

Joachim Wehler

Elliptic Curves and Modular Forms

Release 2.0

October 21, 2026

© Joachim Wehler, 2020-2026

I prepared these notes for the participants of the lectures. The lecture took place at the mathematical department of LMU (Ludwig-Maximilians-Universität) at Munich. The first time during the winter term 2017/18 and then during the winter term 2020/21. Some parts of the notes rely on notes of an unpublished lecture of O. Forster.

Compared to the oral lecture in class these written notes contain some additional material.

I thank all participants - in particular W. Hensgen - for pointing out some errors and for their proposals for improvement. Please report to

wehler@math.lmu.de

any further errors or typos, adding the version of the lecture notes.

Release notes:

- *Release 2.0*: Revised edition.

Contents

Introduction	1
PARI files	3
Part I General Theory	
1 Elliptic functions	9
1.1 The field of elliptic functions	9
1.2 The Weierstrass $\wp$ -function	14
1.3 Abel's theorem	24
2 The modular group $SL(2, \mathbb{Z})$ and Hecke congruence subgroups $\Gamma_0(N)$	31
2.1 The moduli space of complex tori and group actions	31
2.2 Topology of the orbit space of the $SL(2, \mathbb{Z})$ -action	47
2.3 The modular curves $X(\Gamma_0(N))$ as compact Riemann surfaces	66
3 The algebra of modular forms	85
3.1 Modular forms and cusp forms	86
3.2 Eisenstein series generate the algebra of modular forms	94
3.3 Generalization to Hecke congruence subgroups $\Gamma_0(N)$	121
4 Elliptic curves	129
4.1 Embedding complex tori as plane cubic hypersurfaces in $\mathbb{P}^2$	129
4.2 Elliptic curves over an arbitrary field	143
4.3 Elliptic curves over finite fields	173
Part II Advanced Theory	
5 Introduction to Hecke theory and applications	187
5.1 Hecke operators of the modular group and their eigenforms	187
5.2 Petersson scalar product and level change of modular forms	205
5.3 Numerology: Lagrange, Jacobi, Ramanujan, Mordell	224

6	Modular functions and imaginary quadratic fields	235
6.1	Complex multiplication	235
6.2	Modular polynomials	243
6.3	The class number	261
7	Outlook: Modular elliptic curves, monstrous moonshine	275
7.1	The modularity theorem	275
7.2	Monstrous moonshine	282
	List of results	289
	References	293
	Index	297

Introduction

Modular Forms and elliptic curves are a classical domain from mathematics. At least, since the proof of Fermat's last conjecture the domain attracts widespread attention. The domain of Modular Forms integrates at least the three mathematical disciplines Complex Analysis, Algebraic Geometry, and Algebraic Number Theory.

Elliptic curves can be investigated by different mathematical methods.

- Algebraic geometry: Elliptic curves are the zero-sets of polynomials.
- Complex analytic geometry: When considered as Riemann surfaces then elliptic curves are complex tori.
- Arithmetic geometry: Elliptic curves can be defined over different fields, e.g. over $\mathbb{Q}$ or $\mathbb{F}_p$ and over the ring $\mathbb{Z}$.

The algebraic point of view identifies complex elliptic curves with smooth cubic hypersurfaces of $\mathbb{P}^2$. Hence elliptic curves are the first ones in the series of cubic hypersurfaces in complex projective space $\mathbb{P}^n$, $n \geq 2$. The higher dimensional cubics are also challenging examples, see [31, Chap.V, §4], [33].

Chapter 1 introduces elliptic functions as doubly periodic, meromorphic functions in the complex plane $\mathbb{C}$. The period group is a lattice $\Lambda \subset \mathbb{C}$. The main tool to study elliptic function is the residue theorem.

Chapter 2 characterizes elliptic functions as meromorphic functions on the torus $\mathbb{C}/\Lambda$. The classes of biholomorphically equivalent complex tori are the orbits of a group action of the modular group

$$SL(2, \mathbb{Z}) \times \mathbb{H} \rightarrow \mathbb{H}$$

The quotients

$$\Gamma_0(N) \backslash \mathbb{H}$$

of the restricted action of the Hecke congruence subgroups $\Gamma_0(N) \subset SL(2, \mathbb{Z})$ are Riemann surfaces with compactifications $X_0(N)$.

Chapter 3 introduces modular forms and cusp forms as holomorphic functions on $\mathbb{H} \cup \{\infty\}$ with a certain transformation behaviour with respect to the $\Gamma_0(N)$ -action. These functions unveil themselves as meromorphic differential forms on the modular curve $X_0(N)$. The Riemann-Roch theorem computes the dimensions of the vector spaces of modular forms and cusp forms.

Chapter 4 makes a new start from the viewpoint of algebraic geometry: Complex tori embed as elliptic curves into the complex projective plane. They can be represented as non-singular cubic curves, defined as zero set of a Weierstrass polynomial. The main tool to represent a complex torus as an elliptic curve is the Weierstrass $\wp$ -function and its differential equation. Conversely, each complex elliptic curve arises as embedding of a complex torus.

Choosing different fields for the coefficients of the Weierstrass polynomials allows to consider elliptic curves defined over $\mathbb{C}, \mathbb{Q}, \mathbb{Z}$ or even over the finite fields $\mathbb{F}_p$. The focus of the chapter are some relations between complex analytic geometry, algebraic geometry and arithmetic geometry.

Chapter 5 considers families of Hecke operators, which are linear endomorphisms on the finite dimensional vector spaces of modular forms. Each family of Hecke operators acts on a given vector space of modular forms of fixed weight. On the corresponding subspace of cusp forms the family can be diagonalized simultaneously, a result which relies on the Petersson scalar product. As an application the chapter proves some theorems from algebraic number theory related to Lagrange, Jacobi, Ramanujan, Mordell and other mathematicians.

In the second, more advanced part of the lecture notes Chapter 6 deals with deeper applications of the theory of modular forms to algebraic number theory. The chapter investigates the relation between tori with complex multiplication and imaginary quadratic number fields. The main role is played by the modular invariant j . As an application the chapter proves a lower bound of the class number of imaginary quadratic number fields.

The final Chapter 7 gives an outlook to the modularity theorem for elliptic curves, which plays the dominant role in Wiles' proof of the Fermat conjecture. A second section gives an outlook to the moonshine relation between the Fourier coefficients of the modular j -invariant and the dimensions of the irreducible representations of the monster group, which at day culminates in the work of Borcherds.

It is also the aim of these lecture notes to illustrate their results by a series of numerical calculations using a computer algebra system.

PARI files

Chapter 1:

- `Weierstrass_p_function_08`: Laurent expansion of the Weierstrass $\wp$ -function and its derivative, cf. Remark 1.14.

Chapter 2

- `modular_curve_as_covering`: Genus, degree and further parameters of modular curve $X_0(N)$, cf. Remark 2.33.

Chapter 3

- `Congruence_subgroup_19`: Dimension of $M_k(\Gamma_0(p^n))$ via Riemann-Roch, cf. Remark 3.27.

Chapter 4

- `Torus_to>Weierstrass_equation_01`: The plane cubic of an embedded torus, cf. Example 4.4
- `Elliptic_curve_plot_06`: Plot of some plane cubics, cf. Example 4.28.
- `Elliptic_curve_02`: Invariants of some plane cubics, cf. Example 4.28.
- `Elliptic_curve_weierstrass_equation_12`: Global minimal Weierstrass polynomial and its discriminant, cf. Example 4.38.
- `Elliptic_curve_Hasse_estimate_05`: Hasse estimate for number of $\mathbb{F}_p$ -rational points, cf. Remark 4.41.
- `Elliptic_curve_04_02`: Analytic rank of elliptic curves, cf. Remark 4.46.

Chapter 5

- `Hecke_matrix_02`: Hecke operator $T_2 \in \text{End}(S_{28}(\Gamma_1))$, cf. Example 5.15.
- `Congruence_subgroup_20`: Oldforms and newforms for different $\Gamma_0(N)$, cf. Example 5.30.
- `Modular_forms_theta_01`: 4-squares theorem and similar decompositions, cf. Example 5.38.

Chapter 6

- `modular_polynomial_01`: Modular polynomials of different levels, cf. Example 6.12.
- `modular_polynomial_02`: The polynomials Φ_m and their factorization, cf. Example 6.14.

Chapter 7

- `Elliptic_curve_taniyama_10`: Illustration of Wiles' theorem by some numerical examples, cf. Example 7.8.

The PARI-files of these lecture notes are contained in a separate folder, see the link at my homepage. It is recommended to download all files together, because some of them will call other files. After downloading the PARI files into the working directory of PARI a given file can be called and executed by the command

```
\r filename
```

For more information see the PARI manual from the PARI homepage.

Part I
General Theory

The context of Chapter 1 is complex analysis on domains in the complex plane. The base field is the field $\mathbb{C}$ of complex numbers. We develop the theory of elliptic functions, i.e. doubly periodic meromorphic functions. Classical tools are Laurent series and the residue theorem for meromorphic functions. The period group of doubly periodic functions are lattices in $\mathbb{C}$.

The view point of Chapter 2 are compact Riemann surfaces. We continue working with the base field $\mathbb{C}$, but now the domain is a complex manifold. The most simple example of a compact Riemann surface is the Riemann sphere $\mathbb{P}^1(\mathbb{C})$. Meromorphic functions are holomorphic maps into the Riemann sphere. A second class of compact Riemann surfaces are complex tori. The doubly periodic meromorphic functions from Chapter 1 are meromorphic functions on a complex torus. The main tools from the theory of Riemann surfaces are line bundles, divisors and the Theorem of Riemann-Roch. The set of isomorphism classes of complex tori turns out to be the orbit space of the modular group $SL(2, \mathbb{Z})$ acting on the upper half plane $\mathbb{H}$. We show that the compactified orbit space is biholomorphic equivalent to $\mathbb{P}^1(\mathbb{C})$.

Chapter 4 switches from the analytic theory to an algebraic context. We consider the Riemann sphere from an algebraic point of view and study non-singular curves C in the projective space $\mathbb{P}^1(\mathbb{C})$, now equipped with the Zariski topology. These curves are the zero set of one or more homogeneous polynomials. The smallest subfield $k \subset \mathbb{C}$ which comprises their coefficients is the field of definition of C . As a consequence we extend the range of base fields under consideration to the field $\mathbb{Q}$ and to the residue fields $\mathbb{F}_{p^n}$ of $\mathbb{Z}$. The Weierstrass $\wp$ -function and its derivative $\wp'$ from Chapter 1 embed a given torus into $\mathbb{P}^2(\mathbb{C})$. The image is an elliptic curve E . It is the zero set of a Weierstrass polynomial F , a homogeneous polynomial of degree 3. If all coefficients of F are rational and if in addition a $\mathbb{Q}$ -valued point of E exists, then E is a rational elliptic curve. The existence of a rational point $O \in E(\mathbb{Q})$ allows to equip $E(\mathbb{Q})$ with an additive group structure having O as neutral element. Rational elliptic curves have a Weierstrass polynomial F with integer coefficients. We study the reduction E_p of E modulo the primes $p \in \mathbb{Z}$. The sequence $(\text{card } E_p)_{p \text{ prime}}$ encodes important arithmetic properties of a rational elliptic curve.

Chapter 3 takes up the study of the modular group from Chapter 2. We extend the operation of $\Gamma := SL(2, \mathbb{Z})$ on $\mathbb{H}$ to an operation on the vector space of holomorphic functions on $\mathbb{H}$. Those functions which are symmetric with respect to this operation are named resp. modular forms and cusp forms. All of them develop into a Fourier series around the point ∞ . Due to their symmetry modular forms correspond bijectively to meromorphic sections of the line bundle Ω^1 and its tensor products defined on the compactified orbit space $\mathbb{P}^1(\mathbb{C})$. The complex vector spaces of sections are finite dimensional, their dimension can be computed by the theorem of Riemann-Roch. Taking the tensor product as multiplication provides the direct sum of these vector spaces with the structure of a finitely generated graded algebra forms $M_*(\Gamma) = \bigoplus_k M_k(\Gamma)$. The grading is given by the weight of the modular forms. The algebra of modular forms $M_*(\Gamma)$ contains the ideal of cusp forms $S_*(\Gamma) = \bigoplus_k S_k(\Gamma)$. On $M_*(\Gamma)$ acts the family of Hecke operators as an Abelian algebra of endomorphisms. The ideal $S_*(\Gamma)$ is stable with respect to this

action. Each component $S_k(\Gamma)$ has a basis of eigenforms with respect to the Hecke algebra. For each eigenform f the family of eigenvalues reveals remarkable relations from the arithmetic of the Fourier coefficients of f . These relations make up the *magic of modular forms*.

Chapter 1

Elliptic functions

The chapter deals with complex analysis in the plane.

1.1 The field of elliptic functions

The present section deals with complex analysis in the plane, i.e. we study holomorphic and meromorphic functions on domains in the complex plane $\mathbb{C}$.

The trigonometric functions *sin* and *cos* are examples of periodic holomorphic functions, while *tan* and *cot* are periodic meromorphic functions. All periods of these functions are integer multiples of respectively 2π and π . Elliptic functions are meromorphic with period group isomorphic to $\mathbb{Z} \oplus \mathbb{Z}$.

Recall that a meromorphic function on a domain $G \subset \mathbb{C}$ is a holomorphic function

$$f : G \setminus S \rightarrow \mathbb{C}$$

with a discrete closed subset $S \subset G$ of poles of f - but no essential singularities.

Definition 1.1 (Period of a meromorphic function). Consider a meromorphic function f on a domain G . A number $\omega \in \mathbb{C}$ is a *period* of f if

- for all $z \in G$ also $z \pm \omega \in G$
- and for all $z \in G \setminus P$

$$f(z + \omega) = f(z).$$

It only seems that Definition 1.1 excludes the poles from the definition of periodicity. A pole z_0 of f is an isolated singularity of f . Therefore the function f expands into a convergent Laurent series around z_0

$$f(z) = \sum_{n \geq n_0}^{\infty} a_n \cdot (z - z_0)^n$$

with coefficients

$$a_n = \frac{1}{2\pi i} \int_{|z-z_0|=\varepsilon} \frac{f(z)}{(z-z_0)^{n+1}} dz \text{ for suitable } \varepsilon > 0.$$

If f has the period ω then the Laurent expansions of f around z_0 and $z_0 + \omega$ are the same because the integrands attain the same value at z and $z + \omega$. Hence periodicity of f includes also the poles.

Example 1.2 (Periodic meromorphic functions).

1. The functions $\sin, \cos : \mathbb{C} \rightarrow \mathbb{C}$ are holomorphic with period $\omega = 2\pi$.
2. The tangent-function is meromorphic on $\mathbb{C}$ with singularities $(\pi/2) + k\pi$, $k \in \mathbb{Z}$, and period $\omega = \pi$.
3. The exponential function $\mathbb{C} \rightarrow \mathbb{C}^*$, $z \mapsto e^{2\pi iz}$, is holomorphic with period $\omega = 1$.

If a meromorphic function f has a period $\omega \neq 0$ one can ask for the set of all periods. Apparently elements of the form $k\omega$, $k \in \mathbb{Z}$ are also periods of f . How many independent periods of f exist?

Definition 1.3 (Discrete subgroup). A subgroup $\Lambda \subset (\mathbb{C}, +)$ is a *discrete subgroup*, if the subspace topology of Λ is discrete, i.e. if any $g \in \Lambda$ has an open neighbourhood $U \subset \mathbb{C}$ with

$$U \cap \Lambda = \{g\}.$$

Proposition 1.4 (Period group). Consider a non-constant meromorphic function f on a domain $G \subset \mathbb{C}$. The set Λ_f of all periods of f is a discrete subgroup of $\mathbb{C}$. It is named the period group of f .

Proof. i) Apparently $\Lambda_f \subset \mathbb{C}$ is a subgroup of $\mathbb{C}$.

ii) If Λ_f were not discrete then a convergent sequence $(\omega_n)_{n \in \mathbb{N}}$ of pairwise distinct periods $\omega_n \in \Lambda_f$ would exist.

Hence for an arbitrary but fixed $z_0 \in G$, which is not a pole of f , and for all $n \in \mathbb{N}$

$$f(z_0) = f(z_0 + \omega_n).$$

The identity theorem implies that f is constant, a contradiction. $\square$

Proposition 1.5 (Discrete subgroups of $(\mathbb{C}, +)$). *Each discrete subgroup*

$$\Lambda \subset (\mathbb{C}, +)$$

belongs to one of the following types:

1. Rank = 0:

$$\Lambda = \{0\}$$

2. Rank = 1:

$$\Lambda = \mathbb{Z}\omega = \{m \cdot \omega : m \in \mathbb{Z}\}$$

with a suitable element $\omega \in \mathbb{C}^$.*

3. Rank = 2:

$$\Lambda = \mathbb{Z}\omega_1 + \mathbb{Z}\omega_2 = \{m_1 \cdot \omega_1 + m_2 \cdot \omega_2 : m_1, m_2 \in \mathbb{Z}\}$$

with $\omega_1, \omega_2 \in \mathbb{C}$ linearly independent over the base field $\mathbb{R}$. The subgroup Λ is named a lattice.

For a proof see [1, Chap. 7, Sect. 2.1].

Note. If we specify a lattice in the form

$$\Lambda = \mathbb{Z}\omega_1 + \mathbb{Z}\omega_2$$

we will assume that the basis (ω_1, ω_2) is *positively oriented*, i.e.

$$\det \begin{pmatrix} \operatorname{Re} \omega_1 & \operatorname{Re} \omega_2 \\ \operatorname{Im} \omega_1 & \operatorname{Im} \omega_2 \end{pmatrix} > 0.$$

Definition 1.6 (Elliptic function). Consider a non-constant meromorphic function f on a domain $G \subset \mathbb{C}$ with period group Λ_f .

- If $\operatorname{rank} \Lambda_f \geq 1$ then f is named *periodic*.
- If $\operatorname{rank} \Lambda_f = 2$ then f is named *doubly periodic*. Its period group

$$\Lambda_f = \mathbb{Z}\omega_1 + \mathbb{Z}\omega_2$$

is also named the *period lattice* of f , see Figure 1.1. The subset

$$P = \{\lambda_1 \omega_1 + \lambda_2 \omega_2 : 0 \leq \lambda_1, \lambda_2 < 1\}$$

is the *fundamental period parallelogram* of f with respect to the basis (ω_1, ω_2) .

A doubly periodic meromorphic function on $\mathbb{C}$ is an *elliptic function* with respect to a lattice Λ if

$$\Lambda \subset \Lambda_f$$

i.e. if all lattice points $\omega \in \Lambda$ are periods of f .

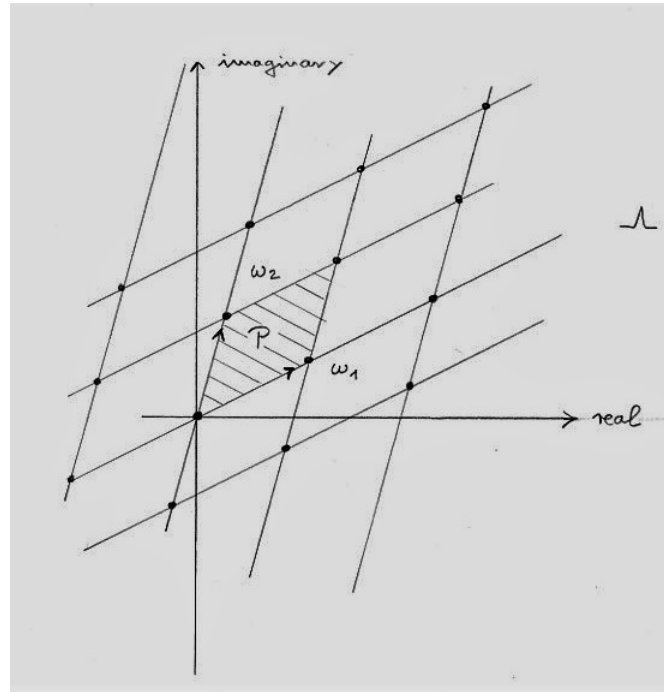


Fig. 1.1 Period lattice Λ

Which elliptic functions do exist? Proposition 1.7 shows that it is not interesting to study the subclass of holomorphic elliptic functions. Therefore we will focus on meromorphic functions with poles. See Proposition 1.8, and its corollaries about the value attainment of elliptic functions.

Proposition 1.7 (Constant elliptic functions). *Any holomorphic elliptic function*

$$f : \mathbb{C} \rightarrow \mathbb{C}$$

is constant.

Proof. An entire function f is holomorphic and attains the maximum of its modulus on its closed period parallelogram which is a compact set. Being a bounded entire function, f is constant according to Liouville's theorem. $\square$

Proposition 1.8 (Residue theorem for elliptic functions). *An elliptic function f with period parallelogram P satisfies*

$$\sum_{\zeta \in P} \text{res}_{\zeta}(f) = 0.$$

Proof. The boundary ∂P comprises only finitely many poles of f . Therefore we can choose a number $a \in \mathbb{C}$ such the boundary of the translated period parallelogram

$$P_a := a + P$$

contains no poles of f . The residue theorem implies

$$\frac{1}{2\pi i} \int_{\partial P_a} f(z) dz = \sum_{\zeta \in P_a} \text{res}_{\zeta}(f) = \sum_{\zeta \in P} \text{res}_{\zeta}(f).$$

The integrand on the left-hand side is doubly periodic. Therefore the integration along opposite sides of the period parallelogram cancels and the whole integral vanishes. $\square$

Corollary 1.9 (Poles and zeros of elliptic functions).

1. *There are no elliptic functions f with one pole of order = 1 modulo the period lattice Λ_f and no other pole mod Λ_f .*
2. *There are no elliptic functions f with only one zero of order = 1 modulo the period lattice Λ_f and no other zero mod Λ_f .*

Proof. 1. A pole of order 1 at $a \in \mathbb{C}$ has $\text{res}_a(f) \neq 0$, which contradicts Proposition 1.8.

2. Apply part 1) to the elliptic function $1/f$. $\square$

Corollary 1.10 (Counting poles and zeros of elliptic functions). *Any non-constant elliptic function f attains modulo each lattice*

$$\Lambda \subset \Lambda_f$$

all values $a \in \mathbb{C} \cup \{\infty\}$ with the same multiplicity. In particular, f has mod Λ the same number of poles and zeros taken with multiplicity.

Proof. The case $a \neq \infty$ reduces to the second claim by considering the function $f - a$. It has the same poles as f . Therefore it suffices to show that f has the same number of zeros and poles, taken with multiplicity. Moreover it suffices to prove the theorem for $\Lambda = \Lambda_f$.

We apply Proposition 1.8 to the meromorphic function f'/f and obtain

$$\sum_{\zeta \in P} \operatorname{res}_{\zeta} \left(\frac{f'}{f} \right) = 0.$$

Each residue evaluates to

$$\operatorname{res}_{\zeta} \left(\frac{f'}{f} \right) = \begin{cases} k & \text{f has a zero of order } k \text{ at } \zeta \\ -k & \text{f has a pole of order } k \text{ at } \zeta \end{cases}$$

□

1.2 The Weierstrass $\wp$ -function

In the present section we consider an arbitrary but fixed lattice

$$\Lambda := \mathbb{Z} \omega_1 + \mathbb{Z} \omega_2.$$

We use the notation $\Lambda' := \Lambda \setminus \{0\}$. Moreover P denotes the fundamental period parallelogram of Λ .

The $\wp$ -function of Λ is a distinguished elliptic function: All elliptic functions of the given lattice Λ derive from $\wp$ and its derivative $\wp'$.

Lemma 1.11 (Lattice constants). *For $k > 2$ the infinite series*

$$G_{\Lambda,k} := \sum_{\omega \in \Lambda'} \frac{1}{\omega^k}$$

is absolutely convergent. Its value is named the lattice constant of Λ .

Proof. We choose the exhaustion of Λ by the sequence $(\Lambda_n)_{n \in \mathbb{N}}$ of disjoint sets of indices of increasing modulus

$$\Lambda_n := \{\mu \cdot \omega_1 + \nu \cdot \omega_2 \in \Lambda : \mu, \nu \in \mathbb{Z}; |\mu|, |\nu| \leq n \text{ and } (|\mu| = n \text{ or } |\nu| = n)\},$$

i.e.

$$\Lambda = \bigcup_{n \in \mathbb{N}} \Lambda_n.$$

Then

$$\text{card } \Lambda_n = 8n.$$

A suitable constant c exists with $|\omega| \geq c \cdot n$ for all $n \in \mathbb{N}$ and for all $\omega \in \Lambda_n$. Therefore

$$\sum_{\omega \in \Lambda_n} \frac{1}{|\omega|^k} \leq \frac{8 \cdot n}{c^k \cdot n^k} \leq \left(\frac{8}{c^k} \right) \cdot \frac{1}{n^{k-1}}.$$

For $k > 2$

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

Therefore the claim follows from

$$\sum_{\omega \in \Lambda'} \frac{1}{\omega^k} = \sum_{n=1}^{\infty} \left(\sum_{\omega \in \Lambda_n} \frac{1}{\omega^k} \right)$$

□

We will study elliptic functions with period lattice Λ . According to Corollary 1.9 a candidate f for the most simple example would have only one pole mod Λ , the pole having order = 2 and residue = 0. The Laurent expansion of f around $z_0 = 0$ would start

$$f(z) = \frac{1}{z^2} + a_1 \cdot z + O(2).$$

Being doubly periodic, f would have poles exactly at the points $\omega \in \Lambda$. The following Theorem 1.12 shows that such a function actually exists.

Theorem 1.12 (Weierstrass $\wp$ -function). *For each lattice Λ the series*

$$\wp(z) := \frac{1}{z^2} + \sum_{\omega \in \Lambda'} \left(\frac{1}{(z - \omega)^2} - \frac{1}{\omega^2} \right)$$

is absolutely and compactly convergent for all $z \in \mathbb{C} \setminus \Lambda$. It defines an even elliptic function $\wp$ with respect to Λ , named the Weierstrass $\wp$ -function of Λ . The pole set of $\wp$ is Λ , each pole has order = 2.

If we consider only the first summand with fixed z then asymptotically

$$\frac{1}{(z - \omega)^2} \sim \frac{1}{\omega^2}.$$

Therefore the series with these summands alone does not converge. We will show that the difference

$$\frac{1}{(z - \omega)^2} - \frac{1}{\omega^2}$$

is proportional to

$$\frac{1}{\omega^3}.$$

Therefore the series $\wp(z)$ converges.

Note: TeX reserves the separate symbol “backslash wp” to denote the Weierstrass $\wp$ -function.

Proof. We recall the definition of compact convergence of a sequence $(f_v)_{v \in \mathbb{N}}$ of meromorphic functions on a domain $G \subset \mathbb{C}$, cf. [68]: For each compact subset $K \subset G$ exists an index $v_0 \in \mathbb{N}$ such that

- for all $v > v_0$ the function f_v has no pole in K
- the sequence $(f_v)_{v > v_0}$ is uniformly convergent on K .

i) *Meromorphic with pole set Λ* : For each compact set $K \subset \mathbb{C}$ only finitely many summands of $\wp$ have a pole in K . We show that the remaining series converges absolutely and uniform on K : Choose an arbitrary but fixed $R > 0$. For

$$|z| \leq R \text{ and } 2R \leq |\omega|$$

holds

$$\frac{|\omega|}{2} - |z| \geq 0.$$

The triangle inequality

$$|z - \omega| = |\omega - z| \geq ||\omega| - |z|| = \left| \frac{|\omega|}{2} - |z| + \frac{|\omega|}{2} \right| \geq \frac{|\omega|}{2}$$

implies the estimate

$$\begin{aligned} \left| \frac{1}{(z - \omega)^2} - \frac{1}{\omega^2} \right| &= \left| \frac{\omega^2 - (z - \omega)^2}{\omega^2 \cdot (z - \omega)^2} \right| \leq \frac{|-\omega^2 + 2z\omega|}{|\omega|^2 \cdot \left(\frac{|\omega|}{2} \right)^2} \leq 4 \cdot \frac{|z|^2 + 2 \cdot |z| \cdot |\omega|}{|\omega|^4} \leq \\ &\leq 4 \cdot \frac{|z|^2}{|\omega|^4} + 8 \cdot \frac{|z|}{|\omega|^3} \leq \frac{4 \cdot R^2}{2 \cdot R \cdot |\omega|^3} + 8 \cdot \frac{R}{|\omega|^3} = 10 \cdot \frac{R}{|\omega|^3} \end{aligned}$$

independent from $|z| < R$.

According to Lemma 1.11 the series

$$\sum_{\omega \in \Lambda'} \frac{1}{\omega^3}$$

converges absolutely. For $|z| \leq R$ we decompose the series with respect to the summation over $\omega \in \Lambda$

$$\left[\frac{1}{z^2} + \sum_{\omega \in \Lambda', |\omega| < 2R} \left(\frac{1}{(z-\omega)^2} - \frac{1}{\omega^2} \right) \right] + \left[\sum_{\omega \in \Lambda, |\omega| \geq 2R} \left(\frac{1}{(z-\omega)^2} - \frac{1}{\omega^2} \right) \right]$$

The first summand is meromorphic with poles exactly at the points $z \in \Lambda \cap D_{2R}(0)$. The second summand converges absolutely and compactly on $D_{2R}(0)$ due to Lemma 1.11, and defines a holomorphic function on $D_{2R}(0)$.

Because R can be chosen arbitrary, the series $\wp$ is compact convergent and defines a meromorphic function on $\mathbb{C}$ with pole set Λ .

ii) *Even function with periods from Λ :*

- $\wp$ is even: It is admissible to rearrange the summation by replacing ω by $-\omega$, because the series is absolute convergent

$$\begin{aligned} \wp(-z) &= \frac{1}{z^2} + \sum_{\omega \in \Lambda'} \left(\frac{1}{(-z-\omega)^2} - \frac{1}{\omega^2} \right) = \frac{1}{z^2} + \sum_{\omega \in \Lambda'} \left(\frac{1}{(-z-(-\omega))^2} - \frac{1}{(-\omega)^2} \right) = \\ &= \frac{1}{z^2} + \sum_{\omega \in \Lambda'} \left(\frac{1}{(z-\omega)^2} - \frac{1}{\omega^2} \right) = \wp(z). \end{aligned}$$

- $\wp$ has periods from Λ : The derivative is

$$\wp'(z) = \frac{-2}{z^3} + \sum_{\omega \in \Lambda'} \frac{-2}{(z-\omega)^3} = -2 \sum_{\omega \in \Lambda} \frac{1}{(z-\omega)^3}$$

Apparently for each fixed $\omega_0 \in \Lambda$

$$\wp'(z + \omega_0) = \wp'(z)$$

after rearranging the summation by replacing ω by $\omega - \omega_0$. As a consequence

$$\wp(z + \omega_j) = \wp(z) + c_j$$

with two suitable constants $c_j \in \mathbb{C}$, $j = 1, 2$.

For $j = 1, 2$ choosing the specific arguments

$$z := -\omega_j/2,$$

and using that $\wp$ is even implies

$$\wp(\omega_j/2) = \wp((- \omega_j/2) + \omega_j) = \wp(-\omega_j/2) + c_j = \wp(\omega_j/2) + c_j$$

Hence

$$c_j = 0$$

□

Lemma 1.13 (Half-periods of a lattice). *Consider a lattice*

$$\Lambda = \mathbb{Z}\omega_1 + \mathbb{Z}\omega_2$$

and denote by

$$Z_{(\omega_1, \omega_2)} := \left\{ \frac{\omega_1}{2}, \frac{\omega_2}{2}, \frac{\omega_1 + \omega_2}{2} \right\}$$

the set of half-periods of its fundamental period parallelogram with respect to (ω_1, ω_2) . The derivative $\wp'$ of the Weierstrass function $\wp$ of Λ has mod Λ

- a pole of order = 3 at $0 \in \mathbb{C}$ and no other poles,
- and exactly three zeros, represented by the points of $Z_{(\omega_1, \omega_2)}$.

Proof. Consider $u \in Z_{(\omega_1, \omega_2)}$. According to Theorem 1.12 the function $\wp$ has no pole at u . The derivative $\wp'$ has the same pole set as $\wp$. Because $\wp$ is even, the function $\wp'$ is an odd elliptic function

$$\wp'(u) = -\wp'(-u) = -\wp'(-u + 2u) = -\wp'(u).$$

Therefore

$$\wp'(u) = 0,$$

i.e. each point of $Z_{(\omega_1, \omega_2)}$ is a zero of $\wp'$ within the fundamental period parallelogram. Hence $\wp'$ has at least three zeros mod Λ .

Moreover $\wp'$ like $\wp$ has a single pole at $0 \in \mathbb{C} \bmod \Lambda$. The pole of $\wp'$ has order = 3. Corollary 1.10 implies that $\wp'$ has exactly three zeros mod Λ . As a consequence, the points of $Z_{(\omega_1, \omega_2)}$ represent exactly the zeros of $\wp'$, and each zero has order = 1 □

Note that the determination of the two zeros mod Λ of $\wp$ is much more delicate [21].

Remark 1.14 (Weierstrass function $\wp$). Figure 1.2 shows the Laurent expansion around $0 \in \mathbb{C}$ of the Weierstrass function $\wp$ and its derivative $\wp'$ for the two lattices Λ

- with basis $(1, \rho = e^{2\pi i/3})$
- and basis $(1, i)$.

See PARI file “Weierstrass_p_function_08”. The numerical calculation confirms that $\wp$ is an even function and that $\wp'$ vanishes at the half-lattice points of Λ .

[illegible][illegible]

Fig. 1.2 Laurent expansion of $\wp$ and $\wp'$ and vanishing at half-lattice points

Theorem 1.12 shows that the $\wp$ -function of the lattice Λ has all points from Λ as periods. Corollary 1.15 shows the other direction: There are no additional periods, i.e. $\Lambda_\wp = \Lambda$.

Corollary 1.15 (Period lattice of $\wp$ and $\wp'$). *The Weierstrass function $\wp$ of the lattice Λ and its derivative $\wp'$ have the period lattice Λ , i.e.*

$$\Lambda = \Lambda_{\wp} = \Lambda_{\wp'},$$

i.e. the only periods of $\wp$ and of $\wp'$ are the points from Λ .

Proof. We know

$$\Lambda \subset \Lambda_{\wp} \subset \Lambda_{\wp'}$$

Here the first inclusion has been shown in Theorem 1.12, while the second inclusion is obvious. We show

$$\Lambda_{\wp'} \subset \Lambda :$$

Let $\omega \in \Lambda_{\wp'}$ be a period of $\wp'$. Then

$$\wp'(\omega/2) = \wp'((\omega/2) - \omega) = \wp'(-\omega/2) = -\wp'(\omega/2).$$

Either $\omega/2$ is not a pole of $\wp'$. Then $\omega/2$ is a zero of $\wp'$ and 1.13 implies $\omega/2 \in \Lambda/2$ or $\omega \in \Lambda$.

Or $\omega/2$ is a pole of $\wp'$. Then Theorem 1.12 implies $\omega/2 \in \Lambda$, and in particular $\omega \in \Lambda$. $\square$

Proposition 1.16 (Coefficients of the Laurent expansion of $\wp$). *The Weierstrass function $\wp$ of Λ has the Laurent expansion around zero*

$$\wp(z) = \frac{1}{z^2} + \sum_{k=1}^{\infty} a_{2k} \cdot z^{2k}$$

with coefficients derived from the lattice constants of Λ

$$a_{2k} = (2k+1) \cdot G_{\Lambda, 2k+2}.$$

Proof. We determine for $\omega \neq 0$ the Taylor series of

$$\frac{1}{(z-\omega)^2} - \frac{1}{\omega^2}$$

around $0 \in \mathbb{C}$: Taking the derivative

$$\frac{1}{(z-\omega)^2} = \frac{d}{dz} \left(\frac{-1}{z-\omega} \right)$$

and expanding into a geometric series

$$\frac{-1}{z-\omega} = \frac{1}{\omega-z} = \frac{1/\omega}{1-(z/\omega)} = (1/\omega) \cdot \sum_{v=0}^{\infty} (1/\omega)^v \cdot z^v$$

shows

$$\frac{1}{(z-\omega)^2} = (1/\omega) \cdot \sum_{v=1}^{\infty} v \cdot (1/\omega)^v \cdot z^{v-1} = \sum_{v=0}^{\infty} (v+1) \cdot \frac{1}{\omega^{v+2}} \cdot z^v$$

and

$$\frac{1}{(z-\omega)^2} - \frac{1}{\omega^2} = \sum_{v=1}^{\infty} (v+1) \frac{1}{\omega^{v+2}} \cdot z^v.$$

Because $\wp$ is even and the series converges absolutely, see Theorem 1.12, we obtain after rearrangement and using that $\wp$ is even

$$\begin{aligned}\wp(z) &= \frac{1}{z^2} + \sum_{\omega \in \Lambda'} \left(\frac{1}{(z-\omega)^2} - \frac{1}{\omega^2} \right) = \frac{1}{z^2} + \sum_{v=1}^{\infty} (v+1) \left(\sum_{\omega \in \Lambda'} \frac{1}{\omega^{v+2}} \right) z^v = \\ &= \frac{1}{z^2} + \sum_{k=1}^{\infty} (2k+1) \left(\sum_{\omega \in \Lambda'} \frac{1}{\omega^{2k+2}} \right) z^{2k}\end{aligned}$$

□

Theorem 1.17 (Differential equation of $\wp$). *The Weierstrass function $\wp$ of the lattice Λ satisfies the ordinary differential equation*

$$\wp'^2 = 4 \cdot \wp^3 - g_2 \cdot \wp - g_3$$

with the constants

$$g_2 := g_{\Lambda,2} := 60 \cdot G_{\Lambda,4}, \quad g_3 := g_{\Lambda,3} := 140 \cdot G_{\Lambda,6}$$

derived from the lattice constants.

Proof. To simplify the notation we omit the index Λ from the lattice constants $G_{\Lambda,k}$. Due to Proposition 1.16 the Laurent expansions start

$$\begin{aligned}\wp(z) &= \frac{1}{z^2} + \sum_{k \geq 1} (2k+1) \cdot G_{2k+2} \cdot z^{2k} = \frac{1}{z^2} + 3 \cdot G_4 \cdot z^2 + 5 \cdot G_6 \cdot z^4 + O(6) \\ \wp'(z) &= -\frac{2}{z^3} + 6 \cdot G_4 \cdot z + 20 \cdot G_6 \cdot z^3 + O(5) \\ \wp'(z)^2 &= \frac{4}{z^6} - \frac{24 \cdot G_4}{z^2} - 80 \cdot G_6 + O(2) \\ \wp(z)^2 &= \frac{1}{z^4} + 6 \cdot G_4 + 10 \cdot G_6 \cdot z^2 + O(4) \\ \wp(z)^3 &= \frac{1}{z^6} + \frac{6 \cdot G_4}{z^2} + 10 \cdot G_6 + \frac{3 \cdot G_4}{z^2} + 5 \cdot G_6 + O(2) = \frac{1}{z^6} + \frac{9 \cdot G_4}{z^2} + 15 \cdot G_6 + O(2)\end{aligned}$$

Therefore all summands of order < 1 of the sum

$$\wp'^2 - 4 \cdot \wp^3 + g_2 \cdot \wp + g_3$$

cancel and

$$\wp'^2 - 4 \cdot \wp^3 + g_2 \cdot \wp + g_3 = O(2).$$

The left-hand side is an elliptic function. The right-hand side shows that the function is even holomorphic and vanishes at $0 \in \mathbb{C}$. According to Proposition 1.7 the left-hand side is constant, i.e. zero, q.e.d.

Theorem 1.18 (Field of elliptic functions). *The field $\mathcal{M}(\Lambda)$ of elliptic functions with respect to the lattice Λ is a field extension of $\mathbb{C}$ of transcendence degree = 1, more precisely*

$$\mathcal{M}(\Lambda) = \mathbb{C}(\wp)[\wp']$$

with $\wp'$ algebraic over the field $\mathbb{C}(\wp)$. The minimal polynomial of $\wp'$ is quadratic

$$F(T) = T^2 - 4 \cdot \wp^3 + g_2 \cdot \wp + g_3 \in \mathbb{C}(\wp)[T]$$

with coefficients the lattice constants

$$g_2 := 60 \cdot G_{\Lambda,4}, \quad g_3 := 140 \cdot G_{\Lambda,6}.$$

Note the equality $\mathbb{C}(\wp)[\wp'] = \mathbb{C}(\wp)(\wp')$ because $\wp'$ is algebraic over $\mathbb{C}(\wp)$.

Proof. i) Due to Theorem 1.17 we have

$$F(\wp') = 0 \in \mathbb{C}(\wp).$$

The derivative $\wp'$ is odd while $\wp$ is even. Therefore $\wp' \notin \mathbb{C}(\wp)$, and the minimal polynomial F must have degree ≥ 2 .

ii) Consider an elliptic function $f \in \mathcal{M}(\Lambda)$, $f \neq 0$. We show $f \in \mathbb{C}(\wp, \wp')$.

- We first consider the case that f has pole set Λ .

The unique pole of $f \bmod \Lambda$ has order $k \geq 2$ due to Corollary 1.9. We prove

$$f \in \mathbb{C}(\wp, \wp')$$

by induction on the pole order k .

If $k = 2$ then for suitable $c \in \mathbb{C}$ the function

$$f - c \cdot \wp$$

has at most a single pole mod Λ of order < 2 . Therefore the function is holomorphic, and even constant due to Corollary 1.7.

For the induction step assume $k > 2$. On one hand, for even $k = 2m$ a suitable constant $c \in \mathbb{C}$ exists such that $f - c \cdot \wp^m$ has only poles of order $< k$. Therefore

$$f - c \cdot \wp^m \in \mathbb{C}(\wp, \wp')$$

by induction assumption. On the other hand, for odd $k = 2m + 1$ a suitable constant $c \in \mathbb{C}$ exists such that

$$f - c \cdot \wp^{m-1} \cdot \wp'$$

has only poles of order $< k$. By induction assumption

$$f - c \cdot \wp^{m-1} \wp' \in \mathbb{C}(\wp, \wp').$$

In both cases

$$f \in \mathbb{C}(\wp, \wp').$$

- Eventually, we consider the case of an arbitrary pole set of f .

Then f has mod Λ only finitely many poles $z_v \in P \setminus \Lambda$, $v = 1, \dots, n$. The function $\wp$ is holomorphic in $P \setminus \Lambda$. Hence for suitable constants $k_1, \dots, k_n \in \mathbb{N}$ the function

$$f(z) \cdot \prod_{v=1}^n (\wp(z) - \wp(z_v))^{k_v}$$

has poles at most in Λ . According to the first part

$$f(z) \cdot \prod_{v=1}^n (\wp(z) - \wp(z_v))^{k_v} \in \mathbb{C}(\wp, \wp').$$

Therefore

$$f \in \mathbb{C}(\wp, \wp')$$

□

Proposition 1.19 (Discriminant). *For a given lattice*

$$\Lambda = \mathbb{Z}\omega_1 + \mathbb{Z}\omega_2 \subset \mathbb{C}$$

the cubic polynomial

$$G_\Lambda(T) := 4T^3 - g_2 \cdot T - g_3 \in \mathbb{C}[T], \quad g_2 := 60 \cdot G_{\Lambda,4}, \quad g_3 := 140 \cdot G_{\Lambda,6}$$

has three pairwise distinct zeros: The values $\wp(\omega) \in \mathbb{C}$ at the half-period points of Λ

$$\omega \in Z_\Lambda := \{\omega_1/2, \omega_2/2, (\omega_1 + \omega_2)/2\}$$

As a consequence, the discriminant

$$\text{discr}(G_\Lambda) = \frac{1}{24} \cdot (g_2^3 - 27 \cdot g_3^2) \in \mathbb{C}$$

of the polynomial $G_\Lambda(T)$ is non-zero, in particular

$$g_2^3 - 27 \cdot g_3^2 \neq 0.$$

Proof. 1. *Pairwise distinct zeros:* The differential equation of $\wp'$ from Theorem 1.17 implies

$$\wp'(\omega)^2 = 4 \cdot \wp(\omega)^3 - g_2 \cdot \wp(\omega) - g_3 = G_\Lambda(\wp(\omega))$$

Due to Lemma 1.13

$$\omega \in Z_\Lambda \implies G_\Lambda(\wp(\omega)) = 0$$

and the function $\wp$ assumes at the half-period $\omega \in Z_\Lambda$ the value $\wp(\omega)$ with multiplicity at least = 2 because $\wp'(\omega) = 0$. In case

$$\wp(\omega) = \wp(\omega') \text{ for } \omega \neq \omega' \in Z_\Lambda$$

the elliptic function $\wp$ would assume the value $\wp(\omega)$ with multiplicity at least = 4. We would obtain a contradiction to Corollary 1.10 that the elliptic function $\wp$ assumes *mod* Λ each value with multiplicity = 2. Hence

$$\wp(\omega) \neq \wp(\omega').$$

2. *Discriminant:* One checks the discriminant formula

$$\text{discr}(G_\Lambda) = \frac{1}{24} \cdot (g_2^3 - 27 \cdot g_3^2) \in \mathbb{C},$$

e.g. see [35, Cor. 3.4]. By definition the discriminant of a polynomial is non-zero iff the zeros of the polynomial are pairwise distinct.

□

1.3 Abel's theorem

The zeros and poles of an elliptic functions cannot be prescribed arbitrarily: Abel's Theorem states a sufficient and necessary condition.

We continue fixing an arbitrary lattice

$$\Lambda := \mathbb{Z} \omega_1 + \mathbb{Z} \omega_2,$$

keeping the notations $\Lambda' := \Lambda \setminus \{0\}$ and P denoting the fundamental period parallelogram of Λ .

Proposition 1.20 (Zero and pole divisor). *Consider an elliptic function and denote by $A = (a_i)_{i=1,\dots,n}$ its family of zeros in P and by $B = (b_i)_{i=1,\dots,n}$ its family of poles in P . Then $n \geq 2$ and*

$$\sum_{i=1,\dots,n} a_i - \sum_{i=1,\dots,n} b_i \in \Lambda.$$

Note that each point p of A and B appears as often as its multiplicity as zero or pole of f indicates. By Corollary 1.9 holds $n \geq 2$, and by Corollary 1.10 both families A and B have the same cardinality.

Proof. We consider a translation

$$P_a := P + a$$

of the fundamental parallelogram such that f has neither zeros nor poles on the boundary ∂P_a . With a counter-clockwise orientation of ∂P_a the residue theorem gives

$$\frac{1}{2\pi i} \int_{\partial P_a} z \cdot \frac{f'(z)}{f(z)} dz = \sum_{p \in P_a} \text{res}_p \left(z \cdot \frac{f'(z)}{f(z)} \right).$$

- Right-hand side: Assume $\text{ord}(f; p) = k \in \mathbb{Z}^*$. Then

$$\begin{cases} p \text{ is a zero of } f \text{ of order } k & \text{if } k > 0 \\ p \text{ is a pole of } f \text{ of order } k & \text{if } k < 0 \end{cases}$$

In a neighbourhood of p

$$\frac{f'(z)}{f(z)} = \frac{k}{z-p} + f_1(z)$$

with f_1 holomorph. As a consequence

$$z \cdot \frac{f'(z)}{f(z)} = (p + (z-p)) \cdot \frac{f'(z)}{f(z)} = \frac{p \cdot k}{z-p} + f_2(z)$$

with f_2 holomorph. Therefore

$$\text{res}_p \left(z \cdot \frac{f'(z)}{f(z)} \right) = p \cdot k$$

and

$$\sum_{p \in P_a} \text{res}_p \left(z \cdot \frac{f'(z)}{f(z)} \right) = \sum_{i=1}^n a_i - \sum_{i=1}^n b_i.$$

- Left-hand side: Denote by A, B, C, D the vertices of P_a , counter-clockwise oriented.

$$\int_{\vec{AB} + \vec{CD}} z \cdot \frac{f'(z)}{f(z)} dz = \int_{\vec{AB}} (z - (z + \omega_2)) \cdot \frac{f'(z)}{f(z)} dz = -\omega_2 \int_{\vec{AB}} \frac{f'(z)}{f(z)} dz = -\omega_2 \cdot [\log f(z)]_A^B$$

Here $\log f$ denotes a branch of the logarithm in a simply-connected neighbourhood of the line AB .

$$\int_{\vec{AB} + \vec{CD}} z \cdot \frac{f'(z)}{f(z)} dz = \omega_2 \cdot (\log f(A) - \log f(B)).$$

Because

$$f(A) = f(B)$$

an integer $m_2 \in \mathbb{Z}$ exists with

$$\log f(A) - \log f(B) = 2\pi i \cdot m_2.$$

A similar argument for the lines $\vec{BC}$ and $\vec{DA}$ provides a second constant $m_1 \in \mathbb{Z}$, such that finally

$$\frac{1}{2\pi i} \cdot \int_{\partial P_a} z \cdot \frac{f'(z)}{f(z)} dz = m_1 \cdot \omega_1 + m_2 \cdot \omega_2 \in \Lambda, \text{ q.e.d.}$$

Theorem 1.21 (Abel's theorem). Consider $n \geq 2$ and two finite families of complex points

$$A := (a_i)_{i=1, \dots, n}, \quad B := (b_i)_{i=1, \dots, n} \subset \mathbb{C}$$

with A and B pairwise disjoint mod Λ . Then are equivalent:

- There exists an elliptic function $f \in \mathcal{M}(\Lambda)$ which has mod Λ the zero set A and the pole set B
- The points of A and B satisfy

$$\sum_{i=1}^n a_i - \sum_{i=1}^n b_i \in \Lambda.$$

Proof. Proposition 1.20 proves that the zero set and the pole set of f satisfy the condition stated above. In order to prove the opposite direction of the theorem assume two sets A and B with the property stated above. Without restriction we may assume

$$\sum_{j=1}^n (a_j - b_j) = 0$$

as an equation, not just a congruence; otherwise replace a_1 by a suitable point which is congruent *mod* Λ . The proof for the existence of a suitable elliptic function f is by induction.

Start of induction $n = 2$: If

$$a_1 - b_1 = b_2 - a_2$$

then consider

$$z_0 := \frac{a_1 + a_2 + b_1 + b_2}{4}.$$

Geometrically the point z_0 is the center of the parallelogram with vertices the points a_1, a_2, b_1, b_2 . We may assume $z_0 = 0$, otherwise we translate all coordinates by $-z_0$. As a consequence, the two conditions state

$$a_1 = -a_2 \text{ and } b_1 = -b_2.$$

Depending on the position relative Λ of a_1 and b_1 we now distinguish the following cases for the choice of f :

- $a_1, b_1 \notin \Lambda$: The even function

$$g(z) := \wp(z) - \wp(a_1)$$

has

- a simple zero at a_1 and at a_2 if $a_1 \not\equiv a_2 \pmod{\Lambda}$, and a zero of order = 2 at $a_1 \equiv a_2$ otherwise
- and a pole of order = 2 at $z = 0$.

An analogous result holds for the function

$$h(z) := \wp(z) - \wp(b_1).$$

Hence the quotient

$$f := \frac{g}{h} = \frac{\wp - \wp(a_1)}{\wp - \wp(b_1)}$$

has zeros exactly the zeros of g and poles exactly the zeros of h , while the poles of g and h at the origin cancel.

- $a_1 \in \Lambda$: Then $b_1 \notin \Lambda$. Set

$$f := \frac{1}{\wp - \wp(b_1)}$$

- $b_1 \in \Lambda$: Then $a_1 \notin \Lambda$. Set

$$f := \wp - \wp(a_1)$$

Hence f is elliptic with prescribed zeros and poles.

Induction step $n \geq 2$, $n \mapsto n+1$: If mod Λ

$$(a_1 - b_1) + \dots + (a_n - b_n) + (a_{n+1} - b_{n+1}) \equiv 0$$

then choose $\tilde{a}_n \in \mathbb{C}$ with

$$\tilde{a}_n \equiv (a_1 - b_1) + \dots + (a_{n-1} - b_{n-1}) + a_n.$$

Then mod Λ

$$(a_1 - b_1) + \dots + (a_{n-1} - b_{n-1}) + (a_n - \tilde{a}_n) \equiv 0$$

and

$$(\tilde{a}_n + a_{n+1}) - (b_n + b_{n+1}) \equiv 0.$$

Depending on the position of $\tilde{a}_n$ relative Λ we distinguish the following cases:

- $\tilde{a}_n \neq a_j$ for all $j = 1, \dots, n$: First we exclude the two exceptional cases $\tilde{a}_n \equiv b_n$ and $\tilde{a}_n \equiv b_{n+1}$: They lead to respectively

$$a_{n+1} \equiv b_{n+1} \text{ and } a_{n+1} \equiv b_n$$

which has been excluded by assumption.

For the remaining general cases, by induction assumption two elliptic functions exist

$$f_1 \in \mathcal{M}(\Lambda)$$

with zeros $(a_1, \dots, a_n)$ and poles $(b_1, \dots, b_{n-1}, \tilde{a}_n)$ and

$$f_2 \in \mathcal{M}(\Lambda)$$

with zeros $(\tilde{a}_n, a_{n+1})$ and poles (b_n, b_{n+1}) .

- $\tilde{a}_n \equiv a_n$: Then

$$(a_1 - b_1) + \dots + (a_{n-1} - b_{n-1}) \equiv 0$$

By induction assumption two elliptic functions exist

$$f_1 \in \mathcal{M}(\Lambda)$$

with zeros $(a_1, \dots, a_{n-1})$ and poles $(b_1, \dots, b_{n-1})$ and

$$f_2 \in \mathcal{M}(\Lambda)$$

with zeros (a_n, a_{n+1}) and poles (b_n, b_{n+1}) .

- $\tilde{a}_n \neq a_n$, $\tilde{a}_n \equiv a_j$ for at least one $j = 1, \dots, n-1$, w.l.o.g. $\tilde{a}_n \equiv a_1$: Then

$$(a_n - b_1)(a_2 - b_2) + \dots + (a_{n-1} - b_{n-1}) \equiv 0$$

and

$$(a_1 + a_{n+1}) - (b_n + b_{n+1}) \equiv 0.$$

Similarly to the previous case by induction assumption two elliptic functions exist

$$f_1 \in \mathcal{M}(\Lambda)$$

with zero set $(a_2, \dots, a_n)$ and pole set $(b_1, \dots, b_{n-1})$ and

$$f_2 \in \mathcal{M}(\Lambda)$$

with zero (a_1, a_{n+1}) and poles (b_n, b_{n+1}) .

In all cases the elliptic function

$$f := f_1 \cdot f_2 \in \mathcal{M}(\Lambda)$$

has the prescribed zeros $(a_1, \dots, a_{n+1})$ and poles $(b_1, \dots, b_{n+1})$, q.e.d.

Corollary 1.22 (Elliptic functions with pole of order = 3). *An elliptic function $f \in \mathcal{M}(\Lambda)$ with a pole mod Λ at $0 \in \Lambda$ of order = 3 and no other poles mod Λ , has exactly three zeros a_1, a_2, a_3 mod Λ . They satisfy*

$$a_1 + a_2 + a_3 = 0 \text{ mod } \Lambda.$$

Corollary 1.22 applies to $f = \wp'$, see Lemma 1.13.

Chapter 2

The modular group $SL(2, \mathbb{Z})$ and Hecke congruence subgroups $\Gamma_0(N)$

We now leave the complex plane. The lengthy case-by-case analysis in the proof of Theorem 1.21 indicates: To consider elliptic functions as functions with a certain period living in the complex plane, is not the best approach. Elliptic functions are better studied in the context of Riemann surfaces as meromorphic functions on a torus.

2.1 The moduli space of complex tori and group actions

An elliptic function f with lattice Λ , as studied in Chapter 1, has as its natural domain of definition not the complex plane $\mathbb{C}$. Due to its periodicity the natural domain of definition for f is the quotient $\mathbb{C}/\Lambda$, a complex torus: Elliptic functions are meromorphic functions on the torus.

Therefore the present section investigates complex analysis on complex tori. After providing the quotient $\mathbb{C}/\Lambda$ with the quotient topology with respect to the canonical projection

$$\mathbb{C} \rightarrow \mathbb{C}/\Lambda$$

one can also introduce a complex structure on the quotient and obtain a complex torus. In addition, the torus inherits the structure of an Abelian group from the additive group $(\mathbb{C}, +)$.

We will prove that the equivalence classes of complex tori up to biholomorphic isomorphisms depend on a single complex parameter. The parameter space can be chosen as quotient of the upper half-plane

$$\mathbb{H} := \{\tau \in \mathbb{C} : \text{Im } \tau > 0\}$$

with respect to the action of the modular group $SL(2, \mathbb{Z})$ as a group of fractional-linear transformations. This group is the first indicator that the theory of elliptic

functions and tori closely links to number theory. Starting with the fact that the integral domain $\mathbb{Z}$ is factorial, i.e. its elements factorize uniquely into prime numbers, we investigate some number theoretic properties of certain subgroups of $SL(2, \mathbb{Z})$.

The basis of a lattice

$$\Lambda = \mathbb{Z}\omega_1 + \mathbb{Z}\omega_2$$

is not uniquely determined. Two positively oriented bases generate the same lattice if and only if they are equivalent modulo $SL(2, \mathbb{Z})$, i.e. iff they belong to the same orbit of the following $SL(2, \mathbb{Z})$ -action:

Lemma 2.1 (Different bases of a given lattice). *Denote by*

$$M := \{(\omega_1, \omega_2) \text{ basis of } \mathbb{R}^2 \text{ with } \det(\omega_1 \ \omega_2) > 0\}$$

the set of positively oriented bases of $\mathbb{R}^2$. The set $\mathcal{R}$ of all lattices $\Lambda \subset \mathbb{C}$ is the set of (left) equivalence classes of M

$$SL(2, \mathbb{Z}) \backslash M$$

with respect to the equivalence relation induced by the map

$$SL(2, \mathbb{Z}) \times M \rightarrow M, \left(\gamma, \begin{pmatrix} \omega_1 \\ \omega_2 \end{pmatrix} \right) \mapsto \begin{pmatrix} \omega'_1 \\ \omega'_2 \end{pmatrix} := \gamma \cdot \begin{pmatrix} \omega_1 \\ \omega_2 \end{pmatrix}.$$

Proof. The invertible matrices $\gamma \in M(2 \times 2, \mathbb{Z})$ have determinant ± 1 . Because both bases are positively oriented we have $\det \gamma = 1$. $\square$

Two tori belonging to different lattices may be biholomorphically equivalent. The following Definition 2.2 introduces a name for the relation between lattices with isomorphic tori.

Definition 2.2 (Similar lattices). Two lattices $\Lambda, \Lambda' \subset \mathbb{C}$ are *similar* if a complex number $a \in \mathbb{C}^*$ exists with

$$\Lambda' = a \cdot \Lambda.$$

Proposition 2.3 (Holomorphic maps between tori). *Consider a holomorphic map*

$$F : \mathbb{C}/\Lambda_1 \rightarrow \mathbb{C}/\Lambda_2$$

between two tori with $F(0) = 0$. Then a uniquely determined number $a \in \mathbb{C}$ exists such that the following diagram commutes

$$\begin{array}{ccc}
\mathbb{C} & \xrightarrow{\mu_a} & \mathbb{C} \\
p_1 \downarrow & & \downarrow p_2 \\
\mathbb{C}/\Lambda_1 & \xrightarrow{F} & \mathbb{C}/\Lambda_2
\end{array}$$

Here μ_a denotes the multiplication by a and p_i , $i = 1, 2$, denote the canonical projections.

In particular, any holomorphic map between tori which fixes the neutral element is a group homomorphism.

Note: Due to the commutativity of the diagram in Proposition 2.3 the multiplier $a \in \mathbb{C}$ satisfies

$$a \cdot \Lambda_1 \subset \Lambda_2.$$

Proof. i) The composition

$$F \circ p_1 : \mathbb{C} \rightarrow \mathbb{C}/\Lambda_2$$

starts from the simply-connected domain $\mathbb{C}$. Therefore it lifts to the total space $\mathbb{C}$ of the covering projection p_2 . It provides a holomorphic map

$$\tilde{F} : \mathbb{C} \rightarrow \mathbb{C},$$

uniquely determined by the condition $\tilde{F}(0) = 0$. Commutativity of the lifting diagram shows: For all $z \in \mathbb{C}$, $\omega \in \Lambda_1$,

$$\tilde{F}(z + \omega) - \tilde{F}(z) \in \Lambda_2.$$

The lattice Λ_2 is discrete and $\tilde{F}$ is continuous. Therefore

$$\tilde{F}(z + \omega) - \tilde{F}(z)$$

does not depend on z . As a consequence for all $z \in \mathbb{C}$

$$\tilde{F}'(z + \omega) - \tilde{F}'(z) = 0.$$

Hence $\tilde{F}'$ is a holomorphic and elliptic. According to Proposition 1.7 the function $\tilde{F}'$ equals a constant $a \in \mathbb{C}$, and for all $z \in \mathbb{C}$

$$\tilde{F}(z) = a \cdot z$$

ii) The uniqueness of a follows from the fact that μ_a induces the zero map F iff for all $z \in \mathbb{C}$ holds $a \cdot z \in \Lambda_2$. Because $\Lambda_2 \subset \mathbb{C}$ is discrete and $a \cdot 0 = 0$ we have $a \cdot z = 0$ for all $z \in \mathbb{C}$, which implies $a = 0$. $\square$

Corollary 2.4 (Lattices of biholomorphically equivalent tori). *Two tori*

$$T_1 = \mathbb{C}/\Lambda_1 \text{ and } T_2 = \mathbb{C}/\Lambda_2$$

are biholomorphically equivalent iff their lattices Λ_1 and Λ_2 are similar.

Proof. i) Assume the existence of a biholomorphic map

$$F : T_1 \rightarrow T_2$$

satisfying without restriction $F(0) = 0$. Proposition 2.3 provides a commutative diagram

$$\begin{array}{ccc} \mathbb{C} & \xrightarrow{\mu_a} & \mathbb{C} \\ p_1 \downarrow & & \downarrow p_2 \\ \mathbb{C}/\Lambda_1 & \xrightarrow{F} & \mathbb{C}/\Lambda_2 \end{array}$$

with μ_a the multiplication by a complex number $a \in \mathbb{C}$, satisfying

$$\mu_a \cdot \Lambda_1 \subset \Lambda_2.$$

Analogously, there exists a complex number $b \in \mathbb{C}$ with μ_b the unique lift of F^{-1} fixing the origin. Then the map

$$\mu_b \circ \mu_a = \mu_{b \cdot a}$$

is the unique lift of

$$id_{\mathbb{C}/\Lambda_1} = F^{-1} \circ F$$

fixing the origin. Hence

$$b \cdot a = 1.$$

In particular

$$a, b \in \mathbb{C}^*$$

and

$$a \cdot \Lambda_1 = \Lambda_2.$$

ii) Assume

$$\Lambda_2 = a \cdot \Lambda_1$$

with $a \in \mathbb{C}^*$. Then the multiplication

$$\mu_a : \mathbb{C} \rightarrow \mathbb{C},$$

induces a biholomorphic map $F : T_1 \rightarrow T_2$. $\square$

Theorem 2.5 (Biholomorphic equivalence classes of complex tori).

1. Any complex torus is biholomorphic equivalent to a torus $T = \mathbb{C}/\Lambda$ with a normalized lattice

$$\Lambda = \mathbb{Z} \cdot 1 + \mathbb{Z} \cdot \tau \subset \mathbb{C}$$

with $\tau \in \mathbb{H}$ in the upper half-plane.

2. The period τ of a normalized lattice is determined up to a fractional linear transformation with integer coefficients. Hence two tori

$$\mathbb{C}/\Lambda_\tau \text{ and } \mathbb{C}/\Lambda_{\tau'}$$

with normalized lattices

$$\Lambda_\tau = \mathbb{Z} \cdot 1 + \mathbb{Z} \cdot \tau \text{ and } \Lambda_{\tau'} = \mathbb{Z} \cdot 1 + \mathbb{Z} \cdot \tau'$$

are biholomorphic equivalent iff

$$\tau' = \frac{a\tau + b}{c\tau + d}$$

for a matrix

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL(2, \mathbb{Z}).$$

Proof. 1. Consider an arbitrary torus $T = \mathbb{C}/\Lambda$ with a lattice satisfying

$$\Lambda = \mathbb{Z}\omega_1 + \mathbb{Z}\omega_2$$

and $\det(\omega_1 \ \omega_2) > 0$. Then

$$\tau := \frac{\omega_2}{\omega_1}$$

satisfies $\text{Im } \tau > 0$ and the lattice Λ is similar to the normalized lattice

$$\Lambda_\tau := \mathbb{Z} \cdot 1 + \mathbb{Z} \cdot \tau$$

due to

$$\frac{1}{\omega_1} \cdot \Lambda = \Lambda_\tau.$$

According to Corollary 2.4 the two tori T and $T' := \mathbb{C}/\Lambda_\tau$ are biholomorphic equivalent.

2. i) Assume

$$\tau' = \gamma \cdot \tau = \frac{a\tau + b}{c\tau + d} \text{ with } \gamma := \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL(2, \mathbb{Z}).$$

Set

$$\begin{pmatrix} \omega_2 \\ \omega_1 \end{pmatrix} := \gamma \cdot \begin{pmatrix} \tau \\ 1 \end{pmatrix}, \text{ i.e. } \omega_2 = a\tau + b, \ \omega_1 = c\tau + d.$$

The base change implies

$$\mathbb{Z} \cdot \omega_1 + \mathbb{Z} \cdot \omega_2 = \mathbb{Z} \cdot 1 + \mathbb{Z} \cdot \tau = \Lambda_\tau.$$

As a consequence

$$\frac{1}{\omega_1} \cdot \Lambda_\tau = \Lambda_{\tau'}$$

because

$$\tau' = \frac{\omega_2}{\omega_1}.$$

The lattices Λ_τ and $\Lambda_{\tau'}$ are similar. Hence the tori $\mathbb{C}/\Lambda_\tau$ and $\mathbb{C}/\Lambda_{\tau'}$ are biholomorphic equivalent according to Corollary 2.4.

ii) For the opposite direction assume that two tori with normalized lattices

$$\mathbb{C}/\Lambda_\tau \text{ and } \mathbb{C}/\Lambda_{\tau'}, \tau, \tau' \in \mathbb{H},$$

are biholomorphic equivalent. According to Corollary 2.4 a number $\alpha \in \mathbb{C}^*$ exists with

$$\Lambda_{\tau'} = \alpha \cdot \Lambda_\tau.$$

As a consequence

$$\Lambda_{\tau'} = \mathbb{Z} \cdot \alpha + \mathbb{Z} \cdot \alpha \cdot \tau.$$

Therefore a matrix

$$\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL(2, \mathbb{Z})$$

exists with

$$\begin{pmatrix} \tau' \\ 1 \end{pmatrix} = \gamma \cdot \begin{pmatrix} \alpha\tau \\ \alpha \end{pmatrix},$$

i. e.

$$\tau' = \frac{a \cdot \alpha\tau + b \cdot \alpha}{c \cdot \alpha\tau + d \cdot \alpha} = \frac{a \cdot \tau + b}{c \cdot \tau + d}$$

□

Theorem 2.5 shows the consequences when two points of $\mathbb{H}$ are related by an element of the group $SL(2, \mathbb{Z})$. We formalize this relation by the concept of a *group action* on a topological space.

Definition 2.6 (Group action).

1. Consider a topological group G with neutral element $e \in G$ and a topological space X . A *left action* of G on X is a continuous map

$$\phi : G \times X \rightarrow X, (g, x) \mapsto gx,$$

also written $g(x) := gx$, satisfying the following properties: For all $g_1, g_2 \in G$ and for all $x \in X$

- $ex = x$
- $g_1(g_2x) = (g_1g_2)x$

2. Consider a group action G on X .

- The action is *faithful* if for all $g \in G$, $g \neq e$, exists $x \in X$ with $gx \neq x$, i.e. no group element different from e acts in a trivial way.
- A *fixed point* of the group action is a point $x \in X$ with $gx = x$ for all $g \in G$. The action is *fixed-point-free* if it has no fixed points.
- The *isotropy group* of a point $x \in X$ is the subgroup

$$G_x := \{g \in G : gx = x\} \subset G.$$

- The *orbit* of $x \in X$ is the subspace

$$\{gx \in X : g \in G\} \subset X$$

Two points of X are equivalent with respect to the group action if they belong to the same orbit. The set of left equivalence classes is the *orbit space* $G \backslash X$, the set of orbits of the action. The canonical map to the orbit set is denoted

$$\pi : X \rightarrow G \backslash X, x \mapsto [x].$$

- A *fundamental domain* $\mathcal{F}$ of the group action is a subset $\mathcal{F} \subset X$ such that the restriction

$$\pi|_{\mathcal{F}} : \mathcal{F} \rightarrow G \backslash X$$

is bijective. Depending on eventual additional properties of the group action one looks for fundamental domains with specific additional properties.

- The action is *transitive* if the whole set X is a single orbit, i.e. if for any two points $x_1, x_2 \in X$ exists $g \in G$ with $x_2 = gx_1$.

Apparently all points of the same orbit have conjugated isotropy groups because

$$G_{gx} = g \cdot G_x \cdot g^{-1}.$$

Theorem 2.5 deals with the specific group action from Definition 2.7.

Definition 2.7 (Action of $SL(2, \mathbb{Z})$ on $\mathbb{H}$). The $SL(2, \mathbb{Z})$ -left action on $\mathbb{H}$ is the map

$$\Phi : SL(2, \mathbb{Z}) \times \mathbb{H} \rightarrow \mathbb{H}, (\gamma, \tau) \mapsto \gamma(\tau) := \frac{a\tau + b}{c\tau + d}, \gamma := \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Remark 2.8 (Action of $SL(2, \mathbb{Z})$ and extensions).

1. *Imaginary part:* If $\tau \in \mathbb{H}$ and

$$\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in GL(2, \mathbb{R})^+ := \{A \in GL(2, \mathbb{C}) : \det A > 0\}$$

then

$$\operatorname{Im} \gamma(\tau) = \det(\gamma) \cdot \frac{\operatorname{Im} \tau}{|c\tau + d|^2}$$

As a consequence

$$\tau \in \mathbb{H} \implies \gamma(\tau) \in \mathbb{H}.$$

Hence the action from Definition 2.7 extends to an action

$$GL(2, \mathbb{R})^+ \times \mathbb{H} \rightarrow \mathbb{H}, \gamma(\tau) := \frac{a\tau + b}{c\tau + d}$$

which apparently restricts to an action of all subgroups of $GL(2, \mathbb{R})^+$.

Due to Theorem 2.5 the biholomorphic equivalence classes of complex tori correspond bijectively to the points of the orbit space $SL(2, \mathbb{Z}) \backslash \mathbb{H}$. Therefore the orbit space is named the *moduli space* of 1-dimensional complex tori.

Note: The orbit space is read “ $\mathbb{H}$ modulo $SL(2, \mathbb{Z})$ ”.

2. *Extending the action of $SL(2, \mathbb{Z})$:* We have

$$z \in \mathbb{R} \implies \gamma(z) \in \mathbb{R} \cup \{\infty\} \text{ and } z \in \mathbb{Q} \implies \gamma(z) \in \mathbb{Q} \cup \{\infty\}$$

Moreover, due to

$$\frac{az + b}{cz + d} = \frac{a + (b/z)}{c + (d/z)}$$

we can extend γ to the argument ∞ by setting

$$\gamma(\infty) := \begin{cases} \infty & \text{if } d \neq 0 \text{ and } c = 0 \\ \frac{a}{c} & \text{otherwise} \end{cases}$$

Set

$$\mathbb{H}^* := \mathbb{H} \cup \mathbb{Q} \cup \{\infty\}$$

Then the extension

$$\Phi : SL(2, \mathbb{Z}) \times \mathbb{H}^* \rightarrow \mathbb{H}^*, (\gamma, z) \mapsto \gamma(z),$$

is well-defined.

Note. The proof of Theorem 2.31 will provide $\mathbb{H}^*$ with a specific topology. It induces on $\mathbb{Q} \subset \mathbb{H}$ a topology different from the subspace topology of $\mathbb{Q} \subset \mathbb{C}$ and induces also specific neighbourhoods of $\infty \in \mathbb{H}^*$.

Definition 2.9 (Modular group, Hecke congruence subgroups and cusps).

1. The *modular group* is the group

$$SL(2, \mathbb{Z})$$

The elements

$$\pm id \in SL(2, \mathbb{Z})$$

are the only elements from $SL(2, \mathbb{Z})$ which act trivially on $\mathbb{H}$.

Note. Some authors like Serre in [56] therefore apply the name modular group to the quotient

$$PSL(2, \mathbb{Z}) := SL(2, \mathbb{Z}) / \{\pm id\}.$$

We will not follow this convention, because sometimes one studies also the action of certain subgroups of $SL(2, \mathbb{Z})$ which do not contain the element $-id$.

2. For each $N \in \mathbb{N}^*$ one defines

$$SL(2, \mathbb{Z}/N\mathbb{Z}) := \{A \in M(2 \times 2, \mathbb{Z}/N\mathbb{Z}) : \det A = 1 \in \mathbb{Z}/N\mathbb{Z}\}$$

and defines componentwise the residue morphism

$$\pi : SL(2, \mathbb{Z}) \rightarrow SL(2, \mathbb{Z}/N\mathbb{Z}), A \mapsto \bar{A}$$

3. For a positive integer $N \in \mathbb{N}^*$ the *Hecke congruence subgroup* of level N is the subgroup

$$\Gamma_0(N) := \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL(2, \mathbb{Z}) : c \equiv 0 \pmod{N} \right\} \subset SL(2, \mathbb{Z})$$

The canonical left action of $SL(2, \mathbb{Z})$ from Remark 2.8 on $\mathbb{H}^*$ restricts to a left action

$$\Phi_N : \Gamma_0(N) \times \mathbb{H}^* \rightarrow \mathbb{H}^*.$$

4. The points $\tau \in \mathbb{H}$ with a non-trivial isotropy group, i.e. $\Gamma_0(N) \supsetneq \{\pm id\}$, are named the *elliptic points* of $\Gamma_0(N)$. Half the order of their isotropy group

$$h_\tau := \frac{\text{ord } \Gamma_0(N)_\tau}{2}$$

is named the *period* of τ . An orbit is named an *elliptic orbit* if at least one point of the orbit - and hence all points - are elliptic.

5. The orbits of the points $\tau \in \mathbb{Q} \cup \{\infty\}$ are named the *cusps* of $\Gamma_0(N)$, the set of cusps is denoted $\text{cusp}(\Gamma_0(N))$. If the cusp s passes through the point

$$\tau_s \in \mathbb{Q} \cup \{\infty\}$$

then one compares two isotropy groups of τ_s ,

- on one hand the isotropy group $SL(2, \mathbb{Z})_{\tau_s}$, a subgroup of the modular group
- and on the other hand the isotropy group $\Gamma_0(N)_{\tau_s}$, a subgroup of the congruence subgroup.

The index

$$h_s := [SL(2, \mathbb{Z})_{\tau_s} : \Gamma_0(N)_{\tau_s}]$$

is independent from the choice of the point $\tau_s \in s$ on the cusp. It is named the *width* (Deutsch: "Breite") of the cusp s .

Note: The duality between elliptic points $\tau \in \mathbb{H}$ and their period versus cusps $s \in \text{cusp}(\Gamma_0(N))$ and their width.

Lemma 2.10 (Width of a cusp). *The isotropy group Γ_∞ at the point ∞ of the modular group is generated by the translation*

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \in \Gamma$$

For the congruence subgroup $\Gamma_0(N) \subset \Gamma$ holds:

- The isotropy group $\Gamma_0(N)_\infty \in \Gamma_0(N)$ is generated by the translation

$$\begin{pmatrix} 1 & N \\ 0 & 1 \end{pmatrix} \in \Gamma_0(N)$$

and N equals the width of each cusp of $\Gamma_0(N)$.

- The width of a cusp can also be computed by transforming the point τ_s to infinity: There exists an element $\gamma \in SL(2, \mathbb{Z})$ with $\gamma(\tau_s) = \infty$. Then

$$h_s = [SL(2, \mathbb{Z})_\infty : (\gamma\{\pm id\}\Gamma_0(N)\gamma^{-1})_\infty]$$

Remark 2.11 (Modular group and congruence subgroups).

1. Apparently,

$$SL(2, \mathbb{Z}) = \Gamma_0(1)$$

2. All congruence subgroups contain the element $-id$.

3. Corollary 2.20 will show that the isotropy groups of all elliptic points from Definition 2.9 are finite groups. But the isotropy group $SL(2, \mathbb{Z})_\infty \subset SL(2, \mathbb{Z})$ is an infinite group, more specific

$$SL(2, \mathbb{Z})_\infty = \langle \pm T \rangle \text{ with translation } T := \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \in SL(2, \mathbb{Z})$$

4. Theorem 2.31 will show the importance to consider the group action not only on the open upper half-plane $\mathbb{H}$ but also on its “closure” $\mathbb{H}^*$.

Example 2.12 (Cusps of $SL(2, \mathbb{Z})$, $\Gamma_0(2)$, $\Gamma_0(4)$ and $\Gamma_0(11)$). See [19, Sect. 3.8], and the PARI commands *mfcusps*, *mfnumcusps*, *mfcuspwidth*.

1. For $N \in \mathbb{N}$ the number $\varepsilon_\infty(\Gamma_0(N))$ of cusps is

$$\sum_{d|N} \phi(\gcd(d, N/d))$$

with the sum extending over all positive divisors d of N , and ϕ denoting the Euler function defined as

$$\Phi(n) := \text{card } (\mathbb{Z}/n\mathbb{Z})^*$$

2. The group $SL(2, \mathbb{Z})$ has a single cusp

- Orbit of the point 0:

$$\mathbb{Q} \cup \{\infty\}$$

with *width* = 1. All rationals $p/q \in \mathbb{Q}$ belong to the $SL(2, \mathbb{Z})$ -orbit of ∞ , see Remark 2.8.

3. The group $\Gamma_0(2)$ has two cusps

- *Orbit of the point 0:*

$$\left\{ \frac{p}{q} \in \mathbb{Q} : (p, q) = 1, q \text{ odd} \right\}$$

with $width = 2$.

- *Orbit of the point $1/2$:*

$$\left\{ \frac{p}{q} \in \mathbb{Q} : (p, q) = 1, p \neq 0, q \text{ even} \right\} \cup \{\infty\}$$

with $width = 1$.

4. The group $\Gamma_0(4)$ has three cusps

- *Orbit of the point 0:*

$$\left\{ \frac{p}{q} \in \mathbb{Q} : (p, q) = 1, q \text{ odd} \right\}$$

with $width = 4$.

- *Orbit of the point $1/2$:*

$$\left\{ \frac{p}{q} \in \mathbb{Q} : (p, q) = 1, p \neq 0, q \text{ even, not divisible by 4} \right\}$$

with $width = 1$.

- *Orbit of the point $1/4$:*

$$\left\{ \frac{p}{q} \in \mathbb{Q} : (p, q) = 1, p \neq 0, q \text{ divisible by 4} \right\} \cup \{\infty\}$$

with $width = 1$.

5. The group $\Gamma_0(11)$ has two cusps

- *Orbit of the point 0:*

$$\left\{ \frac{p}{q} \in \mathbb{Q} : (p, q) = 1, q \text{ not divisible by 11} \right\}$$

with $width = 11$.

- *Orbit of the point $1/11$:*

$$\left\{ \frac{p}{q} \in \mathbb{Q} : (p, q) = 1, p \neq 0, q \text{ divisible by } 11 \right\} \cup \{\infty\}$$

with $width = 1$.

Lemma 2.13 (Divisor lemma). *Consider three integers $a, b, c \in \mathbb{Z}$ with*

$$\gcd(a, b, c) = 1, c \neq 0.$$

Then there exists an integer $r \in \mathbb{Z}$ such that

$$\gcd(a + r \cdot b, c) = 1.$$

Proof. E.g. [38, Kap. II, §3 Lemma]: We claim that

$$r := \prod_{p|c, p \nmid a} p$$

satisfies the claim. Otherwise assume the existence of a prime q satisfying

$$q \mid \gcd(a + r \cdot b, c).$$

- *Case $q \mid a$:* Then $q \mid rb$. Because $q \nmid r$ then $q \mid b$. Therefore

$$\gcd(a, b, c) \neq 1,$$

a contradiction.

- *Case $q \nmid a$:* Then $q \mid r$ by definition of r . Because $q \mid a + rb$ also $q \mid a$, a contradiction, q.e.d.

Definition 2.14 (Level N principal congruence subgroup). For each $N \in \mathbb{N}^*$ the group

$$SL(2, \mathbb{Z})(N) := \left\{ A \in SL(2, \mathbb{Z}) : A \equiv \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \mod N \right\}$$

is named the *level N principal congruence subgroup*

Lemma 2.15 (Exact sequence of congruence subgroups). *For each $N \in \mathbb{N}^*$ exists a canonical exact sequence of Abelian groups*

$$1 \rightarrow \Gamma(N) \rightarrow SL(2, \mathbb{Z}) \xrightarrow{\pi} SL(2, \mathbb{Z}/N\mathbb{Z}) \rightarrow 1$$

Proof. i) *kernel:* Apparently $\ker \pi = \Gamma(N)$.

ii) *Surjectivity:* E.g. see [38, Kap. II, §3 Satz]. Assume a given

$$\bar{A} = \begin{pmatrix} \bar{\alpha} & \bar{\beta} \\ \bar{\gamma} & \bar{\delta} \end{pmatrix} \in SL(2, \mathbb{Z}/N\mathbb{Z})$$

We may assume

$$A = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \in M(2 \times 2, \mathbb{Z}),$$

and w.l.o.g $\gamma \neq 0$, otherwise set $\gamma = N$. The condition about the determinant

$$\det \bar{A} = 1 \in \mathbb{Z}/N\mathbb{Z}$$

implies

$$\alpha\delta - \beta\gamma \equiv 1 \pmod{N},$$

and therefore

$$\gcd(\gamma, \delta, N) = 1.$$

Lemma 2.13 provides an integer $r \in \mathbb{Z}$ satisfying

$$\gcd(\gamma, d) = 1$$

for

$$d := \delta + rN.$$

On one hand, from the determinant

$$\alpha d - \beta\gamma = \alpha\delta - \beta\gamma + \alpha rN = 1 + sN$$

with a suitable integer $s \in \mathbb{Z}$. On the other hand

$$\gcd(\gamma, d) = 1$$

provides an equation

$$\gamma y - dx = s$$

with suitable integers $x, y \in \mathbb{Z}$.

Consider the matrix

$$\tilde{A} := \begin{pmatrix} \alpha + xN & \beta + yN \\ \gamma & d \end{pmatrix} \in M(2 \times 2, \mathbb{Z})$$

Then

$$\pi(\tilde{A}) = \pi(A) = \bar{A}$$

and

$$\det \tilde{A} = (\alpha + xN) \cdot d - (\beta + yN) \cdot \gamma = 1 + sN + N(xd - y\gamma) = 1,$$

which implies $\tilde{A} \in \Gamma$ and finishes the proof of the lemma. $\square$

Proposition 2.16 (Index of Hecke congruence subgroups). *For any integer $N \geq 1$ holds the index formula*

$$[SL(2, \mathbb{Z}) : \Gamma_0(N)] = N \cdot \prod_{p|N} \left(1 + \frac{1}{p}\right)$$

Here the product ranges over all prime factors p of N .

In particular, all Hecke congruence subgroups have finite index in the modular group.

Proof. See also [59, Chap. 1.6].

i) *Multiplicativity of the index formula:* We extend the exact sequence from Lemma 2.15 to the commutative diagram

$$\begin{array}{ccccccccc} 1 & \longrightarrow & \Gamma(N) & \longrightarrow & SL(2, \mathbb{Z}) & \xrightarrow{\pi} & SL(2, \mathbb{Z}/N\mathbb{Z}) & \longrightarrow & 1 \\ & & \uparrow id & & \uparrow & & \uparrow \psi & & \\ 1 & \longrightarrow & \Gamma(N) & \longrightarrow & \Gamma_0(N) & \longrightarrow & \Gamma_0(N)/\Gamma(N) & \longrightarrow & 1 \end{array}$$

There is an induced morphism

$$\psi : \Gamma_0(N)/\Gamma(N) \rightarrow SL(2, \mathbb{Z}/N\mathbb{Z})$$

because $\det \bar{A} = \overline{\det A}$. The map ψ is injective. The image of ψ is

$$\text{im } \psi = \left\{ \begin{pmatrix} \bar{\alpha} & \bar{\beta} \\ 0 & \bar{\alpha}^{-1} \end{pmatrix} \right\} \subset SL(2, \mathbb{Z}/N\mathbb{Z})$$

The injectivity of ψ implies

$$[\Gamma_0(N) : \Gamma(N)] = \text{card}(\text{im } \psi) = \phi(N) \cdot N$$

with the Euler function ϕ . The exact sequence of the first line of the diagram implies

$$[SL(2, \mathbb{Z}) : \Gamma(N)] = \text{card } SL(2, \mathbb{Z}/N\mathbb{Z}).$$

The multiplicativity of the index function implies

$$[SL(2, \mathbb{Z}) : \Gamma(N)] = [SL(2, \mathbb{Z}) : \Gamma_0(N)] \cdot [\Gamma_0(N) : \Gamma(N)]$$

ii) *Factorisation*: The factorisation of integers N into their prime powers p^{e_p}

$$N = \prod_{p|N} p^{e_p}$$

implies the factorisation of the Euler function

$$\phi(N) = \prod_{p|N} \phi(p^{e_p})$$

with the general Euler factor

$$\phi(p^e) = p^e \cdot \left(1 - \frac{1}{p}\right)$$

The Chinese remainder theorem implies the factorisation of the $SL(2, -)$ -groups

$$SL(2, \mathbb{Z}/N\mathbb{Z}) \simeq \prod_{p|N} SL(2, \mathbb{Z}/p^{e_p}\mathbb{Z}).$$

One has

$$\text{card } SL(2, \mathbb{Z}/p^e\mathbb{Z}) = p^{3e} \left(1 - \frac{1}{p^2}\right),$$

see [59, Chap. 1.6] and [19, Exerc. 1.2.3]. For a proof see also [30, Cor. 2.8] and use for the reduction from GL to SL the exact sequence of the determinant

$$1 \rightarrow SL(2, \mathbb{Z}/p^e\mathbb{Z}) \rightarrow GL(2, \mathbb{Z}/p^e\mathbb{Z}) \xrightarrow{\det} (\mathbb{Z}/p^e\mathbb{Z})^* \rightarrow 1.$$

According to part i)

$$[SL(2, \mathbb{Z}) : \Gamma(N)] = \text{card } SL(2, \mathbb{Z}/N\mathbb{Z}),$$

It follows

$$\begin{aligned} [SL(2, \mathbb{Z}) : \Gamma(N)] &= \prod_{p|N} \text{card } SL(2, \mathbb{Z}/p^{e_p}\mathbb{Z}) = \\ &= \prod_{p|N} p^{3e_p} \left(1 - \frac{1}{p^2}\right) = N^2 \cdot \prod_{p|N} p^{e_p} \cdot \left(1 - \frac{1}{p^2}\right) = \\ &= N^2 \cdot \prod_{p|N} p^{e_p} \cdot \left(1 - \frac{1}{p}\right) \left(1 + \frac{1}{p}\right) = N^2 \cdot \phi(N) \cdot \prod_{p|N} \left(1 + \frac{1}{p}\right) \end{aligned}$$

Hence

$$\begin{aligned}
[SL(2, \mathbb{Z}) : \Gamma_0(N)] &= \frac{[SL(2, \mathbb{Z}) : \Gamma(N)]}{[\Gamma_0(N) : \Gamma(N)]} = \\
&= \frac{N^2 \cdot \phi(N) \cdot \prod_{p|N} \left(1 + \frac{1}{p}\right)}{N \cdot \phi(N)} = N \cdot \prod_{p|N} \left(1 + \frac{1}{p}\right)
\end{aligned}$$

□

Example 2.17 (Right cosets of congruence subgroups). See the PARI-command `mfcosets`.

1. The subgroup $\Gamma_0(2)$ has index

$$[SL(2, \mathbb{Z}) : \Gamma_0(2)] = 3$$

with the set of right coset representatives

$$\left\{ id, J := \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, U^2 = \begin{pmatrix} 0 & -1 \\ 1 & -1 \end{pmatrix} \right\}$$

for

$$U := \begin{pmatrix} -1 & 1 \\ -1 & 0 \end{pmatrix}$$

2. The subgroup $\Gamma_0(4)$ has index

$$[SL(2, \mathbb{Z}) : \Gamma_0(4)] = 6$$

with the set of right coset representatives

$$\left\{ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} 0 & -1 \\ 1 & 2 \end{pmatrix}, \begin{pmatrix} 0 & -1 \\ 1 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix} \right\}.$$

2.2 Topology of the orbit space of the $SL(2, \mathbb{Z})$ -action

The section continues the study of the moduli space of complex tori. We denote by

$$\Phi : \Gamma \times \mathbb{H} \rightarrow \mathbb{H}, (\gamma, \tau) \mapsto \gamma(\tau) := \frac{a\tau + b}{c\tau + d}, \gamma := \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

the continuous $SL(2, \mathbb{Z})$ -left action introduced in Remark 2.8. Our final aim is to provide the orbit space

$$Y := SL(2, \mathbb{Z}) \backslash \mathbb{H}$$

with an analytic structure and to compactify the resulting Riemann surface to obtain a compact Riemann surface X . The latter will be named the *modular curve* of the action Φ .

To provide the orbit space of a group action with a suitable topological or differentiable structure is always a subtle and challenging task. In the present context the problem to define a non-singular holomorphic structure is intensified by the fact that the modular group and their congruence subgroups do not operate freely, see Theorem 2.18: There are several elliptic points. In the present context one has to achieve the following tasks:

- Ensure that the quotient topology on Y with respect to the canonical projection

$$p : \mathbb{H} \rightarrow Y$$

is Hausdorff. See Theorem 2.25.

- Introduce a non-singular complex structure on Y also in a neighbourhood of the image of the elliptic points. See Section 2.3, Proposition 2.30.
- Extend the action of the congruence subgroups $\Gamma_0(N)$ on $\mathbb{H}$ to an action on the boundary $\mathbb{H}^* \setminus \mathbb{H}$. See Remark 2.8.
- Extend the Hausdorff topology on Y to a Hausdorff topology on the compactification X .
- Extend the complex structure of Y to a non-singular complex structure on X .

The modular group $SL(2, \mathbb{Z})$ contains the two distinguished elements

- *Reflection at the y-axis and dividing by the squared absolute value:*

$$S := \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \in SL(2, \mathbb{Z}), \quad S(\tau) = -1/\tau = \frac{-\bar{\tau}}{|\tau|^2}.$$

- *Translation by one:*

$$T := \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \in SL(2, \mathbb{Z}), \quad T(\tau) = \tau + 1.$$

Theorem 2.18 (Fundamental domain and elliptic points of the $SL(2, \mathbb{Z})$ -action).

1. Generators: The modular group $SL(2, \mathbb{Z})$ is generated by the two matrices $S, T \in \Gamma$ which satisfy the relations

$$S^4 = id, \quad (ST)^6 = id.$$

2. Fundamental domain: *The orbit space $Y = SL(2, \mathbb{Z}) \backslash \mathbb{H}$ of the action*

$$\Phi : SL(2, \mathbb{Z}) \times \mathbb{H} \rightarrow \mathbb{H}, (\gamma, \tau) \mapsto \gamma(\tau),$$

maps bijectively to the set

$$\mathcal{D} := \{\tau \in \mathbb{H} : |\tau| > 1 \text{ and } -1/2 \leq \operatorname{Re} \tau < 1/2\} \cup \{\tau \in \mathbb{H} : |\tau| = 1 \text{ and } -1/2 \leq \operatorname{Re} \tau \leq 0\},$$

the standard fundamental domain of Φ .

3. Elliptic points: *The action Φ has exactly two elliptic points $\tau_0 \in \mathcal{D}$. Their isotropy groups are cyclic:*

- $\tau_0 = i$ with

$$SL(2, \mathbb{Z})_i = \langle S \rangle \subset SL(2, \mathbb{Z}), \text{ ord } SL(2, \mathbb{Z})_i = 4, h_i = 2$$

- $\tau_0 = \rho := e^{2\pi i/3}$ with

$$SL(2, \mathbb{Z})_\rho = \langle ST \rangle \subset SL(2, \mathbb{Z}), \text{ ord } SL(2, \mathbb{Z})_\rho = 6, h_\rho = 3.$$

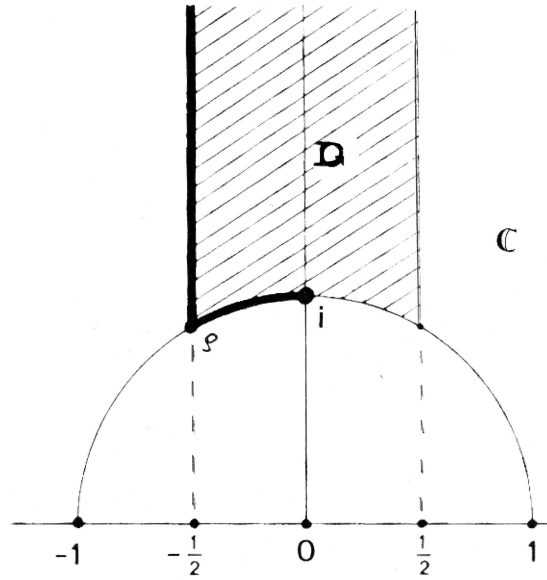


Fig. 2.1 Fundamental domain $\mathcal{D}$ of the modular group $SL(2, \mathbb{Z})$ shaded, from [31]

Figure 2.1 shows the fundamental domain $\mathcal{D}$ of the modular group. In order to exclude pairs of congruent points within $\mathcal{D}$ one has to include from the boundary

only the part marked in bold. This convention is followed by [31] but not by all references. Most references name *fundamental domain* the closure $\overline{\mathcal{D}}$.

Figure 2.2 shows some images of $\mathcal{D}$ under the action of the elements $T, S \in SL(2, \mathbb{Z})$ and their products.

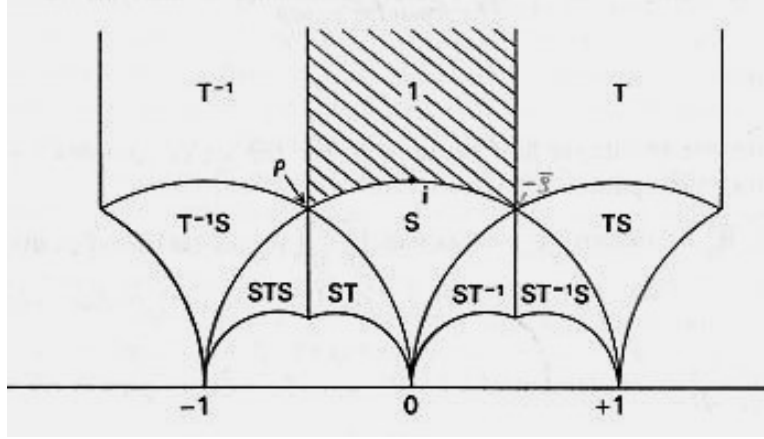


Fig. 2.2 Some translates of the fundamental domain $\mathcal{D}$ of Γ , from [56, Chap. VII, Fig. 1]

Proof. The three parts of the subsequent proof of Theorem 2.18 do not correspond one to one to the three parts of the theorem. Recall that the subgroup

$$\{\pm id\} \subset SL(2, \mathbb{Z})$$

operates on $\mathbb{H}$ in a trivial way and that $S^2 = -id$.

i) *The domain $\mathcal{D}$:* Denote by

$$G := \langle S, T \rangle \subset SL(2, \mathbb{Z})$$

the subgroup generated by the two distinguished elements. We show: Any given point $\tau \in \mathbb{H}$ is equivalent to a point $\tau' \in \mathcal{D} \pmod{G}$.

We choose an element $\gamma_0 \in G$ with $Im \gamma_0(\tau)$ maximal. Such choice is possible because

$$Im \gamma(\tau) = \frac{Im \tau}{|c \cdot \tau + d|^2} \text{ for } \gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in G.$$

The translation $T \in SL(2, \mathbb{Z})$ does not change the imaginary part of its argument. Therefore we may assume

$$-1/2 \leq \operatorname{Re} \gamma_0(\tau) < 1/2.$$

In addition, $|\gamma_0(\tau)| \geq 1$ because otherwise

$$\operatorname{Im} (S \circ \gamma_0)(\tau) = \operatorname{Im} \left(\frac{-\overline{\gamma_0(\tau)}}{|\gamma_0(\tau)|^2} \right) > \operatorname{Im} \gamma_0(\tau).$$

- If $|\gamma_0(\tau)| > 1$ then $\tau' := \gamma_0(\tau)$.
- If $|\gamma_0(\tau)| = 1$ but $0 < \operatorname{Re} \gamma_0(\tau) \leq 1/2$ then

$$\tau' := (S \circ \gamma_0)(\tau).$$

In both cases $\tau' \in \mathcal{D}$.

ii) *Isotropy groups and fundamental domain:* Consider two points $\tau, \tau' \in \mathcal{D}$ with

$$\tau' = \gamma(\tau), \gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL(2, \mathbb{Z}).$$

We show $\tau = \tau'$, and the latter equation implies $\gamma \in SL(2, \mathbb{Z})_\tau$.

We may assume without restriction $\operatorname{Im} \tau' \geq \operatorname{Im} \tau$. Then

$$\operatorname{Im} \tau' = \frac{\operatorname{Im} \tau}{|c \cdot \tau + d|^2}.$$

implies

$$|c \cdot \tau + d| \leq 1.$$

Because $\operatorname{Im} \tau \geq \sqrt{3}/2$ only the cases $c = 0, \pm 1$ are possible:

- *Case 1, $c = 0$:* Then $d = \pm 1$ and $a = d$. Without restriction $a = d = 1$, hence

$$\tau' = \tau + b, \quad b \in \mathbb{Z}.$$

Because $\operatorname{Re} \tau' = (\operatorname{Re} \tau) + b$ we obtain $b = 0$; i.e. $\tau = \tau'$ and

$$\gamma = \pm \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \pm id \in SL(2, \mathbb{Z})_\tau.$$

- *Case 2, $c = 1$:* Then $|\tau + d| \leq 1$, which implies $d \in \{0, \pm 1\}$.

Case 2a, $d = 0$: If $d = 0$ then $b = -1$ and

$$\tau' = \frac{a \cdot \tau - 1}{\tau} = a - (1/\tau), \quad a \in \mathbb{Z}.$$

Due to $\tau \in \mathcal{D}$ the equation

$$|c \cdot \tau + d| = |\tau| \leq 1$$

implies $|\tau| = 1$. Hence

$$1/\tau = \bar{\tau} \text{ and } \tau' + \bar{\tau} = a \in \mathbb{Z}$$

which implies

$$\operatorname{Re} \tau' + \operatorname{Re} \tau = a \in \mathbb{Z} \text{ and } \operatorname{Im} \tau' = \operatorname{Im} \tau.$$

Therefore: Either $a = 0$ and $\tau = i$ which implies

$$\tau' = -(1/\tau) = i = \tau$$

and

$$\gamma = \pm \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \pm S \in SL(2, \mathbb{Z})_i.$$

Or $a = -1$ and $\tau = \rho$ which implies

$$\tau' = -1 - (1/\tau) = \rho = \tau$$

and

$$\gamma = \pm \begin{pmatrix} -1 & -1 \\ 1 & 0 \end{pmatrix} = \pm (ST)^2 \in SL(2, \mathbb{Z})_\rho.$$

Case 2b, $d = \pm 1$: If $d = \pm 1$ then

$$|c \cdot \tau + d| = |\tau \pm 1| \leq 1.$$

Because of $\tau, \tau' \in \mathcal{D}$ the case $d = -1$ with $|\tau - 1| \leq 1$ is excluded. Therefore $d = 1$ leaving

$$|\tau + 1| = |\tau - (-1)| \leq 1.$$

As a consequence, $\tau = \rho$ and

$$ad - bc = a - b = 1 \text{ and } \tau' = \frac{a \cdot \rho + b}{\rho + 1} = a + \rho,$$

hence $a = 0$, $\tau' = \rho = \tau$ and

$$\gamma = \begin{pmatrix} 0 & -1 \\ 1 & 1 \end{pmatrix} = ST \in SL(2, \mathbb{Z})_\rho.$$

- Case 3, $c = -1$: This case reduces to case 2, $c = 1$, by considering $-\gamma_0$.

iii) $G = SL(2, \mathbb{Z})$: Consider an arbitrary element $\gamma \in SL(2, \mathbb{Z})$ and the point

$$\tau_0 := 2 \cdot i \in \mathcal{D}.$$

Due to part i) an element $\gamma_0 \in G$ exists with

$$(\gamma_0 \circ \gamma)(\tau_0) \in \mathcal{D}.$$

Due to part ii) the two points of $\mathcal{D}$ coincide

$$\tau_0 = (\gamma_0 \circ \gamma)(\tau_0) \in \mathcal{D},$$

i.e.

$$\gamma_0 \circ \gamma \in SL(2, \mathbb{Z})_{\tau_0} = \{\pm id\}$$

which implies $\gamma \in G$ $\square$

Remark 2.19 (Presentation of the modular group and fundamental domain of its subgroups).

1. Consider the matrix

$$U := -TS = \begin{pmatrix} -1 & 1 \\ -1 & 0 \end{pmatrix} \in SL(2, \mathbb{Z})$$

which satisfies $U^3 = id$. Apparently also

$$G = \langle S, U \rangle.$$

Therefore Theorem 2.18 implies that the two matrices S and U generate the group $SL(2, \mathbb{Z})$. One can show that all relations of the generators are generated by

$$S^4 = U^3 = S^2 U \cdot (US^2)^{-1} = id,$$

i.e. that

$$SL(2, \mathbb{Z}) = \langle S, U; S^4, U^3, S^2 U \cdot (US^2)^{-1} \rangle$$

is a presentation of $SL(2, \mathbb{Z})$, see [38, Kap. II, §2 Bemerk.].

2. Denote by

$$PSL(2, \mathbb{Z}) := SL(2, \mathbb{Z}) / \{\pm id\}$$

the corresponding *projective-linear group*, which acts faithful on $\mathbb{H}$, and by

$$\pi : SL(2, \mathbb{Z}) \rightarrow PSL(2, \mathbb{Z}), A \mapsto \bar{A},$$

the canonical quotient map. The projective-linear group has the two cyclic subgroups

$$H_1 := \langle \bar{S} \rangle \simeq \mathbb{Z}/2\mathbb{Z} \text{ and } H_2 := \langle \bar{U} \rangle \simeq \mathbb{Z}/3\mathbb{Z}.$$

One can show

$$PSL(2, \mathbb{Z}) = H_1 * H_2 = \langle \bar{S}, \bar{U}; \bar{S}^2, \bar{U}^3 \rangle$$

as the *free product*, see the short argument from [3] and also [56, Chap. VII, § 1, Rem.].

3. For a subgroup $K \subset \Gamma$ with finite index $J := [\Gamma : K] < \infty$ each decomposition into left cosets

$$\Gamma = \dot{\bigcup}_{j=1, \dots, J} \gamma_j \cdot K, \quad \gamma_j \in \Gamma,$$

provides a fundamental domain

$$\mathcal{D}(K) := \dot{\bigcup}_{j=1, \dots, J} \gamma_j^{-1}(\mathcal{D})$$

of the induced K -action, which derives from the standard fundamental domain $\mathcal{D}$ of Γ . In the following we will always choose fundamental domains $\mathcal{D}(K) \subset \mathbb{H}$ which originate as translates of $\mathcal{D}$. These fundamental domains are measurable subsets of $\mathbb{H}$.

Corollary 2.20 (Elliptic points). *For $N \geq 1$ the action*

$$\Phi : \Gamma_0(N) \times \mathbb{H} \rightarrow \mathbb{H}$$

has finitely many elliptic orbits. Each elliptic point $\tau \in \mathbb{H}$ has the finite isotropy group

$$\Gamma_0(N)_\tau = SL(2, \mathbb{Z})_\tau \cap \Gamma_0(N).$$

Proof. Theorem 2.18 shows: The action of $SL(2, \mathbb{Z})$ has exactly two elliptic points $i, \rho \in \mathcal{D}$, and both have finite isotropy groups. The index formula from Proposition 2.16 implies the finiteness of

$$[SL(2, \mathbb{Z}) : \Gamma_0(N)] =: k.$$

According to Remark 2.19 each finite coset decomposition

$$SL(2, \mathbb{Z}) = \dot{\bigcup}_{j=1}^k \gamma_j \cdot \Gamma_0(N), \quad \gamma_j \in SL(2, \mathbb{Z}),$$

provides the fundamental domain of the $\Gamma_0(N)$ -action

$$\mathcal{D}(\Gamma_0(N)) := \dot{\bigcup}_{j=1}^k \gamma_j^{-1}(\mathcal{D})$$

e.g. see Figure 2.3. For each $j = 1, \dots, k$ and each point $\tau \in \gamma_j^{-1}(\mathcal{D})$ the group morphism

$$\Gamma_0(N)_\tau \rightarrow SL(2, \mathbb{Z})_{\gamma_j(\tau)}, \quad \gamma \mapsto \gamma_j \cdot \gamma \cdot \gamma_j^{-1},$$

injects $\Gamma_0(N)_\tau$ into $SL(2, \mathbb{Z})_{\gamma_j(\tau)}$. Hence the elliptic points of $\Gamma_0(N)$ contained in $\mathcal{D}(\Gamma_0(N))$ are contained in the finite set

$$\dot{\bigcup}_{j=1}^k \gamma_j^{-1}(i) \cup \dot{\bigcup}_{j=1}^k \gamma_j^{-1}(\rho)$$

□

Figure 2.3 shows the fundamental domain of $\Gamma_0(13)$. The fundamental domain $\mathcal{D}(\Gamma_0(13))$ splits into 14 disjoint translates of $\mathcal{D}$ according to the index formula from Proposition 2.16

$$[SL(2, \mathbb{Z}) : \Gamma_0(13)] = 14$$

For further explanations of Figure 2.3 see [19, Fig. 3.1 and Exerc. 3.1.4].

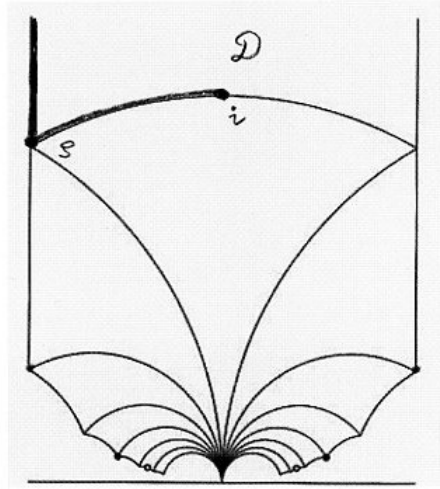


Fig. 2.3 Partition of the fundamental domain of $\Gamma_0(13)$, adapted from [19, Fig. 3.1]

Lemma 2.21 covers the main step in proving the Hausdorff property of the quotient topology with respect to the quotient map

$$\mathbb{H} \rightarrow SL(2, \mathbb{Z}) \backslash \mathbb{H}.$$

Lemma 2.21 (Action of the modular group). *Any two points $\tau_1, \tau_2 \in \mathbb{H}$ have neighbourhoods*

$$U_i \subset \mathbb{H}, \quad i = 1, 2,$$

such that for all $\gamma \in \Gamma$:

$$\gamma(U_1) \cap U_2 \neq \emptyset \implies \gamma(\tau_1) = \tau_2.$$

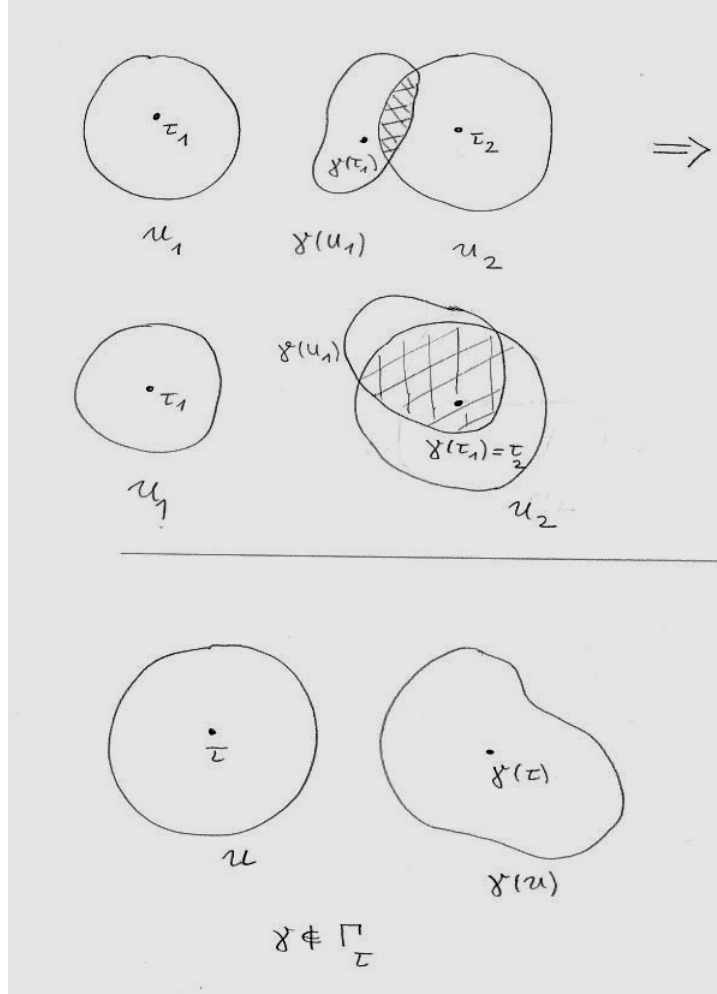


Fig. 2.4 Local behaviour of the action of the modular group, see Lemma 2.21

Figure 2.4 visualizes the content of Lemma 2.21: If an element $\gamma \in SL(2, \mathbb{Z})$ moves a point of the distinguished neighbourhood U_1 of τ_1 to the distinguished neighbourhood U_2 of τ_2 , then γ maps τ_1 to τ_2 .

Proof. For the proof of Lemma 2.21 compare [19, Chap. 2.1].

1. *The isotropy group of $i \in \mathbb{H}$:* We choose two relatively compact neighbourhoods

$$V_j \subset \subset \mathbb{H} \text{ of } \tau_j, \quad j = 1, 2,$$

and show: Only finitely many matrices $\gamma \in SL(2, \mathbb{Z})$ exists with

$$\gamma(V_1) \cap V_2 \neq \emptyset.$$

The idea is to extend the group action of the discrete group $SL(2, \mathbb{Z})$ to the real group $SL(2, \mathbb{R})$, see Remark 2.8,

$$\Phi : SL(2, \mathbb{R}) \times \mathbb{H} \rightarrow \mathbb{H}, (\gamma, \tau) \mapsto \gamma(\tau) := \frac{a\tau + b}{c\tau + d}, \gamma := \begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

which provides a compact isotropy group of the point i

$$SL(2, \mathbb{R})_i = SO(2, \mathbb{R}) :$$

One checks

$$\gamma(i) = i \iff a = d \text{ and } b = -c,$$

hence

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SO(2, \mathbb{R}).$$

The action Φ of $SL(2, \mathbb{R})$ is transitive because any point $\tau \in \mathbb{H}$ belongs to the orbit of i : If

$$\tau = x + iy$$

then $\gamma_\tau(i) = \tau$ for the matrix

$$\gamma_\tau := \begin{pmatrix} \sqrt{y} & \frac{x}{\sqrt{y}} \\ 0 & \frac{1}{\sqrt{y}} \end{pmatrix} = \frac{1}{\sqrt{y}} \cdot \begin{pmatrix} y & x \\ 0 & 1 \end{pmatrix} \in SL(2, \mathbb{R}).$$

As a consequence for any two points $e_1, e_2 \in \mathbb{H}$ and all $\gamma \in SL(2, \mathbb{R})$:

$$\gamma(e_1) = e_2 \iff \gamma \in \gamma_{e_2} \cdot SO(2, \mathbb{R}) \cdot \gamma_{e_1}^{-1}$$

according to the following commutative diagram

$$\begin{array}{ccc} e_1 & \xrightarrow{\gamma} & e_2 \\ \gamma_{e_1}^{-1} \downarrow & & \uparrow \gamma_{e_2} \\ i & \dashrightarrow & i \end{array}$$

If a matrix $\gamma \in SL(2, \mathbb{Z})$ satisfies $\gamma(V_1) \cap V_2 \neq \emptyset$ then also for the closure

$$\gamma(\bar{V}_1) \cap \bar{V}_2 \neq \emptyset$$

and

$$G := \{\gamma \in SL(2, \mathbb{R}) : \gamma(\bar{V}_1) \cap \bar{V}_2 \neq \emptyset\} = \bigcup_{e_1 \in \bar{V}_1, e_2 \in \bar{V}_2} \gamma_{e_2} \cdot SO(2, \mathbb{R}) \cdot \gamma_{e_1}^{-1}.$$

Because γ_e and its inverse γ_e^{-1} depend continuously on $e \in \mathbb{H}$, the union on the right-hand side is the continuous image of the compact set

$$\bar{V}_2 \times SO(2, \mathbb{R}) \times \bar{V}_1$$

under the continuous map

$$\bar{V}_2 \times SO(2, \mathbb{R}) \times \bar{V}_1 \rightarrow SL(2, \mathbb{R}), (e_2, \gamma, e_1) \mapsto \gamma_{e_2} \cdot \gamma \cdot \gamma_{e_1}^{-1},$$

and therefore compact itself. Its intersection

$$G \cap SL(2, \mathbb{Z})$$

with the discrete group $SL(2, \mathbb{Z})$ is finite.

2. *Shrinking V_1 and V_2 :* According to the first part only finitely many $\gamma \in SL(2, \mathbb{Z})$ exist with

$$\gamma(V_1) \cap V_2 \neq \emptyset.$$

In particular, the set

$$G_\Gamma := \{\gamma \in SL(2, \mathbb{Z}) : \gamma(V_1) \cap V_2 \neq \emptyset \text{ and } \gamma(\tau_1) \neq \tau_2\}$$

is finite. We now shrink in finitely many steps the neighbourhoods V_1 and V_2 to neighbourhoods U_1 and U_2 with

$$\gamma(U_1) \cap U_2 = \emptyset \text{ iff } \gamma(\tau_1) \neq \tau_2 :$$

For each $\gamma \in G_\Gamma$ we have

$$\gamma(\tau_1) \neq \tau_2.$$

Because $\mathbb{H}$ is a Hausdorff space we choose for each $\gamma \in G_\Gamma$ two disjoint neighbourhoods of respectively $\gamma(\tau_1)$ and τ_2 in $\mathbb{H}$

$$U_{1,\gamma} \subset \mathbb{H}, U_{2,\gamma} \subset \mathbb{H} \text{ and } U_{1,\gamma} \cap U_{2,\gamma} = \emptyset,$$

and define two neighbourhoods of respectively τ_1 and τ_2 in $\mathbb{H}$

$$U_1 := V_1 \cap \left(\bigcap_{\gamma \in G_\Gamma} \gamma^{-1}(U_{1,\gamma}) \right), U_2 := V_2 \cap \left(\bigcap_{\gamma \in G_\Gamma} U_{2,\gamma} \right).$$

Then for all $\gamma \in G_\Gamma$

$$\gamma(U_1) \cap U_2 = \emptyset.$$

□

The operation of the modular group on the complex upper half-plane is a specific discontinuous group action. There exist different similar but not identical definitions for the latter concept. We use the definition of a *strongly properly discontinuous group action* from [17, Def. 5].

Definition 2.22 (Strongly proper discontinuous group action). A topological group G acts *properly strongly discontinuous* on a topological Hausdorff space X

$$G \times X \rightarrow X$$

if each pair (x, y) of not necessarily distinct points $x, y \in X$ has neighbourhoods

$$U = U(x) \text{ and } V = V(y)$$

such that

$$\{g \in G : g(U) \cap V \neq \emptyset\}$$

is a finite set.

Note: Remark 2.26 will show that a topological group which acts strongly properly discontinuous is necessarily a discrete group.

Proposition 2.23 (Strong properly discontinuous action of the modular group).

The modular group $SL(2, \mathbb{Z})$ acts strongly properly discontinuously on the complex upper half-plane $\mathbb{H}$.

Proof. Consider two points $\tau_1, \tau_2 \in \mathbb{H}$. We choose two neighbourhoods

$$U_1 \text{ of } \tau_1 \text{ and } U_2 \text{ of } \tau_2$$

with the properties from Lemma 2.21. Set

$$G' := \{g \in G : g(U_1) \cap U_2 \neq \emptyset\}$$

According to Lemma 2.21 each $g \in G'$ satisfies

$$g(\tau_1) = \tau_2$$

If $G' \neq \emptyset$ then choose a fixed element $g_0 \in G'$. Any $g \in G'$ satisfies

$$g_0^{-1} \circ g \in G_{\tau_1}$$

Therefore

$$G' = G_{\tau_1} \circ g_0$$

The isotropy group G_{τ_1} is finite, also for the elliptic points of the Γ -action. Hence the set G' is finite, which finishes the proof.

□

Lemma 2.24 is a general result about the orbit space of a group action. It relies on a basic topological fact.

Lemma 2.24 (Openness of the canonical projection onto the orbit space).

Denote by

$$p : \mathbb{H} \rightarrow Y := SL(2, \mathbb{Z}) \backslash \mathbb{H}$$

the canonical projection onto the orbit space of the action

$$\Phi : SL(2, \mathbb{Z}) \times \mathbb{H} \rightarrow \mathbb{H}$$

and provide Y with the corresponding quotient topology. Then p is an open map.

Proof. By definition of the quotient topology a subset $U \subset Y$ is open iff

$$p^{-1}(U) \subset \mathbb{H}$$

is open.

For each arbitrary but fixed $\gamma \in SL(2, \mathbb{Z})$ its orbit map, the restriction

$$\Phi(\gamma, -) : \mathbb{H} \rightarrow \mathbb{H}, \tau \mapsto \gamma(\tau),$$

is continuous with inverse the orbit map of γ^{-1}

$$\Phi(\gamma^{-1}, -) : \mathbb{H} \rightarrow \mathbb{H}, \tau \mapsto \gamma^{-1}(\tau).$$

Hence the orbit map $\Phi(\gamma, -)$ is a homeomorphism.

In order to prove the openness of p we consider an open set $U \subset \mathbb{H}$ and prove the openness of $p(U) \subset Y$: The inverse image

$$p^{-1}(p(U)) = \bigcup_{\gamma \in SL(2, \mathbb{Z})} \gamma(U)$$

is open as union of open sets. Hence

$$p(U) \subset Y$$

is open by definition of the quotient topology.

□

Theorem 2.25 (Hausdorff topology of the orbit space). *The quotient topology on the orbit space*

$$Y := SL(2, \mathbb{Z}) \backslash \mathbb{H}$$

of the $SL(2, \mathbb{Z})$ -action is Hausdorff.

Proof. Consider two distinct points $y_1 \neq y_2 \in Y$.

i) We choose two inverse images

$$\tau_j \in p^{-1}(y_j) \in \mathbb{H}, \quad j = 1, 2$$

Then $\tau_1 \neq \tau_2$ and both points belong to different orbits of Γ . Hence for all $\gamma \in SL(2, \mathbb{Z})$

$$\gamma(\tau_1) \neq \tau_2$$

Proposition 2.23 provides two open neighbourhoods $U_j \subset \mathbb{H}$ of respectively τ_1 and of τ_2 such that the set

$$\{\gamma \in SL(2, \mathbb{Z}) : \gamma(U_1) \cap U_2 \neq \emptyset\}$$

is finite. After finitely many shrinking steps for $U_j, j = 1, 2$ one obtains for all $\gamma \in SL(2, \mathbb{Z})$

$$\gamma(U_1) \cap U_2 = \emptyset$$

which implies

$$p(U_1) \cap p(U_2) = \emptyset.$$

ii) Lemma 2.24 implies that the sets $p(U_j) \subset Y, j = 1, 2$ are open. Hence there are disjoint open neighbourhoods of respectively

$$y_1 = p(\tau_1) \text{ and } y_2 = p(\tau_2)$$

□

With Theorem 2.25 the present section has obtained its goal: The quotient topology on the orbit space $\Gamma \backslash \mathbb{H}$ of the action of the modular group is Hausdorff.

In the rest of this section we make two further steps into the domain of group actions and the topology of quotient spaces of equivalence relations. Remark 2.26 shows how Lemma 2.21 integrates the action of the modular group and the formation of its orbit space $SL(2, \mathbb{Z}) \backslash \mathbb{H}$ into the general theory of proper group actions. The Hausdorff property of the quotient topology on $\Gamma \backslash \mathbb{H}$ - notwithstanding the existence of elliptic points - is the special case of a general result about proper group actions.

Remark 2.26 (The context of proper group actions). Consider the left action of a topological group G on a Hausdorff space X

$$\phi : G \times X \rightarrow X, (g, x) \mapsto g(x)$$

and its induced map

$$\Phi : G \times X \rightarrow X \times X, (g, x) \mapsto (g(x), x)$$

If

$$pr_G : G \times X \rightarrow G$$

denotes the first projection, then each $x \in X$ has the isotropy group

$$G_x = pr_G(\Phi^{-1}((x, x)))$$

1. By definition, the action ϕ is *proper* if and only if Φ is a proper map. See [10, Part III, § 4.1, Def. 1].
2. If the action ϕ is proper, then for each pair of compact sets $K_j \subset X$, $j = 1, 2$, the set

$$\{g \in G : g(K_1) \cap K_2 \neq \emptyset\}$$

is compact.

Proof. The set

$$K_1 \times K_2 \subset X \times X$$

is compact, hence by assumption also its inverse image

$$\Phi^{-1}(K_1 \times K_2) \subset G \times X$$

The latter set projects under the canonical projection

$$pr_G : G \times X \rightarrow G,$$

onto a compact set, namely the set

$$\{g \in G : g(K_1) \cap K_2 \neq \emptyset\}$$

□

3. For a proper action ϕ the quotient topology on the orbit space $G \backslash X$ is Hausdorff. See [10, Part III, § 4.2, Prop. 3].
4. By definition, the action ϕ is *discontinuous* at a point $x \in X$, if the orbit $Gx \subset X$ is a closed discrete set, and the isotropy group $G_x := \{g \in G : gx = x\}$ is finite.
5. For G and X locally compact: If the action ϕ is proper and discontinuous at a point $x \in X$, then it is strong proper discontinuous in the sense of Definition 2.22.

Proof. i) The orbit map of x

$$\phi_x : G \rightarrow X, g \mapsto g(x),$$

is continuous. By definition of discontinuity the orbit Gx is discrete. Hence $x \in Gx$ has an open neighbourhood $U \subset X$ with

$$Gx \cap U = \{x\}$$

The inverse image $\phi_x^{-1}(U) \subset G$ is open and equals the isotropy group

$$\phi_x^{-1}(U) = G_x$$

The latter is finite by definition of discontinuity. Because the topological group G is Hausdorff, one can shrink $\phi_x^{-1}(U)$ to an open neighbourhood V of $e \in G$ with

$$V = \{e\}$$

This shows that G is discrete.

ii) For two points $x, y \in X$ and relatively compact neighbourhoods $U = U(x)$ and $V = V(y)$ the set

$$\overline{U} \times \overline{V} \subset X \times X$$

is compact. Properness of ϕ implies that the set

$$pr_G(\Phi^{-1}(\overline{U} \times \overline{V})) = \{g \in G : g(\overline{U}) \cap \overline{V} \neq \emptyset\}$$

is compact, hence finite because G is discrete. Therefore the action ϕ is strong proper discontinuous.

□

6. If the action ϕ is strong proper discontinuous in the sense of Definition 2.22, then G is discrete and the action is discontinuous and proper. The quotient topology on the orbit space $G \backslash X$ is Hausdorff.

Proof. i) *Discreteness:* Notably, the strong proper discontinuous action ϕ is a proper discontinuous action in the sense of [17, Def. 3]: Each $x \in X$ has an open neighbourhood $U = U(x)$ with

$$\{g \in G : g(U) \cap U \neq \emptyset\}$$

a finite set. According to [17, Sect. 1.3] the group G and all orbits Gx , $x \in X$, are discrete.

ii) *Properness:* Because the action ϕ is strong proper discontinuous, due to [17, Rem. 5] for all compact sets K , $L \subset X$ the inverse image

$$\Phi^{-1}(K \times L) = \{g \in G : g(K) \cap L \neq \emptyset\}$$

is finite, hence in particular compact. Therefore the action of ϕ is proper.

As a consequence the map Φ is closed. Because $G \times \{x\} \subset G \times X$ is closed, also its image

$$\Phi(G \times \{x\}) = \bigcup_{g \in G} \{(g(x), x)\} \subset X \times X$$

is closed. The latter set is homeomorphic to the orbit $Gx \subset X$, which is therefore closed too.

iii) *Discontinuity*: The isotropy group of a point $x \in X$ is

$$G_x = \Phi^{-1}(K \times L)$$

with $K = L = x$. Hence part ii) implies that G_x is finite. Together with the discreteness of the orbit $Gx \subset X$ from part i) and its closedness from part ii) this shows that the action ϕ is discontinuous.

iv) *The Hausdorff property*: The properness of the action ϕ implies according to step 3 that the orbit space $G \backslash X$ is Hausdorff.

□

The action of the modular group on $\mathbb{H}$ and the induced Hausdorff topology on the orbit space

$$\Gamma \backslash \mathbb{H}$$

also relate to a second general question: Which properties of an equivalence relation R on a manifold X ensure that the quotient topology on the set of equivalence classes is Hausdorff?

Lemma 2.27 (Hausdorff criterion for an equivalence relation). *Let X be a topological space and*

$$R \subset X \times X$$

an equivalence relation on X . Then the set of equivalence classes X/R provided with the quotient topology with respect to the canonical map

$$p : X \rightarrow X/R$$

is a Hausdorff space if both of the following conditions are satisfied:

- *The map p is open*
- *The equivalence relation $R \subset X \times X$ is closed.*

Proof. See [20, Chap. VII, 1.6]: Consider $x, y \in X$ with

$$p(x) \neq p(y)$$

as equivalence classes in X/R . Then $(x, y) \notin R$. Because $R \subset X \times X$ is closed, there is a neighbourhood

$$U \times V \subset (X \times X) \setminus R$$

of $(x, y) \in X \times X$. Openness of p implies the existence of neighbourhoods

$$p(U), p(V) \subset X/R$$

of respectively $p(x)$ and $p(y)$ with

$$p(U) \cap p(V) = \emptyset$$

□

With the new concepts and results we obtain two variants of a second proof for the Hausdorff property from Theorem 2.25.

Corollary 2.28 (Hausdorff topology on the orbit space of the modular group).
The quotient topology on the orbit space

$$Y := SL(2, \mathbb{Z}) \backslash \mathbb{H}$$

is Hausdorff.

Proof. Denote by

$$\Phi : SL(2, \mathbb{Z}) \times \mathbb{H} \rightarrow \mathbb{H} \times \mathbb{H}$$

the canonical map of the $SL(2, \mathbb{Z})$ -action. According to Proposition 2.23 the action of $SL(2, \mathbb{Z})$ on $\mathbb{H}$ is strong proper discontinuous and therefore proper due to Remark 2.26.

i) *First variant.* Hence the map Φ is closed and

$$R = \Phi(SL(2, \mathbb{Z}) \times \mathbb{H}) \subset \mathbb{H} \times \mathbb{H}$$

is a closed equivalence relation on $\mathbb{H}$. In addition, Lemma 2.24 states that p is an open map. Then Lemma 2.27 implies the claim.

ii) *Second variant:* According to Remark 2.26, part 3 the orbit space of a proper group action is Hausdorff.

□

Note. Because the orbit space Y is Hausdorff, each orbit

$$p^{-1}(y) \subset \mathbb{H}, y \in Y,$$

of the $SL(2, \mathbb{Z})$ -action is closed.

2.3 The modular curves $X(\Gamma_0(N))$ as compact Riemann surfaces

The elliptic points of the group action are those points which complicate the introduction of a complex structure on the orbit space Y . To overcome this difficulty we make a detailed study of the group action in the neighbourhood of an elliptic point. Fortunately, the isotropy group of the elliptic points turns out to be a group of rotations.

The statement of the following Lemma 2.29 is trivial for non-elliptic points. For an elliptic point $\tau \in \mathbb{H}$ with period h_τ the isotropy group acts like the group of rotations of a disk around the origin with angles a multiple of $2\pi/h_\tau$. This follows from the Lemma of Schwarz.

Lemma 2.29 (Isotropy group acting as group of rotations). *For any point $\tau \in \mathbb{H}$ exist*

- *a neighbourhood $U_\tau \subset \mathbb{H}$ of τ such that for all $\gamma \in SL(2, \mathbb{Z})$*

$$\gamma(U_\tau) \cap U_\tau \neq \emptyset \implies \gamma \in SL(2, \mathbb{Z})_\tau,$$

- *and a fractional linear transformation $\lambda_\tau \in GL(2, \mathbb{C})$ satisfying*

$$\lambda_\tau : U_\tau \xrightarrow{\cong} \Delta_\tau$$

with a disc $\Delta_\tau = \{z \in \Delta : |z| < r_\tau\}$ for a suitable radius $r_\tau > 0$

such that: For all $\gamma \in SL(2, \mathbb{Z})_\tau$ the conjugate map

$$\tilde{\gamma}_\tau := \lambda_\tau \circ \gamma \circ \lambda_\tau^{-1} : \Delta_\tau \rightarrow \Delta_\tau$$

from the commutative diagram

$$\begin{array}{ccc} U_\tau & \xrightarrow{\gamma} & \gamma(U_\tau) \\ \lambda_\tau \downarrow & & \downarrow \lambda_\tau \\ \Delta_\tau & \xrightarrow{\tilde{\gamma}_\tau} & \Delta_\tau \end{array}$$

is a rotation around $0 \in \Delta_\tau$. The angle of rotation is an integer multiple of $2\pi/h_\tau$ with h_τ the period of τ .

Proof. Due to Theorem 2.18 w.l.o.g.

$$\tau \in \{i, \rho\}.$$

Denote by

$$\lambda_\tau : \mathbb{P}^1 \rightarrow \mathbb{P}^1, \lambda_\tau(z) := \frac{z - \tau}{z - \bar{\tau}},$$

the fractional linear automorphism of $\mathbb{P}^1$ with $\lambda(\mathbb{H}) = \Delta$ and

$$\lambda_\tau(\tau) = 0, \lambda_\tau(\bar{\tau}) = \infty.$$

The following diagram commutes for all $\gamma \in SL(2, \mathbb{Z})_\tau$

$$\begin{array}{ccc} \mathbb{H} & \xrightarrow{\gamma} & \mathbb{H} \\ \lambda_\tau \downarrow & & \downarrow \lambda_\tau \\ \Delta & \xrightarrow{\tilde{\gamma}_\tau} & \Delta \end{array}$$

and $\tilde{\gamma}_\tau$ has the fixed point $0 \in \Delta$. The non-trivial isotropy groups of the two fixed points are

$$SL(2, \mathbb{Z})_i = \langle S \rangle \text{ and } SL(2, \mathbb{Z})_\rho = \langle ST \rangle.$$

For $\tau \in \{i, \rho\}$ one checks

$$|\tilde{\gamma}'_\tau(0)| = |\gamma'(\tau)| = 1$$

because in particular

$$|S'(i)| = |(ST)'(\rho)| = 1.$$

The Lemma of Schwarz implies that

$$\tilde{\gamma}_\tau : \Delta \rightarrow \Delta$$

is a rotation

$$\tilde{\gamma}_\tau(z) = e^{k \cdot (2\pi i / h_\tau)} \cdot z, \quad k \in \{0, 1, \dots, h_\tau - 1\}$$

We choose a disk $\Delta_\tau \subset \Delta$ such that the neighbourhood of $\tau \in \mathbb{H}$

$$U_\tau := \lambda_\tau^{-1}(\Delta_\tau)$$

satisfies the properties from Lemma 2.21. Then restricting the above diagram proves the claim.

□

In order to obtain a complex structure on the orbit space, the Hausdorff space

$$Y = SL(2, \mathbb{Z}) \backslash \mathbb{H},$$

we introduce charts on Y with holomorphic transition functions using the canonical projection

$$p : \mathbb{H} \rightarrow Y.$$

Around an elliptic point $\tau \in \mathbb{H}$ the isotropy group $SL(2, \mathbb{Z})_\tau$ acts like a cyclic group of rotations of a disk Δ_τ due to Lemma 2.29. Recall from Theorem 2.18 the periods $h_i = 2$ and $h_\rho = 3$. If τ has the period

$$k := h_\tau = \frac{\text{ord } SL(2, \mathbb{Z})_\tau}{2}$$

then in a neighbourhood Δ_τ of $\lambda_\tau(\tau)$ the function

$$z \mapsto z^k$$

foliates the k different segments of the disk Δ_τ , which are equivalent under the $SL(2, \mathbb{Z})_\tau$ -action, to an open set $W_\tau \subset \mathbb{C}$, see Figure 2.5, and induces a chart

$$p(U_\tau) \xrightarrow{\phi_\tau} W_\tau.$$

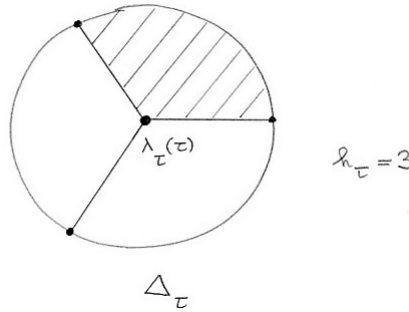


Fig. 2.5 The three equivalent segments of the disk Δ_τ of an elliptic point $\tau \in \mathbb{H}$ with $h_\tau = 3$

Proposition 2.30 (Complex structure of the orbit space). *On the orbit space*

$$Y = SL(2, \mathbb{Z}) \backslash \mathbb{H}$$

exists the complex structure of a Riemann surface with holomorphic canonical projection

$$p : \mathbb{H} \rightarrow Y$$

Note that from a general point of view the open Riemann surface Y is the *coarse moduli space* of complex tori.

Proof. i) *Local homeomorphisms:* For each $\tau \in \mathbb{H}$ choose a neighbourhood $U_\tau \subset \mathbb{H}$ of τ with the properties from Lemma 2.29. Using the notation from the lemma define

$$pow_\tau : \Delta_\tau \rightarrow \mathbb{C}, z \mapsto z^{h_\tau},$$

and set

$$W_\tau := pow_\tau(\Delta_\tau)$$

Eventually, define a chart of Y around $y = p(\tau) \in Y$

$$\phi_\tau : p(U_\tau) \rightarrow W_\tau$$

as the commutative completion of the following diagram

$$\begin{array}{ccc} U_\tau & \xrightarrow{p|_{U_\tau}} & p(U_\tau) \\ \lambda_\tau \downarrow & & \downarrow \phi_\tau \\ \Delta_\tau & \xrightarrow{pow_\tau} & W_\tau \end{array}$$

The map ϕ_τ is

- well-defined and bijective according to Lemma 2.29 because for all $\gamma \in SL(2, \mathbb{Z}) \setminus SL(2, \mathbb{Z})_\tau$

$$\gamma(U_\tau) \cap U_\tau = \emptyset.$$

The different inverse images under $p|_{U_\tau}$ of a given point from $p(U_\tau)$ are identified by $pow_\tau \circ \lambda_\tau$,

- continuous, because p is open,
- open, because λ_τ and pow_τ are open.

As a consequence, ϕ_τ is a homeomorphism.

ii) *Explicit form of the transition functions:* We show that the family

$$\mathfrak{A} := (\phi_\tau : p(U_\tau) \rightarrow W_\tau)_{\tau \in \mathbb{H}}$$

is a complex atlas of Y .

Assume two charts

$$\phi_{\tau_j} : p(U_{\tau_j}) \rightarrow W_{\tau_j}$$

around points $y_j = p(\tau_j) \in Y$, $j = 1, 2$, and assume

$$p(U_{\tau_1}) \cap p(U_{\tau_2}) \neq \emptyset.$$

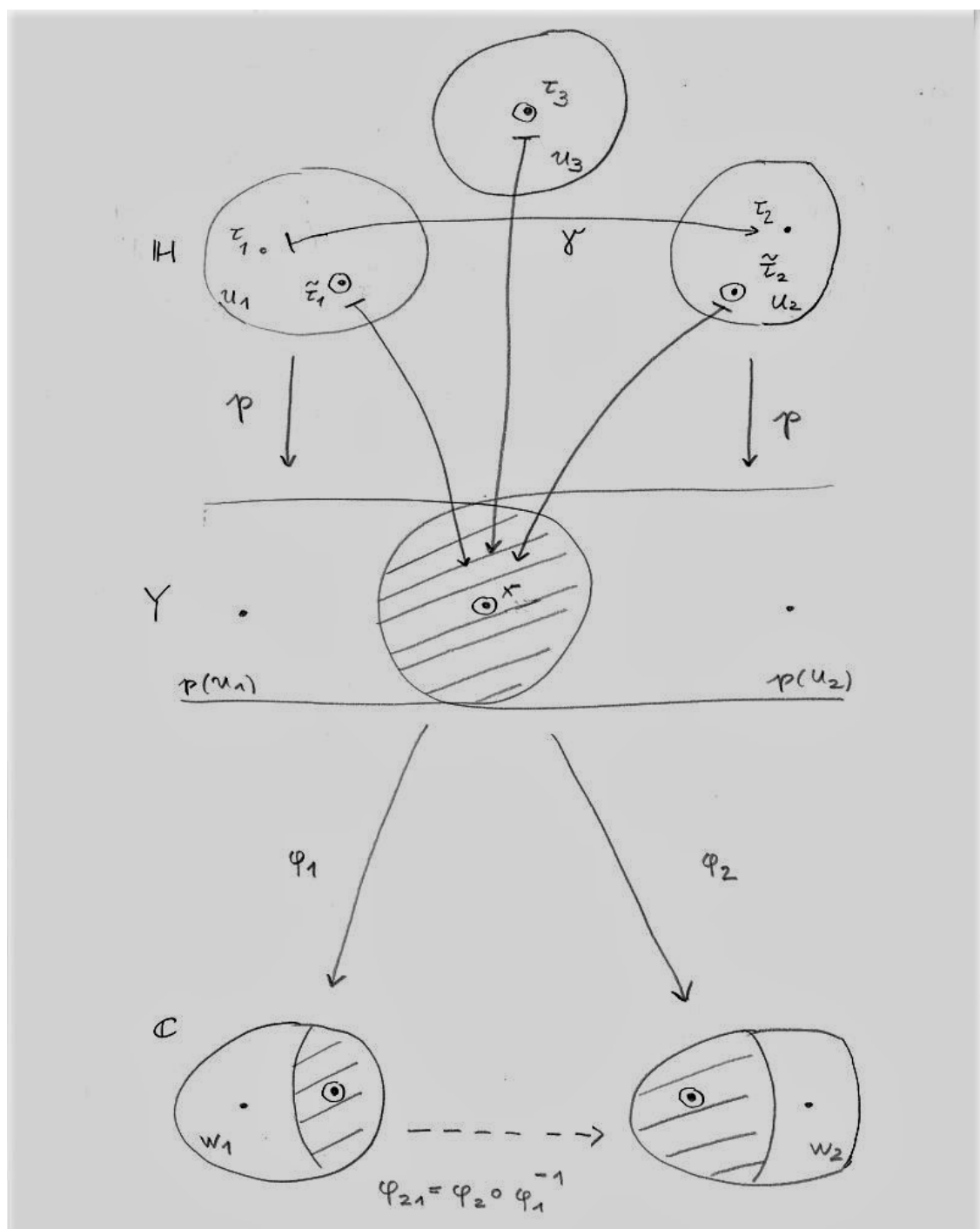
To simplify the subscripts we set

$$U_j := U_{\tau_j}, \phi_j := \phi_{\tau_j}, \lambda_j := \lambda_{\tau_j}, h_j := h_{\tau_j}, W_{\tau_j} := W_j, j = 1, 2.$$

Consider a point

$$x \in p(U_1) \cap p(U_2),$$

see Figure 2.6.

Fig. 2.6 Transition functions of Y

We show that the transition function

$$\phi_{2,1} := \phi_2 \circ \phi_1^{-1}$$

is holomorphic in a neighbourhood of $\phi_1(x) \in W_1 \subset \mathbb{C}$: By definition of ϕ_j two points $\tilde{\tau}_j \in U_j$, $j = 1, 2$, exist with

$$\lambda_i(\tilde{\tau}_j)^{h_j} = \phi_j(x), \quad j = 1, 2.$$

They satisfy $\tilde{\tau}_2 = \gamma(\tilde{\tau}_1)$ for a suitable $\gamma \in SL(2, \mathbb{Z})$. Hence for $q \in \mathbb{C}$ in a neighbourhood of $\phi_1(x)$ the transition functions attains the value

$$\phi_{2,1}(q) = \left((\lambda_2 \circ \gamma \circ \lambda_1^{-1})(q^{1/h_1}) \right)^{h_2}$$

We have to prove: Taking the h_1 -root does not prevent the holomorphy of the map.

iii) *Holomorphy of the transition functions*: We prove the non-trivial case $h_1 > 1$, i.e. we assume that the orbit $x \in X$ is elliptic.

- Case $\phi_1(x) = 0 \in W_1$, i.e. $\tilde{\tau}_1 = \tau_1$:

The point

$$\tau_1 = \tilde{\tau}_1$$

is elliptic. Therefore also $\tilde{\tau}_2 = \gamma(\tilde{\tau}_1) \in U_2$ is elliptic. Because τ_2 is the only elliptic point in U_2 we obtain $\tilde{\tau}_2 = \tau_2$, and

$$\tau_2 = \gamma(\tau_1) \text{ and } h := h_2 = h_1.$$

By construction of λ_j , $j = 1, 2$, the composition of fractional linear transformations

$$\lambda_2 \circ \gamma \circ \lambda_1^{-1}$$

is a fractional linear transformation of $\mathbb{P}^1$ mapping 0 to 0 and ∞ to ∞ . Therefore a matrix

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in GL(2, \mathbb{C})$$

exists with

$$\lambda_2 \circ \gamma \circ \lambda_1^{-1} = A$$

and

$$A \cdot (0 \ 1)^\top = (0 \ 1)^\top, \quad A \cdot (1 \ 0)^\top = (1 \ 0)^\top.$$

With respect to the injection

$$\mathbb{C} \hookrightarrow \mathbb{P}^1, \quad z \mapsto (z : 1),$$

- the vector $(0 \ 1)^\top \in \mathbb{C}^2$ represents the point $(0 : 1) \in \mathbb{P}^1$ and the point $0 \in \mathbb{C}$,

- while the vector $(1 \ 0)^\top \in \mathbb{C}^2$ represents the point $(1 : 0) \in \mathbb{P}^1$ and the point $\infty \in \mathbb{C} \cup \{\infty\}$.

As a consequence

$$A = \begin{pmatrix} a & 0 \\ 0 & d \end{pmatrix}$$

and

$$\phi_{2,1}(q) = \left((\lambda_2 \circ \gamma \circ \lambda_1^{-1})(q^{1/h}) \right)^h = \left(A(q^{1/h}) \right)^h = ((a \cdot q^{1/h})/d)^h = (a/d)^h \cdot q.$$

Therefore $\phi_{2,1}$ is linear, and hence holomorphic in a neighbourhood of $0 = \phi_1(x)$.

- Case $\phi_2(x) = 0 \in W_2$, i.e. $\tilde{\tau}_2 = \tau_2$:

The previous case shows the holomorphy of $\phi_{1,2}$ in a suitable neighbourhood. Because the map is bijective, its inverse $\phi_{2,1}$ is holomorphic.

- Case $\phi_1(x) \in W_1$ arbitrary, i.e. $\tilde{\tau}_1 \in U_1$ arbitrary:

We reduce this case to the other two cases. Choose $\tau_3 \in \mathcal{D}$ with $p(\tau_3) = x$ and consider the chart

$$\phi_{\tau_3} : U_{\tau_3} \rightarrow W_{\tau_3}$$

with the shorthand

$$U_3 := U_{\tau_3}, \phi_3 := \phi_{\tau_3}, W_3 := W_{\tau_3}.$$

In a suitable neighbourhood W of x with

$$W \subset (p(U_1) \cap p(U_2) \cap p(U_3))$$

we have

$$\phi_{1,2} = \phi_{1,3} \circ \phi_{3,2}.$$

The right-hand side is holomorphic due to part i) and ii). Therefore $\phi_{1,2}$ is holomorphic, q.e.d.

□

The important steps in the proof of Proposition 2.30 and its prerequisites:

- The modular group acts strongly properly discontinuous.
- Isotropy groups act as cyclic groups of rotations (Lemma of Schwarz).
- The existence of suitable fractional-linear maps $\lambda : \mathbb{H} \xrightarrow{\cong} \Delta$.

The upper half-plane has a canonical boundary point, the point ∞ . The extension of the group action from Remark 2.8 shows that one has to consider also the rational line $\mathbb{Q}$, but not $\mathbb{R} \setminus \mathbb{Q}$, as part of the boundary.

When extending the Euclidean topology of $\mathbb{H}$ to a topology on $\mathbb{H}^*$, one has to define the neighbourhoods in $\mathbb{H}^*$ of the additional points from

$$\mathbb{Q} \cup \{\infty\}.$$

First, surprisingly as a neighbourhood basis of $\infty \in \mathbb{H}^*$ one does not take the complements of compact sets in $\mathbb{C}$ as one does for the point

$$\infty \in \mathbb{P}^1$$

from the Riemann sphere, alias projective space. Instead one takes the sets

$$\{z \in \mathbb{C} : \text{Im } z > R\} \cup \{\infty\}, \quad R > 0,$$

i.e. the complements of sets with bounded imaginary part. Secondly, one translates neighbourhoods of ∞ to neighbourhoods of a rational $q \in \mathbb{Q}$ by fractional-linear transformations $\gamma \in \Gamma$ with

$$\gamma(\infty) = q.$$

The resulting neighbourhood basis of a point $q \in \mathbb{Q}$ are disks in $\mathbb{H}$ with q a point on the Euclidean boundary of the disk, see Figure 2.7. The second step is determined by the requirement, that the modular group acts as group of homeomorphisms on $\mathbb{H}^*$. The new topology of $\mathbb{H}^*$ induces on $\mathbb{Q}$ the discrete topology. Apparently the latter is finer than the topology induced on $\mathbb{Q}$ by the Euclidean topology of $\mathbb{R}$.

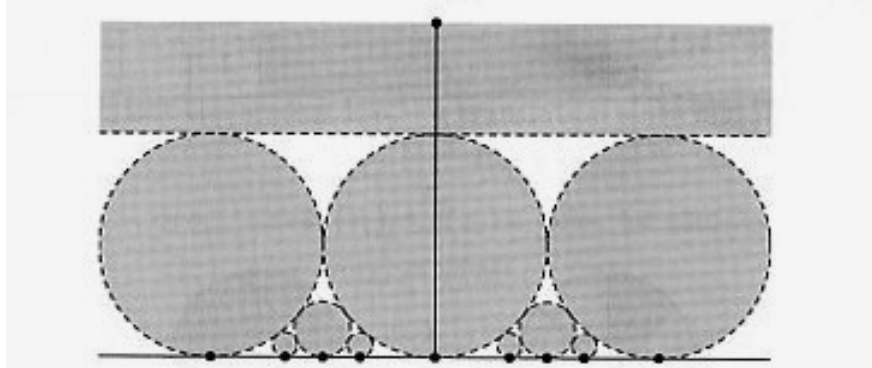


Fig. 2.7 Neighbourhoods of infinity and of some rational points from $\mathbb{H}^*$, from [19, Fig. 2.5]

Theorem 2.31 (The modular curve of the modular group). *Consider the action of the modular group $SL(2, \mathbb{Z})$ on the extended upper half-plane*

$$\mathbb{H}^* := \mathbb{H} \cup \mathbb{Q} \cup \{\infty\}.$$

according to Remark 2.8.

1. On the orbit space

$$X := SL(2, \mathbb{Z}) \backslash \mathbb{H}^*$$

exists the structure of a compact Riemann surface such that the following diagram commutes with canonical holomorphic projections p and p^* and canonical injections in the horizontal direction

$$\begin{array}{ccc} \mathbb{H} & \longrightarrow & \mathbb{H}^* \\ p \downarrow & & \downarrow p^* \\ Y & \longrightarrow & X \end{array}$$

2. The Riemann surface X is biholomorphic equivalent to $\mathbb{P}^1$.

Proof. Each rational number

$$q \in \mathbb{Q} \subset \mathbb{H}^*$$

maps by a fractional linear transformation $\gamma \in SL(2, \mathbb{Z})$ to the distinguished point $\infty \in \mathbb{H}^*$, see Remark 2.8. Hence the topology of $\mathbb{H}^*$ around $q \in \mathbb{H}^*$ can be obtained from the topology of $\mathbb{H}^*$ around $\infty \in \mathbb{H}^*$. Similarly the complex structure of X around $p^*(\infty) \in X$ is obtained from the study of a neighbourhood of the point $\infty \in \mathbb{H}^*$.

i) *Topology on $\mathbb{H}^*$:* For the topology on $\mathbb{H}^*$ one takes as neighbourhood basis of ∞ the sets

$$U_R := \{\tau \in \mathbb{H} : \text{Im } \tau > R\} \cup \{\infty\}, \quad R > 0,$$

and as neighbourhood basis of a rational point

$$q := a/c \in \mathbb{Q}, \quad \gcd(a, c) = 1,$$

the sets

$$\gamma(U_R), \quad R > 0, \quad \gamma = \begin{pmatrix} a & * \\ c & * \end{pmatrix} \in \Gamma,$$

because

$$\gamma(\infty) = q$$

due to Remark 2.8. These neighbourhood bases together with the open sets of $\mathbb{H}$ generate a topology on $\mathbb{H}^*$ with $\mathbb{H} \subset \mathbb{H}^*$ an open subspace.

Note. Because fractional linear transformations from Γ take lines to circles or lines the neighbourhoods of $\gamma(U_R)$ are circles tangent to the real axis at the distinguished

point $a/c \in \mathbb{Q}$. Hence neighbourhoods of ∞ and of rational points are completely different from their counterpart with respect to the Euclidan topology, see Figure 2.7.

ii) *Hausdorff topology on $SL(2, \mathbb{Z}) \backslash \mathbb{H}^*$* : We provide the set

$$X := SL(2, \mathbb{Z}) \backslash \mathbb{H}^*$$

with the quotient topology with respect to the canonical projection

$$p^* : \mathbb{H}^* \rightarrow X.$$

Then p^* is an open map: The proof is the same as the corresponding proof for Lemma 2.24. The proof uses only the fact that the quotient arises from a group action.

In order to show that X is Hausdorff we consider two points

$$x_i = p^*(\tau_i) \in X, \tau_i \in \mathbb{H}^*, i = 1, 2, \text{ and assume } x_1 \neq x_2.$$

We distinguish the following cases:

- $\tau_1 \in \mathbb{H}$, $\tau_2 \in \mathbb{Q} \cup \{\infty\}$: W.l.o.g. $\tau_2 = \infty$. For each

$$\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma \text{ and } \tau \in \mathbb{H}$$

the formula

$$\begin{aligned} \operatorname{Im} \gamma(\tau) &= \frac{\operatorname{Im} \tau}{|c \cdot \tau + d|^2} = \frac{\operatorname{Im} \tau}{(c \cdot \operatorname{Re} \tau + d)^2 + c^2 \cdot (\operatorname{Im} \tau)^2} \leq \\ &\leq \operatorname{Im} \tau \cdot \begin{cases} \frac{1}{d^2} & c = 0 \\ \frac{1}{c^2 \cdot (\operatorname{Im} \tau)^2} & c \neq 0 \end{cases} \leq \begin{cases} \operatorname{Im} \tau & c = 0 \\ \frac{1}{\operatorname{Im} \tau} & c \neq 0 \end{cases} \leq \max \left\{ \operatorname{Im} \tau, \frac{1}{\operatorname{Im} \tau} \right\} \end{aligned}$$

shows

$$\operatorname{Im} \gamma(\tau) \leq \max \left\{ \operatorname{Im} \tau, \frac{1}{\operatorname{Im} \tau} \right\},$$

independently from γ . Therefore a relatively compact neighbourhood

$$U_1 \subset \subset \mathbb{H}$$

of τ_1 and a neighbourhood

$$U_2 \subset \mathbb{H}^*$$

of $\tau_2 = \infty$ exist such that for all $\gamma \in SL(2, \mathbb{Z})$

$$\gamma(U_1) \cap U_2 = \emptyset.$$

As a consequence

$$p^*(U_1) \cap p^*(U_2) = \emptyset.$$

- $\tau_1, \tau_2 \in \mathbb{H}$: This case has been considered already by Corollary 2.28.

iii) *Compactness of X* : Denote by $\overline{\mathcal{D}}^{\mathbb{H}} \subset \mathbb{H}$ the closure with respect to $\mathbb{H}$ of the fundamental domain $\mathcal{D}$ of $SL(2, \mathbb{Z})$. Then

$$\overline{\mathcal{D}} := \overline{\mathcal{D}}^{\mathbb{H}} \cup \{\infty\} \subset \mathbb{H}^*$$

is compact with respect to $\mathbb{H}^*$: Consider an open covering $(U_i)_{i \in I}$ of $\overline{\mathcal{D}}$. If $\infty \in U_{i_0}$ then for suitable $R > 0$:

$$\{\tau \in \overline{\mathcal{D}}^{\mathbb{H}} : \text{Im } \tau > R\} \cup \{\infty\} \subset U_{i_0},$$

and the remaining set

$$\{\tau \in \overline{\mathcal{D}}^{\mathbb{H}} : \text{Im } \tau \leq R\}$$

is compact in $\mathbb{H}$, and therefore covered by a finite subcovering of $(U_i)_{i \in I}$.

As a consequence X is compact as continuous image via p^* of the compact set $\overline{\mathcal{D}} \subset \mathbb{H}^*$. One has

$$Y = X \setminus \{p^*(\infty)\}.$$

Therefore $Y \subset X$ is an open subset.

iv) *Complex structure on X* : In order to define a chart

$$\phi_\infty : p^*(U_\infty) \rightarrow W_\infty$$

around the cusp

$$x = p^*(\infty) \in X$$

we consider the neighbourhood of $\infty \in \mathbb{H}^*$

$$U_\infty := U_2 = \{\tau \in \mathbb{H} : \text{Im } \tau > 2\} \cup \{\infty\}.$$

We define

$$pow_\infty : U_\infty \rightarrow \Delta, \tau \mapsto \begin{cases} e^{2\pi i \tau} & \text{if } \tau \neq \infty \\ 0 & \text{if } \tau = \infty \end{cases}$$

and set

$$W_\infty := pow_\infty(U_\infty) \subset \Delta.$$

For two points $\tau_i \in U_\infty, i = 1, 2$,

$$p^*(\tau_2) = p^*(\tau_1) \iff \tau_2 = \gamma(\tau_1) \text{ for a suitable } \gamma \in SL(2, \mathbb{Z})$$

$$\iff \tau_2 = \tau_1 + m \text{ for a suitable } m \in \mathbb{Z}, \text{ due to Theorem 2.18,}$$

$$\iff \text{pow}_\infty(\tau_2) = \text{pow}_\infty(\tau_1).$$

We define the chart ϕ_∞ as the commutative completion of the diagram

$$\begin{array}{ccc} U_\infty & \xrightarrow{p^*|_{U_\infty}} & p^*(U_\infty) \\ & \searrow \text{pow}_\infty & \downarrow \phi_\infty \\ & & W_\infty \end{array}$$

One shows analogously to the proof of Proposition 2.30 that the map ϕ_∞ is a homeomorphism.

Eventually we have to show that the charts of Y and the chart ϕ_∞ are compatible, It suffices to prove the compatibility of the chart

$$\phi_\infty : p^*(U_\infty) \rightarrow W_\infty$$

around $p^*(\infty)$ with the chart

$$\phi_\tau : p(U_\tau) = p^*(U_\tau) \rightarrow W_\tau$$

around an arbitrary point $p(\tau)$, $\tau \in \mathbb{H}$. We may assume that U_τ does not contain an elliptic point, and therefore that

$$\phi_\tau \circ (p^*|_{U_\tau})$$

is injective and biholomorphic. Hence the transition function of the two charts is holomorphic.

v) *Modular curve* $X \simeq \mathbb{P}^1$: The image of the fundamental domain

$$Y = p(\mathcal{D})$$

is homeomorphic to the plane $\mathbb{C}$: Identify points on the left and right boundary of $\mathcal{D}$ and equivalent points of $\mathcal{D}$ on the left part of the unit circle and on the right part. Therefore the compactification of Y is homeomorphic to the 1-point compactification of $\mathbb{C}$, i.e. X is homeomorphic to S^2 . The only complex structure on the sphere S^2 is the complex projective space $\mathbb{P}^1(\mathbb{C})$, cf. [69].

□

Note. A formal proof for the biholomorphy

$$X \simeq \mathbb{P}^1$$

will be given in Corollary 3.16. This second proof uses the modular invariant j from the theory of modular forms.

Definition 2.32 (Modular curve of $SL(2, \mathbb{Z})$). The compact Riemann surface

$$X = SL(2, \mathbb{Z}) \backslash \mathbb{H}^*$$

from Theorem 2.31 is named the *modular curve* $X(1)$ of $SL(2, \mathbb{Z})$ or simply the *modular curve*.

An analogous construction can be made for each congruence subgroup $\Gamma_0(N) \subset SL(2, \mathbb{Z})$.

Remark 2.33 (Modular curves of the congruence subgroups).

1. *Definition:* Generalizing the construction of the modular curve

$$X(1) = SL(2, \mathbb{Z}) \backslash \mathbb{H}^*$$

from Theorem 2.31 one can introduce for each Hecke congruence subgroup

$$\Gamma_0(N) \subset SL(2, \mathbb{Z}), \quad N \in \mathbb{N},$$

a complex structure on the orbit space

$$X_0(N) := \Gamma_0(N) \backslash \mathbb{H}^* = (\Gamma_0(N) \backslash \mathbb{H}) \dot{\cup} \text{Cusp}(\Gamma_0(N)),$$

see [19, Prop. 2.4.2ff.] The resulting compact Riemann surface $X_0(N)$ is named the *modular curve* of $\Gamma_0(N)$.

2. *Branched covering of the modular curves:* The inclusion of groups

$$\Gamma_0(N) \hookrightarrow \Gamma(1)$$

induces a holomorphic branched covering

$$f : X_0(N) \rightarrow X(1)$$

of compact Riemann surfaces. The covering has degree

$$\deg f = [SL(2, \mathbb{Z}) : \Gamma_0(N)] = N \cdot \prod_{p|N} \left(1 + \frac{1}{p}\right),$$

see [19, Exerc. 3.1.1] and Proposition 2.16.

3. *Genus of modular curves*: The Riemann-Hurwitz formula, e.g., [69], applied to the covering

$$f : X_0(N) \rightarrow X(1) \simeq \mathbb{P}^1$$

computes the genus of the modular curves of the Hecke congruence subgroups as

$$g(X_0(N)) = 1 + \frac{b(f)}{2} - \deg f$$

with

$$b(f) = \sum_{x \in X_0(N)} b(f; x)$$

the total branching order of f .

A detailed investigation of the elliptic points and of the cusps of the action of $\Gamma_0(N)$ shows the *genus formula*

$$g(X_0(N)) = 1 + \frac{\deg f}{12} - \frac{\varepsilon_2}{4} - \frac{\varepsilon_3}{3} - \frac{\varepsilon_\infty}{2}$$

with ε_2 , ε_3 and ε_∞ denoting the number of elliptic points of respectively period 2 and 3 and the number of cusps of the action of $\Gamma_0(N)$, see [19, Theor. 3.1.1].

The ε -constants can be computed as

$$\varepsilon_\infty = \sum_{d|N} \Phi(\gcd(d, N/d)),$$

see Example 2.12, and ε_2 and ε_3 are the numbers of solutions in $\mathbb{Z}/N\mathbb{Z}$ of the quadratic equations respectively

$$x^2 + 1 = 0 \text{ and } x^2 - x + 1 = 0.$$

Using the quadratic residue symbol for $n \in \{-1, -3\}$ and extended to all primes p

$$\left(\frac{-1}{p}\right) := \begin{cases} 0 & \text{if } p = 2 \\ 1 & \text{if } p \equiv 1 \pmod{4} \\ -1 & \text{if } p \equiv 3 \pmod{4} \end{cases}$$

and

$$\left(\frac{-3}{p}\right) := \begin{cases} 0 & \text{if } p = 3 \\ 1 & \text{if } p \equiv 1 \pmod{3} \\ -1 & \text{if } p \equiv 2 \pmod{3} \end{cases}$$

one obtains, see [19, Cor. 3.7.2, Exerc. 3.7.6] and [59, Theor. 1.43]:

$$\varepsilon_2 = \begin{cases} \prod_{p|N} \left(1 + \left(\frac{-1}{p}\right)\right) & \text{if } 4 \nmid N \\ 0 & \text{otherwise} \end{cases}$$

and

$$\varepsilon_3 = \begin{cases} \prod_{p|N} \left(1 + \left(\frac{-3}{p}\right)\right) & \text{if } 9 \nmid N \\ 0 & \text{otherwise} \end{cases}$$

4. *Estimate of the genus $g(X_0(N))$* : As a corollary of the genus formula the constants from part 2) satisfy the estimate

$$2 \cdot g(X_0(N)) - 2 + \frac{\varepsilon_2}{2} + \frac{2\varepsilon_3}{3} + \varepsilon_\infty > 0.$$

5. *Modular curves of prime level*: For a prime $p \in \mathbb{N}$ and $k := p + 1$ the modular curve $X_0(p)$ has the genus

$$g(X_0(p)) = \begin{cases} \left\lfloor \frac{k}{12} \right\rfloor - 1 & \text{if } k \equiv 2 \pmod{12} \\ \left\lfloor \frac{k}{12} \right\rfloor & \text{otherwise} \end{cases}$$

In particular,

$$g(X_0(2)) = 0, \quad g(X_0(11)) = 1, \quad g(X_0(13)) = 0 \text{ and } g(X_0(17)) = 1$$

6. *Numerical example*: See the PARI-file “modular_curve_as_covering”. For the computation of the extended quadratic residue symbol PARI provides the command

$$\text{kronecker}(a, p).$$

The Kronecker symbol is defined for arbitrary integer $a \in \mathbb{Z}$ and arbitrary primes including $p = 2$ by extending the Legendre symbol:

- *Odd prime p* :

$$\text{kronecker}(a, p) := \left(\frac{a}{p}\right) \text{ for odd } p$$

- *Even prime $p = 2$* :

$$\text{kronecker}(a, 2) := \begin{cases} 0 & \text{if } a \text{ even} \\ (-1)^{(a^2-1)/8} & \text{otherwise} \end{cases}$$

The Kronecker symbol satisfies

- *Odd prime p :*

$$\text{kronecker}(-4, p) = \left(\frac{-4}{p} \right) = \left(\frac{2}{p} \right)^2 \cdot \left(\frac{-1}{p} \right) = \left(\frac{-1}{p} \right)$$

- *Even prime $p = 2$:*

$$\text{kronecker}(-4, 2) = 0 = \left(\frac{-1}{2} \right) \text{ and } \text{kronecker}(-3, 2) = -1 = \left(\frac{-3}{2} \right)$$

Hence

$$\varepsilon_2 = \begin{cases} \prod_{p|N} (1 + \text{kronecker}(-4, p)) & \text{if } 4 \nmid N \\ 0 & \text{otherwise} \end{cases}$$

and

$$\varepsilon_3 = \begin{cases} \prod_{p|N} (1 + \text{kronecker}(-3, p)) & \text{if } 9 \nmid N \\ 0 & \text{otherwise} \end{cases}$$

Figure 2.8 shows the invariants of the covering

$$X_0(p) \rightarrow X(1)$$

for all primes $p < 100$.

```

modular_curve_as_covering: Start
modular_curve_as_covering: Modular curve  $X_0(N)$ 
as branched covering of modular curve  $X(1)$ .

(N, genus, d, eps_2, eps_3, eps_inf)=(2, 0, 3, 1, 0, 2)
(N, genus, d, eps_2, eps_3, eps_inf)=(3, 0, 4, 0, 1, 2)
(N, genus, d, eps_2, eps_3, eps_inf)=(5, 0, 6, 2, 0, 2)
(N, genus, d, eps_2, eps_3, eps_inf)=(7, 0, 8, 0, 2, 2)
(N, genus, d, eps_2, eps_3, eps_inf)=(11, 1, 12, 0, 0, 2)
(N, genus, d, eps_2, eps_3, eps_inf)=(13, 0, 14, 2, 2, 2)
(N, genus, d, eps_2, eps_3, eps_inf)=(17, 1, 18, 2, 0, 2)
(N, genus, d, eps_2, eps_3, eps_inf)=(19, 1, 20, 0, 2, 2)
(N, genus, d, eps_2, eps_3, eps_inf)=(23, 2, 24, 0, 0, 2)
(N, genus, d, eps_2, eps_3, eps_inf)=(29, 2, 30, 2, 0, 2)
(N, genus, d, eps_2, eps_3, eps_inf)=(31, 2, 32, 0, 2, 2)
(N, genus, d, eps_2, eps_3, eps_inf)=(37, 2, 38, 2, 2, 2)
(N, genus, d, eps_2, eps_3, eps_inf)=(41, 3, 42, 2, 0, 2)
(N, genus, d, eps_2, eps_3, eps_inf)=(43, 3, 44, 0, 2, 2)
(N, genus, d, eps_2, eps_3, eps_inf)=(47, 4, 48, 0, 0, 2)
(N, genus, d, eps_2, eps_3, eps_inf)=(53, 4, 54, 2, 0, 2)
(N, genus, d, eps_2, eps_3, eps_inf)=(59, 5, 60, 0, 0, 2)
(N, genus, d, eps_2, eps_3, eps_inf)=(61, 4, 62, 2, 2, 2)
(N, genus, d, eps_2, eps_3, eps_inf)=(67, 5, 68, 0, 2, 2)
(N, genus, d, eps_2, eps_3, eps_inf)=(71, 6, 72, 0, 0, 2)
(N, genus, d, eps_2, eps_3, eps_inf)=(73, 5, 74, 2, 2, 2)
(N, genus, d, eps_2, eps_3, eps_inf)=(79, 6, 80, 0, 2, 2)
(N, genus, d, eps_2, eps_3, eps_inf)=(83, 7, 84, 0, 0, 2)
(N, genus, d, eps_2, eps_3, eps_inf)=(89, 7, 90, 2, 0, 2)
(N, genus, d, eps_2, eps_3, eps_inf)=(97, 7, 98, 2, 2, 2)
modular_curve_as_covering: End

```

Fig. 2.8 Modular curves $X_0(N) \rightarrow X(1)$

Chapter 3

The algebra of modular forms

Modular forms are holomorphic functions on $\mathbb{H}$ distinguished by two properties:

1. On one hand, they have the modular group as group of symmetries.
2. On the other hand, they expand into a Fourier series around the point ∞ with rational Fourier coefficients after normalization.

The interplay of these two properties is coined “The magic of modular forms” [71]. Therefrom arises a series of remarkable identities between the Fourier coefficients of modular forms. In many cases these identities imply interesting statements about arithmetic functions. We will see examples in Section 5.1 and Chapter 6.

In general, one introduces modular forms as functions on $\mathbb{H}$ with a certain transformation behaviour with respect to the action of the modular group or its congruence subgroups. The basic examples are the Eisenstein series from Section 3.2. A meromorphic quotient of certain modular forms is the modular invariant j , see Definition 3.15, a key object of the whole theory.

From a more elevated point of view, modular forms are differential forms on the modular curves, which are compact Riemann surfaces, see Theorem 3.17. The Riemann-Roch theorem computes the dimensions of the different vector spaces of modular forms.

From now on we reserve the letter Γ_1 for the modular group

$$\Gamma_1 := SL(2, \mathbb{Z})$$

3.1 Modular forms and cusp forms

Modular forms are holomorphic functions on $\mathbb{H}$

- which transform in a simple manner under the operation of the modular group Γ_1 or its Hecke congruence subgroups $\Gamma_0(N)$
- and extend to $\mathbb{H}^*$ in a certain holomorphic way.

In a second step, such invariant objects can be considered analytical objects on the modular curves $X_0(N)$. Because the modular curves are compact Riemann surfaces one cannot expect interesting objects when restricting to invariant holomorphic *functions*. Instead one has to consider more general holomorphic and meromorphic differential forms.

We start from the general context of representation theory of the group Γ_1 , i.e. we introduce a Γ_1 -action on certain vector spaces of holomorphic differential forms. Proposition 3.1 carries over the action of Γ_1 from an action on functions to an action on differential forms on $\mathbb{H}$. We denote by $\gamma^*\omega$ the pull-back of a differential ω form along a holomorphic map γ .

Proposition 3.1 (Γ_1 acting on differential forms). *Denote by*

$$W^1 := H^0(\mathbb{H}, \Omega_{\mathbb{H}}^1)$$

the vector space of holomorphic differential forms on the upper half-plane $\mathbb{H}$.

1. *Right action on differential forms: The map*

$$\Phi^1 : W^1 \times \Gamma_1 \rightarrow W^1, (\omega, \gamma) \mapsto \Phi^1(\omega, \gamma) := \gamma^*\omega,$$

is a Γ_1 -right action. For

$$\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_1, \omega = f d\tau \in W^1 \text{ and } \tau \in \mathbb{H}$$

holds

$$\Phi^1(\omega, \gamma)(\tau) := (\gamma^*\omega)(\tau) = (f \circ \gamma)(\tau) \frac{d\tau}{(c\tau + d)^2}$$

2. *Action on the tensor product of differential forms: For even $k \in \mathbb{N}$ set*

$$\Omega^{k/2} := \Omega_{\mathbb{H}}^{\otimes k/2} \text{ and } W^{k/2} := H^0(\mathbb{H}, \Omega^{k/2})$$

and

$$\Phi^{k/2} : W^{k/2} \times \Gamma_1 \rightarrow W^{k/2}, (\omega, \gamma) \mapsto \Phi^{k/2}(\omega, \gamma) := \gamma^*\omega.$$

For

$$\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_1, \quad \omega = f(d\tau)^{k/2} \in W^{k/2} \quad \text{and } \tau \in \mathbb{H}$$

holds

$$\Phi^{k/2}(\omega, \gamma)(\tau) := ((\gamma^*)^{k/2} \omega)(\tau) = (f \circ \gamma)(\tau) \frac{(d\tau)^{k/2}}{(c\tau + d)^k}$$

3. The condition of invariance: For even $k \in \mathbb{N}$,

$$\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_1 \quad \text{and } \omega = f(d\tau)^{k/2} \in W^{k/2}$$

holds the equivalence which characterizes invariant differential forms:

$$\gamma^* \omega = \omega \iff f(\tau) = (f \circ \gamma)(\tau) \frac{1}{(c\tau + d)^k} \quad \text{for all } \tau \in \mathbb{H}.$$

Note. In the context of modular forms $\Omega^{k/2}$ denotes the tensor product of Ω^1 not the exterior product.

Proof. 1. We have

$$\Phi^1(\omega, \gamma_1 \cdot \gamma_2) = (\gamma_1 \cdot \gamma_2)^* \omega = \gamma_2^* (\gamma_1^* \omega) = \Phi^1(\Phi^1(\omega, \gamma_1), \gamma_2)$$

Moreover

$$(d\gamma)(\tau) = d \left(\frac{a\tau + b}{c\tau + d} \right) = \frac{a(c\tau + d) - (a\tau + b)c}{(c\tau + d)^2} d\tau = \frac{1}{(c\tau + d)^2} d\tau$$

Therefore

$$(\Phi^1(f d\tau, \gamma))(\tau) := (\gamma^*(f d\tau))(\tau) = f(\gamma(\tau)) (d\gamma)(\tau) = (f \circ \gamma)(\tau) \frac{d\tau}{(c\tau + d)^2}$$

2. For even $k > 2$ the proof is analogous by applying the derivation separately for each tensor factor $d\tau$.

□

Modular forms will not be necessarily invariant under the operation of $\Gamma_0(N)$, but they multiply with a factor of automorphy which is a holomorphic function without zeros. We define the concept in a slightly more general context.

Definition 3.2 (Weight- k operator and factor of automorphy). For functions $f : \mathbb{H} \rightarrow \mathbb{C}$, matrices

$$\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_1,$$

and $even k \in \mathbb{N}^*$ we define the *weight- k operator* on f

$$f[\gamma]_k : \mathbb{N} \rightarrow \mathbb{C}, f[\gamma]_k(\tau) := f(\gamma(\tau)) \cdot \frac{1}{(c\tau + d)^k}$$

For the left group action

$$\Phi : GL(2, \mathbb{R})^+ \times \mathbb{H} \rightarrow \mathbb{H}, (\gamma, \tau) \mapsto \gamma(\tau)$$

the map

$$h : GL(2, \mathbb{R})^+ \times \mathbb{H} \rightarrow \mathbb{H}, h(\gamma, \tau) := c\tau + d,$$

which focuses on the denominator, is named the *factor of automorphy* of Φ .

Proposition 3.1 shows: For even $k \in \mathbb{N}$ the invariant differential forms

$$\omega = f d\tau^{k/2} \in (W^{k/2})_1^\Gamma$$

from the fixed space of the Γ_1 -representation are characterized by holomorphic coefficient functions $f \in \mathcal{O}(\mathbb{H})$ satisfying

$$f = f[\gamma]_k \text{ for all } \gamma \in \Gamma_1$$

For the translation

$$T := \begin{pmatrix} a & b \\ c & d \end{pmatrix} := \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \in \Gamma_0(N)$$

the factors of automorphy satisfy for all $\tau \in \mathbb{H}$

$$\frac{1}{c\tau + d} = 1.$$

Hence invariance under translation means

$$T^*(f (d\tau)^{k/2}) = f (d\tau)^{k/2} \iff f(\tau + 1) = f(\tau) \text{ for all } \tau \in \mathbb{H},$$

i.e. the function f has the period = 1. Therefore f expands into a convergent Fourier series

$$f(\tau) = \hat{f}(q) = \sum_{n=-\infty}^{\infty} a_n \cdot q^n, \quad q = e^{2\pi i \tau},$$

similar to the trigonometric function

$$\cos(2\pi z) = \frac{1}{2} (e^{2\pi i z} + e^{-2\pi i z}).$$

The holomorphic map

$$\mathbb{H} \rightarrow \Delta^*, \tau \mapsto e^{2\pi i \tau},$$

surjects to the punctured unit disk

$$\Delta^* := \{z \in \mathbb{C} : 0 < |z| < 1\}.$$

Here neighbourhoods of $\infty \in \mathbb{H}^*$ with respect to the topology of $\mathbb{H}^*$ from Theorem 2.31 are mapped to neighbourhoods of zero. A meromorphic function on $\mathbb{H}$ with period = 1 induces a corresponding meromorphic function $\hat{f}$ on Δ^* according to the following diagram

$$\begin{array}{ccc} \mathbb{H} & \xrightarrow{f} & \mathbb{C} \\ \downarrow & \nearrow \hat{f} & \\ \Delta^* & & \end{array}$$

Set

$$\hat{f}(q) := f(\tau) = \sum_{n=-\infty}^{\infty} a_n \cdot q^n, \quad q = e^{2\pi i \cdot \tau}.$$

Definition 3.3 (Fourier series). Consider a meromorphic function f on $\mathbb{H}$ invariant with respect to the translation T and denote by

$$\hat{f}(q) := \hat{f}(e^{2\pi i \tau}) := f(\tau)$$

the induced meromorphic function $\hat{f}$ on Δ^* .

1. The function f is named *meromorphic at ∞* if $\hat{f}$ has an isolated singularity at 0, which is a pole or a removable singularity.
2. If f is meromorphic at ∞ then the induced function $\hat{f}$ has a Laurent expansion around 0

$$\hat{f}(q) = \sum_{n=n_0}^{\infty} a_n \cdot q^n, \quad q = e^{2\pi i \tau},$$

with $n_0 \in \mathbb{Z}$ satisfying $a_{n_0} \neq 0$ if $f \neq 0$. The Laurent expansion is named the q -expansion or *Fourier series* of f , and the elements $(a_n)_{n \geq n_0}$ are the *Fourier coefficients* of f .

3. A function f , meromorphic at ∞ , is named *holomorphic at ∞* if $n_0 \in \mathbb{N}$, i.e. if $0 \in \Delta$ is a removable singularity of $\hat{f}$. In this case one defines

$$f(\infty) := \hat{f}(0).$$

Note: For a meromorphic function f on $\mathbb{H}$, invariant with respect to T :

$$\begin{aligned}
f \text{ holomorphic at } \infty &\iff \lim_{\text{Im } \tau \rightarrow \infty} |f(\tau)| < \infty \iff \\
&\iff \exists R > 0 : |f(\tau)| \text{ bounded for } \text{Im } \tau \geq R.
\end{aligned}$$

Lemma 3.4 (Transformation of the factors of automorphy).

1. *The factors of automorphy satisfy the composition rule*

$$h(B \cdot A, \tau) = h(B, A(\tau)) \cdot h(A, \tau); \quad A, B \in GL(2, \mathbb{R})^+, \quad \tau \in \mathbb{H}.$$

2. *Consider the vector space*

$$V := \{f : \mathbb{H} \rightarrow \mathbb{C}\}$$

of complex-valued functions on $\mathbb{H}$. For any fixed $k \in \mathbb{Z}$ the map

$$V \times GL(2, \mathbb{R})^+ \rightarrow V, \quad (f, \gamma) \mapsto f[\gamma]_k,$$

with

$$f[\gamma]_k(\tau) := f(\gamma(\tau)) \cdot h(\gamma, \tau)^{-k} \cdot (\det \gamma)^{k-1}, \quad \tau \in \mathbb{H},$$

is a right group action of $GL(2, \mathbb{R})^+$ on the vector space V .

In the definition of $f[\gamma]_k$ from Lemma 3.4 the exponent of $\det \gamma$ is not uniform in the literature. We use the exponent $(k-1)$ from [19, Exerc. 1.2.11], while [37, Chap. III, §3] uses the exponent $k/2$. The choice of the exponent $(k-1)$ will later simplify Definition 5.11.

Note that in Lemma 3.4 the group $GL(2, \mathbb{R})^+$ is written on the right-hand side of the vector space V .

Proof. Consider $A, B \in GL(2, \mathbb{R})^+$ and $\tau \in \mathbb{H}$.

1.

$$A \cdot \begin{pmatrix} \tau \\ 1 \end{pmatrix} = \begin{pmatrix} A(\tau) \\ 1 \end{pmatrix} \cdot h(A, \tau).$$

Here we distinguish between on the left-hand side the matrix product of A with a vector and on the right-hand side the fractional linear transformation $A(\tau)$.

Applying this identity to the product $B \cdot A$

$$(B \cdot A) \cdot \begin{pmatrix} \tau \\ 1 \end{pmatrix} = \begin{pmatrix} (B \cdot A)(\tau) \\ 1 \end{pmatrix} \cdot h(B \cdot A, \tau).$$

Left-hand side:

$$B \cdot \left(A \cdot \begin{pmatrix} \tau \\ 1 \end{pmatrix} \right) = B \cdot \left(\begin{pmatrix} A(\tau) \\ 1 \end{pmatrix} \cdot h(A, \tau) \right) = \begin{pmatrix} B(A(\tau)) \\ 1 \end{pmatrix} \cdot h(B, A(\tau)) \cdot h(A, \tau).$$

Equating with the right-hand side gives for the lower row

$$h(B \cdot A, \tau) = h(B, A(\tau)) \cdot h(A, \tau),$$

hence for the upper row

$$(B \cdot A)(\tau) = B(A(\tau)).$$

2. Starting with the right-hand side we compute

$$\begin{aligned} ((f[B]_k)[A]_k)(\tau) &= (f[B]_k)(A(\tau)) \cdot h(A, \tau)^{-k} (\det A)^{k-1} = \\ &= f(B(A(\tau))) \cdot h(B, A(\tau))^{-k} \cdot h(A, \tau)^{-k} \cdot (\det B)^{k-1} \cdot (\det A)^{k-1} = \\ &= f((B \cdot A)(\tau)) \cdot h(B \cdot A, \tau)^{-k} \cdot (\det (B \cdot A))^{k-1} = f[B \cdot A]_k(\tau). \end{aligned}$$

Here the penultimate equality holds due the result from part 1) and due to the equality

$$(B \cdot A)(\tau) = B(A(\tau))$$

from the proof of part i). Hence

$$f[B \cdot A]_k = (f[B]_k)[A]_k.$$

□

The Fourier expansion from Definition 3.3 is well-defined for holomorphic functions $f : \mathbb{H} \rightarrow \mathbb{C}$ which are invariant with respect to the translation T . The modular group Γ_1 has only a single cusp, the orbit comprising ∞ and all rationals $q \in \mathbb{Q}$.

But Hecke congruence subgroups $\Gamma_0(N)$, $N \geq 2$, do not have only the single cusp passing through the point ∞ . In addition, they have one or more cusps, each comprising a different subset of $\mathbb{Q}$. Fix a Hecke congruence subgroup $\Gamma_0(N)$, a

weight $k \in \mathbb{Z}$, and a rational $q = \frac{a}{c} \in \mathbb{Q}$. How to extend in a meromorphic or holomorphic way a holomorphic function

$$f : \mathbb{H} \rightarrow \mathbb{C}$$

with

$$f[\gamma]_k = f$$

for all $\gamma \in \Gamma_0(N)$?

Definition 3.5 (Fourier coefficients for congruence subgroups). Let $\Gamma_0(N) \subset \Gamma_1$ be a congruence subgroup of level $N \in \mathbb{N}$ acting on $\mathbb{H}^*$, and $T \in \Gamma_0(N)$ a generator of all translations in $\Gamma_0(N)$. For a cusp $s \in \mathbb{H}^*$ and a rational point on it

$$x_s \in \mathbb{Q} \cap s$$

exists an element

$$\alpha \in \Gamma_1 \text{ with } x_s = \alpha(\infty)$$

Then each meromorphic function f on $\mathbb{H}$, which is invariant under the action of $\Gamma_0(N)$, is in particular invariant under the action of T . Hence $f \circ \alpha^{-1}$ develops into a Fourier series with period the period of T

$$f(\tau \circ \alpha^{-1}) = \sum_{n=n_0}^{\infty} a_n \cdot q^n, \quad q = e^{2\pi i((\tau \circ \alpha^{-1})/h_s)},$$

with $n_0 \in \mathbb{Z}$ and h_q the width of the cusp s . The series is named the *Fourier series of f at the cusp s* .

The fundamental object of the whole theory is the *modular form*. We introduce modular forms of the modular group Γ_1 respectively its Hecke congruence subgroups $\Gamma_0(N)$ as the coefficient functions of the invariant differential forms from Proposition 3.1 which extend to $\infty \in \mathbb{H}^*$. It turns out that modular forms are meromorphic differential forms on the modular curve $X_0(N)$, see Theorem 3.17.

Definition 3.6 (Modular forms and automorphic forms).

1. *Modular forms and automorphic forms of the modular group:* Consider a meromorphic function f on $\mathbb{H}$ and an integer $k \in \mathbb{Z}$.
 - The function f is *weakly modular of weight k* with respect to Γ_1 if for all

$$\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_1$$

and for all $\tau \in \mathbb{H}$

$$f[\gamma]_k = f$$

on the left-hand side using Definition 3.2.

- The function f is an *automorphic form of weight k* with respect to Γ_1 if f is weakly modular of weight k and meromorphic at ∞ .
- The function f is a *modular form of weight k* with respect to Γ_1 if f is weakly modular of weight k , holomorphic on $\mathbb{H}$ and holomorphic at ∞ .
- A modular form f of weight k which vanishes on the cusp of Γ_1 , i.e.

$$f(\infty) = 0,$$

is a *cusp form of weight k* with respect to Γ_1 .

- An *automorphic function* f with respect to Γ_1 is an automorphic form of weight $k = 0$, i.e. satisfying

$$f \circ \gamma = f.$$

Similarly, a *modular function* f is a modular form of weight $k = 0$.

- One denotes by

$$A_k(\Gamma_1) \supset M_k(\Gamma_1) \supset S_k(\Gamma_1)$$

the vector spaces of automorphic forms with respect to Γ_1 of weight k , the subspace of modular forms, and the subspace of cusp forms.

2. *Modular forms of congruence subgroups*: Consider a congruence subgroup $\Gamma_0(N)$ with a positive integer $N \in \mathbb{N}^*$.

- A *modular form* of weight $k \in \mathbb{Z}$, $k \in \mathbb{Z}$, with respect to $\Gamma_0(N)$ is a holomorphic function

$$f : \mathbb{H} \rightarrow \mathbb{C}$$

with the following properties: For all $\gamma \in \Gamma_0(N)$ holds

$$f[\gamma]_k = f,$$

and for all $\alpha \in \Gamma_1$ the function $f[\alpha]_k$ is holomorphic at ∞ in the sense of Definition 3.3.

- A modular form f of weight $k \in \mathbb{Z}$ with respect to $\Gamma_0(N)$ is a *cusp form* of weight $k \in \mathbb{Z}$ with respect to $\Gamma_0(N)$ if for all $\alpha \in \Gamma_1$ the function $f[\alpha]_k$ is holomorphic at ∞ , and its Fourier series in the sense of Definition 3.3 vanishes at $\hat{q} = 0$.

3. *Automorphic forms of congruence subgroups*: An automorphic form of weight $k \in \mathbb{Z}$, with respect to $\Gamma_0(N)$ is a meromorphic function on $\mathbb{H}$ with the following properties: For all $\gamma \in \Gamma_0(N)$ holds

$$f[\gamma]_k = f,$$

and for all $\alpha \in \Gamma_1$ the function $f[\alpha]_k$ is meromorphic at ∞ in the sense of Definition 3.3.

We denote by

$$A_k(\Gamma_0(N)) \supset M_k(\Gamma_0(N)) \supset S_k(\Gamma_0(N))$$

the complex vector space of respectively automorphic forms, modular forms, and cusp forms of weight k with respect to $\Gamma_0(N)$.

Definition 3.6, part 2) and 3) generalize part 1): For $\gamma \in \Gamma_1$ holds $\det \gamma = 1$, hence

$$f[\gamma]_k(\tau) = (f \circ \gamma)(\tau) \cdot \frac{1}{(c\tau + d)^k}$$

Note that Definition 3.6 uses the weight convention from [19, Chap. 1.1]. Some sources use a different weight convention. In the literature also the notation to discriminate between modular *functions* and modular *forms* is not uniform.

As a consequence of Theorem 2.18 and Lemma 3.4 it is sufficient to require in Definition 3.6, part 1) the transformation law only for the two generators S, T of Γ_1 .

Remark 3.7 (Modular forms of odd weight). A modular form of a congruence subgroup $\Gamma_0(N)$, $N \in \mathbb{N}$, of odd weight $k \in \mathbb{Z}$ satisfies for all $\tau \in \mathbb{H}$ and for the particular element

$$\gamma = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \in \Gamma_0(N)$$

the equation

$$f(\tau) = f[\gamma]_k(\tau) = f(\gamma(\tau)) \cdot (-1)^k = f(\tau) \cdot (-1)^k.$$

Therefore $f = 0$ is the only modular function of $\Gamma_0(N)$ of odd weight.

3.2 Eisenstein series generate the algebra of modular forms

For the modular group Γ_1 we will now explicitly construct modular forms G_k of even weight $k \geq 4$, and also a distinguished holomorphic function G_2 , which behaves in a different way under the Γ_1 -action. To achieve the construction we need the convergence of certain infinite series. Lemma 3.8 provides the necessary prerequisites.

The whole section will use the unbranched covering projection

$$q : \mathbb{H} \rightarrow \Delta^*, \quad \tau \mapsto q(\tau) := e^{2\pi i \tau},$$

to study the number theoretic properties of modular forms f in a neighbourhood of ∞ by considering the Fourier coefficients of their Fourier expansion

$$\hat{f}(q) = \sum_{n=0}^{\infty} a_n \cdot q^n$$

Lemma 3.8 (Fourier series and partial fraction). Assume $k \in \mathbb{N}$. Consider $\tau \in \mathbb{C}$ and set

$$q := e^{2\pi i \tau}.$$

1. For $k \geq 2$ and $\text{Im } \tau \neq 0$ holds

$$\sum_{n \in \mathbb{Z}} \frac{1}{(\tau + n)^k} = \frac{(-2\pi i)^k}{(k-1)!} \cdot \begin{cases} \sum_{r=1}^{\infty} r^{k-1} \cdot q^r & \text{Im } \tau > 0 \\ (-1)^k \cdot \sum_{r=1}^{\infty} r^{k-1} \cdot q^{-r} & \text{Im } \tau < 0 \end{cases}$$

2. The double series

$$\sum_{m, d=1}^{\infty} d \cdot q^{md}$$

satisfies the prerequisites of the rearrangement theorem, i.e. there exists a constant K such that for all finite subsets $A, B \subset \mathbb{N}^*$ holds

$$\sum_{m \in A, d \in B} d \cdot |q|^{md} \leq K$$

3. For even $k \geq 2$, $\text{Im } \tau > 0$,

$$\sum_{m \in \mathbb{Z}^*} \left(\sum_{n \in \mathbb{Z}} \frac{1}{(m\tau + n)^k} \right) = \frac{2(2\pi i)^k}{(k-1)!} \cdot \sum_{n=1}^{\infty} \sigma_{k-1}(n) q^n$$

is absolutely convergent.

Here for each $k \geq 2$ the arithmetic function

$$\sigma_{k-1} : \mathbb{N}^* \rightarrow \mathbb{N}^*, \quad \sigma_{k-1}(n) := \sum_{d|n} d^{k-1},$$

adds the $(k-1)$ -powers of the positive integer divisors of n .

Proof. 1. i) We first consider the case $\text{Im } \tau > 0$. The proof compares the partial fraction expansion and the Fourier expansion of the cotangent function.

- Partial fraction, see [68]:

$$\pi \cdot \cot(\pi \tau) = \frac{1}{\tau} + \sum_{n=1}^{\infty} \left(\frac{1}{\tau + n} + \frac{1}{\tau - n} \right)$$

- Fourier series

$$\pi \cdot \cot(\pi \tau) = \pi \frac{\cos(\pi \tau)}{\sin(\pi \tau)} = (-i\pi) \cdot \frac{1+q}{1-q} = -\pi i - 2\pi i \sum_{n=1}^{\infty} q^n.$$

Here the second equality results from the Euler formulas

$$\cos \pi \tau = \frac{1}{2}(e^{i\pi \tau} + e^{-i\pi \tau}) \text{ and } \sin \pi \tau = \frac{1}{2i}(e^{i\pi \tau} - e^{-i\pi \tau})$$

which imply

$$\cot(\pi\tau) = i \cdot \frac{e^{i\pi\tau} + e^{-i\pi\tau}}{e^{i\pi\tau} - e^{-i\pi\tau}} = i \cdot \frac{q+1}{q-1} = (-i) \cdot \frac{1+q}{1-q}$$

The third equality uses the geometric series

$$\frac{1+q}{1-q} = (1+q) \sum_{n=0}^{\infty} q^n = 1 + 2 \sum_{n=1}^{\infty} q^n$$

which is absolutely convergent because $|q| < 1$ due to $\tau \in \mathbb{H}$.

Equating both expansions gives

$$\frac{1}{\tau} + \sum_{n=1}^{\infty} \left(\frac{1}{\tau+n} + \frac{1}{\tau-n} \right) = -\pi i - 2\pi i \cdot \sum_{n=1}^{\infty} q^n$$

Successive $(k-1)$ -times differentiation with respect to τ , using $q^n = (e^{2\pi i\tau})^n$, shows

$$\sum_{n \in \mathbb{Z}} \frac{1}{(\tau+n)^k} = \frac{(-2\pi i)^k}{(k-1)!} \cdot \sum_{n=1}^{\infty} n^{k-1} q^n.$$

ii) For $\text{Im } \tau < 0$ we use

$$(\tau+n)^k = (-1)^k \cdot (-\tau-n)^k$$

Hence

$$\sum_{n \in \mathbb{Z}} \frac{1}{(\tau+n)^k} = (-1)^k \cdot \sum_{n \in \mathbb{Z}} \frac{1}{(-\tau-n)^k} = (-1)^k \cdot \sum_{n \in \mathbb{Z}} \frac{1}{(-\tau+n)^k}$$

Part i) applied to $-\tau$ implies

$$(-1)^k \cdot \sum_{n \in \mathbb{Z}} \frac{1}{(-\tau+n)^k} = \frac{(-2\pi i)^k}{(k-1)!} \cdot (-1)^k \cdot \sum_{r=1}^{\infty} r^{k-1} \cdot q^{-r}$$

which proves the claim.

2. Consider $\tau \in \mathbb{H}$ with $q := e^{2\pi i\tau}$ real and $0 < q < 1$.

For any real $x \in [0, 1[$ the geometric series satisfies

$$\sum_{d=0}^{\infty} x^d = \frac{1}{1-x}$$

Taking the derivative and multiplying by x shows

$$\sum_{d=1}^{\infty} d \cdot x^d = \frac{x}{(1-x)^2}$$

For arbitrary but fixed $m \geq 1$ set $x := q^m < 1$

$$\sum_{m=1}^{\infty} \left(\sum_{d=1}^{\infty} d \cdot q^{md} \right) = \sum_{m=1}^{\infty} \left(\sum_{d=1}^{\infty} d \cdot (q^m)^d \right) = \sum_{m=1}^{\infty} \frac{q^m}{(1-q^m)^2}$$

The last series is convergent because for fixed q

$$\lim_{m \rightarrow \infty} |1 - q^m|^2 = 1.$$

Set

$$K := \sum_{m=1}^{\infty} \frac{q^m}{(1-q^m)^2}$$

As a consequence

$$\sum_{m \in A, d \in B} d \cdot q^{md} \leq K.$$

For arbitrary $\tau \in \mathbb{H}$ we have

$$q := e^{2\pi i \tau} \text{ with } 0 \leq |q| < 1$$

Hence w.l.o.g. we may assume q real with $0 \leq q < 1$, which proves the claim.

3. We split the exterior summation $\sum_{m \in \mathbb{Z}^*}$ into

$$\sum_{m < 0} \text{ where } \operatorname{Im}(m\tau) < 0 \text{ and } \sum_{m > 0} \text{ where } \operatorname{Im}(m\tau) > 0.$$

Due to part 1) for each $m \in \mathbb{Z}^*$ exists the interior series

$$\sum_{n \in \mathbb{Z}} \frac{1}{(m\tau + n)^k} = \frac{(2\pi i)^k}{(k-1)!} \cdot \begin{cases} \sum_{r=1}^{\infty} r^{k-1} \cdot q^{-rm} & \text{if } m < 0 \\ \sum_{r=1}^{\infty} r^{k-1} \cdot q^{rm} & \text{if } m > 0 \end{cases}$$

Hence for even $k \geq 2$

$$\sum_{m \in \mathbb{Z}^*} \left(\sum_{n \in \mathbb{Z}} \frac{1}{(m\tau + n)^k} \right) = 2 \cdot \frac{(2\pi i)^k}{(k-1)!} \cdot \sum_{m=1}^{\infty} \left(\sum_{r=1}^{\infty} r^{k-1} \cdot q^{rm} \right).$$

The latter series is convergent: For $k \geq 4$ any rearrangement of the original double series is permitted. For $k = 2$ the result of part 2) allows to apply the rearrangement theorem to the double series

$$\sum_{m,r=1}^{\infty} r \cdot q^{rm}, |q| < 1.$$

We rearrange the series by changing the indices of summation from (m, r) to $(n = rm, d|n)$:

$$\sum_{m=1}^{\infty} \left(\sum_{r=1}^{\infty} r^{k-1} \cdot q^{rm} \right) = \sum_{n=1}^{\infty} \left(\sum_{d|n} d^{k-1} \cdot q^n \right) = \sum_{n=1}^{\infty} \sigma_{k-1}(n) \cdot q^n.$$

Eventually for all even $k \geq 2$

$$\sum_{m \in \mathbb{Z}^*} \left(\sum_{n \in \mathbb{Z}} \frac{1}{(m\tau + n)^k} \right) = 2 \cdot \frac{(2\pi i)^k}{(k-1)!} \cdot \sum_{n=1}^k \sigma_{k-1}(n) \cdot q^n$$

□

Eisenstein series are the lattice constants of normalized lattices

$$\Lambda_{1,\tau} = \mathbb{Z} \cdot 1 + \mathbb{Z} \cdot \tau$$

when considered as functions of the argument $\tau \in \mathbb{H}$. Eisenstein series are explicit representations of modular forms.

Theorem 3.9 (Eisenstein series, $k \geq 4$). Consider an even integer $k \geq 4$.

1. The Eisenstein series

$$G_k(\tau) := \sum'_{(c,d) \in \mathbb{Z}^2} \frac{1}{(c\tau + d)^k}$$

is absolutely and compactly convergent for $\tau \in \mathbb{H}$.

The series defines a modular form G_k of Γ_1 of weight k with

$$G_k(\infty) = 2\zeta(k).$$

Here

$$\zeta(s) := \sum_{n=1}^{\infty} (1/n^s), \text{ Re } s > 1,$$

denotes the Riemann ζ -function.

2. The Eisenstein series $G_k \in M_k(\Gamma_1)$ has the Fourier expansion

$$G_k(\tau) = 2\zeta(k) + 2 \frac{(2\pi i)^k}{(k-1)!} \sum_{n=1}^{\infty} \sigma_{k-1}(n) \cdot q^n, \quad q = e^{2\pi i \tau}.$$

Note. The attached prime at

$$\sum_{(c,d) \in \mathbb{Z}^2} ' ,$$

indicates to omit the summand with index $(0,0)$.

Proof. i) *Convergence and transformation:* For $\tau \in \mathbb{H}$ consider the normalized lattice

$$\Lambda_\tau := \mathbb{Z} + \mathbb{Z}\tau$$

and for $k \geq 4$ its lattice constants $G_{\Lambda_\tau, k}$. Apparently

$$G_k(\tau) := G_{\Lambda_\tau, k}$$

The series $G_{\Lambda, k}$ is absolute convergent due to Lemma 1.11, hence allows arbitrary rearrangement. With respect to the translation

$$T : \mathbb{H} \rightarrow \mathbb{H}, \tau \mapsto T(\tau) := \tau + 1,$$

we obtain

$$G_k(\tau) = G_k(\tau + 1).$$

With respect to the reflection

$$S : \mathbb{H} \rightarrow \mathbb{H}, \tau \mapsto S(\tau) := \frac{-1}{\tau}$$

we have

$$G_k(S(\tau)) \cdot h(S, \tau)^{-k} = \frac{1}{\tau^k} \cdot G_k(-(1/\tau)) = \frac{1}{\tau^k} \cdot \sum_{(c,d) \in \mathbb{Z}^2} ' \frac{1}{(-c(1/\tau) + d)^k} = G_k(\tau).$$

Because S and T generate the modular group, Lemma 3.4 implies for all $\gamma \in \Gamma_1$ and for all $\tau \in \mathbb{H}$

$$G_k(\tau) = G_k(\gamma(\tau)) \cdot h(\gamma, \tau)^{-k}, \text{ i.e. } G_k = G_k[\gamma]_k$$

ii) *Fourier expansion and holomorphy:* We split the summation over

$$(c, d) \in \mathbb{Z}^2, (c, d) \neq (0, 0),$$

into the simple sum for (c, d) with $c = 0$ and $d \in \mathbb{Z}^*$ and the double sum with $c \in \mathbb{Z}^*$ and $d \in \mathbb{Z}$. Then Lemma 3.8, part 3) implies

$$\begin{aligned} G_k(\tau) &= \sum_{(c,d) \in \mathbb{Z}^2} ' \frac{1}{(c\tau + d)^k} = \sum_{d \in \mathbb{Z}^*} \frac{1}{d^k} + \sum_{c \in \mathbb{Z}^*, d \in \mathbb{Z}} \frac{1}{(c\tau + d)^k} = \\ &= 2\zeta(k) + \frac{2(2\pi i)^k}{(k-1)!} \cdot \sum_{n=1}^{\infty} \sigma_{k-1}(n) q^n, \quad q = e^{2\pi i \cdot \tau}. \end{aligned}$$

The Fourier expansion, a convergent power series, proves the holomorphy of G_k in $\mathbb{H}$ and at ∞ .

□

In order to get rid of the transcendent values in the Fourier expansion of G_k one divides by the value of the Riemann ζ -function.

Definition 3.10 (Normalized Eisenstein series). For even $k \geq 4$ one defines the *normalized Eisenstein series*

$$E_k := \frac{G_k}{2\zeta(k)} \in M_k(\Gamma_1)$$

If it wasn't apparent by then, at this point it is clear that modular functions like the Eisenstein series capture a lot of information from number theory:

Corollary 3.11 (Normalized Eisenstein series). *For each $k \geq 4$ all Fourier coefficients of the normalized Eisenstein series are rational, more precisely*

$$E_k(\tau) = 1 - \frac{2k}{B_k} \cdot \sum_{n=1}^{\infty} \sigma_{k-1}(n) \cdot q^n \in M_k(\Gamma_1), \quad q = e^{2\pi i \cdot \tau}$$

Proof. 1. *Euler relation of Bernoulli numbers:* See [22, Kap. VII, Satz 4.1]. We relate the cotangent function to the Bernoulli numbers $(B_n)_{n \in \mathbb{N}}$. The Bernoulli numbers are the generators of the power series

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

We compute

$$\cot z = \frac{\cos z}{\sin z} = i \cdot \frac{e^{iz} + e^{-iz}}{e^{iz} - e^{-iz}} = i \cdot \frac{e^{2iz} + 1}{e^{2iz} - 1} = i + \frac{2i}{e^{2iz} - 1}$$

and after multiplication by z

$$z \cdot \cot z = iz + \frac{2iz}{e^{2iz} - 1} = iz + 1 - \frac{2iz}{2} + \sum_{\mu=1}^{\infty} \frac{B_{2\mu}}{(2\mu)!} \cdot (2iz)^{2\mu} = 1 + \sum_{\mu=1}^{\infty} (-1)^\mu \frac{2^{2\mu}}{(2\mu)!} \cdot B_{2\mu} \cdot z^{2\mu}$$

where we have used vanishing of all Bernoulli numbers with odd index ≥ 3 and the first values $B_0 = 1$, $B_1 = -1/2$

- On one hand from the definition of the Bernoulli numbers - replacing z by πz -

$$\pi z \cdot \cot(\pi z) = 1 + \sum_{\mu=1}^{\infty} (-1)^\mu \cdot \frac{2^{2\mu}}{(2\mu)!} \cdot \pi^{2\mu} \cdot B_{2\mu} \cdot z^{2\mu}$$

- On the other hand the partial fraction of the cotangent function

$$\pi z \cdot \cot(\pi z) = 1 + z \sum_{v=1}^{\infty} \frac{2z}{z^2 - v^2}$$

and the geometric series

$$\frac{1}{z^2 - v^2} = -\frac{1}{v^2} \cdot \sum_{\mu=0}^{\infty} \left(\frac{z^2}{v^2} \right)^{\mu}$$

imply

$$\begin{aligned} \pi z \cdot \cot(\pi z) &= 1 + 2z^2 \cdot \sum_{v=1}^{\infty} \left(-\frac{1}{v^2} \cdot \sum_{\mu=0}^{\infty} \left(\frac{z^2}{v^2} \right)^{\mu} \right) = \\ &= 1 - 2 \sum_{\mu=1}^{\infty} \left(\sum_{v=1}^{\infty} \frac{1}{v^{2\mu}} \right) z^{2\mu} \end{aligned}$$

Comparing the coefficients of both series representing

$$\pi z \cdot \cot(\pi z)$$

implies the Euler relation

$$2 \cdot \zeta(2\mu) := 2 \cdot \sum_{v=1}^{\infty} \frac{1}{v^{2\mu}} = (-1)^{\mu+1} \frac{2^{2\mu}}{(2\mu)!} \cdot \pi^{2\mu} \cdot B_{2\mu}$$

2. *Rationality of the Fourier coefficients:* The Fourier expansion of the Eisenstein series G_k from Theorem 3.9 and Part 1) show

$$E_k(\tau) := \frac{G_k(\tau)}{2\zeta(k)} = 1 - \frac{2k}{B_k} \cdot \sum_{n=1}^{\infty} \sigma_{k-1}(n) \cdot q^n \in M_k(\Gamma_1).$$

All Bernoulli numbers are rational due to the recursion formula for $N \in \mathbb{N}^*$

$$\sum_{n=0}^N \binom{N+1}{n} \cdot B_n = 0 \text{ with } B_0 = 1.$$

As a consequence, all Fourier coefficients of E_k are rational.

□

Concerning the first values of Bernoulli numbers

$$\left(k, \zeta(k), B_k, -\frac{2k}{B_k} \right)$$

one has

$$\begin{aligned}
 & (2, \pi^2/6, 1/6, -24) \\
 & (4, \pi^4/(2 \cdot 3^2 \cdot 5), -1/30, 240) \\
 & (6, \pi^6/(3^3 \cdot 5 \cdot 7), 1/42, -504) \\
 & (8, \pi^8/(2 \cdot 3^3 \cdot 5^2 \cdot 7), -1/30, 480) \\
 & (10, \pi^{10}/(3^5 \cdot 5 \cdot 7 \cdot 11), 5/66, -264) \\
 & (12, 691 \cdot \pi^{12}/(3^6 \cdot 5^3 \cdot 7^2 \cdot 11 \cdot 13), -691/2730, 65520/691)
 \end{aligned}$$

For the specific indices $k \in \{4, 6, 8\}$ the numerical value of B_k implies that all Fourier coefficients of the corresponding E_k are not only rationals but even integers.

Remark 3.12 (Eisenstein series).

1. *Rationality of the Fourier coefficients:* The result from Corollary 3.11 is remarkable. The invariance of the holomorphic function E_k under the right action of the modular group implies: The Fourier coefficients of these holomorphic functions are not arbitrary numbers from $\mathbb{C}$ but rationals from $\mathbb{Q}$. This rationality is the reason for the “magic of modular forms” revealed by later applications. See also Proposition 3.14.
2. *Weight $k \geq 4$:* For even $k > 2$ the Eisenstein series G_k relate to the lattice constants $G_{\Lambda, k}$ of an arbitrary lattice with positively oriented basis

$$\Lambda = \mathbb{Z}\omega_1 + \mathbb{Z}\omega_2$$

as follows: Set $\tau := \omega_2/\omega_1 \in \mathbb{H}$. Then

$$G_{\Lambda, k} = \frac{1}{\omega_1^k} G_k(\tau).$$

3. *Weight $k = 2$:* Similar to the series

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

in one dimension, in two dimensions the series

$$\sum_{(m,n) \in \mathbb{Z}^2} \frac{1}{(m\tau + n)^2}$$

is not absolutely convergent. Anyhow, in both cases some specific orders of summation provide certain finite values. In the 2-dimensional case we define the Eisenstein series

$$G_2(\tau) := 2 \cdot \sum_{n=1}^{\infty} \frac{1}{n^2} + \sum_{m \in \mathbb{Z}^*} \left(\sum_{n \in \mathbb{Z}} \frac{1}{(m\tau + n)^2} \right), \quad \tau \in \mathbb{H}.$$

Lemma 3.8, part 3) implies that the sum on the right-hand side is convergent with Fourier series

$$G_2(\tau) = 2 \cdot \zeta(2) - 8\pi^2 \cdot \sum_{n=1}^{\infty} \sigma_1(n) \cdot q^n, \quad q = e^{2\pi i \cdot \tau}.$$

The equality

$$\zeta(2) = \pi^2/6$$

implies for the normalized Eisenstein series of weight $k = 2$

$$E_2(\tau) := \frac{G_2(\tau)}{2\zeta(2)} = 1 - 24 \cdot \sum_{n=1}^{\infty} \sigma_1(n) \cdot q^n, \quad q = e^{2\pi i \cdot \tau}.$$

The convergent Fourier representation shows: G_2 is holomorphic on $\mathbb{H} \cup \{\infty\}$ due to Lemma 3.8, part 3), and invariant with respect to translation.

4. *Transformation of G_2* : With respect to the reflection $S \in \Gamma_1$ the Eisenstein series G_2 transforms as

$$(G_2[S]_2)(\tau) := G_2(S\tau) \cdot \frac{1}{\tau^2} = G_2(\tau) - \frac{2\pi i}{\tau}$$

or

$$(G_2[\gamma]_2)(\tau) = G_2(\tau) - 2\pi i c \cdot h(\gamma, \tau)^{-1}$$

for general

$$\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_1,$$

see [38, Kap. III, § 6, Abschn. 1] or [19, Exerc. 1.2.8]. In particular, G_2 is not weakly modular. The lack of modularity is due to the conditional convergence of G_2 .

Proposition 1.19 dealt with the discriminant of the cubic polynomial from the differential equation of the Weierstrass $\wp$ -function of a lattice. Definition 3.13 shows: These discriminants are the values of a modular form.

Definition 3.13 (Discriminant form). The *discriminant form* is the modular form of Γ_1 of weight $k = 12$ defined as

$$\Delta := g_2^3 - 27 \cdot g_3^2 : \mathbb{H} \rightarrow \mathbb{C}$$

derived from the Eisenstein series

$$g_2 := 60 \cdot G_4, \quad g_3 := 140 \cdot G_6.$$

In Definition 3.13 the letters g_2 and g_3 designate modular functions, not the constants from Theorem 1.17, which denote the value at the period τ of a given normalized lattice

$$\Lambda = \mathbb{Z} \cdot 1 + \mathbb{Z} \cdot \tau.$$

Proposition 3.14 (Fourier coefficients of the discriminant form). *All Fourier coefficients $(\tau_n)_{n \in \mathbb{N}}$ of the normalized discriminant form*

$$\frac{\Delta}{(2\pi)^{12}} \in M_{12}(\Gamma_1)$$

are integers:

$$\frac{\Delta(q)}{(2\pi)^{12}} = \sum_{n=0}^{\infty} \tau_n \cdot q^n \in \mathbb{Z}\{q\}, \quad q = e^{2\pi i \cdot \tau},$$

with $\tau_0 = 0$ and $\tau_1 = 1$.

In particular $\Delta \in S_{12}(\Gamma_1)$ is a cusp form of weight 12, and the zero at $\infty \in \mathbb{H}^*$ has order = 1.

The Fourier coefficients τ_n , $n \geq 0$, are the subject of certain of Ramanujan's well-known conjectures, see Definition 5.31.

Proof. i) *Fourier series:* We refer to Definition 3.10 of the normalized Eisenstein series and their Fourier series expansion from Corollary 3.11. Using

$$\zeta(4) = \frac{\pi^4}{90} \text{ and } \zeta(6) = \frac{\pi^6}{945}$$

we obtain

$$E_4 = 3^2 \cdot 5 \cdot \frac{G_4}{\pi^4} = 1 + 240 \cdot \sum_{n=1}^{\infty} \sigma_3(n) \cdot q^n$$

$$E_6 = \frac{3^3 \cdot 5 \cdot 7}{2} \cdot \frac{G_6}{\pi^6} = 1 - 504 \cdot \sum_{n=1}^{\infty} \sigma_5(n) \cdot q^n.$$

We have

$$g_2 = 60 \cdot G_4 = \frac{2^2}{3} \pi^4 \cdot E_4 \text{ and } g_3 = 140 \cdot G_6 = \frac{2^3}{3^3} \pi^6 \cdot E_6.$$

Hence

$$\Delta = g_2^3 - 3^3 g_3^2 = \pi^{12} \cdot \frac{2^6}{3^3} (E_4^3 - E_6^2) = (2\pi)^{12} \cdot \frac{E_4^3 - E_6^2}{2^6 \cdot 3^3} = (2\pi)^{12} \cdot \frac{E_4^3 - E_6^2}{1728}$$

and eventually

$$\frac{\Delta}{(2\pi)^{12}} = \frac{E_4^3 - E_6^2}{2^6 \cdot 3^3} = \frac{E_4^3 - E_6^2}{1728}.$$

We set as shorthand

$$X_k(q) := \sum_{n=1}^{\infty} \sigma_k(n) \cdot q^n$$

and obtain

$$\frac{\Delta}{(2\pi)^{12}} = \frac{(1 + 240 \cdot X_3)^3 - (1 - 504 \cdot X_5)^2}{2^6 \cdot 3^3}$$

When denoting by

$$Q = Q(q)$$

the numerator of the latter fraction, then

$$Q = (3 \cdot 240 \cdot X_3 + 3 \cdot 240^2 \cdot X_3^2 + 240^3 \cdot X_3^3) + (1008 \cdot X_5 - 504^2 \cdot X_5^2)$$

ii) *Fourier coefficients* $\tau_0 = 1, \tau_1 = 1$: In lowest order in q with

$$X_3(q) = q + O(2) \text{ and } X_5(q) = q + O(2)$$

we get

$$Q(q) = 3 \cdot 240 \cdot q + 1008 \cdot q + O(2) = 1728 \cdot q + O(2)$$

which proves $\tau_0 = 0$ and $\tau_1 = 1$.

iii) *Fourier coefficients* $\tau_n \in \mathbb{Z}, n \geq 2$: The remaining part of the proof reduces to the claim that also all other coefficients of the series Q are divisible by

$$1728 = 2^6 \cdot 3^3 = (2^2 \cdot 3)^3 = 12^3$$

We compute

$$Q = (3 \cdot 240 \cdot X_3 + 3 \cdot 240^2 \cdot X_3^2 + 240^3 \cdot X_3^3) + (1008 \cdot X_5 - 504^2 \cdot X_5^2)$$

$$Q \equiv 3 \cdot 240 \cdot X_3 + 1008 \cdot X_5 \equiv 12^2(5 \cdot X_3 + 7 \cdot X_5) \pmod{12^3}$$

To complete the proof we have to show that all coefficients of the power series

$$5 \cdot X_3 + 7 \cdot X_5$$

are divisible by 12. We show even more: For all $d \in \mathbb{Z}$

$$5 \cdot d^3 + 7 \cdot d^5 \equiv 0 \pmod{12}$$

According to the Chinese remainder theorem the claim reduces to the two congruences

$$5 \cdot d^3 + 7 \cdot d^5 \equiv 0 \pmod{3} \text{ and } 5 \cdot d^3 + 7 \cdot d^5 \equiv 0 \pmod{4}$$

or

$$d^5 \equiv d^3 \pmod{3, \text{ mod } 4}.$$

The latter two congruences reduce to

- If $d^3 \not\equiv 0 \pmod{3}$ then $d^2 \equiv 1 \pmod{3}$,
- and: if $d^3 \not\equiv 0 \pmod{4}$ then $d^2 \equiv 1 \pmod{4}$

The congruences have to be checked for

$$d \in \{1, 2\} \subset \mathbb{Z}/3\mathbb{Z} = \{0, 1, 2\} \text{ and } d \in \{1, 3\} \subset \mathbb{Z}/4\mathbb{Z} = \{0, 1, 2, 3\},$$

which can be easily done.

For the relation between the Ramanujan function and the theory of Monstrous Moonshine see Section 7.2. $\square$

Definition 3.15 (Modular invariant j). The *modular invariant* j is the automorphic function, defined as the quotient of two modular forms of weight = 12,

$$j := 1728 \cdot \frac{g_2^3}{\Delta} = 1728 \cdot \frac{g_2^3}{g_2^3 - 27g_3^2}.$$

Corollary 3.16 (Modular curve and modular invariant j). The *modular invariant* is a biholomorphic map

$$j : X(1) \xrightarrow{\sim} \mathbb{P}^1$$

which maps the cusp of $X(1)$ to $\infty \in \mathbb{P}^1$.

Proof. Proposition 1.19 implies that j is a holomorphic function on $\mathbb{H}$ because the denominator Δ has no zeros in $\mathbb{H}$. The Fourier expansion of

$$\frac{\Delta}{(2\pi)^{12}}$$

from Proposition 3.14 and the equality

$$G_4(\infty) = 2\zeta(4) \neq 0$$

from Theorem 3.9 show that j has a pole of order = 1 at the point ∞ . Because j is even an automorphic *function*, it is a meromorphic function on the modular curve, i.e. a holomorphic map

$$j : X(1) \xrightarrow{\sim} \mathbb{P}^1.$$

Because all fibres of j have the same multiplicity = 1, the map j is biholomorphic.

$\square$

We now come back to Proposition 3.1 and elaborate on investigating the action of the modular group $\Gamma_1 = SL(2, \mathbb{Z})$ as an action on differential forms. We set out to show: Modular forms of weight k are those meromorphic differential forms on the modular curves $X_0(N)$ which are multiples of a certain divisor which depends on the level N and the weight k . Hence modular forms are holomorphic sections of certain line bundles on the modular curves. Applying the Riemann-Roch theorem we draw several conclusions. In detail:

- We recall the Riemann-Roch theorem on a compact Riemann surface X of genus $g(X)$ for a divisor $D \in \text{Div}(X)$, and also Serre duality, see [69]:

$$\dim H^0(X, \mathcal{O}_D) - \dim H^1(X, \mathcal{O}_D) = 1 - g(X) + \deg D.$$

We will use also the notation

$$\deg \mathcal{O}_D := \deg D.$$

The canonical divisor $K \in \text{Div}(X)$ has degree

$$\deg K = 2g(X) - 2.$$

Serre-duality allows to replace the cohomology group H^1 by the dual of a cohomology group H^0

$$H^1(X, \mathcal{O}_D) = H^0(X, \mathcal{O}_{K-D})^\vee.$$

Moreover, for each divisor $D \in \text{Div}(X)$ with $\deg D < 0$

$$\dim H^0(X, \mathcal{O}_D) = 0.$$

- We derive a formula for the dimension of the vector spaces $M_k(\Gamma_1)$. Section 3.3 will generalize the computation to $M_k(\Gamma_0(N))$.
- Further on, Chapter 4 will show: Any complex elliptic curve arises from a complex torus. More precisely, the elliptic curve is parametrized by the two meromorphic functions $\wp_\Lambda$ and $\wp'_\Lambda$ defined on a suitable torus $\mathbb{C}/\Lambda$, see Theorem 4.27.

We recall the canonical projection to the modular curve

$$p : \mathbb{H}^* \rightarrow X := \Gamma_1 \backslash \mathbb{H}^*,$$

skipping the $*$ -superscript used in the notation from Theorem 2.31. We introduce the class notation

$$[\tau] = p(\tau) \in X$$

and identify this class and the corresponding point divisor $[\tau] \in \text{Div}(X)$. In addition we use the floor function: For $x \in \mathbb{R}$

$$\lfloor x \rfloor := \max\{n \in \mathbb{Z} : n \leq x\}.$$

Theorem 3.17 (Modular forms as coefficients of meromorphic differential forms on the modular curve). *Consider the modular curve $X = X(1)$ of the modular group Γ_1 and fix an even weight $k \geq 0$.*

1. Differential forms: *The map of vector spaces is injective*

$$\alpha : M_k(\Gamma_1) \rightarrow H^0\left(X, \mathcal{M} \otimes_{\mathcal{O}} \Omega^{\otimes k/2}\right), f \mapsto \omega_f := f \cdot (d\tau)^{k/2}.$$

2. The distinguished divisor: *Consider the divisor*

$$D_k := \left\lfloor \frac{k}{4} \right\rfloor \cdot [i] + \left\lfloor \frac{k}{3} \right\rfloor \cdot [\rho] + \frac{k}{2} \cdot [\infty] \in \text{Div}(X).$$

The attachment

$$U \mapsto \Omega_{\Gamma_1, k}(U), U \subset X \text{ open},$$

with

$$\Omega_{\Gamma_1, k}(U) := \left\{ \omega \in H^0(U, \mathcal{M} \otimes_{\mathcal{O}} \Omega^{\otimes k/2}) : \omega \neq 0 \text{ and } \text{div } \omega \geq -D_k|_U \right\} \cup \{0\}$$

is a sheaf $\Omega_{\Gamma_1, k}$ on X . The sheaf $\Omega_{\Gamma_1, k}$

- *is a subsheaf of the sheaf of meromorphic differential forms on X , i.e.*

$$\Omega_{\Gamma_1, k} \subset \mathcal{M} \otimes_{\mathcal{O}} \Omega^{\otimes k/2}$$

- *and is isomorphic to the line bundle*

$$\mathcal{L}_{\Gamma_1, k} := \mathcal{O}_{(k/2) \cdot K + D_k} \in \text{Pic}(X)$$

of multiples of the divisor $-((k/2) \cdot K + D_k) \in \text{Div}(X)$ with $K \in \text{Div}(X)$ the canonical divisor of X .

3. Meromorphic differential forms on the modular curve: *The image of the map α from part 1) is*

$$\alpha(M_k(\Gamma_1)) = H^0(X, \Omega_{\Gamma_1, k}).$$

In particular, sections of the sheaf $\Omega_{\Gamma_1, k}$, i.e. certain *meromorphic* differential forms, correspond bijectively to *holomorphic* sections of the line bundle $\mathcal{L}_{\Gamma_1, k}$.

Proof. Cf. [54, §4]. The numbers of the different parts of the proof do not correspond one-to-one to the sections of the theorem.

i) *Definition of α* : If $f \in M_k(\Gamma_1)$ then Proposition 3.1 implies that the modular group leaves invariant the differential form

$$f(\tau) (d\tau)^{k/2}$$

on $\mathbb{H}$. Hence the latter defines a meromorphic form ω_f on X with pullback

$$f(\tau) (d\tau)^{k/2} := p^* \omega_f :$$

One considers first the form ω_f outside the three points which are the orbits of the two elliptic points and the cusp. Secondly one extends ω_f holomorphically into these exceptional points by using Riemann's extension theorem for bounded holomorphic functions.

ii) *Distinguished divisor*: The sheaf $\Omega_{\Gamma_1, k}$ is well-defined. The definition of the divisor D_k results from the following relation between the order of f at $t \in \mathbb{H}^*$ and the order of ω_f at $[t] \in X$:

$$\text{ord}_t f = \begin{cases} \text{ord}_{[\infty]} \omega_f + k/2 & t = \infty \\ h_t \cdot \text{ord}_{[t]} \omega_f + (k/2) \cdot (h_t - 1) & t \in \mathbb{H} \end{cases}$$

Here h_t denotes the period of t . The order of a meromorphic differential form

$$\omega \in H^0\left(X, \mathcal{M} \otimes_{\mathcal{O}} \Omega^{\otimes k/2}\right)$$

at a point $x \in X$ is defined as follows: Choose a chart of X around x

$$q : U \rightarrow V \subset \mathbb{C}$$

and represent

$$\omega|_U = f \cdot (dq)^{k/2}$$

Then

$$\text{ord}_x \omega := \text{ord}_x f,$$

independently from the choice of the chart. To prove the order relation for $t \in \{i, \rho, \infty\}$ we use the local form of

$$p : \mathbb{H}^* \rightarrow X$$

and the charts around t from Theorem 2.31 and Proposition 2.30. We denote by τ points from $U_t \subset \mathbb{H}$ and by q points from $W_t \subset \mathbb{C}$:

- For $t = \infty \in \mathbb{H}^*$ we recall the chart pow_∞ around t and the chart ϕ_∞ around the cusp $[\infty]$ from the diagram

$$\begin{array}{ccc}
 U_\infty & \xrightarrow{p|U_\infty} & p(U_\infty) \\
 & \searrow \text{pow}_\infty & \downarrow \phi_\infty \\
 & & W_\infty
 \end{array}$$

For $\tau \in U_\infty$

$$q(\tau) := \text{pow}_\infty(\tau) := \begin{cases} e^{2\pi i \tau} & \tau \neq \infty \\ 0 & \tau = \infty \end{cases}$$

We denote by τ the points on $U_\infty \subset \mathbb{H}$ and by q the points on $W_\infty \subset \Delta \subset \mathbb{C}$. Then

$$dq = 2\pi i q d\tau \text{ and } (dq)^{k/2} = (2\pi i)^{k/2} q^{k/2} (d\tau)^{k/2}$$

which implies

$$(d\tau)^{k/2} = \frac{1}{(2\pi i)^{k/2}} \cdot \frac{(dq)^{k/2}}{q^{k/2}}$$

The order relation relates

$$\text{ord}_\infty f = \text{ord}(\hat{f}; q=0) = \text{ord}(p^* \omega_f; \tau = \infty)$$

to

$$\text{ord}_{[\infty]} \omega_f = \text{ord}((\phi_\infty^{-1})^* \omega_f; q=0) = \text{ord}(h; q=0)$$

with

$$h(q) (dq)^{k/2} := (\phi_\infty^{-1})^* \omega_f$$

The commutative diagram

$$p = \phi^{-1} \circ \text{pow}_\infty \text{ implies } p^* = \text{pow}_\infty^* \circ (\phi_\infty^{-1})^*$$

On one hand

$$p^* \omega_f = f(\tau) (d\tau)^{k/2}$$

On the other hand,

$$\text{pow}_\infty^* ((\phi_\infty^{-1})^* \omega_f) = \text{pow}_\infty^* (h(q) (dq)^{k/2}) = h(q(\tau)) \cdot (2\pi i)^{k/2} \cdot q(\tau)^{k/2} (d\tau)^{k/2}$$

Hence

$$f(\tau) = h(q(\tau)) \cdot (2\pi i)^{k/2} \cdot q(\tau)^{k/2}$$

and

$$\hat{f}(q) = h(q) \cdot (2\pi i)^{k/2} \cdot q^{k/2}$$

which implies

$$\text{ord}(\hat{f}; q=0) = \text{ord}(h; q=0) + k/2$$

or

$$\text{ord}_\infty f = \text{ord}_{[\infty]} \omega_f + k/2$$

- For $t \in \{i, \rho\}$ we recall the chart ϕ_t from the following diagram

$$\begin{array}{ccc} U_t & \xrightarrow{p|U_t} & p(U_t) \\ \lambda_t \downarrow & & \downarrow \phi_t \\ \Delta_t & \xrightarrow{\text{pow}_t} & W_t \end{array}$$

The map

$$\lambda_t : U_t \xrightarrow{\sim} \Delta_t$$

is biholomorphic. For the order calculation we assume w.l.o.g.

$$\lambda_t = \text{id} \text{ and } U_t = \Delta_t$$

The isotropy group I_t acts on U_t by multiplication by an e -th root of unity with

$$e := h_t$$

For $\tau \in U_t$

$$q(\tau) := \text{pow}_t(\tau) := \tau^e$$

Therefore

$$dq = e \cdot \tau^{e-1} d\tau \text{ and } (dq)^{k/2} = e^{k/2} \cdot \tau^{(k/2) \cdot (e-1)} (d\tau)^{k/2}$$

If

$$\omega_f(q) = g(q) (dq)^{k/2}$$

then

$$(p^* \omega_f)(\tau) = \omega_f(q(\tau)) = g(q(\tau)) \cdot e^{k/2} \cdot \tau^{(k/2) \cdot (e-1)} (d\tau)^{k/2} = f(\tau) (d\tau)^{k/2}$$

As a consequence

$$f(\tau) = g(q(\tau)) \cdot \tau^{(k/2) \cdot (e-1)} \cdot e^{k/2}$$

By definition

$$\text{ord}_{[t]} \omega_f = \text{ord}_0 g$$

Hence

$$\text{ord}_t f = e \cdot \text{ord}_0 g + (k/2) \cdot (e-1) = e \cdot \text{ord}_{[t]} \omega_f + (k/2) \cdot (e-1) =$$

$$= h_t \cdot \text{ord}_{[t]} \omega_f + (k/2) \cdot (h_t - 1)$$

Summing up

$$\text{ord}_t f = \text{ord}_{[t]} \omega_f \text{ if } t \in \mathcal{D} \setminus \{i, \rho, \infty\}$$

$$\text{ord}_i f = 2 \cdot \text{ord}_{[i]} \omega_f + k/2$$

$$\text{ord}_\rho f = 3 \cdot \text{ord}_{[\rho]} \omega_f + k$$

$$\text{ord}_\infty f = \text{ord}_{[\infty]} \omega_f + k/2$$

or, because for all $t \in \mathbb{H}^*$ holds $\text{ord}_t f \geq 0$:

$$\text{ord}_{[t]} \omega_f = \text{ord}_t f \geq 0 \text{ if } t \in \mathcal{D} \setminus \{i, \rho, \infty\}$$

$$\text{ord}_{[i]} \omega_f = 1/2 \cdot \text{ord}_i f - k/4 \geq -k/4, \text{ hence } \text{ord}_{[i]} \omega_f \geq -\lfloor k/4 \rfloor$$

$$\text{ord}_{[\rho]} \omega_f = 1/3 \cdot \text{ord}_\rho f - k/3, \text{ hence } \text{ord}_{[\rho]} \omega_f \geq -\lfloor k/3 \rfloor$$

$$\text{ord}_{[\infty]} \omega_f = \text{ord}_\infty f - k/2, \text{ hence } \text{ord}_{[\infty]} \omega_f \geq -k/2$$

which finishes the proof of the order relations. As a consequence

$$\alpha(M_k(\Gamma_1)) \subset \left\{ \omega \in H^0 \left(X, \mathcal{M} \otimes_{\mathcal{O}} \Omega^{\otimes k/2} \right) : \text{div } \omega \geq - \left\lfloor \frac{k}{4} \right\rfloor \cdot [i] - \left\lfloor \frac{k}{3} \right\rfloor \cdot [\rho] - \frac{k}{2} \cdot [\infty] \right\}$$

Conversely, each $\omega \in H^0(X, \Omega_{\Gamma_1, k})$ defines the modular form $f \in M_k(\Gamma_1)$ with

$$f(\tau) (d\tau)^{k/2} = p^* \omega$$

iii) *Line bundle*: It is well known that the sheaves of multiples of a divisor are line bundles. The sheaf of multiples of the sum of two divisors is the tensor product of the sheaves of multiples of each summand, see [69].

□

The formulas in the proof of Theorem 3.17, part 2) show: While a modular form f on $\mathbb{H}^*$ is holomorphic when considered as a function, the representing differential form ω_f on the modular curve X may have poles. Actually, it is the divisor of ω_f which encodes the information about f .

In the language of line bundles, modular forms of weight $= k$ are the sections of the line bundle $\mathcal{L}_{\Gamma_1, k}$ on the modular curve:

$$M_k(\Gamma_1) \simeq H^0(X, \mathcal{L}_{\Gamma_1, k})$$

Example 3.18 (Cusp forms $S_2(\Gamma_1)$).

1. There are no cusp forms of weight $= 2$, i.e.

$$S_2(\Gamma_1) = 0.$$

Proof. Each cusp form

$$f \in S_k(\Gamma_1)$$

is holomorphic on $\mathbb{H}^*$ and vanishes at $\infty \in \mathbb{H}^*$. Hence

$$\text{ord}_\tau f \geq 0, \tau \in \mathbb{H}, \text{ and } \text{ord}_\infty f \geq 1.$$

The order relations from Theorem 3.17 imply

$$\omega_f := \alpha(f) \in H^0(X(1), \Omega_{\Gamma_1, k})$$

satisfies for $\tau \in \mathcal{D} \setminus \{[i], [\rho], [\infty]\}$

$$\text{ord}_{[\tau]} \omega_f = \text{ord}_{[\tau]} f \geq 0,$$

and for $k = 2$ at the elliptic points and the cusp:

$$\text{ord}_{[i]} \omega_f = (1/2)(\text{ord}_i f - 1) \geq -(1/2), \text{ hence } \text{ord}_{[i]} \omega_f \geq 0$$

$$\text{ord}_{[\rho]} \omega_f = (1/3)(\text{ord}_\rho f - 2) \geq -(2/3), \text{ hence } \text{ord}_{[\rho]} \omega_f \geq 0$$

$$\text{ord}_{[\infty]} \omega_f = \text{ord}_\infty f - 1 \geq 0$$

Therefore

$$\Omega_{\Gamma_1, 2} \simeq \Omega$$

and ω_f is holomorphic, i.e.

$$\omega_f \in H^0(X, \Omega)$$

The group $H^0(X, \Omega)$ vanishes because

$$\dim H^0(X, \Omega) = g(X) = g(\mathbb{P}^1) = 0$$

□

2. Corollary 3.28 will show that the latter formula generalizes to Hecke congruence subgroups as the isomorphy of vector spaces:

$$S_2(\Gamma_0(N)) \simeq H^0(X_0(N), \Omega)$$

which implies

$$\dim S_2(\Gamma_0(N)) = g(X_0(N))$$

The calculation from Example 3.18 carries over to arbitrary even weights $k \in \mathbb{Z}$.

Corollary 3.19 (Dimension of $M_k(\Gamma_1)$). *The vector spaces $M_k(\Gamma_1)$ of modular forms of even weight $k \geq 0$ have the following dimensions:*

$$\dim M_k(\Gamma_1) = \begin{cases} \lfloor k/12 \rfloor & k \equiv 2 \pmod{12} \\ \lfloor k/12 \rfloor + 1 & \text{otherwise} \end{cases}$$

For the dimension of the vector space $S_k(\Gamma_1)$ of cusp forms see Lemma 3.23.

Proof. We have

$$M_k(\Gamma_1) \simeq H^0(X, \mathcal{L}_{\Gamma_1, k}).$$

i) For $k = 0$ we have on the compact Riemann surface $X \simeq \mathbb{P}^1$

$$\dim M_0(\Gamma_1) = \dim H^0(X, \mathcal{O}) = 1.$$

Therefore we are left with even $k > 0$:

ii) *Serre duality and Riemann-Roch theorem:* By Serre-duality for each $k > 0$

$$\dim H^1(X, \mathcal{L}_{\Gamma_1, k}) = \dim H^0(X, \mathcal{L}_{\Gamma_1, k}^\vee \otimes_{\mathcal{O}} \mathcal{O}_K) = \dim H^0(X, \mathcal{O}_{-(k/2) \cdot K - D_k + K})$$

Now

$$\begin{aligned} \deg(-(k/2) \cdot K - D_k + K) &= ((k/2) - 2) - \lfloor k/4 \rfloor - \lfloor k/3 \rfloor \leq \\ &\leq ((k/2) - 2) - ((k/4) - 1) - ((k/3) - 1) = k/2 - k/4 - k/3 = k \cdot (1/2 - 1/4 - 1/3) < 0 \end{aligned}$$

As a consequence

$$0 = \dim H^0(X, \mathcal{L}_{\Gamma_1, k}^\vee \otimes_{\mathcal{O}} \mathcal{O}_K) = \dim H^1(X, \mathcal{L}_{\Gamma_1, k})$$

The Riemann-Roch theorem implies for $\mathcal{L}_{\Gamma_1, k} = \mathcal{O}_{(k/2)K + D_k}$

$$\begin{aligned} \dim M_k(\Gamma_1) &= \dim H^0(X, \mathcal{L}_{\Gamma_1, k}) = 1 - g(X) + (k/2)(-2) + \lfloor k/4 \rfloor + \lfloor k/3 \rfloor + k/2 = \\ &= 1 - k + \lfloor k/4 \rfloor + \lfloor k/3 \rfloor + k/2 = 1 - (k/2) + \lfloor k/3 \rfloor + \lfloor k/4 \rfloor \end{aligned}$$

iii) *Evaluating the formula for $\dim M_k(\Gamma_1)$:* We consider the different values of even $k \pmod{12}$ separately. We make the ansatz

$$k = \alpha \cdot 12 + \beta, \quad \beta \in \{0, 2, \dots, 10\}$$

and compute for all values of β

$$1 - (k/2) + \lfloor k/3 \rfloor + \lfloor k/4 \rfloor$$

• Residue $\beta = 0$:

$$1 - 6\alpha + 4\alpha + 3\alpha = 1 + \alpha = 1 + \lfloor k/12 \rfloor$$

• Residue $\beta = 2$:

$$1 - (6\alpha + 1) + \lfloor 4\alpha + 2/3 \rfloor + \lfloor 3\alpha + 2/4 \rfloor = 1 - (6\alpha + 1) + 4\alpha + 3\alpha = \alpha = \lfloor k/12 \rfloor$$

- Residue $\beta \in \{4, 6, 8, 10\}$:

$$\begin{aligned} & 1 - (6\alpha + \beta/2) + \lfloor 4\alpha + \beta/3 \rfloor + \lfloor 3\alpha + \beta/4 \rfloor = \\ & = 1 - 6\alpha - \beta/2 + 4\alpha + \lfloor \beta/3 \rfloor + 3\alpha + \lfloor \beta/4 \rfloor = 1 + \alpha + (-\beta/2 + \lfloor \beta/3 \rfloor + \lfloor \beta/4 \rfloor) = \\ & = 1 + \alpha = 1 + \lfloor k/12 \rfloor \end{aligned}$$

□

In particular

$$\dim M_0 = 1, \dim M_2 = 0, \dim M_4 = \dim M_6 = \dim M_8 = \dim M_{10} = 1, \dim M_{12} = 2.$$

The non-zero vector spaces in the last line are spanned by the Eisenstein series G_4 , G_6 and their products. Only M_{12} contains a non-zero cusp form, namely the discriminant modular form Δ .

Apparently the numerical value $\dim H^0(X, \mathcal{L}_{\Gamma_1, k})$ can be obtained without applying the theorem of Riemann-Roch and Serre duality in the proof of Corollary 3.19: On the modular curve $X \simeq \mathbb{P}^1$ each line bundle $\mathcal{L}$ is determined by its degree m , and for $m \geq 0$

$$\dim H^0(\mathbb{P}^1, \mathcal{L}) = m + 1.$$

We gave the proof in the above form, because Theorem 3.26 will generalize the proof to the modular curves $X_0(N)$ of the Hecke congruence subgroups.

Corollary 3.20 (Weight formula for modular forms). *Consider an even integer $k \geq 4$ and a modular form $f \in M_k(\Gamma_1)$. Then the sum of the orders of f satisfies the following equation:*

$$\text{ord}_\infty f + \frac{1}{2} \text{ord}_i f + \frac{1}{3} \text{ord}_\rho f + \sum'_\tau \text{ord}_\tau f = \frac{k}{12}$$

Here the symbol $\sum'_\tau$ denotes the summation over all points $\tau \in \mathcal{D} \setminus \{i, \rho\}$.

Proof. We recall the relation between the order of $f \in M_k(\Gamma_1)$ at a point τ and the order of the differential form $\omega_f \in H^0(X, \mathcal{M} \otimes_{\mathcal{O}} \Omega^{\otimes k/2})$ at the point $[\tau] \in X$ established during the proof of Theorem 3.17:

$$\text{ord}_\infty f = \text{ord}_{[\infty]} \omega_f + k/2$$

$$\text{ord}_i f = 2 \text{ord}_{[i]} \omega_f + k/2$$

$$\text{ord}_\rho f = 3 \text{ord}_{[\rho]} \omega_f + k$$

$$\text{ord}_\tau f = \text{ord}_{[\tau]} \omega_f, \tau \in \mathcal{D} \setminus \{i, \rho, \infty\}.$$

The modular curve X has genus $g = 0$ according to Corollary 3.16. Therefore the line bundle Ω_X has degree -2 , and the tensor product $\Omega_X^{\otimes k/2}$ has degree

$$= (-2) \cdot k/2 = -k$$

which implies

$$\deg \text{div}(\omega_f) = -k.$$

As a consequence

$$\begin{aligned} & \text{ord}_\infty f + \frac{1}{2} \text{ord}_i f + \frac{1}{3} \text{ord}_\rho f + \sum'_\tau \text{ord}_\tau f = \\ &= \left(\text{ord}_{[\infty]} \omega_f + \frac{k}{2} \right) + \left(\text{ord}_{[i]} \omega_f + \frac{k}{4} \right) + \left(\text{ord}_{[\rho]} \omega_f + \frac{k}{3} \right) + \sum'_\tau \text{ord}_{[\tau]} \omega_f = \\ &= \deg \text{div}(\omega_f) + \frac{k}{2} + \frac{k}{4} + \frac{k}{3} = -k + \frac{k}{2} + \frac{k}{4} + \frac{k}{3} = \frac{k}{12} \end{aligned}$$

□

Corollary 3.21 (Specific values and Fourier coefficients of the modular invariant j).

1. *The modular invariant*

$$j : X(1) \rightarrow \mathbb{P}^1$$

attains at the elliptic points the distinguished values

$$j([i]) = 1728 \text{ and } j([\rho]) = 0.$$

2. *All Fourier coefficients $(b_n)_{n \geq -1}$ of the modular invariant*

$$j(q) = \sum_{n=-1}^{\infty} b_n \cdot q^n$$

are integers, and $b_{-1} = 1$.

Proof. 1. Corollary 3.20 implies $G_4(\rho) = 0$ and $G_6(i) = 0$. The definition

$$j = 1728 \cdot \frac{g_2^3}{g_2^3 - 27g_3^2}$$

implies the values of j at the distinguished orbits.

2. To compute the Fourier expansion of

$$j = 2^6 \cdot 3^3 \cdot \frac{g_2^3}{\Delta}$$

we use the definition

$$g_2 = \frac{2^2}{3} \cdot \pi^4 \cdot E_4, \text{ hence } g_2^3 = \frac{2^6}{3^3} \cdot \pi^{12} \cdot E_4^3,$$

and secondly the formula from Proposition 3.14

$$\frac{\Delta(\tau)}{(2\pi)^{12}} = \sum_{n=1}^{\infty} \tau_n \cdot q^n \in \mathbb{Z}\{q\}, \quad q = e^{2\pi i \cdot \tau}.$$

We obtain

$$\begin{aligned} j &= \left(2^6 \cdot 3^3\right) \cdot \frac{g_2^3}{\Delta} = \left(2^6 \cdot 3^3\right) \cdot \left(\frac{2^6}{3^3} \cdot E_4^3 \cdot \pi^{12}\right) \cdot \frac{1}{\Delta} = \\ &= \frac{E_4^3}{\Delta / (2\pi)^{12}} = \frac{E_4^3}{\sum_{n=1}^{\infty} \tau_n \cdot q^n} = \frac{1}{q} \cdot \frac{E_4^3}{1 + \sum_{n=2}^{\infty} \tau_n \cdot q^{n-1}} = \\ &= \frac{1}{q} \cdot E_4^3 \cdot \sum_{m=0}^{\infty} \left((-1)^m \cdot \sum_{n=2}^{\infty} \tau_n \cdot q^{n-1} \right)^m = \frac{1}{q} \cdot (1 + O(q)) \end{aligned}$$

by applying the formula of the geometric series and using

$$E_4(q) = 1 + O(q)$$

from Corollary 3.11 and the fact, that E_4 has integer Fourier coefficients.

□

Corollary 3.22 (The field of automorphic functions).

1. If an automorphic function $f \in A_0(\Gamma_1)$ is holomorphic on $\mathbb{H}$ with Fourier series

$$f(\tau) = \sum_{n \geq -k} c_n \cdot q^n, \quad c_{-k} \neq 0,$$

for a suitable $k \in \mathbb{N}$, then a polynomial $P(X) \in \mathbb{C}[X]$ of degree k exists with

- $f = P(j)$,
- and all coefficients of $P(X)$ are $\mathbb{Z}$ -linear combinations of the Fourier coefficients $c_{-k}, \dots, c_{-1}, c_0$.

2. The field of automorphic functions equals the field of meromorphic functions on the modular curve X

$$A_0(\Gamma_1) = \mathbb{C}(j).$$

Thanks to Corollary 3.16 the second part of Corollary 3.22 is a disguised version of the well-known fact that the field of meromorphic functions on $\mathbb{P}^1$ is $\mathbb{C}(z)$, the field of rational functions in one complex variable.

Proof. 1. The proof goes by induction on the pole order k at ∞ : If $k = 0$ then f is a modular function, hence equals the constant c_0 . For the induction step $k - 1 \mapsto k$ consider the function

$$g := f - c_{-k} \cdot j^k.$$

The function g has at most a pole of order $= k - 1$ at ∞ . Therefore the induction assumption applies to g .

2. Each automorphic function $f \in A_0(\Gamma_1)$ can be considered a meromorphic function $f \in \mathcal{M}(\mathbb{P}^1)$ on the projective space. Denote by

$$A := \{x_1, \dots, x_n\} \subset \mathbb{H} \cup \{\infty\} \text{ and } B = \{y_1, \dots, y_n\} \subset \mathbb{H} \cup \{\infty\}$$

the respective zero set and pole set of f , counted with multiplicity. Both sets have the same cardinality because a non-constant holomorphic map between compact Riemann surfaces attains each value with the same multiplicity, see [69]. Then

$$g := \prod_{v=1}^n \frac{j - j(x_v)}{j - j(y_v)} \in A_0(\Gamma_1)$$

has the same zero set and pole set as f . The quotient f/g has no poles and zeros on $\mathbb{H} \cup \{\infty\}$. According to part 1) a constant $a \in \mathbb{C}$ exists with

$$a = \frac{f}{g}$$

or

$$f = a \cdot g \in \mathbb{C}(j).$$

□

Lemma 3.23 (Exact sequence for cusp forms). *Denote by*

$$\varepsilon : M_k(\Gamma_1) \rightarrow \mathbb{C}, f \mapsto f(\infty),$$

the evaluation with kernel $S_k(\Gamma_1)$ and by

$$\mu_\Delta : M_{k-12}(\Gamma_1) \rightarrow M_k(\Gamma_1), f \mapsto f \cdot \Delta,$$

the multiplication with the discriminant form. Then for even $k \geq 4$: The following sequence of vector space morphisms is exact

$$0 \rightarrow M_{k-12}(\Gamma_1) \xrightarrow{\mu_\Delta} M_k(\Gamma_1) \xrightarrow{\varepsilon} \mathbb{C} \rightarrow 0.$$

In particular:

- For even $k \geq 4$:

$$M_k(\Gamma_1) = S_k(\Gamma_1) \oplus \mathbb{C} \cdot G_k,$$

and the multiplication by the discriminant modular form Δ defines a vector space isomorphism

$$\mu_\Delta : M_{k-12}(\Gamma_1) \xrightarrow{\sim} S_k(\Gamma_1).$$

- For $k < 0$:

$$\dim M_k = 0$$

As a consequence, the formulas from Corollary 3.19 allow to compute the dimension of $S_k(\Gamma_1)$. A partial generalization of Lemma 3.23 will be proved in Corollary 3.28.

Proof. i) Exactness on the right-hand side: For even $k \geq 4$ the Eisenstein series $G_k \in M_k(\Gamma_1)$ satisfies

$$G_k(\infty) = 2 \cdot \zeta(k) \neq 0.$$

Therefore ε is surjective.

ii) Exactness in the middle: If $f \in M_k(\Gamma_1)$ with $\varepsilon(f) = 0$, i.e.

$$f \in S_k(\Gamma_1)$$

then the quotient

$$g := \frac{f}{\Delta}$$

is holomorphic on $\mathbb{H}$, because Δ has no zero on $\mathbb{H}$ due to Proposition 1.19. In addition g is holomorphic at the point ∞ because f has a zero at ∞ and Δ has a zero of order 1 at ∞ according to Proposition 3.14. Hence $g \in M_{k-12}(\Gamma_1)$. In addition, the composition

$$\varepsilon \circ \mu_\Delta = 0$$

is obvious.

iii) Exactness on the left-hand side: Apparently the map μ_Δ is injective.

iv) As a consequence

$$\dim S_k(\Gamma_1) = 0 \text{ for } k \in \{0, 2, 4, 6, 8, 10\}.$$

Then iteratively for all even $k < 0$

$$\dim M_k(\Gamma_1) = 0$$

because for all even $k \in \mathbb{Z}$

$$M_{k-12}(\Gamma_1) \xrightarrow{\mu_\Delta} S_k(\Gamma_1)$$

is injective.

□

The result $\dim M_k(\Gamma_1) = 0$ if $k < 0$ generalizes to Hecke congruence subgroups:
For $k < 0$

$$\dim M_k(\Gamma_0(N)) = 0,$$

see [?, Prop. 3].

Theorem 3.24 (Algebra of modular forms). *The associative algebra*

$$M_*(\Gamma_1) := \left(\bigoplus_{k \geq 0} M_{2k}(\Gamma_1), +, \cdot \right),$$

named the algebra of modular forms of the modular group Γ_1 , is commutative and freely generated by the two Eisenstein series

$$G_4 \in M_4(\Gamma_1) \text{ and } G_6 \in M_6(\Gamma_1).$$

The subalgebra of cusp forms

$$S_*(\Gamma_1) := \bigoplus_{k \geq 0} S_{2k}(\Gamma_1) \subset M_*(\Gamma_1)$$

is a principal ideal generated by $\Delta \in S_{12}(\Gamma_1)$.

Proof. One checks for $k \neq k'$

$$M_{2k}(\Gamma_1) \cap M_{2k'}(\Gamma_1) = \{0\}.$$

i) *Generated by G_4 and G_6 :* We prove the claim about $M_k(\Gamma_1)$ by reducing successively the weight. The even integer $k \geq 4$ decomposes as

$$k = 4 \cdot \alpha + 6 \cdot \beta$$

with non-negative integers $\alpha, \beta \in \mathbb{Z}$. Hence

$$G_4^\alpha \cdot G_6^\beta \in M_k \text{ and } \varepsilon(G_4^\alpha \cdot G_6^\beta) \neq 0$$

Consider an arbitrary element $f \in M_k$. For suitable $\lambda \in \mathbb{C}$

$$\varepsilon(f - \lambda \cdot G_4^\alpha \cdot G_6^\beta) = (f - \lambda \cdot G_4^\alpha \cdot G_6^\beta)(\infty) = 0.$$

According to Lemma 3.23 an element $g \in M_{k-12}$ exists with

$$f - \lambda \cdot G_4^\alpha \cdot G_6^\beta = \Delta \cdot g.$$

The claim follows by repeating the argument with $g \in M_{k-12}(\Gamma_1)$, and using

$$M_4(\Gamma_1) = \mathbb{C} \cdot G_4 \text{ and } M_6(\Gamma_1) = \mathbb{C} \cdot G_6$$

due to Corollary 3.19.

ii) *Freely generated*: The generators G_4 and G_6 are even algebraically independent. Assume the existence of a polynomial $P \in \mathbb{C}[T_1, T_2]$ with

$$P(G_4, G_6) = 0.$$

We may assume that all monomials of $P(G_4, G_6)$ have the same weight. The case

$$P(G_4, G_6) = a \cdot G_4^m + G_6 \cdot \hat{P}(G_4, G_6), \quad a \in \mathbb{C}^*,$$

is excluded because $G_4(i) \neq 0$, but $G_6(i) = 0$. Analogously the case

$$P(G_4, G_6) = b \cdot G_6^m + G_4 \cdot \hat{P}(G_4, G_6), \quad b \in \mathbb{C}^*,$$

is excluded because $G_6(\rho) \neq 0$, but $G_4(\rho) = 0$. Hence the product $G_4 G_6$ divides $P(G_4, G_6)$, i.e.

$$P(G_4, G_6) = G_4 G_6 \cdot \hat{P}(G_4, G_6) = 0.$$

Iteration of the argument for

$$\hat{P}(G_4, G_6) = 0 \text{ with } \deg \hat{P} < \deg P$$

proves the claim $P = 0$ by descending induction.

iii) *Ideal of cusp forms*: According to Lemma 3.23 any cusp form is a multiple of Δ .
□

3.3 Generalization to Hecke congruence subgroups $\Gamma_0(N)$

In an analogous way to Theorem 3.17, also modular forms of a Hecke congruence subgroup $\Gamma_0(N)$, $N \in \mathbb{N}$, are meromorphic differential forms on the corresponding modular curve $X_0(N)$. But one has to take into account that a congruence subgroup has more than one cusp and that cusps may have distinct width. Therefore the width of the cusps enters as an additional parameter when determining local coordinates around the cusps of the modular curve. In addition, one cannot use a standard atlas like on $X = \mathbb{P}^1$ but one has to deal with atlases in a generic way.

Recall from Remark 2.33 the branched covering

$$f : X_0(N) \rightarrow X(\Gamma_1) \simeq \mathbb{P}^1$$

of degree

$$\deg f = N \cdot \prod_{p|N} \left(1 + \frac{1}{p}\right)$$

and the genus formula

$$g(X_0(N)) = 1 + \frac{\deg f}{12} - \frac{\varepsilon_2}{4} - \frac{\varepsilon_3}{3} - \frac{\varepsilon_\infty}{2}$$

Moreover, we recall the canonical projection from the extended half-plane

$$p : \mathbb{H}^* \rightarrow X_0(N) := \Gamma_0(N) \backslash \mathbb{H}^*$$

The aid to represent modular forms of $M_k(\Gamma_0(N))$, which are defined as functions on $\mathbb{H}^*$, as meromorphic differential forms on the modular curve $X_0(N)$ are the divisors

$$D_k := \left\lfloor \frac{k}{4} \right\rfloor \cdot \sum_{\substack{x \in X_0(N) \\ x \text{ elliptic} \\ h_x=2}} x + \left\lfloor \frac{k}{3} \right\rfloor \cdot \sum_{\substack{x \in X_0(N) \\ x \text{ elliptic} \\ h_x=3}} x + \frac{k}{2} \cdot \sum_{\substack{x \in X_0(N) \\ x \text{ cusp}}} x \in \text{Div}(X_0(N)), \text{ even } k \geq 0.$$

They have degree

$$\deg D_k = \left\lfloor \frac{k}{4} \right\rfloor \cdot \varepsilon_2 + \left\lfloor \frac{k}{3} \right\rfloor \cdot \varepsilon_3 + \frac{k}{2} \cdot \varepsilon_\infty$$

For each fixed even $k \geq 0$ the divisor D_k singles out those meromorphic differential forms from

$$H^0(X_0(N), \mathcal{M} \otimes_{\mathcal{O}} \Omega^{\otimes k/2})$$

which arise from modular forms under the canonical injection

$$\alpha : M_k(\Gamma_0(N)) \rightarrow H^0(X_0(N), \mathcal{M} \otimes_{\mathcal{O}} \Omega^{\otimes k/2}), \quad g \mapsto \omega_g,$$

with ω_g the well-defined differential form with pullback

$$p^* \omega_g(\tau) = g(\tau) (d\tau)^{k/2}.$$

Lemma 3.25 (Order relation). *Consider a Hecke congruence subgroup $\Gamma_0(N)$ and an even weight $k \geq 0$.*

1. Definition of α : The map

$$\alpha : M_k(\Gamma_0(N)) \rightarrow H^0(X_0(N), \mathcal{M} \otimes_{\mathcal{O}} \Omega^{\otimes k/2}), \quad g \mapsto \omega_g,$$

with ω_g the well-defined differential form satisfying

$$p^* \omega_g(\tau) = g(\tau) (d\tau)^{k/2}$$

is well-defined.

2. Order relation: *The order of a modular form g and the order of the meromorphic differential form*

$$\omega_g = \alpha(g)$$

relate as

$$\text{ord}_\tau g = \begin{cases} \text{ord}_{[\tau]} \omega_g + k/2 & [\tau] \text{ cusp} \\ h_\tau \cdot \text{ord}_{[\tau]} \omega_g + k/2 \cdot (h_\tau - 1) & \tau \in \mathbb{H} \end{cases}$$

Proof. Each modular form $g \in M_k(\Gamma_0(N))$ is a holomorphic function on $\mathbb{H}$. The function g defines on $\mathbb{H}$ the $\Gamma_0(N)$ -invariant holomorphic differential form

$$g(\tau) (d\tau)^{k/2}$$

Therefore $g(\tau) (d\tau)^{k/2}$ defines on the open Riemann surface

$$Y_0(N) := p(\mathbb{H})$$

a meromorphic differential form

$$\omega_g \in H^0(Y_0(N), \mathcal{M} \otimes \Omega^{\otimes k/2})$$

with singularities at most the finitely many orbits of elliptic points $\tau \in \mathbb{H}$.

Moreover $\Gamma_0(N)$ has only finitely many cusps, see Example 2.12. For arbitrary $q \in \mathbb{Q}$ and

$$\beta \in \Gamma_1 \text{ with } q = \beta(\infty)$$

each modular form $g \in M_k(\Gamma_0(N))$ has a well-defined order

$$\text{ord}_q g := \text{ord}_\infty g[\beta]_k$$

independent from the choice of β . Hence ω_g extends as meromorphic differential form to the compactified modular curve

$$\omega_g \in H^0(X_0(N), \mathcal{M} \otimes \Omega_{X_0(N)}^{\otimes k/2})$$

A cumbersome calculation by using charts, see [19, Sect. 3.3.], shows that the order relations from the proof of Theorem 3.17 carry over: These formula show that the map α is well-defined. The map α is injective, and via pullback of differential forms also surjective.

□

On a compact Riemann surface line bundles and divisors correspond to each other bijectively, see [69]. Theorem 3.26 introduces the line bundles of meromorphic differential forms induced by the distinguished divisors D_k on the modular curves.

Theorem 3.26 (The line bundle of modular forms). *Consider a Hecke congruence subgroup*

$$\Gamma_0(N), N \in \mathbb{N},$$

and an even weight $k \geq 0$.

On the modular curve $X_0(N)$ with canonical divisor $K \in \text{Div}(X_0(N))$ the line bundle

$$\mathcal{L}_{N,k} := \mathcal{O}_{(k/2) \cdot K + D_k} \in \text{Pic}(X_0(N))$$

corresponding to the divisor

$$D_k := \left\lfloor \frac{k}{4} \right\rfloor \cdot \sum_{\substack{x \in X_0(N) \\ x \text{ elliptic} \\ h_x=2}} x + \left\lfloor \frac{k}{3} \right\rfloor \cdot \sum_{\substack{x \in X_0(N) \\ x \text{ elliptic} \\ h_x=3}} x + \frac{k}{2} \cdot \sum_{\substack{x \in X_0(N) \\ x \text{ cusp}}} x \in \text{Div}(X_0(N))$$

of degree

$$\deg D_k = \left\lfloor \frac{k}{4} \right\rfloor \cdot \varepsilon_2 + \left\lfloor \frac{k}{3} \right\rfloor \cdot \varepsilon_3 + \frac{k}{2} \cdot \varepsilon_\infty \geq 0$$

satisfies:

- Meromorphic differential forms:

$$\mathcal{L}_{N,k} \simeq \Omega_{N,k}$$

with the subsheaf of meromorphic differential forms on $X_0(N)$

$$\Omega_{N,k} \subset \mathcal{M} \otimes_{\mathcal{O}} \Omega^{\otimes k/2}$$

defined by

$$\Omega_{N,k}(U) := \{\omega \in H^0(U, \mathcal{M} \otimes_{\mathcal{O}} \Omega^{\otimes k/2}) : \omega \neq 0 \text{ and } \text{div } \omega \geq -D_k|_U \cup \{0\}\}$$

for each open $U \subset X_0(N)$

- Global sections: *Consider the canonical map*

$$\alpha : M_k(\Gamma_0(N)) \rightarrow H^0(X_0(N), \mathcal{M} \otimes_{\mathcal{O}} \Omega^{\otimes k/2}), g \mapsto \omega_g,$$

with ω_g the unique differential form with

$$p^* \omega_g = g(\tau) (d\tau)^{k/2}.$$

Here $p : \mathbb{H}^* \rightarrow X_0(N)$ denotes the canonical projection. Then

$$\begin{aligned} H^0(X_0(N), \mathcal{L}_{N,k}) &\simeq H^0(X_0(N), \Omega_{N,k}) = \\ &= \text{im} \left[\alpha : M_k(\Gamma_0(N)) \rightarrow H^0(X_0(N), \mathcal{M} \otimes_{\mathcal{O}} \Omega^{\otimes k/2}) \right] \end{aligned}$$

- Cohomology: For $k \geq 2$

$$H^1(X_0(N), \mathcal{L}_{N,k}) = 0$$

- Dimension of the vector space of sections: For $k \geq 2$

$$\dim H^0(X_0(N), \mathcal{L}_{N,k}) = (g(X_0(N)) - 1) \cdot (k - 1) + \deg D_k,$$

and for $k = 0$

$$\mathcal{L}_{N,0} \simeq \mathcal{O}, \text{ hence } \dim H^0(X_0(N), \mathcal{L}_{N,0}) = 1.$$

Proof. We set

$$g := g(X_0(N))$$

- i) Estimating the degree of $(k/2) \cdot K + D_k$: By definition

$$\mathcal{L}_{N,k} = \mathcal{O}_{(k/2) \cdot K + D_k}.$$

Using the estimates

$$\lfloor k/4 \rfloor \geq (k/4) - 1/2 \text{ and } \lfloor k/3 \rfloor \geq (k/3) - 2/3$$

and the genus formula for $X_0(N)$ we obtain for the degree of the corresponding divisor for $k \geq 2$

$$\begin{aligned} \deg((k/2) \cdot K + D_k) &\geq \left(\frac{k}{2} \cdot (2g - 2) \right) + \left(\frac{k-2}{4} \cdot \varepsilon_2 + \frac{k-2}{3} \cdot \varepsilon_3 + \frac{k}{2} \cdot \varepsilon_{\infty} \right) = \\ &= (2g - 2) + \frac{k-2}{2} \cdot \left(2g - 2 + \frac{\varepsilon_2}{2} + \frac{2}{3} \cdot \varepsilon_3 + \varepsilon_{\infty} \right) + \varepsilon_{\infty} = \\ &= (2g - 2) + \frac{k-2}{2} \cdot \left(\frac{\deg f}{6} \right) + \varepsilon_{\infty} > 2g - 2 \end{aligned}$$

- ii) Vanishing of $\dim H^1(X_0(N), \mathcal{L}_{N,k})$: We apply Serre's duality theorem in the form

$$H^1(X_0(N), \mathcal{L}_{N,k}) = H^0(X_0(N), \mathcal{L}_{N,k}^{\vee} \otimes_{\mathcal{O}} \Omega)^{\vee}$$

and estimate for even $k \geq 2$ due to part i)

$$\deg(-(k/2) \cdot K + D_k) + K < (2 - 2g) + (2g - 2) = 0$$

Hence

$$\dim H^1(X_0(N), \mathcal{L}_{N,k}) = \dim H^0(X_0(N), \mathcal{L}_{N,k}^\vee \otimes \mathcal{O} \Omega) = 0$$

iii) *Riemann-Roch theorem for $\mathcal{L}_{N,k}$* : The Riemann-Roch theorem proves the dimension formula

$$\begin{aligned} \dim H^0(X_0(N), \mathcal{L}_{N,k}) &= 1 - g + \deg((k/2) \cdot K + D_k) = \\ &= 1 - g + \deg((k/2) \cdot K) + \deg D_k = (k-1)(g-1) + \deg D_k. \end{aligned}$$

□

Remark 3.27 (Modular forms of Hecke congruence subgroups). Figure 3.1 shows different modular curves as coverings of $\mathbb{P}^1$ and the dimension of their vector spaces of modular forms, see PARI file "Congruence_subgroup_19". The last column shows the numerical value

$$\dim H^0(X(\Gamma_0(N)), \mathcal{L}_{N,k})$$

computed by applying the Riemann-Roch theorem to the line bundle $\mathcal{L}_{N,k}$.

```
Congruence subgroups_19, Start
Modular forms of congruence subgroups Gamma_0(p**n)
Parameter: p_min = 2, p_max = 7, n_max = 4, weight = 4
h0_RR = dimension via Riemann-Roch theorem
(p,n) = (2,0), q = p^n = 1, X(Gamma_0(1)): (deg, genus, dim M_4, h0_RR) = (1,0,1,1)
(p,n) = (2,1), q = p^n = 2, X(Gamma_0(2)): (deg, genus, dim M_4, h0_RR) = (3,0,2,2)
(p,n) = (2,2), q = p^n = 4, X(Gamma_0(4)): (deg, genus, dim M_4, h0_RR) = (6,0,3,3)
(p,n) = (2,3), q = p^n = 8, X(Gamma_0(8)): (deg, genus, dim M_4, h0_RR) = (12,0,5,5)
(p,n) = (2,4), q = p^n = 16, X(Gamma_0(16)): (deg, genus, dim M_4, h0_RR) = (24,0,9,9)
(p,n) = (3,0), q = p^n = 1, X(Gamma_0(1)): (deg, genus, dim M_4, h0_RR) = (1,0,1,1)
(p,n) = (3,1), q = p^n = 3, X(Gamma_0(3)): (deg, genus, dim M_4, h0_RR) = (4,0,2,2)
(p,n) = (3,2), q = p^n = 9, X(Gamma_0(9)): (deg, genus, dim M_4, h0_RR) = (12,0,5,5)
(p,n) = (3,3), q = p^n = 27, X(Gamma_0(27)): (deg, genus, dim M_4, h0_RR) = (36,1,12,12)
(p,n) = (3,4), q = p^n = 81, X(Gamma_0(81)): (deg, genus, dim M_4, h0_RR) = (108,4,33,33)
(p,n) = (5,0), q = p^n = 1, X(Gamma_0(1)): (deg, genus, dim M_4, h0_RR) = (1,0,1,1)
(p,n) = (5,1), q = p^n = 5, X(Gamma_0(5)): (deg, genus, dim M_4, h0_RR) = (6,0,3,3)
(p,n) = (5,2), q = p^n = 25, X(Gamma_0(25)): (deg, genus, dim M_4, h0_RR) = (30,0,11,11)
(p,n) = (5,3), q = p^n = 125, X(Gamma_0(125)): (deg, genus, dim M_4, h0_RR) = (150,8,43,43)
(p,n) = (5,4), q = p^n = 625, X(Gamma_0(625)): (deg, genus, dim M_4, h0_RR) = (750,48,203,203)
(p,n) = (7,0), q = p^n = 1, X(Gamma_0(1)): (deg, genus, dim M_4, h0_RR) = (1,0,1,1)
(p,n) = (7,1), q = p^n = 7, X(Gamma_0(7)): (deg, genus, dim M_4, h0_RR) = (8,0,3,3)
(p,n) = (7,2), q = p^n = 49, X(Gamma_0(49)): (deg, genus, dim M_4, h0_RR) = (56,1,18,18)
(p,n) = (7,3), q = p^n = 343, X(Gamma_0(343)): (deg, genus, dim M_4, h0_RR) = (392,26,105,105)
(p,n) = (7,4), q = p^n = 2401, X(Gamma_0(2401)): (deg, genus, dim M_4, h0_RR) = (2744,201,714,714)
```

Fig. 3.1 Dimension formulas for modular forms of weight $k = 4$ of $\Gamma_0(p^n)$

Apparently

$$M_0(\Gamma_0(N)) \simeq \mathbb{C}, \quad N \in \mathbb{N},$$

which implies

$$\dim M_0(\Gamma_0(N)) = 1 \text{ and } \dim S_0(\Gamma_0(N)) = 0.$$

Corollary 3.28 (Codimension of cusp forms of Hecke congruence subgroups).

Consider a congruence subgroup

$$\Gamma_0(N), \quad N \in \mathbb{N}.$$

1. *For even weight $k \geq 4$ the subspace of cusp forms*

$$S_k(\Gamma_0(N)) \subset M_k(\Gamma_0(N))$$

has codimension $= \varepsilon_\infty$, the number of cusps of $X_0(N)$. In particular, the codimension does not depend on the weight k .

2. *For weight $k = 2$ cusp forms are exactly the holomorphic differential forms on the modular curve. Hence*

$$\dim S_2(\Gamma_0(N)) = g(X_0(N)).$$

For the computation of ε_∞ in the general case see Example 2.12.

Proof. Set

$$g := g(X_0(N))$$

and consider the divisor $D_k \in \text{Div}(X_0(N))$ from Theorem 3.26, and define the divisor

$$D_k^\infty := \sum_{\substack{x \in X_0(N) \\ x \text{ cusp}}} (-1) \cdot x + D_k \in \text{Div}(X_0(N))$$

and the corresponding line bundle

$$\mathcal{L}_{N,k}^\infty = \mathcal{O}_{(k/2) \cdot K + D_k^\infty} \in \text{Pic}(X_0(N))$$

with

$$\deg \mathcal{L}_{N,k}^\infty = -\varepsilon_\infty + \deg \mathcal{L}_{N,k}.$$

The order relations imply for even $k \geq 2$

$$\begin{aligned} H^0(X_0(N), \mathcal{L}_{N,k}^\infty) &= \\ &= \{ \omega_h \in H^0(X_0(N), \Omega_{N,k}) : \text{ord}_x \omega_h \geq (k/2) - 1 \text{ for all cusps } x \in X_0(N) \} = \\ &= \{ h \in M_k(\Gamma_0(N)) : \text{ord}_t h \geq 1 \text{ for all } t \in \mathbb{H}^* \text{ with } [t] \text{ cusp} \} = S_k(\Gamma_0(N)) \end{aligned}$$

1. $k \geq 4$: The formula from the proof of Theorem 3.26 shows for even $k \geq 4$

$$\deg \mathcal{L}_{N,k}^\infty \geq 2g - 2 + \frac{k-2}{2} \cdot \frac{\deg f}{6} > 2g - 2$$

because $\deg f \geq 1$. Hence $H^1(X_0(N), \mathcal{L}_{N,k}^\infty) = 0$ because

$$\deg (\mathcal{L}_{N,k}^\infty)^\vee \otimes \mathcal{O} \Omega^1 < -(2g-2) + (2g-2) = 0,$$

and the Riemann-Roch theorem implies

$$\begin{aligned} \dim S_k(\Gamma_0(N)) &= \dim H^0(X_0(N), \mathcal{L}_{N,k}^\infty) = \\ &= \dim H^0(X_0(N), \mathcal{L}_{N,k}) - \varepsilon_\infty = \dim M_k(\Gamma_0(N)) - \varepsilon_\infty. \end{aligned}$$

2. $k = 2$: The explicit form of D_2^∞ shows

$$D_2^\infty = 0 \in \text{Div}(X_0(N)),$$

which implies

$$\mathcal{L}_{N,2}^\infty = \mathcal{O}_K = \Omega^1$$

and

$$S_2(X_0(N)) = H^0(X_0(N), \Omega)$$

□

Chapter 4

Elliptic curves

The present chapter ties quite closely with Section 2.1 on tori. We now bridge complex analysis and algebraic geometry.

Section 4.1 shows that each complex torus is an algebraic variety in the complex projective space $\mathbb{P}^2$. The variety is the zero set of a cubic polynomial.

Section 4.2 introduces elliptic curves as projective-algebraic varieties of genus = 1 defined over arbitrary fields, in particular over the field $\mathbb{C}$, its subfield $\mathbb{Q}$, and the finite fields $\mathbb{F}_p$ of prime characteristic. We show that the points of elliptic curves correspond bijectively to the smooth zero set of cubic polynomials, the Weierstrass polynomials. They build a bridge between complex analysis and algebraic geometry.

Section 4.3 introduces elliptic curves over the finite field $\mathbb{F}_p$ and their ζ -function.

4.1 Embedding complex tori as plane cubic hypersurfaces in $\mathbb{P}^2$

Generalizing the 1-dimensional projective space $\mathbb{P}^1$ we introduce the higher-dimensional complex projective spaces $\mathbb{P}^n$.

Definition 4.1 (*n -dimensional complex projective space*). For $n \in \mathbb{N}$ the n -dimensional complex *projective space* is the quotient

$$\mathbb{P}^n := (\mathbb{C}^{n+1} \setminus \{0\}) / \sim$$

with respect to the equivalence relation

$$z = (z_0, \dots, z_n) \sim w = (w_0, \dots, w_n) : \Leftrightarrow \exists \lambda \in \mathbb{C}^* : w = \lambda \cdot z$$

The canonical projection onto equivalence classes

$$\pi : \mathbb{C}^{n+1} \setminus \{0\} \rightarrow \mathbb{P}^n, z \mapsto [z].$$

provides $\mathbb{P}^n$ with the quotient topology, named the Euclidean topology of $\mathbb{P}^n$.

For $z = (z_0, \dots, z_n) \in \mathbb{C}^{n+1} \setminus \{0\}$ the expression

$$(z_0 : \dots : z_n) := \pi(z) \in \mathbb{P}^n$$

is named the *homogeneous coordinate* of $\pi(z)$.

One considers $\mathbb{C}^n \subset \mathbb{P}^n$ via the injection

$$\mathbb{C}^n \hookrightarrow \mathbb{P}^n, (z_1, \dots, z_n) \mapsto (1 : z_1 : \dots : z_n).$$

as an open subspace.

Definition 4.2 (Standard atlas of $\mathbb{P}^n$). The *standard atlas* of $\mathbb{P}^n$ is the family

$$\mathcal{U} := (U_i)_{i=0, \dots, n}$$

with the open sets

$$U_i := \{(z_0 : \dots : z_n) \in \mathbb{P}^n : z_i \neq 0\}, i = 0, \dots, n.$$

and for each $i = 0, \dots, n$ as i -th *standard chart* the homeomorphism

$$\phi_i : U_i \rightarrow \mathbb{C}^n, (z_0 : \dots : z_n) \mapsto \left(\frac{z_0}{z_i}, \dots, \frac{z_{i-1}}{z_i}, \hat{1}, \frac{z_{i+1}}{z_i}, \dots, \frac{z_n}{z_i} \right)$$

which is a holomorphic map. Here the “hat” indicates omission of the term. The transformation between two charts ϕ_j and ϕ_i with $i \neq j$ is the map

$$\phi_i \circ \phi_j^{-1} : \phi_j(U_i \cap U_j) \rightarrow \phi_i(U_i \cap U_j),$$

$$\left(\frac{z_0}{z_j}, \dots, \frac{z_{j-1}}{z_j}, \hat{1}, \frac{z_{j+1}}{z_j}, \dots, \frac{z_n}{z_j} \right) \mapsto \left(\frac{z_0}{z_i}, \dots, \frac{z_{i-1}}{z_i}, \hat{1}, \frac{z_{i+1}}{z_i}, \dots, \frac{z_n}{z_i} \right).$$

The standard atlas induces a complex structure on $\mathbb{P}^n$. The latter becomes a complex manifold, generalizing the 1-dimensional projective line $\mathbb{P}^1$ and the 2-dimensional complex projective plane $\mathbb{P}^2$.

Theorem 4.3 shows once more the importance of the Weierstrass $\wp$ -function and its differential equation for $\wp'$ in the theory of complex tori.

Theorem 4.3 (The embedding theorem for tori). *Consider a complex torus*

$$T := \mathbb{C}/\Lambda$$

with lattice

$$\Lambda = \mathbb{Z}\omega_1 + \mathbb{Z}\omega_2.$$

Set

$$g_2 := 60 \cdot G_{\Lambda,4}, \quad g_3 := 140 \cdot G_{\Lambda,6},$$

and denote by $\wp = \wp_\Lambda$ the Weierstrass function of Λ . The map

$$\phi : T \rightarrow \mathbb{P}^2, \quad p \mapsto \begin{cases} (1 : \wp(p) : \wp'(p)) & \text{if } p \neq 0 \\ (0 : 0 : 1) & \text{if } p = 0 \end{cases}$$

maps the torus T biholomorphically onto $\phi(T) = E$, which is the cubic hypersurface

$$E := \{(x_0 : x_1 : x_2) \in \mathbb{P}^2 : F_{\text{hom}}(x_0, x_1, x_2) = 0\}$$

defined by the homogeneous cubic polynomial

$$F_{\text{hom}}(X_0, X_1, X_2) = X_2^2 X_0 - (4X_1^3 - g_2 \cdot X_1 X_0^2 - g_3 \cdot X_0^3) \in \mathbb{C}[X_0, X_1, X_2].$$

The homogeneous polynomial $F_{\text{hom}}(X_0, X_1, X_2) \in \mathbb{C}[X_0, X_1, X_2]$ is the homogenization of the polynomial

$$F(X, Y) = Y^2 - (4X^3 - g_2 X - g_3) \in \mathbb{C}[X, Y].$$

Proof. The subsequent proof of Theorem 4.3 relies heavily on the properties of $\wp$ and $\wp'$ from Section 1.2.

i) *Equation of $\phi(T)$:* The differential equation of the Weierstrass function $\wp$

$$(\wp')^2 = \wp^3 - g_2 \cdot \wp - g_3$$

implies

$$\phi(T \setminus \{0\}) \subset E_{\text{aff}} := \{(x, y) \in \mathbb{C}^2 : y^2 = 4 \cdot x^3 - g_2 \cdot x - g_3\}$$

The closure of the affine algebraic set

$$E_{\text{aff}} \subset \mathbb{C}^2 \subset \mathbb{P}^2$$

is the projective algebraic variety $E \subset \mathbb{P}^2$, see Figure 4.1 with abscissa X and ordinate Y , which is the zero set in $\mathbb{P}^2$ of

$$F_{\text{hom}}(X_0, X_1, X_2),$$

cf. [31, Chap. I, Exerc. 2.9].

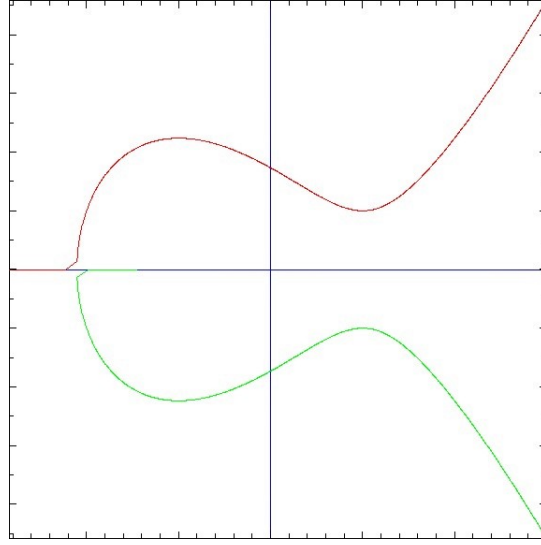


Fig. 4.1 Real points of E_{aff}

ii) *Holomorphy of ϕ* : To check holomorphy in a neighbourhood of $0 \in T$ we choose an open neighbourhood $U \subset T$ of the origin such that $\wp'$ has no zeros in U . On

$$U^* := U \setminus 0$$

one obtains by dividing the homogeneous coordinate by $\wp'$

$$\phi|_{U^*} = (1 : \wp : \wp') = (1/\wp' : \wp/\wp' : 1),$$

which converges with limit

$$\lim_{z \rightarrow 0} \phi(z) = (0 : 0 : 1) := O \in \mathbb{P}^2$$

because $\wp$ and $\wp'$ have poles at 0 of order respectively = 2 and 3. Hence $0 \in T$ is a removable singularity of ϕ .

iii) *Injectivity of ϕ* : The holomorphic map

$$\wp : T \rightarrow \mathbb{P}^1$$

is a branched covering of degree = 2, while the holomorphic map

$$\wp' : T \rightarrow \mathbb{P}^1$$

is a branched covering of degree = 3. To give an indirect proof of the injectivity of ϕ assume two distinct points

$$z \neq w \in T \setminus \{0\}$$

with

$$(\wp(z), \wp'(z)) = (\wp(w), \wp'(w)) \in \mathbb{C}^2.$$

Then

$$\wp(z) = \wp(w) \text{ implies } w = -z$$

due to Corollary 1.10 and because $\wp$ is an even function due to Theorem 1.12.

Moreover

$$\wp'(z) = \wp'(w) = \wp'(-z) = -\wp'(z) \text{ implies } \wp'(z) = 0.$$

Lemma 1.13 implies that

$$z = [\omega/2], \omega \in \Lambda \text{ and } w = -z \in T$$

are two distinct classes of half-period points of Λ . But $\wp$ attains at z the value $\wp(z)$ with multiplicity $= 2$, hence $\wp$ assumes the value $\wp(z) = \wp(w)$ with multiplicity ≥ 4 , a contradiction.

iv) *The image of ϕ* : The only non-affine point is

$$E \setminus E_{aff} = O = \phi(0).$$

Consider an affine point

$$(1 : x : y) \in E.$$

The $\wp$ -function is a non-constant holomorphic map

$$\wp : T \rightarrow \mathbb{P}^1,$$

hence closed and open, and in particular surjective: There exists $z \in T$ with

$$\wp(z) = x.$$

The function $\wp$ is even, therefore also

$$\wp(-z) = x.$$

From

$$y^2 = 4x^3 - g_2 \cdot x - g_3 = 4\wp(z)^3 - g_2 \cdot \wp(z) - g_3 = \wp'(z)^2$$

follows:

- Either

$$y = \wp'(z) \text{ and } \phi(z) = (1 : x : y).$$

- Or

$$y = -\wp'(z) = \wp'(-z) \text{ and } \phi(-z) = (1 : x : y)$$

v) *Smoothness of E :*

- A point $(x_0, y_0) \in \mathbb{C}^2$ is a singular point of E_{aff} iff it satisfies the three equations

$$0 = F(x_0, y_0) = y_0^2 - (4x_0^3 - Ax_0 - B), \quad A := g_2, \quad B := g_3,$$

$$\frac{\partial F}{\partial y}(x_0, y_0) = 2y_0 = 0 \quad \text{and} \quad \frac{\partial F}{\partial x}(x_0, y_0) = -12x_0^2 + A = 0$$

Introducing the cubic polynomial

$$f(x) := 4x^3 - Ax - B \in \mathbb{C}[x]$$

the condition is equivalent to

$$f(x_0) = 0 \quad \text{and} \quad f'(x_0) = 0.$$

The latter condition is equivalent to x_0 being a multiple zero of f , i.e. to the vanishing of the discriminant of f

$$\Delta_f = A^3 - 27B^2.$$

Proposition 1.19 implies $\Delta_f \neq 0$.

- To prove the non-singularity of E at the point $O = (0 : 0 : 1) \in \mathbb{P}^2$ we consider the standard coordinate of $\mathbb{P}^2$ around the point O

$$\phi_2 : U_2 \rightarrow \mathbb{C}^2, \quad (z_0 : z_1 : z_2) \mapsto (u, v) := \left(\frac{z_0}{z_2}, \frac{z_1}{z_2} \right).$$

We have

$$\phi_2(E \cap U_2) = \{(u, v) \in \mathbb{C}^2 : f(u, v) = 0\}$$

with

$$f(u, v) := u - (4 \cdot v^3 - A \cdot u^2 \cdot v - B \cdot u^3).$$

Then the partial derivatives are

$$\frac{\partial f}{\partial u}(u, v) = 1 - (2 \cdot A \cdot u \cdot v - 3 \cdot B \cdot u^2) \quad \text{and} \quad \frac{\partial f}{\partial v}(u, v) = -12 \cdot v^2 + A \cdot u^2$$

hence

$$\nabla f(0, 0) = (1, 0) \neq 0.$$

The non-singular projective algebraic curve $E = \phi(T) \subset \mathbb{P}^2$ is connected, hence a compact Riemann surface with the induced analytic structure as a hypersurface of $\mathbb{P}^2$.

vi) *Biholomorphy:* The map

$$\phi : T \rightarrow E$$

1

Example 4.4 (Embedding tori as plane cubics). The PARI script `Torus_to>Weierstrass_equation_01` computes a cubic equation in $\mathbb{P}^2$ for a given torus $T = \mathbb{C}/\Lambda$, see Figure 4.2. Real values are approximated by rationals.

[illegible]

Fig. 4.2 Embedding tori as plane cubics

The embedding result of Theorem 4.3 is a particular case of the general question: How to obtain holomorphic maps

$$X \rightarrow \mathbb{P}^n$$

with X a compact complex manifold X ?

This question has a general answer, which is not restricted to complex analysis. We recall the general construction in the context of compact Riemann surfaces X , see also [69]. The pivotal objects are the twisted line bundle on $\mathbb{P}^n$ and very ample line bundles on X . In the following we shall not distinguish between a line bundle and the invertible sheaf of its holomorphic sections. We shall employ in both cases the same notation

$$\mathcal{L} \in \text{Pic}(X).$$

Definition 4.5 (The twisted line bundle on $\mathbb{P}^n$). The *twisted line bundle* $\mathcal{O}(1)$ on $\mathbb{P}^n$ is defined with respect to the standard atlas of $\mathbb{P}^n$ by the cocycle

$$g = (g_{ij}) \in Z^1(\mathcal{U}, \mathcal{O}^*)$$

with

$$g_{ij}(z_0 : \dots : z_n) := \frac{z_j}{z_i}, \quad 0 \leq i \neq j \leq n.$$

Proposition 4.6 (Sections of $\mathcal{O}(1)$). The sections of $\mathcal{O}(1)$ on $\mathbb{P}^n$ correspond bijectively to the linear polynomials from $H\text{Pol}(1) \subset \mathbb{C}[z_0, \dots, z_n]$ via the map

$$H\text{Pol}(1) \rightarrow H^0(\mathbb{P}^n, \mathcal{O}(1)), \quad z_k \mapsto s_k, \quad k = 0, \dots, n.$$

Here $s_k \in H^0(\mathbb{P}^n, \mathcal{O}(1))$ is defined with respect to the i -th standard chart by the holomorphic function

$$s_{k,i} : U_i \rightarrow \mathbb{C}, \quad s_{k,i}(z_0 : \dots : z_n) := \frac{z_k}{z_i}$$

Proof. The map defined above on the standard charts provides a section

$$s_k := (s_{k,i})_{i=0, \dots, n} \in H^0(\mathbb{P}^n, \mathcal{O}(1))$$

because

$$g_{ij} \cdot s_{k,j} = \frac{z_j}{z_i} \cdot \frac{z_k}{z_j} = \frac{z_k}{z_i} = s_{k,i}.$$

□

Proposition 4.6 generalizes to sections of the tensor powers

$$\mathcal{O}(k) := \mathcal{O}(1)^{\otimes k}, \quad k \in \mathbb{N},$$

which correspond bijectively to the homogeneous polynomials of degree $= k$, cf. also [69].

In order that a line bundle $\mathcal{L}$ on a Riemann surface X defines a holomorphic map

$$\phi_{\mathcal{L}} : X \rightarrow \mathbb{P}^n$$

to a projective space, the line bundle $\mathcal{L}$ has to be *globally generated*.

Definition 4.7 (Base-point, globally generated line bundle). Consider a Riemann surface X and a line bundle $\mathcal{L}$ on X .

1. A point $x \in X$ is a *base-point* of $\mathcal{L}$ if each section $\sigma \in H^0(X, \mathcal{L})$ vanishes at x , i.e. $\sigma(x) = 0$. Note

$$\sigma(x) = 0 \iff \sigma_x \in \mathfrak{m}_x \mathcal{L}_x$$

with $\mathcal{L}_x$ the stalk of $\mathcal{L}$ at $x \in X$ and

$$\mathfrak{m}_x \subset \mathcal{O}_x$$

the maximal ideal of the stalk $\mathcal{O}_x$ of the structure sheaf at $x \in X$.

2. The line bundle is *globally generated* or *base-point-free* if it has no base-points, i.e. if for each $x \in X$ exists a section $\sigma \in H^0(X, \mathcal{L})$ such that the germ $\sigma_x \in \mathcal{L}_x$ generates the stalk $\mathcal{L}_x$ as $\mathcal{O}_x$ -module. The latter condition is equivalent to $\sigma(x) \neq 0$.

Proposition 4.8 (Holomorphic maps to projective spaces). *Consider a compact Riemann surface X and a globally generated line bundle $\mathcal{L}$ on X . After choosing a basis*

$$(\sigma_i)_{i=0, \dots, n} \in H^0(X, \mathcal{L})$$

the map

$$\phi_{\mathcal{L}} : X \rightarrow \mathbb{P}^n, x \mapsto (\sigma_0(x) : \dots : \sigma_n(x)),$$

is well-defined - i.e. independent from the chosen trivialization of $\mathcal{L}$ - and holomorphic. It induces an isomorphism

$$\phi_{\mathcal{L}}^*(\mathcal{O}(1)) \simeq \mathcal{L}$$

with $\phi_{\mathcal{L}}^(\mathcal{O}(1))$ the pull-back of the twisted line bundle.*

Proof. We set

$$\phi := \phi_{\mathcal{L}}.$$

i) *Definition:* We consider the definition of ϕ in a suitable open neighbourhood U of a given point $x \in X$. On U we choose a trivialization to identify the line bundle $\mathcal{L}|_U$ with the structure sheaf $\mathcal{O}|_U$. Hence sections from $H^0(U, \mathcal{L})$ become holomorphic functions. Because $\mathcal{L}$ is globally generated there is an index $j \in \{0, \dots, n\}$ with

$$\sigma_j(x) \neq 0.$$

Hence the point

$$(\sigma_0(x) : \dots : \sigma_n(x)) \in \mathbb{P}^n$$

is well-defined. It is independent of the choice of the chart and of the isomorphism

$$\mathcal{L}|_U \simeq \mathcal{O}|_U.$$

Apparently the map ϕ is holomorphic on U .

ii) *Pullback of the twisted line bundle:* For each $i = 0, \dots, n$ the sets

$$X_i := \phi^{-1}(U_i) = \{x \in X : (\sigma_i)_x \notin \mathfrak{m}_x \mathcal{L}_x\}$$

form an open covering $(X_i)_{i=0, \dots, n}$ of X because $\mathcal{L}$ is globally generated.

For $i = 0, \dots, n$ the canonical morphisms between the rings of local sections

$$\mathcal{O}(1)(U_i) \rightarrow \phi_*(\mathcal{L})(U_i) = \mathcal{L}(X_i) \text{ induced by } s_i|_{U_i} \mapsto \sigma_i|_{X_i}$$

define a canonical sheaf morphism on $\mathbb{P}^n$

$$\mathcal{O}(1) \rightarrow \phi_* \mathcal{L}$$

The latter corresponds to a sheaf morphism on X

$$\phi^*(\mathcal{O}(1)) \rightarrow \mathcal{L},$$

because the two functors ϕ_* and ϕ^* , direct image and inverse image, are adjoint, see [31, Chap. II, § 5], i.e. there is a canonical group isomorphism

$$\text{Hom}_{\mathcal{O}_X}(\phi^*(\mathcal{O}(1)), \mathcal{L}) \simeq \text{Hom}_{\mathcal{O}_{\mathbb{P}^n}}(\mathcal{O}(1), \phi_* \mathcal{L})$$

The induced morphism

$$\phi^*(\mathcal{O}(1)) \rightarrow \mathcal{L}$$

between the two line bundles is an isomorphism because it is an isomorphism on stalks: If for a given $i \in \{0, \dots, n\}$

$$x \in X_i \subset X \text{ and } y := \phi(x) \in U_i \subset \mathbb{P}^n$$

then according to the definition of the analytic inverse image sheaf

$$(\phi^* \mathcal{O}(1))_x = \mathcal{O}(1)_y \otimes_{\mathcal{O}_{\mathbb{P}^n, y}} \mathcal{O}_{X, x} \rightarrow \mathcal{L}_x, (s_i)_y \otimes 1 \mapsto (\sigma_i)_x.$$

□

Theorem 4.9 provides a criterion to ensure that the map provided by a globally generated line bundle $\mathcal{L}$ is a closed embedding.

Theorem 4.9 (Projective embedding induced by a line bundle). *Consider a compact Riemann surface X and a globally generated line bundle $\mathcal{L}$ on X . Then the induced map*

$$\phi_{\mathcal{L}} : X \rightarrow \mathbb{P}^n$$

is a holomorphic closed embedding iff $\mathcal{L}$ satisfies both of the following properties:

1. *Separating points: For any two distinct points $p \neq q \in X$ exists a section*

$$\sigma \in H^0(X, \mathcal{L}) \text{ with } \sigma(p) \neq 0 \text{ but } \sigma(q) = 0$$

2. *Separating tangent vectors: For all $x \in X$ the map*

$$d_{\mathcal{L},x} : \{\sigma \in H^0(X, \mathcal{L}) : \sigma_x \in \mathfrak{m}_x \mathcal{L}_x\} \rightarrow \mathfrak{m}_x \mathcal{L}_x / \mathfrak{m}_x^2 \mathcal{L}_x, \sigma \mapsto [\sigma_x],$$

is surjective.

For the proof of Theorem 4.9 see [69].

Remark 4.10 (Separating points and tangent vectors). We explain the meaning of the two statements from Theorem 4.9.

1. *Separating points:* The homogeneous coordinates of a given point in $\mathbb{P}^n$ do not fix the value of the components. But they determine whether a component is zero or not zero. Therefore for each pair of points $p \neq q \in X$

$$\phi_{\mathcal{L}}(p) \neq \phi_{\mathcal{L}}(q)$$

if at least one component of the left-hand side is different from zero while the same component on the right-hand side is zero.

2. *Separating tangent vectors:* A holomorphic mapping is an embedding in the neighbourhood of a point $x \in X$ if the Jacobi map at x does not vanish. The latter condition is equivalent to the fact that the push-forward of tangent vectors at x is injective, equivalently the pull-back of cotangent vectors is surjective: We consider the isomorphy of $\mathcal{O}_x$ -modules - respectively of complex vector spaces -

$$\mathfrak{m}_x \mathcal{L}_x / \mathfrak{m}_x^2 \mathcal{L}_x \simeq (\mathfrak{m}_x / \mathfrak{m}_x^2) \otimes_{\mathcal{O}_x} \mathcal{L}_x = T_x^1 \otimes_{\mathcal{O}_x} \mathcal{L}_x$$

with the cotangent space

$$T_x^1 := \mathfrak{m}_x / \mathfrak{m}_x^2$$

of X at x . Then the map

$$d_{\mathcal{L},x} : \{\sigma \in H^0(X, \mathcal{L}) : \sigma_x \in \mathfrak{m}_x \mathcal{L}_x\} \rightarrow T_x^1 \otimes_{\mathcal{O}_x} \mathcal{L}_x$$

is induced by the total differential d_x of holomorphic functions: A section

$$\sigma \in H^0(X, \mathcal{L}) \text{ with } \sigma_x \in \mathfrak{m}_x \mathcal{L}_x$$

factorizes in a suitable neighbourhood U of x as

$$\sigma = f \cdot \sigma_1$$

with a holomorphic function $f \in \mathcal{O}(U)$ satisfying $f_x \in \mathfrak{m}_x$ and a holomorphic section $\sigma_1 \in \mathcal{L}(U)$. Then

$$d_{\mathcal{L},x}(\sigma) = d_x f \otimes \sigma_1 \in T_x^1 \otimes_{\mathcal{O}_x} \mathcal{L}_x$$

$$\begin{array}{ccc}
\{s \in H^0(\mathbb{P}^n, \mathcal{O}(1)) : s(y) = 0\} & \xrightarrow{d_{\mathcal{O}(1),y}} & T_{\mathbb{P}^n,y}^1 \otimes_{\mathcal{O}_{\mathbb{P}^n,y}} \mathcal{O}(1)_y \\
\downarrow \phi^* & & \downarrow \phi_{T^1}^* \\
\{\sigma \in H^0(X, \mathcal{L}) : \sigma(x) = 0\} & \xrightarrow{d_{\mathcal{L},x}} & T_{X,x}^1 \otimes_{\mathcal{O}_{X,x}} \mathcal{L}_x
\end{array}$$

Fig. 4.3 Separating tangent vectors

The commutative diagram from Figure 4.3 with

$$y := \phi(x)$$

and surjective maps

$$d_{\mathcal{O}(1),y} \text{ and } \phi^*$$

shows: The pull-back $\phi_{T^1}^*$ is surjective if and only if $d_{\mathcal{L},x}$ is surjective.

Definition 4.11 (Very ample line bundle). A globally generated line bundle $\mathcal{L}$ on a compact Riemann surface X is *very ample* if the induced map

$$\phi_{\mathcal{L}} : X \rightarrow \mathbb{P}^n$$

is an embedding.

A globally generated line bundle $\mathcal{L}$ on X has enough sections to define a holomorphic map

$$\phi_{\mathcal{L}} : X \rightarrow \mathbb{P}^n.$$

If $\mathcal{L}$ is very ample then there are enough sections that $\phi_{\mathcal{L}}$ is even an embedding. Due to the compactness of X its image under an embedding is always closed.

Notation 4.12. For a line bundle $\mathcal{L}$ on a Riemann surface X and a divisor $D \in \text{Div}(X)$ we denote by

$$\mathcal{L}_D := \mathcal{L} \otimes_{\mathcal{O}} \mathcal{O}_D$$

the sheaf of meromorphic sections of $\mathcal{L}$ which are multiples of the divisor $-D$. The sheaf $\mathcal{L}_D$ is isomorphic to a line bundle.

Consider a line bundle $\mathcal{L}$ on a compact Riemann surface X . Theorem 4.13 states a numerical criterion for the dimension of the vector spaces

$$H^0(X, \mathcal{L}_D), D \in \text{Div}(X),$$

which ensures that the map

$$\phi_{\mathcal{L}} : X \rightarrow \mathbb{P}^n$$

is well-defined and a closed embedding. This criterion is very helpful in the applications because the dimension on the vector spaces in question can often be computed by using the theorem of Riemann-Roch in combination with Serre duality.

Theorem 4.13 (Very-ampleness criterion). *Consider a compact Riemann surface X . For a line bundle $\mathcal{L}$ on X are equivalent:*

- *The line bundle $\mathcal{L}$ is very ample.*
- *For any two not necessarily distinct points $p, q \in X$ the corresponding point divisors*

$$P := 1 \cdot p, Q := 1 \cdot q \in \text{Div}(X)$$

satisfy

$$\dim H^0(X, \mathcal{L}_{-(P+Q)}) = \dim H^0(X, \mathcal{L}) - 2.$$

Proof. i) *Assume the validity of the dimension formula:* The formula implies for any two point divisors $P, Q \in \text{Div}(X)$

$$H^0(X, \mathcal{L}_{-(P+Q)}) \subsetneq H^0(X, \mathcal{L}_{-P}) \subsetneq H^0(X, \mathcal{L})$$

and each proper inclusion has codimension = 1 because it is defined by a single linear equation.

- The equation

$$h^0(X, \mathcal{L}_{-P}) = h^0(X, \mathcal{L}) - 1.$$

states that the kernel of the evaluation

$$H^0(X, \mathcal{L}) \rightarrow \mathcal{L}_p / \mathfrak{m}_p \mathcal{L}_p \simeq \mathbb{C}, \sigma \mapsto [\sigma_p],$$

has codimension = 1. Hence p is not a base-point of $\mathcal{L}$. As a consequence, the line bundle $\mathcal{L}$ is globally generated and

$$\phi_{\mathcal{L}} : X \rightarrow \mathbb{P}^n$$

is well-defined.

- The equation

$$H^0(X, \mathcal{L}_{-(P+Q)}) \subsetneq H^0(X, \mathcal{L}_{-P})$$

implies that for any pair of distinct points $p, q \in X$ there exists a section

$$\sigma \in H^0(X, \mathcal{L}_{-P}) \setminus H^0(X, \mathcal{L}_{-(P+Q)})$$

i.e. satisfying

$$\sigma(p) = 0 \text{ but } \sigma(q) \neq 0.$$

Hence the sheaf $\mathcal{L}$ separates points.

- The dimension formula shows for any point $p \in X$ with point divisor $P \in \text{Div}(X)$

$$H^0(X, \mathcal{L}_{-2P}) \subsetneq H^0(X, \mathcal{L}_{-P})$$

has codimension = 1. Hence there exists a section

$$\sigma \in H^0(X, \mathcal{L}_{-P}) \setminus H^0(X, \mathcal{L}_{-2P})$$

which implies the surjectivity of the composition of the canonical maps

$$H^0(X, \mathcal{L}_{-P}) \rightarrow \mathfrak{m}_{X,p} \mathcal{L}_p \rightarrow \mathfrak{m}_{X,p} \mathcal{L}_p / \mathfrak{m}_{X,p}^2 \mathcal{L}_p$$

due to

$$\dim_{\mathbb{C}}(\mathfrak{m}_{X,p} / \mathfrak{m}_{X,p}^2) = 1.$$

Therefore $\mathcal{L}$ separates tangent vectors.

ii) Assume $\mathcal{L}$ very ample: Theorem 4.9 implies that $\mathcal{L}$ separates points and tangent vectors. Separating points implies for all point divisors $P \neq Q \in \text{Div}(X)$

$$\dim H^0(X, \mathcal{L}_{-(P+Q)}) = \dim H^0(X, \mathcal{L}) - 2.$$

Separating tangent vectors implies for each point divisor $P \in \text{Div}(X)$

$$\dim H^0(X, \mathcal{L}_{-2P}) = \dim H^0(X, \mathcal{L}) - 2.$$

□

Remark 4.14 (Projective embedding of compact Riemann surfaces). As a consequence of the very-ampleness criterion from Theorem 4.13 one can prove: Each compact Riemann surface X embeds holomorphically into a projective space.

1. If $g(X) \geq 2$ then the tri-canonical line bundle $\Omega^{\otimes 3} \in \text{Pic}(X)$ is very ample, see [69].

2. For a complex torus $X = \mathbb{C}/\Lambda$ holds $\Omega \simeq \mathcal{O}$, hence

$$g(X) := \dim H^0(X, \Omega) = 1$$

Consider the divisor

$$D := 3 \cdot 0 \in \text{Div}(X)$$

with the point $0 \in X$ being the origin. Then the line bundle

$$\mathcal{L} := \mathcal{O}_D \in \text{Pic}(X)$$

is very ample according to Theorem 4.13: As a consequence of $\Omega \simeq \mathcal{O}$, for two point divisors

$$P = 1 \cdot p, Q = 1 \cdot q \in \text{Div}(X)$$

holds

$$\dim H^1(X, \mathcal{L}_{-(P+Q)}) = \dim H^0(X, \mathcal{O}_{-D+(P+Q)}) = 0$$

because

$$\deg(-D + (P + Q)) = -3 + 2 = -1 < 0$$

Hence

$$\dim H^0(X, \mathcal{L}_{-(P+Q)}) = 1 - g(X) + \deg \mathcal{L} - 2 = \dim H^0(X, \mathcal{L}) - 2$$

The Riemann-Roch theorem implies

$$\dim H^0(X, \mathcal{L}) = \dim H^0(X, \mathcal{O}_D) = \deg D = 3$$

Apparently, a basis of $H^0(X, \mathcal{O}_D)$ is the family of linearly independent elements

$$(1, \wp, \wp'),$$

which gives a second proof for the embedding from Theorem 4.3.

4.2 Elliptic curves over an arbitrary field

Complex analysis of several variables or complex analytic geometry on one hand, and on the other hand algebraic geometry are mathematical theories with many concepts in common and several concepts in parallel. The counterpart of a 1-dimensional torus from complex analysis is an elliptic curve from algebraic geometry. Before defining elliptic curves in Definition 4.20 we recall some basic definition and results from algebraic geometry.

We compile the following GAGA-dictionary (Géométrie Algébrique et Géométrie Analytique).

Proposition 4.15 (Dictionary: Complex analytic geometry - Algebraic geometry).

<i>Complex analytic geometry</i>	<i>Algebraic geometry</i>
<i>Complex Manifold</i>	<i>Smooth variety</i>
<i>Euclidean topology</i>	<i>Zariski topology</i>
<i>Compact complex manifold</i>	<i>Smooth projective algebraic variety</i>
<i>Compact Riemann surface</i>	<i>Projective algebraic curve</i>
<i>Genus</i>	
<i>Stein manifold</i>	<i>Smooth affine variety</i>
<i>Structure sheaf $\mathcal{O}_X$</i>	
<i>Holomorphic function</i>	<i>Regular function</i>
<i>Meromorphic function</i>	<i>Rational function</i>
<i>Local ring $\mathcal{O}_{X,x}$</i>	
<i>Line bundle</i>	
<i>Very ampleness</i>	
<i>Divisor</i>	
<i>Sheaf Ω_X of differential forms</i>	
<i>Cohomology</i>	
<i>Riemann-Roch Theorem</i>	
<i>Serre Duality</i>	

The table relates concepts from complex analytic geometry on the left-hand side with concepts from algebraic geometry on the right-hand side. But not any compact manifold, and not even any compact Kaehler manifold has a counter part in algebraic geometry. Hence not any instance of a concept on the left-hand side of the table induces a corresponding instance on the right-hand side. On both sides there are objects which cannot be studied by methods from the other side.

In algebraic geometry a variety is assumed to be irreducible. *Irreducibility* corresponds to *connectedness* in the case of manifolds. If one skips the requirement of smoothness one obtains the irreducible reduced complex space as the analogue of a variety.

Remark 4.16 recalls varieties as the basic objects of classical algebraic geometry. Varieties are irreducible by definition, hence they are defined by prime ideals. But different from manifolds varieties may have singularities. Hence varieties are the analogue of reduced and irreducible complex spaces. We distinguish between their field k of definition and the algebraic closure $\bar{k}$ where the points of the variety have their coordinates. As a topological space varieties will be equipped with the Zariski topology.

Recall that a field k is *perfect* if each irreducible polynomial with coefficients from k is separable, i.e. has pairwise distinct roots in an algebraic closure $\bar{k}$. All fields in the present section will be assumed perfect. The relevant examples of perfect fields are the fields k of characteristic $= 0$ or the finite fields $k = \mathbb{F}_p$ of prime characteristic $= p$.

Remark 4.16 (Algebraic variety with structure sheaf). Consider a perfect field k and denote by $\bar{k}$ the algebraic closure of k .

1. *Affine variety:* An *affine algebraic variety* X is the zero set

$$X = V(\mathfrak{a}) \subset \mathbb{A}^n(\bar{k}) = \bar{k}^n$$

of a prime ideal $\mathfrak{a} \subset \bar{k}[X_1, \dots, X_n]$, i.e.

$$X := \{x \in \mathbb{A}^n(\bar{k}) : f(x) = 0 \text{ for all } f \in \mathfrak{a}\}.$$

The ring

$$R := \bar{k}[X_1, \dots, X_n] / \mathfrak{a}$$

is the *affine coordinate ring* of X .

If $\mathfrak{a}$ can be generated by polynomials from $k[X_1, \dots, X_n]$, then X is *defined over* k , denoted X/k . The smallest subfield of $\bar{k}$ with this property is named the *field of definition* of X .

2. *A projective algebraic variety* X is the zero set

$$X = V(\mathfrak{a}) \subset \mathbb{P}^n(\bar{k})$$

of a prime ideal

$$\mathfrak{a} \subset \bar{k}[X_0, \dots, X_n],$$

generated by homogeneous polynomials, i.e.

$$X := \{(x = (x_0 : \dots : x_n) \in \mathbb{P}^n(\bar{k}) : f(x) = 0 \text{ for all homogeneous } f \in \mathfrak{a}\}.$$

3. *Zariski topology*: Closed sets of $\mathbb{A}^n(\bar{k})$ are the affine algebraic subvarieties and their finite unions. The *Zariski topology* on an affine algebraic variety $X \subset \mathbb{A}^n(\bar{k})$ is the subspace topology induced from the Zariski topology of $\mathbb{A}^n(\bar{k})$. A basis of the Zariski topology of X are the sets

$$D_h := \{x \in X : h(x) \neq 0\}, \quad h \in R.$$

Closed sets of $\mathbb{P}^n(\bar{k})$ are the projective algebraic subvarieties. The *Zariski topology* on a projective algebraic variety $X \subset \mathbb{P}^n(\bar{k})$ is the subspace topology induced from the Zariski topology of $\mathbb{P}^n(\bar{k})$.

Non-empty open subsets of an irreducible space are irreducible itself, and are also dense.

4. *Structure sheaf and sheaf of rational functions*: Consider an affine algebraic variety X with coordinate ring R and an open set $U \subset X$. A function

$$f : U \rightarrow \bar{k}$$

is *regular* if it has locally the form

$$f = \frac{g}{h} \text{ with } g, h \in R, \quad h \text{ locally without zeros in } X.$$

Note that the quotient representation of f is only required to hold locally. The sheaf $\mathcal{O}_X$ of *regular functions* on an affine algebraic variety X is the contravariant functor

$$U \mapsto \mathcal{O}_X(U) := \{f : U \rightarrow \bar{k} \mid f \text{ regular}\}, \quad U \subset X \text{ open},$$

with the restriction of functions. The sheaf $\mathcal{O}_X$ is named the *structure sheaf* of X . Sections over a basis open set $D_h \subset X$ form the ring

$$\mathcal{O}(D_h) = R_h := \{g/h^n : g \in R, n \in \mathbb{N}\}$$

Because the ideal

$$\mathfrak{a} \subset \bar{k}[X_1, \dots, X_n]$$

is prime, the quotient R and all rings

$$\mathcal{O}_X(U), \quad U \subset X \text{ open},$$

are integral domains. The sheaf $\mathcal{M}_X$ of *rational functions* on X is the sheafification of the presheaf

$$U \mapsto Q(\mathcal{O}_X(U)), \quad U \subset X \text{ open},$$

with $Q(\mathcal{O}_X(U))$ the quotient field of the integral domain $\mathcal{O}_X(U)$. Sections of $\mathcal{M}_X(U)$ are locally quotients

$$g/h, \quad g, h \in R, \quad h \neq 0.$$

For a non-empty open set $U \subset X$ holds

$$\mathcal{M}(U) = Q(R) =: \bar{k}(X),$$

independent from U . The field $Q(R)$ is named the *function field* $\bar{k}(X)$ of X . The sheaf

$$\mathcal{M}^* \subset \mathcal{M}$$

is the multiplicative subsheaf of non-zero rational functions.

The definitions carry over to projective algebraic varieties, because the latter have an atlas of affine algebraic varieties.

5. *Non-singularity*: Consider an algebraic variety X and a point $x \in X$. The local ring $\mathcal{O}_{X,x}$ at $x \in X$ is the ring of germs of regular functions defined in a neighbourhood of $x \in X$. Its dimension defines the dimension of X at $x \in X$.

The variety X is *non-singular* or *smooth* at $x \in X$ if $\mathcal{O}_{X,x}$ is a *regular* local ring, i.e. if

$$\dim \mathcal{O}_{X,x} = \dim_{\kappa(x)} (\mathfrak{m}_x / \mathfrak{m}_x^2) \quad (\text{Dimension equals cotangent dimension}).$$

Here

$$\kappa(x) := \mathcal{O}_{X,x} / \mathfrak{m}_x$$

denotes the residue field of X at x .

6. *Curve over k* : A projective algebraic *curve* C/k is a 1-dimensional, non-singular projective algebraic variety defined over the field k . If

$$C \subset \mathbb{P}^2(\bar{k})$$

then C/k is a *plane projective algebraic curve*. Its *genus* is defined as

$$g(C) = \dim_{\bar{k}} H^0(C, \Omega),$$

the dimension of the $\bar{k}$ -vector space of *regular differential forms*.

7. *Affine schemes*: From a modern viewpoint the primary property of a variety is not the point set

$$X \subset \mathbb{A}^n(\bar{k}),$$

the primary object is the ideal

$$\mathfrak{a} \subset k[X_1, \dots, X_n]$$

with minimal field of definition k , or its quotient, the ring

$$R := k[X_1, \dots, X_n] / \mathfrak{a}$$

Due to Grothendieck's reformulation of algebraic geometry, the basic concept is for arbitrary rings R the *affine scheme*

$$(X := \operatorname{Spec} R, \mathcal{O}_X)$$

Here $\operatorname{Spec} R$ is a topological space with points the prime ideals of R . The structure sheaf $\mathcal{O}_X$ has as stalks the localizations

$$R_{\mathfrak{p}}, \mathfrak{p} \subset R \text{ prime ideal.}$$

We have included the requirement of non-singularity into our definition of a projective algebraic curve because all curves under consideration will be assumed smooth if not stated otherwise. Note that any plane projective algebraic curve is defined by a single homogeneous polynomial. Its degree is the *degree* of C/k .

Remark 4.16 for X/k involves two fields:

- First, the *field of definition* k is not necessarily algebraically closed. It is the smallest field which contains all coefficients of a set of generators of $I(X)$.
- Secondly, the algebraically closed field $\bar{k}$, used to visualize X from a geometrical viewpoint as the maximal set of its points. Also the condition on smoothness refers to $\bar{k}$.

But the set $X \subset \mathbb{A}^n(\bar{k})$ respectively $X \subset \mathbb{P}^n(\bar{k})$ is not the only point set attached to X/k . Definition 4.17 introduces point sets $X(K)$ attached to all intermediate fields $k \subset K \subset \bar{k}$.

Definition 4.17 (K -valued point of X/k). Consider a projective algebraic variety X/k with

$$X \subset \mathbb{P}^n(\bar{k}).$$

Then for any field

$$k \subset K$$

the set of K -valued points of X is defined as

$$X(K) := X \cap \mathbb{P}^n(K) =$$

$$= \{x = (x_0 : \dots : x_n) \in \mathbb{P}^n(\bar{k}) : x \in X \text{ and } x_i \in K \text{ for all } i = 0, \dots, n\}.$$

Remark 4.18 (GAGA).

1. *Chow's theorem:* Historically, the first GAGA-result is Chow's theorem: Any projective complex analytic manifold $X \subset \mathbb{P}^n$ is the variety of homogeneous polynomials, i.e. if $X \subset \mathbb{P}^n$ is locally defined as the zero set of holomorphic functions, then X can be globally defined by homogenous polynomials.
2. *Analytification of schemes:* To formalize the analogy between complex analytic geometry and complex algebraic geometry Grothendieck uses the language of category theory. On one hand, let $\underline{An}_{\mathbb{C}}$ denote the category of complex spaces and holomorphic maps. On the other hand, let $\underline{Al}_{\mathbb{C}}$ denote the category of schemes of finite type defined over $\mathbb{C}$ and regular maps. The first step attaches to a schema $X \in \underline{Al}_{\mathbb{C}}$ a distinguished complex analytic space $X^{an} \in \underline{An}_{\mathbb{C}}$. The latter is obtained via the functor:

$$\Phi : \underline{An}_{\mathbb{C}} \rightarrow \underline{Ens}$$

attaching to a complex analytic space Y the set

$$\Phi(Y) := \text{Hom}_{\mathbb{C}}(Y, X)$$

of morphisms of ringed spaces. The functor Φ can be represented by a universal object (X^{an}, ϕ) , named the *analytification* of X . The pair comprises a complex analytic space X^{an} and a morphism

$$\phi : X^{an} \rightarrow X,$$

see [50]. As a consequence: For a given complex analytic space Y each morphism

$$f : Y \rightarrow X$$

of ringed spaces induces a unique holomorphic map

$$f^{an} : Y \rightarrow X^{an},$$

such that the following diagram commutes

$$\begin{array}{ccc} & & Y \\ & \swarrow f^{an} & \downarrow f \\ X^{an} & \xrightarrow{\phi} & X \end{array}$$

The analytification $\phi : X^{an} \rightarrow X$ defines for each $x \in X^{an}$ a pull-back

$$\phi_x : \mathcal{O}_{X, \phi(x)} \rightarrow \mathcal{O}_{X^{an}, x}$$

which allows to compare the local ring of *regular* functions on X and the local ring of *holomorphic* functions on X^{an} . As a consequence one obtains comparison theorems between the algebraic properties of the local rings of the structure sheaves of X and X^{an} , see [50, Chap. 2].

In addition, any morphism $f : X \rightarrow Y$ in $\underline{An}_{\mathbb{C}}$ induces a unique holomorphic map

$$f^{an} : X^{an} \rightarrow Y^{an}$$

such that the following diagram commutes

$$\begin{array}{ccc} X^{an} & \xrightarrow{\phi_X} & X \\ f^{an} \downarrow & & \downarrow f \\ Y^{an} & \xrightarrow{\phi_Y} & Y \end{array}$$

Again, one obtains comparison theorems between the algebraic properties of the morphisms

$$f : X \rightarrow Y \text{ and } f^{an} : X^{an} \rightarrow Y^{an},$$

see [50, Chap. 3].

3. Coherent sheaves and proper maps: The morphism

$$\phi : X \rightarrow X^{an}$$

from the analytification of X defines the functor *analytical inverse image* as the pull-back of module sheaves over the structure sheaf:

$$\phi^* : \underline{Sheaf}_{\mathcal{O}_X} \rightarrow \underline{Sheaf}_{\mathcal{O}_{X^{an}}}, \mathcal{F} \mapsto \phi^* \mathcal{F}$$

and

$$f : \mathcal{F} \rightarrow \mathcal{G}, \mapsto \phi^* f : \phi^* \mathcal{F} \rightarrow \phi^* \mathcal{G}.$$

The stalk at a point $x \in X^{an}$ of the inverse image sheaf is easy to calculate

$$(\phi^* \mathcal{F})_x = \mathcal{F}_{\phi(x)} \otimes_{\mathcal{O}_{X, \phi(x)}} \mathcal{O}_{X^{an}, x}$$

Serre investigated in his GAGA-paper [53] for a projective schema $X \in \underline{Al}_{\mathbb{C}}$ and a coherent $\mathcal{O}_X$ -module the relation between the cohomology in the algebraic and the analytic category. Subsequently his results have been generalized to the relative case: For a proper morphism

$$f : X \rightarrow Y$$

in $\underline{Al}_C$ and a coherent $\mathcal{O}_X$ -module sheaf $\mathcal{F}$ the canonical morphisms in $\underline{An}_C$ between the corresponding direct image sheaves

$$(R^j f_* \mathcal{F})^{an} \xrightarrow{\simeq} R^j f_*^{an}(\mathcal{F}^{an}), \quad j \in \mathbb{N},$$

are isomorphisms, see [50, Chap. 4].

The literal analogue of divisors on compact Riemann surfaces are divisors on a projective algebraic curve C/k . A priori, these divisors refer to the points of the curve $C \subset \mathbb{P}^n(\bar{k})$ with coordinates in $\bar{k}$. In a second step we will consider those divisors which are defined over k .

Definition 4.19 (Divisors on projective algebraic curves). Consider a projective algebraic curve C/k with $C \subset \mathbb{P}^n(\bar{k})$.

1. A *divisor* on C/k is formal sum

$$D = \sum_{x \in C} n_x \cdot x$$

with integers $n_x \in \mathbb{Z}$ satisfying $n_x = 0$ for all but finitely many indices. For a point $p \in C$ the corresponding point divisor is denoted

$$P := 1 \cdot p$$

Each divisor $D \in \text{Div}(C)$ has a well-defined degree

$$\deg D := \sum_{x \in C} n_x \in \mathbb{Z}$$

The additive group of divisors on C/k is denoted $\text{Div}(C)$, its subgroup $\text{Div}_0(C) \subset \text{Div}(C)$ contains all divisors of degree = 0.

2. A divisor $D \in \text{Div}(C)$ is a *principal divisor* if $D = 0$ or

$$D = \text{div } f$$

for a rational function $f \in \bar{k}(C)$. Two divisors $D_1, D_2 \in \text{Div}(C)$ are *equivalent*, annotated

$$D_1 \sim D_2,$$

if their difference is a principal divisor. The divisor of a rational function $f \in \bar{k}^*$ has degree

$$\deg(\text{div } f) = 0.$$

Denote by

$$\text{Prin}(C) \subset \text{Div}_0(C)$$

the subgroup of principal divisors. The quotient

$$Cl(C) := Div(C)/Prin(C)$$

is named the *divisor class group* of C . Its subgroup

$$Cl_0(C) := Div_0(C)/Prin(C) \subset Cl(C)$$

is the *class group of degree zero divisors*.

3. A divisor

$$D = \sum_{x \in C} n_x \cdot x \in Div(C)$$

is *defined over k* if it is invariant under the action of the Galois group $G_{\bar{k}/k}$, i.e. if for all $\sigma \in G_{\bar{k}/k}$

$$D = D^\sigma := \sum_{x \in C} n_x \cdot x^\sigma \in Div(C)$$

4. For a divisor $D \in Div(C)$ we denote by $\mathcal{O}_D$ the sheaf of multiples of $-D$, i.e.

$$\mathcal{O}_D(U) := \{f \in \mathcal{M}^*(U) : \text{div } f \geq -D|_U\} \cup \{0\}, \quad U \subset X \text{ open.}$$

For projective-algebraic curves C/k the theorems of Riemann-Roch, Serre duality, the very-ampleness criterion from Theorem 4.13 and the dimension formula

$$\dim H^0(C, \mathcal{O}) = 1$$

hold literally in the same form as on compact Riemann surfaces. Also the canonical divisor $K \in Div(C)$ has degree

$$\deg K = 2g(C) - 2.$$

Hence we will apply these results also in the context of algebraic geometry without further mentioning.

Definition 4.20 (Elliptic curve over k). An *elliptic curve defined over k* is a pair (E, O) with

- a projective algebraic curve E/k of genus $g = 1$
- and a k -valued point $O \in E$

Note. In Definition 4.20 the notation “O” denotes the upper-case letter O, not the digit 0. We denote by

$$O_\infty := 1 \cdot O \in \text{Div}(E)$$

the distinguished point divisor. It is defined over k . If the specific base field k and the k -valued point O are not relevant for the context, we will often denote an elliptic curve just by E .

Note that there exist smooth projective algebraic curves $C/\mathbb{Q}$ of genus $g = 1$ without rational points, i.e. with $C(\mathbb{Q}) = \emptyset$. An example is the *Selmer* curve $C \in \mathbb{P}^2(\mathbb{C})$ defined by the homogeneous cubic equation

$$3X^3 + 4Y^3 + 5Z^3 = 0$$

The formula for the genus g of a hypersurface $F \subset \mathbb{P}^2$ of degree d is

$$g = \frac{(d-1)(d-2)}{2}$$

Hence for the Selmer curve

$$g(C) = \frac{2 \cdot 1}{2} = 1$$

It is not an easy task to show that the Selmer curve has no rational points, see the reference to Selmer’s paper in [60, Chap. I, §2].

The fact $g = 1$ for an elliptic curve E allows a simple application of the Riemann-Roch theorem to compute the vector spaces of multiples of a divisor on E .

Lemma 4.21 (Divisors on an elliptic curve). *Consider an elliptic curve E and a divisor $D \in \text{Div}(E)$. Then*

$$\dim H^0(E, \mathcal{O}_D) = \begin{cases} 0 & \text{if } \deg D < 0 \\ 0 & \text{if } \deg D = 0, D \text{ not principal} \\ 1 & \text{if } \deg D = 0, D \text{ principal} \\ \deg D & \text{if } \deg D > 0 \end{cases}$$

Proof. The Riemann-Roch theorem together with Serre duality imply for any divisor $D \in \text{Div}(E)$

$$\begin{aligned} \dim H^0(E, \mathcal{O}_D) - \dim H^1(E, \mathcal{O}_D) &= \dim H^0(E, \mathcal{O}_D) - \dim H^0(E, \mathcal{O}_{-D+K}) = \\ &= 1 - g + \deg D = \deg D \end{aligned}$$

Because an elliptic curve E has genus $g = 1$, its canonical divisor has degree

$$\deg K = 2g - 2 = 0.$$

As a consequence

$$\deg(-D + K) = \deg(-D) = -\deg D$$

i) Case $\deg D \neq 0$: If $\deg D > 0$ then $\deg(-D) < 0$. As a consequence

$$\dim H^0(E, \mathcal{O}_D) = \begin{cases} \deg D & \text{if } \deg D > 0 \\ 0 & \text{if } \deg D < 0 \end{cases}$$

ii) Case $\deg D = 0$: We consider the sheaf $\mathcal{O}_D$ as a line bundle. Then

$$D \text{ principal} \iff \mathcal{O}_D \simeq \mathcal{O}$$

On one hand, D principal implies

$$H^0(E, \mathcal{O}_D) = \dim H^0(E, \mathcal{O}) = 1$$

On the other hand, for non-principal D the line bundle $\mathcal{O}_D$ is not trivial. Hence each regular section of $s \in H^0(E, \mathcal{O}_D)$ has a zero. If $s \neq 0$, then $\deg D = 0$ implies that s has also at least a pole, contradicting the regularity of s . Therefore $s = 0$.

□

The following Lemma 4.22 will be used in the proof of Theorem 4.23. It indicates why the existence of a k -valued point is requested for an elliptic curve defined over k .

Lemma 4.22 (Divisors defined over a subfield). *Consider a projective algebraic curve C/k and a divisor*

$$D = \sum_{x \in X} n_x \cdot x \in \text{Div}(C)$$

which is defined over k . Then the finite dimensional $\bar{k}$ -vector space

$$H^0(C, \mathcal{O}_D) = \{f \in \bar{k}^*(C) : \text{div } f \geq -D\} \cup \{0\}$$

has a basis of rational functions from $k(C)$, more precisely

$$H^0(C, \mathcal{O}_D) = \bar{k} \otimes_k \{f \in k^*(C) : \text{div } f \geq -D\} \cup \{0\}$$

For the proof see [60, II, Prop. 5.8].

Elliptic curves are by definition subvarieties of a projective space $\mathbb{P}^n(\bar{k})$.

Theorem 4.23 proves that they are even plane curves, i.e. that they embed as hypersurfaces into $\mathbb{P}^2(\bar{k})$.

Theorem 4.23 (Elliptic curves are plane cubics). *For an elliptic curve $(E/k, O)$ exist two rational functions $x, y \in k(E)$ with poles only at O such that the map*

$$\phi : E \rightarrow \mathbb{P}^2(\bar{k}), p \mapsto \begin{cases} (1 : x(p) : y(p)) & p \neq O \\ (0 : 0 : 1) & p = O \end{cases}$$

induces an isomorphism onto the plane cubic $C \subset \mathbb{P}^2(\bar{k})$, which is defined by a homogeneous polynomial $F_{hom} \in k[X_0, X_1, X_2]$ of the form

$$F_{hom}(X_0, X_1, X_2) = X_0X_2^2 + a_1X_0X_1X_2 + a_3X_0^2X_2 - (X_1^3 + a_2X_0X_1^2 + a_4X_0^2X_1 + a_6X_0^3)$$

with five suitable constants $a_i \in k, i = 1, 2, 3, 4, 6$. In particular, C is defined over k and non-singular.

Proof. We construct a very ample line bundle $\mathcal{L} = \mathcal{O}_D$ on E and define $\phi := \phi_{\mathcal{L}}$. Therefore consider the divisor

$$D := 3 \cdot O_{\infty} \in \text{Div}(E).$$

The line bundle

$$\mathcal{L} := \mathcal{O}_D \in \text{Pic}(E)$$

satisfies the very ampleness criterion from Theorem 4.13: The proof follows from the dimension formula in Lemma 4.21 because for all points $p, q \in E$

$$\dim H^0(E, \mathcal{O}_{D-p}) = \dim H^0(E, \mathcal{O}_D) - 1$$

and

$$\dim H^0(E, \mathcal{O}_{D-(p+q)}) = \dim H^0(E, \mathcal{O}_D) - 2.$$

For $n \geq 0$

$$H^0(E, \mathcal{O}_{n \cdot O_{\infty}}) = \{f \in \bar{k}^*(E) : f \text{ has at most a pole at } O \text{ of order } = n\} \cup \{0\}$$

and the dimension formula implies the proper inclusion

$$H^0(E, \mathcal{O}_{3 \cdot O_{\infty}}) \supsetneq H^0(E, \mathcal{O}_{2 \cdot O_{\infty}}) \supsetneq H^0(E, \mathcal{O}_{O_{\infty}}) \simeq H^0(E, \mathcal{O}) = \mathbb{C}.$$

Hence a basis $(1, x, y)$ of

$$H^0(E, \mathcal{L}) = H^0(E, \mathcal{O}_{3 \cdot O_{\infty}})$$

exists with

- $y \in H^0(E, \mathcal{O}_{3 \cdot O_{\infty}}) \subset \bar{k}(E)$ has a pole of order $= 3$ at O and no other pole,
- $x \in H^0(E, \mathcal{O}_{2 \cdot O_{\infty}}) \subset \bar{k}(E)$ has a pole of order $= 2$ at O and no other pole,
- $1 \in H^0(E, \mathcal{O}) = H^0(E, \mathcal{O}_{1 \cdot O_{\infty}}) \subset \bar{k}(E)$ has no poles.

By assumption $O \in E(k)$. Therefore the divisor D is defined over k . According to Lemma 4.22 we may choose even $x, y \in k(E)$. With respect to the basis $(1, x, y)$ the line bundle $\mathcal{L}$ defines the closed embedding

$$\phi := \phi_{\mathcal{L}} : E \rightarrow \mathbb{P}^2(\bar{k}), \quad p \mapsto (1 : x(p) : y(p)),$$

onto a curve C/k in $\mathbb{P}^2(\bar{k})$. The family

$$(1, x, y, x^2, xy, x^3, y^2)$$

of seven rational functions from the 6-dimensional vector space $H^0(E, \mathcal{O}_{6, O_\infty})$ is linearly dependent. Therefore the functions satisfy a linear relation with coefficients from k . Among them the two functions x^3 and y^2 are the only ones having a pole at $O \in E$ of order = 6. Hence their coefficients in the linear relation are not zero. After division we may assume that their coefficients are ± 1 and

$$y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6$$

with coefficients $a_i \in k$. This relation defines the inhomogeneous polynomial

$$F(X, Y) := Y^2 + a_1XY + a_3Y - (X^3 + a_2X^2 + a_4X + a_6) \in k[X, Y]$$

and its homogenization, the polynomial $F_{hom} \in k[X_0, X_1, X_2]$ with

$$F_{hom}(X_0, X_1, X_2) := X_0X_1^2 + a_1X_0X_1X_2 + a_3X_2X_0^2 - (X_1^3 + a_2X_1^2X_0 + a_4X_1X_0^2 + a_6X_0^3).$$

The inhomogeneous variables (X, Y) relate to the homogeneous variables (X_0, X_1, X_2) according to

$$X = \frac{X_1}{X_0} \text{ and } Y = \frac{X_2}{X_0}.$$

The curve $\phi(E)$ is a subvariety of the 1-dimensional cubic hypersurface

$$C := V(F_{hom}) \subset \{(x_0 : x_1 : x_2) \in \mathbb{P}^2(\bar{k}) : F_{hom}(x_0, x_1, x_2) = 0\}.$$

The cubic C intersects the line at infinity

$$\mathbb{P}^1(\bar{k}) \simeq \{(x_0 : x_1 : x_2) \in \mathbb{P}^2(\bar{k}) : x_0 = 0\}$$

in the single point $(0 : 0 : 1)$ with multiplicity = 3. Its affine part is the affine variety of the polynomial $F(X, Y) \in k[X, Y]$

$$\begin{aligned} C_{\text{aff}} &= \{(x_0 : x_1 : x_2) \in \mathbb{P}^2(\bar{k}) : F_{hom}((x_0 : x_1 : x_2)) = 0 \text{ and } x_0 \neq 0\} = \\ &= \{(x, y) \in \mathbb{A}^2(\bar{k}) : F(x, y) = 0\} = V(F). \end{aligned}$$

The polynomial $F \in k[X][Y]$, considered as polynomial in Y with coefficients from the ring $k[X]$, is irreducible: It does not split as

$$F(X, Y) = (Y + b_1)(Y + b_2)$$

with polynomials $b_1, b_2 \in k[X]$. Therefore C_{aff} and also C are irreducible. Hence the non-constant regular map

$$\phi : E \rightarrow C$$

is surjective, cf. [31, Chap. II, Prop. 6.8], an analogue of the theorem about non-constant proper holomorphic maps between Riemann surfaces. As a consequence

$$\phi(E) = C \text{ and } E \simeq C \text{ over } k$$

and C is non-singular.

□

In the following we will identify an abstract elliptic curve $(E/k, O)$ with the plane cubic from Theorem 4.23. We say that $(E/k, O)$, more precisely its affine part, is defined over k by the *Weierstrass equation* $F(X, Y) = 0$.

Definition 4.24 (Weierstrass polynomial and Weierstrass equation, discriminant and j -invariant). A *Weierstrass polynomial* over a field k is a cubic polynomial

$$F(X, Y) = Y^2 + a_1XY + a_3Y - (X^3 + a_2X^2 + a_4X + a_6) \in k[X, Y]$$

with coefficients

$$a_1, a_2, a_3, a_4, a_6 \in k.$$

1. The equation

$$F(X, Y) = 0,$$

i.e.

$$Y^2 + a_1XY + a_3Y = X^3 + a_2X^2 + a_4X + a_6$$

is named the *Weierstrass equation* over k defined by F .

2. The *discriminant* of F is

$$\Delta_F := -b_2^2b_8 - 8b_4^3 - 27b_6^2 + 9b_2b_4b_6$$

with

$$b_2 := a_1^2 + 4a_2, \quad b_4 := 2a_4 + a_1a_3, \quad b_6 := a_3^2 + 4a_6$$

and

$$b_8 := a_1^2a_6 + 4a_2a_6 - a_1a_3a_4 + a_2a_3^2 - a_4^2.$$

3. The *j -invariant* of F is

$$j_F := \frac{c_4^3}{\Delta_F}, \quad c_4 := b_2^2 - 24b_6.$$

Starting from an elliptic curve over a field k the resulting Weierstrass polynomial is not unique. There exist coordinate transformations defined over k which transform a Weierstrass polynomial over k to a different form, in particular to a more simple form.

Proposition 4.25 (Weierstrass equation in short form, discriminant and j -invariant).

Consider a Weierstrass polynomial

$$F(X, Y) = Y^2 + a_1XY + a_3Y - (X^3 + a_2X^2 + a_4X + a_6) \in k[X, Y]$$

1. Short form: If $\text{char } k \neq 2, 3$ then there exists a linear coordinate change

$$\begin{pmatrix} X \\ Y \end{pmatrix} \mapsto \begin{pmatrix} X' \\ Y' \end{pmatrix} = \begin{pmatrix} u^2 \cdot X + r \\ u^3 \cdot Y + s \cdot u^2 \cdot X + t \end{pmatrix}$$

with constants

$$r, s, t \in k, u \in k^*$$

which transforms $F(X, Y)$ to the Weierstrass polynomial in short form

$$F'(X', Y') = Y'^2 - (X'^3 + A \cdot X' + B), \quad A, B \in k$$

with discriminant

$$\Delta_{F'} = -2^4 \cdot (4A^3 + 27B^2) \in k$$

and the same j -invariant

$$j_{F'} = j_F$$

2. Transforming Weierstrass polynomials in short form: Each linear coordinate transformation over k

$$\begin{pmatrix} X' \\ Y' \end{pmatrix} \mapsto \begin{pmatrix} X \\ Y \end{pmatrix} := \begin{pmatrix} u^2 & 0 \\ 0 & u^3 \end{pmatrix} \cdot \begin{pmatrix} X' \\ Y' \end{pmatrix}.$$

with $u \in k^*$ transforms a Weierstrass polynomial in short form

$$F'(X', Y') = Y'^2 - (X'^3 + A' \cdot X' + B') \in k[X', Y']$$

to the Weierstrass polynomial in short form

$$F(X, Y) = Y^2 - (X^3 + A \cdot X + B) \in k[X, Y].$$

such that the numerical constants relate as

$$A = u^4 \cdot A', \quad B = u^6 \cdot B', \quad \Delta_F = u^{12} \cdot \Delta_{F'}.$$

retaining the j -invariant

$$j_F := 1728 \cdot \frac{(4A)^3}{\Delta_F} = 1728 \cdot \frac{(4A')^3}{\Delta_{F'}} =: j_{F'}$$

3. The j -invariant: Two Weierstrass polynomials in short form have the same j -invariant if and only if they are related by an isomorphism as in part 2) defined over $\bar{k}$ - not necessarily over k .

4. For each algebraic number $j_0 \in \bar{k}$ exists an elliptic curve $E/k(j_0)$ with j -invariant equal to j_0 .

For the proof of Proposition 4.25 see [62]: For part 1) cf. Chap. III, §1; for part 2) cf. Chap. III, Rem. 1.3; for part 3) cf. Chap. III, Prop. 1.4(b); for part 4) cf. Chap. III, Prop. 1.4(c).

The discriminant $\Delta_F \in k$ of a Weierstrass polynomial $F(X, Y) \in k[X, Y]$ indicates whether the projective algebraic variety

$$C = V(F_{hom}) \subset \mathbb{P}^2(\bar{k})$$

is non-singular or not. If $\Delta_F \neq 0$ then $(C/k, (0 : 0 : 1))$ is non-singular, i.e. an elliptic curve.

Proposition 4.26 (Singularities of cubics). Assume $\text{char } k \neq 2, 3$ and consider a plane cubic curve

$$C := V(F_{hom}) \subset \mathbb{P}^2(\bar{k})$$

defined by the homogenization of a Weierstrass polynomial in short form

$$F(X, Y) = Y^2 - (X^3 + AX + B) \in k[X, Y]$$

with $A, B \in k$ and discriminant

$$\Delta_F = -16 \cdot (4A^3 + 27B^2).$$

The curve C has at most one singular point. The curve

- is non-singular iff $\Delta_F \neq 0$
- has a node iff $\Delta_F = 0$ and $A \neq 0$
- has a cusp iff $\Delta_F = A = 0$.

For a proof of Proposition 4.26 see [62, Chap. III, Prop. 1.4].

Theorem 4.27 is the converse of Theorem 4.3. it shows that the complex points of an elliptic curve, defined over a subfield of $\mathbb{C}$, can be parametrized by the Weierstrass function $\wp$ and its derivative $\wp'$ of a suitable torus. The proof relies on the result of Theorem 4.23 that elliptic curves are plane cubics which satisfy a Weierstrass equation.

A very coarse analogue of the uniformization is the parametrization of the points of the unit circle by the values of the exponential function $e^{2\pi i \cdot t}$ on the parameter interval $[0, 1]$.

Theorem 4.27 (Uniformization of elliptic curves $E/\mathbb{C}$ by complex tori). *For any elliptic curve E defined over $\mathbb{C}$ exists a torus $T = \mathbb{C}/\Lambda$ and a biholomorphic map*

$$\Phi : T \xrightarrow{\sim} E(\mathbb{C}).$$

Here $E(\mathbb{C}) \subset \mathbb{P}^n(\mathbb{C})$ is provided with the induced Euclidean topology and the analytic structure from the complex manifold $\mathbb{P}^2(\mathbb{C})$.

Proof. Due to Theorem 4.23 we may assume that E is defined as a plane curve, defined by the homogenization of a Weierstrass polynomial in short form

$$F(X, Y') := Y'^2 - (X^3 + a_4 \cdot X + a_6) \in \mathbb{C}[X, Y'].$$

The substitution

$$Y' = (1/2) \cdot Y$$

shows that E can also be defined by the Weierstrass polynomial

$$F(X, Y) = Y^2 - (4 \cdot X^3 - A \cdot X - B)$$

with suitable constants $A, B \in \mathbb{C}$ and discriminant

$$\Delta_F = A^3 - 27 \cdot B^3$$

due to Proposition 4.26. We search for a lattice

$$\Lambda = \mathbb{Z}\omega_1 + \mathbb{Z}\omega_2 \subset \mathbb{C}$$

with lattice constants $G_{\Lambda,4}$ and $G_{\Lambda,6}$ satisfying

$$A = 60 \cdot G_{\Lambda,4} \text{ and } B = 140 \cdot G_{\Lambda,6}.$$

The restriction of the modular invariant

$$j|_{\mathbb{H}} : \mathbb{H} \rightarrow \mathbb{C}, \quad j(\tau) := 1728 \cdot \frac{g_2^3(\tau)}{g_2^3(\tau) - 27g_3^2(\tau)},$$

is surjective according to Corollary 3.16. Hence a number $\tau \in \mathbb{H}$ exists with

$$j(\tau) = 1728 \cdot \frac{A^3}{\Delta_F}.$$

If $g_2(\tau) = 0$ then $\tau = \rho$ according to Corollary 3.20. We choose

$$\Lambda := \Lambda_\rho.$$

Similarly, if $g_3(\tau) = 0$ then $\tau = i$, and we choose

$$\Lambda := \Lambda_i.$$

Otherwise, the two equations

$$j(\tau) = 1728 \cdot \frac{g_2^3(\tau)}{g_2^3(\tau) - 27g_3^2(\tau)} = 1728 \cdot \frac{A^3}{\Delta_F}$$

and

$$\Delta_F = A^3 - 27B^2$$

imply

$$\frac{g_2^3(\tau)}{g_2^3(\tau) - 27g_3^2(\tau)} = \frac{A^3}{A^3 - 27 \cdot B^2}.$$

Clearing the denominators by crosswise multiplication implies

$$\left(\frac{A}{g_2(\tau)} \right)^3 = \left(\frac{B}{g_3(\tau)} \right)^2$$

We choose $\omega_1 \in \mathbb{C}$ with

$$\frac{A}{g_2(\tau)} = \left(\frac{1}{\omega_1} \right)^4 \text{ and } \frac{B}{g_3(\tau)} = \left(\frac{1}{\omega_1} \right)^6.$$

This choice is possible because we may replace ω_1 by $i\omega_1$. In addition, we set

$$\omega_2 := \tau \cdot \omega_1, \text{ i.e. } \tau = \omega_2 / \omega_1.$$

Then (ω_1, ω_2) is a positively oriented basis of the lattice

$$\Lambda := \mathbb{Z}\omega_1 + \mathbb{Z}\omega_2.$$

Its lattice constants are

$$G_{\Lambda,4} = \left(\frac{1}{\omega_1} \right)^4 \cdot G_4(\tau) = \left(\frac{1}{\omega_1} \right)^4 \cdot (1/60) \cdot g_2(\tau) = (1/60) \cdot A$$

and similarly

$$G_{\Lambda,6} = (1/140) \cdot B,$$

i.e.

$$A = g_{\Lambda,2} \text{ and } B = g_{\Lambda,3}.$$

The map

$$\phi : \mathbb{C}/\Lambda \rightarrow \mathbb{P}^2$$

from Theorem 4.3 maps the torus $\mathbb{C}/\Lambda$ biholomorphically to the cubic hypersurface with Weierstrass polynomial

$$F(X, Y) = Y^2 - (4 \cdot X^3 - g_{\Lambda,2} \cdot X - g_{\Lambda,3}) = Y^2 - (4 \cdot X^3 - A \cdot X - B)$$

□

Example 4.28 (Elliptic curves and cubics with a singularity). The following Figures 4.5 - 4.8 show some elliptic curves. The table from Figure 4.4 shows their invariants. The role of the *conductor* will be explained later in Definition 4.39.

Figure 4.9 and Figure 4.10 show the singular cubics

- *node* with its double crossing
- and the *Neil parabola* with its cusp.

Both cubics have a single singularity.

For the computation and the plots see the PARI files `Elliptic_curve_02` and `Elliptic_curve_plot_06`.

Figure	Weierstrass equation	j -invariant	discriminant	conductor
Elliptic curve				
4.5	$Y^2 = X^3 - 3X + 3$	$-2^8 \cdot 3^3 / 5$	$-2^4 \cdot 3^3 \cdot 5$	$2^3 \cdot 3^3 \cdot 5$
4.6	$Y^2 = X^3 + X$	1728	-2^6	2^6
4.7	$Y^2 = X^3 - X$	1728	2^6	2^5
4.8	$Y^2 = X^3 + 3$	0	$-2^4 \cdot 3^5$	$2^4 \cdot 3^5$
	$Y^2 = X^3 - X - 14$	$-2^5 \cdot 3^3 / 661$	$-2^7 \cdot 661$	$2^7 \cdot 661$
	$Y^2 + 5XY + Y = X^3$	$5^3 \cdot 101^3 / (2 \cdot 7^2)$	$2 \cdot 7^2$	$2 \cdot 7$
Singular cubic				
4.9	$Y^2 = X^3 + X^2$	—	0	—
4.10	$Y^2 = X^3$	—	0	—

Fig. 4.4 Numerical values of some cubic curves

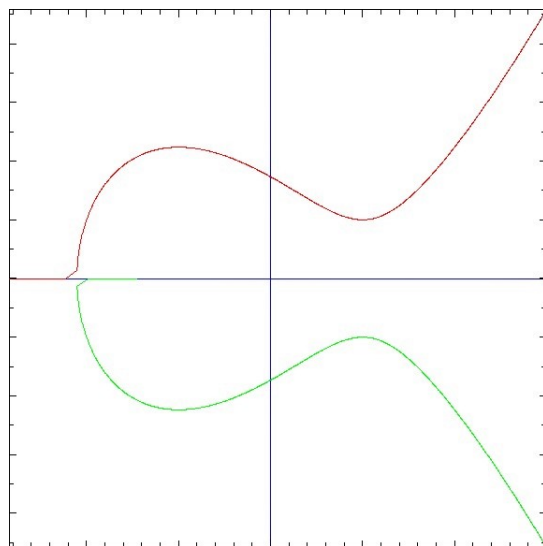


Fig. 4.5 $Y^2 = X^3 - 3X + 3$

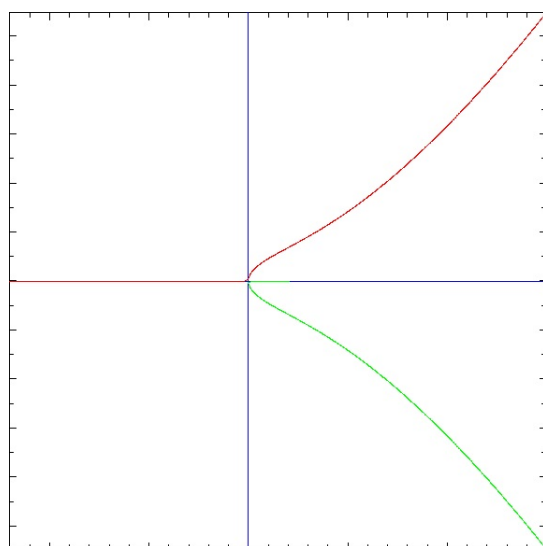


Fig. 4.6 $Y^2 = X^3 + X$

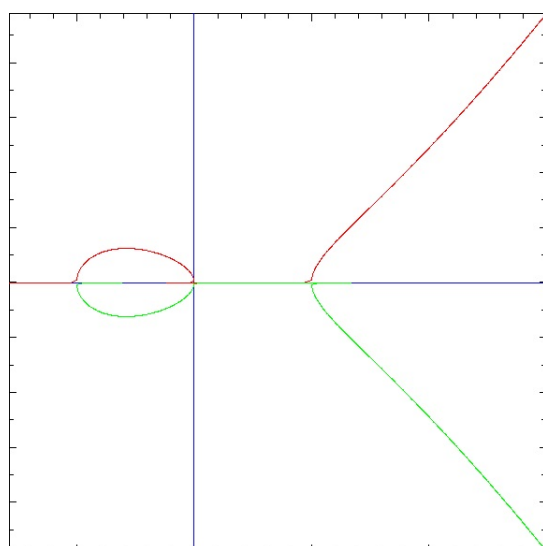


Fig. 4.7 $Y^2 = X^3 - X$

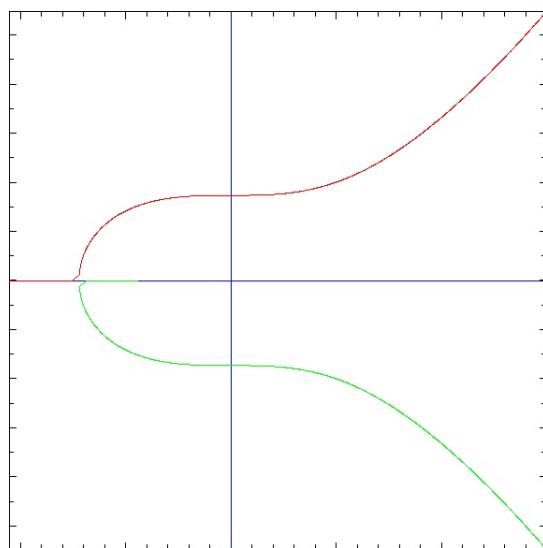


Fig. 4.8 $Y^2 = X^3 + 3$

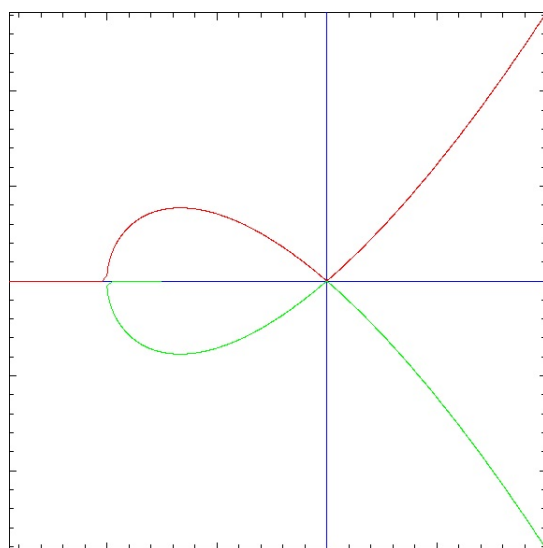


Fig. 4.9 $Y^2 = X^3 + X^2$

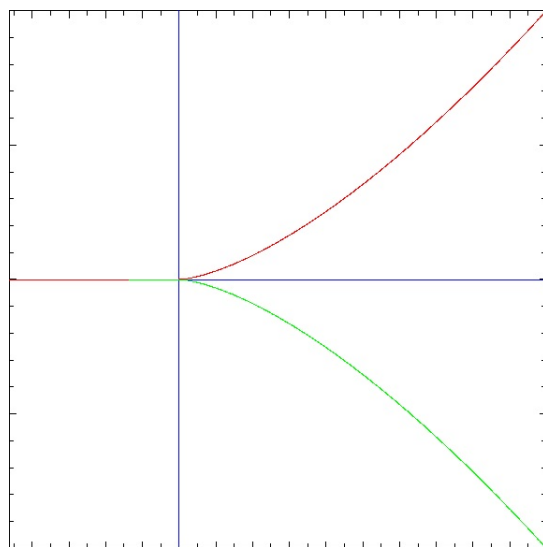


Fig. 4.10 $Y^2 = X^3$

According to Theorem 4.3 a torus is biholomorphic equivalent to an elliptic curve $E/\mathbb{C}$. Hence we expect that the biholomorphic map ϕ carries over the abelian group structure on the torus to a group structure on the elliptic curve. Theorem 4.23 suggests that ϕ maps the neutral elements, i.e. that the distinguished point $O \in E$ becomes the neutral element of the group structure on the elliptic curve. Our next goal is to make explicit the group structure on E .

Lemma 4.29 prepares the proof of Theorem 4.30 and the Definition 4.31 of the group structure of an elliptic curve $(E/k, O)$. The lemma shows that for a pair (p, q) of distinct points from E the divisor

$$D := P - Q \in \text{Div}(E)$$

cannot be a principal divisor. Note that the lemma gives also a second proof of Corollary 1.9.

Lemma 4.29 (No single pole of order one). *Consider an elliptic curve $(E/k, O)$ defined over a field k . There is no rational function $f \in \bar{k}(E)^*$ with*

$$\text{div } f = P - Q$$

with $P, Q \in \text{Div}(E)$ the elementary divisors of two distinct points $p \neq q \in E$.

Proof. Consider the point divisor

$$Q \in \text{Div}(E) \text{ of degree } \deg Q = 1.$$

Lemma 4.21 implies

$$\dim H^0(E, \mathcal{O}_Q) = 1.$$

Hence the inclusion

$$\bar{k} \subset H^0(E, \mathcal{O}_Q)$$

implies

$$\bar{k} = H^0(E, \mathcal{O}_Q).$$

Hence each $f \in H^0(E, \mathcal{O}_Q)$ is constant, and therefore $f = 0$ or

$$\text{div } f = 0 \in \text{Div}(E)$$

which implies $p = q$, a contradiction, q.e.d.

Lemma 4.29 shows: The obstructions for a divisor $D \in \text{Div}_0(E)$ to be a principal divisor are divisors of type $P - Q$ or even of type $P - O_\infty$. A divisor $D \in \text{Div}_0(E)$ is the divisor of a rational function up to a divisor of type

$$P - O_\infty.$$

Because the point $P \in E$ is uniquely determined by D , the obstruction is expressed by P alone.

We use the notation from Definition 4.19, notably the equivalence relation for divisors.

Theorem 4.30 (The class group of degree zero divisors on an elliptic curve).

Consider an elliptic curve $(E/k, O)$.

1. For any divisor $D \in \text{Div}_0(E)$ exists a unique point $p \in E$ with

$$D \sim P - O_\infty$$

Denote the corresponding map by

$$\sigma : \text{Div}_0(E) \rightarrow E, D \mapsto p.$$

2. The map σ is surjective. It induces a bijective map from the group of divisor classes of degree = 0

$$\bar{\sigma} : \text{Cl}_0(E) \xrightarrow{\sim} E$$

with inverse

$$E \xrightarrow{\sim} \text{Cl}_0(E), p \mapsto [P - O_\infty].$$

Proof. 1. Definition of σ : The divisor $D + O_\infty$ has degree = 1. Lemma 4.21 implies

$$\dim H^0(E, \mathcal{O}_{D+O_\infty}) = 1.$$

We choose a non-zero rational function $f \in \bar{k}(E)^*$ with divisor satisfying

$$\text{div } f \geq -(D + O_\infty).$$

Because

$$\deg(\text{div } f) = 0 \text{ but } \deg(-(D + O_\infty)) = -1,$$

there exists a point $p \in E$ with

$$\text{div } f = -(D + O_\infty) + P$$

Then

$$0 \sim -(D + O_\infty) + P \text{ i.e. } D \sim P - O_\infty.$$

If for a different point

$$q \in E, q \neq p,$$

also

$$D \sim Q - O_\infty$$

then

$$0 = D - D \sim P - Q.$$

Hence for a suitable rational function $g \in \bar{k}^*(E)$

$$P - Q = \text{div } g,$$

contradicting Lemma 4.29. Therefore

$$\sigma(D) := p \in E$$

is well-defined.

2. i) *Surjectivity of σ* : Consider an arbitrary point $p \in E$. Set

$$D := P - O_\infty \in \text{Div}_0(E).$$

Then $\sigma(D) = P$ by definition, which proves the surjectivity of σ .

ii) *Induced map $\bar{\sigma}$* :

$$[D_1] = [D_2] \implies P_1 - O_\infty \sim P_2 - O_\infty \implies P_1 \sim P_2 \implies \sigma(D_1) = \sigma(D_2)$$

iii) *Injectivity of $\bar{\sigma}$* :

$$\begin{aligned} \bar{\sigma}([D_1]) = \bar{\sigma}([D_2]) &\implies \sigma(D_1) = \sigma(D_2) \implies \\ &\implies \exists p \in E : D_1 \sim P - O_\infty \text{ and } D_2 \sim P - O_\infty \implies D_1 \sim D_2 \implies [D_1] = [D_2] \end{aligned}$$

□

Definition 4.31 (Group structure of an elliptic curve). Consider an elliptic curve $(E/k, O)$ and denote by $(Cl_0(E), +)$ the group of divisor classes of degree zero. The bijective map from Theorem 4.30

$$\bar{\sigma} : Cl_0(E) \rightarrow E, [P - O_\infty] \mapsto p,$$

carries over the group structure from $(Cl_0(E), +)$ to E : The map

$$\oplus : E \times E \rightarrow E, p \oplus q := \bar{\sigma}([P - O_\infty] + [Q - O_\infty]),$$

defines an Abelian *additive group* $(E, \oplus)$ with neutral element the k -valued point $O \in E(k)$.

Corollary 4.32 (Abel's Theorem for elliptic curves). Consider an elliptic curve $(E/k, O)$. A divisor

$$D = \sum_{p \in E} n_p \cdot P \in \text{Div}_0(E)$$

is a *principal divisor* iff

$$\bigoplus_{p \in E} n_p * p = O \in E.$$

Here the last equation refers to the group structure $\oplus$ on E . The product $n_p * p$ means:

- Taking n_p -times the sum of a point $p \in E$ if $n_p \geq 0$, and
- taking $(-n_p)$ -times the sum of $\ominus p \in E$ if $n_p \leq 0$. Here the symbol $\ominus p \in E$ denotes the negative of p with respect to the group structure.

Proof. Because

$$0 = \deg D = \sum_{p \in E} n_p$$

we have

$$D = D - \sum_{p \in E} n_p \cdot O_\infty = \sum_{p \in E} n_p \cdot (P - O_\infty) \in \text{Div}_0(E)$$

Due to Theorem 4.30, part ii)

$$\begin{aligned} D \text{ principal} &\iff [D] = 0 \iff \overline{\sigma}([D]) = O \in E \iff \\ &\iff \overline{\sigma}\left(\sum_{p \in E} n_p \cdot [P - O_\infty]\right) = O \in E \iff \bigoplus_{p \in E} n_p * \overline{\sigma}([P - O_\infty]) = O \in E \iff \\ &\iff \bigoplus_{p \in E} n_p * p = 0 \in E. \end{aligned}$$

□

Proposition 4.33 (Group structure of tori and elliptic curves). *For a complex torus $T = \mathbb{C}/\Lambda$ the biholomorphic map from Theorem 4.3*

$$\phi : T \rightarrow E(\mathbb{C})$$

onto the complex points of a plane elliptic curve E is an isomorphism of groups.

Proof. 1. i) Denote by $0_T \in T$ the neutral element of the group $(T, \oplus)$. The canonical morphism

$$\rho : (T, \oplus) \rightarrow (\text{Div}_0(T), +), \quad p \mapsto 1 \cdot p - 1 \cdot 0_T,$$

is a group morphism: For $j = 1, 2$ and two points $p_j \in T$ and divisors from $\text{Div}_0(T)$

$$D_j := 1 \cdot p_j - 1 \cdot 0_T \text{ and } D := 1 \cdot (p_1 \oplus p_2) - 1 \cdot 0_T$$

the difference

$$D_1 + D_2 - D = 1 \cdot p_1 + 1 \cdot p_2 - 1 \cdot (p_1 \oplus p_2) - 1 \cdot 0_T \in \text{Div}_0(T)$$

is a principal divisor according to Abel's Theorem 1.21 because

$$p_1 \oplus p_2 \ominus (p_1 \oplus p_2) \ominus 0_T = 0_T \in \Lambda$$

ii) For a point $p \in T$ the image

$$\rho(p) = 1 \cdot p - 1 \cdot 0_T \in \text{Div}_0(T)$$

is a principal divisor iff $p = 0_T$, again due to Abel's Theorem. If we denote by

$$\bar{\rho} : T \rightarrow \text{Cl}_0(T)$$

the composition of ρ with the canonical map $\text{Div}_0(T) \rightarrow \text{Cl}_0(T)$ then

$$\bar{\rho} : T \rightarrow \text{Cl}_0(T)$$

is injective.

iii) Each divisor

$$D = \sum_{a \in A} n_a \cdot a - \sum_{b \in B} n_b \cdot b \in \text{Div}_0(T), \quad A, B \subset T, A \cap B = \emptyset, n_a, n_b > 0,$$

defines a point

$$p := \left(\bigoplus_{a \in A} n_a \cdot a \right) \ominus \left(\bigoplus_{b \in B} n_b \cdot b \right) \in T$$

Then

$$\rho(p) = 1 \cdot p - 1 \cdot 0_T \in \text{Div}_0(T)$$

Due to Abel's Theorem the difference

$$\rho(p) - D = 1 \left(\left(\bigoplus_{a \in A} n_a \cdot a \right) \ominus \left(\bigoplus_{b \in B} n_b \cdot b \right) \right) - 1 \cdot 0_T - \left(\sum_{a \in A} n_a \cdot a - \sum_{b \in B} n_b \cdot b \right) \in \text{Div}_0(T)$$

is a principal divisor because $\ominus 0_T = 0_T \in \Lambda$. Hence ρ is surjective.

iv) Part ii) and iii) imply that ρ is an isomorphism of groups.

2. Due to part 1) and Definition 4.31 the group structures respectively on domain and codomain of

$$\phi : T \rightarrow E(\mathbb{C})$$

are determined by the divisor class groups. Because ϕ is an analytical isomorphism it is a group isomorphism.

□

Remark 4.34 (Group structure of an elliptic curve).

1. *Geometric version:* The group structure from Definition 4.31 of an elliptic curve $(E/k, O)$ can also be obtained by a geometric construction. It relies on the fact that a cubic and a line in $\mathbb{P}^2$ intersect in exactly three points, counted

with multiplicity. Nevertheless, from a geometric point of view the construction seems rather exotic.

See Figure 4.11: To add two points $p, q \in E \setminus \{O\}$ consider the line $L(p, q)$ passing through p and q . The line intersects E in a third point $r \in E$. The second line $L(r, O)$ intersects E in a third point, which is $p \oplus q$. For more details see [69].

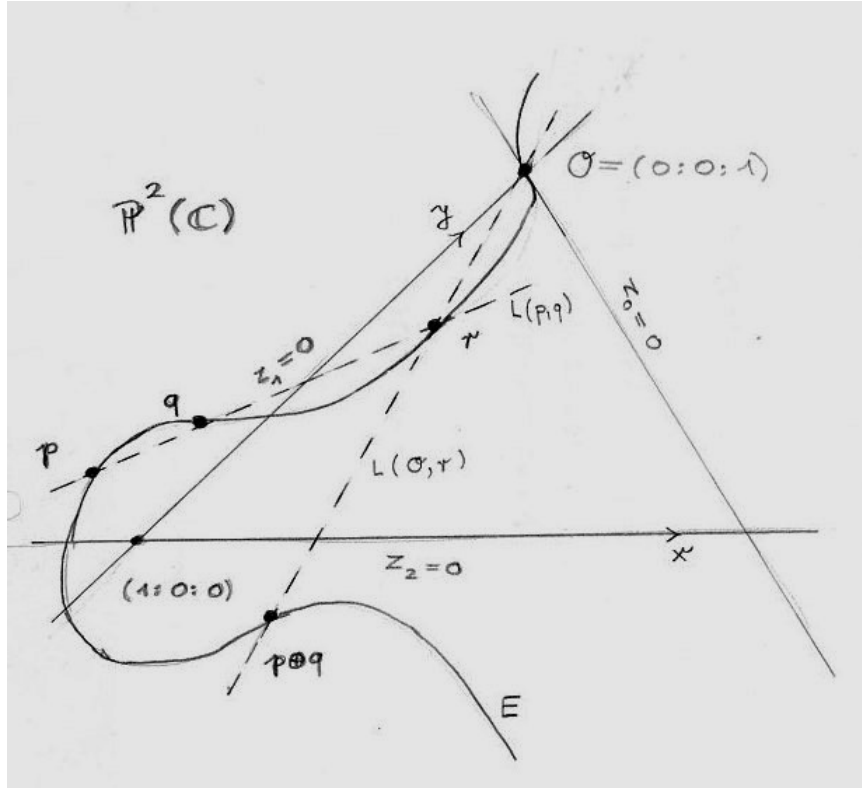


Fig. 4.11 Group law on an elliptic curve (Geometric construction)

2. *Mordells theorem*: The group structure on $E(\overline{\mathbb{Q}})$ provides also the set $E(K)$ of K -valued points with a group structure for any intermediate field

$$\mathbb{Q} \subset K \subset \overline{\mathbb{Q}}.$$

Mordell proved: The Abelian group $E(\mathbb{Q})$ of $\mathbb{Q}$ -valued points of E is finitely generated, [62, Chap. VIII, Theor. 4.1]. As a consequence, $E(\mathbb{Q})$, today names the *Mordell-Weil group* of $E/\mathbb{Q}$, decomposes as the product of an Abelian torsion group and a free Abelian group of finite rank $r < \infty$

$$E(\mathbb{Q}) = E(\mathbb{Q})_{tors} \times \mathbb{Z}^r.$$

Many open questions are related to the rank r .

3. *Falting's theorem*: Elliptic curves have genus $g = 1$. Curves of genus $g > 1$ do no longer carry a group structure. Mordell conjectured: Projective algebraic curves $C/\mathbb{Q}$ with genus $g > 1$ have only finitely many points with rational coordinates, i.e.

$$\text{card } C(\mathbb{Q}) < \infty.$$

The Mordell-conjecture was proved by Faltings, see [64].

Example 4.35 (Group structure of an elliptic curve). Consider an elliptic curve $E/\mathbb{Q}$. The PARI file `Elliptic_curve_torsion_group` investigates the torsion group of $E(\mathbb{Q})$ for the elliptic curve $E/\mathbb{Q}$ with Weierstrass polynomial

$$y^2 = x^3 - 43x + 166$$

see Figure 4.12. It shows that the torsion group $E(\mathbb{Q})_{tors}$ has a factor isomorphic to $\mathbb{Z}/7\mathbb{Z}$. For more information, in particular for the equality

$$E(\mathbb{Q})_{tors} = \mathbb{Z}/7\mathbb{Z}$$

see [60, Chap. VII, Theor. 7.5]. It is conjectured that the rank r of the free part of the group $E(\mathbb{Q})$ equals the analytic rank of $E/\mathbb{Q}$.

```

=====
Elliptic_curve_torsion_group: Start.

Weierstrass polynom.:  $Y^2 = X^3 + (-43X) + (166)$ 
j-invar.: [-1, 1; 2, -7; 3, 3; 13, -1; 43, 3]
discr.: [-1, 1; 2, 19; 13, 1], analytic. rank: 0

[3, 8] is on curve? 1
[3, 8] is torsion point of order 7
2 * [3, 8]: [-5, -16]
-----
[3, -8] is on curve? 1
[3, -8] is torsion point of order 7
2 * [3, -8]: [-5, 16]
-----
[-5, 16] is on curve? 1
[-5, 16] is torsion point of order 7
2 * [-5, 16]: [11, -32]
-----
[-5, -16] is on curve? 1
[-5, -16] is torsion point of order 7
2 * [-5, -16]: [11, 32]
-----
[11, 32] is on curve? 1
[11, 32] is torsion point of order 7
2 * [11, 32]: [3, 8]
-----
[11, -32] is on curve? 1
[11, -32] is torsion point of order 7
2 * [11, -32]: [3, -8]
-----
Elliptic_curve_torsion_group: End

```

Fig. 4.12 The group $E(\mathbb{Q})_{tors}$

4.3 Elliptic curves over finite fields

The present section starts the investigation of elliptic curves defined over $\mathbb{Q}$. Each elliptic curve $E/\mathbb{Q}$ can also be defined by a distinguished Weierstrass polynomial with integer coefficients, called a *global minimal polynomial*, see Definition 4.36 and Remark 4.37.

Entering into the field of arithmetic geometry one may reduce the integer equation modulo any prime $p \in \mathbb{Z}$, each time obtaining an elliptic curve defined over a finite field $\mathbb{F}_p$. Considered from a general point of view: An elliptic curve defined over $\mathbb{Q}$ is a family of not necessarily smooth curves defined over all prime fields of prime characteristic. Or even more general, it is a map to $\text{Spec } \mathbb{Z}$, and all but finitely many fibres are elliptic curves over prime fields. Reduction modulo p is a powerful tool to investigate the arithmetic information encoded in an elliptic curve defined over $\mathbb{Q}$.

Before investigating the family of elliptic curves one has to eliminate redundant prime powers of the discriminant which otherwise would distort the result. The discriminants of the different Weierstrass polynomials of an elliptic curve differ by a non-zero rational. After multiplying with a principal nominator they differ by a non-zero integer. Hence the reduction of the discriminants modulo a prime $p \in \mathbb{Z}$ may differ. To generalize the discriminant criterion from Proposition 4.26 one has to find a unique minimal discriminant. The corresponding Weierstrass polynomial is named *global minimal*.

Definition 4.36 (Global minimal Weierstrass polynomial). For an elliptic curve $E/\mathbb{Q}$ a Weierstrass polynomial $F \in \mathbb{Z}[X, Y]$ with integer coefficients is *global minimal* if its discriminant

$$\Delta_F = \pm \prod_{p \text{ prime}} p^{v(p)}, \quad v(p) \in \mathbb{N}^*,$$

is minimal in the following sense: The elliptic curve $E/\mathbb{Q}$ cannot be defined by a Weierstrass polynomial $F'(X, Y) \in \mathbb{Z}[X, Y]$ with discriminant $\Delta_{F'}$ satisfying for all primes $p \in \mathbb{Z}$

$$\text{ord}_p(\Delta_{F'}) \leq \text{ord}_p(\Delta_F)$$

and for at least one p_0

$$\text{ord}_{p_0}(\Delta_{F'}) < \text{ord}_{p_0}(\Delta_F)$$

“Global minimality” is a global property because it refers to all primes of Δ_F

Remark 4.37 (Global minimal Weierstrass polynomial).

1. *Non-uniqueness of the discriminant:* The Weierstrass polynomial of an elliptic curve, even if is in short form, is not uniquely determined. In general, an elliptic curve does not have a unique discriminant. It is not the discriminant, but the j -invariant which characterizes the elliptic curve.
2. *Global minimal Weierstrass polynomial:* The discriminant of a global minimal Weierstrass polynomial

$$F(X, Y) \in \mathbb{Z}[X, Y]$$

is uniquely determined. The last condition from Definition 4.36 implies for any prime factor $p \neq 2, 3$ of its discriminant

$$\text{ord}_p(\Delta_F) < 12.$$

Each elliptic curve over $k = \mathbb{Q}$ has a global minimal Weierstrass polynomial. The result generalizes to elliptic curves over arbitrary number fields with class number $= 1$, see [62, Chap. VIII, Cor. 8.3] and more specific [35, Theor. 10.3]. A global minimal Weierstrass polynomial is not necessarily in short form.

3. *PARI command:* The PARI command *ellminimalmodel* provides the global minimal model for elliptic curves over number fields k with class number $= 1$.

Example 4.38 (Global minimal Weierstrass polynomial).

1. See PARI-file `Elliptic_curve_weierstrass_equation_12`. Figure 4.13, upper example, considers the Weierstrass polynomial

$$F(X, Y) = Y^2 - (X^3 + 15.625)$$

Its discriminant

$$\Delta_F = -2^4 \cdot 3^3 \cdot 5^{12}$$

is not minimal because

$$5^{12} \mid \Delta_F$$

The global minimal Weierstrass polynomial is

$$F_{\min}(X, Y) = Y^2 - (X^3 + 1)$$

with discriminant

$$\Delta_{F_{\min}} = -2^4 \cdot 3^3.$$

```

=====
Elliptic curve weierstrass_equation_12 = Start.
Elliptic curve parameter = [0, 0, 0, 0, 15625]
Elliptic curve, Weierstrass equation: Y^2 + 0*XY + 0*Y = X^3 + 0*X^2 + 0*X + 15625
Discriminant = -105468750000 = [-1, 1; 2, 4; 3, 3; 5, 12]

Elliptic curve, global minimal Weierstrass equation: Y^2 + 0*XY + 0*Y = X^3 + 0*X^2 + 0*X + 1
Discriminant = -432 = [-1, 1; 2, 4; 3, 3]
=====
Elliptic curve parameter = [0, 0, 0, 1/5, 1]
Elliptic curve, Weierstrass equation: Y^2 + 0*XY + 0*Y = X^3 + 0*X^2 + 1/5*X + 1
Discriminant = -54064/125 = [-1, 1; 2, 4; 5, -3; 31, 1; 109, 1]

Elliptic curve, global minimal Weierstrass equation: Y^2 + 0*XY + 0*Y = X^3 + 0*X^2 + 125*X + 15625
Discriminant = -105593750000 = [-1, 1; 2, 4; 5, 9; 31, 1; 109, 1]
=====
Elliptic curve weierstrass_equation_12 = End.
=====

```

Fig. 4.13 Global minimal Weierstrass polynomial

2. The Weierstrass polynomial

$$F(X, Y) = Y^2 - (X^3 - 595 \cdot X + 5.586) \in \mathbb{Z}[X, Y]$$

with discriminant

$$\Delta_F = 2^{12} \cdot 7^3$$

is global minimal, despite of the exponent 12 of the particular prime factor $p = 2$.

3. Consider the elliptic curve $E/\mathbb{Q}$ with Weierstrass polynomial in short form

$$F(X, Y) = Y^2 - (X^3 + \frac{1}{5} \cdot X + 1) \in \mathbb{Q}[X, Y]$$

which is not global minimal because it has a non-integer coefficient. To obtain a global minimal Weierstrass polynomial, we make the transformation

$$\begin{pmatrix} X \\ Y \end{pmatrix} = \begin{pmatrix} u^{-2} \cdot X' \\ u^{-3} \cdot Y' \end{pmatrix} \text{ with } u := 5$$

and multiply by u^6 . We obtain the Weierstrass equation

$$Y'^2 = X'^3 + \frac{1}{u} \cdot u^{-2} \cdot u^6 \cdot X' + u^6 \text{ i.e. } Y'^2 = X'^3 + 5^3 \cdot X' + 5^6$$

Hence $E/\mathbb{Q}$ is also defined by the Weierstrass polynomial

$$F_{\min}(X, Y) = Y^2 - (X^3 + 5^3 \cdot X + 5^6) \in \mathbb{Z}[X, Y]$$

which is global minimal because

$$\Delta_{F_{\min}} = -2^4 \cdot 5^9 \cdot 31 \cdot 109 \in \mathbb{Z},$$

see Figure 4.13, lower example.

Definition 4.39 (Reduction mod p). Consider an elliptic curve

$$E = (E/\mathbb{Q}, O)$$

with global minimal Weierstrass polynomial $F(X, Y) \in \mathbb{Z}[X, Y]$.

1. Denote by

$$F_{hom} \in \mathbb{Z}[X_0, X_1, X_2]$$

the homogenization of F . For each prime $p \in \mathbb{N}$ reducing mod p the coefficients of F_{hom} provides a Weierstrass polynomial

$$F_{hom,p} \in \mathbb{F}_p[X_0, X_1, X_2]$$

with coefficients from the finite field $\mathbb{F}_p$. The Weierstrass polynomial defines the projective variety

$$E_p := \text{Var}(F_{hom,p}) := \{(x_0 : x_1 : x_2) \in \mathbb{P}^2(\overline{\mathbb{F}}_p) : F_{hom,p}(x_0, x_1, x_2) = 0\} \subset \mathbb{P}^2(\overline{\mathbb{F}}_p),$$

which is named the *reduction mod p of E* .

2. At a prime $p \in \mathbb{Z}$ the elliptic curve E is

- *stable* if (E_p, O) is non-singular, i.e. an elliptic curve over $\mathbb{F}_p$.
- *semistable* if (E_p, O) is singular with a node.
- *unstable* if (E_p, O) is singular with a cusp.

3. The *conductor* $N_{E/\mathbb{Q}} \in \mathbb{N}^*$ of $E/\mathbb{Q}$ is the integer

$$N_{E/\mathbb{Q}} := \prod_{p|\Delta} p^{e_p}$$

with the primes p of the global minimal discriminant Δ with singular reduction E_p . The corresponding prime powers are

$$e_p = 1 \iff E_p \text{ is semistable,}$$

$$e_p = 2 \iff p \geq 5 \text{ and } E_p \text{ is unstable,}$$

$$e_p \geq 2 \iff p = 2 \text{ or } p = 3 \text{ and } E_p \text{ is unstable}$$

For the exact value in the cases $p = 2$ and $p = 3$ see [60, App. C, Prop. 16.2].

According to Proposition 4.26 the distinction between stable and not-stable at a prime p depends on the reduction of the minimal discriminant. The following example 4.40 shows why it is necessary to study elliptic curves by considering global minimal Weierstrass polynomials, not just arbitrary Weierstrass polynomials.

Example 4.40 (Reduction mod p of a global minimal Weierstrass polynomial). Consider the elliptic curve $(E/\mathbb{Q}, O)$ from Example 4.38 defined by the inhomogeneous Weierstrass polynomial

$$F(X, Y) = Y^2 - (X^3 + 15.625) \in \mathbb{Q}[X, Y]$$

with discriminant

$$\Delta_F = -2^4 \cdot 3^3 \cdot 5^{12} \in \mathbb{Z}.$$

Note $15.625 = 5^6$.

i) The Weierstrass polynomial $F \in \mathbb{Z}[X, Y]$ is in short form but it is not global minimal. Reducing its coefficients *mod* $p = 5$ gives the Weierstrass polynomial in short form

$$F_5(X, Y) = Y^2 - X^3 \in \mathbb{F}_5[X, Y]$$

with discriminant

$$\Delta_{F_5} = 0 \in \mathbb{F}_5.$$

Hence the Weierstrass polynomial F_5 defines a singular cubic over $\mathbb{F}_5$ with a cusp. Note that the discriminant Δ_F contains the prime factor 5^{12} .

ii) Therefore we have to apply first the coordinate transformation with the matrix

$$\begin{pmatrix} 1/u^2 & 0 \\ 0 & 1/u^3 \end{pmatrix}, \quad u := 5,$$

which transforms $F(X, Y)$ to the global minimal Weierstrass polynomial

$$F'(X', Y') = Y'^2 - (X'^3 + 1)$$

with discriminant

$$\Delta_{F'} = -2^4 \cdot 3^3 \in \mathbb{Z}.$$

Reducing the coefficients of F' *mod* $p = 5$ creates the Weierstrass polynomial

$$F'_5(X', Y') = Y'^2 - (X'^3 + 1) \in \mathbb{F}_5[X', Y']$$

with discriminant

$$\Delta_{F'_5} = 3 \in \mathbb{F}_5.$$

The reduction of E *mod* $p = 5$ defines an elliptic curve E_5 over $\mathbb{F}_5$.

Remark 4.41 (Hasse estimate).

1. Important properties of an elliptic curve $E/\mathbb{Q}$ are encoded in the function, which attaches to each prime p with stable E_p the family

$$E_p(\mathbb{F}_{p^n}), n \in \mathbb{N}^*,$$

of groups of $\mathbb{F}_{p^n}$ -valued points of the elliptic curve E_p . These groups collect all points from $E_p \subset \mathbb{P}^2(\overline{\mathbb{F}}_p)$ having all coordinates in the finite fields $\mathbb{F}_{p^n}$.

We recall the following result of Hasse for the case $n = 1$:

$$|(1 + p) - \text{card } E_p(\mathbb{F}_p)| \leq 2 \cdot \sqrt{p}.$$

For a proof see [35, Theor. 10.5].

Hasse's result was conjectured by Artin. The value

$$1 + p$$

for comparison is motivated as follows: The field $\mathbb{F}_p$ has p elements. A Weierstrass equation

$$Y^2 = X^3 + AX + B$$

is quadratic in Y . Hence for each $x \in \mathbb{F}_p$ there exist at most two solutions $(x, y) \in \mathbb{A}^2(\mathbb{F}_p)$. Taking into account the point at infinity $O \in E_p(\mathbb{F}_p)$ gives the estimate

$$\text{card } E_p(\mathbb{F}_p) \leq (1 + 2p).$$

In the average, i.e. by choosing the Weierstrass polynomial at random, one could expect a 50% chance to find a solution (x, y) at all. Therefore $(1 + p)$ is a plausible average for $\text{card } E_p(\mathbb{F}_p)$.

One defines for any prime p the difference

$$a_p(E) := (1 + p) - \text{card } E_p(\mathbb{F}_p)$$

between the average number and the actual number of $\mathbb{F}_p$ -valued points of $E/\mathbb{Q}$.

2. See the PARI-file `Elliptic_curve_Hasse_estimate_05` for the numerical example of the Hasse inequality displayed in Figure 4.14.

```

=====
Elliptic curve: y^2 = x^3 + 0*x + (-17)
Discriminant: -124848 = [-1, 1; 2, 4; 3, 3; 17, 2]
j(E): 0
Number of points in E[F_p]:
-----
card E[F_2] = 2, |a_p| = 1, 2 * sqrt(2) = 2.828427125
card E[F_3] = 3, |a_p| = 1, 2 * sqrt(3) = 3.464101615
card E[F_5] = 6, |a_p| = 0, 2 * sqrt(5) = 4.472135955
card E[F_7] = 3, |a_p| = 5, 2 * sqrt(7) = 5.291502622
card E[F_11] = 12, |a_p| = 0, 2 * sqrt(11) = 6.633249581
card E[F_13] = 21, |a_p| = 7, 2 * sqrt(13) = 7.211102551
card E[F_17] = 17, |a_p| = 1, 2 * sqrt(17) = 8.246211251
card E[F_19] = 13, |a_p| = 7, 2 * sqrt(19) = 8.717797887
card E[F_23] = 24, |a_p| = 0, 2 * sqrt(23) = 9.591663047
card E[F_29] = 30, |a_p| = 0, 2 * sqrt(29) = 10.77032961
card E[F_31] = 21, |a_p| = 11, 2 * sqrt(31) = 11.13552873
card E[F_37] = 49, |a_p| = 11, 2 * sqrt(37) = 12.16552506
card E[F_41] = 42, |a_p| = 0, 2 * sqrt(41) = 12.80624847
card E[F_43] = 31, |a_p| = 13, 2 * sqrt(43) = 13.11487705
card E[F_47] = 48, |a_p| = 0, 2 * sqrt(47) = 13.71130920
card E[F_53] = 54, |a_p| = 0, 2 * sqrt(53) = 14.56021978
card E[F_59] = 60, |a_p| = 0, 2 * sqrt(59) = 15.36229150
card E[F_61] = 49, |a_p| = 13, 2 * sqrt(61) = 15.62049935
card E[F_67] = 73, |a_p| = 5, 2 * sqrt(67) = 16.37070554
card E[F_71] = 72, |a_p| = 0, 2 * sqrt(71) = 16.85229955
card E[F_73] = 64, |a_p| = 10, 2 * sqrt(73) = 17.08800749
card E[F_79] = 84, |a_p| = 4, 2 * sqrt(79) = 17.77638883
card E[F_83] = 84, |a_p| = 0, 2 * sqrt(83) = 18.22086716
card E[F_89] = 90, |a_p| = 0, 2 * sqrt(89) = 18.86796226
card E[F_97] = 103, |a_p| = 5, 2 * sqrt(97) = 19.69771560
card E[F_101] = 102, |a_p| = 0, 2 * sqrt(101) = 20.09975124

```

Fig. 4.14 The Hasse estimate

3. The sequence

$$(a_p(E))_{p \text{ prime}}$$

will turn out to be a fundamental characteristic of the elliptic curve $E/\mathbb{Q}$, see Remark 4.44. These numbers determine also the cardinalities

$$\text{card } E_p(\mathbb{F}_{p^n}) = (1 + p) - a_p(E)$$

for $n \geq 2$. The latter cardinalities will serve as the generators of the ζ -function of $E/\mathbb{Q}$, see Definition 4.43.

Historically, the first example of a ζ -function is due to by L. Euler. Later this function was named *Riemann ζ -function*.

Proposition 4.42 (ζ -function and Euler product). *The Riemann ζ -function*

$$\zeta : \{s \in \mathbb{C} : \operatorname{Re} s > 1\} \rightarrow \mathbb{C}, \quad \zeta(s) := \sum_{n=1}^{\infty} \frac{1}{n^s}$$

is holomorphic.

1. Meromorphic extension: *The ζ -function extends to a meromorphic function on $\mathbb{C}$ with a single pole at $s = 1$. The pole has order = 1 and residue = 1.*
2. Functional equation: *It satisfies the functional equation with respect to reflection at the point $s = 1$*

$$\pi^{-s/2} \cdot \Gamma(s/2) \cdot \zeta(s) = \pi^{-(1-s)/2} \cdot \Gamma((1-s)/2) \cdot \zeta(1-s)$$

with the Γ -function,

3. Product representation: *It has the product representation*

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

with the product taken over all primes p .

The Euler-product formula from Proposition 4.42

$$\sum_{n=1}^{\infty} \frac{1}{n^s} = \prod_p \frac{1}{1 - p^{-s}}$$

illustrates the importance of the Riemann ζ -function for investigating the distribution of prime numbers.

Proof. 1. See [70, Teil I, § 4].

2. See reference from part 1).

3. For each $N \in \mathbb{N}$ expand the finite product by using the geometric series and the unique prime factor decomposition of integers

$$\prod_{p < N} \frac{1}{1 - p^{-s}} = \prod_{p < N} \left(1 + \frac{1}{p^s} + \frac{1}{p^{2s}} + \dots \right) = \sum' \frac{1}{n^s}$$

with the sum on the right-hand side taken over all positive integers with all prime factors less than N . The limit $N \rightarrow \infty$ proves the claim.

□

From the viewpoint of the ring $\mathbb{Z}$ the Riemann ζ -function is a global object. Because it considers all primes $p \in \mathbb{Z}$ and collects as a product the local information encoded in the geometric series of each p

$$\frac{1}{1-p^{-s}} = \sum_{n=0}^{\infty} (1/p^{-s})^n$$

We now switch from the primes p of $\mathbb{Z}$ to the elliptic curves $E_p/\mathbb{F}_p$, which are the fibres of a given elliptic curve $E/\mathbb{Q}$ over its stable primes p . First we study the local information, i.e. the information encoded in $E_p/\mathbb{F}_p$. The curve E_p contributes as local information at the prime p the sequence of cardinalities of points with values in the finite fields of characteristic p

$$\text{card } E_p(\mathbb{F}_{p^n}), n \in \mathbb{N}^*.$$

Definition 4.43 (Zeta-function of an elliptic curve over $\mathbb{F}_p$). Consider an elliptic curve $E_p/\mathbb{F}_p$. The ζ -function of E_p is the formal power series with generator derived from the sequence $(\text{card } E_p(\mathbb{F}_{p^n}))_{n \in \mathbb{N}^*}$

$$Z(E_p, T) := \exp \left(\sum_{n=1}^{\infty} \text{card } E_p(\mathbb{F}_{p^n}) \cdot \frac{T^n}{n} \right) \in \mathbb{Q}[[T]].$$

One checks by formal derivation

$$T \cdot \frac{d}{dT} \log Z(E_p, T) = T \cdot \frac{Z'}{Z} = T \cdot \sum_{n=1}^{\infty} (\text{card } E_p(\mathbb{F}_{p^n}) \cdot T^{n-1}) = \sum_{n=1}^{\infty} (\text{card } E_p(\mathbb{F}_{p^n}) \cdot T^n)$$

Remark 4.44 (Properties of the ζ -function of an elliptic curve over $\mathbb{F}_p$). The ζ -function of an elliptic curve $E_p/\mathbb{F}_p$

$$Z(E_p, T) \in \mathbb{Q}[[T]]$$

has the following properties:

1. *Rationality:* The ζ -function is a rational function over the ring $\mathbb{Z}$, more precisely:

$$Z(E_p, T) = \frac{L(E_p, T)}{(1-T) \cdot (1-p \cdot T)} \in \mathbb{Z}(T)$$

with the quadratic polynomial

$$L(E_p, T) = 1 - a_p(E) \cdot T + p \cdot T^2, \quad a_p(E) := (1+p) - \text{card } E_p(\mathbb{F}_p).$$

2. *Functional equation*: The ζ -function satisfies the functional equation

$$Z\left(E_p, \frac{1}{p} \cdot T\right) = Z(E_p, T).$$

3. *Analogue of the Riemann hypothesis*: Both complex zeros $\alpha, \beta \in \mathbb{C}$ of $L(E_p, T)$ have modulus

$$|\alpha| = |\beta| = \sqrt{p}.$$

The proof of these properties is due to Weil, [62, Chap. V, Theor. 2.2].

Weil also stated and proved a generalization for non-singular projective algebraic curves $C_p/\mathbb{F}_p$ of arbitrary genus. If the curve C_p is the reduction mod p of a curve

$$C/\mathbb{Q}$$

then Weil's theorem shows the close relation between the curve C_p and the complex curve $X(C) \subset \mathbb{P}^n$: The degree of the polynomial in the numerator of the ζ -function of C_p is the first Betti number of the compact Riemann surface $X/\mathbb{C}$.

Moreover, Weil made a conjecture for the general case of higher-dimensional non-singular projective algebraic varieties. Deligne proved the Weil conjecture. For an excellent introduction to the whole subject see [31, Appendix C].

The L -series $L(E)$ of an elliptic curve $E/\mathbb{Q}$ collects for each prime $p \in \mathbb{Z}$ the local information of the non-singular reductions E_p . Hence the L -series is a *global* object.

Definition 4.45 (L-series of an elliptic curve over $\mathbb{Q}$). The L -series of an elliptic curve $E/\mathbb{Q}$ is the function

$$L(E, -) : \{s \in \mathbb{C} : \operatorname{Re} s > 3/2\} \rightarrow \mathbb{C}, \quad L(E, s) := \prod_{p \text{ prime}} \frac{1}{Z(E_p, p^{-s})}$$

Remark 4.46 (L-series of an elliptic curve $E/\mathbb{Q}$).

1. The convergence of the L -series and its representation as an absolute convergent Dirichlet series with respect to the variable s follows from Hasse's estimate in Remark 4.41 and the theory of Euler products, see [35, Chap. X, Cor. 10.6]. Hence $L(E, -)$ is a holomorphic function.

Due to the proof of the Shimura-Taniyama-Weil conjecture the L -series of an elliptic curve extends holomorphically to the whole complex plane. There are

several unsolved conjectures which arithmetic information of an elliptic curve is encoded in the L -series. E.g. Birch and Swinnerton-Dyer conjectured

$$\text{rank } E(\mathbb{Q}) = r$$

with r equal to the *analytic rank* defined as the order of the zero of the L -series at $s = 1$

$$r := \text{ord}(L(E); 1).$$

The PARI script `Elliptic_curve_04_2` computes the analytic rank for some elliptic curves $E/\mathbb{Q}$ of higher analytic rank, see Figure 4.15.

```

=====
Elliptic_curve_04: Start.

Min. Weierstrass polynom.: Y^2 = X^3 + (7*X)
Analytic rank: 0

Min. Weierstrass polynom.: Y^2 = X^3 + (3*X)
Analytic rank: 1

Min. Weierstrass polynom.: Y^2 = X^3 + (73*X)
Analytic rank: 2

Min. Weierstrass polynom.: Y^2 = X^3 + (89*X)
Analytic rank: 2

Min. Weierstrass polynom.: Y^2 = X^3 + (-82*X)
Analytic rank: 3

Min. Weierstrass polynom.: Y^2 = X^3 + (-X^2) + (-24649*X) + (1355209)
Analytic rank: 4

Min. Weierstrass polynom.: Y^2 = X^3 + (-856967076*X)
Analytic rank: 4

Min. Weierstrass polynom.: Y^2 = X^3 + (-X^2) + (-2069247973*X) + (36191779888342)
Analytic rank: 6

Elliptic_curve_04: End
=====

```

Fig. 4.15 Analytic rank of several elliptic curves $E/\mathbb{Q}$

For more results and conjectures see [62, App. C, § 16].

2. The two polynomials in the denominator of the ζ -function $Z(E_p, T)$ evaluate at $T = p^{-s}$ as values of the Riemann ζ -function

$$\frac{1}{1 - p^{-s}} = \zeta(s) \text{ and } \frac{1}{1 - p^{1-s}} = \zeta(s-1)$$

They do not provide new information about the elliptic curve.

3. If one compares the L -series $L(E, s)$ of an elliptic curve $E/\mathbb{Q}$ with the Riemann ζ -function $\zeta(s)$ on the level of the two product representations

$$L(E, s) = \prod_p \frac{1}{Z(E_p, p^{-s})} \text{ and } \zeta(s) = \prod_p \frac{1}{1 - p^{-s}}$$

then for each given prime p the local factor

$$Z(E_p, p^{-s}) \text{ corresponds to } \frac{1}{1 - p^{-s}}$$

The geometric series on the right hand side depends only on the prime p , while the corresponding ζ -function on the left-hand side depends on p and on the reduction of the given elliptic curve $E/\mathbb{Q}$.

Part II

Advanced Theory

Chapter 5

Introduction to Hecke theory and applications

Hecke theory investigates the modular algebra $M_*(\Gamma)$ of modular forms by studying the family

$$(T_m)_{m \in \mathbb{N}^*}$$

of Hecke operators, which are endomorphisms of the vector spaces $M_k(\Gamma)$, $k \in \mathbb{N}$, of modular forms of weight k . After introducing a Hermitian scalar product on the subspaces of cusp forms

$$S_k(N) \subset M_k(\Gamma_0(N))$$

the Hecke operators become self-adjoint operators on $S_*(\Gamma)$.

We recall the general notations for respectively the modular group and the modular algebra

$$\Gamma := SL(2, \mathbb{Z}) \text{ and } M_*(\Gamma) := \bigoplus_{k \in \mathbb{N}} M_k(\Gamma)$$

5.1 Hecke operators of the modular group and their eigenforms

The first part of this section follows Serre's approach to derive the Hecke operators from the corresponding *lattice functions*, i.e. to deal first with lattices instead with modular functions. In my opinion the approach via lattice functions is more transparent and needs less prerequisites than the direct approach. After showing the equivalence of functions on lattices and modular forms the results about lattice functions easily carry over to modular forms and provide the classical Hecke operators and their properties.

Notation 5.1 (Lattice and lattice basis). We denote by $\mathcal{L}$ the free Abelian group with generators the lattices $\Lambda \subset \mathbb{C}$ and by $\Lambda(\omega_1, \omega_2)$ the lattice with $\mathbb{Z}$ -basis (ω_1, ω_2) satisfying $\det(\omega_1, \omega_2) > 0$.

Definition 5.2 (Homogeneous lattice function). For $k \in \mathbb{N}^*$ a lattice function

$$F : \mathcal{L} \rightarrow \mathbb{C}$$

is *homogeneous of degree $-k$* if it satisfies for all lattices $\Lambda(\omega_1, \omega_2)$

$$F(\Lambda(\lambda \cdot \omega_1, \lambda \cdot \omega_2)) = \frac{1}{\lambda^k} \cdot F(\Lambda(\omega_1, \omega_2))$$

Lemma 5.3 (Lattice functions and modular forms). *The weakly modular functions of even weight $k \in \mathbb{N}^*$*

$$f : \mathbb{H} \rightarrow \mathbb{C}$$

correspond bijectively to the homogeneous lattice functions of degree $-k$

$$F : \mathcal{L} \rightarrow \mathbb{C}$$

using the conversion

$$F(\Lambda(\omega_1, \omega_2)) = \frac{1}{\omega_2^k} \cdot f\left(\frac{\omega_1}{\omega_2}\right), \det(\omega_1 \ \omega_2) > 0$$

Proposition 5.4 (Isomorphism of Γ -modules). *For even $k \in \mathbb{N}$ denote by $W_k(\Gamma)$ the vector space of weakly modular functions of weight k and by $\mathcal{L}^\vee(-k)$ the vector space of homogeneous lattice functions of degree $-k$.*

1. Weakly modular functions: *The left Γ -action*

$$\Gamma \times W_k(\Gamma) \rightarrow W_k(\Gamma), (\gamma, f) \mapsto f^\gamma,$$

with

$$f^\gamma(\tau) := f(\gamma(\tau)), \tau \in \mathbb{H},$$

introduces a Γ -module structure on $W_k(\Gamma)$.

2. Homogeneous lattice functions: *The left Γ -action*

$$\Gamma \times \mathcal{L}^\vee(-k) \rightarrow \mathcal{L}^\vee(-k), (\gamma, F) \mapsto F^\gamma,$$

with

$$F^\gamma(\Lambda(\tau, 1)) := F(\Lambda(\gamma(\tau), 1)), \tau \in \mathbb{H},$$

introduces a Γ -module structure on $\mathcal{L}^\vee(-k)$.

3. The bijective map from Lemma 5.3

$$W_k(\Gamma) \rightarrow \mathcal{L}^\vee(-k), f \mapsto F,$$

with

$$F(\Lambda(\omega_1, \omega_2)) = \frac{1}{\omega_2^k} \cdot f\left(\frac{\omega_1}{\omega_2}\right), \det(\omega_1 \ \omega_2) > 0$$

is an isomorphism of Γ -modules.

One easily checks that in Proposition 5.4 the resulting lattice function F is homogeneous of degree $-k$:

$$F(\Lambda(\lambda\omega_1, \lambda\omega_2)) := \frac{1}{(\lambda \cdot \omega_2)^k} \cdot f\left(\frac{\lambda\omega_1}{\lambda\omega_2}\right) = \frac{1}{\lambda^k} \cdot \frac{1}{\omega_2^k} \cdot f\left(\frac{\omega_1}{\omega_2}\right) = \frac{1}{\lambda^k} \cdot F(\Lambda(\omega_1, \omega_2))$$

Definition 5.5 is a preliminary definition of the Hecke operators. The operators are not defined on the vector spaces of modular forms but on the Abelian group $\mathcal{L}$ of all lattices in $\mathbb{C}$.

Definition 5.5 (Two lattice maps). We define on $\mathcal{L}$ two types of lattice maps which are endomorphisms of $\mathcal{L}$:

- *Homothetic:*

$$R_\mu : \mathcal{L} \rightarrow \mathcal{L}, \Lambda \mapsto \mu \cdot \Lambda, \mu \in \mathbb{C}^*$$

- *Sublattices:*

$$T_m : \mathcal{L} \rightarrow \mathcal{L}, \Lambda \mapsto \sum_{\substack{\Lambda' \subset \Lambda \\ [\Lambda : \Lambda'] = m}} \Lambda', m \in \mathbb{N}^*$$

A lattice Λ is *finer* than a lattice Λ' , if $\Lambda' \subset \Lambda$ is a sublattice. Like a finer sieve has more and smaller holes so a finer lattice has more and smaller fundamental parallelograms than a coarser lattice. Denote by $[\Lambda : \Lambda']$ the index of the subgroup $\Lambda' \subset \Lambda$. The lattice functions T_m from Definition 5.5 sum up all coarse lattices with fixed index which are contained in the finer lattice.

Before investigating the algebra generated by the maps R_μ and T_m in Proposition 5.10 we first have to determine all sublattices Λ' of a fixed index within a given lattice Λ . These sublattices form the index set in the sum of T_m in Definition 5.5. Therefore we add some remarks about integer matrices.

Definition 5.6 (Integer matrices). For $m \in \mathbb{N}^*$ we denote by

$$\Gamma_m := \left\{ A := \begin{pmatrix} a & b \\ c & d \end{pmatrix} : \det A = m \right\} \subset GL(2, \mathbb{Q})^+$$

the set of integer matrices with fixed invertible determinant $m \in \mathbb{N}^*$. The canonical Γ_1 -left action on Γ_m is the map

$$\Gamma_1 \times \Gamma_m \rightarrow \Gamma_m, (\gamma, \gamma_m) \mapsto \gamma \cdot \gamma_m$$

Proposition 5.7 (complete set of representations). *For the Γ_1 -left action from Definition 5.6 the set*

$$V(m) := \left\{ \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} \in M(2 \times 2, \mathbb{Z}) : 1 \leq a, 0 \leq b \leq d \right\}$$

is a complete rest system of representatives for the orbit space Γ/Γ_m . The cardinality is the sum-of-divisors function

$$\text{card } V(m) = \sigma_1(m) = \sum_{d|m} d$$

In particular for a prime p

$$V(p) = \left\{ \begin{pmatrix} p & 0 \\ 0 & 1 \end{pmatrix} \right\} \cup \left\{ \begin{pmatrix} 1 & b \\ 0 & p \end{pmatrix} : b = 0, 1, \dots, p-1 \right\}$$

and $\text{card } V(p) = p + 1$.

Proof. For the proof see [38, Kap. IV, §2 Satz]. We define a map

$$p : \Gamma_m \rightarrow V(m),$$

which attaches to each $M \in \Gamma_m$ a triangular matrix $p(M)$ belonging to the same orbit. We show that $p(M)$ does not depend on the specific point M but only on the orbit of M

1. *Range $V(m)$:* Consider a fixed but arbitrary element

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_m.$$

We construct a matrix

$$A = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \in \Gamma$$

such that

$$A \cdot M = \begin{pmatrix} \alpha a + \beta c & \alpha b + \beta d \\ \gamma a + \delta c & \gamma b + \delta d \end{pmatrix} \in V(m) :$$

First, in order to remove the entry c of M we choose coprime integers $\gamma, \delta \in \mathbb{Z}$ with

$$\gamma \cdot a + \delta \cdot c = 0.$$

In order to obtain a matrix with $\det = 1$ we choose integers $\alpha, \beta \in \mathbb{Z}$ with

$$\alpha \cdot \delta - \beta \cdot \gamma = 1.$$

Such choice is possible because γ and δ are coprime. Set

$$A_1 := \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \in \Gamma.$$

Then

$$A_1 \cdot M = \begin{pmatrix} a' & b' \\ 0 & d' \end{pmatrix}$$

and

$$\det(A_1 \cdot M) = \det M = a' \cdot d' = m.$$

In particular $d' \neq 0$ and possibly after replacing A_1 by $-A_1$ even $a', d' > 0$.

Secondly, because $T \in \Gamma$, for any $k \in \mathbb{Z}$ also

$$A_2(k) := T^k \cdot A_1 \in \Gamma.$$

Then

$$A_2(k) \cdot M = \begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} a' & b' \\ 0 & d' \end{pmatrix} = \begin{pmatrix} a' & b' + k \cdot d' \\ 0 & d' \end{pmatrix}.$$

We choose $k_0 \in \mathbb{Z}$ such that

$$0 \leq b' + k_0 \cdot d' < d'$$

and set

$$A := A_2(k_0).$$

2. *Dependence on the orbit:* The construction from the previous part determines uniquely the element

$$A \cdot M \in V(m) :$$

Assume two matrices $A_i \in \Gamma$ such that

$$A_i \cdot M =: M_i \in V(m), \quad i = 1, 2.$$

Then

$$M = A_1^{-1} \cdot M_1 = A_2^{-1} \cdot M_2 \text{ or } A_{21} \cdot M_1 = M_2, \quad A_{21} := A_2 \cdot A_1^{-1}.$$

Set

$$A_{21} =: \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \text{ and } M_i =: \begin{pmatrix} a_i & b_i \\ 0 & d_i \end{pmatrix}, \quad i = 1, 2.$$

From

$$\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \cdot \begin{pmatrix} a_1 & b_1 \\ 0 & d_1 \end{pmatrix} = \begin{pmatrix} a_2 & b_2 \\ 0 & d_2 \end{pmatrix}$$

follows:

- $\gamma \cdot a_1 = 0$, hence $\gamma = 0$.
- From $d_2 = \delta \cdot d_1$ follows $\delta > 0$. Together with $\alpha \cdot \delta = 1$ it implies $\alpha = \delta = 1$.
- As a consequence $a_1 = a_2$, $d_1 = d_2$, and $b_1 + \beta \cdot d_1 = b_2$.
- Because $0 \leq b_1, b_2 < d_1 = d_2$ we get $b_1 = b_2$ and $\beta = 0$.

As a consequence

$$A_{21} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \text{ or } A_1 = A_2.$$

3. *Definiton of p* : Part i) and ii) show that

$$p : \Gamma_m \rightarrow V(m), M \mapsto A \cdot M,$$

is well-defined, and $p(M)$ belongs to the same orbit as M .

4. *Surjectivity of p* : The map p is surjective because $V(m) \subset \Gamma_m$. In addition,

$$p(M_1) = p(M_2) \iff \exists A \in \Gamma \text{ with } A \cdot M_1 = M_2$$

$$\iff M_1 \text{ and } M_2 \text{ on the same orbit of } \Phi_m.$$

As a consequence p induces the bijective map

$$\Gamma \backslash \Gamma_m \xrightarrow{\cong} V(m), \bar{M} \mapsto p(M).$$

5. *Cardinality of the orbit space*: The equality

$$\text{card } V(m) = \sigma_m(1)$$

follows at once from the explicit representation of the elements of $V(m)$: Each coefficient $d \geq 0$ satisfies $d|m$ and allows d different coefficients b .

Proposition 5.8 relates these matrices to the sublattices of a given lattice.

Proposition 5.8 (Sublattices of a lattice). *Consider a lattice $\Lambda := \Lambda(\omega_1, \omega_2)$ in $\mathbb{C}$ and a sublattice*

$$\Lambda' := \mathbb{Z}\omega'_1 \oplus \mathbb{Z}\omega'_2 \subset \Lambda(\omega_1, \omega_2)$$

1. *Then exists a matrix $\gamma_m \in \Gamma_m$ with*

$$\begin{pmatrix} \omega'_1 \\ \omega'_2 \end{pmatrix} = \gamma \cdot \begin{pmatrix} \omega_1 \\ \omega_2 \end{pmatrix}$$

2. The matrix $\gamma \in \Gamma_m$ from part 1 is not unique. All matrices with the property from part 1 form the orbit of γ_m under the left Γ_1 -operation from Definition 5.6.
3. Consider a fixed lattice $\Lambda(\omega_1, \omega_2)$. For $m \in \mathbb{N}^*$ and

$$\gamma_m := \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in V(m)$$

define the sublattice

$$\Lambda(a\omega_1 + b\omega_2, d\omega_2) \subset \Lambda(\omega_1, \omega_2)$$

Then the map

$$\Gamma_m \rightarrow \{\Lambda' \subset \Lambda : [\Lambda : \Lambda'] = m\}, \gamma \mapsto \Lambda' := \Lambda(a\omega_1 + b\omega_2, d\omega_2)$$

is bijective.

4. Consider a lattice Λ .

- For all sublattices $\Lambda' \subset \Lambda$ of index $n \in \mathbb{N}$ holds

$$n \cdot \Lambda \subset \Lambda'$$

- For each pair of distinct sublattices $\Lambda_j \subset \Lambda$, $j = 1, 2$ of index a prime p holds

$$\Lambda_1 \cap \Lambda_2 = p \cdot \Lambda$$

Proof. For the Proof of step 1 - step 3 see [56, Chap. 5, §5.2, Lemma 2]. Step 4 follows from step 3:

i) Assume

$$\Lambda = \Lambda(\omega_1, \omega_2), \Lambda' = \Lambda(\omega'_1, \omega'_2)$$

According to step 3 exists

$$\gamma := \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} \in V(n)$$

with

$$\begin{pmatrix} \omega'_1 \\ \omega'_2 \end{pmatrix} = \gamma \cdot \begin{pmatrix} \omega_1 \\ \omega_2 \end{pmatrix}$$

We have to find a matrix $\delta \in \Gamma_m$ with

$$m \cdot \begin{pmatrix} \omega_1 \\ \omega_2 \end{pmatrix} = \delta \cdot \begin{pmatrix} \omega'_1 \\ \omega'_2 \end{pmatrix}$$

equivalently

$$m \cdot \begin{pmatrix} \omega_1 \\ \omega_2 \end{pmatrix} = \delta \cdot \gamma \cdot \begin{pmatrix} \omega_1 \\ \omega_2 \end{pmatrix}$$

equivalently

$$m \cdot \mathbb{1} = \delta \cdot \gamma$$

We define

$$\delta := (\det \gamma) \cdot \gamma^{-1} = \begin{pmatrix} d & -b \\ 0 & a \end{pmatrix} \in \Gamma_n$$

Because $\det \delta = n$ we have $[\Lambda' : \Lambda] = n$.

ii) We use the result from part i) to prove the second claim

$$p\Lambda = \Lambda'_1 \cap \Lambda'_2$$

The investigation of sublattices

$$\Lambda' \subset \Lambda \text{ with } p \cdot \Lambda \subset \Lambda'$$

becomes a matter of linear algebra after the reduction mod p : The sublattice $p\Lambda \subset \Lambda$ has index p^2 , hence as an Abelian group

$$\Lambda/p\Lambda \simeq (\mathbb{Z}/p\mathbb{Z})^2 = k^2$$

with $k = \mathbb{F}_p$ the prime field with characteristic p . Because $[\Lambda : p\Lambda] = p^2$ the quotient

$$W := \Lambda/p\Lambda$$

is a 2-dimensional k -vector space. The two lattices Λ'_j , $j = 1, 2$, induce two lines

$$L_j := \Lambda'_j/p\Lambda \subset W, \quad j = 1, 2$$

passing through the origin

$$\{0\} = p\Lambda/p\Lambda = L_1 \cap L_2$$

of W . Hence

$$p\Lambda = \Lambda'_1 \cap \Lambda'_2$$

□

For two matrices

$$M_j = \begin{pmatrix} a_j & b_j \\ 0 & d_j \end{pmatrix} \in V(n_j), \quad j = 1, 2$$

their product is

$$M_1 \cdot M_2 = \begin{pmatrix} a_1 a_2 & a_1 b_2 + b_1 d_2 \\ 0 & d_1 d_2 \end{pmatrix} \in \Gamma_{n_1 \cdot n_2}$$

We want to show that the product $V(n) \cdot V(m)$ is a complete set of representatives of the Γ -left action on $\Gamma_{n_1 \cdot n_2}$

The rest system of a natural number d is the set

$$\{0, 1, \dots, d-1\}$$

of the different rests modulo d .

Lemma 5.9 (Restsystems for $V(n_1 \cdot n_2)$). *For $j = 1, 2$ consider two coprime numbers*

$$n_j \in \mathbb{N}^*, j = 1, 2.$$

1. *For each factorization $n_j = a_j \cdot d_j$, $j = 1, 2$ the map of restsystems*

$$\phi : \{0, \dots, d_1-1\} \times \{0, \dots, d_2-1\} \rightarrow \{0, \dots, d_1 \cdot d_2-1\}, (b_1, b_2) \mapsto a_1 \cdot b_2 + b_1 \cdot d_2,$$

is bijective.

2. *The map on the product of complete sets of representatives*

$$V(n_1) \times V(n_2) \rightarrow \Gamma_{n_1 \cdot n_2}, (M_1, M_2) \mapsto M_1 \cdot M_2,$$

maps injectively onto a complete set of representatives of $\Gamma \backslash \Gamma_{n_1 \cdot n_2}$.

3.

$$V(n_1) \cdot V(n_2) = V(n_1 \cdot n_2)$$

Proof. 1. It suffices to show that ϕ is injective because its domain and codomain have the same finite cardinality. Therefore assume

$$(\tilde{b}_1, \tilde{b}_2) \in (\mathbb{Z}/d_1\mathbb{Z}) \times (\mathbb{Z}/d_2\mathbb{Z}) \text{ with } \phi((b_1, b_2)) = \phi((\tilde{b}_1, \tilde{b}_2))$$

i.e.

$$a_1 \cdot b_2 + b_1 \cdot d_2 \equiv a_1 \cdot \tilde{b}_2 + \tilde{b}_1 \cdot d_2 \pmod{d_1 d_2}$$

Then

$$a_1(b_2 - \tilde{b}_2) \equiv (\tilde{b}_1 - b_1)d_2 \pmod{d_1 d_2}$$

which implies

$$a_1(b_2 - \tilde{b}_2) \equiv 0 \pmod{d_2}$$

Because

$$\gcd(n_1, n_2) = 1 \implies \gcd(a_1, d_2) = 1$$

therefore

$$b_1 - \tilde{b}_1 \equiv 0 \pmod{d_1}, \text{ i.e. } b_1 \equiv \tilde{b}_1 \pmod{d_1}$$

Both numbers b_1 and $\tilde{b}_1$ belong to the same rest system, which implies

$$b_1 = \tilde{b}_1.$$

Analogously one proves $b_2 = \tilde{b}_2$.

2. In addition to part 1) the coprimality of N_1 and n_2 implies the bijectivity of the map of divisor sets

$$\{a_1 \in \mathbb{N}^* : a_1 | n_1\} \times \{a_2 \in \mathbb{N}^* : a_2 | n_2\} \rightarrow \{a \in \mathbb{N}^* : a | n_1 n_2\}, (a_1, a_2) \mapsto a_1 a_2$$

□

Proposition 5.10 (The algebra of lattice maps). *The lattice functions from Definition 5.5 satisfy the following rules:*

1.

$$R_\lambda \circ R_\mu = R_{\lambda \cdot \mu}$$

2.

$$R_\lambda \circ T_n = T_n \circ R_\lambda$$

3. Commutativity:

$$T_m \circ T_n = T_n \circ T_m \text{ if } (m, n) = 1$$

4. Recursion:

$$T_{p^n} \circ T_p = T_{p^{n+1}} + p \cdot T_{p^{n-1}} \circ R_p \text{ if } p \text{ prime}$$

5. Abelian: *The algebra of lattice functions*

$$\text{span} \langle (T_m)_{m \in \mathbb{N}^*}, (R_\lambda)_{\lambda \in \mathbb{C}^*} \rangle$$

is Abelian. It is generated by $(T_p)_{p \text{ prime}}$ together with $(R_\lambda)_{\lambda \in \mathbb{C}^}$*

Proof. See [56, Chap. VII, §5.1, Prop. 10]

ad step 1: The proof is trivial.

ad step 2: The proof is an easy calculation with the bases of the lattices in question.

ad step 3: Lemma 5.9 provides the complete set of representatives

$$V(m \cdot n) = V(m) \cdot V(n) \in \Gamma_m \cdot \Gamma_n \subset \Gamma_{m \cdot n}$$

Hence

$$T_{m \cdot n} = T_m \circ T_n$$

and the left-hand side is independent from the order of both factors.

ad step 4: Consider a lattice Λ . Both terms

$$(T_{p^n} \circ T_p)\Lambda \text{ and } T_{p^{n+1}}\Lambda$$

are the sum of sublattices $\Lambda'' \subset \Lambda$ with index $[\Lambda : \Lambda''] = p^{n+1}$. While the term

$$\Lambda'' \subset (T_{p^{n-1}} R_p)\Lambda = T_{p^{n-1}}(R_p \Lambda)$$

has index $[\Lambda : \Lambda''] = p^2$. We have to show: For each sublattice

$$\Lambda'' \subset \Lambda$$

which is a summand

$$a \cdot \Lambda'' \subset (T_{p^n} \circ T_p) \Lambda, \quad b \cdot \Lambda'' \subset T_{p^{n+1}} \Lambda, \quad c \cdot \Lambda'' \subset T_{p^{n-1}} (R_p \Lambda)$$

with non-negative integer coefficients $a, b, c \in \mathbb{Z}$ the coefficients satisfy

$$a = b + pc$$

By definition of the map $T_{p^{n+1}}$ each summand of $T_{p^n} \Lambda$ is a lattice Λ'' appearing with multiplicity = 1. Hence $b = 1$, and the claim reduces to

$$a = 1 + pc$$

The idea of the proof is now to distinguish between those Λ'' which are contained in the sublattice $p\Lambda \subset \Lambda$ and those which are not. Recall from Proposition 5.8 that each sublattice $\Lambda' \subset \Lambda$ of index $n \in \mathbb{N}$ contains the lattice $n \cdot \Lambda$.

- $\Lambda'' \subset p\Lambda$: Then $c = 1$ is the multiplicity of Λ'' in Λ . The lattice $T_p \Lambda$ is the sum

$$T_p \Lambda = \Lambda'_1 + \dots + \Lambda'_{p+1}$$

Each of these Λ'_j contains $p\Lambda$, and hence a posteriori Λ'' , which has index p^{n+1} in Λ . Hence Λ'' has multiplicity $a = 1 + p$ in $T_{p^n}(T_p \Lambda)$. In particular

$$a = 1 \text{ which equals } b + pc = 1$$

- $\Lambda'' \not\subset p\Lambda$: Then $c = 0$ because $p\Lambda = R_p \Lambda$ and each sublattice $\Lambda' \subset \Lambda$ of index p contains $p\Lambda$. Further on, there exists a sublattice $\Lambda' \subset \Lambda$ with index = p such that $\Lambda'' \subset \Lambda'$. Then the number a equals the count of such lattices Λ' with

$$\Lambda'' \subset \Lambda' \subset \Lambda$$

If Λ'' were contained in two different sublattices Λ'_j , $j = 1, 2$ of index p , then also

$$\Lambda'' \subset (\Lambda'_1 \cap \Lambda'_2)$$

Due to Proposition 5.8

$$\Lambda'_1 \cap \Lambda'_2 = p \cdot \Lambda$$

Hence $\Lambda'' \subset p\Lambda$, which has been excluded. As a consequence

$$a = 1 \text{ equals } b + pc = 1 + 0 = 1$$

We now carry over the definition of the lattice functions from Definition 5.5 to modular forms by using the correspondence from Lemma 5.3. The result are the Hecke operators.

Definition 5.11 (Hecke operators). For $m \in \mathbb{N}^*$ consider on the dual $\mathcal{L}^\vee$ the map

$$\check{T}_m : \mathcal{L}^\vee \rightarrow \mathbb{C}, F \mapsto \check{T}_m(F) := F \circ T_m$$

and define for even $k \in \mathbb{N}$ the *Hecke operators* as the maps on modular functions of weight k

$$T_m : M_k(\Gamma) \rightarrow M_k(\Gamma), T_m(f) := m^{k-1} \cdot \check{T}_m(F(\Lambda(-, 1)))$$

i.e.

$$T_m(f)(\tau) = m^{k-1} \cdot \check{T}_m(F(\Lambda(\tau, 1))), \tau \in \mathbb{H}.$$

Theorem 5.12 (Transformation of Fourier coefficients under Hecke operators).

1. The Hecke operators from Definition 5.11 are well-defined, i.e. for each even weight k and for all $m \geq 1$

$$f \in M_k(\Gamma) \implies T_m(f) = m^{k-1} \cdot \sum_{\gamma \in V(m)} f^\gamma \cdot \frac{1}{d^k} = \sum_{\gamma \in V(m)} f[\gamma]_k \in M_k(\Gamma)$$

with the notation $f[\gamma]_k$ from Lemma 3.4.

2. For a modular form $f \in M_k(\Gamma)$ with Fourier expansion

$$f(\tau) = \sum_{n=0}^{\infty} a_n \cdot q^n, q = e^{2\pi i \tau},$$

its Hecke transform $T_m f \in M_k(\Gamma)$ has the Fourier expansion

$$(T_m f)(\tau) = \sum_{n=0}^{\infty} b_n \cdot q^n$$

with Fourier coefficients

$$b_n = \sum_{\substack{r > 0 \\ r | (n, m)}} r^{k-1} \cdot a_{nm/r^2}, n \in \mathbb{N},$$

Here the summation is over all common divisors $r \in \mathbb{N}^*$ of the numbers m and n . In particular

$$b_0 = a_0 \cdot \sigma_{k-1}(m) \text{ and } b_1 = a_m$$

3. The ideal of cusp forms is stable under all Hecke operators, i.e. for each $m \in \mathbb{N}^*$

$$T_m(S_*(\Gamma)) \subset S_*(\Gamma).$$

Due to Theorem 5.12, part 1) we will use the same notation T_m for a Hecke operator and its restriction to a single modular space of fixed weight.

Proof. 1. By definition

$$\begin{aligned} T_m(f)(\tau) &:= m^{k-1} \cdot \check{T}_m(F(\Lambda(\tau, 1))) = m^{k-1} \cdot F(T_m(\Lambda(\tau, 1))) = \\ &m^{k-1} \cdot F\left(\sum_{\substack{\Lambda' \\ [\Lambda(\tau, 1):\Lambda']}} \Lambda'\right) = m^{k-1} \cdot \sum_{\substack{\Lambda' \\ [\Lambda(\tau, 1):\Lambda']}} F(\Lambda') = m^{k-1} \cdot \sum_{\gamma \in V(m)} F(\Lambda(\gamma(\tau), 1)) = \\ &m^{k-1} \cdot \sum_{\gamma \in V(m)} F(\Lambda(a\tau + b, d)) = \sum_{\gamma \in V(m)} f(a\tau + b) \cdot d^{-k} \cdot m^{k-1} = \sum_{\gamma \in V(m)} f[\gamma]_k(\tau) \end{aligned}$$

The matrices of the last summations have the form

$$\gamma = \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} \in V(m)$$

Hence

$$T_m(f) \in M_k(\Gamma)$$

which proves the well-definedness of the Hecke operators.

2. To transform the Fourier series of a modular form of weight k

$$f(\tau) = \sum_{n=0}^{\infty} a_n \cdot q^n, \quad q = e^{2\pi i \tau},$$

into the Fourier series of the modular form

$$(f[\gamma]_k)(\tau) = f(\gamma(\tau)) \cdot h(\gamma, \tau)^{-k} \cdot (\det \gamma)^{k-1}, \quad \gamma = \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} \in V(m),$$

we employ the following intermediate results:

- For each summand

$$\exp(2\pi i n \cdot \gamma(\tau)) = \exp\left(2\pi i n \cdot \frac{a\tau + b}{d}\right) = e^{2\pi i n \cdot (a\tau/d)} \cdot e^{2\pi i n \cdot (b/d)}, \quad n \in \mathbb{N},$$

- For the second factor

$$\sum_{0 \leq b < d} e^{2\pi i b \cdot (n/d)} = \begin{cases} d & \text{if } d|n \\ 0 & \text{otherwise} \end{cases}$$

The first case follows from

$$e^{2\pi i b \cdot (n/d)} = 1$$

for each of the d summands with $0 \leq b < d$. The second case follows from the formula of the finite geometric series with argument $e^{2\pi i \cdot (n/d)}$:

$$\sum_{b=0}^{d-1} e^{2\pi i b \cdot (n/d)} = \frac{e^{2\pi i \cdot d \cdot (n/d)} - 1}{e^{2\pi i \cdot (n/d)} - 1} = 0$$

- Simplifying the summand

$$\begin{aligned} f(\gamma(\tau)) &= f\left(\frac{a\tau + b}{d}\right) = \sum_{n=0}^{\infty} a_n \cdot \exp\left(2\pi i n \cdot \frac{a\tau + b}{d}\right) = \\ &= \sum_{n=0}^{\infty} a_n \cdot \exp\left(2\pi i n \cdot \frac{a\tau}{d}\right) \cdot \exp\left(2\pi i n \cdot \frac{b}{d}\right) \end{aligned}$$

Hence

$$\sum_{b=0}^{d-1} f(\gamma(\tau)) = \sum_{n=0}^{\infty} d \cdot a_{nd} \cdot q^{n(m/d)}, \quad q = \exp(2\pi i n \tau),$$

because

$$\begin{aligned} \sum_{b=0}^{d-1} f(\gamma(\tau)) &= \sum_{b=0}^{d-1} \left(\sum_{n=0}^{\infty} a_n \cdot e^{2\pi i n(a\tau)/d} \cdot e^{2\pi i n(b/d)} \right) = \\ &= \sum_{n=0}^{\infty} \left(a_n \cdot e^{2\pi i n(a\tau)/d} \cdot \sum_{b=0}^{d-1} e^{2\pi i n(b/d)} \right) = \sum_{n=0}^{\infty} a_{nd} \cdot e^{2\pi i na\tau} \cdot d = \\ &= d \cdot \sum_{n=0}^{\infty} a_{nd} \cdot q^{na} = d \cdot \sum_{n=0}^{\infty} a_{nd} \cdot q^{n(m/d)} \end{aligned}$$

Due to the second step the third last sum retains only summands with index $= nd$.

- Summing up we have

$$\begin{aligned} (T_m f)(\tau) &= \sum_{\gamma \in V(m)} f[M]_k(\tau) = m^{k-1} \cdot \sum_{\gamma \in V(m)} h(\gamma, \tau)^{-k} \cdot f(\gamma(\tau)) = \\ &= m^{k-1} \sum_{\substack{d>0 \\ d|m}} \frac{1}{d^k} \cdot \sum_{n=0}^{\infty} a_{nd} \cdot q^{n(m/d)} \cdot d = \sum_{\substack{d>0 \\ d|m}} \left(\frac{m}{d}\right)^{k-1} \left(\sum_{n=0}^{\infty} a_{nd} \cdot q^{n(m/d)} \right) = \\ &= \sum_{\substack{d>0 \\ d \in (m, n)}} \left(\frac{m}{d}\right)^{k-1} \left(\sum_{n=0}^{\infty} a_n \cdot q^{(n/d) \cdot (m/d)} \right) \end{aligned}$$

The last step made in the inner sum the substitution

$$nd \mapsto n$$

The subsequent substitution of index pairs

$$(d, n) \mapsto (t, r) \text{ with } t := \frac{n}{d} \cdot \frac{m}{d} \text{ and } r := \frac{m}{d}$$

provides

$$n = \frac{d^2 \cdot t}{m} = \left(\frac{d}{m}\right)^2 tm = \frac{tm}{r^2}.$$

Hence

$$(T_m f)(\tau) = \sum_{t=0}^{\infty} \left(\sum_{\substack{r>0 \\ r|(m,t)}} r^{k-1} \cdot a_{(tm)/r^2} \right) q^t$$

with the inner sum being finite.

3. A cusp form f satisfies $a_0 = 0$. The formula for the Fourier coefficients of $T_m f$ from the previous part shows that $T_m f$ is a cusp form again.

□

The following corollary specializes the general formula from Theorem 5.12. It gives the formula for the Fourier coefficients of the Hecke transform for prime indices.

Corollary 5.13 (Fourier coefficients of Hecke transforms with prime index).

For a Hecke transform T_p with a prime index p holds: For a modular form $f \in M_k(\Gamma)$ with Fourier coefficients $(a_n)_{n \in \mathbb{N}}$ the Hecke transform $T_p f$ has the Fourier coefficients $(b_n)_{n \in \mathbb{N}}$ with

$$b_n = \begin{cases} a_{np} & \text{if } p \nmid n \\ a_{np} + p^{k-1} \cdot a_{n/p} & \text{if } p|n \end{cases}$$

Proof. The claim follows from Theorem 5.12, part 2: For the coefficient b_n the sum

$$b_n = \sum_{\substack{r>0 \\ r|(n,t)}} r^{k-1} \cdot a_{(nt)/r^2}$$

has at most two terms, for $r = 1$ and $r = p$. The corresponding terms are respectively a_{pn} and $a_{p/n}$.

- If $p \nmid n$ only $r = 1$ with the coefficient a_{pn} remains.
- If $p|n$ then in addition also $r = p$ with the coefficient $p^{k-1} \cdot a_{n/p}$ remains.

□

Corollary 5.14 (The discriminant modular form as common eigenvector of Hecke operators). *The discriminant modular form*

$$\Delta \in S_{12}(\Gamma)$$

is a simultaneous eigenform of all Hecke operators

$$T_m \in \text{End}(S_{12}(\Gamma)), \quad m \in \mathbb{N}^*.$$

The Fourier coefficients of the normalized discriminant modular form

$$\frac{\Delta}{(2\pi)^{12}} = \sum_{m=1}^{\infty} \tau_m \cdot q^m$$

are the corresponding eigenvalues, i.e. for all $m \in \mathbb{N}^$*

$$T_m(\Delta) = \tau_m \cdot \Delta.$$

Proof. According to Corollary 3.19 and Lemma 3.23 the vector space $S_{12}(\Gamma)$ of cusp forms of weight $k = 12$ is 1-dimensional with the discriminant modular form Δ from Definition 3.13 as a basis. Hence Δ is an eigenvector of each Hecke operator T_m , $m \geq 1$. We recall the Fourier expansion of the normalized discriminant modular form from Proposition 3.14:

$$\frac{\Delta(\tau)}{(2\pi)^{12}} = \sum_{n=1}^{\infty} \tau_n \cdot q^n = \tau_1 + O(2), \quad \tau_1 = 1$$

If we denote the Fourier series of its Hecke transform by

$$T_m \left(\frac{\Delta(\tau)}{(2\pi)^{12}} \right) = \sum_{n=1}^{\infty} b_n \cdot q^n = b_1 + O(2)$$

then Theorem 5.12 implies for the linear coefficient

$$b_1 = \tau_m$$

which proves

$$T_m \left(\frac{\Delta}{(2\pi)^{12}} \right) = \tau_m \cdot \frac{\Delta}{(2\pi)^{12}} \quad \text{or} \quad T_m(\Delta) = \tau_m \cdot \Delta$$

□

Example 5.15 (Hecke operators). Consider the vector space of cusp forms $S_{28}(\Gamma)$. Due to Theorem 3.19 and Lemma 3.23

$$\dim S_{28}(\Gamma) = 2.$$

The two cuspforms of weight 28

$$f_1 := \Delta \cdot E_4^4 \text{ and } f_2 := \Delta^2 \cdot E_4$$

are linearly independent because

$$\text{ord}(f_1; \infty) = 1 \text{ and } \text{ord}(f_2; \infty) = 2.$$

Hence (f_1, f_2) is a basis of $S_{28}(\Gamma)$. We consider the Hecke operator

$$T_2 : S_{28}(\Gamma) \rightarrow S_{28}(\Gamma).$$

Figure 5.1 shows

- the Fourier series of f_1 and f_2
- the Fourier series of $T_2 f_1$ and $T_2 f_2$
- the Hecke matrix of T_2 with respect to the basis (f_1, f_2)
- the eigenvalues of T_2
- the Fourier coefficients of two corresponding eigenvectors (v_1, v_2) of T_2 , normalized with linear Fourier coefficient = 1,
- the representation of the eigenvectors with respect to (f_1, f_2)

```

=====
Hecke_matrix_02, Start

Hecke operator T_2: S_28(Gamma_0(1)) --> S_28(Gamma_0(1))

Basis of S_28(Gamma_0(1)): f1 = Delta * E_4^4, f2 = Delta^2 * E_4 with Fourier series:

f1_ser: q + 936*q^2 + 331452*q^3 + 53282368*q^4 + 0(q^5)
f2_ser: q^2 + 192*q^3 - 8280*q^4 + 0(q^5)
-----

heck_2_f1_ser: 936*q + 187500096*q^2 + 36142047072*q^3 + 0(q^4)
heck_2_f2_ser: q - 8280*q^2 - 1438020*q^3 + 0(q^4)
-----

Matrix of Hecke operator T_2 with respect to basis (f1,f2): [936, 1; 186624000, -9216]
-----

Eigenvalues: -18713.59859, 10433.59859
-----

v_1: 1.000000000*q - 18713.59859*q^2 - 3441270.930*q^3 + 215981044.4*q^4 + 0(q^5)
v_2: 1.000000000*q + 10433.59859*q^2 + 2154990.930*q^3 - 25357748.36*q^4 + 0(q^5)
-----

v_1 = 1.000000000 * f1 + -19649.59859 * f2
v_2 = 1.000000000 * f1 + 9497.598595 * f2
-----

```

Fig. 5.1 Hecke operator $T_2 \in \text{End}(S_{28}(\Gamma))$

Hence the endomorphism

$$T_2 : S_{28}(\Gamma) \rightarrow S_{28}(\Gamma)$$

has two distinct eigenvalues, and can be diagonalized.

The result from Figure 5.1 has been computed by the PARI script

Hecke_matrix_02.

It is not by chance that in Example 5.15 the Hecke endomorphism $T_2 \in \text{End}(S_{28}(\Gamma))$ can be diagonalized. Theorem 5.17 will show: The Hecke algebra of a given weight is commutative. And Theorem 5.26 will show that the Hecke operators are selfadjoint with respect to a certain Hermitian scalar product on the cusp forms. Hence all Hecke operators on cusp forms can be simultaneously diagonalized with real eigenvalues.

Definition 5.16 (Hecke algebra of weight k). The *Hecke algebra* of even weight $k \in \mathbb{N}$

$$\text{Hecke}_k(\Gamma) \subset \text{End}(M_k(\Gamma))$$

is the subalgebra spanned by the family $(T_m|_M_k(\Gamma))_{m \in \mathbb{N}}$ of Hecke operators.

Theorem 5.17 (The Hecke algebras are commutative). *For each weight $k \in \mathbb{N}$ the Hecke algebra $\text{Hecke}_k(\Gamma)$ is commutative. It has the following properties:*

1. *For coprime $n, m \in \mathbb{N}^*$ Hecke operators are multiplicative:*

$$T_m \circ T_n = T_{m \cdot n}$$

2. *For a prime power p^r , $r \geq 1$, holds*

$$T_{p^r} \circ T_p = T_{p^{r+1}} + p^{k-1} \cdot T_{p^{r-1}}$$

In particular

$$T_{p^r} \in \mathbb{Z}[T_p]$$

3. *The algebra $\text{Hecke}_k(\Gamma)$ is generated by the family T_p , p prime.*

Proof. The proof follows from Proposition 5.10.

□

5.2 Petersson scalar product and level change of modular forms

We now introduce a Hermitian scalar product on the vector space of cusp forms of a given congruence subgroup

$$\Gamma_0(N) \subset \Gamma$$

with the usual notation $\Gamma = SL(2, \mathbb{Z})$ for the modular group.

Definition 5.18 (Hyperbolic volume form). The *hyperbolic volume form* on $\mathbb{H}$ is the differential form

$$d\mu(\tau) := \frac{dx \wedge dy}{y^2}, \quad \tau = x + iy \in \mathbb{H}$$

Lemma 5.19 (Invariance of the hyperbolic volume form). *The hyperbolic volume form is $GL(2, \mathbb{R})^+$ -invariant, i.e. for each $\gamma \in GL(2, \mathbb{R})^+$ the pullback satisfies*

$$\gamma^* d\mu = d\mu$$

Proof. For each

$$\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in GL(2, \mathbb{R})^+$$

we have to show: For all $\tau \in \mathbb{H}$

$$(\gamma^* d\mu)(\tau) = d\mu(\tau) :$$

For the proof one uses the usual coordinate transformation

$$dx \wedge dy = \frac{i}{2} d\tau \wedge d\bar{\tau}$$

which implies

$$d\mu(\tau) = \frac{i}{2} \cdot \frac{d\tau \wedge d\bar{\tau}}{(Im \tau)^2}$$

and

$$(\gamma^* d\mu)(\tau) = \frac{i}{2} \cdot \frac{d\gamma(\tau) \wedge d\gamma(\bar{\tau})}{(Im \gamma(\tau))^2}$$

The proof of Proposition 3.1 and the formulas from Remark 2.8

$$Im \gamma(\tau) = det(\gamma) \cdot \frac{Im \tau}{|c\tau + d|^2} \text{ and } d\gamma(\tau) = det(\gamma) \cdot \frac{d\tau}{(c\tau + d)^2}$$

imply

$$(\gamma^* d\mu)(\tau) = d\mu(\tau)$$

□

Proposition 5.20 (Integration over fundamental domains).

1. The standard fundamental domain $\mathcal{D}$ of Γ has the volume

$$vol \mathcal{D} := \int_{\mathcal{D}} d\mu = \frac{\pi}{3}$$

2. The fundamental domain of a congruence subgroup $\Gamma_0(N)$ has the volume

$$vol \mathcal{D}(\Gamma_0(N)) := \int_{\mathcal{D}(\Gamma_0(N))} d\mu = [\Gamma : \Gamma_0(N)] \cdot vol \mathcal{D} = \frac{\pi}{3} \cdot N \cdot \prod_{p|N} \left(1 + \frac{1}{p}\right),$$

independent from the chosen left coset decomposition of Γ with respect to $\Gamma_0(N)$.

3. Consider a continuous function

$$\Omega : \mathbb{H} \rightarrow \mathbb{C}$$

with fixed group

$$\Gamma_{\Omega} := \{A \in \Gamma : \Omega \circ A = \Omega\}.$$

Then for each subgroup $K \subset \Gamma_{\Omega}$ with $\pm 1 \in K$ and for each pair $(\mathcal{F}_1, \mathcal{F}_2)$ of fundamental domains of K :

$$\int_{\mathcal{F}_1} \Omega \, d\mu = \int_{\mathcal{F}_2} \Omega \, d\mu$$

if at least one of the two integrals is finite.

Proof. 1. From the geometrical description of $\mathcal{D}$, see Figure 2.1:

$$\begin{aligned} \int_{\mathcal{D}} d\mu &= \int_{-1/2}^{1/2} \left(\int_{\sqrt{1-x^2}}^{\infty} \frac{dy}{y^2} \right) dx = \int_{-1/2}^{1/2} \left((-1/y) \Big|_{\sqrt{1-x^2}}^{\infty} \right) dx = \\ &= \int_{-1/2}^{1/2} \frac{dx}{\sqrt{1-x^2}} = \arcsin x \Big|_{-1/2}^{1/2} = \pi/6 - (-\pi/6) = \pi/3 \end{aligned}$$

2. The general formula for

$$\text{vol}(\mathcal{D}(\Gamma_0(N))) = [\Gamma : \Gamma_0(N)] \cdot \text{vol} \mathcal{D}$$

follows from

- the representation of a fundamental domain of $\Gamma_0(N)$ as the disjoint union

$$\mathcal{D}(\Gamma_0(N)) = \bigcup_{j=1}^{[\Gamma:\Gamma_0(N)]} \gamma_j^{-1}(\mathcal{D})$$

of fundamental domains of Γ , see Remark 2.19,

- the invariance of the volume form $d\mu$ according to Lemma 5.19,
- and the index formula Proposition 2.16.

3. For $A \in \Gamma_Q$ the usual transformation formula for integrals and the invariance of the hyperbolic volume form $d\mu$ imply for each measurable set $B \subset \mathbb{H}$

$$\int_B \Omega \, d\mu = \int_{A(B)} \Omega \, d\mu$$

If one splits $\mathbb{H}$ as the disjoint union of the sets $A(\mathcal{F}_2)$, $A \in K$, and observes that ± 1 act as identity on $\mathbb{H}$, which introduces the factor $= 1/2$, then

$$\begin{aligned} \int_{\mathcal{F}_1} \Omega \, d\mu &= (1/2) \cdot \sum_{A \in K} \int_{A(\mathcal{F}_2) \cap \mathcal{F}_1} \Omega \, d\mu = (1/2) \cdot \sum_{A \in K} \int_{\mathcal{F}_2 \cap A^{-1}(\mathcal{F}_1)} \Omega \, d\mu = \\ &= (1/2) \cdot \sum_{A \in K} \int_{A(\mathcal{F}_1) \cap \mathcal{F}_2} \Omega \, d\mu = \int_{\mathcal{F}_2} \Omega \, d\mu \end{aligned}$$

□

Definition 5.21 (Petersson scalar product). For each level $N \in \mathbb{N}^*$ and each even weight $k \in \mathbb{N}$, $k \geq 0$, the *Petersson scalar product* on the space $S_k(\Gamma_0(N))$ of cusp forms is the $\mathbb{C}$ -sesquilinear map

$$\langle -, - \rangle_{\Gamma_0(N)} : S_k(\Gamma_0(N)) \times S_k(\Gamma_0(N)) \rightarrow \mathbb{C},$$

$$\langle f, g \rangle_{\Gamma_0(N)} := \frac{1}{\text{vol } \mathcal{D}(\Gamma_0(N))} \cdot \int_{\mathcal{D}(\Gamma_0(N))} f(\tau) \cdot \overline{g(\tau)} \cdot (\text{Im } \tau)^k d\mu(\tau)$$

Note

$$\text{vol } \mathcal{D}(\Gamma_0(N)) < \infty \text{ because } [\Gamma : \Gamma_0(N)] < \infty.$$

Theorem 5.22 will prove that the Petersson scalar product is well-defined.

Definition 5.21, see [19, Chap. 5, Def. 5.4.1], implies: If

$$\Gamma_1 := \Gamma_0(N_1) \subset \Gamma_2 := \Gamma_0(N_2) \text{ and } f, g \in S_k(\Gamma_2)$$

then

$$\langle f, g \rangle_{\Gamma_1} = \langle f, g \rangle_{\Gamma_2}$$

While [38, Kap. IV, 3.2], which considers only the modular group Γ , employs in Definition 5.21 instead before the integral the factor $= 1$.

Theorem 5.22 (Petersson scalar product). For each congruence subgroup $\Gamma_0(N) \subset \Gamma$, $N \in \mathbb{N}^*$, the map

$$\langle -, - \rangle_{\Gamma_0(N)} : S_k(\Gamma_0(N)) \times S_k(\Gamma_0(N)) \rightarrow \mathbb{C},$$

from Definition 5.21 is well-defined. It is a Hermitian scalar product on the vector space $S_k(\Gamma_0(N))$.

Proof. For the proof see [19, Chap. 5.4]. To abbreviate the integrand in Definition 5.21 we introduce for fixed $f, g \in S_k(\Gamma_0(N))$ the function

$$\Omega : \mathbb{H} \times \mathbb{H} \rightarrow \mathbb{C}, \quad \Omega(\tau) := f(\tau) \cdot \overline{g(\tau)} \cdot (\text{Im } \tau)^k$$

Then

$$\langle f, g \rangle_{\Gamma_0(N)} = \frac{1}{\text{vol } \mathcal{D}(\Gamma_0(N))} \cdot \int_{\mathcal{D}(\Gamma_0(N))} \Omega(\tau) d\mu(\tau)$$

1. *Transformation of ϕ :* For all $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma$ holds $\Omega \circ \gamma = \phi$:

$$\Omega(\gamma(\tau)) := f(\tau) \cdot \overline{g(\tau)} \cdot (\text{Im } \gamma(\tau))^k =$$

$$\begin{aligned}
&= f[\gamma]_k(\tau) \cdot (c\tau + d)^k \cdot \overline{g[\gamma]_k(\tau)} \cdot \overline{(c\tau + d)}^k \cdot (Im \tau)^k \cdot |(c\tau + d)^k|^{-2k} = \\
&= f(\tau) \cdot \overline{g(\tau)} \cdot (Im \tau)^k = \Omega(\tau)
\end{aligned}$$

As a consequence, for any measurable set $A \subset \mathbb{H}$ and $\gamma \in \Gamma$

$$\int_{\gamma A} \Omega d\mu(\tau) = \int_A \Omega \circ \gamma d\mu(\tau)$$

2. *Boundedness of ϕ* : We claim that for cusp forms $f, g \in S(\Gamma_0(N))$ and for all $\gamma \in \Gamma$ the function

$$\Omega \circ \gamma: \mathcal{D} \rightarrow \mathbb{C}$$

is bounded. Because the sets

$$\{\tau \in \mathbb{H} : Im \tau \leq y\}, y > 0,$$

are compact and $\Omega \circ \gamma$ is continuous, we have to investigate only the behaviour on neighbourhoods of ∞ in $\mathbb{H}^*$ of the form

$$\{\tau \in \mathbb{H} : Im \tau > y\}, y > 0$$

We show

$$\lim_{Im \tau \rightarrow \infty} \phi(\gamma(\tau)) = 0$$

For each cusp of $\Gamma_0(N)$ the functions $f[\gamma]_k, g[\gamma]_k$ have the Fourier development

$$(f[\gamma]_k)(\tau) = \sum_{n=1}^{\infty} a_n(f, \gamma) \cdot q_h^n, (g[\gamma]_k)(\tau) = \sum_{n=1}^{\infty} a_n(g, \gamma) \cdot q_h^n, q_h = e^{2\pi i(\tau/h)}$$

with $h \in \mathbb{N}^*$ the width of the cusp. Both Fourier series are of order $O(q_h)$ with respect to the limit $\tau \rightarrow \infty$, because the Fourier series of a cusp form has no constant term. Hence the product $\Omega(f, g) \circ \gamma$ is of order $O(q_h)^2$ and the integrand of the Petersson scalar product is of order

$$O(q_h)^2 \cdot Im(\tau)^k$$

To express $O(q_h)^2$ with respect to the variable $\tau = e^{2\pi i(\tau/h)}$ we compute

$$|q_h| = e^{-2\pi(Im \tau/h)}$$

The term decreases exponentially with $\tau > 0$, which is stronger than the polynomial growth of

$$(c\tau + d)^{2k} \cdot (Im \tau)^k$$

As a consequence

$$\lim_{Im \tau \rightarrow \infty} \Omega(\gamma(\tau)) = 0$$

which finishes the proof of boundedness.

3. *Scalar product*: Apparently the map $\langle -, - \rangle_{\Gamma_0(N)}$ is a positive Hermitian form, hence a scalar product.

□

Note. The proof of Theorem 5.22 shows that the Hermitian scalar product $\langle f, g \rangle_{\Gamma_0(N)}$ is also defined for two modular forms $f, g \in M_k(\Gamma_0(N))$ if at least one of them is a cusp form.

Next we prove that the Hecke operators $(T_n)_{n \in \mathbb{N}^*}$ are self-adjoint with respect to the Petersson scalar product. After the preparations in Lemma 5.23 the proof follows in Proposition 5.24 and Theorem 5.25.

Lemma 5.23 (Complete sets of common representatives for left/right-action).

Consider an arbitrary but fixed prime p , and the left action

$$\Phi_{left} : \Gamma \times \Gamma_p \rightarrow \Gamma_p, (\gamma, M) \mapsto \gamma \cdot M$$

1. Besides the set $V(p)$ from Proposition 5.8 the set of symmetric matrices

$$V_{sym} := \left\{ \begin{pmatrix} p & 0 \\ 0 & 1 \end{pmatrix} \right\} \cup \left\{ \begin{pmatrix} 1 & b \\ b & p+b^2 \end{pmatrix} : 0 \leq b < p \right\}$$

is a complete set of representatives of Φ_{left} too.

2. Consider a complete set $\mathcal{V}$ of representatives of Φ_{left} formed by symmetric matrices. Then also the system of adjoints

$$\mathcal{V}^\# := \{M^\# : M \in \mathcal{V}\}, M^\# := (\det M) \cdot M^{-1},$$

is a complete sets of representatives of Φ_{left} .

Proof. 1. For each $0 \leq b < p$

$$\begin{pmatrix} 1 & b \\ b & p+b^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ b & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & b \\ 0 & p \end{pmatrix}$$

Hence the matrices

$$\begin{pmatrix} 1 & b \\ b & p+b^2 \end{pmatrix} \text{ and } \begin{pmatrix} 1 & b \\ 0 & p \end{pmatrix}$$

belong to the same orbit of Φ_{left} .

2. i) Let

$$\Phi_{right} : \Gamma_p \times \Gamma \rightarrow \Gamma_p, (M, \gamma) \mapsto M \cdot \gamma,$$

be the corresponding right action of Γ . For each complete set $\mathcal{V}$ of representatives of Φ_{left} the set of transposed elements

$$\mathcal{V}^\top := \{v^\top : v \in \mathcal{V}\}$$

is a complete set of representatives of Φ_{right} : Assume $v \in \Gamma_p$. Then also $v^\top \in \Gamma_p$ and

$$v^\top = A \cdot M$$

for suitable

$$A \in \Gamma \text{ and } M \in \mathcal{V}.$$

Hence

$$v = M^\top \cdot A^\top$$

with $A^\top \in \Gamma^\top = \Gamma$.

As a consequence, if $\mathcal{V}$ is formed by symmetric matrices, then $\mathcal{V}^\top$ is also a complete set of representatives for Φ_{right} .

ii) For given $w \in \Gamma_p$ set

$$v := (\det w) \cdot w^{-1} \in \Gamma_p.$$

Then

$$v = w^\sharp \text{ and } \det v = (\det w)^2 \cdot \det w^{-1} = \det w \text{ and } w = (\det v) \cdot v^{-1}$$

Due to part i) exist

$$A \in \Gamma \text{ and } M \in \mathcal{V}$$

with

$$v = M \cdot A$$

Then

$$\det v = \det M$$

and

$$v^{-1} = A^{-1} \cdot M^{-1}$$

and

$$w = v^{-1} \cdot \det v = A^{-1} \cdot ((\det M) \cdot M^{-1})$$

with

$$A^{-1} \in \Gamma^{-1} = \Gamma$$

□

Recall from Remark 2.8 the operation of $GL(2, \mathbb{Q})^+$ and its subgroups on $\mathbb{H}$

Proposition 5.24 (Petersson scalar product). *Consider a weight $k \in \mathbb{N}$ and two cusp forms $f, g \in S_k(\Gamma)$. For two continuous functions*

$$h_1, h_2 : \mathbb{H} \rightarrow \mathbb{P}^1$$

set - generalizing the function from the proof of Theorem 5.22 -

$$\Omega := \Omega(h_1, h_2) : \mathbb{H} \rightarrow \mathbb{P}^1, \quad \Omega(h_1, h_2)(\tau) := h_1(\tau) \cdot \overline{h_2(\tau)} \cdot (Im \tau)^k$$

Then:

1. Transformation formula for subgroups of finite index: *Each subgroup $K \subset \Gamma_\Omega$ with finite index $[\Gamma_\Omega : K] < \infty$ satisfies*

$$\frac{1}{[\Gamma_\Omega : K]} \cdot \int_{\mathcal{D}(K)} \Omega \, d\mu = \int_{\mathcal{D}(\Gamma_\Omega)} \Omega \, d\mu$$

if at least one of the two integrals is finite.

2. Finiteness of the index under conjugation: *The extended principal congruence subgroup of Γ of level a prime p*

$$\tilde{\Gamma}(p) := \{L \in \Gamma : L \equiv \pm 1 \bmod p\}$$

has finite index. With respect to the fixed groups

$$\Gamma_\Omega(f[M]_k, g), \quad M \in V(p), \quad \text{and} \quad \Gamma(\Omega(f, g[M^\sharp]_k))$$

holds the inclusion of subgroups:

$$\tilde{\Gamma}(p) \subset \Gamma(\Omega(f[M]_k, g)) \quad \text{and} \quad M^{-1}\tilde{\Gamma}(p)M \subset \Gamma(\Omega(f[M]_k, g)),$$

and both subgroups have finite index, and analogously for

$$\tilde{\Gamma}(p) \subset \Gamma(\Omega(f, g[M^\sharp]_k)) \quad \text{and} \quad M\tilde{\Gamma}(p)M^{-1} \subset \Gamma(\Omega(f, g[M^\sharp]_k)).$$

3. Invariance of the index under conjugation:

$$[\Gamma : \tilde{\Gamma}(p)] = [\Gamma : M^{-1}\tilde{\Gamma}(p)M]$$

Proof. Cf. [38, Kap. IV, §3.3]. Because k is fixed during the whole computation we simplify the notation $[-] := [-]_k$.

1. Consider a fundamental domain $\mathcal{F}$ of the fixed group $\Gamma(\Omega)$. Then each decomposition of $\Gamma(\Omega)$ into the finite number $[\Gamma(\Omega) : K]$ of left classes with respect to K

$$\Gamma(\Omega) = \dot{\bigcup}_V M_V^{-1} \cdot K$$

provides a fundamental domain of K

$$\mathcal{G} = \dot{\bigcup}_V M_V(\mathcal{F})$$

Proposition 5.20, part 3), applied to each pair $(\mathcal{F}, M_V(\mathcal{F}))$ of fundamental domains of $\Gamma(\Omega)$, implies

$$\int_{M_V \mathcal{F}} \Omega \, d\mu = \int_{\mathcal{F}} \Omega \, d\mu,$$

hence

$$\begin{aligned} \int_{\mathcal{D}(K)} \Omega \, d\mu &= \int_{\mathcal{G}} \Omega \, d\mu = \sum_V \int_{M_V(\mathcal{F})} \Omega \, d\mu = \\ &= \sum_V \int_{\mathcal{F}} \Omega \, d\mu = [\Gamma(\phi) : K] \cdot \int_{\mathcal{D}(\Gamma(\Omega))} \Omega \, d\mu \end{aligned}$$

2. i) According to Proposition 2.16

$$[\Gamma : \Gamma(p)] < \infty \text{ hence also } [\Gamma : \tilde{\Gamma}(p)] < \infty.$$

Each element $L \in \tilde{\Gamma}(p) \subset \Gamma$ with decomposition

$$L = \pm \mathbb{1} + p \cdot A, \quad A \in M(2 \times 2, \mathbb{Z}).$$

has the conjugate

$$L_0 := MLM^{-1} = M(\pm \mathbb{1} + p \cdot A)M^{-1} = \pm \mathbb{1} + M(pA)M^{-1} = \pm \mathbb{1} + MAM^\sharp$$

Hence

$$L_0 M = ML \text{ and } L_0 \in \Gamma$$

which implies

$$(f[M])[L] = (f[L_0])[M] = f[M].$$

Because $\tilde{\Gamma}(p) \subset \Gamma$ also

$$g[L] = g.$$

The two L -invariant elements

$$f[M] \text{ and } g$$

satisfy for $\tau \in \mathbb{H}$ the transformation formula

$$(f[M] \cdot \bar{g})(L(\tau)) = (f[M] \cdot \bar{g})(\tau) \cdot |h(L, \tau)|^{2k} :$$

We compute

$$\begin{aligned} (f[M] \cdot \bar{g})(L(\tau)) &= f[M](L(\tau)) \cdot \bar{g}(L(\tau)) = (f[M][L])(\tau) \cdot h(L, \tau)^k \cdot \bar{g}(L(\tau)) = \\ &= f[M](\tau) \cdot h(L, \tau)^k \cdot \bar{g}[L](\tau) \cdot \overline{h(L, \tau)^k} = f[M](\tau) \cdot \bar{g}(\tau) \cdot |h(L, \tau)|^{2k} \end{aligned}$$

Multiplying by $(Im L(\tau))^k$ both sides of the transformation formula for

$$f[M] \text{ and } g$$

and using the transformation formula for the imaginary part of the argument

$$Im L(\tau) = (Im \tau) \cdot h(L, \tau)^{-2}$$

shows

$$(f[M] \cdot \bar{g})(L(\tau)) \cdot (Im L(\tau))^k = (f[M] \cdot \bar{g})(\tau) \cdot (Im \tau)^k$$

or

$$\Omega(f[M], g) \circ L = \Omega(f[M], g)$$

Because $L \in \tilde{\Gamma}(p)$ is arbitrary we have shown

$$\tilde{\Gamma}(p) \subset \Gamma(\Omega(f[M], g))$$

The finiteness of the index of $\tilde{\Gamma}(p)$ in the fixgroups follow from the finiteness

$$[\Gamma : \tilde{\Gamma}(p)] < \infty$$

ii) For arbitrary $k \in \mathbb{N}$ holds

$$M \cdot \tilde{\Gamma}(kp) \cdot M^{-1} \subset \tilde{\Gamma}(k) :$$

If $L \in \tilde{\Gamma}(kp)$ then

$$M \cdot L \cdot M^\sharp \equiv \pm M \cdot \mathbb{1} \cdot M^\sharp \mod kp = \pm p \cdot \mathbb{1} \mod kp$$

$$M \cdot L \cdot \frac{1}{\det M} \cdot M^\sharp \equiv \pm \mathbb{1} \mod k \text{ i.e. } M \cdot L \cdot M^{-1} \in \tilde{\Gamma}(k).$$

Now setting $k = 1$ shows

$$M\tilde{\Gamma}(p)M^{-1} \subset \tilde{\Gamma},$$

while setting $k = p$ shows

$$M \cdot \tilde{\Gamma}(p^2) \cdot M^{-1} \subset \tilde{\Gamma}(p) \text{ or } \tilde{\Gamma}(p^2) \subset M^{-1} \cdot \tilde{\Gamma}(p) \cdot M$$

As a consequence

$$[\Gamma : M^{-1}\tilde{\Gamma}(p)M] \leq [\Gamma : \tilde{\Gamma}(p^2)] < \infty$$

Finally one checks

$$M^{-1}\tilde{\Gamma}(p)M \subset \Gamma(\Omega(f[M], g)).$$

iii) The claim about the subgroups related to $M^\sharp$ follows from part i) and ii) by interchanging the role of f and g and the role of M and $M^\sharp$.

3. We consider the two subgroups of Γ from part 2)

$$\tilde{\Gamma}(p) \subset \Gamma \text{ and } K := M^{-1}\tilde{\Gamma}(p)M.$$

Consider a fundamental domain $\mathcal{F} \subset \mathbb{H}$ of $\tilde{\Gamma}(p)$. Then $M^{-1}\mathcal{F}$ is a fundamental domain of K : For each $x \in \mathbb{H}$ also $M(x) \in \mathbb{H}$. Hence by assumption exists a unique $\gamma \in \tilde{\Gamma}(p)$ with

$$\gamma(M(x)) \in \mathcal{F}.$$

Then

$$M^{-1}(\gamma(M(x))) = (M^{-1}\gamma M)(x) \in M^{-1}\mathcal{F}$$

The uniqueness of γ implies the uniqueness of $M^{-1}\gamma M$.

We compute the volume of both fundamental domains, hereby using the $GL(2, \mathbb{Q})^+$ -invariance of the hyperbolic volume form $d\mu$ according to Lemma 5.19

$$\begin{aligned} [\Gamma : \tilde{\Gamma}(p)] \cdot \text{vol } \mathcal{D} &= \text{vol } \mathcal{F} = \int_{\mathcal{F}} d\mu = \\ &= \int_{M^{-1}\mathcal{F}} d\mu = \text{vol}(M^{-1}\mathcal{F}) = [\Gamma : K] \cdot \text{vol } \mathcal{D} \end{aligned}$$

Hence

$$[\Gamma : \tilde{\Gamma}(p)] = [\Gamma : K] = [\Gamma : M^{-1}\tilde{\Gamma}(p)M],$$

in particular independent from $M \in \Gamma_p$

□

Theorem 5.25 (Selfadjointness of the Hecke operators). *For each even weight $k \geq 0$ the Hecke operators*

$$T_m \in \text{End}(S_k(\Gamma)), \quad m \in \mathbb{N},$$

are self-adjoint with respect to the Petersson scalar product, i.e. for all $f, g \in S_k(\Gamma)$

$$\langle T_m f, g \rangle = \langle f, T_m g \rangle$$

Proof. Due to Theorem 5.17 it suffices to prove the claim for $m = p$ prime. Due to Lemma 5.23 there exists a complete set $V(p)$ of representatives of the Γ -action

$$\Phi_{\text{left}} : \Gamma \times \Gamma_p \rightarrow \Gamma_p$$

such that also $V(p)^\sharp$, the set of adjoints of the matrices from $V(p)$, is a complete set of representatives.

i) *Transformation formula for the integrand:* To simplify the notations we set for $M \in V(p)$

$$\Omega_{M, \text{left}} := \Omega(f[M]_k, g) \text{ and } \Omega_{M^\sharp, \text{right}} := \Omega(f, g[M^\sharp]_k)$$

Explicit computation with the definition of the $[-]_k$ -operation on $M_k(\Gamma)$ proves

$$\Omega_{M, \text{left}} = \Omega_{M^\sharp, \text{right}} \circ M,$$

see [38, Chap. IV, § 3.5].

ii) *Existence of the integrals:* Denote by

$$\mathcal{F} = \dot{\bigcup}_{\gamma} \gamma(\mathcal{D})$$

the fundamental domain of the fixed group $\tilde{\Gamma}(p)$ resulting from a finite coset representation of Γ with respect to the subgroup $\tilde{\Gamma}(p) \subset \Gamma$. We show for $M \in V(p)$

$$\int_{\mathcal{F}} \Omega_{M, left} d\mu \text{ exists}$$

by proving for $\gamma \in \Gamma$

$$\int_{\gamma(\mathcal{D})} \Omega_{M, left} d\mu \text{ exists :}$$

Similarly to the proof of Theorem 5.22 one checks on $\mathcal{D}$ an analogue of the transformation formula. First,

$$\Omega_{M, left} \circ \gamma = \Omega(f[M]_k, g) \circ \gamma = \Omega((f[M]_k)[\gamma]_k, g[\gamma]_k) = \Omega(f[M \cdot \gamma]_k, g[\gamma]_k)$$

Due to Proposition 5.7 there exists a matrix $A \in \Gamma$ with

$$A \cdot (M \cdot \gamma) = L \text{ with an upper triangular matrix } L \in V(p)$$

Set

$$\gamma_{left} := A^{-1} \in \Gamma,$$

then

$$M \cdot \gamma = \gamma_{left} \cdot L.$$

Secondly, using that f and g are modular forms with respect to Γ

$$\begin{aligned} \Omega(f[M \cdot \gamma]_k, g[\gamma]_k) &= \Omega(f[\gamma_{left} \cdot L]_k, g[\gamma]_k) = \Omega((f[\gamma_{left}])[L]_k, g[\gamma]_k) = \\ &= \Omega(f[M]_k, g) = \Omega_{L, left} \end{aligned}$$

As a consequence

$$\Omega_{M, left} \circ \gamma = \Omega_{L, left}$$

and

$$\int_{\gamma(\mathcal{D})} \Omega_{M, left} d\mu = \int_{\mathcal{D}} \Omega_{M, left} \circ \gamma d\mu = \int_{\mathcal{D}} \Omega_{L, left} d\mu$$

Because L is an upper triangular matrix the Fourier series of the cusp form f can easily be evaluated at $L(\tau)$, $\tau \in \mathbb{H}$, while the automophy factor $h(L, -)$ is a constant and $\det L = p$. An analogous estimate as in the proof of Theorem 5.22, part ii) shows the boundedness of the integrand on $\mathcal{D}$. Hence

$$\int_{\gamma(\mathcal{D})} \Omega_{M, left} d\mu = \int_{\mathcal{D}} \Omega_{L, left} d\mu < \infty$$

iii) *Left-right switching of $[-]_k$* : For given $M \in V(p)$

$$\int_{\mathcal{F}} \Omega_{M, \text{left}} d\mu = [\Gamma(\Omega_{M, \text{left}}) : \tilde{\Gamma}(p)] \cdot \int_{\mathcal{D}(\Gamma(\Omega_{M, \text{left}}))} \Omega_{M, \text{left}} d\mu$$

because

$$\tilde{\Gamma}(p) \subset \Gamma(\Omega_{M, \text{left}})$$

has finite index according to Lemma 5.24, part 2). Concerning the indices Lemma 5.24, part 3) implies

$$[\Gamma : \tilde{\Gamma}(p)] = [\Gamma : M^{-1}\tilde{\Gamma}(p)M]$$

hence also

$$[\Gamma(\Omega_{M, \text{left}}) : \tilde{\Gamma}(p)] = [\Gamma(\Omega_{M, \text{left}}) : M^{-1}\tilde{\Gamma}(p)M]$$

Using the transformation formula from part i) and taking $M^{-1}\mathcal{F}$ as a fundamental domain of $M^{-1}\tilde{\Gamma}(p)M$ the calculation continues

$$\begin{aligned} & [\Gamma(\Omega_{M, \text{left}}) : \tilde{\Gamma}(p)] \cdot \int_{\mathcal{D}(\Gamma(\Omega_{M, \text{left}}))} \Omega_{M, \text{left}} d\mu = \\ &= [\Gamma(\Omega_{M, \text{left}}) : M^{-1}\tilde{\Gamma}(p)M] \cdot \int_{\mathcal{D}(\Gamma(\Omega_{M, \text{left}}))} \Omega_{M, \text{left}} d\mu = \\ & \int_{M^{-1}\mathcal{F}} \Omega_{M, \text{left}} d\mu = \int_{M^{-1}\mathcal{F}} \Omega_{M^\sharp, \text{right}} \circ M d\mu = \int_{\mathcal{F}} \Omega_{M^\sharp, \text{right}} d\mu \end{aligned}$$

iv) *Left-right switching of the Hecke operator:*

$$\langle T_p f, g \rangle = \int_{\mathcal{D}} (T_p f \cdot \bar{g})(\tau) \cdot (Im \tau)^k d\mu = \int_{\mathcal{D}} \Omega(T_p f, g) d\mu$$

For any pair of cusp forms $T_p f$ and g the integrand is invariant under the action of Γ as checked during the proof of Theorem 5.22, part i). Hence continuing

$$\begin{aligned} & \int_{\mathcal{D}} \Omega(T_p f, g) d\mu = \frac{1}{[\Gamma : \tilde{\Gamma}(p)]} \cdot \int_{\mathcal{F}} \Omega(T_p f, g) d\mu = \\ &= \frac{1}{[\Gamma : \tilde{\Gamma}(p)]} \cdot \int_{\mathcal{F}} \sum_{M \in V(p)} \Omega_{M, \text{left}} d\mu = \frac{1}{[\Gamma : \tilde{\Gamma}(p)]} \cdot \sum_{M \in V(p)} \int_{\mathcal{F}} \Omega_{M, \text{left}} d\mu \end{aligned}$$

Using the result from part iii) and the property of $V(p)^\sharp$ the calculation continues

$$\begin{aligned} & \frac{1}{[\Gamma : \tilde{\Gamma}(p)]} \cdot \sum_{M \in V(p)} \int_{\mathcal{F}} \Omega_{M, \text{left}} d\mu = \frac{1}{[\Gamma : \tilde{\Gamma}(p)]} \cdot \sum_{M \in V(p)} \int_{\mathcal{F}} \Omega_{M^\sharp, \text{right}} d\mu = \\ & \frac{1}{[\Gamma : \tilde{\Gamma}(p)]} \cdot \int_{\mathcal{F}} \sum_{M \in V(p)} \Omega_{M^\sharp, \text{right}} d\mu = \frac{1}{[\Gamma : \tilde{\Gamma}(p)]} \cdot \int_{\mathcal{F}} \Omega(f, T_p g) d\mu = \end{aligned}$$

$$\square \quad = \int_{\mathcal{D}} \Omega(f, T_p g) d\mu = \int_{\mathcal{D}} (f \cdot \overline{T_p g})(\tau) \cdot (Im \tau)^k d\mu = \langle f, T_p g \rangle$$

Theorem 5.26 (Eigenforms of the Hecke algebra of the modular group). *For each even weight $k \geq 0$ the cusp space $S_k(\Gamma)$ has a basis of eigenforms of the Hecke algebra $Hecke_k(\Gamma)$.*

Proof. First, Theorem 5.17 implies that $Hecke_k(\Gamma)$ is commutative. Secondly, Theorem 5.25 implies that all Hecke operators T_m , $m \in \mathbb{N}$, are selfadjoint endomorphisms of the unitary vector space $(S_k(\Gamma), \langle -, - \rangle)$. Hence the claim follows due to a result from linear algebra about simultaneous diagonalizability of a family of pairwise commuting, self-adjoint endomorphisms.

□

Remark 5.27 (Hecke theory of congruence subgroups). For each level $N \in \mathbb{N}$ and even weight $k \geq 0$ one can define Hecke operators on modular forms of $M_k(\Gamma_0(N))$. The corresponding Hecke algebras are commutative, and each vector space of cusp forms of fixed weight has an eigenbasis with respect to the family

$$T_m, m \in \mathbb{N} \text{ and } \gcd(N, m) = 1.$$

For a proof see [36, Chap. 4, Theor. 4.22], [19], [35, Chap. IX, §6].

One should not consider a Hecke congruence subgroup $\Gamma_0(N)$ as an isolated object of a fixed level N . Instead one should take into consideration the whole family of Hecke congruence subgroups on different levels

$$\Gamma_0(N), N \in \mathbb{N}$$

For each weight $k \in \mathbb{N}$ and positive integer $m \in \mathbb{N}$ the modular spaces of different levels

$$M_k(\Gamma_0(M)) \text{ and } M_k(\Gamma_0(m \cdot M))$$

are related. Proposition 5.28 constructs explicitly a map which changes levels.

Proposition 5.28 (Increasing the level $M \mapsto m \cdot M$). *Consider a weight $k \in \mathbb{N}$, a level $M \in \mathbb{N}$ and a positive integer $m \in \mathbb{N}$. Set*

$$\mu_m := \begin{pmatrix} m & 0 \\ 0 & 1 \end{pmatrix} \in \Gamma_m$$

The map

$$j : M_k(\Gamma_0(M)) \rightarrow M_k(\Gamma_0(m \cdot M)), f \mapsto f[\mu_m]_k,$$

is a well-defined injection and restricts to an inclusion of cusp forms. The level change j respect cusp forms.

Under the assumptions of Proposition 5.28 one has apparently also the injection

$$M_k(\Gamma_0(M)) \hookrightarrow M_k(\Gamma_0(m \cdot M)).$$

Proof. 1. *Conjugation of $\Gamma_0(M)$:* Denote by

$$\Gamma' := \mu_m^{-1} \cdot \Gamma_0(M) \cdot \mu_m := \{\mu_m^{-1} \cdot \gamma \cdot \mu_m : \gamma \in \Gamma_0(M)\} \subset GL(2, \mathbb{Q})^+$$

the μ_m -conjugate of $\Gamma_0(M)$. If

$$\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_0(M)$$

then

$$\mu_m^{-1} \cdot \gamma \cdot \mu_m = \frac{1}{m} \cdot \begin{pmatrix} 1 & 0 \\ 0 & m \end{pmatrix} \cdot \gamma \cdot \begin{pmatrix} m & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} a & b/m \\ cm & d \end{pmatrix} \in GL(2, \mathbb{Q})^+$$

Claim: For $N := m \cdot M$ holds

$$\Gamma' \cap \Gamma_0(M) = \Gamma_0(N).$$

The inclusion

$$\Gamma' \cap \Gamma_0(M) \subset \Gamma_0(N)$$

is obvious because

$$M|c \implies m \cdot M|c \cdot m, \text{ i.e. } N|c \cdot m$$

To prove the opposite inclusion

$$\Gamma_0(N) \subset \Gamma' \cap \Gamma_0(M) :$$

Each matrix

$$\beta = \begin{pmatrix} a & b \\ cN & d \end{pmatrix} = \begin{pmatrix} a & b \\ cmM & d \end{pmatrix} \in \Gamma_0(N)$$

has the form

$$\beta = \mu_m^{-1} \cdot \gamma \cdot \mu_m \text{ with } \gamma := \begin{pmatrix} a & mb \\ cM & d \end{pmatrix}.$$

Here

$$\det \gamma = \det \beta = 1, \text{ hence } \gamma \in \Gamma_0(M).$$

As a consequence

$$\Gamma_0(N) \subset \Gamma' \text{ and even } \Gamma_0(N) \subset (\Gamma' \cap \Gamma_0(M))$$

2. *Changing right-multiplication to left-multiplication:* Claim: For each $\alpha \in \Gamma$ exists $\alpha_{left} \in \Gamma$ such that

$$\mu_m \cdot \alpha = \alpha_{left} \cdot \begin{pmatrix} * & * \\ 0 & * \end{pmatrix} \in M(2 \times 2, \mathbb{Z}) \cap GL(2, \mathbb{Q})^+.$$

For the proof set

$$M := \mu_m \cdot \alpha \in \Gamma_m$$

By Proposition 5.8 exists a matrix $A \in \Gamma$ such that

$$A \cdot M \in \Gamma_m$$

is triangular. Set

$$\alpha_{left} := A^{-1} \in \Gamma$$

3. *Weak modularity:* Assume $f \in M_k(\Gamma_0(M))$ and set

$$g := f[\mu_m]_k.$$

Due to part 1) a given $\beta \in \Gamma_0(N)$ has the form

$$\beta = \mu_m^{-1} \cdot \gamma \cdot \mu_m$$

with a suitable $\gamma \in \Gamma_0(M)$. Then

$$\begin{aligned} g[\beta]_k &= (f[\mu_m]_k)[\beta]_k = f[\mu_m \cdot \beta]_k = f[\mu_m \cdot \mu_m^{-1} \cdot \gamma \cdot \mu_m]_k = \\ &= f[\gamma \cdot \mu_m]_k = (f[\gamma]_k)[\mu_m]_k = f[\mu_m]_k = g \end{aligned}$$

Hence $f[\mu_m]_k$ is weakly modular with respect to $\Gamma_0(N)$.

4. *Holomorphy at the cusps:* According to the definition one has to show: For each

$$g := f[\mu_m]_k, f \in M_k(\Gamma_0(M)),$$

and each $\alpha \in \Gamma$ the function $g[\alpha]_k$ is holomorphic at $\infty \in \mathbb{H}^*$ according to the definition of $M_k(\Gamma_0(N))$. Due part 2) there exists a matrix

$$\alpha_{left} \in \Gamma$$

such that

$$\mu_m \cdot \alpha = \alpha_{left} \cdot \theta$$

with a triangular matrix

$$\theta = \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} \in M(2 \times 2, \mathbb{Z}) \cap GL(2, \mathbb{Q})^+$$

Hence

$$g[\alpha]_k = f[\mu_m \cdot \alpha]_k = f[\alpha_{left} \cdot \theta]_k = (f[\alpha_{left}]_k)[\theta]_k$$

From the Fourier expansion

$$f[\alpha_{\text{left}}]_k = \sum_{n=0}^{\infty} a_n \cdot \hat{q}^n, \quad \hat{q} = e^{2\pi i \cdot \tau/h}$$

for a suitable $h \in \mathbb{N}$ one obtains as a consequence

$$g[\alpha]_k(\tau) = f[\alpha_{\text{left}}]_k[\theta]_k(\tau) = m^{k-1} \cdot d^{-k} \cdot \sum_{n=0}^{\infty} a_n \cdot e^{2\pi i n \cdot (b/(dh))} \cdot \tilde{q}^n, \quad \tilde{q} = e^{2\pi i (\tau/h) \cdot (a/d)}.$$

Hence $g[\alpha]_k$ is holomorphic at ∞ . And the last Fourier series shows that j maps cusp forms to cusp forms.

□

If we fix the level N then each prime factor $p|N$ defines a level changing

$$(N/p) \mapsto N, \text{ using } \mu_p = \begin{pmatrix} p & 0 \\ 0 & 1 \end{pmatrix} \in \Gamma_p.$$

The subspace of $M_k(\Gamma_0(N))$ generated by the images of the maps

$$j_{p,(N/p),k} : S_k(\Gamma_0(N/p)) \times S_k(\Gamma_0(N/p)) \rightarrow S_k(\Gamma_0(N)), \quad (f, g) \mapsto f + g[\mu_p]_k,$$

for all prime factors $p|N$ is named the *oldspace*, see Definition 5.29.

Definition 5.29 (Oldform and newform of cusp forms). Consider a level $N \in \mathbb{N}^*$ and a weight k .

1. The image

$$\left(\sum_{p|N} \text{im } j_{p,(N/p),k} \right) \subset S_k(\Gamma_0(N))$$

is named the subspace

$$S_k^{\text{old}}(\Gamma_0(N)) \subset S_k(\Gamma_0(N))$$

of *oldforms*. The summation is indexed by the primes p dividing N .

2. The orthogonal complement of $S_k^{\text{old}}(\Gamma_0(N))$ with respect to the Petersson scalar product is named the subspace

$$S_k^{\text{new}}(\Gamma_0(N)) := S_k^{\text{old}}(\Gamma_0(N))^{\perp} \subset S_k(\Gamma_0(N))$$

of *newforms*.

Apparently, by the same mappings one can also define the oldspace of modular forms as a subspace of $M_k(\Gamma_0(N))$. But unfortunately the Petersson scalar product does not extend from the cusp forms to all modular forms. Hence there is no orthogonal complement of subspace of old forms.

Example 5.30 (Oldforms and newforms). See PARI-file `Congruence_subgroup_20`. We investigate the level change for the first weights $k = 0, 2, \dots, 12$.

- *Congruence subgroup $\Gamma_0(2)$:* Figure 5.2 displays for the Hecke congruence subgroup $\Gamma_0(2)$ of level $N = 2$ the splitting of the cusp space into oldforms and newforms.

According to Definition 5.29 one has to consider for the splitting the single prime factor $p = 2$ and the cusp forms $S_k(\Gamma_0(1))$, i.e. the cusp forms of the full modular group $\Gamma = \Gamma_0(1)$. They inject by the identity id and by $2 \cdot id$, which doubles the argument.

For weight $k = 12$ one has the 1-dimensional subspace

$$S_{12}(\Gamma) = \mathbb{C} \cdot \Delta \subset S_{12}(\Gamma_0(2))$$

In addition also the modular form

$$\Delta_2 : \mathbb{H} \rightarrow \mathbb{C}, \tau \mapsto \Delta(2 \cdot \tau)$$

belongs to $S_{12}(\Gamma_0(2))$. The computation shows

$$S_{12}(\Gamma_0(2)) = \text{span}_{\mathbb{C}} \langle \Delta, \Delta_2 \rangle,$$

all cusp forms in $S_{12}(\Gamma_0(2))$ are old forms induced from $S_{12}(\Gamma)$.

- *Congruence subgroup $\Gamma_0(4)$:* Similarly Figure 5.3 considers the modular group $\Gamma_0(4)$ of level $N = 4$. One has to consider the two level changes

$$1 \mapsto 4, \text{ and } 2 \mapsto 4$$

For weight $k = 12$: As expected the oldforms of $\Gamma_0(4)$ comprise the 2-dimensional space induced from the cusp forms of $\Gamma_0(2)$ and the 1-dimensional space induced from the cusp form of Γ . In addition there is a 1-dimensional space of new forms.

- *Congruence subgroup $\Gamma_0(11)$:* The final Figure 5.4 considers the modular group $\Gamma_0(11)$ of level $N = 11$. For the splitting one has to consider the level change $1 \mapsto 11$. For weight $k = 12$: The old forms of $\Gamma_0(11)$ derive from the cusp form Δ via the identity id and via multiplying by 4 the argument. In addition there is an 8-dimensional space of new forms.

```

=====
Congruence subgroups_20, Start

Modular forms, cusp forms, newforms,

Congruence subgroup Gamma_0(2): Modular forms, oldforms, newforms

Index [Gamma:Gamma_0(2)] = 3

Right cosets: [[0, -1; 1, 0], [1, 0; 1, 1], [1, 0; 2, 1]]

Cusps of Gamma_0(2): [0, 1/2]

Weight k: (dim M_k(Gamma_0(2), dim S_k_old(Gamma_0(2), dim S_k_new(Gamma_0(2)))

0: (1, 0, 0)
2: (1, 0, 0)
4: (2, 0, 0)
6: (2, 0, 0)
8: (3, 0, 1)
10: (3, 0, 1)
12: (4, 2, 0)

Congruence subgroups_20, End
=====

```

Fig. 5.2 $\Gamma_0(2)$ oldforms and newforms

```

=====
Congruence subgroups_20, Start

Modular forms, cusp forms, newforms,

Congruence subgroup Gamma_0(4): Modular forms, oldforms, newforms

Index [Gamma:Gamma_0(4)] = 6

Right cosets: [[0, -1; 1, 0], [1, 0; 1, 1], [0, -1; 1, 2], [0, -1; 1, 3], [1, 0; 2, 1], [1, 0; 4, 1]]

Cusps of Gamma_0(4): [0, 1/2, 1/4]

Weight k: (dim M_k(Gamma_0(4), dim S_k_old(Gamma_0(4), dim S_k_new(Gamma_0(4)))

0: (1, 0, 0)
2: (2, 0, 0)
4: (3, 0, 0)
6: (4, 0, 1)
8: (5, 2, 0)
10: (6, 2, 1)
12: (7, 3, 1)

Congruence subgroups_20, End
=====

```

Fig. 5.3 $\Gamma_0(4)$ oldforms and newforms

```

=====
Congruence subgroups_20, Start
Modular forms, cusp forms, newforms,
Congruence subgroup Gamma_0(11): Modular forms, oldforms, newforms
Index [Gamma:Gamma_0(11)] = 12
Right cosets: [[0, -1; 1, 0], [1, 0; 1, 1], [0, -1; 1, 2], [0, -1; 1, 3], [0, -1; 1, 4], [0, -1; 1, 5], [0, -1; 1, 6], [0, -1; 1, 7], [
0, -1; 1, 8], [0, -1; 1, 9], [0, -1; 1, 10], [1, 0; 11, 1]]
Cusps of Gamma_0(11): [0, 1/11]
Weight k: (dim M_k(Gamma_0(11)), dim S_k_old(Gamma_0(11)), dim S_k_new(Gamma_0(11)))
0: (1, 0, 0)
2: (2, 0, 1)
4: (4, 0, 2)
6: (6, 0, 4)
8: (8, 0, 6)
10: (10, 0, 8)
12: (12, 2, 8)
Congruence subgroups_20, End
=====

```

Fig. 5.4 $\Gamma_0(11)$ oldforms and newforms

5.3 Numerology: Lagrange, Jacobi, Ramanujan, Mordell

The present section draws some conclusions from the existence of simultaneous eigenforms of the Hecke operators. First we prove a conjecture of Ramanujan concerning a certain numerical function. Secondly, we derive the four squares theorem.

The first non-trivial cusp form, the modular invariant $\Delta \in S_{12}(\Gamma)$, encodes the Ramanujan τ -function: The values $\tau(n)$, $n \in \mathbb{N}$, are the coefficients of the normalized modular invariant. The question, how the value $\tau(n)$ depends on the prime divisors p of n , was the subject of intense investigation. Decades after Ramanujan made his conjecture on the basis of the first 30 values, it was Deligne who proved the final part of the conjecture about the modulus $|\tau(p)|$ as a consequence of his full proof of the Weyl conjectures.

Definition 5.31 (Ramanujan τ -function). The Fourier coefficients $(\tau_n)_{n \in \mathbb{N}}$ of the normalized discriminant form

$$\frac{\Delta(\tau)}{(2\pi)^{12}} = \sum_{n=1}^{\infty} \tau_n \cdot q^n, \quad q = e^{2\pi i \cdot \tau},$$

define the *Ramanujan τ -function*

$$\tau : \mathbb{N} \rightarrow \mathbb{Z}, n \mapsto \tau_n.$$

According to Proposition 3.14 all Fourier coefficients are integers $\tau(n) \in \mathbb{Z}$.

Computing the first coefficients $\tau(n)$ up to $n = 6$ gives

$$\frac{\Delta(\tau)}{(2\pi)^{12}} = q - 24q^2 + 252q^3 - 1.472q^4 + 4.830q^5 - 6.048q^6 + O(7)$$

From these values one verifies at once the product formula

$$\tau(2) \cdot \tau(3) = -24 \cdot 252 = -6.048 = \tau(6),$$

and as a particular case of the recursion formula

$$\tau(p^2) = \tau(p)^2 - p^{11} \cdot \tau(p^0)$$

the formula for $p = 2$

$$\tau(2^2) = -1.472 = \tau(2)^2 - 2^{11} = (-24)^2 - 2.048.$$

From explicit computation of the first 30 values of $\tau(n)$ Ramanujan [49] conjectured in 1916 several generalizations. Figure 5.5 shows equations (103) and (104) from the original paper with two of Ramanujan's conjectures.

Note. There is a typo in equation (104) referring to the factor 2. It should read

$$|\tau(p)| \leq 2 \cdot p^{11/2}$$

The conjecture was proved inter alia one year later by Mordell [45]. Figure 5.6 shows Mordell's introduction from his paper.

the form $o(n^{\frac{11}{2}})$. For it appears that

$$\sum_1^\infty \frac{\tau(n)}{n^t} = \prod_p \frac{1}{1 - \tau(p)p^{-t} + p^{11-2t}}. \quad (101)$$

This assertion is equivalent to the assertion that, if

$$n = p_1^{a_1} p_2^{a_2} p_3^{a_3} \cdots p_r^{a_r},$$

where $p_1, p_2, \dots, p_r$ are the prime divisors of n , then

$$n^{-\frac{11}{2}} \tau(n) = \frac{\sin(1+a_1)\theta_{p_1}}{\sin \theta_{p_1}} \frac{\sin(1+a_2)\theta_{p_2}}{\sin \theta_{p_2}} \cdots \frac{\sin(1+a_r)\theta_{p_r}}{\sin \theta_{p_r}}, \quad (102)$$

where

$$\cos \theta_p = \frac{1}{2} p^{-\frac{11}{2}} \tau(p).$$

It would follow that, if n and n' are prime to each other, we must have

$$\tau(nn') = \tau(n)\tau(n'). \quad (103)$$

Let us suppose that (102) is true, and also that (as appears to be highly probable)

$$\{2\tau(p)\}^2 \leq p^{11}, \quad (104)$$

Fig. 5.5 Ramanujan's conjecture on $\tau(n)$ from [49]

Mr Mordell, On Mr Ramanujan's Empirical Expansions, etc. 117

On Mr Ramanujan's Empirical Expansions of Modular Functions. By L. J. MORDELL, Birkbeck College, London. (Communicated by Mr G. H. HARDY.)

[Received 14 June 1917.]

In his paper* "On Certain Arithmetical Functions" Mr Ramanujan has found empirically some very interesting results as to the expansions of functions which are practically modular functions. Thus putting

$$\left(\frac{\omega_2}{2\pi}\right)^{12} \Delta(\omega_1, \omega_2) = r [(1-r)(1-r^2)(1-r^3)\dots]^{24} = \sum_{n=1}^{\infty} T(n) r^n,$$

he finds that

$$T(mn) = T(m) T(n) \dots \dots \dots (1)$$

if m and n are prime to each other; and also that

$$\sum_{n=1}^{\infty} \frac{T(n)}{n^s} = \prod 1/(1 - T(p)p^{-s} + p^{11-2s}) \dots \dots \dots (2),$$

where the product refers to the primes 2, 3, 5, 7 He also gives many other results similar to (2).

My attention was directed to these results by Mr Hardy, and I have found that results of this kind are a simple consequence of the properties of modular functions. In the case above

$$\Delta(\omega_1, \omega_2) \quad (r = e^{2\pi i \omega}, \quad \omega = \omega_1/\omega_2)$$

is the well-known modular invariant of dimensions -12 in ω_1, ω_2 , which is unaltered by the substitutions of the homogeneous modular group defined by

Fig. 5.6 Mordell's introduction to [45]

A film about Ramanujan is to be recommended, named “The man who knew infinity” (Deutsch: Die Poesie des Unendlichen).

Today Ramanujan’s product formula and recursion formula can be obtained as a first application of Hecke theory:

Proposition 5.32 (Product and recursion formula of the Ramanujan τ -function).

The Ramanujan τ -function satisfies:

1. Product formula: For all integers $m, n \geq 1$

$$\tau(m) \cdot \tau(n) = \sum_{r|(m,n)} r^{11} \cdot \tau(mn/r^2).$$

In particular, for coprime m and n :

$$\tau(m) \cdot \tau(n) = \tau(m \cdot n).$$

2. Recursion formula: For all primes p and integers $k \geq 1$

$$\tau(p) \cdot \tau(p^k) = \tau(p^{k+1}) + p^{11} \cdot \tau(p^{k-1}).$$

Proof. 1. Consider an arbitrary, but fixed $m \in \mathbb{N}$. Corollary 5.14 from Hecke theory for weight $k = 12$ derives the eigenvalue equation

$$T_m \left(\frac{\Delta}{(2\pi)^{12}} \right) = \tau(m) \cdot \frac{\Delta}{(2\pi)^{12}}$$

When equating for each fixed $n \in \mathbb{N}$ the terms of order n on both sides of the equation the coefficient formula from Theorem 5.12 implies:

$$\sum_{r|(m,n)} r^{11} \cdot \tau(mn/r^2) = \tau(m) \cdot \tau(n).$$

For coprime integers m, n we have $(m, n) = 1$ and the summation reduces to the single term for $r = 1$:

$$\tau(m) \cdot \tau(n) = \tau(m \cdot n)$$

2. We set $m = p^k$ and $n = p$ in the product formula from part 1). Due to

$$(p^k, p) = p$$

we now have two summands, referring to the indices $r = 1$ and $r = p$. We obtain

$$\tau(p^k) \cdot \tau(p) = \tau(p^{k+1}) + p^{11} \cdot \tau(p^{k-1})$$

□

Remark 5.33 (Growth estimation of the Ramanujan τ -function). The final conjecture of Ramanujan about the τ -function is the growth estimation

$$|\tau(p)| \leq 2 \cdot p^{11/2}, \quad p \text{ prime},$$

see Figure 5.5 equation (104) (after correction). This conjecture lies much deeper than Ramanujan's other conjectures. Deligne proved the growth condition of the τ -function as a consequence of his proof of the Weil conjectures in 1974 [57].

The next application of modular forms investigates the modular form θ of weight $= 2$ of the congruence subgroup $\Gamma_0(4) \subset \Gamma$. The Fourier coefficients $r(n)_{n \in \mathbb{N}}$ of θ count how many possibilities exist to represent the integer n as the sum of integer squares. Hence the Fourier series of θ gives the *Four squares theorem*.

Fermat recalls in a letter to a friend from 1659 those results from arithmetic which he considers his most important contribution to this field. One of these results is the claim that any natural number can be written as the sum of four squares. The first proof of this theorem has been given by Lagrange in 1770. The more refined version of Theorem 5.37 is due to Jacobi in 1834. For the history of these results see [51].

Definition 5.34 (Representing integers as sum of squares).]

1. *Multiplicity of square representations:* The number of possibilities to represent a non-negative integer n as the sum of k integer squares defines the arithmetic function

$$r : \mathbb{N}^* \times \mathbb{N}^* \rightarrow \mathbb{N}$$

with

$$r(n, k) := \text{card} \left\{ v = (v_1, \dots, v_k) \in \mathbb{Z}^k : n = v_1^2 + \dots + v_k^2 \right\}.$$

Note that $r(n, k)$ counts k -tuples with positive and negative components as different, and distinguishes also k -tuples with the same components, but in different order.

2. *Generating function:* The series

$$\vartheta(\tau, k) := \sum_{n=0}^{\infty} r(n, k) \cdot q^n, \quad q := e^{2\pi i \tau},$$

which is generated by the sequence $r(n, k)_{n \in \mathbb{N}}$, is named a ϑ -series.

Proposition 5.35 (Generating function as modular form). *The generating function $\theta(-, 4)$ from Definition 5.34 is a modular form of weight 2 of the congruence subgroup $\Gamma_0(4)$, i.e.*

$$\vartheta(-, 4) \in M_2(\Gamma_0(4)).$$

Its Fourier series starts

$$\vartheta(\tau, 4) = 1 + 8q + O(2), \quad q = e^{2\pi i\tau}.$$

Proof. Scetch, see cf. [19, Chap.1, §2; Chap. 4, §9], [38, Chap. III.7 Korollar].

i) The congruence subgroup $\Gamma_0(4)$ has the two generators

$$T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \text{ and } \gamma = \begin{pmatrix} 1 & 0 \\ 4 & 1 \end{pmatrix}$$

ii) The translation T apparently satisfies

$$\vartheta(-, 4)[T]_2 = \vartheta(-, 4)$$

With

$$\vartheta(\tau) := \vartheta(\tau, 1) = \sum_{n=0}^{\infty} r(n, 1) \cdot q^n, \quad q = e^{2\pi i\tau}$$

the product formula

$$\vartheta(\tau, 4) = \vartheta(\tau)^4$$

reduces the weak modularity

$$\vartheta(-, 4)[\gamma]_2 = \vartheta(-, 4)$$

to the transformation formula for ϑ : The difficult part is to prove

$$\vartheta(-1/(4\tau)) = \sqrt{-2i\tau} \cdot \vartheta(\tau),$$

see [19, Chap. IV, 4.9], [38, Kap. III, E.3]. Then follows

$$\begin{aligned} \vartheta\left(\frac{\tau}{4\tau+1}\right) &= \vartheta\left(\frac{-1}{4(-1/(4\tau)-1)}\right) = \sqrt{2i\left(\frac{1}{4\tau}+1\right)} \cdot \vartheta\left(-\frac{1}{4\tau}-1\right) = \\ &= \sqrt{2i\left(\frac{1}{4\tau}+1\right)} \cdot \vartheta\left(-\frac{1}{4\tau}\right) = \sqrt{2i\left(\frac{1}{4\tau}+1\right)} \cdot (-2i\tau) \cdot \vartheta(\tau) = \sqrt{4\tau+1} \cdot \vartheta(\tau), \end{aligned}$$

and exponentiation

$$\vartheta(-, 4) = \vartheta^4$$

shows the transformation formula

$$\vartheta\left(\frac{\tau}{4\tau+1}, 4\right) = (4\tau+1)^2 \cdot \vartheta(\tau, 4) \text{ i.e. } \vartheta(-, 4)[\gamma]_2 = \vartheta(-, 4)$$

iii) The Fourier series

$$\vartheta(\tau) = 1 + r(1, 1) \cdot q + O(q^2) = 1 + 2 \cdot q + O(q^2)$$

shows

$$\vartheta(\tau, 4) = (1 + 2 \cdot q)^4 + O(2) = 1 + 8 \cdot q + O(2)$$

□

Theorem 5.36 (The modular form $G_{2,N} \in M_2(\Gamma_0(N))$). Denote by G_2 the Eisenstein series from Remark 3.12. For each level $N \geq 1$ set

$$\mu_N := \begin{pmatrix} N & 0 \\ 0 & 1 \end{pmatrix} \in M(2 \times 2, \mathbb{Z}) \cap GL(2, \mathbb{Q})^+$$

The function

$$G_{2,N} := G_2 - G_2[\mu_N]_2 : \mathbb{H} \rightarrow \mathbb{C}$$

is a modular function

$$G_{2,N} \in M_2(\Gamma_0(N))$$

Proof. By definition

$$(G_2[\mu_N]_2)(\tau) := G_2(\mu_N(\tau)) \cdot \det \mu_N \cdot h(\mu_N, \tau)^{-2} = N \cdot G_2(N\tau)$$

i) *Holomorphy:* Due to Remark 3.12 the functions G_2 and hence also $G_{2,N}$ are holomorphic in $\mathbb{H} \cup \{\infty\}$.

ii) *Weak modularity:* We have to show for all $\alpha \in \Gamma_0(N)$:

$$G_{2,N}[\alpha]_2 = G_{2,N}.$$

We recall the formula from Remark 3.12

$$G_2[\gamma]_2(\tau) = G_2(\tau) - 2\pi i \cdot c \cdot h(\gamma, \tau)^{-2} \text{ for } \gamma := \alpha = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma$$

We have to compute the transforms of the two summands of $G_{2,N}$.

• $G_2[\alpha]_2$: Set

$$\gamma := \alpha = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_0(N) \subset \Gamma.$$

Then

$$G_2[\alpha]_2(\tau) = G_2(\tau) - 2\pi i \cdot c \cdot h(\alpha, \tau)^{-2}$$

• $(G_2[\mu_N]_2)[\alpha]_2$: Then

$$\mu_N \cdot \alpha = \beta \cdot \mu_N \text{ with } \beta := \begin{pmatrix} a & Nb \\ c/N & d \end{pmatrix}$$

Here $\beta \in \Gamma$ because

$$c \equiv 0 \pmod{N} \implies c/N \in \mathbb{Z}$$

Then

$$(G_2[\beta]_2)(z) = G_2(z) - 2\pi i \cdot (c/N) \cdot h(\beta, z)^{-2}$$

Hence with the argument $z = \mu_N(\tau) = N\tau$

$$(G_2[\beta]_2)(\mu_N(\tau)) = G_2(\mu_N(\tau)) - 2\pi i \cdot (c/N) \cdot h(\beta, \mu_N(\tau))^{-2}$$

Then

$$\begin{aligned} ((G_2[\mu_N]_2)[\alpha]_2)(\tau) &= (G_2[\mu_N \cdot \alpha]_2)(\tau) = (G_2[\beta \cdot \mu_N]_2)(\tau) = \\ &= ((G_2[\beta]_2)[\mu_N]_2)(\tau) = (G_2[\beta]_2)(\mu_N(\tau)) \cdot h(\mu_N, \tau)^{-2} \cdot N = \\ &= (G_2(\mu_N(\tau)) - 2\pi i \cdot (c/N) \cdot h(\beta, \mu_N(\tau))^{-2}) \cdot h(\mu_N, \tau)^{-2} \cdot N = \\ &= G_2(\mu_N(\tau)) \cdot h(\mu_N, \tau)^{-1} \cdot N - 2\pi i \cdot ((c/N) \cdot N) \cdot (h(\beta, \mu_N(\tau)) \cdot h(\mu_N, \tau))^{-2} = \\ &= (G_2[\mu_N]_2)(\tau) - 2\pi i \cdot ((c/N) \cdot N) \cdot (h(\beta, \mu_N(\tau)) \cdot h(\mu_N, \tau))^{-2} \end{aligned}$$

Now

$$h(\beta, \mu_N(\tau)) = h(\beta, N\tau) = (c/N) \cdot N\tau + d = c\tau + d \text{ and } h(\mu_N, \tau) = 1$$

and

$$h(\alpha, \tau) = c\tau + d$$

which implies

$$h(\beta, \mu_N(\tau)) = h(\alpha, \tau)$$

Hence

$$((G_2[\mu_N]_2)[\alpha]_2)(\tau) = (G_2[\mu_N]_2)(\tau) - 2\pi i \cdot c \cdot h(\alpha, \tau)^{-2}$$

As a consequence, the transformation formulas for both summands imply:

$$G_{2,N}[\alpha]_2 = G_2[\alpha]_2 - ((G_2[\mu_N]_2)[\alpha]_2) = G_2 - (G_2[\mu_N]_2) = G_{2,N}$$

iii) Finally one checks the Fourier expansion at the cusps using the Fourier expansion of G_2 from Remark 3.12, see [19, Chap. I, Exerc. 1.2.8.e] and [35, Chap. IX, §3, Example 4].

□

Theorem 5.37 (Four squares theorem). *For any $n \in \mathbb{N}$, $n \geq 1$, there are*

$$r(n, 4) = 8 \cdot \sum_{d|n, 4 \nmid d} d$$

possibilities to represent the non-negative integer n as the sum of 4 integer squares.

In particular for $4 \nmid n$

$$r(n, 4) = 8 \cdot \sigma_1(n) = 8 \cdot \sum_{d|n} d.$$

Proof. Computation within $M_2(\Gamma_0(4))$:

i) *Dimension:* $\dim M_2(\Gamma_0(4)) = 2$ according to Theorem 3.26, see also Figure 5.3.

ii) *Basis:* According to Theorem 5.36 for each $N \geq 2$ the Eisenstein series

$$G_{2,N}(\tau) := G_2(\tau) - N \cdot G_2(N\tau)$$

is a modular form of the congruence subgroup $\Gamma_0(N)$ of weight 2, i.e.

$$G_{2,N}(\tau) \in M_2(\Gamma_0(N)).$$

The Fourier expansion of G_2 from Remark 3.12 provides for the two Eisenstein series $G_{2,2}$ and $G_{2,4}$ the Fourier expansions

$$G_{2,2}(\tau) = -\frac{\pi^2}{3} \left(1 + 24 \cdot \sum_{n=1}^{\infty} \left(\sum_{\substack{d|n \\ d \text{ odd}}} d \right) q^n \right) = -\frac{\pi^2}{3} (1 + 24q + O(2))$$

and

$$G_{2,4}(\tau) = -\pi^2 \left(1 + 8 \cdot \sum_{n=1}^{\infty} \left(\sum_{d|n, 4 \nmid n} d \right) q^n \right) = -\pi^2 (1 + 8q + O(2)).$$

Comparing terms of order 0 and 1 in the Fourier expansions shows that both Eisenstein series are linearly independent. Hence the family $(G_{2,2}, G_{2,4})$ is a basis of $M_2(\Gamma_0(4))$. In particular, it generates $\theta(-, 4)$.

Proposition 5.35 implies by comparing Fourier coefficients of order 0 and 1

$$\theta(-, 4) = -\frac{1}{\pi^2} G_{2,4}.$$

Equating on both sides the Fourier coefficients of order $n \in \mathbb{N}^*$ shows

$$r(n, 4) = 8 \cdot \sum_{d|n, 4 \nmid d} d$$

□

Example 5.38 (Modular forms via ϑ -functions). Figure 5.7 shows for some pairs (k, e) the numbers

$$r(n, k, e).$$

Here $r(n, k, e) \in \mathbb{N}$ is the number of different representations of n as a sum of k summands, each an e -power. These numbers are the Fourier coefficients of suitable modular forms. The case

$$r(n, 4, 2) =: r(n, 4)$$

has been investigated in Theorem 5.37. The figure shows the output of the PARI-script `Modular_forms_theta_01`.

```

=====
Modular_forms_theta_01: Start.
Modular_forms_theta_01: Fourier coefficients of  $r(-,k,e)$  with  $r(n,k,e)$  the number,
of different representations of  $n$  as a sum of  $k$  summands, each an  $e$ -power
parameter: (k,e) = (1,2)
Underlying modular space:  $M_{1/2}(G_0(4, 1))$ 
Generating modular form:  $1 + 2^*q + 2^*q^4 + 2^*q^9 + O(q^{10})$ 
-----
parameter: (k,e) = (2,2)
Underlying modular space:  $M_1(G_0(4, -4))$ 
Generating modular form:  $1 + 4^*q + 4^*q^2 + 4^*q^4 + 8^*q^5 + 4^*q^8 + 4^*q^9 + O(q^{10})$ 
-----
parameter: (k,e) = (3,2)
Underlying modular space:  $M_{3/2}(G_0(4, 1))$ 
Generating modular form:  $1 + 6^*q + 12^*q^2 + 8^*q^3 + 6^*q^4 + 24^*q^5 + 24^*q^6 + 12^*q^8 + 30^*q^9 + O(q^{10})$ 
-----
parameter: (k,e) = (4,2)
Underlying modular space:  $M_2(G_0(4, 1))$ 
Generating modular form:  $1 + 8^*q + 24^*q^2 + 32^*q^3 + 24^*q^4 + 48^*q^5 + 96^*q^6 + 64^*q^7 + 24^*q^8 + 104^*q^9 + O(q^{10})$ 
-----
parameter: (k,e) = (8,2)
Underlying modular space:  $M_4(G_0(4, 1))$ 
Generating modular form:  $1 + 16^*q + 112^*q^2 + 448^*q^3 + 1136^*q^4 + 2016^*q^5 + 3136^*q^6 + 5504^*q^7 + 9328^*q^8 + 12112^*q^9 + O(q^{10})$ 
-----
Modular_forms_theta_01: End

```

Fig. 5.7 Modular functions to compute $r(-,k,e)$

Chapter 6

Modular functions and imaginary quadratic fields

Analytic number theory investigates and proves results from number theory by methods from complex analysis. A typical example is the proof of the class number formula of an imaginary quadratic field using the theory of modular functions.

Modular functions provide a sequence of finite-dimensional complex representations of the modular group $\Gamma_1 = SL(2, \mathbb{Z})$, depending on the weight $k \in \mathbb{N}$.

The modular j -invariant parametrizes the set of complex tori by the values of the affine part of the modular curve $X(\Gamma_1) \simeq \mathbb{P}$. All complex analytic properties of a given torus with normalized period lattice $(1, \tau)$ are encoded in the value $j(\tau) \in \mathbb{C}$ of the modular j -invariant. But $j(\tau)$ is too coarse to discriminate between the arithmetic properties of the different members of a class of biholomorphic equivalent tori.

Section 6.2 restricts to tori with complex multiplication and asks: Which arithmetic information can be obtained when considering besides $j(\tau)$ also the values

$$j(\alpha(\tau)), \alpha \in \Gamma_{m, \text{prim}}, m \in \mathbb{N}^*?$$

General references for the present chapter are the lecture notes by Zagier [72, Chap. 6.1] and the essay by Serre [54].

6.1 Complex multiplication

As a first arithmetic property of complex tori and elliptic curves we observe: Complex tori with normalized lattice generated by $(1, \tau)$ relate for certain numbers

$\tau \in \mathbb{H}$ to an imaginary quadratic field. How to characterize tori with those arithmetic properties? The answer is given by Theorem 6.6.

As a general reference for number fields we recommend [24], [42], [34].

Definition 6.1 (Number field).

1. A field K is a *number field* if

$$\mathbb{Q} \subset K \subset \mathbb{C} \text{ with degree } \deg K := [K : \mathbb{Q}] < \infty.$$

If $\deg K = 2$ then K is a *quadratic number field*.

2. The *Galois group* $\text{Gal}(K/\mathbb{Q})$ of a number field K is the group of all field automorphism of K which pointwise fix the elements of $\mathbb{Q}$.

Remark 6.2 (Number field). Consider a number field K of degree

$$n = [K : \mathbb{Q}]$$

1. Denote by

$$(\alpha_j)_{j=1, \dots, n}$$

a $\mathbb{Q}$ -basis of K . For each $\alpha \in K$ the multiplication

$$\mu_\alpha : K \rightarrow K, x \mapsto \alpha \cdot x,$$

is determined by the values

$$\mu_\alpha(\alpha_j) = \sum_{k=1, \dots, n} \alpha_{jk} \cdot \alpha_k, \quad j = 1, \dots, n.$$

For $\alpha \in K$ *norm* and *trace* are respectively defined as

$$N_K(\alpha) := \det(\alpha_{jk}) \in \mathbb{Q} \text{ and } \text{tr}_K(\alpha) := \sum_{j=1, \dots, n} \alpha_{jj} \in \mathbb{Q}$$

2. The *discriminant* of K with respect to the basis $(\alpha_j)_{j=1, \dots, n}$ is defined as

$$\Delta_K := \det(\text{tr}_K(\alpha_j \cdot \alpha_k)) \in \mathbb{Q}$$

The discriminant depends on the choice of the basis: It transforms with the square of the determinant of the base change matrix.

3. The ring of *algebraic integers* in K , the *number ring* of K , is the ring of integers

$$\mathcal{O}_K := \{x \in K : x \text{ integral over } \mathbb{Z}\}$$

It is a lattice in K , and K is the quotient field of $\mathcal{O}_K$.

The ring $\mathcal{O}_K$ is the prototype of a *Dedekind ring*, i.e. $\mathcal{O}_K$ is a Noetherian, 1-dimensional integral domain, integrally closed in its quotient field.

4. An *order* of K is a subring $\Lambda \subset \mathcal{O}_K$ which is a free $\mathbb{Z}$ -module of rank $= \deg K$.
5. The number field K has a $\mathbb{Q}$ -basis $(\alpha_j)_{j=1,\dots,n}$ by algebraic integers $\alpha_j \in \mathcal{O}_K$. Each element $\alpha \in K$ has a unique representation

$$\alpha = \frac{\mathbb{Z}\alpha_1 + \dots + \mathbb{Z}\alpha_n}{\Delta_K}$$

6. For a quadratic field $K = \mathbb{Q}(\sqrt{d})$ with a squarefree integer $d \in \mathbb{Z}$ one has

$$\Delta_K = \begin{cases} 4 \cdot d & \text{if } d \equiv 2, 3 \pmod{4} \\ d & \text{if } d \equiv 1 \pmod{4} \end{cases}$$

and

$$\mathcal{O}_K = \mathbb{Z} \oplus \mathbb{Z} \left[\frac{\Delta_K + \sqrt{\Delta_K}}{2} \right] = \mathbb{Z} \oplus \mathbb{Z} \cdot \begin{cases} \sqrt{d} & d \equiv 2, 3 \pmod{4} \\ \frac{1 + \sqrt{d}}{2} & d \equiv 1 \pmod{4} \end{cases}$$

7. Each number field K has a *primitive element*, i.e.

$$K = \mathbb{Q}(\alpha)$$

with a suitable $\alpha \in K$. Each primitive element satisfies a polynomial equation

$$P(\alpha) = 0 \text{ with } P(X) \in \mathbb{Q}[X] \text{ and } \deg P = \deg K.$$

8. Each quadratic number field has the form

$$K = \mathbb{Q}(\sqrt{d}) \text{ with } d \in \mathbb{Z} \text{ squarefree.}$$

If $d < 0$ then K is named *imaginary quadratic*, if $d > 0$ then K is named *real quadratic*.

9. The Galois group $Gal(K/\mathbb{Q})$ of a quadratic number field $K = \mathbb{Q}(\sqrt{d})$ equals the group

$$\mathbb{Z}_2 \simeq \{\sigma_0, \sigma_1\}$$

The generator

$$\sigma_1 : K \rightarrow K$$

is

$$\sigma_1(z) = \begin{cases} -z & d > 0 \\ \bar{z} & d < 0 \end{cases}$$

while

$$\sigma_0 = id$$

Norm and trace of an element $z \in K$ are obtained from the two elementary symmetric functions

$$N(z) = \sigma_0(z) \cdot \sigma_1(z) \text{ and } Tr(z) = \sigma_0(z) + \sigma_1(z)$$

Note. Due to historical reasons the symbol $\mathcal{O}$ has at least two different meanings in mathematics. In the context of complex analysis $\mathcal{O}$ also denotes the complex structure sheaf.

Definition 6.3 (Endomorphisms of a torus). An *endomorphism* f of a torus

$$(T, 0) = \mathbb{C}/\Lambda$$

is a holomorphic map

$$f : (T, 0) \rightarrow (T, 0)$$

with $f(0) = 0$.

Recall from Proposition 2.3 that each endomorphism f of a torus fits into a commutative diagram

$$\begin{array}{ccc} \mathbb{C} & \xrightarrow{\mu_\alpha} & \mathbb{C} \\ p \downarrow & & \downarrow p \\ \mathbb{C}/\Lambda & \xrightarrow{f} & \mathbb{C}/\Lambda \end{array}$$

with a unique multiplication map μ_α determined by an element $\alpha = \alpha_f \in \mathbb{C}$: For all $z \in \mathbb{C}$

$$\mu_\alpha(z) = \alpha \cdot z$$

Proposition 6.4 (Endomorphism ring). *The set of endomorphisms of a torus*

$$T = \mathbb{C}/\Lambda$$

constitutes with respect to addition and composition a ring $\text{End}(T)$, the endomorphism ring of T . The map

$$\text{End}(T) \rightarrow \{\alpha \in \mathbb{C} : \alpha \cdot \Lambda \subset \Lambda\}, f \mapsto \alpha_f,$$

is an isomorphism of rings.

If not stated otherwise by the above ring isomorphism we identify the endomorphism ring $\text{End}(T)$ with a subring of $\mathbb{C}$. Because each torus T is an Abelian group its endomorphism ring $\text{End}(T)$ comprises at least the group $\mathbb{Z}$, the scalar multiplication by integer numbers:

$$\mathbb{Z} \subset \text{End}(T) \subset \mathbb{C}.$$

Which intermediate rings are possible? The case

$$\text{End}(T) = \mathbb{Z}$$

is the general case, justifying a specific name for the exceptional tori with additional endomorphisms.

Definition 6.5 (Complex multiplication (CM)).

1. A complex number $\tau \in \mathbb{H}$ is a *CM-point* and its class $[\tau] \in X(\Gamma)$ in the modular curve is a *CM-class*, if the corresponding torus

$$T = \mathbb{C}/\Lambda, \Lambda = \mathbb{Z} \oplus \mathbb{Z} \cdot \tau,$$

satisfies

$$\mathbb{Z} \subsetneq \text{End}(T)$$

2. The values $j(\tau) \in \mathbb{C}$ at *CM-points* $\tau \in \mathbb{H}$ are named *singular moduli* or *singular values* of the modular invariant j .

Theorem 6.6 states a remarkable relation between tori and number fields. It reveals the arithmetic structure encoded in a complex torus with complex multiplication.

Theorem 6.6 (Complex multiplication and imaginary quadratic fields).

1. A point τ is a CM-point $\iff \mathbb{Q}(\tau)$ is an imaginary quadratic number field.

2. For CM-point $\tau \in \mathbb{H}$ the corresponding torus

$$T := \mathbb{C}/\Lambda_\tau, \quad \Lambda_\tau := \mathbb{Z} + \mathbb{Z} \cdot \tau,$$

has a non-trivial ring of endomorphisms

$$\mathbb{Z} \subsetneq \text{End}(T) \subset \mathbb{C}$$

which is an order of the imaginary quadratic number field $\mathbb{Q}(\tau)$.

According to Theorem 6.6 a singular value $j(\tau) \in \mathbb{C}$ defines a quadratic imaginary number field $K := \mathbb{Q}(\tau)$. The ring of endomorphism of the torus T with lattice $\Lambda(1, \tau)$ is an order in the ring of integers $\mathcal{O}_K \subset K$. Note that any order of an imaginary quadratic number field is a lattice of rank = 2.

Proof. 1. i) Consider a CM-point $\tau \in \mathbb{H}$. By definition exists $\alpha \in \mathbb{C} \setminus \mathbb{Z}$ with

$$\alpha \cdot \Lambda_\tau \subset \Lambda_\tau.$$

We obtain a matrix

$$\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M(2 \times 2, \mathbb{Z})$$

with

$$\alpha \cdot \begin{pmatrix} 1 \\ \tau \end{pmatrix} = \gamma \cdot \begin{pmatrix} 1 \\ \tau \end{pmatrix}.$$

Then

$$\alpha = a + b \cdot \tau, \quad \alpha \cdot \tau = c + d \cdot \tau,$$

which implies $b \neq 0$, because $\alpha \notin \mathbb{Z}$, and

$$b \cdot \tau^2 + (a - d) \cdot \tau - c = 0.$$

Because $\tau \in \mathbb{H}$ is not a real number, the number field $\mathbb{Q}(\tau)$ is imaginary quadratic.

ii) Assume that $\tau \in \mathbb{H}$ belongs to an imaginary quadratic number field. Then for suitable integers $a, b, c \in \mathbb{Z}$ with $a \neq 0$ because $\tau \in \mathbb{H}$

$$a \cdot \tau^2 + b \cdot \tau + c = 0.$$

From

$$(a\tau) \cdot \begin{pmatrix} 1 \\ \tau \end{pmatrix} = \begin{pmatrix} 0 & a \\ -c & -b \end{pmatrix} \cdot \begin{pmatrix} 1 \\ \tau \end{pmatrix}$$

follows $a\tau \in \text{End}(T)$, $\alpha\tau \notin \mathbb{Z}$, and T has complex multiplication,

2. Assume that T has complex multiplication and consider an endomorphism $f \in \text{End}(T)$ operating as multiplication μ_α with a complex number α . Alike to the proof of part 1, i) we obtain a matrix

$$\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M(2 \times 2, \mathbb{Z})$$

with

$$\alpha \cdot \begin{pmatrix} 1 \\ \tau \end{pmatrix} = \gamma \cdot \begin{pmatrix} 1 \\ \tau \end{pmatrix}$$

Therefore α is an eigenvalue of γ and satisfies the eigenvalue equation, which is an integral equation:

$$\alpha^2 - (d + a) \cdot \alpha + ab - cd = 0.$$

As a consequence $\alpha \in \mathcal{O}_K$, which implies

$$\mathbb{Z} \subsetneq \text{End}(T) \subset \mathcal{O}_K$$

Because $\mathcal{O}_K$ is free of rank = 2 the same holds true for $\text{End}(T)$, and the subring $\text{End}(T)$ is a sublattice of $\mathcal{O}_K$, i.e. an order

□

We consider in detail the two elliptic points of the modular group Γ , see Theorem 2.18:

Example 6.7 (CM-points).

1. *Elliptic point i :* The point $i \in \mathbb{H}$ determines the normalized lattice of *Gaussian integers*

$$\Lambda_i = \mathbb{Z} + \mathbb{Z} \cdot i$$

The torus $T = \mathbb{C}/\Lambda_i$ has the endomorphism ring

$$\text{End}(T) = \Lambda_i$$

The endomorphism ring

$$\text{End}(T) = \Lambda_i$$

equals the ring

$$\mathcal{O}_K \subset K := \mathbb{Q}(i) = \mathbb{Q}(\sqrt{-1})$$

of algebraic integers from K . Recall from Corollary 3.21

$$j(i) = 1728 = 2^6 \cdot 3^3 \in \mathbb{Z}.$$

2. *Elliptic point $\rho = e^{2\pi i/3}$:* The point $\rho \in \mathbb{H}$ determines the normalized lattice

$$\Lambda_\rho = \mathbb{Z} + \mathbb{Z} \cdot \rho$$

The torus $T = \mathbb{C}/\Lambda_\rho$ has the endomorphism ring

$$\text{End}(T) = \Lambda_\rho$$

The result is similar to the first example: The endomorphism ring

$$\text{End}(T) = \Lambda_\rho$$

equals the ring

$$\mathcal{O}_K \subset K := \mathbb{Q}(\rho) = \mathbb{Q}(\sqrt{-3})$$

of algebraic integers from K . Recall from Corollary 3.21

$$j(\rho) = 0 \in \mathbb{Z}.$$

Also the element $2\rho \in \mathbb{H}$ belongs to K . The torus

$$T = \mathbb{C}/\Lambda_{2\rho}$$

with the normalized lattice

$$\Lambda_{2\rho} = \mathbb{Z} + \mathbb{Z} \cdot 2\rho$$

has as endomorphisms the multiplication with arbitrary elements

$$\alpha \in \Lambda_{2\rho}.$$

For the proof one uses the equation

$$\rho^2 + \rho + 1 = 0.$$

The endomorphism ring is

$$\text{End}(T) = \mathcal{O}$$

with

$$\mathcal{O} = \Lambda_{2\rho} \subsetneq \mathcal{O}_K$$

is a proper order of K . Note $j(\rho) = 0 \in \mathbb{Z}$. A PARI-calculation shows

$$j(2\rho) = 54000 \in \mathbb{Z}.$$

6.2 Modular polynomials

The present section takes up the study of imaginary quadratic fields K from the viewpoint of modular forms. The main result is Theorem 6.18 about the singular values of the modular invariant j . The theorem shows that each singular value $j(\tau) \in \mathbb{C}$ is an algebraic number. For the proof one has to find a number $m \in \mathbb{N}^*$ and a polynomial

$$\Phi_m \in \mathbb{Z}[X] \text{ with } \Phi_m(j(\tau)) = 0$$

The proof relies on Theorem 6.6 which shows that the field $K = \mathbb{Q}(\tau)$ is an imaginary quadratic number field. To construct the polynomial Φ_m Definition 6.11 introduces the *modular polynomials*. The modular polynomials are a family of polynomials $(F_m(X))_{m \in \mathbb{N}^*}$ which extract properties of the j -invariant $j \in \mathcal{M}(\mathbb{H})$ from its transformation behaviour with respect to the different operations of the modular group Γ_1 . The structure of this section follows the essay of Borel in [9, Chap. III].

First we complete our investigation of the integer matrices Γ_m from Section 5.1. Different from Proposition 5.7 which considers the Γ_1 -left action on the matrices from Γ_m we now consider a subset $\Gamma_{m, \text{prim}} \subset \Gamma_m$ and a complete set of representations $V(m, \text{prim})$ of the Γ_1 -action on $\Gamma_{m, \text{prim}}$. See Definition 6.8, Proposition 6.9 and Corollary 6.10.

Definition 6.8 (Integer matrices with coprime entries). We define the set

$$\Gamma_{m, \text{prim}} := \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_m : \gcd(a, b, c, d) = 1 \right\} \subset \Gamma_m$$

of matrices with coprime entries.

Proposition 6.9 (Action on $\Gamma_{m, \text{prim}}$).

1. For each $m \in \mathbb{N}^*$ the Γ_1 -left action on Γ_m restricts to a Γ_1 -left action on $\Gamma_{m, \text{prim}}$

$$\Gamma_1 \times \Gamma_{m, \text{prim}} \rightarrow \Gamma_{m, \text{prim}}$$

A complete set of representatives of the Γ_1 -left action on $\Gamma_{m, \text{prim}}$ is the set of matrices

$$V(m, \text{prim}) := \left\{ \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} \in V(m) : \gcd(a, b, d) = 1 \right\} \subset V(m)$$

2. The cardinality of $V(m, \text{prim})$ is the value of the Dedekind ψ -function

$$\text{card } V(m, \text{prim}) = \psi(m) := m \cdot \prod_{p|m} \left(1 + \frac{1}{p} \right)$$

the product taken over all prime divisors of m .

Proof. 1. For the proof consider any

$$A = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \in \Gamma \text{ and } M = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_{m, \text{prim}}.$$

The ideal generated by the coefficients of the product matrix $A \cdot M$

$$I := (\alpha a + \beta c, \alpha b + \beta d, \gamma a + \delta c, \gamma b + \delta d) \subset \mathbb{Z}$$

contains the element

$$\delta \cdot (\alpha a + \beta c) + (-\beta) \cdot (\gamma a + \delta c) = (\alpha \delta - \beta \gamma) a + (\delta \beta - \beta \delta) c = a$$

and similar for b, c, d with the other generators of I . Therefore

$$(1) = (a, b, c, d) \subset I$$

which implies

$$(1) = I$$

Hence the coefficients of the product matrix $A \cdot M$ are coprime, i.e.

$$\Phi_m(A, M) = A \cdot M \in \Gamma_{m, \text{prim}}$$

2. For the proof see [40, Chap. 5.1].

□

Corollary 6.10 (Complete representatives of the orbit set). *For each $\gamma \in \Gamma$:*

1. *The set*

$$V(m) \cdot \gamma \subset \Gamma_m$$

is a complete set of representatives of the orbit set $\Gamma \backslash \Gamma_m$, i.e. with respect to the canonical projection

$$\Gamma_m \rightarrow \Gamma \backslash \Gamma_m, M \mapsto [M],$$

holds the equality of sets

$$\Gamma \backslash \Gamma_m = \{[M] : M \in V(m)\} = \{[M \cdot \gamma] : M \in V(m)\}.$$

2. *The set*

$$V(m, \text{prim}) \cdot \gamma \subset \Gamma_{m, \text{prim}}$$

is a complete set of representatives of the orbit set $\Gamma \backslash \Gamma_{m, \text{prim}}$.

Proof. 1. Consider a given $\gamma \in \Gamma$. We show: For arbitrary but fixed $M \in V(m)$ exist an element

$$M_1 \in V(m)$$

and a unique matrix $A \in \Gamma$ such that

$$A \cdot M = M_1 \cdot \gamma.$$

According to Proposition 5.7, applied to $M \cdot \gamma^{-1} \in \Gamma_m$, a unique element $A \in \Gamma$ exists with

$$M_1 := A \cdot (M \cdot \gamma^{-1}) \in V(m).$$

Then

$$M_1 \cdot \gamma = A \cdot M.$$

Hence $M_1 \cdot \gamma$ lies on the same orbit as M . As a consequence, the map

$$V(m) \rightarrow V(m), [M] \mapsto [M \cdot \gamma],$$

is well-defined and injective.

2. Analogously.

□

There is the following analogy to the family of Hecke operators $(T_m)_{m \in \mathbb{N}^*}$: The latter extract the Fourier coefficients of a given modular form $f \in M_k(\Gamma_1)$ from the transformation behaviour of f with respect to the operation of the modular group Γ_1

$$\Gamma_1 \times \Gamma_m \rightarrow \Gamma_m$$

on the set of matrices Γ_m . The Hecke operators are the family of linear maps with

$$T_m(f) := \sum_{M \in V(m)} f[M]_k$$

defined as the weighted sum of over a complete system $V(m)$ of representatives.

Similarly the modular polynomials are the family

$$F_m(X) = \prod_{M \in V(m, \text{prim})} (X - j \circ M) \in \mathcal{M}(\mathbb{H})[X]$$

defined as the product over the complete set $V(m, \text{prim})$ of representatives from Proposition 5.8. Each modular polynomial $F_m(X)$ is a normed polynomial of degree $\psi(m)$ in X . Its coefficients are polynomials in the modular invariant j .

Definition 6.11 (Modular polynomials). Consider $m \in \mathbb{N}^*$ and the complete set $V(m)$ of representatives of the orbit space $\Gamma_1 \backslash \Gamma_m$. The m -th modular polynomial is the polynomial

$$F_m(X) := \prod_{M \in V(m, \text{prim})} (X - j_M) \in \mathcal{M}(\mathbb{H})[X].$$

Note. The modular invariant j is constant along the orbits of the action of the modular group Γ_1 . Therefore the value of j_M in Definition 6.11 does not depend on the choice of a complete system of representatives of $\Gamma \backslash \Gamma_{m,prim}$

The field $\mathbb{Q}(j)$ is the function field of rational functions on the modular curve

$$X(\Gamma_1) \simeq \mathbb{P}^1$$

Each modular polynomial $F_m(X) \in K[X]$ is a normed polynomial of degree $\sigma_1(m)$ with coefficients from the function field $\mathbb{Q}(j)$. From this point of view the factorized form of Definition 6.11 defines the modular polynomial $F_m(X)$ as the product of its zeros, which are the functions j_M , $M \in V(m, prim)$.

Example 6.12 (Modular polynomials). The PARI-script `modular_polynomial_01` computes for several m the modular polynomials F_m , see Figure 6.1.

```

=====
modular_polynomial_01: Start

Modular polynomial F_2 : x^3 + (-j^2 + 1488*j - 162000)*x^2 + (1488*j^2 + 40773375*j + 8748000
000)*x + (j^3 - 162000*j^2 + 8748000000*j - 157464000000000)
-----
Modular polynomial F_3 : x^4 + (-j^3 + 2232*j^2 - 1069956*j + 36864000)*x^3 + (2232*j^3 + 2587
918086*j^2 + 8900222976000*j + 452984832000000)*x^2 + (-1069956*j^3 + 8900222976000*j^2 - 7708
45966336000000*j + 185542587187200000000)*x + (j^4 + 36864000*j^3 + 452984832000000*j^2 + 185
5425871872000000000*j)
-----
modular_polynomial_01: End
=====

```

Fig. 6.1 Modular polynomials

1. $m = 2$:

$$\psi(2) = 3$$

and

$$\begin{aligned}
 F_2(X, j) = & X^3 + \\
 & + (-j^2 + 1488*j - 162000)*X^2 + \\
 & + (1488*j^2 + 40773375*j + 8748000000)*X + \\
 & + (j^3 - 162000*j^2 + 8748000000*j - 157464000000000)
 \end{aligned}$$

2. $m = 3$:

$$\psi(3) = 4$$

and

$$F_3(X, j) = X^4 +$$

$$\begin{aligned}
& +(-j^3 + 2232*j^2 - 1069956*j + 36864000)*X^3 + \\
& +(2232*j^3 + 2587918086*j^2 + 8900222976000*j + 452984832000000)*X^2 + \\
& +(-1069956*j^3 + 8900222976000*j^2 - 770845966336000000*j + 1855425871872000000000)*X + \\
& +(j^4 + 36864000*j^3 + 452984832000000*j^2 + 1855425871872000000000*j)
\end{aligned}$$

The roots of the modular polynomial $F_m(X) \in \mathcal{M}(\mathbb{H})[X]$ are class invariants $j_M \in V(m, \text{prim})$ of the operation

$$\Gamma_1 \times \Gamma_{m, \text{prim}} \rightarrow \Gamma_{m, \text{prim}}$$

i.e. they are constant along the orbits of the action. We now take a closer look on the coefficients of the modular polynomials. We denote by

$$s_\mu, \quad j = 1, \dots, \psi(m),$$

the elementary symmetric functions in $\psi(m)$ variables. Then up to sign the functions

$$s_\mu(j_1, \dots, j_{\psi(m)}) \in \mathcal{M}(\mathbb{H}), \quad j = 1, \dots, \psi(m),$$

are the coefficients of the modular polynomial $F_m(X)$. E.g., the last function $s_{\psi(m)}$ provides the constant term

$$(-1)^{\psi(m)} s_{\psi(m)}(j_1, \dots, j_{\psi(m)}) = (-1)^{\psi(m)} \cdot j_1 \cdot \dots \cdot j_{\psi(m)},$$

and the first function s_1 provides the coefficient of the second-highest term $X^{\psi(m)-1}$

$$(-1) \cdot s_1(j_1, \dots, j_{\psi(m)}) = -(j_1 + \dots + j_{\psi(m)})$$

Theorem 6.13 proves: The coefficients of the modular polynomials are polynomials in the modular invariant j with integer coefficients, i.e. the coefficients belong to the subring

$$\mathbb{Z}[j] \subset \mathcal{M}(\mathbb{H}).$$

See Example 6.12 to exemplify the last statement.

Theorem 6.13 (Coefficients of the modular polynomials). *For each $m \in \mathbb{N}^*$*

$$F_m(X) = \prod_{M \in V(m, \text{prim})} (X - j_M) \in \mathbb{Z}[j][X]$$

i.e. the coefficients of $F_m(X)$ are polynomials in the modular invariant j with integer coefficients.

Proof. For arbitrary but fixed index $\mu = 1, \dots, \psi(m)$ we consider the coefficient

$$c := c_\mu := s_\mu(j_1, \dots, j_{\psi(m)}) \in \mathcal{M}(\mathbb{H})$$

of the modular polynomial

$$F_m(X) \in \mathcal{M}(\mathbb{H})[X].$$

The coefficient $c \in \mathcal{M}(\mathbb{H})$ is a meromorphic function. Step by step the proof will show that this function is a polynomial in the modular invariant j with integer coefficients, i.e. $c \in \mathbb{Z}[j]$.

i) *Weakly modularity of the function c* : According to Corollary 6.10 right multiplication with an element $\gamma \in \Gamma_1$ permutes the orbits of the Γ_1 -left operation Φ_m

$$\{[M_1], \dots, [M_{\psi(m)}]\} = \{[M_1 \cdot \gamma], \dots, [M_{\psi(m)} \cdot \gamma]\}.$$

Hence right multiplication does not change the value of the elementary symmetric functions: For all $j = 1, \dots, \psi(m)$

$$s_\mu(j_1, \dots, j_{\psi(m)}) = s_\mu(j_1 \circ \gamma, \dots, j_{\psi(m)} \circ \gamma).$$

Therefore c is weakly modular of weight $k = 0$.

ii) *Holomorphic on $\mathbb{H}$* : The modular invariant j is holomorphic on $\mathbb{H}$ according to Corollary 3.16. Therefore also the function c is holomorphic.

iii) *Meromorphic at ∞* : The coefficient c is weakly modular according to part i). Hence c has a Fourier expansion

$$c(\tau) = \sum_{v \in \mathbb{Z}} a_v \cdot q^v, \quad q = e^{2\pi i \tau}.$$

We have to show that the singularity $q = 0$ is a pole of c . For the proof we estimate the growth of c : According to Corollary 3.21 the modular invariant has the Fourier expansion

$$j(\tau) = \frac{1}{q} + P(q)$$

with P a convergent power series with integer coefficients. For an upper triangular matrix

$$M = \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} \in V(m, \text{prim})$$

the Fourier series of the class invariant $j \circ M$ derives from the Fourier series of j by substituting the variable $q = e^{2\pi i \tau}$ by

$$e^{2\pi i M(\tau)} = e^{2\pi i(b/d)} \cdot q^{(a/d)}$$

The equality

$$a \cdot d = m$$

implies

$$e^{2\pi i \alpha(\tau)} = \zeta_m^{ab} \cdot (q^{1/m})^{a^2} \text{ with } \zeta_m := e^{2\pi i/m}$$

a primitive m -th root of unity. In particular

$$|(j \circ \alpha)(\tau)| \leq \frac{1}{|q|^{a/d}} + O(1)$$

for a single class invariant. Hence for suitable constants k, K

$$|c(\tau)| \leq K \cdot \frac{1}{|q|^k}$$

for small $q \in \Delta \setminus \{0\}$. The estimate proves that $j \circ \alpha$ has a pole at $\tau = \infty$.

iv) *Integrality* $c \in \mathcal{O}_K[j]$: Due to the part already proved, Corollary 3.22 implies that the function c is a polynomial with complex coefficients in the modular invariant, i.e.

$$c \in \mathbb{C}[j].$$

We show that the coefficients are even algebraic integers from the cyclotomic field of m -th roots

$$K := \mathbb{Q}(\zeta_m) \subset \mathbb{C}$$

i.e.

$$c \in \mathcal{O}_K[j]$$

For the proof we recall from part iii) the substitution of q by

$$\zeta_m^{ab} \cdot (q^{1/m})^{a^2}.$$

As a consequence the class invariant $j \circ \alpha$ has a Fourier expansion with respect to $q^{1/m}$, namely

$$j(\alpha(\tau)) = \frac{1}{(q^{1/m})^{a^2} \cdot \zeta_m^{ab}} + P\left((q^{1/m})^{a^2} \cdot \zeta_m^{ab}\right)$$

with P a convergent power series in the variable $q^{1/m}$. Therefore the coefficients of $c \in \mathbb{C}[j]$ are algebraic integers in $\mathcal{O}_K$, i.e.

$$c \in \mathcal{O}_K[j]$$

v) *Integrality* $c \in \mathbb{Z}[j]$: We sharpen the result from part iv) and show

$$c \in \mathbb{Z}[j]$$

i.e. the algebraic integers from part iv) are even integers. The operation of the Galois group of the m -th cyclotomic field

$$\text{Gal}(K/\mathbb{Q}) \simeq (\mathbb{Z}/m\mathbb{Z})^*$$

extends to an operation on the set $\{j_1, \dots, j_{\psi(m)}\}$ of class invariants

$$\text{Gal}(K/\mathbb{Q}) \times \{j_1, \dots, j_{\psi(m)}\} \rightarrow \{j_1, \dots, j_{\psi(m)}\}, (\chi, j_\mu) \mapsto j_\mu^\chi,$$

which is defined as follows: Assume the automorphism $\chi \in \text{Gal}(K/\mathbb{Q})$ maps

$$\zeta_m \mapsto \chi(\zeta_m) := \zeta_m^e \text{ with } (e, m) = 1,$$

and let j_μ be the class invariant of the matrix

$$M = \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} \in V(m, \text{prim})$$

Then we define j_μ^χ as the class invariant of the matrix

$$M^\chi = \begin{pmatrix} a & b^\chi \\ 0 & d \end{pmatrix}$$

with

$$0 \leq b^\chi < d \text{ and } b^\chi \equiv e \cdot b \pmod{d}.$$

Because the polynomial

$$c := s_\mu(j_1, \dots, j_{\psi(m)}) \in \mathcal{O}_K[j]$$

is symmetric in the class invariants $j_1, \dots, j_{\psi(m)}$ its coefficients from $\mathcal{O}_K$ are invariant under the operation of $\text{Gal}(K/\mathbb{Q})$. Hence they belong to the fixed field $\mathbb{Q}$ of $\text{Gal}(K/\mathbb{Q})$ and even to $\mathcal{O}_K \cap \mathbb{Q} = \mathbb{Z}$, i.e.

$$c \in \mathbb{Z}[j]$$

□

We now evaluate the modular polynomials

$$F_m(X) \in \mathbb{Z}[j][X]$$

on the diagonal $\{X = j\}$, i.e. we consider the polynomials of the single variable j

$$\Phi_m(j) := F_m(j, j) = \prod_{M \in V(m, \text{prim})} (j - j_M) \in \mathbb{Z}[j].$$

Example 6.14 (The polynomials $\Phi_m(j) \in \mathbb{Z}[j]$). Taking up Example 6.12 Figure 6.2 shows some polynomials Φ_m . The PARI script `Modular_polynomial_02` computes their factorization:

1. $m = 2$:

$$\Phi_2(j) = F_2(j, j) = -j^4 + 2978*j^3 + 40449375*j^2 + 17496000000*j - 15746400000000$$

splits into the factors

$$j - 8000, j - 1728, (j + 3375)^2$$

2. $m = 3$:

$$\begin{aligned} \Phi_3(j) = F_3(j, j) = & -j^6 + 4464*j^5 + 2585778176*j^4 + 17800519680000*j^3 \\ & - 769939996672000000*j^2 + 371085174374400000000*j \end{aligned}$$

splits into the factors

$$j - 54000, (j - 8000)^2, j, (j + 32768)^2$$

3. The example $\Phi_5(j)$ shows: In general $\Phi_m(j)$, m prime, does not split into linear factors over $\mathbb{Z}$.

```
modular_polynomial_02: Start

Polynomial Phi_2 : -j^4 + 2978*j^3 + 40449375*j^2 + 17496000000*j - 15746400000000
factor 1 = [j - 8000, 1]
factor 2 = [j - 1728, 1]
factor 3 = [j + 3375, 2]
-----
Polynomial Phi_3 : -j^6 + 4464*j^5 + 2585778176*j^4 + 17800519680000*j^3 - 769939996672000000*j^2 + 371085174374400000000*j
factor 1 = [j - 54000, 1]
factor 2 = [j - 8000, 2]
factor 3 = [j, 1]
factor 4 = [j + 32768, 2]
-----
Polynomial Phi_5 : -j^10 + 7440*j^9 + 1665990262720*j^8 + 215757860427776000*j^7 - 440440798293848579637248*j^6
+ 53797234800359280738891202560*j^5 + 4726025910884027749483397649530880*j^4 + 7366996272355613764702158779590
9017600*j^3 - 250688456991364600842741491417948646014976*j^2 + 106548661606848850900840320546713018302464000*j
+ 141359947154721358697753474691071362751004672000
factor 1 = [j - 287496, 2]
factor 2 = [j - 1728, 2]
factor 3 = [j + 32768, 2]
factor 4 = [j + 884736, 2]
factor 5 = [j^2 - 1264000*j - 681472000, 1]
-----
modular_polynomial_02: End
```

Fig. 6.2 Some polynomials Φ_m and their factorization

Lemma 6.15 (The leading coefficient of the modular polynomials). *If $m \geq 2$ is not a square, then the polynomial*

$$\Phi_m(j) \in \mathbb{Z}[j]$$

has leading coefficient ± 1 .

Example 6.12 illustrates Lemma 6.15 by

$$\Phi_2(j) = -j^4 + O(3), \Phi_3(j) = -j^6 + O(5) \text{ and } \Phi_5(j) = -j^{10} + O(9).$$

Proof. The function

$$\Phi_m(j) : \mathbb{H} \rightarrow \mathbb{C}$$

is an automorphic function. Consider its Fourier expansion

$$\Phi_m(j)(\tau) = \frac{c_k}{q^k} + \frac{c_{k-1}}{q^{k-1}} + \text{higher terms}$$

Because j has a pole at $\tau = \infty$ of order $= 1$, the coefficient c_k of the Fourier series is also the leading coefficient of Φ_m as polynomial in j . The coefficient is the product of the highest negative coefficients of the factors

$$j - j_\mu, \mu = 1, \dots, \psi(m)$$

with $j_\mu = j \circ M$ for a suitable representative $M \in V(m, \text{prim})$. For arbitrary but fixed index μ the representation

$$j(\tau) = \frac{1}{q} + P(q)$$

with a convergent power series P implies

$$j_\mu(\tau) = (j \circ M)(\tau) = j\left(\frac{a\tau + b}{d}\right) = \frac{e^{-2\pi i(b/d)}}{q^{a/d}} + P\left(q^{a/d} \cdot e^{2\pi i(b/d)}\right)$$

with

$$M = \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} \in V(m, \text{prim}).$$

The leading terms of $j - j_\mu$ are

$$\frac{1}{q} - \frac{e^{-2\pi i(b/d)}}{q^{a/d}}$$

The assumption that $m = a \cdot d$ is not a square implies $a \neq d$ equivalently $a/d \neq 1$. Hence the two summands have different order in q and cannot cancel. Exactly one of them is highest negative term in the Fourier expansion of $j - j_\mu$. Apparently its coefficient is a root of unity, and the same holds for the coefficient c_k of the product

$$\Phi = \prod_{M \in V(m, \text{prim})} (j - j_M).$$

Because $c_k \in \mathbb{Z}$ we have $c_k = \pm 1$

□

We recall a simple version of the elementary divisor theorem.

Proposition 6.16 (Elementary divisor theorem).

1. For each matrix

$$M \in \Gamma_m, m \in \mathbb{N}^*,$$

exist two matrices $\gamma_1, \gamma_2 \in \Gamma$ with

$$\gamma_1 \cdot B \cdot \gamma_2 = \begin{pmatrix} e_1 & 0 \\ 0 & e_2 \end{pmatrix} \text{ satisfying the divisor relation } e_1 | e_2.$$

2. For each pair of primitive matrices

$$M_1, M_2 \in \Gamma_{m, \text{prim}}, m \in \mathbb{N}^*,$$

exist two matrices $\gamma_1, \gamma_2 \in \Gamma$ with

$$M_2 = \gamma_1 \cdot M_1 \cdot \gamma_2.$$

Proof. The Smith normal-form theorem [44, Chap. V, Theor. 1.2] holds over a principal ideal domain like $\mathbb{Z}$. It provides for a matrix $M \in M(2 \times 2, \mathbb{Z})$ two matrices $A, B \in GL(2, \mathbb{Z})$ with

$$A \cdot M \cdot B = \begin{pmatrix} d_1 & 0 \\ 0 & d_2 \end{pmatrix} \in M(2 \times 2, \mathbb{Z})$$

Up to permutation the numbers d_1, d_2 are uniquely determined and satisfy the divisibility relation

$$d_1 | d_2$$

Because $\det A = \det B = \pm 1$ we may assume w.l.o.g.

$$\det A = \det B = 1$$

For a primitive matrix $M \in V(m, \text{prim})$ the numbers d_1 and d_2 are coprime, which implies $d_1 = 1, d_2 = \det M = m$ and

$$A \cdot M \cdot B = \begin{pmatrix} 1 & 0 \\ 0 & m \end{pmatrix}$$

We apply the result to M_1 and M_2 and obtain corresponding matrices such that

$$A_1 \cdot M_1 \cdot B_1 = A_2 \cdot M_2 \cdot B_2$$

or

$$(A_2^{-1} A_1) \cdot M_1 = M_2 \cdot (B_2 B_1^{-1})$$

Set

$$\gamma_1 := (A_2^{-1}A_1) \text{ and } \gamma_2 := (B_2B_1^{-1})$$

□

According to Theorem 6.13 the modular polynomial

$$F_m(j, X) := F_m(X) = \prod_{\mu=1}^{\psi(m)} (X - j_\mu) \in \mathbb{Z}[j, X]$$

is a polynomial in the two variables j and X with integer coefficients. It expands as

$$F_m(j, X) = \sum_{n=0}^N \left(\sum_{r+s=n} c_{rs} \cdot j^r \cdot X^s \right), \quad c_{rs} \in \mathbb{Z}.$$

Proposition 6.17 will show that $F_m(j, X)$ is symmetric in both variables, i.e. the coefficients satisfy

$$c_{rs} = c_{sr}.$$

As a consequence, the highest exponent of j in $F_m(j, X)$ is $\psi(m)$.

Proposition 6.17 (Symmetry of the modular polynomials). *The modular polynomials*

$$F_m(j, X) \in \mathbb{Z}[j, X], \quad m \in \mathbb{N}^*,$$

are symmetric with respect to j and X , i.e.

$$F_m(j, X) = F_m(X, j).$$

Proof. We set $R := \mathbb{Z}[j]$ with quotient field

$$K := \mathcal{Q}(R) = \mathbb{Q}(j).$$

i) *Irreducibility of $F_m(X) \in R[X]$ over K :* By definition the roots of the modular polynomial

$$F_m(X) = \prod_{M \in V(m, \text{prim})} (X - j_M) \in R[X]$$

are the class invariants

$$j_M = j \circ M, \quad M \in V(m, \text{prim}).$$

They are pairwise distinct which can be shown by comparing their Fourier coefficients. To prove that $F_m(X)$ is an irreducible polynomial we show that the Galois group

$$\text{Gal}(L/K)$$

of the splitting field

$$L = K(j_1, \dots, j_{\psi(m)})$$

of the polynomial $F_m(X) \in R[X]$ acts transitively on the roots. Then no root belongs to the fixed field K .

Elements of the Galois group can be obtained as follows: Any matrix $\gamma \in \Gamma_1$ defines the automorphism

$$L = K(j_1, \dots, j_{\psi(m)}) \rightarrow L, f \mapsto f \circ \gamma.$$

Consider two class invariants

$$j_\mu, j_\nu, 1 \leq \mu, \nu \leq \psi(m)$$

corresponding to the matrices

$$M_\mu, M_\nu \in V(m, \text{prim})$$

One has to find $\gamma \in \Gamma_1$ such that

$$j_\nu = j_\mu \circ \gamma$$

i.e. one has to find $\gamma, \gamma' \in \Gamma_1$ such that

$$\gamma' \cdot M_\nu = M_\mu \cdot \gamma,$$

which is achieved by Proposition 6.16, part 2). Therefore $\text{Gal}(L/K)$ operates transitively on the roots of $F_m(X)$, and the latter polynomial is irreducible over K .

ii) *Common root of $F_m(j, X)$ and $F_m(X, j)$* : We consider the two polynomials

$$f(X) := F_m(j, X) \in R[X] \text{ and } g(X) := F_m(X, j) \in R[X]$$

For any matrix $M \in V(m, \text{prim})$ the class invariant $j_M = j \circ M$ is a root of the modular polynomial $f(X) \in R[X]$. In particular for

$$M := \begin{pmatrix} 1 & 0 \\ 0 & m \end{pmatrix} \in V_{m, \text{prim}}$$

holds for all $\tau \in \mathbb{H}$

$$0 = f(j_M(\tau)) = F_m(j(\tau), j_M(\tau)) = F_m\left(j(\tau), j\left(\frac{\tau}{m}\right)\right)$$

Replacing the argument τ by $M \cdot m$ shows for all $\tau \in \mathbb{H}$

$$0 = F_m(j(\tau \cdot m), j(\tau)) = F_m(j_{M'}(\tau), j(\tau)) = g(j_{M'}(\tau)) \text{ i.e. } g(j_{M'}) = 0$$

with

$$M' := \begin{pmatrix} m & 0 \\ 0 & 1 \end{pmatrix} \in V_{m, \text{prim}}$$

Also $f(j_{M'}) = 0$ because $M' \in V_{m, \text{prim}}$. Hence both polynomials $f(X), g(X) \in R[X]$ have the common root $j_{M'}$.

iii) *Symmetry*: Due to Lemma 6.15 and part i) the polynomial $f(X) \in K[X]$ is irreducible with leading coefficient ± 1 . Hence $f(X)$ is the minimal polynomial of the field extension $K \subset K[j_\beta]$, and $g(X)$ is a multiple of $f(X)$, i.e.

$$g(X) = h(X) \cdot f(X)$$

with a polynomial $h(X) \in K[X]$. According to Theorem 6.13 both polynomials $f(X)$ and $g(X)$ have their coefficients already in the ring $R = \mathbb{Z}[j]$. The ring R is factorial as a ring of polynomials over the factorial ring $\mathbb{Z}$. The Lemma of Gauss implies that also the polynomial $h[X]$ has all coefficients from R , and the equation

$$g(X) = h(X) \cdot f(X)$$

holds in $R[X]$. Set

$$H(j, X) := h(X) \in R[X] = \mathbb{Z}[j][X].$$

Then

$$F_m(X, j) = H(j, X) \cdot F_m(j, X) \in \mathbb{Z}[j, X].$$

Considered as an equation in the ring $\mathbb{Z}[j, X]$ of polynomials in two variables the last equation is invariant when interchanging the role of j and X . Hence also

$$F_m(j, X) = H(X, j) \cdot F_m(X, j) \in \mathbb{Z}[j, X].$$

As a consequence

$$F_m(j, X) = H(X, j) \cdot H(j, X) \cdot F_m(j, X),$$

$$H(X, j) \cdot H(j, X) = 1,$$

and

$$H(X, j) = H(j, X) = \pm 1.$$

In case $H(X, j) = -1$ the previous equation

$$F_m(j, X) = H(X, j) \cdot F_m(X, j)$$

would imply

$$F_m(X, j) = -F_m(j, X) \text{ and } F_m(j, j) = -F_m(j, j).$$

Then

$$F_m(j, j) = 0$$

which contradicts the fact that $\Phi_m(j) = F_m(j, j)$ has leading coefficient ± 1 according to Lemma 6.15. Therefore

$$H(X, j) = 1$$

and

$$F_m(j, X) = F_m(X, j)$$

□

Theorem 6.13, Lemma 6.15 and Proposition 6.17 correspond to Borel's theorem from [9, Chap. III, §4, Theor. 1].

We now derive from the theory of modular polynomials that all singular values $j(\tau) \in \mathbb{C}$ are algebraic integers.

Theorem 6.18 (Singular values of the modular j -invariant). *For each CM-point $\tau \in \mathbb{H}$ the j -value $j(\tau) \in \mathbb{C}$ is an algebraic integer. More precisely: There exists $m \in \mathbb{N}^*$ such that $j(\tau)$ is a root of the polynomial $\Phi_m(X) \in \mathbb{Z}[X]$.*

Proof. Consider a CM-point $\tau \in \mathbb{H}$. Theorem 6.6 implies that τ belongs to an imaginary quadratic field

$$K = \mathbb{Q}(\sqrt{d})$$

with a squarefree $d \in \mathbb{Z}$, $d < 0$. The subring

$$\mathcal{O}_K \subset K$$

of algebraic integers of K is a free $\mathbb{Z}$ -module of rank = 2

$$\mathcal{O}_K = \mathbb{Z} \oplus \mathbb{Z} \cdot z$$

with a $\mathbb{Z}$ -basis $(z, 1)$, $z \in \mathbb{H}$. For the admissible values of z see Remark 6.2.

i) *Matrices $\alpha \in \Gamma_{m, \text{prim}}$ fixing z :* For each matrix $\alpha \in \Gamma_{m, \text{prim}}$ holds

$$\alpha(\tau) = \tau \iff j(\alpha(\tau)) = j(\tau)$$

After having found a number $m \in \mathbb{N}^*$ and a corresponding matrix $\alpha \in \Gamma_{m, \text{prim}}$ we replace α by a matrix $M \in V(m, \text{prim})$ from the same orbit under the left action of Γ_1 . Then $j_M(\tau)$ is a root of the polynomial $\Phi_m(X) \in \mathbb{Z}[X]$.

We now determine a pair

$$(m, \alpha) \in \mathbb{N}^* \times \Gamma_{m, \text{prim}} \text{ with } \alpha(z) = z$$

The element

$$\xi := \begin{cases} \sqrt{d} & d < -1 \\ 1+i & d = -1 \end{cases}$$

is an algebraic integer. Its norm

$$N(\xi) = \xi \bar{\xi} \in \mathbb{Z}$$

is squarefree. Multiplication with $\xi \in \mathcal{O}_K$ defines an injective $\mathbb{Z}$ -linear endomorphism

$$\mathcal{O}_K \rightarrow \mathcal{O}_K, x \mapsto \xi \cdot x.$$

Denote by

$$\alpha = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M(2 \times 2, \mathbb{Z}) \cap GL(2, \mathbb{R})^+$$

its matrix with respect to the basis $(z, 1)$ of $\mathcal{O}_K$. Define

$$m := \det \alpha \in \mathbb{N}^*$$

Then

$$\xi \cdot \begin{pmatrix} z \\ 1 \end{pmatrix} = \alpha \cdot \begin{pmatrix} z \\ 1 \end{pmatrix}$$

with multiplication in the ring $\mathcal{O}_K$ on the left-hand side, and matrix multiplication on the right-hand side. In components

$$\begin{pmatrix} \xi \cdot z \\ \xi \cdot 1 \end{pmatrix} = \begin{pmatrix} az + b \\ cz + d \end{pmatrix}$$

- The first form shows: The matrix

$$\alpha \in M(2 \times 2, \mathbb{Z})$$

which represents the $\mathbb{Z}$ -linear multiplication by ξ has the eigenvalue $\xi \in \mathbb{C} \setminus \mathbb{R}$ and by conjugation $\alpha = \bar{\alpha}$ also the eigenvalue $\bar{\xi} \neq \xi$. Because d is squarefree, also $\xi \cdot \bar{\xi}$ is squarefree which implies

$$(a, b, c, d) = 1$$

and therefore

$$\alpha \in \Gamma_{m, \text{prim}}$$

due to Definition 6.8.

- The component form shows: Considered as a fractional linear transformation

$$\mathbb{H} \rightarrow \mathbb{H}, w \mapsto \alpha(w) := \frac{a \cdot w + b}{c \cdot w + d}$$

the matrix $\alpha \in GL(2, \mathbb{R})^+$ satisfies

$$\alpha(z) = \frac{a \cdot z + b}{c \cdot z + d} = \frac{\xi \cdot z}{\xi \cdot 1} = z$$

and fixes z .

ii) $j(z)$ is an algebraic integer: Proposition 6.9 implies the existence of a matrix

$$M \in V(m, \text{prim}), \mu \in \{1, \dots, \psi(m)\},$$

with the same Γ_1 -orbit as α with respect to the left action

$$\Phi_{m, \text{prim}} : \Gamma_1 \times \Gamma_{m, \text{prim}} \rightarrow \Gamma_{m, \text{prim}}, (A, M) \mapsto A \cdot M,$$

w.l.o.g. $\mu = 1$. Then

$$j \circ M = j \circ \alpha$$

because the value of j is constant along each orbit of the Γ -operation. Due to part i)

$$j(z) = j(\alpha(z)) = j(M(z)) = j_M(z)$$

which implies: The function

$$\Phi_m(j) = \prod_{M \in V(m, \text{prim})} (j - j_M) \in \mathbb{Z}[j]$$

vanishes at $z \in \mathbb{H}$:

$$\Phi_m(j)(z) = \Phi_m(j(z)) = 0.$$

The polynomial

$$\Phi_m(j) \in \mathbb{Z}[j]$$

has integer coefficients, and leading coefficient ± 1 due to Lemma 6.15. Hence

$$0 = \Phi_m(j(z))$$

proves that $j(z)$ is an algebraic integer. In particular, the subring

$$\mathbb{Z}[j(z)] \subset \mathbb{C}$$

is integral over the ring $\mathbb{Z}$.

iii) $j(\tau)$ is integral over $\mathbb{Z}[j(z)]$: The given element

$$\tau \in K = \mathbb{Q} + \mathbb{Q} \cdot z$$

has a representation

$$\tau = \beta(z) := \frac{az + b}{d}$$

with a primitive matrix

$$\beta = \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} \in \Gamma_{m', \text{prim}}, \quad m' := a \cdot d.$$

Alike to the argument from part ii) the Γ_1 -orbit of β passes through a matrix $M' \in V(m', \text{prim})$ which proves

$$j(\tau) = j(\beta(z)) = (j \circ M')(z) = j_{M'}(z).$$

Theorem 6.13 and Lemma 6.15 imply that $j_{M'}$ is integral over the ring $\mathbb{Z}[j]$. Hence

$$j(\tau) = j_{M'}(z)$$

is integral over $\mathbb{Z}[j(z)]$.

iv) $j(\tau)$ is an algebraic integer: When combining part iii) and part ii) the transitivity of integral dependence proves the integrality of $j(\tau)$ over $\mathbb{Z}$, i.e. τ is an algebraic integer.

□

Example 6.19 (Singular values of the modular j -function). Continuing Example 6.14 we consider a CM -class $[\tau] \in X(\Gamma_1)$ and choose the level

$$m = 2.$$

1. According to Example 6.14 the polynomial

$$\Phi_2(j) \in \mathbb{Z}[j]$$

has degree = 4 and factorizes with the three factors

$$j - 8000, \quad j - 1728, \quad (j + 3375)^2$$

2. Which property is expressed by these factors, how are the values $j \in \{8000, 1728, -3375\}$ related?

According to Theorem 6.18 we determine a complete set of representatives

$$\alpha_j \in \Gamma_{2, \text{prim}}, \quad j = 1, 2, 3,$$

with fixed points $\alpha_j(z_j) = z_j$, $z_j \in \mathbb{H}$:

•

$$\alpha_1 = \begin{pmatrix} 1 & -2 \\ 1 & 0 \end{pmatrix} \text{ has fixed point } z_1 = \frac{1 + i\sqrt{7}}{2}$$

which defines the lattice

$$\Lambda_1 := \left\langle 1, \frac{1 + \sqrt{-7}}{2} \right\rangle = \mathcal{O}_{K_1} \subset K_1 := \mathbb{Q}(\sqrt{-7})$$

of the torus $T_1 := \mathbb{C}/\Lambda_1$.

•

$$\alpha_2 = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \text{ has fixed point } z_2 = i$$

which defines the lattice

$$\Lambda_2 := \langle 1, i \rangle = \mathcal{O}_{K_2} \subset K_2 := \mathbb{Q}(\sqrt{-1})$$

of the torus $T_2 := \mathbb{C}/\Lambda_2$.

•

$$\alpha_3 = \begin{pmatrix} 0 & -1 \\ 2 & 0 \end{pmatrix} \text{ has fixed point } z_3 = i\sqrt{2}$$

which defines the lattice

$$\Lambda_3 := \langle 1, i\sqrt{2} \rangle = \mathcal{O}_{K_3} \subset K_3 := \mathbb{Q}(\sqrt{-2})$$

of the torus $T_3 := \mathbb{C}/\Lambda_3$.

3. Hence the three resulting fields K_j , $j = 1, 2, 3$ are imaginary quadratic. Due to Theorem 6.6 each of the three fields relates to a torus with complex multiplication. In addition Theorem 6.18 shows that the tori have their j -values from

$$\{8000, 1728, -3375\}$$

We already know $j(z_2) = j(i) = 1728$. A PARI-computation completes at once the remaining values

$$j(z_1) = -3375 \text{ and } j(z_3) = 8000.$$

One knows that all three fields have class number $h_K = 1$, which will later confirm Corollary 6.34.

6.3 The class number

A further application of the modular j -invariant to number theory is the investigation of imaginary quadratic number fields and their class number.

Of course the class number is a concept which refers to number fields and their rings of integers. Nevertheless, for those students who are familiar with complex analysis we start with an analogue from the theory of Riemann surfaces, cf. [69]. A *divisor* is a concept which captures the order of poles and zeros of meromorphic functions on a Riemann surface X . A divisor is locally defined as the equivalence class of meromorphic functions modulo holomorphic functions without zeros. This idea is made precise by the divisor sequence on X .

Remark 6.20 (Divisor sequence and Picard group). Consider a compact Riemann surface X .

1. The *divisor sequence* on X is the exact sequence of sheaves

$$1 \rightarrow \mathcal{O}^* \rightarrow \mathcal{M}^* \rightarrow \mathcal{D} \rightarrow 0$$

Here the sheaf of divisors $\mathcal{D}$ is defined as the quotient of the sheaf $\mathcal{M}^*$ of locally non-zero meromorphic functions modulo the subsheaf $\mathcal{O}^*$ of holomorphic functions without zeros. On a compact Riemann surface the induced long exact sequence of vector spaces is

$$1 \rightarrow \mathbb{C}^* \rightarrow \mathcal{M}^*(X) \rightarrow \text{Div}(X) \rightarrow \text{Cl}(X) \rightarrow 0$$

Because

$$H^0(X, \mathcal{O}^*) = \mathbb{C}^*, \text{Div}(X) := H^0(X, \mathcal{D}), \text{Cl}(X) := \frac{\text{Div}(X)}{\text{im}[\mathcal{M}^*(X) \rightarrow \text{Div}(X)]}$$

There are no further terms because

$$H^1(X, \mathcal{M}^*) = 0$$

which is a result from the theory of compact Riemann surfaces. It implies the isomorphism

$$\text{Cl}(X) \xrightarrow{\cong} H^1(X, \mathcal{O}^*) = \text{Pic}(X)$$

between the group of divisor classes and the Picard group of isomorphism classes of holomorphic line bundles on X .

2. *Picard group:* The kernel of the degree map

$$\deg : \text{Pic}(X) \rightarrow \mathbb{Z}$$

is the subgroup $\text{Pic}_0(X)$ of all line bundles, or divisors, of degree zero. The corresponding exact sequence of Abelian groups

$$0 \rightarrow \text{Pic}_0(X) \rightarrow \text{Pic} \rightarrow \text{NS}(X) \rightarrow 0$$

with *Nero-Severi group* as quotient

$$NS(X) := Pic(X)/Pic_0(X)$$

splits:

$$Pic(X) = Pic_0(X) \oplus NS(X)$$

Here the degree zero part of $Pic(X)$ is a finitely generated free Abelian group, its rank depends on the genus g of the compact Riemann surface X

$$Pic_0(X) = \mathbb{Z}^{2g},$$

while the discrete part of is free Abelian of rank 1

$$NS(X) = \mathbb{Z}$$

For number fields the analogue of divisors are fractional ideals, see [4, Chap. 9].

Definition 6.21 (Fractional ideal). Consider a number field K with ring of integers $\mathcal{O}_K \subset K$.

1. A *fractional ideal* of K is defined as an $\mathcal{O}_K$ -submodule $\mathfrak{a} \subset K$ for which there exists an element $d \in \mathcal{O}_K$, $d \neq 0$, with

$$d \cdot \mathfrak{a} \subset \mathcal{O}_K.$$

Hence a fractional ideal is a subset of the form

$$\mathfrak{a} = \frac{\tilde{a}}{d} \subset K$$

with an ideal $\tilde{a} \subset \mathcal{O}_K$. The element d is named a *common denominator* of the fractional ideal $\mathfrak{a}$, and the ideal $\tilde{a} \subset \mathcal{O}_K$ is named an *integral ideal* of K .

2. Fractional ideals generated by a single element $\alpha \in K$

$$\mathfrak{a} = (\alpha) := \mathcal{O}_K \cdot \alpha$$

are named *principal* fractional ideals.

3. For two fractional ideals $\mathfrak{a}, \mathfrak{b} \subset K$ the *product*

$$\mathfrak{a} \cdot \mathfrak{b}$$

is the fractional ideal generated by all products $a \cdot b$, $a \in \mathfrak{a}$, $b \in \mathfrak{b}$.

4. For a fractional ideal $\mathfrak{a} \neq \{0\}$ the set

$$\mathfrak{a}^{-1} := \{x \in K : x \cdot \mathfrak{a} \subset \mathcal{O}_K\}$$

is named the *inverse* of $\mathfrak{a}$.

Remark 6.22 (Fractional ideals). Consider a number field K with number ring $\mathcal{O}_K$.

1. Each non-zero ideal of $\mathcal{O}_K$ and each non-zero fractional ideal of K is a lattice of rank $= \deg K$. The fractional ideals of K are the finitely generated $\mathcal{O}_K$ -submodules of K .
2. Because $\mathcal{O}_K$ is a Dedekind ring, for a fractional ideal

$$\mathfrak{a} \neq \{0\} \subset K$$

the inverse $\mathfrak{a}^{-1}$ is also a fractional ideal, see [4, Theor. 9.8]. It satisfies

$$\mathfrak{a} \cdot \mathfrak{a}^{-1} = (1)$$

3. With respect to multiplication the set of fractional ideals is an Abelian group $J(K)$. Its quotient by the subgroup of principal fractional ideals is named the *ideal class group* or simply *class group* of K

$$Cl(K) := J(K)/im[K^* \rightarrow J(K)]$$

The canonical map is

$$K^* \rightarrow J(K), x \mapsto (x).$$

Hence two fractional ideals

$$\mathfrak{a}_1, \mathfrak{a}_2 \in J(K)$$

define the same class in $Cl(K)$ iff

$$\mathfrak{a}_2 = a \cdot \mathfrak{a}_1 \text{ for a suitable } a \in K^*.$$

As a consequence there is an exact sequence of multiplicative Abelian groups

$$1 \rightarrow \mathcal{O}_K^* \rightarrow K^* \rightarrow J(K) \rightarrow Cl(K) \rightarrow 1$$

which is completely analogue to the long exact divisor sequence on a compact Riemann surface from Definition 6.20.

4. Comparing the exact sequence of a number field with the exact divisor sequence of a Riemann surface from Remark 6.20 one easily observes the analogy between corresponding objects:

Compact Riemann surface X	Number field K
$\mathcal{O}_X^*$	$\mathcal{O}_K^*$
$\mathcal{M}_X^*$	K^*
$Div(X)$	$J(K)$
$Cl(X) = Pic(X)$	$Cl(K)$

Definition 6.23 (Class number of a number field). The *class number* of a number field K is the cardinality

$$h_K := \text{card } Cl(K)$$

of the set of classes of fractional ideals from K .

A number ring $\mathcal{O}_K$ is a principal ideal domain iff

$$h_K = 1$$

The class group $Cl(K)$ of an algebraic number field is always a finite Abelian group. We now set up to proof for imaginary quadratic number fields K the finiteness of the class number h_K . Herefore we characterize each ideal class of K by a distinguished fractional ideal. To motivate Definition 6.24 and the proofs of Proposition 6.25 and Proposition 6.26 see [70, §8, Satz 1]. The background is the theory of quadratic forms. For an introduction with focus on the historical development see [51].

Definition 6.24 (Reduced fractional ideal). Consider an algebraic number field K and the class of a fractional ideal $\mathfrak{a} \subset K$. The ideal is *reduced* if it has the form

$$\mathfrak{a} = \mathbb{Z}a \oplus \mathbb{Z}\omega_1, \text{ with } a > 0 \text{ and } \omega_1 = s \cdot t \cdot w \text{ with}$$

- $s \in \mathbb{Z}$
- $\tau := \frac{\omega_1}{a}$ satisfying $-\frac{1}{2} \cdot \text{Re } \tau \leq \frac{1}{2}$ and $|\tau| \geq 1$,
- and $\text{Re } \tau \geq 0$ for $|\tau| = 1$.

Proposition 6.25. *Each ideal class in a quadratic imaginary number field contains exactly one reduced ideal,*

Proof. Denote by $D < 0$ the discriminant of K and consider the reduced ideal

$$\mathfrak{a} = \mathbb{Z}a + \mathbb{Z} \cdot \frac{b + \sqrt{D}}{2}$$

It satisfies

$$|b| \leq a \text{ and } a|N(\omega_1)| = \frac{b^2 - D}{4}$$

Hence for suitable $c \in \mathbb{Z}$ holds

$$b^2 - D = 4 \cdot ac \text{ or } D = b^2 - 4ac < 0$$

While

$$\left| \frac{\omega_1}{a} \right| \geq 1 \implies N(\omega_1) \geq a^2 \implies b^2 - \Delta \geq 4a^2 \implies -D \geq 3a^2$$

Proposition 6.26. *Proposition[Finiteness of the class number] The class number h_K of an imaginary quadratic number field K is finite.*

Proof. Denote by $D < 0$ the discriminant of K and consider a reduced fractional ideal of K

$$\mathfrak{a} = \mathbb{Z} \cdot a \oplus \mathbb{Z} \cdot \frac{b + \sqrt{D}}{2}$$

Then

$$|b| \leq a \text{ and } a|N(\omega_1)| \frac{b^2 - D}{4} < 0$$

implies the existence of $c \in \mathbb{Z}$ satisfying $b^2 - D = 4ac$, $D = b^2 - 4ac < 0$.

$$\left| \frac{\omega_1}{a} \right| \geq 1 \implies N(\omega_1) \geq a^2 \implies b^2 - D \geq 4a^2 \implies -D \geq 3a^2 \implies 0 < a \leq \sqrt{\frac{|D|}{3}}$$

as an upper bound for a . As a consequence the class number is finite.

□

For the finiteness of the class number see also [42, Chap. 5, Cor. 2] and [34, Chap. 12, §2].

Example 6.27 (Class number).

1. There are exactly 9 imaginary quadratic fields

$$K = \mathbb{Q}(\sqrt{-D}), \quad D > 0 \text{ squarefree integer,}$$

with class number $h_K = 1$. They belong to the values

$$D \in \{1, 2, 3, 7, 11, 19, 43, 67, 163\}.$$

Recall that for $D = 1$ the field $K := \mathbb{Q}(i)$, the ring of integers $\mathcal{O}_K \subset K$, is made up by the Gauss' integers

$$a + i \cdot b \text{ with } a, b \in \mathbb{Z}$$

2. The imaginary quadratic field

$$K = \mathbb{Q}(\sqrt{-5})$$

has class number $h_K = 2$. The class of the ideal

$$(2, 1 + \sqrt{-5}) \subset \mathcal{O}_K$$

is not principal. It generates $Cl(K)$, see [42, Chap. 5].

Remark 6.28 (Dedekind ζ -function of a number field). Let K be an algebraic number field.

1. The ring of integers $\mathcal{O}_K \subset K$ is a Dedekind ring. Hence each ideal of $\mathcal{O}_K$ uniquely factorizes into a finite number of prime ideals from $\mathcal{O}_K$. As a consequence the *Dedekind ζ -function of K*

$$\zeta_K(s) := \sum_{\mathfrak{a}} \frac{1}{N(\mathfrak{a})^s} \text{ with } \operatorname{Re} s > 1,$$

equals the Euler product

$$\sum_{\mathfrak{a}} \frac{1}{N(\mathfrak{a})^s} = \prod_{\mathfrak{p}} \frac{1}{1 - N(\mathfrak{p})^{-s}}$$

Here the sum on the left-hand side considers all non-zero ideals of $\mathcal{O}_K$ and their norm, and the product on the right-hand side considers the norm of all prime ideals of $\mathcal{O}_K$, see [70, §11].

2. The Dedekind ζ -function of K extends holomorphically to $\mathbb{C} \setminus \{1\}$. The pole at $s = 1$ is of first order with residuum

$$\operatorname{res}(\zeta_K; 1) = \frac{2^{r_1} (2\pi)^{r_2}}{w \cdot \sqrt{|D|}} \cdot h_K \cdot R$$

The formula combines the characteristic parameters of K : r_1 the number of real embeddings, r_2 the number of pairs of complex embeddings, w the number of roots of unity, D the discriminant, h_K the class number, and R the regulator, see [46, Kap. VII, Kor. 5.11].

3. For the special case of the imaginary quadratic number field

$$K = \mathbb{Q}(\sqrt{-5}) \text{ with discriminant } D = -20$$

the class number h_K can be obtained via explicit computation of the Dirichlet character χ_D . The function ζ_K factorizes as the product of the Riemann ζ -function and the L -series of χ_D as

$$\zeta_K(s) = \zeta(s) \cdot L(s; \chi_D)$$

see [70, §11, Kor. zu Satz 1]. One knows

$$\text{res}(\zeta; 1) = 1$$

Moreover, the discriminant $D = -20$ is a *fundamental* discriminant, see [70, §5 Def.]. Explicit computation of the L -series shows

$$h_K = -\frac{w/2}{|D|} \cdot \sum_{n=1}^{|D|-1} n \cdot \chi_D(n) = -\frac{1}{20} \cdot \sum_{n=1}^{|D|-1} n \cdot \chi_D(n)$$

The values of

$$\chi_D(n), \quad n = 1, \dots, 19,$$

derive from the Legendre symbol, see [70, §5 Def.]:

$$\chi_D(p) := \left(\frac{D}{p} \right), \quad p \text{ prime}$$

$$\chi_D(2) := \begin{cases} 0 & D \equiv 0 \pmod{4} \\ 1 & D \equiv 1 \pmod{8} \\ -1 & D \equiv 5 \pmod{8} \end{cases}$$

$$\chi_D(p_1^{n_1} \cdot \dots \cdot p_k^{n_k}) := \chi_D(p_1)^{n_1} \cdot \dots \cdot \chi_D(p_k)^{n_k}$$

The only non-zero values for the indices n from the interval $[1, 19]$ are

$$\chi_D(n) = 1 \iff n \in \{1, 3, 7, 9\} \text{ and } \chi_D(n) = -1 \iff n \in \{11, 13, 17, 19\}$$

Hence

$$h_K = (-1/20) \cdot (1 + 3 + 7 + 9 - 11 - 13 - 17 - 19) = (-1/20) \cdot (20 - 60) = 2$$

We now apply the theory of complex multiplication to the investigation of the class group of a quadratic imaginary number field K . The goal is to derive an estimate for the class number h_K . Herefore we consider the ring of integers $\mathcal{O}_K \subset K$ and determine the isomorphism classes of complex elliptic curves E with

$$\text{End}(E) \simeq \mathcal{O}_K$$

The proof uses Theorem 4.3 on the uniformization of complex elliptic curves by complex tori.

Proposition 6.29 (Ideal classes and tori with given endomorphism ring). *For each imaginary quadratic field K with ring of integers $\mathcal{O}_K \subset K$ exists a bijection*

$$\text{Cl}(K) \xrightarrow{\sim} \{[T] : T \text{ torus with } \text{End } T = \mathcal{O}_K\}$$

between the ideal class group $Cl(K)$ and the set of classes of biholomorphically equivalent complex tori with endomorphism ring $\mathcal{O}_K$. The bijection maps

Ideal class $[\mathfrak{a}] \mapsto$ torus equivalence class $[T = \mathbb{C}/\Lambda(1, \tau)]$ with $\Lambda(1, \tau) \in [\mathfrak{a}]$

Its inverse maps

torus equivalence class $[T = \mathbb{C}/\Lambda(1, \tau)] \mapsto$ Ideal class $[\Lambda(1, \tau)]$

Proof. 1. *Well-definedness of the map:* Starting with a fractional ideal $\mathfrak{a} \subset K$ there exists a common denominator $d \in \mathcal{O}_K$ with $d \cdot \mathfrak{a} \subset \mathcal{O}_K$, according to Definition 6.21. Then

$$[d \cdot \mathfrak{a}] = [\mathfrak{a}] \in Cl(K)$$

and

$$\Lambda := d \cdot \mathfrak{a} \subset \mathcal{O}_K$$

is a lattice. Due to Proposition 6.4 the torus

$$T := \mathbb{C}/\Lambda$$

has the endomorphism ring

$$End(T) = \{\alpha \in \mathbb{C} : \alpha \cdot \Lambda \subset \Lambda\}$$

Because Λ is an $\mathcal{O}_K$ -module one obtains

$$\mathcal{O}_K \subset End(T).$$

For the opposite inclusion consider a number $\alpha \in End(T)$, i.e. $\alpha \cdot \Lambda \subset \Lambda$. The normalization $\mathbb{Z} + \mathbb{Z} \cdot \tau = \Lambda$ implies

$$\alpha \cdot 1 = \alpha \in \Lambda.$$

2. *Well-definedness in the opposite direction:* We start with a torus class $[T]$. We may assume $T = \mathbb{C}/\Lambda$ with a normalized lattice

$$\Lambda = \Lambda(1, \tau) \text{ satisfying } End(T) = \mathcal{O}_K$$

We have to show that $\Lambda \subset K$ is a fractional ideal. After choosing a $\mathbb{Z}$ -basis $(1, \omega)$ of $\mathcal{O}_K$ the assumption

$$\mathcal{O}_K = End(T) = \{\alpha \in \mathbb{C} : \alpha \cdot \Lambda(1, \tau) \subset \Lambda(1, \tau)\}$$

the multiplication by ω

$$\Lambda(1, \tau) \xrightarrow{\mu_\omega} \Lambda(1, \tau)$$

provides a matrix

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M(2 \times 2, \mathbb{Z})$$

with

$$\omega \cdot \begin{pmatrix} 1 \\ \tau \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot \begin{pmatrix} 1 \\ \tau \end{pmatrix}.$$

The last vector equation combines the two scalar equations

$$\omega = a + b \cdot \tau \text{ and } \omega \cdot \tau = c + d \cdot \tau$$

- The second equation implies

$$\tau = \frac{c}{\omega - d},$$

hence $\Lambda(1, \tau) \subset Q(\mathcal{O}_K) = K$.

- The lattice $\Lambda(1, \tau) \subset K$ is an $\mathcal{O}_K$ -module because both equations taken together imply

$$1, \omega, \tau, \omega \cdot \tau \in \Lambda(1, \tau)$$

- Finally the first equation implies

$$b \cdot \Lambda(1, \tau) \subset \mathbb{Z} + \mathbb{Z} \cdot b \cdot \tau \subset \mathbb{Z} \oplus \mathbb{Z} \cdot \omega = \mathcal{O}_K$$

As a consequence $\Lambda(1, \tau) \subset K$ is a fractional ideal of K .

3. *Bijectivity*: By construction the two maps from part i) and ii) are inverse to each other.

□

Corollary 6.30 (Ideal classes and elliptic curves with given endomorphism ring).

For each imaginary quadratic field K with ring of integers $\mathcal{O}_K \subset K$ exists a bijection

$$Cl(K) \xrightarrow{\sim} \{[E] : E \text{ elliptic curve with } \text{End}(E) \simeq \mathcal{O}_K\}$$

between the ideal class group $Cl(K)$ and the set of classes of biholomorphically equivalent elliptic curves with endomorphism ring $\mathcal{O}_K$.

Proof. The proof follows from the uniformization of elliptic curves $E/\mathbb{C}$ by complex tori T according to Theorem 4.27 and by applying Proposition 6.29.

□

Proposition 6.31 (Extending the Galois group). For an algebraic number field K each automorphism from $Gal(K/\mathbb{Q})$ extends to the Galois group $Gal(\mathbb{C}/\mathbb{Q})$.

Note. The extension in Proposition 6.31 is not canonical.

Proof. A given automorphism $s \in \text{Gal}(K/\mathbb{Q})$ extends in two steps:

1. Extension from K to $\overline{\mathbb{Q}}$: The proof uses Zorn's Lemma. For a given automorphism $s : K \rightarrow K$ from $\text{Gal}(K/\mathbb{Q})$ we consider the family

$$\mathcal{F} := \{(M, \rho) : K \subset M \subset \overline{\mathbb{Q}}, \rho : M \rightarrow M \text{ automorphism with } \rho|_K = s\}$$

of intermediate fields M and extensions $\rho : M \rightarrow \overline{\mathbb{Q}}$ of s , equipped with the partial order

$$(M_1, \rho_1) \leq (M_2, \rho_2) \iff M_1 \subset M_2 \text{ and } \rho_2|_{M_1} = \rho_1$$

Each ascending chain $(M_j, \rho_j)_{j \in J}$ has a maximal bound

$$(M^\infty, \rho^\infty)$$

defined as

$$M^\infty := \bigcup_{j \in J} M_j \text{ and } \tau^\infty : M^\infty \rightarrow M^\infty, \tau^\infty(x) := \tau_j(x) \text{ if } x \in M_j \text{ for suitable } j \in J$$

Note that all elements of M^∞ are algebraic, hence $M^\infty \subset \overline{\mathbb{Q}}$. Now Zorn's Lemma provides a maximal element (P, ϕ) of $\mathcal{F}$.

We prove $P = \overline{\mathbb{Q}}$ by an indirect proof: Assume $P \subsetneq \overline{\mathbb{Q}}$. Then choose an element $\alpha \in \overline{\mathbb{Q}} \setminus P$ and consider the intermediate field $P(\alpha)$ with

$$P \subsetneq P(\alpha) \subset \overline{\mathbb{Q}}$$

and the extension of s

$$\rho : P(\alpha) \rightarrow P(\alpha) \text{ defined by } \rho|_P := \phi, \rho(\alpha) := \alpha$$

Then $(P(\alpha), \rho) > (P, \phi)$, a contradiction.

2. Extension from $\overline{\mathbb{Q}}$ to $\mathbb{C}$: The field extension $\mathbb{C}/\overline{\mathbb{Q}}$ is a pure transcendental extension. After choosing a transcendence basis of $\mathbb{C}$ over $\overline{\mathbb{Q}}$ each automorphism from $\text{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ extends to an element of $\text{Gal}(\mathbb{C}/\overline{\mathbb{Q}})$ which a posteriori is an element of $\text{Gal}(\mathbb{C}/\mathbb{Q})$. Here one uses again Zorn's lemma to prove the existence of a transcendence basis.

□

Theorem 6.32 (Estimating the class number of imaginary quadratic number fields). *The class number of an imaginary quadratic number field K satisfies the estimate*

$$\deg K := [K : \mathbb{Q}] \leq h_K$$

See [61, Chap. II, Rem. 2.2.1].

Proof. Theorem 6.6 provides a uniquely determined CM-point $\tau \in \mathbb{H}$ with $K = \mathbb{Q}(j(\tau))$. According to Corollary 6.30 there exists an elliptic curve $E = E/\mathbb{C}$ with $\text{End}(E) \simeq \mathcal{O}_K$.

The elliptic curve $E \subset \mathbb{P}^2(\mathbb{C})$ is the variety $E = \text{Var}(F)$ of a homogenous polynomial

$$F \in \mathbb{C}[X_0, X_1, X_2]$$

Hence each automorphism $s \in \text{Gal}(\mathbb{C}/\mathbb{Q})$ operates on the coefficients of F . It transforms F into

$$F^s \in \mathbb{C}[X_0, X_1, X_2]$$

and E into an elliptic curve $E^s := \text{Var}(F^s)$. The invariant $j(E) \in \mathbb{C}$ is a rational function of the coefficients of F . Hence s transforms $j(E)$ into the invariant

$$j(E^s) = j(E)^s$$

For an endomorphism $\phi \in \text{End}(E) = \mathcal{O}_K$ holds $\phi^s \in \text{End}(E^s)$, hence

$$\text{End}(E) \simeq \text{End}(E^s) \simeq \mathcal{O}_K$$

The invariant $j(E)$ determines the isomorphy class of the elliptic curve E , while $j(E^s) = j(E)^s$ determines the isomorphy class of E^s .

According to Theorem 6.18 the field extension $\mathbb{Q}(j(\tau))/\mathbb{Q}$ is algebraic. Therefore it has a finite degree

$$m := [\mathbb{Q}(j(\tau)) : \mathbb{Q}]$$

Denote by L the splitting field of $\mathbb{Q}(j(\tau))$. Then there are m elements $s_1, \dots, s_m \in \text{Gal}(L/\mathbb{Q})$ - and a posteriori in $\text{Gal}(\mathbb{C}/\mathbb{Q})$ due to Proposition 6.31 - with pairwise distinct values $s_k(j(\tau)), k = 1, \dots, m$. Hence the corresponding, pairwise non-isomorphic elliptic curves

$$E^{s_k}, k = 1, \dots, m,$$

with endomorphism ring $\mathcal{O}_K$ determine m distinct isomorphy classes. According to Corollary 6.30 these isomorphy classes determine m pairwise distinct ideal classes in $Cl(K)$, which implies

$$[\mathbb{Q}(j(\tau)) : \mathbb{Q}] = m \leq h_K$$

□

Lemma 6.33 shows, that the only order of $\mathcal{O}_K$ which is isomorphic to a subring of $\text{End } T$ is the subring itself.

Lemma 6.33 (Rigidity of orders). *Consider an imaginary quadratic number field K , an order $\Lambda \subset \mathcal{O}_K$ and a subring $R \subset K$. If $\Lambda \simeq R$ as ring isomorphism then $\Lambda = R$.*

Proof. For the proof assume $\Lambda = \mathbb{Z} + \mathbb{Z} \cdot \tau$ and consider a ring isomorphism

$$f : \Lambda \rightarrow R$$

For suitable $a, b \in \mathbb{Z}$ holds

$$\tau^2 = a \cdot \tau + b.$$

Then

$$\begin{aligned} f(\tau)^2 &= f(a \cdot \tau + b) = a \cdot f(\tau) + b \implies \\ \implies (f(\tau) - \tau) \cdot (f(\tau) + \tau) &= f(\tau)^2 - \tau^2 = a \cdot f(\tau) + b - \tau^2 = \\ &= a \cdot f(\tau) + b - a \cdot \tau - b = a \cdot (f(\tau) - \tau) \end{aligned}$$

Hence

$$f(\tau) = \tau \text{ or } f(\tau) = a - \tau$$

In both cases

$$R = f(\mathbb{Z} + \mathbb{Z} \cdot \tau) = \mathbb{Z} + \mathbb{Z} \cdot f(\tau) = \mathbb{Z} + \mathbb{Z} \cdot \tau.$$

□

Corollary 6.34 (Integer values of the modular invariant). *Consider an imaginary quadratic number field K with class number $h_K = 1$ and number ring*

$$\mathcal{O}_K = \mathbb{Z} + \mathbb{Z} \cdot \tau$$

Then $j(\tau) \in \mathbb{Z}$.

Proof. The proof follows from Theorem 6.18 and Theorem 6.32.

□

Theorem 6.32 brings together quite a series of results from complex analysis, algebraic geometry and number theory:

- Both, complex tori and complex elliptic curves are the same mathematical object. The Weierstrass $\wp$ -function transforms the view on it from complex analysis to the view from algebraic geometry.
- The action of the discrete modular group $SL(2, \mathbb{Z})$ and its congruence subgroups on the upper half-plane have compactified orbit spaces, the modular curves. They are compact Riemann surfaces X with the modular forms as differential forms on X of bounded order pole order.

- The arithmetic of number fields and their Dedekind rings of integers on one hand. And on the other hand lattice theory and complex multiplication on complex tori. The theory of modular forms may serve as a bridge to the class field theory of imaginary quadratic number fields.

Chapter 7

Outlook: Modular elliptic curves, monstrous moonshine

7.1 The modularity theorem

The relation between elliptic curves $E/\mathbb{Q}$ and modular forms is given by Wiles' theorem, the proof of the Shimura-Taniyama-Weil conjecture for semistable elliptic curves. These are those elliptic curves which have for each prime p a stable or semistable reduction $E/\mathbb{F}_p$ according Definition 4.39. Shortly after the proof has been generalized to all elliptic curves $E/\mathbb{Q}$ by Breuil, Conrad, Diamond and Taylor in [11]. It is now named the *modularity theorem*.

The case of semistable elliptic curves was sufficient for Wiles's proof of the Fermat conjecture.

The modularity theorem relates for each elliptic curve $E/\mathbb{Q}$ the numbers

$$a_p(E) := 1 + p - \text{card } E_p(\mathbb{F}_p), \quad p \text{ prime},$$

to the eigenvalues of Hecke operators acting on a suitable cusp form as common eigenform. The numbers $a_p(E)$ measure the deviation of the numbers of $\mathbb{F}_p$ -valued points of the reduction $E/\mathbb{F}_p$ from the average value, see Remark 4.41.

To contextualize the last modularity result we start with an analogue - not for the cubics of elliptic curves but for *quadratics*.

Remark 7.1 (Law of quadratic reciprocity). The *Legendre symbol* considers an odd prime $p \in \mathbb{N}$, and determines for each integer $d \in \mathbb{Z}$ whether d is a square mod p . The Legendre symbol is defined for odd prime p as

$$\left(\frac{d}{p}\right) := \begin{cases} +1 & d \not\equiv 0 \pmod{p} \text{ and } d \text{ quadratic rest mod } p \\ -1 & d \not\equiv 0 \pmod{p} \text{ and } d \text{ not quadratic rest mod } p \\ 0 & d \equiv 0 \pmod{p} \end{cases}$$

The law of quadratic reciprocity relates the Legendre symbol of two odd primes $p \neq q$ as

$$\left(\frac{p}{q}\right) \cdot \left(\frac{q}{p}\right) = (-1)^{\frac{p-1}{2} \cdot \frac{q-1}{2}}$$

It has the two complements for odd prime p

$$\left(\frac{-1}{p}\right) = (-1)^{(p-1)/2} = \begin{cases} 1 & p \equiv 1 \pmod{4} \\ -1 & p \equiv 3 \pmod{4} \end{cases}$$

$$\left(\frac{2}{p}\right) = (-1)^{(p^2-1)/8} = \begin{cases} 1 & p \equiv \pm 1 \pmod{8} \\ -1 & p \equiv \pm 3 \pmod{8} \end{cases}$$

Remark 7.2 (Reduction of quadrics modulo primes).

1. Consider an arbitrary, but fixed integer $d \neq 0$ and the quadratic polynomial

$$F(X) := X^2 - d \in \mathbb{Z}[X]$$

which defines the affine quadric $Q/\mathbb{Q}$

$$\{x \in \mathbb{C} : F(x) = 0\}.$$

For each prime p we consider the reduction *mod* p

$$F_p(X) := X^2 - d_p \in \mathbb{F}_p[X]$$

and the set $Q(\mathbb{F}_p)$ of points of Q with coordinates from $\mathbb{F}_p$.

Analogously to Remark 4.41 for elliptic curves we define

$$a_p(Q) := \text{card } Q(\mathbb{F}_p) - 1.$$

The value 1 in the definition is the average number of solutions *mod* p . Apparently, $a_p(Q)$ equals for an odd prime p the *Legendre symbol*

$$a_p(Q) = \left(\frac{d}{p}\right) \text{ while } a_2(Q) = 1 - 1 = 0$$

The values extend to $a_n(Q)$, $n \in \mathbb{N}$, and satisfy the product rule

$$a_{nm}(Q) = a_n(Q) \cdot a_m(Q)$$

like the numerator of the Legendre symbol. As a consequence $a_n(Q) = 0$ for each even n . Due to the Law of Quadratic Reciprocity the Legendre symbol depends only on the residue class $p \pmod{4d}$.

2. We now represent the sequence $(a_p(Q))_{p \text{ prime}}$ as a sequence of eigenvalues of a simultaneous eigenvector for a suitable sequence of linear endomorphisms $(T_p)_{p \text{ prime}}$. Our aim is to formulate an analogue of the modular group, of modular functions and of Hecke operators:

Set

$$N = N_Q := 4|d|,$$

and consider the multiplicative group

$$G := (\mathbb{Z}/N\mathbb{Z})^*,$$

and the complex vector space of functions defined on G

$$V := \{f : G \rightarrow \mathbb{C}\}.$$

Eventually, as an analogue to the Hecke operators we introduce the family of linear endomorphisms $(\hat{T}_p)_{p \text{ prime}}$ of V defined as

$$(\hat{T}_p f)(n) := \begin{cases} f(p \cdot n) & p \nmid N \\ 0 & p \mid N \end{cases}$$

for $n \in G$. We claim: All endomorphism $(\hat{T}_p)_{p \text{ prime}}$ have a simultaneous eigenvector, namely the distinguished function

$$f_Q : G \rightarrow \mathbb{C}, f(n) := a_n(Q).$$

Note that f_Q is well-defined on the residue classes of G because the Legendre symbol depends on the residue class $p \bmod 4d$. Computing the eigenvalues we get for all $n \in G$

$$\hat{T}_p(f_Q)(n) = \begin{cases} f_Q(p \cdot n) = a_{pn}(Q) = a_p(Q) \cdot a_n(Q) & p \nmid N \\ 0 & p \mid N \end{cases}$$

As a consequence, for all primes $p \in \mathbb{Z}$

$$\hat{T}_p(f_Q) = a_p(Q) \cdot f_Q.$$

Theorem 7.3 (Wiles' modularity theorem). *The Shimura-Taniyama-Weil conjecture is true.*

For Wiles' proof of Theorem 7.3 in the stable and the semistable case see the symposium [16]. Concerning the remaining cases subsequently proved by several authors see [11].

When comparing rational elliptic curves $E/\mathbb{Q}$ and cusp forms $f \in S_k(\Gamma_0(N))$ one compares the L -series of both mathematical objects. For the elliptic curve $E/\mathbb{Q}$ recall the L -series $L(E, -)$ arising from the numbers $(a_p(E))_{p \text{ prime}}$, see Definitions 4.43 and 4.45, and also Remark 4.44.

For the cusp form f see Definition 7.4.

Definition 7.4 (L -series of a cusp form). Consider a level $N \in \mathbb{N}^*$ and a weight $k \in \mathbb{N}$. For a cusp form $f \in S_k(\Gamma_0(N))$ with Fourier series

$$f(\tau) = \sum_{n=1}^{\infty} c_n \cdot q^n, \quad q := e^{2\pi i \tau},$$

one defines its L -series

$$L(f, -) : \{s \in \mathbb{C} : \operatorname{Re} s > 1 + (k/2)\} \rightarrow \mathbb{C}, \quad L(f, s) := \sum_{n=1}^{\infty} \frac{c_n}{n^s}$$

Remark 7.5 (L -series of a cusp form).

1. The L -series $L(f, -)$ of a cusp form $f \in S_k(\Gamma_0(N))$ from Definition 7.4 is well-defined. One obtains $L(f, -)$ from the Mellin transform of f , see [35, Chap. VIII, §5].
2. The L -series $L(f, -)$ of $f \in S_k(\Gamma_0(N))$ extends to a holomorphic function on $\mathbb{C}$ and satisfies the functional equation with respect to the reflection at k

$$\frac{1}{(2\pi)^s} \cdot \Gamma(s) \cdot L(f, s) = (-1)^{k/2} \cdot \frac{1}{(2\pi)^{k-s}} \cdot \Gamma(k-s) \cdot L(f, k-s)$$

see [35, Chap. VIII, Theor. 8.1]

The textbook [19] does not prove Theorem 7.3. But it gives an excellent introduction to the problem and it highlights different views onto the theorem.

Theorem 7.6 (Versions of the modularity theorem).

1. Complex elliptic curve: *Each elliptic curve $E/\mathbb{C}$ with rational modular invariant*

$$j_E = \frac{1728 \cdot g_2^3}{g_3^3 - 27 \cdot g_2^2} \in \mathbb{Q}$$

is uniformized by a modular curve, i.e. there exists an integer $N \in \mathbb{N}$ and a surjective holomorphic map

$$X_0(N) \rightarrow E/\mathbb{C},$$

see [19, Theor. 2.5.1].

2. Uniformisation of elliptic curves: *Each elliptic curve $E/\mathbb{Q}$ is uniformized by a modular curve defined over $\mathbb{Q}$, i.e. there exists an integer $N \in \mathbb{N}$ and a surjective regular map over $\mathbb{Q}$*

$$X_0(N)/\mathbb{Q} \rightarrow E/\mathbb{Q},$$

see [19, Theor. 7.7.2].

3. Jacobi torus: *For each elliptic curve $E/\mathbb{C}$ with rational modular invariant*

$$j_E = \frac{1728 \cdot g_2^3}{g_2^3 - 27 \cdot g_3^2} \in \mathbb{Q}$$

exists an integer $N \in \mathbb{N}$ and a surjective holomorphic map

$$\text{Jac}(X_0(N)) \rightarrow E,$$

from the Jacobi torus of a modular curve $X_0(N)$, $N \in \mathbb{N}$, which is a group homomorphism, see [19, Theor. 6.1.3].

4. Cusp form, version with coefficients: *For each elliptic curve $E/\mathbb{Q}$ with conductor $N_E \in \mathbb{N}$ exists a new form*

$$f(q) = \sum_{n=1}^{\infty} a_n(f) \cdot q^n \in S_2(\Gamma_0(N_E))$$

satisfying for all primes $p \in \mathbb{N}$

$$a_p(f) = a_p(E),$$

see [19, Theor. 8.8.1].

Recall from Definition 4.39 for an elliptic curve $E/\mathbb{Q}$ with global minimal Weierstrass polynomial $F \in \mathbb{Z}[X, Y]$ the conductor

$$\text{cond } E = \prod_p p^{e_p} \in \mathbb{Z},$$

having the same prime factors p as the discriminant Δ_F and exponents $e_p \in \mathbb{N}^$ deriving from the type of singularity of the reduction $E(\mathbb{F}_p)$.*

5. Cusp form, version with L -series: *For each elliptic curve $E/\mathbb{Q}$ with conductor $N_E \in \mathbb{N}$ exists a new form*

$$f(q) = \sum_{n=1}^{\infty} a_n(f) \cdot q^n \in S_2(\Gamma_0(N_E))$$

with

$$L(f, -) = L(E, -),$$

see [19, Theor. 8.8.3]

Remark 7.7 (Forerunner of the Shimura-Taniyama-Weil Conjecture).

1. It is known that a non-constant holomorphic map

$$X_0(N) \rightarrow E/\mathbb{C}$$

as required in Theorem 7.6 is defined by two automorphic functions, see Definition 3.6,

$$u, v \in A_0(\Gamma_0(N))$$

which satisfy a Weierstrass equation of $E/\mathbb{Q}$.

2. If the level N is minimal such that a non-constant holomorphic map

$$X_0(N) \rightarrow E(\mathbb{C})$$

exists, then the map is induced by a new form $f \in S_2^{\text{new}}(\Gamma_0(N))$.

3. If the new form $f \in S_2^{\text{new}}(\Gamma_0(N))$ from the previous step has integer-valued Fourier coefficients then

$$L(f, -) = L(E, -)$$

and N is the conductor of E .

4. The Fermat conjecture once stated that the Fermat equation

$$X^n + Y^n = Z^n, \quad n \in \mathbb{N}, \quad n \geq 3,$$

has no non-trivial integer solutions.

For the particular exponents $n = 3, 4$ the Fermat conjecture was proved by Euler. In addition it was known: The general case follows from the case of prime exponents $n = l \geq 5$. Before 1995 the conjecture was open for general primes $l \geq 5$.

Assuming for an indirect proof of the Fermat conjecture the existence of a non-trivial solution

$$(a, b, c) \in \mathbb{Z}, \quad \gcd(a, b, c) = 1,$$

of the Fermat equation with prime exponent $l \geq 5$ Frey proposed to study the elliptic curve $E_{FH}/\mathbb{Q}$ with Weierstrass equation

$$Y^2 = X \cdot (X - a^l) \cdot (X - c^l),$$

today named the *Frey-Hellegouarch* curve. The relation between non-trivial solutions of the Fermat equation and points on elliptic curves had already been considered by Hellegouarch, see [32]. The curve E_{FH} has discriminant

$$\Delta = 2^4 \cdot (abc)^{2l}$$

and modular j -invariant

$$j = \frac{c_4^3}{\Delta} \text{ with } c_4 = 2^4 \cdot (a^{2l} - (ac)^l + c^{2l})$$

The global minimal Weierstrass polynomial of E_{FH} has conductor

$$N = \prod_{p|abc} p$$

Serre and Ribet proved that the existence of E_{FH} would contradict the validity of the Shimura-Taniyama-Weil conjecture. Hence Wiles' modularity theorem, Theorem 7.3, implies that the Frey-Hellegouarche curve E_{FH} does not exist. Hence the Fermat conjecture is true.

A reference for all issues of this remark is [35, Chap. XII, § 1 and 4].

Example 7.8 (Modularity theorem). Figure 7.1 shows different elliptic curves $E/\mathbb{Q}$ defined by a minimal Weierstrass polynomial, their conductor N , and a sample of coefficients of their L -series and of the Fourier coefficients of the new-forms in

$$S_2(\Gamma_0(N)) = H^0(X_0(N), \Omega),$$

see Corollary 3.28. The numerical result from the PARI script `Elliptic_curve_taniyama_10` confirms the Modularity Theorem that a rational elliptic curves arises from a modular form.

```

Case 1: Global minimal model:  $Y^2 + (Y) = X^3 + (-X)$ 
j(E) = 110592/37, minimal discr.: 37 = Mat([37, 1]), conductor: 37 = Mat([37, 1])
Genus  $X_0(37) = \dim S_2(\Gamma_0(37))$ : 2
L-series:  $q - 2*q^2 - 3*q^3 + 2*q^4 - 2*q^5 + 6*q^6 - q^7 + 6*q^9 + 4*q^{10} - 5*q^{11} - 6*q^{12} - 2*q^{13} + O(q^{14})$ 

Old forms:  $S_{\text{old}}^2(\Gamma_0(37))$  has dimension: 0
New forms:  $S_{\text{new}}^2(\Gamma_0(37))$  has dimension: 2
basis_1:  $q + q^3 - 2*q^4 - q^7 - 2*q^9 + 3*q^{11} - 2*q^{12} - 4*q^{13} + O(q^{14})$ 
basis_2:  $q - 2*q^2 - 3*q^3 + 2*q^4 - 2*q^5 + 6*q^6 - q^7 + 6*q^9 + 4*q^{10} - 5*q^{11} - 6*q^{12} - 2*q^{13} + O(q^{14})$ 
-----
Case 2: Global minimal model:  $Y^2 + (YX) + (Y) = X^3 + (-X)$ 
j(E) = -15625/28, minimal discr.: -28 = [-1, 1; 2, 2; 7, 1], conductor: 14 = [2, 1; 7, 1]
Genus  $X_0(14) = \dim S_2(\Gamma_0(14))$ : 1
L-series:  $q - q^2 - 2*q^3 + q^4 + 2*q^6 + q^7 - q^8 + q^9 - 2*q^{12} - 4*q^{13} + O(q^{14})$ 
Old forms:  $S_{\text{old}}^2(\Gamma_0(14))$  has dimension: 0
New forms:  $S_{\text{new}}^2(\Gamma_0(14))$  has dimension: 1
basis_1:  $q - q^2 - 2*q^3 + q^4 + 2*q^6 + q^7 - q^8 + q^9 - 2*q^{12} - 4*q^{13} + O(q^{14})$ 
-----
Case 3: Global minimal model:  $Y^2 + (YX) + (Y) = X^3 + (-11*X) + (12)$ 
j(E) = 128787625/98, minimal discr.: 98 = [2, 1; 7, 2], conductor: 14 = [2, 1; 7, 1]
Genus  $X_0(14) = \dim S_2(\Gamma_0(14))$ : 1
L-series:  $q - q^2 - 2*q^3 + q^4 + 2*q^6 + q^7 - q^8 + q^9 - 2*q^{12} - 4*q^{13} + O(q^{14})$ 
Old forms:  $S_{\text{old}}^2(\Gamma_0(14))$  has dimension: 0
New forms:  $S_{\text{new}}^2(\Gamma_0(14))$  has dimension: 1
basis_1:  $q - q^2 - 2*q^3 + q^4 + 2*q^6 + q^7 - q^8 + q^9 - 2*q^{12} - 4*q^{13} + O(q^{14})$ 

```

Fig. 7.1 Relation between rational elliptic curves and new forms

7.2 Monstrous moonshine

What is “monstrous moonshine”? Monstrous moonshine in mathematics is a synonym for the relation between modular forms on one side and finite simple groups on the other side.

That type of moonshine was detected in 1979, then formalized as a conjecture, proved by Borcherds in 1992, and finally honored by a Fields Medal awarded to him in 1998.

For an introduction see [8] and [65], on a more advanced level see [26] and [27].

Definition 7.9 (Simple group). A group G is *simple* if it has no normal, proper subgroup $H \subset G$, i.e. satisfying

$$H \neq \{e\} \text{ and } H \neq G$$

If a finite group G has a proper normal subgroup $H \subsetneq G$ then one may investigate G by studying the smaller groups H and G/H . Proposition 7.10 generalizes this construction. It formalizes in which sense simple groups are the building blocks of all finite groups.

Proposition 7.10 (Jordan-Hölder theorem). *Each finite group G has a finite, strictly increasing sequence of subgroups, a composition series,*

$$G_0 = \{e\} \subset G_1 \subset \dots \subset G_n = G$$

such that for each $i = 1, \dots, n$

$$G_{i-1} \subset G_i$$

is a normal subgroup with simple quotient group

$$Q_i := G_i / G_{i-1}$$

All composition series of G have the same length. Each two composition series of G have - up to permutation - the same simple quotients counted with multiplicity.

For a proof see [40, Chap. IV, §4].

Example 7.11 (Composition series). The cyclic groups C_{12} has the following different composition series

$$\{0\} \subset C_3 \subset C_6 \subset C_{12}$$

$$\{0\} \subset C_2 \subset C_6 \subset C_{12}$$

$$\{0\} \subset C_2 \subset C_4 \subset C_{12}$$

Each composition series of C_{12} has the same family of simple quotients (C_2, C_2, C_3) up to permutation.

Definition 7.12 (Group representation). Consider a group G .

1. A G -action on the vector space V

$$G \times V \rightarrow V, (g, v) \mapsto g.v,$$

is a group morphism to the automorphisms of the vector space V

$$\rho : G \rightarrow GL(V)$$

As a shorthand one calls the vector space a G -module.

2. Each group representation

$$\rho : G \rightarrow GL(V)$$

with a finite dimensional vector space V induces a group representation on the 1-dimensional vector space $\mathbb{C}$, which is named the *trace* of ρ

$$\text{trace } \rho : G \rightarrow \mathbb{C}^* = GL(\mathbb{C}), \chi_\rho(g) := \text{trace } \rho(g),$$

The G -representation for monstrous moonshine is infinite-dimensional. But it splits as the direct sum of finite-dimensional representations. For a finite-dimensional representation the trace of an element $g \in G$ only depends on the *conjugacy class* of g_0

$$[g_0] := \{g \cdot g_0 \cdot g^{-1} : g \in G\}$$

because invariance of the trace under cyclic permutation shows for $g_0, g \in G$

$$\begin{aligned} \text{trace } \rho(g \cdot g_0 \cdot g^{-1}) &= \text{trace}(\rho(g) \cdot \rho(g_0) \cdot \rho(g)^{-1}) = \\ &= (\text{trace}(\rho(g_0) \cdot \rho(g)^{-1} \cdot \rho(g))) = \text{trace } \rho(g_0) \end{aligned}$$

Apparently the trace of the neutral element equals the dimension of the representation

$$\text{trace } \rho(e) = \dim V \in \mathbb{N}.$$

Theorem 7.13 (Complete reducibility). *Consider a finite group G .*

1. *Any G -module is completely reducible, i.e.*

$$V = \bigoplus_{i \in I} V_i$$

with irreducible G -modules $V_i, i \in I$.

2. *The irreducible G -modules correspond bijectively to the conjugacy classes of G .*

Proof. See [55, §2, Theor. 7].

Remark 7.14 (Classification of finite simple groups).

1. Each finite simple group falls into exactly one of the following classes
 - a set of 18 infinite sequences of groups
 - a set of 26 groups, named the *sporadics*.
2. To the first class belongs the sequence $(A_n)_{n \geq 5}$ of the group of alternating permutations of n elements, the sequence $C(p)_{p \text{ prime}}$ of cyclic groups with p elements, and a series of Lie type groups like $(PSL(n, \mathbb{F}_{p^k}))_n$.
3. The sporadics can be distinguished by their order. The next-to-biggest sporadic is Fischer's *baby monster* $\mathbb{B}$ with

$$\text{ord } \mathbb{B} = 2^{41} \cdot 3^{13} \cdot 5^6 \cdot 7^2 \cdot 11 \cdot 13 \cdot 17 \cdot 19 \cdot 23 \cdot 31 \cdot 47 \sim 4 \cdot 10^{33}$$

4. The largest sporadic is the *monster* $\mathbb{M}$ with

$$\text{ord } \mathbb{M} = 2^{46} \cdot 3^{20} \cdot 5^9 \cdot 7^6 \cdot 11^2 \cdot 13^3 \cdot 17 \cdot 19 \cdot 23 \cdot 29 \cdot 31 \cdot 41 \cdot 47 \cdot 59 \cdot 71 \sim 8 \cdot 10^{53}$$

The monster has 194 conjugacy classes and therefore - due to Theorem 7.13 - exactly 194 irreducible representations. Their dimensions are well-known and tabulated, cf. [58].

Example 7.15 (Monstrous Moonshine). Figure 7.2 shows for all indices $n \leq 18$

- on the left-hand side the Fourier coefficients $c(n)$ of the q -expansion of the modular j -invariant
- and on the right-hand side for the monster $\mathbb{M}$ the dimensions $a(n)$ of the irreducible, complex $\mathbb{M}$ -modules V_n ordered by dimension.

n: -1	1
n: 0	744
n: 1	196884
n: 2	21493760
n: 3	864299970
n: 4	20245856256
n: 5	333202640600
n: 6	4252023300096
n: 7	44656994071935
n: 8	401490886656000
n: 9	3176440229784420
n: 10	22567393309593600
n: 11	146211911499519294
n: 12	874313719685775360
n: 13	4872010111798142520
n: 14	25497827389410525184
n: 15	126142916465781843075
n: 16	593121772421445058560
n: 17	2662842413150775245160
n: 18	11459912788444786513920

Coeff. of q -expansion of modular j -invariant

n	a(n)
1	1
2	196883
3	21296876
4	842609326
5	18538750076
6	19360062527
7	293553734298
8	3879214937598
9	36173193327999
10	125510727015275
11	190292345709543
12	222879856734249
13	1044868466775133
14	1109944460516150
15	2374124840062976
16	8980616927734375
17	8980616927734375
18	15178147608537368

Dimension of irred. represent. of monster $\mathbb{M}$

Fig. 7.2 Monstrous Moonshine

What later was named “monstrous moonshine” started with the observations from Figure 7.2, line 1 and 2,

$$c(-1) = a(1) = 1 \text{ and } c(1) = 196.884 = 1 + 196.883 = a(1) + a(2)$$

by McKay in 1978. The formula states

$$c(-1) = a(1) \text{ and } c(1) = a(1) + a(2)$$

relating two coefficients of the q -expansion of the modular j -invariant to the dimensions of the two lowest irreducible $\mathbb{M}$ -modules. As Borcherds remarks “*moonshine* is not a poetic term referring to light from the moon. It means foolish or crazy idea (Deutsch: Verrückte Idee).”

Definition 7.16 (Moonshine group and ”Hauptmodul”).

1. *Moonshine group*: A discrete subgroup $G \subset SL(2, \mathbb{R})$ is a group of *moonshine type* if for a suitable level $N \in \mathbb{N}^*$

$$\Gamma_0(N) \subset G$$

and for a translation

$$\gamma = \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix} \in G \iff b \in \mathbb{Z},$$

i.e. G contains all integer-valued translations and no other translations.

2. *Hauptmodul*: The orbit space of the canonical action of a group G of moonshine type on the extended upper half plane $\mathbb{H}^*$

$$\phi : G \times \mathbb{H}^* \rightarrow \mathbb{H}^*$$

is a compact Riemann surface $X(G)$, the *modular curve* of G . For G with

$$\text{genus } X(G) = 0$$

the unique meromorphic function $J_G \in \mathcal{M}(X(G))$ with Fourier expansion normalized as

$$J_G(\tau) = \frac{1}{q} + \sum_{n=1}^{\infty} a_n \cdot q^n, \quad q = e^{2\pi i \cdot \tau},$$

is the *Hauptmodul* of G .

Moonshine groups $G \subset SL(2, \mathbb{R})$ are *congruent* to the modular group Γ in the following sense: The intersection

$$H := G \cap \Gamma$$

satisfies the equality of indices

$$[G : H] = [\Gamma : H].$$

In particular, the index is finite because

$$[\Gamma : H] \leq [\Gamma : \Gamma_0(N)] < \infty.$$

For a moonshine group G of genus $X(G) = 0$ the modular curve $X(G)$ is $\mathbb{P}^1$, and J_G is a distinguished generator of the field of meromorphic functions

$$\mathcal{M}(X(G)) = \mathbb{C}(J_G).$$

Example 7.17 (Hauptmodul).

- The Hauptmodul of the modular group Γ is the normalized modular j -invariant

$$J_\Gamma(q) = j(q) - 744 = \frac{1}{q} + 196'884 \cdot q + 21'493'760 \cdot q^2 + 864'299'970 \cdot q^3 + O(4)$$

- The Hecke congruence subgroup $\Gamma_0(2)$ has the Hauptmodul

$$J_{\Gamma_0(2)} = \frac{1}{q} + 276 \cdot q - 2'048 \cdot q^2 + 11'202 \cdot q^3 + O(4)$$

- The normalizer of $\Gamma_0(2)$ in $SL(2, \mathbb{R})$, the moonshine group

$$\Gamma_0(2)_+ \subset SL(2, \mathbb{R})$$

has the Hauptmodul

$$J_{\Gamma_0(2)_+} = \frac{1}{q} + 4'372 \cdot q + 96'256 \cdot q^2 + O(3)$$

Remark 7.18 (Forerunner of the Monstrous Moonshine Theorem).

1. *Thompson conjecture:* Thompson conjectured the existence of a graded, infinite-dimensional $\mathbb{M}$ -module

$$V^\natural = \bigoplus_{n \in \mathbb{N}} V_n^\natural$$

satisfying: The graded components $V_n^\natural$ split as the direct sum of irreducible $\mathbb{M}$ -modules with a certain multiplicity. The $\mathbb{M}$ -module $V^\natural$ is characterized by the family of series, later named *McKay-Thompson series*

$$T_{[g]}(\tau) = \sum_{n \in \mathbb{N}} \text{trace}(g|V_n^\natural) \cdot q^n, \quad g \in \mathbb{M},$$

see [66].

2. *Conway-Norton conjecture*: Each McKay-Thompson series $T_{[g]}$ is the Hauptmodul of a moonshine group $\Gamma_{[g]}$ with modular curve of genus $= 0$.

Thompson decoded the data from Figure 7.2 as follows:

$$c(2) = 21'493'760 = a(1) + a(2) + a(3)$$

and subsequently

$$c(3) = 2 \cdot a(1) + 2 \cdot a(2) + a(3) + a(4)$$

$$c(4) = 3 \cdot a(1) + 3 \cdot a(2) + a(3) + 2 \cdot a(4) + a(5)$$

$$c(5) = 4 \cdot a(1) + 5 \cdot a(2) + 3 \cdot a(3) + 2 \cdot a(4) + a(5) + a(6) + a(7)$$

If the irreducible representations ρ_i , $i \in \mathbb{N}$, of $\mathbb{M}$ are ordered according to increasing dimension then the first summands of the representation from Remark 7.18 are

$$(V_0, \rho_1), (V_1 = \{0\}), (V_2, \rho_1 \oplus \rho_{196'883}), (V_3, \rho_1 \oplus \rho_{196'883} \oplus \rho_{21'296'876})$$

The series generated by the dimensions $(\dim V_n)_{n \in \mathbb{N}}$ has the Fourier expansion

$$\begin{aligned} \sum_{n=-1}^{\infty} (\dim V_{n+1}) \cdot q^n &= J(\tau) = \\ &= \frac{1}{q} + 196'884 \cdot q + 21'493'760 \cdot q^2 + 21'493'760 \cdot q^3 + O(4) \end{aligned}$$

with $J(\tau) := j(\tau) - 744$, $q = e^{2\pi i \tau}$.

The open problem was to find explicitly a representation with the properties from the Conway-Norton conjecture, see Remark 7.18. The final solution came from Borcherds. He proved in 1992 Theorem 7.19, see [6].

Theorem 7.19 (Monstrous Moonshine). *The monster group $\mathbb{M}$ has a representation as automorphism group of the monster vertex algebra $V^\natural$. The representation satisfies the properties of the Conway-Norton conjecture.*

List of results

Part I. General Theory

Chapter 1. Elliptic functions

Focus: Complex analysis

Holomorphic elliptic functions are constant (Prop. 1.7)

Sum of residues vanishes (Prop. 1.8)

Elliptic functions attain each value with equal multiplicity (Cor. 1.10)

Weierstrass $\wp$ -function of a lattice (Theor. 1.12)

Differential equation of $\wp$ (Theor. 1.17)

The Field $\mathbb{C}(\wp)[\wp']$ of elliptic functions (Theor. 1.18)

Elliptic functions with prescribed zeros and poles (Theor. 1.21)

Chapter 2. Modular group and congruence subgroups

Focus: Complex analysis, Topology

Holomorphic maps between complex tori (Prop. 2.3)

Moduli space of complex tori (Theor. 2.5)

Fundamental domain of the $SL(2, \mathbb{Z})$ -action (Theor. 2.18)

Proper discontinuity of the $SL(2, \mathbb{Z})$ -action (Lem. 2.21)

The orbit space of the $SL(2, \mathbb{Z})$ -action is Hausdorff (Theor. 2.25, Cor. 2.28)

Complex structure of the orbit space of the $SL(2, \mathbb{Z})$ -action (Prop. 2.30)

The modular curve as compact Riemann surface (Theor. 2.31)

The modular curves of the Hecke congruence subgroups (Rem. 2.33)

Chapter 3. The algebra of modular forms

Focus: Complex analysis

$SL(2, \mathbb{Z})$ -invariant differential forms (Prop. 3.1)

Eisenstein series (Theor. 3.9)

Modular curve and modular invariant j (Cor. 3.16)

Modular forms as meromorphic diff. forms on the modular curve (Theor. 3.17)

The field $\mathbb{C}(j)$ of automorphic functions (Cor. 3.22)

The algebra of modular forms and the ideal of cusp forms (Theor. 3.24)

The line bundle of modular forms (Theor. 3.26)

The codimension of cusp forms (Cor. 3.28)

Chapter 4. Elliptic curves

Focus: Complex analysis, algebraic geometry, arithmetic geometry

Embedding tori (Theor. 4.3, Rem. 4.14)

Very-ampleness criterion (Theor. 4.13)

Elliptic curves are plane cubics (Theor. 4.23)

Elliptic curves over $\mathbb{C}$ are tori (Theor. 4.27)

The group of classes of degree zero divisors of an elliptic curve (Theor. 4.30)

Abel's theorem for elliptic curves (Theor. 4.32)

Chapter 5. Introduction to Hecke theory and applications

Focus: Complex analysis, number theory

Fourier coefficients of Hecke transforms (Theor. 5.12)

The discriminant Δ as simultaneous eigenform (Cor. 5.14)

Commutativity of the Hecke algebra (Theor. 5.17)

Petersson scalar product (Theor. 5.22)

Self-adjointness of the Hecke operators of the modular group (Theor. 5.25)

Eigenforms of the Hecke algebra of the modular group (Theor. 5.26)

Properties of the Ramanujan τ -function (Prop. 5.32)

The modular forms $G_{2,N} \in M_2(\Gamma_0(N))$ (Theor. 5.36)

The 4-squares theorem (Theor. 5.37)

Part II. Advanced Theory

Chapter 6. Application to imaginary quadratic fields

Focus: Complex analysis, algebraic number theory

Complex multiplication and imaginary quadratic fields (Theor. 6.6)

Coefficients of the modular polynomials from $\mathbb{Z}[j]$ (Theor. 6.13)

Singular values $j(\tau) \in \mathbb{C}$ of CM -points are algebraic integers (Theor. 6.18)

The ideal class group and tori with complex multiplication (Prop. 6.29)

Estimating the class number of an imaginary quadratic number field (Theor. 6.32)

Chapter 7. Modular elliptic curves, monstrous moonshine

Focus: Complex analysis, algebraic geometry, group theory

The modularity theorem (Theor. 7.3)

The Monstrous Moonshine theorem (Theor. 7.19)

References

References for these notes are grouped as

- Modular Forms: [9], [12] [16] [18] [21] [36] [38] [41] [45] [47] [48] [49] [54] [56] [59] [?] [71]
- Complex Analysis: [1] [5] [23] [29]
- Algebraic Geometry: [25] [31] [53] [60] [62]
- Number Theory: [14] [24] [28] [34] [42] [51]
- Online platform: [63] (textbook level), [43] (research level)

The basic textbook is [19], for an elementary survey see [2].

1. Ahlfors, Lars: Complex Analysis. 2nd edition, McGraw-Hill, Tokyo (1966)
2. Alfes-Neumann, Claudia: Modulformen. Fundamentale Werkzeuge der Mathematik. Springer, Wiesbaden (2020)
3. Alperin, Roger C.: $PSL_2(\mathbb{Z}) = \mathbb{Z}_2 * \mathbb{Z}_3$. The American Mathematical Monthly. Vol. 100 (1993), 385-386
4. Atiyah, M. F.; Macdonald, I. G.: Introduction to Commutative Algebra. Addison Wesley (1969)
5. Ben-Zvi, David, D.: Moduli Spaces. <https://www.ma.utexas.edu/users/benzvi/math/pcm0178.pdf>
6. Borcherds, Richard E.: Monstrous Moonshine and Monstrous Lie Superalgebras. Invent. Math. 109, (1992), p. 405-444
7. Borcherds, Richard E.: Proceedings of the I.C.M., Vol. I (Berlin, 1998). Doc. Math. 1998, Extra Vol. I, 607-615, <http://math.berkeley.edu/~reb/papers/icm98/icm98.pdf> Call 8.7.2020
8. Borcherds, Richard E.: What is ... The Monster? Notices of the AMS, 49,7 (2002), p. 1076-1077
9. Borel, A.; Chowla, S.; Herz, C.; Iwasawa, K.; Serre, J-P.: Seminar on Complex Multiplication. Lecture Notes in Mathematics 21. Springer, Berlin (1966)
10. Bourbaki, Nicolas: Elements of Mathematics. General Topology, Chapters 1-4. Springer (1989)

11. Breuil, Christophe; Conrad, Brian; Diamond, Fred; Taylor, Richard: On the Modularity of Elliptic Curves over $\mathbb{Q}$: Wild 3-adic exercises. *Journal of the American Mathematical Society*. 14(4), p. 843-939 (2001)
12. Chenevier, Gaëtan: Introduction aux Formes Modulaires. Internet (2015)
13. Cheng, Miranda: Moonshine-I.
<https://www.youtube.com/watch?v=HB5dCXQIWqM> Call 8.7.2020
14. Cohen, Henri: A Course in Computational Algebraic Number Theory. Springer, Berlin (1993)
15. Conway, J.H.; Norton, S.P.: Monstrous Moonshine. *Bull. London. Math. Soc.* 11, (1979), 308-339
16. Cornell, Gary; Silvermann, Joseph H.; Stevens, Glenn (Eds.): *Modular Forms and Fermat's Last Theorem*. Springer, New York (1997)
17. Deo, Satya; Varadarajan, K.: Discrete groups and discontinuous actions. *Rocky Mountain Journal of Mathematics*. 27,2 (1997)
18. Darmon, Henri: A proof of the full Shimura-Taniyama-Weil conjecture is announced. *Notices Americ. Math. Soc.* 46,11 (1999), p. 1397-1401
19. Diamond, Fred; Shurman, Jerry: *A First Course in Modular Forms*. Springer, New York (2005)
20. Dugundji, James: *Topology*. Allyn and Bacon, Boston (1966)
21. Eichler, Martin; Zagier, Don: On the Zeros of the Weierstrass $\wp$ -Function. *Mathematische Annalen* 258, (1982), p. 399-407
22. Fischer, Wolfgang; Lieb, Ingo: *Funktionentheorie*. Vieweg, Braunschweig (1980)
23. Forster, Otto: *Riemannsche Flächen*. Springer, Berlin (1977)
24. Forster, Otto: *Algorithmische Zahlentheorie*. 2. Auflage Springer Spektrum (2015)
25. Fulton, William: *Algebraic Curves. An Introduction to Algebraic Geometry*. The Benjamin/Cummings Publishing Company (1969)
26. Gannon, Terry: Monstrous Moonshine: The first twenty-five years. *Bull. London Math. Soc.* 38, p. 1-33 (2006)
27. Gannon, Terry: *Moonshine beyond the Monster. The bridge connecting Algebra, Modular Forms and Physics*, Cambridge University Press, Cambridge (2006)
28. Gross, Benedict: The arithmetic of elliptic curves - An update. *Arabian Journal of Science and Engineering* 1 (2009), p. 95-103
<http://www.math.harvard.edu/~gross/preprints/ell2.pdf>
29. Gunning, Robert C.: *Lectures on Riemann Surfaces*. Princeton University Press (1966)
30. Han, Juncheol: The general linear group over a ring. *Bull. Korean Math. Soc.* 43 (2006), No. 3, pp. 619-626
31. Hartshorne, Robin: *Algebraic Geometry*. Springer, New York (1977)
32. Y. Hellegouarch: Points d'ordre $2p^h$ sur les courbes elliptiques. *Acta Arithmetica* 26 (1972), 253-263
33. Huybrechts, Daniel: The geometry of cubic hypersurfaces. (2020)
<http://www.math.uni-bonn.de/people/huybrech/Notes.pdf> Call 15.9.2020
34. Ireland, Kenneth; Rosen, Michael: *A Classical Introduction to Modern Number Theory*. Springer, New York (1982)
35. Knapp, Anthony: *Elliptic Curves*. Princeton University Press. Princeton (1992)
36. Kilford, Lloyd James: *Modular forms: a classical and computational introduction*. Imperial College Press, 2nd ed. (2015)
37. Koblitz, Neal: *Introduction to Elliptic Curves and Modular Forms*. Springer, New York, 2nd ed. (1993)
38. Koecher, Max; Krieg, Aloys: *Elliptische Funktionen und Modulformen*. 2. Aufl. Springer, Berlin (2007)
39. Lang, Serge: *Algebra*. Addison-Wesley, Reading, 7th ed. (1977)
40. Lang, Serge: *Elliptic Functions*. 2nd edition Springer, Berlin (1987)
41. Lang, Serge: *Introduction to Modular Forms*. Springer, Berlin (1995)
42. Marcus, Daniel A.: *Number Fields*. Springer, New York (1977)
43. mathoverflow. <https://mathoverflow.net/>

44. Mines, Ray; Richmann, Fred; Ruitenburg, Wim: A Course in Constructive Algebra. Springer, New York (1988)
45. Mordell, Louis Joel: On Mr. Ramanujan's empirical expansion of modular functions. Proc. Cambridge Philosophical Society, **19** (1917), p. 117-124
46. Neukirch, Jürgen: Algebraische Zahlentheorie. Springer (1992)
47. Ono, Ken: The last words of a genius. Notices Americ. Math. Soc. **57**, 11 (2010), p.1410-1419
48. Panchichkine, Alexei: Formes Modulaire et Courbes Elliptiques. <https://www-fourier.ujf-grenoble.fr/~panchish/06ensl.pdf>
49. Ramanujan, Srinivasa: On certain arithmetical functions. Trans. Cambridge Philosophical Society. **22**(9) (1916), p. 159-184
50. Raynaud, Michèle: Géométrie Algébrique et Géométrie Analytique. Séminaire de Géométrie Algébrique du Bois Marie 1960-61, Exposé XII. <https://arxiv.org/pdf/math/0206203.pdf> Call June 2020
51. Scharlau, Winfried; Opolka Hans: Von Fermat bis Minkowski. Eine Vorlesung über Zahlentheorie. Springer, Berlin (1980)
52. Schultz, Dan: Notes on Modular Forms. <https://faculty.math.illinois.edu/~%7eschult25/ModFormNotes.pdf> Call 27.10.2020
53. Serre, Jean-Pierre: Géométrie Algébrique et Géométrie Analytique. Annales de l'Institut Fourier. IV (1955-56), p. 1-42
54. Serre, Jean-Pierre: II Modular Forms. (1957/58). In: [9].
55. Serre, Jean-Pierre: Représentations linéaires des groupes finis. Herman, Paris (1967)
56. Serre, Jean-Pierre: A Course in Arithmetic. Springer, New York (1973)
57. Serre, Jean-Pierre: Valeurs propre des endomorphismes de Frobenius [d'après P. Deligne]. Exposé 446, Séminaire N. Bourbaki (1974)
58. Schmidt, Eric M.: Table of n , $a(n)$ for $n = 1..194$ <https://oeis.org/A001379/b001379.txt> Call 6.10.2026
59. Shimura, Goto: Introduction to the Arithmetic Theory of Automorphic Functions. Iwanami Shoten and Princeton University Press, o. O. (1971)
60. Silverman, Joseph: The Arithmetic of Elliptic Curves. Springer, New York (1986)
61. Silverman, Joseph: Advanced Topics in the Arithmetic of Elliptic Curves. Springer, New York (1994)
62. Silverman, Joseph; Tate, John: Rational points on Elliptic Curves. Springer, New York (1992)
63. stackexchange. <https://math.stackexchange.com/>
64. Szpiro, Lucien: Séminaire sur les pinceaux arithmétiques: La conjecture de Mordell. Astérisque 127. Société Mathématiques de France, Paris (1985)
65. Tatitscheff, Valdo: A short introduction to Monstrous Moonshine. Preprint (2019) https://www.researchgate.net/publication/331008545_A_short_introduction_to_Monstrous_Moonshine
66. Thompson, J. G.: Some Numerology between the Fischer-Griess monster and the Elliptic Modular Function. Bull. Lond. Math. Soc. **11** (1979), 352-353
67. van der Geer, Gerard: Siegel Modular Forms. arXiv:math/0605346 [math.AG]
68. Wehler, Joachim: Complex Analysis. Lecture Notes (2019) http://www.mathematik.uni-muenchen.de/~%7ewehler/20181218_Funktionentheorie_Script.pdf
69. Wehler, Joachim: Riemann Surfaces. Lecture Notes (2020) http://www.mathematik.uni-muenchen.de/~%7ewehler/20190530_RiemannSurfacesScript.pdf
70. Zagier, Don: Zetafunktionen und quadratische Körper. Eine Einführung in die höhere Zahlentheorie. Springer, Berlin (1981)
71. Zagier, Don: <https://www.youtube.com/watch?v=zKt5L0ggZ3o> (2015)
72. Zagier, Don: Elliptic Modular Forms and Their Applications. https://people.mpim-bonn.mpg.de/zagier/files/doi/10.1007/978-3-540-74119-0_1/fulltext.pdf Call 20.9.2026

Index

- K -valued point, 148, 149
- L -series, 182
- ϑ -series, 228, 232
- ζ -function
 - Euler product, 180
 - of elliptic curve, 181
 - of number field, 267
- Abel's theorem, 26, 167
- algebraic variety, 145
- automorphic form, 92
- class group
 - and elliptic curves, 270
 - and tori, 268
- class number
 - definition, 265
 - estimate, 271
 - example, 266
 - finiteness, 266
 - general class number formula, 267
- compact Riemann surface
 - projective embedding, 142
- complex analysis - algebraic geometry
 - dictionary, 143
- composition series, 283
- congruence subgroup
 - definition, 39
 - fundamental domain, 53
 - modular curve, 79
 - order relation, 121, 122
- Conway-Norton conjecture, 288
- cubic curve
 - singularity, 159, 162
- cusp, 39, 41
- cusp
 - width, 40
- cusp form
 - codimension, 127
 - definition, 92
 - differential form, 112
 - dimension formula, 114
 - exact sequence, 118
- Dedekind ring, 237
- degree
 - plane curve, 148
- discrete subgroup, 10, 11
- discriminant form
 - definition, 103
 - eigenform of Hecke operators, 202
 - Fourier coefficients, 104
- divisor
 - degree, 151
 - exact sequence, 262
- divisor lemma, 43
- doubly periodic, 11
- Eisenstein series
 - definition, 98
 - normalization, 100
 - relation to lattice constants, 102
- elementary divisor theorem, 253
- elliptic curve
 - ζ -function, 181
 - L -series, 182
 - Abel's theorem, 167
 - as plane curve, 154
 - as torus, 160
 - definition, 152
 - divisor, 153
 - divisor class group, 166
 - group structure, 166–169, 171
 - L -series, 182

- reduction mod p , 176, 177
- Weierstrass equation, 158
- elliptic function
 - definition, 11
 - holomorphic, 12
 - pole, 13
 - poles, 29
 - poles and zeros, 13, 25
 - residue theorem, 13
- elliptic functions
 - field, 22
- elliptic point, 39, 54
- equivalence criterion
 - Hausdorff quotient, 64
- factor of automorphy, 88, 90
- Falting's theorem, 169
- Fermat conjecture, 280
- field of definition, 148
- finite simple groups
 - classification, 284
- four squares theorem, 231
- Fourier coefficients
 - definition, 89
 - of Hecke transforms, 198
- Fourier expansion, 89
- fractional ideal, 263
- Frey-Helleguarch curve, 280
- fundamental domain
 - construction, 49
 - definition, 36
 - volume, 206
- GAGA, 149
- Galois group
 - definition, 236
 - lifting to $\mathbb{C}$, 270
- group action, 36
- group representation
 - complete reducibility, 283, 284
 - on integer matrices, 243
- Hasse estimate, 178
- Hauptmodul, 286, 287
- Hecke algebra
 - commutativity, 205
 - definition, 204
- Hecke congruence subgroup
 - changing level, 218
 - index, 45
- Hecke operator
 - transforming Fourier coefficients, 198
- Hecke operators
 - definition, 198
- eigenforms, 202, 218
- example, 202
- for congruence subgroups, 218
- Fourier coefficients, 198
- self-adjointness, 215
- transforming Fourier coefficients, 201
- holomorphic at ∞ , 89
- holomorphic map
 - defined by a line bundle, 137
 - embedding into projective space, 138
 - map to $\mathbb{P}^n$, 137
- hyperbolic volume form
 - definition, 205
 - invariance, 205
- imaginary quadratic number field
 - definition, 236
 - relation with complex multiplication, 240
- integer matrices
 - complete set of representatives, 190, 210, 244
- integers as sum of squares, 228
- invariant differential forms, 86
- isotropy group, 36
- Jacobi torus, 278
- Jordan-Hölder theorem, 283
- Kronecker symbol, 81
- L-series
 - elliptic curve, 182
 - is holomorphic, 278
 - of a cusp form, 278
- lattice
 - basis, 32
 - definition, 11
 - lattice constant, 14
 - normalization, 34
 - period lattice, 11
 - similarity, 32
- lattice function
 - definition, 189
 - homogenous, 188
 - relation to modular forms, 188
- lattice map
 - algebra, 196
- Legendre symbol, 81
- line bundle
 - base-point, 136
 - base-point free, 136
 - globally generated, 136, 140
 - separating points, 139
 - separating tangent lines, 139

- very ample, 140
 - very ampleness criterion, 141
- local ring, 145
- McKay-Thompson conjecture, 287
- meromorphic at ∞ , 89
- meromorphic function
 - doubly periodic, 11
 - period, 9, 10
- modular curve
 - construction, 74
 - definition, 79
- modular form
 - $G_{2,N}$, 230
 - algebra of modular forms, 120
 - as differential form on the modular curve, 108
 - definition, 92
 - dimension formula, 114, 126
 - Fourier series, 89, 91
 - newform, 221, 278
 - odd weight, 94
 - oldform, 221
 - q-expansion, 89
 - section of a line bundle, 124
 - weight formula, 115
- modular function, 93
- modular group
 - action of isotropy groups, 66
 - action on $\mathbb{H}$, 38
 - definition, 39
 - fundamental domain, 48
 - orbit space, 68
 - presentation, 53
 - properly discontinuous action, 55
 - strongly properly discontinuous group action, 59
- modular invariant
 - and modular curve, 106
 - definition, 106
 - elliptic point, 116
 - field of automorphic functions, 117
 - singular value, 239, 257, 260, 273
- modular polynomial
 - coefficients, 247
 - definition, 245
 - diagonal evaluation, 251
 - example, 246
 - symmetry, 254
- modularity theorem, 277, 278, 281
- moduli space, 68
- monstrous moonshine
 - Borcherd's theorem, 288
 - definition, 285
 - type, 286
- Mordell's theorem, 169
- newform
 - definition, 221
 - example, 222
- number field
 - algebraic integer, 236
 - class group, 268
 - class number, 265
 - definition, 236
 - discriminant, 236
 - fractional ideal, 263
 - ideal class group, 264
 - imaginary quadratic, 236
 - norm, 236
 - order, 236
 - trace, 236
- number ring
 - definition, 236
- oldform
 - definition, 221
 - example, 222
- orbit space
 - complex charts, 68
 - group action, 36
 - Hausdorff topology, 65
- period group, 10
- period parallelogram, 11
- Petersson scalar product
 - definition, 208
 - properties, 208, 211
- plane projective algebraic curve, 145
- positively oriented basis, 32
- principal congruence subgroup
 - definition, 43
 - exact sequence, 44
- projective algebraic curve, 145
- projective algebraic variety, 145
- projective space
 - definition, 129
 - sections of $\mathcal{O}(1)$, 136
 - standard atlas, 130
 - twisted line bundle, 136
- proper group action, 62
- Quadratic reciprocity, 275
- Ramanujan τ -function
 - definition, 224
 - growth estimate, 228
 - product and recursion formulas, 227

- reduced ideal, 265
- regular function, 145
- regular local ring, 145
- Riemann ζ -function, 98
- simple group, 282
- smooth point, 145
- strongly properly discontinuous group action, 59
- structure sheaf, 146
- sublattices, 192
- theorem
 - Abel's, 167
- torus
 - complex multiplication, 239, 240
 - embedding as elliptic curve, 131, 135
 - endomorphism, 238
 - endomorphism ring, 239
 - equivalence relation of lattices, 34
 - holomorphic map, 32
 - ring of endomorphisms, 239
 - twisted line bundle, 136
- Weierstrass $\wp$ -function
 - definition, 15
 - differential equation, 21
 - lattice half-periods, 18
 - Laurent expansion, 18, 20
 - period lattice, 19
 - pole set, 15
- Weierstrass equation, 157
- Weierstrass polynomial
 - definition, 157
 - global minimal, 173, 174
- weight- k operator, 87
- Zariski topology, 145