Analytic Number Theory

Walter Bergweiler

CAU Kiel

Summer term 2018

Version of July 6, 2018

Contents

An overview

Number theory can be seen as a branch of mathematics where one mainly studies the integers and their properties. It is perhaps surprising that when doing so methods from analysis, which is concerned with real and complex numbers, are often useful. Analytic number theory deals with such applications of analysis to number theory.

We begin by studying the connection between a sequence (a_n) of integers and the power series

$$\sum_{n=0}^{\infty} a_n x^n,$$

which is called the generating function of the sequence. The first example we consider here are the Fibonacci numbers

$$1,\ 1,\ 2,\ 3,\ 5,\ 8,\ 13,\ \ldots.$$

Next we consider in how many ways a natural number n can be written as a sum of smaller natural numbers. We obtain information about the number $p(n)$ of such "partitions" by studying the associated generating function.

The main part of the lecture will then be concerned with prime numbers. In particular, we shall prove that the number $\pi(n)$ of prime numbers less than or equal to n satisfies

$$\lim_{n\to\infty} \frac{\pi(n)\cdot\log n}{n} = 1.$$

This result is known as the prime number theorem. The proof will use the Riemann ζ-function which for $s \in \mathbb{C}$ with $\operatorname{Re} s > 1$ is defined by

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

In particular, we will see that the prime numbers are closely related to the zeros of ζ.

1 Generating functions

We begin with an example. The Fibonacci sequence $(F_n)_{n\geq 0}$ is defined by $F_0 = 0$, $F_1 = 1$ and $F_n = F_{n-1} + F_{n-2}$ for $n \geq 2$. Thus $F_2 = 1$, $F_3 = 2$, $F_4 = 3$, $F_5 = 5$, $F_6 = 8$, etc.

Without worrying about convergence we consider the power series

$$f(x) = \sum_{n=0}^{\infty} F_n x^n.$$

Then

$$xf(x) = \sum_{n=0}^{\infty} F_n x^{n+1} = \sum_{n=1}^{\infty} F_{n-1} x^n$$

and

$$x^2 f(x) = \sum_{n=0}^{\infty} F_n x^{n+2} = \sum_{n=2}^{\infty} F_{n-2} x^n.$$

Thus

$$\begin{aligned}(1-x-x^2)f(x) &= f(x) - xf(x) - x^2 f(x)\\ &= F_0 + F_1 x - F_0 x + \sum_{n=2}^{\infty}(F_n - F_{n-1} - F_{n-2})x^2\\ &= x.\end{aligned}$$

So far this computation was purely formal. But now we can argue that the function f given by

$$f(x) = \frac{x}{1-x-x^2}$$

has a power series expansion which converges in some disk around 0 and that the above computation shows that its coefficients are given by the Fibonacci numbers.

To compute the coefficients, we note that

$$1 - x - x^2 = -(x-a)(x-b)$$

with

$$a = \frac{-1+\sqrt{5}}{2} \quad \text{and} \quad b = \frac{-1-\sqrt{5}}{2}.$$

Thus we have the partial fraction decomposition

$$f(x) = \frac{\alpha}{x-a} + \frac{\beta}{x-b}$$

with

$$\alpha = \lim_{x\to a}(x-a)f(x) = \lim_{x\to a}\frac{x}{-(x-b)} = \frac{a}{b-a} = -\frac{a}{\sqrt{5}}$$

and

$$\beta = \lim_{x\to b}(x-b)f(x) = \frac{b}{a-b} = \frac{b}{\sqrt{5}}.$$

It follows that if $|x| < \min\{|a|, |b|\}$, then

$$\begin{aligned}f(x) &= -\frac{\alpha}{a}\cdot\frac{1}{1-\frac{x}{a}} - \frac{\beta}{b}\cdot\frac{1}{1-\frac{x}{b}}\\ &= -\frac{\alpha}{a}\sum_{n=0}^{\infty}\left(\frac{x}{a}\right)^n - \frac{\beta}{b}\sum_{n=0}^{\infty}\left(\frac{x}{b}\right)^n\\ &= \sum_{n=0}^{\infty}\left(-\frac{\alpha}{a}\cdot\frac{1}{a^n} - \frac{\beta}{b}\cdot\frac{1}{b^n}\right)x^n.\end{aligned}$$

Thus

$$F_n = -\frac{\alpha}{a}\cdot\frac{1}{a^n} - \frac{\beta}{b}\cdot\frac{1}{b^n}$$

for all $n \in \mathbb{N}_0 = \{0, 1, 2, \dots\}$. Since

$$\frac{1}{a} = \frac{2}{\sqrt{5}-1} = \frac{2(\sqrt{5}+1)}{(\sqrt{5}-1)(\sqrt{5}+1)} = \frac{\sqrt{5}+1}{2}$$

and

$$\frac{1}{b} = \frac{-2}{\sqrt{5}+1} = \frac{1-\sqrt{5}}{2}$$

we conclude that

$$F_n = -\frac{\alpha}{a}\left(\frac{1}{a}\right)^n - \frac{\beta}{b}\left(\frac{1}{b}\right)^n = \frac{1}{\sqrt{5}}\left(\frac{\sqrt{5}+1}{2}\right)^n - \frac{1}{\sqrt{5}}\left(\frac{1-\sqrt{5}}{2}\right)^n.$$

This formula for F_n is named after Moivre and Binet.

Once the formula is known, it is not difficult to prove it by induction. The point is, however, to find the formula, and this is where the approach using power series is useful.

It is remarkable that while the F_n are all integers, the above formula for them involves the irrational number $\sqrt{5}$. In fact, there are similar integer recursions where the corresponding formula involves complex numbers. This happens when instead of the polynomial $1 - x - x^2$ in the above argument one has a polynomial with non-real zeros.

Note that

$$\left|\frac{1}{\sqrt{5}}\left(\frac{1-\sqrt{5}}{2}\right)^n\right| < \frac{1}{2}$$

for all $n \in \mathbb{N}$. This shows that F_n is the integer closest to the term

$$\frac{1}{\sqrt{5}}\left(\frac{\sqrt{5}+1}{2}\right)^n.$$

Since

$$\left(\frac{1-\sqrt{5}}{2}\right)^n \to 0$$

as $n \to \infty$, the above term is actually very close to an integer for large n. For example, $F_{10} = 55$ and

$$\frac{1}{\sqrt{5}}\left(\frac{\sqrt{5}+1}{2}\right)^{10} = 55.0036\dots.$$

Definition 1.1. Let $(a_n)_{n\geq 0}$ be a sequence of real (or complex) numbers. Then the power series

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

is called the *generating function* of (a_n).

The idea is to obtain information about the sequence (a_n) from properties of the function f. In general one cannot expect to obtain an explicit formula as we did in the case of the Fibonacci numbers. However, for example one may be able to obtain estimates or information about the asymptotic behavior. To illustrate that, we note that the radius of convergence r of the generating function of the Fibonacci numbers is given by

$$r = \min\{|a|, |b|\} = |a| = \frac{\sqrt{5}-1}{2}.$$

On the other hand, we have

$$\frac{1}{r} = \limsup_{n\to\infty} \sqrt[n]{F_n}.$$

For a given $\varepsilon > 0$ we thus obtain

$$F_n \leq \left(\frac{1}{r} + \varepsilon\right)^n = \left(\frac{\sqrt{5}+1}{2} + \varepsilon\right)^n$$

for large n.

We consider some further examples.

Example 1.1. Let $m \in \mathbb{N}$ and

$$f(x) = (x + x^2 + x^3 + x^4 + x^5 + x^6)^m = \sum_{n=m}^{6m} p_{nm} x^n.$$

Then

$$f(x) = \left(\sum_{i_1=1}^{6} x^{i_1}\right) \cdot \left(\sum_{i_2=1}^{6} x^{i_2}\right) \cdot \ldots \cdot \left(\sum_{i_m=1}^{6} x^{i_m}\right) = \sum_{i_1,\ldots,i_m=1}^{6} x^{i_1+i_2+\ldots+i_m}.$$

Thus

$$p_{nm} = \operatorname{card}\left\{(i_1,\ldots,i_m) \in \{1,\ldots,6\}^m : \sum_{j=1}^{m} i_j = n\right\}.$$

Here and in the following $\operatorname{card} X$ denotes the cardinality (i.e., the number of elements) of a finite set X.

The interpretation is as follows. We throw m dice. Overall there are 6^m possible outcomes, and p_{nm} is the number of outcomes for which the number of points is n.

We consider a curious application of this idea. In order to do so, note that

$$x + x^2 + \ldots + x^6 = x\frac{x^6-1}{x-1} = x(x+1)(x^2+x+1)(x^2-x+1)$$

and thus

$$\begin{aligned}\left(x + x^2 + \ldots + x^6\right)^2 &= \left(x(x+1)(x^2+x+1)\right) \cdot \left(x(x+1)(x^2+x+1)(x^2-x+1)^2\right) \\ &= \left(x + 2x^2 + 2x^3 + x^4\right) \cdot \left(x + x^3 + x^4 + x^5 + x^6 + x^8\right).\end{aligned}$$

The interpretation of this equation is the following. Take a (regular) dice, but replace the number 5 and 6 on it by 2 and 3. So it will have the numbers 1 and 4 once and the numbers 2 and 3 twice. Take a second (regular) dice and replace the number 2 by 8. So it will have the numbers 1, 3, 4, 5, 6 and 8 once.

Then throwing these two modified dice will lead to the same result, as throwing two regular dice, i.e., given n we will obtain n points with the same probability.

Example 1.2. We consider in how many ways an amount of money can be split into coins. For simplicity we first consider the case where there are only coins of 1, 2 and 3 cents, but the method described can be extended to the general case.

The generating function is given by

$$f(x) = \frac{1}{(1-x)(1-x^2)(1-x^3)},$$

since (for $|x| < 1$) we have

$$f(x) = \left(\sum_{k=0}^{\infty} x^k\right) \cdot \left(\sum_{l=0}^{\infty} x^{2l}\right) \cdot \left(\sum_{m=0}^{\infty} x^{3m}\right) = \sum_{k,l,m=0}^{\infty} x^{k+2l+3m} = \sum_{n=0}^{\infty} c_n x^n$$

where

$$c_n = \operatorname{card}\big\{(k,l,m) \in \mathbb{N}_0^3 \colon k + 2l + 3m = n\big\}$$

is the number of possibilities to split n cents into coins of 1, 2 and 3 cents. We have

$$f(x) = \frac{1}{(1-x)^3(1+x)(1+x+x^2)}.$$

Performing a partial fraction decomposition and combining the terms with only simple poles in a suitable way we find that

$$f(x) = \frac{1}{6} \cdot \frac{1}{(1-x)^3} + \frac{1}{4} \cdot \frac{1}{(1-x)^2} + \frac{1}{4} \cdot \frac{1}{1-x^2} + \frac{1}{3} \cdot \frac{1}{1-x^3}.$$

Since

$$\frac{1}{(1-x)^2} = \frac{d}{dx}\left(\frac{1}{1-x}\right) = \frac{d}{dx}\left(\sum_{n=0}^{\infty} x^n\right) = \sum_{n=0}^{\infty}(n+1)x^n$$

and

$$\frac{1}{(1-x)^3} = \frac{1}{2} \cdot \frac{d}{dx}\left(\frac{1}{(1-x)^2}\right) = \sum_{n=0}^{\infty} \frac{(n+2)(n+1)}{2} x^n$$

we conclude that

$$c_n = \frac{(n+2)(n+1)}{12} + \frac{n+1}{4} + \begin{cases} \frac{1}{4}, & \text{if } 2 \mid n \text{ and } 3 \nmid n, \\ \frac{1}{3}, & \text{if } 3 \mid n \text{ and } 2 \nmid n, \\ \frac{1}{4} + \frac{1}{3}, & \text{if } 6 \mid n, \\ 0, & \text{otherwise.} \end{cases}$$

Here $m \mid n$, with integers m and n, means that m is a divisor of n; that is, $n = km$ for some $k \in \mathbb{Z}$.

Since

$$\frac{(n+2)(n+1)}{12} + \frac{n+1}{4} = \frac{n^2}{12} + \frac{n}{2} + \frac{5}{12}$$

and $c_n \in \mathbb{N}$, the above formula can also be written as

$$c_n = \left\lfloor \frac{n^2}{12} + \frac{n}{2} + 1 \right\rfloor \quad \text{or} \quad c_n = \left\lceil \frac{n^2}{12} + \frac{n}{2} + \frac{5}{12} \right\rceil .$$

Here, for $x \in \mathbb{R}$, we put

$$\lfloor x \rfloor = \max \{n \in \mathbb{Z} \colon n \leq x\} \quad \text{and} \quad \lceil x \rceil = \min \{n \in \mathbb{Z} \colon n \geq x\} .$$

Now we discuss the general case where we have k coins with denominations $w_1, \ldots, w_k$. Then the generating function takes the form

$$f(x) = \frac{1}{(1-x^{w_1}) \cdot (1-x^{w_2}) \cdot \ldots \cdot (1-x^{w_k})}.$$

It will, in general, be hopeless to write down the partial fraction decomposition of f, and thus the coefficients c_n in the expression

$$f(x) = \sum_{n=0}^{\infty} c_n x^n$$

explicitly. But we will be able to determine the asymptotic behaviors of c_n as $n \to \infty$.

For simplicity we assume that $w_1 = 1$. Then the denominator of f has a zero of order k at 1, but all other zeros have order less than k. Denoting by $x_1 = 1, x_2, \ldots, x_t$ the (possibly complex) zeros of the denominator of f we obtain a representation

$$f(x) = \sum_{j=1}^{k} \frac{\beta_{1j}}{(1-x)^j} + \sum_{s=2}^{t} \sum_{j=1}^{m_s} \frac{\beta_{sj}}{\left(1 - \dfrac{x}{x_s}\right)^j}$$

where $m_s \in \mathbb{N}$ and $m_s \leq k-1$ for $2 \leq s \leq t$ and $\beta_{sj} \in \mathbb{C}$ for $1 \leq s \leq t$ and $1 \leq j \leq m_s$, with $m_1 = k$.

For $|y| < 1$ and $j \in \mathbb{N}$ we have

$$\begin{aligned} \frac{1}{(1-y)^j} &= \frac{1}{(j-1)!} \cdot \frac{d^{j-1}}{dy^{j-1}} \left(\frac{1}{1-y} \right) \\ &= \frac{1}{(j-1)!} \sum_{n=0}^{\infty} (n+j)(n+j-1) \cdots (n+1) y^n . \end{aligned}$$

This implies that

$$c_n = \frac{\beta_{1k}}{(k-1)!}(n+k)(n+k-1)\cdots(n+1) + d_n \ ,$$

where d_n is a sum of terms of the form

$$\frac{\beta_{sj}}{(j-1)!}(n+j)(n+j-1)\cdots(n+1)\left(\frac{1}{x_s}\right)^n$$

with $j \leq k-1$. Since $|x_s| = 1$ for all s we obtain

$$|d_n| \leq An^{k-1}$$

with some constant A, for all $n \in \mathbb{N}$. We deduce that

$$c_n = \frac{\beta_{1k}}{(k-1)!}n^k + e_n \quad \text{with} \quad |e_n| \leq Bn^{k-1}$$

for some constant B. Here

$$\beta_{1k} = \lim_{x\to 1}(1-x)^k f(x) = \lim_{x\to 1}\frac{(1-x)(1-x)\cdots(1-x)}{(1-x^{w_1})(1-x^{w_2})\cdots(1-x^{w_k})}.$$

Since

$$\lim_{x\to 1}\frac{1-x}{1-x^w} = \lim_{x\to 1}\frac{-1}{-wx^{w-1}} = \frac{1}{w}$$

for $w \in \mathbb{R}\setminus\{0\}$ by l'Hospital's rule, we find that

$$\beta_{1k} = \frac{1}{w_1 \cdot w_2 \cdot \ldots \cdot w_k}.$$

Altogether we thus have

$$c_n = \frac{1}{w_1 \cdot w_2 \cdot \ldots \cdot w_k \cdot (k-1)!} \cdot n^k + e_n$$

where $|e_n| \leq Bn^{k-1}$.

A convenient way to write such asymptotic relations is given by the Landau symbols. For sequences (a_n) and (b_n), with $b_n > 0$ for large n, we write

$$a_n = \mathcal{O}(b_n) \quad \text{as} \quad n \to \infty \quad \text{if} \quad \limsup_{n\to\infty}\frac{|a_n|}{b_n} < \infty$$

and

$$a_n = o(b_n) \quad \text{as} \quad n \to \infty \quad \text{if} \quad \lim_{n\to\infty}\frac{a_n}{b_n} = 0.$$

The notation $a_n = \mathcal{O}(b_n)$ there means that there exist $n_0 \in \mathbb{N}$ and $K \geq 0$ with $|a_n| \leq Kb_n$ for $n \geq n_0$.

With

$$C = \frac{1}{w_1 \cdot w_2 \cdot \ldots \cdot w_k \cdot (k-1)!}$$

the formula for the coefficients c_n can thus be written in the form

$$c_n = C \cdot n^k + \mathcal{O}\left(n^{k-1}\right) = \left(C + \mathcal{O}\left(\frac{1}{n}\right)\right) n^k.$$

We also use the notation

$$a_n \sim b_n \quad \text{if} \quad \lim_{n\to\infty} \frac{a_n}{b_n} = 1.$$

Equivalently, $a_n = (1+o(1))b_n$ as $n \to \infty$. We thus have

$$c_n \sim C \cdot n^k \quad \text{as} \quad n \to \infty.$$

Example 1.3. We consider the following question. Does there exist a partition of $\mathbb{N}_0$ into two disjoint subsets A and B such that for every $n \in \mathbb{N}$ we have

$$\operatorname{card}\big\{(a,a') \in A^2 \colon a + a' = n,\ a \neq a'\big\} = \operatorname{card}\big\{(b,b') \in B^2 \colon b + b' = n,\ b \neq b'\big\}?$$

Let us assume that there exists such a partition with, say, $0 \in A$. Then $1 \in B$, since otherwise $1 = 0 + 1$ would be a representation of 1 as a sum of two elements in A, but no such representation would exist with elements in B.

Next we find that $2 \in B$, since otherwise $2 = 0 + 2$ would be a sum of elements in A, but no such representation would exist with elements in B. Analogous considerations lead to $3 \in A$, $4 \in B$, $5 \in A$, $6 \in A$, ... It is unclear, however, what the pattern is.

We consider, for $|x| < 1$, the generating functions

$$f(x) = \sum_{n\in A} x^n \quad \text{and} \quad g(x) = \sum_{n \in B} x^n.$$

Then

$$f(x) + g(x) = \sum_{n=0}^{\infty} x^n = \frac{1}{1-x}.$$

Put

$$c_n = \operatorname{card}\big\{(a,a') \in A^2 \colon a + a' = n,\ a \neq a'\big\}.$$

Then

$$f(x)^2 - f\big(x^2\big) = \sum_{n=0}^{\infty} c_n x^n.$$

Indeed, if we write

$$f(x)^2 = \sum_{n=0}^{\infty} d_n x^n,$$

then

$$d_n = \operatorname{card}\big\{(a, a') \in A^2,\ a + a' = n\big\}$$

and we have $d_n = c_n$ if n is not of the form $n = 2a$ for some $a \in A$, and $d_n = c_n + 1$ otherwise.

The same argument yields that

$$g(x)^2 - g\big(x^2\big) = \sum_{n=0}^{\infty} c_n x^n$$

and hence that

$$f(x)^2 - f\big(x^2\big) = g(x)^2 - g\big(x^2\big)\,.$$

This yields

$$\begin{aligned} f\big(x^2\big) - g\big(x^2\big) &= f(x)^2 - g(x)^2 \\ &= (f(x) - g(x)) \cdot (f(x) + g(x)) \\ &= (f(x) - g(x)) \cdot \frac{1}{1-x}. \end{aligned}$$

Thus

$$f(x) - g(x) = (1-x)\left(f\big(x^2\big) - g\big(x^2\big)\right).$$

Inductively we obtain

$$\begin{aligned} f(x) - g(x) &= (1-x)(1-x^2)\left(f\big(x^4\big) - g\big(x^4\big)\right) \\ &= \left(\prod_{m=0}^{n-1}\left(1 - x^{2^m}\right)\right)\left(f\big(x^{2^n}\big) - g\big(x^{2^n}\big)\right). \end{aligned}$$

Since

$$\lim_{n\to\infty} f\big(x^{2^n}\big) = \lim_{y\to 0} f(y) = f(0) = 1$$

and

$$\lim_{n\to\infty} g\big(x^{2^n}\big) = g(0) = 0$$

this yields

$$f(x) - g(x) = \lim_{n\to\infty} \prod_{m=0}^{n}\left(1 - x^{2^m}\right).$$

The product on the right hand side has the expansion

$$\prod_{m=0}^{n}\left(1 - x^{2^m}\right) = 1 + \sum_{0 \le m_1 < m_2 < \cdots < m_\ell \le n} (-1)^\ell x^{2^{m_1} + \ldots + 2^{m_\ell}}.$$

Now the binary (or base 2) representation of a natural number k has the form

$$k = \sum_{i=0}^{\infty} \varepsilon_i 2^i\ ,$$

with $\varepsilon_i \in \{0,1\}$ for all i and $\varepsilon_i = 1$ for only finitely many i. This implies that every $k \in \mathbb{N}$ has a unique representation

$$k = \sum_{j=1}^{\ell} 2^{m_j} \quad \text{with} \quad 0 \le m_1 < m_2 < \cdots < m_\ell.$$

Here $\ell = \ell(k)$ says how often the digit 1 occurs in the binary representation of k. Thus $\ell(k)$ is the number of i for which $\varepsilon_i = 1$.

Putting $\ell(0) = 0$ we thus obtain

$$\prod_{m=0}^{n} \left(1 - x^{2^m}\right) = \sum_{k=0}^{2^{n+1}-1} (-1)^{\ell(k)} x^k = \sum_{\substack{k=0 \\ \ell(k) \text{ even}}}^{2^{n+1}-1} x^k - \sum_{\substack{k=0 \\ \ell(k) \text{ odd}}}^{2^{n+1}-1} x^k.$$

We conclude that

$$A = \{k \in \mathbb{N}_0 \colon \ell(k) \text{ even}\} \quad \text{and} \quad B = \{k \in \mathbb{N}_0 \colon \ell(k) \text{ odd}\}$$

have the required properties. Moreover, up to interchanging them, they are the only sets with these properties.

At the end of the last example we considered a limit of the form

$$\lim_{k\to\infty} \prod_{j=1}^{k} a_j.$$

We will consider such limits more systematically now. A naive definition for the convergence of the infinite product $\prod_{j=1}^{\infty} a_j$ would be the existence of the limit $\lim_{k\to\infty} \prod_{j=1}^{k} a_j$. This definition would have two disadvantages:

- If $a_n = 0$ for some n, then $\lim_{k\to\infty} \prod_{j=1}^{k} a_j = 0$, regardless of the behavior of a_j as $j \to \infty$;
- $\lim_{k\to\infty} \prod_{j=1}^{k} a_j = 0$ is possible even if $a_j \ne 0$ for all j. For example, this happens for $a_j = j/(j+1)$.

Definition 1.2. Let (a_j) be a sequence in $\mathbb{C}$. Then $\prod_{j=1}^{\infty} a_j$ is called *convergent*, if there exists $N \in \mathbb{N}$ with $a_j \ne 0$ for $j \ge N$ and if $\lim_{k\to\infty} \prod_{j=N}^{k} a_j$ exists and $\lim_{k\to\infty} \prod_{j=N}^{k} a_j \ne 0$. In this case we put

$$\prod_{j=1}^{\infty} a_j := a_1 \cdot a_2 \cdot \ldots \cdot a_{N-1} \cdot \lim_{k\to\infty} \prod_{j=N}^{k} a_j.$$

For a convergent infinite product $\prod_{j=1}^{\infty} a_j$ we easily see that $\prod_{j=1}^{\infty} a_j = 0$ if and only if there exists $j \in \mathbb{N}$ with $a_j = 0$. For $k > N$ we have

$$a_k = \frac{\prod_{j=N}^{k} a_j}{\prod_{j=N}^{k-1} a_j}$$

and thus $a_k \to 1$ if the product $\prod_{j=1}^{\infty} a_j$ converges. This necessary condition for convergence is not sufficient, as shown by the example $a_j = j/(j+1)$ already considered.

The condition $a_k \to 1$ for infinite products corresponds to the condition $a_k \to 0$ for infinite series $\sum_{j=1}^{\infty} a_j$. The analogue of the Cauchy criterion for infinite series says that the infinite product $\prod_{j=1}^{\infty} a_j$ converges if and only if for every $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that

$$\left| \prod_{j=m}^{n} a_j - 1 \right| < \varepsilon \quad \text{for } n > m \geq N.$$

Since $a_j \to 1$ for a convergent infinite product $\prod_{j=1}^{\infty} a_j$ we write the factors a_j in the form $a_j = 1 + c_j$.

The product $\prod_{j=1}^{\infty}(1+c_j)$ is called *absolutely convergent* if $\prod_{j=1}^{\infty}(1+|c_j|)$ converges. Since

$$\left| \prod_{j=m}^{n}(1+c_j) - 1 \right| \leq \prod_{j=m}^{n}(1+|c_j|) - 1$$

for $n \geq m$ we see that absolutely convergent infinite products are convergent. Moreover, the sequence

$$\left(\prod_{j=1}^{n}(1+|c_j|) \right)_{n \in \mathbb{N}}$$

is non-decreasing, so it converges if and only if it is bounded.

Theorem 1.1. *An infinite product $\prod_{j=1}^{n}(1+c_j)$ converges absolutely if and only if the series $\sum_{j=1}^{\infty} c_j$ converges absolutely.*

Proof. Since $\log(1+x) \leq x$ for $x > -1$ we have

$$\log\left(\prod_{j=1}^{n}(1+|c_j|) \right) = \sum_{j=1}^{n} \log(1+|c_j|) \leq \sum_{j=1}^{n} |c_j|$$

so that the absolute convergence of the series implies that of the infinite product. Suppose now that the infinite product converges absolutely. Then $c_j \to 0$ and thus there exists N with $|c_j| \leq 1$ for $j \geq N$. Since $x \leq 2\log(1+x)$ for $0 \leq x \leq 1$ we deduce that

$$\sum_{j=N}^{n} |c_j| \leq 2 \sum_{j=N}^{n} \log(1+|c_j|) = 2\log\left(\prod_{j=N}^{n}(1+|c_j|) \right)$$

for $n \geq N$, from which the conclusion follows. □

The above considerations extend to infinite products of functions. Definitions like (locally) uniform convergence can be generalized to infinite products in an obvious way.

Let (f_n) be a sequence of functions which are holomorphic in a domain D. The Weierstraß theorem says that if (f_n) converges locally uniformly in D, then the limit function $f = \lim_{n\to\infty} f_n$ is also holomorphic in D. Moreover, $f_n' \to f'$ locally uniformly.

The following theorem is the analogous result for infinite products.

Theorem 1.2. *Let D be a domain and let (g_j) be a sequence of functions holomorphic in D. Suppose that $\sum_{j=1}^{\infty} |g_j|$ converges locally uniformly in D. Then*

$$f(z) := \prod_{j=1}^{\infty}(1 + g_j(z))$$

converges locally uniformly in D and defines a holomorphic function $f\colon D \to \mathbb{C}$.

Moreover, for $z \in D$ we have $f(z) = 0$ if and only if there exists $j \in \mathbb{N}$ such that $1 + g_j(z) = 0$.

2 Partitions

We saw in Example 1.2 that the coefficients c_n in

$$f(x) = \frac{1}{(1-x)(1-x^2)(1-x^3)} = \sum_{n=0}^{\infty} c_n x^n$$

count in how many ways n can be represented as a sum with summands from $\{1, 2, 3\}$. If we allow summands from $\{1, 2, \dots, N\}$ with $N \in \mathbb{N}$, the corresponding generating function f is given by

$$f(x) = \frac{1}{\prod_{k=1}^{N}(1-x^k)} = \prod_{k=1}^{N} \frac{1}{1-x^k}.$$

Allowing arbitrary summands from $\mathbb{N}$ leads to the generating function

$$F(x) = \frac{1}{\prod_{k=1}^{\infty}(1-x^k)} = \prod_{k=1}^{\infty} \frac{1}{1-x^k}.$$

Note that the product converges absolutely if $|x| < 1$, since the geometric series $\sum_{k=1}^{\infty} x^k$ converges absolutely for such x.

For $0 < \rho < 1$ the convergence is uniform for $|x| \le \rho$. This also holds for $x \in \mathbb{C}$. Theorem 1.2 thus yields that the function F defined by

$$F(z) = \frac{1}{\prod_{k=1}^{\infty}(1-z^k)} = \prod_{k=1}^{\infty} \frac{1}{1-z^k}$$

is holomorphic in $\mathbb{D} := \{z \in \mathbb{C}\colon |z| < 1\}$. Moreover, F has no zeros in $\mathbb{D}$.

For $n \in \mathbb{N}$ we call a representation $n = n_1 + n_2 + \cdots + n_k$ with $n_1, n_2, \ldots, n_k \in \mathbb{N}$ satisfying $n_1 \geq n_2 \geq \cdots \geq n_k$ a *partition* of n. The number of different partitions of n is denoted by $p(n)$ and the function $p\colon \mathbb{N} \to \mathbb{N}$ is called the *partition function*. The condition $n_1 \geq n_2 \geq \cdots \geq n_k$ means, for example, that $5 = 3 + 1 + 1 = 1 + 3 + 1 = 1 + 1 + 3$ counts only as one partition of 5. The partitions of 5 are given by

$$\begin{aligned} &1+1+1+1+1+1,\\ &2+1+1+1,\\ &2+2+1,\\ &3+1+1,\\ &3+2,\\ &4+1,\\ &5. \end{aligned}$$

Thus $p(5) = 7$.

Putting $p(0) = 1$ we collect the above considerations in the following theorem.

Theorem 2.1. *For $z \in \mathbb{D}$ we have*

$$F(z) = \prod_{k=1}^{\infty} \frac{1}{1 - z^k} = \sum_{n=0}^{\infty} p(n) z^n.$$

We will use this later to get estimates for $p(n)$. But before we consider some variants of $p(n)$. We denote by $p_o(n)$ the number of partitions of n where all summands are odd and by $p_d(n)$ the number of partitions where all summands are distinct. For $n = 5$, the partitions $1+1+1+1+1$, $3+1+1$, 5 contribute to $p_o(5)$ while $3+2$, $4+1$, 5 contribute to $p_d(5)$. Thus $p_o(5) = p_d(5) = 3$.

Theorem 2.2. $p_o(n) = p_d(n)$ *for all* $n \in \mathbb{N}$.

Proof. We put $p_o(0) = p_d(0) = 1$. The same considerations that led to Theorem 2.1 show that

$$\prod_{k=1}^{\infty} \frac{1}{1 - z^{2k-1}} = \sum_{n=0}^{\infty} p_o(n) z^n$$

for $z \in \mathbb{D}$. We also have

$$\prod_{k=1}^{\infty} \left(1 + z^k\right) = \sum_{n=0}^{\infty} p_d(n) z^n$$

for $z \in \mathbb{D}$. To see this, we note that for $N \in \mathbb{N}$ we have

$$\prod_{k=1}^{N} \left(1 + z^k\right) = \sum_{A \subset \{1,\ldots,N\}} z^{\sum_{m \in A} m},$$

where we put $\sum_{m\in\emptyset} m = 0$ and $z^0 = 1$. The claim follows since for $n \in \mathbb{N}$ there are precisely $p_d(n)$ subsets A of $\mathbb{N}$ (or of $\{1,\dots,n\}$) such that $\sum_{m\in A} m = n$.

To prove the theorem it thus suffices to show that

$$\prod_{k=1}^{\infty} \frac{1}{1-z^{2k-1}} = \prod_{m=1}^{\infty} (1+z^m)\,.$$

But this follows since

$$\begin{aligned}\prod_{m=1}^{\infty}(1+z^m) &= \prod_{m=1}^{\infty}\frac{1-z^{2m}}{1-z^m}\\ &= \prod_{m=1}^{\infty}\left(1-z^{2m}\right)\cdot\prod_{m=1}^{\infty}\frac{1}{1-z^m}\\ &= \left(1\cdot\left(1-z^2\right)\cdot 1\cdot\left(1-z^4\right)\cdot\ldots\right)\cdot\left(\frac{1}{1-z}\cdot\frac{1}{1-z^2}\cdot\frac{1}{1-z^3}\cdot\frac{1}{1-z^4}\cdot\ldots\right)\\ &= \frac{1}{1-z}\cdot 1\cdot\frac{1}{1-z^3}\cdot 1\cdot\ldots\\ &= \prod_{k=1}^{\infty}\frac{1}{1-z^{2k-1}}.\end{aligned}$$

□

Our next aim is to obtain upper bounds for the number $p(n)$.

Lemma 2.1. *Let $x \in \mathbb{R}$, $|x| < 1$. Then*

$$\log F(x) = \sum_{j=1}^{\infty}\frac{1}{j}\cdot\frac{x^j}{1-x^j}.$$

Remark. It is clear from the definition that $F(x) > 0$ for $|x| < 1$ so that $\log F(x)$ is defined.

Proof of Lemma 2.1. For $|x| < 1$ we have

$$\log\frac{1}{1-x} = -\log(1-x) = -\sum_{j=1}^{\infty}\frac{(-1)^{j+1}}{j}(-x)^j = \sum_{j=1}^{\infty}\frac{x^j}{j}$$

and thus

$$\begin{aligned}\log F(x) &= \log\left(\prod_{k=1}^{\infty}\frac{1}{1-x^k}\right)\\ &= \sum_{k=1}^{\infty}\log\frac{1}{1-x^k}\\ &= \sum_{k=1}^{\infty}\sum_{j=1}^{\infty}\frac{1}{j}\cdot x^{kj}\end{aligned}$$

$$= \sum_{j=1}^{\infty} \frac{1}{j} \sum_{k=1}^{\infty} x^{kj}$$
$$= \sum_{j=1}^{\infty} \frac{1}{j} \cdot \frac{x^j}{1-x^j}. \qquad \square$$

Remark. We will use the equation

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}.$$

There are many proofs of this formula. Some of them will be discussed in the exercises.

Lemma 2.2. *Let $0 < x < 1$. Then*

$$\log F(x) \leq \frac{\pi^2}{6} \cdot \frac{x}{1-x}.$$

Proof. Bernoulli's inequality says that

$$y^j - 1 \geq j(y-1)$$

for $y > 1$ and $j \in \mathbb{N}$. Hence

$$\frac{x^j}{1-x^j} = \frac{1}{\left(\frac{1}{x}\right)^j - 1} \leq \frac{1}{j\left(\frac{1}{x} - 1\right)} = \frac{x}{j(1-x)}.$$

Lemma 2.1 now yields that

$$\log F(x) \leq \sum_{j=1}^{\infty} \frac{1}{j^2} \cdot \frac{x}{1-x} = \frac{\pi^2}{6} \cdot \frac{x}{1-x}. \qquad \square$$

Theorem 2.3.

$$p(n) \leq \exp\left(\pi\sqrt{\frac{2n}{3}}\right)$$

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

Proof. For $0 < x < 1$ we have

$$p(n)x^n \leq F(x)$$

and thus

$$\log p(n) \leq \log F(x) - n \log x$$
$$= \log F(x) + n \log\left(1 + \frac{1-x}{x}\right)$$
$$\leq \frac{\pi^2}{6} \cdot \frac{x}{1-x} + n \cdot \frac{1-x}{x}.$$

We now choose x such that the right hand side becomes minimal. To do so we put $y = (1-x)/x$ and note that the function $h\colon (0,\infty) \to \mathbb{R}$,

$$h(y) = \frac{\pi^2}{6} \cdot \frac{1}{y} + ny,$$

attains its minimum for $y = \pi/\sqrt{6n}$, since

$$h'(y) = n - \frac{\pi^2}{6} \cdot \frac{1}{y^2}.$$

We deduce that

$$\log p(n) \le \frac{\pi^2}{6} \cdot \frac{\sqrt{6n}}{\pi} + n \cdot \frac{\pi}{\sqrt{6n}} = \frac{2\pi}{\sqrt{6}}\sqrt{n} = \pi\sqrt{\frac{2n}{3}}. \qquad \square$$

We also give an elementary lower bound for the partition function.

Theorem 2.4. *Let $n \in \mathbb{N}$, $n \ge 2$. Then*

$$p(n) \ge 2^{\lfloor\sqrt{n}\rfloor} = \exp\left(\lfloor\sqrt{n}\rfloor \cdot \log 2\right).$$

Proof. Let $\{x_1, \dots, x_k\}$ be any subset of $\{1, 2, \dots, \lfloor\sqrt{n}\rfloor\}$. Then

$$\sum_{j=1}^{k} x_j \le \sum_{j=1}^{\lfloor\sqrt{n}\rfloor} j = \frac{\lfloor\sqrt{n}\rfloor\left(\lfloor\sqrt{n}\rfloor + 1\right)}{2}.$$

Hence

$$x_{k+1} := n - \sum_{j=1}^{k} x_j \ge n - \frac{\lfloor\sqrt{n}\rfloor\left(\lfloor\sqrt{n}\rfloor + 1\right)}{2}.$$

For $n \ge 10$ we have

$$\begin{aligned} n - \frac{\lfloor\sqrt{n}\rfloor\left(\lfloor\sqrt{n}\rfloor + 1\right)}{2} &\ge n - \frac{\sqrt{n}(\sqrt{n}+1)}{2} \\ &= \sqrt{n}\sqrt{n} - \frac{\sqrt{n}(\sqrt{n}+1)}{2} \\ &= \frac{\sqrt{n}-1}{2}\sqrt{n} \\ &> \lfloor\sqrt{n}\rfloor \end{aligned}$$

and thus $x_{k+1} > x_j$ for $1 \le j \le k$. Hence $x_1, x_2, \dots, x_k, x_{k+1}$ yield a partition of n. Since $\{1, 2, \dots, \lfloor\sqrt{n}\rfloor\}$ has $2^{\lfloor\sqrt{n}\rfloor}$ subsets the conclusion follows for $n \ge 10$. For $n \le 8$ the claim follows directly by observing that $p(2) = 2$, $p(3) = 3$ and $p(n) \ge p(4) = 5 \ge 4 = 2^{\lfloor\sqrt{n}\rfloor}$ for $4 \le n \le 8$. Finally, $p(9) = 30 \ge 8 = 2^3$. $\square$

Remark. Theorem 2.3 and 2.4 yield constants c_1 and c_2 such that

$$\exp(c_1\sqrt{n}) \leq p(n) \leq \exp(c_2\sqrt{n}).$$

So these theorems determine the order of magnitude with which $p(n)$ tends to ∞. It turns out that the upper bound given by Theorem 2.3 is much closer to the true value of $p(n)$ than the lower bound. In fact, we have

$$p(n) \sim \frac{e^{\pi\sqrt{2n/3}}}{4\sqrt{3}n} \quad \text{as} \quad n \to \infty.$$

So in the asymptotic behavior as $n \to \infty$ the upper bound is off only by a factor $4\sqrt{3}n$, which is very small compared to the exponential term.

We will not prove this asymptotic behavior of $p(n)$. However, we will discuss the method used with a simpler example. This example is given by the exponential function. We have

$$e^x = \sum_{n=0}^{\infty} \frac{1}{n!}x^n$$

for $x \in \mathbb{R}$ (or $x \in \mathbb{C}$). Thus the exponential function is the generating function of the sequence $(1/n!)_{n\geq 0}$. As in the proof of Theorem 2.3 we obtain an upper bound for $1/n!$ and thus a lower bound for $n!$ by noting that

$$\frac{1}{n!}x^n \leq e^x$$

for all $x > 0$ and all $n \in \mathbb{N}$. We write this as

$$\frac{1}{n!} \leq \frac{e^x}{x^n} =: h(x)$$

and choose x such that the right hand side becomes minimal. Since

$$h'(x) = \frac{e^x x^n - e^x n x^{n-1}}{x^{2n}} = (x-n)\frac{e^x}{x^{n+1}}$$

we see that the minimum is attained for $x = n$ and thus

$$\frac{1}{n!} \leq h(n) = \frac{e^n}{n^n} = \left(\frac{e}{n}\right)^n$$

so that

$$n! \geq \left(\frac{n}{e}\right)^n.$$

The following result known as Stirling's formula shows that asymptotically $n!$ is only slightly larger than the right hand side.

Theorem 2.5 (Stirling's formula)**.**

$$n! \sim \sqrt{2\pi n}\left(\frac{n}{e}\right)^n \quad \textit{as } n \to \infty.$$

Before we start with the proof we recall that if

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

is a power series with radius of convergence R, and if $0 < r < R \leq \infty$, then

$$a_n = \frac{1}{2\pi i} \int_{|z|=r} \frac{f(z)}{z^{n+1}} dz.$$

This is a standard result of complex function theory.

Recall here that for a continuous function $g\colon \{z \in \mathbb{C}\colon |z| = r\} \to \mathbb{C}$ the integral

$$\int_{|z|=r} g(z)dz$$

is defined as the integral of g along the curve $\gamma\colon [-\pi, \pi] \to \mathbb{C}$, $\gamma(t) = re^{it}$, that is,

$$\int_{|z|=r} g(z)dz = \int_{\gamma} g(z)dz = \int_{-\pi}^{\pi} g(\gamma(t))\gamma'(t)dt = i\int_{-\pi}^{\pi} g(re^{it})re^{it}dt.$$

The above formula for the coefficients a_n thus takes the form

$$a_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{f(re^{it})}{r^n} e^{-int} dt.$$

We may interpret a power series in $z = re^{it}$ also as a Fourier series in t and the formula for a_n is that for the Fourier coefficients.

The above representation of the coefficients also yields the estimate

$$|a_n|r^n \leq \max_{|z|=r} |f(z)|.$$

We had used this already for $a_n = p(n)$ and $a_n = 1/n!$, but in these cases it is obvious since $a_n \geq 0$ for all $n \in \mathbb{N}_0$ and $\max_{|z|=r} |f(z)| = f(r)$. This estimate of a_n is known as Cauchy's inequality.

The integral

$$\int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi}$$

will also be used in the proof. It plays an important role in probability theory.

Proof of Theorem 2.5. Since

$$e^z = \sum_{n=0}^{\infty} \frac{1}{n!} z^n$$

the formula for the Taylor coefficients discussed above yields that

$$\frac{1}{n!} = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{e^{re^{it}}}{r^n} e^{-int} dt = \frac{1}{2\pi r^n} \int_{-\pi}^{\pi} \exp\bigl(re^{it} - int\bigr)\, dt.$$

We have

$$\begin{aligned} e^{it} &= 1 + it + \frac{1}{2}(it)^2 + \frac{1}{6}(it)^3 + \mathcal{O}(t^4) \\ &= 1 + it - \frac{1}{2}t^2 - \frac{1}{6}it^3 + \mathcal{O}(t^4) \end{aligned}$$

and thus

$$re^{it} = r + irt - \frac{1}{2}rt^2 - \frac{1}{6}irt^3 + \mathcal{O}(t^4)$$

as $t \to 0$. Therefore it seems reasonable to choose $r = n$ in the above formula for $1/n!$, since then the coefficients of t in the exponent cancel.

We need the above expansion of e^{it} with an explicit error term and thus note that there exists $\delta > 0$ such that

$$e^{it} = 1 + it - \frac{1}{2}t^2 + R(t)t^3$$

with

$$|R(t)| \le \frac{1}{5} \quad \text{for} \quad |t| \le \delta.$$

Choosing $r = n$ in the above integral we write

$$\begin{aligned} \int_{-\pi}^{\pi} \exp\bigl(ne^{it} - int\bigr)\, dt &= \int_{|t|\le\delta} \exp\bigl(ne^{it} - int\bigr)\, dt + \int_{\delta\le|t|\le\pi} \exp\bigl(ne^{it} - int\bigr)\, dt \\ &=: I_1 + I_2. \end{aligned}$$

We have

$$\begin{aligned} I_1 &= \int_{-\delta}^{\delta} \exp\left(n\left(1 + it - \frac{1}{2}t^2 + R(t)t^3\right) - int\right) dt \\ &= e^n \int_{-\delta}^{\delta} \exp\left(-\frac{1}{2}nt^2 + nR(t)t^3\right) dt \\ &= e^n \sqrt{\frac{2}{n}} \int_{-\sqrt{n/2}\delta}^{\sqrt{n/2}\delta} \exp\left(-s^2 + 2\sqrt{\frac{2}{n}} R\left(\sqrt{\frac{2}{n}}s\right) s^3\right) ds \end{aligned}$$

and thus, with $s_n = \sqrt{n/2}\delta$,

$$\sqrt{\frac{n}{2}} e^{-n} I_1 = \int_{-s_n}^{s_n} \exp\left(-s^2 + 2\delta R\left(\frac{\delta s}{s_n}\right) \frac{s^3}{s_n}\right) ds.$$

We may assume that $\delta \le 1$. For $|s| \le s_n$ we then have

$$2\delta R\left(\frac{\delta s}{s_n}\right) \frac{s^3}{s_n} \le \frac{2\delta}{5} s^2 \le \frac{1}{2}s^2$$

and thus

$$-s^2 + 2\delta R\left(\frac{\delta s}{s_n}\right) \frac{s^3}{s_n} \le -\frac{1}{2}s^2.$$

The theorem about dominated convergence now yields, since $s_n \to \infty$, that

$$\sqrt{\frac{n}{2}} e^{-n} I_1 \to \int_{-\infty}^{\infty} e^{-s^2} ds = \sqrt{\pi}.$$

Hence

$$I_1 \sim \frac{\sqrt{2\pi}}{\sqrt{n}} e^n$$

as $n \to \infty$.

We also have

$$\begin{aligned} I_2 &\leq \int_{\delta \leq |t| \leq \pi} \left| \exp\left(ne^{it} - int\right) \right| dt \\ &= \int_{\delta \leq |t| \leq \pi} \exp\left(n \operatorname{Re}(e^{it})\right) dt \\ &= \int_{\delta \leq |t| \leq \pi} \exp(n \cos t)\, dt \\ &\leq (2\pi - \delta) e^{n \cos \delta} \\ &= o\left(\frac{e^n}{\sqrt{n}}\right) \end{aligned}$$

and thus

$$|I_2| = o(I_1) \quad \text{as} \quad n \to \infty.$$

Altogether we find that

$$\frac{1}{n!} = \frac{1}{2\pi n^n}(I_1 + I_2) \sim \frac{1}{2\pi n^n} I_1 \sim \frac{1}{2\pi n^n} \frac{\sqrt{2\pi}}{\sqrt{n}} e^n = \frac{1}{\sqrt{2\pi n}} \left(\frac{e}{n}\right)^n$$

as $n \to \infty$, from which the conclusion follows. □

The method used to prove Stirling's formula applies in more general situations. Suppose for simplicity that

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

The point is to choose r in the formula

$$a_n = \frac{1}{2\pi i} \int_{|z|=r} \frac{e^{g(z)}}{z^{n+1}} dz = \frac{1}{2\pi} \int_{-\pi}^{\pi} \exp\left(g\left(re^{it}\right) - n \log r - nit\right) dt$$

in such a way that r is a critical point of $h(z) = e^{g(z)}/z^n$. One also calls this a saddle point and thus the method is also known as the saddle point method. This means that

$$0 = h'(r) = e^{g(r)} \left(\frac{g'(r)}{r^n} - \frac{n}{r^{n+1}} \right)$$

and thus $g'(r) = n/r$. With
$$a(r) := rg'(r)$$
and
$$b(r) := ra'(r) = rg'(r) + r^2 g''(r)$$
we have
$$\begin{aligned} g\big(re^{it}\big) &= g(r) + g'(r)r\left(e^{it} - 1\right) + \frac{1}{2}g''(r)r^2\left(e^{it} - 1\right)^2 + \mathcal{O}(t^3) \\ &= g(r) + g'(r)rit - \frac{1}{2}\left(g'(r)r + g''(r)r^2\right)t^2 + \mathcal{O}(t^3) \\ &= g(r) + a(r)it - \frac{1}{2}b(r)t^2 + \mathcal{O}(t^3) \end{aligned}$$
as $t \to 0$. Choosing $r = r_n$ such that $a(r_n) = n$ we thus have
$$g\big(r_n e^{it}\big) - nit = g(r_n) - \frac{1}{2}b(r_n)t^2 + \mathcal{O}(t^3)$$
and thus
$$\begin{aligned} a_n &= \frac{1}{2\pi}\int_{-\pi}^{\pi} \exp\bigg(g(r_n) - n\log r_n - \frac{1}{2}b(r_n)t^2 + \mathcal{O}(t^3)\bigg)\,dt \\ &= \frac{e^{g(r_n)}}{2\pi r_n^n}\int_{-\pi}^{\pi}\exp\bigg(-\frac{1}{2}b(r_n)t^2 + \mathcal{O}(t^3)\bigg)\,dt. \end{aligned}$$
If the $\mathcal{O}(t^3)$-term can be controlled the last integral is asymptotic to
$$\int_{-\pi}^{\pi}\exp\bigg(-\frac{1}{2}b(r_n)t^2\bigg)\,dt \sim \sqrt{\frac{2\pi}{b(r_n)}}.$$
Overall we then obtain
$$a_n \sim \frac{1}{\sqrt{2\pi b(r_n)}} \cdot \frac{e^{g(r_n)}}{r_n^n}.$$
This can be compared with the estimate
$$|a_n| \le \frac{e^{g(r_n)}}{r_n^n}$$
obtained directly from Cauchy's inequality
$$|a_n| r^n \le \max_{|z|=r} |f(z)|.$$

We now discuss the result obtained when the method is applied to the partition function. First we note here (without proof) that Lemma 2.2, which says that
$$\log F(x) \le \frac{\pi^2}{6} \cdot \frac{x}{1-x}$$

can be improved to an asymptotic equation

$$\log F(x) = \frac{\pi^2}{-6\log x} + \frac{1}{2}\log\frac{1-x}{2\pi} + o(1)$$

as $x \to 1$. Note here that $-\log x \sim 1-x$ as $x \to 1$. Using $x = x_n = 1 - \pi/\sqrt{6n}$ as in the proof of Theorem 2.3 improves the conclusion of this theorem to

$$\begin{aligned}\log p(n) &\le \log F(x_n) - n\log x_n \\ &= \frac{\pi^2}{6\cdot\frac{\pi}{\sqrt{6n}}} + \frac{1}{2}\log\frac{\pi}{2\pi\sqrt{6n}} + n\frac{\pi}{\sqrt{6n}} + o(1) \\ &= \pi\sqrt{\frac{2n}{3}} - \frac{1}{2}\log(2\sqrt{6n}) + o(1)\end{aligned}$$

and thus

$$p(n) \le (1+o(1))\frac{\exp\left(\pi\sqrt{\frac{2}{3}n}\right)}{\sqrt{2}\cdot\sqrt[4]{6}\cdot\sqrt[4]{n}}.$$

It can be shown that with $g(r) = \log F(r)$ and $a(r) = rg'(r)$ and $b(r) = ra'(r)$ we have

$$a(r) \sim \frac{\pi^2}{6(\log r)^2} \quad \text{and} \quad b(r) \sim \frac{\pi^2}{-3(\log r)^3}$$

as $r \to 1$. Hence

$$b(x_n) \sim \frac{\pi^2}{3\left(\frac{\pi}{\sqrt{6n}}\right)^3} = \frac{2n\sqrt{6n}}{\pi}$$

and thus

$$\sqrt{2\pi b(x_n)} \sim \sqrt{4n\sqrt{6n}} = 2\sqrt[4]{6}n^{3/4}.$$

The saddle point methods then yields that

$$\begin{aligned}p(n) &\sim \frac{1}{\sqrt{2\pi b(x_n)}}\cdot\frac{F(x_n)}{x_n^n} \\ &\sim \frac{1}{2\cdot\sqrt[4]{6}\cdot n^{3/4}}\cdot\frac{\exp\left(\pi\sqrt{\frac{2}{3}n}\right)}{\sqrt{2}\cdot\sqrt[4]{6}\cdot\sqrt[4]{n}} \\ &= \frac{\exp\left(\pi\sqrt{\frac{2}{3}n}\right)}{4\sqrt{3}\cdot n}.\end{aligned}$$

Of course, this does not constitute a proof. But a proof can be given along these lines. Essentially, it consists of giving rigorous bounds in the $\mathcal{O}$-terms and o-terms arising.

3 Prime numbers

Let $\mathbb{P} = \{2, 3, 5, \dots\}$ be the set of prime numbers. For $x > 0$ we put

$$\pi(x) = \operatorname{card}(\mathbb{P} \cap [0, x]) .$$

Thus $\pi(x)$ is the number of primes less than or equal to x. We will be concerned with the behavior of $\pi(x)$ as $x \to \infty$. The main result is the following.

Theorem 3.1 (Prime number theorem)**.** *As $x \to \infty$,*

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

This result was conjectured by Gauß and Legendre already around 1800. It was proved first in 1896 independently by Hadamard and de la Vallée-Poussin, using methods from complex analysis. We will prove this result in section 5, after introducing the necessary tools from complex analysis in section 4. While some key ideas of the original proofs are still present, we can draw on a number of simplifications and modifications that have been obtained since then.

In the present section we discuss some elementary properties of prime numbers which in part are also preparations for the proof of Theorem 3.1.

We assume familiarity with some basic facts of elementary number theory such as the unique factorizability of natural numbers into prime numbers.

Our first result says that $\pi(x) \to \infty$ as $x \to \infty$.

Theorem 3.2. *There are infinitely many prime numbers.*

Proof. Suppose that $\mathbb{P}$ is finite, say $\mathbb{P} = \{p_1, p_2, \dots, p_N\}$. Put

$$m = p_1 \cdot p_2 \cdot \ldots \cdot p_N + 1.$$

Then $p_j \nmid m$ for all $j \in \{1, \dots, N\}$. But this implies that m is a prime number, a contradiction. □

It turns out that instead of considering $\pi(x)$ it is often easier to consider

$$\theta(x) = \sum_{\substack{p \in \mathbb{P} \\ p \leq x}} \log p.$$

Sums over all $p \in \mathbb{P}$ satisfying $p \leq x$ or some other restriction will occur frequently. We will omit the condition $p \in \mathbb{P}$ in the following, but will use the variable p only if we sum over prime numbers. Thus we simply write

$$\theta(x) = \sum_{p \leq x} \log p.$$

The connection between $\pi(x)$ and $\theta(x)$ is given by the following result.

Lemma 3.1. *As $x \to \infty$,*

$$\pi(x) \sim \frac{\theta(x)}{\log x}.$$

The prime number theorem is thus equivalent to the statement that $\theta(x) \sim x$ as $x \to \infty$. The next result gives a weaker estimate.

Lemma 3.2. *There exist positive constants a and b such that*

$$ax \leq \theta(x) \leq bx$$

for $x \geq 2$. Moreover,

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

as $x \to \infty$.

Combining these lemmas yields the following result obtained by Tchebychev around 1850. It is weaker than the prime number theorem, but it shows that the function $x/\log x$ gives the right order of magnitude for $\pi(x)$.

Theorem 3.3. *There exist positive constants c and d such that*

$$c \cdot \frac{x}{\log x} \leq \pi(x) \leq d \cdot \frac{x}{\log x}$$

for $x \geq 2$.

The rest of the section consists of the proofs of Lemma 3.1 and 3.2.

Proof of Lemma 3.2. Let $n \in \mathbb{N}$ and write

$$n! = \prod_{p \leq n} p^{e(p)}.$$

Thus $e(p)$ is the exponent of the highest power of p that divides $n!$. Now $\lfloor \frac{n}{p} \rfloor$ of the numbers $1, 2, \ldots, n$ are divisible by p. Analogously, $\lfloor \frac{n}{p^2} \rfloor$ of them are divisible by p^2. In general, $\lfloor \frac{n}{p^k} \rfloor$ of them are divisible by p^k, for $k \in \mathbb{N}$. Thus

$$e(p) = \sum_{k \geq 1} \left\lfloor \frac{n}{p^k} \right\rfloor .$$

Here the sum is finite since $\lfloor \frac{n}{p^k} \rfloor = 0$ if $p^k > n$. This condition is equivalent to $k > (\log n)/(\log p)$. We deduce that

$$\begin{aligned}
\log n! &= \log\left(\prod_{p \leq n} p^{e(p)}\right) \\
&= \sum_{p \leq n} e(p) \log p \\
&= \sum_{p \leq n} \sum_{k \geq 1} \left\lfloor \frac{n}{p^k} \right\rfloor \log p \\
&= \sum_{p \leq n} \left\lfloor \frac{n}{p} \right\rfloor \log p + \sum_{p \leq n} \sum_{k \geq 2} \left\lfloor \frac{n}{p^k} \right\rfloor \log p.
\end{aligned}$$

Now

$$\sum_{k\geq 2}\left\lfloor \frac{n}{p^k}\right\rfloor \leq \sum_{k\geq 2}\frac{n}{p^k} = \frac{n}{p^2}\sum_{j=0}^{\infty}\frac{1}{p^j} = \frac{n}{p^2}\frac{1}{1-\dfrac{1}{p}} = \frac{n}{p(p-1)}$$

and thus

$$\sum_{p\leq n}\sum_{k\geq 2}\left\lfloor \frac{n}{p^k}\right\rfloor \log p \leq n\sum_{p\leq n}\frac{\log p}{p(p-1)} \leq n\sum_{j=2}^{\infty}\frac{\log j}{j(j-1)}.$$

The last series converges since $\log j \leq \sqrt{j}$ for large j and thus

$$\frac{\log j}{j(j-1)} \leq \frac{\sqrt{j}}{j(j-1)} \leq \frac{2}{j^{3/2}}$$

for large j. Hence

$$\sum_{p\leq n}\sum_{k\geq 2}\left\lfloor \frac{n}{p^k}\right\rfloor \log p = \mathcal{O}(n).$$

Stirling's formula yields that

$$\begin{aligned}\log n! &= \log\Big((1+o(1))\left(\frac{n}{e}\right)^n\sqrt{2\pi n}\Big)\\ &= n\log\Big(\frac{n}{e}\Big) + \frac{1}{2}\log(2\pi n) + o(1)\\ &= n\log n - n + \frac{1}{2}\log n + \mathcal{O}(1)\\ &= n\log n + \mathcal{O}(n).\end{aligned}$$

We conclude that

$$\sum_{p\leq n}\left\lfloor \frac{n}{p}\right\rfloor \log p = n\log n + \mathcal{O}(n).$$

Replacing n by $2n$ yields that

$$\sum_{p\leq 2n}\left\lfloor \frac{2n}{p}\right\rfloor \log p = 2n\log(2n) + \mathcal{O}(n) = 2n\log n + \mathcal{O}(n).$$

Subtracting the previous inequality, multiplied by factor 2, from this and noting that $\left\lfloor \frac{n}{p}\right\rfloor = 0$ if $p > n$ we obtain

$$\sum_{p\leq 2n}\left(\left\lfloor \frac{2n}{p}\right\rfloor - 2\left\lfloor \frac{n}{p}\right\rfloor\right)\log p = \mathcal{O}(n).$$

Now $\lfloor 2x\rfloor - 2\lfloor x\rfloor \geq 0$ for all $x \geq 0$ and $\lfloor 2x\rfloor - 2\lfloor x\rfloor = 1$ for $\frac{1}{2} \leq x < 1$. Noting also that $2n \notin \mathbb{P}$ we find that

$$\begin{aligned}\theta(2n) - \theta(n) &= \sum_{n<p<2n}\log p\\ &= \sum_{n<p<2n}\left(\left\lfloor \frac{2n}{p}\right\rfloor - 2\left\lfloor \frac{n}{p}\right\rfloor\right)\log p\\ &= \mathcal{O}(n).\end{aligned}$$

We have proven this for $n \in \mathbb{N}$, but it is easy to see that we may replace $n \in \mathbb{N}$ by $x \in \mathbb{R}$ here. In fact, there is at most one prime number between $2\lfloor x \rfloor$ and $2x$, so

$$\theta(2x) - \theta(2\lfloor x \rfloor) \leq \log 2x = \mathcal{O}(x).$$

Since $\theta(x) = \theta(\lfloor x \rfloor)$ we thus have

$$\theta(2x) - \theta(x) = \theta(2\lfloor x \rfloor) - \theta(\lfloor x \rfloor) + \mathcal{O}(x) = \mathcal{O}(x),$$

say

$$\theta(2x) - \theta(x) \leq Kx$$

for $x > 0$ with a positive constant K. We deduce that

$$\begin{aligned}
\theta(x) &= \theta(x) - \theta\Big(\frac{x}{2}\Big) + \theta\Big(\frac{x}{2}\Big) \\
&\leq K\frac{x}{2} + \theta\Big(\frac{x}{2}\Big) \\
&= K\frac{x}{2} + \theta\Big(\frac{x}{2}\Big) - \theta\Big(\frac{x}{4}\Big) + \theta\Big(\frac{x}{4}\Big) \\
&\leq K\frac{x}{2} + K\frac{x}{4} + \theta\Big(\frac{x}{4}\Big).
\end{aligned}$$

Inductively we obtain

$$\theta(x) \leq Kx \sum_{j=1}^{k-1} \frac{1}{2^j} + \theta\Big(\frac{x}{2^k}\Big) \leq Kx + \theta\Big(\frac{x}{2^k}\Big) = Kx,$$

if k is chosen such that $x/2^k < 2$. This is the upper bound for $\theta(x)$ that was claimed in the conclusion.

Recall the inequality

$$\sum_{p \leq n} \left\lfloor \frac{n}{p} \right\rfloor \log p = n \log n + \mathcal{O}(n)$$

proved above. Since

$$0 \leq n \sum_{p \leq n} \frac{\log p}{p} - \sum_{p \leq n} \left\lfloor \frac{n}{p} \right\rfloor \log p \leq \sum_{p \leq n} \log p = \theta(n) = O(n)$$

this yields that

$$\sum_{p \leq n} \frac{\log p}{p} = \frac{1}{n} \sum_{p \leq n} \left\lfloor \frac{n}{p} \right\rfloor \log p + \mathcal{O}(1) = \log n + \mathcal{O}(1).$$

Again we may replace $n \in \mathbb{N}$ by $x \in \mathbb{R}$ here and thus obtain the second conclusion of Lemma 3.2.

It remains to prove the lower bound for $\theta(x)$. In order to do so, let $0 < \alpha < 1$. Then, by what we just proved,

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

Here the $\mathcal{O}(1)$-term depends neither on x nor on α, as long as $\alpha x \geq 2$. Choosing α sufficiently small we obtain

$$\sum_{\alpha x < p \leq x} \frac{\log p}{p} \geq 1$$

for $x \geq 2/\alpha$. On the other hand,

$$\sum_{\alpha x < p \leq x} \frac{\log p}{p} \leq \frac{1}{\alpha x} \sum_{\alpha x < p \leq x} \log p \leq \frac{\theta(x)}{\alpha x}.$$

Hence $\theta(x) \geq \alpha x$. □

Proof of Lemma 3.1. First we note that

$$\theta(x) = \sum_{p \leq x} \log p \leq \sum_{p \leq x} \log x = \pi(x) \log x.$$

To obtain an estimate in the opposite direction, let $\varepsilon > 0$. Then

$$\begin{aligned} \theta(x) &\geq \sum_{x^{1-\varepsilon} < p \leq x} \log p \\ &\geq \sum_{x^{1-\varepsilon} < p \leq x} \log\left(x^{1-\varepsilon}\right) \\ &= (1-\varepsilon) \log x \left(\pi(x) - \pi\left(x^{1-\varepsilon}\right)\right). \end{aligned}$$

Lemma 3.2 yields that

$$\pi\left(x^{1-\varepsilon}\right) \leq x^{1-\varepsilon} = \frac{x}{x^{\varepsilon}} \leq \frac{\theta(x)}{a x^{\varepsilon}}.$$

Combining the last two estimates we find that

$$\pi(x) \leq \frac{\theta(x)}{(1-\varepsilon)\log x} + \pi\left(x^{1-\varepsilon}\right) = \left(\frac{1}{1-\varepsilon} + \frac{\log x}{a x^{\varepsilon}}\right) \frac{\theta(x)}{\log x} \leq \frac{1+\varepsilon}{1-\varepsilon} \cdot \frac{\theta(x)}{\log x}$$

for large x. As this holds for every $\varepsilon > 0$ we see altogether that

$$\frac{\theta(x)}{\log x} \leq \pi(x) \leq (1 + o(1)) \frac{\theta(x)}{\log x}$$

as $x \to \infty$. □

4 The Riemann zeta function

The main tool to prove the prime number theorem is a holomorphic function introduced by Riemann. In order to define it we recall that for $a > 0$ and $z \in \mathbb{C}$ the complex power a^z is defined by

$$a^z = \exp(z \log a).$$

For fixed a the function $z \mapsto a^z$ is entire; that is, it is holomorphic in $\mathbb{C}$.

We note that

$$|a^z| = \exp((\operatorname{Re} z) \cdot \log a) = a^{\operatorname{Re} z}.$$

This implies that the series

$$\sum_{n=1}^{\infty} \frac{1}{n^z}$$

converges for $\operatorname{Re} z > 1$. For $t \in \mathbb{R}$ we put

$$\mathbb{H}_t := \{z \in \mathbb{C} \colon \operatorname{Re} z > t\}.$$

Thus the above series converges in $\mathbb{H}_1$.

Definition 4.1. The function $\zeta \colon \mathbb{H}_1 \to \mathbb{C}$,

$$\zeta(z) = \sum_{n=1}^{\infty} \frac{1}{n^z},$$

is called the *Riemann zeta function*.

The series defining ζ actually converges locally uniformly in $\mathbb{H}_1$ and thus, by Weierstraß's theorem, ζ is holomorphic in $\mathbb{H}_1$.

In the theory of the Riemann zeta function, the complex variable is traditionally denoted by $s = \sigma + it$. However, we shall use the notation $z = x + iy$ for complex numbers that is usually used in complex analysis.

Theorem 4.1. *The function*

$$z \mapsto \zeta(z) - \frac{1}{z-1}$$

has a holomorphic continuation to $\mathbb{H}_0$.

Proof. For $z \in \mathbb{H}_1$ we have

$$\begin{aligned}
\zeta(z) - \frac{1}{z-1} &= \sum_{n=1}^{\infty} \frac{1}{n^z} - \int_1^{\infty} \frac{1}{x^z} dx \\
&= \sum_{n=1}^{\infty} \left(\frac{1}{n^z} - \int_n^{n+1} \frac{1}{x^z} dx \right) \\
&= \sum_{n=1}^{\infty} \int_n^{n+1} \left(\frac{1}{n^z} - \frac{1}{x^z} \right) dx.
\end{aligned}$$

Let $x \in [n, n+1]$. Then

$$\begin{aligned}\left|\frac{1}{n^z} - \frac{1}{x^z}\right| &= \left|\int_n^x \frac{d}{dt}\left(\frac{1}{t^z}\right) dt\right| \\ &= \left|\int_n^x \frac{-z}{t^{z+1}} dt\right| \\ &\le (x-n) \max_{t \in [n,x]} \left|\frac{-z}{t^{z+1}}\right| \\ &= (x-n)|z| \max_{t \in [n,x]} \frac{1}{t^{\operatorname{Re} z+1}} \\ &\le \frac{|z|}{n^{\operatorname{Re} z+1}}.\end{aligned}$$

It follows that

$$\left|\int_n^{n+1} \left(\frac{1}{n^z} - \frac{1}{x^z}\right) dx\right| \le \frac{|z|}{n^{\operatorname{Re} z+1}}.$$

Thus the series for $\zeta(z) - 1/(z-1)$ given above converges locally uniformly in $\mathbb{H}_0$ and hence, by Weierstraß's theorem, it represents a function holomorphic in $\mathbb{H}_0$. □

Remark. Theorem 4.1 yields that ζ has a meromorphic continuation to $\mathbb{H}_0$ with only one pole at 1. This pole is simple and the residue of ζ at this pole is 1.

The following result already indicates a connection between the Riemann zeta function and prime numbers.

Theorem 4.2. *Let $z \in \mathbb{H}_1$. Then*

$$\zeta(z) = \prod_{p \in \mathbb{P}} \frac{1}{1 - \frac{1}{p^z}}.$$

Here the infinite product converges locally uniformly in $\mathbb{H}_1$.

Proof. The locally uniform convergence of the product follows from Theorem 1.2 since the series

$$\sum_{p \in \mathbb{P}} \frac{1}{p^z}$$

converges locally uniformly. Let $p_1 = 2, p_2 = 3, p_3 = 5, p_4 = 7, \dots$ so that $\mathbb{P} = \{p_k \colon k \in \mathbb{N}\}$. For $N \in \mathbb{N}$ we then have

$$\prod_{k=1}^N \frac{1}{1 - \frac{1}{p_k^z}} = \prod_{k=1}^N \sum_{j=0}^\infty \left(\frac{1}{p_k^z}\right)^j = \prod_{k=1}^N \sum_{j=0}^\infty \frac{1}{p_k^{jz}} = \sum_{n \in S_N} \frac{1}{n^z},$$

where S_N is the set of all natural numbers (including 1) for which no primes other than $p_1, \dots, p_N$ occur in the prime number factorization. Letting N tend to infinity we obtain the conclusion. □

The product representation of ζ given by Theorem 4.2 is due to (and named after) Euler. An immediate consequence of it is the following result.

Theorem 4.3. *$\zeta(z) \neq 0$ for $\operatorname{Re} z > 1$.*

One of the most famous unsolved problems is the following conjecture.

Riemann hypothesis. *Let $z \in \mathbb{H}_0$ with $\zeta(z) = 0$. Then $\operatorname{Re} z = \frac{1}{2}$.*

The Riemann hypothesis thus says that all zeros of ζ in the right half-plane lie on the so-called critical line $\{z \in \mathbb{C}\colon \operatorname{Re} z = \frac{1}{2}\}$. A proof of this conjecture would have important consequences, some of which we will discuss later. For the proof of the prime number theorem a weaker result will suffice, however.

Theorem 4.4. *$\zeta(z) \neq 0$ for $\operatorname{Re} z = 1$.*

Before proving Theorem 4.4 we note that if D is a simply connected domain and $f\colon D \to \mathbb{C}$ is a holomorphic function which has no zeros, then f has the form $f = e^g$ with a holomorphic function $g\colon D \to \mathbb{C}$. To see this, we fix $z_0 \in D$, take w_0 with $e^{w_0} = f(z_0)$ and put

$$g(z) = \int_{z_0}^{z} \frac{f'(t)}{f(t)} dt + w_0,$$

noting that the integral does not depend on the path of integration since D is simply connected. A computation then shows that

$$\frac{d}{dz}\left(f(z)e^{-g(z)}\right) = 0$$

for all $z \in D$. Since also $f(z_0)e^{-g(z_0)} = 1$ this yields $f = e^g$.

We also write $g = \log f$, noting that g is unique only up to a constant $2\pi i k$ with $k \in \mathbb{Z}$. The equation $|e^g| = e^{\operatorname{Re} g}$ takes the form $\log |f| = \operatorname{Re}(\log f)$.

By Theorem 4.3 we can thus define a holomorphic function $\log \zeta\colon \mathbb{H}_1 \to \mathbb{C}$. Since $\zeta(x) > 1$ for $x > 1$ we may take the branch where $\log \zeta(x) > 0$ for $x > 1$.

Theorem 4.5. *Let $z \in \mathbb{H}_1$. Then*

$$\log \zeta(z) = \sum_{n=1}^{\infty} \frac{a_n}{n^z}$$

where $a_n = 1/k$ if $n = p^k$ for some $p \in \mathbb{P}$ and some $k \in \mathbb{N}$, and $a_n = 0$ otherwise.

Proof. By the Euler product representation we have, for $x > 1$,

$$\log \zeta(x) = \log\left(\prod_{p\in\mathbb{P}} \frac{1}{1 - \frac{1}{p^x}}\right) = \sum_{p\in\mathbb{P}} -\log\left(1 - \frac{1}{p^x}\right) = \sum_{p\in\mathbb{P}} \sum_{k=1}^{\infty} \frac{1}{k} \cdot \frac{1}{p^{kx}}.$$

This is the required formula if $z = x > 1$. But since the series on the right hand side converges locally uniformly in $\mathbb{H}_1$ and thus represents a holomorphic function there, the formula holds for all $z \in \mathbb{H}_1$ by the identity theorem. □

Proof of Theorem 4.4. Suppose that $\zeta(z_0) = 0$ where $z_0 = 1 + i\alpha$ with $\alpha \in \mathbb{R}$. Note that ζ has a pole at 1 by Theorem 4.1, so $\zeta(z) \to \infty$ as $z \to 1$. Thus $\alpha \neq 0$.

We consider the auxiliary function f given by

$$f(z) = \zeta(z)^3 \zeta(z - i\alpha)^4 \zeta(z - i2\alpha).$$

This function is meromorphic in $\mathbb{H}_0$. Poles can occur only at the poles of one of the factors; that is at 1, $1 + i\alpha$ and $1 + i2\alpha$. However, since $\zeta(z)^3$ has a pole of order 3 at $z = 1$ while $\zeta(z - i\alpha)^4$ has a zero of order at least 4 at $z = 1$, and $\zeta(z - i2\alpha)$ is holomorphic there, we see that f actually has a zero at $z = 1$. Thus $f(x) \to 0$ and hence

$$\log|f(x)| \to -\infty$$

as $x \to 1$.

By Theorem 4.5 we have

$$\log|\zeta(z)| = \operatorname{Re}\left(\sum_{n=1}^{\infty} a_n n^{-z}\right) = \sum_{n=1}^{\infty} a_n \operatorname{Re}\big(n^{-z}\big)$$

for $z \in \mathbb{H}_1$. Writing $z = x + iy$ we have

$$n^{-z} = \exp((x + iy)\log n) = n^{-x} e^{-iy \log n}$$

and thus

$$\operatorname{Re}\big(n^{-z}\big) = n^{-x}\cos(-y\log n) = n^{-x}\cos(y \log n).$$

Hence

$$\log|\zeta(x + iy)| = \sum_{n=1}^{\infty} a_n n^{-x}\cos(y \log n).$$

This yields that

$$\begin{aligned} \log|f(x)| &= 3\log|\zeta(x)| + 4\log|\zeta(x - i\alpha)| + \log|\zeta(z - i2\alpha)| \\ &= \sum_{n=1}^{\infty} a_n n^{-x}\left(3 + 4\cos(\alpha\log n) + \cos(2\alpha \log n)\right). \end{aligned}$$

For $t \in \mathbb{R}$ we have $\cos(2t) = 2\cos^2 t - 1$ and hence

$$3 + 4\cos t + \cos(2t) = 2 + 4\cos t + 2\cos^2 t = 2(1 + \cos t)^2 \geq 0.$$

It follows that $\log|f(x)| \geq 0$ for $x > 1$, contradicting the relation $\log|f(x)| \to -\infty$ as $x \to 1$ obtained earlier. □

5 Proof of the prime number theorem

First we use the series representation of $\log\zeta$ obtained in Theorem 4.5 to obtain such a representation for ζ'/ζ. Note that $(\log\zeta)' = \zeta'/\zeta$. Thus ζ'/ζ is also called the logarithmic derivative of ζ.

Theorem 5.1. *For $z \in \mathbb{H}_1$ we have*

$$-\frac{\zeta'(z)}{\zeta(z)} = \sum_{n=1}^{\infty} \frac{a_n \log n}{n^z} = \sum_{p\in\mathbb{P}} \sum_{k=1}^{\infty} \frac{\log p}{p^{kz}}.$$

Proof. The equation $(\log \zeta)'(z) = \zeta'(z)/\zeta(z)$ already mentioned follows for $z \in \mathbb{R}$, $z > 1$, by the standard chain rule from (real) analysis. But then it holds for $z \in \mathbb{H}_1$ by the identity theorem. Since

$$\frac{d}{dz}\left(\frac{1}{n^z}\right) = \frac{d}{dz}\exp(-z\log n) = -\log n \cdot \exp(-z\log n) = -\frac{\log n}{n^z}$$

the first equation follows from Theorem 4.5 and Weierstraß's theorem. If $n = p^k$ with $p \in \mathbb{P}$ and $k \in \mathbb{N}$, then

$$a_n \log n = \frac{1}{k}\log\left(p^k\right) = \log p.$$

Since $a_n = 0$ otherwise the second equation also follows. □

For us it will be convenient to consider only those terms in the series for ζ'/ζ that correspond to $k = 1$. Thus we consider the function $\Phi\colon \mathbb{H}_1 \to \mathbb{C}$,

$$\Phi(z) = \sum_{p\in\mathbb{P}} \frac{\log p}{p^z}.$$

We thus have

$$\Phi(z) + \frac{\zeta'(z)}{\zeta(z)} = -\sum_{p\in\mathbb{P}} \sum_{k=2}^{\infty} \frac{\log p}{p^{kz}}.$$

Lemma 5.1. *The function $\Phi + \zeta'/\zeta$ has a holomorphic continuation to $\mathbb{H}_{1/2}$.*

Proof. For $\operatorname{Re} z > 0$ we have

$$\sum_{k=2}^{\infty} \frac{1}{p^{kz}} = \sum_{k=2}^{\infty}\left(\frac{1}{p^z}\right)^k = \left(\frac{1}{p^z}\right)^2 \sum_{j=0}^{\infty}\left(\frac{1}{p^z}\right)^j = \left(\frac{1}{p^z}\right)^2 \frac{1}{1-\frac{1}{p^z}} = \frac{1}{p^z\left(p^z-1\right)}.$$

Thus the above series for $\Phi + \zeta'/\zeta$ can be rewritten as

$$\Phi(z) + \frac{\zeta'(z)}{\zeta(z)} = -\sum_{p\in\mathbb{P}} \frac{\log p}{p^z\left(p^z-1\right)}.$$

Let $\varepsilon > 0$. For large p we have $\log p < p^{\varepsilon/2}$ and $p^{1/2+\varepsilon} > 2p^{1/2+\varepsilon/2} > p^{1/2+\varepsilon/2} + 1$. For $\operatorname{Re} z > \frac{1}{2} + \varepsilon$ hence

$$\left|\frac{\log p}{p^z\left(p^z-1\right)}\right| \leq \frac{\log p}{p^{\operatorname{Re} z}\left(p^{\operatorname{Re} z}-1\right)} \leq \frac{p^{\varepsilon/2}}{2p^{1/2+\varepsilon}p^{1/2+\varepsilon/2}} = \frac{1}{2p^{1+\varepsilon}}.$$

Thus the series for $\Phi + \zeta'/\zeta$ converges locally uniformly in $\mathbb{H}_{1/2}$. Hence it is holomorphic there by Weierstraß's theorem. □

By Theorem 4.1, ζ has a simple pole at 0. This implies that the function h given by $h(z) = (z-1)\zeta(z)$ extends holomorphically to a neighborhood of 1 with $h(1) \neq 0$. Since

$$\frac{h'(z)}{h(z)} = \frac{1}{z-1} + \frac{\zeta'(z)}{\zeta(z)}$$

and since $\zeta(z) \neq 0$ if $\operatorname{Re} z = 1$ we see that the function given by

$$z \mapsto \frac{1}{z-1} + \frac{\zeta'(z)}{\zeta(z)}$$

can be continued analytically to some domain containing the closure $\overline{\mathbb{H}_1}$ of $\mathbb{H}_1$. By Lemma 5.1 the same applies to the function

$$z \mapsto \Phi(z) - \frac{1}{z-1}.$$

We summarize this as follows.

Lemma 5.2. *Let* $\Phi\colon \mathbb{H}_1 \to \mathbb{C}$,

$$\Phi(z) = \sum_{p \in \mathbb{P}} \frac{\log p}{p^z}.$$

Then the function defined by

$$z \mapsto \Phi(z) - \frac{1}{z-1}$$

has a holomorphic continuation to a domain containing $\overline{\mathbb{H}_1}$.

The following lemma connects the function Φ of this lemma with the function θ defined by

$$\theta(x) = \sum_{p \leq x} \log p$$

considered in Section 4.

Lemma 5.3. *For* $\operatorname{Re} z > 1$ *we have*

$$\Phi(z) = z \int_1^\infty \frac{\theta(t)}{t^{z+1}} dt.$$

This lemma is a special case of the following result.

Lemma 5.4. *Let* (a_n) *be a sequence of non-negative real numbers and* $r \geq 0$ *such that the series*

$$f(z) = \sum_{n=1}^{\infty} \frac{a_n}{n^z}$$

converges for $z \in \mathbb{H}_r$. *Define* $\varphi\colon \mathbb{R} \to \mathbb{R}$,

$$\varphi(x) = \sum_{n \le x} a_n = \sum_{n=1}^{\lfloor x \rfloor} a_n,$$

with $\varphi(x) = 0$ *for* $x < 1$. *Then*

$$f(z) = z \int_1^\infty \frac{\varphi(t)}{t^{z+1}} dt$$

for $z \in \mathbb{H}_r$.

Proof. Let $z \in \mathbb{H}_r$ and $n \in \mathbb{N}$. With $\varphi(0) = 0$ we have

$$\begin{aligned}
\sum_{n=1}^{N} \frac{a_n}{n^z} &= \sum_{n=1}^{N} \frac{\varphi(n) - \varphi(n-1)}{n^z} \\
&= \sum_{n=1}^{N} \frac{\varphi(n)}{n^z} - \sum_{n=0}^{N-1} \frac{\varphi(n)}{(n+1)^z} \\
&= \sum_{n=1}^{N-1} \varphi(n) \left(\frac{1}{n^z} - \frac{1}{(n+1)^z} \right) + \frac{\varphi(N)}{N^z} \\
&= \sum_{n=1}^{N-1} \varphi(n) \int_n^{n+1} \frac{z}{t^{z+1}} dt + \frac{\varphi(N)}{N^z} \\
&= z \int_1^N \frac{\varphi(t)}{t^{z+1}} dt + \frac{\varphi(N)}{N^z}.
\end{aligned}$$

For $z = x > r \ge 0$ the left hand side converges as $N \to \infty$. As both terms on the right hand side are non-negative this yields that the integral $\int_1^\infty \varphi(t)/t^{x+1} dt$ converges for $x > r$ and that the limit $\lim_{t\to\infty} \varphi(t)/t^x$ exists. But this implies that this limit is actually equal to 0. In fact, since $|\varphi(t)/t^z| = \varphi(t)/t^{\operatorname{Re} z}$ we have $\lim_{t\to\infty} \varphi(t)/t^z = 0$ if $\operatorname{Re} z > r$, and the conclusion follows. □

To prove the prime number theorem, we also need the following result. It will be proved (and also put into some context) later.

Theorem 5.2. *Let* $f\colon [0, \infty) \to \mathbb{C}$ *be bounded and measurable. Let* $g\colon \mathbb{H}_0 \to \mathbb{C}$ *be the Laplace transform of* f*; that is,*

$$g(z) = \int_0^\infty f(t) e^{-zt} dt.$$

Suppose that g *has a holomorphic continuation to domain containing* $\overline{\mathbb{H}_0}$. *Then the integral*

$$\int_0^\infty f(t) dt = \lim_{T\to\infty} \int_0^T f(t) dt$$

exists and

$$\int_0^\infty f(t) dt = g(0) = \lim_{z\to 0} g(z).$$

Remark. Since $|e^{-zt}| = e^{-t\operatorname{Re} z}$ for $t \in \mathbb{R}$ and $z \in \mathbb{C}$ and since f is bounded it follows from the comparison test that the integral defining g exists for $\operatorname{Re} z > 0$. It is also not difficult to see that g is holomorphic; cf. the proof of Theorem 5.2 that will be given later. The essential hypothesis is that g can be continued holomorphically.

Proof of the prime number theorem (Theorem 3.1). By Lemma 3.1 the conclusion is equivalent to

$$\theta(x) \sim x$$

as $x \to \infty$. It follows from Lemma 5.3 with a substitution that

$$\Phi(z) = z \int_0^\infty \theta(e^s) e^{-sz} ds$$

for $\operatorname{Re} z > 1$. Thus

$$\frac{\Phi(z+1)}{z+1} = \int_0^\infty \theta(e^s) e^{-s} e^{-sz} ds$$

for $\operatorname{Re} z > 0$. Since

$$\frac{1}{z} = \int_0^\infty e^{-sz} ds$$

for $\operatorname{Re} z > 0$ we obtain

$$\frac{\Phi(z+1)}{z+1} - \frac{1}{z} = \int_0^\infty \left(\theta(e^s) e^{-s} - 1\right) e^{-sz} ds.$$

Since

$$\frac{\Phi(z+1)}{z+1} - \frac{1}{z} = \frac{1}{z+1}\left(\Phi(z+1) - \frac{1}{z} - 1\right)$$

we deduce from Lemma 5.2 that the function defined by the left hand side has a holomorphic continuation to a domain containing $\overline{\mathbb{H}_0}$. Theorem 5.2 now implies that the integral

$$\int_0^\infty \left(\theta(e^s) e^{-s} - 1\right) ds = \int_1^\infty \left(\frac{\theta(t)}{t} - 1\right) \frac{dt}{t}$$

converges. Here we used that the integrand is bounded since $\theta(x) = \mathcal{O}(x)$ by Lemma 3.2.

The convergence of this integral now implies that $\theta(x) \sim x$ as $x \to \infty$. In fact, if $\theta(x) > (1+2\varepsilon)x$ for some $\varepsilon > 0$, then

$$\int_x^{(1+\varepsilon)x} \left(\frac{\theta(t)}{t} - 1\right) \frac{dt}{t} \geq \left(\frac{(1+2\varepsilon)x}{(1+\varepsilon)x} - 1\right) \int_x^{(1+\varepsilon)x} \frac{dt}{t} = \frac{\varepsilon}{1+\varepsilon} \log(1+\varepsilon) > 0.$$

Since the integral converges, this cannot happen for arbitrarily large x. Similarly, if $\theta(x) < (1-2\varepsilon)x$, then

$$\int_{(1-\varepsilon)x}^{x} \left(\frac{\theta(t)}{t} - 1\right) \frac{dt}{t} \leq \left(\frac{(1-2\varepsilon)x}{(1-\varepsilon)x} - 1\right) \int_{(1-\varepsilon)x}^{x} \frac{dt}{t} = \frac{\varepsilon}{1-\varepsilon} \log(1-\varepsilon) < 0.$$

Hence $(1-2\varepsilon)x \leq \theta(x) \leq (1+2\varepsilon)x$ for all large x and thus $\theta(x) \sim x$ as $x \to \infty$. $\square$

The proof of Theorem 5.2 was postponed and will be given now.

Proof of Theorem 5.2. For $T > 0$ we consider the function $g_T \colon \mathbb{C} \to \mathbb{C}$,

$$g_T(z) = \int_0^T f(t) e^{-zt} dt.$$

This function is holomorphic (and thus entire). One way to see this is an application of Morera's theorem which says that if D is a domain and $F \colon D \to \mathbb{C}$ is a continuous function satisfying $\int_{\partial\Delta} F(z) dz = 0$ for every triangle $\Delta \subset D$, then F is holomorphic. And we have

$$\int_{\partial\Delta} g_T(z) dz = \int_0^T f(t) \int_{\partial\Delta} e^{-zt} dz\, dt = 0$$

by Cauchy's integral theorem. The same argument also shows that g is holomorphic. Alternatively, this follows from Weierstraß's theorem.

We have to prove that

$$\lim_{T\to\infty} g_T(0) = g(0).$$

In order to do so we consider, for some large R, the function $h = h_{T,R}$ given by

$$h(z) = (g(z) - g_T(z))\, e^{zT} \left(1 + \frac{z^2}{R^2}\right).$$

Recall here that, by hypothesis, g and hence h are holomorphic in some domain D containing $\overline{\mathbb{H}_0}$. Note also that

$$h(0) = g(0) - g_T(0).$$

Given $\varepsilon > 0$ we then have to show that if T is sufficiently large and R is suitably chosen, then $|h(0)| < \varepsilon$.

Cauchy's integral formula yields that if γ is a simple closed curve in D which surrounds 0 once, then

$$h(0) = \int_\gamma \frac{h(z)}{z} dz.$$

We choose γ as the boundary of $\{z \in \mathbb{C} \colon |z| \le R,\ \operatorname{Re} z \ge -\delta\}$ for some $\delta > 0$, noting that γ is contained in D if δ is sufficiently small. By γ_+ and γ_- we denote the parts of γ that are in $\mathbb{H}_0$ and $\mathbb{C} \setminus \mathbb{H}_0$, respectively; see Figure 1.

We first estimate the integral over γ_+. Let $x = \operatorname{Re} z > 0$. With

$$M = \sup_{t\ge 0} |f(t)|$$

we then have

$$|g(z) - g_T(z)| = \left| \int_T^\infty f(t) e^{-zt} dt \right| \le M \int_T^\infty e^{-xt} dt = \frac{M e^{-xT}}{x}.$$

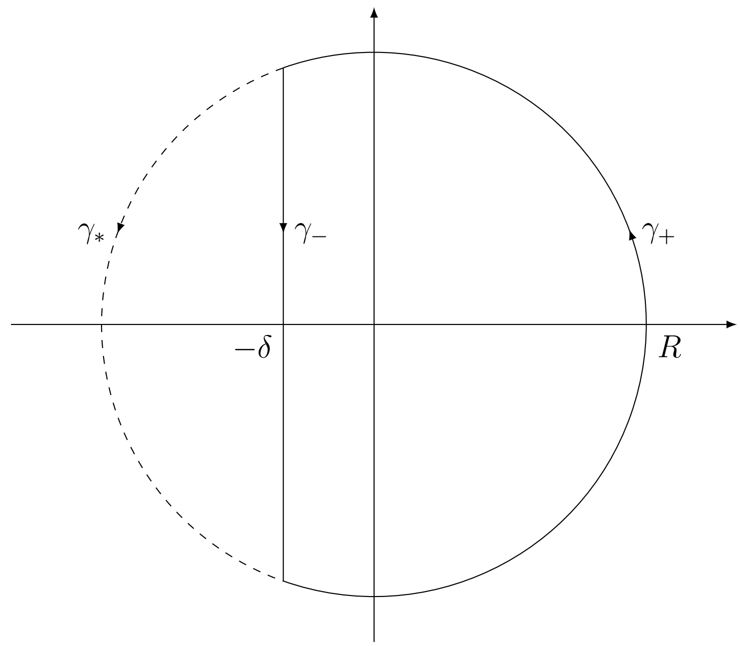

Figure 1: Paths of integration used in the proof of Theorem 5.2.

If $|z| = R$ and thus $R^2 = |z|^2 = z\bar{z}$, then

$$\left| e^{zT}\left(1+\frac{z^2}{R^2}\right)\right| = \left| e^{zT}\left(1+\frac{z^2}{z\bar{z}}\right)\frac{\bar{z}}{R}\right| = \left| e^{zT}\left(\frac{\bar{z}}{R}+\frac{z}{R}\right)\right| = e^{xT}\left|\frac{\bar{z}+z}{R}\right| = \frac{e^{xT}2x}{R}.$$

For z in γ_+ we thus have

$$\left|\frac{h(z)}{z}\right| = |g(z)-g_T(z)|\cdot\left| e^{zT}\left(1+\frac{z^2}{R^2}\right)\right|\cdot\frac{1}{R} = \frac{Me^{-xT}}{x}\cdot\frac{e^{xT}2x}{R}\cdot\frac{1}{R} = \frac{2M}{R^2}.$$

Hence

$$\left|\int_{\gamma_+}\frac{h(z)}{z}dz\right| \le \pi R\frac{2M}{R^2} = \frac{2\pi M}{R} < \varepsilon,$$

if R is chosen sufficiently large.

To estimate the integral over γ_- we write

$$h(z) = g(z)e^{zT}\left(1+\frac{z^2}{R^2}\right) - g_T(z)e^{zT}\left(1+\frac{z^2}{R^2}\right) = u(z)-v(z)$$

and consider the integrals over u and v separately. Since v is entire, in the integral over v we may replace γ_- by the semicircle γ_* connecting iR with $-iR$. In fact we could replace γ_- by any curve in the left half-plane which has the same starting and end point as γ_-.

For $x = \operatorname{Re} z < 0$ we have

$$|g_T(z)| = \left|\int_0^T f(t)e^{-zt}dt\right| \le M\int_0^T e^{-xt}dt = M\left(\frac{e^{-xT}}{|x|}-\frac{1}{|x|}\right) \le \frac{Me^{-xT}}{|x|}.$$

For $x = \operatorname{Re} z < 0$ and $|z| = R$ we thus find as before that

$$\left|\frac{v(z)}{z}\right| \le \frac{Me^{-xT}}{|x|} \cdot \frac{e^{xT}2|x|}{R^2} = \frac{2M}{R^2}$$

and thus

$$\left|\int_{\gamma_-} \frac{v(z)}{z}dz\right| = \left|\int_{\gamma_*} \frac{v(z)}{z}dz\right| \le \pi R\frac{2M}{R^2} = \frac{2\pi M}{R} < \varepsilon.$$

Note that so far the estimates are independent of T. We have only required that R is large. Now we fix this R and a corresponding value of δ. Then there exists $K > 0$ such that

$$\left|g(z)\left(1+\frac{z^2}{R^2}\right)\frac{1}{z}\right| \le K$$

for z in γ_-. Hence

$$\left|\int_{\gamma_-} \frac{u(z)}{z}dz\right| = \left|\int_{\gamma_-} e^{zT}g(z)\left(1+\frac{z^2}{R^2}\right)\frac{1}{z}dz\right| \le K\int_{\gamma_-} \left|e^{zT}\right| |dz|.$$

Since $|e^{zT}| \le 1$ and $|e^{zT}| \to 0$ as $T \to \infty$ if $\operatorname{Re} z < 0$, the theorem about dominated convergence implies that

$$\int_{\gamma_-} \frac{u(z)}{z}dz \to 0 \quad \text{as } T \to \infty.$$

For large T, say $T \ge T_R$, we thus have

$$\left|\int_{\gamma_-} \frac{u(z)}{z}dz\right| < \varepsilon.$$

Altogether we find that

$$\begin{aligned} |h(0)| &= \left|\frac{1}{2\pi i}\int_{\gamma} \frac{h(z)}{z}dz\right| \\ &\le \frac{1}{2\pi}\left(\left|\int_{\gamma_+} \frac{h(z)}{z}dz\right| + \left|\int_{\gamma_-} \frac{v(z)}{z}dz\right| + \left|\int_{\gamma_-} \frac{u(z)}{z}dz\right|\right) \\ &< \frac{1}{2\pi}3\varepsilon < \varepsilon \end{aligned}$$

for $T \ge T_R$. Here we first choose R large to estimate the first two integrals and then take T large for the last integral. □

Remark. Let

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

be a power series with radius of convergence 1. Abel's theorem for power series (German: Abelscher Grenzwertsatz) says that if the series $\sum_{n=0}^{\infty} a_n$ converges, then

$$\sum_{n=0}^{\infty} a_n = \lim_{x \to 1} f(x).$$

In 1897, Alfred Tauber showed conversely that if $\lim_{x\to 1} f(x)$ exists and if $a_n = o(1/n)$ as $n \to \infty$, then the series $\sum_{n=0}^{\infty} a_n$ converges. The condition $a_n = o(1/n)$ was weakened to $a_n = \mathcal{O}(1/n)$ by Littlewood. Further generalizations were obtained by Hardy and Littlewood, who also coined the expression "Tauberian theorem" for results of this type.

Instead of functions defined by infinite series one may also consider functions defined by integrals (or other limit processes). A Tauberian theorem is thus a theorem of the following type:

Let f be defined in a domain D by some series or integral. Suppose that f has a continuous (or holomorphic) continuation to some point $z_0 \in \partial D$. Suppose also that some further hypothesis are satisfied. Then the expression defining f converges also in z_0.

In this sense Theorem 5.2 is a Tauberian theorem.

We have proved the prime number theorem which says that

$$\pi(x) \sim \frac{x}{\log x}$$

as $x \to \infty$. One may ask how good the approximation of $\pi(x)$ by $x/\log x$ actually is. It turns out to be better to consider the integral logarithm

$$\mathrm{Li}(x) := \int_2^x \frac{dt}{\log t}$$

instead of $x/\log x$. Note that

$$\mathrm{Li}(x) \sim \frac{x}{\log x}$$

as $x \to \infty$. This follows directly by l'Hospital's rule. Alternatively, integration by parts yields that

$$\mathrm{Li}(x) = \frac{x}{\log x} - \frac{2}{\log 2} + \int_2^x \frac{dt}{(\log t)^2}$$

and since

$$\int_2^x \frac{dt}{(\log t)^2} = \int_2^{\sqrt{x}} \frac{dt}{(\log t)^2} + \int_{\sqrt{x}}^x \frac{dt}{(\log t)^2} \le \frac{\sqrt{x}-2}{(\log 2)^2} + \frac{x-\sqrt{x}}{(\log\sqrt{x})^2} \le 4\sqrt{x} + \frac{4x}{(\log x)^2}$$

we thus obtain

$$\mathrm{Li}(x) = \frac{x}{\log x} + \mathcal{O}\left(\frac{x}{(\log x)^2}\right) = \left(1 + \mathcal{O}\left(\frac{1}{\log x}\right)\right)\frac{x}{\log x}$$

as $x \to \infty$. Using integration by parts again we obtain

$$\begin{aligned}\mathrm{Li}(x) &= \frac{x}{\log x} - \frac{2}{\log 2} + \int_2^x \frac{dt}{(\log t)^2} \\ &= \frac{x}{\log x} - \frac{2}{\log 2} + \frac{x}{(\log x)^2} - \frac{2}{(\log 2)^2} + 2\int_2^x \frac{dt}{(\log t)^3},\end{aligned}$$

leading to

$$\mathrm{Li}(x) = \frac{x}{\log x} + \frac{x}{(\log x)^2} + \mathcal{O}\left(\frac{x}{(\log x)^3}\right)$$

as $x \to \infty$.

The prime number theorem is thus equivalent to

$$\pi(x) \sim \mathrm{Li}(x)$$

as $x \to \infty$. This has been improved by estimating the difference between $\pi(x)$ and $\mathrm{Li}(x)$. For example, it is known that

$$\pi(x) = \mathrm{Li}(x) + \mathcal{O}\Big(x \exp\Big(-c\sqrt{\log x}\Big)\Big)$$

for some positive constant. This was already shown by de la Vallée Poussin in 1899. Some (slight) improvements have been obtained since then.

Instead of $\mathrm{Li}(x)$ one also considers the principal value

$$\mathrm{li}(x) = \int_0^x \frac{dt}{\log t} := \lim_{\varepsilon \to 0}\left(\int_0^{1-\varepsilon} \frac{dt}{\log t} + \int_{1+\varepsilon}^x \frac{dt}{\log t}\right).$$

Thus

$$\mathrm{Li}(x) = \mathrm{li}(x) - \mathrm{li}(2) = \mathrm{li}(x) - 1.0451\ldots.$$

Numerical data suggest that $\pi(x) < \mathrm{li}(x)$ for all x. However, Littlewood showed in 1914 that $\pi(x) - \mathrm{li}(x)$ changes its sign infinitely often. He did not give an estimate where the first sign change occurs, but his student Skewes showed in 1955 that this happens for some x less than

$$e^{e^{e^{7.705}}} < 10^{10^{10^{10^3}}}.$$

Since then this number has been improved to about 1.398×10^{316}, and it is conjectured that the first change of sign happens about there.

6 Perron's formula

In the proof of the prime number theorem we worked with

$$\theta(x) = \sum_{p \le x} \log p$$

instead of $\pi(x)$. It turns out to be useful to consider also

$$\psi(x) = \sum_{\substack{p \in \mathbb{P}, k \in \mathbb{N} \\ p^k \le x}} \log p.$$

As before we will often omit the condition $p \in \mathbb{P}$ in such sums, but reserve the variable p for prime numbers. We have

$$\psi(x) = \sum_{k=1}^{\infty} \sum_{p^k \le x} \log p = \sum_{k=1}^{\infty} \sum_{p \le x^{1/k}} \log p = \sum_{k=1}^{\infty} \theta\left(x^{1/k}\right).$$

Here the sum is finite since $\theta(x^{1/k}) = 0$ if $x^{1/k} < 2$ which is the case for $k > (\log x)/\log 2$. Since $\theta(x)$ is non-decreasing we thus have

$$\sum_{k=2}^{\infty} \theta\left(x^{1/k}\right) = \theta\left(\sqrt{x}\right) + \sum_{k=3}^{\infty} \theta(x^{1/k}) \le \theta\left(\sqrt{x}\right) + \frac{\log x}{\log 2}\theta\left(x^{1/3}\right).$$

Since $\theta(x) = \mathcal{O}(x)$ by Lemma 3.2 we conclude that

$$\psi(x) = \theta(x) + \mathcal{O}\left(\sqrt{x}\right).$$

For example, this implies that the prime number theorem is also equivalent to $\psi(x) \sim x$.

In the proof of the prime number theorem we considered

$$\Phi(z) = \sum_{p \in \mathbb{P}} \frac{\log p}{p^z}$$

which is closely related to $\theta(x)$; cf. Lemma 5.3. Similarly, $\psi(x)$ is related to the logarithmic derivative ζ'/ζ of ζ which, by Theorem 5.1, satisfies

$$-\frac{\zeta'(z)}{\zeta(z)} = \sum_{k=1}^{\infty} \sum_{p \in \mathbb{P}} \frac{\log p}{p^{kz}}$$

for $z \in \mathbb{H}_1$. We define $\Lambda \colon \mathbb{R} \to \mathbb{R}$ by $\Lambda(x) = \log p$ if x has the form $x = p^n$ for some $p \in \mathbb{P}$ and $n \in \mathbb{N}$, and $\Lambda(x) = 0$ otherwise. Then

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

for $z \in \mathbb{H}_1$ and

$$\psi(x) = \sum_{n \le x} \Lambda(n).$$

The analogue of Lemma 5.3, which also follows immediately from Lemma 5.4, says that

$$-\frac{\zeta'(z)}{\zeta(z)} = z \int_1^{\infty} \frac{\psi(t)}{t^{z+1}} dt.$$

Using this equation instead of Lemma 5.3 one could avoid introducing the function Φ in the proof of the prime number theorem altogether and work with ζ'/ζ directly.

We shall also consider

$$\psi_0(x) := \sum_{n<x} \Lambda(n) + \frac{1}{2}\Lambda(x) = \frac{1}{2}\left(\lim_{y\to x-}\psi(y) + \lim_{y\to x+}\psi(y)\right).$$

Note that $\psi_0(x) = \psi(x)$ unless x has the form $x = p^n$ with $p \in \mathbb{P}$ and $n \in \mathbb{N}$.

We will express $\psi_0(x)$ in terms of the ζ-function. In order to do so we write, for $c \in \mathbb{R}$ and a function $g\colon \{z \in \mathbb{C}\colon \operatorname{Re} z = c\} \to \mathbb{C}$,

$$\int_{c-i\infty}^{c+i\infty} g(z)dz := \lim_{T\to\infty} \int_{[c-iT,c+iT]} g(z)dz = \lim_{T\to\infty} i\int_{-T}^{T} g(c+it)dt.$$

Theorem 6.1. *Let $c > 1$ and $x > 1$. Then*

$$\psi_0(x) = -\frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} \frac{\zeta'(z)}{\zeta(z)}\frac{x^z}{z}dz.$$

In order to prove this we will combine Theorem 5.1 with the following result.

Theorem 6.2. *Let $a > 0$ and $c > 0$. Then*

$$\frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} \frac{a^z}{z}dz = \delta(a) := \begin{cases} 0 & \text{if } 0 < a < 1, \\ \frac{1}{2} & \text{if } a = 1, \\ 1 & \text{if } a > 1. \end{cases}$$

More precisely, if $T > 0$, then

$$\left|\frac{1}{2\pi i}\int_{c-iT}^{c+iT} \frac{a^z}{z}dz - \delta(a)\right| \le \begin{cases} a^c \cdot \min\left\{1, \dfrac{1}{T\log a}\right\} & \text{if } a \neq 1, \\ \dfrac{c}{T} & \text{if } a = 1. \end{cases}$$

Theorem 6.1 and 6.2 are known as *Perron's formula*. The idea in the proof of Theorem 6.1 is to insert the series for ζ'/ζ given in Theorem 5.1 on the right hand side of the formula there and then interchange the order of integration and summation so that Theorem 6.2 can be applied to each term in the sum. The error estimate in Theorem 6.2 will be needed to justify this interchange of integration and summation.

Proof. First we consider the case that $a > 1$ so that $\delta(a) = 1$. Let $R > 0$ and let $\gamma = \gamma_1 + \gamma_2 + \gamma_3 + \gamma_4$ be the positively oriented boundary of the rectangle with vertices $c \pm iT$ and $-R \pm iT$; see Figure 2.

By the residue theorem,

$$\frac{1}{2\pi i}\int_\gamma \frac{a^z}{z}dz = \operatorname{res}\left(\frac{a^z}{z}, 0\right) = a^0 = 1.$$

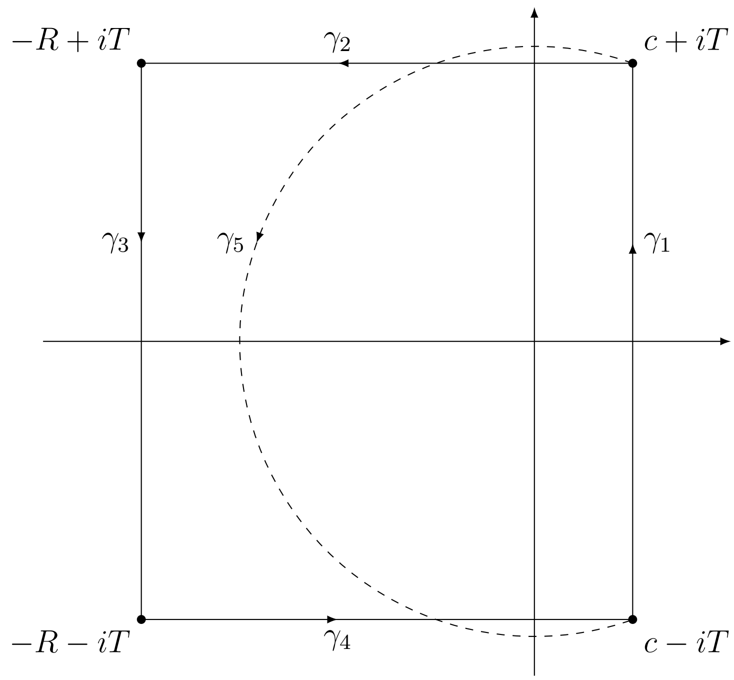

Figure 2: Path of integration used in the proof of Perron's formula.

Also,

$$\left|\int_{\gamma_4} \frac{a^z}{z} dz\right| \leq \int_{-R}^{c} \left|\frac{a^{x-iT}}{x-iT}\right| dx \leq \frac{1}{T}\int_{-R}^{c} a^x dx = \frac{1}{T\log a}\left(a^c - a^{-R}\right).$$

By the same argument,

$$\left|\int_{\gamma_2} \frac{a^z}{z} dz\right| \leq \frac{1}{T\log a}\left(a^c - a^{-R}\right).$$

We also have

$$\left|\int_{\gamma_3} \frac{a^z}{z} dz\right| \leq 2T \max_{|y|\leq T}\left|\frac{a^{-R+iy}}{-R+iy}\right| = \frac{2T}{R}a^{-R}.$$

Thus

$$\left|\frac{1}{2\pi i}\int_{\gamma_1} \frac{a^z}{z} dz - 1\right| \leq \frac{1}{2\pi}\sum_{k=2}^{4}\left|\int_{\gamma_k} \frac{a^z}{z} dz\right| \leq \frac{1}{\pi T\log a}\left(a^c - a^{-R}\right) + \frac{T}{\pi R}a^{-R}.$$

With $R \to \infty$ we thus obtain

$$\left|\frac{1}{2\pi i}\int_{\gamma_1} \frac{a^z}{z} dz - 1\right| \leq \frac{a^c}{\pi T\log a}.$$

Alternatively, we may also consider $\gamma' = \gamma_1 + \gamma_5$ instead of γ, where γ_5 is part of the circle of radius $|c+iT|$ around 0; see Figure 2. This yields that

$$\left|\frac{1}{2\pi i}\int_{\gamma_1} \frac{a^z}{z} dz - 1\right| \leq \frac{1}{2\pi}\left|\int_{\gamma_5} \frac{a^z}{z} dz\right| \leq |c+iT| \cdot \max_{|z|=|c+iT|}\left|\frac{a^z}{z}\right| = a^c.$$

Combining the last two estimates we obtain the conclusion if $a > 1$.

If $a < 1$, we choose $R > c$ and γ as the (negatively) oriented boundary of the rectangle with $c \pm iT$ and $R \pm iT$, and $\gamma' = \gamma_1 + \gamma_6$ where γ_6 is the other part of the circle of radius $|c + iT|$. Now

$$\int_\gamma \frac{a^z}{z} dz = \int_{\gamma'} \frac{a^z}{z} dz = 0$$

by Cauchy's integral theorem, and very similar estimates as before yield the conclusion in this case.

Finally, if $a = 1$, then

$$\begin{aligned}
\frac{1}{2\pi i} \int_{c-iT}^{c+iT} \frac{a^z}{z} dz &= \frac{1}{2\pi i} \int_{c-iT}^{c+iT} \frac{dz}{z} \\
&= \frac{1}{2\pi} \int_{-T}^{T} \frac{dt}{c + it} \\
&= \frac{1}{2\pi} \left(\int_0^T \frac{dt}{c + it} + \int_0^T \frac{dt}{c - it} \right) \\
&= \frac{1}{2\pi} \left(\int_0^T \frac{c - it}{c^2 + t^2} dt + \int_0^T \frac{c + it}{c^2 + t^2} dt \right) \\
&= \frac{1}{\pi} \int_0^T \frac{c}{c^2 + t^2} dt.
\end{aligned}$$

Since

$$\int_0^\infty \frac{c}{c^2 + t^2} dt = \int_0^\infty \frac{ds}{1 + s^2} = \frac{\pi}{2}$$

we obtain

$$0 \leq \frac{1}{2} - \frac{1}{\pi} \int_0^T \frac{c}{c^2 + t^2} dt = \frac{1}{\pi} \int_T^\infty \frac{c}{c^2 + t^2} dt \leq \frac{1}{\pi} \int_T^\infty \frac{c}{t^2} dt = \frac{c}{\pi T}.$$

This yields the conclusion. □

Proof of Theorem 6.1. By Theorem 6.2 we have

$$\begin{aligned}
-\frac{1}{2\pi i} \int_{c-iT}^{c+iT} \frac{\zeta'(z)}{\zeta(z)} \frac{x^z}{z} dz &= \frac{1}{2\pi i} \int_{c-iT}^{c+iT} \sum_{n=1}^\infty \frac{\Lambda(n)}{n^z} \frac{x^z}{z} dz \\
&= \sum_{n=1}^\infty \Lambda(n) \frac{1}{2\pi i} \int_{c-iT}^{c+iT} \left(\frac{x}{n}\right)^z \frac{dz}{z} \\
&= \sum_{n=1}^\infty \Lambda(n) \left(\delta\left(\frac{x}{n}\right) + R(n) \right) \\
&= \psi_0(x) + \sum_{n=1}^\infty \Lambda(n) R(n),
\end{aligned}$$

where $|R(n)| \leq c/T$ if $n = x$ and

$$|R(n)| \leq \left(\frac{x}{n}\right)^c \min\left\{1, \frac{1}{T\log(x/n)}\right\}$$

otherwise. Note that the series converges since

$$|\Lambda(n)R(n)| \leq \log n \cdot \left(\frac{x}{n}\right)^c = x^c \cdot \frac{\log n}{n^c}$$

for $n > x$ and since $c > 1$. The theorem about dominated convergence now yields that

$$\sum_{n=1}^{\infty} \Lambda(n)R(n) \to 0$$

as $T \to \infty$. □

Remark. The same proof shows that if the series

$$f(z) = \sum_{n=1}^{\infty} \frac{a_n}{n^z}$$

converges absolutely for $z \in \mathbb{H}_r$, if

$$\varphi(x) = \sum_{n \leq x} a_n = \sum_{n=1}^{\lfloor x \rfloor} a_n,$$

and if $\varphi_0(x)$ is obtained from $\varphi(x)$ in the same way that $\psi_0(x)$ was obtained from $\psi(x)$, then

$$\varphi_0(x) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} f(z) \frac{x^z}{z} dz$$

for $c > r$.

Our aim is to compute the integral occurring in Theorem 6.1 by the residue theorem. This will express $\psi_0(x)$ in terms of the poles of the integrand which, besides the origin, are precisely the zeros of ζ. In order to do so, we shall show first that ζ can be continued to a function meromorphic in the plane.

7 Bernoulli numbers

In order to extend ζ to a function meromorphic in $\mathbb{C}$ we require some properties of the function $z \mapsto z/(e^z - 1)$. In this section we collect not only those properties that we need for the meromorphic continuation of ζ, but also give some results of independent interest. In particular, we relate the Laurent coefficients of this function to the values of ζ at the even integers.

Definition 7.1. The coefficients B_k in the expansion

$$\frac{z}{e^z - 1} = \sum_{k=0}^{\infty} \frac{B_k}{k!} z^k$$

are called *Bernoulli numbers.*

The B_n can be computed recursively. Indeed, noting that

$$\begin{aligned} z &= (e^z - 1) \sum_{k=0}^{\infty} \frac{B_k}{k!} z^k \\ &= \sum_{j=1}^{\infty} \frac{1}{j!} z^j \sum_{k=0}^{\infty} \frac{B_k}{k!} z^k \\ &= \sum_{n=1}^{\infty} \left(\sum_{k=0}^{n-1} \frac{1}{(n-k)!} \frac{B_k}{k!} \right) z^n \\ &= \sum_{n=1}^{\infty} \frac{1}{n!} \left(\sum_{k=0}^{n-1} \binom{n}{k} B_k \right) z^n \end{aligned}$$

we find that $B_0 = 1$ and

$$\sum_{k=0}^{n-1} \binom{n}{k} B_k = 0$$

for $n \geq 2$. Replacing n by $n+1$ in this formula and solving for B_n we obtain the following result.

Lemma 7.1. *Let $n \in \mathbb{N}$. Then*

$$B_n = -\frac{1}{n+1} \sum_{k=0}^{n-1} \binom{n+1}{k} B_k.$$

Using this recursion formula we easily find that $B_0 = 1$, $B_1 = -\frac{1}{2}$ and $B_2 = \frac{1}{6}$.

The following results show that the Laurent coefficients of the cotangent are closely related to the Bernoulli numbers.

Lemma 7.2. *Let $z \in \mathbb{C} \setminus \{2\pi i k \colon k \in \mathbb{Z}\}$. Then*

$$\frac{1}{e^z - 1} + \frac{1}{2} = \frac{1}{2} i \cot\left(\frac{1}{2} i z\right).$$

Proof. We have

$$\begin{aligned}
\frac{1}{2}i\cot\left(\frac{1}{2}iz\right) &= \frac{1}{2}i\cdot\frac{\cos\left(\frac{1}{2}iz\right)}{\sin\left(\frac{1}{2}iz\right)}\\
&= \frac{1}{2}i\cdot\frac{\frac{1}{2}\left(e^{i(iz/2)}+e^{-i(iz/2)}\right)}{\frac{1}{2i}\left(e^{i(iz/2)}-e^{-i(iz/2)}\right)}\\
&= -\frac{1}{2}\cdot\frac{e^{-z/2}+e^{z/2}}{e^{-z/2}-e^{z/2}}\\
&= -\frac{1}{2}\cdot\frac{1+e^z}{1-e^z}\\
&= \frac{1}{2}\cdot\frac{2+e^z-1}{e^z-1}\\
&= \frac{1}{e^z-1}+\frac{1}{2}.
\end{aligned}$$

□

Lemma 7.3. *The Laurent series of cotangent around* 0 *is*

$$\cot z = \sum_{k=0}^{\infty}(-1)^k\frac{2^{2k}B_{2k}}{(2k)!}z^{2k-1} = \frac{1}{z}+\sum_{k=1}^{\infty}(-1)^k\frac{2^{2k}B_{2k}}{(2k)!}z^{2k-1}.$$

Proof. Since the cotangent is odd with a simple pole at 0 we have an expansion

$$\cot z = \sum_{k=0}^{\infty}c_k z^{2k-1}$$

with certain coefficients c_k. Lemma 7.2 now yields that

$$\begin{aligned}
\sum_{k=0}^{\infty}\frac{B_k}{k!}z^k &= \frac{z}{e^z-1}\\
&= \frac{1}{2}iz\cot\left(\frac{1}{2}iz\right)-\frac{1}{2}z\\
&= \frac{1}{2}iz\sum_{k=0}^{\infty}c_k\left(\frac{1}{2}iz\right)^{2k-1}-\frac{1}{2}z\\
&= \sum_{k=0}^{\infty}c_k\left(\frac{1}{2}i\right)^{2k}z^{2k}-\frac{1}{2}z\\
&= \sum_{k=0}^{\infty}c_k\frac{(-1)^k}{2^{2k}}z^{2k}-\frac{1}{2}z
\end{aligned}$$

so that the conclusion follows by comparing coefficients. □

The argument in the proof also shows that $B_k = 0$ if k is odd and $k \geq 3$.

Lemma 7.4. *Let* $z \in \mathbb{C}\setminus\mathbb{Z}$. *Then*

$$\pi\cot(\pi z) = \frac{1}{z}+\sum_{n\in\mathbb{Z}\setminus\{0\}}\left(\frac{1}{z-n}+\frac{1}{n}\right) = \frac{1}{z}+\sum_{n=1}^{\infty}\frac{2z}{z^2-n^2}.$$

Proof. To see that the first series converges locally uniformly in $\mathbb{C}$ we only have to note that if $|z| \le R$ and $|n| \ge 2R$, then

$$\left|\frac{1}{z-n} + \frac{1}{n}\right| = \frac{|z|}{|(z-n)n|} \le \frac{R}{(|n|-|z|)|n|} \le \frac{2R}{n^2}.$$

The second series is obtained by combining the terms for n and $-n$, noting that

$$\frac{1}{z-n} + \frac{1}{z+n} = \frac{2z}{z^2-n^2}.$$

Let

$$f(z) = \frac{1}{z} + \sum_{n \in \mathbb{Z}\setminus\{0\}} \left(\frac{1}{z-n} + \frac{1}{n}\right) - \pi\cot(\pi z)$$

be the difference of the right and left hand side of the equation claimed. Since $\pi\cot(\pi z)$ has simple poles at the integers with residue 1, and no other poles, it follows that f is entire. We have to prove that $f = 0$.

Since

$$\frac{d}{dz}\pi\cot(\pi z) = \pi\frac{d}{dz}\frac{\cos(\pi z)}{\sin(\pi z)} = \pi\frac{-\pi\sin^2(\pi z) - \pi\cos^2(\pi z)}{\sin^2(\pi z)} = -\frac{\pi^2}{\sin^2(\pi z)}$$

and

$$\frac{d}{dz}\left(\frac{1}{z} + \sum_{n \in \mathbb{Z}\setminus\{0\}}\left(\frac{1}{z-n}+\frac{1}{n}\right)\right) = -\sum_{n\in\mathbb{Z}}\frac{1}{(z-n)^2}$$

we find that

$$f'(z) = \frac{\pi^2}{\sin^2(\pi z)} - \sum_{n\in\mathbb{Z}}\frac{1}{(z-n)^2}.$$

It was proved in the exercises that the right hand side is equal to 0, in the course of the proof of the identity

$$\sum_{n=1}^{\infty}\frac{1}{n^2} = \frac{\pi^2}{6}.$$

For completeness we repeat the argument here.

First we note that $g := f'$ is periodic with period 1. It thus suffices to prove that g is bounded in $\{z\colon 0 \le \operatorname{Re} z \le 1\}$. So let $z = x + iy$ with $0 \le x \le 1$. Then

$$\left|\frac{1}{(z-n)^2}\right| = \frac{1}{(x-n)^2+y^2} \le \begin{cases} \dfrac{1}{(n-1)^2+y^2} & \text{if } n \ge 1, \\ \dfrac{1}{n^2+y^2} & \text{if } n \le 0. \end{cases}$$

Thus

$$\left|\sum_{n\in\mathbb{Z}}\frac{1}{(z-n)^2}\right| \le \frac{2}{y^2} + 2\sum_{n=1}^{\infty}\frac{1}{n^2+y^2}$$

and the theorem about dominated convergence yields that

$$\sum_{n\in\mathbb{Z}} \frac{1}{(z-n)^2} \to 0 \quad \text{as } |y| = |\operatorname{Im} z| \to \infty.$$

We also have

$$\frac{\pi^2}{\sin^2(\pi z)} \to 0 \quad \text{as } |\operatorname{Im} z| \to \infty$$

and hence $g(z) \to 0$ as $|\operatorname{Im} z| \to \infty$. By Liouville's theorem, g is constant and in fact $g = f' = 0$. Thus f is constant. Since f is odd, this yields that $f = 0$. $\square$

Lemmas 7.3 and 7.4 yield that for $m \in \mathbb{N}$ we have

$$\begin{aligned}
(-1)^m \pi^{2m} 2^{2m} \frac{B_{2m}}{2m} &= \frac{d^{2m-1}}{dz^{2m-1}} \left(\pi \cot(\pi z) - \frac{1}{z} \right) \bigg|_{z=0} \\
&= \frac{d^{2m-1}}{dz^{2m-1}} \sum_{n\in\mathbb{Z}\setminus\{0\}} \left(\frac{1}{z-n} + \frac{1}{n} \right) \bigg|_{z=0} \\
&= -(2m-1)! \sum_{n\in\mathbb{Z}\setminus\{0\}} \frac{1}{(z-n)^{2m}} \bigg|_{z=0} \\
&= -2(2m-1)!\, \zeta(2m).
\end{aligned}$$

We thus have the following result.

Theorem 7.1. *Let $m \in \mathbb{N}$. Then*

$$\zeta(2m) = \sum_{n=1}^{\infty} \frac{1}{n^{2m}} = (-1)^{m+1} \frac{2^{2m-1}\pi^{2m} B_{2m}}{(2m)!}.$$

For $m = 1$ this is the formula mentioned already before Lemma 2.2 and in the proof of Lemma 7.4.

The following lemma will be used to obtain a meromorphic continuation of ζ.

Lemma 7.5. *Let $z \in \mathbb{C} \setminus \{2\pi i k \colon k \in \mathbb{Z}\}$. Then*

$$\frac{1}{e^z - 1} - \frac{1}{z} + \frac{1}{2} = 2z \sum_{n=1}^{\infty} \frac{1}{z^2 + 4\pi^2 n^2}.$$

Proof. By Lemma 7.2 we have, with $w = iz/2\pi$ and thus $z = -2\pi i w$,

$$\begin{aligned}
\frac{1}{e^z - 1} - \frac{1}{z} + \frac{1}{2} &= \frac{1}{2} i \cot\left(\frac{1}{2} iz\right) - \frac{1}{z} \\
&= \frac{1}{2\pi} i \left(\pi \cot(\pi w) - \frac{1}{w}\right) \\
&= \frac{1}{2\pi} i \sum_{n=1}^{\infty} \frac{2w}{w^2 - n^2} \\
&= \frac{iw}{\pi} \sum_{n=1}^{\infty} \frac{1}{-\frac{z^2}{4\pi^2} - n^2} \\
&= -\frac{z}{2\pi^2} \sum_{n=1}^{\infty} \frac{1}{-\frac{z^2}{4\pi^2} - n^2} \\
&= 2z \sum_{n=1}^{\infty} \frac{1}{z^2 + 4\pi^2 n^2}.
\end{aligned}$$

□

8 The functional equation of the zeta function

In Definition 4.1 we defined $\zeta(z)$ by the series

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

which converges for $\operatorname{Re} z > 1$. Theorem 4.1 says that ζ extends to a function meromorphic in $\mathbb{H}_0 = \{z\colon \operatorname{Re} z > 0\}$ and in fact that the function

$$z \mapsto \zeta(z) - \frac{1}{1 - z}$$

has a holomorphic continuation to $\mathbb{H}_0$. In this section we show that this function actually extends to an entire function, i.e., a function holomorphic in $\mathbb{C}$.

In order to do so we introduce the Gamma function.

Definition 8.1. The function

$$\Gamma\colon \mathbb{H}_0 \to \mathbb{C}, \quad \Gamma(z) = \int_0^{\infty} t^{z-1} e^{-t} dt$$

is called *Gamma function*.

It is easy to see that the integral converges for $\operatorname{Re} z > 0$. It is also not difficult to show that Γ is holomorphic. For example, this can be done using Morera's theorem as in the proof of Theorem 5.2.

Let $0 < r < R < \infty$. Integration by parts yields that

$$\int_r^R t^{z-1} e^{-t} dt = \frac{1}{z} t^z e^{-t} \Big|_{t=r}^{t=R} + \frac{1}{z} \int_r^R t^z e^{-t} dt.$$

For $\operatorname{Re} z > 0$ we conclude with $r \to 0$ and $R \to \infty$ that

$$\Gamma(z) = \frac{1}{z}\Gamma(z+1)$$

and hence

$$\Gamma(z+1) = z\Gamma(z).$$

Induction yields that

$$\Gamma(z+n) = z(z+1) \cdot \ldots \cdot (z+n-1)\Gamma(z)$$

and thus

$$\Gamma(z) = \frac{1}{z(z+1) \cdot \ldots \cdot (z+n-1)}\Gamma(z+n)$$

for $z \in \mathbb{H}_0$ and $n \in \mathbb{N}$. However, the right hand side of the last expression is meromorphic in $\mathbb{H}_{-n}$. Hence we obtain the following result.

Theorem 8.1. *The Gamma function extends to a function meromorphic in $\mathbb{C}$ and satisfies the functional equation $\Gamma(z+1) = z\Gamma(z)$.*

We have

$$\Gamma(1) = \int_0^\infty e^{-t} dt = \lim_{R\to\infty} \left(-e^{-t}\right)\Big|_{t=0}^{t=R} = 1$$

and hence

$$\Gamma(n) = 1 \cdot 2 \cdot \ldots \cdot (n-1) \cdot \Gamma(1) = (n-1)!$$

for $n \in \mathbb{N}$. There are various equivalent definitions of the Gamma function. Here we only mention one of them.

Theorem 8.2. *Let $z \in \mathbb{C} \setminus \{0, -1, -2, \dots\}$. Then*

$$\Gamma(z) = \lim_{n\to\infty} \frac{n^z}{z} \prod_{k=1}^{n} \frac{k}{k+z} = \lim_{n\to\infty} n^z \frac{1 \cdot 2 \cdot \dots \cdot n}{z(z+1)\cdots(z+n)}.$$

The proof is based on the equation

$$\int_0^n \left(1 - \frac{t}{n}\right)^n t^{z-1} dt = n^z \int_0^1 (1-u)^n u^{z-1} du = \frac{n^z}{z} \prod_{k=1}^{n} \frac{k}{k+z}$$

which is proved by induction. We omit the details.

In particular it follows from Theorem 8.2 that the Gamma function has no zeros, and simple poles at $0, -1, -2, \dots$.

Theorem 8.3. *Let $z \in \mathbb{H}_1$. Then*

$$\Gamma(z)\zeta(z) = \int_0^\infty \frac{t^{z-1}}{e^t - 1} dt.$$

Proof. The substitution $t = ns$ yields that

$$\Gamma(z)n^{-z} = \int_0^\infty t^{z-1} n^{-z} e^{-t} dt = \int_0^\infty s^{z-1} e^{-ns} ds$$

for $n \in \mathbb{N}$. It follows that

$$\begin{aligned}\Gamma(z)\zeta(z) &= \sum_{n=1}^\infty \Gamma(z)n^{-z} \\ &= \sum_{n=1}^\infty \int_0^\infty s^{z-1} n^{-z} e^{-s} ds \\ &= \sum_{n=1}^\infty \int_0^\infty t^{z-1} e^{-nt} dt \\ &= \int_0^\infty t^{z-1} \sum_{n=1}^\infty e^{-nt} dt.\end{aligned}$$

To justify interchanging integration and summation, and to simplify the last integral, we note that

$$\sum_{n=1}^\infty e^{-nt} = \sum_{n=1}^\infty \left(e^{-t}\right)^n = \frac{e^{-t}}{1-e^{-t}} = \frac{1}{e^t - 1}.$$

Inserting this above yields the formula claimed, and since

$$t \mapsto \frac{t^{\operatorname{Re} z - 1}}{e^t - 1}$$

is an integrable majorant, this also justifies interchanging integration and summation. ☐

We will use Theorem 8.3 to show that $\Gamma \cdot \zeta$ has a meromorphic extension to $\mathbb{C}$. Of course, this implies that ζ also extends to a function meromorphic in $\mathbb{C}$.

Theorem 8.4. *The Riemann zeta function has a meromorphic extension to* $\mathbb{C}$.

Proof. The function $t \mapsto 1/(e^t - 1)$ occurring in the integrand of the formula

$$\Gamma(z)\zeta(z) = \int_0^\infty \frac{t^{z-1}}{e^t - 1} dt$$

from Theorem 8.3 was studied in the previous section. Lemmas 7.2 and 7.5 show in particular that we have an expansion

$$\frac{1}{e^z - 1} - \frac{1}{z} = \sum_{k=0}^\infty a_k z^k$$

convergent in some neighborhood of 0 with certain coefficients a_k and $a_0 = -\frac{1}{2}$. In fact, the series converges for $|z| < 2\pi$ and we have $a_k = B_k/(k+1)!$.

In order to obtain an analytic continuation of $\Gamma \cdot \zeta$, we note that the above integral from Theorem 8.3 converges at ∞ for all $z \in \mathbb{C}$. And for $n \in \mathbb{N}_0$ and $\operatorname{Re} z > 1$ we have

$$\begin{aligned}\int_0^1 \frac{t^{z-1}}{e^t - 1} dt &= \int_0^1 \left(\frac{1}{e^t - 1} - \frac{1}{t} - \sum_{k=0}^{n} a_k t^k + \frac{1}{t} + \sum_{k=0}^{n} a_k t^k \right) t^{z-1} dt \\ &= \int_0^1 \left(\frac{1}{e^t - 1} - \frac{1}{t} - \sum_{k=0}^{n} a_k t^k \right) t^{z-1} dt + \int_0^1 \left(t^{z-2} + \sum_{k=0}^{n} a_k t^{k+z-1} \right) dt \\ &= \int_0^1 \left(\frac{1}{e^t - 1} - \frac{1}{t} - \sum_{k=0}^{n} a_k t^k \right) t^{z-1} dt + \frac{1}{z-1} + \sum_{k=0}^{n} \frac{a_k}{z+k}.\end{aligned}$$

Since

$$\frac{1}{e^t - 1} - \frac{1}{t} - \sum_{k=0}^{n} a_k t^k = \mathcal{O}\big(t^{n+1}\big)$$

as $t \to 0$ we see that the integral on the right hand side actually converges for $\operatorname{Re} z > -n-1$. Moreover, the same argument used to show that Γ is holomorphic in $\mathbb{H}_0$ shows that the integral is a holomorphic function of z for $\operatorname{Re} z > -n-1$. Since $n \in \mathbb{N}$ is arbitrary the conclusion follows. $\square$

We consider the above formula for $n = 0$ in more detail. Since $a_0 = -\frac{1}{2}$ and

$$\Gamma(z)\zeta(z) = \int_0^1 \frac{t^{z-1}}{e^t - 1} dt + \int_1^\infty \frac{t^{z-1}}{e^t - 1} dt$$

for $\operatorname{Re} z > 1$ we have

$$\Gamma(z)\zeta(z) = \int_0^1 \left(\frac{1}{e^t - 1} - \frac{1}{t} + \frac{1}{2} \right) t^{z-1} dt + \frac{1}{z-1} - \frac{1}{2z} + \int_1^\infty \frac{t^{z-1}}{e^t - 1} dt$$

for $\operatorname{Re} z > -1$. Now

$$\int_1^\infty \left(-\frac{1}{t} \right) t^{z-1} dt = -\int_1^\infty t^{z-2} dt = \frac{1}{z-1}$$

for $\operatorname{Re} z < 1$ and

$$\int_1^\infty \frac{1}{2} t^{z-1} dt = -\frac{1}{2z}$$

for $\operatorname{Re} z < 0$. Inserting this in the last formula we obtain the following result.

Theorem 8.5. *If* $-1 < \operatorname{Re} z < 0$, *then*

$$\Gamma(z)\zeta(z) = \int_0^\infty \left(\frac{1}{e^t - 1} - \frac{1}{t} + \frac{1}{2} \right) t^{z-1} dt.$$

We now insert Lemma 7.5 into this formula and obtain, for $-1 < \operatorname{Re} z < 0$,

$$\Gamma(z)\zeta(z) = \int_0^\infty 2t \sum_{n=1}^\infty \frac{1}{t^2 + 4\pi^2 n^2} t^{z-1} dt = 2 \int_0^\infty \sum_{n=1}^\infty \frac{t^z}{t^2 + 4\pi^2 n^2} dt.$$

Similarly as in the proof of Theorem 8.3 the theorem about dominated convergence can be used to show that we can interchange integration and summation. Hence

$$\begin{aligned}\Gamma(z)\zeta(z) &= 2 \sum_{n=1}^\infty \int_0^\infty \frac{t^z}{t^2 + 4\pi^2 n^2} dt \\ &= 2 \sum_{n=1}^\infty (2\pi n)^{z-1} \int_0^\infty \frac{s^z}{s^2+1} ds \\ &= 2(2\pi)^{z-1} \zeta(1-z) \int_0^\infty \frac{s^z}{s^2+1} ds.\end{aligned}$$

The value of the integral on the right hand side is given by the following lemma.

Lemma 8.1. *Let $-1 < \alpha < 0$. Then*

$$\int_0^\infty \frac{t^\alpha}{t^2+1} dt = \frac{\pi}{2\cos\left(\frac{\pi}{2}\alpha\right)}.$$

Sketch of proof. The computation of this integral is in fact a standard application of the residue theorem. We consider

$$f(z) = \frac{z^\alpha}{z^2+1}$$

where

$$z^\alpha = \exp(\alpha \log z) = \exp(\alpha(\log|z| + i \arg z)) = |z|^\alpha \exp(i \arg z)$$

with $0 < \arg z < 2\pi$. We integrate f over the path $\gamma = \gamma_1 + \gamma_2 + \gamma_3 + \gamma_4$ sketched in Figure 3.

The residue theorem yields that

$$\int_\gamma f(z) dz = 2\pi i (\operatorname{res}(f, i) + \operatorname{res}(f, -i)).$$

We have

$$\int_{\gamma_2} f(z) dz \to \int_r^R \frac{t^\alpha}{t^2+1} dt \quad \text{as } \varepsilon \to 0$$

and

$$\int_{\gamma_4} f(z) dz \to -e^{i2\alpha\pi} \int_r^R \frac{t^\alpha}{t^2+1} dt \quad \text{as } \varepsilon \to 0.$$

Standard estimates show that

$$\int_{\gamma_1} f(z) dz \to 0 \quad \text{as } r \to 0$$

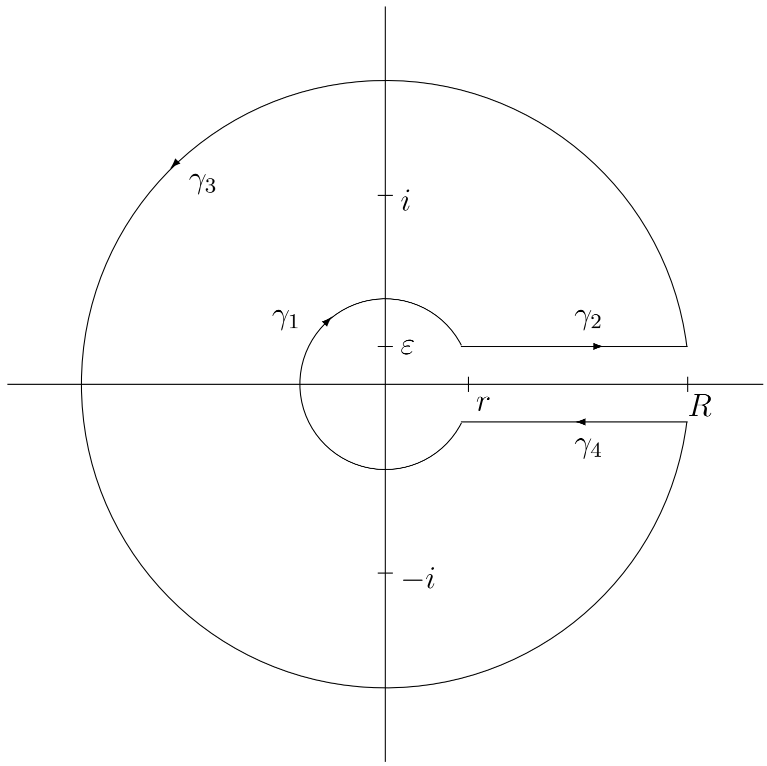

Figure 3: Path of integration in the proof of Lemma 8.1.

and

$$\int_{\gamma_3} f(z)dz \to 0 \quad \text{as } R \to \infty.$$

Moreover,

$$\operatorname{res}(f, i) = \lim_{z\to i}(z-i)f(z) = \lim_{z\to i}\frac{z^\alpha}{z+i} = \frac{e^{i\alpha\pi/2}}{2i}$$

and

$$\operatorname{res}(f, -i) = \lim_{z\to -i}(z+i)f(z) = \lim_{z\to -i}\frac{z^\alpha}{z-i} = \frac{e^{i3\alpha\pi/2}}{-2i}.$$

Altogether we thus have

$$\left(1-e^{i2\alpha\pi}\right)\int_0^\infty \frac{t^\alpha}{t^2+1}dt = 2\pi i\left(\frac{e^{i\alpha\pi/2}}{2i}+\frac{e^{i3\alpha\pi/2}}{-2i}\right) = \pi\left(e^{i\alpha\pi/2}-e^{i3\alpha\pi/2}\right)$$

and hence

$$\begin{aligned}\int_0^\infty \frac{t^\alpha}{t^2+1}dt &= \pi\frac{e^{i\alpha\pi/2}-e^{i3\alpha\pi/2}}{1-e^{i2\alpha\pi}}\\ &= \pi\frac{e^{-i\alpha\pi/2}-e^{i\alpha\pi/2}}{e^{-i\alpha\pi}-e^{i\alpha\pi}}\\ &= \pi\frac{e^{-i\alpha\pi/2}-e^{i\alpha\pi/2}}{e^{-i\alpha\pi}-e^{i\alpha\pi}}\end{aligned}$$

$$= \pi \frac{\sin\left(\frac{\pi}{2}\alpha\right)}{\sin(\alpha\pi)}$$
$$= \frac{\pi}{2\cos\left(\frac{\pi}{2}\alpha\right)} \qquad \square$$

Combining Lemma 8.1 with the formula before this lemma we find that

$$\Gamma(z)\zeta(z) = 2(2\pi)^{z-1}\zeta(1-z)\frac{\pi}{2\cos\left(\frac{\pi}{2}z\right)} = \frac{2^{z-1}\pi^z}{\cos\left(\frac{\pi}{2}z\right)}\zeta(1-z).$$

First this only holds for $-1 < \operatorname{Re} z < 0$, but since both sides are meromorphic in $\mathbb{C}$ it actually holds for all $z \in \mathbb{C}$ (assigning the value ∞ to a function where it has a pole).

We have thus obtained the following result.

Theorem 8.6. *The Riemann zeta function satisfies the functional equation*

$$\zeta(1-z) = 2^{1-z}\pi^{-z}\Gamma(z)\zeta(z)\cos\left(\frac{\pi}{2}z\right).$$

Replacing z by $1-z$ and noting that $\cos(\pi(1-z)/2) = \sin(\pi z/2)$ we can also write this equation as

$$\zeta(z) = 2^z\pi^{z-1}\Gamma(1-z)\zeta(1-z)\sin\left(\frac{\pi}{2}z\right).$$

There are many variants of this equation, some of which we will discuss below. However, we already note one simple consequence.

Theorem 8.7. *Let* $\operatorname{Re} z \leq 0$. *Then* $\zeta(z) = 0$ *if and only if* z *has the form* $z = -2n$ *for some* $n \in \mathbb{N}$.

Proof. Put $w = 1-z$ so that $z = 1-w$. Then $\operatorname{Re} w \geq 1$ and

$$\zeta(z) = 2^z\pi^{z-1}\Gamma(w)\zeta(w)\sin\left(\frac{\pi}{2}z\right).$$

Suppose first that $z \neq 0$ so that $w \neq 1$. By Theorems 4.3 and 4.4 we have $\zeta(w) \neq 0$ and by Theorem 8.2 (and the remark following it) we have $\Gamma(w) \neq 0$. Hence $\zeta(z) = 0$ if and only if $\sin\left(\frac{\pi}{2}z\right) = 0$. This is the case if and only if z is an even integer. Since we assumed that $\operatorname{Re} z \leq 0$ and $z \neq 0$ we conclude that z has the form $z = -2n$ with some $n \in \mathbb{N}$.

To see that $\zeta(0) \neq 0$ we note that the pole of $\zeta(1-z)$ at $z = 0$ and the zero of $\sin\left(\frac{\pi}{2}z\right)$ at this point cancel. $\square$

Using the functional equation and the values of ζ at the even positive integers one may also compute the value of ζ at the negative odd integers.

Theorem 8.8. *Let* $m \in \mathbb{N}$. *Then*

$$\zeta(1-2m) = -\frac{B_{2m}}{2m}.$$

Proof. Using Theorems 7.1 and 8.6 we see that

$$\begin{aligned}\zeta(1-2m) &= 2^{1-2m}\pi^{-2m}\Gamma(2m)\zeta(2m)\cos(m\pi)\\ &= 2^{1-2m}\pi^{-2m}(2m-1)!\,(-1)^{m+1}\frac{2^{2m-1}\pi^{2m}B_{2m}}{(2m)!}(-1)^m\\ &= -\frac{B_{2m}}{2m}.\end{aligned}$$

□

For example, it follows from Theorem 8.8 that $\zeta(-1) = -1/12$.

Remark. The Indian mathematician Ramanujan, in a letter to Hardy, wrote that

$$1+2+3+4+\ldots = -\frac{1}{12}.$$

Hardy recognized that this equation does make sense if the left hand side is interpreted as $\zeta(-1)$.

We briefly mention that there is a branch of analysis, called summability theory (German: Limitierungstheorie), which studies under which hypothesis one can assign a meaningful sum to a divergent series (or a limit to a divergent sequence). We have already seen one method to do so. To a series $\sum_{n=0}^{\infty} a_n$ one associates the value

$$\lim_{x\to 1}\sum_{n=0}^{\infty} a_n x^n.$$

Abel's limit theorem says that for a convergent series $\sum_{n=0}^{\infty} a_n$ this gives the "correct" value of the sum. For example, this method associates to the divergent series $1-2+3-4-\ldots$ the value

$$\lim_{x\to 1}\sum_{n=1}^{\infty}(-1)^{n+1}nx^n = \lim_{x\to 1}\frac{x}{(1+x)^2} = \frac{1}{4}.$$

In the exercises we showed that

$$\left(1-2^{1-z}\right)\zeta(z) = \sum_{n=1}^{\infty}\frac{(-1)^{n+1}}{n^z}.$$

Inserting $z=-1$, and identifying the sum of the divergent series on the right hand side with the value given by Abel's method, yields that

$$-3\zeta(-1) = \sum_{n=1}^{\infty}(-1)^{n+1}n = \frac{1}{4}$$

and hence that

$$1+2+3+4+\ldots = \zeta(-1) = -\frac{1}{3}(1-2+3-4-\ldots) = -\frac{1}{3}\cdot\frac{1}{4} = -\frac{1}{12}.$$

Of course, here the equality signs have to be understood in the sense explained above. Ramanujan, in his letter to Hardy, actually described two methods to assign a finite value to this series, one method being the one described above.

One may write the equation in Theorem 8.6 using the formulas

$$\Gamma(z)\Gamma(1-z) = \frac{\pi}{\sin(\pi z)}$$

and

$$\Gamma(z)\Gamma\left(z+\frac{1}{2}\right) = 2^{1-2z}\sqrt{\pi}\Gamma(2z)$$

which are proved in the exercises. (Actually, the first formula follows by multiplying the formula of Theorem 8.6 with the formula after this theorem, but there are much more direct ways to deduce it.) Some computations will lead to the symmetric form

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

Definition 8.2. The function ξ defined by

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

is called *Riemann xi function*.

Theorem 8.9. *The function ξ is entire and satisfies $\xi(z) = \xi(1-z)$ for all $z \in \mathbb{C}$.*

Proof. The equation $\xi(z) = \xi(1-z)$ follows immediately from the equation before Definition 8.2.

Poles can only occur at the pole of $\zeta(z)$ which is at 1 and at the poles of $\Gamma\left(\frac{1}{2}z\right)$, which are at $0, -2, -4, \ldots$. The pole at 1 is canceled by the term $z-1$, the pole at 0 by the term z, and those at $-2, -4, \ldots$ by the zeros of ζ given by Theorem 8.7. $\square$

9 The explicit formula

As already mentioned, we will use the residue theorem and Theorem 6.1 to express $\psi_0(x)$ in terms of the zeros of ζ. The result is as follows.

Theorem 9.1. *Let $x > 1$. Then*

$$\psi_0(x) = x - \sum_{\substack{\rho:\,\zeta(\rho)=0 \\ 0<\operatorname{Re}\rho<1}} \frac{x^\rho}{\rho} - \frac{\zeta'(0)}{\zeta(0)} - \frac{1}{2}\log\left(1-\frac{1}{x^2}\right).$$

Here the zeros of ζ are counted according to multiplicity; that is, a zero of multiplicity m appears m times in the sum. We remark, however, that it is conjectured that all zeros of ζ are simple. We also mention that it can be shown that $\zeta'(0)/\zeta(0) = \log(2\pi)$.

We note that if $\rho = \beta + i\gamma$ is a zero of ζ, then so is $\bar\rho = \beta - i\gamma$. Assuming the Riemann hypothesis we have $\beta = \frac{1}{2}$ for all ρ occurring in the sum in Theorem 9.1. A short computation shows that

$$\frac{x^\rho}{\rho} + \frac{x^{\bar\rho}}{\bar\rho} = 4\sqrt{x}\frac{\cos(\gamma\log x) + 2\gamma\sin(\gamma\log x)}{1+4\gamma^2}.$$

Thus – in some sense – this sum is similar to a Fourier series. Figure 4 shows the approximation obtained by considering only the first 30 pairs of zeros in the sum. As $\psi(x) \sim x$ as $x \to \infty$, the line $y = x$ is also shown.

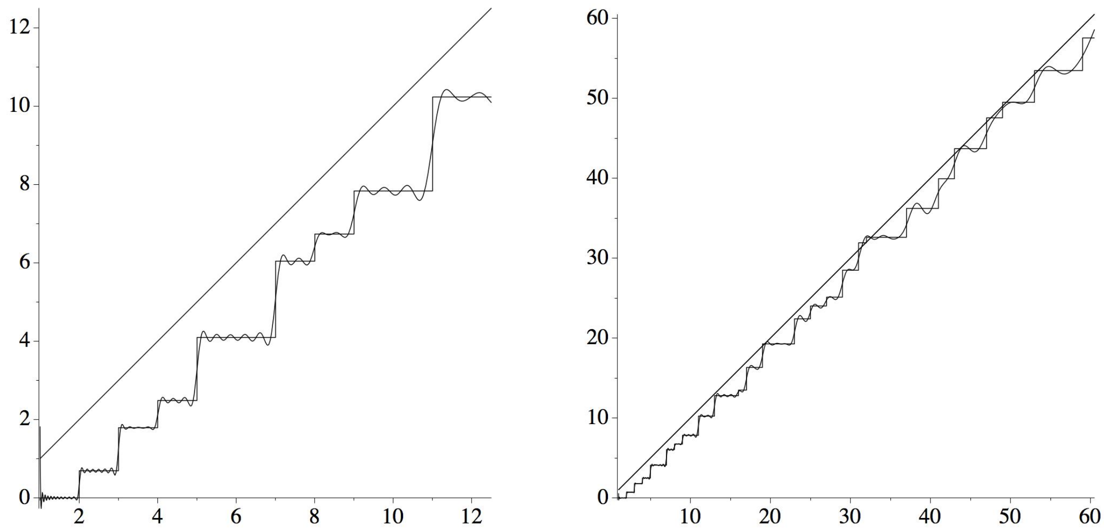

Figure 4: Approximation of $\psi(x)$ with the explicit formula.

Proof of Theorem 9.1, Part 1. We take $c > 1$, $T > 0$, $N \in \mathbb{N}$, $R = 2N + 1$ and, as in the proof of Theorem 6.2, we consider the path $\gamma = \gamma_1 + \gamma_2 + \gamma_3 + \gamma_4$ which parametrizes the boundary of the rectangle with vertices at $c \pm iT$ and $-R \pm iT$; cf. Figure 2. Put

$$f(z) = \frac{\zeta'(z)}{\zeta(z)} \cdot \frac{x^z}{z}.$$

The residue theorem says that

$$\frac{1}{2\pi i} \int_\gamma f(z) dz = \sum_{z \in P} \mathrm{res}(f, z),$$

where P denotes the set of poles of f in the interior of the rectangle. Here we assume that T is chosen such that there are no poles of f on γ.

The poles of f are at the zeros of ζ, at the pole of ζ at 1, and at 0. We have

$$\mathrm{res}(f, 0) = \lim_{z \to 0} z f(z) = \frac{\zeta'(0)}{\zeta(0)}.$$

Since 1 is a simple pole of ζ we have

$$\frac{\zeta'(z)}{\zeta(z)} = -\frac{1}{z-1} + \mathcal{O}(1)$$

as $z \to 1$; cf. also the discussion after Lemma 5.1. It follows that

$$\mathrm{res}(f, 1) = \lim_{z \to 1} (z - 1) f(z) = -\lim_{z \to 1} \frac{x^z}{z} = -x.$$

Finally, if ρ is a zero of ζ of multiplicity m, then

$$\frac{\zeta'(z)}{\zeta(z)} = \frac{m}{z-\rho} + \mathcal{O}(1)$$

as $z \to \rho$ and thus

$$\operatorname{res}(f,\rho) = \lim_{z\to\rho}(z-\rho)f(z) = m\lim_{z\to\rho}\frac{x^z}{z} = m\frac{x^\rho}{\rho}.$$

By Theorem 8.7 and Theorems 4.3 and 4.4 the only zeros of ζ outside the strip $\{z\colon 0<\operatorname{Re} z<1\}$ are at the points $-2k$ with $k\in\mathbb{N}$. They are all simple so that

$$\operatorname{res}(f,-2k) = \frac{x^{-2k}}{-2k} = -\frac{1}{2k}\left(\frac{1}{x^2}\right)^k.$$

We conclude that

$$\begin{aligned}\frac{1}{2\pi i}\int_\gamma f(z)dz &= \operatorname{res}(f,1) + \sum_{\substack{\rho\colon \zeta(\rho)=0\\ 0<\operatorname{Re}\rho<1\\ |\operatorname{Im}\rho|<T}} \operatorname{res}(f,\rho) + \operatorname{res}(f,0) + \sum_{k=1}^N \operatorname{res}(f,-2k)\\ &= -x + \sum_{\substack{\rho\colon \zeta(\rho)=0\\ 0<\operatorname{Re}\rho<1\\ |\operatorname{Im}\rho|<T}} \frac{x^\rho}{\rho} + \frac{\zeta'(0)}{\zeta(0)} - \frac{1}{2}\sum_{k=1}^N \frac{1}{k}\left(\frac{1}{x^2}\right)^k.\end{aligned}$$

Now

$$\lim_{T\to\infty}\frac{1}{2\pi i}\int_{\gamma_1} f(z)dz = -\psi_0(x)$$

by Theorem 8.1 and

$$-\sum_{k=1}^\infty \frac{1}{k}\left(\frac{1}{x^2}\right)^k = \sum_{k=1}^\infty \frac{(-1)^{k+1}}{k}\left(-\frac{1}{x^2}\right)^k = \log\left(1-\frac{1}{x^2}\right).$$

The conclusion thus follows if we show that, for $j\in\{2,3,4\}$,

$$\int_{\gamma_j} f(z)dz \to 0$$

as $T\to\infty$ and $R=2N+1\to\infty$. In order to do so we will first prove several lemmas. □

In order to estimate the integrals $\int_{\gamma_j} f(z)dz$ for $j\in\{2,3,4\}$ one has to find an upper bound for $\zeta'(z)/\zeta(z)$ for z on these curves. This is comparatively easy if $\operatorname{Re} z\le -1$. The result is as follows.

Lemma 9.1. *There exists a constant C such that if $z \in \mathbb{C}$ satisfies $\operatorname{Re} z \leq -1$ and $|z - 2n| \geq 1$ for all $n \in \mathbb{N}$, then*

$$\left|\frac{\zeta'(z)}{\zeta(z)}\right| \leq \log|z| + C.$$

The idea is to reduce the case that $\operatorname{Re} z \leq -1$ to the case that $\operatorname{Re} z \geq 2$ by the functional equation. For the latter case we have a simple estimate.

Lemma 9.2. *There exists a constant K such that*

$$\left|\frac{\zeta'(z)}{\zeta(z)}\right| \leq K$$

for $\operatorname{Re} z \geq 2$.

Proof. The result is a simple consequence of Theorem 5.1 (and the discussion before Theorem 6.1) which yield that

$$\left|\frac{\zeta'(z)}{\zeta(z)}\right| = \left|\sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^z}\right| \leq \sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^2} \leq \sum_{n=1}^{\infty} \frac{\log n}{n^2} < \infty. \qquad \square$$

Since we apply the functional equation we will also need an estimate for the Gamma function.

Lemma 9.3. *There exists a constant C such that*

$$\left|\frac{\Gamma'(z)}{\Gamma(z)}\right| \leq \log|z| + C$$

for $\operatorname{Re} z \geq 2$.

Proof. Theorem 8.1 yields that

$$\frac{\Gamma'(z)}{\Gamma(z)} = \lim_{n\to\infty} \left(\log n - \frac{1}{z} - \sum_{k=1}^{n} \frac{1}{k+z}\right).$$

As shown in the exercises, the limit

$$\gamma = \lim_{n\to\infty} \left(\sum_{k=1}^{n} \frac{1}{k} - \log n\right)$$

exists. The limit γ is known as the Euler-Mascheroni constant. It follows that

$$\frac{\Gamma'(z)}{\Gamma(z)} = \lim_{n\to\infty} \left(\sum_{k=1}^{n} \left(\frac{1}{k} - \frac{1}{k+z}\right) - \frac{1}{z}\right) - \gamma = \sum_{k=1}^{\infty} \frac{z}{k(k+z)} - \frac{1}{z} - \gamma$$

For $\operatorname{Re} z \geq 2$ we have

$$\left|\frac{1}{k+z}\right| \leq \min\left\{\frac{1}{|z|}, \frac{1}{k}\right\}$$

and with $n = \lfloor |z| \rfloor$ we thus find that

$$\begin{aligned}\left|\sum_{k=1}^{\infty} \frac{z}{k(k+z)}\right| &\leq \sum_{k=1}^{n} \frac{1}{k} + \sum_{k=n+1}^{\infty} \frac{|z|}{k^2} \\ &\leq \log n + \gamma + o(1) + |z| \int_n^{\infty} \frac{dx}{x^2} \\ &= \log n + \frac{|z|}{n} + \gamma + o(1) \\ &= \log |z| + \mathcal{O}(1)\end{aligned}$$

as $|z| \to \infty$, from which the conclusion follows. □

The functional equation of the Gamma function yields that

$$\frac{\Gamma'(z+n)}{\Gamma(z+n)} = \frac{\Gamma'(z)}{\Gamma(z)} + \sum_{k=0}^{n-1} \frac{1}{z+k}$$

for $n \in \mathbb{N}$. Lemma 9.3 thus implies the following result.

Lemma 9.4. *Let $a \in \mathbb{R}$ and $\delta > 0$. Then there exists $C > 0$ such that*

$$\left|\frac{\Gamma'(z)}{\Gamma(z)}\right| \leq \log |z| + C$$

if $\operatorname{Re} z \geq a$ and $|z - n| \geq \delta$ for all $n \in \{0, -1, -2, \dots\}$.

Proof of Lemma 9.1. We deduce from the form of the functional equation given after Theorem 8.6 that

$$\frac{\zeta'(z)}{\zeta(z)} = \log(2\pi) - \frac{\Gamma'(1-z)}{\Gamma(1-z)} - \frac{\zeta'(1-z)}{\zeta(1-z)} + \frac{\pi}{2} \cot\left(\frac{\pi}{2} z\right).$$

Put $w = 1 - z$. The condition $\operatorname{Re} z \leq -1$ then takes the form $\operatorname{Re} w \geq 2$. The condition that $|z - 2n| \geq 1$ for all $n \in \mathbb{N}$ implies that $|e^{-i\pi z} - 1| \geq \delta$ for some positive constant δ.

Lemma 7.2 yields that

$$\left|\cot\left(\frac{\pi}{2} z\right)\right| = \left|1 + \frac{2}{e^{-i\pi z} - 1}\right| \leq 1 + \frac{2}{\delta}$$

and by Lemma 9.2 we have

$$\left|\frac{\zeta'(1-z)}{\zeta(1-z)}\right| = \left|\frac{\zeta'(w)}{\zeta(w)}\right| \leq K.$$

Finally

$$\left|\frac{\Gamma'(1-z)}{\Gamma(1-z)}\right| = \left|\frac{\Gamma'(w)}{\Gamma(w)}\right| \leq \log |w| + C = \log |z - 1| + C = \log |z| + \mathcal{O}(1)$$

by Lemma 9.3. Combining these estimates yields the conclusion. □

Proof of Theorem 9.1, Part 2. Let f and $\gamma = \gamma_1+\gamma_2+\gamma_3+\gamma_4$ be as before. We have to show that $\int_{\gamma_j} f(z)dz \to 0$ as $T \to \infty$ and $R = 2N+1 \to \infty$, for $j \in \{2,3,4\}$. Lemma 9.1 implies that

$$\begin{aligned}\left|\int_{\gamma_3} f(z)dz\right| &\le 2T \max_{z\in\gamma_3} |f(z)| \\ &= 2T \max_{z\in\gamma_3} \left|\frac{\zeta'(z)}{\zeta(z)}\cdot\frac{1}{z}\right| \\ &\le 2Tx^{-R} \max_{z\in\gamma_3} \frac{\log|z|+C}{|z|} \\ &\le 2Tx^{-R}\end{aligned}$$

for large R. Thus $\int_{\gamma_3} f(z)dz \to 0$ as $R = 2N+1 \to \infty$.

It also follows that the integral of f over those parts of γ_2 and γ_4 which lie in $\{z\colon \operatorname{Re} z \le -1\}$ tend to 0 as $T \to \infty$. Indeed, for the corresponding part of γ_2 we find that

$$\begin{aligned}\left|\int_{-1+iT}^{-R+iT} f(z)dz\right| &\le \int_{-R}^{-1} \left|\frac{\zeta'(s+iT)}{\zeta(s+iT)}\right| \cdot \frac{x^s}{|s+iT|}ds \\ &\le \int_{-R}^{-1} (\log|s+iT|+C)\frac{x^s}{|s+iT|}ds \\ &\le \frac{\log T + C}{T}\int_{-\infty}^{-1} x^s ds \\ &= \frac{\log T + C}{Tx\log x},\end{aligned}$$

and the estimate for the corresponding part for γ_4 follows by symmetry.

The estimate of the integrals over the remaining parts of γ_2 and γ_4 is more delicate. Here one has to choose T such that γ_2 and γ_4 are not too close to a zero of ζ. Again we may restrict to the integral over the corresponding part of γ_2. We thus see that the proof of Theorem 9.1 will be completed with the following lemma. □

Lemma 9.5. *There exists a sequence of T-values which tends to ∞ such that, as $T \to \infty$ through this sequence,*

$$\int_{-1+iT}^{c+iT} \frac{\zeta'(z)}{\zeta(z)} \cdot \frac{x^z}{z} dz \to 0.$$

The proof of Lemma 9.5 will be postponed to section 11. First we will have to provide some general results about entire functions and their zeros in section 10.

Before doing so, we make some further comments on the relation of $\pi(x)$ and $\psi(x)$ to the zeros of ζ. The main interest is in the asymptotic behavior of $\pi(x)$ and $\psi(x)$ as $x \to \infty$. Here the constant $\zeta'(0)/\zeta(0)$ and the term $\log(1-1/x^2)$ in the explicit formula do not matter. (Note that $\log(1-1/x^2) \to 0$ as $x \to \infty$.) Instead of considering the sum over all zeros of ζ in the *critical strip* $\{z\colon 0 < \operatorname{Re} z < 1\}$ one now considers only those zeros ρ for which $|\operatorname{Im}\rho| \le T$.

An error estimate in Theorem 6.1 as well as the estimates of the integrals $\int_{\gamma_j} f(z)dz$ that we made above (and will make in the proof of Lemma 9.5 in section 11) lead to

$$\psi(x) = x - \sum_{\substack{\rho\colon \zeta(\rho)=0\\ 0<\operatorname{Re}\rho<1\\ |\operatorname{Im}\rho|\le T}} \frac{x^\rho}{\rho} + \mathcal{O}\Big(\frac{x}{T}(\log xT)^2\Big) + \mathcal{O}(\log x).$$

Note that we can take $\psi(x)$ instead of $\psi_0(x)$ here since $0 \le \psi(x) - \psi_0(x) \le \frac{1}{2}\log x$. We will see later (in Theorem 10.4) that

$$\sum_{\substack{\rho\colon \zeta(\rho)=0\\ 0<\operatorname{Re}\rho<1\\ |\operatorname{Im}\rho|\le T}} \frac{1}{|\rho|} = \mathcal{O}\big((\log T)^2\big)\,.$$

Suppose now there exists $\lambda \in \big[\frac{1}{2}, 1\big)$ such that $\operatorname{Re}\rho \le \lambda$ for all zeros ρ of ζ. Then

$$\Bigg|\sum_{\substack{\rho\colon \zeta(\rho)=0\\ 0<\operatorname{Re}\rho<1\\ |\operatorname{Im}\rho|\le T}} \frac{x^\rho}{\rho}\Bigg| \le \sum_{\substack{\rho\colon \zeta(\rho)=0\\ 0<\operatorname{Re}\rho<1\\ |\operatorname{Im}\rho|\le T}} \frac{x^{\operatorname{Re}\rho}}{|\rho|} \le x^\lambda \sum_{\substack{\rho\colon \zeta(\rho)=0\\ 0<\operatorname{Re}\rho<1\\ |\operatorname{Im}\rho|\le T}} \frac{1}{|\rho|} = \mathcal{O}\big(x^\lambda(\log T)^2\big)\,.$$

Choosing $T = x$ then leads to

$$\psi(x) = x + \mathcal{O}\big(x^\lambda(\log x)^2\big)\,.$$

In particular, assuming that the Riemann hypothesis is true so that $\lambda = \frac{1}{2}$ we find that

$$\psi(x) = x + \mathcal{O}\big(\sqrt{x}(\log x)^2\big)\,.$$

One can also write the last equation in terms of $\pi(x)$. As shown in the exercises, this equation is equivalent to

$$\pi(x) = \operatorname{Li}(x) + \mathcal{O}\big(\sqrt{x}\log x\big)\,.$$

In turn it can be shown that if one (and hence both) of the last two equations hold, then the Riemann hypothesis is true. So these two statements are actually equivalent.

Without assumption of the Riemann hypothesis, much less is known. At the end of Section 5 we mentioned that

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

The key step in the proof is to show that there is a positive constant c_0 such that if $\zeta(x+iy) = 0$ for some $x \in (0,1)$, then $y < 1 - c_0/\log x$. Arguments similar to the ones sketched above then give the conclusion. Quite generally, the knowledge of a larger zero-free region of ζ would lead to an improved estimate for the difference of $\pi(x)$ and $\operatorname{Li}(x)$.

10 Entire functions of finite order

Let f be an entire function. We say that f is of finite order if there exists $\alpha > 0$ such that

$$|f(z)| = \mathcal{O}(\exp(|z|^\alpha))$$

as $|z| \to \infty$. The infimum of all these α is called the *order* of f and denoted by $\rho(f)$.

Example 10.1. Let $f(z) = \exp(z^d)$ with $d \in \mathbb{N}$. Then $\rho(f) = d$.

Example 10.2. The function f given by $f(z) = \exp(\exp z)$ does not have finite order.

Example 10.3. Let

$$f(z) = \cos\sqrt{z} = \frac{1}{2}\left(\exp\left(i\sqrt{z}\right) + \exp\left(-i\sqrt{z}\right)\right) = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k)!} z^k.$$

Then

$$|f(z)| \le \exp\sqrt{|z|} \quad \text{and} \quad f(-r) \ge \frac{1}{2}\exp\sqrt{r}$$

and hence $\rho(f) = \frac{1}{2}$.

Theorem 10.1. *The Riemann xi function has order* 1. *More precisely,*

$$\log|\xi(z)| \le \frac{1}{2}|z|\log|z| + \mathcal{O}(|z|)$$

as $|z| \to \infty$.

Proof. The proof of Theorem 4.1 shows that if $\operatorname{Re} z > 0$, then

$$\left|\zeta(z) - \frac{1}{z-1}\right| \le \sum_{n=1}^{\infty}\left|\int_n^{n+1}\left(\frac{1}{n^z} - \frac{1}{x^z}\right)dx\right| \le \sum_{n=1}^{\infty}\frac{|z|}{n^{\operatorname{Re} z+1}} = |z|\zeta(1+\operatorname{Re} z).$$

It follows that there exists a constant C such that $|\zeta(z)| \le C|z|$ if $|z| \ge 2$ and $\operatorname{Re} z \ge \frac{1}{2}$.

The definition of Γ yields that $|\Gamma(z)| \le \Gamma(\operatorname{Re} z)$ for $\operatorname{Re} z > 0$. By Stirling's formula we have

$$\log\Gamma(x) \le \log\Gamma(\lceil x\rceil) = \lceil x\rceil\log\lceil x\rceil + O(\lceil x\rceil) = x\log x + \mathcal{O}(x)$$

as $x \to \infty$. It follows that

$$\log\Gamma(|z|) \le |z|\log|z| + \mathcal{O}(|z|)$$

as $|z| \to \infty$, $\operatorname{Re} z > 0$. Since, by definition,

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

we deduce that

$$\begin{aligned}\log|\xi(z)| &\le \log|z| + \log|z-1| - \frac{1}{2}(\log\pi)\operatorname{Re}(z) + \log\left|\Gamma\left(\frac{1}{2}z\right)\right| + \log|\zeta(z)| \\ &\le \frac{1}{2}|z|\log|z| + \mathcal{O}(|z|)\end{aligned}$$

as $|z| \to \infty$, $\operatorname{Re} z \ge \frac{1}{2}$. Since $\xi(1-z) = \xi(z)$ we obtain the same estimate for $\operatorname{Re} z \le \frac{1}{2}$. This implies that the order of ξ is at most 1. The above reasoning also shows that $\xi(x) = \frac{1}{2}x\log x + \mathcal{O}(x)$ as $x \to \infty$ and thus in fact $\rho(\xi) = 1$. □

Theorem 10.2 (Jensen's formula)**.** *Let f be entire and $r > 0$. Suppose that $f(0) \ne 0$ and $f(z) \ne 0$ for $|z| = r$. Let $a_1, \dots, a_m$ be the zeros of f in $\{z\colon |z| < r\}$, counted according to multiplicity. Then*

$$\log|f(0)| = \frac{1}{2\pi}\int_0^{2\pi} \log|f(re^{i\theta})|d\theta - \sum_{j=1}^m \log\frac{r}{|a_j|}.$$

Proof. Suppose first that f has no zeros in $\{z\colon |z| \le r\}$. Then f has the form $f = e^g$ with some function g holomorphic in a domain containing $\{z\colon |z| \le r\}$ and we have $\log|f| = \operatorname{Re} g$. Thus

$$\begin{aligned}\log|f(0)| &= \operatorname{Re} g(0) \\ &= \operatorname{Re}\left(\frac{1}{2\pi i}\int_{|z|=r}\frac{g(z)}{z}dz\right) \\ &= \operatorname{Re}\left(\frac{1}{2\pi}\int_0^{2\pi} g(re^{i\theta})|d\theta\right) \\ &= \frac{1}{2\pi}\int_0^{2\pi}\operatorname{Re} g(re^{i\theta})d\theta \\ &= \frac{1}{2\pi}\int_0^{2\pi}\log f(re^{i\theta})|d\theta.\end{aligned}$$

Suppose now that f has zeros $a_1, \dots, a_m$. For $a \in \mathbb{C}$ with $|a| < r$ we consider

$$\varphi_a(z) = \frac{r(z-a)}{r^2 - \overline{a}z}.$$

It is easy to see that $|\varphi_a(z)| = 1$ for $|z| = r$. It follows that

$$h(z) = \frac{f(z)}{\prod_{j=1}^m \varphi_{a_j}(z)}$$

defines a meromorphic function which has no zeros in $\{z\colon |z| \le r\}$ and which satisfies $|h(z)| = |f(z)|$ for $|z| = r$. Hence, by the special case already proved,

$$\begin{aligned}\frac{1}{2\pi}\int_0^{2\pi} \log|f(re^{i\theta})|d\theta &= \frac{1}{2\pi}\int_0^{2\pi} \log h(re^{i\theta})|d\theta \\ &= \log|h(0)| \\ &= \log \frac{f(0)}{\prod_{j=1}^m |\varphi_{a_j}(0)|} \\ &= \log|f(0)| - \sum_{j=1}^m \log\frac{|a_j|}{r},\end{aligned}$$

from which the conclusion follows immediately. □

We denote by $n(r, f)$ the number of zeros of an entire function f in $\{z\colon |z| \le r\}$.

Theorem 10.3. *Let f be entire with $f(0) \neq 0$. Then*

$$n(r, f) \le \max_{|z|=er} \log|f(z)| - \log|f(0)|.$$

Proof. Let $a_1, a_2, \ldots$ be the zeros of f. Jensen's formula yields that

$$n(r, f) \le \sum_{|a_j|\le r} \log\frac{er}{|a_j|} = \frac{1}{2\pi}\int_0^{2\pi} \log|f(ere^{i\theta})|d\theta - \log|f(0)|,$$

from which the conclusion immediately follows. □

One consequence is the following result that was stated earlier.

Theorem 10.4. *As $T \to \infty$,*

$$\sum_{\substack{\rho\colon \zeta(\rho)=0 \\ 0<\operatorname{Re}\rho<1 \\ |\operatorname{Im}\rho|\le T}} \frac{1}{|\rho|} = \mathcal{O}\big((\log T)^2\big)\,.$$

Proof. Recall that the zeros of ζ in the critical strip are precisely those of ξ. Theorems 10.1 and 10.3 imply that

$$n(r, \xi) \le \max_{|z|=er} \log|\xi(z)| - \log|\xi(0)| \le (1 + o(1))\frac{e}{2} r \log r = \mathcal{O}(r \log r)$$

as $r \to \infty$. For $T > 1$ choose k such that $2^{k-1} < |1 + iT| = \sqrt{T^2+1} \le 2^k$. Then

$$\begin{aligned}
\sum_{\substack{\rho:\ \zeta(\rho)=0 \\ 0<\operatorname{Re}\rho<1 \\ |\operatorname{Im}\rho|\le T}} \frac{1}{|\rho|} &\le \sum_{j=1}^{k} \sum_{\substack{\rho:\ \xi(\rho)=0 \\ 2^{j-1}\le|\rho|\le 2^j}} \frac{1}{|\rho|} + \mathcal{O}(1) \\
&\le \sum_{j=1}^{k} n(2^j,\xi)\frac{1}{2^{j-1}} + \mathcal{O}(1) \\
&= \mathcal{O}\left(\sum_{j=1}^{k} \log(2^j)\right) \\
&= \mathcal{O}\left(\sum_{j=1}^{k} j\right) \\
&= \mathcal{O}\big(k^2\big) \\
&= \mathcal{O}\big((\log T)^2\big)\,. \qquad \square
\end{aligned}$$

Theorem 10.5. *Let f be an entire function of finite order and let $\mu > \rho(f)$. Let $a_1, a_2, \ldots$ be the zeros of f in $\mathbb{C}\setminus\{0\}$, counted according to multiplicity. Then*

$$\sum_{j=1}^{\infty} \frac{1}{|a_j|^{\mu}} < \infty.$$

Proof. Without loss of generality we may assume that $f(0) \ne 0$. Choose $\varepsilon > 0$ with $\rho(f) + \varepsilon < \mu$. By Theorem 10.3 we have

$$n(r,f) \le \max_{|z|=er} \log|f(z)| - \log|f(0)| \le r^{\rho(f)+\varepsilon}$$

for large r. For large $K \in \mathbb{N}$ we thus have

$$\begin{aligned}
\sum_{|a_j|\ge 2^K} \frac{1}{|a_j|^{\mu}} &= \sum_{k=K}^{\infty} \sum_{2^k\le|a_j|<2^{k+1}} \frac{1}{|a_j|^{\mu}} \\
&\le \sum_{k=K}^{\infty} n(2^{k+1}) \frac{1}{|2^k|^{\mu}} \\
&\le \sum_{k=K}^{\infty} 2^{(k+1)(\rho(f)+\varepsilon)-k\mu} \\
&= 2^{\rho(f)+\varepsilon} \sum_{k=K}^{\infty} 2^{k(\rho(f)+\varepsilon-\mu)} < \infty. \qquad \square
\end{aligned}$$

We will give a representation of the Riemann xi function as an infinite product over its zeros. Such a representation is provided by the Hadamard factorization theorem for general entire functions of finite order. Since $\rho(\xi) = 1$ we restrict here to the case that the order is less than 2.

Theorem 10.6. *Let f be an entire of finite order less than 2 and let $a_1, a_2, \ldots$ be the zeros of f in $\mathbb{C} \setminus \{0\}$. If f has a zero at the origin, let m denote its multiplicity, and put $m = 0$ if $f(0) \neq 0$. Then there exists $A, B \in \mathbb{C}$ such that*

$$f(z) = e^{A+Bz} z^m \prod_k \left(1 - \frac{z}{a_k}\right) e^{z/a_k}$$

Sketch of proof. It is not difficult to see that

$$|(1-w)e^w| \leq |w|^2$$

for $|w| \leq 1$. It follows that

$$\left| \left(1 - \frac{z}{a_k}\right) e^{z/a_k} \right| \leq \frac{|z|^2}{|a_k|^2}$$

for $|a_k| \geq |z|$. Together with Theorem 10.5 this implies that the infinite product converges and defines an entire function whose zeros are the zeros of f in $\mathbb{C} \setminus \{0\}$. Hence the function h defined by

$$h(z) = \frac{f(z)}{z^m \prod_k \left(1 - \frac{z}{a_k}\right) e^{z/a_k}}$$

is an entire function without zeros and thus has the form $h = e^g$ with an entire function g.

It remains to show that g has the form $g(z) = A + Bz$. We omit this argument here. □

In particular, it follows from Theorems 10.1 and 10.6 that the Riemann xi function has a representation

$$\xi(z) = e^{A+Bz} \sum_{\substack{\rho:\, \zeta(\rho)=0 \\ 0<\operatorname{Re}\rho<1}} \left(1 - \frac{z}{\rho}\right) e^{z/\rho}.$$

Since

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

and

$$\lim_{z\to 1}(z-1)\zeta(z) = 1$$

by Theorem 4.1 we obtain, using that $\Gamma(z)\Gamma(1-z) = \pi/\sin(\pi z)$ implies that $\Gamma(\frac{1}{2}) = \sqrt{\pi}$,

$$e^A = \xi(0) = \xi(1) = \frac{1}{2}\pi^{-1/2}\Gamma\left(\frac{1}{2}\right) = \frac{1}{2}.$$

The computation of B is a little more involved and omitted here. We have

$$B = -\frac{1}{2}\gamma - 1 + \frac{1}{2}\log(4\pi).$$

Taking logarithmic derivatives in the product representation of ξ yields

$$\frac{\xi'(z)}{\xi(z)} = B + \sum_{\substack{\rho\colon \zeta(\rho)=0 \\ 0<\operatorname{Re}\rho<1}} \left(\frac{1}{z-\rho} + \frac{1}{\rho}\right).$$

Since

$$\frac{\xi'(z)}{\xi(z)} = \frac{1}{z} + \frac{1}{z-1} - \frac{1}{2}\log\pi + \frac{1}{2}\frac{\Gamma'\left(\frac{1}{2}z\right)}{\Gamma\left(\frac{1}{2}z\right)} + \frac{\zeta'(z)}{\zeta(z)},$$

the bound for Γ'/Γ given in Lemma 9.4 yields the following result.

Theorem 10.7. *Let $R \in \mathbb{R}$. Then, for $|\operatorname{Re} z| \leq R$,*

$$\frac{\zeta'(z)}{\zeta(z)} = \sum_{\substack{\rho\colon \zeta(\rho)=0 \\ 0<\operatorname{Re}\rho<1}} \left(\frac{1}{z-\rho} + \frac{1}{\rho}\right) + \mathcal{O}(\log|\operatorname{Im} z|)$$

as $|\operatorname{Im} z| \to \infty$.

11 The number of zeros in the critical strip

Besides the trivial zeros at the negative even integers, all zeros of the Riemann zeta function are in the critical strip $\{z\colon 0 < \operatorname{Re} z < 1\}$. For $T > 0$ we denote by $N(T)$ the number of zeros of ζ in the rectangle $\{z\colon 0 < \operatorname{Re} z < 1,\ 0 < \operatorname{Im} z \leq T\}$.

Theorem 10.1 and Jensen's formula yield that

$$N(T) = \mathcal{O}(T\log T)$$

as $T \to \infty$. In order to estimates the integrals occurring in Theorem 9.1, one has to choose the path of integration such that it is not too close to a zero of ζ. This is possible by the following result.

Theorem 11.1. *As $T \to \infty$,*

$$N(T+1) - N(T-1) = \mathcal{O}(\log T).$$

The proof will use the following two results.

Lemma 11.1. *Let $R > 0$. Then $\log|\Gamma(z)\cos(\pi z/2)| = \mathcal{O}(\log|\operatorname{Im} z|)$ as $|\operatorname{Im} z| \to \infty$, uniformly for $|\operatorname{Re} z| \leq R$.*

Proof. The "standard" proof of this result uses Stirling's formula, which describes the asymptotics of the Gamma function not only at the positive integers, but also for complex arguments.

Since we have not proved this version of Stirling's formula, we give an alternative argument. The equation $\Gamma(z)\Gamma(1-z) = \pi/\sin(\pi z)$ implies that

$$\begin{aligned}
\left|\Gamma\left(\tfrac{1}{2}+iy\right)\right|^2 &= \Gamma\left(\tfrac{1}{2}+iy\right)\overline{\Gamma\left(\tfrac{1}{2}+iy\right)} \\
&= \Gamma\left(\tfrac{1}{2}+iy\right)\Gamma\left(\tfrac{1}{2}-iy\right) \\
&= \frac{\pi}{\sin\left(\pi\left(\tfrac{1}{2}+iy\right)\right)} \\
&= \frac{\pi}{\cos(i\pi y)} \\
&= \frac{\pi}{\cosh(\pi y))} \\
&\sim \frac{\pi}{\exp(\pi|y|)} \\
&\sim \frac{\pi}{\left|\cos\left(\frac{\pi}{2}\left(\frac{1}{2}+iy\right)\right)\right|^2}
\end{aligned}$$

as $|y| \to \infty$. Thus $g(z) := \Gamma(z)\cos(\pi z/2)$ satisfies $|g(\frac{1}{2}+iy)| \to \sqrt{\pi}$ as $|y| \to \infty$. By Lemma 9.4, and since $\tan z \to \pm i$ as $\operatorname{Im} z \to \pm\infty$, we have

$$\left|\frac{g'(z)}{g(z)}\right| = \left|\frac{\Gamma'(z)}{\Gamma(z)} - \frac{\pi}{2}\tan\left(\frac{\pi z}{2}\right)\right| \le \log|z| + \mathcal{O}(1) = \log|\operatorname{Im} z| + \mathcal{O}(1)$$

as $|\operatorname{Im} z| \to \infty$, uniformly for $|\operatorname{Re} z| \le R$. This implies that if $|x| \le R$, then

$$\log|g(x+iy)| \le \log\left|g\left(\frac{1}{2}+iy\right)\right| + \left|x-\frac{1}{2}\right|\log|y| + \mathcal{O}(1) = \mathcal{O}(\log|y|)$$

as $|y| \to \infty$. □

Lemma 11.2. *Let $R > 0$. Then there exists $M > 0$ such that $|\zeta(z)| \le |z|^M$ if $|\operatorname{Im} z| \ge 2$ and $|\operatorname{Re} z| \le R$.*

Proof. It was shown in the proof of Theorem 10.1 that there exists a constant C such that $|\zeta(z)| \le C|z|$ for $|z| \ge 2$ and $\operatorname{Re} z \ge \frac{1}{2}$. For $\operatorname{Re} z < \frac{1}{2}$ we use the functional equation

$$\zeta(1-z) = 2^{1-z}\pi^{-z}\Gamma(z)\zeta(z)\cos\left(\frac{\pi}{2}z\right)$$

which, replacing z by $1-z$, may also be written as

$$\zeta(z) = 2^z\pi^{z-1}\Gamma(1-z)\zeta(1-z)\cos\left(\frac{\pi}{2}(1-z)\right).$$

By Lemma 11.1 there exists $K > 0$ such that

$$\left|\Gamma(1-z)\cos\left(\frac{\pi}{2}(1-z)\right)\right| \le |z|^K$$

if $|\operatorname{Im} z| \geq 1$ and $|\operatorname{Re}(1-z)| \leq R+1$. The latter condition clearly holds if $|\operatorname{Re} z| \leq R$. Altogether we thus have

$$|\zeta(z)| \leq 2^R \pi^{R-1} C \cdot |1-z| \cdot |z|^K$$

if $|\operatorname{Im} z| \geq 1$, $|z-1| \geq 2$ and $-R \leq \operatorname{Re} z \leq \frac{1}{2}$. The conclusion easily follows. □

Proof of Theorem 11.1. We consider the function f given by $f(z) = \zeta(2+iT+z)$. Since

$$\{z\colon 0 < \operatorname{Re} z < 1,\, T-1 \leq \operatorname{Im} z \leq T+1\} \subset \{z\colon |z-(2+iT)| \leq 3\}$$

we have

$$N(T+1) - N(T-1) \leq n(3,f).$$

By Lemma 11.2 there exists a constant C such that

$$\log|\zeta(z)| \leq C \log(\operatorname{Im} z) \leq C\log(T+3e) \leq 2C\log T$$

for $|z-(2+iT)| \leq 3e$. Jensen's formula thus yields that

$$n(3,f) = 2C\log T - \log|f(0)| = 2C\log T - \log|\zeta(2+iT)|.$$

Since

$$|\zeta(2+iT)| = \left|\sum_{n=1}^{\infty} \frac{1}{n^{2+iT}}\right| \geq 1 - \sum_{n=2}^{\infty} \left|\frac{1}{n^{2+iT}}\right| = 1 - \sum_{n=2}^{\infty} \frac{1}{n^2} = 1 - \frac{\pi^2}{6} > \frac{1}{4}$$

we thus have

$$N(T+1) - N(T-1) \leq n(3,f) \leq 2C\log T + \log 4.$$ □

We will use Theorem 11.1 to refine the estimate for ζ'/ζ given at the end of section 10.

Theorem 11.2. *Let $R \in \mathbb{R}$. Then, for $z = x+iy$ with $|x| \leq R$,*

$$\frac{\zeta'(z)}{\zeta(z)} = \sum_{\substack{\rho\colon \zeta(\rho)=0 \\ 0<\operatorname{Re}\rho<1 \\ |y-\operatorname{Im}\rho|\leq 1}} \frac{1}{z-\rho} + \mathcal{O}(\log|y|)$$

as $|y| \to \infty$.

Proof. Since $\zeta(2+iy) = \mathcal{O}(1)$ by Lemma 9.2, Theorem 10.5 yields that

$$\frac{\zeta'(z)}{\zeta(z)} = \frac{\zeta'(z)}{\zeta(z)} - \frac{\zeta'(2+iy)}{\zeta(2+iy)} + \mathcal{O}(1) = \sum_{\substack{\rho\colon \zeta(\rho)=0 \\ 0<\operatorname{Re}\rho<1}} \left(\frac{1}{z-\rho} - \frac{1}{2+iy-\rho}\right) + \mathcal{O}(\log|y|)$$

Now $|2+iy-\rho| \geq \operatorname{Re}(2+iy-\rho) = 2-\rho \geq 1$ and hence

$$\sum_{\substack{\rho:\, \zeta(\rho)=0 \\ 0<\operatorname{Re}\rho<1 \\ |y-\operatorname{Im}\rho|\leq 1}} \frac{1}{2+iy-\rho} = \mathcal{O}(\log|y|)$$

by Theorem 11.1. We also have

$$\frac{1}{z-\rho} - \frac{1}{2+iy-\rho} = \frac{2+iy-z}{(z-\rho)(2+iy-\rho)} = \frac{2-x}{(z-\rho)(2+iy-\rho)}$$

and thus

$$\left|\frac{1}{z-\rho} - \frac{1}{2+iy-\rho}\right| \leq \frac{2-x}{|\operatorname{Im}(z-\rho)|\cdot|\operatorname{Im}(2+iy-\rho)|} \leq \frac{2}{|y-\operatorname{Im}(\rho)|^2}.$$

Theorem 11.1 now yields that there exists a constant C such that $N(T+1)-N(T) \leq C\log(2+|T|)$ for all $T \in \mathbb{R}$. Hence

$$\begin{aligned}
\left|\sum_{\substack{\rho:\, \zeta(\rho)=0 \\ 0<\operatorname{Re}\rho<1 \\ |y-\operatorname{Im}\rho|>1}} \left(\frac{1}{z-\rho} - \frac{1}{2+iy-\rho}\right)\right| &\leq \sum_{k\in\mathbb{Z}\setminus\{0,-1\}} \sum_{\substack{\rho:\, \zeta(\rho)=0 \\ 0<\operatorname{Re}\rho<1 \\ k\leq\operatorname{Im}\rho-y\leq k+1}} \frac{2}{|y-\operatorname{Im}(\rho)|^2}. \\
&\leq \sum_{k\in\mathbb{Z}\setminus\{0,-1\}} (N(y+k+1)-N(y+k))\frac{2}{k^2} \\
&\leq 2C \sum_{k\in\mathbb{Z}\setminus\{0,-1\}} \frac{\log|y+k|}{k^2}.
\end{aligned}$$

As $\log|y+k| \leq \log(2\max\{|y|,|k|\} = \log 2\cdot\max\{\log|y|,\log|k|\}$ this implies that

$$\left|\sum_{\substack{\rho:\, \zeta(\rho)=0 \\ 0<\operatorname{Re}\rho<1 \\ |y-\operatorname{Im}\rho|>1}} \left(\frac{1}{z-\rho} - \frac{1}{2+iy-\rho}\right)\right| = \mathcal{O}(\log|y|).$$

Combining the above estimates yields the conclusion. □

We can now complete the proof of Theorem 9.1. As shown in section 9, it remains to prove Lemma 9.5 which says that

$$\int_{-1+iT}^{c+iT} \frac{\zeta'(z)}{\zeta(z)} \cdot \frac{x^z}{z} dz \to 0$$

as $T \to \infty$ through a suitably chosen sequence of values. Clearly, this follows from the following result.

Lemma 11.3. *Let $c > 0$. Then for all large $S > 0$ there exists $T \in [S-1, S+1]$ such that if $\operatorname{Im} z = T$ and $-1 \leq \operatorname{Re} z \leq c$, then*

$$\left|\frac{\zeta'(z)}{\zeta(z)}\right| = \mathcal{O}\big((\log T)^2\big)$$

as $S \to \infty$.

Proof. Let n_S be the number of zeros ρ of ζ which satisfy $S - 1 \leq \operatorname{Im} z \leq S + 1$. Theorem 11.1 says that $n_S = \mathcal{O}(\log S)$ as $S \to \infty$.

The imaginary parts of these zeros divide the interval $[S-1, S+1]$ in at most $n_S + 1$ subintervals. One of these subintervals has length at least $2/(n_S + 1)$ and hence there exists $T \in [S-1, S+1]$ such that

$$|z - \rho| \geq |\operatorname{Im}(z-\rho)| = |T - \operatorname{Im}\rho| \geq \frac{1}{n_S + 1}$$

whenever $\operatorname{Im} z = T$. Hence

$$\left|\sum_{\substack{\rho:\ \zeta(\rho)=0 \\ 0<\operatorname{Re}\rho<1 \\ |T-\operatorname{Im}\rho|\leq 1}} \frac{1}{z-\rho}\right| \leq (N(T+1) - N(T-1))(n_S+1) = \mathcal{O}\big((\log T)^2\big)\,.$$

The conclusion now follows from Theorem 11.2. □

To complete the proof of Lemma 9.5 we only have to note that if T is chosen according to Lemma 11.3, then

$$\left|\int_{-1+iT}^{c+iT} \frac{\zeta'(z)}{\zeta(z)} \cdot \frac{x^z}{z} dz\right| = \mathcal{O}\left((\log T)^2 \frac{x^c}{T}\right) = o(1)$$

as $T \to \infty$.

References

[1] Jörg Brüdern, Analytische Zahlentheorie. Springer, Berlin, 1995.

[2] Harold Davenport, Multiplicative Number Theory. Second Edition. Springer, Berlin, 1980. (Also: Lecture Notes by Andrew Granville based on this book.)

[3] Peter D. Lax and Lawrence Zalcman, Complex Proofs of Real Theorems. American Mathematical Society, Providence, RI, 2012.

[4] Donald J. Newman, Analytic Number Theory. Springer, New York, NY, 1998.

[5] Giovanni Sansone and Johan Gerretsen, Lectures on the Theory of Functions of a Complex Variable. P. Noordhoff, Groningen, 1960.

[6] Barry Simon, A comprehensive course in Analysis. Part 2b: Advanced Complex Analysis. American Mathematical Society, Providence, RI, 2015.