therefore the truth of RIEMANN's hypothesis seems highly probable.

"It is evident that the primes are randomly distributed
but, unfortunately, we don't know what 'random' means."
R.C. VAUGHAN

Bibliography

[1] CESARO, *Démonstration élémentaire et généralisation de quelques théorèmes de M. Berger*, Mathesis **1** (1881), 99-102

[2] H. CRAMÉR, *On the order of magnitude of the difference between consecutive primes*, Acta Arith. **2** (1936), 23-46

[3] H. DELANGE, *Sur de formules de Atle Selberg*, Acta Arith. **19** (1971), 105-146

[4] A. DENJOY, *L'Hypothèse de Riemann sur la distribution des zéros de $\zeta(s)$, reliée à la théorie des probabilites*, C.R.Acad. Sci. Paris **192** (1931), 656-658

[5] J.-M. DESHOUILLERS, F. DRESS, G. TENENBAUM, *Lois de rèpartition des diviseurs 1*, Acta Arith. **23** (1979), 273-285

[6] P.D.T.A. ELLIOTT, *Probabilistic Number Theory I, II*, Springer 1979

[7] P. ERDÖS, M. KAC, *On the Gaussian law of errors in the theory of additive functions*, Proc. Nat. Acad. Sci. USA **25** (1939), 206-207

[8] P. ERDÖS, A. WINTNER, *Additive arithmetical functions and statistical independence*, Amer. J. Math. **61** (1939), 713-721

[9] W.J. FELLER, *An introduction to probability theory and its applications*, John Wiley 1950

[10] C.F. GAUSS, *Collected Works*, Teubner 1917

[11] G. HALÁSZ, *Über die Mittelwerte multiplikativer zahlentheoretischer Funktionen*, Acta Math. Acad. Sci. Hung. **19** (1968), 365-403

[12] H. HALBERSTAM, K.F. ROTH, *Sequences*, Oxford 1966

[13] G.H. HARDY, S. RAMANUJAN, *The normal number of prime factors of a number n*, Quart. J. Math. (Oxford) **48** (1917), 76-92

[14] G.H. Hardy, E.M. Wright, *An introduction to the theory of numbers*, Oxford 1938

[15] E. Hlawka, *Theorie der Gleichverteilung*, BIB Mannheim 1979

[16] J. Jacod, P. Protter, *Probability Essentials*, Springer 2000

[17] M. Kac, *Statistical Independence in Probability, Analysis and Number Theory*, Carus Mathematical Monographs, John Wiley 1959

[18] J. Kubilius, *Probabilistic Methods in the Theory of Numbers*, AMS Monographs 1964

[19] J. Kubilius, *Estimation of the central moment for strongly additive arithmetic functions*, Lietovsk. Mat. Sb. **23** (1983), 110-117 (in Russian)

[20] E. Landau, *Handbuch der Lehre von der Verteilung der Primzahlen*, Teubner 1909

[21] S. Lang, *Complex Analysis*, Springer 1977

[22] M. Mendès France, G. Tenenbaum, *The Prime Numbers and their distribution*, AMS 2000

[23] W. Narkiewicz, *The development of prime number theory*, Springer 2000

[24] A. Rényi, P. Turán, *On a theorem of Erdös-Kac*, Acta Arith. **4** (1958), 71-84

[25] L.G. Sathe, *On a problem of Hardy and Ramanujan on the distribution of integers having a given number of prime factors I, II*, J. Indian Math. Soc. **17** (1953), 63-141; **18** (1954), 27-81

[26] A. Schinzel, *Generalisation of a theorem of B.S.K.R. Somayajulu on the Euler's function $\varphi(n)$*, Ganita **5** (1954), 123-128

[27] M.R. Schroeder, *Number theory in science and communication*, Springer 1997, 3rd ed.

[28] W. Schwarz, J. Spilker, *Arithmetical Functions*, London Math. Soc. Lecture Notes 184, Cambridge 1994

[29] A. Selberg, *Note on a paper by L.G. Sathe*, J. Indian Math. Soc. B. **18** (1954), 83-87

[30] G. Tenenbaum, *Introduction to analytic and probabilistic number theory*, Cambridge University Press 1995

[31] P. Turán, *On a theorem of Hardy and Ramanujan*, J. London Math. Soc. **9** (1934), 274-276

[32] H. Weyl, *Über die Gleichverteilung von Zahlen mod Eins*, Math. Annalen **77** (1916), 313-352

[33] E. Wirsing, *Das asymptotische Verhalten von Summen über multiplikative Funktionen II*, Acta Math. Acad. Sci. Hung. **18**, 411-467