DR. JOHANNES SPRANG

ALGEBRAIC NUMBER THEORY II

UNIVERSITÄT DUISBURG-ESSEN

fakultät für mathematik - universität duisburg-essen

https://homepages.uni-regensburg.de/ spj54141/SS21.html

First printing, March 2023

Contents

1 Introduction and Overview

There are some typical topics for a second course on Algebraic Number Theory, including *Tate's thesis* and *class field theory*. Both of these topics are closely related to the theory of Hecke L-functions; while *Tate's thesis* gives a purely Fourier-theoretic approach to the analytic continuation and the functional equation of such functions, *class field theory* allows one to relate Hecke characters to one-dimensional Galois representations. Nevertheless, both topics cover just a very particular aspect of L-functions.

Before we continue with mathematics, let me briefly recall the content of the Indian parable *blind men and an elephant*.

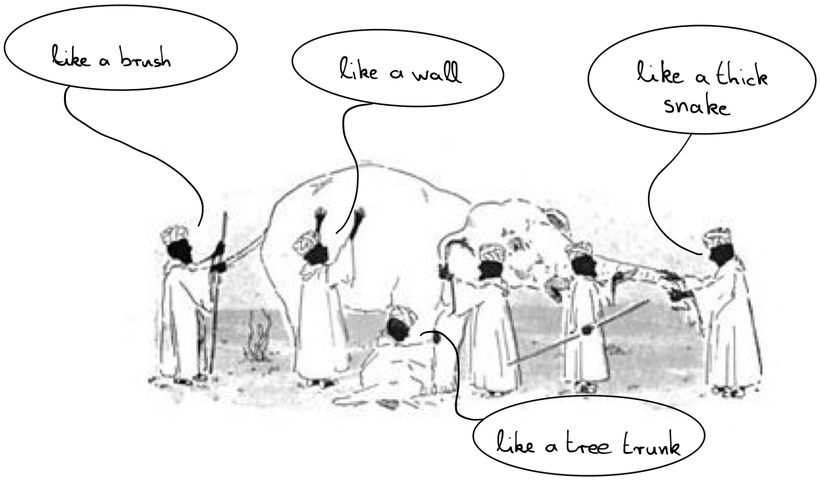


Figure 1.1: The blind men and the elephant; source: Wikipedia

> A group of blind men heard that a strange animal, called an elephant, had been brought to the town, but none of them were aware of its shape and form. Out of curiosity, they said: "We must inspect and know it by touch, of which we are capable". So, they sought it out, and when they found it they groped about it. The first person, whose hand landed on the trunk, said, "This being is like a thick snake". For another one whose hand reached its ear, it seemed like a kind of fan. As for another person, whose hand was upon its leg, said, the elephant is a pillar like a tree-trunk. The blind man who placed his hand upon its side said the

> elephant, "is a wall". Another who felt its tail, described it as a rope. The last felt its tusk, stating the elephant is that which is hard, smooth and like a spear.[1]

So, based on their experience of touching just a very particular part of the elephant's body they form their minds on what an elephant looks like. While everyone is right in a sense, no one will get a complete picture of the elephant.

Something similar happens, if you ask number theorists about L-functions. All of them will agree that they are of central importance for number theory. But if you ask ten number theorists about the most important aspect of L-functions, you will probably get ten different answers.

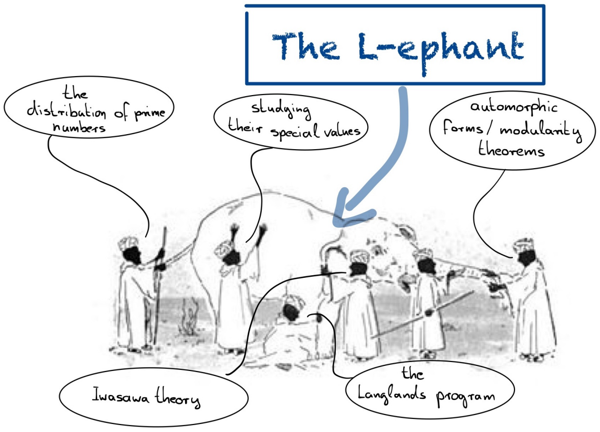

The L-ephant
the distribution of prime numbers
studying their special values
automorphic forms / modularity theorems
Iwasawa theory
the Langlands program

Figure 1.2: The L-ephant.

In this lecture, I will try to give you an idea about many different aspects of L-functions. Of course, I won't be able to go into as much depth as if I had focused on one single aspect. On the other hand, I think it makes more sense to get a vague picture of the entire L-ephant, than to understand its right leg in detail. Now, you might argue that certain aspects of the theory of L-functions are rather analytic, e.g., functional equations, distribution of primes, etc. But if there is one thing we can learn from the Indian parable, it is to be *open-minded* and try to understand different aspects of something we are interested in. That fits perfectly with the spirit of the University of Duisburg-Essen and its slogan ***Offen** im Denken*.

1.1 *Overview*

We will start with the Riemann zeta function and its basic properties. Afterwards, we will briefly say something about the importance of the Riemann zeta function for the distribution of primes. We will

[1] source: "Blind men and the elephant." Wikipedia: The Free Encyclopedia. Wikimedia Foundation, Inc. 31 March 2021.,

prove a weak form of the prime number theorem and briefly indicate the deeper relationship between the Riemann zeta function and prime numbers. Afterwards, we will turn our attention to the special values of the Riemann zeta function. First, we will prove Euler's formula which gives an explicit formula for the values of the Riemann zeta function at the positive even integers. Afterwards, we will briefly discuss odd zeta values, which are much more mysterious. Here, we will prove the irrationality of $\zeta(3)$, which is due to Apéry.

Afterwards, we will turn our attention to cyclotomic fields. In a first step, we will prove the Theorem of Kronecker-Weber which classifies all abelian extensions of $\mathbb{Q}$ and hence can be seen as a very explicit instance of class field theory for the base field $\mathbb{Q}$. We will introduce Dirichlet L-functions in this context. Then, we will discuss the analytic continuation and the functional equation of Dirichlet L-functions in a rather Fourier-theoretic way. This will make you familiar with the main ideas of *Tate's thesis* in a particular case. Then, we will express the Dedekind zeta function of abelian extensions of $\mathbb{Q}$ in terms of Dirichlet L-functions. If time permits, we will prove Kummer's criterion which gives a beautiful relation between special values of L-functions and class groups of cyclotomic fields.

At the end of the term, we will introduce Hecke characters for general number fields. We will give a small overview of how they relate to class field theory and Tate's thesis. Since we have already treated both topics in the more elementary case of the base field $\mathbb{Q}$, we will not go into details here.

2 The Riemann zeta function

In this chapter, we will define the Riemann zeta function. First, we will discuss its basic properties, afterwards we will discuss the relevance of the Riemann zeta function for the distribution of prime numbers. Finally, we will prove Euler's Theorem about the values of the Riemann zeta function at the even positive integers and discuss the irrationality of $\zeta(3)$.

2.1 Basic properties of the Riemann zeta function

The *Riemann zeta function* is defined for $s \in \mathbb{C}$ with $\operatorname{Re}(s) > 1$ by the formula

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

Here, we define $n^s := \exp(s \log n)$.

Lemma 2.1.1. *The series defining the Riemann zeta function converges absolutely and defines a holomorphic function in the half-plane* $\operatorname{Re}(s) > 1$.

Proof. For a real number $\delta > 0$ and $s \in \mathbb{C}$ with $\operatorname{Re}(s) \geq 1 + \delta$, the series

$$\sum_{n=1}^{\infty} \frac{1}{|n^s|}$$

admits the convergent majorant $\sum_{n=1}^{\infty} \frac{1}{n^{1+\delta}}$, i.e., the series defining the Riemann zeta function converges absolutely and uniformly [1] on the domain $\{s \in \mathbb{C} \mid \operatorname{Re}(s) \geq 1 + \delta\}$. Now, recall from complex analysis that a uniform limit of holomorphic functions is again holomorphic, see for example Theorem III.1.3 in Freitag–Busam[2]. Since $\delta > 0$ was arbitrary, the claim follows. □

Next, we will prove that the Riemann zeta function admits an *Euler product*. We will prove this in slightly greater generality. A function $f\colon \mathbb{N} \to \mathbb{C}$ is called *completely multiplicative* if it satisfies $f(1) = 1$ and $f(n \cdot m) = f(n) \cdot f(m)$ for all $n, m \in \mathbb{N}$.[3]

[1] Recall that a sequence of functions $(f_n(s))_n$ with $f_n\colon U \to \mathbb{C}$ converges uniformly to f if and only if for each $\epsilon > 0$ there exists an $N \in \mathbb{N}$ such that $|f(x) - f_n(x)| < \epsilon$ for all $x \in U$ and $n \geq N$. A series of complex functions converges uniformly if and only if its partial sums converge uniformly.

[2] Eberhard Freitag and Rolf Busam. *Funktionentheorie*. Springer-Verlag, Berlin, 1993. ISBN 3-540-50618-7

[3] In other words, it is a homomorphism of monoids $f\colon (\mathbb{N}, \cdot) \to (\mathbb{C}, \cdot)$

Lemma 2.1.2. *Let $f\colon \mathbb{N} \to \mathbb{C}$ be a completely multiplicative function for which the series $\sum_{n=1}^{\infty} f(n)$ converges absolutely. Then*

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

where p runs through the set of all primes.[4]

Proof. The assumptions imply that $|f(n)| < 1$ for $n \geq 2$. Indeed, if we had $|f(n)| \geq 1$ for some $n \geq 2$ then $|f(n^k)| \geq 1$ for every $k \geq 1$ contradicting the absolute convergence of the sum $\sum_{n=1}^{\infty} f(n)$. In particular, we have $|f(p)| < 1$ for every prime p and obtain the geometric series

$$(1 - f(p))^{-1} = \sum_{k=0}^{\infty} f(p)^k.$$

Using the complete multiplicativity of f and the unique factorization in $\mathbb{Z}$, we obtain for every positive integer N the identity

$$\prod_{p \leq N} (1 - f(p))^{-1} = \sum_{\substack{n = p_1^{\alpha_1} \ldots p_m^{\alpha_m} \\ p_i \leq N}} f(n) = \sum_{n=1}^{\infty} f(n) - \sum_{\substack{n \\ p \mid n \text{ for some } p > N}} f(n).$$

Now it follows

$$\left| \sum_{n=1}^{\infty} f(n) - \prod_{p \leq N} (1 - f(p))^{-1} \right| \leq \sum_{\substack{n \\ p \mid n \text{ for some } p > N}} |f(n)| \leq \sum_{n > N} |f(n)|.$$

The latter sum tends to zero as $N \to \infty$ by the absolute convergence of $\sum_{n=1}^{\infty} f(n)$ and the result follows. □

Corollary 2.1.3. *The Riemann zeta function admits the following product formula for $s \in \mathbb{C}$ with* $\mathrm{Re}(s) > 1$

$$\zeta(s) = \prod_{p} \frac{1}{1 - p^{-s}},$$

where p runs through the set of all prime numbers. This formula is called the Euler product *of the Riemann zeta function. For a prime p, the term $\frac{1}{1-p^{-s}}$ is called the* Euler factor *at p.*

Proof. We apply the previous lemma to the completely multiplicative function $n \mapsto \frac{1}{n^s}$. □

2.1.1 *The Gamma function*

As we will see, the Gamma function will play an important role for proving the functional equation of the Riemann zeta function. The Gamma function is defined for $z \in \mathbb{C}$ with $\mathrm{Re}(z) > 0$ by

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

[4] Recall that an infinite product $\prod_{i=1}^{\infty} a_i$ of complex numbers is said to converge if the sequence of partial products $P_n = \prod_{i=1}^{n} a_i$ has a non-zero limit.

To prove that the Gamma function is holomorphic, we recall the following result from complex analysis:

Lemma 2.1.4 (Leibniz rule)**.** *Let $U \subseteq \mathbb{C}$ open and $a, b \in \mathbb{R}$ with $a < b$. Suppose that $f\colon [a,b] \times U \to \mathbb{C}$ is a continuous function, which is holomorphic for every $t \in [a,b]$. Then the function*

$$z \mapsto \int_a^b f(z,t)dt$$

is holomorphic on U.

Proof. We refer to Lemma II.3.3 in Freitag–Busam[5]. □

Using the Leibniz rule it is not difficult to prove that the Gamma function is holomorphic.

Lemma 2.1.5. *The integral defining the Gamma function converges absolutely for* $\operatorname{Re}(z) > 0$*, where it represents a holomorphic function.*

Proof. We split the integral into two parts

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

and discuss both integrals separately. Note that we have the equality

$$|t^{z-1} e^{-t}| = t^{x-1} e^{-t}$$

for $x = \operatorname{Re}(z)$. For any real number $x_0 > 0$, we find a constant $C > 0$ such that $t^{x-1} \leq C e^{t/2}$ for all $0 < x \leq x_0$ and all $t \geq 1$. This estimate together with the existence of

$$\int_1^\infty e^{-t/2} dt$$

shows the absolute convergence of the second integral. For the absolute convergence of the first integral, we use the estimation $|t^{z-1} e^{-t}| < t^{x-1}$ for $t > 0$ and the existence of

$$\int_0^1 t^{x-1} dt \quad \text{for } x > 0.$$

The above estimates show that the functions

$$f_n(z) := \int_{1/n}^n e^{-t} t^z \frac{dt}{t}$$

converge uniformly to the Gamma function. Each of the functions f_n is holomorphic by the Leibniz rule. Therefore, the Gamma function is holomorphic as a uniform limit of holomorphic functions. □

Using integration by parts, it is not difficult to prove the following result.

[5] Eberhard Freitag and Rolf Busam. *Funktionentheorie*. Springer-Verlag, Berlin, 1993. ISBN 3-540-50618-7

Lemma 2.1.6. *The Gamma function satisfies for all $z \in \mathbb{C}$ with $\mathrm{Re}(z) > 0$ the functional equation*

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

In particular, we have for a positive integer n the formula $\Gamma(n) = (n-1)!$.

Proof. We will prove this in the exercises. □

We can use the functional equation to extend the Gamma function to $\mathbb{C} \setminus \mathbb{Z}_{\leq 0}$.

Lemma 2.1.7. *The Gamma function extends to a meromorphic function on all of $\mathbb{C}$ with simple poles at all non-positive integers and residue*[6]

$$\mathrm{res}_{z=-n}\Gamma(z) = \frac{(-1)^n}{n!} \quad \text{for } n \in \mathbb{Z}_{\geq 0}.$$

Proof. We will prove this result in the exercises. □

For later reference, we will prove the completion formula for the Gamma function.

Proposition 2.1.8 (Completion formula)**.** *For all $z \in \mathbb{C} \setminus \mathbb{Z}$ we have*

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

Proof. Both of the functions $\Gamma(z)\Gamma(1-z)$ and $\frac{\pi}{\sin \pi z}$ have only simple poles at the integers. Let us first compute their residues. For $n \in \mathbb{N}_0$, we have

$$\mathrm{Res}_{z=-n}\,\Gamma(z)\Gamma(1-z) = \Gamma(1+n)\,\mathrm{Res}_{z=-n}\,\Gamma(z) = (-1)^n$$

and similarly one proves

$$\mathrm{Res}_{z=n}\,\Gamma(z)\Gamma(1-z) = (-1)^n.$$

Thus, the formula $\mathrm{Res}_{z=n}\,\Gamma(z)\Gamma(1-z) = (-1)^n$ holds for all integers. The leading term of the Taylor expansion of $\sin \pi z$ at $z = n$ is $(-1)^n\pi$. We deduce the formula

$$\mathrm{Res}_{z=n}\,\frac{\pi}{\sin \pi z} = (-1)^n.$$

Since both functions $\Gamma(z)\Gamma(1-z)$ and $\frac{\pi}{\sin \pi z}$ have only simple poles with the same residues, we deduce that

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

extends to an entire function[7] on $\mathbb{C}$. Now, the strategy is to apply Liouville's Theorem[8] to deduce that h is constant. Therefore, we need to prove the boundedness of h. Let us first prove that the function $h(z)$ is 'periodic up to sign', i.e.,

$$h(z+1) = -h(z).$$

[6] Recall: The *residue* of a meromorphic function f at $z_0 \in \mathbb{C}$ is given by the term a_{-1} in its Laurent expansion

$$f(z) = \sum_{k \in \mathbb{Z}} a_k (z-z_0)^k.$$

If f has at most a simple pole in z_0, we can compute the residue as follows:

$$\mathrm{Res}_{z=z_0}\, f(z) = \lim_{z \to z_0} (z-z_0) f(z).$$

[7] An *entire* function is a function which is holomorphic the whole complex plane.

[8] Recall, that Liouville's Theorem says that a bounded entire function is constant.

Of course, the function $\frac{\pi}{\sin \pi z}$ is periodic up to sign and the same property for $\Gamma(z)\Gamma(1-z)$ follows from the following computation:

$$\begin{aligned}\Gamma(z+1)\Gamma(1-(z+1)) &= z\Gamma(z)\Gamma(-z) \\ &= -\Gamma(z)(-z)\Gamma(-z) = -\Gamma(z)\Gamma(1-z).\end{aligned}$$

We conclude that also h is 'periodic up to sign'. In particular, the function $|h(z)|$ is a periodic function. Thus, it suffices to prove that $h(z)$ is bounded on the vertical strip

$$V_0 = \{z \in \mathbb{C} \mid 0 \leq \operatorname{Re}(z) < 1\}.$$

Indeed, let us first remark that $h(z)$ is bounded on the compact set[9]

$$V_0 \cap \{|\operatorname{Im}(z)| \leq 1\}.$$

For the boundedness of h on $V_0 \cap \{|\operatorname{Im}(z)| > 1\}$ it suffices to prove that both functions Γ and $\frac{\pi}{\sin \pi z}$ are bounded on this set. It is not difficult to see that $\frac{\pi}{\sin \pi z}$ is bounded on $V_0 \cap \{|\operatorname{Im}(z)| > 1\}$, so let us turn our attention to $\Gamma(z)$. For $z \in V_0$ with $\operatorname{Im} z > 1$, we have

$$|\Gamma(z)| = \frac{|\Gamma(z+1)|}{|z|} \leq |\Gamma(z+1)| \leq \int_0^\infty e^{-t} t^{\operatorname{Re}(z+1)} \frac{dt}{t} = \Gamma(\operatorname{Re}(z+1)).$$

Now, observe that the function $\Gamma(\operatorname{Re}(z+1))$ is bounded since Γ is bounded on the compact interval $[1,2]$.

Thus, we have shown that the function

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

is an entire bounded function. By Liouville's Theorem, h has to be constant. To conclude that $h = 0$, let us observe

$$h(-z) = -h(z).$$

This equation implies $h(0) = 0$ and hence $h = 0$. □

Corollary 2.1.9. *We have* $\Gamma(1/2) = \sqrt{\pi}$*, and for* $n \in \mathbb{N}$

$$\Gamma\left(\frac{1}{2} + n\right) = \sqrt{\pi} \prod_{k=0}^{n-1} \left(k + \frac{1}{2}\right).$$

Proof. The first formula follows immediately from Proposition 2.1.8:

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

The second formula follows from the first formula using the functional equation

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

□

[9] Recall that every continuous function on a compact set is bounded.

Outlook

[10] In the upcoming lectures, will see many interesting aspects of the Riemann zeta function. But also the Gamma function is an interesting function. We have already seen the formula

$$\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}.$$

This formula together with the functional equation gives us the values of the Gamma function at all rational numbers with denominator 2. In particular, $\Gamma(1/2)$ is a transcendental number. One might ask about the nature of other values $\Gamma(\frac{\bullet}{d})$ at rational numbers with denominator $d > 2$. Surprisingly, the nature of these values is closely related to periods of elliptic curves and abelian varieties. Let us give a simple example. It is not so difficult to compute the following integral:

$$\Omega := \int_1^\infty \frac{1}{\sqrt{x^3 - x}} = \frac{\Gamma(\frac{1}{4})^2}{2^{3/2}\pi^{1/2}}.$$

Of course, this formulas doesn't look very interesting at first glance. But, it has the following interesting arithmetic interpretation. The equation

$$E \colon y^2 = x^3 - x,$$

is an example of an (affine) elliptic curves with complex multiplication. Such elliptic curves play an important role in arithmetic geometry. Now, observe that the right hand side of the defining equation of the elliptic curve E appears in the above integral formula for Ω. More precisely, it can be shown that the differential form $\omega := dx/y$ is an example of a *global differential form* on the above elliptic curve E. This gives the following re-interpretation of the above integral formula:

$$\int_1^\infty \frac{dx}{y} = \int_1^\infty \frac{1}{\sqrt{x^3 - x}} = \frac{\Gamma(\frac{1}{4})^2}{2^{3/2}\pi^{1/2}}.$$

In algebraic geometry, such integrals are called *period integrals* and their values are called *periods*[11]. Thus, the innocent looking integral formula turns out to give an interesting relation between the Gamma value $\Gamma(1/4)$, π and the period of an elliptic curve with complex multiplication. This is only the tip of the iceberg; there are much more general relations between Gamma values and periods (e.g. the Chowla-Selberg formula). This arithmetic interpretation of Gamma values can finally be used to prove deep transcendence results for Gamma values, for example:

Theorem (Chudnovsky)**.** *The values $\Gamma(\frac{1}{4})$ and π are algebraically independent. In particular, $\Gamma(\frac{1}{4})$ is transcendental.*

[10] At the end of the section, we will often give an outlook on interesting topics. Reading these parts of the lecture notes is voluntary. They are neither relevant for understanding the upcoming lectures nor for the final exam.

[11] Periods on a d-dimensional smooth and proper variety X over $\mathbb{Q}$ are defined by integrating an algebraic differential forms ω of degree i along a cycle $C \in H_i(X(\mathbb{C}), \mathbb{Z})$, i.e.,

$$\int_C \omega.$$

In our case, the path

$$\gamma := \{t \in [1, \infty) \mid (t, \sqrt{t^3 - t})\}$$

represents a non-trivial element $\gamma \in H_1(E(\mathbb{C}), \mathbb{Z})$ and so Ω is indeed a period in the above sense.

2.2 The functional equation

Our next goal is to extend the Riemann zeta function to a meromorphic function on $\mathbb{C} \setminus \{1\}$ and to prove its functional equation. Let us start with some facts about rapidly decreasing functions.

2.2.1 The classical theta function

As a preparation for the proof of the functional equation, we will introduce the classical theta function and prove that it is rapidly decreasing in the following sense:

Definition 2.2.1. Let $D \subseteq \mathbb{R}$ be an unbounded subset. A function $f\colon D \to \mathbb{C}$ is called *rapidly decreasing* if for every positive integer $N \in \mathbb{N}$ we have[12] $|t|^N|f(t)| \to 0$ as $|t| \to \infty$. For $D = \mathbb{N}$, we will call such a function a *rapidly decreasing sequence*.

Example 2.2.2. The following functions are examples of rapidly decreasing functions:

(a) The function $f(t) := e^{-t}$ is rapidly decreasing on $D = \mathbb{R}_{>0}$.

(b) The function $n \mapsto e^{-n^2}$ is rapidly decreasing on $D = \mathbb{Z}$.

(c) The function $f(t) := t^{-2021}$ is not rapidly decreasing on $D = [1, \infty)$.

In the proof of the analytic continuation and the functional equation, the *classical theta function* $\theta\colon \mathbb{R}_{>0} \to \mathbb{R}$ given by

$$\theta(t) := \sum_{n\in\mathbb{Z}} e^{-\pi n^2 t} = 1 + 2\sum_{n=1}^{\infty} e^{-\pi n^2 t}$$

will play an important role. Note that this sequence converges absolutely[13] for any real number $t \in \mathbb{R}_{>0}$. The values of the classical theta series are always ≥ 1. Of course, this implies that the theta function is not rapidly decreasing, but the following Lemma shows that the closely related function

$$\omega(t) := \frac{1}{2}(\theta(t) - 1) = \sum_{n=1}^{\infty} e^{-\pi n^2 t}$$

is rapidly decreasing.

Lemma 2.2.3. *The function $\omega(t)$ is rapidly decreasing on $[1, \infty)$.*

Proof. We will prove this in the exercises. □

For later reference, let us record the following elementary property of rapidly decreasing functions:

[12] Alternatively, one can demand for any positive integer $N \in \mathbb{N}$ that $t \mapsto |t|^N|f(t)|$ is bounded on $D \cap [c, \infty)$ for all sufficiently large real numbers c.

[13] Indeed, we have $e^{-n^2 t} = (e^{-nt})^n$. For sufficiently large n, we have $e^{-nt} < 1$ and the series can be estimated by a convergent geometric series.

Lemma 2.2.4. *Let $f\colon [1,\infty) \to \mathbb{R}$ be a continuous function which is rapidly decreasing. Then, for any complex number $s \in \mathbb{C}$, the integral*

$$g(s) := \int_1^\infty f(t)t^s dt$$

converges absolutely and defines a holomorphic function on $\mathbb{C}$.

Proof. Since f is rapidly decreasing, there exists a constant $C > 0$ such that

$$|f(t)|t^{\mathrm{Re}(s)+2} \leq C$$

for all $t \geq 1$, i.e.,

$$|f(t)||t^s| \leq \frac{C}{t^2}.$$

Now, the absolute convergence follows from the convergence of the integral $\int_1^\infty \frac{1}{t^2} dt$. The function g is holomorphic since it is the uniform limit of the holomorphic functions[14]

$$g_n(s) := \int_1^n f(t)t^s dt.$$

□

2.2.2 The functional equation

The functional equation of the Riemann zeta function will follow from the following transformation behaviour of the theta series

$$\theta(t) = \frac{1}{\sqrt{t}}\theta(1/t). \tag{2.1}$$

For the moment, we postpone the proof of (2.1) and deduce the functional equation of the Riemann zeta function from the transformation behaviour of the theta series.

Theorem 2.2.5. *Let us define the* completed Riemann zeta function *as*

$$\xi(s) := \pi^{-s/2}\Gamma(s/2)\zeta(s).$$

The completed zeta function admits a holomorphic continuation to $\mathbb{C} \setminus \{0,1\}$ with simple poles at $s = 0$ and $s = 1$ and satisfies the functional equation

$$\xi(s) = \xi(1-s).$$

Proof. In a first step, let us relate the completed Riemann zeta function

[14] Here, we use the Leibniz rule, i.e., Lemma 2.1.4.

to the theta function. For $s \in \mathbb{C}$ with $\operatorname{Re}(s) > 1$, we have

$$\begin{aligned}\xi(s) = \pi^{-s/2}\Gamma(s/2)\zeta(s) &= \sum_{n=1}^{\infty} \frac{1}{n^s} \int_0^\infty t^{s/2}\pi^{-s/2}e^{-t}\frac{dt}{t} \\ &= \sum_{n=1}^{\infty} \int_0^\infty \left(\frac{t}{n^2\pi}\right)^{s/2} e^{-t}\frac{dt}{t} \\ &\overset{[15]}{=} \sum_{n=1}^{\infty} \int_0^\infty t^{s/2}e^{-\pi t n^2}\frac{dt}{t} \\ &\overset{[16]}{=} \int_0^\infty t^{s/2}\omega(t)\frac{dt}{t} \qquad (2.2)\end{aligned}$$

This will be helpful, since $\theta(t)$ satisfies a nice functional equation which will imply the corresponding functional equation for $\xi(s)$. Let us express the functional equation for $\theta(t)$ in terms of the function $\omega(t) = \frac{1}{2}(\theta(t) - 1)$. Using the functional equation (2.1), we get

$$\begin{aligned}\omega(1/t) = \frac{1}{2}(\theta(1/t) - 1) &= \frac{1}{2}(\sqrt{t}\theta(t) - 1) \\ &= \frac{1}{2}(\sqrt{t}(1 + 2\omega(t)) - 1) = \sqrt{t}\omega(t) + \frac{\sqrt{t}}{2} - \frac{1}{2}.\end{aligned}$$

The strategy is now to use the formula (2.2), i.e.,

$$\xi(s) = \int_0^\infty t^{s/2}\omega(t)\frac{dt}{t}$$

to prove both, the functional equation and the meromorphic continuation. Unfortunately, the integral on the right hand side does not converge for general $s \in \mathbb{C}$. By Lemma 2.2.4, for general $s \in \mathbb{C}$, the convergence at ∞ is not problematic, since $\omega(t)$ is a rapidly decreasing function. But, for $s \in \mathbb{C}$ with $\operatorname{Re}(s) \leq 1$, the integral does not converge absolutely near 0. So let us split the integral into a problematic part and an unproblematic part:

$$\int_0^\infty t^{s/2}\omega(t)\frac{dt}{t} = \underbrace{\int_0^1 t^{s/2}\omega(t)\frac{dt}{t}}_{\text{converges only for } \operatorname{Re}(s)>1} + \underbrace{\int_1^\infty t^{s/2}\omega(t)\frac{dt}{t}}_{\text{converges for all } s\in\mathbb{C}}.$$

Luckily, we can use the transformation behaviour of ω and the substitution $t \mapsto 1/t$ to write the problematic part in a more convenient

[15] Here, we have substituted $\frac{t}{n^2\pi}$ by t. Maybe, you have already wondered why we use the logarithmic differential $\frac{dt}{t}$ instead of dt. One reason is that the logarithmic differential $\frac{dt}{t}$ is invariant under substitutions of the form $t \mapsto c \cdot t$ for a constant $c \in \mathbb{R}$.

[16] Here, we have used the absolute convergence to interchange integration and summation. More precisely, we have used the following fact from analysis: If f_n is a sequence of Lebesgue measurable functions and if $\sum \int |f_n| < \infty$ or $\int \sum |f_n| < \infty$, then

$$\sum \int f_n = \int \sum f_n.$$

way:

$$\begin{aligned}\int_0^1 t^{s/2}\omega(t)\frac{dt}{t} &= \int_1^\infty \omega(1/t)t^{-s/2}\frac{dt}{t}\\ &= \int_1^\infty \left(\sqrt{t}\omega(t)+\frac{\sqrt{t}}{2}-\frac{1}{2}\right)t^{-s/2}\frac{dt}{t}\\ &= \int_1^\infty \omega(t)t^{\frac{1-s}{2}}\frac{dt}{t}+\frac{1}{2}\int_1^\infty t^{\frac{-1-s}{2}}dt-\frac{1}{2}\int_1^\infty t^{-1-\frac{s}{2}}dt\\ &= \underbrace{\int_1^\infty \omega(t)t^{\frac{1-s}{2}}\frac{dt}{t}}_{\text{converges for all } s\in\mathbb{C}} -\frac{1}{1-s}-\frac{1}{s}.\end{aligned}$$

By combining what we have shown above, we obtain

$$\xi(s)=\int_1^\infty t^{s/2}\omega(t)\frac{dt}{t}+\underbrace{\int_1^\infty \omega(t)t^{\frac{1-s}{2}}\frac{dt}{t}}_{\text{converges for all } s\in\mathbb{C}} -\frac{1}{1-s}-\frac{1}{s}. \tag{2.3}$$

The right hand side of this equation is a meromorphic function on $\mathbb{C}\setminus\{0,1\}$ with simple poles at $s=0$ and $s=1$. Thus, we have succeeded to find a meromorphic continuation of $\xi(s)$. The functional equation follows from the fact that the right hand side of (2.3) is invariant under the substitution $s\mapsto 1-s$. □

Using the fact that $\Gamma(s)$ is a non-vanishing meromorphic function on $\mathbb{C}$ with simple poles at the non-positive integers, we deduce immediately the following Corollary.

Corollary 2.2.6. *The Riemann zeta function $\zeta(s)$ admits a holomorphic continuation to $\mathbb{C}\setminus\{1\}$ with a simple pole at $s=1$ with residue 1.*

Proof. We have seen in the Exercises that the Gamma function $\Gamma(s)$ is a non-vanishing holomorphic function on $\mathbb{C}\setminus\{0,-1,-2,\dots\}$ with simple poles at $s=0,-1,-2,\dots$. Thus, $\frac{1}{\Gamma(s)}$ is an entire holomorphic function on $\mathbb{C}$. By Theorem 2.2.5, the Riemann zeta function

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

is holomorphic on $\mathbb{C}\setminus\{0,1\}$ with at most simple poles at $s=0$ and $s=1$. The singularity at $s=0$ is removable, since both $\Gamma(s/2)$ and ξ have simple poles at $s=0$, so the poles cancel. Hence, $\zeta(s)$ extends to $\mathbb{C}\setminus\{1\}$ with a simple pole at $s=1$ with residue

$$\begin{aligned}\operatorname{Res}_{s=1}\zeta(s) &= \lim_{\epsilon\to 0}\epsilon\zeta(1+\epsilon)\\ &= \lim_{\epsilon\to 0}\epsilon\xi(1+\epsilon)\frac{\pi^{(1+\epsilon)/2}}{\Gamma((1+\epsilon)/2)} = \operatorname{Res}_{s=1}\xi(s)\frac{\sqrt{\pi}}{\Gamma(1/2)}=1.\end{aligned}$$

□

In the proof of Theorem 2.2.5, we have written the completed Riemann zeta function as an integral

$$\xi(s) = \int_0^\infty t^{s/2}\omega(t)\frac{dt}{t}.$$

This is a special case of the following definition:

Definition 2.2.7. For a given function $f\colon \mathbb{R}_{>0} \to \mathbb{R}$, we define its Mellin transform by the formula

$$M_f(s) := \int_0^\infty f(t)t^s\frac{dt}{t},$$

whenever the integral exists.

Outlook

The classical theta function

$$\theta(t) := \sum_{n\in\mathbb{Z}} e^{-\pi n^2 t} = 1 + 2\sum_{n=1}^{\infty} e^{-\pi n^2 t}$$

admits various generalizations which appear in different branches of mathematics, e.g., number theory, algebraic geometry, mathematical physics and analysis. Let us indicate the relation to algebraic geometry. One can define the following generalization of the classical theta function. For $z \in \mathbb{Z}$ and $\tau \in \mathbb{H} = \{\tau \in \mathbb{C} \mid \mathrm{Im}(\tau) > 0\}$, we define the *Jacobi theta function* as

$$\Theta(\tau, z) := \sum_{n\in\mathbb{Z}} e^{\pi i n^2\tau + 2\pi i n z}.$$

This series converges absolutely for all $(\tau, z) \in \mathbb{H} \times \mathbb{C}$ and is holomorphic in τ and z. Indeed, this is a generalization of our classical theta function

$$\theta(t) := \Theta(it, 0).$$

Before we describe the relevance of the Jacobi theta functions for algebraic geometry, let us make a brief detour to (complex) elliptic curves. For fixed $\tau \in \mathbb{H}$, let us write Λ_τ for the subgroup $\langle 1, \tau\rangle$ of $(\mathbb{C}, +)$ generated by 1 and τ. Such subgroups are called *lattices* in $\mathbb{C}$. The quotient

$$\mathbb{C}/\Lambda_\tau$$

turns out to be a complex manifold of dimension 1. Even better, it can be shown that $\mathbb{C}/\Lambda_\tau$ are the $\mathbb{C}$-valued points of an elliptic curve [17] defined over $\mathbb{C}$. Thus, it is not only a complex manifold but a complex manifold which 'comes from an algebraic variety'. Conversely, it can be shown that the $\mathbb{C}$-valued points of any elliptic curve over $\mathbb{C}$ are

[17] If you have never seen an elliptic curve, you can think about an elliptic curve (over $\mathbb{C}$) as a curve given by the vanishing locus of the equation

$$y^2 = x^3 - Ax - B$$

for certain $A, B \in \mathbb{C}$ with the property that the polynomial $x^3 - Ax - B$ has only simple roots.

isomorphic to $\mathbb{C}/\Lambda_\tau$ for some $\tau \in \mathbb{C}$. This correspondence is very important for studying isomorphism classes of elliptic curves[18].

Let us now come back to the Jacobi theta function. As we have indicated above, we can associate to $\tau \in \mathbb{H}$ a complex elliptic curve $\mathbb{C}/\Lambda_\tau$. One might hope, that $\Theta(\tau, \cdot)$ is a Λ_τ-periodic function on $\mathbb{C}$. If this were the case, one would obtain a well-defined function on the elliptic curve $\mathbb{C}/\Lambda_\tau$. Unfortunately, it turns out that the Jacobi theta function is not Λ_τ-periodic. Nevertheless, it satisfies a nice transformation behaviour for maps of the form $z \mapsto z + \lambda$ for $\lambda \in \Lambda_\tau$. If one makes this transformation behaviour explicit, it turns out that the Jacobi theta function (and certain generalizations) give an explicit description of sections of line bundles on elliptic curves. This plays an important role in the study of elliptic curves, their line bundles and their cohomology. This can even be generalized to abelian varieties which can be seen as higher dimensional generalizations of elliptic curves. Finally, let us observe that elliptic curves and abelian varieties have many interesting applications to number theory. This closes the circle and we are back in the world of number theory where we belong to, at least in this lecture.

2.3 Fourier Theory

In this section, we deduce the functional equation of the classical theta series

$$\theta(t) := \sum_{n\in\mathbb{Z}} e^{-\pi n^2 t} = 1 + 2\sum_{n=1}^{\infty} e^{-\pi n^2 t}.$$

We will give a purely Fourier-theoretic proof. So let us start with recalling[19] some basic facts from Fourier analysis. Let us write $C^\infty(\mathbb{R})$ for the set of all infinite differentiable complex-valued functions on $\mathbb{R}$.

Definition 2.3.1. The space of *Schwartz* functions consists of all $f \in C^\infty(\mathbb{R})$ such that all derivatives $f^{(n)}$ are rapidly decreasing on $\mathbb{R}$. We will write $\mathcal{S}(\mathbb{R})$ for the space of all Schwartz functions on $\mathbb{R}$. For a Schwartz function $f \in \mathcal{S}(\mathbb{R})$ let us define its *Fourier transform* as

$$\widehat{f}(x) := \int_{\mathbb{R}} f(y)e^{-2\pi i x y}dy.$$

An important example for a function in the Schwartz space is the function $f(x) = e^{-\pi x^2}$. In the exercises, we will verify that it is indeed rapidly decreasing. This function has the important property that it is its own Fourier transform. More generally, we have:

Lemma 2.3.2. *For $t \in \mathbb{R}_{>0}$ let us consider the function $f_t(x) = e^{-\pi t x^2}$. We have*

$$\widehat{f_t} = \frac{1}{\sqrt{t}} f_{1/t}.$$

[18] Finally, this leads to an explicit description of the $\mathbb{C}$-valued points of certain moduli spaces of elliptic curves.

[19] By the way, don't be afraid if you are not familiar with Fourier analysis. We will recall all relevant statements and definitions.

Proof. This will be shown in the exercises. □

We will use the following Theorem without proof:

Theorem 2.3.3 (Fourier Inversion Theorem on $\mathbb{R}$). *For $f \in \mathcal{S}(\mathbb{R})$, we have $\widehat{f} \in \mathcal{S}(\mathbb{R})$ and the Fourier inversion formula holds:*

$$\widehat{\widehat{f}}(x) = f(-x)$$

Proof. For a proof, we refer to Theorem 2.2.14. in Loukas Grafakos' book on 'Classical Fourier Analysis'[20]. □

The functional equation of the classical theta series will follow from the Poisson summation formula:

Theorem 2.3.4 (Poisson summation formula). *For $f \in \mathcal{S}(\mathbb{R})$, we have*

$$\sum_{k \in \mathbb{Z}} f(k) = \sum_{k \in \mathbb{Z}} \widehat{f}(k).$$

Before we give the proof of the Poisson summation formula, let us explain how this implies the functional equation of θ.

Corollary 2.3.5. *The classical theta series satisfies the functional equation:*

$$\theta(t) = \frac{1}{\sqrt{t}} \theta(1/t).$$

Proof. Lemma 2.3.2 gives for the function $f_t(x) = e^{-\pi t x^2}$ the following explicit formula for the Fourier transform

$$\widehat{f_t} = \frac{1}{\sqrt{t}} f_{1/t}.$$

The functional equation of the theta function follows now immediately from the Poisson summation formula:

$$\theta(t) = \sum_{k \in \mathbb{Z}} f_t(k) = \sum_{k \in \mathbb{Z}} \widehat{f_t}(k) = \sum_{k \in \mathbb{Z}} \frac{1}{\sqrt{t}} f_{1/t}(k) = \frac{1}{\sqrt{t}} \theta(1/t).$$

□

For the proof of the Poisson summation formula, we will need the Fourier expansion of periodic functions, i.e., functions $f \colon \mathbb{R} \to \mathbb{C}$ with $f(x+1) = f(x)$. We will identify such functions with functions on $S^1 := \mathbb{R}/\mathbb{Z}$. We will also use the following Theorem form Fourier Analysis without proof.

Theorem 2.3.6 (Fourier Expansion on S^1). *For $f \in C^\infty(\mathbb{R}/\mathbb{Z})$ and $x \in \mathbb{R}$, we have*

$$f(x) = \sum_{k \in \mathbb{Z}} c_k(f) e^{2\pi i k x}, \tag{2.4}$$

where $c_k(f) := \int_0^1 f(t) e^{-2\pi i k t} dt$. The sum (2.4) converges absolutely and uniformly in x. The elements of the sequence $(c_k(f))_{k \in \mathbb{Z}}$ are called Fourier coefficients *and form a rapidly decreasing sequence on $\mathbb{Z}$.*

[20] Loukas Grafakos. *Classical Fourier analysis*, volume 249 of *Graduate Texts in Mathematics*. Springer, New York, second edition, 2008. ISBN 978-0-387-09431-1

Proof. For a proof, we refer to Theorem 2.2.14. in Loukas Grafakos' book on 'Classical Fourier Analysis'[21]. □

Let us now deduce the Poisson summation formula using Fourier theory:

Proof of Poisson summation. For a given Schwartz function $f \in \mathcal{S}(\mathbb{R})$ let us define

$$F(x) := \sum_{k\in\mathbb{Z}} f(x+k).$$

Since f and all its derivatives are rapidly decreasing, this sum converges absolutely and defines a smooth function on $\mathbb{R}$. Furthermore, the function F is periodic:

$$F(x+1) = \sum_{k\in\mathbb{Z}} f(x+k+1) = \sum_{k\in\mathbb{Z}} f(x+k) = F(x).$$

Thus, by Theorem 2.3.6, it admits a Fourier expansion

$$F(x) = \sum_{k\in\mathbb{Z}} c_k(F)e^{2\pi ikx},$$

with $c_k(F) := \int_0^1 F(t)e^{-2\pi ikt}dt$. The result follows from the following computation:

$$\begin{aligned}\sum_{k\in\mathbb{Z}} f(k) = F(0) &= \sum_{k\in\mathbb{Z}} \int_0^1 F(t)e^{-2\pi ikt}dt \\ &= \sum_{k\in\mathbb{Z}} \int_0^1 \sum_{l\in\mathbb{Z}} f(t+l)e^{-2\pi ikt}dt \\ &= \sum_{k\in\mathbb{Z}} \sum_{l\in\mathbb{Z}} \int_l^{l+1} f(t)e^{-2\pi ik(t-l)}dt \\ &= \sum_{k\in\mathbb{Z}} \int_{\mathbb{R}} f(t)e^{-2\pi ikt}dt = \sum_{k\in\mathbb{Z}} \widehat{f}(k).\end{aligned}$$

Here, we have interchanged summation and integration by absolute convergence, compare footnote[16]. □

Outlook

Fourier analysis is everywhere. Perhaps, you might have already seen examples of Fourier expansions for periodic functions in terms of the sin and cos functions. In Theorem 2.3.6, we have seen the following formula for a periodic function f

$$f(x) = \sum_{k\in\mathbb{Z}} c_k(f)e^{2\pi ikx},$$

[21] Loukas Grafakos. *Classical Fourier analysis*, volume 249 of *Graduate Texts in Mathematics*. Springer, New York, second edition, 2008. ISBN 978-0-387-09431-1

where $c_k(f) := \int_0^1 f(t)e^{-2\pi ikt}dt$. By writing $e^{2\pi ikx} = \cos(2\pi ikx) + i\sin(2\pi ikx)$, we obtain the following version of the Fourier expansion formula

$$f(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty}(a_k\cos(2\pi ikx) + b_k\sin(2\pi ikx)), \qquad (2.5)$$

with a_k and b_k defined for $k \geq 1$ by the formula

$$\begin{aligned} a_k &= c_k + c_{-k} \\ b_k &= i(c_k - c_{-k}). \end{aligned}$$

This allows one to write periodic functions as an infinite linear combination of simple trigonometric functions. The above Fourier expansion formula in (2.5) has the advantage that the coefficients b_k vanish if f is real-valued. The fact that the sequence of Fourier coefficients a_k and b_k (respectively c_k) is rapidly decreasing has important applications. One often obtains a quiet good approximation of f by considering only the truncated sequences[22]

$$\frac{a_0}{2} + \sum_{k=1}^{N}(a_k\cos(2\pi ikx) + b_k\sin(2\pi ikx)), \qquad \text{for some } N.$$

This has many applications, even outside of mathematics. Many important modern developments would not exist without Fourier Analysis. Just to mention a few of them: efficient compression of data (mp3, mp4, jpg), bandpass filters, image processing, face recognition, etc.. This is also a good place to recommend the following nice video[23] about a long forgotten machine – the Harmonic Analyzer.

2.4 *Chebyshev bounds for primes*

In the following two sections, we will study the asymptotic distribution of prime numbers. Although, no exact formula for the number of primes

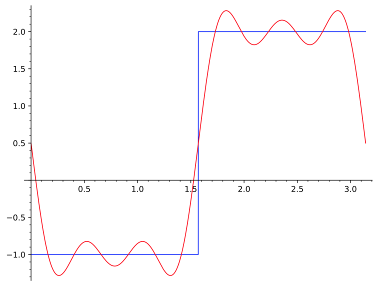
$$\pi(x) := \sum_{p \leq x} 1$$

less than a given positive real number x is known, this functions satisfies a quite regular pattern asymptotically[24]. The aim of the following two sections is to prove the Prime Number Theorem (PNT), which says

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

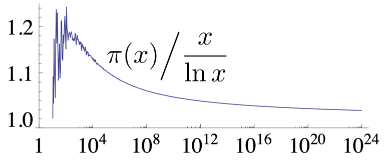
Let us briefly indicate the history of the Prime Number Theorem. It has been known since Euclid that $\pi(x) \to \infty$ as $x \to \infty$. Euler was able to proof $\pi(x)/x \to 0$ as $x \to \infty$. In other words, as x increases, the

[22] Approximation of a piecewise linear function:



[23] Here is a link to the YouTube video about the Harmonic Analyzer:

[24] The following plot shows the function $\frac{\pi(x)}{x/\ln x}$.



prime numbers become rarer. Around 1850, Chebyshev found more quantitative description for the growth of the function $\pi(x)$. He was able to prove that there exist positive real numbers a and b such that

$$a\frac{x}{\log x} \leq \pi(x) \leq b\frac{x}{\log x}.$$

Furthermore, he could also show that the limit

$$\lim_{x\to\infty} \frac{\pi(x)}{x/\log x}$$

is equal to 1 if it exists. Finally, Hadamard and de la Vallée-Poussin (1896) succeeded to prove independently the Prime Number Theorem, i.e.,

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

In this lecture, we will prove the Chebyshev bounds for $\pi(x)$. This proof will not involve the Riemann zeta function and is rather elementary. In the next lecture, we will prove the PNT using the Riemann zeta function. Let us start with the following elementary estimate for the least common multiple of the first n positive integers.

Lemma 2.4.1. *Let $d_k := lcm(1,2,\dots,k)$ be the least common multiple of the first k positive integers. For a positive integer n we have the estimate*

$$d_{2n+1} > 4^n.$$

Proof. The value of the integral

$$I := \int_0^1 x^n(1-x)^n dx$$

can be bounded by $0 < I < 4^{-n}$, since we have the estimate $0 < x(1-x) \leq 1/4$ for $0 < x \leq 1$. On the other hand, by expanding the product in the integrand, we get

$$x^n(1-x)^n = a_n x^n + a_{n+1}x^{n+1} + \cdots + a_{2n}x^{2n},$$

for suitable integers $a_n, \dots, a_{2n}$. We can compute the integral explicitly in terms of the coefficients $a_n, \dots, a_{2n}$:

$$I = \frac{a_n}{n+1} + \cdots + \frac{a_{2n}}{2n+1}.$$

This implies that $d_{2n+1}I$ is a positive integer. In particular, we have $d_{2n+1}I \geq 1$. Together with the estimate $I < 4^{-n}$, we obtain the desired estimate:

$$d_{2n+1} \geq \frac{1}{I} > 4^n.$$

□

Lemma 2.4.2. *For each integer $n \geq 2$, we can bound the product of all primes less than n from above by*

$$\prod_{p \leq n} p < 4^n.$$

Proof. We will prove the statement by induction over n. The statement is obviously true for $n = 2, 3$. For an integer $n \geq 4$, let us assume that we have already shown the estimate

$$\prod_{p \leq k} p < 4^k$$

for all integers $k \leq n - 1$. Our aim is to prove the statement for n. If n is even, we have

$$\prod_{p \leq n} p = \prod_{p \leq n-1} p < 4^{n-1} < 4^n,$$

so the statement is obviously true and we may assume that $n = 2m - 1$ for some $m \geq 3$. By dividing the product into two parts, we get by the induction hypothesis the estimate

$$\prod_{p \leq n} p = \prod_{p \leq m} p \prod_{m < p \leq 2m-1} p \leq 4^m \prod_{m < p \leq 2m-1} p.$$

Now, let us observe that all primes of the last product divide the factorials in the numerator of the binomial coefficient

$$\binom{2m-1}{m} = \frac{(2m-1)!}{m!(m-1)!}$$

but they do not divide the factorials in the denominator. Hence, we get $\prod_{m<p\leq 2m-1} p \mid \binom{2m-1}{m}$ which proves

$$\prod_{m < p \leq 2m-1} p \leq \binom{2m-1}{m}.$$

The binomial theorem gives

$$2\binom{2m-1}{m} = \binom{2m-1}{m} + \binom{2m-1}{m-1} < \sum_{k=0}^{2m-1} \binom{2m-1}{k} = (1+1)^{2m-1}.$$

and allows us to estimate the binomial coefficient $\binom{2m-1}{m} < 2^{2m-2} = 4^{m-1}$. Thus, we get

$$\prod_{p \leq n} p \leq 4^m \binom{2m-1}{m} < 4^m 4^{m-1} = 4^n.$$

□

Theorem 2.4.3 (Chebyshev Bounds)**.** *For all $x \geq 3$ we have*

$$a\frac{x}{\log x} < \pi(x) < b\frac{x}{\log x}$$

for $a = \frac{1}{2}\log 2$ and $b = 6\log 2$.

Proof. For $3 \leq x < 6$, the estimates are checked by a straightforward computation. Thus, we may assume $x \geq 6$. Choose n such that $2n + 1 \leq x \leq 2n + 3$ and let us write

$$d_{2n+1} := \operatorname{lcm}(1, \ldots, 2n+1) = p_1^{\alpha_1} \ldots p_s^{\alpha_s}$$

for the prime decomposition of d_{2n+1} with $s = \pi(2n+1)$. First note, that $p_i^{\alpha_i} \leq 2n + 1$ for all $1 \leq i \leq s$, as each $p_i^{\alpha_i}$ must appear on the list $1, 2, \ldots, 2n + 1$. Therefore,

$$d_{2n+1} \leq (2n+1)^s.$$

On the other hand, we have seen in Lemma 2.4.1 that $d_{2n+1} > 4^n$ and deduce

$$4^n < (2n+1)^s.$$

Taking logarithms gives for $x \geq 6$

$$\begin{aligned}\pi(x) \geq \pi(2n+1) = s > \frac{2n\log(2)}{\log(2n+1)} &> \frac{(x-3)\log(2)}{\log x} \\ &\geq \frac{(x/2)\log(2)}{\log x} = a\frac{x}{\log x}.\end{aligned}$$

The estimate from above follows from the following computation

$$\begin{aligned}\pi(x) &= \sum_{p \leq x} 1 = \sum_{p \leq \sqrt{x}} 1 + \sum_{\sqrt{x} < p \leq x} 1 \\ &< \pi(\sqrt{x}) + \sum_{\sqrt{x} < p \leq x} \frac{\log p}{\log \sqrt{x}} < \sqrt{x} + \frac{2}{\log x}\sum_{p \leq x} \log p \\ &= \sqrt{x} + \frac{2}{\log x}\log\prod_{p \leq x} p \overset{[25]}{\leq} \sqrt{x} + \frac{4x\log 2}{\log x} \overset{[26]}{\leq} 6\log 2\frac{x}{\log x}.\end{aligned}$$

Here, we have used Lemma 2.4.2 and the estimate $\sqrt{x} \leq \frac{2x\log 2}{\log x}$ for $x \geq 6$, see [25] and [26]. □

Outlook

In 1845, Joseph Bertrand conjectured the following statement about primes.

Conjecture (Bertrand's Postulate)**.** *For each positive integer n, there is at least one prime in the interval $(n, 2n]$.*

[25] Here, we use Lemma 2.4.2.

[26] This follows from the estimate $\sqrt{x} \leq \frac{2x\log 2}{\log x}$ for $x \geq 6$.

This conjecture is now a Theorem and has been proven by Chebyshev in 1850. Indeed, Bertrand's Postulate follows from the better estimate

$$a\frac{x}{\log x} \leq \pi(x) \leq b\frac{x}{\log x}.$$

with $a \approx 0.92129$ and $b \approx 1.10555$ obtained by Chebyshev. In the exercises, we will give an elementary proof of Bertrand's Postulate.

2.5 The Prime Number Theorem

In the last section, we have already seen the Chebyshev bounds for the distribution of prime numbers

$$a\frac{x}{\log x} < \pi(x) < b\frac{x}{\log x}.$$

In this section, we will sketch the proof of the Prime Number Theorem, i.e.,

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

In particular, we want to explain why the zero-free region of the Riemann zeta function plays an important role in the proof of the PNT. It will be convenient to introduce the following notation. For two functions $f, g\colon \mathbb{R}_{>0} \to \mathbb{R}$, we will write

$$f(x) \sim g(x), \qquad \text{as } x \to \infty$$

if and only if $\frac{f(x)}{g(x)} \to 1$ as $x \to \infty$. With this notation, the Prime Number Theorem can be formulated as follows.

Theorem 2.5.1 (Prime Number Theorem)**.** *We have*

$$\pi(x) \sim \frac{x}{\log x} \qquad \textit{as } x \to \infty.$$

Before we give the proof, we will need a few auxiliary results. Let us define the function

$$\Phi(s) := \sum_p \frac{\log p}{p^s}.$$

It is easily checked that the series defining $\Phi(s)$ is absolutely convergent for $\mathrm{Re}(s) > 1$ and defines a holomorphic function in this half-plane. Next, we want to prove the non-vanishing of $\zeta(s)$ on $\mathrm{Re}(s) \geq 1$. First, we recall the following fact from complex analysis.

Lemma 2.5.2. *Let $f\colon U \to \mathbb{C}$ be a meromorphic function on an open subset $U \subseteq \mathbb{C}$. Then, $\frac{\partial}{\partial s}\log f(s)$ has at most simple poles on U. Furthermore, if $\mu \in \mathbb{Z}$ is the order*[27] *of f at $s_0 \in U$ then, the residue at $s_0 \in U$ is given by*

$$\mathrm{Res}_{s=s_0} \frac{\partial}{\partial s}\log f(s) = \mathrm{Res}_{s=s_0} \frac{f'(s)}{f(s)} = \mu.$$

[27] Recall that the order of a meromorphic function at $s_0 \in U$ is μ if the Laurent expansion of f near s_0 starts at $k = \mu$, i.e.,

$$f(s) = \sum_{k=\mu}^{\infty} a_k(s-s_0)^k, \qquad \text{with } a_\mu \neq 0.$$

Proof. Let $s_0 \in U$ and consider the Laurent expansion of f at s_0, i.e.,

$$f(s) = \sum_{k=\mu}^{\infty} a_k (s-s_0)^k, \quad \text{with } a_\mu \neq 0.$$

The claim follows from computing the leading terms in the Laurent expansion of f' and $1/f$ near $s_0 \in U$. They are given by the following formulas[28]

$$f'(s) = \sum_{k=\mu}^{\infty} a_k \cdot k(s-s_0)^{k-1} = a_\mu \cdot \mu \cdot (s-s_0)^{\mu-1} + O((s-s_0)^{\mu}),.$$

and

$$\frac{1}{f(s)} = \frac{1}{a_\mu}(s-s_0)^{-\mu} + O((s-s_0)^{-\mu+1})$$

Thus, we get

$$\frac{f'(s)}{f(s)} = \mu(s-s_0)^{-1} + O(1)$$

which proves that f'/f has at most a simple pole at $s_0 \in U$ with $\mathrm{Res}_{s=s_0} \frac{f'(s)}{f(s)} = \mu$. □

We are now ready to prove the non-vanishing of $\zeta(s)$ for $\mathrm{Re}(s) \geq 1$.

Proposition 2.5.3 (Non-vanishing of $\zeta(s)$ on $\mathrm{Re}(s) \geq 1$)**.**

(a) The function $\Phi(s)$ extends meromorphically to $\mathrm{Re}(s) > \frac{1}{2}$*. The poles of $\Phi(s)$ are all simple and located at $s = 1$ and at the zeroes of $\zeta(s)$. The pole of $\Phi(s)$ at $s = 1$ has residue* 1*.*

(b) The Riemann zeta function does not vanish on the half-plane $\mathrm{Re}(s) \geq 1$*.*

Proof. For $\mathrm{Re}(s) > 1$ the Euler product formula implies the non-vanishing of $\zeta(s)$. Furthermore, the Euler product gives the following formula[29]

$$\begin{aligned} -\frac{\zeta'(s)}{\zeta(s)} &= -\frac{d}{ds}\log\zeta(s) = \sum_p \frac{d}{ds}\log(1-p^{-s}) = \sum_p \frac{\log p}{p^s - 1} \\ &= \sum_p \frac{\log p}{p^s}\left(1 + \frac{1}{p^s-1}\right) = \Phi(s) + \sum_p \frac{\log p}{p^s(p^s-1)}. \end{aligned} \tag{2.6}$$

It is not difficult to check[30] that the final sum converges for $\mathrm{Re}(s) > \frac{1}{2}$. Since we have already seen that $\zeta(s)$ extends meromorphically to the entire complex plane with a simple pole at $s = 1$, we deduce from (2.6) the statement (a) in combination with Lemma 2.5.2.

It remains to show that $\zeta(s)$ does not vanish on the line given by $\mathrm{Re}(s) = 1$. For a given real number $\alpha \neq 0$, let us write μ for the order of vanishing of $\zeta(s)$ at $s = 1 + i\alpha$ and ν for the order of $\zeta(s)$ at $s = 1 + 2i\alpha$. From Lemma 2.5.2, we conclude

$$\mathrm{Res}_{s=1}\frac{\zeta'(s)}{\zeta(s)} = -1, \quad \mathrm{Res}_{s=1+i\alpha}\frac{\zeta'(s)}{\zeta(s)} = \mu, \quad \mathrm{Res}_{s=1+2i\alpha}\frac{\zeta'(s)}{\zeta(s)} = \nu.$$

[28] This is a good place to remind you of the Landau big-O symbol. For two functions f and g with g non-zero and $s_0 \in \mathbb{C}$, we write

$$f(s) = O(g(s)) \quad \text{as } s \to s_0$$

if and only if

$$\limsup_{s\to s_0}\left|\frac{f(s)}{g(s)}\right| < \infty.$$

This means that $|f(s)|$ grows at most like a constant times $|g(s)|$ as $s \to s_0$.

[29] Here, the choice of the branch of the logarithm does not play a role, since we take the derivative.

[30] For example, as follows: For all primes p, we have

$$\frac{\log p}{|p^s(p^s-1)|} \leq \frac{\log p}{(p-1)^{2\,\mathrm{Re}(s)}}.$$

For every $\epsilon > 0$ and all sufficiently large primes p, we have $\log p < p^\epsilon$. This gives

$$\frac{\log p}{|p^s(p^s-1)|} \leq \frac{1}{(p-1)^{2\,\mathrm{Re}(s)-\epsilon}},$$

and the absolute convergence for $\mathrm{Re}(s) > \frac{1}{2} + \epsilon$ follows from the convergence of

$$\sum_{n\geq 1}\frac{1}{n^{2\,\mathrm{Re}(s)-\epsilon}}.$$

Since $\epsilon > 0$ was arbitrary, we deduce the desired convergence.

The sum $\sum_p \frac{\log p}{p^s(p^s-1)}$ is holomorphic on the half-plane $\mathrm{Re}(s) > \frac{1}{2}$, so by (2.6) the residue of Φ at $s = s_0$ coincides with the residue of $-\frac{\zeta'(s)}{\zeta(s)}$ at $s = s_0$ for all $s_0 \in \mathbb{C}$ with $\mathrm{Re}(s_0) > 1/2$.

So, we obtain

$$\lim_{\epsilon \searrow 0} \epsilon\Phi(1+\epsilon) = 1, \quad \lim_{\epsilon \searrow 0} \epsilon\Phi(1+i\alpha+\epsilon) = -\mu, \quad \lim_{\epsilon \searrow 0} \epsilon\Phi(1+2i\alpha+\epsilon) = -\nu.$$

Because of $\Phi(\bar{s}) = \overline{\Phi(s)}$, we get also the residues at $s = 1 - i\alpha$ and $s = 1 - 2i\alpha$, i.e.,

$$\lim_{\epsilon \searrow 0} \epsilon\Phi(1+\epsilon) = 1, \quad \lim_{\epsilon \searrow 0} \epsilon\Phi(1 \pm i\alpha+\epsilon) = -\mu, \quad \lim_{\epsilon \searrow 0} \epsilon\Phi(1 \pm 2i\alpha+\epsilon) = -\nu.$$

In particular, this gives

$$\lim_{\epsilon \searrow 0} \sum_{r=-2}^{2} \binom{4}{2+r} \epsilon\Phi(1+\epsilon+ir\alpha) = \binom{4}{2} - 2\binom{4}{3}\mu - 2\binom{4}{4}\nu.$$

The binomial theorem implies for any $\epsilon > 0$

$$0 < \sum_p \frac{\log p}{p^{1+\epsilon}} (p^{i\alpha/2} + p^{-i\alpha/2})^4 = \sum_{r=-2}^{2} \binom{4}{2+r} \Phi(1+\epsilon+ir\alpha).$$

This positivity of the latter term shows $\sum_{r=-2}^{2} \binom{4}{2+r} \epsilon\Phi(1+\epsilon+ir\alpha) > 0$ and hence

$$6 - 8\mu - 2\nu \geq 0.$$

Since μ and ν are non-negative, we deduce $\mu = 0$. Since μ was the order of vanishing of ζ at $s = 1 + i\alpha$ and α was arbitrary, we deduce that $\zeta(s)$ does not vanish for $\mathrm{Re}(s) = 1$. □

We will use the following purely analytic Theorem without proof.

Theorem 2.5.4. *Let $f\colon \mathbb{R}_{>0} \to \mathbb{C}$ be a bounded and continuous function. Suppose that the function $g(z) = \int_0^\infty f(t)e^{-zt}dt$ which is initially only defined for $\mathrm{Re}(z) > 0$ extends holomorphically to $\mathrm{Re}(z) \geq 0$. Then $\int_0^\infty f(t)dt$ exists and we have*

$$\int_0^\infty f(t)dt = g(0).$$

Proof. We refer the interested reader to Page 707 in Zagier's paper on Newman's proof of the PNT[31] for a proof. □

Let us introduce the function

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

By combining the analytic Theorem 2.5.4 with the non-vanishing of $\zeta(s)$ on $\mathrm{Re}(s) \geq 1$, we will prove:

[31] D. Zagier. Newman's short proof of the prime number theorem. *Amer. Math. Monthly*, 104(8):705–708, 1997. ISSN 0002-9890. DOI: 10.2307/2975232. URL https://doi.org/10.2307/2975232

Corollary 2.5.5. *The following integral is convergent:*

$$\int_1^\infty \frac{\vartheta(x)-x}{x^2}dx. \tag{2.7}$$

Proof. Let us first observe for $\mathrm{Re}(s)>0$, we have the equation

$$\begin{aligned}\Phi(s+1) &= \sum_p \frac{\log p}{p^{s+1}} \\ &= (s+1)\int_1^\infty \frac{\vartheta(x)}{x^{s+2}}dx = (s+1)\int_0^\infty e^{-(s+1)t}\vartheta(e^t)dt. \end{aligned} \tag{2.8}$$

We want to apply Theorem 2.5.4 to the continuous function

$$f(t) := \vartheta(e^t)e^{-t}-1.$$

Thus, let us consider the integral

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

For $\mathrm{Re}(s)>0$, we can rewrite this integral using (2.8) as follows

$$g(s) = \int_0^\infty f(t)e^{-st}dt = \int_0^\infty \vartheta(e^t)e^{-(s+1)t} - e^{-st}dt = \frac{\Phi(s+1)}{(s+1)} - \frac{1}{s}.$$

Let us check the assumptions of Theorem 2.5.4. We have to check that $f(t)$ is bounded on $\mathbb{R}_{>0}$ and that $g(s)$ extends to $\mathrm{Re}(s)\geq 0$. The boundedness of f follows by taking logarithms of the estimate in Lemma 2.4.2:

$$|f(t)| \leq \frac{\vartheta(e^t)}{e^t} = \frac{1}{e^t}\log\prod_{p\leq e^t} p \leq \frac{1}{e^t}(2\log 2)e^t \leq 2\log 2. \tag{2.9}$$

Let us now check that the function $g(s)$ extends to $\mathrm{Re}(s)\geq 0$. The function $g(s)$ is initially defined only for $\mathrm{Re}(s)>0$. By Proposition 2.5.3, we know that $\Phi(s+1)$ is meromorphic on $\mathrm{Re}(s)\geq 0$ and its only pole[32] is a simple pole at $s=0$ with residue 1. Since $\frac{1}{s}$ does also have a pole with residue 1 at $s=1$, we deduce that $g(s) = \frac{\Phi(s+1)}{s+1} - \frac{1}{s}$ extends to a holomorphic[33] function on the half-plane $\mathrm{Re}(s)\geq 0$. Hence $g(s)$ extends holomorphically to $\mathrm{Re}(s)\geq 0$ and we can apply Theorem 2.5.4 to obtain the convergence of the integral

$$\begin{aligned} g(0) = \int_0^\infty f(t)dt &= \int_0^\infty \vartheta(e^t)e^{-t}-1dt \\ &= \int_0^\infty \frac{\vartheta(e^t)-e^t}{e^t}dt = \int_1^\infty \frac{\vartheta(x)-x}{x^2}dx. \end{aligned}$$

Here, we have made the substitution $x=e^t$ in the last step. □

We are now well-prepared to prove the following Theorem which will then imply the Prime Number Theorem.

[32] Here, the non-vanishing of $\zeta(s)$ on $\mathrm{Re}(s)\geq 1$ enters the argument. If $\zeta(s)$ had zeroes in $\mathrm{Re}(s)\geq 1$ the function $\Phi(s)$ would have further poles and we could not apply Theorem 2.5.4.

[33] Note, that the functions $\frac{\Phi(s+1)}{s+1}$ and $\frac{1}{s}$ have only simple poles with the same residue on $\mathrm{Re}(s)\geq 0$. So, the poles in the difference 'cancel' and $g(s)$ extends holomorphicaly to $\mathrm{Re}(s)\geq 0$.

Theorem 2.5.6. *We have $\vartheta(x) \sim x$ as $x \to \infty$.*

Proof. We prove the theorem by contradiction. Let us first assume, that for some $\lambda > 1$ there are arbitrary large real numbers y with $\vartheta(y) \geq \lambda y$. For any such y we have

$$\int_y^{\lambda y} \frac{\vartheta(x) - x}{x^2} dx \geq \int_y^{\lambda y} \frac{\lambda y - x}{x^2} dx = \int_1^{\lambda} \frac{\lambda - x}{x^2} dx > 0.$$

The right hand side is a positive real number which is independent of y. Since there are arbitrarily large values y with $\vartheta(y) \geq \lambda y$, this contradicts the convergence of the integral

$$\int_1^{\infty} \frac{\vartheta(x) - x}{x^2} dx.$$

Next, let us assume that for some $\lambda < 1$ there are arbitrary large real numbers y with $\vartheta(y) \leq \lambda y$. By a similar argument, we obtain

$$\int_{\lambda y}^{y} \frac{\vartheta(x) - x}{x^2} dx \leq \int_{\lambda y}^{y} \frac{\lambda y - x}{x^2} dx = \int_{\lambda}^{1} \frac{\lambda - x}{x^2} dx < 0.$$

The right hand side is a negative real number which is independent of y. Again this would contradict the convergence of the integral (2.7). □

We are now ready to prove the Prime Number Theorem:

Proof of the Prime Number Theorem. The Prime Number Theorem is an easy consequence of the asymptotic $\vartheta(x) \sim x$ since we can estimate $\pi(x)$ in terms of $\vartheta(x)$. Let us start with an upper bound:

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

For the lower bound, let us fix a positive real number $\epsilon > 0$:

$$\begin{aligned}\vartheta(x) \geq \sum_{x^{1-\epsilon} \leq p \leq x} \log p \geq \sum_{x^{1-\epsilon} \leq p \leq x} (1-\epsilon) \log x \\ = (1-\epsilon) \log x \pi(x) - (1-\epsilon) \log x \pi(x^{1-\epsilon}).\end{aligned}$$

The Chebyshev bounds give $(1-\epsilon) \log x \pi(x^{1-\epsilon}) \leq 6 \log 2 x^{1-\epsilon}$ and we obtain

$$\vartheta(x) \geq (1-\epsilon) \log x \pi(x) - 6 \log 2 x^{1-\epsilon}.$$

Combining both estimates gives for any $\epsilon > 0$

$$\frac{\vartheta(x)}{(1-\epsilon)\log x} + 6 \log 2 \frac{x^{1-\epsilon}}{(1-\epsilon)\log x} \geq \pi(x) \geq \frac{\vartheta(x)}{\log x}.$$

Dividing by $\frac{x}{\log x}$ and passing to the limit $x \to \infty$ shows

$$\frac{1}{1-\epsilon} \geq \lim_{x \to \infty} \frac{\pi(x)}{\frac{x}{\log x}} \geq 1.$$

Since this holds for any $\epsilon > 0$, we get the desired asymptotic formula

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

□

Outlook

In this lecture, we have seen that the non-vanishing of $\zeta(s)$ for $\operatorname{Re}(s) \geq 1$ can be used to prove the Prime Number Theorem. The PNT is equivalent to the following statement[34]

$$\pi(x) = \frac{x}{\log x} + o\left(\frac{x}{\log x}\right), \quad \text{as } x \to \infty.$$

It is a natural question if one can improve the error term in the PNT. Indeed, there is a much stronger relation between the non-vanishing region of the Riemann zeta function and the distribution of prime numbers. For refinements of the PNT it is more convenient to introduce the *integral logarithm*

$$\operatorname{Li}(x) := \int_2^x \frac{1}{\log t} dt.$$

It is not difficult to see that $\operatorname{Li}(x) \sim \frac{x}{\log x}$, so we can reformulate the PNT as follows

$$\pi(x) = \operatorname{Li}(x) + o\left(\operatorname{Li}(x)\right), \quad \text{as } x \to \infty.$$

Using the non-vanishing of the Riemann zeta function on $\operatorname{Re}(s) \geq 1$ and a slightly more refined argument, Vallée Poussin was able to give a more precise estimate for the error term

$$\pi(x) = \operatorname{Li}(x) + O(xe^{-a\sqrt{\log x}}), \quad \text{as } x \to \infty,$$

for some positive real number a. It was Bernhard Riemann who observed that one could refine this error term considerably if one knew that the Riemann zeta function is non-vanishing for $\operatorname{Re}(s) > \frac{1}{2}$. We have already seen that the Riemann zeta function does not have any zeroes for $\operatorname{Re}(s) \geq 1$. Using the functional equation, it is not difficult to find all zeroes in the region $\operatorname{Re}(s) \leq 1$ (see exercises). The remaining strip $0 < \operatorname{Re}(s) < 1$ is called *the critical strip* and the Riemann Hypothesis claims that all zeroes in this strip lie in the central line $\operatorname{Re}(s) = \frac{1}{2}$. This lead to one of the most important conjectures in mathematics.

Conjecture (The Riemann Hypothesis). *All zeroes of the Riemann zeta function in the strip* $0 < \operatorname{Re}(s) < 1$ *lie on the line* $\operatorname{Re}(s) = \frac{1}{2}$.

[34] This is a good place to remind you of the Landau little-o symbol. For two functions f and g with g non-zero and $x_0 \in \mathbb{R} \cup \{\pm\infty\}$, we write

$$f(x) = o(g(x)) \quad \text{as } x \to x_0$$

if and only if

$$\limsup_{x\to x_0} \left|\frac{f(x)}{g(x)}\right| = 0.$$

This means that $|f|$ grows much slower than $|g|$ if we approach x_0.

It can be shown that the Riemann Hypothesis is indeed equivalent to the following refinement of the Prime Number Theorem:

$$\pi(x) = \mathrm{Li}(x) + O(\sqrt{x} \log x).$$

Thus, the asymptotic distribution of the prime numbers is intimately related to the zero-free region of the Riemann zeta function.

2.6 Euler's Formula

In this section, we will compute the values of the Riemann zeta function at the even positive integers. The question about the values of the Riemann zeta function has its origin in the 17-th century. In 1644, the Italian mathematician Pietro Mengoli raised the question of the value of the convergent series

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

Afterwards, mainly mathematicians of the city Basel worked on this problem and it became popular under the name *Basel problem*. In his groundbreaking work[35] Leonhard Euler solved this problem finally in 1735. He did not only prove the formula [35]

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

but computed more generally for all positive integers k the value of the series

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

The aim of this section is to prove Euler's remarkable formula

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

More generally, we will introduce the *Hurwitz zeta function* and prove an explicit formula for all its values at the negative integers.

2.6.1 Bernoulli numbers and Bernoulli polynomials

The Bernoulli numbers and more generally Bernoulli polynomials occur everywhere in mathematics. They also play an important role in studying the special values of the Riemann zeta function.

Definition 2.6.1. The *Bernoulli numbers* B_n for $n \geq 0$ are defined by the formula

$$\frac{t}{e^t - 1} = \sum_{n=0}^{\infty} B_n \frac{t^n}{n!}.$$

The first Bernoulli numbers are $B_0 = 1$, $B_1 = -\frac{1}{2}$, $B_2 = \frac{1}{6}$, $B_3 = 0$, $B_4 = -\frac{1}{30}$, $B_5 = 0$. All these values are rational, and except for B_1 all Bernoulli numbers of odd index vanish.

Lemma 2.6.2. *The Bernoulli numbers are all rational and $B_{2n+1} = 0$ for $n \geq 1$.*

Proof. The formal power series $\frac{e^t-1}{t}$ has coefficients in $\mathbb{Q}$, i.e.,

$$\frac{e^t - 1}{t} \in \mathbb{Q}[\![t]\!].$$

Its leading term is 1 and hence a unit in $\mathbb{Q}$. Thus, the inverse of $\frac{e^t-1}{t}$ exists in the formal power series ring $\mathbb{Q}[\![t]\!]$ and we get

$$\frac{t}{e^t - 1} = \sum_{n=0}^{\infty} B_n \frac{t^n}{n!} \in \mathbb{Q}[\![t]\!],$$

and hence the rationality of the Bernoulli numbers. The claim $B_{2n+1} = 0$ for $n \geq 1$ is equivalent to

$$\frac{t}{e^t - 1} - B_1 t$$

being an even function. We have

$$\begin{aligned}\frac{t}{e^t - 1} - B_1 t &= \frac{t}{e^t - 1} + \frac{t}{2} \\ &= \frac{2t + t(e^t - 1)}{2(e^t - 1)} \\ &= \frac{t(e^t + 1)}{2(e^t - 1)}\end{aligned}$$

and thus after expansion with $e^{-t/2}$

$$\frac{t}{e^t - 1} - B_1 t = \frac{t}{2} \frac{e^{t/2} + e^{-t/2}}{e^{t/2} - e^{-t/2}}. \tag{2.10}$$

Finally, the following direct computation shows that $g(t) := \frac{t}{e^t-1} - B_1 t$ is an even function:

$$g(-t) = \frac{-t}{2} \frac{e^{-t/2} + e^{t/2}}{e^{-t/2} - e^{t/2}} = \frac{t}{2} \frac{e^{t/2} + e^{-t/2}}{e^{t/2} - e^{-t/2}} = g(t).$$

Thus, we have shown $B_{2n+1} = 0$ for $n \geq 1$. □

We shall also need the Bernoulli polynomials.

Definition 2.6.3. The n-th Bernoulli polynomial $B_n(X)$ for $n \in \mathbb{Z}_{\geq 0}$ is defined by

$$\frac{te^{Xt}}{e^t - 1} = \sum_{n=1}^{\infty} B_n(X) \frac{t^n}{n!}.$$

Here, we view $e^{Xt} = \sum_{k=0}^{\infty} \frac{(Xt)^k}{k!}$ as an element in the power series ring $R[\![t]\!]$ over the polynomial ring $R = \mathbb{Q}[X]$.

Of course, the Bernoulli polynomials evaluated at zero give the Bernoulli numbers, i.e., $B_n = B_n(0)$. The following Lemma gives a more general connection between the Bernoulli numbers and the Bernoulli polynomials.

Lemma 2.6.4. *We have*

$$B_n(X) = \sum_{i=0}^{n} \binom{n}{i} B_i X^{n-i}$$

and

$$B_n(1-X) = (-1)^n B_n(X).$$

Proof. This follows since the generating function $\frac{te^{Xt}}{e^t-1}$ of the Bernoulli polynomials is the product of

$$\frac{t}{e^t-1} = \sum_{i=0}^{\infty} B_i \frac{t^i}{i!}$$

and

$$e^{Xt} = \sum_{j=0}^{\infty} X^j \frac{t^j}{j!}.$$

The proof of the second equation follows from comparing the coefficients of

$$\sum_{n=1}^{\infty} B_n(1-X)\frac{t^n}{n!} = \frac{te^{(1-X)t}}{e^t-1} = \frac{te^{-Xt}}{1-e^{-t}} = \frac{(-t)e^{X(-t)}}{e^{-t}-1} = \sum_{n=1}^{\infty} B_n(X)\frac{(-t)^n}{n!}.$$

□

2.6.2 *Values of the Hurwitz zeta function*

In the following, we will compute the values of the Riemann zeta function at the negative integers. More generally, we will compute the values of the *Hurwitz zeta function*, which is defined for $0 < x \leq 1$ and $s \in \mathbb{C}$ with $\mathrm{Re}(s) > 1$ by

$$\zeta(s,x) := \sum_{n=0}^{\infty} \frac{1}{(n+x)^s}.$$

Note, that the Riemann zeta function is a special case of the Hurwitz zeta function:

$$\zeta(s) = \zeta(s,1).$$

We already know that the Riemann zeta function admits a meromorphic continuation to $\mathbb{C}$. The following result generalizes this to the Hurwitz zeta function:

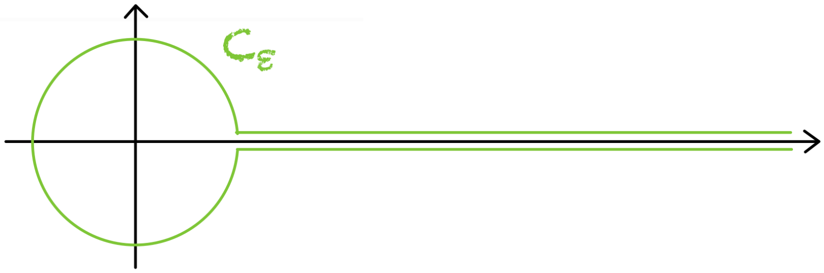
Theorem 2.6.5. *For $0 < x \leq 1$, the Hurwitz zeta function admits a meromorphic continuation to all of $\mathbb{C}$ with a simple pole at $s = 1$ of residue 1 and satisfies*

$$\zeta(1-n,x) = -\frac{B_n(x)}{n} \quad \textit{for } n \in \mathbb{N}.$$

Proof. Set $F(t) := \frac{te^{(1-x)t}}{e^t-1}$ and consider the integral

$$H(s) := \int F(z) z^{s-2} dz,$$

where the integral is over the following path



which consist of the positive real axis (top side), a circle C_ϵ around 0 of radius ϵ, and the positive real axis (bottom side). We define $z^s := \exp(s\log(z))$, where we take the branch of the logarithm with branch cut along $\mathbb{R}_{>0}$ which is given by $\log t$ on the top side of the real axis and by $\log(t) + 2\pi i$ on the bottom side. Then, the integral defining $H(s)$ converges absolutely and locally uniformly for all $s \in \mathbb{C}$. Hence, $H(s)$ is a holomorphic function defined on the whole complex plane. By our choice of the branch of the logarithm, we may write

$$\begin{aligned} H(s) = \int F(z) z^{s-2} dz = -\int_\epsilon^\infty F(t)\exp((s-2)\log t) dz + \int_{C_\epsilon} F(z) z^{s-2} dz \\ + \int_\epsilon^\infty F(t)\exp((s-2)(\log t + 2\pi i)) dt \\ =(e^{2\pi i s}-1)\int_\epsilon^\infty F(t) t^{s-2} dt + \int_{C_\epsilon} F(z) z^{s-2} dz. \end{aligned}$$

For a moment, let us assume $\mathrm{Re}(s) > 1$. Then $\int_{C_\epsilon} \to 0$ as $\epsilon \to 0$, so

$$\begin{aligned} H(s) &= (e^{2\pi i s}-1)\int_0^\infty F(t) t^{s-2} dt \\ &= (e^{2\pi i s}-1)\int_0^\infty t^{s-1}\frac{e^{(1-x)t}}{e^t-1} dt \\ &= (e^{2\pi i s}-1)\int_0^\infty t^{s-1}\sum_{m=0}^\infty e^{-(x+m)t} dt \\ &= (e^{2\pi i s}-1)\sum_{m=0}^\infty \frac{1}{(x+m)^s}\Gamma(s) \\ &= (e^{2\pi i s}-1)\zeta(s,x)\Gamma(s). \end{aligned}$$

In particular, we get for $s \in \mathbb{C}$ with $\mathrm{Re}(s) > 1$

$$\zeta(s,x) = \frac{H(s)}{(e^{2\pi i s} - 1)\Gamma(s)}. \tag{2.11}$$

The right hand side of this equation is a holomorphic function on $\mathbb{C} \setminus \{1\}$ with a simple pole at $s = 1$. In particular, this equation provides a meromorphic continuation of $\zeta(s,x)$. Let us now assume that $s = 1 - n$ is an integer. Then, by Cauchy's integral theorem and Lemma 2.6.4, we get[36]

$$\begin{aligned} H(1-n) &= \int_{C_\epsilon} F(z) z^{-n-1} dz \\ &= (2\pi i)\frac{B_n(1-x)}{n!} = (2\pi i)(-1)^n \frac{B_n(x)}{n!}. \end{aligned} \tag{2.12}$$

We have already seen that the Gamma function has simple poles at all non-positive integers with

$$\mathrm{Res}_{s=1-n}\Gamma(s) = \frac{(-1)^{n-1}}{(n-1)!}.$$

We deduce

$$\begin{aligned} \lim_{s \to 1-n} (e^{2\pi i s} - 1)\Gamma(s) &= \lim_{s \to 1-n} (2\pi i(s-(1-n)) + O((s-(1-n))^2))\Gamma(s) \\ &= 2\pi i \cdot \mathrm{Res}_{s=1-n}\Gamma(s) = \frac{2\pi i(-1)^{n-1}}{(n-1)!}. \end{aligned} \tag{2.13}$$

Combining (2.11), (2.12) and (2.13) gives the desired equality

$$\zeta(1-n,x) = -\frac{B_n(x)}{n}.$$

□

Putting $x = 1$ in the above formula for the Hurwitz zeta values gives.

Corollary 2.6.6. *For $n \geq 1$, we have*

$$\zeta(1-n) = -(-1)^n \frac{B_n}{n}.$$

In particular, we have $\zeta(0) = -1/2$ and for any integer $n \geq 2$

$$\zeta(1-n) = -\frac{B_n}{n}.$$

Proof. For $x = 1$, we get

$$\zeta(1-n) = \zeta(1-n,1) = -\frac{B_n(1)}{n} = -(-1)^n \frac{B_n}{n}.$$

Since $B_1 = -1/2$ and $B_{2n+1} = 0$ for $n \geq 1$, we deduce the second claim. □

[36] For the integer $s = 1 - n$, the two integrals from ∞ to ϵ and from ϵ to ∞ cancel, so only the integral over C_ϵ survives.

2.6.3 *Euler's formula*

By combining the explicit formula for the negative zeta values with the functional equation, we deduce Euler's formula.

Corollary 2.6.7 (Euler's formula)**.** *For $n \geq 1$, we have*

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

Proof. Let us first express the Riemann zeta function in terms of the completed Riemann zeta function and apply the functional equation for ξ to get

$$\begin{aligned}\zeta(2n) = \frac{\pi^n}{\Gamma(n)}\xi(1-2n) &= \frac{\pi^n}{\Gamma(n)}\pi^{-\frac{1-2n}{2}}\Gamma\left(\frac{1-2n}{2}\right)\zeta(1-2n)\\ &= \frac{\pi^{2n}}{\sqrt{\pi}\Gamma(n)}\Gamma\left(\frac{1-2n}{2}\right)\zeta(1-2n).\end{aligned}$$

So, let us first compute the Gamma values using the functional equation $\Gamma(z+1) = z\Gamma(z)$ and Corollary 2.1.9

$$\Gamma\left(\frac{1-2n}{2}\right) = \frac{\Gamma(1/2)}{\prod_{k=1}^{n}\left(\frac{1}{2}-k\right)} = (-1)^n 2^n \frac{\sqrt{\pi}}{\prod_{k=1}^{n}(2k-1)},$$

and

$$\frac{1}{\Gamma(n)} = \frac{1}{\prod_{k=1}^{n-1}k} = \frac{2^{n-1}}{\prod_{k=1}^{n-1}2k}.$$

Thus, we obtain

$$\begin{aligned}\zeta(2n) = \pi^{2n}(-1)^n 2^{2n-1}&\frac{1}{\prod_{k=1}^{n-1}2k\prod_{k=1}^{n}(2k-1)}\zeta(1-2n)\\ &= (-1)^n\frac{(2\pi)^{2n}}{2(2n-1)!}\zeta(1-2n).\end{aligned}$$

Finally, we conclude using the explicit formula for $\zeta(1-2n) = -\frac{B_{2n}}{2n}$

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

□

Outlook

Euler's formula for the values of the Riemann zeta function at the positive even integers says that $\zeta(2n)$ is a rational number times $(2\pi i)^n$ and leads immediately to the following two questions:

(a) What can be said about the odd values of the Riemann zeta function?

(b) Is the appearance of $(2\pi i)^n$ in Euler's formula a coincidence?

The first question concerns the values of the Riemann zeta function at the odd positive integers, while the second question is about the even zeta values. This already indicates that the values of the Riemann zeta function at the integers fall naturally into two classes, depending on the parity of the integer in the argument. Let us explain, why it is much more difficult to say something about the values of the Riemann zeta function at the odd positive integers. Let us start recalling how we were able to prove Euler's formula. We started with the formula $\zeta(1-n) = -\frac{B_n}{n}$ for the Riemann zeta function at the negative integers. This formula does indeed hold for all integers $n \geq 0$ independent of the parity of n. Then, we used the functional equation

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

to relate the value at $s = n$ to the value at $s = 1-n$. The point is now the following. If n is even, the Gamma function $\Gamma\left(\frac{1-s}{2}\right)$ is defined at $s = n$, so we can divide by this Γ-value and compute $\zeta(n)$. On the other hand, if $n \geq 2$ is odd then $\frac{1-n}{2}$ is a negative integer and the Γ-value $\Gamma\left(\frac{1-s}{2}\right)$ is not defined[37] at $s = n$. Thus, the poles of the Gamma function are the reason for the different nature of the even respectively odd zeta values. This lead Deligne to the following definition[38].

Definition. The value $\zeta(n)$ of the Riemann zeta function at an integer n is called *critical* if neither of the Gamma factors $\Gamma\left(\frac{s}{2}\right)$ and $\Gamma\left(\frac{1-s}{2}\right)$ has a pole at $s = n$. Otherwise, we call the value *non-critical*.

Thus, the *critical* values of the Riemann zeta function are exactly the values corresponding to positive even integers $\zeta(2n)$ and the corresponding values on the other side of the functional equation $\zeta(1-2n)$. Deligne formulated a very general and deep conjecture for all critical values of 'motivic' L-functions. Roughly, the expectation is that critical L-values are always algebraic up to powers of certain explicit periods[39]. In the case of the Riemann zeta function this conjecture predicts[40] that $\zeta(2n)$ is algebraic up to $(2\pi i)^{2n}$, so Euler's formula can be seen as a 'proof' for the Deligne conjecture for the Riemann zeta function. For more general L-functions, this conjecture is vastly open. Nevertheless, the Deligne conjecture can be seen as a possible answer to the second question at the beginning of this section. At the end of the next section, we will say something about the non-critical zeta values, i.e., the odd zeta values $\zeta(3), \zeta(5), \ldots$.

[37] Recall that the Gamma function has poles at all non-positive integers.

[38] He defined it much more generally, for all values of a motivic L-function $L(M,s)$ at an integer n.

[39] Here, period is meant in the sense of footnote [11]

[40] To realize $2\pi i$ as a period, we can use Cauchy's integral formula

$$2\pi i = \int_\gamma \frac{dt}{t},$$

Here, $\gamma \in H_1(\mathbb{G}_m(\mathbb{C}), \mathbb{Z})$ is the generator of the first homology of $\mathbb{G}_m(\mathbb{C}) = \mathbb{C}^\times$ which is a counter-clockwise circle around $0 \in \mathbb{C}$. Furthermore, observe that the differential form $\frac{dt}{t}$ is an algebraic Kähler differential $\frac{dt}{t} \in \Gamma(\mathbb{G}_m, \Omega^1_{\mathbb{G}_m})$.

2.7 A proof that Euler missed

Euler's formula raises immediately the question about the values of the values of the Riemann zeta function at the odd positive integers. These values are much more mysterious and purely understood. It is not expected that there is a closed and explicit formula for the odd zeta values in terms of basic mathematical constants such as Euler's formula. Thus, it is a natural question if one can say something about the structure of these values, e.g., if they are rational, irrational, algebraic, or transcendental.

For quite a long time, nothing has been known about the structure of the odd zeta values. So, it was a mathematical sensation when Apéry succeeded to prove the irrationality of $\zeta(3)$, in 1979. The following citation is attributed to the famous mathematician Carl Ludwig Siegel:

> „Man kann diesen Beweis nur wie einen Kristall vor sich her tragen.“

It is very surprising that Apéry's proof has not been found earlier, since it is quite elementary. This is summarized in a very concise way in the following quotation of the mathematician van der Poorten, who said about Apéry's proof:

> „A proof that Euler missed...“

Up to today, $\zeta(3)$ is the only particular of the infinite many numbers

$$\zeta(2n+1) \quad \text{für } n \geq 1$$

for which we can prove its irrationality. Nevertheless, there are certain asymptotic results about the irrationality of odd zeta values. In this lecture, we will present a nice version of Apéry's proof which goes back to Frits Beukers. When we established the Chebyshev bounds for the function $\pi(x)$, we have already seen the sequence

$$d_n := \mathrm{lcm}(1, \dots, n)$$

and deduced upper bounds for it. Equipped with the full strength of the Prime Number Theorem, we can improve this bound.

Proposition 2.7.1. *For all sufficiently large integers n, we have*[41]

$$d_n = \mathrm{lcm}(1, \dots, n) < 3^n.$$

Proof. Exercise. □

Before we proceed with the details of Apery's proof, let us outline the strategy. We will construct two integer sequences $(A_n)_{n\geq 0}$ and $(B_n)_{n\geq 0}$ with the following property:[42]

[41] With slightly more effort, one can prove $d_n \sim e^n$ as $n \to \infty$.

[42] We can think about this inequality as follows: The sequence $(\frac{A_n}{B_n})_n$ of rational numbers provides a very good rational approximation of $\zeta(3)$. It is a general principle in transcendental number theory that a real number is irrational or transcendental if it has very good approximations by rational numbers.

$$0 < |A_n + B_n\zeta(3)| \to 0 \text{ für } n \to \infty.$$

Let us now assume that $\zeta(3) = \frac{a}{b}$ was a rational number. Then

$$(|bA_n + aB_n|)_{n\geq 0}$$

is a non-zero sequence of integers which tends to zero as $n \to \infty$. This is an obvious contradiction.

For the construction of the sequences $(A_n)_{n\geq 0}$ and $(B_n)_{n\geq 0}$ we will need a few auxiliary results:

Proposition 2.7.2. *For $r, s \in \mathbb{N}_0$ let us consider the integral*[43]

$$I_{r,s} := \int_0^1 \int_0^1 -\frac{\log(xy)}{1-xy} x^r y^s dxdy.$$

(a) For $r > s$ we get

$$d_r^3 \cdot I_{r,s} \in \mathbb{Z}.$$

Here, we recall $d_r = \mathrm{lcm}(1, \dots, r)$.

(b) For $r = s$ we have

$$I_{r,r} = 2\left(\zeta(3) - \sum_{k=1}^{r} \frac{1}{k^3}\right).$$

In particular, $d_r^3 I_{r,r} \in \mathbb{Z} + \zeta(3)\mathbb{Z}$.

Proof. By partial integration, we obtain for $k \geq 0$:

$$\begin{aligned}\int_0^1 \log(x) x^{r+k} dx &= \lim_{\epsilon\to 0} \int_\epsilon^1 \log(x) x^{r+k} dx \\ &= \frac{1}{r+k+1} \lim_{\epsilon\to 0} \left([x^{r+k+1}\log(x)]_\epsilon^1 - \int_\epsilon^1 \frac{1}{x} \cdot x^{r+k+1} dx \right) \\ &= \frac{-1}{(r+k+1)^2}.\end{aligned} \tag{2.14}$$

Using (2.14) and the geometric series gives

$$\begin{aligned} I_{r,s} &= \int_0^1 \int_0^1 -\frac{\log(xy)}{1-xy} x^r y^s dxdy \\ &= -\int_0^1 \left(\sum_{k=0}^{\infty} \int_0^1 \log(xy) x^{r+k} y^{s+k} dx \right) dy \\ &= -\int_0^1 \left(\sum_{k=0}^{\infty} \log(y) y^{s+k} \int_0^1 x^{r+k} dx + y^{s+k} \int_0^1 \log(x) x^{r+k} dx \right) dy \\ &= -\sum_{k=0}^{\infty} \int_0^1 \left(\frac{y^{s+k}\log y}{r+k+1} - \frac{y^{s+k}}{(r+k+1)^2} \right) dy. \end{aligned}$$

By applying (2.14) to the first of the two integrals, we obtain

$$I_{r,s} = \sum_{k=0}^{\infty} \left(\frac{1}{(r+k+1)(s+k+1)^2} + \frac{1}{(r+k+1)^2(s+k+1)} \right). \tag{2.15}$$

[43] Let us observe that $I_{r,s} = I_{s,r}$, so we may without loss of generality assume that $r \geq s$.

a) For $r > s$ we can rewrite the summand of (2.15) as follows

$$\begin{aligned}
&\frac{1}{(r+k+1)(s+k+1)^2} + \frac{1}{(r+k+1)^2(s+k+1)} \\
&= \frac{1}{(r+k+1)(s+k+1)} \left(\frac{1}{s+k+1} + \frac{1}{r+k+1} \right) \\
&= \frac{1}{r-s} \frac{(r+k+1)-(s+k+1)}{(r+k+1)(s+k+1)} \left(\frac{1}{s+k+1} + \frac{1}{r+k+1} \right) \\
&= \frac{1}{r-s} \left(\frac{1}{s+k+1} - \frac{1}{r+k+1} \right) \left(\frac{1}{s+k+1} + \frac{1}{r+k+1} \right) \\
&= \frac{1}{r-s} \left(\frac{1}{(s+k+1)^2} - \frac{1}{(r+k+1)^2} \right).
\end{aligned}$$

Substituting this into (2.15) yields

$$I_{r,s} = \sum_{k=0}^{\infty} \frac{1}{r-s} \left(\frac{1}{(s+k+1)^2} - \frac{1}{(r+k+1)^2} \right) = \frac{1}{r-s} \sum_{k=1}^{r-s} \frac{1}{(s+k)^2}.$$

For $1 \leq k \leq r-s$ the integer $s+k$ is contained in the set $\{1, \ldots, r\}$. Thus, $\mathrm{lcm}(1, \ldots, r)$ is a multiple of $s+k$ and we obtain

$$d_r^2 \frac{1}{(s+k)^2} \in \mathbb{Z}.$$

Since $1 \leq r-s \leq r$, we also get

$$d_r \frac{1}{r-s} \in \mathbb{Z}.$$

Altogether, we find

$$d_r^3 I_{r,s} = \frac{d_r}{r-s} \sum_{k=1}^{r-s} \frac{d_r^2}{(s+k)^2} \in \mathbb{Z}.$$

(b) For $r = s$ we can write (2.15) as

$$\begin{aligned}
I_{r,r} &= 2 \sum_{k=0}^{\infty} \left(\frac{1}{(r+k+1)^3} \right) \\
&= 2 \left(\sum_{k=1}^{\infty} \frac{1}{k^3} - \sum_{k=1}^{r} \frac{1}{k^3} \right) = 2 \left(\zeta(3) - \sum_{k=1}^{r} \frac{1}{k^3} \right).
\end{aligned}$$

Because k^3 is a divisor of d_r^3, the claim about the integrality of the linear combination follows. □

For a positive integer n, we define the polynomial $Q_n := X^n(1-X)^n$ and set

$$P_n := \frac{1}{n!} Q_n^{(n)}.$$

Here, we write $Q_n^{(n)}$ for the polynomial obtained by taking the n-th formal derivative of the polynomial Q_n. Note, that P_n is a polynomial of degree n with integral coefficients[44].

[44] It is easily checked that the coefficients of n-th formal derivative of a polynomial in $\mathbb{Z}[X]$ are divisible by $n!$.

Corollary 2.7.3. *There are sequences $(A_n)_{n\in\mathbb{N}}$ and $(B_n)_{n\in\mathbb{N}}$ of integers with the property*

$$d_n^3 \int_0^1 \int_0^1 -\frac{\log(xy)}{1-xy} P_n(x)P_n(y)dxdx = A_n + B_n\zeta(3).$$

Proof. Let us define the integers $a_{r,s}$ for $0 \le r,s \le n$ as coefficients of the polynomial $P_n(X)P_n(Y) \in \mathbb{Z}[X,Y]$:

$$P_n(X)P_n(Y) = \sum_{r=0}^{n}\sum_{s=0}^{n} a_{r,s}X^rY^s.$$

We set

$$B_n := d_n^3 \cdot 2 \cdot (a_{0,0} + a_{1,1} + \cdots + a_{n,n})$$

and

$$A_n := d_n^3 \left(\sum_{0\le s<r\le n} 2a_{r,s}I_{r,s} \right) - 2\sum_{r=1}^{n} a_{r,r} \sum_{k=1}^{r} \frac{d_n^3}{k^3}.$$

Let us observe that A_n and B_n are integers. Proposition 2.7.2 implies:

$$d_n^3 \int_0^1 \int_0^1 -\frac{\log(xy)}{1-xy} P_n(x)P_n(y)dxdx = A_n + B_n\zeta(3).$$

□

We are now ready to prove the irrationality of $\zeta(3)$:

Theorem 2.7.4 (Apéry). *$\zeta(3)$ is irrational.*

Proof. We have already constructed integers A_n and B_n for each $n \in \mathbb{N}$ such that

$$d_n^3 \int_0^1 \int_0^1 -\frac{\log(xy)}{1-xy} P_n(x)P_n(y)dxdx = A_n + B_n\zeta(3).$$

Our next aim is to prove that

$$0 < |A_n + B_n\zeta(3)| \to 0 \text{ mit } n \to \infty.$$

In order to estimate the integral, the following observation will be useful:

Claim: We have

$$\int_0^1 \int_0^1 -\frac{\log(xy)}{1-xy} P_n(x)P_n(y)dxdy = \int_0^1 \int_0^1 \int_0^1 \frac{Q_n(u)Q_n(v)Q_n(w)}{(1-(1-uv)w)^{n+1}}$$

with $Q_n(x) = x^n(1-x)^n$.
Proof of the claim: Exercise
Next, we want to establish an estimate for the integral

$$\int_0^1 \int_0^1 \int_0^1 \frac{Q_n(u)Q_n(v)Q_n(w)}{(1-(1-uv)w)^{n+1}} dudvdw. \tag{2.16}$$

Since the integrand is strictly positive on the interior of the cube $[0,1]^3$, we obtain

$$0 < \int_0^1 \int_0^1 \int_0^1 \frac{Q_n(u)Q_n(v)Q_n(w)}{(1-(1-uv)w)^{n+1}} dudvdw.$$

On the other hand, a straightforward argument shows

$$\frac{u(1-u)v(1-v)w(1-w)}{1-(1-uv)w} \leq (\sqrt{2}-1)^4 \text{ für } 0 \leq u,v,w \leq 1.$$

This proves an upper estimate for the intergal (2.16):

$$\begin{aligned}\int_0^1 \int_0^1 \int_0^1 & \frac{Q_n(u)Q_n(v)Q_n(w)}{(1-(1-uv)w)^{n+1}} dudvdw \\ & \leq (\sqrt{2}-1)^{4n} \int_0^1 \int_0^1 \int_0^1 \frac{1}{(1-(1-uv)w)} dudvdw \\ & = (\sqrt{2}-1)^{4n} \int_0^1 \int_0^1 -\frac{\log(uv)}{1-uv} dudv = (\sqrt{2}-1)^{4n} 2\zeta(3).\end{aligned}$$

Combining Corollary 2.7.3 with the above claim gives

$$0 < |A_n + B_n\zeta(3)| \leq d_n^3(\sqrt{2}-1)^{4n} 2\zeta(3).$$

Proposition 2.7.1 implies, using $3^3 \cdot (\sqrt{2}-1)^4 \approx 0,79 < 1$, that the right hand side converges for $n \to \infty$ to 0. If $\zeta(3) = \frac{a}{b}$ was rational, we would get

$$0 < |bA_n + B_n a| \to 0 \text{ für } n \to \infty$$

a non-zero sequence of integers which converges to zero – a contradiction. □

Outlook

In the following, let us indicate what is known (and what is not known) about the odd zeta values. A first naive guess, having Euler's formula in mind, would be that $\zeta(n)/(2\pi i)^n$ is always rational. But indeed this is not expected to be true. Grothendieck's period conjecture, a deep conjecture in arithmetic geometry, would imply that

$$2\pi i, \zeta(3), \zeta(5), \ldots, \zeta(2n+1), \ldots$$

are algebraically independent. In particular, we would not expect a similar formula as Euler's formula for the odd zeta values. At the same time, Grothendieck's Period Conjecture would imply the transcendence of all odd zeta values. Unfortunately, we are even far away from proving the transcendence of a single particular odd zeta value. Indeed, even worse, we do not know the irrationality of a single

$\zeta(2n+1)$ for $n \geq 2$. Nevertheless, Wadim Zudilin succeeded to prove that at least one of the numbers

$$\zeta(5), \zeta(7), \zeta(9), \zeta(11)$$

is irrational. A major breakthrough has been made by Rivoal and Ball-Rivoal in 2001. They were able to prove that infinitely many odd zeta values are irrational. More precisely, they succeeded to prove:

Theorem (Rivoal, Ball–Rivoal, 2001)**.** *For $\epsilon > 0$ and all sufficiently large odd integers s, the $\mathbb{Q}$-vector space spanned by*

$$\zeta(3), \zeta(5), \ldots, \zeta(s)$$

has at least dimension $\frac{(1-\epsilon)}{1+\sqrt{2}} \log s$. In particular, there are at least $\frac{(1-\epsilon)}{1+\sqrt{2}} \log s$ irrational numbers in the list

$$\zeta(3), \zeta(5), \ldots, \zeta(s).$$

This already proves that there are infinitely many odd zeta values. On the other hand, $\log s$ is a very slowly growing function. In a joint work with Stefan Fischler and Wadim Zudilin, we were able to improve this lower bound considerably:

Theorem (Fischler–S.–Zudilin, 2018)**.** *For $\epsilon > 0$ and s sufficiently large, at least*

$$2^{(1-\epsilon)\frac{\log s}{\log \log s}}$$

of the numbers

$$\zeta(3), \zeta(5), \ldots, \zeta(s),$$

are irrational.

Building on our method, this has been further improved recently by Li Lai and Pin Yu:

Theorem (Lai-Yu,2020)**.** *For $\epsilon > 0$ and s sufficiently large, at least*

$$(c_0 - \epsilon)\sqrt{\frac{s}{\log s}}, \quad c_0 \approx 1.192507..$$

of the numbers

$$\zeta(3), \zeta(5), \ldots, \zeta(s),$$

are irrational.

Surprisingly, the above results predict that there are quite many irrational odd zeta values, but the only particular odd zeta value for which we can prove irrationality remains $\zeta(3)$.

3 The Kronecker-Weber Theorem

In this chapter, we will classify all abelian extensions of $\mathbb{Q}$, i.e. Galois extensions of $\mathbb{Q}$ with abelian Galois group. More precisely, the Theorem of Kronecker-Weber says that every abelian extension of $\mathbb{Q}$ is contained in a cyclotomic extension, i.e., in some field $\mathbb{Q}(\zeta_n)$ for a primitive n-th root of unity. This result can be seen as a very explicit version of global class field theory for the ground field $\mathbb{Q}$. For the proof, we will introduce local fields and first classify all abelian extensions of $\mathbb{Q}_p$ for all primes p, where $\mathbb{Q}_p$ is the field of p-adic numbers. Finally, we will deduce the Kronecker-Weber Theorem from the corresponding local statement.

3.1 Absolute values and valuation rings

In this section, we define absolute values, valuations and discuss their relation to discrete valuation rings.

Definition 3.1.1. An *absolute value* on a field k is a map $|\cdot|\colon k \to \mathbb{R}_{\geq 0}$ such that for all $x, y \in k$:

(a) $|x| \geq 0$ and ($|x| = 0 \Leftrightarrow x = 0$),

(b) $|xy| = |x||y|$

(c) $|x+y| \leq |x| + |y|$.

The absolute value is called *non-Archimedean* if and only if it satisfies the strict triangle inequality $|x+y| \leq \max(|x|, |y|)$, otherwise it will be called *Archimedean*. The pair $(k, |\cdot|)$ will be called a *valued field*. An absolute value $|\cdot|$ on k defines a metric $d(x,y) := |x-y|$ on k and hence a topology[1]. Two absolute values are called *equivalent* if the define they same topology on k.

Of course, we have the following trivial absolute value on any field k:

Example 3.1.2. For a given field k, the map $|\cdot|\colon k \to \mathbb{R}_{\geq 0}$ defined by

$$|0| := 0, \quad |x| := 1 \text{ for all } x \in k^\times$$

[1] A basis of open neighbourhoods for this topology is given by the open balls

$$B_\epsilon(\alpha) := \{x \in k \mid |\alpha - x| < \epsilon\}.$$

So, a subset $U \in k$ is open if and only if for every $x \in U$ there exists an open ball around x which is contained in U

is an absolute value. It is called the *trivial* absolute value.

In the following, we will usually ignore the trivial absolute value. Of course, a *non-trivial* absolute value is an absolute value which is not the *trivial* absolute value. Let us give the following important examples for absolute values on $\mathbb{Q}$.

Example 3.1.3. The usual absolute value $|\cdot|$ defines an absolute value on $\mathbb{Q}$. Furthermore, for each prime $p \in \mathbb{Z}$, the map

$$|\cdot|_p \colon \mathbb{Q} \to \mathbb{R}, \quad x \mapsto |x|_p := p^{-v_p(x)},$$

is an example for a non-Archimedean absolute value. Here, v_p is defined by $v_p(x) = n$ if $x = p^n \frac{a}{b}$ with $n \in \mathbb{Z}$ and integers a and b which are co-prime to p.

The following criterion is useful if one wants to decide if two absolute values are equivalent:

Proposition 3.1.4. *For two non-trivial absolute values $|\cdot|_1$ and $|\cdot|_2$ on a field k, the following are equivalent:*

(a) $|\cdot|_1$ is equivalent to $|\cdot|_2$,

(b) For all $x \in k$, we have $|x|_1 < 1 \Rightarrow |x|_2 < 1$,

(c) There exists a positive real number s such that for all $x \in k$: $|x|_1 = |x|_2^s$.

Proof. $(a) \Rightarrow (b)$ Let us assume that $|\cdot|_1$ and $|\cdot|_2$ define the same topology. For an absolute value $|\cdot|$ on k and $x \in k$, we have

$$|x| < 1 \Leftrightarrow (|x^k|)_k \text{ is a zero-sequence in the topology defined by } |\cdot|.$$

Thus, for all $x \in k$, we have $|x|_1 < 1 \Leftrightarrow |x|_2 < 1$. This proves $(a) \Rightarrow (b)$. $(b) \Rightarrow (c)$ Let us fix[2] $y \in k$ with $|y|_1 > 1$. Let us define $s := \frac{\log |y|_1}{\log |y|_2}$. We want to show that any $x \in k^\times$ satisfies[3]

$$|x|_1 = |x|_2^s.$$

For each $x \in k^\times$, we can find a real number α with $|x|_1 = |y|_1^\alpha$. Let us choose a sequence $\left(\frac{m_i}{n_i}\right)_i$ of rational numbers with $m_i \in \mathbb{Z}$, $n_i \in \mathbb{N}$ which converges from above to α, i.e., $\frac{m_i}{n_i} \searrow \alpha$ as $i \to \infty$. Then, we have for all $i \geq 0$

$$|x|_1 = |y|_1^\alpha < |y|_1^{m_i/n_i}.$$

This implies

$$\left|\frac{x^{n_i}}{y^{m_i}}\right|_1 < 1 \quad \text{for all } i \geq 0.$$

Now (b) yields

$$\left|\frac{x^{n_i}}{y^{m_i}}\right|_2 < 1 \quad \text{for all } i \geq 0,$$

[2] Such a y exists, because $|\cdot|_1$ is non-trivial. Indeed, for any $z \in k$ with $|z| \neq 0, 1$ we have either $|z| > 1$ or $|z^{-1}| > 1$

[3] Of course, the statement of (c) for $x = 0$ holds trivially. So, we may assume $x \in k^\times$.

which gives for all $i \geq 0$

$$|x|_2 < |y|_2^{m_i/n_i}.$$

Passing to the limit gives $|x|_2 \leq |y|_2^{\alpha}$. Analogously, by choosing a sequence of rationals which converges to α form below, we deduce $|x|_2 \geq |y|_2^{\alpha}$. Thus, we have shown that

$$|x|_2 = |y|_2^{\alpha}.$$

Since $x \in k^{\times}$ was arbitrary, (c) follows from the following computation

$$|x|_1 = |y|_1^{\alpha} = |y|_2^{\alpha \frac{\log|y|_1}{\log|y|_2}} = |x|_2^{\frac{\log|y|_1}{\log|y|_2}} = |x|_2^{s}.$$

$(c) \Rightarrow (a)$ For $\epsilon > 0$, $x \in k$ and $i \in \{0,1\}$, let us consider the following ball

$$B^i_\epsilon(x) := \{y \in k \mid |x-y|_i < \epsilon\}.$$

Then (c) implies $B^1_\epsilon(x) \subseteq B^2_{\epsilon^{1/s}}(x)$ and $B^2_\epsilon(x) \subseteq B^1_{\epsilon^s}(x)$. Since $\{B^i_\epsilon(x) \mid x \in k, \epsilon > 0\}$ is a basis for the topology induced by $|\cdot|_i$, we get (a). □

Sometimes, it is more convenient to work with the following additive version of a non-Archimedean absolute value:

Definition 3.1.5. An (additive) valuation on a field k is a map

$$v\colon k \to \mathbb{R} \cup \{\infty\}$$

satsifying

(a) $v(x) = \infty \Leftrightarrow x = 0$,

(b) $v(xy) = v(x) + v(y)$, and

(c) $v(x+y) \geq \min(v(x), v(y))$.

The valuation is called *discrete*[4] if and only if $v(k^{\times}) = \frac{1}{t}\mathbb{Z}$ for some $t \in \mathbb{R}_{>0}$. Furthermore, it is called normalized if and only if $v(k^{\times}) = \mathbb{Z}$.

Of course, the datum of an additive valuation is equivalent to the datum of a non-Archimedean absolute value:

Lemma/Definition 3.1.6. For a non-Archimedean valued field $(k, |\cdot|)$, define $v(\cdot) := -\log|\cdot|$. This gives an additive valuation on k which will be called the *exponential valuation associated with* $|\cdot|$. Conversely, given an additive valuation v on a field k and a real number $q > 1$, $|\cdot| := q^{-v(\cdot)}$ is a non-Archimedean absolute value on k whose equivalence class does not depend on the choice of q. We will say that a valued field $(k, |\cdot|)$ is *discretely valued* if its associated exponential valuation is discrete.

[4] Note that v is discrete if and only if the subspace topology on $v(k^{\times}) \subseteq \mathbb{R}$ is discrete. This explains the terminology.

Proof. The only non-trivial claim, that different choices of q give equivalent absolute values, follows from Proposition 3.1.4. □

Next, let us turn our attention to examples of absolute values on number fields. Let us recall from 'Algebraic Number Theory I' that each prime ideal $\mathfrak{p}\mathcal{O}_K$ of a number field K induces a discrete valuation

$$v_{\mathfrak{p}} \colon K \to \mathbb{Z} \cup \{\infty\},$$

which is given by $v_{\mathfrak{p}}(0) = \infty$ and for non-zero x by the exponents in the unique prime-decomposition of the fractional ideal (x):

$$(x) = \prod_{\mathfrak{p}} \mathfrak{p}^{v_{\mathfrak{p}}(x)}.$$

The following definition gives examples of absolute values for a given number field K and generalizes the Example 3.1.3.

Definition 3.1.7. Let K be a number field with ring of integers $\mathcal{O}_K$.

(a) For a prime ideal $\mathfrak{p}$ of a number field K, let us define the *normalized absolute value*[5] by the following formula

$$|x|_{\mathfrak{p}} := (N\mathfrak{p})^{-v_{\mathfrak{p}}(x)},$$

where $N\mathfrak{p} := [\mathcal{O}_K : \mathfrak{p}]$ denotes the norm of $\mathfrak{p}$. The absolute value $|\cdot|_{\mathfrak{p}}$ is non-Archimedean.

(b) For a real or complex embedding σ, let us define the absolute value

$$|x|_{\sigma} := |\sigma(x)|,$$

where $|\cdot|$ is the usual absolute value on $\mathbb{R}$ respectively on $\mathbb{C}$.

For completeness, let us mention the following Theorem of Ostrowski which shows that the above definition gives all absolute values of a number field up to equivalence.

Theorem 3.1.8 (Ostrowski)**.** *Any non-trivial non-Archimedean absolute value of a number field K is equivalent to $|\cdot|_{\mathfrak{p}}$ for a prime ideal $\mathfrak{p}$. Any* Archimedean *absolute value of K is equivalent to $|\cdot|_{\sigma}$ coming from a real or complex embedding.*

Proof. We will prove this result for $K = \mathbb{Q}$ in the exercises. The general case is not much more complicated. □

Next, we want to study non-Archimedean absolute values from a more algebraic perspective. Let us recall that a *discrete valuation ring* is a principal ideal domain with exactly one non-zero maximal ideal. A generator π of the maximal ideal $\mathfrak{m}$ will be called *uniformizer*, i.e., we have $\mathfrak{m} = (\pi)$.

[5] This defines an absolute value by the definition of a 'discrete valuation'.

Proposition 3.1.9. *Let $(k, |\cdot|)$ be a non-Archimedean (non-trivially) valued field with associated exponential valuation v. The set*

$$A := \{x \in k \mid v(x) \geq 0\} = \{x \in k \mid |x| \leq 1\}$$

is a local sub-ring of k with maximal ideal

$$\mathfrak{m} := \{x \in k \mid v(x) > 0\} = \{x \in k \mid |x| < 1\},$$

and group of units

$$A^\times = \{x \in k \mid v(x) = 0\} = \{x \in k \mid |x| = 1\}.$$

If the valuation v is discrete then A is a discrete valuation ring.

Proof. From the property that $|\cdot|$ is non-Archimedean, it follows immediately that A is a sub-ring of k and that $\mathfrak{m}$ is an ideal in A. The fact that A is a local ring with maximal ideal $\mathfrak{m}$ follows[6] from the following claim:

$$A^\times = \{x \in k \mid v(x) = 0\}. \tag{3.1}$$

"$\subseteq$": For $a \in A^\times$ there exists $b \in A$ such that $a \cdot b = 1$. We deduce $0 = v(1) = v(a) + v(b)$ and together with $v(a), v(b) \geq 0$, we obtain $v(a) = 0$.
"$\supseteq$": Every $x \in k$ with $v(x) = 0$ is contained in $A \setminus \{0\}$. But the inverse $x^{-1} \in k$ is also contained in A, since $v(x^{-1}) = -v(x) = 0$. Thus, $x \in A^\times$. This finishes the proof of (3.1).
It remains to prove the last statement of the Proposition: Let us assume that v is discrete. We want to prove that A is a discrete valuation ring. Since v is discrete, we have $v(k^\times) = \frac{1}{t}\mathbb{Z}$ and $v' := t \cdot v$ is a normalized discrete valuation of k with

$$A = \{x \in k \mid v'(x) \geq 0\}.$$

Let $\pi \in A$ be an element with $v'(\pi) = 1$. Let $0 \neq \mathfrak{a} \subseteq A$ an ideal and set

$$n := \min\{v'(a) \mid a \in \mathfrak{a}\}.$$

Claim: $\mathfrak{a} = (\pi^n)$.
"$\subseteq$": For $x \in \mathfrak{a}$, we have $v'(x) \geq n$ by the definition of n. This implies $v'\left(\frac{x}{\pi^n}\right) \geq 0$, and we get $\frac{x}{\pi^n} \in A$. But this means $x \in (\pi^n)$.
"$\supseteq$": By the definition of n, there exists an element $x \in \mathfrak{a}$ with $v'(x) = n = v'(\pi^n)$. This implies $v'\left(\frac{x}{\pi^n}\right) = 0$ and hence $\frac{x}{\pi^n} \in A^\times$. We get $(\pi^n) = (x) \subseteq \mathfrak{a}$, as desired.
Since $\mathfrak{a}$ was an arbitrary non-zero ideal, we deduce that all ideals of A are of the form (0) or (π^n) for an integer $n \geq 0$. In particular, A is a PID with a unique maximal ideal (π). □

Corollary 3.1.10. *For an integral domain A, the following are equivalent:*

[6] Recall from Commutative Algebra that a ring A is local if and only if $A \setminus A^\times$ is an ideal.

(a) A is a discrete valuation ring,

(b) the quotient field k of A admits a discrete valuation v such that $A = \{x \in k \mid v(x) \geq 0\}$.

Proof. $(a) \Rightarrow (b)$ For a DVR A with maximal ideal $\mathfrak{m} = (\pi)$ define $v(0) := \infty$ and $v(x) := \max\{n \in \mathbb{Z} \mid x \in \pi^n A\}$. It is easily checked that this gives a discrete valuation on A.
$(b) \Rightarrow (a)$ This has been shown in Proposition 3.1.9. □

3.2 Completions

Recall from analysis that one defines the real numbers by completion of $\mathbb{Q}$ with respect to the usual absolute value on $\mathbb{Q}$. This procedure can be applied to any absolute value.

Definition 3.2.1. A valued field $(k, |\cdot|)$ is called *complete* if and only if all Cauchy sequences[7] have a limit in k.

The following Theorem shows that every valued field has a unique completion:

Theorem 3.2.2. *For every valued field $(k, |\cdot|)$ there is a complete valued field $(\widehat{k}, \widehat{|\cdot|})$ with the following properties:*

(a) $k \subseteq \widehat{k}$ and $\widehat{|\cdot|}$ extends $|\cdot|$,

(b) k is dense in $\widehat{k}$.

The completion is unique up to isomorphism, i.e., any other complete valued field satisfying (a) and (b) is isomorphic to $(\widehat{k}, \widehat{|\cdot|})$. The field $(\widehat{k}, \widehat{|\cdot|})$ is called the completion of $(k, |\cdot|)$.

Proof. We sketch the proof. We define

$$\mathcal{R} := \{(x_i)_{i\in\mathbb{N}} \mid (x_i)_{i\in\mathbb{N}} \text{ is a Cauchy sequence in } (k, |\cdot|)\}$$

as the set of Cauchy sequences in k. It is easily checked that point-wise addition and multiplication of sequences defines a ring-structure on $\mathcal{R}$ and that

$$\mathfrak{m} := \{(x_i)_{i\in\mathbb{N}} \mid (x_i)_{i\in\mathbb{N}} \text{ is a zero-sequence in } (k, |\cdot|)\}$$

is a maximal ideal of $\mathcal{R}$. We define $\widehat{k} := \mathcal{R}/\mathfrak{m}$. By mapping $x \in k$ to the constant sequence $(x)_{i\in\mathbb{N}}$, we get a map

$$k \to \mathcal{R} \twoheadrightarrow \widehat{k}.$$

The absolute value $\widehat{|\cdot|}$ on $\widehat{k}$ can be defined as follows

$$\widehat{|(x_i)_i \mod \mathfrak{m}|} := \lim_i |x_i|.$$

Now, it is not difficult to check that $(\widehat{k}, \widehat{|\cdot|})$ satisfies the desired properties. □

[7] A Cauchy sequence in a valued field $(k, |\cdot|)$ is a sequence $(x_i)_i \subseteq k$ such that for any $\epsilon > 0$ there exists a positive integer N such that for all $n, m \geq N$: $|x_n - x_m| < \epsilon$.

Definition 3.2.3. For a prime ideal $\mathfrak{p}$ of a number field K, we will write $(\widehat{K}_{\mathfrak{p}}, |\cdot|_{\mathfrak{p}})$ for the completion of $(K, |\cdot|_{\mathfrak{p}})$. The associated exponential valuation $v_{\mathfrak{p}}$ is discrete, and we will write $\mathcal{O}_{K,\mathfrak{p}}$ (or sometimes just $\mathcal{O}_{\mathfrak{p}}$) for the associated complete discrete valuation ring. In particular, we obtain for a rational prime p the *field of p-adic numbers* $(\mathbb{Q}_p, |\cdot|_p)$ as completion of $(\mathbb{Q}, |\cdot|_p)$. The discrete valuation ring of $\mathbb{Q}_p$ will be denoted by $\mathbb{Z}_p$ and will be called *ring of p-adic integers*.

Example 3.2.4. Let p be a prime and consider the discrete valuation ring $\mathbb{Z}_p$ in the p-adic completion $\mathbb{Q}_p$ of $\mathbb{Q}$ with respect $|\cdot|_p$. The intersection of $\mathbb{Q}$ with $\mathbb{Z}_p$ is

$$\mathbb{Z}_p \cap \mathbb{Q} = \{x \in \mathbb{Q} \mid v_p(x) \geq 0\} = \mathbb{Z}_{(p)}, \tag{3.2}$$

where we write $\mathbb{Z}_{(p)}$ for the localization of $\mathbb{Z}$ at the prime ideal (p). The equation (3.2) also shows that the localization $\mathbb{Z}_{(p)}$ is the discrete valuation ring of the valued field $(\mathbb{Q}, |\cdot|_p)$. The residue field of this discrete valuation ring is

$$\mathbb{Z}_{(p)}/p\mathbb{Z}_{(p)} \cong \mathbb{Z}/p\mathbb{Z}.$$

In the following, let us assume that $(k, |\cdot|)$ is a discretely valued field[8].

Proposition 3.2.5. *Let us assume that $(k, |\cdot|)$ is a discretely valued field with valuation ring A and maximal ideal $\mathfrak{m}$. Let $(\widehat{k}, \widehat{|\cdot|})$ be the completion of $(k, |\cdot|)$ with valuation ring $\widehat{A}$ and maximal ideal $\widehat{\mathfrak{m}}$. Then for $n \geq 1$, the subsets $A \subseteq \widehat{A}$ and $\mathfrak{m}^n \subseteq \widehat{\mathfrak{m}}^n$ are dense and induce an isomorphism*

$$A/\mathfrak{m}^n \xrightarrow{\sim} \widehat{A}/\widehat{\mathfrak{m}}^n. \tag{3.3}$$

Furthermore, we have $v(k^\times) = \widehat{v}(\widehat{k}^\times)$, where v and $\widehat{v}$ denote the exponential valuations of $|\cdot|$ and $\widehat{|\cdot|}$.

Proof. Let v be the exponential valuation of $|\cdot|$ on k. Without loss of generality, we may assume that v is normalized. By the definition of $(\widehat{k}, \widehat{|\cdot|})$, the additive valuation $\widehat{v}(x) = \lim_n v(x_n)$ for a Cauchy sequence $x = (x_n)_n$ is the exponential valuation associated to $\widehat{|\cdot|}$ on $\widehat{k}$. In particular, $\widehat{v}(\widehat{k}^\times)$ is contained in the closure of $v(k^\times) \subseteq \mathbb{R}$ in $\mathbb{R}$. Since $v(k^\times)$ is discrete, we get $v(k^\times) = \widehat{v}(\widehat{k}^\times)$. In particular, $\widehat{v}$ is a discrete valuation and the inclusions $A \subseteq \widehat{A}$ and $\mathfrak{m}^n \subseteq \widehat{\mathfrak{m}}^n$ are dense. We claim that the composition

$$\varphi\colon A \hookrightarrow \widehat{A} \twoheadrightarrow \widehat{A}/\widehat{\mathfrak{m}}^n$$

is surjective. Since $A \subseteq \widehat{A}$ is dense, also the image $\operatorname{im}(\varphi) \subseteq \widehat{A}/\widehat{\mathfrak{m}}^n$ is dense in the quotient topology. The equation

$$\widehat{\mathfrak{m}}^n = \{x \in \widehat{A} \mid \widehat{v}(x) > n-1\}$$

[8] i.e., a field with a non-Archimedean valuation whose exponential valuation is discrete.

shows that the ideal $\widehat{\mathfrak{m}}^n$ is open in $\widehat{A}$. Thus, the quotient topology on $\widehat{A}/\widehat{\mathfrak{m}}^n$ is the discrete topology. Since every subset of a discrete topological space is closed, we deduce that $\operatorname{im}(\varphi)$ is closed. Thus, $\operatorname{im}(\varphi)$ is dense and closed in $\widehat{A}/\widehat{\mathfrak{m}}^n$ and we get $\operatorname{im}(\varphi) = \widehat{A}/\widehat{\mathfrak{m}}^n$. This proves the surjectivity of φ. From $\ker\varphi = A \cap \widehat{\mathfrak{m}}^n$, we deduce the isomorphism (3.3). □

Let $(k, \widehat{|\cdot|})$ be a complete discretely valued field with valuation ring A and maximal ideal $\mathfrak{m}$. Since the absolute value is non-Archimedean[9], it satisfies the strict triangle inequality

$$|x+y| \leq \max(|x|, |y|).$$

In such fields, we have the following useful property.

Lemma 3.2.6. *Let $(k, \widehat{|\cdot|})$ be a complete discretely valued field. If $(a_n)_{n\geq 1} \subseteq k$ is a zero-sequence in k then*[10]

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

converges in k.

Proof. This will be proven in the exercises. □

Example 3.2.7. For every prime p, and every $k \geq 0$ the prime power p^k divides $n!$ for $n \geq p^k$. Thus, $(n!)_{n\geq 1}$ is a zero-sequence in $\mathbb{Q}_p$ for any prime p. In particular, the series

$$\sum_{n=1}^{\infty} n!$$

converges[11] for every prime p in the valued field $\mathbb{Q}_p$.

Let us recall that a discrete valuation ring is a principal ideal domain with exactly one maximal ideal. The following result is useful for an explicit description of elements in complete discretely valued fields.

Proposition 3.2.8. *Let A be a complete discrete valuation ring with maximal ideal $\mathfrak{m}$, uniformizer π and normalized valuation v. Furthermore, let $R \subseteq A$ be a system of representatives for $A/\mathfrak{m}$. Then any $x \in k \setminus \{0\}$ can be written uniquely in the form*

$$x = \pi^m \sum_{k=0}^{\infty} a_k \pi^k$$

with $m = v(x)$, $a_i \in R$ and $a_0 \neq 0$.

Proof. The element $u := \pi^{-m}x$ satisfies

$$v(u) = v(\pi^{-m}) + v(x) = -m + m = 0,$$

[9] Recall that a discretely valued field is by definition a field with a non-Archimedean absolute value whose associated exponential valuation is discrete.

[10] Of course, this does not hold in Archimedean fields like $\mathbb{R}$, e.g., $(n^{-1})_{n\geq 1}$ is a zero sequence but $\sum_{n=1}^{\infty} n^{-1}$ diverges.

[11] By the way, it is an unsolved conjecture that the value of this sequence in $\mathbb{Q}_p$ is transcendental, i.e., not algebraic over $\mathbb{Q}$.

and hence it is a unit in A. We prove by induction on n that u admits a unique representation of the form

$$u = a_0 + a_1\pi + a_2\pi^2 + \cdots + a_{n-1}\pi^{n-1} + b_n\pi^n \tag{3.4}$$

with $a_i \in R$ for $i = 0, \ldots, n-1$, $a_0 \neq 0$ and $b_n \in A$. For $n = 1$, this follows from the fact that R is a system of representatives for $A/\mathfrak{m}$. Now, let us assume that we already have found a unique representation for u of the form (3.4). Then, let $a_n \in R$ be the unique representative for b_n. Thus, we can write

$$u = a_0 + a_1\pi + a_2\pi^2 + \cdots + a_{n-1}\pi^{n-1} + a_n\pi^n + b_{n+1}\pi^{n+1}$$

with $b_{n+1} \in A$ which is uniquely determined by

$$b_{n+1} = \frac{u - (a_0 + a_1\pi + a_2\pi^2 + \cdots + a_{n-1}\pi^{n-1} + a_n\pi^n)}{\pi^{n+1}}.$$

The coefficients $a_0, \ldots, a_{n-1}$ are uniquely determined by the induction hypothesis. This shows that for every $n \geq 1$ there is a unique representation of u of the form (3.4). Thus, we obtain the uniquely determined series

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

which converges to u, by the previous Lemma. □

For a complete discrete valuation ring A with maximal ideal $\mathfrak{m}$, the projections

$$A/\mathfrak{m}^m \twoheadrightarrow A/\mathfrak{m}^n,$$

for $m \geq n$, form a projective system of ring homomorphisms and we can form the projective limit

$$\varprojlim_n A/\mathfrak{m}^n := \{(x_n)_{n\geq 1} \mid x_n \equiv x_{n+1} \mod \mathfrak{m}^n \text{ for all } n \geq 1.\} \subseteq \prod_{n\geq 1} A/\mathfrak{m}^n.$$

The projective limit carries a canonical ring structure given by the component-wise addition and multiplication. Furthermore, it inherits the subspace topology from the product topology of $\prod_{n\geq 1} A/\mathfrak{m}^n$, where each factor $A/\mathfrak{m}^n$ is equipped with the discrete topology[12].

Proposition 3.2.9. *Let A be a complete discrete valuation ring with maximal ideal $\mathfrak{m}$. We equip A with the topology given by some absolute value associated to the canonical valuation v on A. Then, the canonical map*

$$\varphi\colon A \to \varprojlim_n A/\mathfrak{m}^n, \quad x \mapsto (x \mod \mathfrak{m}^n)_{n\geq 1}$$

is an isomorphism of topological rings.

[12] Of course, the above can be summarized by saying that we take the limit in the category of topological rings, where $A/\mathfrak{m}^n$ is equipped with the discrete topology.

Proof. Of course, φ is a homomorphism of rings. It is injective since we have

$$\ker \varphi = \cap_{n\geq 0}\mathfrak{m}^n = \{0\}.$$

On the other hand, it is surjective by Proposition 3.2.8. It remains to check that φ is a homeomorphism of topological spaces. A basis of neighbourhoods of zero for the product topology on $\prod_{n\geq 1} A/\mathfrak{m}^n$ is given by the family of subsets

$$\mathcal{B} := (\mathcal{B}_N)_{N\geq 1}$$

with $\mathcal{B}_N := \prod_{n=1}^N\{0\} \times \prod_{n>N} A/\mathfrak{m}^n$. Since $\varprojlim_n A/\mathfrak{m}^n$ is equipped with the subspace topology, a basis of neighbourhoods for $0 \in \varprojlim_n A/\mathfrak{m}^n$ is given by the family $\left(\mathcal{B}_N \cap \varprojlim_n A/\mathfrak{m}^n\right)_{N\geq 1}$. On the other hand, a basis of neighbourhoods of $0 \in A$ on A with the topology induced by the normalized valuation v is given by $(\mathfrak{m}^N)_{N\geq 1}$. Now, the following equation shows that φ identifies both bases of neighbourhoods of zero:

$$\varphi(\mathfrak{m}^N) = \mathcal{B}_N \cap \varprojlim_n A/\mathfrak{m}^n, \quad \text{for all } N \geq 1.$$

Since a basis of zero determines uniquely the topology on any topological ring, we deduce that φ is a homeomorphism. □

Example 3.2.10. If we apply the above Proposition to the discrete valuation ring $\mathbb{Z}_p$ of $(\mathbb{Q}_p, |\cdot|_p)$, then we obtain

$$\mathbb{Z}_p \cong \varprojlim_n \mathbb{Z}_p/p^n\mathbb{Z}_p.$$

On the other hand, we have already seen in Proposition 3.2.5 and Example 3.2.4 that

$$\mathbb{Z}_p/p^n\mathbb{Z}_p \cong \mathbb{Z}_{(p)}/p^n\mathbb{Z}_{(p)} \cong \mathbb{Z}/p^n\mathbb{Z},$$

and obtain

$$\mathbb{Z}_p \cong \varprojlim_n \mathbb{Z}/p^n\mathbb{Z}.$$

Thus, we have two equivalent ways of thinking about $\mathbb{Z}_p$. Either as the valuation ring in the completion of $\mathbb{Q}$ with respect to $|\cdot|_p$, or as the limit $\varprojlim_n \mathbb{Z}/p^n\mathbb{Z}$. Both descriptions are useful on there own.

Outlook: Local fields

Ostrowski's Theorem provides a classification of all absolute values on a given number field. Indeed it is possible to characterize all possible fields which can arise in this way as completions of number fields.

Definition 3.2.11. A *local field* is a field k with a non-trivial absolute value $|\cdot|$ such that the topology induced by $|\cdot|$ is locally compact, i.e., for every $\alpha \in k$ there is an $\epsilon > 0$ such that the closed ball

$$B_{\leq\epsilon}(\alpha) := \{x \in k \mid |x - \alpha| \leq \epsilon\}$$

is compact.

Of course, $\mathbb{R}$ and $\mathbb{C}$ with the usual absolute values are examples of local fields since all closed balls are compact in the usual topologies of $\mathbb{R}$ and $\mathbb{C}$. The field $\mathbb{Q}$ with the usual (Archimedean) absolute value is not a local field, since no closed ball $B_{\leq\epsilon}(x) \subseteq \mathbb{Q}$ is compact; such balls are always missing limit points. Indeed, one can show that any local field has to be complete. If we ignore the local fields of positive characteristic then the local fields of characteristic zero are exactly the fields which we obtain as completions of number fields.

More precisely, the following Theorem implies that $\mathbb{R}$, $\mathbb{C}$ and the finite extensions of $\mathbb{Q}_p$ are all local fields of characteristic zero.

Theorem 3.2.12. *The only Archimedean local fields are $\mathbb{R}$ and $\mathbb{C}$. The non-Archimedean local fields of characteristic zero are exactly the finite extensions of $\mathbb{Q}_p$.*

Proof. For a proof, see for example Chapter II, (5.2) in Neukirch's book on Algberaic Number Theory[13], but note that Neukirch's definition of a local field excludes the Archimedean fields. □

3.3 *Hensel's Lemma*

The aim of this section is to prove Hensel's Lemma. This is a very useful tool to lift decompositions of separable polynomials from the residue field of a complete discrete valuation ring to the discrete valuation ring (DVR). Let us fix the following notation for this section:

Notation 3.3.1. In this section, let A be a complete discrete valuation ring with fraction field K, maximal ideal $\mathfrak{m}$, uniformizer π and residue field κ.

Definition 3.3.2. Notation as in 3.3.1. A polynomial $f = a_0 + \cdots + a_n X^n \in A[X]$ is called *primitive* if and only if the ideal generated by its coefficients $(a_0, \ldots, a_n)$ is the whole ring A.[14]

Hensel's Lemma allows us to lift decompositions of polynomials from $\kappa[X]$ to $A[X]$.

Theorem 3.3.3 (Hensel's Lemma). *Let $f \in A[X]$ be a primitive polynomial over a complete[15] discrete valuation ring. Assume that the reduction $\overline{f} := f$*

[13] Jürgen Neukirch. *Algebraic number theory*, volume 322 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1999. ISBN 3-540-65399-6. DOI: 10.1007/978-3-662-03983-0. URL https://doi.org/10.1007/978-3-662-03983-0. Translated from the 1992 German original and with a note by Norbert Schappacher, With a foreword by G. Harder

[14] For a DVR, this is equivalent to $f \not\equiv 0 \mod \mathfrak{m}$.

[15] Completeness is important here. Let us consider the discrete valuation ring $\mathbb{Z}_{(5)}$, which is the localization of $\mathbb{Z}$ at the prime ideal (5), and the polynomial $f = X^2 + 1 \in \mathbb{Z}_{(5)}[X]$. The reduction $\overline{f} = X^2 + 1 \in \mathbb{F}_5[X]$ splits as $\overline{f} = X^2 + 1 = (X-2)(X-3) \in \mathbb{F}_5$. But f does not split in $\mathbb{Z}_{(5)}$ since $\mathbb{Q}$ does not contain i and $-i$.

mod $\mathfrak{m} \in \kappa[X]$ *decomposes into a product of co-prime polynomials*

$$\overline{f} = \overline{g}\overline{h}, \quad \overline{g}, \overline{h} \in \kappa[X].$$

Then, there exist $g, h \in A[X]$ such that:

- g mod $\mathfrak{m} = \overline{g}$ *and* h mod $\mathfrak{m} = \overline{h}$ *in* $\kappa[X]$*,*
- $\deg(g) = \deg(\overline{g})$*,*
- $f = g \cdot h$ *in* $A[X]$*.*

Proof. Let us define $d := \deg f$ and $m := \deg \overline{g}$. Then

$$\deg \overline{h} = \deg \overline{f} - \deg \overline{g} \leq d - m.$$

Let us choose lifts $g_0, h_0 \in A[X]$ of $\overline{g}$ and $\overline{h}$ such that

$$\begin{aligned} g_0 &= g \bmod \mathfrak{m}, &&\text{and } \deg g_0 = m \\ h_0 &= h \bmod \mathfrak{m}, &&\text{and } \deg h_0 \leq d - m. \end{aligned}$$

Since $\overline{g}$ and $\overline{h}$ are co-prime, we can find $a, b \in A[X]$ such that

$$a g_0 + b h_0 \equiv 1 \mod \mathfrak{m}.$$

In the following, we will lift the decomposition $\overline{f} = \overline{g}\overline{h}$ successively modulo higher and higher powers of $\mathfrak{m} = (\pi)$. More precisely, we will prove the following claim by induction on n.
Claim: For any positive integer n, there exist polynomials

$$p_1, p_2, \ldots, p_{n-1} \in A[X] \text{ and } q_1, q_2, \ldots, q_{n-1} \in A[X]$$

with $\deg p_i < m$ and $\deg q_i \leq d - m$ such that

$$f \equiv g_{n-1} h_{n-1} \mod \pi^n, \tag{3.5}$$

where

$$g_{n-1} := g_0 + \pi p_1 + \ldots \pi^{n-1} p_{n-1}, \quad h_{n-1} := h_0 + \pi q_1 + \ldots \pi^{n-1} q_{n-1}.$$

Proof of the claim: For $n = 1$, we have $f \equiv g_0 h_0 \mod \pi$ by the assumptions. Let us assume that we have already constructed g_{n-1} and h_{n-1} with the desired property. Our aim is to construct g_n and h_n. By the induction hypothesis, we have $f - g_{n-1}h_{n-1} \in \pi^n A$ and we define

$$f_n := \pi^{-n}(f - g_{n-1}h_{n-1}) \in A[X].$$

Because of $\deg g_0 = \deg(g_0 \bmod \mathfrak{m})$, the leading term of g_0 is contained in $A \setminus \mathfrak{m}$, and hence it is a unit. By division with remainder, we find $p_n, q_n' \in A[X]$ such that

$$b f_n = q_n' g_0 + p_n$$

with

$$\deg p_n < \deg g_0 = m. \tag{3.6}$$

We deduce

$$\begin{aligned} g_0(af_n + h_0 q_n') + h_0 p_n &= g_0 a f_n + h_0(q_n' g_0 + p_n) \\ &= g_0 a f_n + h_0 b f_n \equiv f_n \mod \mathfrak{m}. \end{aligned} \tag{3.7}$$

We define q_n as the polynomial obtained from $af_n + h_0 q_n'$ by removing all coefficients which are divisible by π. In particular, we have $\deg(q_n \bmod \mathfrak{m}) = \deg q_n$. Thus, we get from equation (3.7)

$$g_0 q_n \equiv f_n - h_0 p_n \mod \mathfrak{m}. \tag{3.8}$$

Now, by equation (3.8) we obtain

$$m + \deg q_n = \deg(f_n - h_0 p_n \bmod \mathfrak{m}).$$

Since $\deg(f_n \bmod \mathfrak{m}) \leq d$ and $\deg(h_0 p_n \bmod \mathfrak{m}) < (d-m) + m = d$, we deduce from this equality

$$\deg q_n \leq d - m.$$

It remains to prove $f = g_n h_n \mod \mathfrak{m}^{n+1}$, but this follows from the following computation

$$\begin{aligned} f - g_n h_n &= f - (g_{n-1} + \pi^n p_n)(h_{n-1} + \pi^n q_n) \\ &= f - g_{n-1} h_{n-1} - \pi^n (p_n h_{n-1} + q_n g_{n-1} + \pi^n p_n q_n) \\ &= \pi^n (\underbrace{f_n - p_n h_{n-1} - q_n g_{n-1}}_{\equiv f_n - p_n h_0 - q_n g_0 \overset{(3.8)}{\equiv} 0 \mod \mathfrak{m}} - \pi^n p_n q_n) \equiv 0 \mod \mathfrak{m}^{n+1}. \end{aligned}$$

This finishes the proof of the Claim.

Now, the statement of Hensel's Lemma follows from defining

$$g := \lim_n g_n, \quad h := \lim_n h_n.$$

Note, that this limit exists in $A[X]$, since A is complete[16], g_n and h_n have bounded degrees and $\pi^n \to 0$ as $n \to \infty$. Furthermore, we have $g_n \bmod \mathfrak{m} = g_0$, $h_n \bmod \mathfrak{m} = h_0$, and we deduce from $\deg g_n = \deg \overline{g}$ that $\deg g = m = \deg \overline{g}$. By (3.5), we obtain

$$f = g \cdot h$$

as desired. □

We have the following variant of Hensel's Lemma which allows us to find unique lifts of simple roots in the residue field.

[16] The ring of polynomials over a complete discrete valuation ring is in general not necessarily complete with respect to the $\mathfrak{m}$-adic topology, i.e., the topology for which $(\mathfrak{m}^k A[X])_k$ forms a basis of neighbourhoods of zero. For example, the sequence $(f_k)_k$ given by

$$f_k = 1 + \pi X + \pi^2 X^2 + \cdots + \pi^k X^k \in A[X]$$

is a Cauchy sequence in $A[X]$ but it does not converge. On the other hand, the subspace of polynomials of degeree $\leq N$ for a fixed integer N is complete.

Corollary 3.3.4. *Let $f \in A[X]$ be a monic polynomial with reduction $\overline{f} \in \kappa[X]$ and $\overline{\alpha} \in \kappa$ a simple root of $\overline{f}$ in κ. Then, there exists a unique root $\alpha \in A$ of f such that $\overline{\alpha} = \alpha \mod \mathfrak{p}$.*

Proof. We consider $\overline{f} = \overline{g}\overline{h}$ with $\overline{g} := X - \overline{\alpha}$. Since $\overline{\alpha}$ is a simple root, the polynomials $\overline{g}$ and $\overline{h}$ are co-prime. Thus, Hensel's Lemma gives a decomposition $f = gh$ with $\deg g = 1$, hence $g = X - \alpha$ for some $\alpha \in A$ with $\alpha \mod \mathfrak{m} = \overline{\alpha}$. This shows the existence of a root α of f lifting $\overline{\alpha}$. The uniqueness follows from the fact that $\overline{\alpha}$ is a simple root. □

For a polynomial $f = a_0 + a_1 X + \cdots + a_n X^n \in K[X]$, let us define

$$|f| := \max\{|a_0|, |a_1|, \ldots, |a_n|\}.$$

Another useful application of Hensel's Lemma is the following:

Corollary 3.3.5. *Let $f = a_0 + a_1 X + \cdots + a_n X^n \in K[X]$ be a polynomial over the fraction field K of a complete DVR A.*

(a) If f is irreducible in $K[X]$, than $|f| = \max\{|a_0|, |a_n|\}$.

(b) If f is irreducible, monic and $f(0) = a_0 \in A$ then $f \in A[X]$.

Proof. (a) After multiplication with a suitable element from K, we may assume without loss of generality that $|f| = 1$. This is equivalent to $f \in A[X]$ and $\overline{f} \neq 0$ in $\kappa[X]$. Let $r := \min_{0 \leq i \leq n}\{i : |a_i| = 1\}$ be the index of the least non-vanishing coefficient of $\overline{f} \in \kappa[X]$, i.e.,

$$\overline{f} = \underbrace{X^r}_{=:\overline{h}} \underbrace{(\overline{a}_r + \overline{a}_{r+1} X + \cdots + \overline{a}_n X^{n-r})}_{=:\overline{g}} \in \kappa[X].$$

If we had $\max\{|a_0|, |a_n|\} < 1$, this would imply $0 < r < n$. Thus, $\deg \overline{g}, \deg \overline{h} > 0$. The polynomial $\overline{g} = X^r$ is obviously co-prime to $\overline{h}$ since $\overline{a}_r \neq 0$. Thus, we could apply Hensel's Lemma to get a non-trivial decomposition of f and we would obtain a contradiction to the irreducibility of f. This proves $\max\{|a_0|, |a_n|\} = 1$ and hence (a).
(b) is an easy consequence of (a). Indeed, the statement (a) implies $|f| = \max\{|a_0|, |a_n|\}$. By the assumptions, we have $|a_n| = 1$ and $|a_0| = |f(0)| \leq 1$. Hence, we get $|f| \leq 1$ but this means that all coefficients are in A. □

Another consequence of Hensel's Lemma is the following fact about the integral closure of a complete discrete valuation ring:

Corollary 3.3.6. *Let L/K be a finite field extension of the fraction field K of the complete DVR A. Let us denote by $B := \overline{A}^L$ the integral closure of A in L, then*

$$B = \{\alpha \in L \mid N_{L/K}(\alpha) \in A\},$$

where $N_{L/K}$ denotes the field norm $N_{L/K} \colon L \to K$.

Proof. Let us start with the following observation. Let $\alpha \in L$ with normalized minimal polynomial

$$f = a_0 + a_1 X + \cdots + a_{n-1} X^{n-1} + X^n \in K[X].$$

By the basic properties of norms[17], we have

$$N_{L/K}(\alpha) = N_{K(\alpha)/K}(\alpha)^{[L:K(\alpha)]}$$

and $N_{K(\alpha)/K}(\alpha) = (-1)^n a_0$. We deduce

$$N_{L/K}(\alpha) = (-1)^{n[L:K(\alpha)]} a_0^{[L:K(\alpha)]}. \tag{3.9}$$

$\subseteq$: If we apply (3.9) to an integral element $\alpha \in B$, we get $N_{L/K}(\alpha) \in A$. $\supseteq$: Conversely, let us assume that $N_{L/K}(\alpha) \in A$. By (3.9), $a_0 \in K$ is integral and hence an element of A. According to Corollary 3.3.5(b), we have $f \in A[X]$. Thus, the normalized minimal polynomial of α has coefficients in A and we get $\alpha \in \overline{A}^L = B$. □

Finally, we will implicitly use Hensel's Lemma when we want to extend a given absolute value to a larger field.

Theorem 3.3.7 (Extension of absolute values)**.** *Let $(K, |\cdot|)$ be a complete and discretely valued field and L/K an algebraic field extension. Then, there is a unique absolute value $|\cdot|_L$ on L such that $|\cdot|_L$ extends $|\cdot|$, i.e., for all $x \in K$, we have $|x|_L = |x|$. The valuation ring of $|\cdot|_L$ is the integral closure of A in L. Furthermore, if L/K is finite of degree n then $|\cdot|_L$ is complete, discrete and satisfies*

$$|\cdot|_L = |N_{L/K}(\cdot)|^{\frac{1}{n}}.$$

Proof. First, observe that it is enough to prove existence and uniqueness for *finite* field extensions. Indeed, if we have already proven the existence and uniqueness for all finite field extensions, then for a given $\alpha \in L$ we choose a finite sub-extension F/K containing α and define

$$|\alpha|_L := |\alpha|_F.$$

By the uniqueness of $|\cdot|_F$, this does not depend on F and hence it is well-defined. For any given $\alpha, \beta \in L$, we can choose a finite extension F containing α and β and check that $|\cdot|_L$ is indeed an absolute value:

$$\begin{aligned} |\alpha|_L &= |\alpha|_F = 0 \Leftrightarrow \alpha = 0, \\ |\alpha\beta|_L &= |\alpha\beta|_F = |\alpha|_F |\beta|_F = |\alpha|_L |\beta|_L, \\ |\alpha + \beta|_L &= |\alpha + \beta|_F \leq \max(|\alpha|_F, |\beta|_F) = \max(|\alpha|_L, |\beta|_L). \end{aligned}$$

Thus, we may without loss of generality assume that L/K is finite.
Existence: For a finite extension L/K, the formula for $|\cdot|_L$ is already suggested by the statement of the Theorem. For $\alpha \in L$ set

$$|\alpha|_L := |N_{L/K}(\cdot)|^{\frac{1}{n}}.$$

[17] Recall the following statements about norms (e.g., ANT 1):

(a) For any finite field extensions $L/M/K$ and $\alpha \in M$, we have $N_{L/K}(\alpha) = N_{M/K}(\alpha)^{[L:M]}$.

(b) Let L/K be finite and

$$P = a_0 + \cdots + a_{n-1} X^{n-1} + X^n$$

be the minimal polynomial of $\alpha \in K$. Then $N_{K(\alpha)/K} = (-1)^n a_0$.

We have to check that this defines an absolute value on L. Let $\alpha, \beta \in L$. Obviously, we have $|\alpha|_L = 0 \Leftrightarrow \alpha = 0$ and by the multiplicativity of norms $|\alpha\beta|_L = |\alpha|_L|\beta|_L$ for $\alpha, \beta \in L$. Furthermore, the explicit formula for $|\alpha|_L$ shows, that the associated exponential valuation of $|\cdot|_L$ is discrete. It remains to check the non-Archimedean triangle inequality. We define $B := \overline{A}^L$ as the integral closure of A in L. By Corollary 3.3.6 and the definition of $|\cdot|_L$, we have for all $x \in L$

$$|x|_L \leq 1 \Leftrightarrow |N_{L/K}(x)| \leq 1 \Leftrightarrow N_{L/K}(x) \in A \Leftrightarrow x \in B. \tag{3.10}$$

In particular, we have for all $x \in L$ the equivalence $|x|_L \leq 1 \Leftrightarrow |x+1|_L \leq 1$. Applying this to $x = \frac{\alpha}{\beta}$ with $\beta \in L^\times$, we obtain after multiplication with $|\beta|_L$ the equivalence

$$|\alpha|_L \leq |\beta|_L \Leftrightarrow |\alpha + \beta|_L \leq |\beta|_L.$$

Since we can interchange the roles of α and β, we also deduce

$$|\beta|_L \leq |\alpha|_L \Leftrightarrow |\alpha + \beta|_L \leq |\alpha|_L.$$

Combining both inequalities gives

$$|\alpha + \beta|_L \leq \max(|\alpha|_L, |\beta|_L)$$

as desired. This proves the existence of $|\cdot|_L$. The claim that B is the valuation ring of $(L, |\cdot|)$ follows from equation (3.10).
Uniqueness: Let $|\cdot|_L$ be the absolute value on L which has been constructed above. In particular, we already know that $B := \overline{A}^L$ is the discrete valuation ring of $|\cdot|_L$. Let $|\cdot|_L'$ be a second absolute value on L which extends $|\cdot|$. We want to prove that both absolute values coincide. Let us denote by B' the valuation ring of $|\cdot|'$ and by $\mathfrak{m}'$ its maximal ideal, i.e.,

$$B' := \{x \in L \mid |x|_L' \leq 1\}, \quad \mathfrak{m}' := \{x \in L \mid |x|_L' < 1\}.$$

Assume there was an element $\alpha \in B \setminus B'$. Let $f = a_0 + \cdots + a_{n-1}X^{n-1} + X^n \in K[X]$ be the normalized minimal polynomial of α, i.e.,

$$\alpha^n = -a_0 - a_1\alpha - \cdots - a_{n-1}\alpha^{n-1}. \tag{3.11}$$

Of course, $a_0, \ldots, a_{n-1} \in A$ because $\alpha \in B$ is integral. Since $|\cdot|_L'$ extends $|\cdot|$, we get $A \subseteq B'$ and deduce $a_0, \ldots, a_{n-1} \in B'$. By assumption, we have $\alpha \notin B'$, i.e., $|\alpha|_L' > 1$, which implies $\alpha^{-1} \in \mathfrak{m}'$. If we multiply (3.11) by α^{-n}, we get

$$1 = \underbrace{-a_0\alpha^{-n} - a_1\alpha^{-n+1} - a_{n-1}\alpha^{-1}}_{\in \mathfrak{m}'}.$$

which leads to the contradiction that 1 is contained in the maximal ideal $\mathfrak{m}'$ of B'. Thus, we have shown $B \subseteq B'$. Since $B \cap \mathfrak{m}'$ is a non-zero prime ideal of B and since B is a discrete valuation ring, we deduce $B \cap \mathfrak{m}' = \mathfrak{m}$. This implies $\mathfrak{m} \subseteq \mathfrak{m}'$. In terms of the absolute values, this means for all $\alpha \in L$

$$|\alpha|_L < 1 \Rightarrow \alpha \in \mathfrak{m} \Rightarrow \alpha \in \mathfrak{m}' \Rightarrow |\alpha|'_L < 1.$$

Now, it follows from 3.1.4 that $|\cdot|_L = |\cdot|'^s_L$ for some positive real number s. Since both absolute values extend $|\cdot|$ on K, we deduce $s = 1$ and the uniqueness follows. The completeness of L with respect to $|\cdot|_L$ follows from the fact that any finite-dimensional normed K-vector space over $(K, |\cdot|)$ is complete, see the next Definition and the following Lemma. □

Definition 3.3.8. Let $(K, |\cdot|)$ be a complete discretely valued field and let V be a finite dimensional vector space. A *norm* on V is a map $\|\cdot\| \colon V \to \mathbb{R}_{\geq 0}$ such that for all $\alpha \in K$ and $v, w \in V$

(a) $\|v\| = 0$ if and only if $v = 0$,

(b) $\|\alpha v\| = |\alpha| \|v\|$,

(c) $\|v + w\| \leq \|v\| + \|w\|$.

Two norms $\|\cdot\|$ and $\|\cdot\|'$ are called equivalent if and only if there exist positive real numbers $c, C \in \mathbb{R}$ such that $c\|\cdot\| \leq \|\cdot\|' \leq C\|\cdot\|$.

We have the following Lemma.

Lemma 3.3.9. *Let V be a finite dimensional vector space over a complete valued field $(K, |\cdot|)$. Up to equivalence there is exactly one norm on V and V is complete with respect to this norm.*

Proof. For the existence, let us choose a basis $v_1, \ldots, v_n$ of V and define the maximum norm

$$\|\sum_{i=1}^{n} \alpha_i v_i\|_\infty := \max_{1 \leq i \leq n} |\alpha_i|.$$

It is easily checked that the maximum norm is a norm and that V is complete with respect to $\|\cdot\|_\infty$. For the uniqueness, let $\|\cdot\|$ be an arbitrary norm on V. We prove that $\|\cdot\|$ is equivalent to $\|\cdot\|_\infty$. For an arbitrary vector $v = \sum_{i=1}^{n} \alpha_i v_i$, we have

$$\|v\| \leq \sum_{i=1}^{n} |\alpha_i| \|v_i\| \leq C\|v\|_\infty,$$

for $C := \max_{1 \leq i \leq n} \|v_i\|$. The existence of the lower bound c is left as an exercise. □

An important Corollary of the unique extension of absolute values is the following result.

Corollary 3.3.10. *Let $(K, |\cdot|)$ be a complete discretely valued field and L/K a finite Galois extension with unique extension $|\cdot|_L$ of $|\cdot|$. Then, we have $|\sigma(\alpha)|_L = |\alpha|_L$ for all $\alpha \in L$ and all $\sigma \in Gal(L/K)$.*

Proof. We define $|\cdot|_L' := |\cdot|_L \circ \sigma$. This gives an absolute value on L extending $|\cdot|$. By uniqueness, we get $|\cdot|_L' = |\cdot|_L$. □

For later reference, let us formulate the following additive version of Theorem 3.3.7.

Corollary 3.3.11. *Let $(K, |\cdot|)$ be a complete and discretely valued field with exponential valuation v. For any algebraic field extension L/K, there is a unique additive valuation v_L on L which extends v. If L/K is finite then the value group $v(K^\times)$ of K is of finite index in the value group $v_L(L^\times)$ of L.*

Proof. This is an immediate restatement of Theorem 3.3.7 in terms of exponential valuations. □

3.4 *Ramification and completion*

In this section, we will recall facts about ramification of Dedekind domains from Algebraic Number Theory 1. Afterwards, we will put them into new perspectives by relating them to valuations. Finally, we will compare the ramification of Dedekind rings to the ramification in their completion.

In this section, let us fix the following notation.

Notation 3.4.1. Let A be a Dedekind domain with fraction field K and let L be a finite field extension of K. We denote the integral closure of A in L by $B := \overline{A}^L$. Thus, we have the following setup:

$$\begin{tikzcd} L & \supseteq & B & & \kappa(\mathfrak{q}) \\ K \arrow[u, hook] & \supseteq & A \arrow[u, hook] & & \kappa(\mathfrak{p}) \arrow[u, hook] \end{tikzcd}$$

The residue field of a prime ideal $\mathfrak{p} \subseteq A$ (resp. $\mathfrak{q} \subseteq B$) will be denoted by $\kappa(\mathfrak{p})$ (resp. $\kappa(\mathfrak{q})$).

Let us first recall the following facts about ramification which should be familiar from Algebraic Number Theory 1. For a non-zero prime ideal $\mathfrak{p} \subseteq A$, let us consider its unique prime decomposition in B:

$$\mathfrak{p}B = \prod_{\mathfrak{q} \subseteq B \text{ prime}} \mathfrak{q}^{e_\mathfrak{q}}.$$

For a prime $\mathfrak{q}$, the number $e_\mathfrak{q}$ is the *ramification index*[18] of $\mathfrak{q}$ over $\mathfrak{p}$. Furthermore, let us write

$$f_\mathfrak{q} = f(\mathfrak{q}/\mathfrak{p}) := [\kappa(\mathfrak{q}) : \kappa(\mathfrak{p})]$$

[18] If we want to emphasize the dependence of $\mathfrak{p}$, we will sometimes write $e(\mathfrak{q}/\mathfrak{p})$ for $e_\mathfrak{q}$.

for the *inertia degree*, i.e., the degree $[\kappa(\mathfrak{q}) : \kappa(\mathfrak{p})]$ of the extension of residue fields. Let us recall that the prime $\mathfrak{q}$ is called *unramified* over $\mathfrak{p}$ if and only if $e(\mathfrak{q}/\mathfrak{p}) = 1$ and $\kappa(\mathfrak{q})/\kappa(\mathfrak{p})$ is a separable field extension[19]. Furthermore, we always have

$$[L : K] = \sum_{\mathfrak{q}|\mathfrak{p}} f_{\mathfrak{q}} e_{\mathfrak{q}}. \tag{3.12}$$

Let us now assume that L/K is a Galois extension with Galois group $G := \mathrm{Gal}(L/K)$. In this case, one has $e_{\mathfrak{q}} = e_{\mathfrak{q}'}$ and $f_{\mathfrak{q}} = f_{\mathfrak{q}'}$ for all $\mathfrak{q}, \mathfrak{q}'$ over a fixed prime $\mathfrak{p}$ of A. Let us denote the *decomposition group* of a prime $\mathfrak{q}$ over $\mathfrak{p}$ by

$$G_{\mathfrak{q}} := \{\sigma \in G \mid \sigma(\mathfrak{q}) = \mathfrak{q}\}.$$

In the above situation, we have a surjection $G_{\mathfrak{q}} \twoheadrightarrow \mathrm{Gal}(\kappa(\mathfrak{q})/\kappa(\mathfrak{p}))$ and the *inertia group* $I_{\mathfrak{q}}$ of $\mathfrak{q}$ is the kernel of this surjection, i.e., we have a short exact sequence

$$1 \to I_{\mathfrak{q}} \to G_{\mathfrak{q}} \twoheadrightarrow \mathrm{Gal}(\kappa(\mathfrak{q})/\kappa(\mathfrak{p})) \to 1. \tag{3.13}$$

Let us also recall that L/K is unramified at $\mathfrak{q}$ if and only if the inertia group $I_{\mathfrak{q}}$ is trivial. Finally, let us observe that the ramification index can be expressed as the index of the value groups of the corresponding valuations.

Lemma 3.4.2. *Let $\mathfrak{q} \subseteq B$ be a prime over $\mathfrak{p} \subseteq A$ and let us write $v_{\mathfrak{q}}$ for the normalized valuations associated to $\mathfrak{q}$. Then*

$$e(\mathfrak{q}/\mathfrak{p}) = [v_{\mathfrak{q}}(L^\times) : v_{\mathfrak{q}}(K^\times)].$$

Proof. This follows immediately from the definition of the ramification index and the normalized valuation[20]. More precisely, the normalized value group $v_{\mathfrak{q}}(L^\times)$ is $\mathbb{Z}$. For any $x \in K^\times$, the fractional ideal $(x)_K \subseteq K$ with decomposition

$$(x)_K = \prod_{\mathfrak{p}} \mathfrak{p}^{v_{\mathfrak{p}}(x)}$$

has in L the decomposition

$$(x)_L = \prod_{\mathfrak{p}} \prod_{\mathfrak{q}|\mathfrak{p}} \mathfrak{q}^{v_{\mathfrak{p}}(x) e(\mathfrak{q}/\mathfrak{p})}.$$

Since the normalized valuation $\nu_{\mathfrak{p}}(x)$ assumes every integer for a suitable $x \in K^\times$, this proves the assertion $[v_{\mathfrak{q}}(L^\times) : v_{\mathfrak{q}}(K^\times)] = e(\mathfrak{q}/\mathfrak{p})$. □

Let us also formulate the following special case of the previous Lemma when A and B are complete discrete valuation rings[21].

[19] In this lecture, the second condition will be satisfied automatically. For a finite extension of $\mathbb{Q}_p$, all residue fields are finite and hence perfect, i.e., any finite extension is separable.

[20] Here, recall the definition of the normalized valuation of a prime ideal in a Dedekind domain. For $x \in K$ we define $v_{\mathfrak{p}}(x)$ as the exponent of $\mathfrak{p}$ in the unique prime decomposition $(x) = \prod_{\mathfrak{p}} \mathfrak{p}^{v_{\mathfrak{p}}(x)}$. On the other hand, the ramification index $e_{\mathfrak{q}}$ is the exponent of $\mathfrak{q}$ in the prime decomposition of $\mathfrak{p} = \prod_{\mathfrak{q}} \mathfrak{q}^{e_{\mathfrak{q}}}$.

[21] Here, it is a good point to recall that every discrete valuation ring is a Dedekind domain. Thus, everything you know about Dedekind rings applies in particular to DVRs.

Corollary 3.4.3. *Let L/K be a finite extension of a complete discretely valued field $(K, |\cdot|)$ with exponential valuation v_K. Then, we have*

$$e_{L/K} := e(\mathfrak{m}_L/\mathfrak{m}_K) = [v_L(L^\times) : v_K(K^\times)],$$

where v_L is the unique extension of v_K to L and $\mathfrak{m}_K$ respectively $\mathfrak{m}_L$ are the maximal ideals in the discrete valuation rings A (resp. B) of K (resp. L).

Proof. We have shown the previous Lemma only for the normalized valuation and not for arbitrary valuations. So, we only have to relate the normalized valuation of the maximal ideals to the given exponential valuations. That's easily done. The exponential valuation v_L of $|\cdot|_L$ differs from the normalized valuation $v_{\mathfrak{m}_L}$ only by multiplication with a positive real number $t \in \mathbb{R}_{>0}$, i.e.

$$v_L = t \cdot v_{\mathfrak{m}_L}.$$

Now, the claim follows from the previous Lemma

$$\begin{aligned} e(\mathfrak{m}_L/\mathfrak{m}_K) &= [v_{\mathfrak{m}_L}(L^\times) : v_{\mathfrak{m}_L}(K^\times)] \\ &= [v_L(L^\times) : v_L(K^\times)] = [v_L(L^\times) : v_K(K^\times)]. \end{aligned}$$

□

The following result relates the ramification of Dedekind rings to the corresponding ramification in the completion.

Theorem 3.4.4. *Let $\mathfrak{p} \in A$ be an ideal with factorization $\mathfrak{p}B = \prod_{\mathfrak{q}|\mathfrak{p}} \mathfrak{q}^{e_\mathfrak{q}}$. Let $K_\mathfrak{p}$ be the completion of K with respect to $|\cdot|_\mathfrak{p}$ and let $\widehat{\mathfrak{p}}$ be the maximal ideal in its complete discrete valuation ring $\widehat{A}_\mathfrak{p}$. Similarly, denote for $\mathfrak{q} \subseteq B$ by $L_\mathfrak{q}$ the completion of L with respect to $|\cdot|_\mathfrak{q}$ and $\widehat{\mathfrak{q}}$ the maximal ideal in its valuation ring $\widehat{B}_\mathfrak{q}$. Then, the following holds:*

(a) Each $L_\mathfrak{q}$ is a finite extension of $K_\mathfrak{p}$ with $[L_\mathfrak{q} : K_\mathfrak{p}] \leq [L : K]$.

(b) Each $\widehat{\mathfrak{q}}$ is the unique prime of $\widehat{B}_\mathfrak{q}$ lying over $\widehat{\mathfrak{p}}$.

(c) Each $\widehat{\mathfrak{q}}$ has ramification index $e_{\widehat{\mathfrak{q}}} = e_\mathfrak{q}$ and residue field degree $f_{\widehat{\mathfrak{q}}} = f_\mathfrak{q}$.

(d) $[L_\mathfrak{q} : K_\mathfrak{p}] = e_\mathfrak{q} f_\mathfrak{q}$.

(e) If L/K is Galois then each $L_\mathfrak{q}/K_\mathfrak{p}$ is Galois and we have isomorphisms of decomposition groups $G_\mathfrak{q} \cong G_{\widehat{\mathfrak{q}}} = Gal(L_\mathfrak{q}/K_\mathfrak{p})$ and inertia groups $I_\mathfrak{q} \cong I_{\widehat{\mathfrak{q}}}$.

Proof. (a) First, note that $K \hookrightarrow L$ induces[22] injections $K_\mathfrak{p} \hookrightarrow L_\mathfrak{q}$ and $\widehat{A}_\mathfrak{p} \hookrightarrow \widehat{B}_\mathfrak{q}$. Since any K-basis $b_1, \ldots, b_m$ of L spans $L_\mathfrak{q}$ as $K_\mathfrak{p}$-vector space, we obtain $[L_\mathfrak{q} : K_\mathfrak{p}] \leq [L : K]$.
(b) The valuation rings $\widehat{A}_\mathfrak{p}$ of $K_\mathfrak{p}$ and $\widehat{B}_\mathfrak{q}$ of $L_\mathfrak{q}$ are complete DVRs, hence they only have one non-zero prime ideal and the claim follows.

[22] For example, this can be seen as follows. Since $v_\mathfrak{q}$ extends $v_\mathfrak{p}$, the absolute value $|\cdot|_\mathfrak{q}$ extends $|\cdot|_\mathfrak{p}$. Hence, a Cauchy sequence respect to $|\cdot|_\mathfrak{p}$ is also a Cauchy sequence with respect to $|\cdot|_\mathfrak{q}$. The inclusion of the completions is now given by $[(x_n)_n] \mapsto [(x_n)_n]$.

(c) Let us write $v_{\mathfrak{q}}$ and $v_{\widehat{\mathfrak{q}}}$ for the normalized valuations associated to $\mathfrak{q}$ and $\widehat{\mathfrak{q}}$. By Proposition 3.2.5, the value group of a discretely valued field does not change under completion, i.e., we have $v_{\mathfrak{q}}(K^{\times}) = v_{\widehat{\mathfrak{q}}}(K_{\mathfrak{p}}^{\times})$ and $v_{\mathfrak{q}}(L^{\times}) = v_{\widehat{\mathfrak{q}}}(L_{\mathfrak{q}}^{\times})$. By Lemma 3.4.2, we have $e_{\mathfrak{q}} = [v_{\mathfrak{q}}(L^{\times}) : v_{\mathfrak{q}}(K^{\times})]$ and deduce

$$e_{\widehat{\mathfrak{q}}} = [v_{\widehat{\mathfrak{q}}}(L_{\mathfrak{q}}^{\times}) : v_{\widehat{\mathfrak{q}}}(K_{\mathfrak{p}}^{\times})] = [v_{\mathfrak{q}}(L^{\times}) : v_{\mathfrak{q}}(K^{\times})] = e_{\mathfrak{q}}.$$

By Proposition 3.2.5, the residue field of a DVR coincides with the residue field of its completion. This shows

$$f_{\mathfrak{q}} = [\kappa(\mathfrak{q}) : \kappa(\mathfrak{p})] = [\kappa(\widehat{\mathfrak{q}}) : \kappa(\widehat{\mathfrak{p}})] = f_{\widehat{\mathfrak{q}}}.$$

(d) follows from combining (b) and (c).
(e) Now, we assume that L/K is Galois. Each $\sigma \in G_{\mathfrak{q}}$ acts on L and respects the valuation $v_{\mathfrak{q}}$, since it fixes $\mathfrak{q}$.[23] Thus σ induces an automorphism of $L_{\mathfrak{q}}$ and fixes $K_{\mathfrak{p}}$. We get a group homomorphism

$$\varphi \colon G_{\mathfrak{q}} \to \mathrm{Aut}_{K_{\mathfrak{p}}}(L_{\mathfrak{q}}).$$

This map is injective: If $\varphi(\sigma)$ acts trivially on $L_{\mathfrak{q}}$, then it also acts trivially on $L \subseteq L_{\mathfrak{q}}$, so $\ker \varphi$ is trivial. On the other hand, we have

$$e_{\mathfrak{q}} f_{\mathfrak{q}} = |G_{\mathfrak{q}}| \leq |\mathrm{Aut}_{K_{\mathfrak{p}}}(L_{\mathfrak{q}})| \leq [L_{\mathfrak{q}} : K_{\mathfrak{p}}] = e_{\mathfrak{q}} f_{\mathfrak{q}},$$

so we have everywhere equality. In particular, φ is surjective and we have $|\mathrm{Aut}_{K_{\mathfrak{p}}}(L_{\mathfrak{q}})| = [L_{\mathfrak{q}} : K_{\mathfrak{p}}]$. The last equality implies that $L_{\mathfrak{q}}/K_{\mathfrak{p}}$ is a Galois extension. There is only one prime $\widehat{\mathfrak{q}}$ of the complete discrete valuation ring $\widehat{B}_{\widehat{\mathfrak{q}}}$ and this prime is necessarily fixed by every $\sigma \in \mathrm{Gal}(L_{\mathfrak{q}}/K_{\mathfrak{p}})$, so $\mathrm{Gal}(L_{\mathfrak{q}}/K_{\mathfrak{p}}) \cong G_{\widehat{\mathfrak{q}}}$. The inertia groups $I_{\mathfrak{q}}$ and $I_{\widehat{\mathfrak{q}}}$ have both $e_{\mathfrak{q}}$ elements, and φ restricts to an injective homomorphism $I_{\mathfrak{q}} \to I_{\widehat{\mathfrak{q}}}$, so also the inertia groups are isomorphic. □

Outlook

By Ostrowski's Theorem, a complete list of non-trivial absolute values on $\mathbb{Q}$ up to equivalence is given by the usual absolute value $|\cdot|$ together with the non-Archimedean absolute values $|\cdot|_p$ for each prime p.

The completion $\mathbb{R}$ of $\mathbb{Q}$ is not algebraically closed but the degree 2 extension $\mathbb{C}/\mathbb{R}$ is algebraically closed. Furthermore, $\mathbb{C}$ is complete with respect to the unique extension $|\cdot|$ of $\mathbb{R}$. So, $\mathbb{C}$ is the smallest field extension of $\mathbb{Q}$ which is algebraically closed, complete and extends the usual absolute value $|\cdot|$ on $\mathbb{Q}$. It is a natural question to ask for a field with analogous properties for the non-Archimedean absolute values.

A first natural step is to consider an algebraic closure $\overline{\mathbb{Q}}_p$ of $\mathbb{Q}_p$. By Theorem 3.3.7, there is a unique absolute value $|\cdot|_p$ on $\overline{\mathbb{Q}}_p$ extending the absolute value on $\mathbb{Q}_p$. Unfortunately, it turns out that $\overline{\mathbb{Q}}_p$ is, in contrary to the algebraic closure of $\mathbb{R}$, not complete[24]. Let us denote the completion of $(\overline{\mathbb{Q}}_p, |\cdot|_p)$ by $(\mathbb{C}_p, |\cdot|_p)$.

[23] More precisely, for $x \in L$ with $(x) = \prod_{\mathfrak{P}} \mathfrak{P}^{n_{\mathfrak{P}}}$ we have $(\sigma(x)) = \prod_{\mathfrak{P}} \sigma(\mathfrak{P})^{n_{\mathfrak{P}}}$. Since $\sigma \in G_{\mathfrak{q}}$ fixes $\mathfrak{q}$, we deduce that the multiplicity of $\mathfrak{q}$ in (x) is the same as the multiplicity of $\mathfrak{q}$ in $(\sigma(x))$.

[24] This also shows that we can not drop the hypothesis that L/K is finite in the last assertion about completeness of Theorem 3.3.7

Theorem. *The field $(\mathbb{C}_p, |\cdot|_p)$ is complete and algebraically closed.*

Thus, $\mathbb{C}_p$ can be seen as a kind of p-adic analogue of the field of complex numbers. Of course, we can form its valuation ring

$$\mathcal{O}_{\mathbb{C}_p} := \{x \in \mathbb{C}_p : |x|_p \leq 1\}$$

with its maximal ideal $\mathfrak{m} := \{x \in \mathbb{C}_p : |x|_p < 1\}$. The associated additive valuation v_p on $\mathbb{C}_p$ is not discrete. Indeed, there are elements in $\mathbb{C}_p$ with arbitrary small additive valuation, e.g., $v_p(\sqrt[n]{p}) = \frac{1}{n}$. Thus, $\mathcal{O}_{\mathbb{C}_p}$ is not a discrete valuation ring. Indeed, $\mathcal{O}_{\mathbb{C}_p}$ is a rather pathological example of a local ring. For example, one can even show that it is non-Noetherian and its maximal ideal satisfies

$$\mathfrak{m}^2 = \mathfrak{m}.$$

3.5 *Unramified and totally ramified extensions of $\mathbb{Q}_p$*

In this section, we will prove structure theorems about unramified and totally ramified extensions of finite extensions of $\mathbb{Q}_p$. For this section, let us fix the following notation.

Notation 3.5.1. Let us write $|\cdot|_p$ for the usual absolute value (i.e., $|p|_p = \frac{1}{p}$) and v_p for the normalized valuation on $\mathbb{Q}_p$. Let L/K be finite extensions of $\mathbb{Q}_p$ and equip them with the absolute value given by the unique extension of $|\cdot|_p$. We will again write $|\cdot|_p$ (resp. v_p) for this unique extension on L and K. The discrete valuation ring of K (resp. L) will be denoted by $\mathcal{O}_K$ (resp. $\mathcal{O}_L$), its maximal ideal by $\mathfrak{m}_K$ (resp. $\mathfrak{m}_L$) and its residue field by κ_K (resp. κ_L), i.e.,

$$\begin{tikzcd}
L \arrow[r, phantom, "\supseteq"] & \mathcal{O}_L \arrow[r, two heads] & \mathcal{O}_L/\mathfrak{m}_L = \kappa_L \\
K \arrow[u, hook] \arrow[r, phantom, "\supseteq"] & \mathcal{O}_K \arrow[u, hook] \arrow[r, two heads] & \mathcal{O}_K/\mathfrak{m}_K = \kappa_K. \arrow[u, hook]
\end{tikzcd}$$

Let us make the following definition.

Definition 3.5.2. Let us write $e := [v_p(L^\times) : v_p(K^\times)]$ for the ramification index of L/K. We call the field extension L/K

(a) *unramified* if and only if $e = 1$,

(b) *tamely ramified* if and only if $p \nmid e$,

(c) *totally ramified* if and only if $e = [L : K]$.

Let us first study all unramified extensions of $\mathbb{Q}_p$.

Lemma 3.5.3. *Let $P \in \mathcal{O}_K[X]$ be a normalized polynomial such that $P \bmod \mathfrak{m}_K \in \kappa_K[X]$ is separable. If $L = K(\alpha)$ for some root α of P then L/K is unramified.*

Proof. Exercises. □

The following result shows that all unramified extensions are 'cyclotomic', i.e., generated by roots of unity. More precisely, we have.

Theorem 3.5.4 (Unramified extensions)**.** *The extension L/K is unramified if and only if $L = K(\zeta_{q^n-1})$ for some $n \geq 1$, where ζ_{q^n-1} denotes a primitive (q^n-1)th root of unity and q is the cardinality of the residue field κ_K of K. If this is the case, then $n = [\kappa_L : \kappa_K]$ is the degree of the extension of residue fields κ_L/κ_K.*

Proof. Let us first observe that the extension $K(\zeta_{q^n-1})/K$ is unramified; the polynomial is normalized and the reduction of the polynomial $P = X^{q^n-1} - 1$ in $\kappa_K[X]$ is separable, hence Lemma 3.5.3 shows that $K(\zeta_{q^n-1})/K$ is unramified.

Let L/K be an arbitrary finite unramified extension with corresponding extension κ_L/κ_K of residue fields. We want to prove that L is generated by a primitive q^n-1-root of unity. Since L/K is unramified, we have $e = 1$ and

$$[L : K] = f = [\kappa_L : \kappa_K].$$

The group of units $\kappa_L^\times$ is cyclic of order $q^n - 1$ where $n = [\kappa_L : \kappa_K]$. Let $\overline{\alpha}$ be a generator of $\kappa_L^\times$. The polynomial $X^{q^n-1} - 1 \in \kappa_K[X]$ is separable and has $\overline{\alpha}$ as a root. Hence, by Hensel's Lemma, there is a unique root $\alpha \in L$ of

$$X^{q^n-1} - 1 \in L[X]$$

lifting $\overline{\alpha}$. Since α is a root of $X^{q^n-1} - 1$ and its reduction has order $q^n - 1$, we deduce that α is a primitive (q^n-1)-root of unity. We obtain

$$[K(\alpha) : K] \geq [\kappa_K(\overline{\alpha}) : \kappa_K] = [\kappa_L : \kappa_K] = [L : K].$$

On the other hand, we have $K \subseteq K(\alpha) \subseteq L$ and, by degree reasons, we deduce $L = K(\alpha)$. Thus, L is generated by the primitive (q^n-1)-root of unity α and we have

$$[L : K] = n = [\kappa_L : \kappa_K].$$

□

Corollary 3.5.5. *For each positive integer n, there is a unique unramified extension of degree n over K.*

Proof. We denote by q the cardinality of the residue field of K. By Theorem 3.5.4, the unramified extensions of K are exactly the extensions $L = K(\zeta_{q^n-1})/K$ and we have $n = [\kappa_L : \kappa_K] = [L : K]$. □

The totally ramified extensions of $\mathbb{Q}_p$ are intimately related to Eisenstein polynomials. Recall the definition of an Eisenstein polynomial:

Definition 3.5.6. A monic polynomial

$$P = X^n + a_{n-1}X^{n-1} + \cdots + a_1X + a_0 \in \mathcal{O}_K[X]$$

is an *Eisenstein polynomial* if $a_i \in \mathfrak{m}_K$ for $0 \leq i < n$ and $a_0 \notin \mathfrak{m}_K^2$. Note that a_0 is then a uniformizer for A.

We recall the Eisenstein irreducibility criterion.

Lemma 3.5.7 (Eisenstein criterion)**.** *Let $P \in \mathcal{O}_K[X]$ be an Eisenstein polynomial. Then f is irreducible in both $\mathcal{O}_K[X]$ and $K[X]$.*

Proof. See lecture Algebra I. □

Theorem 3.5.8 (Totally ramified extensions)**.** *The extension L/K is totally ramified if and only if $L = K(\pi_L)$ where π_L is the root of an Eisenstein polynomial $P \in \mathcal{O}_K[X]$. If this is the case then π_L is a uniformizer for the discrete valuation ring $\mathcal{O}_L$ of L.*

Proof. Let us assume that L/K is totally ramified of degree $e = [L : K]$ and $\pi_K \in \mathcal{O}_K$ (resp. $\pi_L \in \mathcal{O}_L$) is a uniformizer of K (resp. L). According to Corollary 3.4.3, we have

$$v_p(L^\times) = \frac{1}{e}v_p(K^\times).$$

Since π_K generates $v_p(K^\times)$ and π_L generates $v_p(L^\times)$, we get $ev_p(\pi_L) = v_p(\pi_K)$, or written multiplicatively $|\pi_L|_p^e = |\pi_K|_p$. For all conjugates π_L' of π_L, we have according to Corollary 3.3.10 $|\pi_L|_p = |\pi_L'|_p$. Since each coefficient a_i of the minimal polynomial of π_L

$$P = \prod_{\pi_L'}(X - \pi_L') = a_0 + a_1X + \cdots + a_{e-1}X^{n-1} + X^n \in \mathcal{O}_K[X],$$

(here, π_L' runs over the conjugates of π_L) is a symmetric polynomials of degree $n-i$ in the conjugates π_L' of π_L, we get $|a_i|_p < 1$ and hence $a_i \in \mathfrak{m}_K$. Furthermore, we get for a_0 the formula

$$|a_0|_p = \prod_{\pi_L'} |\pi_L'|_p = |\pi_L|_p^n.$$

This implies $n \geq e$, otherwise a_0 can not[25] be an element of K. On the other hand, we have $n \leq e$ since $K(\pi_L) \subseteq L$. Thus, we have $e = n$ and deduce

$$|a_0|_p = \prod_{\pi_L'} |\pi_L'|_p = |\pi_L|_p^e = |\pi_K|_p. \tag{3.14}$$

Hence, $a_0 \in \mathfrak{m}_K \setminus \mathfrak{m}_K^2$ and we conclude that P is an Eisenstein polynomial.

Conversely, let us assume that π_L is the root of an Eisenstein polynomial $P \in \mathcal{O}_K[X]$ of degree e with constant term a_0. The same reasoning

[25] Since $0 < \nu_p(a_0) = n\nu_p(\pi_L) \in n\nu_p(L^\times) \in \frac{n}{e}\nu_p(K^\times)$, n must be a positive multiple of e.

as in equation (3.14), shows that $|a_0|_p = |\pi_L|_p^e$. Since P is Eisenstein, we have $|a_0|_p = |\pi_K|_p$ and conclude

$$|\pi_L|_p^e = |\pi_K|_p.$$

In particular, L/K has ramification index e. Together with $[L : K] = \deg P = e$, we deduce that L/K is totally ramified. □

Theorem 3.5.9 (Tamely totally ramified extensions)**.** *The extension L/K is tamely and totally ramified if and only if $L = K(\pi_K^{1/e})$ for some uniformizer $\pi_K \in \mathcal{O}_K$ and a positive integer e with $p \nmid e$.*

Proof. Let us assume that L/K is tamely and totally ramified extension of degree e and let us denote by π_K and π_L the uniformizers of K and L. Since L/K is totally ramified, we get $|\pi_K|_p = |\pi_L|_p^e$. This implies that there is a unit $u \in \mathcal{O}_L^\times$ such that

$$u\pi_L^e = \pi_K.$$

Since L/K is totally ramified, L and K have the same residue field $\kappa_K = \kappa_L$. Thus, we can change the uniformizer π_K in such a way that $u \equiv 1 \mod \mathfrak{m}_L$. Thus, let us without loss of generality assume $u \equiv 1 \mod \mathfrak{m}_L$. Let us define $g := X^e - u \in \mathcal{O}_L[X]$. Because L/K is tamely ramified, we have $p \nmid e$. Hence, the reduction $\overline{g} = X^e - 1$ of g is separable. This allows us to apply Hensel's Lemma to find a root β of $\overline{g}$ lifting the root $1 \in \kappa_L$ of $\overline{g}$. We claim that $\alpha := \beta\pi_L$ is a eth root of π_K. Indeed, we have

$$\alpha^e = \beta^e \pi_L^e = u\pi_L^e = \pi_K.$$

Thus, we get $K(\alpha) = K(\pi_L^{1/e}) \subseteq L$. On the other hand, we must have equality since both L/K and $K(\alpha)/K$ are totally ramified of degree e. □

Outlook

The notion of ramification appears in different fields of mathematics, for example in the theory of Riemann surfaces. A Riemann surface is a 1-dimensional connected complex manifold, i.e., roughly it is a geometric object which looks locally like an open subset of $\mathbb{C}$. A morphism of Riemann surfaces X and Y is a map $f\colon X \to Y$ which is locally given by a holomorphic map. The notion of ramification appears in the theory of Riemann surfaces usually in form of the following definition.

Lemma/Definition 3.5.10. For a non-constant morphism $f\colon X \to Y$ of Riemann surfaces and a point $x \in X$, there exists a positive integer e and coordinates z at x and z' at $f(x)$ such that f is locally given in these coordinates by $z \mapsto z^e$. More precisely, we find neighbourhoods U of x and V of $f(x)$ together with charts

$$\varphi\colon U \to U' \subseteq \mathbb{C}, \quad \psi\colon V \to V' \subseteq \mathbb{C}$$

with $x \mapsto 0 \in \mathbb{C}$ and $f(x) \mapsto 0 \in \mathbb{C}$ such that the following diagram commutes

$$\begin{tikzcd}
X \arrow[r] & Y \\
U \arrow[u, hook] \arrow[r] \arrow[d, "\cong"] & V \arrow[u, hook] \arrow[d, "\cong"] \\
U' \arrow[r] & V' \\
z \arrow[r, mapsto] & z^e.
\end{tikzcd}$$

The number e is does only depend on f and x and is called the ramification index of f at x. Let us write $\mathcal{O}_X(U)$ (resp. $\mathcal{O}_Y(V)$) for the holomorphic functions on an open subset $U \subseteq X$ (resp. $V \subseteq Y$). By pre-composition with $f|_U\colon U \to V$, we obtain a pull-back map on holomorphic functions

$$f^*\colon \mathcal{O}_Y(V) \to \mathcal{O}_X(U), \quad g \mapsto g \circ f.$$

We may view the coordinate z' on V as a holomorphic function $z'\colon V \to \mathbb{C}$. By the definition of the ramification index, the pull-back of z' gives in terms of the coordinate z the function $(f^*z')(z) = z^e$. Using the chart ψ, we can view every holomorphic function g on V as a function defined on an open neighbourhood of 0 in $\mathbb{C}$. Since a holomorphic function in a neighbourhood of 0 is uniquely determined by its power series expansion, we obtain an injective map

$$\mathcal{O}_Y(V) \to \mathbb{C}[\![z']\!].$$

Similarly, using the chart φ we obtain an injective map

$$\mathcal{O}_X(U) \to \mathbb{C}[\![z]\!].$$

Since f^* identifies z' with z^e, we can summarize the above discussion by a commutative diagram

$$\begin{tikzcd}
\mathcal{O}_Y(V) \arrow[r, "f^*"] \arrow[d, hook] & \mathcal{O}_X(U) \arrow[d, hook] \\
\mathbb{C}[\![z']\!] \arrow[r] & \mathbb{C}[\![z]\!]. \\
z' \arrow[r, mapsto] & z^e.
\end{tikzcd}$$

We are now in a situation, where we can relate the ramification of Riemann surfaces to the ramification theory of Dedekind rings. Therefore, let us observe that the ring

$$A := \mathbb{C}[\![z']\!]$$

is a local domain with exactly one non-zero prime ideal ideal $(z') \subseteq \mathbb{C}[\![z']\!]$. Hence, it is a discrete valuation ring. The associated discrete valuation is given by $v(0) = \infty$ and

$$v(\sum_{k=0}^{\infty} c_k z'^k) := \min\{k \geq 0 \mid c_k \neq 0\}.$$

The maximal ideal of this ring is $\mathfrak{m}_A := z'\mathbb{C}[\![z']\!]$ and we write $K = \mathbb{C}(\!(z')\!)$ for its fraction field. By the same argumet for z instead of z', we obtain a discretely valued field $L = \mathbb{C}(\!(z)\!)$ with discrete valuation ring $B = \mathbb{C}[\![z]\!]$ and maximal ideal $\mathfrak{m}_L := (z) \subseteq B$. The map $z' \mapsto z^e$ induces an extension of valued fields $K \subseteq L$. Furthermore, we have $\mathfrak{m}_K \cdot B = z^e B = \mathfrak{m}_L^e$. Hence, the extension L/K is a ramified extension of degree e. This relates the ramification of Riemann surfaces to the ramification theory of Dedekind domains. Using the inclusion

$$\mathcal{O}_X(U) \hookrightarrow \mathbb{C}[\![z]\!],$$

induced by $x \in X$ and the local coordinate z near x, we obtain a valuation $v_x \colon \mathcal{O}_X(U) \to \mathbb{Z}$. This valuation has a quite concrete interpretation. The value $v_x(g)$ is exactly the vanishing order of the function g at $x \in X$. Of course, it extends to the fraction field of $\mathcal{O}_X(U)$ which is the field of meromorphic functions $\mathcal{M}_X(U)$ on U and gives the order of a meromorphic function at $x \in X$.

In a certain sense, the above picture describes the local situation at a point $x \in X$. Let us briefly outline the global picture. The meromorphic functions on a Riemann surface form a field $\mathcal{M}_X(X)$. As we have seen above, every point $x \in X$ gives us a valuation v_x on $\mathcal{M}_X(X)$ and it can be shown that the completion of $\mathcal{M}_X(X)$ with respect to v_x is exactly the field $\mathbb{C}(\!(z)\!)$. By pull-back of meromorphic functions, a non-constant morphism of Riemann surfaces $X \to Y$ gives a field extension $\mathcal{M}_Y(Y) \subseteq \mathcal{M}_X(X)$ of the fields of meromorphic functions. It turns out, that the theory of (compact) Riemann surfaces behaves quite similar to the theory of number fields. In this analogy, the number field corresponds to the field of meromorphic functions and the primes correspond to points of the Riemann surface.

The above discussion can be used to think about points of a Riemann surface in at least two different ways:

(a) Points are like primes.

(b) Points are like valuations.

If one takes this seriously, than each of the above ways of thinking about points leads to a rich geometric theory. (a) leads to Algebraic Geometry[26], while (b) leads to Rigid Analytic Geometry.

3.6 Krasner's Lemma

In this section, we will prove a small Lemma due to Krasner which has some surprising consequences. For example, it can be used to prove that there are only finitely many extensions of $\mathbb{Q}_p$ of bounded degree. Of course, this is not true for number fields like $\mathbb{Q}$. This already indicates that it might be much easier to classify all finite abelian extensions of $\mathbb{Q}_p$ and motivates the 'local' approach towards the Kronecker-Weber Theorem.

Notation 3.6.1. Let us fix an algebraic closure $\overline{\mathbb{Q}}_p$ of $\mathbb{Q}_p$ and write $|\cdot|_p$ for the unique extension of the p-adic absolute value on $\mathbb{Q}_p$.

Definition 3.6.2. Let K be a finite extension of $\mathbb{Q}_p$ in $\overline{\mathbb{Q}}_p$. For $\alpha \in \overline{\mathbb{Q}}_p$ let us denote by $\alpha = \alpha_1, \alpha_2, \dots, \alpha_n$ the conjugates of α over K, i.e., the roots of the minimal polynomial of α over K. We say that $\beta \in \overline{\mathbb{Q}}_p$ *belongs to α over K* if $|\beta - \alpha|_p < |\beta - \alpha_i|_p$ for all $2 \leq i \leq n$.[27]

Our next aim is to study tamely and totally ramified extensions of $\mathbb{Q}_p$.

Theorem 3.6.3 (Krasner's Lemma)**.** *Let $\mathbb{Q}_p \subseteq K \subseteq \overline{\mathbb{Q}}_p$ be a finite extension of $\mathbb{Q}_p$ and $\alpha, \beta \in \overline{\mathbb{Q}}_p$. If β belongs to α over K then $K(\alpha) \subseteq K(\beta)$.*

Proof. Suppose that $\alpha \notin K(\beta)$. Then, there is an embedding $\sigma \in \mathrm{Gal}(\overline{\mathbb{Q}}_p/K(\beta))$ for which $\sigma(\alpha) \neq \alpha$, i.e., $\sigma(\alpha) = \alpha_i$ for some $2 \leq i \leq n$. By Corollary 3.3.10, we have

$$|\beta - \alpha|_p = |\sigma(\beta - \alpha)|_p = |\sigma(\beta) - \sigma(\alpha)|_p = |\beta - \alpha_i|_p,$$

but this contradicts the hypothesis that β belongs to α. □

Although Krasner's Lemma is an easy consequence of the Galois invariance of the p-adic absolute value, it has many interesting consequences. Let us recall that we have introduced for a polynomial $f = a_0 + a_1 X + \cdots + a_n X^n \in \mathcal{O}_K[X]$ the notation

$$|f| = \max_{i=0,\dots,n} |a_i|_p.$$

Krasner's Lemma is the main ingredient to prove the following Proposition:

Proposition 3.6.4. *Let $f \in \mathcal{O}_K[X]$ be a normalized irreducible polynomial with roots $\alpha_1, \dots, \alpha_n \in \overline{\mathbb{Q}}_p$. For any $g \in \mathcal{O}_K[X]$ satisfying*

$$|f - g| < \frac{\min_{i \neq j} |\alpha_i - \alpha_j|_p^n}{2^n}$$

[26] At least, for compact Riemann surfaces one can make this vague analogy between valuations and prime ideals very precise.

[27] So β is closer to α than to any conjugate of α over K.

and for any root $\alpha \in \overline{\mathbb{Q}}_p$ of f, there exists a unique root $\beta \in \overline{\mathbb{Q}}_p$ of g with $K(\alpha) = K(\beta)$. In particular, g is irreducible and has the same splitting field as f.

Proof. Exercises. □

The above Proposition can be used to prove that there are only finitely many extensions of $\mathbb{Q}_p$ of a fixed degree. More precisely, we have.

Corollary 3.6.5. *For a given finite extension K of $\mathbb{Q}_p$ and a positive integer n, there are only finitely many extensions of K of degree n in $\overline{\mathbb{Q}}_p$.*

Proof. Exercise. □

Furthermore, we can use the above Proposition to prove that all finite extensions of $\mathbb{Q}_p$ are obtained as completions of number fields.

Corollary 3.6.6. *Let $\hat{K}/\mathbb{Q}_p$ be a finite extensions of $\mathbb{Q}_p$. Then, there exists a finite extension K of $\mathbb{Q}$ and a prime $\mathfrak{p}$ over p such that $\hat{K}$ is the completion of K at $|\cdot|_{\mathfrak{p}}$.*

Proof. Exercises. □

3.7 Infinite Galois theory

In this section, we give a very brief introduction to infinite Galois theory. Let us first recall that the Galois group of a (not necessarily finite) Galois extension L/K is defined as the group of all K-linear field automorphisms of L, i.e., $\mathrm{Gal}(L/K) := \mathrm{Aut}_K(L)$. We recall the following Galois correspondence for finite Galois extensions.

Theorem 3.7.1 (Finite Galois correspondence)**.** *Let L/K be a finite Galois extension of fields with Galois group $G = Gal(L/K)$. There is an inclusion reversing bijection*

$$\begin{tikzcd}
\{\textit{subextensions } L/F/K\} \arrow[r] & \{\textit{subgroups } H \subseteq G\} \\
F \arrow[r, mapsto] & Gal(L/F) \\
L^H & H. \arrow[l, mapsto]
\end{tikzcd}$$

Under this correspondence, a normal field extension F/K corresponds to a normal subgroup $H \subseteq G$ and we have $Gal(F/K) \cong G/H$.

Next, we want to generalize the Galois correspondence to arbitrary (not necessarily finite) Galois extensions. The maps $F \mapsto \mathrm{Gal}(L/F)$ and $H \mapsto L^H$ from the finite Galois correspondence are still defined for infinite Galois groups. Unfortunately, it turns out that the Galois correspondence fails in general for infinite extensions as the following example shows.

Example 3.7.2. For a prime p, we have the degree two Galois extension $\mathbb{Q}(\sqrt{p})/\mathbb{Q}$ with Galois group $\mathrm{Gal}(\mathbb{Q}(\sqrt{p})/\mathbb{Q}) \cong \mathbb{F}_2$, where $1 \in \mathbb{F}_2$ corresponds to the map

$$\mathbb{Q}(\sqrt{p}) \to \mathbb{Q}(\sqrt{p}), \quad \sqrt{p} \mapsto -\sqrt{p}.$$

Let us now take the compositum of all these fields $L = \mathbb{Q}(\sqrt{p} \mid \text{prime})$. Of course, this is again a Galois extension of $\mathbb{Q}$ of infinite degree. It is easily checked that the Galois group of this extension is

$$\mathrm{Gal}(L/\mathbb{Q}) = \prod_{p \text{ prime}} \mathbb{F}_2.$$

This group contains uncountably many index 2 subgroups[28]. On the other hand, there are only countably many quadratic extensions of $\mathbb{Q}$. Thus, the Galois correspondence in the above form can not hold for infinite Galois extensions.

The above example shows that there are in general way to many subgroups of an infinite Galois group to make the Galois correspondence work. Thus, we need a way to single out the 'relevant' subgroups of $\mathrm{Gal}(L/K)$. As a motivation, let us come back to our example:

Example 3.7.3. Let us again consider the Galois extension $L = \mathbb{Q}(\sqrt{p} \mid \text{prime})$ over $\mathbb{Q}$. Every finite Galois sub-extension of $L/\mathbb{Q}$ is contained in an extension of the form

$$F = \mathbb{Q}(p \mid p \in J),$$

where J is a finite subset of the set $\mathbb{P}$ of all primes. The corresponding Galois group of F is given by the subgroup

$$\mathrm{Gal}(L/F) = \prod_{p \in \mathbb{P}\setminus J} \mathbb{F}_2 \times \prod_{p \in J} \{0\}.$$

Note, that the subgroups of the above kind form a basis for the product topology of $0 \in \prod_p \mathbb{F}_2$, when we equip $\mathbb{F}_2$ with the discrete topology. Thus, we can single out the 'relevant' subgroups corresponding to all finite sub-extensions by introducing a suitable topology on $\mathrm{Gal}(L/\mathbb{Q})$. In this case, a subgroup $U \subseteq \mathrm{Gal}(L/\mathbb{Q})$ corresponds to a finite extension $L/\mathbb{Q}$ if and only if it is an open subgroup with respect to the product topology on $\mathrm{Gal}(L/\mathbb{Q}) = \prod_{p \text{ prime}} \mathbb{F}_2$.

We want to use the above example as a motivation to find the 'relevant' subgroups corresponding to sub-extensions. It was Krull who observed that this can be done by introducing a suitable topology on the Galois group. Let us recall that a *topological group* is a group G equipped with a topology such that the multiplication

$$m \colon G \times G \to G$$

[28] An index 2 subgroup of $V := \prod_{p \text{ prime}} \mathbb{F}_2$ corresponds to a surjection

$$\prod_{p \text{ prime}} \mathbb{F}_2 \twoheadrightarrow \mathbb{F}_2.$$

Note, that every non-zero element of the $\mathbb{F}_2$ vector space

$$V^* = \mathrm{Hom}_{\mathbb{F}_2}\left(\prod_{p \text{ prime}} \mathbb{F}_2, \mathbb{F}_2\right)$$

corresponds to such a surjection. Now, the claim follows from the fact that V^* is uncountable.

and the inversion

$$i\colon G \to G, \quad g \mapsto g^{-1}$$

are continuous maps. We define the following topology on a Galois group.

Definition 3.7.4. Let L/K be a Galois extension of fields. We equip $\mathrm{Gal}(L/K)$ with the topology given by the following basis of open neighbourhoods. For $\sigma \in \mathrm{Gal}(L/K)$ a family of open neighbourhoods is given by the family of cosets $(\sigma \cdot \mathrm{Gal}(L/F))_F$, where F runs over all finite sub-extensions of K. This topology is called *Krull topology*.

Let us first observe, that for a finite Galois extension L/K the Krull topology on the Galois group $\mathrm{Gal}(L/K)$ is the discrete topology; indeed, every singleton $\{\sigma\}$ is open since it is of the form $\sigma\,\mathrm{Gal}(L/L)$ for the finite extension L over K. In particular, every subgroup in a finite Galois group is open and closed with respect to the Krull topology. Thus, the following Theorem is really a generalization of the finite Galois correspondence.

Theorem 3.7.5 (Galois correspondence)**.** *Let L/K be a Galois extension of fields with Galois group $G = Gal(L/K)$. There is an inclusion reversing bijection*

$$\begin{tikzcd}
\{\text{subextensions } L/F/K\} \arrow[r] & \{\textbf{closed} \text{ subgroups } H \subseteq G\} \\
F \arrow[r, mapsto] & Gal(L/F) \\
L^H & H. \arrow[l, mapsto]
\end{tikzcd}$$

Under this correspondence, a normal field extension F/K corresponds to a normal subgroup $H \subseteq G$ and we have $Gal(F/K) \cong G/H$. The ***finite*** *field extensions F/K correspond exactly to the* ***open*** *subgroups of G.*

Proof. For a proof, see Theorem 7.2 in Milne's notes on Fields and Galois Theory[29]. □

A pro-finite group is a topological group that is isomorphic to an inverse limit of an inverse system of discrete[30] finite groups. Let us explain these notions: Recall that a *directed set* is a non-empty set I together with a reflexive and transitive binary relation $\leq$ with the property that for any $i, j \in I$, there is an element $k \in I$ such that $i \leq k$ and $j \leq k$. An *inverse system* of discrete finite groups consists of a directed set $(I, \leq)$, a collection of discrete finite groups $\mathcal{G} = \{G_i \mid i \in I\}$ and transition maps $f_i^j\colon G_j \to G_i$ whenever $i \leq j$ such that $f_i^i = \mathrm{id}_{G_i}$ and the collection satisfies the composition property $f_i^j \circ f_j^k = f_i^k$. The inverse limit of the inverse system $((I, \leq), \mathcal{G}, (f_i^j)_{i,j})$ can now be described explicitly as[31]

[29] James S. Milne. Fields and galois theory (v4.61), 2020. Available at www.jmilne.org/math/

[30] I.e., we equip these finite groups with the discrete topology.

[31] Of course, it also satisfies a universal property, see Exercise sheet 8.

$$\varprojlim_i G_i = \left\{ (g_i)_{i\in I} \in \prod_{i\in I} G_i \mid f_i^j(g_j) = g_i \forall i \leq j \right\}.$$

We equip $\varprojlim_i G_i$ with the subspace topology of the product $\prod_i G_i$.

Example 3.7.6. The positive integers $(\mathbb{N}, |)$ with the binary operation $n \mid m$ given by divisibility form a directed set (we will check this in the Exercises). For $n \mid m$, we have canonical projections

$$f_n^m \colon \mathbb{Z}/m\mathbb{Z} \twoheadrightarrow \mathbb{Z}/n\mathbb{Z}.$$

This gives a directed system $((\mathbb{N}, |), (\mathbb{Z}/n\mathbb{Z})_n, (f_n^m)_{m,n})$ of finite groups. The associated profinite group

$$\varprojlim_{n\in(\mathbb{N},|)} \mathbb{Z}/n\mathbb{Z}$$

is denoted by $\widehat{\mathbb{Z}}$.

By Tychonoff's Theorem, this product is compact and $\varprojlim_i G_i$ is a closed subset of this compact topological group. Hence, every profinite topological group is compact. Even better, we have the following purely topological characterization of pro-finite groups:

Theorem 3.7.7. *A topological group G is pro-finite if and only if G is compact, Hausdorff*[32] *and totally disconnected*[33]*. If G is pro-finite, then we have a canonical isomorphism of topological groups*

$$G \xrightarrow{\sim} \varprojlim_i G/U,$$

where U runs over all open normal subgroups of finite index.

Proof. See, for example Theorem 2.1.3 in the book Profinite Groups by L. Ribes and P. Zalesskii[34]. □

Now, one can check that the Krull topology is a profinite group.

Theorem 3.7.8. *The Krull topology makes $Gal(L/K)$ a compact Hausdorff and totally disconnected topological group.*

Proof. See Proposition 7.8 in Milne's notes[35]. □

Applying the above characterization of pro-finite groups to $\mathrm{Gal}(L/K)$, we obtain.

Theorem 3.7.9. *For any Galois extension L/K, we have an isomorphism of topological groups*

$$Gal(L/K) \xrightarrow{\sim} \varprojlim_F Gal(F/K),$$

[32] Recall that a topological space is Hausdorff if and only if for each two distinct points there are neighbourhoods of each of the two points which are distinct.

[33] A topological space is called *totally disconnected* if and only if the only connected subsets are singletons.

[34] L. Ribes and P. Zalesskii. *Profinite Groups*. A Series of Modern Surveys in Mathematics. Springer Berlin Heidelberg, 2010. ISBN 9783642016424

[35] James S. Milne. Fields and galois theory (v4.61), 2020. Available at www.jmilne.org/math/

where F runs through all finite Galois extensions of K in L. More precisely, the directed set is the set

$$I = \{F \mid F \textit{ finite Galois subextension}\}$$

with the binary relation $\subseteq$ given by inclusion of subfields. For $F_1, F_2 \in I$ with $F_1 \subseteq F_2$, the transition maps is given by the canonical surjection

$$Gal(F_2/K) \twoheadrightarrow Gal(F_1/K).$$

Proof. This is a direct consequence of Theorem 3.7.8 and Theorem 3.7.7. □

As an example, let us compute the Galois group of the field obtained by adjoining all p-power roots of unity to $\mathbb{Q}_p$:

Example 3.7.10. Let p be a prime and consider the field extension $L = \mathbb{Q}_p(\zeta_{p^j} \mid j \in \mathbb{N})$ over $\mathbb{Q}_p$. By Theorem 3.7.9, we have

$$\mathrm{Gal}(L/\mathbb{Q}_p) = \varprojlim_{j \in (\mathbb{N}, \leq)} \mathrm{Gal}(\mathbb{Q}_p(\zeta_{p^j})/\mathbb{Q}_p).$$

The Galois groups $\mathrm{Gal}(\mathbb{Q}_p(\zeta_{p^j})/\mathbb{Q}_p)$ can be computed explicitly

$$(\mathbb{Z}/p^j\mathbb{Z})^\times \xrightarrow{\sim} \mathrm{Gal}(\mathbb{Q}_p(\zeta_{p^j})/\mathbb{Q}_p), \quad k \mapsto (\zeta_{p^j} \mapsto \zeta_{p^j}^k)$$

and we obtain

$$\mathrm{Gal}(L/\mathbb{Q}_p) = \varprojlim_{j \in (\mathbb{N}, \leq)} \mathrm{Gal}(\mathbb{Q}_p(\zeta_{p^j})/\mathbb{Q}_p) \cong \varprojlim_{j \in (\mathbb{N}, \leq)} (\mathbb{Z}/p^j\mathbb{Z})^\times \cong \mathbb{Z}_p^\times.$$

Similarly, we can compute the Galois group of the field obtained by adjoining all roots of unity of prime-to-p order to $\mathbb{Q}_p$:

Example 3.7.11. Let p be a prime and consider the field extension $K := \mathbb{Q}_p(\zeta_n \mid n \text{ prime to } p)$ over $\mathbb{Q}_p$. We obtain

$$\mathrm{Gal}(K/\mathbb{Q}_p) = \varprojlim_{n \in (\mathbb{N}, |) \text{ prime to } p} \mathrm{Gal}(\mathbb{Q}_p(\zeta_n)/\mathbb{Q}_p).$$

For n which is prime to p, all the extensions $\mathbb{Q}_p(\zeta_n)/\mathbb{Q}_p$ are unramified. By Theorem 3.5.4, we get

$$\mathrm{Gal}(K/\mathbb{Q}_p) = \varprojlim_{m \in (\mathbb{N}, |)} \mathrm{Gal}(\mathbb{Q}_p(\zeta_{p^m-1})/\mathbb{Q}_p).$$

For $m \in \mathbb{N}$, we have

$$\mathbb{Z}/m\mathbb{Z} \xrightarrow{\sim} \mathrm{Gal}(\mathbb{Q}_p(\zeta_{p^m-1})/\mathbb{Q}_p, \quad k \mapsto \left(\zeta_{p^m-1} \mapsto \zeta_{p^m-1}^{p^k}\right).$$

Thus, we obtain

$$\mathrm{Gal}(K/\mathbb{Q}_p) \cong \varprojlim_{m \in (\mathbb{N}, |)} \mathbb{Z}/m\mathbb{Z} \cong \widehat{\mathbb{Z}}.$$

Sometimes, it is useful to compute an inverse limit using a *co-final directed subset* $(J, \leq) \subseteq (I, \leq)$, i.e., a subset $J \subseteq I$ which is directed with the binary operation $\leq$ on I and is *co-final* in the following sense

$$\forall i \in I \exists j \in J \text{ such that } i \leq j.$$

For a co-final directed subset $(J, \leq) \subseteq (I, \leq)$, it is not difficult to check that we have an isomorphism

$$\varprojlim_{i \in (I, \leq)} G_i \cong \varprojlim_{j \in (J, \leq)} G_j.$$

For example, the even positive integers form a co-final subset in $(\mathbb{N}, \leq)$ and we have

$$\mathbb{Z}_p = \varprojlim_{n \in (\mathbb{N}, \leq)} \mathbb{Z}/p^k\mathbb{Z} = \varprojlim_{n \in (2\mathbb{N}, \leq)} \mathbb{Z}/p^k\mathbb{Z} = \varprojlim_{n \in (\mathbb{N}, \leq)} \mathbb{Z}/p^{2k}\mathbb{Z}.$$

3.8 Kummer theory

In this section, we give a brief introduction to Kummer theory. Kummer theory provides an explicit description of all cyclic extensions of degree n for arbitrary fields which are of characteristic prime to n and contain all n-th roots of unity.

So let us first fix this setup for the following section.

Notation 3.8.1. Let n be a positive integer and K be an arbitrary field such that the characteristic of K does not divide n. Let us fix a separable closure $\overline{K}$ of K. Let us furthermore assume that K contains all n-th roots of unity $\mu_n(\overline{K})$.

To motivate Kummer theory, let us start with the following observation. For any field K as above and any $\alpha \in K$, the field $K(\sqrt[n]{\alpha})$ obtained by adjoining any n-th root of K is a splitting field[36] for the polynomial $f = X^n - \alpha$. By our assumption that n is not divisible by the characteristic of K, the polynomial $f \in K[X]$ is separable. Hence, L/K is a Galois extension. Furthermore, we have an injective homomorphism

$$\begin{aligned} \mathrm{Gal}(L/K) &\longrightarrow \mu_n(\overline{K}), \\ \sigma &\longmapsto \frac{\sigma(\sqrt[n]{\alpha})}{\sqrt[n]{\alpha}}. \end{aligned}$$

This homomorphism is an isomorphism if and only if $X^n - \alpha$ is irreducible[37]. In particular, since any subgroup of a cyclic group is again cyclic the Galois group of the extension L/K is cyclic. We can summarize the above discussion as follows. Adjoining a n-th root of an arbitrary element of K gives us a cyclic[38] Galois extension of K. One

[36] Since K contains all n-th roots of unity, the field $K(\sqrt[n]{\alpha})$ contains all roots of f.

[37] Recall from algebra that the Galois group of the splitting field of a polynomial acts transitively on the roots if and only if the polynomial is irreducible.

[38] Here and in the following, we use the terminology 'cyclic Galois extension' or 'cyclic field extension' as a short form for 'Galois extension with a cyclic Galois group'

aspect of Kummer theory is to prove the converse. In particular, Kummer theory says that any cyclic extension of K of degree n is obtained by adjoining an n-th root of a suitable element of K.

For the proof, we will need the following statement.

Proposition 3.8.2 (Linear independence of automorphisms)**.** *Let L/K be a finite extension of fields. Then $Aut_K(L)$ is a linearly independent subset of the L-vector space of all K-linear maps $L \to L$.*

Proof. Suppose the set $\mathrm{Aut}_K(L)$ is linearly dependent. Then, there is a non-trivial linear combination of minimal length r

$$\varphi = c_1\sigma_1 + \cdots + c_r\sigma_r = 0$$

with $c_i \in L^\times$ and pairwise distinct automorphisms $\sigma_i \in \mathrm{Aut}_K(L)$ for $i = 1, \ldots, r$. Of course, such a linear combination can not have length 1, i.e., $r > 1$. Since $\sigma_1 \neq \sigma_r$, we find an element $\alpha \in L$ with $\sigma_1(\alpha) \neq \sigma_r(\alpha)$. We have $\varphi(\beta) = 0$ for any $\beta \in L$, and hence we get

$$\varphi(\alpha\beta) - \sigma_1(\alpha)\varphi(\beta) = 0.$$

But the latter relation is of the form

$$c_2'\sigma_2 + \cdots + c_r'\sigma_r = 0$$

for $c_i' := c_i\sigma_i(\alpha) - c_i\sigma_1(\alpha)$. Hence we have found a shorter linear dependence relation in $\mathrm{Aut}_K(L)$. Note, that the latter relation is non-trivial since $c_r \neq 0$ by our choice of α. This gives a contradiction to the minimality of r and hence $\mathrm{Aut}_K(L)$ is linearly independent. □

We have the following Corollary.

Corollary 3.8.3. *Let L/K be a finite Galois extension of fields with a cyclic Galois group $Gal(L/K) = \langle\sigma\rangle$ of order n. For every n-th root of unity $\zeta_n \in \mu_n(L)$ of L, there is an element $x \in L$ such that*

$$\sigma(x) = \zeta_n x.$$

Proof. The automorphism σ is a K-linear map on L with characteristic polynomial $X^n - 1 \in K[X]$. By the linear independence of automorphisms, the polynomial $X^n - 1$ must be its minimal polynomial, since the set $\{1, \sigma, \sigma^2, \ldots, \sigma^{n-1}\}$ is linearly independent. Therefore, $\zeta_n \in L$ is an eigenvalue of σ and we deduce that there exists an eigenvector $x \in L$ such that

$$\sigma(x) = \zeta_n x.$$

□

Using this corollary, we can now prove that any cyclic extension of K is obtained by adjoining an n-th root of K.

Corollary 3.8.4. *Let L/K is a cyclic extension of degree n which not divisible by the characteristic of K and assume that K contains all n-th roots of unity. Then $L = K(\sqrt[n]{\alpha})$ for some $\alpha \in K$.*

Proof. Let L/K be a cyclic Galois extension with Galois group $\mathrm{Gal}(L/K) = \langle \sigma \rangle$. By Corollary 3.8.3, we find for a primitive n-th root $\zeta_n \in K$ an element $x \in L$ such that

$$\sigma(x) = \zeta_n x.$$

We have

$$\sigma(x^n) = \sigma(x)^n = \zeta_n^n x^n = x^n.$$

Thus, $\alpha := x^n$ is invariant under $\langle \sigma \rangle = \mathrm{Gal}(L/K)$ and hence it is contained in K by Galois Theory. Moreover, the orbit of x under the Galois action $x, x\zeta_n, \ldots, x\zeta_n^{n-1}$ has length n since ζ_n was a primitive n-th root of unity, so $[K(x) : K] = n$. Because $K(x)$ is a sub-field of the degree n extension L/K, we must have equality, i.e., $L = K(x)$ for the n-th root x of $\alpha \in K$. □

Let us now introduce the Kummer pairing.

Lemma/Definition 3.8.5. Let K be a field as in Notation 3.8.1. We have a well-defined bilinear pairing

$$\begin{aligned} \langle \cdot, \cdot \rangle \colon \mathrm{Gal}(\overline{K}/K) \times (K^\times / K^{\times n}) &\longrightarrow \mu_n(K) \\ (\sigma, \alpha) &\longmapsto \tfrac{\sigma(\sqrt[n]{\alpha})}{\sqrt[n]{\alpha}}. \end{aligned}$$

This pairing is called the *Kummer pairing*.

Proof. Let us first show that the pairing does not depend on the choice of the choosen n-th root of α. Let $x, y \in \overline{K}$ be two roots of α. Then $x = \zeta_n y$ for some n-th root of unity ζ_n which is fixed by $\mathrm{Gal}(\overline{K}/K)$. Hence we get

$$\frac{\sigma(x)}{x} = \frac{\sigma(\zeta_n y)}{\zeta_n y} = \frac{\zeta_n \sigma(y)}{\zeta_n y} = \frac{\sigma(y)}{y}.$$

In order to prove that the Kummer pairing is well-defined modulo $K^{\times n}$ in the second component, it suffices to show

$$\langle \sigma, \alpha^n \rangle = 1, \quad \forall \alpha \in K^\times, \sigma \in \mathrm{Gal}(\overline{K}/K).$$

But of course, we can pick $\alpha \in K$ as an n-th root of α^n, hence we get

$$\langle \sigma, \alpha^n \rangle = \frac{\sigma(\alpha)}{\alpha} = 1.$$

By its definition, the Kummer pairing is bilinear. □

We are now ready to prove the main theorem of this section.

Theorem 3.8.6 (Kummer Theory)**.** *Let K be a field containing all n-th roots of unity $\mu_n(\overline{K})$ where n is a positive integer which not divisible by the characteristic of K. Then the Kummer pairing induces an isomorphism*

$$\begin{aligned}\Phi\colon K^\times/K^{\times n} &\longrightarrow Hom_{cont}(Gal(\overline{K}/K), \mu_n(K))\\ \alpha &\longmapsto (\sigma \mapsto \langle \sigma, \alpha\rangle),\end{aligned}$$

where $Hom_{cont}(Gal(\overline{K}/K), \mu_n(K))$ denotes the group of all continuous group homomorphisms from the absolute Galois group of K to the discrete abelian group$\mu_n(K)$ of all n-th roots of unity.

Proof. Injectivity: For $\alpha \in K^\times \setminus K^{\times n}$, the extension $K(\sqrt[n]{\alpha})$ is non-trivial and hence some $\sigma \in \mathrm{Gal}(\overline{K}/K)$ will act non-trivially. For such an element $\sigma \in \mathrm{Gal}(\overline{K}/K)$, we have $\langle\sigma, \alpha\rangle \neq 1$, so $\alpha \notin \ker\Phi$. This proves the injectivity of Φ.
Surjectivity: Now, let $\varphi\colon\ \mathrm{Gal}(\overline{K}/K) \to \mu_n(K)$ be a homomorphism with an image of order d. Let us write $H := \ker\varphi$ and denote by $L := \overline{K}^H$ the fixed field of φ, so φ factors as

$$\mathrm{Gal}(\overline{K}/K) \twoheadrightarrow \mathrm{Gal}(\overline{K}/K)/H \xrightarrow{\sim} \mu_d(K)$$

By the continuity of φ, we deduce that H is an open subgroup in the Krull topology and hence L/K is finite. More precisely, we have $\mathrm{Gal}(L/K) = \mathrm{Gal}(\overline{K}/K)/H \cong \mathbb{Z}/d\mathbb{Z}$, so L/K is a cyclic extension of degree d. Corollary 3.8.4 shows that $L = L(\sqrt[d]{\alpha})$ for some $\alpha \in K$. In this way, we obtain a morphism

$$\Phi(\alpha)\colon\ \mathrm{Gal}(\overline{K}/K) \twoheadrightarrow \mathrm{Gal}(\overline{K}/K)/H \xrightarrow{\sim} \mu_d(K).$$

At this point, we do not yet know that $\Phi(\alpha)$ does coincide with the given map φ. But for any $m \in (\mathbb{Z}/d\mathbb{Z})^\times$ and $e := n/d$, we obtain a continuous homomorphism

$$\Phi(\alpha^{em}).$$

For distinct choices of $m \in (\mathbb{Z}/d\mathbb{Z})^\times$, these maps are distinct since each choice of m gives different classes in $\alpha^{em} \in K^\times/K^{\times n}$ and Φ is injective. On the other hand, there are exactly $\#(\mathbb{Z}/d\mathbb{Z})^\times = \#\,\mathrm{Aut}(\mathbb{Z}/d\mathbb{Z})$ distinct homomorphisms of the form

$$\mathrm{Gal}(\overline{K}/K) \twoheadrightarrow \mathrm{Gal}(\overline{K}/K)/H \xrightarrow{\sim} \mu_n(K).$$

Thus, we have $\varphi = \Phi(\alpha^{me})$ for some $m \in (\mathbb{Z}/d\mathbb{Z})^\times$. This proves the surjectivity. □

Let us consider the following example:

Example 3.8.7. For $K = \mathbb{Q}$ and $n = 2$, the 2nd roots of unity $\{\pm 1\}$ are contained in $\mathbb{Q}$ and the characteristic of $\mathbb{Q}$ does not divide 2. Hence, we may apply Kummer's Theorem and obtain

$$\mathbb{Q}^\times/\mathbb{Q}^{\times 2} \cong \mathrm{Hom}_{\mathrm{cont}}(\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}), \{\pm 1\}).$$

Each element of $\mathbb{Q}^\times/\mathbb{Q}^{\times 2}$ is determined by a unique square-free integer d and with this identification, the Kummer pairing is given by

$$\{d \in \mathbb{Z} \mid d \text{ square-free}\} \xrightarrow{\sim} \mathrm{Hom}_{\mathrm{cont}}(\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}), \{\pm 1\}).$$

which is given by

$$d \mapsto \left(\sigma \mapsto \frac{\sigma(\sqrt{d})}{\sqrt{d}}\right).$$

Thus, Kummer Theory for $n = 2$ and $K = \mathbb{Q}$ reflects the fact that the quadratic extensions of $\mathbb{Q}$ are precisely the fields obtained by adjoining a square root of a square-free integer.

In the case, where the n-th roots of unity are not contained in K, the following lemma will bes useful.

Lemma 3.8.8. *Let p be a prime and F be a field of characteristic prime to p. Let $L = F(\zeta_p, \sqrt[p]{\alpha})$ for some $\alpha \in F(\zeta_p)^\times$. Define the homomorphism*

$$\omega\colon \mathit{Gal}(F(\zeta_p)/F) \to (\mathbb{Z}/p\mathbb{Z})^\times,$$

by $\sigma(\zeta_p) = \zeta_p^{\omega(\sigma)}$. If L/F is abelian then $\frac{\sigma(\alpha)}{\alpha^{\omega(\sigma)}} \in F(\zeta_p)^{\times p}$ for all $\sigma \in \mathit{Gal}(F(\zeta_p)/F)$.

Proof. Let $G = \mathrm{Gal}(L/F)$, $H = \mathrm{Gal}(L/F(\zeta_p)) \subseteq G$ and let A be the subgroup of $F(\zeta_p)^\times/F(\zeta_p)^{\times p}$ generated by α. The Kummer pairing induces a bilinear pairing $H \times A \to \mu_p(\overline{K})$ that is compatible with the Galois action of $\mathrm{Gal}(F(\zeta_p)/F) \cong G/H$

$$\langle h, \alpha^{\omega(\sigma)}\rangle = \langle h, \alpha\rangle^{\omega(\sigma)} = \sigma(\langle h, \alpha\rangle) = \langle h, \sigma(\alpha)\rangle,$$

for all $\sigma \in \mathrm{Gal}(F(\zeta_p)/F)$ and $h \in H$. The isomorphism Φ induced by the Kummer pairing is injective, so $\alpha^{\omega(\sigma)} \equiv \sigma(\alpha) \mod F(\zeta_p)^{\times p}$. □

Proposition 3.8.9. *For $p > 2$ no extension of $\mathbb{Q}_p$ has Galois group isomorphic to $(\mathbb{Z}/p\mathbb{Z})^k$ for $k \geq 3$.*

Proof. It is enough to prove that there is no Galois extension with Galois group isomorphic to $(\mathbb{Z}/p\mathbb{Z})^3$. Suppose there were a extension $K/\mathbb{Q}_p$ with $\mathrm{Gal}(K/\mathbb{Q}_p) \cong (\mathbb{Z}/p\mathbb{Z})^3$. Let us choose the uniformizer[39] $\pi := \zeta_p - 1$ of the field $\mathbb{Q}_p(\zeta_p)$.
Claim 1: There exists a subgroup $A \subseteq U_1/U_1^p$ where $U_1 := \{u \equiv 1 \mod \pi\}$ is the group of principal units of the field $\mathbb{Q}_p(\zeta_p)$.

[39] The element $\pi := \zeta_p - 1$ is the root of the Eisenstein polynomial

$$\frac{(X+1)^p - 1}{X} \in \mathbb{Q}_p[X],$$

and hence it is a uniformizer of $\mathbb{Q}_p(\zeta_p)$ by Theorem 3.5.8.

Proof of Claim 1: The field $K/\mathbb{Q}_p$ is linearly disjoint from $\mathbb{Q}_p(\zeta_p)/\mathbb{Q}$, since the latter has degree $p-1$ which is prime to p. By Kummer theory, there is a subgroup $A \subseteq \mathbb{Q}_p(\zeta_p)^\times/\mathbb{Q}_p(\zeta_p)^{\times p}$ isomorphic to $(\mathbb{Z}/p\mathbb{Z})^3$, for which $K(\zeta_p) \cong \mathbb{Q}(\zeta_p, A^{1/p})$, where $A^{1/p} = \{\sqrt[p]{\alpha} \mid \alpha \in A\}$. Our aim is to prove that we may without loss of generality assume that $A \subseteq U_1/U_1^p$. The extension $\mathbb{Q}(\zeta_p, A^{1/p})$ is the compositum of two linearly disjoint abelian extensions, hence it is also abelian. Now, Lemma 3.8.8 implies for any $\alpha \in A$

$$\frac{\sigma(\alpha)}{\alpha^{\omega(\sigma)}} \in \mathbb{Q}_p(\zeta_p)^{\times p}, \tag{3.15}$$

where $\sigma \in G := \mathrm{Gal}(\mathbb{Q}_p(\zeta_p)/\mathbb{Q}_p)$ and $\omega\colon G \xrightarrow{\sim} (\mathbb{Z}/p\mathbb{Z})^\times$ is given by $\sigma(\zeta_p) = \zeta_p^{\omega(\sigma)}$. Let us write v_π for the normalized valuation on $\mathbb{Q}_p(\zeta_p)$ associated to π. We want to compute the valuation of an arbitrary element $\alpha \in A$. Note, that it is only well-defined $\mod p$ since α is only defined up to multiplication of elements from $\mathbb{Q}_p(\zeta_p)^{\times p}$. For each $\alpha \in A$, we have according to (3.15) the formula

$$v_\pi(a) = v_\pi(\sigma(\alpha)) \equiv \omega(\sigma) v_\pi(\alpha) \mod p.$$

Thus, we have $(1-\omega(\sigma))v_\pi(\alpha) \equiv 0 \mod p$ for all $\sigma \in G$. Since $\omega(\sigma)$ runs through all elements in $(\mathbb{Z}/p\mathbb{Z})^\times$, this implies $v_\pi(\alpha) \equiv 0 \mod p$. On the other hand, α is only defined up to multiplication with a pth power in $\mathbb{Q}_p(\zeta_p)^\times$. Thus, we may without loss of generality multiply α by $\pi^{-v_\pi(\alpha)}$ and obtain a representative $\alpha \in \mathbb{Q}_p(\zeta_p)^\times$ with $v_\pi(\alpha) = 0$. Furthermore, $\mu_{p-1}(\mathbb{Q}_p)^p = \mu_{p-1}(\mathbb{Q}_p)$ so every $(p-1)$-root of unity is a p-th power. It follows from Exercise 2 on Sheet 6, that the $(p-1)$-roots of unity form a system of representatives for the units in the residue field of a totally ramified extension. Thus, after multiplication with a suitable element from $\mu_{p-1}(\mathbb{Q}_p(\zeta_p))$, we may furthermore assume $\alpha \equiv 1 \mod \pi$. Summarizing the above discussion shows that we may assume $A \subseteq U_1/U_1^p$, where $U_1 := \{u \equiv 1 \mod \pi\}$ are the principal units in the field $\mathbb{Q}_p(\zeta_p)$. This shows the Claim 1.

Claim 2: Any $u \in U_1$ can be written as

$$u = \zeta_p^b u_2$$

for suitable $0 \le b \le p-1$ and $u_2 \in U_2 := \{u \equiv 1 \mod \pi^2\}$.
Proof of Claim 2: Let us write $u = 1 + b\pi + O(\pi^2)$. Since we have $\zeta_p = 1+\pi$, we get

$$\zeta_p^{-b} = 1 - b\pi + O(\pi^2).$$

This implies

$$\zeta_p^{-b} u = 1 + O(\pi^2).$$

Thus setting $u_2 := \zeta_p^{-b} u$ proves Claim 2.
Now, we will show that the elements of A are of a very particular form.

Claim 3: Every $\alpha \in A$ can be written in the form

$$\alpha = \zeta_p^b(1 + c\pi^p + O(\pi^{p+1}))$$

for suitable integers b and c.
Proof of Claim 3: We start with the following observation

$$\frac{\sigma(\pi)}{\pi} = \frac{\sigma(\zeta_p - 1)}{\zeta_p - 1} = \frac{\zeta_p^{\omega(\sigma)} - 1}{\zeta_p - 1} = \zeta_p^{\omega(\sigma)-1} + \cdots + \zeta_p + 1 \equiv \omega(\sigma) \mod \pi.$$

This shows

$$\sigma(\pi) \equiv \omega(\sigma)\pi \mod \pi^2. \tag{3.16}$$

Applying Claim 2 to $\alpha \in A$ gives

$$\alpha = \zeta_p^b(1 + c\pi^e + O(\pi^{e+1}))$$

for some integers c, b and $e \geq 2$. Together with (3.16), this allows us to $\sigma(\alpha)$ as follows

$$\sigma(\alpha) = \zeta_p^{\omega(\sigma)b}(1 + c\omega(\sigma)^e\pi^e + O(\pi^{e+1})).$$

On the other hand, we have

$$\alpha^{\omega(\sigma)} = \zeta_p^{\omega(\sigma)b}(1 + c\omega(\sigma)\pi^e + O(\pi^{e+1})).$$

A straightforward computation shows $U_1^p \subseteq U_{p+1}$ and since we have

$$\frac{\sigma(\alpha)}{a^{\omega(\sigma)}} \in U_1^p,$$

we deduce by comparing the coefficients in the π-adic expansion of $\sigma(\alpha)$ and $\alpha^{\omega(\sigma)}$ the congruence

$$\omega(\sigma) = \omega(\sigma)^e$$

for every $\sigma \in G$. But this implies $e \equiv 1 \mod p - 1$ and since $e \geq 2$, we get $e \geq p$. This proves the Claim 3.
Now, Claim 3 shows that A is contained in the subgroup of U_1/U_1^p generated by ζ_p and $(1 + \pi^p)$. This is an abelian group of exponent p generated by two elements, hence isomorphic to a subgroup of $(\mathbb{Z}/p\mathbb{Z})^2$. This contradicts $A \cong (\mathbb{Z}/p\mathbb{Z})^3$. □

Outlook

We have seen in this section that not every finite group can appear as a Galois group of $\mathbb{Q}_p$. More precisely, for an odd prime p, the groups $(\mathbb{Z}/p\mathbb{Z})^k$ for $k \geq 3$ can not be realized as a Galois group over $\mathbb{Q}_p$. The question whether every finite group appears as a Galois group of some Galois extension over the field of rational numbers $\mathbb{Q}$ is much more difficult and widely open. Indeed, it is expected that each finite group appears as such a Galois group:

Conjecture (Inverse Problem of Galois Theory)**.** *For every finite group G there exists a Galois extension $K/\mathbb{Q}$ with $G \cong Gal(K/\mathbb{Q})$.*

This again shows that extensions of global fields are much more complicated to understand than there local counterparts. For example, the Galois groups $\mathrm{Gal}(\overline{\mathbb{Q}}_p/\mathbb{Q}_p)$ are (topologically) finitely presented and it is possible to give explicit generators and relations. The absolute Galois group $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ is not (topologically) finitely generated and no 'explicit' description is known. Nevertheless, it contains many important arithmetic information and appears everywhere in (algebraic) number theory.

3.9 *Proof of the local Kronecker–Weber Theorem*

In this section, we will combine the results of the previous weeks to prove the local Kronecker-Weber Theorem, i.e., we want to prove the following Theorem.

Theorem 3.9.1. *For any finite abelian extension $K/\mathbb{Q}_p$ there exists a positive integer n such that $K \subseteq \mathbb{Q}_p(\zeta_n)$.*

We claim that it suffices to prove the Kronecker-Weber Theorem for all finite abelian extensions of prime power degree. Indeed, let us assume that $K/\mathbb{Q}_p$ is a finite abelian extension with Galois group $G = \mathrm{Gal}(L/\mathbb{Q}_p)$. By the structure theorem for finite abelian groups, G decomposes as

$$G = G_1 \times \cdots \times G_r,$$

where each G_i is a cyclic group of prime power degree. For $1 \leq j \leq r$, let us define

$$G^{(j)} := \prod_{\substack{i=1 \\ i \neq j}}^{r} G_i.$$

and the fixed field $K_j := K^{G^{(j)}}$. The field K_j is a cyclic extension of $\mathbb{Q}_p$ of prime power degree. We have

$$K = K_1 \ldots K_r.$$

Let us assume for a moment that we already know the local Kronecker-Weber Theorem for all cyclic extensions of prime power degree. In particular, we can find for any $1 \leq j \leq r$ a positive integer n_j such that

$$K_j \subseteq \mathbb{Q}_p(\zeta_{n_j}).$$

If we set $n := \mathrm{lcm}(n_1, \ldots, n_r)$ then we get

$$K = K_1 \ldots K_r \subseteq \mathbb{Q}_p(\zeta_{n_1}) \ldots \mathbb{Q}_p(\zeta_{n_r}) = \mathbb{Q}_p(\zeta_n)$$

and the local Kronecker–Weber Theorem would follow. Thus, it suffices to prove the local Kronecker–Weber Theorem for cyclic extensions $K/\mathbb{Q}_p$ of prime power degree l^r for all primes l and $r \geq 1$. We distinguish between $l = p$ and $l \neq p$. Let us start with the case $l \neq p$.

Proposition 3.9.2. *Let $K/\mathbb{Q}_p$ be a cyclic extension of degree l^r for some prime $l \neq p$. Then there exists a positive integer n such that $K \subseteq \mathbb{Q}_p(\zeta_n)$.*

Proof. Let $I \subseteq \mathrm{Gal}(K/\mathbb{Q}_p)$ be the inertia group of $K/\mathbb{Q}_p$ and set $F := K^I$. Then $F/\mathbb{Q}_p$ is unramified while K/F is totally and tamely ramified. By the structure theorem of totally ramified extensions (Theorem 3.5.8), we get

$$K = F(\sqrt[e]{\pi}) \tag{3.17}$$

for some uniformizer π of F. Since $F/\mathbb{Q}_p$ is unramified, we have $(p) = (\pi)$ in the valuation ring $\mathcal{O}_F$ of F and deduce that

$$\pi = -up$$

for some unit $u \in A^\times$. From (3.17), we deduce

$$K \subseteq \mathbb{Q}_p(\sqrt[e]{-p})F(\sqrt[e]{u})$$

The element $\sqrt[e]{u}$ is a root of the normalized polynomial $X^e - u$ whose reduction is separable, hence by Lemma 3.5.3 the extension $F(\sqrt[e]{u})/F$ is unramified. Since $F/\mathbb{Q}_p$ is also unramified, we deduce that $F(\sqrt[e]{u})/\mathbb{Q}_p$ is unramified. By the structure theorem for unramified extensions, we deduce $F(\sqrt[e]{u}) = \mathbb{Q}_p(\zeta_k)$ for $k = p^{[F(\sqrt[e]{u}):\mathbb{Q}_p]} - 1$. It remains to show that $\mathbb{Q}_p(\sqrt[e]{-p})$ is also contained in a cyclotomic extension. But note that $\mathbb{Q}_p(\sqrt[e]{-p})/\mathbb{Q}_p$ is a sub-extension of an abelian extension, hence it is also an abelian Galois extension. In particular, it contains all eth roots of $-p$, i.e.,

$$\sqrt[e]{-p}, \zeta_e \cdot \sqrt[e]{-p}, \ldots, \zeta_e^{e-1} \cdot \sqrt[e]{-p} \in \mathbb{Q}_p(\sqrt[e]{-p}).$$

So, it does also contain all e-th roots of unity and we obtain the tower $\mathbb{Q}_p \subseteq \mathbb{Q}_p(\zeta_e) \subseteq \mathbb{Q}_p(\sqrt[e]{-p})$. Since $p \nmid e$, we deduce from Lemma 3.5.3 that $\mathbb{Q}_p \subseteq \mathbb{Q}_p(\zeta_e)$ is an unramified extension. On the other hand, we already know that $\mathbb{Q}_p(\sqrt[e]{-p})/\mathbb{Q}_p$ is totally ramified. Thus, we must have $\mathbb{Q}_p(\zeta_e) = \mathbb{Q}_p$ and deduce[40] that $e \mid (p-1)$ for . This gives $\mathbb{Q}_p(\sqrt[e]{-p}) \subseteq \mathbb{Q}_p((-p)^{1/(p-1)})$. But we will prove in the Exercises the equality

$$\mathbb{Q}_p((-p)^{1/(p-1)}) = \mathbb{Q}_p(\zeta_p).$$

Combining everything gives

$$K \subseteq \mathbb{Q}_p(\sqrt[e]{-p})F(\sqrt[e]{u}) \subseteq \mathbb{Q}_p(\zeta_p)\mathbb{Q}_p(\zeta_k) \subseteq \mathbb{Q}_p(\zeta_{pk}),$$

which proves the statement of the theorem. □

[40] We have seen in Exercise 2 and 4 of Sheet 6 that the roots of unity $\mu_e(K)$ for $p \nmid e$ map isomorphically to the roots of unity of κ_K. For $K = \mathbb{Q}_p$, we deduce that $e \mid (p-1)$ since $p-1$ is the order of the groups of unity in $\mathbb{F}_p$.

In order to finish the proof of the local Kronecker-Weber Theorem, we have to deal with the case $l = p$.

Proposition 3.9.3. *Let $K/\mathbb{Q}_p$ be a cyclic extension of degree p^r. Then there exists a positive integer n such that $K \subseteq \mathbb{Q}_p(\zeta_n)$.*

Proof. Let us first assume $p \neq 2$. Before we start the proof, let us think about obvious candidates of cyclotomic fields which might contain such a degree p^r extension of $\mathbb{Q}_p$. The first example which comes to mind is the unique unramified extension of $\mathbb{Q}_p$ of degree p^r, namely $\mathbb{Q}_p(\zeta_{p^{p^r}-1})$. On the other hand, the abelian extension $\mathbb{Q}_p(\zeta_{p^{r+1}})$ of $\mathbb{Q}_p$ is a totally ramified Galois extension of degree $p^r(p-1)$. This can be seen easily by observing that the minimal polynomial of $\zeta_{p^{r+1}} - 1$ is the polynomial

$$\frac{(X+1)^{p^{r+1}}-1}{(X+1)^{p^r}-1} = (X+1)^{p^r(p-1)} + (X+1)^{p^r(p-2)} + \cdots + (X+1)^{p^r} + 1$$

which is easily seen[41] to be an Eisenstein polynomial. In particular, there is a unique totally ramified abelian sub-extension of $\mathbb{Q}_p(\zeta_{p^{r+1}})$ of degree p^r. Thus, we have two obvious candidates and our goal is to prove that our given field K is contained in their composite, i.e., in the field $\mathbb{Q}_p(\zeta_m)$ for $m = p^{r+1}(p^{p^r}-1)$. The field $\mathbb{Q}_p(\zeta_m)$ is the compositum of the linearly disjoint fields $\mathbb{Q}_p(\zeta_{p^{r+1}})$ and $\mathbb{Q}_p(\zeta_{p^{p^r}-1})$, hence its Galois group is isomorphic to

$$\mathrm{Gal}(\mathbb{Q}_p(\zeta_m)/\mathbb{Q}_p) \cong (\mathbb{Z}/(p-1)\mathbb{Z}) \times (\mathbb{Z}/p^r\mathbb{Z}) \times (\mathbb{Z}/p^r\mathbb{Z}).$$

Let us assume that K is not contained in this field. Then, we would have

$$\mathrm{Gal}(K(\zeta_m)/\mathbb{Q}_p) \cong (\mathbb{Z}/(p-1)\mathbb{Z}) \times (\mathbb{Z}/p^r\mathbb{Z}) \times (\mathbb{Z}/p^r\mathbb{Z}) \times (\mathbb{Z}/p^s\mathbb{Z}),$$

for some $1 < s \leq r$. It follows that the group $\mathrm{Gal}(K(\zeta_m)/\mathbb{Q}_p)$ has a quotient which is isomorphic to $(\mathbb{Z}/p\mathbb{Z})^3$. Thus, there would be a finite Galois extension of $\mathbb{Q}_p$ with Galois group $(\mathbb{Z}/p\mathbb{Z})^3$. By Kummer theory, such an extension can not exist, see Proposition 3.8.9.

Let us now indicate the proof for the prime $p = 2$. The strategy is similar, but for $p = 2$ we have to adapt the argument since there are abelian extensions of $\mathbb{Q}_2$ with Galois group $(\mathbb{Z}/2\mathbb{Z})^3$, namely the extension $\mathbb{Q}_2(\zeta_{24})$. We want to prove that K is contained in $\mathbb{Q}_2(\zeta_m)$ with $m = (2^{2^r}-1)2^{r+2}$ which has Galois group

$$\mathrm{Gal}(\mathbb{Q}_2(\zeta_m)/\mathbb{Q}_2) \cong (\mathbb{Z}/2\mathbb{Z}) \times (\mathbb{Z}/2^r\mathbb{Z})^2.$$

If K is not contained in $\mathbb{Q}_2(\zeta_m)$, then the Galois group $K(\zeta_m)$ has to be of the following form:

$$\mathrm{Gal}(K(\zeta_m)/\mathbb{Q}_2) \cong \begin{cases} \mathbb{Z}/2\mathbb{Z} \times (\mathbb{Z}/2^r\mathbb{Z})^2 \times \mathbb{Z}/2^s\mathbb{Z} & \text{with } 1 \leq s \leq r, \text{ or} \\ (\mathbb{Z}/2^r\mathbb{Z})^2 \times \mathbb{Z}/2^s\mathbb{Z} & \text{with } 2 \leq s \leq r. \end{cases}$$

[41] The constant term is p and the higher coefficients are all divisible by p since p divides $\binom{p^i}{k}$ for $i \geq 1$ and $1 \leq k < p^i$.

Thus, it admits either a quotient isomorphic to $(\mathbb{Z}/2\mathbb{Z})^4$ or to $(\mathbb{Z}/4\mathbb{Z})^3$. By the next Lemma, both groups can not be realised as a Galois group over $\mathbb{Q}_2$. □

Lemma 3.9.4. *No extension of $\mathbb{Q}_2$ has Galois group isomorphic to $(\mathbb{Z}/2\mathbb{Z})^4$ or to $(\mathbb{Z}/4\mathbb{Z})^3$.*

Proof. This will be proven in the exercises. □

Let us define the maximal abelian extension K^{ab} on a field K as the compositum of all finite abelian extensions in a fixed algebraic closure of K. The local Kronecker-Weber Theorem allows us to give an explicit description of the Galois group of the maximal abelian extension of $\mathbb{Q}_p$:

Corollary 3.9.5. *For a prime p, we have $\mathbb{Q}_p^{ab} = \mathbb{Q}_p(\zeta_n \mid n \in \mathbb{N})$ and we get an (explicit) isomorphism*

$$\widehat{\mathbb{Z}} \times \mathbb{Z}_p^\times \xrightarrow{\sim} Gal(\mathbb{Q}_p^{ab}/\mathbb{Q}_p).$$

Proof. The local Kronecker-Weber Theorem says that any abelian extension of $\mathbb{Q}_p$ is contained in a cyclotomic extension $\mathbb{Q}_p(\zeta_n)$. This proves $\mathbb{Q}_p^{ab} = \mathbb{Q}_p(\zeta_n \mid n \in \mathbb{N})$. We can write this in a more convenient way as follows

$$\mathbb{Q}_p(\zeta_n \mid n \in \mathbb{N}) = \mathbb{Q}_p(\zeta_n \mid n \text{ prime to } p)\mathbb{Q}_p(\zeta_{p^j} \mid j \in \mathbb{N}).$$

Since both fields on the right hand side are linearly disjoint, we get

$$\mathrm{Gal}(\mathbb{Q}_p^{ab}/\mathbb{Q}_p) = \mathrm{Gal}(\mathbb{Q}_p(\zeta_n \mid n \text{ prime to } p)/\mathbb{Q}_p) \times \mathrm{Gal}(\mathbb{Q}_p(\zeta_{p^j} \mid j \in \mathbb{N})/\mathbb{Q}_p).$$

Now, the isomorphism follows from the explicit isomorphisms of Example 3.7.10 and Example 3.7.11. □

Outlook

Using the local Kronecker-Weber Theorem, we were able to prove an isomorphism

$$\widehat{\mathbb{Z}} \times \mathbb{Z}_p^\times \xrightarrow{\sim} \mathrm{Gal}(\mathbb{Q}_p^{ab}/\mathbb{Q}_p). \tag{3.18}$$

On the other hand, we have $\mathbb{Q}_p^\times = p^{\mathbb{Z}} \times \mathbb{Z}_p^\times \cong \mathbb{Z} \times \mathbb{Z}_p^\times$. The group $\widehat{\mathbb{Z}} \times \mathbb{Z}_p^\times$ is the pro-finite completion of the group $\mathbb{Z} \times \mathbb{Z}_p^\times$. Here, the pro-finite completion of an abstract group G is the group

$$\widehat{G} := \varprojlim_U G/U$$

where U runs through all normal subgroups of finite index in G. Thus, combining the isomorphism $\mathbb{Q}_p^\times \cong \mathbb{Z} \times \mathbb{Z}_p^\times$ with, we obtain an explicit isomorphism

$$\widehat{\mathbb{Q}_p^\times} \xrightarrow{\sim} \mathrm{Gal}(\mathbb{Q}_p^{ab}/\mathbb{Q}_p).$$

So, it turns out that the Galois group of the maximal abelian extension of $\mathbb{Q}_p$ can be described in an explicit way by the pro-finite completion of the units of $\mathbb{Q}_p$. Surprisingly, this holds for any finite extension of $\mathbb{Q}_p$ (more generally for local fields):

Theorem 3.9.6 (Local class field Theory)**.** *Let K be a finite extension of $\mathbb{Q}_p$, then there exists an explicit isomorphism*

$$\widehat{K^\times} \xrightarrow{\sim} Gal(K^{ab}/K).$$

So local class field theory allows us to study the Galois group of the maximal abelian extension of K purely in terms of the units in K. The units in K are easily studied, on sheet 6, exercise 4, we have given an explicit description of the group $K^\times$.

3.10 The global Kronecker–Weber Theorem

In this section, we want to prove the Global Kronecker–Weber Theorem:

Theorem 3.10.1 (Global Kronecker–Weber Theorem)**.** *Let $K/\mathbb{Q}$ be a finite abelian field extension. There exists a positive integer n such that $K \subseteq \mathbb{Q}(\zeta_n)$, where ζ_n is a primitive n-th root of unity.*

Let us recall the following result from Algebraic Number Theory I:

Theorem 3.10.2 (Hermite–Minkowski)**.** *Let K be a number field of degree n and with discriminant d_K. Then*

$$\sqrt{|d_K|} \geq \frac{n^n}{n!}\left(\frac{\pi}{4}\right)^{n/2}.$$

In particular, there are no non-trivial everywhere unramified extensions of $\mathbb{Q}$.

Proof. Algebraic Number Theory I. □

Corollary 3.10.3. *Let $K/\mathbb{Q}$ be a finite abelian Galois extension. Then $Gal(K/\mathbb{Q})$ is generated by the inertia groups*[42] *$I_p(K/\mathbb{Q})$ where p runs through all the primes.*

Proof. Let $H \subseteq \mathrm{Gal}(K/\mathbb{Q})$ be the subgroup generated by all the inertia groups $I_p(K/\mathbb{Q})$. The fixed field $L := K^H$ is fixed by each $I_p(K/\mathbb{Q})$ hence it is everywhere unramified. The Hermite–Minkowski Theorem implies $L = \mathbb{Q}$ and hence $H = \mathrm{Gal}(K/\mathbb{Q})$. □

Furthermore, we will need some facts about the prime decomposition in cyclotomic fields. First, let us recall that a Frobenius element of a prime $\mathfrak{P} \subseteq \mathcal{O}_L$ over $\mathfrak{p} \subseteq \mathcal{O}_K$ in a Galois extension L/K of number fields is an element $\mathrm{Frob}_{\mathfrak{P}}$ of the decomposition group

[42] Since $K/\mathbb{Q}$ is abelian, the inertia groups $I_{\mathfrak{p}}$ of a prime ideal $\mathfrak{p} \subseteq \mathcal{O}_K$ over p does only depend on p.

$G_{\mathfrak{P}} \subseteq G = \mathrm{Gal}(L/K)$ which maps to the Frobenius morphism $(x \mapsto x^{\#\kappa(\mathfrak{p})}) \in \mathrm{Gal}(\kappa(\mathfrak{P})/\kappa(\mathfrak{P}))$ of the residue field extension under

$$G_{\mathfrak{P}} \twoheadrightarrow \mathrm{Gal}(\kappa(\mathfrak{P})/\kappa(\mathfrak{P})).$$

The kernel of the latter map is the inertia group $I_{\mathfrak{P}}$. So, it is only well-defined up to multiplication by $I_{\mathfrak{P}}$. In particular, $\mathrm{Frob}_{\mathfrak{P}}$ is uniquely determined if $\mathfrak{P}$ is unramified over $\mathfrak{p}$. Furthermore, recall from ANT 1 that the sets of Frobenius elements as well as the decomposition and inertia groups of two primes $\mathfrak{P}$ and $\mathfrak{P}'$ over $\mathfrak{p}$ are conjugate in $\mathrm{Gal}(L/K)$. In particular, if L/K is abelian then the set of Frobenius elements, the decomposition group and the inertia group do only depend on $\mathfrak{p}$ and not on the chosen prime above $\mathfrak{p}$.

Theorem 3.10.4 (Cyclotomic Extensions)**.** *Let n be a positive integer and ζ_n a primitive n-th root of unity in an algebraic closure $\overline{\mathbb{Q}}$ of $\mathbb{Q}$. We have*

(a) *$\mathbb{Q}(\zeta_n)$ is a Galois extension of $\mathbb{Q}$ with Galois group*

$$\omega_n\colon \mathit{Gal}(\mathbb{Q}(\zeta_n)/\mathbb{Q}) \xrightarrow{\sim} (\mathbb{Z}/n\mathbb{Z})^\times, \quad \sigma(\zeta_n) = \zeta_n^{\omega_n(\sigma)}.$$

(b) *The ring of integers of $\mathbb{Q}(\zeta_n)$ is $\mathbb{Z}[\zeta_n]$.*

(c) *The extension $\mathbb{Q}(\zeta_n)/\mathbb{Q}$ is ramified at p if and only if*

$$\begin{cases} p \mid n & \text{if } p \neq 2 \\ 4 \mid n & \text{if } p = 2. \end{cases}$$

If $n = p^k$ then $p = (\zeta_{p^k} - 1)^{[\mathbb{Q}(\zeta_{p^k}):\mathbb{Q}]}$, in particular, p is totally ramified in $\mathbb{Q}(\zeta_{p^k})/\mathbb{Q}$.

(d) *For a prime p with $p \nmid n$, let us write $\mathrm{Frob}_p \in \mathit{Gal}(\mathbb{Q}(\zeta_n)/\mathbb{Q})$ for the Frobenius element*[43] *at p. We have $\omega_n(\mathrm{Frob}_p) = p + n\mathbb{Z}$*

(e) *For a prime p, let us write $n = p^{v_p(n)}m$ with $m := n/p^{v_p(n)}$. Then the inertia group $I_p(\mathbb{Q}(\zeta_n)/\mathbb{Q})$ of $\mathit{Gal}(\mathbb{Q}(\zeta_n)/\mathbb{Q})$ corresponds to the group $(\mathbb{Z}/p^{v_p(n)}\mathbb{Z})^\times$ under the isomorphism*

$$\begin{aligned} \mathit{Gal}(\mathbb{Q}(\zeta_n)/\mathbb{Q}) &\cong \mathit{Gal}(\mathbb{Q}(\zeta_m)/\mathbb{Q}) \times \mathit{Gal}(\mathbb{Q}(\zeta_{p^{v_p(n)}})/\mathbb{Q}) \\ &\cong (\mathbb{Z}/m\mathbb{Z})^\times \times (\mathbb{Z}/p^{v_p(n)}\mathbb{Z})^\times. \end{aligned}$$

(f) *For each prime p, let us write $n = p^{v_p(n)}m$ with $m := n/p^{v_p(n)}$ and define*

$$\overline{p} := p + m\mathbb{Z} \in (\mathbb{Z}/m\mathbb{Z})^\times$$

and

$$f_p := \mathrm{ord}(\overline{p}).$$

[43] By the discussion preceding this Theorem, Frob_p is well-defined since p is unramified and the extension is abelian.

Then, the prime p decomposes in $\mathbb{Z}[\zeta_n]$ as follows:

$$p\mathbb{Z}[\zeta_n] = (\mathfrak{p}_1 \dots \mathfrak{p}_r)^{\varphi(p^{v_p(n)})}$$

with $f(\mathfrak{p}_i/p) = f_p$ and $e(\mathfrak{p}/p) = \varphi(p^{v_p(n)})$, where $\varphi(p^{v_p(n)})$ denotes the Euler totient function of $p^{v_p(n)}$.

Proof. (a), (b), (c) and (d) have been shown in the lecture ANT 1 in the winter term. (e) follows from the fact that p is unramified in $\mathbb{Q}(\zeta_m)/\mathbb{Q}$, while it is totally ramified in $\mathbb{Q}(\zeta_{p^{v_p(n)}})/\mathbb{Q}$. So, the inertia group of $\mathrm{Gal}(\mathbb{Q}(\zeta_{p^{v_p(n)}})/\mathbb{Q})$ is the whole Galois group.
(f) is a consequence of the previous claims. Indeed, by (e) the ramification index at p is given by

$$e_p = |I_p| = \#(\mathbb{Z}/p^{v_p(n)}\mathbb{Z})^\times = \varphi(p^{v_p(n)}).$$

Again, since $\mathbb{Q}(\zeta_m)/\mathbb{Q}$ is unramified while $\mathbb{Q}(\zeta_{v_p(n)})/\mathbb{Q}$ is totally ramified, the inertia degree f_p coincides with the inertia degree of p in $\mathbb{Q}(\zeta_m)/\mathbb{Q}$. This is given by the order of the decomposition group, i.e.,

$$f_p = \mathrm{ord}(\mathrm{Frob}_p).$$

By (d), we have $\mathrm{ord}(\mathrm{Frob}_p) = \mathrm{ord}(\overline{p})$ for $\overline{p} = p + m\mathbb{Z}$. □

We can now prove the Global Kronecker-Weber Theorem:

Proof. Let $K/\mathbb{Q}$ be a finite abelian extension with Galois group $G = \mathrm{Gal}(K/\mathbb{Q})$. For any ramified rational prime p we pick[44] a prime $\mathfrak{p} \subseteq \mathcal{O}_K$ over p. By Theorem 3.4.4 (e), we have $\mathrm{Gal}(K_\mathfrak{p}/\mathbb{Q}_p) \cong G_\mathfrak{p} \subseteq G$. Hence, the local extension $K_\mathfrak{p}/\mathbb{Q}_p$ is abelian and the local Kronecker–Weber Theorem implies that $K_\mathfrak{p} \subseteq \mathbb{Q}_p(\zeta_{m_p})$ for some positive integer m_p. If p is unramified in K then $K_\mathfrak{p}/\mathbb{Q}_p$ is unramified and we may in this case assume that m_p is co-prime to p. Let $e_p := \nu_p(m_p)$ and put $m = \prod_p p^{e_p}$, which is a finite product since only finitely primes ramify in $K/\mathbb{Q}$. We set $L := K(\zeta_m)$. We want to show $L = \mathbb{Q}(\zeta_m)$, which then implies $K \subseteq \mathbb{Q}(\zeta_m)$.

The field $L = K \cdot \mathbb{Q}(\zeta_m)$ is a compositum of Galois extensions of $\mathbb{Q}$, and therefore Galois over $\mathbb{Q}$. Its Galois group is isomorphic to a subgroup[45] of $\mathrm{Gal}(\mathbb{Q}(\zeta_m)/\mathbb{Q}) \times \mathrm{Gal}(K/\mathbb{Q})$ and hence is an abelian group. Now, let $\mathfrak{P} \subseteq \mathcal{O}_L$ and $\mathfrak{p} \subseteq \mathcal{O}_K$ be a primes such that $\mathfrak{P}|\mathfrak{p}|p$. The completion of L at $\mathfrak{P}$ is given by

$$L_\mathfrak{P} = K_\mathfrak{p}(\zeta_m) = \mathbb{Q}_p(\zeta_m, \zeta_{m_p}) = \mathbb{Q}_p(\zeta_{\mathrm{lcm}(m,m_p)}).$$

By Theorem 3.10.4 and Theorem 3.4.4 (e), the inertia groups of $\mathbb{Q}_p(\zeta_m)$ and $\mathbb{Q}_p(\zeta_{\mathrm{lcm}(m,m_p)})$ are both isomorphic to $(\mathbb{Z}/p^{e_p}\mathbb{Z})^\times$, i.e.,

$$I_p(\mathbb{Q}_p(\zeta_m)/\mathbb{Q}_p) \cong (\mathbb{Z}/p^{e_p}\mathbb{Z})^\times \cong I_p(\mathbb{Q}_p(\zeta_{\mathrm{lcm}(m,m_p)})/\mathbb{Q}_p).$$

[44] Since $K/\mathbb{Q}$ is Galois, all primes above p are ramified with the same ramification index. In the following, it does not matter which prime we pick.

[45] More precisely, we recall from Algebra: For two finite Galois extensions L_1, L_2 of K with Galois groups $G_1 = \mathrm{Gal}(L_1/K)$ and $G_2 = \mathrm{Gal}(L_2/K)$, the Galois group $\mathrm{Gal}(L_1L_2/K)$ is isomorphic to

$$\{(\sigma, \tau) \in G_1 \times G_2 : \sigma|_{K_1 \cap K_2} = \tau|_{K_1 \cap K_2}\}.$$

On the other hand, the tower $\mathbb{Q}_p(\zeta_m) \subseteq L \subseteq \mathbb{Q}_p(\zeta_{\mathrm{lcm}(m,m_p)})$ gives rise to a surjection of inertia groups

$$I_p(\mathbb{Q}_p(\zeta_{\mathrm{lcm}(m,m_p)})/\mathbb{Q}_p) \twoheadrightarrow I_p(L_{\mathfrak{P}}/\mathbb{Q}_p) \twoheadrightarrow I_p(\mathbb{Q}_p(\zeta_m)/\mathbb{Q}_p).$$

Thus, we deduce $|I_p(L_{\mathfrak{P}}/\mathbb{Q}_p)| = \varphi(p^{e_p})$ where φ is Euler's totient function. By Theorem 3.10.3 and since $L/\mathbb{Q}$ is abelian, the map

$$\bigoplus_p I_p(L_{\mathfrak{P}}/\mathbb{Q}_p) = \bigoplus_p I_p(L/\mathbb{Q}) \twoheadrightarrow \mathrm{Gal}(L/\mathbb{Q})$$

induced by the inclusions $I_p(L/\mathbb{Q}) \subseteq \mathrm{Gal}(L/\mathbb{Q})$ is surjective. We deduce

$$|\,\mathrm{Gal}(L/\mathbb{Q})| \leq \prod_p |I_p(L_{\mathfrak{P}}/\mathbb{Q}_p)| = \prod_p \varphi(p^{e_p}) = \varphi(m).$$

But the subextension $\mathbb{Q}(\zeta_m)/\mathbb{Q}$ of $L/\mathbb{Q}$ has already degree $\varphi(m)$ and we conclude $L = \mathbb{Q}(\zeta_m)$. This implies $K \subseteq \mathbb{Q}(\zeta_m)$ and the global Kronecker–Weber Theorem holds. □

Previously, we have already introduced the profinite group

$$\widehat{\mathbb{Z}} := \varprojlim_{n \in (\mathbb{N},|)} \mathbb{Z}/n\mathbb{Z}.$$

Let us observe that the finite groups $(\mathbb{Z}/n\mathbb{Z}, +)$ have an additional multiplicative structure and form a ring. The transition maps $\mathbb{Z}/m\mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$ for $n|m$ are all ring homomorphisms and thus we obtain a natural ring structure on

$$\widehat{\mathbb{Z}} := \varprojlim_{n \in (\mathbb{N},|)} \mathbb{Z}/n\mathbb{Z}.$$

In particular, it makes sense to talk about the units $\widehat{\mathbb{Z}}^\times$ of $\widehat{\mathbb{Z}}$. Note, that we can also describe the units in this ring as follows:

$$\widehat{\mathbb{Z}}^\times = \varprojlim_{n \in (\mathbb{N},|)} (\mathbb{Z}/n\mathbb{Z})^\times.$$

We equip them with the topology induced by the inverse limit. The following Corollary gives an explicit description of $\mathrm{Gal}(\mathbb{Q}^{ab}/\mathbb{Q})$:

Corollary 3.10.5. *There is a canonical isomorphism of pro-finite groups*

$$Gal(\mathbb{Q}^{ab}/\mathbb{Q}) \xrightarrow{\sim} \widehat{\mathbb{Z}}^\times.$$

Proof. The global Kronecker-Weber Theorem implies that the cyclotomic extensions $\mathbb{Q}(\zeta_n)$ form a cofinal set in the directed set of all finite abelian Galois extension with respect to inclusion. By Theorem 3.7.9, we have

$$\mathrm{Gal}(\mathbb{Q}^{ab}/\mathbb{Q}) \xrightarrow{\sim} \varprojlim_n \mathrm{Gal}(\mathbb{Q}(\zeta_n)/\mathbb{Q}).$$

By Theorem 3.10.4, we have canonical isomorphism $\omega_n \colon \operatorname{Gal}(\mathbb{Q}(\zeta_n)/\mathbb{Q}) \to (\mathbb{Z}/n\mathbb{Z})^\times$. They are compatible with the transition maps and we get

$$\operatorname{Gal}(\mathbb{Q}^{ab}/\mathbb{Q}) \xrightarrow{\sim} \varprojlim_n \operatorname{Gal}(\mathbb{Q}(\zeta_n)/\mathbb{Q}) \cong \varprojlim_n (\mathbb{Z}/n\mathbb{Z})^\times = \widehat{\mathbb{Z}}^\times.$$

□

Outlook

For a general number field K, we will see later the definition of the ring of adeles $\mathbb{A}_K$. Its group of units $\mathbb{I}_K := \mathrm{GL}_1(\mathbb{A}_K)$ is called the group of ideles. In the case $K = \mathbb{Q}$, one can show that this group is isomorphic to $\mathbb{I}_\mathbb{Q} \cong \mathbb{R}_{>0} \times \mathbb{Q}^\times \times \widehat{\mathbb{Z}}^\times$. The Kronecker-Weber Theorem can be restated as follows. There is a canonical surjection

$$\mathbb{I}_\mathbb{Q}/\mathbb{Q}^\times \twoheadrightarrow \operatorname{Gal}(\mathbb{Q}^{ab}/\mathbb{Q})$$

with kernel $\mathbb{R}_{>0}$. It is this formulation which can be generalized to arbitrary number fields and leads to the general statements of global class field theory.

3.11 Dirichlet L-functions

In the next lectures, we want to explain the importance of the Kronecker–Weber Theorem (and more generally class field theory) to the theory of L-functions. In this section, we will define Dirichlet characters and Dirichlet L-functions. As an application of these L-functions, we sketch Dirichlet's proof that there are infinitely many primes in each arithmetic progression and deduce the Chebotarev density theorem for abelian extensions of $\mathbb{Q}$.

Definition 3.11.1. Let d be a positive integer. A *Dirichlet character modulo d* is a group homomorphism

$$\chi \colon (\mathbb{Z}/d\mathbb{Z})^\times \to \mathbb{C}^\times.$$

For $d \mid D$ and a Dirichlet character χ modulo d, we obtain a Dirichlet character modulo D by pre-composition with the canonical projection $(\mathbb{Z}/D\mathbb{Z})^\times \to (\mathbb{Z}/d\mathbb{Z})^\times$, i.e.,

$$(\mathbb{Z}/D\mathbb{Z})^\times \to (\mathbb{Z}/d\mathbb{Z})^\times \to \mathbb{C}^\times. \tag{3.19}$$

Given a Dirichlet character χ modulo D, the *conductor d* of χ is the smallest divisor of D such that χ factors as in (3.19). We call a Dirichlet character χ modulo D *primitive* if D is its conductor. The Dirichlet character which is constant 1 on $(\mathbb{Z}/D\mathbb{Z})^\times$ will be called the trivial Dirichlet character modulo D and will be denoted by $\mathbb{1}$.

We can associate an L-function to any such Dirichlet character as follows. Given a Dirichlet character χ modulo D, we may extend it to a function on the integers as follows:

$$\chi\colon \mathbb{Z} \to \mathbb{C}, \quad n \mapsto \chi(n) := \begin{cases} \chi(n \mod D) & \text{if } (n, D) = 1 \\ 0 & \text{if } (n, D) \neq 1. \end{cases}$$

Now, we define

$$L(\chi, s) := \sum_{n \geq 1} \frac{\chi(n)}{n^s}.$$

It is immediately checked that $L(\chi, s)$ converges absolutely and locally uniformly on $\mathrm{Re}(s) > 1$ to a holomorphic function. Furthermore, it admits an Euler product:

Lemma 3.11.2. *For* $\mathrm{Re}(s) > 1$, *we have the formula*

$$L(\chi, s) = \prod_p (1 - \chi(p) p^{-s})^{-1}.$$

Proof. Since $n \mapsto \chi(n) n^{-s}$ is a multiplicative function, we may apply Lemma 2.1.2. □

Note, that the Euler factors of the Dirichlet L-function associated to the trivial character $\mathbb{1}\colon (\mathbb{Z}/D\mathbb{Z})^\times \to \mathbb{C}^\times$ differs only at finitely many primes from the Euler factors of the Riemann zeta function. More precisely,

$$L(\mathbb{1}, s) = \zeta(s) \prod_{p \mid D} (1 - p^{-s}). \tag{3.20}$$

For a given Dirichlet character χ modulo D, let us observe that $\chi\colon \mathbb{Z} \to \mathbb{C}$ is periodic with period length D. This is the crucial point in relating Dirichlet L-functions to Hurwitz zeta functions:

Lemma 3.11.3. *Let χ be a Dirichlet character modulo D. For* $\mathrm{Re}(s) > 1$, *we have*

$$L(\chi, s) = D^{-s} \sum_{d=1}^{D} \chi(d) \zeta(s, d/D).$$

Proof.

$$\begin{aligned} L(\chi, s) &= \sum_{d=1}^{D} \sum_{n \geq 0} \frac{\chi(d)}{(d + n \cdot D)^s} \\ &= D^{-s} \sum_{d=1}^{D} \chi(d) \sum_{n \geq 0} \frac{1}{(n + d/D)^s} = D^{-s} \sum_{d=1}^{D} \chi(d) \zeta(s, d/D). \end{aligned}$$

□

This allows us to deduce many interesting properties of Dirichlet L-functions from the corresponding properties of Hurwitz zeta functions:

Theorem 3.11.4. *Let χ be a Dirichlet character modulo D. The Dirichlet L-function admits a meromorphic continuation to $\mathbb{C}$. If $\chi = \mathbb{1}$ is the trivial character modulo D, then $L(\mathbb{1}, s)$ is meromorphic with a simple pole at $s = 1$. If χ is a non-trivial character, the Dirichlet L-function is an entire function. The values at the negative integers are given by the formula*

$$L(\chi, 1-n) = -\frac{B_{n,\chi}}{n}, \quad \text{for } n \in \mathbb{N},$$

where

$$B_{n,\chi} := D^{n-1} \sum_{d=1}^{D} \chi(d) B_n(d/D).$$

In particular, the values of $L(\chi, s)$ at $s = 1-n$ are all algebraic.

Proof. This follows immediately from Theorem 2.6.5, where we have shown that $\zeta(s, x)$ for $0 < x \leq 1$ admits a meromorphic continuation to $\mathbb{C}$ with a simple pole at $s = 1$ of residue 1 and satisfies

$$\zeta(1-n, x) = -\frac{B_n(x)}{n} \quad \text{for } n \in \mathbb{N}.$$

The claim about $L(\chi, s)$ being entire for χ non-trivial, follows from

$$\mathrm{Res}_{s=1} L(\chi, s) = D^{-1} \sum_{d=1}^{D} \chi(d) \,\mathrm{Res}_{s=1} \zeta(s, d/D) = D^{-1} \underbrace{\sum_{d=1}^{D} \chi(d)}_{=0}.$$

Here, we use the fact that

$$\sum_{d=1}^{D} \chi(d) = 0,$$

which is left as an exercise. □

Using Dirichlet L-functions, one can prove Dirichlet's Theorem about primes in arithmetic progressions. Let us first introduce the notion of a Dirichlet density:

Definition 3.11.5. Let $S \subseteq T \subseteq \mathbb{N}$ be sets of positive integers. We define the Dirichlet density of S in T as

$$\lim_{s \to 1^+} \frac{\sum_{n \in S} n^{-s}}{\sum_{n \in T} n^{-s}}$$

if the limit exists.

Lemma 3.11.6. *We have*

$$\sum_{p \in \mathbb{P}} \frac{1}{p^s} = \log \zeta(s) + O(1) = -\log(s-1) + O(1), \quad \textit{as } s \to 1.$$

In particular, the Dirichlet density of a subset $S \subseteq \mathbb{P}$ is given by

$$\lim_{s\to 1^+} \frac{\sum_{p\in S} \frac{1}{p^s}}{-\log(s-1)},$$

if the limit exists. Furthermore, the Dirichlet density of finite subsets $S \subseteq \mathbb{P}$ is zero.

Proof. Since the Riemann zeta function has a simple pole at $s = 1$ of residue 1, we deduce

$$\log \zeta(s) = -\log(s-1) + O(1), \quad \text{as } s \to 1^+.$$

On the other hand, we have

$$\log \zeta(s) = \sum_{p\in\mathbb{P}} -\log(1-p^{-s}) = \sum_{p\in\mathbb{P}} \frac{1}{p^s} + \underbrace{\sum_{p\in}\sum_{n\geq 2} \frac{1}{np^{ns}}}_{=O(1) \text{ as } s\to 1^+}.$$

The claims about Dirichlet densities are immediate consequences of this formula. □

The following Lemma plays a key role in the proof of Dirichlet's theorem:

Lemma 3.11.7 ((Orthogonality relation)). *For $a, b \in \mathbb{Z}/D\mathbb{Z}$, we have*

$$\frac{1}{\varphi(D)} \sum_{\chi} \overline{\chi(a)}\chi(b) = \begin{cases} 1 & a = b \\ 0 & a \neq b, \end{cases}$$

where φ denotes Euler's totien function and χ runs over all Dirichlet characters modulo D.

Proof. Exercise. □

Theorem 3.11.8. *Let a and D be co-prime positive integers, then there are infinitely many primes of the form $a + nD$. More precisely, the set*

$$\{p \in \mathbb{P} \mid p \equiv a \mod D\}$$

has Dirichlet density $1/\varphi(D)$ in $\mathbb{P}$, where φ denotes Euler's totient function.

Proof. We sketch the proof and divide it into several sub-claims:
Claim 1: We have

$$\log L(\chi, s) = \sum_{p\in\mathbb{P}} \frac{\chi(p)}{p^s} + O(1), \quad \text{as } s \to 1.$$

Proof of the Claim 1: This follows easily from the following computation:

$$\begin{aligned}\log L(\chi,s) = \log \prod_{p\in\mathbb{P}}(1-\chi(p)p^{-s})^{-1} &= \sum_{p\in\mathbb{P}}\sum_{n=1}^{\infty}\frac{\chi(p)^n}{np^{ns}}\\ &= \sum_{p\in\mathbb{P}}\frac{\chi(p)}{p^s} + \underbrace{\sum_{p\in\mathbb{P}}\sum_{n=2}^{\infty}\frac{\chi(p)^n}{np^{ns}}}_{O(1)\text{ as } s\to 1}.\end{aligned}$$

Claim 2: For each non-trivial Dirichlet character χ, we have $L(\chi,1)\neq 0$.
Proof of the Claim 2: This will be shown later.

We are now ready to prove the Theorem. By Claim 2 and Lemma **??**, we have

$$\frac{1}{\varphi(D)}\sum_{\chi}\overline{\chi}(a)\log L(\chi,s) = -\frac{1}{\varphi(D)}\log(s-1)+O(1), \qquad \text{as } s\to 1. \tag{3.21}$$

On the other hand, by Lemma 3.11.7, equation (3.20) and Claim 1, we have

$$\frac{1}{\varphi(D)}\sum_{\chi}\overline{\chi}(a)\log L(\chi,s) = \sum_{p\in\mathbb{P}_a}\frac{1}{p^s}+O(1),$$

where we write $\mathbb{P}_a := \{p\in\mathbb{P} \mid p\equiv a \mod D\}$. Dividing the latter equation by $-\log(s-1)$ and observing (3.21) gives

$$\lim_{s\to 1^+}\frac{\sum_{p\in\mathbb{P}_a}\frac{1}{p^s}}{-\log(s-1)} = \frac{1}{\varphi(D)}.$$

□

Thus, in the sense of Dirichlet densities, about $1/\varphi(D)$ of the primes are congruent a modulo D. In Chapter 2.5, we have used the non-vanishing of $\zeta(s)$ on $\mathrm{Re}(s)\geq 1$ to deduce an asymptotic formula for the prime counting function. By combining the methods of Chapter 2.5 with the idea of the proof of Dirichlet's Theorem, one can prove:

Theorem 3.11.9. *For co-prime positive integers a and D, we have*

$$\sum_{\substack{p\leq x\\ p\equiv a \mod D}} 1 \sim \frac{1}{\varphi(D)}\frac{x}{\log x} \qquad \textit{as } x\to\infty.$$

Proof. We only sketch the argument. We define

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

The function $\vartheta_{\mathbb{1}}$ for the trivial character $\mathbb{1}$ differs from the function ϑ defined in Chapter 2.5 only in finitely many terms. So, we deduce from Corollary 2.5.5 the convergence of the integral

$$\int_1^{\infty}\frac{\vartheta_{\mathbb{1}}(x)-x}{x^2}dx.$$

Similarly, one proves for every non-trivial character χ the convergence of

$$\int_1^\infty \frac{\vartheta_\chi(x)}{x^2}dx.$$

Now, we define

$$\vartheta_a(x) = \sum_{\substack{p\leq x\\ p\equiv a \mod D}} \log p,$$

and observe by the orthogonality relation

$$\vartheta_a(x) = \frac{1}{\varphi(D)}\sum_\chi \overline{\chi(a)}\vartheta_\chi(x).$$

The convergence of the above integrals implies

$$\int_1^\infty \frac{\vartheta_a(x)-x}{x^2}dx < \infty.$$

By a similar argument as in Theorem 2.5.6 one deduces from the convergence of this integral the asymptotic formula

$$\vartheta_a(x) \sim \frac{1}{\varphi(x)}x \quad \text{as } x\to\infty.$$

And finally, we can use this formula to prove

$$\sum_{\substack{p\leq x\\ p\equiv a \mod D}} 1 \sim \frac{1}{\varphi(D)}\frac{x}{\log x} \quad \text{as } x\to\infty,$$

by the same argument as in the Prime Number Theorem. □

We can restate this result in terms of natural densities:

Definition 3.11.10. Let $S\subseteq T\subseteq \mathbb{N}$ be sets of positive integers. We define the natural density of S in T as

$$\lim_{x\to\infty}\frac{\sum_{\substack{n\in S\\ n\leq x}} 1}{\sum_{\substack{n\in T\\ n\leq x}} 1}$$

if the limit exists.

Thus, we obtain:

Corollary 3.11.11. *For co-prime positive integers a and D the natural density of $\{p\in\mathbb{P} \mid p\equiv a \mod D\}$ in the set of all primes $\mathbb{P}$ is $1/\varphi(D)$.*

An immediate consequence of Dirichlet's result on primes in arithmetic progressions together with the Kronecker-Weber Theorem is the following special case of Chebotarev's Density Theorem for abelian extensions of $\mathbb{Q}$:

Corollary 3.11.12. *Let $K/\mathbb{Q}$ be a finite abelian extension and $g \in Gal(K/\mathbb{Q})$. The natural and Dirichlet density of the set*

$$\{p \in \mathbb{P} \mid p \text{ unramified and } \mathrm{Frob}_p = g\}$$

in the set of all primes $\mathbb{P}$ is $1/\# Gal(K/\mathbb{Q})$.

Proof. By the Kronecker-Weber Theorem, there is a positive integer D such that $K \subseteq \mathbb{Q}(\zeta_D)$. Let us denote by F_p the pre-image of $\mathrm{Frob}_p(K/\mathbb{Q}) \in G = \mathrm{Gal}(K/\mathbb{Q})$ under

$$\mathrm{Gal}(\mathbb{Q}(\zeta_D)/\mathbb{Q}) \twoheadrightarrow G.$$

Since the Frobenius element of $\mathbb{Q}(\zeta_D)$ at p maps to the Frobenius element of K at p and since F_p has exactly $\frac{\#\mathrm{Gal}(\mathbb{Q}(\zeta_D)/\mathbb{Q})}{\#\mathrm{Gal}(K/\mathbb{Q})}$ elements, it suffices to prove the claim for $K = \mathbb{Q}(\zeta_D)$. Let us denote by $a + D\mathbb{Z}$ the image of $g \in G$ under the isomorphism in Theorem 3.10.4 (a)

$$\mathrm{Gal}(\mathbb{Q}(\zeta_D)/\mathbb{Q}) \xrightarrow{\sim} (\mathbb{Z}/D\mathbb{Z})^\times.$$

By Theorem 3.10.4, the image of $\mathrm{Frob}_p \in G$ under this isomorphism is $p + D\mathbb{Z}$. Thus, we get

$$\{p \in \mathbb{P} \mid p \text{ unramified and } \mathrm{Frob}_p = g\} = \{p \in \mathbb{P} \mid p \equiv a \mod D\},$$

and the result follows from Dirichlet's Theorem for primes in arithmetic progressions. □

3.12 Dirichlet characters as Galois representations

In this section, we will use the Kronecker–Weber Theorem to relate Dirichlet characters to 1-dimensional Galois representations.

We have introduced Dirichlet characters as group homomorphisms $(\mathbb{Z}/D\mathbb{Z})^\times \to \mathbb{C}^\times$. A more convenient way to view Dirichlet characters is provided by the pro-finite group

$$\widehat{\mathbb{Z}}^\times = \varprojlim_{d \in (\mathbb{N}, |)} (\mathbb{Z}/d\mathbb{Z})^\times.$$

For every $d \in \mathbb{N}$, we have canonical projections

$$\widehat{\mathbb{Z}}^\times \twoheadrightarrow (\mathbb{Z}/d\mathbb{Z})^\times.$$

This allows us to associate to every Dirichlet character a continuous homomorphism $\widehat{\mathbb{Z}}^\times \to \mathbb{C}^\times$. The following Lemma shows that the continuous homomorphisms $\widehat{\mathbb{Z}}^\times \to \mathbb{C}^\times$ are precisely the Dirichlet characters.

Lemma 3.12.1. *Every continuous group homomorphism $\chi\colon \widehat{\mathbb{Z}}^\times \to \mathbb{C}^\times$ factors through*

$$\chi\colon \widehat{\mathbb{Z}}^\times \to (\mathbb{Z}/d\mathbb{Z})^\times \to \mathbb{C}^\times$$

for some $d \in \mathbb{N}$. In particular, we can identify the set of all primitive Dirichlet characters with $Hom_{cont}(\widehat{\mathbb{Z}}^\times, \mathbb{C}^\times)$.

Proof. Exercise. □

Let us now relate the Kronecker-Weber Theorem to Dirichlet characters. The Kronecker-Weber Theorem gave us the explicit description $\mathbb{Q}^{ab} = \mathbb{Q}(\zeta_n \mid n \in \mathbb{N})$ for the maximal abelian field extension of $\mathbb{Q}$. This allowed us to compute the Galois group of $\mathbb{Q}^{ab}/\mathbb{Q}$. More precisely, we have an explicit isomorphism

$$\mathrm{Gal}(\mathbb{Q}^{ab}/\mathbb{Q}) \cong \widehat{\mathbb{Z}}^\times = \varprojlim_{n\in(\mathbb{N},|)} (\mathbb{Z}/n\mathbb{Z})^\times.$$

Using this isomorphism, we get:

Corollary 3.12.2. *We have an isomorphism*

$$Hom_{cont}(Gal(\overline{\mathbb{Q}}/\mathbb{Q}), \mathbb{C}^\times) \xrightarrow{\sim} \{\chi \text{ primitive Dirichlet characters}\}, \quad \rho \mapsto \chi_\rho.$$

Proof. It suffices to show that

$$\mathrm{Hom}_{cont}(\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}), \mathbb{C}^\times) \cong \mathrm{Hom}_{cont}(\mathrm{Gal}(\mathbb{Q}^{ab}/\mathbb{Q}), \mathbb{C}^\times),$$

but since $\mathbb{C}^\times$ is an abelian group, any continuous group homomorphism $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \mathbb{C}^\times$ factors through the Galois group of some abelian Galois extension $K/\mathbb{Q}$. Hence, it factors through the maximal such extension which is $\mathbb{Q}^{ab}$. □

The continuous homomorphisms

$$\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \mathbb{C}^\times = \mathrm{GL}_1(\mathbb{C})$$

are exactly the 1-dimensional complex Galois representations. Next week, we will study such complex Galois representations more carefully. In particular, we will associate an L-functions to each complex Galois representation.

Outlook

This identification of continuous group homomorphisms $\mathrm{Gal}(\mathbb{Q}^{ab}/\mathbb{Q}) \to \mathbb{C}^\times$ with primitive Dirichlet characters looks quite innocent, but it is the starting point of many interesting and deep conjectures in modern number theory. Indeed, the left hand side of the identification

$$\mathrm{Hom}_{cont}(\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}), \mathbb{C}^\times) \xrightarrow{\sim} \{\chi \text{ primitive Dirichlet characters}\},$$

can be seen as the set of all 1-dimensional (complex) Galois representations. Galois representations play an important role in modern number theory for several reasons: The absolute Galois group of a number field itself contains many interesting information but it is a quite complicated object. Studying the representations of a group is a general approach to gain a better understanding of the group itself. Furthermore, Galois representations appear naturally in arithmetic geometry (usually with $\mathbb{Q}_l$ coefficients for a prime l). Many interesting L-functions can be described in a natural way as L-functions associated to such Galois representations. In the next section, we will outline the definition of an L-function associated to a (complex) Galois representation, so called Artin L-functions. Unfortunately, it is notoriously difficult to understand the analytic properties, like functional equations or meromorphic continuation, of such L-functions defined by representations. In particular, we will see that in the simple case of 1-dimensional (complex) Galois representations of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ we can identify the Artin L-function of the 1-dimensional representation ρ with the Dirichlet L-function of χ_ρ. We have already seen that it is not so difficult to prove analytic properties of Dirichlet L-functions. This is a general theme in the theory of L-functions. The Langland's program can be seen as a conjectural generalization of this to higher dimensional representations: One wants to associate to every L-function of a Galois representations a certain 'automorphic' L-function which is more accessible from an analytic point of view. For one-dimensional complex Galois representations of $\mathbb{Q}$, the associated 'automorphic' L-functions are exactly the Dirichlet L-functions. Thus, the Kronecker-Weber Theorem can be seen as the motivating example for the Langland's program.

3.13 *Complex representations*

In this section, we outline the definition of Artin L-functions. These L-functions are associated to certain complex representations. We will skip most of the proofs and cite several results from representation theory.

Definition 3.13.1. Let G be a (topological) group and V a finite dimensional $\mathbb{C}$-vector space. Let us denote by $\mathrm{GL}(V)$ the set of all $\mathbb{C}$-linear automorphisms of V. A *(complex) representation* (ρ, V) of G is a group homomorphism

$$\rho\colon G \to \mathrm{GL}(V).$$

In other words, it is a $\mathbb{C}$-linear action of G on V. If G is a topological group, we call such a representation *continuous* if the map

$$G \times V \to V, \quad (g, v) \mapsto \rho(g)(v)$$

is continuous (here, V is equipped with the usual topology induced by any isomorphism $V \cong \mathbb{C}^n$). A homomorphism $\varphi\colon (\rho', W) \to (\rho, V)$ of representations $\rho'\colon G \to \mathrm{GL}(W)$ to $\rho\colon G \to \mathrm{GL}(V)$ is a $\mathbb{C}$-linear map $\varphi\colon W \to V$ which is compatible with the G-actions, i.e.,

$$\varphi(\rho(g)(w)) = \rho'(g)(\varphi(w)).$$

A sub-representation is a linear subspace $W \subseteq V$ such that $\rho(g)|_W \in \mathrm{GL}(W)$. A representation ρ is called *irreducible* if there is no non-zero proper sub-representation.

Let us give some examples:

Example 3.13.2.(a) The *trivial* representation of a finite group G on a finite dimensional $\mathbb{C}$-vector space G is the representation

$$G \to \mathrm{GL}(V), \quad g \mapsto \mathrm{id}_V.$$

Since any $\mathbb{C}$-linear subspace is a sub-representation, this representation is irreducible if and only if $\dim V = 1$. For $V = \mathbb{C}$, we will often write $\mathbb{1}_G$ for the trivial representation $\mathbb{1}_G\colon G \to \mathrm{GL}(\mathbb{C})$.

(b) For a finite group G, let us consider the $\mathbb{C}$-vector space $\mathbb{C}[G] := \bigoplus_{g\in G} \mathbb{C}$ with the standard basis $(e_g)_{g\in G}$. The *(right) regular representation* of a group G is $\mathbb{C}[G]$ together with

$$\rho\colon G \to \mathrm{GL}(\mathbb{C}[G]), \quad g \mapsto (e_h \mapsto e_{gh}).$$

For a non-trivial group G, this representation is never irreducible since the kernel of the map

$$\mathbb{C}[G] \to \mathbb{C}, \quad \sum_{g\in G} a_g e_g \mapsto \sum_{g\in G} a_g$$

is a non-trivial proper sub-representation.

(c) The identity $\mathrm{GL}(V) \to \mathrm{GL}(V)$ is a representation of $\mathrm{GL}(V)$.

(d) Given two representations $\rho_1\colon G \to V_1$ and $\rho_2\colon G \to V_2$ we can form their direct sum

$$\rho_1 \oplus \rho_2\colon G \to \mathrm{GL}(V_1) \times \mathrm{GL}(V_2) \subseteq \mathrm{GL}(V_1 \oplus V_2), \quad g \mapsto (\rho_1(g), \rho_2(g)).$$

The irreducible representations can be seen as the building blocks for all finite dimensional (complex) representations of a finite group:

Theorem 3.13.3 (Maschke's Theorem)**.** *Let G be a finite group and ρ a representation on a finite dimensional $\mathbb{C}$-vector space. Then ρ is isomorphic to a finite direct sum of irreducible representations of G.*

Proof. See §1.4, Theorem 2 in Serre's book 'Linear representations of finite groups'[46]. □

[46] Jean-Pierre Serre. *Linear representations of finite groups*. Graduate Texts in Mathematics, Vol. 42. Springer-Verlag, New York-Heidelberg, 1977. ISBN 0-387-90190-6. Translated from the second French edition by Leonard L. Scott

Next, let us study the functoriality of representations under homomorphisms of groups $f\colon H \to G$. Of course, given a representation $\rho\colon G \to \mathrm{GL}(V)$, we obtain a representation of H by pre-composition with H, i.e.,

$$f^*\rho := \rho \circ f\colon H \to \mathrm{GL}(V).$$

Often, the morphism $f\colon H \to G$ is just the inclusion of a subgroup. In this case, we will call $f^*\rho$ the *restriction of* ρ *to* H and we will also denote it by $\mathrm{Res}^G_H \rho := f^*\rho$. Conversely, we have the following construction:

Definition 3.13.4. Let $H \subseteq G$ be a subgroup of a finite group G and $\rho\colon H \to \mathrm{GL}(V)$ a representation of H. We define

$$\mathrm{Ind}^H_G \rho := \{\Phi \in \mathrm{Map}(G,V) \mid \Phi(xh^{-1}) = h\Phi(x) \text{ for all } h \in H, x \in G\}.$$

as the representation of G with the $\mathbb{C}$-linear G-action $(g\Phi)(x) := \Phi(g^{-1}x)$. This representation is called *the induced representation* of ρ.

Example 3.13.5. Let G be a finite group and consider the trivial subgroup $H := \{e\} \subseteq G$ together with the trivial representation $\mathbb{1}_H$ of H. In this case, the induced representation turns out to be the regular representation of G, indeed

$$\mathrm{Ind}^G_{\{e\}} \mathbb{1}_H = \mathrm{Hom}_{Set}(G,\mathbb{C}) \cong \bigoplus_{g\in G} \mathbb{C} = \mathbb{C}[G],$$

and it is easily checked that the G-action on $\mathrm{Ind}^G_{\{e\}} \mathbb{1}_H$ coincides with the G-action on the regular representation $\mathbb{C}[G]$.

For finite groups $H \subseteq G$, we have constructed functors

$$\mathrm{Res}^G_H\colon \mathrm{Rep}_G \to \mathrm{Rep}_H$$

and

$$\mathrm{Ind}^H_G\colon \mathrm{Rep}_H \to \mathrm{Rep}_G,$$

where Rep_G (resp. Rep_H) denotes the category of complex representations of G (resp. H). The following Theorem shows that these functors are adjoint:

Lemma 3.13.6 (Frobenius Reciprocity)**.** *Let $H \subseteq G$ be finite groups, ρ a representation of G and π a representation of H, then*

$$Hom_{\mathrm{Rep}_G}(\mathrm{Ind}^H_G \pi, \rho) \cong Hom_{\mathrm{Rep}_H}(\pi, \mathrm{Res}^G_H \rho).$$

Proof. Exercise. □

Finally, let us state Schur's Lemma:

Lemma 3.13.7. *Let (ρ, V) be an irreducible finite-dimensional representation of a finite group. Then*

$$End_{\mathrm{Rep}_G}(\rho) = \mathbb{C}.$$

Proof. See §2.2, Proposition 4 in Serre's book 'Linear representations of finite groups'[47]. □

Let us observe that Schur's Lemma implies for two irreducible complex representations ρ and $\tilde{\rho}$ of a finite group G:

$$\mathrm{Hom}_{\mathrm{Rep}_G}(\rho, \tilde{\rho}) = \begin{cases} 0 & \rho \ncong \tilde{\rho} \\ \mathbb{C} & \rho \cong \tilde{\rho}. \end{cases}$$

Using Frobenius reciprocity and Schur's Lemma, we get:

Corollary 3.13.8. *Let G be a finite group, then*

$$\mathbb{C}[G] \cong \bigoplus_{(\rho,W) \text{ irreducible rep. of } G} (\rho, W)^{\dim W},$$

where the direct sum runs over all isomorphism classes of irreducible representations of G.

Proof. By Maschke's Theorem, we already know that

$$\mathbb{C}[G] \cong \bigoplus_{(\rho,W) \text{ irreducible rep. of } G} (\rho, W)^{n_W},$$

where (ρ, W) runs over the isomorphism classes of irreducible representations of G and n_W are certain non-negative integers. It remains to show $n_W = \dim_{\mathbb{C}} W$. By Schur's Lemma, we have

$$n_W = \dim_{\mathbb{C}} \mathrm{Hom}_{\mathrm{Rep}_G}(\mathbb{C}[G], \rho).$$

On the other hand, we have seen in Example 3.13.5 that $\mathbb{C}[G] = \mathrm{Ind}_G^{\{e\}} \mathbb{1}_{\{e\}}$ where $\mathbb{1}_{\{e\}}$ denotes the trivial representation. Together with Frobenius reciprociy, we get

$$\begin{aligned} n_W &= \dim_{\mathbb{C}} \mathrm{Hom}_{\mathrm{Rep}_G}(\mathrm{Ind}_G^{\{e\}} \mathbb{1}_{\{e\}}, \rho) \\ &= \dim_{\mathbb{C}} \mathrm{Hom}_{\mathbb{C}}(\mathbb{1}_{\{e\}}, \mathrm{Res}_{\{e\}}^G \rho) = \dim_{\mathbb{C}} \mathrm{Hom}_{\mathbb{C}}(\mathbb{C}, W) = \dim_{\mathbb{C}} W. \end{aligned}$$

□

3.14 *Artin L-functions*

The aim of this section is to associate an L-function to every complex Galois representation of number fields, so called *Artin L-functions*. Such L-functions associated to Galois representations are in general rather difficult to study from an analytic point of view. In the case of 1-dimensional Galois representations of the absolute Galois group of $\mathbb{C}$, the Kronecker-Weber Theorem allows us to compare Artin L-functions to Dirichlet L-functions.

[47] Jean-Pierre Serre. *Linear representations of finite groups*. Graduate Texts in Mathematics, Vol. 42. Springer-Verlag, New York-Heidelberg, 1977. ISBN 0-387-90190-6. Translated from the second French edition by Leonard L. Scott

Let us now introduce Artin L-functions. Let L/K be a finite Galois extension of number fields with Galois group $G = \mathrm{Gal}(L/K)$. For a prime $\mathfrak{P} \subseteq \mathcal{O}_L$ over $\mathfrak{p} \subseteq \mathcal{O}_K$ let us denote by $I_{\mathfrak{P}} \subseteq G_{\mathfrak{P}}$ the inertia and the decomposition group. Let us denote by $\mathrm{Frob}_{\mathfrak{P}} \in G_{\mathfrak{P}}$ any element mapping to the Frobenius morphism of the field extension $\kappa(\mathfrak{P})/\kappa(\mathfrak{p})$ under

$$G_{\mathfrak{P}} \twoheadrightarrow \mathrm{Gal}(\kappa(\mathfrak{P})/\kappa(\mathfrak{p})).$$

Any other choice of such a Frobenius element differs by an element of the inertia group $I_{\mathfrak{P}}$. Let $\rho\colon \mathrm{Gal}(L/K) \to V$ be a finite-dimensional representation of $\mathrm{Gal}(L/K)$. Le us write $V^{I_{\mathfrak{P}}}$ for the fixed-points of the subgroup $I_{\mathfrak{P}}$ under the action given by ρ. Thus, the group $I_{\mathfrak{P}}$ acts trivially on $V^{I_{\mathfrak{P}}}$ and hence, the $\mathbb{C}$-linear automorphism

$$\rho(\mathrm{Frob}_{\mathfrak{P}})\colon V^{I_{\mathfrak{P}}} \to V^{I_{\mathfrak{P}}}$$

does not depend upon the choice of the Frobenius element $\mathrm{Frob}_{\mathfrak{P}}$. Furthermore, let us observe that the sets $\{\mathrm{Frob}_{\mathfrak{P}}\}$ and $\{\mathrm{Frob}_{\mathfrak{P}'}\}$ of Frobenius elements at $\mathfrak{P}$ respectively $\mathfrak{P}'$, as well as the decomposition groups and inertia groups, for two prime ideal $\mathfrak{P}, \mathfrak{P}' \mid \mathfrak{p}$ are conjugate inside of $\mathrm{Gal}(L/K)$. Since the determinant of an element of $\mathrm{GL}(V)$ does only depend on the conjugacy class in $\mathrm{GL}(V)$, for $s \in \mathbb{C}$, the complex number

$$\det(1 - \rho(\mathrm{Frob}_{\mathfrak{P}})N\mathfrak{p}^{-s}; V^{I_{\mathfrak{P}}})$$

is well-defined and does only depend on $\mathfrak{p}$. This allows us to make the following definition:

Lemma/Definition 3.14.1 (Artin L-function)**.** Let L/K be a finite Galois extension of number fields with Galois group $G = \mathrm{Gal}(L/K)$ and

$$\rho\colon G \to \mathrm{GL}(V)$$

a finite dimensional complex representation of G. The Artin L-function of ρ is defined by

$$\mathcal{L}(L/K, \rho, s) := \prod_{\mathfrak{p} \subseteq \mathcal{O}_K} \frac{1}{\det(1 - \rho(\mathrm{Frob}_{\mathfrak{P}})N\mathfrak{p}^{-s}; V^{I_{\mathfrak{P}}})},$$

here $\mathfrak{p}$ runs through all non-zero prime ideals of K, $\mathfrak{P}$ is a chosen prime above $\mathfrak{p}$ and $\mathrm{Frob}_{\mathfrak{P}} \in G_{\mathfrak{P}}$ is a Frobenius element in the decomposition group $G_{\mathfrak{P}}$ of $\mathfrak{P}$. The product converges absolutely and locally uniformly on $\mathrm{Re}(s) > 1$ to a holomorphic function.

Proof. By the comment preceding the definition, the local factors are well-defined. Let us briefly address the question of convergence. It is enough to prove that for any $\epsilon > 0$ the product converges absolutely and uniformly for every real numbers $s > 1+\epsilon$. For a prime $\mathfrak{p}$ of $\mathcal{O}_K$, the determinant in the local factor

$$\det(1-\rho(\mathrm{Frob}_{\mathfrak{P}})N\mathfrak{p}^{-s}; V^{I_{\mathfrak{P}}})$$

coincides with $X^n P_{\rho(\mathrm{Frob}_{\mathfrak{P}})}(X^{-1})$ evaluated at $X = N\mathfrak{p}^{-s}$ where $n = \dim_{\mathbb{C}} V^{I_{\mathfrak{P}}}$ and $P_{\rho(\mathrm{Frob}_{\mathfrak{P}})}(X) \in \mathbb{C}[X]$ denotes the characteristic polynomial of $\rho(\mathrm{Frob}_{\mathfrak{P}})\colon V^{I_{\mathfrak{P}}} \to V^{I_{\mathfrak{P}}}$. Since $\rho(\mathrm{Frob}_{\mathfrak{P}})$ is of finite order, the characteristic polynomial factors as

$$P_{\rho(\mathrm{Frob}_{\mathfrak{P}})}(X) = \prod_{i=1}^{n}(X-\epsilon_i)$$

for certain roots of unity ϵ_i. So,

$$\det(1-\rho(\mathrm{Frob}_{\mathfrak{P}})N\mathfrak{p}^{-s}; V^{I_{\mathfrak{P}}}) = \prod_{i=1}^{n}(1-\epsilon_i N\mathfrak{p}^{-s}).$$

For $s > 1$, we have $0 < N\mathfrak{p}^{-s} < 1$ and hence, we can estimate

$$|1-\epsilon_i N\mathfrak{p}^{-s}| \geq 1 - N\mathfrak{p}^{-s}.$$

This gives the estimate

$$\det(1-\rho(\mathrm{Frob}_{\mathfrak{P}})N\mathfrak{p}^{-s}; V^{I_{\mathfrak{P}}})^{-1} \leq \prod_{i=1}^{n}(1-N\mathfrak{p}^{-s}).$$

The convergence follows now from the fact that the Dedekind zeta function admits a convergent Euler product for $\mathrm{Re}(s) > 1$. Let us first recall the definition of the Dedekind zeta function:

$$\zeta_K(s) := \sum_{\mathfrak{p}\subseteq\mathcal{O}_K} \frac{1}{N\mathfrak{p}^s}.$$

This function converges absolutely and uniformly for $\mathrm{Re}(s) > 1$ and admits an Euler product

$$\zeta_K(s) = \prod_{\mathfrak{p}}(1-N\mathfrak{p}^{-s})^{-1}.$$

This has been proven in ANT 1. See also Neukrirch's book on Algebraic Number Theory[48], Ch. VII, Thm. (5.2). □

The following Theorem discusses some important properties of Artin L-functions.

Theorem 3.14.2. *Let L/K be a Galois extension of number fields with Galois group $G = Gal(L/K)$. Then*

(a) ***(Trivial representation)*** *For the trivial representation we have $\mathcal{L}(L/K, \mathbb{1}, s) = \zeta_K(s)$.*

[48] Jürgen Neukirch. *Algebraic number theory*, volume 322 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1999. ISBN 3-540-65399-6. DOI: 10.1007/978-3-662-03983-0. URL https://doi.org/10.1007/978-3-662-03983-0. Translated from the 1992 German original and with a note by Norbert Schappacher, With a foreword by G. Harder

(b) ***(Additivity)*** *If ρ_1 and ρ_2 are two finite dimensional representations of G then*
$$\mathcal{L}(L/K, \rho_1 \oplus \rho_2, s) = \mathcal{L}(L/K, \rho_1, s)\mathcal{L}(L/K, \rho_2, s).$$

(c) ***(Restriction)*** *If L'/K is Galois with $L' \supseteq L \supseteq K$ and ρ is a finite dimensional representation of G then $\mathcal{L}(L'/K, f^*\rho, s) = \mathcal{L}(L/K, \rho, s)$ where $f\colon Gal(L'/K) \twoheadrightarrow Gal(L/K) = G$.*

(d) ***(Induction)*** *If K'/K is Galois with $L \supseteq K' \supseteq K$ and ρ is a representation of $H = Gal(L/K')$ then $\mathcal{L}(L/K', \rho, s) = \mathcal{L}(L/K, \operatorname{Ind}_G^H \rho, s)$*

Proof. The proofs are not difficult, but we refer to Neukirch's book[49] on Algebraic Number Theory, Ch. VII, Thm. (10.4). □

Let us observe that the statement (b) of the previous Theorem says that the Artin L-function does not really depend on the field L.

An important Corollary is the following result:

Corollary 3.14.3. *For a Galois extension of number fields L/K with Galois group $G = Gal(L/K)$ we have*
$$\zeta_L(s) = \zeta_K(s) \prod_{(\rho,V)\neq \mathbb{1}_G \text{ irreducible}} \mathcal{L}(L/K, \rho, s)^{\dim_{\mathbb{C}} V}.$$

Proof. According to Theorem 3.14.2 (a) we have
$$\zeta_L(s) = \mathcal{L}(L/L, \mathbb{1}, s).$$
Now, using Theorem 3.14.2 (d) gives
$$\zeta_L(s) = \mathcal{L}(L/K, \operatorname{Ind}_G^{\{e\}} \mathbb{1}, s).$$
On the other hand, we have already seen that
$$\operatorname{Ind}_G^{\{e\}} \mathbb{1}_G = \mathbb{C}[G] = \bigoplus_{(\rho,V)\text{irreducible}} \rho^{\dim_{\mathbb{C}} V}.$$
Finally, we obtain the statement of the Corollary using Theorem 3.14.2 (b). □

In general, it is rather difficult to prove analytic properties of L-functions associated to Galois representations. Artin formulated the following conjecture which is still vastly open:

Conjecture (Artin conjecture)**.** *Let L/K be a Galois extension of number fields and $\rho\colon Gal(L/K) \to GL(V)$ a non-trivial irreducible representation then $\mathcal{L}(L/K, \rho, s)$ extends to a holomorphic function on $\mathbb{C}$.*

The Kronecker-Weber Theorem allows us to compare 1-dimensional Galois representations of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ to Dirichlet L-functions:

[49] Jürgen Neukirch. *Algebraic number theory*, volume 322 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1999. ISBN 3-540-65399-6. DOI: 10.1007/978-3-662-03983-0. URL https://doi.org/10.1007/978-3-662-03983-0. Translated from the 1992 German original and with a note by Norbert Schappacher, With a foreword by G. Harder

Theorem 3.14.4. *Let $L/\mathbb{Q}$ be a Galois extension of number fields and*

$$\rho\colon \mathit{Gal}(L/\mathbb{Q}) \to GL(\mathbb{C}) = \mathbb{C}^\times$$

a 1*-dimensional Galois representation. Then, there exists a primitive Dirichlet character χ_ρ of conductor d such that*

$$\mathcal{L}(L/\mathbb{Q}, \rho, s) = L(\chi_\rho, s).$$

Proof. Since $\mathbb{C}^\times$ is an abelian group the homomorphism ρ factors through the abelian Galois group

$$\mathrm{Gal}(\overline{\mathbb{Q}}^{\ker\rho}/\mathbb{Q}) \cong \mathrm{im}(\rho) \subseteq \mathbb{C}^\times$$

of the fixed field of $\ker\rho$ over $\mathbb{Q}$. By the Kronecker-Weber Theorem any such abelian extension is contained in a cyclotomic field. By Theorem 3.14.2 (b), we may thus without loss of generality assume that $L = \mathbb{Q}(\zeta_d)$, where d is the minimal positive integers such that $\mathbb{Q}(\zeta_d)$ contains $\overline{\mathbb{Q}}^{\ker\rho}$. By the minimality of d, the homomorphism

$$\chi_\rho\colon (\mathbb{Z}/d\mathbb{Z})^\times \cong \mathrm{Gal}(\mathbb{Q}(\zeta_d)/\mathbb{Q}) \xrightarrow{\rho} \mathbb{C}^\times$$

is a primitive Dirichlet character of conductor d. It remains to show that $L(\chi_\rho, s) = \mathcal{L}(L/\mathbb{Q}, \rho, s)$. Since both L-functions are given for $\mathrm{Re}(s) > 1$ by an Euler product it suffices to prove

$$\det(1 - \rho(\mathrm{Frob}_\mathfrak{P})p^{-s}; \mathbb{C}^{I_p}) = (1 - \chi_\rho(p)p^{-s})$$

for all primes p, where $\mathfrak{P} \subseteq \mathbb{Z}[\zeta_d]$ denotes a prime ideal above p. If $p \mid d$ then $\chi_\rho(p) = 0$ and the right hand side is 1. On the other hand, p is ramified in $\mathbb{Q}(\zeta_d)$. By Theorem 3.10.4 (e), the inertia group I_p corresponds to the subgroup $(\mathbb{Z}/p^{v_p(d)}\mathbb{Z})^\times$ of $(\mathbb{Z}/d\mathbb{Z})^\times$. The image of this group under ρ is non-trivial, otherwise $\overline{\mathbb{Q}}^{\ker\rho}$ would be contained in $\mathbb{Q}(\zeta_{d'})$ with $d' = d/p^{v_p(d)}$. Hence, $\mathbb{C}^{I_p} = \{0\}$ and we conclude that also the Euler factor in the right hand side of the above equation is trivial for $p \mid d$. If $p \nmid d$ then p is unramified at p and the Frobenius morphism $\mathrm{Frob}_\mathfrak{P}$ at p corresponds to $p + d\mathbb{Z}$ under $(\mathbb{Z}/d\mathbb{Z})^\times \cong \mathrm{Gal}(\mathbb{Q}(\zeta_d)/\mathbb{Q})$. Thus $\rho(\mathrm{Frob}_\mathfrak{P}) = \chi_\rho(p)$ and we deduce

$$\det(1 - \rho(\mathrm{Frob}_\mathfrak{P})p^{-s}; \mathbb{C}^{I_p}) = \det(1 - \chi_\rho(p)p^{-s}; \mathbb{C}) = (1 - \chi_\rho(p)p^{-s})$$

as desired. □

We have already seen that Dirichlet L-functions of a non-trivial Dirichlet character are entire. Thus, the Kronecker-Weber Theorem proves the Artin conjecture for all 1-dimensional Galois representations over $\mathbb{Q}$. Furthermore, we can now prove the non-vanishing of $L(\chi, 1)$ for any non-trivial Dirichlet character. This has been used in the proof of Dirichlet's Theorem:

Corollary 3.14.5. *For any non-trivial Dirichlet character χ of modulus D, we have $L(\chi, 1) \neq 0$.*

Proof. By the previous Theorem together with Corollary 3.14.3, we see

$$\frac{\zeta_K(s)}{\zeta(s)} = \prod_{\chi \neq \mathbb{1}} L(\chi, s)$$

where $K = \mathbb{Q}(\zeta_D)$ and χ runs through all non-trivial Dirichlet characters modulo D. By the analytic class number formula (proven in ANT 1), we know that the zeta functions on the right hand side have simple poles and that the residues of $\zeta(s)$ and $\zeta_K(s)$ at $s = 1$ are non-zero. Hence, the limit $s \to 1$ of the right hand side exists and is non-zero. Since the functions in the product on the right hand side are all holomorphic at $s = 1$, non of them can vanish at $s = 1$ and the result follows. □

4 Tate's thesis

In the first chapter, we have studied the basic properties of the Riemann zeta function. In particular, we have used the Poisson summation formula to prove its functional equation. At the end of the second chapter, we have introduced Artin L-functions which encode many interesting information about number fields but whose analytic properties are rather difficult to study. In the case of Artin L-functions of 1-dimensional complex Galois representations we could use the Kronecker-Weber Theorem to relate them to Dirichlet L-functions. We have already seen that Dirichlet L-functions admit an analytic continuation to the entire complex plane. The aim of this chapter is to prove functional equations for Dirichlet L-functions using Fourier analysis on the ring of adeles. This approach to functional equations (more generally for Hecke L-functions[1]) goes back to the PhD thesis of John Tate. Of course, the functional equation of such L-functions has been known previously, but the proofs where tedious and not very conceptional. It was Tate's insight that the functional equation of Hecke L-functions is a consequence of the Poisson summation formula on the ring of adeles. We will explain Tate's approach in the case of Dirichlet characters, i.e., for $K = \mathbb{Q}$. Although the general case is not much more difficult, we belief that the key ideas can be explained better in the slightly simpler case $K = \mathbb{Q}$.

4.1 Harmonic analysis on locally compact abelian groups

We have already seen several instances of the Fourier inversion formula, e.g. for the groups $(\mathbb{R}, +)$ and $(\mathbb{R}/\mathbb{Z}, +)$. These groups belong to an important class of topological abelian groups called locally compact abelian groups. Locally compact abelian groups provide the correct framework for doing Fourier theory in a more general context. The aim of this lecture is to introduce locally compact abelian groups and to formulate Fourier theory (Harmonic Analysis) in this general context.

Recall that a topological group is a group G equipped with a topol-

[1] Hecke L-functions are generalizations of Dirichlet L-functions to more general number fields.

ogy such that the multiplication

$$m\colon G\times G\to G$$

and the inversion

$$i\colon G\to G$$

are continuous maps.

Definition 4.1.1. A *locally compact abelian group*, sometimes we will use the abbreviation *(LCA)*, is a topological group such that the multiplication is commutative and the underlying topological space is Hausdorff and locally compact, i.e. every $g\in G$ admits a compact neighbourhood.

We have already seen a plenty of examples for such groups:

Example 4.1.2. The following are examples (resp. non-examples) of locally compact abelian groups:

(a) $(\mathbb{R},+)$, $(\mathbb{Z},+)$ and $(\mathbb{T},\cdot)$ where $\mathbb{T}:=\{z\in\mathbb{C}:|z|=1\}\cong\mathbb{R}/\mathbb{Z}$ are examples of locally compact abelian groups, if we equip them with their usual topology.

(b) Every finite abelian group with the discrete topology is LCA.

(c) Every pro-finite abelian group is locally compact[2]. In particular, $\mathrm{Gal}(\mathbb{Q}^{ab}/\mathbb{Q})\cong\widehat{\mathbb{Z}}^\times$ is a locally compact abelian group.

(d) $(\mathbb{Q}_p,+)$ with the topology induced by the p-adic absolute value $|\cdot|_p$ is LCA, since every $x\in\mathbb{Q}_p$ admits the compact neighbourhood $x+\mathbb{Z}_p$.

(e) The group $(\mathbb{Q},+)$ with the subspace topology of $\mathbb{R}$ ist **not** locally compact.

Our next aim is to define the notion of a Haar measure on a locally compact abelian group. Therefore, let us first recall some definitions from measure theory: Recall that the Borel sigma algebra $\mathcal{B}(X)$ on a topological space X is the smallest σ-algebra[3] on X containing all the open subsets. A *Borel measure* $\mu\colon\mathcal{B}(X)\to\mathbb{R}_{\geq 0}\cup\infty$ is a measure[4] on the Borel σ-algebra of X. On $(\mathbb{R},+)$ the Lebesgue measure provides a very nice Borel measure which is invariant under translations. The following definition can be seen as a generalization of the Lebesgue measure to arbitrary locally compact abelian groups:

Definition 4.1.3. Let G be a locally compact abelian group. A *Haar measure* on G is a Borel measure satisfying the following conditions[5]:

(a) μ is *inner regular*, i.e., for any $A\in\mathcal{B}(G)$, we have

$$\mu(A)=\sup\{\mu(K)\mid K\subseteq A\text{ compact }\}.$$

[2] Recall that pro-finite groups are exactly the compact totally disconnected Hausdorff groups

[3] A σ-algebra is a set $\mathcal{A}$ of subsets of X s.t.:

(a) $X\in\mathcal{A}$,

(b) $\mathcal{A}$ is closed under complements,

(c) $\mathcal{A}$ is closed under countable unions.

[4] A measure on a set X with a σ-algebra $\mathcal{A}$ is a function $\mu\colon\mathcal{A}\to\mathbb{R}\cup\{\infty\}$ s.t.

- $\mu(\emptyset)=0$
- μ is σ-additive, i.e., $\mu(\bigcup_{n=1}^\infty E_i)=\sum_{i=1}^\infty\mu(E_i)$ for any countable family of pairwise disjoint Borel sets $E_i\in\mathcal{A}$.

[5] A measure satisfying $(a),(b)$ and (c) is usually called *Radon measure*. Radon measures are a class of Borel measures with many good properties. So, we can think about a Haar measure as a 'nice' Borel measure which is translation invariant.

(b) μ is *outer regular*, i.e., for any $A \in \mathcal{B}(G)$, we have

$$\mu(A) = \inf\{\mu(U) \mid A \subseteq U \text{ open }\}.$$

(c) μ is *locally finite*, i.e., for any compact set $K \subseteq G$ we have $\mu(K) < \infty$.

(d) μ is *translation invariant*, i.e., for any $g \in G$ and $X \in \mathcal{B}(G)$, we have

$$\mu(g + X) = \mu(X).$$

An important example of a Haar measure on $\mathbb{R}$ is given by the Lebesgue measure. It is the unique Haar measure on $\mathbb{R}$ with the property $\mu([0,1]) = 1$. For a general LCA we do not have such distinguished subsets to normalize our measure. Nevertheless, we have:

Theorem 4.1.4. *For any locally compact abelian group G, there exists a Haar measure on G. The Haar measure is uniquely determined up to multiplication with a positive real number.*

Proof. We refer to Theorem 1-8 in the book 'Fourier Analysis on Number Fields'[6] for a proof. □

In some cases, there will be a canonical choice of normalization. For example, for compact LCAs:

Corollary 4.1.5. *If G is a LCA which is compact then there is a unique normalized Haar measure, i.e.,* $\mu(G) = 1$.

Next, we would like to define the Pontryagin dual of a locally compact abelian group. We recall that the set of continuous functions $C(X,Y)$ between topological spaces X and Y can be equipped with the compact-open topology, which is the topology defined by the subbase $\{V(K,U)\}_{K,U}$ with

$$V(K,U) := \{f \in C(X,Y) \mid f(K) \subseteq U\}$$

where U runs over all open subsets of Y and K over the compact subsets of X. The following definition plays an important role for locally compact abelian groups:

Definition 4.1.6. Let G be a locally compact abelian group.

(a) A *character* of a locally compact abelian group G is a continuous homomorphism

$$\chi\colon G \to \mathbb{T} := \{z \in \mathbb{C} : |z| = 1\}.$$

The character $g \mapsto 1$ for all $g \in G$ is called the *trivial character*.

(b) The *Pontryagin dual* of G is the abelian topological group $G^\vee := \mathrm{Hom}_{cts}(G, \mathbb{T})$ with the subspace topology given by the compact-open topology via $G^\vee \subseteq C(G, \mathbb{T})$.

[6] D. Ramakrishnan and R. J. Valenza. *Fourier analysis on number fields*, volume 186 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1999. ISBN 0-387-98436-4. DOI: 10.1007/978-1-4757-3085-2. URL https://doi.org/10.1007/978-1-4757-3085-2

Let us note, that Dirichlet characters fit in this general context of characters of locally abelian groups:

Example 4.1.7. For $G = \widehat{\mathbb{Z}}^{\times}$, the group of characters $G^{\vee}$ consists exactly of the primitive Dirichlet characters.

With this definition, one can prove the following result:

Proposition 4.1.8. *The Pontyagin dual $G^{\vee}$ of a locally compact abelian group G is a locally compact abelian group.*

Proof. We do not prove this in full generality. In all cases of our interest, there will be an explicit description of $G^{\vee}$ which proves that $G^{\vee}$ is a locally compact abelian group. For the general case, we refer to Theorem (23.15) in 'Abstract Harmonic Analysis' by Hewitt and Ross[7]. □

Let following observation about integration of characters will be useful:

Lemma 4.1.9. *For a compact abelian group $(G, \cdot)$, a character χ of G and a Haar measure μ on G, we have*

$$\int_G \chi(g) d\mu(g) = \begin{cases} \mu(G) & \text{if } \chi \text{ is trivial} \\ 0 & \text{if } \chi \text{ is non-trivial} \end{cases}.$$

Proof. Exercise. □

For a locally compact abelian group G, we have a canonical evaluation map $G \to G^{\vee\vee}$ given by $g \mapsto (\chi \mapsto \chi(g))$.

Theorem 4.1.10. *The canonical evaluation map $G \to G^{\vee\vee}$ is a homeomorphism of locally compact abelian groups.*

Proof. Actually, this is not so difficult to prove. Nevertheless, we refer to Theorem (24.2) in 'Abstract Harmonic Analysis' by Hewitt and Ross[8]. □

Let us now formulate Fourier theory in the general context of locally compact abelian groups and discuss some important examples. For the following, we fix a locally compact abelian group G and a Haar measure μ on G.

Definition 4.1.11. We say that two measurable complex valued functions f and h *agree almost everywhere* if the set $\{x \in G \mid f(x) \neq h(x)\}$ has measure 0. This defines an equivalence relation $\sim$ on the space of measurable complex functions. For $p \geq 1$, consider the complex vector space

$$L^p(G) := \left\{ f \colon G \to \mathbb{C} \mid f \text{ measurable s.t. } \int_G |f|^p d\mu < \infty \right\} / \sim$$

[7] E. Hewitt and K. A. Ross. *Abstract harmonic analysis. Vol. I*, volume 115 of *Grundlehren der Mathematischen Wissenschaften*. Springer-Verlag, Berlin-New York, 2nd edition, 1979. ISBN 3-540-09434-2

[8] E. Hewitt and K. A. Ross. *Abstract harmonic analysis. Vol. I*, volume 115 of *Grundlehren der Mathematischen Wissenschaften*. Springer-Verlag, Berlin-New York, 2nd edition, 1979. ISBN 3-540-09434-2

and define the norm $\|\cdot\|_p$ on $L^p(G)$ by

$$\|f\|_p := \left(\int_G |f|^p d\mu\right)^{1/p}.$$

Note, that the norm $\|f\|_2$ is induced by the scalar product[9].

$$\langle f, g\rangle_{L^2} := \int_G f\overline{g}d\mu,$$

In this general setup, we can define the Fourier transform of a measurable function as follows:

Theorem 4.1.12. *Let G be a locally compact abelian group with a Haar measure μ.*

(a) For $f \in L^2(G) \cap L^1(G)$, the Fourier transform

$$\widehat{f}(\chi) := \int_G f(x)\overline{\chi}(x)d\mu(x).$$

gives a well-defined map

$$L^2(G) \cap L^1(G) \to L^2(G^\vee), \quad f \mapsto \widehat{f}.$$

There is a unique Haar measure $\mu^\vee$ on $G^\vee$, called the dual Haar measure *of μ, such that $\|f\|_{L^2(G)} = \|\widehat{f}\|_{L^2(G^\vee)}$ for all $f \in L^2(G) \cap L^1(G)$.*

(b) The Fourier transform of (a) extends to a well-defined isometry

$$L^2(G) \to L^2(G^\vee)$$

such that for all $f \in L^2(G)$ we have

$$\widehat{\widehat{f}}(x) = f(-x)$$

almost everywhere.

Proof. For a proof, we refer to Theorem 3-26 in 'Fourier Analysis on Number Fields'[10]. □

4.2 Local Fourier Analysis

By Ostrowski's Theorem (Theorem 3.1.8), we have seen that the equivalence classes of absolute values on $\mathbb{Q}$ are precisely given by the absolute values $|\cdot|_\infty$ and $\mathbb{Q}_p$ for the primes $p \in \mathbb{P}$. We will call an equivalence class of absolute values on a number field a *place*. We will identify the set of places for $\mathbb{Q}$ with $\mathbb{P} \cup \{\infty\}$. The additive groups of the corresponding completions $\mathbb{R}$ and $\mathbb{Q}_p$ for $p \in \mathbb{P}$ are all examples of locally compact abelian groups. In this section, we will make the general statements about Harmonic Analysis explicit in the case

[9] The vector space $L^2(G)$ with this scalar product is a Hilbert space, i.e., it is a vector space with a scalar product such that it is complete with respect to the topology induced by the scalar product.

[10] D. Ramakrishnan and R. J. Valenza. *Fourier analysis on number fields*, volume 186 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1999. ISBN 0-387-98436-4. DOI: 10.1007/978-1-4757-3085-2. URL https://doi.org/10.1007/978-1-4757-3085-2

$(\mathbb{R}, +)$ and $(\mathbb{Q}_p, +)$. In particular, we will see that $\mathbb{R}$ and $\mathbb{Q}_p$ are examples of *self-dual* (or *auto-dual*) locally compact abelian groups, i.e., they satisfy $G \cong G^\vee$. In general, there is no canonical choice of such an *auto-duality isomorphism*. Nevertheless, we will fix a rather natural choice in the cases $\mathbb{R}$ and $\mathbb{Q}_p$.

4.2.1 *Fourier Theory on* $\mathbb{R}$

In a first step, we want to prove the *auto-duality* of $\mathbb{R}$, i.e., we want to show $\mathbb{R} \cong \mathbb{R}^\vee$. Let us start with fixing an explicit character in $\mathbb{R}^\vee$.

Definition 4.2.1. Let us define the character $e_\infty \in \mathbb{R}^\vee = \mathrm{Hom}_{cts}(\mathbb{R}, \mathbb{T})$ by

$$e_\infty \colon \mathbb{R} \to \mathbb{T}, \quad x \mapsto e_\infty(x) := \exp(2\pi i x).$$

The next proposition shows that every character of $\mathbb{R}$ is given by scaling e_∞. More precisely:

Proposition 4.2.2. *We have an explicit auto-duality isomorphism*

$$\mathbb{R} \xrightarrow{\sim} \mathbb{R}^\vee = Hom_{cts}(\mathbb{R}, \mathbb{T}), \quad y \mapsto (x \mapsto e_\infty(x \cdot y)).$$

Proof. The map in the statement gives an injective group homomorphism

$$\Psi \colon \mathbb{R} \to \mathbb{R}^\vee.$$

It remains to check that Ψ is surjective and a homeomorphism.
For the surjectivity, we will use the lifting property of coverings from algebraic topology:
Fact: Let us fix $t \in \mathbb{R}$ and write t_0 for its image in $\mathbb{T}$ under the exponential map, i.e., $t_0 := \exp(2\pi i t) \in \mathbb{T}$. For any continuous map $f \colon \mathbb{R} \to \mathbb{T}$ with $f(0) = t_0$, there is a unique lift $\tilde{f} \colon \mathbb{R} \to \mathbb{R}$ with $\tilde{f}(0) = t$ such that the following diagram commutes:

$$\begin{tikzcd}
 & \mathbb{R} \arrow[d, "\exp(2\pi i -)"] \\
\mathbb{R} \arrow[ur, dotted, "\tilde{f}"] \arrow[r, "f"] & \mathbb{T}.
\end{tikzcd}$$

Claim: The map Ψ is surjective.
Proof of the Claim: Let $f \in \mathrm{Hom}_{cts}(\mathbb{R}, \mathbb{T})$. By the lifting property, we find a unique lift $\tilde{f}$ with $\tilde{f}(0) = 0$. For each $\alpha \in \mathbb{R}$, applying the lifting property to both sides of the equation

$$f(\alpha + x) = f(\alpha) f(x).$$

shows $\tilde{f}(\alpha + x) = \tilde{f}(\alpha) + \tilde{f}(x)$ for all $x \in \mathbb{R}$. Thus, the lift $\tilde{f}$ is a continuous homomorphism

$$\mathbb{R} \to \mathbb{R}.$$

But every continuous homomorphism $\tilde{f}\colon \mathbb{R} \to \mathbb{R}$ is uniquely determined[11] by $\tilde{f}(1)$, namely $\tilde{f}(x) = \tilde{f}(1) \cdot x$. Setting $y := \tilde{f}(1)$ proves $\Psi(y) = f$. This shows the surjectivity of Ψ.
A basis of neighbourhoods of $1 \in \mathbb{R}^\vee$ for the compact-open topology on $\mathbb{R}^\vee$ is given by

$$\{f \in \mathbb{R}^\vee \mid f([-1,1]) \subseteq \exp(2\pi i(-\epsilon,\epsilon))\}_{\epsilon>0}.$$

It is straight forward to check that this basis corresponds under the above bijection to the family of subsets

$$\{y \in \mathbb{R} : |y| < \epsilon\}_{\epsilon>0}$$

which is a basis for the topology on $\mathbb{R}$. Thus, Ψ is an isomorphism of locally compact abelian groups. □

The above auto-duality isomorphism $\mathbb{R} \to \mathbb{R}^\vee$ is not canonical. It depends on the choice of a character e_∞ and is characterized by $1 \mapsto e_\infty$. For any other non-trivial character χ, we would similarly get an auto-duality isomorphism $\mathbb{R} \to \mathbb{R}^\vee$ by $1 \mapsto \chi$. Nevertheless, the choice e_∞ is in some sense distinguished. Recall from the last section that for every fixed Haar measure μ on $\mathbb{R}$, there is a unique 'dual' Haar measure $\mu^\vee$ on $\mathbb{R}^\vee$ such that the Fourier transform $L^2(\mathbb{R}) \to L^2(\mathbb{R}^\vee)$ is an isometry. On $\mathbb{R}$, we have a distinguished Haar measure, namely the Lebesgue measure. In principle, we could take the any of the isomorphisms $\mathbb{R} \cong \mathbb{R}^\vee$ to make the Fourier transform $L^2(\mathbb{R}) \to L^2(\mathbb{R}^\vee)$ explicit, but the choice $1 \mapsto e_\infty$ has the advantage that the 'dual' Haar measure of the Lebesgue measure is again the Lebesgue measure[12].

In the following, we will fix the Lebesgue measure μ_∞ as the Haar measure on $\mathbb{R}$ and the auto-duality isomorphism

$$\mathbb{R} \to \mathbb{R}^\vee, \quad y \mapsto (x \mapsto e_\infty(xy)).$$

With this identification, the Theorem 4.1.12 reads in the special case $G = \mathbb{R}$ as follows:

Corollary 4.2.3. *Let us equip $\mathbb{R}$ with the Lebesgue measure μ_∞.*

(a) *For $f \in L^2(\mathbb{R}) \cap L^1(\mathbb{R})$, the* Fourier transform

$$\widehat{f}(y) := \int_{\mathbb{R}} f(x)e_\infty(-xy)d\mu_\infty(x).$$

gives a well-defined map

$$L^2(\mathbb{R}) \cap L^1(\mathbb{R}) \to L^2(\mathbb{R}), \quad f \mapsto \widehat{f},$$

such that $\|f\|_{L^2(\mathbb{R})} = \|\widehat{f}\|_{L^2(\mathbb{R})}$ for all $f \in L^2(\mathbb{R}) \cap L^1(\mathbb{R})$.

[11] $\tilde{f}(1)$ determines $\tilde{f}\colon \mathbb{Q} \to \mathbb{R}$ by $\tilde{f}(n/m) = n\tilde{f}(1)/m$. Since $\mathbb{Q}$ is dense in $\mathbb{R}$, we deduce that $\tilde{f}(1)$ determines $\tilde{f}$.

[12] Actually, this does not uniquely determine e_∞. The choice $e_\infty^{-1}(x) = \exp(-2\pi i x)$ does also identify the dual of the Lebesgue measure with the Lebesgue measure. So it would be equally fine to choose e_∞^{-1}.

(b) *The Fourier transform of* (a) *extends to a well-defined isometry*

$$L^2(\mathbb{R}) \to L^2(\mathbb{R})$$

such that for all $f \in L^2(\mathbb{R})$ *we have*

$$\widehat{\widehat{f}}(x) = f(-x)$$

almost everywhere.

Proof. Follows immediately from Theorem 4.1.12 using

$$\Psi \colon \mathbb{R} \xrightarrow{\sim} \mathbb{R}^\vee, \quad y \mapsto (x \mapsto e_\infty(xy))$$

and $\mu_\infty = \Psi^* \mu_\infty^\vee$. □

In Definition 2.3.1, we have defined a certain class of smooth functions which behaves particularly nice under the Fourier transformation, namely the Schwartz functions $\mathcal{S}(\mathbb{R})$. It is not difficult to check that $\mathcal{S}(\mathbb{R})$ is contained in $L^2(\mathbb{R}) \cap L^1(\mathbb{R})$. The advantage of the space $\mathcal{S}(\mathbb{R})$ is that it is stable under the Fourier transform, i.e., for $f \in \mathcal{S}(\mathbb{R})$ we have $\widehat{f} \in \mathcal{S}(\mathbb{R})$. With this observation, Theorem 2.3.3 turns out to be a special case of Theorem 4.1.12.

4.2.2 Fourier Theory on $\mathbb{Q}_p$

Let us fix a prime p. Similarly to the case of $\mathbb{R}$, we want to prove an auto-duality isomorphism $\mathbb{Q}_p \cong \mathbb{Q}_p^\vee$. As in the case of $\mathbb{R}$, we first want to define a character $e_p \in \mathbb{Q}_p^\vee$.

For the definition of e_p, let us first observe the following isomorphism of groups[13]

$$\mathbb{Q}_p/\mathbb{Z}_p = \bigcup_{n \geq 1} p^{-n}\mathbb{Z}_p/\mathbb{Z}_p \cong \bigcup_{n \geq 1} p^{-n}\mathbb{Z}/\mathbb{Z}.$$

The latter group (with the discrete topology) admits a continuous group homomorphism

$$\mathbb{Q}_p/\mathbb{Z}_p \cong \bigcup_{n \geq 1} p^{-n}\mathbb{Z}/\mathbb{Z} \hookrightarrow \mathbb{R}/\mathbb{Z}.$$

Definition 4.2.4. Let us define the character $e_p \in \mathbb{Q}_p^\vee$ as follows[14]

$$e_p \colon \mathbb{Q}_p/\mathbb{Z}_p \to \mathbb{R}/\mathbb{Z} \xrightarrow{\exp(-2\pi i(\cdot))} \mathbb{T}.$$

Again, all other characters of $\mathbb{Q}_p$ are obtained by scaling e_p:

Proposition 4.2.5. *We have the following isomorphism of locally compact groups*

$$\mathbb{Q}_p \xrightarrow{\sim} \mathbb{Q}_p^\vee = Hom_{cts}(\mathbb{Q}_p, \mathbb{T}), \quad y \mapsto (x \mapsto e_p(x \cdot y)).$$

[13] Probably it is better to write $\varinjlim_n p^{-n}\mathbb{Z}/\mathbb{Z}$ instead of $\bigcup_n p^{-n}\mathbb{Z}/\mathbb{Z}$; on the other hand, it is perfectly fine for us to think about $\bigcup_{n \geq 1} p^{-n}\mathbb{Z}/\mathbb{Z}$ as a subset of $\mathbb{Q}_p/\mathbb{Z}_p$. Observe, that it carries the discrete topology.

[14] If you prefer an explicit description of this map:

$$e_p\left(\sum_{j=-N}^{\infty} a_j p^j\right) := \exp\left(-\sum_{j=-N}^{-1} 2\pi i a_j p^j\right)$$

Proof. See Exercises.

□

As in the case of the real numbers, we see that $\mathbb{Q}_p$ is (non-canonically) isomorphic to its own Pontryagin dual. Again, the choice of a measure allows us to justify the above choice of an auto-duality isomorphism $\mathbb{Q}_p \cong \mathbb{Q}_p^\vee$. In the case of $\mathbb{Q}_p$, let us choose the unique Haar measure μ_p on $\mathbb{Q}_p$ such that $\mu_p(\mathbb{Z}_p) = 1$. We will see in the exercises, that the above choice has the property that $\mu_p^\vee$ is identified with μ_p under the isomorphism

$$\mathbb{Q}_p \to \mathbb{Q}_p^\vee, \quad y \mapsto (x \mapsto e_p(x \cdot y)).$$

In the following let us fix this isomorphism to identify $\mathbb{Q}_p^\vee$ with $\mathbb{Q}_p$.

Corollary 4.2.6. *Let us equip $\mathbb{Q}_p$ with the unique measure μ_p such that $\mu_p(\mathbb{Z}_p) = 1$.*

(a) For $f \in L^2(\mathbb{Q}_p) \cap L^1(\mathbb{Q}_p)$, the Fourier transform

$$\widehat{f}(y) := \int_{\mathbb{Q}_p} f(x) e_p(-xy) d\mu(x).$$

gives a well-defined map

$$L^2(\mathbb{Q}_p) \cap L^1(\mathbb{Q}_p) \to L^2(\mathbb{Q}_p), \quad f \mapsto \widehat{f},$$

such that $\|f\|_{L^2(\mathbb{Q}_p)} = \|\widehat{f}\|_{L^2(\mathbb{Q}_p)}$ for all $f \in L^2(\mathbb{Q}_p) \cap L^1(\mathbb{Q}_p)$.

(b) The Fourier transform of (a) extends to a well-defined isometry

$$L^2(\mathbb{Q}_p) \to L^2(\mathbb{Q}_p)$$

such that for all $f \in L^2(\mathbb{Q}_p)$ we have

$$\widehat{\widehat{f}}(x) = f(-x)$$

almost everywhere.

Proof. Follows immediately from Theorem 4.1.12 using

$$\mathbb{Q}_p \xrightarrow{\sim} \mathbb{Q}_p^\vee, \quad y \mapsto (x \mapsto e_p(xy)).$$

□

In the case of $G = \mathbb{R}$ we have identified a suitable subspace of functions in $L^2(\mathbb{R}) \cap L^1(\mathbb{R})$ which is stable under the Fourier transform, namely the Schwartz functions. Let us introduce a similar class of functions in the non-Archimedean world:

Definition 4.2.7. A Schwartz-Bruhat function on $\mathbb{Q}_p$ is a function $f \colon \mathbb{Q}_p \to \mathbb{C}$ which is locally constant with compact support. Let us denote by $\mathcal{S}(\mathbb{Q}_p)$ the space of all Schwartz-Bruhat functions on $\mathbb{Q}_p$.

A Schwartz-Bruhat function f is obviously contained in $L^2(\mathbb{Q}_p) \cap L^1(\mathbb{Q}_p)$, so its Fourier transform

$$\widehat{f}(y) := \int_{\mathbb{Q}_p} f(x) e_p(-xy) d\mu(x)$$

exists. Let us check that the Fourier transform of a Schwartz-Bruhat function is again a Schwartz-Bruhat function:

Proposition 4.2.8. *For $f \in \mathcal{S}(\mathbb{Q}_p)$, we have $\widehat{f} \in \mathcal{S}(\mathbb{Q}_p)$ and*

$$\widehat{\widehat{f}}(x) = f(-x).$$

Proof. The open subsets of $\mathbb{Q}_p$ are exactly the subsets of the form $a + p^k\mathbb{Z}_p$ for $a \in \mathbb{Q}_p$ and $k \in \mathbb{Z}$, so every Schwartz-Bruhat function is a finite linear combination of characteristic functions $\mathbb{1}_{a+p^k\mathbb{Z}_p}$.
A straightforward computation shows the following formulas for general Schwartz-Bruhat functions:
Claim 1: For $f \in \mathcal{S}(\mathbb{Q}_p)$ we have:

(a) For $a \in \mathbb{Q}_p$ and $g(x) := e_p(ax)f(x)$ we have $\widehat{g}(x) = \widehat{f}(x-a)$.

(b) For $a \in \mathbb{Q}_p$ and $g(x) := f(x-a)$ we have $\widehat{g}(x) = \widehat{f}(x)e_p(-ax)$.

(c) For $\lambda \in \mathbb{Q}_p^\times$ and $g(x) := f(\lambda x)$ we have $\widehat{g}(x) = \frac{1}{|\lambda|_p}\widehat{f}(\frac{x}{\lambda})$.

Proof of Claim 1: These are straightforward computations. For example, let us prove (a):

$$\widehat{g}(x) = \int_{\mathbb{Q}_p} g(y)e_p(-xy)d\mu_p(y) = \int_{\mathbb{Q}_p} f(y)e_p(ay)e_p(-xy)d\mu_p(y) = \widehat{f}(x-a).$$

The properties (b) and (c) follow similarly.

Since $x \mapsto e_p(ax)$ is itself a locally constant function, it suffices by Claim 1 to prove that the Fourier transform $\widehat{\mathbb{1}}_{\mathbb{Z}_p}$ is again a Schwartz-Bruhat function. Indeed, we have:
Claim 2: The function $\mathbb{1}_{\mathbb{Z}_p}$ is its own Fourier transform.
Proof of Claim 2: We compute

$$\widehat{\mathbb{1}}_{\mathbb{Z}_p}(x) = \int_{\mathbb{Z}_p} e_p(-xy)d\mu_p(y)$$

On the other hand, the function $y \mapsto e_p(-xy)$ is a character on $\mathbb{Z}_p$ and it is trivial if and only if $x \in \mathbb{Z}_p$. We deduce from Lemma 4.1.9 that

$$\widehat{\mathbb{1}}_{\mathbb{Z}_p}(x) = \int_{\mathbb{Z}_p} e_p(-xy)d\mu_p(y) = \mathbb{1}_{\mathbb{Z}_p}(x).$$

This finishes the proof that $\widehat{f} \in \mathcal{S}(\mathbb{Q}_p)$ for $f \in \mathcal{S}(\mathbb{Q}_p)$. The Fourier inversion formula follows from the general Theorem 4.1.12 on Fourier theory on locally compact abelian groups. □

4.3 Adeles and Ideles

In the last section, we have formulated Fourier theory for locally compact abelian groups. We would like to do Fourier theory at all completions at the same time. A first naive approach would be to consider the product

$$\mathbb{R} \times \prod_{p \text{ prime}} \mathbb{Q}_p.$$

Unfortunately, this turns out to be a topological group which is not locally compact. Indeed, we have:

Lemma 4.3.1. *Let $(X_i)_{i\in I}$ a family of locally compact topological spaces then $\prod_{i\in I} X_i$ is locally compact if and only if X_i is compact for almost all $i \in I$.*

Proof. See Exercise 1 on Sheet 13. □

To fix this, we introduce the following notion of a restricted product:

Definition 4.3.2. Let $(X_i)_{i\in I}$ be a family of topological spaces and let $(U_i)_{i\in I}$ be a family of open subsets $U_i \subseteq X_i$. The *restricted product* $\prod_{i\in I}(X_i, U_i)$ is the topological space

$$\prod_{i\in I}(X_i, U_i) := \{(x_i)_{i\in I} \in \prod_{i\in I} X_i \mid x_i \in U_i \text{ for almost all } i \in I\}$$

equipped with the topology given by the basis of open sets

$$\mathcal{B} := \{\prod_i V_i \mid V_i \subseteq X_i \text{ open for all } i \text{ and } V_i = U_i \text{ for almost all } i \in I\}.$$

Example 4.3.3. Let $(X_i)_{i\in I}$ be a family of topological spaces and let $(U_i)_{i\in I}$ be a family of open subsets $U_i \subseteq X_i$.

(a) If I is finite then $\prod_{i\in I}(X_i, U_i) = \prod_{i\in I} X_i$.

(b) If $U_i = X_i$ for almost all $i \in I$ then $\prod_{i\in I}(X_i, U_i) = \prod_{i\in I} X_i$.

(c) If X_i is a sequence of vector spaces equipped with the discrete topology and $U_i := \{0\}$ then $\prod_{i\in I}(X_i, U_i) = \bigoplus_{i\in I} X_i$.

Our main application of this construction will be in the following situation. Let $(G_i)_{i\in I}$ be countable family of locally compact abelian groups and $(H_i)_{i\in I}$ a family of compact open subgroups $H_i \subseteq G_i$. Of course, we have that

$$G := \prod_{i\in I}(G_i, H_i)$$

is again a locally compact abelian group. For locally compactness, let us observe that for a given $x \in \prod_{i\in I}(G_i, H_i)$ the open neighbourhood $x + \prod_{i\in I} H_i$ of x is compact. It is easily checked that the multiplication and inversion are continuous. For doing Fourier theory on the restricted product, we have to choose a Haar measure on G. The following result shows that the choices of Haar measure μ_i on G_i induce a unique Haar measure on G:

Proposition 4.3.4. *Let $(G_i)_{i\in I}$ a countable family of locally compact abelian groups with compact open subgroups $(H_i)_{i\in I}$. Furthermore for $i \in I$, let μ_i be the unique Haar measure on G_i such that $\mu_i(H_i) = 1$. Then there is a unique Haar measure μ on $G := \prod_{i\in I}(G_i, H_i)$ such that:*

(a) For every finite set $S \subseteq I$ the restriction of μ to $G_S := \prod_{j\in S} G_j \times \prod_{i\in I\setminus S} H_i \subseteq G$ is the product measure.

(b) For any family of integrable continuous functions $(f_i)_{i\in I}$ such that $f_i|_{H_i} = 1$ for almost all $i \in I$ the product

$$f(g) := \prod_{i\in I} f_i(g)$$

for $g \in G$ is well-defined and defines a continuous function on G. Moreover, we have

$$\int_G f(g)d\mu(g) = \prod_{i\in I}\int_{G_i} f_i(g_i)d\mu_i(g_i)$$

and the function f is in $L^1(G)$ if and only if the product on the right hand side has a finite value.

Proof. We sketch the proof:
(a): For every finite set $S \subseteq I$ there is a unique product measure μ_S on $G_S := \prod_{j\in S} G_i \times \prod_{i\in I\setminus S} H_i$. It is easily checked that μ_S is a Haar measure; indeed, by measure theory the product measure μ_S is a Radon measure (i.e., inner and outer regular and $\mu_S(K) < \infty$ for compact subsets). Furthermore, it is translation invariant since all μ_i are translation invariant. For $S \subseteq T$ the measure μ_T restricts to μ_S under $G_S \subseteq G_T$. Thus, we get a unique measure μ on $G = \varinjlim_S G_S$ restricting to μ_S for every finite S. The measure μ is a Haar measure since all the measures μ_S are Haar measures on G_S.
(b) The product

$$f(g) = \prod_{i\in I} f_i(g_i)$$

is a finite product since $g_i \in H_i$ and $f_i|_{H_i} = 1$ for almost all $i \in I$. The continuity follows from the continuity of f_i because a base of G can be chosen by subsets of the form $\prod_{j\in T} U_j \times \prod_{i\in I\setminus T} H_i$, where T is a finite set containing all the $i \in I$ with $f_i|_{H_i} \neq 1$ and U_j is open in G_j; so f can be computed locally by a finite product of continuous functions. By the construction of μ, we have that f is integrable if and only if

$$\lim_S \int_{G_S} f(g_S)d\mu_S(g_S) < \infty, \tag{4.1}$$

where the limit is taken over larger and larger S. Since μ_S is the product measure on G_S, we have

$$\int_{G_S} f(g_S)d\mu_S(g_S) = \prod_{j\in S}\int_{G_j} f(g_j)d\mu_j(_j) \cdot \prod_{i\in I\setminus S}\underbrace{\int_{H_i} f(h_i)d\mu_i(h_i)}_{=1 \text{ for almost all } i}.$$

Without loss of generality, we may assume that all the S in the limit contain the indices $i \in I$ with $f_i|_{H_i} \neq 1$. We get

$$\lim_S \int_{G_S} f(g_S) d\mu_S(g_S) = \lim_S \prod_{j \in S} \int_{G_j} f(g_j) d\mu_j(g_j) = \prod_{i \in I} \int_{G_j} f(g_j) d\mu_j(g_j)$$

and by (4.1) the product on the right hand side has finite value if and only if f is integrable. □

Let us now define the ring of adeles of the field $\mathbb{Q}$:

Definition 4.3.5. Let $\mathbb{Q}$ be a number field. The ring of finite adeles is the topological ring given by the restricted product

$$\mathbb{A}_{\mathbb{Q},\mathrm{fin}} := \prod_{p \text{ prime}} (\mathbb{Q}_p, \mathbb{Z}_p)$$

where p runs over all primes. The ring of adeles is the topological ring

$$\mathbb{A}_{\mathbb{Q}} := \mathbb{R} \times \prod_{p \text{ prime}} (\mathbb{Q}_p, \mathbb{Z}_p),$$

where $\mathbb{R}$ carries the usual topology. The ring of adeles can be defined for more general number fields. In this lecture, we will mainly restrict our attention to the case $K = \mathbb{Q}$. Hence, we will often drop the subscript $\mathbb{Q}$ from the notation and simply write $\mathbb{A}$ (respectively $\mathbb{A}_{\mathrm{fin}}$) for the ring of (finite) adeles.

Note, that the ring of finite adeles is a locally compact topological ring since it is the product of the two locally compact topological rings $\mathbb{A}_{\mathrm{fin}}$ and $\mathbb{R}$. Let us now study the structure of the adeles more carefully. Since ever rational number has negative valuation at only finitely many primes, we get a well-defined injective ring homomorphism by the diagonal embedding

$$\mathbb{Q} \hookrightarrow \mathbb{A}, \quad x \mapsto (x)_{\nu \in \mathbb{P} \cup \{\infty\}}$$

and similarly for $\mathbb{A}_{\mathrm{fin}}$

$$\mathbb{Q} \hookrightarrow \mathbb{A}_{\mathrm{fin}}, \quad x \mapsto (x)_{\nu \in \mathbb{P}}.$$

Theorem 4.3.6. *We have the following properties for the (finite) adeles:*

(a) By the diagonal embedding, $\mathbb{Q} \subseteq \mathbb{A}$ is a discrete subgroup.

(b) The quotient $\mathbb{A}/\mathbb{Q}$ is compact.

(c) We have $\mathbb{A}_{fin} \cong \mathbb{Q} \otimes_{\mathbb{Z}} \widehat{\mathbb{Z}}$ and $\mathbb{A} \cong \mathbb{Q} \otimes_{\mathbb{Z}} \left(\mathbb{R} \times \widehat{\mathbb{Z}}\right)$.

(d) $\mathbb{Q}$ is dense in $\mathbb{A}_{fin}$.

Proof. (a) The open subset $U = (-\frac{1}{2}, \frac{1}{2}) \times \prod_{p\in\mathbb{P}} \mathbb{Z}_p$ satisfies $U \cap \mathbb{Q} = \{0\}$: Indeed, let $x \in U \cap \mathbb{Q}$. Then $|x|_p \leq 1$ for all $p \in \mathbb{P}$, hence $x \in \mathbb{Z}$. On the other hand, we have $|x|_\infty < 1/2$ and the only integer with this property is $x = 0$. This proves that $\{0\} \subseteq \mathbb{Q}$ is open and hence $\mathbb{Q}$ inherts the discrete topology via the inclusion $\mathbb{Q} \subseteq \mathbb{A}$.
(b) For the compactness of $\mathbb{A}/\mathbb{Q}$ it suffices to prove that $W = [0,1) \times \prod_p \mathbb{Z}_p$ is a set of representatives for $\mathbb{A}/\mathbb{Q}$. Indeed, this implies that $\mathbb{A}/\mathbb{Q}$ is compact as the image of the compact set $[0,1] \times \prod_p \mathbb{Z}_p$ under the continuous map $\mathbb{A} \to \mathbb{A}/\mathbb{Q}$. So it suffices to prove the following:
Claim 1: Every $x = (x_v)_v \in \mathbb{A}$ can be written uniquely as $x = q + w$ with $q \in \mathbb{Q}$ and $w \in W$.
Proof of Claim 1: For $x = (x_v)_v \in \mathbb{A}$, there is a finite set of primes S such that for all $p \in \mathbb{P} \setminus S$ we have $x_p \in \mathbb{Z}_p$. For $p \in S$, let us write

$$x_p = \sum_{j\geq -N}^{\infty} a_j p^j.$$

Then $r_p := \sum_{j\geq -N}^{-1} a_j p^j \in \mathbb{Q}$ and $x_p - r_p \in \mathbb{Z}_p$. For a second prime number $l \neq p$, we have

$$|r_p|_l \leq \max_{-N\leq j\leq -1} |a_j p^j|_l \leq 1.$$

So subtracting r_p from x does not destroy the integrality at the other places. Thus, with $r := \sum_{p\in S} r_p \in \mathbb{Q}$, we get $x - r \in \mathbb{R} \times \prod_{p\in\mathbb{P}} \mathbb{Z}_p$. Subtracting the integer $z := \lfloor x_\infty - r \rfloor$ from $x - r$ gives

$$w := x - r - z \in W.$$

So we obtain $x = w + q$ with $q := r + z \in \mathbb{Q}$ and $w \in W$. The uniqueness of the decomposition follows from $W \cap \mathbb{Q} = \{0\}$.
(c) It suffices to prove $\mathbb{A}_{\text{fin}} \cong \mathbb{Q} \otimes_{\mathbb{Z}} \widehat{\mathbb{Z}}$. The inclusion $\widehat{\mathbb{Z}} \subseteq \mathbb{A}_{\text{fin}}$ induces an injection

$$\mathbb{Q} \otimes_{\mathbb{Z}} \widehat{\mathbb{Z}} \hookrightarrow \mathbb{A}_{\text{fin}}.$$

This map is surjective since for any $x = (x_p)_p \in \mathbb{A}_{\text{fin}}$ there is a sequence $(n_p)_{p\in\mathbb{P}}$ of non-negative integers such that $n_p = 0$ for almost all $p \in \mathbb{P}$ and $z_p := x_p \cdot p^{n_p} \in \mathbb{Z}_p$. Then $z := (z_p)_p \in \widehat{\mathbb{Z}} = \prod_{p\in\mathbb{P}} \mathbb{Z}_p$ and for $n = \prod_p p^{n_p}$ we get

$$\frac{1}{n} \otimes z \mapsto x$$

under $\mathbb{Q} \otimes_{\mathbb{Z}} \widehat{\mathbb{Z}} \hookrightarrow \mathbb{A}_{\text{fin}}$.
(d) By (c) it suffices to prove that $\mathbb{Z}$ is dense in $\widehat{\mathbb{Z}}$. But any open subset in $\widehat{\mathbb{Z}}$ is of the form $z + N\widehat{\mathbb{Z}}$ for some $z \in \widehat{\mathbb{Z}}$ and $N \in \mathbb{N}$. We have to prove that there is a $n \in \mathbb{Z}$ with $n \in z + N\widehat{\mathbb{Z}}$. For this, we can take any $n \in \mathbb{Z}$ mapping to $z + N\widehat{\mathbb{Z}}$ under

$$\mathbb{Z} \twoheadrightarrow \mathbb{Z}/N\mathbb{Z} = \widehat{\mathbb{Z}}/N\widehat{\mathbb{Z}}.$$

□

Finally, we define the group of ideles as the units in the ring of adeles:

Definition 4.3.7. The topological group of *ideles* is the group of units $\mathbb{I} := \mathbb{A}^\times$ with the topology induced by the subspace topology[15] of the embedding

$$\mathbb{I} = \mathbb{A}^\times \to \mathbb{A} \times \mathbb{A}, \quad x \mapsto (x, x^{-1}).$$

We have $\mathbb{A}^\times = \mathbb{R}^\times \times \mathbb{A}_{\text{fin}}^\times$ and call $\mathbb{I}_{\text{fin}} := \mathbb{A}_{\text{fin}}^\times$ the group of finite ideles.

Observe that every element of $\mathbb{R}^\times \times \prod_{p \in \mathbb{P}} (\mathbb{Q}_p^\times, \mathbb{Z}_p^\times) \subseteq \mathbb{A}$ is a unit in the ring of adeles, so we get an inclusion

$$\mathbb{R}^\times \times \prod_{p \in \mathbb{P}} (\mathbb{Q}_p^\times, \mathbb{Z}_p^\times) \subseteq \mathbb{I}.$$

Lemma 4.3.8. *The inclusion*

$$\mathbb{R}^\times \times \prod_{p \in \mathbb{P}} (\mathbb{Q}_p^\times, \mathbb{Z}_p^\times) \subseteq \mathbb{I}$$

is an equality and induces an isomorphism of topological groups if we equip the left hand side with the restricted product topology. In particular, $\mathbb{I}$ is a locally compact abelian group.

Proof. Exercise 2 on Sheet 13. □

Let us now define the absolute value on the group of ideles. In the following, we equip $\mathbb{Q}_p$ with the absolute value $|\cdot|_p$ on $\mathbb{Q}_p$ which is normalized by $|p|_p = p^{-1}$ and $\mathbb{R}$ with the usual absolute value $|\cdot|_\infty$.

Definition 4.3.9. The absolute value of an *idele* $x \in \mathbb{I}$ is defined by

$$|x| := \prod_{v \in \mathbb{P} \cup \{\infty\}} |x_v|_v.$$

This product is well-defined as $|x_v|_v \neq 1$ for only finitely many factors. We define

$$\mathbb{I}_1 := \{x \in \mathbb{I} \mid |x| = 1\}.$$

By the diagonal embedding, we get an embedding

$$\mathbb{Q}^\times \hookrightarrow \mathbb{I}.$$

Lemma 4.3.10. *For $x \in \mathbb{Q}^\times$, we have*

$$|x| = \prod_{v \in \mathbb{P} \cup \infty} |x|_v = 1.$$

In particular, $\mathbb{Q}^\times \subseteq \mathbb{I}_1$ under the diagonal embedding $\mathbb{Q}^\times \to \mathbb{I}, x \mapsto (x)_v$.

[15] Note that this topology is not the same as the topology induced by the embedding $\mathbb{I} \subseteq \mathbb{A}$.

Proof. This is a straightforward computation if $x = \pm \prod_{p \in \mathbb{P}} p^{\alpha_p}$ with $\alpha_p \in \mathbb{Z}$ then $|x|_p = p^{-\alpha_p}$ and $|x|_\infty = \prod_{p \in \mathbb{P}} p^{\alpha_p}$. So taking the product over all places gives 1. □

Theorem 4.3.11. *We have the following properties for the group of ideles.*

(a) By the diagonal embedding, $\mathbb{Q}^\times \subseteq \mathbb{I}$ is a discrete subgroup.

(b) The quotient $\mathbb{I}_1/\mathbb{Q}^\times$ is compact.

(c) We have $\mathbb{I}_1/\mathbb{Q}^\times \cong \widehat{\mathbb{Z}}^\times$.

(d) The absolute value induces an isomorphism of topological groups $\mathbb{I}/\mathbb{Q}^\times \cong \mathbb{R}_{>0} \times \widehat{\mathbb{Z}}^\times$.

Proof. (a) The open subgroup $U = (\frac{1}{2}, \frac{3}{2}) \times \prod_{p \in \mathbb{P}} \mathbb{Z}_p^\times$ satisfies $U \cap \mathbb{Q}^\times = \{1\}$: Indeed, for $x \in U \cap \mathbb{Q}^\times$ we have at all primes $|x|_p = 1$ so $x \in \{\pm 1\}$. The condition x_∞ rules out the possibility $x = -1$. Thus, $\{1\}$ is an open neighbourhood of $1 \in \mathbb{Q}^\times$ which proves (a).
(c) Let us first observe that $x = (x_v)_v \in \mathbb{I}_1$ satisfies $x_\infty \in \mathbb{Q}$. Indeed, we have

$$|x|_\infty = \prod_p |x_p|^{-1}$$

and the right hand side is a finite product of rational numbers. So, the map

$$\mathbb{I}_1 \to \widehat{\mathbb{Z}}^\times, \quad (x_v)_v \mapsto (x_p/x_\infty)_p$$

is a well-defined continuous map and contains $\mathbb{Q}^\times$ in the kernel, so it induces

$$\mathbb{I}_1/\mathbb{Q}^\times \to \widehat{\mathbb{Z}}^\times.$$

The inverse to this map is given by

$$\widehat{\mathbb{Z}}^\times = \prod_p \mathbb{Z}_p^\times \to \mathbb{I}_1/\mathbb{Q}^\times, \quad z = (z_p)_p \mapsto (1, z)\mathbb{Q}^\times.$$

(b) Follows from (c) and the compactness of the pro-finite group $\widehat{\mathbb{Z}}^\times$.
(d) We have a short exact sequence

$$1 \to \mathbb{I}_1/\mathbb{Q}^\times \to \mathbb{I}/\mathbb{Q}^\times \to \mathbb{R}_{>0} \to 1$$

which is split by $\mathbb{R}_{>0} \to \mathbb{I}$, $r \mapsto (r, 1)\mathbb{Q}^\times \in (\mathbb{R}^\times \times \mathbb{A}_{\text{fin}}^\times)/\mathbb{Q}^\times$. □

By combining the previous result with the Kronecker-Weber Theorem gives:

Corollary 4.3.12. *We have an isomorphism*

$$\mathbb{I}_1/\mathbb{Q}^\times \xrightarrow{\sim} Gal(\mathbb{Q}^{ab}/\mathbb{Q}).$$

Furthermore, we can identify the Pontryagin dual of $\mathbb{I}_1/\mathbb{Q}$ with the group of all primitive Dirichlet characters:

$$(\mathbb{I}_1/\mathbb{Q})^\vee = \mathrm{Hom}_{cts}(\mathbb{I}_1/\mathbb{Q}, \mathbb{C}^\times) \cong \{\text{primitive Dirichlet characters}\}.$$

These two points of view admit a natural generalization to other number fields different from $\mathbb{Q}$ and provide the natural framework for class field theory for general number fields.

Outlook

The ring of adeles can be defined for an arbitrary of number field K:

Definition 4.3.13. Let K be a number field. The ring of finite adeles of K is given by the following restricted product

$$\mathbb{A}_{K,\mathrm{fin}} := \prod_{\mathfrak{p}\ \text{prime}} (K_\mathfrak{p}, \mathcal{O}_{K_\mathfrak{p}})$$

where $\mathfrak{p}$ runs over all non-zero prime ideals and $K_\mathfrak{p}$ is the completion of K at $\mathfrak{p}$ and $\mathcal{O}_{K_\mathfrak{p}}$ is the discrete valuation ring of $K_\mathfrak{p}$. The ring of adeles is the topological ring

$$\mathbb{A}_K := \mathbb{R}^r \times \mathbb{C}^s \times \prod_{\mathfrak{p}\subseteq\ \text{prime}} (K_\mathfrak{p}, \mathcal{O}_{K_\mathfrak{p}}),$$

where $\mathbb{R}$ and $\mathbb{C}$ carry their usual topology, r denotes the number of real embeddings and s denotes the number of pairs of complex embeddings. The group of *ideles* $\mathbb{I}_K := \mathbb{A}_K^\times$ is equipped with the topology induced by

$$\mathbb{A}_K^\times \subseteq \mathbb{A}_K \times \mathbb{A}_K, x \mapsto (x, x^{-1}).$$

The ideles provide a natural framework for class field theory for general number fields. The diagonal embedding $K^\times \to \mathbb{I}_K$ identifies $K^\times$ as a discrete subgroup of $\mathbb{I}_K$ and the quotient

$$C_K := \mathbb{I}_K/K^\times$$

is called the *idele class group* of K. For a finite abelian extension L/K one can use the local norms between the completions of L and K to define a norm map

$$\mathrm{Nm}_{L/K} \colon C_L \to C_K$$

between the idele class groups of L and K. Class field theory provides for every number field K a map called *Artin reciprocity map*

$$\Phi_K \colon C_K \to \mathrm{Gal}(K^{ab}/K)$$

with the following properties:

Theorem 4.3.14 (Global class field theory)**.** *Let us fix an algebraic closure $\overline{K}$ of K. The Artin reciprocity map satisfies the following properties:*

(a) For every finite abelian extension L/K, Φ_K induces an isomorphism

$$\Phi_{L/K}\colon C_K/\operatorname{Nm}_{L/K} C_L \xrightarrow{\sim} Gal(L/K).$$

(b) The open subgroups of C_K are in bijection with the norm groups $\operatorname{Nm}_{L/K} L$ where L runs through all finite abelian extensions of K. In particular, we have

$$\varprojlim_U C_K/U \xrightarrow{\sim} Gal(K^{ab}/K),$$

where U runs through all open subgroups of C_K.

Note that C_K can be defined purely in terms of K. So global class field theory says in particular that the Galois group of the maximal abelian extension can be described purely in terms of data of K. In the case $K = \mathbb{Q}$, we have already seen that $\mathbb{I}/\mathbb{Q}^\times \cong \mathbb{R}_{>0} \times \widehat{\mathbb{Z}}^\times$. The Artin reciprocity map is in this case exactly the map

$$C_{\mathbb{Q}} = \mathbb{I}/\mathbb{Q}^\times \twoheadrightarrow \widehat{\mathbb{Z}}^\times \cong \operatorname{Gal}(\mathbb{Q}^{ab}/\mathbb{Q})$$

induced by the Kronecker-Weber theorem.

4.4 Fourier theory on the adeles

Last week, we have studied Fourier theory for general locally compact abelian groups and applied the theory to the additive groups of the local fields $\mathbb{R}$ and $\mathbb{Q}_p$. In this section, we want to do Fourier theory at all these completions at the same time. The locally compact ring of adeles provides a natural framework for this.

In the following, let us write $\mathbb{Q}_\nu$ for the completion of $\mathbb{Q}$ at $\nu \in \mathbb{P} \cup \{\infty\}$ with respect to $|\cdot|_\nu$, where $|\cdot|_p$ and $|\cdot|_\infty$ are the usual absolute values on $\mathbb{Q}$. In particular, we have $\mathbb{Q}_\infty = \mathbb{R}$ for $\nu = \infty$.

Let us briefly recall the discussion about the local Fourier analysis on $\mathbb{Q}_\nu$. For every $\nu \in \mathbb{P} \cup \{\infty\}$, we have fixed a Haar measure on μ_ν normalized by $\mu_\nu([0,1]) = 1$ for $\nu = \infty$ (i.e., the Lebesgue measure) and by $\mu_p(\mathbb{Z}_p) = 1$ for $\nu = p$. Furthermore, we have seen that the Pontryagin dual of $(\mathbb{Q}_\nu, +)$ is isomorphic to $(\mathbb{Q}_\nu, +)$. We have fixed one such isomorphism for every $\nu \in \mathbb{P} \cup \{\infty\}$, namely: For $\nu = p \in \mathbb{P}$, we fixed

$$\mathbb{Q}_p \xrightarrow{\sim} \mathbb{Q}_p^\vee, \quad y \mapsto (x \mapsto e_p(x \cdot y))$$

and for $\nu = \infty$ we fixed

$$\mathbb{R} \xrightarrow{\sim} \mathbb{R}^\vee, \quad y \mapsto (x \mapsto e_\infty(x \cdot y) = \exp(2\pi i x \cdot y)).$$

These choices have the advantage that our normalized measure μ_v is identified with the Fourier dual $\mu_v^\vee$ measure under these isomorphisms.

By patching the local characters e_v together, we obtain:

Lemma 4.4.1. *The map*

$$e\colon \mathbb{A} \to \mathbb{T}, \quad x = (x_v)_v \mapsto e(x) := \prod_v e_v(x_v)$$

gives a well-defined character of the locally compact group $(\mathbb{A}, +)$.

Proof. This follows from Exercise 4 (a) on Sheet 13. □

Similarly, as in the local case we can use this character to prove the following auto-duality isomorphism for the locally compact group $(\mathbb{A}, +)$.

Proposition 4.4.2. *We have an isomorphism of locally compact abelian groups:*

$$\mathbb{A} \to \mathbb{A}^\vee, \quad y \mapsto (x \mapsto e(x \cdot y)).$$

Proof. The map $y \mapsto (x \mapsto e(x \cdot y))$ gives a continuous and injective group homomorphism. Exercise 4 on Sheet 13 shows that ever character $\chi \in \mathbb{A}^\vee$ can be written as a product $\chi = \chi_\infty \prod_{p \in \mathbb{P}} \chi_p$ with $\chi_\infty \in \mathbb{R}^\vee$, $\chi_p \in \mathbb{Q}_p^\vee$ and $\chi_p|_{\mathbb{Z}_p} = 1$ for almost all $p \in \mathbb{P}$. By Proposition 4.2.2 and Proposition 4.2.5, there exists a unique $y_v \in \mathbb{Q}_v$ for every v such that

$$\chi_v(-) = e_v(- \cdot y_v).$$

Furthermore, $y_p \in \mathbb{Z}_p$ for almost all $p \in \mathbb{P}$ since $\chi_p|_{\mathbb{Z}_p} = 1$. This shows $y = (y_v)_v \in \mathbb{A}$ and proves the surjectivity. It is easily checked that this map is also open and the result follows. □

From now on, let us fix the Haar measure μ on $\mathbb{A}$ determined by the local Haar measures μ_v, see Proposition 4.3.4. From the general Fourier theory of locally compact abelian groups, we get:

Corollary 4.4.3. *Let us equip* $(\mathbb{A}, +)$ *with the Haar measure* μ *determined by the local Haar measures* μ_v *for* $v \in \mathbb{P} \cup \{\infty\}$.

(a) For $f \in L^2(\mathbb{A}) \cap L^1(\mathbb{A})$, *the Fourier transform*

$$\widehat{f}(y) := \int_{\mathbb{A}} f(x) e(-xy) d\mu(x).$$

gives a well-defined map

$$L^2(\mathbb{A}) \cap L^1(\mathbb{A}) \to L^2(\mathbb{A}), \quad f \mapsto \widehat{f},$$

such that $\|f\|_{L^2(\mathbb{A})} = \|\widehat{f}\|_{L^2(\mathbb{A})}$ *for all* $f \in L^2(\mathbb{A}) \cap L^1(\mathbb{A})$.

(b) *The Fourier transform of* (a) *extends to a well-defined isometry*

$$L^2(\mathbb{A}) \to L^2(\mathbb{A})$$

such that for all $f \in L^2(\mathbb{A})$ *we have*

$$\widehat{\widehat{f}}(x) = f(-x)$$

almost everywhere.

Proof. Follows immediately from Theorem 4.1.12 using

$$\mathbb{A} \xrightarrow{\sim} \mathbb{A}^\vee, \quad y \mapsto (x \mapsto e(xy)).$$

□

As in the local cases, there will be a class of Schwartz functions in $L^1(\mathbb{A}) \cap L^2(\mathbb{A})$ which is stable under the Fourier transform and satisfies the Fourier inversion formula.

Definition 4.4.4. A *simple Schwartz-Bruhat function* on $\mathbb{A}$ is a function $f\colon \mathbb{A} \to \mathbb{C}$ of the form

$$f = \prod_{\nu \in \mathbb{P} \cap \{\infty\}} f_\nu$$

where $f_\nu \in \mathcal{S}(\mathbb{Q}_\nu)$ for $\nu \in \mathbb{P} \cup \{\infty\}$ is a Schwartz(-Bruhat) function on $\mathbb{Q}_\nu$ and $f_p = \mathbb{1}_{\mathbb{Z}_p}$ for almost all $p \in \mathbb{P}$. A *Schwartz-Bruhat function* on $\mathbb{A}$ is a function $f\colon \mathbb{A} \to \mathbb{C}$ which is a finite linear combination of simple Schwartz-Bruhat functions. We will write $\mathcal{S}(\mathbb{A})$ for the space of Schwartz-Bruhat functions on $\mathbb{A}$.

Let us observe the following Lemma:

Lemma 4.4.5. *Every Schwartz-Bruhat function* f *on* $\mathbb{A}$ *is a finite linear combination of functions of the form*

$$f(x) = f_\infty(x_\infty) \cdot \mathbb{1}_{a+N\widehat{\mathbb{Z}}}(x_{fin}) = f_\infty(x_\infty) \cdot \prod_{p \in \mathbb{P}} \mathbb{1}_{a_p + N\mathbb{Z}_p}(x_p),$$

where $f_\infty \in \mathcal{S}(\mathbb{R})$, $a = (a_p)_p \in \mathbb{A}_{fin}$ *and* $N \in \mathbb{Z}$. *Note that* $a_p + N\mathbb{Z}_p = \mathbb{Z}_p$ *for almost all* $p \in \mathbb{P}$.

Proof. Every compact open subset of $\mathbb{A}_{\mathrm{fin}}$ is a finite disjoint union of sets of the form $a + N\widehat{Z}$. This proves that every Schwartz-Bruhat function can be written as a finite linear combinations of functions of the form $f_\infty \cdot \mathbb{1}_{a+N\widehat{\mathbb{Z}}}$. On the other hand, we have $a + N\widehat{\mathbb{Z}} = \prod_{p \in \mathbb{P}} a_p + N\mathbb{Z}_p$ by the Chinese Remainder Theorem. This proves

$$\mathbb{1}_{a+N\widehat{\mathbb{Z}}} = \prod_{p \in \mathbb{P}} \mathbb{1}_{a_p+N\mathbb{Z}_p}.$$

□

Proposition 4.4.6. *For a simple Schwartz-Bruhat function $f = \prod_{v \in \mathbb{P} \cap \{\infty\}} f_v$, we have*

$$\widehat{f} = \prod_{v \in \mathbb{P} \cap \{\infty\}} \widehat{f_v}.$$

In particular, we have $\widehat{f} \in \mathcal{S}(\mathbb{A})$ for $f \in \mathcal{S}(\mathbb{A})$ and the Fourier inversion formula holds, i.e.,

$$\widehat{\widehat{f}}(x) = f(-x).$$

Proof. We have

$$\begin{aligned}\widehat{f}(x) = \int_{\mathbb{A}} f(y) e(-xy) d\mu_{\mathbb{A}}(y) &= \int_{\mathbb{A}} \prod_v f_v(y_v) e_v(-x_v y_v) d\mu_{\mathbb{A}}(y)) \\ &= \prod_v \int_{\mathbb{A}} f_v(y_v) e_v(-x_v y_v) d\mu_{\mathbb{A}}(y)) = \prod_v \widehat{f_v}(x_v).\end{aligned}$$

For each v, $\widehat{f_v}$ is again a Schwartz(-Bruhat) function by Proposition 4.2.8 and Theorem 2.3.3. Furthermore, we have $\widehat{\mathbb{1}_{\mathbb{Z}_p}} = \mathbb{1}_{\mathbb{Z}_p}$ so $\widehat{f_p} = \mathbb{1}_{\mathbb{Z}_p}$ for almost all $p \in \mathbb{P}$. Thus, $\widehat{f}$ is again a simple Schwartz-Bruhat function. Since every Schwartz-Bruhat function on $\mathbb{A}$ is a finite linear combination of simple Schwartz-Bruhat functions, we deduce that the Fourier transform of a Schwartz-Bruhat function is again a Schwartz-Bruhat function. The statement about Fourier inversion follows from the general Theorem on Harmonic analysis. □

4.5 Adelic zeta functions

In the following, we want to imitate the proof of the functional equation for the Riemann zeta function. Recall the following main steps in the proof:

- We used Fourier theory to prove the classical Poisson summation formula.
- We used the Poisson summation formula to prove the functional equation
$$\theta(t) = \frac{1}{\sqrt{t}}\theta(1/t).$$
of the classical theta function θ.
- We wrote the Riemann zeta function as a Mellin transform of the classical theta function
$$\xi(s) = \int_0^\infty \frac{1}{2}(\theta(x) - 1)x^{s/2}\frac{dx}{x}.$$
- We divided the Mellin integral in two parts
$$\xi(s) = \int_0^1 \frac{1}{2}(\theta(x) - 1)x^{s/2}\frac{dx}{x} + \int_1^\infty \frac{1}{2}(\theta(x) - 1)x^{s/2}\frac{dx}{x}.$$

The second integral extends to all $s \in \mathbb{C}$. Using the functional equation, we were able to extend the first integral to all of $\mathbb{C}$ with simple poles at $s = 0$ and $s = 1$ and the functional equation follows from the functional equation of θ.

In the following, we want to follow these steps in the adelic setup. Let us first prove an adelic version of the Poisson summation formula:

Theorem 4.5.1 (The adelic Poisson summation formula)**.** *For every $f \in \mathcal{S}(\mathbb{A})$, we have*

$$\sum_{q\in\mathbb{Q}} f(q) = \sum_{q\in\mathbb{Q}} \widehat{f}(q)$$

and both series are absolutely convergent.

Proof. By Lemma 4.4.5, it is enough to prove the statement for $f = f_\infty \mathbb{1}_{a+N\widehat{\mathbb{Z}}}$. Since $\mathbb{Q}$ is dense in $\mathbb{A}_{\mathrm{fin}}$, we may without loss of generality assume that $a \in \mathbb{Q}$. We get

$$\begin{aligned}\sum_{q\in\mathbb{Q}} f(q) &= \sum_{q\in(a+N\widehat{\mathbb{Z}})\cap\mathbb{Q}} f_\infty(q) \\ &= \sum_{q\in N\widehat{\mathbb{Z}}\cap\mathbb{Q}} f_\infty(q-a) = \sum_{q\in N\mathbb{Z}} f_\infty(q-a) = \sum_{q\in\mathbb{Z}} f_\infty(Nq-a). \qquad (4.2)\end{aligned}$$

Since f_∞ is a Schwartz function on $\mathbb{R}$, the last series converges absolutely. Since this holds for arbitrary Schwartz-Bruhat functions, it also shows the absolute convergence of the series on the right hand side of the statement.

Claim 1: For $a \in \mathbb{Q}$, $\lambda \in \mathbb{Q}^\times$ and $h \in \mathcal{S}(\mathbb{A})$, we define

$$g(x) := h(\lambda(x-a)).$$

Then $\widehat{g}(x) = e(-ax)\widehat{h}(\lambda^{-1}x)$.

Proof: It is enough to prove this for simple Schwartz-Bruhat functions $h = \prod_\nu h_\nu$. For such functions, we have $\widehat{g} = \prod_\nu \widehat{g}_\nu$ and the claim follows from the local computation

$$\widehat{g}_\nu(x_\nu) = e_\nu(-ax_\nu)|N|_\nu \widehat{f}_\nu(\lambda^{-1}x_\nu),$$

together with the product formula $\prod_\nu |a|_\nu = 1$. For the local computation, see also Claim 1 in the proof of Proposition 4.2.8.

Using Claim 1, we obtain:

Claim 2: Let $\lambda \in \mathbb{Q}^\times$ and $a \in \mathbb{Q}$. If the Poisson summation formula holds for $h \in \mathcal{S}(\mathbb{A})$ then it also holds for $g \in \mathcal{S}(\mathbb{A})$ with $g(x) := h(\lambda(x-a))$.

Proof of Claim 2: Using claim 1, we compute:

$$\sum_{q\in\mathbb{Q}} \widehat{g}(q) = \sum_{q\in\mathbb{Q}} \underbrace{e(-aq)}_{=1} \widehat{h}(\lambda^{-1}q) = \sum_{q\in\mathbb{Q}} \widehat{h}(q)$$
$$= \sum_{q\in\mathbb{Q}} h(q) = \sum_{q\in\mathbb{Q}} h(\lambda(q-a)) = \sum_{q\in\mathbb{Q}} g(q).$$

Now, we observe that $\mathbb{1}_{a+N\widehat{\mathbb{Z}}}(x) = \mathbb{1}_{\widehat{\mathbb{Z}}}(\frac{1}{N}(x-a))$. By Claim 2, we can thus reduce the proof of the general Poisson summation formula to the case $f = f_\infty \mathbb{1}_{\widehat{\mathbb{Z}}}$. In this case, we have $\widehat{f} = \widehat{f}_\infty \mathbb{1}_{\widehat{\mathbb{Z}}}$. Finally, equation (4.2) for f gives

$$\sum_{q\in\mathbb{Q}} f(q) = \sum_{q\in\mathbb{Z}} f_\infty(q),$$

while (4.2) gives

$$\sum_{q\in\mathbb{Q}} \widehat{f}(q) = \sum_{q\in\mathbb{Z}} \widehat{f}_\infty(q).$$

The statement follows now from the classical Poisson summation formula for $\mathbb{R}$, namely

$$\sum_{q\in\mathbb{Z}} f_\infty(q) = \sum_{q\in\mathbb{Z}} \widehat{f}_\infty(q).$$

□

Definition 4.5.2. For a Schwartz-Bruhat function $f \in \mathcal{S}(\mathbb{A})$ let us define the *adelic theta function* $E(f)\colon \mathbb{I} \to \mathbb{C}$ by

$$E(f)(x) := |x|^{1/2} \sum_{q\in\mathbb{Q}^\times} f(qx).$$

for $x \in \mathbb{I}$.

Lemma 4.5.3. *For $f \in \mathcal{S}(\mathbb{A})$, we have:*

(a) The series defining $E(f)$ converges absolutely and locally uniformly on $\mathbb{I}$ and give a well-defined function on $\mathbb{I}/\mathbb{Q}^\times$.

(b) For every integer $n \geq 2$, there exists a positive constant C_n such that

$$|E(f)(x)| \leq C_n |x|^{-n}, \quad \textit{for all } x \in \mathbb{I} \textit{ with } |x| > 1.$$

(c) For all $x \in \mathbb{I}$:

$$E(f)(x) = E(\widehat{f})(\frac{1}{x}) + |x|^{-1/2}\widehat{f}(0) - |x|^{1/2} f(0).$$

Proof. By Lemma 4.4.5, it is enough to prove the Lemma for simple Schwartz-Bruhat functions of the form

$$f = \mathbb{1}_{a+N\widehat{\mathbb{Z}}} \cdot f_\infty$$

with $a \in \mathbb{A}$, $N \in \mathbb{Z}_{>0}$ and $f_\infty \in \mathcal{S}(\mathbb{R})$. Since $\mathbb{Q}$ is dense in $\mathbb{A}_{\mathrm{fin}}$, we may assume $a \in \mathbb{Q}$. We have

$$\sum_{q\in\mathbb{Q}^\times} f(qx_{\mathrm{fin}}) = \sum_{\substack{q\in\mathbb{Q}^\times \\ qx_{\mathrm{fin}}\in a+N\widehat{\mathbb{Z}}}} f_\infty(qx) = \sum_{\substack{q\in\mathbb{Q}^\times \\ qx\in N\widehat{\mathbb{Z}}}} f_\infty(qx+a).$$

So by replacing $f_\infty(x)$ by $f_\infty(x+a)$, we may without loss of generality assume that $a=0$, i.e., it suffices to prove the statements $(a)-(c)$ for simple Schwartz-Bruhat functions of the form $f = \mathbb{1}_{N\widehat{\mathbb{Z}}} \cdot f_\infty$.

(a) Since $f_\infty \in \mathcal{S}(\mathbb{R})$ we find some positive real number C such that $|f_\infty(x)| \leq C(1+|x|)^{-2}$ for every $x \in \mathbb{R}$. So we get the estimate

$$\sum_{q\in\mathbb{Q}^\times} |f(qx)| \leq C \sum_{\substack{q\in\mathbb{Q}^\times \\ qx_{\mathrm{fin}}\in N\widehat{\mathbb{Z}}}} (1+|qx_\infty|_\infty)^2 .$$

We show that the right hand side converges locally on subsets of the form $U := (\alpha,\beta) \times \frac{1}{M}\widehat{\mathbb{Z}}^\times \subseteq \mathbb{R}^\times \times \mathbb{I}_{\mathrm{fin}}$, where $(\alpha,\beta) \subseteq \mathbb{R}$ is an open interval which is contained in $\mathbb{R}^\times$. For $x \in U$ and $q \in \mathbb{Q}^\times$, we have $qx_{\mathrm{fin}} \in N\widehat{\mathbb{Z}}$ if and only if $q \in NM\widehat{\mathbb{Z}} \cap \mathbb{Q}^\times = NM\mathbb{Z} \setminus \{0\}$. So we get for $x \in U$

$$\sum_{q\in\mathbb{Q}^\times} |f(qx)| \leq C \sum_{\substack{q\in\mathbb{Q}^\times \\ qx_{\mathrm{fin}}\in N\widehat{\mathbb{Z}}}} (1+|qx_\infty|_\infty)^2 = C \sum_{q\in NM\mathbb{Z}\setminus\{0\}} (1+|qx_\infty|_\infty)^2 .$$

The last sum converges uniformly for $x_\infty \in (\alpha,\beta)$. This proves that the series in the statement converges locally uniformly on $\mathbb{I}$. Furthermore, $E(f)$ is obviously invariant under $\mathbb{Q}^\times$ and we obtain a continuous function on $\mathbb{I}/\mathbb{Q}^\times$.

(b) Next, we prove the estimate in (b), i.e., we prove for $f = \mathbb{1}_{N\widehat{\mathbb{Z}}} \cdot f_\infty$ and $n \geq 2$ that there exists a $C_n \in \mathbb{R}_{>0}$ s.t. for all $x \in \mathbb{I}$ with $|x| \geq 1$

$$|E(f)| \leq C_n |x|^{-n}.$$

So let $x \in \mathbb{I}$ with $|x| \geq 1$. Since $\mathbb{Q}^\times$ is dense in $\mathbb{I}_{\mathrm{fin}}$, we may without loss of generality assume $x_{\mathrm{fin}} \in \mathbb{Q}^\times$. For $q \in \mathbb{Q}$ we have $f(qx) = 0$ if $qx_{\mathrm{fin}} \notin N\widehat{\mathbb{Z}}$. On the other hand, $qx_{\mathrm{fin}} \in N\widehat{\mathbb{Z}}$ and $|x| \geq 1$ imply $|qx_\infty|_\infty \geq N$. Since $f_\infty \in \mathcal{S}(\mathbb{R})$ we can find for any $n \geq 2$ a positive constant c_n such that $f_\infty(qx_\infty) < c_n |qx_\infty|^{-n}$ for all $|qx_\infty|_\infty \geq N$. The above discussion implies

$$\begin{aligned}\sum_{q\in\mathbb{Q}^\times}|f(qx)| &\le c_n \sum_{\substack{q\in\mathbb{Q}^\times\\ qx_{\text{fin}}\in N\widehat{\mathbb{Z}}}} |qx_\infty|_\infty^{-n}\\ &= c_n \sum_{\substack{q\in\mathbb{Q}^\times\\ qx_{\text{fin}}\in N\widehat{\mathbb{Z}}}} |qx|^{-n}|qx_{\text{fin}}|_{\text{fin}}^{n}\\ &= c_n|x|^{-n} \sum_{\substack{q\in\mathbb{Q}^\times\\ qx_{\text{fin}}\in N\widehat{\mathbb{Z}}}} |qx_{\text{fin}}|_{\text{fin}}^{n}\\ &= c_n|x|^{-n} \underbrace{\sum_{k\in N\mathbb{Z}} |k|^{-n}}_{<\infty}.\end{aligned}$$

(c) For $x\in\mathbb{I}$ let us compute the Fourier transform of $g(y) := f(xy)$:

$$\widehat{g}(y) = \int_{\mathbb{A}} f(xz)e(-zy)d\mu(z) = |x|^{-1}\widehat{f}(\frac{y}{x}).$$

Applying now the Poisson summation formula to g gives

$$\begin{aligned}E(f)(x) &= |x|^{\frac{1}{2}} \sum_{q\in\mathbb{Q}^\times} f(qx) = |x|^{\frac{1}{2}} \sum_{q\in\mathbb{Q}} g(q) - |x|^{\frac{1}{2}} f(0)\\ &= |x|^{\frac{1}{2}} \sum_{q\in\mathbb{Q}} \widehat{g}(q) - |x|^{\frac{1}{2}} f(0) = |x|^{-\frac{1}{2}} \sum_{q\in\mathbb{Q}} \widehat{f}(\frac{q}{x}) - |x|^{\frac{1}{2}} f(0)\\ &= E(f)(\frac{1}{x}) + |x|^{\frac{1}{2}} f(0) - |x|^{-\frac{1}{2}}\widehat{f}(0).\end{aligned}$$

□

Next, we want to associate a zeta function to an arbitrary Schwartz-Bruhat function. Since the ideles can be written as a restricted product we get from Proposition 4.3.4 a unique Haar measure on the ideles from the local Haar measures. More concretely, let us define $\mu^\times := \mu^\times_\infty \times \mu^\times_{\text{fin}}$, where $\mu^\times_\infty$ is the Haar measure $\frac{dt}{|t|}$ on $\mathbb{R}^\times$ and $\mu^\times_{\text{fin}}$ is the unique Haar measure on $\mathbb{I}_{\text{fin}}$ normalized by

$$\mu^\times_{\text{fin}}(\widehat{\mathbb{Z}}^\times) = 1.$$

With this fixed Haar measure, we define the following adelic zeta function:

Definition 4.5.4. For $f\in\mathcal{S}(\mathbb{A})$ and $\chi\colon \widehat{\mathbb{Z}}^\times \to \mathbb{C}$ a Dirichlet character, let us define the adelic zeta function of f as[16]

$$\zeta(f,\chi,s) := \int_{\mathbb{I}} f(x)\chi(x)|x|^s d\mu^\times_{\text{fin}}(x).$$

Here, we view χ as a character on $\mathbb{I}$ via the projection

$$\mathbb{I} \to \mathbb{I}/\mathbb{Q}^\times \cong \widehat{\mathbb{Z}}^\times \times \mathbb{R}_{>0} \to \widehat{\mathbb{Z}}^\times,$$

see Theorem 4.3.11.

[16] It can be shown that any continuous homomorphism $\chi\colon \mathbb{I}/\mathbb{Q}^\times\mathbb{C}^\times$ is of the form $\chi(x) = \chi_0(x)\cdot|x_\infty|_\infty^s$ for a unique $s\in\mathbb{C}$ and a Dirichlet character χ_0. Thus, we could equally well write $\zeta(f,\chi)$ for $\zeta(f,\chi_0,s)$.

For the following, let us recall that $F := \mathbb{R}_{>0} \times \widehat{\mathbb{Z}}^\times \subseteq \mathbb{I}$ maps isomorphically to $\mathbb{I}/\mathbb{Q}^\times$ under the quotient map, see Theorem 4.3.11 (d). By abuse of notation, let us denote the induced Haar measure on $\mathbb{I}/\mathbb{Q}^\times$ obtained by restriction of $\mu^\times$ to $F \cong \mathbb{I}/\mathbb{Q}^\times$ again by $\mu^\times$. As in the case of the Riemann zeta function, we can write $\zeta(f,\chi,s)$ as a 'adelic Mellin tranform' of the theta function $E(f)(x)$.

Lemma 4.5.5. *The integral defining $\zeta(f,\chi,s)$ converges locally uniformly for* $\mathrm{Re}(s) > 1$ *and we have for such s*

$$\zeta(f,\chi,s) = \int_{\mathbb{I}/\mathbb{Q}^\times} E(f)(x)\chi(x)|x|^{s-1/2} d\mu^\times(x).$$

Proof. Since f is a Schwartz-Bruhat function, $|f(x)|$ can be estimated by $C \cdot \mathbb{1}_{N^{-1}\widehat{\mathbb{Z}}}(x_{\mathrm{fin}})(1+|x_\infty|_\infty^N)^{-1}$ for a suitable constant $C \in \mathbb{R}_{>0}$ and a positive integer $N > \mathrm{Re}(s)$. The claimed convergence follows from the estimate

$$\int_{\mathbb{I}} |f(x)||x|^s d\mu^\times(x) \leq C \int_{\mathbb{I}_{\mathrm{fin}}} |x_{\mathrm{fin}}|_{\mathrm{fin}}^s \cdot \mathbb{1}_{N^{-1}\widehat{\mathbb{Z}}}(x_{\mathrm{fin}}) d\mu_{\mathrm{fin}}^\times(x_{\mathrm{fin}}) \\ \times \int_{\mathbb{R}^\times} (1+|x_\infty|_\infty^N)^{-1}|x_\infty|^s \frac{dx_\infty}{|x_\infty|}$$

and the convergence of the integrals on the right hand side; for the convergence of the integral over the finite ideles, observe

$$\begin{aligned}\int_{\mathbb{I}_{\mathrm{fin}}} |x_{\mathrm{fin}}|_{\mathrm{fin}}^s \cdot \mathbb{1}_{N^{-1}\widehat{\mathbb{Z}}}(x_{\mathrm{fin}}) d\mu_{\mathrm{fin}}^\times(x_{\mathrm{fin}}) &= N^s \int_{\mathbb{I}_{\mathrm{fin}}} |x_{\mathrm{fin}}|_{\mathrm{fin}}^s \cdot \mathbb{1}_{\widehat{\mathbb{Z}}}(x_{\mathrm{fin}}) d\mu_{\mathrm{fin}}^\times(x_{\mathrm{fin}}) \\ &= N^s \sum_{k\in\mathbb{N}} \int_{k\widehat{\mathbb{Z}}^\times} |x_{\mathrm{fin}}|_{\mathrm{fin}}^s d\mu_{\mathrm{fin}}^\times(x_{\mathrm{fin}}) \\ &= N^s \sum_{k\in\mathbb{N}} k^{-s} \underbrace{\int_{\widehat{\mathbb{Z}}^\times} |x_{\mathrm{fin}}|_{\mathrm{fin}}^s d\mu_{\mathrm{fin}}^\times(x_{\mathrm{fin}})}_{=1} < \infty.\end{aligned}$$

These estimates hold locally uniformly for $\mathrm{Re}(s) > 1$ and the statement about convergence follows. The set $F := \mathbb{R}_{>0} \times \widehat{\mathbb{Z}}^\times$ is by Theorem 4.3.11 (d) a set of representatives for $\mathbb{I}/\mathbb{Q}^\times$. By absolute convergence, we can compute for $\mathrm{Re}(s) > 1$:

$$\begin{aligned}\zeta(f,\chi,s) &= \int_{\mathbb{I}} f(x)\chi(x)|x|^s d\mu^\times(x) = \sum_{q\in\mathbb{Q}} \int_{qF} f(x)\chi(x)|x|^s d\mu^\times(x) \\ &= \sum_{q\in\mathbb{Q}} \int_F f(qx)\chi(qx)|qx|^s d\mu^\times(x) = \int_{\mathbb{I}/\mathbb{Q}^\times} E(f)(x)\chi(x)|x|^{s-1/2} d\mu^\times(x)\end{aligned}$$

□

Finally, we can use the functional equation of the adelic theta function $E(f)(x)$ to prove the analytic continuation and functional equation of the adelic zeta function:

Theorem 4.5.6. *The adelic zeta function $\zeta(f,\chi,s)$ admits a holomorphic continuation to $\mathbb{C}\setminus\{0,1\}$ with at most simple poles at $s=0,1$ with residues*

$$\operatorname{Res}_{s=0}\zeta(f,\chi,s)=\begin{cases}-f(0) & \chi \text{ is trivial}\\ 0 & \chi \text{ is non-trivial}\end{cases}$$

$$\operatorname{Res}_{s=1}\zeta(f,\chi,s)=\begin{cases}\widehat{f}(0) & \chi \text{ is trivial}\\ 0 & \chi \text{ is non-trivial}\end{cases}.$$

Furthermore, it satisfies the functional equation

$$\zeta(f,\chi,s)=\zeta(\widehat{f},\overline{\chi},1-s).$$

Proof. The set $\{1\}$ has measure zero in $\mathbb{R}_{>0}$, so $\mathbb{I}_1/\mathbb{Q}^\times$ is a set of measure zero in $\mathbb{I}/\mathbb{Q}^\times$. Therefore, we can decompose the above integral into a problematic part and an unproblematic part:

$$\begin{aligned}\zeta(f,\chi,s)&=\int_{\mathbb{I}/\mathbb{Q}^\times}E(f)(x)\chi(x)|x|^{s-1/2}d\mu^\times(x)\\ &=\underbrace{\int_{|\cdot|<1}E(f)(x)\chi(x)|x|^{s-1/2}d\mu^\times(x)}_{\text{problematic part}}+\underbrace{\int_{|\cdot|>1}E(f)(x)\chi(x)|x|^{s-1/2}d\mu^\times(x)}_{\text{unproblematic part}}.\end{aligned}$$

Let us first check that the unproblematic integral converges for all $s\in\mathbb{C}$. By Lemma 4.5.3 we find for every positive integer n a constant C_n such that $|E(f)(x)|<C_n|x|^{-n}$ for $|x|>1$. Thus, we get

$$\begin{aligned}\int_{|x|>1}|E(f)(x)||x|^{\operatorname{Re}s-\frac{1}{2}}d\mu^\times(x)&\le C_n\int_{|x|>1}|x|^{\operatorname{Re}s-\frac{1}{2}-n}d\mu^\times(x)\\ &=C_n\int_1^\infty t^{\operatorname{Re}s-\frac{1}{2}-n}\frac{dt}{t}\cdot\underbrace{\int_{\widehat{\mathbb{Z}}^\times}1d\mu^\times_{\text{fin}}(x)}_{=1}\\ &=C_n\int_1^\infty t^{\operatorname{Re}s-\frac{1}{2}-n}\frac{dt}{t}.\end{aligned}$$

The last integral converges for $\operatorname{Re}(s)<n-\frac{1}{2}$. Since n was arbitrary, we get the desired convergence.
As in the proof of the functional equation for the Riemann zeta function, we will use the functional equation from Lemma 4.5.3

$$E(f)(x)=E(\widehat{f})(\frac{1}{x})+|x|^{-1/2}\widehat{f}(0)-|x|^{1/2}f(0).$$

to extend the problematic integral meromorphically. Using this functional equation, we can rewrite the problematic integral as follows:

$$\begin{aligned}\int_{|\cdot|<1}E(f)(x)\chi(x)|x|^{s-1/2}d\mu^\times(x)&=\int_{|\cdot|<1}E(\widehat{f})\left(\frac{1}{x}\right)\chi(x)|x|^{s-1/2}d\mu^\times(x)\\ &+\widehat{f}(0)\int_{|\cdot|<1}\chi(x)|x|^{s-1}d\mu^\times(x)+f(0)\int_{|\cdot|<1}\chi(x)|x|^{s}d\mu^\times(x).\end{aligned}$$

The first integral is

$$\int_{|\cdot|>1} E(\widehat{f})(x)\overline{\chi(x)}|x|^{1/2-s}d\mu^\times(x),$$

while the second and third integral are given by

$$\begin{aligned}\int_{|\cdot|<1} \chi(x)|x|^{s-1}d\mu^\times(x) &= \int_0^1 t^{s-1}\frac{dt}{t}\cdot\int_{\widehat{\mathbb{Z}}^\times}\chi(x)d\mu^\times_{\text{fin}}(x)\\ &= \frac{1}{s-1}\begin{cases}0 & \text{if } \chi \text{ is non-trivial}\\ 1 & \text{if } \chi \text{ is trivial}\end{cases}\end{aligned}$$

and similarly

$$\int_{|\cdot|<1} \chi(x)|x|^{s}d\mu^\times(x) = \frac{1}{s}\begin{cases}0 & \text{if } \chi \text{ is non-trivial}\\ 1 & \text{if } \chi \text{ is trivial}\end{cases}.$$

Thus, we may write the adelic zeta function as follows:

$$\begin{aligned}\zeta(f,\chi,s) &= \int_{\mathbb{I}/\mathbb{Q}^\times} E(f)(x)\chi(x)|x|^{s-1/2}d\mu^\times(x)\\ &= \int_{|\cdot|>1} E(f)(x)\chi(x)|x|^{s-1/2}d\mu^\times(x)\\ &\quad+\int_{|\cdot|>1} E(\widehat{f})(x)\overline{\chi(x)}|x|^{1/2-s}d\mu^\times(x)\\ &\quad-\left(\frac{\widehat{f}(0)}{1-s}+\frac{f(0)}{s}\right)\begin{cases}0 & \text{if } \chi \text{ is non-trivial}\\ 1 & \text{if } \chi \text{ is trivial}\end{cases}\end{aligned}$$

This formula proves that $\zeta(f,\chi,s)$ admits an analytic continuation with at most simple poles at $s=0$ and $s=1$ and residues $\mathrm{Res}_{s=0}\,\zeta(f,\chi,s) = f(0)$ and $\mathrm{Res}_{s=1}\,\zeta(f,\chi,s) = -\widehat{f}(0)$. The functional equation follows by comparing this formula to the corresponding formula for $\zeta(\widehat{f},\overline{\chi},1-s)$. □

4.6 *Functional equation for Dirichlet L-functions*

In this section, we prove the functional equation for Dirichlet L-functions. Therefore, we want to apply the abstract functional equation for adelic zeta functions for a suitable choice of Schwartz-Bruhat function $f\in\mathcal{S}(\mathbb{A})$. It will turn out that the following choice gives the desired result:

Definition 4.6.1. For a Dirichlet character χ of conductor D let us define S to be the set of primes dividing D. Furthermore, let us define depending on the parity of χ the quantity $\epsilon\in\{0,1\}$ by

$$\chi(-1) = (-1)^\epsilon.$$

Furthermore, we define

- For $p \in S$, we set $f_p := p^k(1-1/p)\mathbb{1}_{1+p^k\mathbb{Z}_p}$ with $k := \nu_p(D)$.
- For $p \in \mathbb{P} \setminus S$, we set $f_p := \mathbb{1}_{\mathbb{Z}_p}$.
- For $\nu = \infty$, we set $f_\infty(x) := x^\epsilon \exp(-\pi^2 x)$.

This defines a Schwartz-Bruhat function $f := \prod_{\nu \in \mathbb{P}\cup\{\infty\}} f_\nu \in \mathcal{S}(\mathbb{A})$.

Theorem 4.6.2. *Let χ be a Dirichlet character of conductor D and choose the Schwartz-Bruhat function as in Definition 4.6.1. Then*

$$\zeta(f, \overline{\chi}, s) = L_\infty(\chi, s)L(\chi, s)$$

where $L_\infty(\chi, s)$ is the Gamma factor given by

$$L_\infty(\chi, s) = \Gamma\left(\frac{s+\epsilon}{2}\right)\pi^{-\frac{s+\epsilon}{2}}.$$

In particular, the Dirichlet L-function satisfies the functional

$$L_\infty(\chi, s)L(\chi, s) = (-i)^\epsilon D^{-s}\mathcal{G}(\chi)L_\infty(\overline{\chi}, 1-s)L(\overline{\chi}, 1-s)$$

with[17]

$$\mathcal{G}(\chi) = \varphi(D)\int_{\frac{1}{D}\widehat{\mathbb{Z}}^\times} \chi(x)e_{fin}(-x)d\mu_{fin}^\times(x)$$

and $e_{fin} = \prod_{p\in\mathbb{P}} e_p$. Furthermore, we have $|\mathcal{G}(\chi)| = \sqrt{D}$.

Proof. For $\nu \in \mathbb{P} \cup \{\infty\}$, let us write χ_ν for the local character

$$\mathbb{Q}_\nu^\times \to \mathbb{I}/\mathbb{Q}^\times \to \mathbb{C}^\times,$$

given by the inclusion $\mathbb{Q}_\nu^\times \to \mathbb{I}$, $x \mapsto (1, \dots, 1, x, 1, \dots)$. With the choice of f from Definition 4.6.1, we can write the adelic zeta function as a product of local integrals

$$\zeta(f, \overline{\chi}, s) = \int_{\mathbb{I}} f(x)\chi(x)|x|^s d\mu^\times(x) = \prod_{\nu\in\mathbb{P}\cup\{\infty\}} \int_{\mathbb{Q}_\nu^\times} f_\nu(x_\nu)\overline{\chi}_\nu(x_\nu)|x|_\nu^s d\mu_\nu^\times.$$

where $d\mu_\nu^\times$ denotes the Haar measure $dt/|t|$ for $\nu = \infty$ and the unique Haar measure on $\mathbb{Q}_p^\times$ with $\mu_p^\times(\mathbb{Z}_p^\times) = 1$ for $\nu = p \in \mathbb{P}$. Our aim is to identify the local integrals with the Euler factors of the Dirichlet L-function. Let us start with $p \in \mathbb{P} \setminus S$. For such a prime, we have $f_p = \mathbb{1}_{\mathbb{Z}_p}$ and $\chi_p|_{\mathbb{Z}_p^\times} = 1$ and hence we compute

$$\begin{aligned}&\int_{\mathbb{Q}_p^\times} f_p(x)\overline{\chi}_p(x)|x|_p^s d\mu_p^\times(x) = \int_{\mathbb{Z}_p\setminus\{0\}} \overline{\chi}_p(x)|x|_p^s d\mu_p^\times(x)\\ &= \sum_{j=0}^\infty \int_{p^j\mathbb{Z}_p^\times} \overline{\chi}_p(x)p^{-js}d\mu_p^\times(x) = \sum_{j=0}^\infty \overline{\chi}_p(p)^j p^{-js} = \frac{1}{1-p^{-s}\chi_p(1/p)}.\end{aligned}$$

Finally, let us observe that

$$\chi_p(1/p) = \chi(1, \dots, 1, 1/p, 1, \dots) = \chi(p, \dots, p, 1, p, \dots).$$

[17] Here, $\mu_{\text{fin}}^\times$ denotes the 'restricted product' measure on $\mathbb{I}_{\text{fin}}$.

The latter element is contained in $\widehat{\mathbb{Z}}^\times$ and hence $\chi_p(1/p) = \chi(p)$ since $p \in \mathbb{P} \setminus \{S\}$ and χ factors over the projection to $\prod_{p\in S} \mathbb{Z}_p^\times$.

For $p \in S$, we have $f_p = p^k(1-1/p)\mathbb{1}_{1+p^k\mathbb{Z}_p}$ and $\chi_p(1+p^k\mathbb{Z}_p) = 1$ with $k := \nu_p(D)$. We get

$$\begin{aligned}\int_{\mathbb{Q}_p^\times} f_p(x)\overline{\chi}_p(x)|x|_p^s d\mu_p^\times(x) &= p^k(1-1/p)\int_{1+p^k\mathbb{Z}_p} \overline{\chi}_p(x)|x|_p^s d\mu_p^\times(x) \\ &= p^k(1-1/p)\int_{1+p^k\mathbb{Z}_p} d\mu_p^\times(x) = 1\end{aligned}$$

Finally, the Archimedean place gives the contribution

$$\begin{aligned}&\int_{\mathbb{R}^\times} e^{-\pi x^2} x^\epsilon \underbrace{\overline{\chi}_\infty(x)}_{\mathrm{sign}(x)} |x|^s \frac{dx}{|x|} = \int_{\mathbb{R}^\times} e^{-\pi x^2}|x|^{s+\epsilon}\frac{dx}{|x|} \\ &\quad = \int_{\mathbb{R}^\times} e^{-\pi x^2}(x^2)^{\frac{s+\epsilon}{2}}\frac{dx}{|x|} = \int_{\mathbb{R}_{>0}} e^{-\pi t} t^{\frac{s+\epsilon}{2}}\frac{dt}{t} = \Gamma\left(\frac{s+\epsilon}{2}\right)\pi^{-\frac{s+\epsilon}{2}}.\end{aligned}$$

Combining everything gives for $\mathrm{Re}(s) > 1$ the desired equality

$$\zeta(f,\overline{\chi},s) = L_\infty(\chi,s)L(\chi,s).$$

Let us now deduce the functional equation of the Dirichlet L-function from the abstract adelic functional equation:

$$\zeta(f,\overline{\chi},s) = \zeta(\widehat{f},\chi,1-s).$$

Therefore, we have to compute the local zeta integrals

$$\int_{\mathbb{Q}_\nu^\times} \widehat{f}_\nu(x)\chi_\nu(x)|x|_\nu^{1-s}d\mu_\nu^\times(x)$$

for the Fourier transform of f_ν for $\nu \in \mathbb{P}\cup\{\infty\}$. For $p \in \mathbb{P}\setminus S$, we have $f_p = \mathbb{1}_{\mathbb{Z}_p}$ and hence $\widehat{f}_p = \mathbb{1}_p$. The local zeta integral is then given by a similar computation as above by

$$\int_{\mathbb{Q}_p^\times} \widehat{f}_p(x)\chi_p(x)|x|_p^{1-s}d\mu_p^\times(x) = \frac{1}{1-\chi(p)p^{-(1-s)}}.$$

For $\nu = \infty$ and $\epsilon = 0$, we get $\widehat{f_\infty} = f_\infty$. For $\nu = \infty$ and $\epsilon = 1$, we compute

$$\begin{aligned}\widehat{f_\infty}(x) &= \int_{\mathbb{R}} y e^{-\pi y^2} e^{-2\pi i x y} dy = -\frac{1}{2\pi i}\frac{\partial}{\partial x}\int_{\mathbb{R}^\times} e^{-\pi y^2}e^{-2\pi i x y}dy \\ &= -\frac{1}{2\pi i}\frac{\partial}{\partial x}e^{-\pi x^2} = -ixe^{-\pi x^2} = -if_\infty(x).\end{aligned}$$

So, for $\nu = \infty$, we have $\widehat{f_\infty} = (-i)^\epsilon f_\infty$ and the above computation shows

$$\int_{\mathbb{R}^\times} \widehat{f_\infty}(x)\chi_\infty(x)|x|_\infty^{1-s}d\mu_\infty^\times(x) = (-i)^\epsilon L_\infty(\overline{\chi},1-s).$$

For $p \in S$, we have $f_p = p^k(1-1/p)\mathbb{1}_{1+p^k\mathbb{Z}_p}$ and its Fourier transform is readily computed;

$$\begin{aligned}\widehat{\mathbb{1}}_{1+p^k\mathbb{Z}_p}(x) &= p^k\left(1-\frac{1}{p}\right)\int_{1+p^k\mathbb{Z}_p} e_p(-xy)d\mu_p(y)\\ &= p^k\left(1-\frac{1}{p}\right)p^{-k}e_p(-x)\int_{\mathbb{Z}_p} e_p(-p^kxy)d\mu_p(y)\\ &= \left(1-\frac{1}{p}\right)e_p(-x)\mathbb{1}_{p^{-k}\mathbb{Z}_p}(x).\end{aligned}$$

We start computing the local zeta integral:

$$\begin{aligned}\int_{\mathbb{Q}_p^\times}\widehat{f}_p(x)\chi_p(x)|x|_p^{1-s}d\mu_p^\times(x) &= \left(1-\frac{1}{p}\right)\int_{p^{-k}\mathbb{Z}_p\setminus\{0\}} e_p(-x)\chi_p(x)|x|^{1-s}d\mu_p^\times(x)\\ &= \left(1-\frac{1}{p}\right)\sum_{j=-k}^{\infty} p^{js-j}\int_{p^j\mathbb{Z}_p^\times} e_p(-x)\chi_p(x)d\mu_p^\times(x) \quad (4.3)\end{aligned}$$

Claim: For $j > -k$, we have

$$\int_{p^j\mathbb{Z}_p^\times} e_p(-x)\chi_p(x)d\mu_p^\times(x) = 0.$$

Proof of the Claim: For $j \geq 0$, we have $e_p(x) = 1$ for all $x \in p^j\mathbb{Z}_p^\times$ and hence

$$\int_{p^j\mathbb{Z}_p^\times} e_p(-x)\chi_p(x)d\mu_p^\times(x) = \int_{p^j\mathbb{Z}_p^\times}\chi_p(x)d\mu_p^\times(x)$$

vanishes since it is an integral over a non-trivial character on a compact group, see Lemma 4.1.9. On the other hand, for $-k < j < 0$, the character χ_p is non-trivial on $1+p^{k-1}\mathbb{Z}_p$. So, let us choose a $\lambda \in 1+p^{k-1}\mathbb{Z}_p$ with $\chi(\lambda) \neq 1$. Since $j > -k$, we have $x\lambda \equiv x \mod \mathbb{Z}_p$ for every $x \in p^j\mathbb{Z}_p^\times$. This implies $e_p(-\lambda x) = e_p(-x)$ and we get

$$\begin{aligned}\chi_p(\lambda)\int_{p^j\mathbb{Z}_p^\times} e_p(-x)\chi_p(x)d\mu_p^\times(x) &= \int_{p^j\mathbb{Z}_p^\times} e_p(-\lambda x)\chi_p(\lambda x)d\mu_p^\times(x)\\ &= \int_{p^j\mathbb{Z}_p^\times} e_p(-x)\chi_p(x)d\mu_p^\times(x).\end{aligned}$$

From $\chi(\lambda) \neq 1$, we deduce

$$\int_{p^j\mathbb{Z}_p^\times} e_p(-x)\chi_p(x)d\mu_p^\times(x) = 0$$

as desired. This proves the claim.

Using the above claim in equation (4.3) gives for $p \in S$

$$\int_{\mathbb{Q}_p^\times}\widehat{f}_p(x)\chi_p(x)|x|_p^{1-s}d\mu_p^\times(x) = \left(p^k - p^{k-1}\right)p^{-sk}\int_{p^{-k}\mathbb{Z}_p^\times} e_p(-x)\chi_p(x)d\mu_p^\times(x).$$

Combining the above computations of the local zeta integrals for $\widehat{f}$ and taking the product over all v gives

$$\prod_v \int_{\mathbb{Q}_v^\times} \widehat{f}_v(x)\chi_v(x)|x|_v^{1-s} d\mu_v^\times(x) = (-i)^\epsilon D^{-s}\mathcal{G}(\chi)L_\infty(\overline{\chi}, 1-s)L(\overline{\chi}, 1-s)$$

and the functional equation follows. It remains to compute the absolute value of $\mathcal{G}(\chi)$: Applying the functional equation twice shows $|\mathcal{G}(\chi)\mathcal{G}(\overline{\chi})| = D$. On the other hand, we have

$$\begin{aligned}\overline{\mathcal{G}(\chi)} &= \varphi(D)\int_{\frac{1}{D}\widehat{\mathbb{Z}}^\times} \overline{\chi(x)e_{\text{fin}}(-x)} d\mu_{\text{fin}}^\times(x) \\ &= \varphi(D)\overline{\chi}(-1)\int_{\frac{1}{D}\widehat{\mathbb{Z}}^\times} \overline{\chi}(x)e_{\text{fin}}(-x)d\mu_{\text{fin}}^\times(x) = \overline{\chi}(-1)\mathcal{G}(\overline{\chi}).\end{aligned}$$

So $\mathcal{G}(\overline{\chi})$ and $\overline{\mathcal{G}(\chi)}$ only differ by a sign and we conclude $|\mathcal{G}(\chi)| = \sqrt{D}$. □

Outlook

The approach of Tate's thesis works for more general number fields. For a general number field K, we define a Hecke character as a continuous homomorphism

$$\chi\colon \mathbb{I}_K/K^\times \to \mathbb{C}^\times.$$

Such homomorphisms to $\mathbb{C}^\times$ generalize characters and are often called *quasi-characters*. We have seen in the exercises that every such homomorphism for $K = \mathbb{Q}$ is of the form

$$\frac{\chi_0}{|\cdot|^s}$$

for some $s \in \mathbb{C}$ and χ_0 a Dirichlet character. So a Hecke character for $K = \mathbb{Q}$ generalizes the concept of Dirichlet characters. This point of view is quite convenient; if $\mathrm{Re}(s) > 1$ such a Hecke character for $K = \mathbb{Q}$ allows us to write the Dirichlet L-function in the convenient way

$$L(\chi_0, s) = \sum_{n\geq 1} \chi(n).$$

So in some sense, we do not distinguish between the part of the character $|\cdot|^s$ coming from the Archimedean part of the ideles and the part χ_0 coming form the finite ideles. Among all these Hecke characters there are certain *algebraic Hecke characters*. In the case of $K = \mathbb{Q}$, these are the Hecke characters corresponding to $\frac{\chi_0}{|\cdot|^n}$ for $n \in \mathbb{Z}$. Furthermore, it is useful to introduce the notion of a *dual* Hecke character. In the case $K = \mathbb{Q}$, this is the Hecke character $\chi^\vee := \frac{\overline{\chi}_0}{|\cdot|^{1-s}}$. So it is exactly the Hecke character on the other side of the adelic functional equation, i.e., if we write $\zeta(f,\chi) := \zeta(f,\chi_0,s)$ we may write the adelic functional equation as

$$\zeta(f,\chi) = \zeta(\widehat{f},\chi^\vee).$$

This point of view generalizes to arbitrary number fields. The adelic zeta function can be defined for arbitrary number fields and satisfies a nice functional equation:

$$\zeta(f,\chi) = \zeta(f,\chi^\vee).$$

For a suitable choice of adelic Schwartz-Bruhat functions, we can relate the adelic zeta function $\zeta(f,\chi)$ to classical Hecke L-functions. In particular, we obtain the functional equation for all Dedekind zeta functions. Furthermore, one can use Tate's thesis to give a conceptional proof of the analytic class number formula:

$$\mathrm{Res}_{s=1}\,\zeta_K(s) = \frac{2^r(2\pi)^s h_K \mathrm{Reg}_K}{e_K\sqrt{|d_K|}},$$

where r (resp.) s are the number of real (resp. pairs of complex) embeddings of K, h_K is the class number, Reg_K the regulator, e_k the number of roots of unity in K and d_K the discriminant of K. The point is that the residues in the adelic functional equation are given for general number fields by

$$-f(0)\mathrm{Vol}(C_K^1) \quad \text{resp. } \widehat{f}(0)\mathrm{Vol}(C_K^1),$$

where C_K^1 is the idele class group $\mathbb{I}_{1,K}/K^\times$ and $\mathrm{Vol}(C_K^1)$ is the volume of this compact group with respect to the measure induced by the normalized measure on the ideles. It turns out that

$$\mathrm{Vol}(C_K^1) = \frac{2^r(2\pi)^s h_K \mathrm{Reg}_K}{e_K\sqrt{|d_K|}}$$

which proves the analytic class number formula for Dedekind zeta functions.

5 Towards Iwasawa Theory

In this final chapter of the lecture notes, we will prove Kummer's criterion relating the p-divisibility of zeta values to the p-divisibility of the class group of the number field $\mathbb{Q}(\zeta_p)$. The proof will combine many of the aspects of L-functions we have seen in this lecture. In particular, it involves generalized Bernoulli numbers, the functional equation of Dirichlet L-functions, the explicit formula for the values of Dirichlet L-functions and the analytic class number formula.

5.1 More on generalized Bernoulli numbers

For the proof of Kummer's criterion, we will need to study generalized Bernoulli numbers more carefully.

Recall that, for a Dirichlet character χ modulo D, we have defined the generalized Bernoulli numbers as linear combinations of values of the Bernoulli polynomials, more precisely:

$$B_{n,\chi} := D^{n-1} \sum_{d=1}^{D} \chi(d) B_n(d/D).$$

Together with the generating function[1] of the Bernoulli polynomials this gives immediately the following generating function for the generalized Bernoulli numbers:

$$\sum_{d=1}^{D} \chi(d) \frac{te^{dt}}{e^{Dt}-1} = \sum_{n=1}^{\infty} B_{n,\chi} \frac{t^n}{n!}.$$

It will be useful to introduce Bernoulli polynomials twisted by a Dirichlet character:

Definition 5.1.1. For a Dirichlet character χ modulo D, let us define the *generalized Bernoulli polynomials* by the generating series

$$\sum_{d=1}^{D} \chi(d) \frac{te^{(d+X)t}}{e^{Dt}-1} = \sum_{n=1}^{\infty} B_{n,\chi}(X) \frac{t^n}{n!}.$$

In Exercises 3 and 4 on Sheet 4, we have already seen that there is a close relation between power sums $\sum_{i=1}^{n} i^k$ and the Bernoulli numbers.

[1] Recall

$$\frac{te^{Xt}}{e^t-1} = \sum_{n=1}^{\infty} B_n(X) \frac{t^n}{n!}.$$

The generlized Bernoulli numbers are related to the following twisted power sums:

Proposition 5.1.2. *Let k be a positive integer and χ a Dirichlet character modulo D. The generalized Bernoulli polynomials satisfy the following formula*

$$\sum_{d=1}^{Dk} \chi(d)d^n = \frac{1}{n+1}\left(B_{n+1,\chi}(Dk) - B_{n+1,\chi}\right).$$

Proof. Let us write $f(t,X)$ for the generating function of the generalized Bernoulli polynomials and consider

$$f(T,X+D)-f(T,X) = (e^{(X+D)t}-e^{Xt})\sum_{d=1}^{D}\frac{\chi(d)te^{dt}}{e^{Dt}-1} = \sum_{d=1}^{D}\chi(d)te^{(d+X)t}.$$

Comparing the coefficients of $t^{n+1}/(n+1)!$ on both sides of the equality shows

$$B_{n+1,\chi}(X+f) - B_{n+1,\chi}(X) = (n+1)\sum_{d=1}^{D}\chi(d)(d+X)^n.$$

Evaluating at $X = jD$ for $0 \leq j < k$ and summing over these values gives

$$B_{n+1,\chi}(kD) - B_{n+1,\chi}(0) = (n+1)\sum_{j=0}^{k-1}\sum_{d=1}^{D}\chi(d)(d+jD)^n = (n+1)\sum_{d=1}^{Dk}\chi(d)d^n.$$

□

We have already seen in one of the question hours that the Bernoulli numbers satisfy certain p-adic congruences, *the Kummer congruences.*

Theorem 5.1.3 (Kummer)**.** *Let p be a prime and m,n be positive integers satisfying $n \equiv m \not\equiv 0 \mod p-1$. If $m \equiv n \mod p^k$ then*

$$(1-p^{m-1})\frac{B_m}{m} \equiv (1-p^{n-1})\frac{B_n}{n} \mod p^{k+1}.$$

Proof. The proof is not too difficult, but we have to skip it due to time reasons. □

In the following, let p be an odd prime. In the next section, we want to relate the class number of $\mathbb{Q}(\zeta_p)$ to the p-adic properties of zeta values. For this, the Teichmüller character of $\mathbb{Q}(\zeta_p)$ plays an important role. In Exercise 2 of Sheet 6, we have shown that there is a unique group homomorphism

$$\omega\colon (\mathbb{Z}/p\mathbb{Z})^\times \to \mu_{p-1}(\mathbb{Z}_p) \subseteq \mathbb{Z}_p^\times$$

satisfying $\omega(x) \mod p = x$. Let us choose an isomorphism $\mu_{p-1}(\mathbb{Z}_p) \cong \mu_{p-1}(\mathbb{C})$. Note that this isomorphism is equivalent to the choice of a prime ideal $\mathfrak{p}$ over p in $\mathbb{Q}(\mu_{p-1}(\mathbb{C}))$; indeed, the choice of an embedding $\mathbb{Q}(\mu_{p-1}(\mathbb{C})) \hookrightarrow \mathbb{Q}_p$ is equivalent to the choice of a prime ideal $\mathfrak{p}$ above p. Using this isomorphism, we may view ω as a Dirichlet character

$$\omega\colon (\mathbb{Z}/p\mathbb{Z})^\times \to \mathbb{C}^\times.$$

In the following, we will fix an isomorphism $\mu_{p-1}(\mathbb{Z}_p) \cong \mu_{p-1}(\mathbb{C})$ and regard ω as a Dirichlet character.

Corollary 5.1.4. *For a positive integer k with $1+k \not\equiv 0 \mod p-1$, we have*

$$B_{1,\omega^k} \equiv B_{1+k} \mod \mathfrak{p},$$

and for $1+k \equiv 0 \mod p-1$, we get

$$pB_{1,\omega^k} \equiv -1 \mod \mathfrak{p}.$$

Proof. A straightforward computations with the generating function of the generalized Bernoulli polynomials shows for an arbitrary Dirichlet character χ

$$B_{n,\omega^k}(X) = \frac{1}{n}\sum_{k=0}^{n}\binom{n}{k}B_{n-k,\chi}\cdot X^k. \tag{5.1}$$

In particular, we obtain for $n=2$ and $\chi = \omega^k$ the equation $B_{2,\omega^k}(X) = B_{2,\omega^k} + 2B_{1,\omega^k}X + B_{0,\omega^k}$. Proposition 5.1.2 gives

$$\sum_{d=1}^{p}\omega^k(d)d = \frac{1}{2}\left(2B_{1,\omega^k}\cdot p + B_{0,\omega^k}\cdot p^2\right).$$

This implies[2]

$$B_{1,\omega^k} \equiv \frac{1}{p}\sum_{d=1}^{p}\omega^k(d)d \mod p. \tag{5.2}$$

By the defining property of the Teichmüller character, we have

$$\omega^k(d) \equiv d^k \mod \mathfrak{p}$$

and hence[3] $\omega^{kp}(d) \equiv d^{kp} \mod p\cdot\mathfrak{p}$. Together with (5.2) this yields

$$B_{1,\omega^k} \equiv \frac{1}{p}\sum_{d=1}^{p}d^{pk+1} \mod \mathfrak{p}.$$

If $1+k \equiv 0 \mod p-1$, each $1 \le d < p$ satisfies $d^{pk+1} \equiv 1 \mod p$ and we get $pB_{1,\omega^{p-2}} \equiv -1 \mod \mathfrak{p}$. For $k+1 \not\equiv 0 \mod p-1$, the same argument as above using (5.1) and Proposition 5.1.2 shows

$$B_{k+1} \equiv \frac{1}{p}\sum_{d=1}^{p}d^{pk+1} \mod p.$$

Comparing the congruences for B_{k+1} and B_{1,ω^k} proves the statement of the Corollary.[4] □

[2] Note that $p \neq 2$ by the assumption $k+1 \not\equiv 0 \mod p-1$.

[3] Using that the Binomial coefficients $\binom{p}{i}$ are divisible by p for $1 \le i \le p-1$, it is not difficult to check that $x \equiv y \mod \mathfrak{p}$ implies $x^p \equiv y^p \mod \mathfrak{p}\cdot p$.

[4] The congruence in the statement can also be deduced using general Kummer congruences between generalized Bernoulli numbers.

5.2 The Kummer criterion

We are now well-prepared to prove Kummer's criterion. Let us first recall the class number formula for the Dedekind zeta function

$$\zeta_K(s) = \sum_{\mathfrak{a} \subseteq \mathcal{O}_K} \frac{1}{N\mathfrak{a}^s}$$

of a number field K from Algebraic Number Theory 1:

Theorem 5.2.1 (Analytic class number formula)**.** *For a number field K we have*

$$\operatorname{Res}_{s=1} \zeta_K(s) = \frac{2^r (2\pi)^s h_K \operatorname{Reg}_K}{e_K \sqrt{|d_K|}},$$

where r (resp.) s are the number of real (resp. pairs of complex) embeddings of K, h_K is the class number, Reg_K the regulator, e_k the number of roots of unity in K and d_K the discriminant of K.

On the other hand, we can use Theorem **??** and the functional equation to get the following explicit formula for the value of the Dirichlet L-function at $s = 1$ for an odd[5] Dirichlet character χ of conductor D

$$L(\chi, 1) = \pi i \frac{\mathcal{G}(\chi)}{D} B_{1,\overline{\chi}}.$$

Let us recall that the Dedekind zeta function of an abelian extension of $\mathbb{Q}$ can be expressed as a product of Dirichlet L-functions. More precisely, we obtain the following Corollary of Theorem 3.14.2 and Theorem 3.14.4:

Corollary 5.2.2. *Let K be an abelian extension of $\mathbb{Q}$ and $\mathbb{Q}(\zeta_D)$ the smallest cyclotomic extension containing K. Let us denote by*

$$X := \{\chi \in Gal(\mathbb{Q}(\zeta_D)/\mathbb{Q})^\vee : \chi|_{Gal(\mathbb{Q}(\zeta_D)/K)} = 1\}$$

the set of characters of $Gal(\mathbb{Q}(\zeta_D)/\mathbb{Q})$ which are trivial on $Gal(\mathbb{Q}(\zeta_D)/K)$. Using $Gal(\mathbb{Q}(\zeta_D)/\mathbb{Q}) \cong (\mathbb{Z}/D\mathbb{Z})^\times$ we will view X as a subset of the Dirichlet characters modulo D. Then[6]

$$\zeta_K(s) = \prod_{\chi \in X} L(\chi, s).$$

Finally, we will use the following facts about cyclotomic fields without proof[7]:

Theorem 5.2.3. *For a prime $p > 2$, let us consider the cyclotomic extension $K = \mathbb{Q}(\zeta_p)$ and its maximal totally real subfield $K^+ := \mathbb{Q}(\zeta_p + \zeta_p^{-1})$. We have:*

(a) The class number h_{K^+} of K^+ divides the class number h_K of K.

[5] i.e., $\chi(-1) = -1$

[6] Here, we view each $\chi \in X$ as a *primitive* Dirichlet character of a certain conductor $d \mid D$.

[7] Some of these properties will be discussed in my lecture about 'Iwasawa theory and cyclotomic fields' in the winter term

(b) The regulators of K respectively K^+ satisfy:

$$\frac{\mathrm{Reg}_K}{\mathrm{Reg}_{K^+}} = 2^{\frac{p-3}{2}}.$$

(c) $p \mid \frac{h_K}{h_{K^+}}$ if and only if $p \mid h_K$.

(d) $|d_K| = p^{p-2}$ and $|d_{K^+}| = p^{\frac{p-3}{2}}$.

Proof. We will treat some of these results in the winter term. □

We are now well-prepared for the proof of the following Theorem:

Theorem 5.2.4 (Kummer criterion)**.** *Let p be an odd prime and consider the class group h_K of the cyclotomic field $K = \mathbb{Q}(\zeta_p)$. Then $p \mid h_K$ if and only if p divides one of the zeta values*

$$\zeta(1-2n) = -\frac{B_{2n}}{2n},$$

for $2 \leq 2n < p-1$.

Proof. Let us write $K = \mathbb{Q}(\zeta_p)$ and $K^+ = \mathbb{Q}(\zeta_p + \zeta_p^{-1})$. The idea is to compare the following two explicit formulas for $\zeta_K(s)/\zeta_{K^+}(s)$ at $s=1$: The first one comes from the analytic class number formula

$$\frac{\zeta_K(s)}{\zeta_{K^+}(s)}\Big|_{s=1} = \pi^{\frac{p-1}{2}} \frac{h_K}{h_{K^+}} \frac{\mathrm{Reg}_K}{\mathrm{Reg}_{K^+}} \sqrt{\frac{|d_K|}{|d_{K^+}|}} \frac{e_K}{e_{K^+}}.$$

The second formula comes from the explicit decomposition into Dirichlet L-functions and the explicit formula for these values:

$$\frac{\zeta_K(s)}{\zeta_{K^+}(s)}\Big|_{s=1} = \prod_{\chi(-1)=-1} L(\chi,1) = (\pi i)^{\frac{p-1}{2}} \prod_{\chi(-1)=-1} \frac{\mathcal{G}(\chi)}{p} B_{1,\overline{\chi}},$$

where χ runs over the odd Dirichlet characters of conductor p. By taking absolute values and comparing both equations, we obtain

$$\frac{h_K}{h_{K^+}} = \frac{|\mathrm{Reg}_{K^+}|}{|\mathrm{Reg}_K|} \sqrt{\frac{|d_{K^+}|}{|d_K|}} \frac{e_{K^+}}{e_K} \prod_{\chi(-1)=-1} \frac{|\mathcal{G}(\chi)|}{p} |B_{1,\overline{\chi}}|.$$

In Theorem 4.6.2, we have already seen that $|\mathcal{G}(\chi)| = \sqrt{p}$. The field K^+ is totally real, so it only contains the roots of unity $\{\pm 1\}$ and we get $e_{K^+} = 2$. The field K contains the roots of unity $\{\pm\zeta_p \mid \zeta_p \in \mu_p(\overline{\mathbb{Q}})\}$ and hence $e_K = 2p$. By Theorem 5.2.3, we have the following equations

$$\frac{\mathrm{Reg}_K}{\mathrm{Reg}_{K^+}} = 2^{\frac{p-3}{2}}, \quad \frac{|d_K|}{|d_{K^+}|} = p^{\frac{p-1}{2}}.$$

Combining this with the above equation for $\frac{h_K}{h_{K^+}}$ gives

$$\frac{h_K}{h_{K^+}} = 2^{-\frac{p-3}{2}} p \prod_{\chi(-1)=-1} |B_{1,\overline{\chi}}|.$$

Now, observe that ω is an odd Dirichlet character of conductor p and induces an isomorphism of cyclic groups $(\mathbb{Z}/p\mathbb{Z})^\times \cong \mu_{p-1}(\mathbb{C})$. Thus, the odd Dirichlet characters of conductor p are exactly the odd powers $\omega, \omega^3, \dots, \omega^{p-2}$. Furthermore, by Theorem 5.1.4, we have for $k+1 \not\equiv 0 \mod p-1$ the congruence

$$B_{1,\omega^k} \equiv B_{1+k} \mod \mathfrak{p},$$

and for $k = p-2$ we get $|pB_{1,\omega^{p-2}}| \equiv 1 \mod \mathfrak{p}$. This implies

$$\frac{h_K}{h_{K^+}} \equiv 2^{-\frac{p-3}{2}} \prod_{j=1}^{\frac{p-3}{2}} |B_{2j}| \mod p.$$

So $\frac{h_K}{h_{K^+}}$ is divisible by p if and only if p divides at least one B_{2j} for $1 \leq j \leq \frac{p-3}{2}$. On the other hand, by Theorem 5.2.3 $\frac{h_K}{h_{K^+}}$ is an integer and it is divisible by p if and only if h_K is divisible by p. Thus, we have shown that p divides h_K if and only if p divides one of the Bernoulli numbers

$$B_2, B_4, \dots, B_{p-3}.$$

The Theorem follows now from the explicit formula for the values of the Riemann zeta function

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

observing that p does not divide any of the denominators $2, 4, \dots, p-3$. □

If you ever tried to compute the class number of a number field by hands you will certainly know to appreciate Kummer's criterion. For example, we have $B_{12} = -\frac{691}{2730}$ so the prime $p = 691$ divides B_12 and we deduce that the class number of $\mathbb{Q}(\zeta_{691})$ is divisible by 691.

Outlook

As often in mathematics, the answer to a mathematical problem usually raises many new questions. For example, Kummer's criterion immediately raises many questions about the relation of cyclotomic fields and p-adic properties of zeta values:

- Can we generalize Kummer's criterion to cyclotomic fields of the form $\mathbb{Q}(\zeta_{p^n})$?
- What does it mean in terms of class groups that p divides a particular zeta value $\zeta(2k)$?
- We have seen that $p \mid \frac{h_K}{h_{K^+}}$ is equivalent to $p \mid h_K$. Is it possible that $p \mid h_{K^+}$?

- Are there infinitely many primes such that p does (not) divide the class group of $\mathbb{Q}(\zeta_p)$?
- The Kummer congruences allow us to interpolate the Riemann zeta function p-adically. What is the relation between the p-adic zeta function and the p-divisibility of class groups of cyclotomic fields?
- and many more ...

Some of these questions have surprising answers and some of them are even unsolved conjectures. So Kummer's criterion can be seen as the tip of the iceberg of many deep relations between p-adic L-functions and class groups in cyclotomic extensions[8]. This area of number theory is usually summarized under the term Iwasawa theory. If you are interested to learn more about the p-adic aspect of L-functions and their arithmetic interpretation then I would like to cordially invite you to my lecture on Iwasawa theory in the winter semester.

[8] Or more generally $\mathbb{Z}_p$-extensions

6 *Epilogue*

I hope that I could give you some ideas about various different aspects of L-functions. Of course, this area of mathematics is far too extensive to ever fully understand everything about L-functions. And of course, there are also plenty of aspects of L-functions we have never said a word about in this lecture. In some sense, we have only touched certain particular parts of a baby L-ephant. Nevertheless, I hope that I could give you an impression about the existence of many different facets of L-functions.

Acknowledgement

I would like to thank all participants of this course. Special thanks go to Manuel Hoff for the excellent organization and presentation of the exercises. Furthermore, I would also like to thank him for many valuable corrections and suggestions. Many thanks also to Florian Leptien, Federico Mocchetti, Niklas Müller and Reinier Sorgdrager for pointing out many typos and corrections in these notes and for all the interesting contributions in the weekly question hour.

Bibliography

Eberhard Freitag and Rolf Busam. *Funktionentheorie*. Springer-Verlag, Berlin, 1993. ISBN 3-540-50618-7.

Loukas Grafakos. *Classical Fourier analysis*, volume 249 of *Graduate Texts in Mathematics*. Springer, New York, second edition, 2008. ISBN 978-0-387-09431-1.

E. Hewitt and K. A. Ross. *Abstract harmonic analysis. Vol. I*, volume 115 of *Grundlehren der Mathematischen Wissenschaften*. Springer-Verlag, Berlin-New York, 2nd edition, 1979. ISBN 3-540-09434-2.

James S. Milne. Fields and galois theory (v4.61), 2020. Available at www.jmilne.org/math/.

Jürgen Neukirch. *Algebraic number theory*, volume 322 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1999. ISBN 3-540-65399-6. DOI: 10.1007/978-3-662-03983-0. URL https://doi.org/10.1007/978-3-662-03983-0. Translated from the 1992 German original and with a note by Norbert Schappacher, With a foreword by G. Harder.

D. Ramakrishnan and R. J. Valenza. *Fourier analysis on number fields*, volume 186 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1999. ISBN 0-387-98436-4. DOI: 10.1007/978-1-4757-3085-2. URL https://doi.org/10.1007/978-1-4757-3085-2.

L. Ribes and P. Zalesskii. *Profinite Groups*. A Series of Modern Surveys in Mathematics. Springer Berlin Heidelberg, 2010. ISBN 9783642016424.

Jean-Pierre Serre. *Linear representations of finite groups*. Graduate Texts in Mathematics, Vol. 42. Springer-Verlag, New York-Heidelberg, 1977. ISBN 0-387-90190-6. Translated from the second French edition by Leonard L. Scott.

D. Zagier. Newman's short proof of the prime number theorem. *Amer. Math. Monthly*, 104(8):705–708, 1997. ISSN 0002-9890. DOI: 10.2307/2975232. URL https://doi.org/10.2307/2975232.