Analytic and Combinatorial Number Theory

Martin Klazar

May 26, 2023

Contents

Introduction

These lecture notes of the course I was teaching in the summer term 2022/23, i.e., in the winter and spring of 2023, have as their theme the use of analytic methods in handling discrete, combinatorial, number-theoretic and algebraic structures. Chapter 1 contains results on algebraic numbers, with the highlight being the proof of transcendence of π. In Chapter 2 we explain the possibly first main result in analytic number theory, the theorem of P. Dirichlet that for any coprime natural numbers a and m, the arithmetic progression

$$a,\ a+m,\ a+2m,\ \dots$$

contains infinitely many prime numbers. Chapter 3 is devoted to analytical methods in asymptotic combinatorial enumeration. In Chapter 4 we explain the elementary proof of the Prime Number Theorem, which is the theorem saying that the number of prime numbers in the set $\{1,2,\dots,n\}$ is for n going to infinity asymptotically $n/\log n$.

What distinguishes these lecture notes is that we emphasize and state clearly notions and theorems in mathematical analysis by which each number-theoretic or combinatorial result was obtained. We label these analytical tools by capital Latin letters as Theorem A, Corollary B and so on and state them explicitly, but usually we do not prove them.

We use the following notation for numeric domains. $\mathbb{N}=\{1,2,\dots\}$ are natural numbers, $\mathbb{N}_0=\{0,1,\dots\}$ are nonnegative integers, $\mathbb{Z}=\{\dots,-1,0,1\dots\}$ is the ordered integral domain of integers, $\mathbb{Q}=\{\frac{m}{n}\mid m,n\in\mathbb{Z},n\neq 0\}$ is the ordered field of fractions (rational numbers), $\mathbb{R}$ denotes the complete ordered field of real numbers, which is the completion of $\mathbb{Q}$, and $\mathbb{C}$ denotes the field of complex numbers, which is the algebraically closed quadratic extension

$$\mathbb{C}=\mathbb{R}[\mathrm{i}]=\mathbb{R}[\sqrt{-1}]$$

of $\mathbb{R}$. There are the inclusions or embeddings

$$\mathbb{N}\subset\mathbb{N}_0\subset\mathbb{Z}\subset\mathbb{Q}\subset\mathbb{R}\subset\mathbb{C}\,.$$

For example, the last one is the embedding $a\mapsto a+0\mathrm{i}$. In fact, we begin our notes with the *analytical* proof of the mentioned fundamental fact that the field $\mathbb{C}$ is algebraically closed.

In Prague, June 2023 M. Klazar

Chapter 1

Algebraicity, irrationality and transcendence

Let X and Y be sets. The notation $X \subset Y$ means that X is a *subset of* Y, for any x it holds that if $x \in X$ then $x \in Y$. By writing $f\colon X \to Y$ we say that f is a *map (function, mapping) from* X *to* Y: $f \subset X \times Y$ and for every $x \in X$ there is a unique $y \in Y$ such that $(x, y) \in f$, which is written as $f(x) = y$. We call X the *definition domain* of f and for $Z \subset X$ define the *image of* Z *by* f to be the set

$$f[Z] := \{f(x) \mid x \in Z\} \subset Y\ .$$

Let $K \subset L$ be an extension of fields (i.e., K is a subfield of L) and $a \in L$. We say that a is *irrational (over* K*)* if $a \notin K$ (this is non-standard terminology). The element a is *algebraic (over* K*)* if

$$\sum_{j=0}^{n} b_j a^j = 0_L$$

for some $b_j \in K$, not all of them 0_K, where $n \in \mathbb{N}$. The minimum such n is called the *degree* $\deg(a) = \deg_K(a) \in \mathbb{N}$ *of* a *(over* K*)*. In other words, $p(a) = 0_L$ for a nonzero polynomial $p(x) \in K[x]$, so a is a root of a nonzero polynomial with coefficients in K. For example, if $a \in K$ then a is algebraic and $\deg(a) = 1$. We say that a is *transcendental (over* K*)* if it is not algebraic over K. Often algebraicity, irrationality and transcendence, without anything else said, refer to the extension of fields $\mathbb{Q} \subset \mathbb{C}$.

By *polynomials* $p(x)$ in $K[x]$, where K is a field, we mean formal expressions

$$p(x) = \sum_{j=0}^{n} a_j x^j$$

where $n \in \mathbb{N}_0$, the *coefficients* $a_j \in K$ and $a_n \neq 0_K$. We say that the number $n =: \deg p(x) \in \mathbb{N}_0$ is the *degree* of the polynomial. The *zero polynomial*, denoted by $\equiv 0$, has an empty list of coefficients and is the additively neutral element in the ring, in fact an integral domain, $K[x]$ of all polynomials. We assume that the reader is familiar with the arithmetic of this ring. We set $\deg(\equiv 0) := -\infty$. Then

$$\forall\, p,\, q \in K[x]\, \big(\deg(pq) = \deg p + \deg q \wedge \deg(p+q) \geq \min(\{\deg p,\, \deg q\}\big)\ ,$$

with equality holding when $\deg p \neq \deg q$. We associate with every polynomial $p \in K[x]$ the *polynomial map*

$$p\colon K \to K,\ a \mapsto p(a)\ ,$$

defined by substituting for the variable x in $p(x) = \sum_{j=0}^{n} a_j x^j$ the element a and then performing the arithmetic operations in K described by the expression. If $p(x)$ is $\equiv 0$, the map is the constant map sending everything to 0_K. A nonzero polynomial $p = p(x)$ is *constant* if $\deg p = 0$. Else it is *non-constant* and $\deg p \geq 1$. The polynomial map of any constant polynomial is obviously a constant map. The *minimum polynomial* of an element $a \in L$ algebraic over K is the unique monic (i.e., with the leading coefficient 1_K) polynomial $p(x)$ in $K[x]$ with the minimum degree such that $p(a) = 0_L$.

1.1 The Fundamental Theorem of Algebra

The following theorem belongs to basic results in mathematics. Therefore we start with it our lecture notes.

Theorem 1.1 (FTAlg) *For every non-constant polynomial $p(x)$ in $\mathbb{C}[x]$ there is a number $\alpha \in \mathbb{C}$ such that*

$$p(\alpha) = 0\ .$$

In other words, every non-constant complex polynomial has a complex root.

Constant polynomials have no root and the zero polynomial has each complex number as its root. It is not hard to show that every nonzero polynomial $p = p(x)$ in $K[x]$ has at most $\deg p$ roots.

We begin by exposing main results in mathematical analysis involved in our proof of FTAlg, and then proceed to it. Let $X \subset \mathbb{C}$ be a set of complex numbers. Recall that X is *compact* if every sequence $(u_n) \subset X$ has a convergent subsequence with a limit in X. It is well known that X is compact iff it is closed and bounded. A map $f\colon X \to \mathbb{C}$ is *continuous* if for every sequence $(u_n) \subset X$ with $\lim u_n = x_0 \in X$ also $\lim f(u_n) = f(x_0)$.

Theorem A (continuity and compactness) *If $X \subset \mathbb{C}$ is a compact set and $f\colon X \to \mathbb{C}$ is a continuous map then the image $f[X] \subset \mathbb{C}$ is compact.*

Corollary B (attaining extrema) *If $X \subset \mathbb{C}$ is a compact set and $f\colon X \to \mathbb{R}$ is a continuous map then there are elements $x_0, x_1 \in X$ such that*

$$\forall\, x \in X\, \big(f(x_0) \leq f(x) \leq f(x_1)\big)\ .$$

Thus f attains on X at x_0 and x_1 its global minimum and global maximum, respectively. Indeed, if X and f are as stated then by Theorem A the subset $f[X] \subset \mathbb{R}$ is compact, so it is bounded and closed and contains both its infimum and its supremum.

We review the definition of connectedness. We say that a set $X \subset \mathbb{C}$ is *disconnected* if there exists two open (or closed) sets $A, B \subset \mathbb{C}$ such that

$$(X \subset A \cup B) \wedge (A \cap X \neq \emptyset \neq X \cap B) \wedge (A \cap B \cap X = \emptyset)\ .$$

In this situation we say that A *and* B *cut* X. Else, if such sets A and B do not exist, X is *connected.*

Theorem C (connectedness and compactness) *If $X \subset \mathbb{C}$ is a connected set and $f\colon X \to \mathbb{C}$ is a continuous map then the image $f[X] \subset \mathbb{C}$ is connected.*

Using suprema of real sets, one can show that every real interval $I \subset \mathbb{R}$ is a connected set (recall that for us $\mathbb{R} \subset \mathbb{C}$) and that, in fact, connected subsets of $\mathbb{R}$ are exactly the real intervals.

Corollary D (Bolzano–Cauchy) *If $a < b$ are real numbers and $f\colon [a, b] \to \mathbb{R}$ is a continuous function such that $f(a)f(b) < 0$ then*

$$\exists\, c \in (a,\, b)\, \big(f(c) = 0\big)\ .$$

Indeed, by Theorem C the image $f[I]$, where $I := [a, b]$, is connected and so is a real interval. Since the values $f(a), f(b) \in f[I]$ and 0 lies between them, $0 \in f[I]$.

Thus we can define, for example, the real function

$$\sqrt{\cdot}\colon [0,\, +\infty) \to [0,\, +\infty),\ c \mapsto \sqrt{c}\ ,$$

by taking $\sqrt{c}$ to be the unique real solution $x \geq 0$ of the equation $x^2 - c = 0$. It is easy to see that the square root function is continuous:

$$\sqrt{x} - \sqrt{y} = \frac{x - y}{\sqrt{x} + \sqrt{y}}\ .$$

Using continuity of polynomials and continuity of composite functions composed of continuous functions, we deduce the following.

Corollary E (two continuous maps) *The two functions $f^{\pm}\colon [-1, 1] \to \mathbb{C}$, defined by*

$$f^{\pm}(t) := t \pm \mathrm{i}\sqrt{1 - t^2}\ ,$$

are continuous.

The proof of FTAlg we are going to present uses connectedness of the complex unit circle

$$S := \{z \in \mathbb{C} \mid |z| = 1\}\ .$$

To show it we need the next simple lemma.

Lemma F (union of connected sets) *If $X, Y \subset \mathbb{C}$ are connected sets such that $X \cap Y \neq \emptyset$ then their union $X \cup Y$ is connected.*

Now to prove that S is connected we write it as the union

$$S = f^{-}[I] \cup f^{+}[I]$$

where $I = [-1, 1]$ and the two functions $f^{\pm}$ are defined above. They are continuous by Corollary E, their images $f^{\pm}[I]$ are connected by Theorem C and the images intersect in the points ± 1. By Lemma F, S is connected.

After these analytical preparations we prove FTAlg.

Our proof of Theorem 1.1. It has the *Reduction Step* and the *n-th Root Step*. In the former step we reduce FTAlg to proving that every binomial $x^n - u$, $n \in \mathbb{N}$ and $u \in \mathbb{C}$, has a root. In the latter step we show that this is the case.

• *Reduction Step.* Thus we assume that for every $u \in \mathbb{C}$ and every $n \in \mathbb{N}$ there is a $v \in \mathbb{C}$ such that $v^n = u$, and deduce from it that every non-constant complex polynomial $p(x)$ has a root. We write $p(x)$ as

$$p(x) = \sum_{j=0}^{n} a_j x^j = x^n \left(a_n + \sum_{j=0}^{n-1} \frac{a_j}{x^{n-j}} \right) ,$$

where $n \in \mathbb{N}$, $a_j \in \mathbb{C}$ and $a_n \neq 0$, and see that $\lim_{|x| \to +\infty} |p(x)| = +\infty$. This means that

$$\forall\, c > 0 \; \exists\, d > 0 \; \big(x \in \mathbb{C} \wedge |x| > d \Rightarrow |p(x)| > c \big) \; .$$

So we can take a real constant $d > 0$ such that $|x| > d \Rightarrow |p(x)| > |p(0)| = |a_0|$. By Corollary B, the function $|p(x)|$ attains on the closed disc

$$D := \{ x \in \mathbb{C} \mid |x| \leq d \}$$

on some $\mu \in D$ its minimum value $|p(\mu)|$. Since $0 \in D$ and

$$\forall\, x \in \mathbb{C} \setminus D \; \big(|p(\mu)| \leq |p(0)| < |p(x)| \big) \; ,$$

we see that $|p(\mu)|$ is a global minimum of $|p(x)|$ on $\mathbb{C}$. We show that in fact $|p(\mu)| = 0$. Thus $p(\mu) = 0$ and μ is a root of $p(x)$.

We assume for contradiction that $|p(\mu)| > 0$. We replace the variable x with the variable y defined by the relation $x = x - \mu + \mu = y + \mu$ and, using the Binomial Theorem, transform the polynomial $p(x)$ in the polynomial $q(y)$ with the same degree and leading coefficient:

$$p(x) = \sum_{j=0}^{n} a_j x^j = \sum_{j=0}^{n} a_j (y + \mu)^j = \sum_{j=0}^{n} b_j y^j =: q(y) \; ,$$

where $b_j \in \mathbb{C}$, $b_n = a_n \neq 0$ and $b_0 = q(0) = p(\mu) \neq 0$. To obtain a contradiction, we order $q(y)$ from the lowest powers of y as

$$q(y) = b_0 + b_m y^m + \sum_{j=m+1}^{n} b_j y^j =: b_0 + b_m y^m + r(y)$$

with $m \in \mathbb{N}$, $b_0 \neq 0$ and $b_m \neq 0$ (if $m = n$ we set $r(y)$ to be $\equiv 0$) and show that there is a $\nu \in \mathbb{C}$ such that $|q(\nu)| < |b_0|$. This will be a contradiction because $|b_0| = |q(0)| = |p(\mu)|$ is the global minimum of $|q(y)| = |p(x)|$ on $\mathbb{C}$.

Using the assumption about n-th roots we set $\alpha := (-b_0/b_m)^{1/m}$. We note that for $y \to 0$ it is true that

$$r(y) = O(y^{m+1}) = o(y^m) \ ,$$

and thus we may take a sufficiently small $\delta \in (0, 1)$ such that with $\nu := \delta\alpha$, $|r(\nu)| \leq \delta^m |b_0|/2$. Then

$$\begin{array}{rcl} |q(\nu)| & = & |b_0 + b_m(\delta\alpha)^m + r(\nu)| \\ & = & |b_0(1 - \delta^m) + r(\nu)| \leq |b_0|(1 - \delta^m) + |r(\nu)| \\ & \leq & |b_0|(1 - \delta^m/2) < |b_0| = |q(0)| = |p(\mu)| \end{array}$$

is the announced contradiction. So $|p(\mu)| = 0$ and μ is a root of $p(x)$.

• *n-th Root Step.* We prove that for every $n \in \mathbb{N}$ and every complex number u, the equation

$$x^n = u$$

has at least one solution $x \in \mathbb{C}$. We apply two simplifications: we may assume that (i) n is odd and that (ii) $u \in S$ (i.e., $|u| = 1$). Simplification (i) follows from the facts that every n can be written (uniquely) as the product $n = 2^k n'$ with $k \in \mathbb{N}_0$ and odd $n' \in \mathbb{N}$, and that the equation is solvable for any $u = c + d\mathrm{i}$ if $n = 2$: if $u \neq 0$, for two appropriate choices of signs in

$$a = \frac{\pm\sqrt{\sqrt{c^2 + d^2} + c}}{\sqrt{2}} \quad \text{and} \quad b = \frac{\pm\sqrt{\sqrt{c^2 + d^2} - c}}{\sqrt{2}}$$

we get two different solutions $a + b\mathrm{i} \in \mathbb{C}$ of the equation

$$(a + b\mathrm{i})^2 = a^2 - b^2 + 2ab\mathrm{i} = c + d\mathrm{i} \ .$$

For $u = 0$ there is of course the unique solution $a + b\mathrm{i} = 0$. Simplification (ii) follows from assuming that $u \neq 0$ and dividing the equation by $|u|$:

$$\left(x/|u|^{1/n}\right)^n = u/|u| \in S \ .$$

Here the real n-th root $|u|^{1/n}$ exists due to Corollary D.

Thus we need to show that the map

$$f(x) = x^n \colon S \to S \ ,$$

where n is odd, is onto. We assume for contradiction that there is an $\alpha \in S\backslash f[S]$. Then, since n is odd, also $-\alpha \in S \setminus f[S]$. We denote by $H^{\pm}$ the two open halfplanes in $\mathbb{C}$ determined by the line going through α and $-\alpha$. It follows (since n is odd) that

$$(H^- \cup H^+) \cap S = S \setminus \{\alpha,\, -\alpha\},\ \{-1,\, 1\} \subset f[S] \subset S$$

and that -1 and 1 lie in different halfplanes H^- and H^+. We see that H^- and H^+ cut $f[S]$ and that $f[S]$ is disconnected. But this contradicts the fact that by Theorem C $f[S]$ is connected as a continuous image of the connected set S. Thus $f[S] = S$ and every number $u \in S$ has an n-th root. Our proof of FTAlg is complete. □

1.2 Liouville's inequality

First examples of transcendental numbers, i.e. real (or complex) numbers that are not roots of any nonzero rational polynomial, were found by the French mathematician Joseph Liouville (1809–1882). His construction makes use of the following *Liouville's inequality* bounding from below distances between an irrational algebraic number and fractions. A transcendental number is then easily obtained as an irrational sum of a rapidly converging series of rational summands: the partial sums approximate the sum too well and Liouville's inequality is violated.

Theorem 1.2 (J. Liouville, 1844) *Let $\alpha \in \mathbb{R} \setminus \mathbb{Q}$ be an irrational algebraic real number and $n := \deg(\alpha) \geq 2$. There is a real constant $c = c(\alpha) > 0$ such that for every fraction $p/q \in \mathbb{Q}$ with $q \in \mathbb{N}$,*

$$\left|\alpha - \frac{p}{q}\right| \geq \frac{c}{q^n}\ .$$

Corollary 1.3 (transcendental number λ) *The real number*

$$\lambda := \sum_{n=1}^{\infty} 10^{-n!} = 0.110001000000000000000001000\ldots$$

is transcendental.

Of course, the base 10 can be replaced with any integer $m \geq 2$.

In the proofs we use two results from analysis. For the first one we review the definition of the derivative of a real function at a point. Recall that for $a \in \mathbb{R}$ and $M \subset \mathbb{R}$, the number a is a *limit point* of the set M if

$$\forall\, k\ \exists\, b \in M\ \big(0 < |a - b| \leq 1/k\big)\ .$$

In other words, there is a sequence $(b_n) \subset M \setminus \{a\}$ such that $\lim b_n = a$. Now for $a \in M \subset \mathbb{R}$, where a is a limit point of M, and a function $f\colon M \to \mathbb{R}$, the *(finite) derivative of f at a* is the number $f'(a) \in \mathbb{R}$ such that for every k there is an n such that

$$b \in M \wedge 0 < |b-a| \le 1/n \Rightarrow \left|\frac{f(b)-f(a)}{b-a} - f'(a)\right| \le 1/k\ .$$

We also allow that $f'(a) = \pm\infty$ and leave the formulation of the corresponding definition to the reader as an exercise. The next theorem is well known.

Theorem G (Lagrange's MVT) *Let $a < b$ be real numbers and $f\colon [a,b] \to \mathbb{R}$ be a continuous function such that for every $c \in (a,b)$ the derivative $f'(c)$ exists (it may be $\pm\infty$). Then*

$$\exists\, c \in (a,\, b) \left(f'(c) = \frac{f(b)-f(a)}{b-a}\right)\ .$$

Also the formula for sums of geometric series is well known.

Proposition H (geometric series) *For every $u \in \mathbb{C}$ with $|u| < 1$ and $n \in \mathbb{N}_0$,*

$$u^n + u^{n+1} + \cdots = \frac{u^n}{1-u}\ .$$

An analytic proof of Theorem 1.2. Let α and n be as stated. We take the minimum polynomial $r(x) \in \mathbb{Q}[x]$ of α. Multiplying it by an integer we get rid of the denominators in the coefficients and may assume that $r(x)$ is integral,

$$r(x) = a_n x^n + \cdots + a_1 x + a_0$$

with $a_j \in \mathbb{Z}$ and $a_n \ne 0$. We set $I := [\alpha-1, \alpha+1]$ and consider an arbitrary fraction $\frac{p}{q} \in \mathbb{Q}$ with $q \in \mathbb{N}$. If it lies outside I then, trivially,

$$\left|\alpha - \frac{p}{q}\right| > 1 \ge \frac{1}{q^n}\ .$$

Let $\frac{p}{q} \in I$ and, for example, $\frac{p}{q} < \alpha$ (the case that $\frac{p}{q} > \alpha$ is treated similarly). We set $a := \frac{p}{q}$, $b := \alpha$ and regard $r = r(x)$ as a real function $r\colon [a,b] \to \mathbb{R}$. (This is the polynomial function associated with $r(x)$ when we regard it as a polynomial in $\mathbb{R}[x]$.) By Theorem G there is a real number $\beta \in (a,b) \subset I$ such that

$$r'(\beta) = \frac{r(\alpha) - r(p/q)}{\alpha - p/q}\ .$$

Now $r(\alpha) = 0$ and

$$|r(p/q)| = \frac{|a_n p^n + \cdots + a_1 p q^{n-1} + a_0 q^n|}{q^n} \ge \frac{1}{q^n}$$

because the integer $a_np^n + \cdots + a_1pq^{n-1} + a_0q^n \neq 0$. Indeed, by the minimality of the degree of $r(x)$, $r(p/q) \neq 0$. (If $r(p/q) = 0$ then $r(x)/(x-p/q)$ is a rational polynomial with degree $n-1$ and root α.) Thus

$$\left|\alpha - \frac{p}{q}\right| \geq \frac{1/|r'(\beta)|}{q^n} \ .$$

Since $|r'|\colon I \to \mathbb{R}$ is continuous and the interval $I \subset \mathbb{R}$ (or $\subset \mathbb{C}$) is compact, by Corollary B we can set $d > 0$ to be the maximum value of $|r'(x)|$ on I. We set $c := \min(\{1, 1/d\}) > 0$ and get Liouville's inequality:

$$\left|\alpha - \frac{p}{q}\right| \geq \frac{\min(\{1,\ 1/|r'(\beta)|\})}{q^n} \geq \frac{\min(\{1,\ 1/d\})}{q^n} = \frac{c}{q^n} \ .$$

□

One can find the previous proof, for example, in the monograph [1, pp. 1–2] of Alan Baker (1939–2018). Leafing through the book I see that the example with the number λ is there too, only in [1] λ is called ξ.

Proof of Corollary 1.3. Clearly, λ is irrational. Suppose for the contrary that it is algebraic. Let $n := \deg \lambda \geq 2$ and let $c := c(\lambda) > 0$ be the constant for λ in Liouville's inequality. For $m \in \mathbb{N}$ we consider the m-th partial sum of the series defining λ:

$$\sum_{j=1}^{m} 10^{-j!} = \frac{p_m}{q_m} = \frac{p_m}{10^{m!}}$$

for some $p_m \in \mathbb{N}$. Then, by Proposition H,

$$\left|\lambda - \frac{p_m}{q_m}\right| = \sum_{j=m+1}^{\infty} 10^{-j!} \leq \frac{1/10^{m!\cdot(m+1)}}{1 - 1/10^{m!}} \leq \frac{10}{9} \cdot \frac{1}{q_m^{m+1}} = \frac{10}{9q_m} \cdot \frac{1}{q_m^m} \ .$$

Any m so large that $10/(9q_m) = 1/(9 \cdot 10^{m!-1}) < c$ produces violation of Liouville's inequality. Hence λ is transcendental. □

The Soviet-Belarusian number theorist Vladimir G. Sprindzhuk (1936–1987) gave in [10, p. 14] a different proof of Liouville's inequality. We present it next, in a simplified form; it seems that Theorem 1.2 can be proven purely algebraically, without analysis. But it only seems.

An "algebraic" proof of Theorem 1.2. Let α, n and $r(x)$ be as before. By Theorem 1.1 (here analysis hides) we have the factorization

$$r(x) = a(x-\alpha_1)(x-\alpha_2)\ldots(x-\alpha_n),\ \alpha_j \in \mathbb{C} \ ,$$

where $a \in \mathbb{N}$ and $\alpha = \alpha_1$. We denote by $d \geq 0$ the maximum distance $|\alpha - \alpha_j|$ for $j = 2, 3, \ldots, n$. (One can show that the roots α_j are all distinct but we do not need it.) Let $\frac{p}{q}$ be an arbitrary fraction with $q \in \mathbb{N}$. Like in the first, analytical proof we distinguish two cases: $|\alpha - \frac{p}{q}| \ge 1$ and $|\alpha - \frac{p}{q}| \le 1$. As we know, in the former case Liouville's inequality holds trivially:

$$\left|\alpha - \frac{p}{q}\right| \ge 1 \ge \frac{1}{q^n}\ .$$

In the latter case $|\alpha - \frac{p}{q}| \le 1$ the triangle inequality implies that $|\alpha_j - \frac{p}{q}| \le 1 + d$ for every $j = 2, 3, \dots, n$. As we know from the first proof, $|r(\frac{p}{q})| \ge q^{-n}$. Thus

$$\left|\alpha - \frac{p}{q}\right| = \frac{|r(p/q)|}{a \cdot |\alpha_2 - p/q| \cdot \ldots \cdot |\alpha_n - p/q|} \ge \frac{(1+d)^{1-n}/a}{q^n}\ .$$

We set $c := \min(\{1, (1+d)^{1-n}/a\})$ and get Liouville's inequality. □

In reality the former proof is analytically simpler than the latter because it does not rely on FTAlg, whereas the latter does.

1.3 e is transcendental

The transcendence of the number $\mathrm{e} = 2.71828\dots$ was proven first by the French mathematician Charles Hermite (1822–1901).

Theorem 1.4 (Ch. Hermite, 1873) *The number* e *is transcendental.*

We first give analytical tools used in the proof. One easily proves by integration by parts the following identity.

Theorem I (values of the gamma function) *For every $n \in \mathbb{N}_0$,*

$$\int_0^{+\infty} x^n \mathrm{e}^{-x}\,\mathrm{d}x = n!\ .$$

The substitution $y = x - m$ and the identity $\mathrm{e}^{x+y} = \mathrm{e}^x\mathrm{e}^y$, $x, y \in \mathbb{R}$, yield the next result.

Proposition J (shifting by m) *For every $m \in \mathbb{N}_0$ and every polynomial $p(x)$ in $\mathbb{Z}[x]$,*

$$\mathrm{e}^m \int_m^{+\infty} p(x)\mathrm{e}^{-x}\,\mathrm{d}x = \int_0^{+\infty} p(x+m)\mathrm{e}^{-x}\,\mathrm{d}x\ .$$

We present a proof of Hermite's theorem devised by David Hilbert (1862–1943) in 1893.

Hilbert's proof of Theorem 1.4 ([5]). Suppose for the contrary that e is algebraic. It follow that for some $n \in \mathbb{N}_0$ and $q_j \in \mathbb{Z}$ with $q_0 \ne 0$,

$$q_0 + q_1\mathrm{e} + \cdots + q_n\mathrm{e}^n = 0\ .$$

For $m \in \mathbb{N}$ we take the integral polynomials

$$p_m(x) := x^m \big((x-1)(x-2)\dots(x-n)\big)^{m+1} .$$

Using Proposition J and the additivity of Riemann integrals we get the identity

$$\begin{aligned} 0 &= \big(q_0 + q_1\mathrm{e} + \cdots + q_n\mathrm{e}^n\big) \int_0^{+\infty} p_m(x)\mathrm{e}^{-x}\,\mathrm{d}x \\ &= \sum_{j=0}^{n} q_j\mathrm{e}^j \int_0^j p_m(x)\mathrm{e}^{-x}\,\mathrm{d}x + \sum_{j=0}^{n} q_j \int_0^{+\infty} p_m(x+j)\mathrm{e}^{-x}\,\mathrm{d}x \\ &=: A(m) + B(m) . \end{aligned}$$

It is not hard to show that for every $m \in \mathbb{N}$,

$$|A(m)| \le c^m$$

for a constant $c > 1$. As for $B(m)$, Theorem I and the linearity of Riemann integrals imply that for every $m \in \mathbb{N}$,

$$B(m) = q_0(-1)^{n(m+1)}(n!)^{m+1} \cdot m! + b_m \cdot (m+1)!$$

for some $b_m \in \mathbb{Z}$. Thus $B(m)$ are integers divisible by $m!$. The crucial observation is that

$$B(m) = 0 \;\Rightarrow\; m+1 \text{ divides } q_0 \cdot (n!)^{m+1} .$$

The conclusion of this implication does not hold, for example, when $m+1$ is coprime with $q_0 \cdot n!$. It follows that $|B(m)| \ge m!$ for infinitely many m. But since always $A(m) + B(m) = 0$ and $A(m)$ is bounded exponentially in m, we get a contradiction. □

1.4 Hilbert's proofs of transcendence of e and π

Here is my translation of the full article [5] of D. Hilbert.

On the transcendence of the numbers e and π

By

David Hilbert in Königsberg in Prussia

———

Let us suppose that the number e satisfies the degree n equation

$$a + a_1e + a_2e^2 + \cdots + a_ne^n = 0 ,$$

whose coefficients a, a_1, $\dots$, a_n are integers. If the left side of this equation is multiplied by the integral

$$\int_0^\infty = \int_0^\infty z^\rho[(z-1)(z-2)\cdots(z-n)]^{\rho+1}e^{-z}dz ,$$

where ρ denotes a positive integer, the expression

$$a\int_0^\infty + a_1 e\int_0^\infty + a_2 e\int_0^\infty + \cdots + a_n e^n\int_0^\infty$$

arises and this expression splits in the sum of the next two expressions:

$$\begin{aligned} P_1 &= a\int_0^\infty + a_1 e\int_1^\infty + a_2 e^2\int_2^\infty + \cdots + a_n e^n\int_n^\infty \ , \\ P_2 &= \qquad\quad a_1 e\int_0^1 + a_2 e^2\int_0^2 + \cdots + a_n e^n\int_0^n \ . \end{aligned}$$

The formula

$$\int_0^\infty z^\rho e^{-z} dz = \rho!$$

shows that the integral $\int_0^\infty$ is an integer divisible by $\rho!$ and it also easily follows, when one applies the respective substitutions $z = z'+1$, $z = z'+2, \ldots, z = z'+n$, that

$$e\int_1^\infty,\ e^2\int_2^\infty,\ \ldots,\ e^n\int_n^\infty,$$

are integers divisible by $(\rho+1)!$. So P_1 is an integer divisible by $\rho!$ and, as we see, modulo $\rho+1$ the congruence

(1) $$\frac{P_1}{\rho!} \equiv \pm a(n!)^{\rho+1} \qquad (\rho{+}1)$$

holds.

On the other hand, when K, resp. k, denotes the maximum that the functions

$$z(z-1)(z-2)\ldots(z-n)\ ,$$

resp.

$$(z-1)(z-2)\ldots(z-n)e^{-z}\ ,$$

attain in absolute value on the interval from $z = 0$ to $z = n$, it holds that

$$\left|\int_0^1\right| < kK^\rho,\ \left|\int_0^2\right| < 2kK^\rho,\ \ldots\ \left|\int_0^n\right| < nkK^\rho$$

and from this it follows, when we set for brevity

$$\kappa = \{|a_1 e| + 2|a_2 e^2| + \cdots + |a_n e^n|\}k\ ,$$

the inequality

(2) $$|P_2| < \kappa K^\rho\,.$$

Now one takes a positive integer ρ which, *first*, is divisible by the integer $a.n!$ and for which, *second*, $\kappa\frac{K^\rho}{\rho!} < 1$ is. Then $\frac{P_1}{\rho!}$ is, as a consequence of the congruence (1), an integer that is not divisible by $\rho{+}1$ and is necessarily different from 0. Moreover this $\frac{P_2}{\rho!}$ is, as a consequence of the inequality (2) and when taken in absolute value, smaller than 1, therefore the equality

$$\frac{P_1}{\rho!} + \frac{P_2}{\rho!} = 0$$

is impossible.

Let π be an algebraic number, namely let the number $a_1 = i\pi$ satisfy a degree n equation with integral coefficients. If we denote by $\alpha_2, \ldots, \alpha_n$ other roots of the equation, then, since $1 + e^{\pi i}$ has value 0, also the expression

$$(1 + e^{\alpha_1})(1 + e^{\alpha_2}) \cdots (1 + e^{\alpha_n}) = 1 + e^{\beta_1} + e^{\beta_2} + \cdots + e^{\beta_N}$$

has to have value 0. Therefrom, as one easily sees, the N exponents $\beta_1, \ldots, \beta_N$ are roots of a degree N equation with integral coefficients. If moreover, say, the M exponents $\beta_1, \ldots, \beta_M$ are different from 0, while the other vanish, then these M exponents $\beta_1, \ldots, \beta_M$ are roots of a degree M equation of the form

$$f(z) = bz^M + b_1 z^{M-1} + \cdots + b_M = 0$$

whose coefficients are also integers and in which especially the last coefficient b_M differs from zero. The above expression then becomes the form

$$a + e^{\beta_1} + e^{\beta_2} + \cdots + e^{\beta_M} + ,$$

where a is a positive integer.

One multiplies this expression by the integral

$$\int_0^\infty = \int_0^\infty z^\rho [g(z)]^{\rho+1} e^{-z} dz \ ,$$

where ρ denotes again a positive integer and where we set for brevity $g(z) = b^M f(z)$; then one gets

$$a \int_0^\infty + e^{\beta_1} \int_0^\infty + e^{\beta_2} \int_0^\infty + \cdots + e^{\beta_M} \int_0^\infty$$

and this expression splits in the sum of the next two expressions:

$$\begin{aligned} P_1 &= a \int_0^\infty + e^{\beta_1} \int_{\beta_1}^\infty + e^{\beta_2} \int_{\beta_2}^\infty + \cdots + e^{\beta_M} \int_{\beta_M}^\infty \ , \\ P_2 &= \qquad\quad e^{\beta_1} \int_0^{\beta_1} + e^{\beta_2} \int_0^{\beta_2} + \cdots + e^{\beta_M} \int_0^{\beta_M} \ , \end{aligned}$$

where generally the integral $\int_{\beta_i}^\infty$ in the complex z-plane is taken from the point $z = \beta_i$ along a line parallel to the real axis up to $+\infty$, and the integral $\int_0^{\beta_i}$ is taken from the point $z = 0$ along the joining straight segment up to the point $z = \beta_i$.

The integral $\int_0^\infty$ is again an integer divisible by $\rho!$ and, as one sees, modulo $\rho + 1$ the congruence

$$\frac{1}{\rho!}\int_0^\infty \equiv b^{\rho M+M} b_M^{\rho+1} \qquad (\rho + 1)$$

holds. By means of the substitution $z = z' + \beta_i$ and by $g(\beta_i) = 0$ one obtains further that

$$e^{\beta_i}\int_{\beta_i}^\infty = \int_0^\infty (z' + \beta_i)^\rho [g(z' + \beta_i)]^{\rho+1} e^{-z'} dz' = (\rho + 1)! G(\beta_i)\,,$$

where $G(\beta_i)$ is an integral function[1] in β_i, with degree in β_i below the number $\rho M + M$ and with all coefficients divisible by $b^{\rho M+M}$. Since $\beta_1, \ldots, \beta_M$ are roots of the integral equation $f(z) = 0$ and therewith after multiplication by the first coefficient b turn in *algebraic integers*,

$$G(\beta_1) + G(\beta_2) + \cdots + G(\beta_M)$$

is necessarily an *integer*. It follows from this that the expression P_1 is an integer divisible by $\rho!$ and modulo $\rho + 1$ the congruence

$$(3) \qquad \frac{P_1}{\rho!} \equiv ab^{\rho M+M} b_M^{\rho+1} \qquad (\rho + 1)$$

holds.

On the other hand, when K, resp. k, denotes the maximum attained in absolute value by the function $zg(z)$, resp. $g(z)e^{-z}$, on the straight integration segments between $z = 0$ and $z = \beta_i$, then

$$\left|\int_0^{\beta_i}\right| < |\beta_i| k K^\rho \qquad (i = 1, 2, \ldots, M)$$

and therefrom it follows, when one sets for brevity

$$\kappa = \left\{\left|\beta_1 e^{\beta_1}\right| + \left|\beta_2 e^{\beta_2}\right| + \cdots + \left|\beta_M e^{\beta_M}\right|\right\} k\,,$$

the inequality

$$(4) \qquad |P_2| < \kappa K^\rho\,.$$

Now one takes a positive integer ρ which, *first*, is divisible by abb_M and for which, *second*, $\kappa\frac{K^\rho}{\rho!} < 1$ is. Then $\frac{P_1}{\rho!}$ is, as a consequence of the congruence (3), an integer not divisible by $\rho + 1$ and thus necessarily different from 0, and since

[1] integral polynomial

moreover $\frac{P_2}{\rho!}$ is, as a consequence of the inequality (4), in absolute value less than 1, the equality

$$\frac{P_1}{\rho!} + \frac{P_2}{\rho!} = 0$$

is impossible.

It is easy to see how by continuing in the way we passed one can equally easily prove also the general theorem of Lindemann on the exponential function.

Königsberg in Prussia, January 5th, 1893.

1.5 Comments on the two previous proofs

Chapter 2

Dirichlet's theorem on prime numbers in AP

The abbreviation AP stands for "arithmetic progression(s)". These are sets $X \subset \mathbb{Z}$ of the form
$$X = \{a + jm \mid j \in I\}$$
where $a \in \mathbb{Z}$, $m \in \mathbb{N}$ and I is a finite or infinite interval of integers. An interval in $\mathbb{Z}$ is any set $I \subset \mathbb{Z}$ such that for any three integers $k < l < m$ with $k, m \in X$ also $l \in X$. The positive integer m is the *common difference* of the progression. In the analytic proof of Dirichlet's theorem in Section 2.2 we will use the notation

$$(a \bmod m) := \{a + nm \mid n \in \mathbb{Z}\}$$

($a \in \mathbb{Z}$, $m \in \mathbb{N}$) and $\mathbb{Z}/m\mathbb{Z} := \{(a \bmod m) \mid a \in [m]\}$.

"Dirichlet's theorem ..." refers to the following famous result in which, more precisely in Dirichlet's proof of which, Analytic Number Theory was born.

Theorem 2.1 (P. Dirichlet, 1837) *Suppose that $a, m \in \mathbb{N}$, say $1 \le a < m$, are coprime numbers. Then the infinite* AP

$$a + m,\ a + 2m,\ a + 3m,\ \ldots$$

contains infinitely many prime numbers.

In other words, for coprime a and m there exist infinitely many primes p of the form $p \equiv a \pmod{m}$. If a and m are not coprime then the AP clearly contains no prime number. The German mathematician (Johann) Peter (Gustav Lejeune) Dirichlet (1805–1859)[1] proved the theorem in [3]. His argument, as given in [3], was complete only in the case when the modulus (common difference) m is a prime number.

[1] His family was of French origin and the name in fact says "le jeune de Richelet", a lad from Richelet.

In the next section we present an interesting elementary argument of P. Erdős that proves some (unfortunately, only finitely many) cases of Dirichlet's theorem. In the section after that we give an analytic proof of full Dirichlet's theorem.

2.1 An elementary argument of P. Erdős

For $m \in \mathbb{N}$ we define the quantity

$$\sigma = \sigma(m) := \sum_{\substack{p \le m \\ (p,\, m)=1}} \frac{1}{p}\ ,$$

so σ is the sum of reciprocals of the prime numbers not exceeding m that do not divide it. If the sum is empty, which happens only for $m = 1$ and 2, we define it as 0. For example,

$$\sigma(6) = \frac{1}{5} \ \text{ and } \ \sigma(7) = \frac{1}{2} + \frac{1}{3} + \frac{1}{5} = \frac{31}{30}\ (> 1)\ .$$

In 1935 the prodigious Hungarian mathematician Paul (Pál) Erdős (1913–1996) gave in [4] an elementary proof for Dirichlet's theorem for moduli m such that $\sigma(m) < 1$.

Theorem 2.2 (P. Erdős, 1935) *Suppose that $a, m \in \mathbb{N}$, say $1 \le a < m$, are coprime numbers such that $\sigma(m) < 1$. Then there exist infinitely many prime numbers $p \equiv a \pmod{m}$.*

Good news is that the argument of P. Erdős is completely elementary, free of analysis and very nice. Bad news is that the set of $m \in \mathbb{N}$ with $\sigma(m) < 1$ is finite. It was determined by P. Moree in [7]:

$$\begin{aligned}\{m \in \mathbb{N} \mid \sigma(m) < 1\} \quad = \quad & \{1,\ 2,\ 3,\ 4,\ 5,\ 6,\ 8,\ 9,\ 10,\ 12,\ 14,\ 15,\ 16,\ 18,\ 20,\ 21, \\ & 22,\ 24,\ 26,\ 28,\ 30,\ 36,\ 40,\ 42,\ 48,\ 50,\ 54,\ 60,\ 66,\ 70, \\ & 72,\ 78,\ 84,\ 90,\ 96,\ 102,\ 108,\ 114,\ 120,\ 126,\ 132, \\ & 138,\ 150,\ 156,\ 168,\ 180,\ 210,\ 240,\ 270,\ 300,\ 330 \\ & 390,\ 420,\ 630,\ 840\}\ .\end{aligned}$$

We note that although 7 is outside the set, $14 = 2 \cdot 7$ lies in it and therefore the modulus $m = 7$ *is* covered by the argument of P. Erdős. In this way all moduli $m = 1, 2, \ldots, 28$ are covered by the argument and $m = 29$ is the smallest one that escapes it.

Open Problem. *Extend the argument of P. Erdős, which we are going to present now, so that it covers some moduli m such that no multiple of m is in the set, for example $m = 29$.*

We turn to the argument of PE. We rework and expand our explanation of it we gave 13 years ago in [6]. If $A\subset\mathbb{Z}$ is a finite set and $d\in\mathbb{N}$ then

$$A(d):=\#\{x\in A\mid d\,|\,x\}\ ,$$

i.e., $A(d)$ counts the multiples of d in A. For a prime number p and $n\in\mathbb{Z}\setminus\{0\}$ we define the *p-adic order* of n by

$$\mathrm{ord}_p(n):=\max(\{k\in\mathbb{N}_0\mid p^k\,|\,n\})\ ,$$

and we set $\mathrm{ord}_p(0):=+\infty$ for any prime p.

Lemma 2.3 *Let $a\in\mathbb{Z}$, $d,m,n\in\mathbb{N}$, d and m be coprime, and let $A\subset\mathbb{Z}$ be the* AP

$$A:=\{a+m,\,a+2m,\,\dots,\,a+nm\}\ .$$

Then the following hold.

1. *$A(d)=\lfloor n/d\rfloor$ or $\lfloor n/d\rfloor+1$.*
2. *The latter case holds for $A(d)$ iff*

   $$\exists\,j\in\{1,\,2,\,\dots,\,r(n,\,d)\}\ \big(d\,|\,(a+jm)\big)$$

   where $r(n,d)=n-d\lfloor n/d\rfloor$ is the remainder for the division of n by d.
3. *If $a=0$ and $m=1$ then the latter case does not occur and $A(d)=\lfloor n/d\rfloor$.*
4. *It holds that*

   $$\mathrm{ord}_p\big(\textstyle\prod_{x\in A}x\big)=\sum_{i\ge1}A(p^i)\ .$$

   It actually holds for any finite set $A\subset\mathbb{Z}$.

Proof. Let a, d, m, n and A be as stated. We note that $(j,k\in\mathbb{Z})$ if $j\not\equiv k\pmod d$ then also $a+jm\not\equiv a+km\pmod d$. Hence we observe that for every interval $I\subset\mathbb{Z}$ with length $|I|=d$ there is exactly one $j\in I$ such that $a+jm\equiv0\pmod d$ and that for $|I|<d$ there is at most one such j. We consider the partition

$$[n]=\{1,\,2,\,\dots,\,n\}=\underbrace{\{1,\,2,\,\dots,\,r(n,\,d)\}}_{I_0}\cup I_1\cup I_2\cup\dots\cup I_{\lfloor n/d\rfloor}$$

where $I_0<I_1<I_2<\dots<I_{\lfloor n/d\rfloor}$ are intervals in $\mathbb{Z}$ and each I_i, $i\ge1$, has length d. By the observation we see that

$$A(d)=\lfloor n/d\rfloor+\delta,\ \ \delta\in\{0,\,1\}\ ,$$

where $\delta=1$ iff $\exists\,j\in I_0$ with $d\,|\,(a+jm)$. Thus we deduce parts 1–3; for part 3 note that if $a=0$ and $m=1$ then for no $j\in I_0$ the number d divides the number $a+jm=j<d$.

In part 4 we assume that $A \subset \mathbb{Z}$ is any finite set. For $0 \in A$ the stated equality holds, $\mathrm{ord}_p(0) = +\infty$ and $A(p^i) \geq 1$ for every $i \in \mathbb{N}$. For $0 \notin A$ we use additivity of p-adic orders ($\mathrm{ord}_p(kl) = \mathrm{ord}_p(k) + \mathrm{ord}_p(l)$ for every $k, l \in \mathbb{Z}$), double-count the pairs in the finite set

$$S := \{(i,\, x) \in \mathbb{N} \times A \mid p^i \,|\, x\}$$

and get that $\mathrm{ord}_p\big(\prod_{x \in A} x\big) = \sum_{x \in A} \mathrm{ord}_p(x)$ indeed equals

$$\textstyle\sum_{x \in A} \#\{i \in \mathbb{N} \mid (i,\, x) \in S\} = \sum_{i \in \mathbb{N}} \#\{x \in A \mid (i,\, x) \in S\} = \sum_{i \geq 1} A(p^i)\;.$$

□

Parts 3 and 4 of the lemma imply the well known formula

$$\mathrm{ord}_p(n!) = \sum_{i \geq 1} \lfloor n/p^i \rfloor$$

due originally to Adrien-Marie Legendre (1752–1833). Parts 1, 3 and 4 show that every binomial coefficient is a natural number. Namely, for every $n, m \in \mathbb{N}$, $n \leq m$, and prime p the p-adic order of

$$\binom{m}{n} = \frac{m(m-1)\dots(m-n+1)}{n!} =: \frac{N}{D}$$

equals $\mathrm{ord}_p(N) - \mathrm{ord}_p(D) = \sum_{i \geq 1}(A(p^i) - B(p^i)) = \sum_{i \geq 1} \delta(i)$, for the n-term APs

$$A = \{m-n+1,\, m-n+2,\, \dots,\, m\} \;\text{ and }\; B = \{1,\, 2,\, \dots,\, n\}$$

and with $\delta(i) \in \{0, 1\}$, thus $\mathrm{ord}_p(\binom{m}{n}) \geq 0$. Below we will see that part 2 is crucial for the functioning of PE's argument.

We extend p-adic order to fractions: if $m/n \in \mathbb{Q}$ is nonzero, we set

$$\mathrm{ord}_p(m/n) := \mathrm{ord}_p(m) - \mathrm{ord}_p(n) \in \mathbb{Z}\;.$$

Again $\mathrm{ord}_p(0/n) := +\infty$.

We begin the proper argument of PE. Let an $m \in \mathbb{N}$ with $m \geq 3$ be the given modulus and an $a \in \mathbb{N}$ coprime to m, $1 \leq a < m$, be the given coprime residue class. We are going to show that if $\sigma(m) < 1$ then there exist infinitely many primes $p \equiv a \pmod{m}$. Let $p_1 < p_2 < \dots < p_h < m$ be all prime numbers not exceeding m and not dividing it. Since $m \geq 3$, $h \geq 1$. We set $P := p_1 p_2 \dots p_h$ and take the unique numbers $q_1, \dots, q_h \in \mathbb{N}$ such that $1 \leq q_i < m$ and $q_i p_i \equiv a \pmod{m}$. Thus every q_i is coprime to m. For $n \in \mathbb{N}$ and divisible by P we define

$$P_n(a,\, m) := \frac{(a+m)(a+2m)\dots(a+nm)}{n!} \in \mathbb{Q}$$

and

$$Q_n(a,\, m) := \frac{P_n(a,\, m)}{P_{n/p_1}(q_1,\, m) P_{n/p_2}(q_2,\, m) \dots P_{n/p_h}(q_h,\, m)} \in \mathbb{Q}\;.$$

Proposition 2.4 (properties of $P_n(a,m)$ and $Q_n(a,m)$) *Let m, a, p_i, q_i and P be as above, recall that $\sigma = \sigma(m) = \sum_{i=1}^{h} 1/p_i$, let $n \in \mathbb{N}$ be a multiple of P and let p be a prime number. The following hold.*

1. *For $n \to \infty$ on multiples of P,*
$$Q_n(a,\, m) = m^{(1-\sigma)n+o(n)}\ .$$

2. *Let p and m be coprime and $k := \mathrm{ord}_p(P_n(a,m))$. Then $1 \le p^k < (n+1)m$.*

3. *If $p \,|\, m$ and $\sigma \le 1$ then $\mathrm{ord}_p(Q_n(a,m)) \le 0$.*

4. *Suppose that $n \ge m$ and that $p > \sqrt{(n+1)m}$. Then*
$$\mathrm{ord}_p(P_n(a,m)) \in \{0,1\} \ \ \textit{and}\ \ \forall\, i \in [h]\,\big(\mathrm{ord}_p(P_{n/p_i}(q_i,m)) \in \{0,1\}\big)\ .$$

5. *With the assumptions of part 4, if in addition $p \not\equiv a \pmod m$ and $\mathrm{ord}_p(P_n(a,m)) = 1$ then*
$$\exists\, i \in [h]\,\big(\mathrm{ord}_p(P_{n/p_i}(q_i,m)) = 1\big)\ .$$

Proof. 1. For every $j \in [n]$, $jm < a + jm < (j+1)m$. We multiply these $2n$ inequalities, divide the resulting pair of inequalities by $n!$ and get that
$$m^n < P_n(a,\, m) < (n+1)m^n\ .$$
So $P_n(a,m) = m^{n+o(n)}$ for $n \to \infty$. We substitute this asymptotics in the definition of $Q_n(a,m)$ and obtain part 1.

2. Let p, m and k be as stated. Consider the APs
$$A := \{a+m,\, a+2m,\, \dots,\, a+nm\}\ \text{ and }\ B := \{1,\, 2,\, \dots,\, n\}\ .$$
As above, by parts 1, 3 and 4 of Lemma 2.3 we have that
$$k = \mathrm{ord}_p\big(\textstyle\prod_{x\in A} x\big) - \mathrm{ord}_p\big(\prod_{x\in B} x\big) = \sum_{i\ge 1}(A(p^i) - B(p^i)) = \sum_{i\ge 1}\delta(i)$$
where $\delta(i) \in \{0,1\}$ for every $i \in \mathbb{N}$ and $\delta(i) = 0$ if $p^i \ge (n+1)m$ (since $\max(A), \max(B) < (n+1)m$). We obtain part 2.

3. Now p does not divide any of the factors $a + jm$ and $q_i + jm$ in the numerators of the fractions $P_n(a,m)$ and $P_{n/p_i}(q_i,m)$ and therefore
$$\mathrm{ord}_p\big(Q_n(a,\, m)\big) = -\mathrm{ord}_p\big(\tfrac{n!}{n/p_1!\, n/p_2! \dots n/p_h!}\big) \le 0$$
because $\frac{n}{p_1} + \frac{n}{p_2} + \cdots + \frac{n}{p_h} \le n$ — the displayed ratio of factorials is a natural number as it is a multiple of a multinomial coefficient.

4. Let n, m and p be as stated. Then $p \ge \sqrt{(m+1)m} > m$ and p is coprime to m. For $i = 0, 1, \dots, h$ we again consider the APs
$$A_i := \{q_i + m,\, q_i + 2m,\, \dots,\, q_i + (n/p_i)\cdot m\}\ \text{ and }\ B := \{1,\, 2,\, \dots,\, n\}$$

where $q_0 := a$ and $p_0 := 1$. Let $k_i := \mathrm{ord}_p(P_{n/p_i}(q_i, m))$. Again by parts 1, 3 and 4 of Lemma 2.3,

$$k_i = \mathrm{ord}_p\big(\textstyle\prod_{x\in A_i} x\big) - \mathrm{ord}_p\big(\prod_{x\in B} x\big) = \sum_{l\ge 1}(A_i(p^l) - B(p^l)) = \sum_{l\ge 1}\delta_i(l)$$

where $\delta_i(l) \in \{0,1\}$ for every $i = 0, 1, \ldots, h$ and every $l \in \mathbb{N}$. But now $\delta_i(l) = 0$ for $l \ge 2$ because $p^2 > (n+1)m > \max(A_i)$, $i = 0, 1, \ldots, h$. We easily get part 4.

5. We assume that n, m and p are as in part 4 and that in addition $p \not\equiv a \pmod m$ and $\mathrm{ord}_p(P_n(a, m)) = 1$. By parts 2 and 4 of Lemma 2.3 there is a $j \in [r(n, p)]$ such that p divides $a + jm$. Thus $1 \le j \le r := n - p\lfloor \frac{n}{p} \rfloor$ and

$$a + jm = pb,\ b \in \mathbb{N}\ .$$

Clearly, b and m are coprime. Since $1 \le j < p$, $1 \le a < m$ and $p > m$, we see that $b = (a + jm)/p \in [1, m)$. But $b > 1$ because p is not a modulo m. So $1 < b < m$ and since b and m are coprime, there is an $i \in [h]$ such that $p_i \mid b$. We write that $b = p_i c$, $c \in \mathbb{N}$. By the above definitions we have that $a = p_i q_i + tm$ for some $t \in \mathbb{Z}$. From $p_i, q_i > 0$ and $a < m$ it follows that $t \le 0$. We substitute in the last displayed equation the expressions for b and a and get that

$$p_i q_i + (t + j)m = pp_i c\ .$$

It follows that $p_i \mid (t+j)$ and we write that $t+j = p_i j'$ for some $j' \in \mathbb{Z}$. Canceling p_i we get that

$$q_i + j'm = pc\ .$$

If we show that $j' \in [r(n/p_i, p)]$, we are done because then by parts 2 and 4 of Lemma 2.3 it holds that

$$\mathrm{ord}_p(P_{n/p_i}(q_i,\, m)) = 1\ .$$

The first inequality in $1 \le j' \le r' := \frac{n}{p_i} - p\lfloor \frac{n/p_i}{p} \rfloor$ clearly holds because $j' = (pc - q_i)/m > (m - m)/m = 0$. Suppose for contradiction that $r' < j'$. Then

$$0 \le p_i r' < p_i j' = t + j \le j \le r \ \text{ and }\ 0 \le p_i r' < r$$

where r is the remainder for division of n by p. But $p_i r' = n - p \cdot p_i \lfloor \frac{n/p_i}{p} \rfloor$, which implies that $p_i r'$ is the same remainder. Thus $p_i r' = r$, which contradicts the last displayed inequality. □

Proof of Theorem 2.2. Let $a, m \in \mathbb{N}$, $m \ge 3$ and $1 \le a < m$, be coprime numbers such that $\sigma = \sigma(m) = \sum_{i=1}^{h} \frac{1}{p_i} < 1$ and let $P = p_1 p_2 \ldots p_h$ be the product of all primes not exceeding m and not dividing it. Then for $n \to \infty$ (with $n \ge m$) via multiples of P we have the relations

$$\begin{aligned} m^{(1-\sigma)n + o(n)} = Q_n(a,\, m) &= \prod_p p^{\mathrm{ord}_p(Q_n(a,\, m))} \\ &\le \prod_{p \le \sqrt{(n+1)m}} (n+1)m \prod_{\substack{p < (n+1)m \\ p \equiv a \pmod m}} p\ . \end{aligned}$$

The first equality holds by part 1 of the last proposition. We get the second equality by taking the prime factorization of $Q_n(a,m)$. Finally, for $n \geq m$ we cover all prime numbers p by the following five sets:

$$\begin{aligned}
&X_1 \cup \cdots \cup X_5 := \{p \mid p \geq (n+1)m\} \cup \\
&\cup \{p \mid \sqrt{(n+1)m} < p < (n+1)m \wedge p \equiv a \pmod{m}\} \cup \\
&\cup \{p \mid \sqrt{(n+1)m} < p < (n+1)m \wedge p \not\equiv a \pmod{m}\} \cup \\
&\cup \{p \mid (p,\, m) = 1 \wedge p \leq \sqrt{(n+1)m}\} \cup \{p \mid p \,|\, m\}\ .
\end{aligned}$$

By part 2 of Proposition 2.4, $p \in X_1 \Rightarrow \mathrm{ord}_p(Q_n(a,m)) = 0$ and corresponding factors in the factorization are bounded by ≤ 1. For $p \in X_2$ corresponding factors are bounded, due to part 4, by the second product on the right-hand side. For $p \in X_3$ corresponding factors are bounded, due to parts 4 and 5, by ≤ 1. For $p \in X_4$ corresponding factors are bounded, due to part 2, by the first product. By part 3, $p \in X_5 \Rightarrow \mathrm{ord}_p(Q_n(a,m)) \leq 0$ and corresponding factors are bounded by ≤ 1.

As for the first product,

$$\prod_{p \leq \sqrt{(n+1)m}} (n+1)m \leq \big((n+1)m\big)^{\sqrt{(n+1)m}} = m^{o(n)}\ .$$

Dividing the above displayed bound on $Q_n(a,m)$ by this one we therefore get, for $n \to \infty$ on multiples of P, the lower bound

$$\prod_{\substack{p<(n+1)m \\ p\equiv a \pmod{m}}} p > m^{(1-\sigma)n+o(n)} \to +\infty \quad (\sigma \in [0,\, 1))\ .$$

Thus the set of primes p such that $p \equiv a \pmod{m}$ is infinite. □

2.2 H. N. Shapiro's analytic proof

We prove Theorem 2.1 in a stronger form given in Theorem 2.30 below: if $a, m \in \mathbb{N}$ are coprime numbers then for all real $x > 1$,

$$\sum_{\substack{p \leq x \\ p\equiv a \pmod{m}}} \frac{\log p}{p} = \frac{\log x}{\varphi(m)} + O(1)$$

where p denotes prime numbers and the empty sum is defined as 0. For $x \to +\infty$ the right side goes to $+\infty$ with and Theorem 2.1 follows. The proof is due to Harold N. Shapiro (1922–2013) in [9], its characteristic feature is Proposition 2.29 which elaborates a consequence of the counterfactual that $L(\chi) = 0$.

Recall that *Euler's totient function* $\varphi\colon \mathbb{N}\to\mathbb{N}$ counts coprime residue classes, so that

$$\varphi(n) = |\{m\in[n] \mid (m,\, n)=1\}| = n(1-1/p_1)(1-1/p_2)\dots(1-1/p_k)$$

where $p_1<p_2<\dots<p_k$ are (all) prime divisors of n.

We say that a function $f\colon \mathbb{N}\to\mathbb{C}$ is *completely multiplicative* if $f(mn)=f(m)f(n)$ for every $m,n\in\mathbb{N}$. It is *strongly bounded* if

$$\exists\, c>0\;\forall\,\text{finite interval } I\subset\mathbb{N}\left(\left|\textstyle\sum_{x\in I} f(x)\right|\le c\right).$$

Strong boundedness implies boundedness. The function f is strongly bounded iff the stated inequality holds for every interval $I=[n]$. Clearly, if f is completely multiplicative and bounded then $|f(n)|\le 1$ for every n. For $a,m\in\mathbb{N}$ the characteristic (indicator) functions

$$\chi_m,\, \mathbb{I}_{a,\,m}\colon \mathbb{N}\to\{0,\,1\}$$

are defined by

$$\chi_m(n)=1 \iff (n,\,m)=1 \;\text{ and }\; \mathbb{I}_{a,\,m}(n)=1 \iff n\equiv a \pmod{m}$$

— χ_m is the indicator function of numbers coprime to m and $\mathbb{I}_{a,m}$ of numbers congruent to a modulo m. Note that χ_m is completely multiplicative but that it is not strongly bounded. $\mathbb{I}_{a,m}$ is not strongly bounded either.

The first of the two hearts of this proof, in fact of any proof, of Dirichlet's theorem is the next partition of $\mathbb{I}_{a,m}$ in a linear combination of completely multiplicative functions that are also strongly bounded, except one. Interestingly, even though $\mathbb{I}_{a,m}$ has values just 0 and 1, functions in the linear combination are complex-valued.

Theorem 2.5 (complex partitions of $\mathbb{I}_{a,m}$) *For every $m\in\mathbb{N}$ there exists a finite set $D=D(m)$ of functions $\chi\colon\mathbb{N}\to\mathbb{C}$ with the following properties.*

1. *$\chi_m\in D$ and every $\chi\in D$ is completely multiplicative and, except χ_m, strongly bounded. Also, $\chi\in D\Rightarrow\overline{\chi}\in D$.*
2. *For any $a\in\mathbb{N}$ coprime to m and for any $\chi\in D$ there is a coefficient $c_{a,\chi}\in\mathbb{C}$ such that, for any $a'\equiv 1 \pmod{m}$,*

$$c_{a,\,\chi_m}=c_{a',\,\chi}=\frac{1}{\varphi(m)} \;\text{ and }\; \mathbb{I}_{a,\,m}=\sum_{\chi\in D(m)} c_{a,\,\chi}\cdot\chi\,.$$

The functions in $D(m)$ are called *Dirichlet characters (modulo m)* and we obtain them from so called characters of (finite Abelian) groups. The real-valued χ in $D(m)$ (for example χ_m) are called *real Dirichlet characters*, and else they are *non-real Dirichlet characters.*

For the proof of Theorem 2.5 we need several results on characters. A *group* means in this proof always a finite Abelian group

$$G = (G,\, 1_G,\, \cdot)\ ,$$

in multiplicative notation. A *character* χ of G is a group homomorphism

$$\chi\colon (G,\, 1_G,\, \cdot) \to (\mathbb{C}^\times,\, 1,\, \cdot)$$

from G to the multiplicative group of nonzero complex numbers. So $\chi\colon G \to \mathbb{C} \setminus \{0\}$, $\chi(ab) = \chi(a)\chi(b)$ for every $a, b \in G$ and $\chi(1_G) = 1$. We denote the set of characters of G by G^*. Every set G^* contains the *principal character* χ_0 such that $\chi_0(g) = 1$ for every $g \in G$.

Proposition 2.6 (existence of characters) *If $G \subset H$ is an extension of groups with cyclic factor group H/G of order $n \in \mathbb{N}$, then for every character $\chi \in G^*$ there exist exactly n distinct characters $\psi \in H^*$ such that*

$$\psi \,|\, G = \chi\ ,$$

and every extension of χ to a character of H is one of them.

Proof. Let aG be a generator of H/G: $a \in H$, $a^n = b \in G$ and every $c \in H$ has the unique *standard expression*

$$c = a^i g \ \text{ with } \ 0 \le i < n \wedge g \in G\ .$$

Let $\psi \in H^*$ and $\chi \in G^*$ be such that $\psi \,|\, G = \chi$. Then, with the standard expression for c,

$$\psi(c) = \psi(a)^i \psi(g) = \psi(a)^i \chi(g)\ .$$

Also, $\psi(a)^n = \psi(a^n) = \psi(b) = \chi(b)$ and $\psi(a)$ is an n-th root of $\chi(b)$. So we associate with each of the n numbers $\alpha \in \mathbb{C}$, $\alpha^n = \chi(b)$, the function $\psi_\alpha\colon H \to \mathbb{C}^\times$ defined by

$$\psi_\alpha(c) := \alpha^i \chi(g)$$

for the standard expression $c = a^i g$. Since $\psi_\alpha(a) = \alpha$, these n maps are distinct and the above ψ is one of them. It remain to show that each $\psi = \psi_\alpha$ is a character of H. Let $c_j = a^{i_j} g_j$ with $j = 1, 2$ be two standard expressions. If $0 \le i_1 + i_2 < n$ then

$$\psi(c_1 c_2) = \psi(a^{i_1+i_2} g_1 g_2) = \alpha^{i_1+i_2} \chi(g_1 g_2) = \alpha^{i_1} \chi(g_1) \alpha^{i_2} \chi(g_2) = \psi(c_1)\psi(c_2)\ .$$

Else $i_1 + i_2 = n + j$ with $0 \le j < n$. But then also

$$\begin{aligned} &\psi(c_1 c_2) = \psi(a^n a^j g_1 g_2) = \psi(a^j b g_1 g_2) = \alpha^j \chi(b) \chi(g_1) \chi(g_2) \\ &= \alpha^{j+n} \chi(g_1) \chi(g_2) = \alpha^{i_1} \chi(g_1) \alpha^{i_2} \chi(g_2) = \psi(c_1)\psi(c_2)\ . \end{aligned}$$

So $\psi_\alpha \in H^*$. □

Corollary 2.7 ($|G^*| = |G|$) *For every group G we have that $|G^*| = |G|$, there are as many characters of G as G has elements.*

Proof. Let $G = (G, 1_G, \cdot)$ be any group. We set $G_0 = (\{1_G\}, 1_G, \cdot)$ to be the trivial subgroup of G. For $n \in \mathbb{N}_0$, if the subgroups G_0, G_1, $\dots$, G_n of G have been already defined and still $G_n \neq G$, we take any new element $g \in G \setminus G_n$ and set

$$G_{n+1} := \langle G_n \cup \{g\} \rangle \ ,$$

i.e., G_{n+1} is the subgroup of G generated by the stated union. Since G is finite there is an n such that $G_n = G$. We get a chain of subgroups

$$\{1_G\} = G_0 \subset G_1 \subset \dots \subset G_n = G$$

such that every factor-group G_i/G_{i-1}, $i \in [n]$, is cyclic with order $n_i := |G_i/G_{i-1}| \geq 2$. For $n = 0$ trivially $|G^*| = |G|$ because then $G = G_0 = G_n = \{1_G\}$ and $G^* = \{\chi_0\}$. We assume that $n \geq 1$. We express the cardinality $|G|$ by the cancelative product

$$|G| = |G_0| \cdot \prod_{i=1}^{n} \frac{|G_i|}{|G_{i-1}|} = \prod_{i=1}^{n} |G_i/G_{i-1}| = \prod_{i=1}^{n} n_i \ .$$

By the previous proposition, the map

$$G_i^* \ni \psi \mapsto \psi \,|\, G_{i-1} = \chi \in G_{i-1}^*$$

is n_i-to-1 and onto. Thus also $|G_i^*|/|G_{i-1}^*| = n_i$. As we know, $|G_0^*| = 1$. Thus

$$|G^*| = |G_n^*| = |G_0^*| \cdot \prod_{i=1}^{n} \frac{|G_i^*|}{|G_{i-1}^*|} = \prod_{i=1}^{n} n_i = |G| \ .$$

□

Corollary 2.8 (detecting $a \neq 1_G$) *For every group G and every element a in $G \setminus \{1_G\}$ there is a character $\chi \in G^*$ such that $\chi(a) \neq 1$.*

Proof. In the chain of subgroups of G in the previous proof we set $G_1 = \langle \{1_G\} \cup \{a\} \rangle = \langle \{a\} \rangle$. Then G_1 is a cyclic group generated by a and $|G_1| = n_1 \geq 2$. So there exists a $\psi \in G_1^*$ such that $\psi(a) \neq 1$. Using Proposition 2.6 we extend ψ along the chain to a character of G and get the desired χ. □

Proposition 2.9 (1st orthogonal relation) *For every group G and every character $\chi \in G^*$,*

$$\sum_{a \in G} \chi(a) = \begin{cases} |G| & \dots \quad \chi = \chi_0 \ \textit{and} \\ 0 & \dots \quad \chi \neq \chi_0 \ . \end{cases}$$

Proof. For $\chi = \chi_0$ the claim holds trivially. If $\chi \neq \chi_0$ then there is a $b \in G$ with $\chi(b) \neq 1$. But then

$$S := \sum_{a \in G} \chi(a) = \sum_{a \in G} \chi(ab) = \sum_{a \in G} \chi(a)\chi(b) = S \cdot \chi(b)$$

and $S = 0$. □

For a group G we define on its characters G^* a binary operation $\odot$:

$$\chi \odot \psi \colon G \to \mathbb{C}^\times,\ (\chi \odot \psi)(g) := \chi(g)\psi(g),\ g \in G\ .$$

It is easy to check that $\chi \odot \psi \in G^*$.

Proposition 2.10 (on $\odot$) *For every group G,*

$$G^* = (G^*,\, \chi_0,\, \odot)$$

is a group.

Proof. Since for every $\chi \in G^*$ and every $g \in G$ we have that $(\chi_0 \in G^*)$

$$(\chi \odot \chi_0)(g) = \chi(g)\chi_0(g) = \chi(g) = \chi_0(g)\chi(g) = (\chi_0 \odot \chi)(g)\ ,$$

$\chi \odot \chi_0 = \chi = \chi_0 \odot \chi$ and χ_0 is a neutral element with respect to $\odot$. In the same way we check associativity and commutativity of $\odot$. If $\chi, \psi, \theta \in G^*$ then for every $g \in G$,

$$[(\chi \odot \psi) \odot \theta](g) = \cdots = \chi(g)\psi(g)\theta(g) = \cdots = [\chi \odot (\psi \odot \theta)](g)$$

and

$$[\chi \odot \psi](g) = \chi(g)\psi(g) = \psi(g)\chi(g) = [\psi \odot \chi](g)\ .$$

So $(\chi \odot \psi) \odot \theta = \chi \odot (\psi \odot \theta)$ and $\chi \odot \psi = \psi \odot \chi$. For any $\chi \in G^*$ we define $\chi^{-1} \colon G \to \mathbb{C}^\times$ by $(\chi^{-1})(g) := 1/\chi(g)$. It is easy to check that $\chi^{-1} \in G^*$. Since

$$[\chi \odot \chi^{-1}](g) = \chi(g) \cdot (1/\chi(g)) = 1$$

for every $g \in G$, $\chi \odot \chi^{-1} = \chi_0$ and χ^{-1} is an inverse of χ in the group G^*. □

Proposition 2.11 (2nd orthogonality relation) *For any group G and any element $g \in G$,*

$$\sum_{\chi \in G^*} \chi(g) = \left\{ \begin{array}{lll} |G^*| = |G| & \ldots & g = 1_G \ \textit{and} \\ 0 & \ldots & g \neq 1_G\ . \end{array} \right.$$

Proof. For $g = 1_G$ the claim is trivial. For $g \neq 1_G$ there is by Corollary 2.8 a character $\psi \in G^*$ such that $\psi(g) \neq 1$. Then, like before,

$$S := \sum_{\chi \in G^*} \chi(g) = \sum_{\chi \in G^*} (\chi \odot \psi)(g) = \sum_{\chi \in G^*} \chi(g)\psi(g) = S \cdot \psi(g)$$

and $S = 0$. The second equality holds due to the previous proposition because in the group G^* the map $\chi \mapsto \chi \odot \psi$ is a permutation of G^*. □

Proposition 2.12 (roots of unity) $\forall\, G\ \exists\, n\ \forall\, \chi \in G^*\ \forall\, g \in G\ \big(\chi(g)^n = 1\big)$ *— values of characters are roots of unity and* $\forall\, G\ \forall\, \chi \in G^*\ \forall\, g \in G\ \big(|\chi(g)| = 1\big)$.

Proof. We set $n := |G|$ to the order of G. Since $g^n = 1_G$ for every $g \in G$, for every $\chi \in G^*$ and every $g \in G$ one has that $\chi(g)^n = \chi(g^n) = \chi(1_G) = 1$. □

We define Dirichlet characters.

Proof of Theorem 2.5. Let $m \in \mathbb{N}$. We consider the multiplicative and $\varphi(m)$-element group

$$G(m) := ((\mathbb{Z}/m\mathbb{Z})^\times,\ (1 \bmod m),\ \cdot)$$

of residues classes modulo m coprime to m. We associate to each of the $\varphi(m)$ characters $\chi' \in G(m)^*$ a map $\chi\colon \mathbb{N} \to \mathbb{C}$ by setting for $n \in \mathbb{N}$ its value to

$$\chi(n) := \left\{ \begin{array}{lll} \chi'\big((n \bmod m)\big) & \ldots & (n,\, m) = 1 \ \text{ and} \\ 0 & \ldots & (n,\, m) > 1\ . \end{array} \right.$$

We define $D(m) := \{\chi \mid \chi' \in G(m)^*\}$.

1. We have $\chi_m \in D(m)$ because χ_m arises from the principal character $\chi_0' \in G(m)^*$.

Let $\chi \in D(m)$ and $k, l \in \mathbb{N}$. If one of k and l is not coprime to m then nor is kl and $\chi(kl) = 0 = 0 = \chi(k)\chi(l)$. So let $(k, m) = (l, m) = 1$. Then

$$\begin{aligned} \chi(kl) &= \chi'\big((kl \bmod m)\big) = \chi'\big((k \bmod m) \cdot (l \bmod m)\big) \\ &= \chi'\big((k \bmod m)\big) \cdot \chi'\big((l \bmod m)\big) = \chi(k)\chi(l)\ . \end{aligned}$$

Let $\chi \in D(m) \setminus \{\chi_m\}$, so χ arises from a non-principal $\chi' \in G(m)^*$, and let $I \subset \mathbb{N}$ be a nonempty finite interval. We show that

$$|\textstyle\sum_{n \in I} \chi(n)| \le \varphi(m) - 1\ .$$

We split $I = I_0 \cup I_1 \cup \cdots \cup I_k$ in the intervals $I_0 < I_1 < \cdots < I_k$ such that each I_i with $i \ge 1$ has length m and $|I_0| < m$. Numbers in each I_i are mutually non-congruent modulo m. Therefore for $i \ge 1$,

$$\textstyle\sum_{n \in I_i} \chi(n) = \sum_{g \in G(m)^*} \chi'(g) = 0$$

by Proposition 2.9. As for $\sum_{n\in I_0}\chi(n)$, if I_0 contains all $\varphi(m)$ residues modulo m coprime to m then also $\sum_{n\in I_0}\chi(n)=0$, else

$$|\textstyle\sum_{n\in I_0}\chi(n)|\le 1+1+\cdots+1\le\varphi(m)-1$$

by Proposition 2.12 and the triangle inequality. Altogether we get the stated bound on $|\sum_{n\in I}\chi(n)|$.

Finally, by the previous proposition we know that every nonzero value of every $\chi\in D(m)$ lies on the unit circle $S\subset\mathbb{C}$. Since $z\in S\Rightarrow 1/z=\overline{z}$, if $\chi\in D(m)$ arises from $\chi'\in G(m)^*$ then $\overline{\chi}$ arises from $(\chi')^{-1}\in G(m)^*$ and $\overline{\chi}\in D(m)$.

2. For $a\in\mathbb{N}$ coprime to m and for $\chi\in D(m)$ arising from $\chi'\in G(m)^*$ we define

$$c_{a,\,\chi}:=\frac{\chi'\big((a\bmod m)^{-1}\big)}{\varphi(m)}\ .$$

Then, for any $a'\equiv 1\pmod m$,

$$c_{a,\,\chi_m}=\frac{\chi_0'\big((a\bmod m)^{-1}\big)}{\varphi(m)}=\frac{1}{\varphi(m)}\ \text{ and }\ c_{a',\,\chi}=\frac{\chi'(1_{G(m)})}{\varphi(m)}=\frac{1}{\varphi(m)}\ .$$

Finally, let $a\in\mathbb{N}$ be coprime to m and $n\in\mathbb{N}$ be arbitrary. Then by Proposition 2.11 and the definition of $c_{a,\chi}$, the value $\big(\sum_{\chi\in D(m)}c_{a,\chi}\cdot\chi\big)(n)$ equals

$$\begin{aligned}&\frac{1}{\varphi(m)}\sum_{\chi'\in G(m)^*}\chi'\big((a\bmod m)^{-1}\big)\cdot\chi'\big((n\bmod m)\big)=\\&=\frac{1}{\varphi(m)}\sum_{\chi'\in G(m)^*}\chi'\big((a\bmod m)^{-1}\cdot(n\bmod m)\big)=\\&=\left\{\begin{array}{lll}1&\dots&n\equiv a\pmod m\ \text{ and}\\0&\dots&\text{else}\ ,\end{array}\right.\\&=\mathbb{I}_{a,\,m}(n)\ .\end{aligned}$$

□

We proceed to the next part of the proof which uses the *von Mangoldt function* $\Lambda\colon\mathbb{N}\to[0,+\infty)$. It has values $\Lambda(n)=\log p$ if $n=p^k$ for some $k\in\mathbb{N}$ and prime p, and $\Lambda(n)=0$ else.

Proposition 2.13 (**Λ and** log) *For every $n\in\mathbb{N}$,*

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

Proof. If $n=p_1^{a_1}p_2^{a_2}\dots p_k^{a_k}$ is the prime factorization of n then

$$\sum_{d\,|\,n}\Lambda(d)=\sum_{i=1}^{k}\sum_{j=1}^{a_i}\log p_i=\sum_{i=1}^{k}a_i\log p_i=\log\big(p_1^{a_1}p_2^{a_2}\dots p_k^{a_k}\big)=\log n\ .$$

□

Proposition 2.14 (on $\sum \log p$) *For every real $x > 1$,*

$$\sum_{p \le x} \log p < (2 \log 2)x\ .$$

Proof. We deduce for $n \in \mathbb{N}$ the two inequalities

$$\prod_{n+1 < p \le 2n+1} p \le \binom{2n+1}{n} = \frac{(2n+1)!}{n!(n+1)!} < 4^n\ .$$

The first one follows at once from the divisibility of the binomial coefficient by every prime in the stated range. The second inequality is a corollary of the binomial expansion $2 \cdot 4^n = (1+1)^{2n+1} = \sum_{i=0}^{n} \binom{2n+1}{i}$. Thus for every $n \in \mathbb{N}$ we have the estimate

$$\sum_{n+1 < p \le 2n+1} \log p < (2 \log 2)n\ .$$

We use it to prove by induction on $m \in \mathbb{N}$ the bound

$$\sum_{p \le m} \log p < (2 \log 2)m\ .$$

For $m = 1$ and 2 it clearly holds (with the empty sum defined as 0). For even $m > 2$ the bound holds too by induction because the sum is the same as for $m - 1$. Let $m = 2n + 1 > 2$ be odd. Then

$$\begin{aligned} \sum_{p \le m} \log p &= \sum_{p \le n+1} \log p + \sum_{n+1 < p \le 2n+1} \log p < (2 \log 2)(n+1) + (2 \log 2)n \\ &= (2 \log 2)m \end{aligned}$$

where we bounded the last sum by the above estimate and the sum before it by induction. We set $m := \lfloor x \rfloor$ and have the stated inequality. □

Proposition 2.15 (on $\sum \Lambda(n)$) *For every real $x > 1$,*

$$\sum_{n \le x} \Lambda(n) < 3x\ .$$

Proof. It is not hard to compute that the maximum value of $(\log x)/\sqrt{x}$ for $x > 1$ equals $2/\mathrm{e}$. Using this and the bound in the previous proposition we get that

$$\begin{aligned} \sum_{n \le x} \Lambda(n) &= \sum_{p \le x} \log p + \sum_{p^k \le x,\, k \ge 2} \log p < (2 \log 2)x + (2 \log 2) \sum_{k=2}^{\lfloor \log x / \log 2 \rfloor} x^{1/k} \\ &\le (2 \log 2)x + 2\sqrt{x} \log x \le (2 \log 2 + 4/\mathrm{e})x < 3x\ . \end{aligned}$$

□

Proposition 2.16 (on $\sum \log n$) *For all real $x > 1$,*

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

Proof. This follows from the integral estimates

$$\int_1^{\lfloor x \rfloor} \log t \ \mathrm{d}t \le \sum_{n \le x} \log n \le \int_2^{\lfloor x \rfloor + 1} \log t \ \mathrm{d}t$$

and from the antiderivative $\int \log t = t \log t - t$. □

Proposition 2.17 (on $\sum (\log p)/p$) *For all real $x > 1$,*

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

Proof. We have that

$$\begin{aligned}
x \log x + O(x) &\overset{\text{Prop. 2.16 and 2.13}}{=} \sum_{n \le x} \log n = \sum_{n \le x} \sum_{d \,|\, n} \Lambda(d) \\
&\overset{\text{swapping } \sum\text{s}}{=} \sum_{d \le x} \lfloor x/d \rfloor \Lambda(d) \\
&\overset{\lfloor \alpha \rfloor = \alpha - \{\alpha\}}{=} x \sum_{d \le x} \frac{\Lambda(d)}{d} + \delta \sum_{d \le x} \Lambda(d), \ \ \delta \in [-1, 0] \ , \\
&\overset{\text{def. of } \Lambda, \text{ Prop. 2.15}}{=} x \Bigg(\sum_{p \le x} \frac{\log p}{p} + \sum_{p^k \le x, \, k \ge 2} \frac{\log p}{p^k} \Bigg) + O(x) \\
&\overset{\sum_{n, \, k \ge 2} (\log n)/n^k \text{ conv.}}{=} x \sum_{p \le x} \frac{\log p}{p} + O(x) \ .
\end{aligned}$$

Dividing by x we get the stated result. □

A new actor enters the stage, for any function $f \colon \mathbb{N} \to \mathbb{C}$ we define the series

$$L(f) := \sum_{n=1}^{\infty} \frac{f(n)}{n} \ .$$

Its convergence for strongly bounded f follows from the next useful inequality.

Proposition 2.18 (Abel's inequality) *Let $n \in \mathbb{N}$, for $i \in [n]$ let $a_i \in \mathbb{C}$ and $b_i \in \mathbb{R}$ satisfy $b_1 \ge b_2 \ge \cdots \ge b_n \ge 0$, and let $A_i := a_1 + a_2 + \cdots + a_i$. Then*

$$|a_1 b_1 + a_2 b_2 + \cdots + a_n b_n| \le \max_{1 \le i \le n} |A_i| \cdot b_1 \ .$$

Proof. We set $A_0 = b_0 = b_{n+1} := 0$. Then

$$\begin{aligned}\left|\sum_{k=1}^{n} a_k b_k\right| &= \left|\sum_{k=1}^{n}(A_k - A_{k-1})b_k = \right|\sum_{k=0}^{n} A_k b_k - \sum_{k=0}^{n} A_k b_{k+1}\left|\right. \\ &= \left|\sum_{k=1}^{n} A_k(b_k - b_{k+1})\right| \le \sum_{k=1}^{n} |A_k|(b_k - b_{k+1}) \\ &\le \max_{1\le k\le n} |A_k| \sum_{k=1}^{n}(b_k - b_{k+1}) = \max_{1\le k\le n} |A_k| \cdot b_1 \ .\end{aligned}$$

□

Proposition 2.19 (Dirichlet's convergence criterion) *Let* $(a_n) \subset \mathbb{C}$ *and* $(b_n) \subset \mathbb{R}$ *be two sequences such that* (i) (a_n) *is strongly bounded, in the sense that for some* $c > 0$ *one has* $|\sum_{n\in I} a_n| \le c$ *for every finite interval* $I \subset \mathbb{N}$, (ii) $b_1 \ge b_2 \ge \cdots \ge 0$ *and* (iii) $\lim b_n = 0$. *Then the series*

$$\sum_{n=1}^{\infty} a_n b_n$$

converges, possibly non-absolutely.

Proof. Let $c > 0$ be such that $|\sum_{n\in I} a_n| \le c$ holds for every finite interval $I \subset \mathbb{N}$. For a given $\varepsilon > 0$ we take a large $n_0 \in \mathbb{N}$ such that $0 \le c \cdot b_{n_0} < \varepsilon$. Then for every $m > n \ge n_0$,

$$\left|\sum_{k=1}^{m} a_k b_k - \sum_{k=1}^{n} a_k b_k\right| = \left|\sum_{k=n+1}^{m} a_k b_k\right| \overset{\text{Abel's ineq.}}{\le} c b_{n+1} < c b_{n_0} < \varepsilon \ .$$

Thus the sequence of partial sums of $\sum_{n\ge1} a_n b_n$ is Cauchy and this series converges. □

Corollary 2.20 (on $L(f)$) *If* $f\colon \mathbb{N} \to \mathbb{C}$ *is strongly bounded then the series*

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

converges, possibly non-absolutely.

Proof. Use the previous criterion with $a_n := f(n)$ and $b_n := 1/n$. □

We come to the second heart of the proof of Dirichlet's theorem. Namely, L-series of non-principal Dirichlet characters have nonzero sums.

Theorem 2.21 **($L(\chi) \neq 0$)** *For every $m \in \mathbb{N}$ and every Dirichlet character $\chi \in D(m)$, $\chi \neq \chi_m$, the sum*

$$L(\chi) = \sum_{n=1}^{\infty} \frac{\chi(n)}{n} \neq 0 \ .$$

To prove this theorem we need several auxiliary results.

Proposition 2.22 (AGM inequality) *If $n \in \mathbb{N}$ and $a_1, a_2, \dots, a_n$ are non-negative real numbers then*

$$\frac{a_1 + a_2 + \cdots + a_n}{n} \geq (a_1 a_2 \dots a_n)^{1/n} \ .$$

Proof. Derivatives show that $\mathrm{e}^{x-1} \geq x$ for every $x \geq 0$. Let $a := (a_1 + a_2 + \cdots + a_n)/n > 0$ (for $a = 0$ all $a_i = 0$ and the AGM inequality holds trivially) and let $x_i := a_i/a$. We multiply the n inequalities $\mathrm{e}^{x_i - 1} \geq x_i$ and get that

$$1 = \mathrm{e}^{(a_1 + a_2 + \cdots + a_n)/a - n} \geq x_1 x_2 \dots x_n = \frac{a_1 a_2 \dots a_n}{a^n} \ ,$$

which after rearrangement gives the AGM inequality. □

Proposition 2.23 (a non-increasing sequence) *If $n \in \mathbb{N}$, $t \in [0, 1)$ and*

$$b_n := \frac{1}{n(1-t)} - \frac{t^n}{1 - t^n}$$

then

$$1 \geq b_1 \geq b_2 \geq \cdots \geq 0 \ .$$

Proof. Since $b_n \to 0$, it suffices to prove that $b_n - b_{n+1} \geq 0$ for every n. Now $(1-t)(b_n - b_{n+1})$ equals

$$A := \frac{1}{n(n+1)} - \frac{t^n}{(1 + t + t^2 + \cdots + t^{n-1})(1 + t + t^2 + \cdots + t^n)} \ .$$

Note that $t^n = t^{(n-1)/2} \cdot t^{n/2} \cdot t^{1/2}$ and that $0 \leq t^{1/2} < 1$. By the AGM inequality,

$$\frac{1 + t + t^2 + \cdots + t^{n-1}}{n} \geq t^{(n-1)/2} \ \text{ and } \ \frac{1 + t + t^2 + \cdots + t^n}{n+1} \geq t^{n/2} \ .$$

It follows that $A \geq 0$, as we need. □

We prove Theorem 2.21 first only for real Dirichlet characters, but in a more general setting. If $f \colon \mathbb{N} \to \mathbb{C}$ is completely multiplicative then $f(1)^2 = f(1^2) = f(1)$ and $f(1) = 0$ or 1, and in the former case $f(n) = f(n)f(1) = 0$ for every $n \in \mathbb{N}$. We assume that this degenerate case does not occur.

Theorem 2.24 $(L(f)\neq 0)$ *For any function $f\colon \mathbb{N}\to\mathbb{R}$ that is completely multiplicative with $f(1)=1$ and strongly bounded, the sum*

$$L(f)=\sum_{n=1}^{\infty}\frac{f(n)}{n}\neq 0\ .$$

Proof. We assume for contradiction that $\sum_{n\ge 1} f(n)/n=0$ and for $t\in[0,1)$ take the non-negative and non-increasing sequence $(b_n)=(b_n(t))$ defined in the previous proposition. Then for $t\in[0,1)$,

$$\begin{aligned}
O(1) &\overset{f\text{ is s. b., Props. 2.18, 2.19, 2.23}}{=} & -\sum_{n=1}^{\infty} f(n)b_n\\
&\overset{\text{def. of } b_n}{=} & \sum_{n=1}^{\infty}\frac{f(n)t^n}{1-t^n}-\frac{1}{1-t}\sum_{n=1}^{\infty}\frac{f(n)}{n}\\
&\overset{L(f)=0,\text{ sum of GS}}{=} & \sum_{n=1}^{\infty} f(n)\sum_{k=1}^{\infty} t^{kn}\\
&\overset{\text{swapping }\sum\text{s by AC}}{=} & \sum_{m=1}^{\infty} t^m\sum_{n\,|\,m} f(n)\\
&\overset{f\text{ is c. m., } f(1)=1,\ m=p_1^{a_1}\dots p_r^{a_r}}{=} & \sum_{m=1}^{\infty} t^m\prod_{i=1}^{r}\sum_{j=0}^{a_i} f(p_i)^j\\
&\ge & \frac{3}{4}\sum_{m=p^2} t^m\to+\infty\ \text{ for }\ t\to 1^-\ ,
\end{aligned}$$

which is a contradiction. We justify the last inequality. First note that $|f(p)|\le 1$ for every prime p because f is completely multiplicative and (strongly) bounded. If $f(p_i)\neq 1$, resp. $=1$, then $\sum_{j=0}^{a_i} f(p_i)^j=(1-f(p_i)^{a_i+1})/(1-f(p_i))$, resp. $=a_i+1$. Always $\sum_{j=0}^{a_i} f(p_i)^j\ge 0$ and $\prod_{i=1}^{r}\dots\ge 0$. If $m=p_1^2$ then $\prod_{i=1}^{r}\dots=1+f(p_1)+f(p_1)^2=(\frac{1}{2}+f(p_1))^2+\frac{3}{4}\ge\frac{3}{4}$. Thus we get the last displayed inequality. □

It is time to introduce the Möbius function $\mu\colon\mathbb{N}\to\{-1,0,1\}$. It has the values $\mu(1)=1$,

$$\mu(p_1p_2\dots p_k)=(-1)^k$$

(the primes p_i are distinct), and $\mu(n)=0$ if n is not a product of distinct primes.

Proposition 2.25 (on μ) *The Möbius function has the following properties.*

1. *For every $n\in\mathbb{N}\setminus\{1\}$,*

$$\sum_{d\,|\,n}\mu(d)=0\ .$$

For $n=1$ this sum equals 1.

2. *If two functions $f, g\colon \mathbb{N} \to \mathbb{C}$ are related by $f(n) = \sum_{d \mid n} g(d)$ then for every $n \in \mathbb{N}$,*

$$g(n) = \sum_{d \mid n} \mu(n/d) f(d) = \sum_{kl=n} \mu(k) f(l) \ .$$

3. *For every $n \in \mathbb{N} \setminus \{1\}$ and real $x > 0$,*

$$\sum_{d \mid n} \mu(d) \log(x/d) = \Lambda(n) \ .$$

For $n = 1$ this sum equals $\log x$.

Proof. 1. For $n = 1$ the sum equals 1 trivially. Let $n = p_1^{a_1} p_2^{a_2} \dots p_k^{a_k} > 1$, so $k \ge 1$. Then

$$\sum_{d \mid n} \mu(d) = \sum_{X \subset [k]} (-1)^{|X|} = \sum_{i=0}^{k} \binom{k}{i} (-1)^i = (1-1)^k = 0 \ .$$

2. Let f and g be as stated and let $n \in \mathbb{N}$. Then

$$\sum_{d \mid n} \mu(n/d) f(d) = \sum_{ab=n} \mu(a) f(b) = \sum_{acd=n} \mu(a) g(c) = \sum_{c \mid n} g(c) \sum_{a \mid n/c} \mu(a) = g(n) \ .$$

In the second equality we used the assumed relation between f and g, and in the last equality we used part 1.

3. For $n = 1$ it holds trivially. Let $n > 1$. Then by part 1 the displayed sum equals

$$(\log x) \sum_{d \mid n} \mu(d) - \sum_{d \mid n} \mu(d) \log d = - \sum_{d \mid n} \mu(d) \log d \ .$$

But inverting by part 2 the relation $\log n = \sum_{d \mid n} \Lambda(d)$ of Proposition 2.13 we get, using again part 1, that also

$$\Lambda(n) = \sum_{d \mid n} \mu(d) \log(n/d) = (\log n) \sum_{d \mid n} \mu(d) - \sum_{d \mid n} \mu(d) \log d = - \sum_{d \mid n} \mu(d) \log d \ .$$

□

Proposition 2.26 (if $L(\chi) \neq 0$ 1) *Let $m \in \mathbb{N}$ and $\chi \in D(m)$ with $\chi \neq \chi_m$. If $L(\chi) \neq 0$ then for all $x > 1$,*

$$\sum_{n \le x} \frac{\chi(n) \Lambda(n)}{n} = O(1) \ .$$

Proof. Let m and χ be as stated and let $L(\chi) \neq 0$. Then for $x > 1$,

$$\begin{aligned}
O(1) &\overset{\chi \text{ is s. b., Prop. 2.18}}{=} \sum_{n\le x}\frac{\chi(n)\log n}{n}\\
&\overset{\text{Prop. 2.13}}{=} \sum_{n\le x}\frac{\chi(n)}{n}\sum_{d\,|\,n}\Lambda(d)\\
&\overset{\chi \text{ is c. m., swapping } \sum \text{s}}{=} \sum_{d\le x}\frac{\chi(d)\Lambda(d)}{d}\sum_{e\le x/d}\frac{\chi(e)}{e}\\
&\overset{\chi \text{ is s. b., Prop. 2.18}}{=} \sum_{d\le x}\frac{\chi(d)\Lambda(d)}{d}\big(L(\chi)-O(d/x)\big)\\
&\overset{\chi \text{ is bounded}}{=} L(\chi)\sum_{d\le x}\frac{\chi(d)\Lambda(d)}{d} - O(1/x)\sum_{d\le x}\Lambda(d)\\
&\overset{\text{Prop. 2.15}}{=} L(\chi)\sum_{d\le x}\frac{\chi(d)\Lambda(d)}{d} + O(1)\ .
\end{aligned}$$

Dividing by the sum $L(\chi)$ we get the stated result. □

Proposition 2.27 (if $L(\chi) \neq 0$ 2) *Let $m \in \mathbb{N}$ and $\chi \in D(m)$ with $\chi \neq \chi_m$. If $L(\chi) \neq 0$ then for all $x > 1$,*

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

Proof. Let m and χ be as stated and let $L(\chi) \neq 0$. Then for $x > 1$,

$$\begin{aligned}
O(1) &\overset{\text{Prop. 2.26}}{=} \sum_{d\le x}\frac{\chi(d)\Lambda(d)}{d}\\
&\overset{\text{def. of } \Lambda}{=} \sum_{p\le x}\frac{\chi(p)\log p}{p} + \sum_{p^k\le x,\, k\ge 2}\frac{\chi(p^k)\log p}{p^k}\\
&\overset{\chi \text{ is b., } \sum_{n,\,k\ge 2}(\log n)/n^k \text{ conv.}}{=} \sum_{p\le x}\frac{\chi(p)\log p}{p} + O(1)
\end{aligned}$$

and $\sum_{p\le x}(\chi(p)\log p)/p = O(1)$ as claimed. □

Proposition 2.28 (for χ_m) *Let $m \in \mathbb{N}$. Then for $x > 1$,*

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

Proof. By Proposition 2.17, for $x > 1$ the sum equals

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

□

For any function $f\colon \mathbb{N} \to \mathbb{C}$ and any real $x > 1$ we define the finite sum

$$S(f,\, x) := \sum_{n \le x} \frac{f(n)\Lambda(n)}{n}\ .$$

Proposition 2.29 (on $S(\chi, x)$) *For every $m \in \mathbb{N}$ and every $\chi \in D(m)$ the following hold.*

1. $\chi = \chi_m \Rightarrow S(\chi, x) = \log x + O(1)$ *for all real* $x > 1$.
2. $\chi \ne \chi_m$ *and* $L(\chi) \ne 0 \Rightarrow S(\chi, x) = O(1)$ *for all real* $x > 1$.
3. $\chi \ne \chi_m$ *and* $L(\chi) = 0 \Rightarrow S(\chi, x) = -\log x + O(1)$ *for all real* $x > 1$.

Proof. 1. By Proposition 2.28, for $x > 1$ the sum $S(\chi_m, x)$ equals

$$\sum_{p \le x} \frac{\chi_m(p) \log p}{p} + \sum_{p^k \le x,\, k \ge 2} \frac{\chi_m(p^k) \log p}{p^k} = \log x + O(1)$$

because the series $\sum_{n,k \ge 2} (\log n)/n^k$ (absolutely) converges.

2. This was proved in Proposition 2.26.

3. We suppose that $\chi \in D(m)$ is not χ_m and that $L(\chi) = 0$. Then for every $x > 1$,

$$\begin{aligned}
S(\chi,\, x) &\overset{\text{part 3 of Prop. 2.25}}{=} -\log x + \sum_{n \le x} \frac{\chi(n)}{n} \sum_{d \,|\, n} \mu(d) \log(x/d) \\
&\overset{\chi \text{ is c. m., swapping } \sum\text{s}}{=} -\log x + \sum_{d \le x} \frac{\chi(d)\mu(d) \log(x/d)}{d} \sum_{e \le x/d} \frac{\chi(e)}{e} \\
&\overset{\text{Prop. 2.18, } \chi \text{ is s. b.}}{=} -\log x + \sum_{d \le x} (\ldots)\big(L(\chi) - O(d/x)\big) \\
&\overset{|\chi(d)\mu(d)| \le 1,\ L(\chi) = 0}{=} -\log x + O\bigg(\frac{1}{x} \cdot \sum_{d \le x} \log(x/d)\bigg) = -\log x + O(1)
\end{aligned}$$

as $\sum_{d \le x} \log(x/d) = \lfloor x \rfloor \log x - \sum_{d \le x} \log d = O(x)$ by Proposition 2.16. □

Proof of Theorem 2.21. If $\chi \in D(m) \setminus \{\chi_m\}$ is real, $L(\chi) \ne 0$ was proven in Theorem 2.24. We assume for contradiction that $\psi \in D(m) \setminus \{\chi_m\}$ is a non-real Dirichlet character with $L(\psi) = 0$. It follows from part 1 of Theorem 2.5 and from the definition of $L(f)$ that $\overline{\psi} \in D(m)\setminus\{\chi_m\}$ and that also $L(\overline{\psi}) = 0$. Since $\psi \neq \overline{\psi}$,

$$N := \#\{\chi \in D(m) \setminus \{\chi_m\} \mid L(\chi) = 0\} \geq 2 \;.$$

But for $x > 1$,

$$\begin{aligned}
0 &\overset{\Lambda \text{ is nonneg.}}{\leq} \sum_{\substack{n\leq x \\ n\equiv 1 \pmod m}} \frac{\Lambda(n)}{n} = \sum_{n\leq x} \frac{\mathbb{I}_{1,\,m}(n)\Lambda(n)}{n} \\
&\overset{\text{part 2 of Thm. 2.5}}{=} \sum_{n\leq x} \frac{1}{\varphi(m)} \sum_{\chi\in D(m)} \frac{\chi(n)\Lambda(n)}{n} \\
&\overset{\text{Prop. 2.29, swapping }\sum\text{s}}{=} \sum_{\chi\in D(m)} \frac{S(\chi,\, x)}{\varphi(m)} = \frac{(1-N)\log x}{\varphi(m)} + O(1) \\
&\overset{\text{since } N\geq 2}{<} -\frac{\log x}{\varphi(m)} + O(1)
\end{aligned}$$

and $0 < -(\log x)/\varphi(m) + O(1)$ is contradictory for $x \to +\infty$. □

Now we can prove Dirichlet's theorem on prime numbers in AP in the following asymptotic form.

Theorem 2.30 (stronger Dirichlet's theorem) *If $a, m \in \mathbb{N}$ are coprime then for all real $x > 1$,*

$$\sum_{\substack{p\leq x \\ p\equiv a \pmod m}} \frac{\log p}{p} = \frac{\log x}{\varphi(m)} + O(1) \;.$$

Proof. Let $a, m \in \mathbb{N}$ be coprime numbers. Then for all $x > 1$,

$$\begin{aligned}
\sum_{\substack{p\leq x \\ p\equiv a \pmod m}} \frac{\log p}{p} &\overset{\mathbb{I}_{a,\,m}}{=} \sum_{p\leq x} \frac{\mathbb{I}_{a,\,m}(p)\log p}{p} \\
&\overset{\text{part 2 of Thm. 2.5}}{=} \sum_{p\leq x}\sum_{\chi\in D(m)} c_{a,\,\chi}\cdot\chi(p)\cdot\frac{\log p}{p} \\
&\overset{\text{swapping }\sum\text{s}}{=} \sum_{\chi\in D(m)} c_{a,\,\chi} \sum_{p\leq x} \frac{\chi(p)\log p}{p} \\
&\overset{\text{part 2 of Thm. 2.5}}{=} \frac{1}{\varphi(m)} \sum_{p\leq x} \frac{\chi_m(p)\log p}{p} + \\
&\qquad + \sum_{\substack{\chi\in D(m) \\ \chi\neq\chi_m}} c_{a,\,\chi} \sum_{p\leq x} \frac{\chi(p)\log p}{p} \\
&\overset{\text{Thm. 2.21, Prop. 2.27 and 2.28}}{=} \frac{\log x}{\varphi(m)} + O(1) \;.
\end{aligned}$$

□

2.3 Remarks on the previous proof

Chapter 3

An elementary proof of PNT

Chapter 4

Stirling's asymptotic formula by Newton's integral

Chapter 5

Counting graphs by multivariate Cauchy's formula

Chapter 6

The Jacobi identity

Chapter 7

“The function $\zeta(s)$ and the Dirichlet series related to it”

We review Chapter I with this heading in the classical tract [11] of E. C. Titchmarsh (and D. R. Heath-Brown) on the zeta function $\zeta(s)$, and then in Section 7.2 comment on used analytic tools. We keep the original numbering of formulas in [11] and the formulas themselves, slightly adapt notation, and paraphrase and often directly quote Titchmarsh’s words.

7.1 Chapter I in [11]

1.1. Definition of $\zeta(s)$. The first two formulas are

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s} \tag{1.1.1}$$

and

$$\zeta(s) = \prod_p \left(1 - \frac{1}{p^s}\right)^{-1}. \tag{1.1.2}$$

Here $n \in \mathbb{N} = \{1, 2, \dots\}$ and p runs through primes. Either can be taken as a definition of $\zeta(s)$; $s \in \mathbb{C}$ with $s = \sigma + it$. The Dirichlet series (1.1.1) converges for $\sigma > 1$ and uniformly converges for $\sigma > 1 + \delta$, $\delta > 0$ (**Comment 7.2.1**). “It therefore defines an analytic function $\zeta(s)$, regular for $\sigma > 1$.” (**Comment 7.2.2**)

The infinite product (1.1.2) absolutely converges for $\sigma > 1$, for so does

$$\sum_p \left|\frac{1}{p^s}\right| = \sum_p \frac{1}{p^\sigma}$$

(**Comment 7.2.3**). We see that

$$\zeta(s) = \prod_p \left(1 + \frac{1}{p^s} + \frac{1}{p^{2s}} + \dots\right) .$$

Unique prime factorization and formal multiplication yield (1.1.1). To prove rigorously that (1.1.2) holds for $\sigma > 1$, we take only finitely many factors. Multiplication of absolutely convergent series (**Comment 7.2.4**) leads to

$$\prod_{p \le P} \left(1 + \frac{1}{p^s} + \frac{1}{p^{2s}} + \dots\right) = 1 + \frac{1}{n_1^s} + \frac{1}{n_2^s} + \dots$$

where each n_i has all prime factors $\le P$. Since every $n \le P$ has this form, if $\zeta(s)$ is defined by (1.1.1) then

$$\begin{aligned}\left|\zeta(s) - \prod_{p \le P}\left(1 - \frac{1}{p^s}\right)\right| &= \left|\zeta(s) - 1 - \frac{1}{n_1^s} - \frac{1}{n_2^s} - \dots\right| \\ &\le \frac{1}{(P+1)^s} + \frac{1}{(P+2)^s} + \dots .\end{aligned}$$

“This tends to 0 as $P \to \infty$, if $\sigma > 1$; and (1.1.2) follows.” This fundamental identity is due to Euler.

“Since a convergent infinite product of non-zero factors is not zero, we deduce that $\zeta(s)$ *has no zeros for* $\sigma > 1$ (**Comment 7.2.5**). This may be proved directly as follows.” For $\sigma > 1$ one has that

$$\left(1 - \frac{1}{2^s}\right)\left(1 - \frac{1}{3^s}\right) \dots \left(1 - \frac{1}{P^s}\right)\zeta(s) = 1 + \frac{1}{m_1^s} + \frac{1}{m_2^s} + \dots$$

where each m_i has all prime factors $> P$. Hence

$$\left|\left(1 - \frac{1}{2^s}\right)\left(1 - \frac{1}{3^s}\right) \dots \left(1 - \frac{1}{P^s}\right)\zeta(s)\right| \ge 1 - \frac{1}{(P+1)^s} - \frac{1}{(P+2)^s} - \dots > 0$$

if P is large enough. So $|\zeta(s)| > 0$.

Recall that $\pi(x)$ denotes the number of primes $\le x$. “We can transform (1.1.2) into a relation between $\zeta(s)$ and $\pi(x)$; for if $\sigma > 1$,

$$\begin{aligned}\log \zeta(s) &= -\sum_p \log\left(1 - \frac{1}{p^s}\right) = -\sum_{n=2}^{\infty} \{\pi(n) - \pi(n-1)\} \log\left(1 - \frac{1}{n^s}\right) \\ &= -\sum_{n=2}^{\infty} \pi(n) \left\{ \log\left(1 - \frac{1}{n^s}\right) - \log\left(1 - \frac{1}{(n+1)^s}\right)\right\} \\ &= \sum_{n=2}^{\infty} \int_n^{n+1} \frac{s}{x(x^s - 1)}\, dx = s \int_s^{\infty} \frac{\pi(x)}{x(x^s - 1)}\, dx . \qquad (1.1.3)\end{aligned}$$

The rearrangement of the series is justified (**Comment 7.2.6**) since $\pi(n) \leq n$ and

$$\log(1 - n^{-s}) = O(n^{-\sigma}) \text{ .''}$$

Multiplying in

$$\frac{1}{\zeta(s)} = \prod_p \left(1 - \frac{1}{p^s}\right)$$

we get that

$$\frac{1}{\zeta(s)} = \sum_{n=1}^{\infty} \frac{\mu(n)}{n^s} \quad (\sigma > 1) \tag{1.1.4}$$

where μ is the Möbius function. "'The process is easily justified as in the case of $\zeta(s)$." (**Comment 7.2.7**)

We have that

$$\textstyle\sum_{d\,|\,q} \mu(d) = 1 \ (q = 1), \quad 0 \ (q > 1) \ . \tag{1.1.5}$$

This follows from the identity

$$1 = \sum_{m=1}^{\infty} \frac{1}{m^s} \sum_{n=1}^{\infty} \frac{\mu(n)}{n^s} = \sum_{q=1}^{\infty} \frac{1}{q^s} \sum_{d\,|\,q} \mu(d) \ .$$

It also gives the Möbius inversion

$$g(q) = \textstyle\sum_{d\,|\,q} f(d) \tag{1.1.6}$$

and

$$f(q) = \sum_{d\,|\,q} \mu\left(\frac{q}{d}\right) g(d) \tag{1.1.7}$$

where $g, f \colon \mathbb{N} \to \mathbb{C}$.

7.2 Analysis in Chapter I

Comment 7.2.1 (convergence of the series $\zeta(s)$)

The domain of convergence of $\zeta(s)$, $s = \sigma + it$, follows from the computation

$$|n^{-s}| = \big|\exp(-\sigma \log n)\big| \cdot \big|\exp\big(i(-t)\log n\big)\big| = n^{-\sigma} \cdot 1 = n^{-\sigma}$$

and from the domain of convergence of $\zeta(s)$ for real s. It is in fact absolute convergence.

Comment 7.2.2 ($\zeta(s)$ is an analytic function)

Comment 7.2.3 (convergence of infinite products)

Comment 7.2.4 (multiplying AC series)

Comment 7.2.5 ($\zeta(s) \neq 0$ for $\sigma > 1$)

Comment 7.2.6 (rearrangement of a series)

Comment 7.2.7 (Dirichlet series for $\mu(n)$)

Chapter 8

“The analytic character of ζ(s), and the functional equation”

We treat Chapter II in [11] with this heading in the same way as Chapter I, and again in Section 8.2 comment on used analytic tools.

8.1 Chapter II in [11]

2.1. Analytic continuation and the functional equation, first method. “We have next to inquire whether the analytic function $\zeta(s)$ can be continued beyond this region.” (i.e. beyond $\sigma > 1$). “The result is

THEOREM 2.1. *The function $\zeta(s)$ is regular for all values of s except $s = 1$, where there is a simple pole with residue 1. It satisfies the functional equation*

$$\zeta(s) = 2^s \pi^{s-1} \sin \tfrac{1}{2} s\pi \Gamma(1-s)\zeta(1-s).\text{”} \tag{2.1.1}$$

(**Comment 8.2.1**) “We shall first give a proof depending on the following summation formula.

Let $\phi(x)$ be any function with a continuous derivative in the interval $[a, b]$. Then, if $[x]$ denotes the greatest integer not exceeding x,

$$\sum_{a<n\leq b} \phi(n) = \int_a^b \phi(x)\,dx + \int_a^b (x - [x] - \tfrac{1}{2})\phi'(x)\,dx + \\ + (a - [a] - \tfrac{1}{2})\phi(a) - (b - [b] - \tfrac{1}{2})\phi(b). \tag{2.1.2}$$

(**Comment 8.2.2**) Since the formula is plainly additive with respect to the interval $(a,b]$ it suffices to suppose that $n \le a < b \le n+1$. One then has

$$\int_a^b (x-n-\tfrac{1}{2})\phi'(x)\,dx = (b-n-\tfrac{1}{2})\phi(b) - (a-n-\tfrac{1}{2})\phi(a) - \int_a^b \phi(x)\,dx,$$

on integrating by parts (**Comment 8.2.3**). Thus the right hand side of (2.1.2) reduces to $(b-[n])\phi(b)$. This vanishes unless $b = n+1$, in which case it is $\phi(n+1)$, as required."

We apply the summation formula to $\phi(n) = n^{-s}$, $s \ne 1$, and $a, b \in \mathbb{N}$ and get that

$$\sum_{n=a+1}^{b} \frac{1}{n^s} = \frac{b^{1-s} - a^{1-s}}{1-s} - s\int_a^b \frac{x-[x]-\frac{1}{2}}{x^{s+1}}\,dx + \tfrac{1}{2}(b^{-s} - a^{-s}). \qquad (2.1.3)$$

We assume that $\sigma > 1$, $a = 1$ and $b \to \infty$. We add 1 to each side and get that

$$\zeta(s) = s\int_1^\infty \frac{[x]-x+\frac{1}{2}}{x^{s+1}}\,dx + \frac{1}{s-1} + \frac{1}{2}. \qquad (2.1.4)$$

"Since $[x]-x+\frac{1}{2}$ is bounded, this integral is convergent for $\sigma > 0$, and uniformly convergent in any finite region to the right of $\sigma = 0$. It therefore defines an analytic function of s, regular for $\sigma > 0$. The right-hand side therefore provides the analytic continuation of $\zeta(s)$ up to $\sigma = 0$, and there is clearly a simple pole at $s = 1$ with residue 1." (**Comment 8.2.4**)

If $0 < \sigma < 1$ then

$$\int_0^1 \frac{[x]-x}{x^{s+1}}\,dx = -\int_0^1 x^{-s}\,dx = \frac{1}{s-1}, \quad \frac{s}{2}\int_1^\infty \frac{dx}{x^{s+1}} = \frac{1}{2},$$

and (2.1.4) turns in

$$\zeta(s) = s\int_0^\infty \frac{[x]-x}{x^{s+1}}\,dx \quad (0 < \sigma < 1). \qquad (2.1.5)$$

(**Comment 8.2.5**) "Actually (2.1.4) gives the analytic continuation of $\zeta(s)$ for $\sigma > -1$; for if

$$f(x) = [x] - x + \tfrac{1}{2}, \quad f_1(x) = \int_1^x f(y)\,dy,$$

then $f_1(x)$ is also bounded, since, as is easily seen,

$$\int_k^{k+1} f(y)\,dy = 0$$

for any integer k. Hence

$$\int_{x_1}^{x_2} \frac{f(x)}{x^{s+1}}\,dx = \left[\frac{f_1(x)}{x^{s+1}}\right]_{x_1}^{x_2} + (s+1)\int_{x_1}^{x_2} \frac{f_1(x)}{x^{s+1}},$$

(**Comment 8.2.6**) which tends to 0 as $x_1 \to \infty$, $x_2 \to \infty$ if $\sigma > -1$. Hence the integral in (2.1.4) is convergent for $\sigma > -1$. Also it is easily verified that

$$s\int_0^1 \frac{[x]-x+\frac{1}{2}}{x^{s+1}}\,dx = \frac{1}{s-1}+\frac{1}{2} \quad (\sigma < 0).$$

Hence

$$\zeta(s) = s\int_0^\infty \frac{[x]-x+\frac{1}{2}}{x^{s+1}}\,dx \quad (-1<\sigma<0).\text{”} \tag{2.1.6}$$

(**Comment 8.2.7**) We use the expansion into Fourier series

$$[x]-x+\tfrac{1}{2} = \sum_{n=1}^\infty \frac{\sin\, 2n\pi x}{n\pi} \tag{2.1.7}$$

$(x \notin \mathbb{Z})$ (**Comment 8.2.8**). We substitute it in (2.1.6), integrate term by term and get that

$$\begin{aligned}\zeta(s) &= \frac{s}{\pi}\sum_{n=1}^\infty \frac{1}{n}\int_0^\infty \frac{\sin 2n\pi x}{x^{s+1}}\,dx\\ &= \frac{s}{\pi}\sum_{n=1}^\infty \frac{(2n\pi)^s}{n}\int_0^\infty \frac{\sin y}{y^{s+1}}\,dy\\ &= \frac{s}{\pi}(2\pi)^s\{-\Gamma(-s)\}\sin\tfrac{1}{2}s\pi\zeta(1-s),\end{aligned}$$

which is (2.1.1) (**Comment 8.2.9**). It primarily holds for $-1<\sigma<0$. “Here, however, the right-hand side is analytic for all values of s such that $\sigma<0$. It therefore provides the analytic continuation of $\zeta(s)$ over the remainder of the plane, and there are no singularities other than the pole already encountered at $s=1$. (**Comment 8.2.10**)

We have still to justify the term-by-term integration. Since the series (2.1.7) is boundedly convergent, term-by-term integration over any finite range is permissible. It is therefore sufficient to prove that

$$\lim_{\lambda\to\infty}\sum_{n=1}^\infty \frac{1}{n}\int_\lambda^\infty \frac{\sin 2n\pi x}{x^{s+1}}\,dx = 0 \quad (-1<\sigma<0).$$

Now

$$\begin{aligned}\int_\lambda^\infty \frac{\sin 2n\pi x}{x^{s+1}}\,dx &= \left[-\frac{\cos 2n\pi x}{2n\pi x^{s+1}}\right]_\lambda^\infty - \frac{s+1}{2n\pi}\int_\lambda^\infty \frac{\cos 2n\pi x}{x^{s+2}}\,dx\\ &= O\!\left(\frac{1}{n\lambda^{\sigma+1}}\right)+O\!\left(\frac{1}{n}\int_\lambda^\infty \frac{dx}{x^{\sigma+2}}\right) = O\!\left(\frac{1}{n\lambda^{\sigma+1}}\right),\end{aligned}$$

and the desired result clearly follows.” (**Comment 8.2.11**)

Alternative forms of (2.1.1).

8.2 Analysis in Chapter II

Comment 8.2.1 (the functional equation (2.1.1))

Comment 8.2.2 (summation formula (2.1.2))

Comment 8.2.3 (integration by parts 1)

Comment 8.2.4 (extending $\zeta(s)$ by (2.1.4))

Comment 8.2.5 (derivation of (2.1.5))

Comment 8.2.6 (integration by parts 2)

Comment 8.2.7 (derivation of (2.1.6))

Comment 8.2.8 (Fourier series)

Comment 8.2.9 (where does $-\Gamma(-s)$ come from)

Comment 8.2.10 (extending $\zeta(s)$ to $\mathbb{C} \setminus \{1\}$)

Comment 8.2.11 (justifying the term-by-term $\int$)

Bibliography

[1] A. Baker, *Transcendental Number Theory*, Cambridge University Press, Cambridge, UK 1975

[2] R. P. Crease, *The Great Equations*, Robinson, London 2009

[3] G. Lejeune Dirichlet, Beweis des Satzes, dass jede unbegrenzte arithmetische Progression, deren erstes Glied und Differenz ganze Zahlen ohne gemeinschaftlichen Factor sind, unendlich viele Primzahlen enthält, *Abhandlungen der Königlich Preussischen Akademie der Wissenschaften von 1837*, 45–81

[4] P. Erdős, Über die Primzahlen gewisser arithmetischer Reihen, *Math. Z.* **39** (1935), 473–491

[5] D. Hilbert, Ueber die Transcendenz der Zahlen e und π, *Math. Annalen* **43** (1893), 216–219

[6] M. Klazar, *Analytic and Combinatorial Number Theory* II, KAM-DIMATIA Series 2010-969, 2010, 46 pp.

[7] P. Moree, Bertrand's postulate for primes in arithmetical progressions, *Computers Math. Aplic.* **26** (1993), 35–43

[8] B. Sambale, Dirichlets Primzahlsatz, 2022, 12 pp., `https://www.iazd.uni-hannover.de/fileadmin/iazd/sambale/teach/dirichlet.pdf`

[9] H. N. Shapiro, On primes in arithmetic progression. II, *Ann. of Math.* (2) **52** (1950), 231–243

[10] V. G. Sprindzhuk, *Klassicheskie Diofantovy Uravneniia ot Dvuch Neizvestnykh*, Nauka, Moskva 1982 (Classical Diophantine Equations with two Unknowns, in Russian)

[11] E. C. Tirchmarsh, *The Theory of the Riemann Zeta-function*, Clarendon Press, Oxford 1986 (2nd edition, revised by D. R. Heath-Brown)