FACULTAD DE MATEMÁTICAS
DEPARTAMENTOS DE ÁLGEBRA Y ANÁLISIS

SPACES OF MODULAR FORMS, MODULAR CURVES AND DIMENSIONS

Jesús López Sánchez

Supervised by:

Sara Arias de Reyna Domínguez
Juan Arias de Reyna Martínez

FACULTAD DE MATEMÁTICAS
DEPARTAMENTOS DE ÁLGEBRA Y ANÁLISIS

Spaces of modular forms, modular curves and dimensions

Jesús López Sánchez

Memory presented as part of the requirements to obtain a Master's degree in Mathematics from the University of Seville.

Supervisors: Sara Arias de Reyna Domínguez
Juan Arias de Reyna Martínez

June 20, 2019

Contents

Abstract

The Modularity Theorem states that all rational elliptic curve arise from modular forms. In 1995, Andrew Wiles proved a special case of this theorem (then known as the Taniyama–Shimura conjecture) for semistable elliptic curves, completing the proof of Fermat's Last Theorem after some 350 years. Later, Christophe Breuil, Brian Conrad, Fred Diamond and Richard Taylor extended Wiles's techniques to prove completely the Modularity Theorem. In this work we explain a complex analytic version of this notable theorem.

Translation to Spanish

El Teorema de Modularidad afirma que todas las curvas elípticas racionales surgen de formas modulares. En 1995, Andrew Wiles probó un caso especial de este teorema (entonces conocido como la conjetura de Taniyama–Shimura) para curvas elípticas semiestables, completando así la prueba del Último Teorema de Fermat después de unos 350 años. Más tarde, Christophe Breuil, Brian Conrad, Fred Diamond and Richard Taylor extendieron las técnicas de Wiles para probar completamente el Teorema de Modularidad. En este trabajo nosotros explicamos una versión analítica compleja de este notable teorema.

Symbols

$\mathbb{Z}$	set of integer numbers
$\mathbb{Z}^+$	set of positive integer numbers
$\mathbb{Q}$	set of rational numbers
$\mathbb{R}$	set of real numbers
$\mathbb{C}$	set of complex numbers
$\mathbb{C}^*$	set of nonzero complex numbers
$\widehat{\mathbb{C}}$	Riemann sphere
$\mathbb{H}$	upper half plane
$\mathbb{D}$	unit disc
$\operatorname{Re} z$	real part of a complex number z
$\operatorname{Im} z$	imaginary part of a complex number z
$\lvert z\rvert$	absolute value of a complex number z
$D(z,\varepsilon)$	$\{w \in \mathbb{C} \mid \lvert w-z\rvert < \varepsilon\}$
$\dot{D}(z,\varepsilon)$	$\{w \in \mathbb{C} \mid 0 < \lvert w-z\rvert < \varepsilon\}$
$\overline{D}(z,\varepsilon)$	$\{w \in \mathbb{C} \mid \lvert w-z\rvert \leq \varepsilon\}$
$\emptyset$	empty set
$int(A)$	interior of a subset A
A'	derived of a subset A
$\overline{A}$	clousure of a subset A
$Y(\Gamma)$	noncompact modular curve $\Gamma\backslash\mathbb{H}$
$X(\Gamma)$	compact modular curve $\Gamma\backslash\mathbb{H}^*$ ($\mathbb{H}^* = \mathbb{H} \cup \mathbb{Q} \cup \{\infty\}$)
$\mathcal{A}_k(\Gamma)$	set of automorphic forms of weight k with respect to Γ
$\mathcal{M}_k(\Gamma)$	set of modular forms of weight k with respect to Γ
$\mathcal{S}_k(\Gamma)$	set of cusp forms of weight k with respect to Γ
$\mathcal{O}_X(U)$	set of holomorphic functions on U
$\mathcal{M}_X(U)$	set of meromorphic functions on U
$\mathcal{M}^{(n)}(X)$	set of meromorphic differentials on X of degree n
$\operatorname{Div}(X)$	set of divisors on X

For a commutative ring R with unity we write

$$\begin{aligned}
\mathrm{M}_2(R) &= \text{set of square matrices of degree 2 with coefficients in } R,\\
\mathrm{GL}_2(R) &= \{\alpha \in \mathrm{M}_2(R) \mid \det(\alpha) \in R^*\},\\
\mathrm{SL}_2(R) &= \{\alpha \in \mathrm{M}_2(R) \mid \det(\alpha) = 1\},
\end{aligned}$$

where R^* is the group of invertible elements in R.

Introduction

> *Cubum autem in duos cubos, aut quadratoquadratum in duos quadratoquadratos, et generaliter nullam in infinitum ultra quadratum potestatum in duos ejusdem nominis fas est dividere: cujes rei demonstrationem mirabilem sane detexi. Hanc marginis exiguitas non caperet.*
>
> Pierre de Fermat, 1637

The French mathematician Pierre de Fermat (17 August 1601, Beaumont-de-Lomagne, France – 12 January 1665, Castres, France) wrote this note in the margin of his copy of Diophantus's Arithmetica stating that the equation

$$x^n + y^n = z^n, \quad n \in \mathbb{Z}^+, \tag{1}$$

has no solutions in positive integers if n is greater than 2. He also claimed to have a marvelous proof of this statement, but this was never published (the previous note was published in 1670 by his older son after his death). This statement became known over time as Fermat's Last Theorem (FLT), since it was the last of Fermat's asserted theorems to remain unproved.

Figure 1: Pierre de Fermat

Let $(\mathrm{FLT})_n$ denote the following statement:

The equation (1) has no solutions in positive integers.

It is easy to prove that $(\mathrm{FLT})_n$ implies $(\mathrm{FLT})_{kn}$ for any positive integer k, so it suffices to prove

$$(\mathrm{FLT})_4 \quad \text{and} \quad (\mathrm{FLT})_\ell \text{ for any prime } \ell > 2,$$

to conclude that Fermat's Last Theorem is true. Fermat proved $(\mathrm{FLT})_4$ by showing that the equation

$$x^4 + y^4 = z^2$$

has no solutions in positive integers. Its proof can be consulted in [vdP96, p.3]. As a consequence, (FLT) reduced to the following question:

Is $(\mathrm{FLT})_\ell$ true for any prime $\ell > 2$?

History of Fermat's Last Theorem

During the next three centuries (18-20th), giving an answer to this question would become one of the most difficult mathematical problems to resolve. The first complete proof of the case $(\mathrm{FLT})_3$ was given by Gauss [IR90, p.284] (Euler gave a proof of $(\mathrm{FLT})_3$ in 1753, but this contains an alleged error). Gauss's proof leads to a strategy that succeeds for other values of ℓ as well. Peter Dirichlet and Adrien Legendre proved independently $(\mathrm{FLT})_5$ in 1825, and Gabriel Lamé settled $(\mathrm{FLT})_7$ in 1839.

The work of Sophie Germain

In 1823, Sophie Germain proved that if $q = 2\ell + 1$ is also prime number, then the equation

$$x^\ell + y^\ell = z^\ell \tag{2}$$

has no solutions in positive integers with $\ell \nmid xyz$. Germain's theorem was the first really general result on Fermat's Last Theorem, since the previous results only considered Fermat's equation for a specific exponent.

At this point, the study of Fermat's Last Theorem was divided into two cases,

- the first case involved showing that the equation (2) has no solutions in positive integers with $\ell \nmid xyz$, and

- the second case involved showing that the equation (2) has no solutions in positive integers with $\ell \mid xyz$.

The proposed proof of Gabriel Lamé

On 1 March, 1847, Lamé informed the Parisian Académie des Sciences that he had resolved the general case. The basic idea of his proof consisted in working with cyclotomic integers,

$$\mathbb{Z}[\zeta_\ell] = \{a_0 + a_1\zeta_\ell + \cdots + a_{\ell-2}\zeta_\ell^{\ell-2} \mid a_i \in \mathbb{Z}\}, \quad \text{where } \zeta_\ell = e^{2\pi i/\ell}.$$

Using these numbers, we can write

$$x^\ell + y^\ell = (x+y)(x+\zeta_\ell y)(x+\zeta_\ell^2 y)\cdots(x+\zeta_\ell^{\ell-1}y),$$

and therefore, Fermat's equation assumes the form

$$(x+y)(x+\zeta_\ell y)(x+\zeta_\ell^2 y)\cdots(x+\zeta_\ell^{\ell-1}y) = z^n.$$

As this product of numbers without common factors (assume $\gcd(x,y)=1$) is an ℓ-th power, Lamé thought that each number would be an ℓ-th power (he implicitly assumed that unique factorization into products of primes also held for cyclomotic integers, but $\mathbb{Z}[\zeta_\ell]$ is a UFD if and only if $\ell < 23$) and proceeded with an argument showing necessarily one of x or y to be zero.

The work of Ernst Kummer

The mathematician Ernst Kummer formalized this argument of Lamé. He began studying the ideal class group of $\mathbb{Q}(\zeta_\ell)$, which is a finite group that measures how far $\mathbb{Z}[\zeta_\ell]$ is from being a unique factorization domain (for example, $\mathbb{Z}[\zeta_\ell]$ is a UFD if and only if $h_\ell = 1$, where h_ℓ denotes the order of the ideal class group of $\mathbb{Q}(\zeta_\ell)$). Between 1847 and 1853, he published some masterful papers. In these papers he defined regular prime numbers (a prime number ℓ is called regular if $\ell \nmid h_\ell$; otherwise, ℓ is called irregular) and proved the following theorem:

Theorem 0.1 (Kummer). *Let ℓ be a odd prime number. Then*

1. (FLT)$_\ell$ *is true if ℓ is regular,*
2. *ℓ is regular if and only if ℓ does not divide the numerator of B_i for any even $2 \le i \le (\ell-3)/2$, where B_i are the Bernoulli numbers,*

$$\frac{z}{e^z - 1} = \sum_{n=0}^{\infty} (B_n/n!)z^n.$$

The unique inconvenient of this result is that there exist infinitely many irregular primes (and therefore, Fermat's Last Theorem was not proved).

Computational studies

In 1954, Harry Vandiver used a SWAC computer to prove

$$(\mathrm{FLT})_\ell \quad \text{for all primes } \ell \text{ up to } 2521.$$

By 1978, Samuel Wagstaff had extended this to all primes ℓ less than 125.000. Before Wiles, $(\mathrm{FLT})_\ell$ had been proved for all primes ℓ less than four million.

Taniyama–Shimura conjecture

On the other hand, in the middle of the 20th century, the Japanese mathematicians Yutaka Taniyama and Goro Shimura observed a possible relation between two apparently distinct branches of mathematics, elliptic curves and modular forms. This possible relation was formalized later by Shimura, giving rise to what we know now as Modularity Theorem (then known as Taniyama-Shimura conjecture):

All rational elliptic curves arise from modular forms.

Ribet's theorem for Frey curves

In 1985, Gerhard Frey observed a link between Fermat's equation and the modularity theorem (then still a conjecture). If there exist positive integers a, b, c such that

$$a^\ell + b^\ell = c^\ell,$$

then the semistable elliptic curve (it has square-free conductor)

$$y^2 = x(x - a^\ell)(x + b^\ell) \qquad \text{[Frey curve]}$$

would have such unusual properties that it was unlikely to be modular. A year later, Kenneth Ribet proved that this curve is definitely not modular. His strategy consisted in showing that if the Frey curve is associated to a modular form, then it must be associated to one of weight 2 and level 2. No cuspidal eigenforms of this kind exist, giving the desired contradiction.

Wiles's proof of Fermat's Last Theorem

The British mathematician Andrew Wiles published in 1995 a proof of the Modularity Theorem (then still known as the Taniyama–Shimura conjecture) for semistable elliptic curves [Wil95, TW95]. Due to the previous works of Frey and Ribet, the Modularity Theorem for semistable elliptic curves implied Fermat's Last Theorem, since if there exist positive integers a, b, c such that $a^\ell + b^\ell = c^\ell$, then the semistable elliptic curve

$$y^2 = x(x - a^\ell)(x + b^\ell)$$

would not be modular, contradicting the Modularity Theorem for semistable elliptic curves proved by Andrew Wiles.

Note. A nice historical overview of Fermat's Last Theorem, together with notes and remarks is given in [vdP96].

The Modularity Theorem

An elliptic curve over $\mathbb{Q}$ is a nonsingular cubic equation of the form

$$E:\, y^2 = 4x^3 - c_2x - c_3, \quad c_2, c_3 \in \mathbb{Q} \quad (c_2^3 - 27c_3^2 \neq 0).$$

A modular form is simply a holomorphic function on the complex upper half plane that satisfies "certain" transformation and holomorphy conditions. The original version of the Modularity Theorem that was proved by Andrew Wiles, Christophe Breuil, Brian Conrad, Fred Diamond and Richard Taylor is the following:

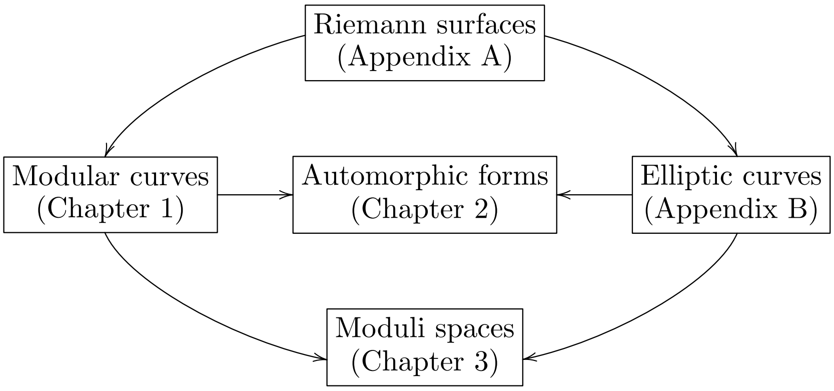
> Each Galois representation $\rho_{E,\ell}$ associated to an elliptic curve E over $\mathbb{Q}$ arises from a Galois representation $\rho_{f,\ell}$ associated to a modular form f,
>
> $$\rho_{E,\ell} \sim \rho_{f,\ell}$$

In this work we explain an (equivalent) version the Modularity Theorem that relates rational elliptic curves and modular curves as Riemann surfaces.

Theorem 0.2 (Modularity Theorem, Complex analytic version). *Let E be a complex elliptic curve with $j(E) \in \mathbb{Q}$. Then for some positive integer N there exists a surjective morphism of Riemann surfaces from the modular curve $X_0(N)$ to the elliptic curve E,*

$$X_0(N) \longrightarrow E.$$

To understand this affordable version of the theorem we have to introduce:



Outline of the work

In Chapter 1 we introduce the modular group, its congruence subgroups and the modular curves, which are quotient spaces of the upper half plane by the action of a congruence subgroup of the modular group. Furthermore, we show these curves are Riemann surfaces that can be compactified.

In Chapter 2 we introduce the automorphic, modular and cusp forms. They are (meromorphic) holomorphic functions on the upper half plane that satisfy certain transformation and (meromorphy) holomorphy conditions. We comment on the dimension formulas of the vector spaces of modular and cusp forms, and we conclude with two interesting applications.

In Chapter 3 we introduce moduli spaces (isomorphism classes of complex elliptic curves enhanced by associated torsion data) for some modular curves and we explain the complex analytic version of the Modularity Theorem we have mentioned above.

In Appendix A we recall all the theory of Riemann surfaces that we need in this work: Holomorphic maps, meromorphic differentials, divisors and the Riemann-Roch Theorem.

In Appendix B we explain the principal results over compact Riemann surfaces of genus equal to 1. These Riemann surfaces are called complex elliptic curves for reasons to be explained in this chapter.

Comments on Bibliography

Most of the content of this work has been extracted from the book [DS05]. This book explains (equivalent) distinct versions of the Modularity Theorem. Other books about this topic we have used are [Apo90, Miy06, Ser73, Shi71].

A good reference book about Riemann surfaces we have used in Appendix is [Mir95]. Also, we have used this magnificent book [Sil09] about elliptic curves in Appendix B.

The results of complex analysis we utilize in this work are really basic. The reader can consult them in any of these two books [Ahl78, SS03].

Chapter 1

Modular curves

In this first chapter we introduce the modular curves which are quotient spaces of the upper half plane by the action of a congruence subgroup Γ of $\mathrm{SL}_2(\mathbb{Z})$. We show these curves are Riemann surfaces (see Appendix A) that can be compactified. The theory of compact Riemann surfaces allows us to calculate the topological genus of these compactified curves.

1.1 The modular group

The modular group is

$$\mathrm{SL}_2(\mathbb{Z}) = \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathrm{M}_2(\mathbb{Z}) \,|\, ad - bc = 1 \right\}.$$

Lemma 1.1. *The modular group* $\mathrm{SL}_2(\mathbb{Z})$ *is generated by the two matrices*

$$T = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \quad \textit{and} \quad S = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}.$$

Proof: Let Γ be the subgroup of $\mathrm{SL}_2(\mathbb{Z})$ generated by these matrices and $\alpha = \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right] \in \mathrm{SL}_2(\mathbb{Z})$. Below, we describe an algorithm to compute $\gamma \in \Gamma$ such that $\alpha\gamma \in \Gamma$ (and therefore we will be able to conclude that $\alpha \in \Gamma$).

First observe that we can suppose without loss of generality that $c \neq 0$, since otherwise $\alpha = \left[\begin{smallmatrix} \pm 1 & b \\ 0 & \pm 1 \end{smallmatrix}\right] \in \Gamma$. Indeed,

$$T^n = \begin{bmatrix} 1 & n \\ 0 & 1 \end{bmatrix} \quad \text{and} \quad S^2T^n = -T^n = -\begin{bmatrix} 1 & n \\ 0 & 1 \end{bmatrix}, \quad \forall n \in \mathbb{Z}.$$

The identity

$$\alpha T^n = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 1 & n \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} a' & b' \\ c & nc + d \end{bmatrix}, \quad n \in \mathbb{Z},$$

shows that there exists a matrix $\gamma_1 \in \Gamma$ such that $\alpha\gamma_1$ has bottom row

$$(c', d') = (c, nc + d), \quad \text{with } |d'| \leq |c'|/2.$$

On the other hand, the identity

$$\alpha\gamma_1 S = \begin{bmatrix} a' & b' \\ c' & d' \end{bmatrix} \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} b' & -a' \\ d' & -c' \end{bmatrix}, \quad n \in \mathbb{Z},$$

shows that this process can be iterated (a finite number of times) to find a matrix $\gamma \in \Gamma$ such that $\alpha\gamma$ has bottom row $(0, \pm 1)$, and therefore $\alpha\gamma \in \Gamma$. □

Each element of the modular group induces naturally an automorphism of the Riemann sphere, the fractional linear transformation

$$\alpha(z) = \frac{az + b}{cz + d}, \quad \forall\, z \in \widehat{\mathbb{C}}, \quad \text{for } \alpha = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathrm{SL}_2(\mathbb{Z}).$$

The identity matrix I and its negative $-I$ both induce the identity map. It is not difficult to prove that two matrices $\alpha, \alpha' \in \mathrm{SL}_2(\mathbb{Z})$ induce the same transformation if and only if $\alpha' = \pm\alpha$. Furthermore, note that

$$(\alpha\alpha')(\tau) = \alpha(\alpha'(\tau)), \quad \forall\, z \in \widehat{\mathbb{C}}, \quad \forall\, \alpha, \alpha' \in \mathrm{SL}_2(\mathbb{Z}).$$

Therefore, the modular group acts on the Riemann sphere. The subgroup of transformations defined by the modular group is generated by the two maps induced by the two matrix generators,

$$T(\tau) = \tau + 1 \quad \text{and} \quad S(\tau) = -1/\tau.$$

The upper half plane is

$$\mathbb{H} = \{\tau \in \mathbb{C} \,|\, \mathrm{Im}\,(\tau) > 0\}.$$

The formula

$$\mathrm{Im}\,(\alpha(\tau)) = \frac{\mathrm{Im}\,(\tau)}{|c\tau + d|^2}, \quad \forall\, \tau \in \mathbb{H}, \quad \forall\, \alpha = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathrm{SL}_2(\mathbb{Z}) \qquad (1.1)$$

shows that each element of the modular group maps the upper half plane back to itself. Hence the modular group acts also on the upper half plane.

Definition 1.2. *A matrix* $\alpha = \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right] \in \mathrm{SL}_2(\mathbb{Z})$*,* $\alpha \neq \pm I$*, is called elliptic, parabolic, or hyperbolic if the absolute value of its trace,* $|a+d|$*, is less than* 2*, equal to* 2*, or greater than* 2*, respectively.*

To see the geometrical meaning of this classification, we have to study the fixed points of the induced transformation on the Riemann sphere.

If $c = 0$, then α is of the form

$$\pm \begin{bmatrix} 1 & m \\ 0 & 1 \end{bmatrix}, \quad \text{for some } m \in \mathbb{Z} \setminus \{0\},$$

so α is parabolic and its unique fixed point is $\infty \in \widehat{\mathbb{C}}$. If $c \neq 0$, then $\infty \in \widehat{\mathbb{C}}$ can not be a fixed point of α, since $\alpha(\infty) = a/c$. In this case, observe that the fixed points of α satisfy the quadratic equation $cz^2 + (d - a)z - b = 0$. As the discriminant of this equation is

$$(d - a)^2 + 4cb = (d + a)^2 - 4,$$

the fixed points of α are two conjugate complex numbers, a real number, or two distinct real numbers if α is elliptic, parabolic or hyperbolic, respectively.

Theorem 1.3. *A matrix $\alpha = \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right] \in \mathrm{SL}_2(\mathbb{Z})$, $\alpha \neq \pm I$, is characterized by its fixed points in $\widehat{\mathbb{C}}$ as follows:*

- ▷ *α is elliptic if and only if α has two fixed points τ and $\overline{\tau}$, with $\tau \in \mathbb{H}$.*
- ▷ *α is parabolic if and only if α has only one fixed point in $\mathbb{R} \cup \{\infty\}$.*
- ▷ *α is hyperbolic if and only if α has two distinct fixed points in $\mathbb{R}$.*

1.2 Congruence subgroups

The principal congruence subgroups are

$$\Gamma(N) = \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathrm{SL}_2(\mathbb{Z}) \mid \begin{bmatrix} a & b \\ c & d \end{bmatrix} \equiv \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \pmod N \right\}, \quad N \in \mathbb{Z}^+.$$

The matrix congruence is interpreted by entries, i.e.,

$$\begin{aligned} a \equiv 1 \pmod N, \quad b \equiv 0 \pmod N, \\ c \equiv 0 \pmod N, \quad d \equiv 1 \pmod N. \end{aligned}$$

Let us consider the canonical homomorphism from $\mathrm{SL}_2(\mathbb{Z})$ to $\mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})$,

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathrm{SL}_2(\mathbb{Z}) \mapsto \begin{bmatrix} \bar{a} & \bar{b} \\ \bar{c} & \bar{d} \end{bmatrix} \in \mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z}).$$

By definition, its kernel is $\Gamma(N)$. So $\Gamma(N)$ is a normal subgroup of $\mathrm{SL}_2(\mathbb{Z})$.

Let $\begin{bmatrix} \bar{a} & \bar{b} \\ \bar{c} & \bar{d} \end{bmatrix}$ be a matrix of $\mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})$. As

$$ad - bc \equiv 1 \pmod N,$$

we deduce that $\gcd(c,d,N)=1$. Using the Chinese remainder theorem it is possible to prove that there exist integers c_1 and d_1 such that

$$c_1 \equiv c \pmod{N}, \quad d_1 \equiv d \pmod{N} \quad \text{and} \quad \gcd(c_1,d_1)=1.$$

Let $g=\gcd(c,d)$. If $c \neq 0$, consider the system

$$\begin{cases} t \equiv 1 \pmod{p}, & \text{with } p \text{ prime, } p|g, \\ t \equiv 0 \pmod{p}, & \text{with } p \text{ prime, } p|c,\ p \nmid g, \end{cases}$$

and define $c_1 = c$ and $d_1 = d + tN$. If $c = 0$, then necessarily $d \neq 0$ (unless $N=1$, but this case is trivial). Therefore, consider the system

$$\begin{cases} s \equiv 1 \pmod{p}, & \text{with } p \text{ prime, } p|g, \\ s \equiv 0 \pmod{p}, & \text{with } p \text{ prime, } p|d,\ p \nmid g, \end{cases}$$

and define $c_1 = c + sN$ and $d = d_1$. These two cases prove the statement.

Let k be the unique integer such that

$$ad_1 - bc_1 = 1 + kN.$$

As $\gcd(c_1,d_1)=1$, there exist integers a_1 and b_1 such that $a_1d_1 - b_1c_1 = -k$. Letting

$$a_2 = a + a_1N, \quad b_2 = b + b_1N,$$
$$c_2 = c_1, \quad d_2 = d_1,$$

we obtain that

$$\begin{bmatrix} a_2 & b_2 \\ c_2 & d_2 \end{bmatrix} \in \mathrm{SL}_2(\mathbb{Z}) \quad \text{and} \quad \begin{bmatrix} a_2 & b_2 \\ c_2 & d_2 \end{bmatrix} \equiv \begin{bmatrix} a & b \\ c & d \end{bmatrix} \pmod{N}.$$

Thus we can conclude that this homomorphism is surjective. Therefore, it induces an isomorphism

$$\mathrm{SL}_2(\mathbb{Z})/\Gamma(N) \xrightarrow{\sim} \mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z}).$$

As a consequence, note that $\Gamma(N)$ has finite index in $\mathrm{SL}_2(\mathbb{Z})$. In fact, it is possible to compute that

$$[\mathrm{SL}_2(\mathbb{Z}) : \Gamma(N)] = N^3 \prod_{p|N} \left(1 - \frac{1}{p^2}\right),$$

where the product is taken over all prime divisors of N [Miy06, p.105].

Definition 1.4. *A subgroup Γ of $\mathrm{SL}_2(\mathbb{Z})$ is a congruence subgroup if there exists $N \in \mathbb{Z}^+$ such that $\Gamma(N) \subset \Gamma$, in which case Γ is called congruence subgroup of level N.*

Remarks 1.5.

▷ Each congruence subgroup Γ has finite index in $\mathrm{SL}_2(\mathbb{Z})$, since

$$\Gamma(N) \subset \Gamma, \quad \text{for some } N \in \mathbb{Z}^+.$$

▷ Each congruence subgroup Γ contains a translation matrix of the form

$$\begin{bmatrix} 1 & h \\ 0 & 1 \end{bmatrix} : \tau \to \tau + h, \quad h \in \mathbb{Z}^+.$$

Let

$$h_\Gamma = \min\{h \in \mathbb{Z}^+ \mid \begin{bmatrix} 1 & h \\ 0 & 1 \end{bmatrix} \in \Gamma\}$$

and let us fix a positive integer N such that $\Gamma(N) \subset \Gamma$. As

$$\begin{bmatrix} 1 & N \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & h_\Gamma \\ 0 & 1 \end{bmatrix}^q \begin{bmatrix} 1 & r \\ 0 & 1 \end{bmatrix}, \quad \text{whenever } N = qh_\Gamma + r,\ q, r \in \mathbb{Z},$$

we deduce that h_Γ necessarily divides N.

The reason why these subgroups are called congruence subgroups is justified in the following lemma which describes the congruence subgroups of level N.

Lemma 1.6. *Let Γ be a subgroup of $\mathrm{SL}_2(\mathbb{Z})$ and $N \in \mathbb{Z}^+$. The following conditions are equivalent:*

1. *Γ is a congruence subgroup of level N.*
2. *There exist $\gamma_1, \ldots, \gamma_d \in \mathrm{SL}_2(\mathbb{Z})$ such that $\Gamma = \bigcup_{j=1}^d \Gamma(N)\gamma_j$, i.e.,*

$$\Gamma = \bigcup_{j=1}^{d} \{\gamma \in \mathrm{SL}_2(\mathbb{Z}) \mid \gamma \equiv \gamma_j \pmod N\}.$$

Proof: If $\Gamma(N) \subset \Gamma$, then $\Gamma(N)$ has finite index in Γ. Reciprocally, if there exist $\gamma_1, \ldots, \gamma_d \in \mathrm{SL}_2(\mathbb{Z})$ such that $\Gamma = \bigcup_{j=1}^d \Gamma(N)\gamma_j$, then

$$I \equiv \gamma_j \pmod{N}, \quad \text{for some } j = 1, \ldots, d,$$

where I is the identity matrix. So $\Gamma(N) = \Gamma(N)\gamma_j \subset \Gamma$.

□

An immediate consequence of the fact that $\Gamma(N)$ is normal in $\mathrm{SL}_2(\mathbb{Z})$ is that any conjugated subgroup of a congruence subgroup of level N is also a congruence subgroup of level N.

Lemma 1.7. *Let Γ be a congruence subgroup of $\mathrm{SL}_2(\mathbb{Z})$, $N \in \mathbb{Z}^+$ and $\alpha \in \mathrm{SL}_2(\mathbb{Z})$. If Γ is a congruence subgroup of level N, then*

$$\alpha^{-1}\Gamma\alpha = \{\alpha^{-1}\gamma\alpha \,|\, \gamma \in \Gamma\}$$

is also a congruence subgroup of level N.

Proof: As $\Gamma(N) \subseteq \Gamma$, we deduce that

$$\Gamma(N) = \alpha^{-1}\Gamma(N)\alpha \subseteq \alpha^{-1}\Gamma\alpha,$$

since $\Gamma(N)$ is a normal subgroup of $\mathrm{SL}_2(\mathbb{Z})$.

□

In addition to the principal congruence subgroups, the most important congruence subgroups are

$$\Gamma_0(N) = \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathrm{SL}_2(\mathbb{Z}) \,|\, \begin{bmatrix} a & b \\ c & d \end{bmatrix} \equiv \begin{bmatrix} * & * \\ 0 & * \end{bmatrix} \pmod N \right\}, \quad N \in \mathbb{Z}^+,$$

and

$$\Gamma_1(N) = \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathrm{SL}_2(\mathbb{Z}) \,|\, \begin{bmatrix} a & b \\ c & d \end{bmatrix} \equiv \begin{bmatrix} 1 & * \\ 0 & 1 \end{bmatrix} \pmod N \right\}, \quad N \in \mathbb{Z}^+,$$

where $*$ means “unspecified”. Note that

$$\Gamma(N) \subset \Gamma_1(N) \subset \Gamma_0(N) \subset \mathrm{SL}_2(\mathbb{Z}), \quad \forall\, N \in \mathbb{Z}^+.$$

In the special case $N = 1$,

$$\Gamma(1) = \Gamma_1(1) = \Gamma_0(1) = \mathrm{SL}_2(\mathbb{Z}).$$

The map $\Gamma_1(N) \to \mathbb{Z}/N\mathbb{Z}$,

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \Gamma_1(N) \mapsto \bar{b} \in \mathbb{Z}/N\mathbb{Z},$$

is surjective, since

$$\begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix} \in \Gamma_1(N), \quad \forall\, k \in \mathbb{Z}.$$

Its kernel is $\Gamma(N)$. Therefore, $\Gamma(N)$ is also a normal subgroup of $\Gamma_1(N)$ and

$$\Gamma_1(N)/\Gamma(N) \cong \mathbb{Z}/N\mathbb{Z}, \quad \text{with} \quad [\Gamma_1(N) : \Gamma(N)] = N.$$

Analogously, the map $\Gamma_0(N) \to (\mathbb{Z}/N\mathbb{Z})^*$,

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \Gamma_0(N) \mapsto \overline{d} \in (\mathbb{Z}/N\mathbb{Z})^*.$$

is surjective as well, since for each $\overline{d} \in (\mathbb{Z}/N\mathbb{Z})^*$ there exist integers e_d and k_d such that

$$e_d d = 1 + k_d N.$$

Therefore

$$\begin{bmatrix} e_d & k_d \\ N & d \end{bmatrix} \in \Gamma_0(N), \quad \forall \overline{d} \in (\mathbb{Z}/N\mathbb{Z})^*.$$

Further, its kernel is $\Gamma_1(N)$. So $\Gamma_1(N)$ is a normal subgroup of $\Gamma_0(N)$ and

$$\Gamma_0(N)/\Gamma_1(N) \cong (\mathbb{Z}/N\mathbb{Z})^*, \quad \text{with} \quad [\Gamma_0(N) : \Gamma_1(N)] = \phi(N),$$

where ϕ is the Euler's totient function from number theory,

$$\phi(N) = |(\mathbb{Z}/N\mathbb{Z})^*| = N \prod_{p|N} \left(1 - \frac{1}{p}\right).$$

Finally, taking into account all the above, we can deduce that

$$[\mathrm{SL}_2(\mathbb{Z}) : \Gamma_0(N)] = N \prod_{p|N} \left(1 + \frac{1}{p}\right),$$

where the product is taken over all prime divisors of N.

1.3 The Riemann surfaces $Y(\Gamma) = \Gamma \backslash \mathbb{H}$

Let Γ be a congruence subgroup of $\mathrm{SL}_2(\mathbb{Z})$ acting on the upper half plane. The modular curve $Y(\Gamma)$ is defined as the quotient space of orbits under Γ,

$$Y(\Gamma) = \Gamma \backslash \mathbb{H} = \{\Gamma\tau \mid \tau \in \mathbb{H}\}.$$

The topology of this space is induced by the natural projection $\pi : \mathbb{H} \to Y(\Gamma)$,

$$\pi(\tau) = \Gamma\tau, \quad \forall \tau \in \mathbb{H},$$

that is, a subset of $Y(\Gamma)$ is open if its inverse image under π is open in $\mathbb{H}$. This makes π an open map, since

$$\pi^{-1}(\pi(U)) = \bigcup_{\gamma \in \Gamma} \gamma(U), \quad \forall U \subset \mathbb{H}.$$

Observe that π is also a continuous map by definition. As a consequence, the modular curve $Y(\Gamma)$ is a connected topological space, since $\pi(\mathbb{H}) = Y(\Gamma)$.

The modular curves for $\Gamma(N)$, $\Gamma_0(N)$ and $\Gamma_1(N)$ are denoted

$$Y(N) = \Gamma(N) \backslash \mathbb{H}, \quad Y_0(N) = \Gamma_0(N) \backslash \mathbb{H} \quad \text{and} \quad Y_1(N) = \Gamma_1(N) \backslash \mathbb{H}.$$

In this section we show that $Y(\Gamma)$ can be made into a Riemann surface. The reader can consult the theory of Riemann surfaces in Appendix A. One of the main results we need to carry out this task is that the action of the modular group on the upper half plane is properly discontinuous, i.e., any two points in $\mathbb{H}$ have neighbourhoods small enough so that each transformation of the modular group taking one point away from the other also takes its neighbourhood away from the other's.

Proposition 1.8. *Let $\tau_1, \tau_2 \in \mathbb{H}$. Then there exist open neighbourhoods U_1 of τ_1 and U_2 of τ_2 in $\mathbb{H}$, with the following property:*

$$\text{For all } \alpha \in \mathrm{SL}_2(\mathbb{Z}),\text{ if } \alpha(U_1) \cap U_2 \neq \emptyset,\text{ then } \alpha(\tau_1) = \tau_2.$$

Proof: Let U_1' and U_2' be any open neighbourhoods of τ_1 and τ_2, respectively, with compact closure in $\mathbb{H}$. Consider the intersection

$$\alpha(U_1') \cap U_2', \quad \alpha \in \mathrm{SL}_2(\mathbb{Z}).$$

We claim that this intersection is empty for all but finitely many $\alpha \in \mathrm{SL}_2(\mathbb{Z})$.

Note that for all but finitely many pairs $(c,d) \in \mathbb{Z}^2$, with $\gcd(c,d) = 1$,

$$\sup\{\mathrm{Im}\,(\alpha(\tau)) \mid \alpha = \begin{bmatrix} * & * \\ c & d \end{bmatrix} \in \mathrm{SL}_2(\mathbb{Z}), \tau \in U_1'\} < \inf\{\mathrm{Im}\,(\tau) \mid \tau \in U_2'\}$$

holds (making the intersection empty), since

$$\frac{\mathrm{Im}\,(\tau)}{|c\tau + d|^2} \leq \min\{\frac{1}{c^2 y_1}, \frac{Y_1}{(c\mathrm{Re}\,(\tau) + d)^2}\}, \quad \forall\, \tau \in U_1', \quad \forall\, \alpha = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathrm{SL}_2(\mathbb{Z}),$$

where

$$y_1 = \inf\{\mathrm{Im}\,(\tau) \mid \tau \in U_1'\} \quad \text{and} \quad Y_1 = \sup\{\mathrm{Im}\,(\tau) \mid \tau \in U_1'\}.$$

Furthermore, for each pair $(c,d) \in \mathbb{Z}^2$, with $\gcd(c,d) = 1$, the matrices $\alpha \in \mathrm{SL}_2(\mathbb{Z})$ with bottom row (c,d) are

$$\begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \quad k \in \mathbb{Z},$$

where $(a,b) \in \mathbb{Z}^2$ is any particular pair such that $ad - bc = 1$. As

$$\alpha(U_1') \cap U_2' = \left(\begin{bmatrix} a & b \\ c & d \end{bmatrix} (U_1') + k \right) \cap U_2',$$

we deduce that the intersection is empty for all but finitely many $\alpha \in \mathrm{SL}_2(\mathbb{Z})$ with bottom row (c,d). Therefore, combining these two remarks we can conclude that there are only finitely many $\alpha \in \mathrm{SL}_2(\mathbb{Z})$ such that

$$\alpha(U_1') \cap U_2' \neq \emptyset,$$

as claimed.

Finally, let F be the finite set $\{\alpha \in \mathrm{SL}_2(\mathbb{Z}) \,|\, \alpha(U_1') \cap U_2' \neq \emptyset, \alpha(\tau_1) \neq \tau_2\}$. For each $\alpha \in F$ there exist disjoint open neighbourhoods $U_{1,\alpha}$ of $\alpha(\tau_1)$ and $U_{2,\alpha}$ of τ_2 in $\mathbb{H}$, since $\alpha(\tau_1) \neq \tau_2$. Define

$$U_1 = U_1' \cap \left(\bigcap_{\alpha \in F} \alpha^{-1}(U_{1,\alpha}) \right), \quad \text{an open neighbourhood of } \tau_1 \text{ in } \mathbb{H},$$

and

$$U_2 = U_2' \cap \left(\bigcap_{\alpha \in F} U_{2,\alpha} \right), \quad \text{an open neighbourhood of } \tau_2 \text{ in } \mathbb{H}.$$

Observe that these open neighbourhoods satisfy the desired property, i.e., if there exists $\alpha \in \mathrm{SL}_2(\mathbb{Z})$ such that $\alpha(U_1) \cap U_2 \neq \emptyset$, then necessarily $\alpha \notin F$, since otherwise

$$\alpha(U_1) \cap U_2 \subseteq U_{1,\alpha} \cap U_{2,\alpha} = \emptyset.$$

□

The first immediate consequence of this result is the following corollary.

Corollary 1.9. *The modular curve $Y(\Gamma)$ is a Hausdorff topological space.*

Proof: Let $\pi(\tau_1)$, $\pi(\tau_2)$ be distinct points in $Y(\Gamma)$. By Proposition 1.8 there exist open neighbourhoods U_1 of τ_1 and U_2 of τ_2 in $\mathbb{H}$ such that

$$\gamma(U_1) \cap U_2 = \emptyset, \quad \forall \gamma \in \Gamma.$$

As a consequence,

$$\pi(U_1) \quad \text{and} \quad \pi(U_2)$$

are disjoint open subsets of $Y(\Gamma)$ containing $\pi(\tau_1)$ and $\pi(\tau_2)$, respectively.

□

1.3.1 Elliptic points

To define charts on the curve $Y(\Gamma)$ we have to prove that the isotropy subgroups of Γ are finite cyclic. This requirement will bee more clearly seen in the next subsection where we study the action of an isotropy subgroup of Γ on the upper half plane.

Definition 1.10. *Let Γ_τ denote the isotropy subgroup of a point $\tau \in \mathbb{H}$, i.e., the τ-fixing subgroup of Γ,*

$$\Gamma_\tau = \{\gamma \in \Gamma \,|\, \gamma(\tau) = \tau\}.$$

The point τ is an elliptic point of Γ if there exists $\gamma \in \Gamma_\tau$ such that $\gamma \neq \pm I$, i.e., if Γ_τ defines a nontrivial subgroup of transformations.

Remarks 1.11.

- ▷ Let $\tau \in \mathbb{H}$. Note that τ is an elliptic point of Γ if and only if the containment of matrix groups
$$\{\pm I\} \subset \{\pm I\}\Gamma_\tau$$
is proper.

- ▷ Let $\tau, \tau' \in \mathbb{H}$ be Γ-equivalent points, $\tau' = \gamma(\tau)$ for some $\gamma \in \Gamma$. Then their isotropic subgroups are conjugated subgroups, since
$$\Gamma_{\tau'} = \gamma\Gamma_\tau\gamma^{-1}$$
Therefore, if τ is an elliptic point of Γ, then so is τ', and as a consequence it makes sense to say that the corresponding point $\pi(\tau) \in Y(\Gamma)$ is elliptic.

Definition 1.12. *A connected subset F of $\mathbb{H}$ is a fundamental domain for Γ if it satisfies the following three conditions:*

- ▷ $\mathbb{H} = \bigcup_{\gamma\in\Gamma} \gamma(F)$
- ▷ $F = \overline{G}$, *where* $G = int(F)$
- ▷ $\gamma(G) \cap G = \emptyset$ *for all* $\gamma \in \Gamma$, $\gamma \neq \pm I$.

Let F be the connected subset of $\mathbb{H}$
$$\{\tau \in \mathbb{H} \,|\, |\mathrm{Re}\,(\tau)| \leq 1/2, |\tau| \geq 1\}.$$
We represent it in Figure 1.1. The points $i, \rho = e^{2\pi i/3}$ and $\rho + 1$ are special. The following two results prove that F is a fundamental domain for $\mathrm{SL}_2(\mathbb{Z})$. In general, it is possible to prove that there exists a fundamental domain for any congruence subgroup Γ of $\mathrm{SL}_2(\mathbb{Z})$ [Miy06, p.22].

Lemma 1.13. *Let π be the natural projection from $\mathbb{H}$ to $Y(1) = \mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H}$,*
$$\pi(\tau) = \mathrm{SL}_2(\mathbb{Z})\tau, \quad \forall\, \tau \in \mathbb{H}.$$
Each point $\tau \in \mathbb{H}$ is $\mathrm{SL}_2(\mathbb{Z})$-equivalent to some point in F, i.e., $\pi(F) = Y(1)$.

Proof: Let us describe an algorithm to compute some $\tau' \in F$ such that
$$\mathrm{SL}_2(\mathbb{Z})\tau = \mathrm{SL}_2(\mathbb{Z})\tau'.$$
First apply repeatedly one of the matrices
$$\begin{bmatrix} 1 & \pm 1 \\ 0 & 1 \end{bmatrix} : \tau \mapsto \tau \pm 1$$

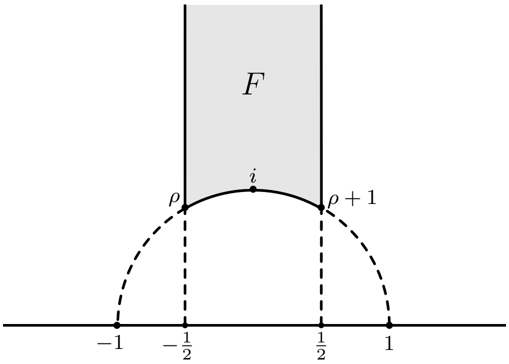


Figure 1.1: The fundamental domain for $\mathrm{SL}_2(\mathbb{Z})$

to translate τ into the vertical trip $\{z \in \mathbb{C} \,|\, |\mathrm{Re}\,(z)| \leq 1/2\}$ and replace τ by this transform. Now, if $\tau \notin F$, then necessarily $|\tau| < 1$. So

$$\mathrm{Im}\,(-1/\tau) = \mathrm{Im}\,(-\bar{\tau}/|\tau|^2) = \mathrm{Im}\,(\tau/|\tau|^2) > \mathrm{Im}\,(\tau).$$

Then replace τ by $\left[\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right](\tau) = -1/\tau$ and repeat the same process again.

The formula (1.1) shows that this algorithm finalizes with some $\tau \in F$ because there are only finitely many pairs $(c, d) \in \mathbb{Z}^2$ such that $|c\tau + d| < 1$. □

Note that the projection $\pi : F \to Y(1)$ is not injective. The translation $\left[\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right] : \tau \mapsto \tau + 1$ identifies the two boundary rays and the inversion $\left[\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right] : \tau \mapsto -1/\tau$ identifies the two halves of the boundary circular arc. But these boundary identifications are the only ones that exist in F.

Theorem 1.14. *Let τ_1 and τ_2 be points in F such that*

$$\tau_2 = \alpha(\tau_1), \quad \textit{for some } \alpha = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathrm{SL}_2(\mathbb{Z}).$$

If τ_1 and τ_2 are distinct points, then either

▷ $\mathrm{Re}\,(\tau_1) = \pm 1/2$ *and* $\tau_2 = \tau_1 \mp 1$ *or*

▷ $|\tau_1| = 1$ *and* $\tau_2 = -1/\tau_1$*.*

Otherwise, τ_1 and τ_2 must be both equal to i, ρ, or $\rho + 1$, unless $\alpha = \pm I$.

Proof: By symmetry, we can suppose without loss of generality that

$$\mathrm{Im}\,(\tau_1) \leq \mathrm{Im}\,(\tau_2),$$

or equivalently, $|c\tau_1 + d|^2 \leq 1$. Further, as $\tau_1 \in F$, $\mathrm{Im}\,(\tau_1) \geq \sqrt{3}/2$. So

$$|c|\sqrt{3}/2 \leq |c|\mathrm{Im}\,(\tau_1) = |\mathrm{Im}\,(c\tau_1 + d)| \leq |c\tau_1 + d| \leq 1,$$

As a consequence of this inequality, we can deduce that necessarily $|c| \leq 1$.

If $c = 0$, then $\alpha = \left[\begin{smallmatrix} \pm 1 & b \\ 0 & \pm 1 \end{smallmatrix}\right]$. Since $-1/2 \leq \mathrm{Re}\,(\tau_1), \mathrm{Re}\,(\tau_2) \leq 1/2$, this implies either $b = 0$ and $\alpha = \pm I$ or $|b| = 1$ and $\tau_2 = \tau_1 \pm b$, in which case one of the numbers $\mathrm{Re}\,(\tau_1)$ and $\mathrm{Re}\,(\tau_2)$ must be equal to $-1/2$ and the other to $1/2$.

If $c \neq 0$, then we can suppose that $c = 1$, since $\tau_2 = \alpha(\tau_1) = (-\alpha)(\tau_1)$. The condition $|c\tau_1 + d|^2 \leq 1$ is equivalent to

$$(\mathrm{Re}\,(\tau_1) + d)^2 + (\mathrm{Im}\,(\tau_1))^2 \leq 1.$$

So

$$(\mathrm{Re}\,(\tau_1) + d)^2 \leq 1 - (\mathrm{Im}\,(\tau_1))^2 \leq 1 - 3/4 = 1/4,$$

implying $|\mathrm{Re}\,(\tau_1) + d| \leq 1/2$. Note that this inequality forces that $|d| \leq 1$.

If $c = 1$ and $|d| = 1$, then $|\mathrm{Re}\,(\tau_1) + d| = 1/2$. So the preceding inequality implies that $|c\tau_1 + d| = 1$ and $\mathrm{Im}\,(\tau_1) = \mathrm{Im}\,(\tau_2) = \sqrt{3}/2$ (i.e., $\tau_1, \tau_2 \in \{\rho, \rho+1\}$). The case $d = 1$ forces that $\alpha = \left[\begin{smallmatrix} a & a-1 \\ 1 & 1 \end{smallmatrix}\right]$ and $\tau_2 = -1/(\tau_1+1)+a$, so either $a = 0$ and $\tau_2 = \tau_1 = \rho$ or $a = 1$, $\tau_1 = \rho$ and $\tau_2 = \rho + 1$ ($a = -1$ is not possible). And the case $d = -1$ forces that $\alpha = \left[\begin{smallmatrix} a & -(a+1) \\ 1 & -1 \end{smallmatrix}\right]$ and $\tau_2 = 1/(\tau_1 - 1) + a$, so either $a = 0$ and $\tau_2 = \tau_1 = \rho + 1$ or $a = -1$, $\tau_2 = \rho$ and $\tau_1 = \rho + 1$ ($a = 1$ is not possible).

If $c = 1$ and $d = 0$, then $\alpha = \left[\begin{smallmatrix} a & -1 \\ 1 & 0 \end{smallmatrix}\right]$ and the condition $|c\tau_1 + d| \leq 1$ becomes $|\tau_1| \leq 1$, so in fact $|\tau_1| = 1$ (since $|\tau_1| \geq 1$) and $\mathrm{Im}\,(\tau_1) = \mathrm{Im}\,(\tau_2)$. This implies either $a = 0$ and $\tau_2 = -1/\tau_1$ or $|a| = 1$ and $\tau_2 = -1/\tau_1 + a$, in which case τ_1 and τ_2 must be both equal to ρ or $\rho + 1$.

□

We deduce from this theorem that the only elliptic points of $\mathrm{SL}_2(\mathbb{Z})$ in F are i, ρ and $\rho + 1$ with isotropy subgroups

- $\mathrm{SL}_2(\mathbb{Z})_i = \left\langle \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \right\rangle$, a cyclic subgroup of order 4,
- $\mathrm{SL}_2(\mathbb{Z})_\rho = \left\langle \begin{bmatrix} 0 & -1 \\ 1 & 1 \end{bmatrix} \right\rangle$, a cyclic subgroup of order 6,
- $\mathrm{SL}_2(\mathbb{Z})_{\rho+1} = \left\langle \begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix} \right\rangle$, a cyclic subgroup of order 6.

Corollary 1.15. *The modular curve* $\mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H}$ *has two elliptic points,*

$$\mathrm{SL}_2(\mathbb{Z})i \quad \textit{and} \quad \mathrm{SL}_2(\mathbb{Z})\rho.$$

Therefore, for each $\tau \in \mathbb{H}$ *its isotropy subgroup* $\mathrm{SL}_2(\mathbb{Z})_\tau$ *is finite cyclic.*

Proof: Let $\tau \in \mathbb{H}$ be an elliptic point of $\mathrm{SL}_2(\mathbb{Z})$. By Lemma 1.13 there exists $\alpha \in \mathrm{SL}_2(\mathbb{Z})$ such that $\alpha(\tau) \in F$ and by Theorem 1.14 $\alpha(\tau)$ must be equal to i, ρ or $\rho + 1$, since $\alpha(\tau)$ is also an elliptic point of $\mathrm{SL}_2(\mathbb{Z})$.

□

Corollary 1.16. *The modular curve $Y(\Gamma)$ has only finitely many elliptic points. Furthermore, for each $\tau \in \mathbb{H}$ its isotropy subgroup Γ_τ is finite cyclic.*

Proof: Let $\mathrm{SL}_2(\mathbb{Z}) = \bigcup_{j=1}^{d} \Gamma\alpha_j$. As Γ_τ is a subgroup of $\mathrm{SL}_2(\mathbb{Z})_\tau$, $\forall\, \tau \in \mathbb{H}$, the elliptic points of $Y(\Gamma)$ are a subset of $E_\Gamma = \{\Gamma\alpha_j(i), \Gamma\alpha_j(\rho) : 1 \leq j \leq d\}$. For the second statement, recall that a subgroup of a cyclic group is cyclic.

□

Corollary 1.17. *Let $\alpha \in \mathrm{SL}_2(\mathbb{Z})$, $\alpha \neq \pm I$. If α is a elliptic matrix, i.e., $\alpha(\tau) = \tau$, with $\tau \in \mathbb{H}$, then α is conjugate to some of the following matrices:*

$$\begin{bmatrix} 0 & 1 \\ -1 & -1 \end{bmatrix}^{\pm 1}, \quad \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}^{\pm 1}, \quad \begin{bmatrix} 0 & -1 \\ 1 & 1 \end{bmatrix}^{\pm 1}.$$

As a consequence, observe that the matrix α must have order $3, 4$ or 6.

Proof: By Corollary 1.15 there exists $\beta \in \mathrm{SL}_2(\mathbb{Z})$ such that $\beta(\tau)$ is equal to i or ρ, since τ is an elliptic point of $\mathrm{SL}_2(\mathbb{Z})$. Therefore,

$$\mathrm{SL}_2(\mathbb{Z})_\tau = \beta^{-1}\mathrm{SL}_2(\mathbb{Z})_i\beta \quad \text{or} \quad \mathrm{SL}_2(\mathbb{Z})_\tau = \beta^{-1}\mathrm{SL}_2(\mathbb{Z})_\rho\beta.$$

□

Corollary 1.18. *The modular curves $Y(N)$, with $N > 1$, do not have elliptic points.*

Proof: Let us suppose that there exists $\gamma \in \Gamma(N)$, $\gamma \neq \pm I$, such that $\gamma(\tau) = \tau$ for some $\tau \in \mathbb{H}$. Then γ must be conjugate to some of the six matrices $\alpha_1, \ldots, \alpha_6$ of Corollary 1.17, i.e.,

$$\gamma = \beta\alpha_j\beta^{-1}, \quad \text{for some } \beta \in \mathrm{SL}_2(\mathbb{Z}).$$

As $\Gamma(N)$ is a normal subgroup of $\mathrm{SL}_2(\mathbb{Z})$, we deduce that

$$\alpha_j = \beta^{-1}\gamma\beta \in \beta^{-1}\Gamma(N)\beta = \Gamma(N),$$

which is evidently a contradiction, since $\alpha_j \notin \Gamma(N)$, $\forall\, j = 1, \ldots, 6$ $(N > 1)$. Hence, the modular curves $Y(N)$, with $N > 1$, do not have elliptic points.

□

1.3.2 Complex charts

Each point $\tau \in \mathbb{H}$ has an associated positive integer,

$$h_\tau = h_{\tau,\Gamma} = |\{\pm I\}\Gamma_\tau/\{\pm I\}| = \begin{cases} |\Gamma_\tau|/2 & \text{if } -I \in \Gamma_\tau, \\ |\Gamma_\tau| & \text{if } -I \notin \Gamma_\tau. \end{cases}$$

This h_τ is called the period of τ with respect to Γ for reasons to be explained. Its definition counts correctly the τ-fixing transformations induced by Γ. As a consequence, observe that

$$h_\tau > 1 \text{ if and only if } \tau \text{ is an elliptic point of } \Gamma.$$

Let $\alpha \in \mathrm{SL}_2(\mathbb{Z})$. Then the isotropy subgroups Γ_τ and $(\alpha\Gamma\alpha^{-1})_{\alpha(\tau)}$ are conjugated subgroups, since

$$(\alpha\Gamma\alpha^{-1})_{\alpha(\tau)} = \alpha\Gamma_\tau\alpha^{-1}.$$

Hence the period of $\alpha(\tau)$ under $\alpha\Gamma\alpha^{-1}$ is equal to the period of τ under Γ. This proves in particular that the period of $\pi(\tau) \in Y(\Gamma)$ is also well defined.

Examples 1.19.

1. The period of a point $\tau \in \mathbb{H}$ with respect to $\Gamma(N)$, $N > 1$, is $h_\tau = 1$.

2. The periods of the points $i, \rho \in \mathbb{H}$ with respect to $\mathrm{SL}_2(\mathbb{Z})$ are

$$h_i = 2 \quad \text{and} \quad h_\rho = 3.$$

The following corollary which will be necessary to define charts on $Y(\Gamma)$ is another immediate consequence of Proposition 1.8.

Corollary 1.20. *Each point $\tau \in \mathbb{H}$ has an open neighbourhood U in $\mathbb{H}$, with the following property:*

$$\textit{For all } \gamma \in \Gamma, \textit{ if } \gamma(U) \cap U \neq \emptyset, \textit{ then } \gamma \in \Gamma_\tau.$$

Such an open neighbourhood has no elliptic points of Γ except possibly τ.

Proof: Let $\tau_i = \tau$, $i = 1, 2$. By Proposition 1.8, there exist open neighbourhoods U_1 of τ_1 and U_2 of τ_2 in $\mathbb{H}$ such that

$$\text{for all } \alpha \in \mathrm{SL}_2(\mathbb{Z}), \text{ if } \alpha(U_1) \cap U_2 \neq \emptyset, \text{ then } \alpha(\tau_1) = \tau_2.$$

Define $U = U_1 \cap U_2$. Note that

$$\text{for all } \gamma \in \Gamma, \text{ if } \gamma(U) \cap U \neq \emptyset, \text{ then } \gamma \in \Gamma_\tau.$$

Let us now suppose that there exists $\gamma \in \Gamma$, $\gamma \neq \pm I$, such that $\gamma(\tau') = \tau'$, for some $\tau' \in U$. Then $\gamma(U) \cap U \neq \emptyset$, implying $\gamma \in \Gamma_\tau$. But γ has only one fixed point in $\mathbb{H}$ (elliptic transformation), so necessarily $\tau = \tau'$.

□

Let $\tau \in \mathbb{H}$. Define the matrix $\delta_\tau = \left[\begin{smallmatrix} 1 & -\tau \\ 1 & -\bar{\tau} \end{smallmatrix}\right] \in \mathrm{GL}_2(\mathbb{C})$ and observe that the induced transformation on the Riemann sphere

$$\delta_\tau(z) = \frac{z - \tau}{z - \bar{\tau}}, \quad \forall\, z \in \widehat{\mathbb{C}},$$

satisfies

$$\delta_\tau(\tau) = 0, \quad \delta_\tau(\bar{\tau}) = \infty \quad \text{and} \quad \delta_\tau(\widehat{\mathbb{R}}) = \partial\mathbb{D}, \quad \text{where } \widehat{\mathbb{R}} = \mathbb{R} \cup \{\infty\}.$$

As a consequence, it follows that δ_τ induces an isomorphism from $\mathbb{H}$ to $\mathbb{D}$.

Let Γ_{δ_τ} denote the matrix subgroup

$$\delta_\tau \Gamma_\tau \delta_\tau^{-1} = (\delta_\tau \Gamma \delta_\tau^{-1})_0 \subset \mathrm{SL}_2(\mathbb{C}).$$

The action of this subgroup on $\mathbb{D}$ is equivalent to the action of Γ_τ on $\mathbb{H}$,

$$\begin{aligned} \tau_1, \tau_2 \in \mathbb{H} \text{ are } \Gamma_\tau\text{-equivalent } &\Leftrightarrow \ \tau_2 \in \Gamma_\tau \tau_1 \\ &\Leftrightarrow \ \delta_\tau(\tau_2) \in (\delta_\tau \Gamma_\tau \delta_\tau^{-1})(\delta_\tau(\tau_1)) \\ &\Leftrightarrow \ \delta_\tau(\tau_1), \delta_\tau(\tau_2) \in \mathbb{D} \text{ are } \Gamma_{\delta_\tau}\text{-equivalent.} \end{aligned}$$

Let us consider the subgroup of transformations defined by Γ_{δ_τ},

$$\delta_\tau \{\pm I\} \Gamma_\tau \delta_\tau^{-1} / \{\pm I\},$$

which must be finite cyclic of order h_τ by Proposition 1.16. As this subgroup of transformations fixes 0 and ∞, it consists of maps of the form

$$z \mapsto az, \quad a \in \mathbb{C},$$

and since the subgroup is finite cyclic of order h_τ, these must be rotations through angular multiples of $2\pi/h_\tau$ about the origin,

$$z \mapsto e^{2\pi i k/h_\tau} z, \quad k \in \mathbb{Z}.$$

Taking into account this, we can easily describe the action of Γ_{δ_τ} on $\mathbb{D}$,

$$z_1, z_2 \in \mathbb{D} \text{ are } \Gamma_{\delta_\tau}\text{-equivalent } \Leftrightarrow \ z_2 = e^{2\pi i k/h_\tau} z_1, \ \text{ for some } k \in \mathbb{Z}.$$

Observe that each circular sector of angle $2\pi/h_\tau$ in which has been divided the unit disc corresponds in the upper half plane with a "triangle" formed by two circular arcs (including lines) orthogonal to the real line and a segment of real line (possibly unbounded). It is for this reason that we say that δ_τ is straightening neighbourhoods of τ to neighbourhoods of the origin. The figure 1.2 illustrates representative two cases, $\Gamma = \mathrm{SL}_2(\mathbb{Z})$ and $\tau = i, \rho$.

The previous description of the action of Γ_{δ_τ} on $\mathbb{D}$ suggests to consider the wrapping map $\lambda_\tau : \mathbb{D} \to \mathbb{D}$,

$$\lambda_\tau(z) = z^{h_\tau}, \quad \forall\, z \in \mathbb{D}.$$

Observe that this map allow us to write that

$$\tau_1, \tau_2 \in \mathbb{H} \text{ are } \Gamma_\tau\text{-equivalent } \Leftrightarrow \ \lambda_\tau(\delta_\tau(\tau_1)) = \lambda_\tau(\delta_\tau(\tau_2)).$$

Let us construct a chart on $Y(\Gamma)$ about the point $\pi(\tau)$. By Corollary 1.20 there exists an open neighbourhood U of τ in $\mathbb{H}$ with the following properties:

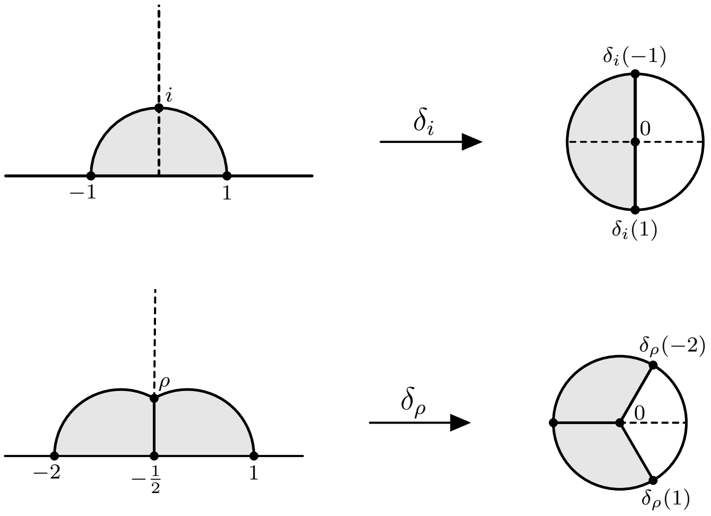


Figure 1.2: $\Gamma = \mathrm{SL}_2(\mathbb{Z})$ and $\tau = i, \rho$

- For all $\gamma \in \Gamma$, if $\gamma(U) \cap U \neq \emptyset$, then $\gamma \in \Gamma_\tau$.
- U has no elliptic points of Γ except possibly τ.

Define $\psi_\tau : U \to \mathbb{C}$ as

$$\psi_\tau(\tau') = \lambda_\tau(\delta_\tau(\tau')), \quad \forall\, \tau' \in U,$$

and let $V = \psi_\tau(U)$. Then for any points $\tau_1, \tau_2 \in U$,

$$\pi(\tau_1) = \pi(\tau_2) \;\Leftrightarrow\; \tau_1 \in \Gamma\tau_2 \;\Leftrightarrow\; \tau_1 \in \Gamma_\tau\tau_2 \;\Leftrightarrow\; \psi_\tau(\tau_1) = \psi_\tau(\tau_2),$$

i.e., the projection π and the wrapping ψ_τ identify the same points of U. As a consequence, there exists a bijection $\phi_\tau : \pi(U) \to V$ making the diagram

$$\begin{tikzcd} & U \arrow[dl, "\pi"'] \arrow[dr, "\psi_\tau"] & \\ \pi(U) \arrow[rr, "\phi_\tau"] & & V \end{tikzcd}$$

commutative. Further, as π and ψ_τ are both continuous and open maps, we deduce that this bijection $\phi_\tau : \pi(U) \to V$ is in fact a homeomorphism.

Caution: This complex chart depend on the open neighbourhood U of τ.

Below, we show that all these charts determine an atlas on the curve $Y(\Gamma)$. Let $\phi_1 : \pi(U_1) \to V_1$ and $\phi_2 : \pi(U_2) \to V_2$ be two charts such that

$$\phi_1 = \phi_{\tau_1}, \quad \phi_2 = \phi_{\tau_2} \quad \text{and} \quad \pi(U_1) \cap \pi(U_2) \neq \emptyset.$$

Consider the commutative diagram

$$\begin{tikzcd}
 & \pi(U_1)\cap\pi(U_2) \arrow[dr, "\phi_2"] & \\
V_{1,2} \arrow[ur, "\phi_1^{-1}"] \arrow[rr, "\phi_{2,1}"] & & V_{2,1}
\end{tikzcd}$$

where $\phi_{2,1} = \phi_2 \circ \phi_1^{-1}$, $V_{1,2} = \phi_1(\pi(U_1) \cap \pi(U_2))$ and $V_{2,1} = \phi_2(\pi(U_1) \cap \pi(U_2))$. We have to prove that $\phi_{2,1}$ is holomorphic at $\phi_1(x)$ for all $x \in \pi(U_1) \cap \pi(U_2)$. Put $x = \pi(\tilde{\tau}_1) = \pi(\tilde{\tau}_2)$ with $\tilde{\tau}_1 \in U_1$, $\tilde{\tau}_2 \in U_2$ and $\tilde{\tau}_2 = \gamma\tilde{\tau}_1$ for some $\gamma \in \Gamma$. Let $U_{1,2} = U_1 \cap \gamma^{-1}(U_2)$, an open neighbourhood of $\tilde{\tau}_1$ in $\mathbb{H}$. Since π is open, its projection $\pi(U_{1,2})$ is an open neighbourhood of x in $\pi(U_1) \cap \pi(U_2)$.

An input point $q = \phi_1(x')$ to $\phi_{2,1}$ in $\phi_1(\pi(U_{1,2}))$ is of the form

$$q = \phi_1(\pi(\tau')) = \psi_1(\tau') = (\delta_1(\tau'))^{h_1}, \quad \text{for some } \tau' \in U_{1,2},$$

where $\delta_1 = \delta_{\tau_1}$ and h_1 is the period of τ_1. So the corresponding output is

$$\begin{aligned}\phi_2(x') &= \phi_2(\pi(\gamma(\tau'))) = \psi_2(\gamma(\tau')) \quad (\text{since } \gamma(\tau') \in U_2)\\ &= (\delta_2(\gamma(\tau')))^{h_2} = ((\delta_2\gamma\delta_1^{-1})(\delta_1(\tau')))^{h_2},\end{aligned}$$

where $\delta_2 = \delta_{\tau_2}$ and h_2 is the period of τ_2.

This calculation shows that the only case possible where the transition function might not be holomorphic at $\phi_1(x)$ is when $\delta_1(\tilde{\tau}_1) = 0$ and $h_1 > 1$, i.e., when $\tau_1 = \tilde{\tau}_1$ and τ_1 is an elliptic point of Γ. In this case, $\tilde{\tau}_2 = \gamma(\tau_1)$ would also be an elliptic point of Γ with the same period, implying $\tau_2 = \tilde{\tau}_2$ (recall that U_2 has no elliptic points except possibly τ_2 by construction). Therefore,

$$\delta_2\gamma\delta_1^{-1} = \begin{bmatrix} \alpha & 0 \\ 0 & \beta \end{bmatrix}, \quad \text{for some } \alpha, \beta \in \mathbb{C}^*,$$

since

$$0 \xmapsto{\delta_1^{-1}} \tau_1 \xmapsto{\gamma} \tau_2 \xmapsto{\delta_2} 0 \quad \text{and} \quad \infty \xmapsto{\delta_1^{-1}} \overline{\tau}_1 \xmapsto{\gamma} \overline{\tau}_2 \xmapsto{\delta_2} \infty.$$

As a consequence, the formula for $\phi_{2,1}$ becomes

$$\phi_{2,1}(q) = \left(\begin{bmatrix} \alpha & 0 \\ 0 & \beta \end{bmatrix} (q^{1/h_1}) \right)^{h_1} = (\alpha/\beta)^{h_1} q, \quad \forall\, q \in \phi_1(\pi(U_{1,2})),$$

Observe that this proves that transition function is holomorphic at $\phi_1(x) = 0$.

The modular curve $Y(\Gamma)$ is now a (noncompact) Riemann surface.

1.4 The Riemann surfaces $X(\Gamma) = \Gamma\backslash\mathbb{H}^*$

In this last section of the chapter we show that the Riemann surface $Y(\Gamma)$ can be compactified. The resulting compact Riemann surface is denoted $X(\Gamma)$. The Riemann-Hurwitz formula A.28 will allow us to calculate its genus.

1.4.1 Cusps

To compactify the modular curve $Y(\Gamma) = \Gamma\backslash\mathbb{H}$, define

$$\mathbb{H}^* = \mathbb{H} \cup \widehat{\mathbb{Q}}, \quad \text{where} \quad \widehat{\mathbb{Q}} = \mathbb{Q} \cup \{\infty\},$$

and consider the following topology on $\mathbb{H}^*$:

- A fundamental system of neighbourhoods at $\tau \in \mathbb{H}$ is formed by the usual open neighbourhoods of τ in $\mathbb{H}$.
- A fundamental system of neighbourhoods at $s \in \widehat{\mathbb{Q}}$ is formed by the subsets

  $$\alpha(\mathcal{N}_M \cup \{\infty\}) : \quad M > 0,\ \alpha \in \mathrm{SL}_2(\mathbb{Z}),\ \alpha(\infty) = s,$$

  where $\mathcal{N}_M = \{\tau \in \mathbb{H} \mid \mathrm{Im}\,(\tau) > M\}$ for any real number $M > 0$ ($\mathcal{N} := \mathcal{N}_1$).

As fractional linear transformations are conformal and take circles to circles, if $\alpha(\infty) \in \mathbb{Q}$, then $\alpha(\mathcal{N}_M \cup \{\infty\})$ is a disc tangent to the real line at $\alpha(\infty)$. The formula (1.1) allow us to compute that the radius of this disc is $1/2c^2M$, where c is the lower left entry of the matrix α. Therefore,

$$\alpha(\mathcal{N}_M \cup \{\infty\}) \cap (\mathcal{N} \cup \{\infty\}) = \emptyset, \quad \forall\, M \geq 1.$$

If $\alpha(\infty) = \infty$, then α is a translation matrix and $\alpha(\mathcal{N}_M \cup \{\infty\}) = \mathcal{N}_M \cup \{\infty\}$, since the isotropy subgroup of ∞ in $\mathrm{SL}_2(\mathbb{Z})$ is

$$\mathrm{SL}_2(\mathbb{Z})_\infty = \left\{ \pm \begin{bmatrix} 1 & m \\ 0 & 1 \end{bmatrix} \mid m \in \mathbb{Z} \right\}.$$

Lemma 1.21. *Let $\alpha \in \mathrm{SL}_2(\mathbb{Z})$. Then the following conditions are equivalent:*

1. *$\alpha \in \mathrm{SL}_2(\mathbb{Z})_\infty$*
2. *$\alpha(\mathcal{N}_M) = \mathcal{N}_M$, for any $M > 0$.*
3. *$\alpha(\mathcal{N}_M) \cap \mathcal{N}_M \neq \emptyset$, for some $M \geq 1$.*

As a consequence, observe that $\mathcal{N}$ does not contains elliptic points of $\mathrm{SL}_2(\mathbb{Z})$.

The proof of this lemma is easy and is left as an exercise to the reader. Below, the Figure 1.3 shows $\mathcal{N} \cup \{\infty\}$ and some of its $\mathrm{SL}_2(\mathbb{Z})$-translates.

Note that $\mathbb{H}^*$ is a Hausdorff topological space with respect to this topology. The modular group acts on this space via fractional linear transformations.

Let Γ be a congruence subgroup of $\mathrm{SL}_2(\mathbb{Z})$ acting on $\mathbb{H}^*$. The compact modular curve $X(\Gamma)$ is defined as the quotient space of orbits under Γ,

$$X(\Gamma) = \Gamma\backslash\mathbb{H}^* = Y(\Gamma) \cup \Gamma\backslash\widehat{\mathbb{Q}}.$$

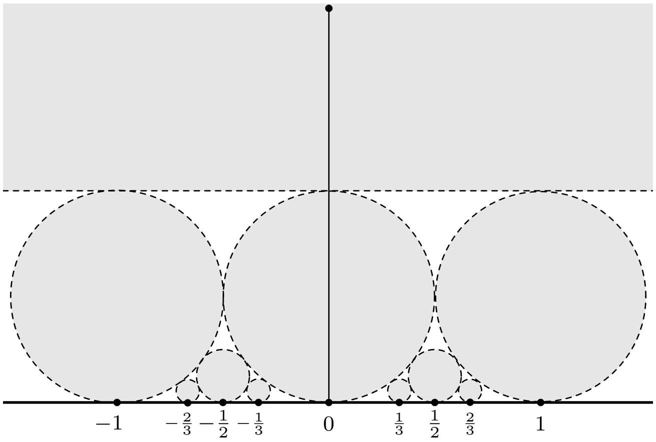


Figure 1.3: Neighbourhoods of ∞ and of some rational points

As in Section 1.3, the topology of this quotient space is induced by the natural projection $\pi : \mathbb{H}^* \to X(\Gamma)$ which is an open continuous map. The Γ-equivalence classes of points in $\widehat{\mathbb{Q}}$ are also called the cusps of $X(\Gamma)$.

The (compact) modular curves for $\Gamma(N)$, $\Gamma_0(N)$ and $\Gamma_1(N)$ are denoted

$$X(N) = \Gamma(N)\backslash\mathbb{H}, \quad X_0(N) = \Gamma_0(N)\backslash\mathbb{H} \quad \text{and} \quad X_1(N) = \Gamma_1(N)\backslash\mathbb{H}.$$

Lemma 1.22. *The modular curve* $\mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H}^*$ *has only one cusp,* $\mathrm{SL}_2(\mathbb{Z})\infty$.

Proof: Let $s = a/c$ be a rational number in reduced form, $\gcd(a, c) = 1$. By the Bézout's identity, there exist integers b and d such that $ad - bc = 1$. Then

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathrm{SL}_2(\mathbb{Z}) \quad \text{and} \quad \begin{bmatrix} a & b \\ c & d \end{bmatrix}(\infty) = s.$$

□

Corollary 1.23. *The modular curve* $X(\Gamma)$ *has only finitely many cusps.*

Proof: Let $\mathrm{SL}_2(\mathbb{Z}) = \bigcup_{j=1}^{d} \Gamma\alpha_j$. For each $s \in \widehat{\mathbb{Q}}$ there exists $\alpha \in \mathrm{SL}_2(\mathbb{Z})$ such that $\alpha(\infty) = s$. So s is Γ-equivalent to $\alpha_j(\infty)$, for some $j = 1, \ldots, d$.

□

Theorem 1.24. *The modular curve* $X(\Gamma)$ *is Hausdorff, connected and compact.*

Proof: Let $x_1, x_2 \in X(\Gamma)$ be distinct points. To prove that $X(\Gamma)$ is Hausdorff, we have to find disjoint open neighbourhoods of these two points.

Note that the case $x_1 = \Gamma\tau_1$, $x_2 = \Gamma\tau_2$, with $\tau_1, \tau_2 \in \mathbb{H}$, was already proved in Corollary 1.9, since the natural inclusion $Y(\Gamma) \hookrightarrow X(\Gamma)$ is an open map.

Suppose that $x_1 = \Gamma s_1$, $x_2 = \Gamma\tau_2$, with $s_1 \in \mathbb{Q} \cup \{\infty\}$ and $\tau_2 \in \mathbb{H}$. Then $s_1 = \alpha_1(\infty)$ for some $\alpha_1 \in \mathrm{SL}_2(\mathbb{Z})$. Let U_2 be any open neighbourhood of τ_2 with compact closure in $\mathbb{H}$. The inequality

$$\mathrm{Im}\,(\alpha(\tau)) \leq \max\{\mathrm{Im}\,(\tau), 1/\mathrm{Im}\,(\tau)\}, \quad \forall\,\tau \in \mathbb{H}, \forall\,\alpha \in \mathrm{SL}_2(\mathbb{Z}),$$

shows that there exists $M > 0$ such that $\alpha(\overline{U_2}) \cap \mathcal{N}_M = \emptyset$, $\forall\,\alpha \in \mathrm{SL}_2(\mathbb{Z})$. Let $U_1 = \alpha_1(\mathcal{N}_M \cup \{\infty\})$. Then

$$\pi(U_1) \quad \text{and} \quad \pi(U_2)$$

are disjoint open subsets of the curve $X(\Gamma)$ containing x_1 and x_2, respectively.

Suppose now that $x_1 = \Gamma s_1$, $x_2 = \Gamma s_2$, with $s_1, s_2 \in \mathbb{Q} \cup \{\infty\}$. Then $s_1 = \alpha_1(\infty)$, $s_2 = \alpha_2(\infty)$ for some $\alpha_1, \alpha_2 \in \mathrm{SL}_2(\mathbb{Z})$. Let

$$U_1 = \alpha_1(\mathcal{N} \cup \{\infty\}) \quad \text{and} \quad U_2 = \alpha_2(\mathcal{N} \cup \{\infty\}).$$

Then $\pi(U_1)$ and $\pi(U_2)$ be must disjoint open subsets of $X(\Gamma)$, since otherwise there exists $\gamma \in \Gamma$ such that $\alpha_2^{-1}\gamma\alpha_1(\mathcal{N} \cup \{\infty\}) \cap (\mathcal{N} \cup \{\infty\}) \neq \emptyset$, i.e., $\alpha_2^{-1}\gamma\alpha_1 \in \mathrm{SL}_2(\mathbb{Z})_\infty$, but this is not possible since the points x_1 and x_2 are distinct. These three cases prove that the curve $X(\Gamma)$ is Hausdorff.

Let $\mathbb{H}^* = G_1 \cup G_2$ be a disjoint union of two open subsets. Then

$$\mathbb{H} = (G_1 \cap \mathbb{H}) \cup (G_2 \cap \mathbb{H}).$$

But $\mathbb{H}$ is a connected subset and the subsets $G_1 \cap \mathbb{H}$ and $G_2 \cap \mathbb{H}$ are open, so this implies that either $\mathbb{H} \subset G_1$ and $G_2 = \emptyset$ or $G_1 = \emptyset$ and $\mathbb{H} \subset G_2$. Thus we conclude that $\mathbb{H}^*$ is connected and therefore so is the curve $X(\Gamma)$.

For compactness, let $\mathrm{SL}_2(\mathbb{Z}) = \bigcup_{j=1}^{d} \Gamma\alpha_j$. As $F^* = F \cup \{\infty\}$ is compact subset of $\mathbb{H}^*$ and

$$\mathbb{H}^* = \mathrm{SL}_2(\mathbb{Z})F^* = \bigcup_{j=1}^{d} \Gamma\alpha_j(F^*),$$

we deduce that the curve modular $X(\Gamma)$ is a finite union of compact subsets,

$$X(\Gamma) = \bigcup_{j=1}^{d} \pi(\gamma_j(F^*)).$$

□

1.4.2 Complex charts

Each point $s \in \widehat{\mathbb{Q}}$ has an associated positive integer

$$h_s = h_{s,\Gamma} = |\mathrm{SL}_2(\mathbb{Z})_\infty / (\delta_s\{\pm I\}\Gamma\delta_s^{-1})_\infty|,$$

where $\delta_s \in \mathrm{SL}_2(\mathbb{Z})$ takes s to ∞. This h_s is called the width of s with respect to Γ for reasons to be explained. As the subgroup $\mathrm{SL}_2(\mathbb{Z})_\infty = \{\pm I\}\langle\left[\begin{smallmatrix}1&1\\0&1\end{smallmatrix}\right]\rangle$ is infinite cyclic as a group of transformations, the width of s is characterized by the conditions

$$\{\pm I\}(\delta_s\Gamma\delta_s^{-1})_\infty = \{\pm I\}\left\langle\begin{bmatrix}1 & h\\ 0 & 1\end{bmatrix}\right\rangle, \quad \text{with } h > 0.$$

Observe that the width of s is independent of the matrix δ_s taking s to ∞, since

$$h_s = |\mathrm{SL}_2(\mathbb{Z})_s / \{\pm I\}\Gamma_s|.$$

Let $\alpha \in \mathrm{SL}_2(\mathbb{Z})$. Then the isotropy subgroups $\mathrm{SL}_2(\mathbb{Z})_s$ and $\mathrm{SL}_2(\mathbb{Z})_{\alpha(s)}$ are conjugated subgroups, $\mathrm{SL}_2(\mathbb{Z})_{\alpha(s)} = \alpha\mathrm{SL}_2(\mathbb{Z})_s\alpha^{-1}$. Further,

$$\{\pm I\}(\alpha\Gamma\alpha^{-1})_{\alpha(s)} = \alpha\{\pm I\}\Gamma_s\alpha^{-1}.$$

Hence the width of $\alpha(s)$ under $\alpha\Gamma\alpha^{-1}$ is equal to the width of s under Γ. This proves in particular that the width of $\pi(s) \in X(\Gamma)$ is also well defined.

Examples 1.25.

1. The width of a point $s \in \widehat{\mathbb{Q}}$ with respect to $\mathrm{SL}_2(\mathbb{Z})$ is $h_s = 1$.
2. More generally, the width of a point $s \in \widehat{\mathbb{Q}}$ with respect to $\Gamma(N)$ is $h_s = N$.

Let us construct a chart on $X(\Gamma)$ about the point $\pi(s)$. Define

$$U = \delta_s^{-1}(\mathcal{N} \cup \{\infty\}).$$

Note that this open neighbourhood of s in $\mathbb{H}^*$ has the following property:

For all $\gamma \in \Gamma$, if $\gamma(U) \cap U \neq \emptyset$, then $\gamma \in \Gamma_s$.

As a consequence,

$$z_1, z_2 \in U \text{ are } \Gamma\text{-equivalent } \Leftrightarrow\ z_1, z_2 \in U \text{ are } \Gamma_s\text{-equivalent.}$$

Consider the subgroup

$$\{\pm I\}(\delta_s\Gamma_s\delta_s^{-1}) = \{\pm I\}(\delta_s\Gamma\delta_s^{-1})_\infty = \{\pm I\}\left\langle\begin{bmatrix}1 & h_s\\ 0 & 1\end{bmatrix}\right\rangle.$$

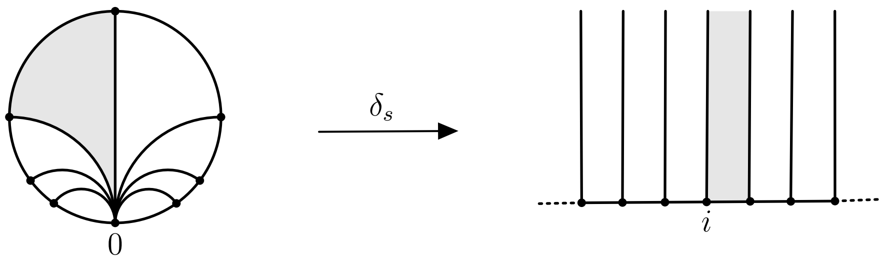


Figure 1.4: $\Gamma = \mathrm{SL}_2(\mathbb{Z})$, $s = 0$ and $\delta_s = \left[\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right]$

Then for any points $\tau_1, \tau_2 \in U$,

$$\begin{aligned} \pi(\tau_1) = \pi(\tau_2) &\Leftrightarrow \tau_1 \in \Gamma_s \tau_2 \\ &\Leftrightarrow \delta_s(\tau_1) \in \delta_s \Gamma_s \delta_s^{-1}(\delta_s(\tau_2)) \\ &\Leftrightarrow \delta_s(\tau_1) = \delta_s(\tau_2) + m h_s, \quad \text{for some } m \in \mathbb{Z}. \end{aligned}$$

This shows that the width of a cusp is the number of unit vertical strips in $\mathcal{N}$ that are distinct under isotropy. The Figure 1.4 illustrates the situation. Observe that each unit vertical strip in $\mathcal{N}$ corresponds in U with a "triangle" formed by three circular arcs (including lines). Hence the width of a cusp is also the number of such triangles that are not identified under isotropy. This time δ_s is straightening neighbourhoods of s to neighbourhoods of ∞.

Define $\psi_s : U \to \mathbb{C}$ as

$$\psi_s(\tau) = \lambda_s(\delta_s(\tau)), \quad \forall \, \tau \in U,$$

where λ_s is the h_s-periodic wrapping map $\lambda_s(z) = e^{2\pi i z/h_s}$ ($\lambda_s(\infty) := 0$). Taking into account the above, it is immediate to check that this map and the projection π identify the same points of U. Let $V = \psi_s(U)$ $(= (e^{-2\pi/h_s})\mathbb{D})$. Then there exists a bijection $\phi_s : \pi(U) \to V$ making the following diagram

$$\begin{tikzcd} & U \arrow[dl, "\pi"'] \arrow[dr, "\psi_s"] & \\ \pi(U) \arrow[rr, "\phi_s"] & & V \end{tikzcd}$$

commutative. Further, as π and ψ_s are both continuous and open maps, we deduce that this bijection $\phi_s : \pi(U) \to V$ is in fact a homeomorphism.

Caution: This chart complex depend on the matrix δ_s taking s to ∞.

Below, we show that all these complex charts together with the ones constructed in the previous section determine an atlas on the curve $X(\Gamma)$.

Let $\phi_1 : \pi(U_1) \to V_1$ and $\phi_2 : \pi(U_2) \to V_2$ be two charts such that

$$\pi(U_1) \cap \pi(U_2) \neq \emptyset.$$

Consider the commutative diagram

$$\begin{tikzcd}
 & \pi(U_1) \cap \pi(U_2) \arrow[dr, "\phi_2"] & \\
V_{1,2} \arrow[ur, "\phi_1^{-1}"] \arrow[rr, "\phi_{2,1}"] & & V_{2,1}
\end{tikzcd}$$

where $\phi_{2,1} = \phi_2 \circ \phi_1^{-1}$, $V_{1,2} = \phi_1(\pi(U_1) \cap \pi(U_2))$ and $V_{2,1} = \phi_2(\pi(U_1) \cap \pi(U_2))$. We have to prove that $\phi_{2,1}$ is holomorphic at $\phi_1(x)$ for all $x \in \pi(U_1) \cap \pi(U_2)$. Put $x = \pi(\tilde{\tau_1}) = \pi(\tilde{\tau_2})$ with $\tilde{\tau_1} \in U_1$, $\tilde{\tau_2} \in U_2$ and $\tilde{\tau_2} = \gamma\tilde{\tau_1}$ for some $\gamma \in \Gamma$. Let $U_{1,2} = U_1 \cap \gamma^{-1}(U_2)$, an open neighbourhood of $\tilde{\tau_1}$ in $\mathbb{H}^*$. Since π is open, its projection $\pi(U_{1,2})$ is an open neighbourhood of x in $\pi(U_1) \cap \pi(U_2)$. Note that the case $\phi_1 = \phi_{\tau_1}$ and $\phi_2 = \phi_{\tau_2}$, with $\tau_1, \tau_2 \in \mathbb{H}$, was already proved in the previous section.

Suppose that $\phi_1 = \phi_{\tau_1}$, with $\tau_1 \in \mathbb{H}$, and $\phi_2 = \phi_{s_2}$, with $s_2 \in \mathbb{Q} \cup \{\infty\}$. As before, an input point $q = \phi_1(x')$ to $\phi_{2,1}$ in $\phi_1(\pi(U_{1,2}))$ is of the form

$$q = \phi_1(\pi(\tau')) = \psi_1(\tau') = (\delta_1(\tau'))^{h_1}, \quad \text{for some } \tau' \in U_{1,2},$$

where $\delta_1 = \delta_{\tau_1}$ and h_1 is the period of τ_1. So the corresponding output is

$$\begin{aligned}\phi_2(x') &= \phi_2(\pi(\gamma(\tau'))) = \psi_2(\gamma(\tau')) \quad (\text{since } \gamma(\tau') \in U_2) \\ &= \exp(2\pi i \delta_2(\gamma(\tau'))/h_2) = \exp(2\pi i \delta_2 \gamma \delta_1^{-1}(\delta_1(\tau'))/h_2),\end{aligned}$$

where $\delta_2 : s_2 \mapsto \infty$ is the straightening map of ϕ_2 and h_2 is the width of s_2. As a consequence, observe that the only case possible where the transition function might not be holomorphic at $\phi_1(x)$ is when $\delta_1(\tilde{\tau_1}) = 0$ and $h_1 > 0$, i.e., $\tau_1 = \tilde{\tau_1}$ and τ_1 is an elliptic point of Γ. But this case is not possible, since otherwise $\delta_2(\gamma(\tau_1)) \in \mathcal{N}$ would also be an elliptic point of Γ, which is an contradiction.

This argument also covers the case $\phi_1 = \phi_{s_1}$, with $s_1 \in \mathbb{Q} \cup \{\infty\}$, and $\phi_2 = \phi_{\tau_2}$, with $\tau_2 \in \mathbb{H}$, since the inverse of a holomorphic biyection is also holomorphic.

Suppose now that $\phi_1 = \phi_{s_1}$ and $\phi_2 = \phi_{s_2}$, with $s_1, s_2 \in \mathbb{Q} \cup \{\infty\}$. Let $\delta_1 : s_1 \mapsto \infty$ and $\delta_2 : s_2 \mapsto \infty$ be the corresponding straightening maps. As $\pi(U_1) \cap \pi(U_2) \neq \emptyset$, there exists $\gamma \in \Gamma$ such that $\delta_2 \gamma \delta_1^{-1}$ is a translation,

$$\delta_2 \gamma \delta_1^{-1} = \pm \begin{bmatrix} 1 & m \\ 0 & 1 \end{bmatrix}, \quad \text{for some } m \in \mathbb{Z}.$$

As a consequence,

$$\gamma(s_1) = \gamma \delta_1^{-1}(\infty) = \pm \delta_2^{-1} \begin{bmatrix} 1 & m \\ 0 & 1 \end{bmatrix} (\infty) = s_2.$$

In this case, an input point $q = \phi_1(x')$ to $\phi_{2,1}$ in $\phi_1(\pi(U_{1,2}))$ is of the form

$$q = \phi_1(\pi(\tau')) = \psi_1(\tau') = \exp(2\pi i \delta_1(\tau)/h_1), \quad \text{for some } \tau' \in U_{1,2},$$

where h_1 is the width of s_1. So the corresponding output is

$$\begin{aligned} \phi_2(x') &= \phi_2(\pi(\gamma(\tau'))) = \psi_2(\gamma(\tau')) \quad (\text{since } \gamma(\tau') \in U_2) \\ &= \exp(2\pi i \delta_2 \gamma \delta_1^{-1}(\delta_1(\tau'))/h_1) \\ &= \exp(2\pi i(\delta_1(\tau') + m)/h_1) = \exp(2\pi i m/h_1) q \end{aligned}$$

Observe that this proves that transition function is holomorphic at $\phi_1(x)$.

The modular curve $X(\Gamma)$ is now a compact Riemann surface. Figure 1.5 summarizes the complex charts of $X(\Gamma)$ for future references.

<table>
<tr><td colspan="2">$\pi : \mathbb{H}^* \to X(\Gamma)$ is the natural projection.
$U \subset \mathbb{H}^*$ is a neighbourhood containing at most one elliptic point or cusp.
The complex chart $\phi : \pi(U) \to V$ satisfies $\phi \circ \pi = \psi$,
where $\psi : U \to V$ is a composition $\psi = \lambda \circ \delta$.</td></tr>
<tr><td>About $\tau \in \mathbb{H}$:</td><td>About $s \in \mathbb{Q} \cup \{\infty\}$:</td></tr>
<tr><td>The straightening map is $z = \delta(\tau')$,
where $\delta = \begin{bmatrix} 1 & -\tau \\ 1 & -\bar{\tau} \end{bmatrix}$, $\delta(\tau) = 0$.
$\delta(U)$ is a neighbourhood of 0.</td><td>The straightening map is $z = \delta(\tau')$,
where $\delta \in \mathrm{SL}_2(\mathbb{Z})$, $\delta(s) = \infty$.
$\delta(U)$ is a neighbourhood of ∞.</td></tr>
<tr><td>The wrapping map is $q = \lambda(z)$
where $\lambda(z) = z^h$, $\lambda(0) = 0$,
with period $h = |\{\pm I\}\Gamma_\tau/\{\pm I\}|$.
$V = \lambda(\delta(U))$ is a neighbourhood of 0.</td><td>The wrapping map is $q = \lambda(z)$
where $\lambda(z) = e^{2\pi i z/h}$, $\lambda(\infty) = 0$,
with width $h = |\mathrm{SL}_2(\mathbb{Z})_s/\{\pm I\}\Gamma_s|$.
$V = \lambda(\delta(U))$ is a neighbourhood of 0.</td></tr>
</table>

Figure 1.5: Complex charts on $X(\Gamma)$

1.4.3 Genus

Let Γ_1, Γ_2 be congruence subgroups of $\mathrm{SL}_2(\mathbb{Z})$ such that $\Gamma_1 \subset \Gamma_2$.

Theorem 1.26. *The natural projection of the corresponding modular curves*

$$F : X(\Gamma_1) \to X(\Gamma_2), \quad \Gamma_1\tau \mapsto \Gamma_2\tau,$$

is a surjective morphism of Riemann surfaces. Its degree is

$$\deg(F) = |\{\pm I\}\Gamma_2 : \{\pm I\}\Gamma_1| = \begin{cases} [\Gamma_2 : \Gamma_1]/2 & \text{if } -I \in \Gamma_2 \text{ and } -I \notin \Gamma_1, \\ [\Gamma_2 : \Gamma_1] & \text{otherwise.} \end{cases}$$

Proof: Consider the commutative diagram

$$\begin{tikzcd}
 & U \arrow[r, "id"] \arrow[dl, "\delta"'] \arrow[d, "\pi_1"'] & U \arrow[d, "\pi_2"] \arrow[dr, "\delta"] & \\
\delta(U) \arrow[dr, "\lambda_1"'] & \pi_1(U) \arrow[r, "F"] \arrow[d, "\phi_1"'] & \pi_2(U) \arrow[d, "\phi_2"] & \delta(U) \arrow[dl, "\lambda_2"] \\
 & V_1 \arrow[r, "F_{\text{local}}"] & V_2 &
\end{tikzcd}$$

where $\phi_j : \pi_j(U) \to V_j$ is a chart on $X(\Gamma_j)$, for $j = 1, 2$. Observe that

$$F_{\text{local}} \circ \psi_1 = \psi_2, \quad \text{where } \psi_j = \lambda_j \circ \delta.$$

Let $\delta = \delta_\tau$, with $\tau \in \mathbb{H}$. Then $\lambda_1(z) = z^{h_1}$, $\lambda_2(z) = z^{h_2}$ and the local map is

$$q \in V_1 \mapsto q^{h_2/h_1} \in V_2,$$

where $h_j = |\{\pm I\}\Gamma_{j,\tau}|/2$, for $j = 1, 2$. Further, as $h_j \in \{1, 2, 3\}$ and $h_2/h_1 \in \mathbb{Z}$, we deduce that the ramification index of F at $\pi_1(\tau)$ is

$$\begin{aligned} e_{\pi_1(\tau)}(F) = h_2/h_1 &= \begin{cases} h_2 & \text{if } \tau \text{ is an elliptic point of } \Gamma_2 \text{ and not of } \Gamma_1, \\ 1 & \text{otherwise,} \end{cases} \\ &= |\{\pm I\}\Gamma_{2,\tau} : \{\pm I\}\Gamma_{1,\tau}|. \end{aligned}$$

Let $\delta : s \mapsto \infty$, with $s \in \widehat{\mathbb{Q}}$. Then $\lambda_1(z) = e^{2\pi i z/h_1}$, $\lambda_2(z) = e^{2\pi i z/h_2}$ and the local map is

$$q \in V_1 \mapsto q^{h_1/h_2} \in V_2,$$

where $h_j = [\mathrm{SL}_2(\mathbb{Z}) : \{\pm I\}\Gamma_{j,s}]$, for $j = 1, 2$. Further, as $h_1/h_2 \in \mathbb{Z}$, we deduce that the ramification index of F at $\pi_1(s)$ is

$$e_{\pi_1(s)}(F) = h_1/h_2 = |\{\pm I\}\Gamma_{2,s} : \{\pm I\}\Gamma_{1,s}|.$$

As a consequence, $\pi_1(s)$ is a ramification point of F if and only if $h_1 > h_2$.

This proves that the natural projection is a morphism of Riemann surfaces. To compute its degree, let $\{\pm I\}\Gamma_2 = \bigcup_{j=1}^{d} \{\pm I\}\Gamma_1\gamma_j$, where γ_j are coset representatives. Then the inverse image of a nonelliptic point $\pi_2(\tau) \in X(\Gamma_2)$ is

$$F^{-1}(\pi_2(\tau)) = \{\pi_1(\gamma_1(\tau)), \ldots, \pi_1(\gamma_d(\tau))\}.$$

This shows that

$$\deg(F) = |\{\pm I\}\Gamma_2 : \{\pm I\}\Gamma_1|,$$

since

$$\deg(F) = \sum_{x \in F^{-1}(y)} e_x(F), \quad \forall\, y \in X(\Gamma_2).$$

□

To calculate the genus of $X(\Gamma)$, specialize to $\Gamma_1 = \Gamma$ and $\Gamma_2 = \mathrm{SL}_2(\mathbb{Z})$.

Theorem 1.27. *Let Γ be a congruence subgroup of $\mathrm{SL}_2(\mathbb{Z})$, $F : X(\Gamma) \to X(1)$ the natural projection and $d = \deg(F)$. Let e_2 and e_3 denote the number of elliptic points of period 2 and 3 of $X(\Gamma)$, and e_∞ the number of cusps of $X(\Gamma)$. Then*

$$2g - 2 = \frac{d}{6} - \frac{\varepsilon_2}{2} - \frac{2\varepsilon_3}{3} - \varepsilon_\infty.$$

where g is the genus of $X(\Gamma)$. As a consequence,

$$g = 1 + \frac{d}{12} - \frac{\varepsilon_2}{4} - \frac{\varepsilon_3}{3} - \frac{\varepsilon_\infty}{2}.$$

Proof: Let $y_2 = \mathrm{SL}_2(\mathbb{Z})i$, $y_3 = \mathrm{SL}_2(\mathbb{Z})\rho$ and $y_\infty = \mathrm{SL}_2(\mathbb{Z})\infty$ be the elliptic point of period 2, the elliptic point of period 3 and the cusp of $X(1)$. Since the elliptic points of period h of $X(\Gamma)$ are in $F^{-1}(y_h)$, for $h = 2, 3$,

$$d = \sum_{x \in F^{-1}(y_h)} e_x(F) = h \cdot \big(|f^{-1}(y_h)| - \varepsilon_h\big) + 1 \cdot \varepsilon_h.$$

Using these equalities twice we obtain that

$$\sum_{x \in F^{-1}(y_h)} (e_x(F) - 1) = (h-1) \cdot \big(|f^{-1}(y_h)| - \varepsilon_h\big) = \frac{h-1}{h}(d - \varepsilon_h).$$

Also,

$$\sum_{x \in F^{-1}(y_\infty)} (e_x(F) - 1) = d - \varepsilon_\infty.$$

Therefore, the Riemann-Hurwitz formula A.28 shows that

$$\begin{aligned} 2g - 2 &= -2d + \frac{1}{2}(d - \varepsilon_2) + \frac{2}{3}(d - \varepsilon_3) + (d - \varepsilon_\infty) \\ &= \frac{d}{6} - \frac{\varepsilon_2}{2} - \frac{2\varepsilon_3}{3} - \varepsilon_\infty, \quad \text{since the genus of } X(1) \text{ is equal to } 0. \end{aligned}$$

□

Chapter 2

Automorphic, modular and cusp forms

In this second chapter we introduce the $\mathbb{C}$-vector spaces of automorphic, modular and cusp forms. As mentioned at the introduction, modular forms play a special role in the proof of Fermat's Last Theorem. They are holomorphic functions on the upper half plane that satisfy certain transformation and holomorphy conditions. We comment on the dimension formulas of the spaces of modular and cusp forms, and we conclude with two interesting applications:

- Transformation law of the Dedekind eta function

- Four squares problem

2.1 Basic definitions

Let $f : \mathbb{H} \to \mathbb{C}$ be a function which is $\mathrm{SL}_2(\mathbb{Z})$-invariant, i.e.,

$$f(\tau) = f(\alpha(\tau)), \quad \forall\, \tau \in \mathbb{H}, \quad \forall\, \alpha \in \mathrm{SL}_2(\mathbb{Z}).$$

If f is holomorphic, then its derivative satisfies the functional equation

$$f'(\tau) = \frac{1}{(c\tau + d)^2} f'(\alpha(\tau)), \quad \forall\, \tau \in \mathbb{H}, \quad \forall\, \alpha = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathrm{SL}_2(\mathbb{Z}).$$

Raising both sides of this functional equation to a positive integer k, we have

$$(f'(\tau))^k = \frac{1}{(c\tau + d)^{2k}} (f'(\alpha(\tau)))^k, \quad \forall\, \tau \in \mathbb{H}, \quad \forall\, \alpha = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathrm{SL}_2(\mathbb{Z}).$$

Thus we conclude that $(f')^k$ is also $\mathrm{SL}_2(\mathbb{Z})$-invariant up to a factor that depends on the variable $\tau \in \mathbb{H}$ and on the matrix $\alpha \in \mathrm{SL}_2(\mathbb{Z})$, $1/(c\tau + d)^{2k}$.

Taking into account this argument, the definitions that we introduce below are logic and natural.

The factor of automorphy $j(\alpha,\cdot):\mathbb{H}\to\mathbb{C}$ associated to $\alpha=\begin{bmatrix} a & b\\ c & d\end{bmatrix}\in \mathrm{SL}_2(\mathbb{Z})$, is defined as

$$j(\alpha,\tau)=c\tau+d,\quad \forall\,\tau\in\mathbb{H}.$$

The following lemma states the basic properties of the factor of automorphy. Its proof is really easy and is left as an exercise to the reader.

Lemma 2.1. *Let $\alpha,\alpha'\in\mathrm{SL}_2(\mathbb{Z})$ and $\tau\in\mathbb{H}$. Then:*

1. $(\alpha\alpha')(\tau)=\alpha(\alpha'(\tau))$
2. $j(\alpha\alpha',\tau)=j(\alpha,\alpha'(\tau))j(\alpha',\tau)$
3. $\mathrm{Im}\,(\alpha(\tau))=\mathrm{Im}\,(\tau)/|j(\alpha,\tau)|^2$
4. $d\alpha(\tau)/d\tau=1/j(\alpha,\tau)^2$

Let $k\in\mathbb{Z}$. The weight-k operator associated to $\alpha=\begin{bmatrix} a & b\\ c & d\end{bmatrix}\in\mathrm{SL}_2(\mathbb{Z})$,

$$[\alpha]_k:\mathcal{M}(\mathbb{H})\to\mathcal{M}(\mathbb{H}),$$

is defined as

$$(f[\alpha]_k)(\tau)=j(\alpha,\tau)^{-k}f(\alpha(\tau)),\quad \forall\,\tau\in\mathbb{H},\quad \forall\,f\in\mathcal{M}(\mathbb{H}).$$

As $f(\alpha(\cdot))$ is a meromorphic function and $j(\alpha,\cdot)$ is a holomorphic function without zeros, we deduce that the weight-k operator is well-defined, i.e.,

$$j(\alpha,\cdot)^{-k}f(\alpha(\cdot))$$

is also a meromorphic function. Further, note that $[\alpha]_k$ is a linear operator,

$$(\lambda f+\beta g)[\alpha]_k=\lambda(f[\alpha]_k)+\beta(g[\alpha]_k),\quad \forall\,f,g\in\mathcal{M}(\mathbb{H}),\ \forall\,\lambda,\beta\in\mathbb{C}.$$

The chosen notation to denote the image of f under $[\alpha]_k$ is not usual, the maps are normally written on the left of the argument in mathematics. The reason why we write the weight-k operator on the right of the argument is justified in the following lemma.

Lemma 2.2. *Let $\alpha,\alpha'\in\mathrm{SL}_2(\mathbb{Z})$ and $k\in\mathbb{Z}$. Then*

$$[\alpha\alpha']_k=[\alpha]_k[\alpha']_k\ \ (\textit{equality of operators}).$$

Proof: Let $f \in \mathcal{M}(\mathbb{H})$. Then

$$\begin{aligned}(f[\alpha\alpha']_k)(\tau) &= j(\alpha\alpha',\tau)f((\alpha\alpha')(\tau))\\ &= j(\alpha',\tau)^{-k}j(\alpha,\alpha'(\tau))^{-k}f(\alpha(\alpha'(\tau)))\\ &= j(\alpha',\tau)^{-k}(f[\alpha]_k)(\alpha'(\tau)) = ((f[\alpha]_k)[\alpha']_k)(\tau), \quad \forall\, \tau \in \mathbb{H}.\end{aligned}$$

Definition 2.3. *Let Γ be a congruence subgroup of $\mathrm{SL}_2(\mathbb{Z})$ and $k \in \mathbb{Z}$. A meromorphic function $f : \mathbb{H} \to \widehat{\mathbb{C}}$ is weight-k invariant under Γ if*

$$f[\gamma]_k = f, \quad \forall\, \gamma \in \Gamma.$$

Remarks 2.4.

▷ If f is weight-k invariant under Γ, then f is $h\mathbb{Z}$-periodic, i.e.,

$$f(\tau) = f(\tau + mh), \quad \forall\, \tau \in \mathbb{H}, \quad \forall\, m \in \mathbb{Z},$$

where

$$h = h_\Gamma = \min\{h \in \mathbb{Z}^+ \mid \begin{bmatrix} 1 & h \\ 0 & 1 \end{bmatrix} \in \Gamma\}.$$

▷ If f is weight-k invariant under Γ, then its zeros and poles are Γ-invariants, since the factor of automorphy is a holomorphic function without zeros.

▷ If f is weight-k invariant under a generating set S of Γ,

$$f[\gamma]_k = f, \quad \forall\, \gamma \in S,$$

then f is also weight-k invariant under Γ. This is because

$$1 = j(\alpha, \alpha^{-1}(\tau))j(\alpha^{-1}, \tau), \quad \forall \tau \in \mathbb{H}, \quad \forall\, \alpha \in \mathrm{SL}_2(\mathbb{Z}),$$

and

$$[\alpha_1 \cdots \alpha_n]_k = [\alpha_1]_k \cdots [\alpha_n]_k, \quad \forall\, \alpha_1, \ldots, \alpha_n \in \mathrm{SL}_2(\mathbb{Z}).$$

▷ The functions that are weight-k invariant under Γ form a $\mathbb{C}$-vector space. If f, g are weight-k invariant under Γ and $\lambda, \beta \in \mathbb{C}$, then

$$\lambda f + \beta g \quad \text{is weight-}k\text{ invariant under }\Gamma.$$

▷ If f is weight-k invariant under Γ and g is weight-l invariant under Γ, then

- fg is weight-$(k + l)$ invariant under Γ, and
- f/g $(g \neq 0)$ is weight-$(k - l)$ invariant under Γ.

- ▷ If f is weight-k invariant under Γ, $-I \in \Gamma$ and k is odd, then f is the zero function, since
$$f(\tau) = (-1)^k f(\tau), \quad \forall\, \tau \in \mathbb{H}.$$

- ▷ Let $\alpha \in \mathrm{SL}_2(\mathbb{Z})$. By Lemma 1.7 $\alpha^{-1}\Gamma\alpha$ is also a congruence subgroup. If f is weight-k invariant under Γ, then the meromorphic function
$$(f[\alpha]_k)(\tau) = j(\alpha, \tau)^{-k} f(\alpha(\tau)), \quad \forall\, \tau \in \mathbb{H},$$
is weight-k invariant under $\alpha^{-1}\Gamma\alpha$. Indeed,
$$\begin{aligned} f[\alpha]_k[\alpha^{-1}\gamma\alpha]_k &= f[\alpha]_k[\alpha^{-1}]_k[\gamma]_k[\alpha]_k \\ &= f[\alpha^{-1}\alpha]_k[\gamma]_k[\alpha]_k = f[\gamma]_k[\alpha]_k = f[\alpha]_k, \quad \forall\, \gamma \in \Gamma. \end{aligned}$$

Let us now develop in detail the definition of automorphic form, modular form and cusp form of weight $k \in \mathbb{Z}$ with respect to a congruence subgroup Γ. Let $f : \mathbb{H} \to \widehat{\mathbb{C}}$ be a meromorphic function which is weight-k invariant under Γ, $h = h_\Gamma$ and $\dot{\mathbb{D}}$ the punctured unit disc,
$$\dot{\mathbb{D}} = \{z \in \mathbb{C} \mid 0 < |z| < 1\}.$$
As f is $h\mathbb{Z}$-periodic, the function $g : \dot{\mathbb{D}} \to \widehat{\mathbb{C}}$,
$$g(e^{2\pi i\tau/h}) = f(\tau), \quad \forall\, \tau \in \mathbb{H}$$
is well-defined. Then:

> We say that the function f is meromorphic (resp. holomorphic) at ∞ if the function g associated to f is meromorphic (resp. holomorphic) at 0.

Note that if f is meromorphic at ∞, then g has a Laurent series,
$$g(q_h) = \sum_{n\in\mathbb{Z}} a_n q_h^n, \quad \text{where } q_h = e^{2\pi i\tau/h},$$
which has finitely many nonzero negative terms. We refer to this series as the Fourier series of f. The order of f at ∞ is defined as
$$\nu_\infty(f) = \min\{n \in \mathbb{Z} \mid a_n \neq 0\},$$
except when $f = 0$, in which case $\nu_\infty(f) = +\infty$.

Let $\alpha \in \mathrm{SL}_2(\mathbb{Z})$. Then $\alpha^{-1}\Gamma\alpha$ is also a congruence subgroup and the meromorphic function
$$(f[\alpha]_k)(\tau) = j(\alpha, \tau)^{-k} f(\alpha(\tau)), \quad \forall\, \tau \in \mathbb{H},$$
is weight-k invariant under $\alpha^{-1}\Gamma\alpha$. Therefore, it makes sense to ask yourself if $f[\alpha]_k$ is meromorphic (resp. holomorphic) at ∞.

Definition 2.5. *Let Γ be a congruence subgroup of $\mathrm{SL}_2(\mathbb{Z})$ and $k \in \mathbb{Z}$. A meromorphic function $f : \mathbb{H} \to \widehat{\mathbb{C}}$ is an automorphic form of weight k with respect to Γ if*

- ▷ *f is weight-k invariant under Γ, and*
- ▷ *$f[\alpha]_k$ is meromorphic at ∞, for all $\alpha \in \mathrm{SL}_2(\mathbb{Z})$.*

The zero function is evidently an automorphic form of any weight $k \in \mathbb{Z}$ with respect to $\mathrm{SL}_2(\mathbb{Z})$. The constant functions are also automorphic forms of weight 0 with respect to $\mathrm{SL}_2(\mathbb{Z})$. In subsection 2.1.1 we present some nontrivial examples of automorphic forms.

The set of automorphic forms of weight k with respect to Γ is denoted $\mathcal{A}_k(\Gamma)$. It is a $\mathbb{C}$-vector space, since if $f, g \in \mathcal{A}_k(\Gamma)$ and $\lambda, \beta \in \mathbb{C}$, then $\alpha f + \beta g \in \mathcal{A}_k(\Gamma)$.

The second condition of the definition, $f[\alpha]_k$ *is meromorphic at* ∞, must be interpreted as a condition of meromorphy at the cusps $s = \alpha(\infty)$ of Γ. Note that it only needs to be checked for finitely many coset representatives α_j in any decomposition $\mathrm{SL}_2(\mathbb{Z}) = \bigcup_{j=1}^{d} \Gamma\alpha_j$, since

$$\alpha_j^{-1}\Gamma\alpha_j = (\gamma\alpha_j)^{-1}\Gamma(\gamma\alpha_j)$$

and

$$f[\gamma\alpha_j]_k = f[\gamma]_k[\alpha_j]_k = f[\alpha_j]_k, \quad \forall\, \gamma \in \Gamma.$$

Let $f : \mathbb{H} \to \widehat{\mathbb{C}}$ be an automorphic form of weight k with respect to Γ. The order of f at a cusp $s \in \widehat{\mathbb{Q}} = \mathbb{Q} \cup \{\infty\}$ is defined as

$$\nu_s(f) = \nu_\infty(f[\alpha]_k), \quad \text{where } \alpha \in \mathrm{SL}_2(\mathbb{Z}),\ \alpha(\infty) = s.$$

We have to prove that this definition is independent of the chosen matrix α.

Let h be the smallest positive integer such that $\left[\begin{smallmatrix} 1 & h \\ 0 & 1 \end{smallmatrix}\right] \in \alpha^{-1}\Gamma\alpha$. Consider the Fourier series of $f[\alpha]_k$,

$$(f[\alpha]_k)(\tau) = \sum_{n \in \mathbb{Z}} a_n q_h^n, \quad q_h = e^{2\pi i\tau/h}.$$

By definition

$$\nu_\infty(f[\alpha]_k) = \min\{n \in \mathbb{Z} \,|\, a_n \neq 0\}.$$

Furthermore, the matrices of $\mathrm{SL}_2(\mathbb{Z})$ that take ∞ to s are

$$\pm\alpha\beta, \quad \text{with } \beta = \begin{bmatrix} 1 & j \\ 0 & 1 \end{bmatrix},\ j \in \mathbb{Z},$$

and $(f[\pm\alpha\beta]_k)(\tau) = (\pm 1)^k (f[\alpha]_k)(\tau + j)$, $\forall\, \tau \in \mathbb{H}$. Therefore,

$$(f[\pm\alpha\beta]_k)(\tau) = (\pm 1)^k \sum_{n \in \mathbb{Z}} a_n \mu_h^{nj} q_h^n, \quad q_h = e^{2\pi i\tau/h},$$

where μ_h is the complex h-th root of the unity $e^{2\pi i/h}$ $(e^{2\pi i(\tau+j)/h} = \mu_h^j q_h)$.

Thus we can conclude that the definition of the order of f at the cusp s is independent of the chosen matrix $\alpha \in \mathrm{SL}_2(\mathbb{Z})$ with $\alpha(\infty) = s$.

Note that the order of f is also well-defined on the modular curve $X(\Gamma)$, i.e., if $s \in \widehat{\mathbb{Q}}$ and $\gamma \in \Gamma$, then

$$\nu_s(f) = \nu_{\gamma(s)}(f).$$

The modular forms are defined the same way as automorphic forms except with holomorphy in place of meromorphy.

Definition 2.6. *Let Γ be a congruence subgroup of $\mathrm{SL}_2(\mathbb{Z})$ and $k \in \mathbb{Z}$. A holomorphic function $f : \mathbb{H} \to \widehat{\mathbb{C}}$ is a modular form of weight k with respect to Γ if*

- ▷ *f is weight-k invariant under Γ, and*
- ▷ *$f[\alpha]_k$ is holomorphic at ∞, for all $\alpha \in \mathrm{SL}_2(\mathbb{Z})$.*

If in addition,

- ▷ *$a_0 = 0$ in the Fourier series of $f[\alpha]_k$, for all $\alpha \in \mathrm{SL}_2(\mathbb{Z})$,*

then f is a cusp form of weight k with respect to Γ.

The sets of modular and cusp forms of weight k with respect to Γ are denoted

$$\mathcal{M}_k(\Gamma) \quad \text{and} \quad \mathcal{S}_k(\Gamma),$$

respectively. Note that both sets are vector subspaces of $\mathcal{A}_k(\Gamma)$, since

$$\mathcal{M}_k(\Gamma) = \{f \in \mathcal{A}_k(\Gamma) \,|\, f \text{ is holomorphic and } \nu_\infty(f[\alpha]_k) \geq 0, \forall\, \alpha \in \mathrm{SL}_2(\mathbb{Z})\}$$

and

$$\mathcal{S}_k(\Gamma) = \{f \in \mathcal{A}_k(\Gamma) \,|\, f \text{ is holomorphic and } \nu_\infty(f[\alpha]_k) \geq 1, \forall\, \alpha \in \mathrm{SL}_2(\mathbb{Z})\}.$$

As a consequence,

$$\mathcal{S}_k(\Gamma) \subset \mathcal{M}_k(\Gamma) \subset \mathcal{A}_k(\Gamma) \subset \mathcal{M}(\mathbb{H}).$$

Remarks 2.7.

- ▷ Let Γ_1, Γ_2 be congruence subgroups of $\mathrm{SL}_2(\mathbb{Z})$ such that $\Gamma_1 \subset \Gamma_2$. Then

$$\mathcal{S}_k(\Gamma_2) \subset \mathcal{S}_k(\Gamma_1), \quad \mathcal{M}_k(\Gamma_2) \subset \mathcal{M}_k(\Gamma_1) \quad \text{and} \quad \mathcal{A}_k(\Gamma_2) \subset \mathcal{A}_k(\Gamma_1).$$

- ▷ Let $\alpha \in \mathrm{SL}_2(\mathbb{Z})$. The weight-$k$ operator associated to the matrix α induces the following isomorphisms of $\mathbb{C}$-vector spaces:

$$\mathcal{A}_k(\Gamma) \cong \mathcal{A}_k(\alpha^{-1}\Gamma\alpha) \quad \mathcal{M}_k(\Gamma) \cong \mathcal{M}_k(\alpha^{-1}\Gamma\alpha) \quad \mathcal{S}_k(\Gamma) \cong \mathcal{S}_k(\alpha^{-1}\Gamma\alpha)$$

- ▷ If $f \in \mathcal{A}_k(\Gamma)$ (resp. $\mathcal{M}_k(\Gamma)$ or $\mathcal{S}_k(\Gamma)$) and $g \in \mathcal{A}_l(\Gamma)$ (resp. $\mathcal{M}_l(\Gamma)$ or $\mathcal{S}_l(\Gamma)$), then $fg \in \mathcal{A}_{k+l}(\Gamma)$ (resp. $\mathcal{M}_{k+l}(\Gamma)$ or $\mathcal{S}_{k+l}(\Gamma)$). Indeed,

  - fg is weight-$(k+l)$ invariant under Γ, and
  - the Fourier series of $(fg)[\alpha]_{(k+l)}$, with $\alpha \in \mathrm{SL}_2(\mathbb{Z})$, is

$$(fg)[\alpha]_{(k+l)} = \sum_{n\in\mathbb{Z}} (\sum_{s+j=n} a_s b_j) q_h^n, \quad q_h = e^{2\pi i\tau/h},$$

  where

$$f[\alpha]_k = \sum_{s\in\mathbb{Z}} a_s q_h^s, \quad g[\alpha]_l = \sum_{j\in\mathbb{Z}} b_j q_h^j \quad \text{and} \quad h = h_{(\alpha^{-1}\Gamma\alpha)}.$$

  Thus the direct sums

$$\mathcal{A}(\Gamma) = \bigoplus_{k\in\mathbb{Z}} \mathcal{A}_k(\Gamma), \quad \mathcal{M}(\Gamma) = \bigoplus_{k\in\mathbb{Z}} \mathcal{M}_k(\Gamma) \quad \text{and} \quad \mathcal{S}(\Gamma) = \bigoplus_{k\in\mathbb{Z}} \mathcal{S}_k(\Gamma)$$

  forms graded rings.

- ▷ Let $\mathcal{R} = \mathcal{A}, \mathcal{M}$ or $\mathcal{S}$. For each $\alpha \in \mathrm{SL}_2(\mathbb{Z})$, the map

$$\sum_{k\in\mathbb{Z}} f_k \in \mathcal{R}(\Gamma) \mapsto \sum_{k\in\mathbb{Z}} f_k[\alpha]_k \in \mathcal{R}(\alpha^{-1}\Gamma\alpha)$$

  is an isomorphism of graded rings.

- ▷ If $f \in \mathcal{M}_k(\Gamma)$ and $g \in \mathcal{S}_l(\Gamma)$, then $fg \in \mathcal{S}_{k+l}(\Gamma)$. As a consequence, observe that $\mathcal{S}(\Gamma)$ is a graded ideal of $\mathcal{M}(\Gamma)$, since

$$\mathcal{M}(\Gamma)\mathcal{S}(\Gamma) \subset \mathcal{S}(\Gamma) \quad \text{and} \quad \mathcal{S}(\Gamma) = \bigoplus_{k\in\mathbb{Z}} (\mathcal{M}_k(\Gamma) \cap \mathcal{S}(\Gamma)).$$

- ▷ If $f \in \mathcal{A}_k(\Gamma)$ and $g \in \mathcal{A}_l(\Gamma)$ $(g \neq 0)$, then $f/g \in \mathcal{A}_{k-l}(\Gamma)$. Indeed,

$$1/g \in \mathcal{A}_{-l}(\Gamma) \quad \text{and} \quad f/g = f(1/g) \in \mathcal{A}_{k-l}(\Gamma).$$

  In particular, note that

$$\mathcal{A}_k(\Gamma) = f\mathcal{A}_0(\Gamma) = \{ff_0 \mid f_0 \in \mathcal{A}_0(\Gamma)\}, \quad \text{whenever } f \neq 0.$$

▷ If $f, g \in \mathcal{A}_0(\Gamma)$ and $\lambda, \beta \in \mathbb{C}$, then $\lambda f + \beta g$, fg, $1/f\,(f \neq 0) \in \mathcal{A}_0(\Gamma)$. This implies that $\mathcal{A}_0(\Gamma)$ is a $\mathbb{C}$-algebra that presents structure of field. Moreover, the map

$$f \in \mathcal{A}_0(\Gamma) \mapsto F \in \mathcal{M}(X(\Gamma)), \quad F(x) = \begin{cases} f(\tau) & \text{if } x = \Gamma\tau, \\ f(s) & \text{if } x = \Gamma s, \end{cases}$$

defines an isomorphism of $\mathbb{C}$-algebras.

▷ If $-I \in \Gamma$ and k is odd, then $\mathcal{S}_k(\Gamma) = \mathcal{M}_k(\Gamma) = \mathcal{A}_k(\Gamma) = \{0\}$, since the only function that is weigh-k invariant under Γ is the zero function.

Theorem 2.8. *Let Γ be a congruence subgroup of $\mathrm{SL}_2(\mathbb{Z})$ of level N, $k \in \mathbb{Z}$ and $f : \mathbb{H} \to \mathbb{C}$ a holomorphic function. If f is weight-k invariant under Γ, f is holomorphic at ∞,*

$$f(\tau) = \sum_{n=0}^{\infty} a_n q_N^n, \quad q_N = e^{2\pi i\tau/N}, \quad \forall\, \tau \in \mathbb{H},$$

and in addition, there exist positive constants C and r such that

$$|a_n| \leq Cn^r, \quad \forall\, n \geq 1,$$

then $f[\alpha]_k$ is holomorphic at ∞, $\forall\, \alpha \in \mathrm{SL}_2(\mathbb{Z})$. As a consequence, $f \in \mathcal{M}_k(\Gamma)$.

Proof: Let $\alpha \in \mathrm{SL}_2(\mathbb{Z})$. The function

$$(f[\alpha]_k)(\tau) = j(\alpha, \tau)^{-k} f(\alpha(\tau)), \quad \forall\, \tau \in \mathbb{H},$$

is holomorphic and weight-k invariant under $\alpha^{-1}\Gamma\alpha$, so it has an expansion

$$(f[\alpha]_k)(\tau) = \sum_{n \in \mathbb{Z}} a'_n q_N^n, \quad \forall\, \tau \in \mathbb{H}.$$

Therefore, to prove that $f[\alpha]_k$ is holomorphic at ∞, it suffices to see that

$$\lim_{q_N \to 0} ((f[\alpha]_k)(\tau) \cdot q_N) = 0. \tag{2.1}$$

Let us suppose that we have proved that there exist constants $C_0, C_1 > 0$ satisfying the following property:

$$|f(\tau)| \leq C_0 + C_1/y^r, \quad \text{as } y \to \infty \quad (\tau = x + iy \in \mathbb{H}). \tag{2.2}$$

If $\alpha(\infty) = \infty$, then (2.1) is immediate, since $\alpha = \pm\left[\begin{smallmatrix} 1 & m \\ 0 & 1 \end{smallmatrix}\right]$, with $m \in \mathbb{Z}$, and

$$(f[\alpha]_k(\tau)) = (\pm 1)^k f(\tau + m) = (\pm 1)^k \mu_N^m \sum_{n=0}^{\infty} a_n q_N^n, \quad \forall\, \tau \in \mathbb{H},$$

where $\mu_N = e^{2\pi i/N}$. Otherwise, we have $c \neq 0$ and

$$\begin{aligned}|(f[\alpha]_k)(\tau)| &= |f(\alpha(\tau))||c\tau+d|^{-k}\\ &\leq \left(C_0 + C_1 \mathrm{Im}\,(\alpha(\tau))^{-r}\right)|c\tau+d|^{-k}\\ &= \left(C_0 + C_1|c\tau+d|^{2r}y^{-r}\right)|c\tau+d|^{-k}, \quad \text{as } y \to \infty,\end{aligned}$$

where c and d are the lower entries of the matrix α. If we assume that

$$0 \leq x \leq N \quad ((f[\alpha]_k)(\tau) = (f[\alpha]_k)(\tau+N)),$$

then $|c\tau+d|$ grows as y, and as a consequence, there exists a constant $C_2 > 0$ such that

$$|(f[\alpha]_k)(\tau)| \leq C_2 y^{r-k}, \quad \text{as } y \to \infty.$$

Using this estimation we obtain (2.1), since $y = (N/2\pi)\log(1/|q_N|)$ and

$$|(f[\alpha]_k)(\tau)\cdot q_N| \leq C_2(N/2\pi)^{r-k}\log(1/|q_N|)^{r-k}|q_N| \to 0, \quad \text{as } q_N \to 0.$$

To conclude this proof we need to prove that there exist constants $C_0, C_1 > 0$ satisfying the property (2.2). By hypothesis,

$$|f(\tau)| \leq |a_0| + C\sum_{n=1}^{\infty} n^r e^{-2\pi ny/N}, \quad \forall\, \tau = x+iy \in \mathbb{H}.$$

Let $g_y : \mathbb{R} \to \mathbb{R}$ be the function

$$g_y(t) = t^r e^{-2\pi ty/N}, \quad \forall\, t \in \mathbb{R}.$$

As its derivative is

$$g_y'(t) = \left(rt^{r-1} - \frac{2\pi y}{N}t^r\right)e^{-2\pi ty/N}, \quad \forall\, t \in \mathbb{R},$$

we deduce that this function decreases monotonically on the interval $[\frac{rN}{2\pi y}, \infty)$. Therefore,

$$\begin{aligned}|f(\tau)| &\leq |a_0| + C\Big(g_y(1) + \sum_{n=2}^{\infty} g_y(n)\Big)\\ &\leq |a_0| + C\Big(1 + \sum_{n=2}^{\infty}\int_{n-1}^{n} g_y(t)\,dt\Big) \quad [rN/2\pi < y]\\ &\leq |a_0| + C + C\int_0^{\infty} t^r e^{-2\pi ty/N}\,dt \quad [t = 2\pi ty/N]\\ &= |a_0| + C + C(N/2\pi)^{r+1}1/y^{r+1}\int_0^{\infty} t^r e^{-t}\,dt, \quad \text{as } y \to \infty.\end{aligned}$$

Letting

$$C_0 = |a_0| + C \quad \text{and} \quad C_1 = C(N/2\pi)^{r+1} \int_0^\infty t^r e^{-t}\, dt,$$

we obtain that

$$|f(\tau)| \leq C_0 + C_1/y^r, \quad \text{as } y \to \infty \quad (\tau = x + iy \in \mathbb{H}).$$

□

The condition of holomorphy at the cusps, $f[\alpha]_k$ *is holomorphic at* ∞, keeps the vector spaces of modular and cusp forms finite-dimensional.

Theorem 2.9. *Let k be an even integer. Let Γ be a congruence subgroup of $\mathrm{SL}_2(\mathbb{Z})$, g the genus of $X(\Gamma)$, ε_2 the number of elliptic points with period 2, ε_3 the number of elliptic points with period 3 and ε_∞ the number of cusps. Then*

$$\dim(\mathcal{M}_k(\Gamma)) = \begin{cases} (k-1)(g-1) + \left\lfloor \frac{k}{4} \right\rfloor \varepsilon_2 + \left\lfloor \frac{k}{3} \right\rfloor \varepsilon_3 + \varepsilon_\infty, & \text{if } k \geq 2, \\ 1, & \text{if } k = 0, \\ 0, & \text{if } k < 0, \end{cases}$$

and

$$\dim(\mathcal{S}_k(\Gamma)) = \begin{cases} (k-1)(g-1) + \left\lfloor \frac{k}{4} \right\rfloor \varepsilon_2 + \left\lfloor \frac{k}{3} \right\rfloor \varepsilon_3 + (\frac{k}{2} - 1)\varepsilon_\infty, & \text{if } k \geq 4, \\ g, & \text{if } k = 2, \\ 0, & \text{if } k \leq 0. \end{cases}$$

Proof: [DS05, p.86-88]

□

Let $s \in \widehat{\mathbb{Q}}$ and $\alpha \in \mathrm{SL}_2(\mathbb{Z})$ taking ∞ to s. Recall that the width $h \in \mathbb{Z}^+$ of s with respect to a congruence subgroup Γ satisfies the condition

$$\{\pm I\}(\Gamma_\alpha)_\infty = \{\pm I\} \left\langle \begin{bmatrix} 1 & h \\ 0 & 1 \end{bmatrix} \right\rangle, \quad \text{where } \Gamma_\alpha = \alpha^{-1}\Gamma\alpha.$$

Therefore, this implies that

$$(\Gamma_\alpha)_\infty = \{\pm I\} \left\langle \begin{bmatrix} 1 & h \\ 0 & 1 \end{bmatrix} \right\rangle, \ (\Gamma_\alpha)_\infty = \left\langle \begin{bmatrix} 1 & h \\ 0 & 1 \end{bmatrix} \right\rangle \ \text{or} \ (\Gamma_\alpha)_\infty = \left\langle - \begin{bmatrix} 1 & h \\ 0 & 1 \end{bmatrix} \right\rangle,$$

since the negative identity matrix $-I$ might not be in the subgroup $\alpha^{-1}\Gamma\alpha$.

The cusp $\pi(s) \in X(\Gamma)$ is called a *regular cusp* of Γ if

$$(\Gamma_\alpha)_\infty = \{\pm I\} \left\langle \begin{bmatrix} 1 & h \\ 0 & 1 \end{bmatrix} \right\rangle \quad \text{or} \quad (\Gamma_\alpha)_\infty = \left\langle \begin{bmatrix} 1 & h \\ 0 & 1 \end{bmatrix} \right\rangle.$$

Otherwise, $\pi(s)$ is called *irregular* cusp of Γ,

$$(\Gamma_\alpha)_\infty = \left\langle - \begin{bmatrix} 1 & h \\ 0 & 1 \end{bmatrix} \right\rangle.$$

Note that this definition is independent of the chosen matrix $\alpha \in \mathrm{SL}_2(\mathbb{Z})$.

Theorem 2.10. *Let k be an odd integer. Let Γ be a congruence subgroup of $\mathrm{SL}_2(\mathbb{Z})$, g the genus of $X(\Gamma)$, ε_3 the number of elliptic points with period 3, ε_∞^{reg} the number of regular cusps and ε_∞^{irr} the number of irregular cusps. If $-I \notin \Gamma$, then*

$$\dim(\mathcal{M}_k(\Gamma)) = \begin{cases} (k-1)(g-1) + \left\lfloor \frac{k}{3} \right\rfloor \varepsilon_3 + \frac{k}{2}\varepsilon_\infty^{reg} + \frac{k-1}{2}\varepsilon_\infty^{irr}, & \text{if } k \geq 3, \\ 0, & \text{if } k < 0, \end{cases}$$

and

$$\dim(\mathcal{S}_k(\Gamma)) = \begin{cases} (k-1)(g-1) + \left\lfloor \frac{k}{3} \right\rfloor \varepsilon_3 + \frac{k-2}{2}\varepsilon_\infty^{reg} + \frac{k-1}{2}\varepsilon_\infty^{irr}, & \text{if } k \geq 3, \\ 0, & \text{if } k < 0. \end{cases}$$

If $\varepsilon_\infty^{reg} > 2g-2$, then $\dim(\mathcal{M}_1(\Gamma)) = \varepsilon_\infty^{reg}/2$ and $\dim(\mathcal{S}_1(\Gamma)) = 0$. If $\varepsilon_\infty^{reg} \leq 2g-2$, then $\dim(\mathcal{M}_1(\Gamma)) \geq \varepsilon_\infty^{reg}/2$ and $\dim(\mathcal{S}_1(\Gamma)) = \dim(\mathcal{M}_1(\Gamma)) - \varepsilon_\infty^{reg}/2$.

Proof: [DS05, p.90-91]

□

The demonstration of these *dimension formulas* uses the Riemann-Roch Theorem A.44. Specifically, the following consequence:

> Let X be a compact Riemann surface of genus g. If D is a divisor on X such that $\deg(D) > 2g-2$, then
>
> $$\dim L(D) = \deg(D) - g + 1.$$

We introduce some applications of these dimension formulas in Section 2.2.

Remark 2.11. Observe that we have not given dimension formulas for the vector spaces

$$\mathcal{M}_1(\Gamma) \quad \text{and} \quad \mathcal{S}_1(\Gamma), \quad \text{when } \varepsilon_\infty^{reg} \leq 2g-2.$$

Determining dimension formulas for these spaces is an open problem.

2.1.1 Eisenstein series for $\mathrm{SL}_2(\mathbb{Z})$

Let $k \geq 3$ be an even integer. The Eisenstein series of weight k for $\mathrm{SL}_2(\mathbb{Z})$ is defined as

$$G_k(\tau) = \sum_{(c,d)}{}' \frac{1}{(c\tau+d)^k}, \quad \forall\, \tau \in \mathbb{H},$$

where the primed summation means to sum over the pairs $(c,d) \in \mathbb{Z}^2 \backslash \{(0,0)\}$. If we let Λ_τ denote the lattice

$$\tau\mathbb{Z} + \mathbb{Z} = \{c\tau + d \,|\, c,d \in \mathbb{Z}\}, \quad \tau \in \mathbb{H},$$

then another expression for the Eisenstein series of weight k is

$$G_k(\tau) = \sum_{w\in\Lambda_\tau}{}' \frac{1}{w^k}, \quad \forall\, \tau \in \mathbb{H},$$

where the primed summation means now to sum over the points $w \in \Lambda_\tau \backslash \{0\}$.

The following lemma is an important technical result. We will use it to prove that the Eisenstein series define holomorphic functions on the upper half plane.

Lemma 2.12. *Let Λ be a lattice in $\mathbb{C}$ and $r > 0$. Then the series*

$$\sum_{w\in\Lambda}{}' \frac{1}{|w|^r} < \infty \quad \textit{if and only if} \quad r > 2.$$

Proof: Let (w_1, w_2) be a basis of Λ. For each $k \in \mathbb{Z}^+$, define

$$A_k = \{mw_1 + nw_2 \,|\, (m,n) \in \mathbb{Z}^2, |m| + |n| = k\} \quad \text{and}$$
$$S_k = \{xw_1 + yw_2 \,|\, (x,y) \in \mathbb{R}^2, |x| + |y| = k\}.$$

Then

$$A_k \subset S_k, \quad S_k = kS_1 \quad \text{and} \quad |A_k| = 4k, \quad \forall\, k \in \mathbb{Z}^+.$$

Using this, we obtain that

$$\sum_{w\in\Lambda}{}' \frac{1}{|w|^r} = \sum_{k=1}^{\infty} \sum_{w\in A_k} \frac{1}{|w|^r}, \tag{2.3}$$

$$\frac{4}{C^r} \sum_{k=1}^{\infty} \frac{1}{k^{r-1}} \leq \sum_{k=1}^{\infty} \sum_{w\in A_k} \frac{1}{|w|^r} \leq \frac{4}{c^r} \sum_{k=1}^{\infty} \frac{1}{k^{r-1}}, \tag{2.4}$$

where

$$C = \max_{z\in S_1} |z| \qquad \text{and} \qquad c = \min_{z\in S_1} |z|.$$

As the series of positive terms

$$\sum_{k=1}^{\infty} \frac{1}{k^{r-1}} < \infty \quad \text{if and only if} \quad r > 2,$$

the proof of the lemma is an immediate consequence of (2.3) and (2.4). □

Theorem 2.13. *The Eisenstein series*

$$G_k(\tau) = \sum_{(c,d)}{}' \frac{1}{(c\tau + d)^k}, \quad \textit{with } k \geq 3,$$

defines a holomorphic function on the upper half plane.

Proof: We show that the series converges uniformly on the subset

$$\Omega(a,b) = \{\tau \in \mathbb{H} \mid |\operatorname{Re}\tau| \leq a, \operatorname{Im}\tau \geq b\}, \quad a, b > 0.$$

Let $\tau \in \mathbb{H}$, $\tau = x + iy$, and $c, d \in \mathbb{Z}$. Then

$$|c\tau + d|^2 = c^2x^2 + d^2 + 2cdx + c^2y^2.$$

Choose $\delta > 0$ such that $\frac{a^2}{a^2+b^2} < \delta^2 < 1$ and rewrite the anterior equality as

$$|c\tau + d|^2 = \left[y^2 + \left(1 - \frac{1}{\delta^2}\right)x^2\right]c^2 + \left(\delta d + \frac{cx}{\delta}\right)^2 + (1-\delta^2)d^2.$$

If $\tau \in \Omega(a,b)$ (i.e., $|x| \leq a$ and $y \geq b$), we obtain that

$$\begin{aligned} |c\tau + d|^2 &\geq \left[y^2 + \left(1 - \frac{1}{\delta^2}\right)x^2\right]c^2 + (1-\delta^2)d^2 \\ &\geq \left[b^2 + \left(1 - \frac{1}{\delta^2}\right)a^2\right]c^2 + (1-\delta^2)d^2. \end{aligned}$$

As the coefficients of c^2 and d^2 are both positive, there exists $\varepsilon > 0$ such that

$$|c\tau + d|^2 \geq \varepsilon^2(c^2 + d^2), \quad \forall\, c, d \in \mathbb{Z},$$

or equivalently,

$$\frac{1}{|c\tau + d|} \leq \frac{1}{\varepsilon}\frac{1}{|ci + d|}, \quad \forall\,(c,d) \in \mathbb{Z}^2, (c,d) \neq (0,0). \tag{2.5}$$

Furthermore, by Lemma 2.12

$$\sum_{(c,d)}{}' \frac{1}{|ci + d|^k} < \infty,$$

so it suffices to apply the Weierstrass M-test to deduce that the series

$$\sum_{(c,d)}{}' \frac{1}{(c\tau + d)^k}$$

converges uniformly on the subset $\Omega(a,b)$.

□

We only have defined Eisenstein series of weight $k \geq 3$ even. If $k \geq 3$ is odd, then G_k is also a holomorphic function, but in this case it is the zero function, since the terms corresponding to $(c,d), -(c,d) \in \mathbb{Z}^2\backslash\{(0,0)\}$,

$$\frac{1}{(c\tau + d)^k} \quad \text{and} \quad \frac{1}{(-1)^k(c\tau + d)^k},$$

cancel out - are cancelled.

Below, we show that G_k is a modular form of weight k with respect to $\mathrm{SL}_2(\mathbb{Z})$. Let $\alpha = \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right] \in \mathrm{SL}_2(\mathbb{Z})$. Then

$$\begin{aligned} G_k(\alpha(\tau)) &= \sum_{(c',d')}{}' \frac{1}{(c'(\alpha(\tau)) + d')^k} \\ &= (c\tau + d)^k \sum_{(c',d')}{}' \frac{1}{((c'a + d'c)(\tau) + (c'b + d'd))^k}, \quad \forall \, \tau \in \mathbb{H}. \end{aligned}$$

Furthermore, as the map $(c', d') \mapsto (c', d')\left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right] = (c'a + d'c, c'b + d'd)$ is a bijection from $\mathbb{Z}^2 \setminus \{(0, 0)\}$ to itself, we deduce that the right side of this equality is $(c\tau+d)^k G_k(\tau)$. Therefore, G_k is weight-k invariant under $\mathrm{SL}_2(\mathbb{Z})$,

$$G_k[\alpha]_k = G_k, \quad \forall \, \alpha \in \mathrm{SL}_2(\mathbb{Z}).$$

To compute the Fourier series of G_k, we use the following two identities of the cotangent function:

$$\pi \cot(\pi z) = \frac{1}{z} + \sum_{d=1}^{\infty} \left(\frac{1}{z - d} + \frac{1}{z + d} \right), \quad \forall \, z \in \mathbb{C} \setminus \mathbb{Z} \tag{2.6}$$

$$\pi \cot(\pi \tau) = \pi i \frac{q + 1}{q - 1} = \pi i - 2\pi i \sum_{m=0}^{\infty} q^m, \quad q = e^{2\pi i \tau}, \ \forall \, \tau \in \mathbb{H} \tag{2.7}$$

The first identity is a partial fraction decomposition of the meromorphic function $\pi \cot(\pi z)$. We omit its proof since it is not trivial. It can be consulted in [SS03, p.142]. The second identity follows from the expansion

$$\frac{1}{q - 1} = -\sum_{m=0}^{\infty} q^m, \quad \forall \, q \in \mathbb{C}, \, |q| < 1.$$

Equating these two expressions for $\pi \cot(\pi \tau)$ and differentiating $k-1$ times with respect to τ, we obtain that

$$\sum_{d \in \mathbb{Z}} \frac{1}{(\tau + d)^k} = \frac{(-2\pi i)^k}{(k - 1)!} \sum_{m=1}^{\infty} m^{k-1} q^m, \quad \forall \, \tau \in \mathbb{H}, \quad k \geq 2. \tag{2.8}$$

For $k \geq 3$ even,

$$\begin{aligned} \sum_{(c,d)}{}' \frac{1}{(c\tau + d)^k} &= \sum_{d \neq 0} \frac{1}{d^k} + 2 \sum_{c=1}^{\infty} \left(\sum_{d \in \mathbb{Z}} \frac{1}{(c\tau + d)^k} \right) \\ &= 2\zeta(k) + 2 \frac{(2\pi i)^k}{(k - 1)!} \sum_{c=1}^{\infty} \sum_{m=1}^{\infty} m^{k-1} q^{cm}, \quad \forall \, \tau \in \mathbb{H}, \end{aligned}$$

where ζ denotes the Riemann zeta function,

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}, \quad \forall\, s \in \mathbb{C},\, \mathrm{Re}\,(s) > 1.$$

Rearranging the terms of the last double series gives the Fourier series,

$$G_k(\tau) = 2\zeta(k) + 2\frac{(2\pi i)^k}{(k-1)!} \sum_{n=1}^{\infty} \sigma_{k-1}(n) q^n, \quad \forall\, \tau \in \mathbb{H}, \tag{2.9}$$

where the coefficients of the series are the values of the arithmetic function

$$\sigma_{k-1}(n) = \sum_{\substack{m|n \\ m>0}} m^{k-1}, \quad \forall\, n \geq 1.$$

This proves that G_k is a modular form of weight k with respect to $\mathrm{SL}_2(\mathbb{Z})$.

Remark 2.14. For each $0 < r < 1$, observe that

$$\begin{aligned} \sum_{c=1}^{\infty} \sum_{m=1}^{\infty} m^{k-1} |q|^{cm} &= \sum_{m=1}^{\infty} m^{k-1} \frac{|q|^m}{1-|q|^m} \\ &\leq \frac{1}{1-r} \sum_{m=1}^{\infty} m^{k-1} r^m < \infty, \quad \forall\, q \in r\mathbb{D}. \end{aligned}$$

Therefore, the doubles series

$$\sum_{c=1}^{\infty} \sum_{m=1}^{\infty} m^{k-1} q^{cm}, \quad q \in \mathbb{D}, \quad k \in \mathbb{Z},$$

converges uniformly on compact subsets of the unit disc.

Remark 2.15. There is another way of proving that G_k is holomorphic at ∞ without having to compute its Fourier series. Let $a = 1/2$ and $b = 1$. By the proof of Theorem 2.13 there exists a positive constant C such that

$$|G_k(\tau)| \leq \sum_{(c,d)}{}' \frac{1}{|c\tau + d|^k} \leq C \sum_{(c,d)}{}' \frac{1}{|ci + d|^k}, \quad \forall\, \tau \in \Omega(a, b),$$

and since G_k is $\mathbb{Z}$-periodic

$$|G_k(\tau)| \leq C \sum_{(c,d)}{}' \frac{1}{|ci + d|^k}, \quad \forall\, \tau = x + iy \in \mathbb{H},\, y \geq 1.$$

Hence G_k is bounded as $\mathrm{Im}\,(\tau) \to \infty$, and the function g associated to G_k,

$$g(q) = G_k(\tau), \quad q = e^{2\pi i \tau}, \quad \forall\, \tau \in \mathbb{H},$$

has a removable singularity at $q = 0$.

For an nontrivial example of cusp form of weight 12 with respect to $\mathrm{SL}_2(\mathbb{Z})$, let $g_2, g_3 : \mathbb{H} \to \mathbb{C}$ be the functions

$$g_2(\tau) = 60G_4(\tau), \quad g_3(\tau) = 140G_6(\tau), \quad \forall\, \tau \in \mathbb{H}.$$

and define the discriminant function

$$\Delta : \mathbb{H} \to \mathbb{C}, \quad \Delta(z) = g_2(\tau)^3 - 27g_3(\tau)^2, \quad \forall\, \tau \in \mathbb{H}$$

If we let $\wp_\tau$ denote the Weierstrass $\wp$-function for the lattice $\Lambda_\tau = \tau\mathbb{Z} + \mathbb{Z}$, then the nonsingular cubic equation satisfied by $\wp_\tau$ and $\wp'_\tau$ is

$$(\wp'_\tau)^2 = 4\wp_\tau^3 - g_2(\tau)\wp_\tau - g_3(\tau),$$

since

$$g_2(\tau) = 60 \sum_{w\in\Lambda_\tau}{}' \frac{1}{w^4}, \quad g_3(\tau) = 140 \sum_{w\in\Lambda_\tau}{}' \frac{1}{w^6}, \quad \forall\, \tau \in \mathbb{H}.$$

Therefore, the discriminant function is nonvanishing on $\mathbb{H}$ by Theorem B.7,

$$\Delta(\tau) \neq 0, \quad \forall\, \tau \in \mathbb{H}.$$

As $g_2 \in \mathcal{M}_4(\mathrm{SL}_2(\mathbb{Z}))$ and $g_3 \in \mathcal{M}_6(\mathrm{SL}_2(\mathbb{Z}))$, we deduce that $\Delta \in \mathcal{M}_{12}(\mathrm{SL}_2(\mathbb{Z}))$. Furthermore, using the expansions (2.9) and the identities

$$\zeta(4) = \sum_{n=1}^{\infty} \frac{1}{n^4} = \frac{\pi^4}{90} \quad \text{and} \quad \zeta(6) = \sum_{n=1}^{\infty} \frac{1}{n^6} = \frac{\pi^6}{945}, \tag{2.10}$$

we obtain that

$$\Delta(\tau) = \pi^{12}(2^{12}q + \cdots), \quad q = e^{2\pi i\tau}, \quad \forall\, \tau \in \mathbb{H}.$$

Thus we conclude that Δ is cusp form of weight 12 with respect to $\mathrm{SL}_2(\mathbb{Z})$.

Remark 2.16. The identities in (2.10) are well-known. They can be derived from (2.6) by taking Laurent series around $z = 0$ to both sides and equating the coefficients of z^3 and z^5.

The modular function $j : \mathbb{H} \to \mathbb{C}$ is defined as

$$j(\tau) = 1728 \frac{(g_2(\tau))^3}{\Delta(\tau)}, \quad \forall\, \tau \in \mathbb{H}.$$

Observe that it is a holomorphic function since the discriminant function Δ does not have zeros in $\mathbb{H}$. Furthermore, as the numerator and denominator are modular forms of weight 12 with respect to $\mathrm{SL}_2(\mathbb{Z})$, we deduce that the function j is an automorphic form of weight 0 with respect to $\mathrm{SL}_2(\mathbb{Z})$.

The coefficient 1728 normalizes its Fourier series to

$$j(\tau) = \frac{(2\pi)^{12} + \cdots}{(2\pi)^{12}q + \cdots} = \frac{1}{q} + \sum_{n=0}^{\infty} a_n q^n, \quad q = e^{2\pi i \tau}, \quad \forall \tau \in \mathbb{H}. \tag{2.11}$$

It is possible to prove that the coefficients a_n are integer numbers [Apo90, p.21].

The following theorem states that any automorphic form of weight 0 with respect to $\mathrm{SL}_2(\mathbb{Z})$ is a rational expression in the modular function j, i.e.,

$$\mathcal{A}_0(\mathrm{SL}_2(\mathbb{Z})) = \mathbb{C}(j).$$

Theorem 2.17. *The modular function j generates the field of meromorphic functions on the modular curve $X(\mathrm{SL}_2(\mathbb{Z}))$.*

Before proving this theorem, observe that as a consequence of (2.11), the modular function $j : X(\mathrm{SL}_2(\mathbb{Z})) \to \widehat{\mathbb{C}}$ has a simple pole at $x_\infty = \mathrm{SL}_2(\mathbb{Z})\infty$.

Proof: Let $f : X(\mathrm{SL}_2(\mathbb{Z})) \to \widehat{\mathbb{C}}$ be a nonconstant meromorphic function. We can suppose without loss of generality that f has neither a zero nor a pole at x_∞, since if f has a zero or a pole at x_∞, then we can replace it by the function

$$fj^e, \quad \text{where } e = \mathrm{ord}_{x_\infty}(f).$$

Define the function

$$g : X(\mathrm{SL}_2(\mathbb{Z})) \to \widehat{\mathbb{C}}, \quad g(\tau) = \frac{\prod_{i=1}^{n}(j(\tau) - j(z_i))}{\prod_{i=1}^{m}(j(\tau) - j(p_i))}, \quad \forall \tau \in \mathbb{H},$$

where $z_1, \ldots z_n$ and $p_1, \ldots p_m$ are the zeros and poles of f, respectively, listed with multiplicity. Observe that g has the same zeros and poles as f, since by Theorem A.35 we have

$$\sum_{x \in X(\mathrm{SL}_2(\mathbb{Z}))} \mathrm{ord}_x(f) = n - m = 0 \quad (g(\infty) \in \mathbb{C}^*).$$

Therefore, the function f/g has not zeros and poles, and as a consequence, it must be constant by Corollary A.14. This proves that

$$f = cg \in \mathbb{C}(j), \quad \text{for some constant } c \in \mathbb{C}.$$

□

Remarks 2.18.

▷ The modular function $j : X(\mathrm{SL}_2(\mathbb{Z})) \to \widehat{\mathbb{C}}$ is a isomorphism of Riemann surfaces, since j has only a simple pole at $\mathrm{SL}_2(\mathbb{Z})\infty$ (Theorem A.34). As a consequence, the genus of $X(\mathrm{SL}_2(\mathbb{Z}))$ is equal to 0.

- ▷ The restriction of the modular function j to the fundamental domain F for $\mathrm{SL}_2(\mathbb{Z})$ defined in Subsection 1.3.1,

$$F = \{\tau \in \mathbb{H} \,|\, |\mathrm{Re}\,(\tau)| \leq 1/2, |\tau| \geq 1\},$$

is a surjective function (Lemma 1.13).

- ▷ The derivative of the modular function j satisfies the functional equation

$$j' = (j')[\alpha]_2, \quad \forall\, \alpha \in \mathrm{SL}_2(\mathbb{Z}).$$

Furthermore, observe that j' is meromorphic at ∞, since

$$j'(\tau) = -2\pi i \Big(\frac{1}{q} - \sum_{n=0}^{\infty} n a_n q^n\Big), \quad q = e^{2\pi i \tau}, \quad \forall\, \tau \in \mathbb{H}.$$

Therefore, the spaces of automorphic forms $\mathcal{A}_{2k}(\mathrm{SL}_2(\mathbb{Z}))$, with $k \in \mathbb{Z}$, contain nonzero elements.

- ▷ More generally, if Γ is a congruence subgroup that does not contain the matrix $-I$, then the spaces of automorphic forms $\mathcal{A}_k(\Gamma)$, with $k \in \mathbb{Z}$, are not trivial [DS05, p.91-92].

2.1.2 More examples of modular forms

The Eisenstein series of weight 2 for $\mathrm{SL}_2(\mathbb{Z})$ is defined as

$$G_2(\tau) = \sum_{c \in \mathbb{Z}} \sum_{d \in \mathbb{Z}'_c} \frac{1}{(c\tau + d)^2}, \quad \forall\, \tau \in \mathbb{H},$$

where $\mathbb{Z}'_c = \mathbb{Z} \setminus \{0\}$ if $c = 0$ and $\mathbb{Z}'_c = \mathbb{Z}$ otherwise. This double series does not converges absolutely (Lemma 2.12), but if we sum in the indicated order, we have a convergent series.

Theorem 2.19. *The Eisenstein series of weight 2 defines a holomorphic function on the upper half plane. Furthermore,*

$$G_2(\tau) = 2\zeta(2) - 8\pi^2 \sum_{n=1}^{\infty} \sigma(n) q^n, \quad q = e^{2\pi i \tau}, \quad \forall\, \tau \in \mathbb{H},$$

where the coefficients of the series are the values of the arithmetic function

$$\sigma(n) = \sum_{\substack{d|n \\ d>0}} d, \quad \forall\, n \geq 1.$$

Proof: As in the previous section, using (2.8) we have

$$\begin{aligned}\sum_{c\in\mathbb{Z}}\sum_{d\in\mathbb{Z}'_c}\frac{1}{(c\tau+d)^2} &= 2\zeta(2)+2\sum_{c=1}^{\infty}\left(\sum_{d\in\mathbb{Z}}\frac{1}{(c\tau+d)^2}\right)\\ &= 2\zeta(2)+2(-2\pi i)^2\sum_{c=1}^{\infty}\sum_{m=1}^{\infty}mq^{cm},\quad \forall\,\tau\in\mathbb{H}.\end{aligned}$$

Therefore, G_2 defines a holomorphic function on $\mathbb{H}$, since the double series

$$\sum_{c=1}^{\infty}\sum_{m=1}^{\infty}mq^{cm},\quad q\in\mathbb{D},$$

converge uniformly on compact subsets of the unit disc (see Remark 2.14). To obtain its Fourier series it suffices to rearrange the terms of this double series.

□

An immediate consequence of this theorem is that the Eisenstein series of weight 2 is $\mathbb{Z}$-periodic, i.e.,

$$G_2(\tau)=G_2(\tau+m),\quad \forall\,\tau\in\mathbb{H},\quad \forall\,m\in\mathbb{Z},$$

and therefore it is invariant under the operator $[T]_2$, where $T=\left[\begin{smallmatrix}1&1\\0&1\end{smallmatrix}\right]$. However, if we let $S=\left[\begin{smallmatrix}0&-1\\1&0\end{smallmatrix}\right]$, then

$$\begin{aligned}(G_2[S]_2)(\tau) &= \tau^{-2}\sum_{c\in\mathbb{Z}}\sum_{d\in\mathbb{Z}'_c}\frac{1}{(c(1/\tau)+d)^2}\\ &= \sum_{c\in\mathbb{Z}}\sum_{d\in\mathbb{Z}'_c}\frac{1}{(d\tau+c)^2}=2\zeta(2)+\sum_{d\in\mathbb{Z}}\sum_{c\neq0}\frac{1}{(c\tau+d)^2},\quad \forall\,\tau\in\mathbb{H},\end{aligned}$$

which differs from $G_2(\tau)=2\zeta(2)+\sum_{c\neq0}\sum_{d\in\mathbb{Z}}(c\tau+d)^{-2}$ in the order of summation.

Lemma 2.20. *Let $\tau\in\mathbb{H}$. Then*

$$\sum_{c\neq0}\sum_{d\in\mathbb{Z}}\frac{1}{(c\tau+d)(c\tau+d+1)}=0.$$

Proof: Using partial fractions, we have

$$\sum_{c\neq0}\sum_{d\in\mathbb{Z}}\frac{1}{(c\tau+d)(c\tau+d+1)}=\sum_{c\neq0}\sum_{d\in\mathbb{Z}}\left(\frac{1}{(c\tau+d)}-\frac{1}{(c\tau+d+1)}\right)$$

and

$$\sum_{d\in\mathbb{Z}}\left(\frac{1}{(c\tau+d)}-\frac{1}{(c\tau+d+1)}\right)=\lim_{N\to\infty}\sum_{d=-N}^{N-1}\left(\frac{1}{(c\tau+d)}-\frac{1}{(c\tau+d+1)}\right).$$

As the finite sum is telescoping, we deduce that

$$\sum_{d=-N}^{N-1}\left(\frac{1}{(c\tau+d)}-\frac{1}{(c\tau+d+1)}\right)=\frac{1}{(c\tau-N)}-\frac{1}{(c\tau+N)}\overset{N\to\infty}{\longrightarrow}0.$$

□

As a consequence of this lemma, observe that

$$\begin{aligned}G_2(\tau)&=G_2(\tau)-\sum_{c\neq 0}\sum_{d\in\mathbb{Z}}\frac{1}{(c\tau+d)(c\tau+d+1)}\\&=2\zeta(2)+\sum_{c\neq 0}\sum_{d\in\mathbb{Z}}\frac{1}{(c\tau+d)^2(c\tau+d+1)},\quad\forall\tau\in\mathbb{H}.\end{aligned}$$

where the double sum is now absolutely convergent (recall the proof of Theorem 2.13). Changing the order of summation in the double series and separating the convergence terms back out, we obtain that

$$\begin{aligned}G_2(\tau)&=2\zeta(2)+\sum_{d\in\mathbb{Z}}\sum_{c\neq 0}\frac{1}{(c\tau+d)^2(c\tau+d+1)}\\&=(G_2[S]_2)(\tau)-\sum_{d\in\mathbb{Z}}\sum_{c\neq 0}\frac{1}{(c\tau+d)(c\tau+d+1)},\quad\forall\tau\in\mathbb{H},\end{aligned}$$

The error term is

$$\begin{aligned}-\sum_{d\in\mathbb{Z}}\sum_{c\neq 0}\frac{1}{(c\tau+d)(c\tau+d+1)}&=-\lim_{N\to\infty}\sum_{d=-N}^{N-1}\sum_{c\neq 0}\frac{1}{(c\tau+d)(c\tau+d+1)}\\&=-\lim_{N\to\infty}\sum_{c\neq 0}\sum_{d=-N}^{N-1}\left(\frac{1}{c\tau+d}-\frac{1}{c\tau+d+1}\right)\\&=-\lim_{N\to\infty}\frac{1}{\tau}\sum_{c\neq 0}\left(\frac{1}{c+(-N/\tau)}+\frac{1}{(-N/\tau)-c}\right)\end{aligned}$$

since the finite series is telescoping. Using the identities 2.6 and 2.7 of the cotangent function (Subsection 2.1.1), we can write

$$\frac{1}{\tau}\sum_{c\neq 0}\left(\frac{1}{c+(-N/\tau)}+\frac{1}{(-N/\tau)-c}\right)=\frac{2\pi i}{\tau}+\frac{2}{N}-\frac{4\pi i}{\tau}\sum_{m=0}^{\infty}(e^{-2\pi i N/\tau})^m,$$

and therefore

$$\lim_{N\to\infty}\frac{1}{\tau}\sum_{c\neq 0}\left(\frac{1}{c+(-N/\tau)}+\frac{1}{(-N/\tau)-c}\right)=\frac{2\pi i}{\tau}.$$

Theorem 2.21. *The Eisenstein series of weight* 2 *satisfies the functional equation*

$$G_2[\alpha]_2(\tau) = G_2(\tau) - \frac{2\pi i c}{j(\alpha,\tau)}, \quad \forall\,\tau \in \mathbb{H}, \quad \forall\,\alpha = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathrm{SL}_2(\mathbb{Z}).$$

Proof: Since $\mathrm{SL}_2(\mathbb{Z})$ is generated by the matrices T and S (Lemma 1.1), it suffices to prove the following assertion: Suppose that G_2 satisfies the equation for two matrices $\alpha_1 = \left[\begin{smallmatrix} a_1 & b_1 \\ c_1 & d_1 \end{smallmatrix}\right]$, $\alpha_2 = \left[\begin{smallmatrix} a_2 & b_2 \\ c_2 & d_2 \end{smallmatrix}\right] \in \mathrm{SL}_2(\mathbb{Z})$,

$$G_2[\alpha_1]_2(\tau) = G_2(\tau) - \frac{2\pi i c_1}{j(\alpha_1,\tau)}, \quad \forall\,\tau \in \mathbb{H} \tag{2.12}$$

and

$$G_2[\alpha_2]_2(\tau) = G_2(\tau) - \frac{2\pi i c_2}{j(\alpha_2,\tau)}, \quad \forall\,\tau \in \mathbb{H}. \tag{2.13}$$

Then G_2 also satisfies the equation for the inverse $\alpha_1^{-1} = \left[\begin{smallmatrix} d_1 & -b_1 \\ -c_1 & a_1 \end{smallmatrix}\right]$ and the product

$$\alpha_1\alpha_2 = \begin{bmatrix} a_1a_2 + b_1c_2 & a_1b_2 + b_1d_2 \\ c_1a_2 + d_1c_2 & c_1b_2 + d_1d_2 \end{bmatrix}.$$

Applying the operator $[\alpha_1^{-1}]_2$ to both sides of the equation (2.12) and using that $1 = j(\alpha_1, \alpha_1^{-1}(\tau))j(\alpha_1^{-1},\tau)$, we obtain that

$$G_2(\tau) = (G_2[\alpha^{-1}]_2)(\tau) + \frac{2\pi i(-c_1)}{j(\alpha_1^{-1},\tau)}, \quad \forall\,\tau \in \mathbb{H}.$$

As a consequence, observe that G_2 satisfies the equation for the inverse α_1^{-1}.

Now applying the operator $[\alpha_2]_2$ to both sides of the equation (2.12) and using that $j(\alpha_1\alpha_2,\tau) = j(\alpha_1,\alpha_2(\tau))j(\alpha_2,\tau)$, we obtain that

$$\begin{aligned}(G_2[\alpha_1\alpha_2]_2)(\tau) &= (G_2[\alpha_2]_2)(\tau) - \frac{2\pi i c_1}{j(\alpha_1\alpha_2,\tau)j(\alpha_2,\tau)} \\ &= G_2(\tau) - \frac{2\pi i c_2}{j(\alpha_2,\tau)} - \frac{2\pi i c_1}{j(\alpha_1\alpha_2,\tau)j(\alpha_2,\tau)} \quad [(2.13)] \\ &= G_2(\tau) - \frac{2\pi i(c_1a_2 + d_1c_2)}{j(\alpha_1\alpha_2,\tau)}\frac{j(\alpha_2,\tau)}{j(\alpha_2,\tau)} \quad [1 + c_2b_2 = a_2d_2] \\ &= G_2(\tau) - \frac{2\pi i(c_1a_2 + d_1c_2)}{j(\alpha_1\alpha_2,\tau)}, \quad \forall\,\tau \in \mathbb{H}.\end{aligned}$$

Therefore, observe that G_2 also satisfies the equation for the product $\alpha_1\alpha_2$. □

For any $N \in \mathbb{Z}^+$, define $G_{2,N} : \mathbb{H} \to \mathbb{C}$ as

$$G_{2,N}(\tau) = G_2(\tau) - NG_2(N\tau), \quad \forall\,\tau \in \mathbb{H}.$$

Then this holomorphic function is a example of modular form of weight 2 with respect to $\Gamma_0(N)$, since

$$\begin{aligned}(G_{2,N}[\alpha]_2)(\tau) &= G_2(\tau) - \frac{2\pi i c' N}{j(\alpha,\tau)} - \frac{N}{j(\alpha,\tau)^2} G_2(N\alpha(\tau)) \\ &= G_2(\tau) - N\left(\frac{2\pi i c'}{j(\alpha', N\tau)} + G_2[\alpha']_2(N\tau)\right) \quad \left(\alpha' = \left[\begin{smallmatrix} a & Nb \\ c' & d \end{smallmatrix}\right]\right) \\ &= G_2(\tau) - NG_2(N\tau), \quad \forall\, \tau \in \mathbb{H}, \quad \forall\, \alpha = \begin{bmatrix} a & b \\ c'N & d \end{bmatrix} \in \Gamma_0(N),\end{aligned}$$

and its Fourier series is

$$G_{2,N}(\tau) = 2\zeta(2)(1-N) - 8\pi^2 \sum_{n=1}^{\infty} \sigma_{1,N}(n) q^n, \quad q = e^{2\pi i \tau}, \quad \forall\, \tau \in \mathbb{H},$$

where the coefficients of the series are the values of the arithmetic function

$$\sigma_{1,N}(n) = \sum_{\substack{0<d|n \\ N\nmid d}} d, \quad \forall\, n \geq 1.$$

2.2 Some applications

2.2.1 Transformation law of the Dedekind eta function

The Dedekind eta function $\eta : \mathbb{H} \to \mathbb{C}$ is defined as the infinite product

$$\eta(\tau) = q_{24} \prod_{n=1}^{\infty} (1-q^n), \quad q_{24} = e^{2\pi i \tau/24}, \quad q = e^{2\pi i \tau}, \quad \forall\, \tau \in \mathbb{H}.$$

Lemma 2.22. *The series*

$$\sum_{n=1}^{\infty} \log(1-q^n), \quad q \in \mathbb{D},$$

converges uniformly on compact subsets of the unit disc. As a consequence, the function η defines a holomorphic function on the upper half plane.

Proof: Let $0 < r < 1$. Recall that

$$\log(1+z) = \sum_{n=0}^{\infty} \frac{(-1)^n z^{n+1}}{n+1}, \quad \forall\, z \in \mathbb{C},\ |z| < 1.$$

Using this Taylor series, we obtain the estimation

$$|\log(1+z)| \leq \sum_{n=0}^{\infty} |z|^{n+1} \leq \frac{|z|}{1-r}, \quad \forall\, z \in \mathbb{C},\ |z| < r.$$

Therefore,

$$\sum_{n=1}^{\infty} |\log(1-q^n)| \leq \frac{1}{1-r}\sum_{n=1}^{\infty} r^n = \frac{r}{(1-r)^2}, \quad \forall\, q \in \mathbb{D},\ |q|<r.$$

□

Note that the function η is nonvanishing on $\mathbb{H}$, since $q_{24} \neq 0$ and

$$\prod_{n=1}^{\infty}(1-q^n) = e^{\sum_{n=1}^{\infty}\log(1-q^n)} \neq 0, \quad \forall\, \tau \in \mathbb{H}.$$

Theorem 2.23. *The Dedekind eta function satisfies the transformation law*

$$\eta(-1/\tau) = \sqrt{-i\tau}\,\eta(\tau), \quad \forall\, \tau \in \mathbb{H},$$

where $\sqrt{}$ *is the principal branch of the multivalued function* $z^{1/2}$.

Proof: The logarithmic derivative of η is

$$\begin{aligned}\frac{d}{d\tau}\log(\eta(\tau)) &= \frac{\pi i}{12} - 2\pi i \sum_{d=1}^{\infty}\frac{dq^d}{1-q^d} = \frac{\pi i}{12} - 2\pi i \sum_{d=1}^{\infty} d \sum_{m=1}^{\infty} q^{dm} \\ &= \frac{\pi i}{12} - 2\pi i \sum_{m=1}^{\infty}\sum_{d=1}^{\infty} dq^{dm} = \frac{\pi i}{12} - 2\pi i \sum_{n=1}^{\infty}\left(\sum_{0<d|n} d\right) q^n \\ &= \frac{\pi i}{12}E_2(\tau), \quad \forall\, \tau \in \mathbb{H},\end{aligned}$$

where $E_2 : \mathbb{H} \to \mathbb{C}$ is the normalized Eisenstein series of weight 2,

$$E_2(\tau) = \frac{G_2(\tau)}{2\zeta(2)} = 1 - 24\sum_{n=1}^{\infty}\sigma(n)q^n, \quad q = e^{2\pi i\tau}, \quad \forall\, \tau \in \mathbb{H}.$$

Therefore,

$$\frac{d}{d\tau}\log(\eta(-1/\tau)) = \frac{\pi i}{12}\tau^{-2}E_2(-1/\tau), \quad \forall\, \tau \in \mathbb{H},$$

and

$$\frac{d}{d\tau}\log(\sqrt{-i\tau}\,\eta(\tau)) = \frac{1}{2\tau} + \frac{\pi i}{12}E_2(\tau) = \frac{\pi i}{12}\left(E_2(\tau) + \frac{12}{2\pi i\tau}\right), \quad \forall\, \tau \in \mathbb{H}.$$

Using Theorem 2.21 with $\alpha = \left[\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right]$, we obtain that

$$\tau^{-2}E_2(-1/\tau) = E_2(\tau) + \frac{12}{2\pi i\tau}, \quad \forall\, \tau \in \mathbb{H},$$

so

$$\frac{d}{d\tau}\log(\eta(-1/\tau)) - \frac{d}{d\tau}\log(\sqrt{-i\tau}\eta(\tau)) = 0, \quad \forall\, \tau \in \mathbb{H}.$$

As a consequence, observe that there exists a constant $c \in \mathbb{C}$ such that

$$\eta(-1/\tau) = c\sqrt{-i\tau}\eta(\tau), \quad \forall\, \tau \in \mathbb{H}.$$

Letting $\tau = i$, we conclude that this constant c must be equal to 1, since

$$\eta(i) = \eta(-1/i) = c\sqrt{-i^2}\eta(i) = c\eta(i).$$

□

The function $\eta^{24} : \mathbb{H} \to \mathbb{C}$,

$$\eta^{24}(\tau) = q\prod_{n=1}^{\infty}(1-q^n)^{24}, \quad q = e^{2\pi i\tau}, \quad \forall\, \tau \in \mathbb{H},$$

is a cusp form of weight 12 with respect to $\mathrm{SL}_2(\mathbb{Z})$, since

$$\eta^{24}(\tau+1) = \eta^{24}(\tau), \quad \eta^{24}(-1/\tau) = \tau^{12}\eta^{24}(\tau), \quad \forall\, \tau \in \mathbb{H},$$

and

$$\lim_{\mathrm{Im}\,(\tau)\to\infty} \eta^{24}(\tau) = 0.$$

As the function Δ is also a cusp form of weight 12 with respect to $\mathrm{SL}_2(\mathbb{Z})$ and $\dim(\mathcal{S}_{12}(\mathrm{SL}_2(\mathbb{Z}))) = 1$ by Theorem 2.9, we deduce that

$$\Delta = c\eta^{24}, \quad \text{for some constant } c \in \mathbb{C}.$$

Equating the coefficients of their Fourier series, we obtain the identity

$$\Delta = (2\pi)^{12}\eta^{24}.$$

2.2.2 Four squares problem

Let us consider the following questions:

> Can each positive integer n be written as a sum of four squares,
>
> $$n = z_1^2 + z_2^2 + z_3^2 + z_4^2, \quad z_i \in \mathbb{Z}?$$
>
> In which case, in how many ways can n be written as a sum of four squares? Is it possible to determine an explicit formula?

The answers to these questions are affirmative, i.e, each positive integer n can be written as a sum of four squares, and in addition, there exists an explicit formula that counts the representation number of n by four squares.

The proof that we present here uses basically the fact of that the functions

$$G_{2,2}(\tau) = G_2(\tau) - 2G_2(2\tau) \quad \text{and} \quad G_{2,4}(\tau) = G_2(\tau) - 4G_2(N\tau)$$

(see Subsection 2.1.2) form a basis of the space of modular forms $\mathcal{M}_2(\Gamma_0(4))$.

Define

$$r(n,k) = \#\{z \in \mathbb{Z}^k \,|\, n = z_1^2 + \cdots + z_k^2\}, \quad k \geq 1,$$

and consider the generating function of the representation numbers,

$$\theta(\tau, k) = \sum_{n=0}^{\infty} r(n,k) q^n, \quad q = e^{2\pi i \tau}, \quad \tau \in \mathbb{H}.$$

This series defines a holomorphic function on the upper half plane, since it converges uniformly on the subsets $\Omega(b) = \{\tau \in \mathbb{H} \,|\, \operatorname{Im} \tau \geq b\}$, $b > 0$. Indeed,

$$\begin{aligned} \sum_{n=0}^{\infty} r(n,k)|q|^n &= \sum_{n=0}^{\infty} r(n,k)(e^{-2\pi y})^n \\ &\leq \sum_{n=0}^{\infty} (1 + 2\sqrt{n})^k (e^{-2\pi b})^n < \infty, \quad \forall\, \tau = x + iy \in \Omega(b). \end{aligned}$$

Furthermore, as

$$r(n,k) = \sum_{s+j=n} r(s,k_1) r(j,k_2), \quad k_1 + k_2 = k,$$

we deduce that

$$\theta(\tau, k_1)\theta(\tau, k_2) = \theta(\tau, k_1 + k_2), \quad \forall\, \tau \in \mathbb{H}.$$

Let $\theta : \mathbb{H} \to \mathbb{C}$ be the function

$$\theta(\tau) = \theta(\tau, 1), \quad \forall\, \tau \in \mathbb{H}.$$

The following lemma summarizes the principal properties of this holomorphic function. Its proof is immediate and it is left as an exercise to the reader.

Lemma 2.24. *Let k be a positive integer. Then:*

1. *θ is $\mathbb{Z}$-periodic*

2. *$\theta(\tau)^k = \theta(\tau, k)$, $\forall\, \tau \in \mathbb{H}$*

3. $\theta(\tau) = \sum_{n=0}^{\infty} r(n,1)q^n = \sum_{n\in\mathbb{Z}} e^{2\pi i n^2\tau}$, $q = e^{2\pi i\tau}$, $\forall\,\tau \in \mathbb{H}$.

This last expression for θ allows us to determine another transformation law.

Theorem 2.25. *The function θ satisfies the transformation law*

$$\theta(-1/2\tau) = \sqrt{-i\tau}\,\theta(\tau/2), \quad \forall\,\tau \in \mathbb{H},$$

where $\sqrt{}$ is the principal branch of the multivalued function $z^{1/2}$.

Proof: Observe that it suffices to prove this transformation law for $\tau = it$, with $t > 0$, since the two sides of the equation are holomorphic functions on $\mathbb{H}$. Let $f, f_t : \mathbb{R} \to \mathbb{R}$ be the functions

$$f(x) = e^{-\pi x^2} \quad \text{and} \quad f_t(x) = e^{-\pi t x^2}.$$

Recall that the Fourier transform of the function f is itself, i.e.,

$$\int_{-\infty}^{+\infty} e^{-\pi x^2} e^{-2\pi i x\xi} dx = e^{-\pi\xi^2}, \quad \forall\,\xi \in \mathbb{R}.$$

The change of variables $x = t^{1/2}x$ in the integral shows that the Fourier transform of the function f_t is

$$\widehat{f_t}(\xi) = t^{-1/2} e^{-\pi\xi^2/t}, \quad \forall\,\xi \in \mathbb{R}.$$

As a consequence, using the Poisson summation formula, we obtain that

$$\sum_{n\in\mathbb{Z}} e^{-\pi t n^2} = \sum_{n\in\mathbb{Z}} t^{-1/2} e^{-\pi n^2/t}, \quad \forall\,t > 0,$$

or equivalently,

$$\theta(\tau/2) = \frac{1}{\sqrt{-i\tau}}\theta(-1/2\tau), \quad \tau = it, \quad \forall\,t > 0.$$

□

Note that the transformation law of the function θ can also be written as

$$\theta(-1/4\tau) = \sqrt{-2i\tau}\,\theta(\tau), \quad \forall\,\tau \in \mathbb{H}. \tag{2.14}$$

The matrix $\left[\begin{smallmatrix} 0 & -1 \\ 4 & 0 \end{smallmatrix}\right]$ taking τ to $-1/4\tau$ is not in $\mathrm{SL}_2(\mathbb{Z})$, but the product

$$\begin{bmatrix} 0 & 1/4 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & -1 \\ 4 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 4 & 1 \end{bmatrix}$$

taking τ to $\tau/(4\tau+1)$, is in $\mathrm{SL}_2(\mathbb{Z})$. Applying the corresponding succession of transformations and using (2.14) twice, we obtain that

$$\begin{aligned}\theta\left(\frac{\tau}{4\tau+1}\right) &= \theta\left(-\frac{1}{4(-1/4\tau-1)}\right)\\ &= \sqrt{2i\left(\frac{1}{4\tau}+1\right)}\theta\Big(-\frac{1}{4\tau}\Big) \qquad \left[\theta\Big(-\frac{1}{4\tau}\Big)=\theta\Big(-\frac{1}{4\tau}-1\Big)\right]\\ &= \sqrt{2i\left(\frac{1}{4\tau}+1\right)(-2i\tau)}\theta(\tau) = \sqrt{4\tau+1}\theta(\tau), \quad \forall\, \tau\in\mathbb{H},\end{aligned}$$

where in the penultimate equality we have used tacitly that

$$\sqrt{2i(1/4\tau+1)}\sqrt{-2i\tau} = \sqrt{2i(1/4\tau+1)(-2i\tau)}, \quad \forall\, \tau\in\mathbb{H}.$$

As a consequence, observe that

$$\theta(\gamma(\tau),4) = (c\tau+d)^2\theta(\tau,4), \quad \text{for } \gamma=\pm\begin{bmatrix}1&1\\0&1\end{bmatrix} \text{ and } \pm\begin{bmatrix}1&0\\4&1\end{bmatrix}.$$

Lemma 2.26. *The subgroup* $\Gamma_{\theta,4}$ *of* $\mathrm{SL}_2(\mathbb{Z})$ *generated by the matrices*

$$\pm\begin{bmatrix}1&1\\0&1\end{bmatrix} \quad \textit{and} \quad \pm\begin{bmatrix}1&0\\4&1\end{bmatrix}$$

is

$$\Gamma_0(4) = \left\{\begin{bmatrix}a&b\\c&d\end{bmatrix}\in\mathrm{SL}_2(\mathbb{Z}) \,\middle|\, \begin{bmatrix}a&b\\c&d\end{bmatrix}\equiv\begin{bmatrix}*&*\\0&*\end{bmatrix} \pmod 4\right\}.$$

Proof: Let $\alpha = \left[\begin{smallmatrix}a&b\\c&d\end{smallmatrix}\right] \in \Gamma_0(4)$. Below, we describe an algorithm to compute $\gamma\in\Gamma_{\theta,4}$ such that $\alpha\gamma\in\Gamma_{\theta,4}$ (and therefore $\alpha\in\Gamma_{\theta,4}$).

We can suppose without loss of generality that $c\neq 0$, since otherwise $\alpha = \left[\begin{smallmatrix}\pm1&b\\0&\pm1\end{smallmatrix}\right]\in\Gamma_{\theta,4}$. The identity

$$\begin{bmatrix}a&b\\c&d\end{bmatrix}\begin{bmatrix}1&n\\0&1\end{bmatrix} = \begin{bmatrix}a'&b'\\c&nc+d\end{bmatrix}, \quad n\in\mathbb{Z},$$

shows that there exists a matrix $\gamma_1\in\Gamma_{\theta,4}$ such that $\alpha\gamma_1\in\Gamma_0(4)$ has bottom row

$$(c',d') = (c, nc+d), \quad \text{with } |d'| < |c'|/2$$

(the inequality is clearly strict because of properties of c and d modulo 4). On the other hand, the identity

$$\begin{bmatrix}a'&b'\\c'&d'\end{bmatrix}\begin{bmatrix}1&0\\4n&1\end{bmatrix} = \begin{bmatrix}a''&b''\\c'+4nd'&d'\end{bmatrix}, \quad n\in\mathbb{Z},$$

shows that there exists a matrix $\gamma_2 \in \Gamma_{\theta,4}$ such that $\alpha\gamma_1\gamma_2 \in \Gamma_0(4)$ has bottom row

$$(c'', d'') = (c' + 4nd', d'), \quad \text{with } |c''| < 2|d''|$$

(the inequality is again strict because of properties of c' and d' modulo 4). As a consequence, observe that these two multiplications allow us to reduce strictly the absolute value of lower left entry of the matrix α, since

$$|c''| < 2|d''| = 2|d'| < |c'| = |c|.$$

Repeating this argument whenever the lower left entry of the resulting matrix is nonzero (a finite number of times), we deduce that there exists a matrix $\gamma \in \Gamma_{\theta,4}$ such that $\alpha\gamma \in \Gamma_0(4)$ has bottom row $(0, \pm 1)$, and therefore

$$\alpha\gamma \in \Gamma_{\theta,4}.$$

□

Using Theorem 2.8 we can conclude that $\theta(\cdot, 4) \in \mathcal{M}_2(\Gamma_0(4))$, since

$$\theta(\tau, 4) = \sum_{n=0}^{\infty} r(n, 4) q_4^{4n}, \quad q_4 = e^{2\pi i\tau/4}, \quad \forall\, \tau \in \mathbb{H},$$

and

$$r(n, 4) \leq (1 + 2\sqrt{n})^4, \quad \forall\, n \geq 1.$$

Theorem 2.27. *The representation number of a positive integer n by four squares is*

$$r(n, 4) = 8 \sum_{\substack{0<d|n \\ 4\nmid d}} d.$$

As a consequence, observe that n can be written as a sum of four squares.

Proof: Consider the modular forms $G_{2,2}, G_{2,4} \in \mathcal{M}_2(\Gamma_0(4))$,

$$G_{2,2}(\tau) = -\frac{\pi^2}{3}\left(1 + 24 \sum_{n=1}^{\infty} \sigma_{1,2}(n) q^n\right), \quad \forall\, \tau \in \mathbb{H}$$

and

$$G_{2,4}(\tau) = -\pi^2 \left(1 + 8 \sum_{n=1}^{\infty} \sigma_{1,4}(n) q^n\right), \quad \forall\, \tau \in \mathbb{H}$$

The subset $\{G_{2,2}, G_{2,4}\}$ forms a basis of the vector space $\mathcal{M}_2(\Gamma_0(4))$, since it is linearly independent and $\dim(\mathcal{M}_2(\Gamma_0(4))) = 2$ (compute the dimension of this space using [DS05, p.107]). Therefore, the function $\theta(\cdot, 4)$ can be expressed in a unique way as a linear combination of these modular forms,

$$\theta(\cdot, 4) = aG_{2,2} + bG_{2,4}, \quad a, b \in \mathbb{C}.$$

The expansions

$$\begin{aligned} \theta(\tau, 4) &= 1 + 8q + \cdots, \\ -\frac{3}{\pi^2} G_{2,2}(\tau) &= 1 + 24q + \cdots, \\ -\frac{1}{\pi^2} G_{2,4}(\tau) &= 1 + 8q + \cdots, \end{aligned}$$

show that $\theta(\cdot, 4) = -(1/\pi^2)G_{2,4}$. Equating the Fourier coefficients, we obtain that the representation number of a positive integer n as a sum of four squares is

$$r(n, 4) = 8\sigma_{1,4}(n) = 8 \sum_{\substack{0<d|n \\ 4\nmid d}} d.$$

□

Chapter 3

Modularity Theorem

In this third chapter we introduce moduli spaces for the modular curves

$$Y_0(N),\ \ Y_1(N),\ \text{ and }\ Y(N),\ \ N \in \mathbb{Z}^+,$$

and we explain a complex analytic version of the Modularity Theorem which is equivalent to the original version that was proved about twenty years ago.

3.1 Weil pairing

Let Λ be a lattice in $\mathbb{C}$ with basis (w_1, w_2),

$$\Lambda = w_1\mathbb{Z} + w_2\mathbb{Z} \quad (\text{assume } w_1/w_2 \in \mathbb{H}),$$

and N a positive integer. Consider the multiply-by-N map

$$[N] : \mathbb{C}/\Lambda \to \mathbb{C}/\Lambda, \quad z + \Lambda \mapsto Nz + \Lambda.$$

This map is a holomorphic group homomorphism (Remark B.6). Its kernel is the set of N-torsion points of $\mathbb{C}/\Lambda$,

$$\ker[N] = \{P \in \mathbb{C}/\Lambda \,|\, [N]P = 0\} = \langle w_1/N + \Lambda\rangle + \langle w_2/N + \Lambda\rangle,$$

a subgroup isomorphic to $\mathbb{Z}/N\mathbb{Z} \times \mathbb{Z}/N\mathbb{Z}$.

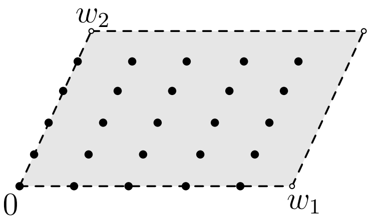


Figure 3.1: ker[5]: the 5-torsion points of $\mathbb{C}/\Lambda$

Let E denote the torus $\mathbb{C}/\Lambda$, $E[N]$ its subgroup of N-torsion points and $\boldsymbol{\mu}_N$ the cyclic group of the complex N-th roots of unity,

$$\boldsymbol{\mu}_N = \{z \in \mathbb{C} \,|\, z^N = 1\} = \langle e^{2\pi i/N} \rangle.$$

The Weil pairing $e_N : E[N] \times E[N] \to \boldsymbol{\mu}_N$ is defined as

$$e_N(P, Q) = e^{2\pi i \det(\gamma)/N}, \quad \forall\, (P, Q) \in E[N] \times E[N],$$

where the matrix $\gamma \in \mathrm{M}_2(\mathbb{Z}/N\mathbb{Z})$ is determined by the condition

$$\begin{bmatrix} P \\ Q \end{bmatrix} = \gamma \begin{bmatrix} w_1/N + \Lambda \\ w_2/N + \Lambda \end{bmatrix} = \begin{bmatrix} k_1(P) & k_2(P) \\ k_1(Q) & k_2(Q) \end{bmatrix} \begin{bmatrix} w_1/N + \Lambda \\ w_2/N + \Lambda \end{bmatrix}. \tag{3.1}$$

Even though the determinant of the matrix γ is defined only modulo N, the root $e^{2\pi i \det(\gamma)/N}$ is well-defined, since the function $e^{2\pi i z/N}$ is $N\mathbb{Z}$-periodic.

The Weil pairing is independent of the chosen basis (w_1, w_2) of the lattice Λ. Let (w_1', w_2') be another basis of Λ, with $w_1'/w_2' \in \mathbb{H}$, and $\gamma' \in \mathrm{M}_2(\mathbb{Z}/N\mathbb{Z})$ satisfying the condition

$$\begin{bmatrix} P \\ Q \end{bmatrix} = \gamma' \begin{bmatrix} w_1'/N + \Lambda \\ w_2'/N + \Lambda \end{bmatrix}.$$

Then

$$\begin{bmatrix} w_1' \\ w_2' \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}, \quad \text{for some } \alpha = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathrm{SL}_2(\mathbb{Z}),$$

and

$$\begin{bmatrix} P \\ Q \end{bmatrix} = \gamma' \begin{bmatrix} w_1'/N + \Lambda \\ w_2'/N + \Lambda \end{bmatrix} = \gamma' \begin{bmatrix} \bar{a} & \bar{b} \\ \bar{c} & \bar{d} \end{bmatrix} \begin{bmatrix} w_1/N + \Lambda \\ w_2/N + \Lambda \end{bmatrix}.$$

As a consequence,

$$\gamma = \gamma' \begin{bmatrix} \bar{a} & \bar{b} \\ \bar{c} & \bar{d} \end{bmatrix}, \quad \text{and therefore } \det(\gamma) = \det(\gamma').$$

Remark 3.1. If P and Q generate the group $E[N]$, then the matrix γ satisfying the condition (3.1) is invertible, i.e., $\gamma \in \mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z})$. Therefore,

$$e_N(P, Q) = e^{2\pi i \det(\gamma)/N}$$

is a primitive complex N-th root of unity, since $\det(\gamma) \in (\mathbb{Z}/N\mathbb{Z})^*$

The following lemma states the principal properties of the Weil pairing. We leave to the reader as an exercise its proof, which is easy and routine.

Lemma 3.2. *The Weil e_N-pairing has the following properties:*

1. It is bilinear,

$$\begin{aligned} e_N(P_1+P_2,Q) &= e_N(P_1,Q)e_N(P_2,Q), \quad \forall P_1,P_2,Q \in E[N], \\ e_N(P,Q_1+Q_2) &= e_N(P,Q_1)e_N(P,Q_2), \quad \forall P_1,P_2,Q \in E[N]. \end{aligned}$$

2. It is alternating,

$$e_N(P,P)=1, \quad \forall\, P \in E[N].$$

In particular, observe that

$$e_N(P,Q)=e_N(Q,P)^{-1}, \quad \forall\, P,Q \in E[N].$$

3. It is nondegenerate,

$$\textit{if } P \in E[N] \textit{ and } e_N(P,Q)=1,\ \forall\, Q \in E[N], \textit{ then } P=0.$$

4. It is compatible with N, i.e., for any positive integer d, the diagram

$$\begin{tikzcd} E[dN]\times E[dN] \arrow[r, "e_{dN}(\cdot,\cdot)"] \arrow[d, "d(\cdot,\cdot)"'] & \boldsymbol{\mu}_{dN} \arrow[d, ".d"] \\ E[N]\times E[N] \arrow[r, "e_N(\cdot,\cdot)"] & \boldsymbol{\mu}_N \end{tikzcd}$$

commutes, where the vertical maps are suitable multiplications by d.

Let Λ' be another lattice in $\mathbb{C}$. Suppose that the torus $\mathbb{C}/\Lambda$ and $\mathbb{C}/\Lambda'$ are isomorphic, i.e., there exists $\alpha \in \mathbb{C}$ such that $\alpha\Lambda = \Lambda'$ (Corollary B.4).

Theorem 3.3. *The isomorphism of Riemann surfaces*

$$z+\Lambda \in \mathbb{C}/\Lambda \overset{F}{\longmapsto} \alpha z + \Lambda' \in \mathbb{C}/\Lambda'$$

preserves the Weil e_N-pairing, $e_N(P,Q)=e_N(F(P),F(Q))$, $\forall\, P,Q \in E[N]$.

Let E' denote the torus $\mathbb{C}/\Lambda'$ and $E'[N]$ its subgroup of N-torsion points. Before proving this theorem, recall that F is also an isomorphism of groups, since $F(0+\Lambda)=0+\Lambda'$ (Remark B.6). Therefore, $F(P) \in E'[N]$, $\forall\, P \in E[N]$.

Proof: Let $(w_1', w_2') = (\alpha w_1, \alpha w_2)$, a basis of the lattice Λ'. If $\gamma \in \mathrm{M}_2(\mathbb{Z}/N\mathbb{Z})$ satisfies the condition

$$\begin{bmatrix} P \\ Q \end{bmatrix} = \gamma \begin{bmatrix} w_1/N+\Lambda \\ w_2/N+\Lambda \end{bmatrix},$$

then applying the isomorphism F to both sides of the equation gives

$$\begin{bmatrix} F(P) \\ F(Q) \end{bmatrix} = \gamma \begin{bmatrix} w_1'/N+\Lambda \\ w_2'/N+\Lambda \end{bmatrix}.$$

Observe that this proves that $e_N(P,Q)=e_N(F(P),F(Q))$, $\forall\, P,Q \in E[N]$. □

3.2 Moduli spaces for modular curves

Recall that a complex elliptic curve is a compact Riemann surface of genus 1. The reader can consult the results over complex elliptic curves in Appendix B. Since any complex elliptic curve E is isomorphic to a complex torus $\mathbb{C}/\Lambda$, in this section we let the term complex elliptic curve be a synonym for complex torus.

Let $E = \mathbb{C}/\Lambda$ and $E' = \mathbb{C}/\Lambda'$ be two complex elliptic curves. If E and E' are isomorphic, then there always exists a holomorphic group isomorphism

$$F : E \to E', \quad F(z + \Lambda) = \alpha z + \Lambda', \quad \text{with } \alpha \in \mathbb{C},\ \alpha\Lambda = \Lambda'.$$

We are interested in these isomorphisms since they preserve the group structures on the complex elliptic curves. Therefore, to simplify in this section, we also assume that the term "isomorphism" always means holomorphic group isomorphism.

Let N be a positive integer:

▷ An *enhanced elliptic curve for* $\Gamma_0(N)$ is an ordered pair (E, C) where E is a complex elliptic curve and C is a cyclic subgroup of E of order N. Two such pairs (E, C) and (E', C') are equivalent, denoted $(E, C) \sim (E', C')$, if there exists some isomorphism $E \to E'$ taking C to C'. The set of equivalence classes is denoted

$$S_0(N) = \{\text{enhanced elliptic curves for } \Gamma_0(N)\}/\sim .$$

An element of $S_0(N)$ is denoted $[E, C]$, the square brackets $[\,]$ connoting equivalence class.

▷ An *enhanced elliptic curve for* $\Gamma_1(N)$ is an ordered pair (E, P) where E is a complex elliptic curve and P is a point of E of order N. Two such pairs (E, P) and (E', P') are equivalent, denoted $(E, P) \sim (E', P')$, if there exists some isomorphism $E \to E'$ taking P to P'. The set of equivalence classes is denoted

$$S_1(N) = \{\text{enhanced elliptic curves for } \Gamma_1(N)\}/\sim .$$

An element of $S_1(N)$ is denoted $[E, P]$.

▷ An *enhanced elliptic curve for* $\Gamma(N)$ is an ordered pair $(E, (P, Q))$ where E is a complex elliptic curve and (P, Q) is a pair of points of E that generates the N-torsion subgroup $E[N]$ with Weil pairing $e_N(P, Q) = e^{2\pi i/N}$. Two such pairs $(E, (P, Q))$ and $(E', (P', Q'))$ are equivalent, denoted

$$(E, (P, Q)) \sim (E', (P', Q')),$$

if there exists some isomorphism $E \to E'$ taking P to P' and Q to Q'. The set of equivalence classes is denoted

$$S(N) = \{\text{enhanced elliptic curves for } \Gamma(N)\}/\sim .$$

An element of $S(N)$ is denoted $[E, (P, Q)]$.

The sets $S_0(N)$, $S_1(N)$ and $S(N)$ are *moduli spaces* of isomorphism classes of complex elliptic curves enhanced by associated N-torsion data. Observe that when $N = 1$, the three moduli spaces reduce to the isomorphism class of complex elliptic curves, since the N-torsion data do not play any role.

For each $\tau \in \mathbb{H}$, let E_τ denote the elliptic curve $\mathbb{C}/\Lambda_\tau$, where $\Lambda_\tau = \tau\mathbb{Z} + \mathbb{Z}$.

Lemma 3.4. *Let $E = \mathbb{C}/\Lambda$ be a complex elliptic curve. Then there exists $\tau \in \mathbb{H}$ such that E is isomorphic to E_τ as Riemann surfaces.*

Proof: Let (w_1, w_2) be a basis of Λ, with $w_1/w_2 \in \mathbb{H}$. Then

$$(1/w_2)\Lambda = \Lambda_\tau, \quad \text{where } \tau = w_1/w_2.$$

As a consequence, the map

$$z + \Lambda \in E \mapsto (1/w_2)z + \Lambda_\tau \in E_\tau$$

is an isomorphism between the complex elliptic curves E and E_τ.

□

Theorem 3.5. *Let N be a positive integer.*

1. *The moduli space for $\Gamma_0(N)$ is*

$$S_0(N) = \{[E_\tau, \langle 1/N + \Lambda_\tau\rangle] \mid \tau \in \mathbb{H}\}.$$

Two points $[E_\tau, \langle 1/N+\Lambda_\tau\rangle]$ and $[E_{\tau'}, \langle 1/N+\Lambda_{\tau'}\rangle]$ are equal if and only if $\Gamma_0(N)\tau = \Gamma_0(N)\tau'$. Thus there exists a bijection

$$\psi_0 : S_0(N) \to Y_0(N), \quad [E_\tau, \langle 1/N + \Lambda_\tau\rangle] \mapsto \Gamma_0(N)\tau.$$

2. *The moduli space for $\Gamma_1(N)$ is*

$$S_1(N) = \{[E_\tau, 1/N + \Lambda_\tau] \mid \tau \in \mathbb{H}\}.$$

Two points $[E_\tau, 1/N + \Lambda_\tau]$ and $[E_{\tau'}, 1/N + \Lambda_{\tau'}]$ are equal if and only if $\Gamma_1(N)\tau = \Gamma_1(N)\tau'$. Thus there exists a bijection

$$\psi_1 : S_1(N) \to Y_1(N), \quad [E_\tau, 1/N + \Lambda_\tau] \mapsto \Gamma_1(N)\tau.$$

3. *The moduli space for $\Gamma(N)$ is*

$$S(N) = \{[E_\tau, (\tau/N + \Lambda_\tau, 1/N + \Lambda_\tau)] \mid \tau \in \mathbb{H}\}.$$

Two points $[E_\tau, (\tau/N+\Lambda_\tau, 1/N+\Lambda_\tau)]$ and $[E_{\tau'}, (\tau'/N+\Lambda_{\tau'}, 1/N+\Lambda_{\tau'})]$ are equal if and only if $\Gamma(N)\tau = \Gamma(N)\tau'$. Thus there exists a bijection

$$\psi : S(N) \to Y(N), \quad [E_\tau, (\tau/N + \Lambda_\tau, 1/N + \Lambda_\tau)] \mapsto \Gamma(N)\tau.$$

Proof: We only prove the third case that is perhaps the more complicated. Let $[E,(P,Q)]$ be a point of $S(N)$. Since E is isomorphic to $E_{\tau'}$ for some $\tau' \in \mathbb{H}$ (Lemma 3.4), we can suppose without loss of generality that $E = E_{\tau'}$. Thus

$$P = (a\tau' + b)/N + \Lambda_{\tau'} \quad \text{and} \quad Q = (c\tau' + d)/N + \Lambda_{\tau'}, \quad \text{with } a,b,c,d \in \mathbb{Z}.$$

As P and Q generate $E_{\tau'}[N]$ and have Weil pairing $e_N(P,Q) = e^{2\pi i/N}$, the matrix $\alpha = \left[\begin{smallmatrix} \bar{a} & \bar{b} \\ \bar{c} & \bar{d} \end{smallmatrix}\right] \in \mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})$. Indeed,

$$\begin{bmatrix} P \\ Q \end{bmatrix} = \alpha \begin{bmatrix} \tau'/N + \Lambda_{\tau'} \\ 1/N + \Lambda_{\tau'} \end{bmatrix} \quad \text{and} \quad e_N(P,Q) = e^{2\pi i \det(\alpha)/N}.$$

In Section 1.2 we proved that $\mathrm{SL}_2(\mathbb{Z})$ surjects to $\mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})$, therefore we can also suppose that $\left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right] \in \mathrm{SL}_2(\mathbb{Z})$ (since this does not affect P and Q). Let $\tau = \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right](\tau') \in \mathbb{H}$ and $m = c\tau' + d$. Then $m\tau = a\tau' + b$, so

$$m\Lambda_\tau = m(\tau\mathbb{Z} + \mathbb{Z}) = (a\tau' + b)\mathbb{Z} + (c\tau' + d)\mathbb{Z} = \tau'\mathbb{Z} + \mathbb{Z} = \Lambda_{\tau'},$$

$$m(\tau/N + \Lambda_\tau) = P \quad \text{and} \quad m(1/N + \Lambda_\tau) = Q.$$

This proves that $[E_{\tau'},(P,Q)]$ and $[E_\tau,(\tau/N+\Lambda_\tau, 1/N+\Lambda_\tau)]$ are equal points.

Suppose now that two points $\tau,\tau' \in \mathbb{H}$ are $\Gamma(N)$-equivalent, i.e.,

$$\tau = \gamma(\tau'), \quad \text{for some } \gamma = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \Gamma(N).$$

Letting $m = c\tau' + d$ as before, we obtain that $m\Lambda_\tau = \Lambda_{\tau'}$,

$$m(1/N + \Lambda_\tau) = 1/N + \Lambda_{\tau'} \quad \text{and} \quad m(1/N + \Lambda_\tau) = 1/N + \Lambda_{\tau'},$$

since $\gamma \equiv I \pmod{N}$ (the matrix congruence is interpreted by entries).

Reciprocally, suppose that

$$[E_\tau,(\tau/N + \Lambda_\tau, 1/N + \Lambda_\tau)] \quad \text{and} \quad [E_{\tau'},(\tau'/N + \Lambda_{\tau'}, 1/N + \Lambda_{\tau'})],$$

with $\tau,\tau' \in \mathbb{H}$, are equal points. Then for some $m \in \mathbb{C}$, $m\Lambda_\tau = \Lambda_{\tau'}$,

$$m(\tau/N + \Lambda_\tau) = \tau'/N + \Lambda_{\tau'} \quad \text{and} \quad m(1/N + \Lambda_\tau) = 1/N + \Lambda_{\tau'},$$

As a consequence of the equality $m\Lambda_\tau = \Lambda_{\tau'}$, observe that

$$\begin{bmatrix} m\tau \\ m \end{bmatrix} = \gamma \begin{bmatrix} \tau' \\ 1 \end{bmatrix}, \quad \text{for some } \gamma = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathrm{SL}_2(\mathbb{Z}),$$

so in particular $m = c\tau' + d$. Using this the other equalities become

$$(a\tau' + b)/N + \Lambda_{\tau'} = \tau'/N + \Lambda_{\tau'} \quad \text{and} \quad (c\tau' + d)/N + \Lambda_{\tau'} = 1/N + \Lambda_{\tau'},$$

showing that $\gamma \equiv I \pmod{N}$. Therefore, $\Gamma(N)\tau = \Gamma(N)\tau'$, since $\tau = \gamma(\tau')$. □

In the special case $N = 1$, the previous theorem shows that the space of isomorphism classes of complex elliptic curves parametrizes the modular curve

$$\mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H} = Y_0(1) = Y_1(1) = Y(1).$$

Recall that the modular function

$$j : \mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H} \to \mathbb{C}$$

is a bijective function (see Remark 2.18). Therefore, we can associate to each isomorphism class of elliptic curves a complex number, the value of j at the corresponding orbit $\mathrm{SL}_2(\mathbb{Z})\tau \in \mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H}$. This value is also associated to any elliptic curve E in the isomorphism class and is denoted $j(E)$.

3.3 Complex analytic version of the Modularity Theorem

The complex analytic version of the Modularity Theorem states that the elliptic curves with rational j-values come from the modular curves

$$X_0(N), \quad N \in \mathbb{Z}^+,$$

via surjective morphisms of Riemann surfaces.

Theorem 3.6 (Modularity Theorem, Complex analytic version)**.** *Let E be a complex elliptic curve with $j(E) \in \mathbb{Q}$. Then for some positive integer N there exists a surjective morphism of Riemann surfaces from the modular curve $X_0(N)$ to the elliptic curve E,*

$$X_0(N) \longrightarrow E.$$

The surjection in the theorem is called a *modular parametrization of E of level N*. Observe that if Γ is a congruence subgroup of $\mathrm{SL}_2(\mathbb{Z})$ such that $\Gamma \subset \Gamma_0(N)$ (e.g., $\Gamma(N)$ or $\Gamma_1(N)$), then the composition of the natural projection from $X(\Gamma)$ to $X_0(N)$ with a modular parametrization of E,

$$X(\Gamma) \longrightarrow X_0(N) \longrightarrow E$$

is also a surjective morphism of Riemann surfaces (Theorem 1.26).

Appendix A

Riemann surfaces

A Riemann surface is simply a nonempty connected Hausdorff topological space endowed with an equivalence class of complex atlases. In this chapter we introduce all the theory of Riemann surfaces we need in this work:

- Holomorphic maps

- Meromorphic differentials

- Divisors and Riemann-Roch Theorem

A good reference book about Riemann surfaces is for example [Mir95].

A.1 Basic definitions

Let X be a topological space. We want X to behave locally as an open subset of the complex plane, thus it will allow us to define complex coordinates at each point of X.

Definition A.1. *A complex chart, or simply chart, on X is a homeomorphism $\phi : U \to V$, where $U \subset X$ is an open subset of X and $V \subset \mathbb{C}$ is an open subset of $\mathbb{C}$.*

A complex chart $\phi : U \to V$ is denoted by (U, ϕ). The open subset U is called domain of the chart ϕ. Furthermore, the chart ϕ is centered at $p \in U$ if $\phi(p) = 0$.

Examples A.2.

1. In the euclidean plane, consider any open subset $U \subset \mathbb{R}^2$. The map $\phi_U : U \to U$, $\phi_U(x, y) = x + iy$, is a chart on $\mathbb{R}^2$.

2. In the Riemann sphere, consider the open subsets $U_1 = \mathbb{C}$, $U_2 = \widehat{\mathbb{C}} \setminus \{0\}$. The maps $\phi_i : U_i \to \mathbb{C}$,
$$\phi_1(z) = z, \quad \phi_2(z) = \frac{1}{z} \quad (\text{where } \frac{1}{\infty} := 0),$$
are charts on $\widehat{\mathbb{C}}$.

Two complex charts $\phi_1 : U_1 \to V_1$ and $\phi_2 : U_2 \to V_2$ on X are compatible if either $U_1 \cap U_2 = \emptyset$ or $U_1 \cap U_2 \neq \emptyset$ and
$$\phi_2 \circ \phi_1^{-1} : \underbrace{\phi_1(U_1 \cap U_2)}_{V_{1,2}} \to \underbrace{\phi_2(U_1 \cap U_2)}_{V_{2,1}}$$
is holomorphic. Note that the definition is symmetric, i.e., if $U_1 \cap U_2 \neq \emptyset$ and
$$\phi_2 \circ \phi_1^{-1} : \phi_1(U_1 \cap U_2) \to \phi_2(U_1 \cap U_2)$$
is holomorphic, then
$$\phi_1 \circ \phi_2^{-1} : \phi_2(U_1 \cap U_2) \to \phi_1(U_1 \cap U_2)$$
is holomorphic. As a consequence,
$$(\phi_2 \circ \phi_1^{-1})'(z) \neq 0, \quad \forall\, z \in \phi_1(U_1 \cap U_2).$$
The function $\phi_2 \circ \phi_1^{-1}$ is called transition function between the two charts.

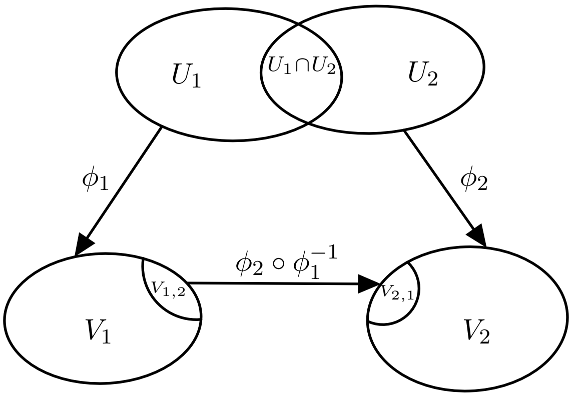


Figure A.1: Transition function

Definition A.3. *A complex atlas, or simply atlas, $\mathcal{A}$ on X is a collection*
$$\{\phi_\alpha : U_\alpha \to V_\alpha\}$$
of pairwise compatible complex charts whose domains cover X, i.e., $X = \bigcup U_\alpha$.

Examples A.4.

1. In the euclidean plane, the collection of charts
$$\mathcal{A} = \{\phi_U : U \to U \,|\, U \subset \mathbb{R}^2 \text{ is an open subset}\}$$
is an atlas on $\mathbb{R}^2$.

2. In the Riemann sphere, the collection of charts
$$\mathcal{A} = \{\phi_i : U_i \to \mathbb{C} \,|\, i = 1, 2\}$$
is an atlas on $\widehat{\mathbb{C}}$. Indeed, $U_1 \cap U_2 = \mathbb{C}^*$ and $\phi_2 \circ \phi_1^{-1}(z) = 1/z$, $\forall\, z \in \mathbb{C}^*$.

Two complex atlases on X are equivalent if every chart of one is compatible with every chart of the other. Note that two complex atlases are equivalent if and only if their union is also a complex atlas. Moreover, every complex atlas $\mathcal{A}$ on X is contained in a unique maximal complex atlas $\mathcal{A}^*$ on X which consists of all charts on X that are compatible with every chart of $\mathcal{A}$. Therefore, two complex atlases are equivalent if and only if they are both contained in the same maximal complex atlas.

Definition A.5. *A complex structure on X is an equivalence class of complex atlases on X, or equivalently, a maximal complex atlas on X.*

As any complex atlas on X determines a unique complex structure on X, this will be the usual way to define a complex structure on X.

Definition A.6. *A Riemann surface is a nonempty connected Hausdorff topological space endowed with a complex structure.*

Convention. If X is a Riemann surface, then by a chart on X we always mean a chart belonging to the maximal atlas of the complex structure on X.

Remarks A.7.

- ▷ A domain in a Riemann surface is a nonempty connected open subset. Note that these subsets inherit naturally the structure of Riemann surface.
- ▷ Each point of a Riemann surface has an open neighbourhood which is homeomorphic to an open disc of the complex plane. As a consequence, the topological local properties of the euclidean plane are preserved. For example, any Riemann surface is locally path-connected, locally compact, locally contractible, locally metrizable…
- ▷ According to Radó's Theorem [NR11, p.71], every Riemann surface is second-countable, i.e., there exists a countable base for its topology.

- ▷ Classically, a compact Riemann surface is called closed while a noncompact surface is called open. It is important to note that there exist notable differences between the theory of compact Riemann surfaces and that of noncompact Riemann surfaces.

- ▷ Every Riemann surface is an orientable connected 2-dimensional smooth manifold, since any complex atlas is an oriented smooth atlas. Therefore, every compact Riemann surface is diffeomorphic to a torus with g holes, for some unique integer $g \geq 0$. This integer g is called the genus of the Riemann surface and is a topological invariant.

Examples A.8.

1. The euclidean plane endowed with the complex structure determined by the atlas $\mathcal{A} = \{\phi_{\mathbb{R}^2} : \mathbb{R}^2 \to \mathbb{R}^2\}$ is a noncompact Riemann surface called Complex Plane. It is denoted by $\mathbb{C}$.

2. The Riemann sphere endowed with the complex structure determined by the atlas $\mathcal{A} = \{\phi_i : U_i \to \mathbb{C} \,|\, i = 1, 2\}$ is a compact Riemann surface called Riemann Sphere. It is denoted by $\widehat{\mathbb{C}}$.

A.2 Morphisms

Let X, Y, Z be Riemann surfaces.

Definition A.9. *Let U be a nonempty open subset of X.*

- ▷ *A map $F : U \to Y$ is holomorphic at $p \in U$ if there exist charts $\phi_1 : U_1 \to V_1$ on X, with $p \in U_1$, and $\phi_2 : U_2 \to V_2$ on Y, with $F(p) \in U_2$, such that the composition*

$$\phi_2 \circ F \circ \phi_1^{-1}$$

  is holomorphic at $\phi_1(p)$.

- ▷ *A map $F : U \to Y$ is holomorphic on a nonempty open subset $V \subseteq U$ if F is holomorphic at each point of V.*

- ▷ *A map $F : X \to Y$ is a morphism if F is holomorphic on X.*

Examples A.10.

1. The holomorphic functions on a nonempty open subset of $\mathbb{C}$.

2. The fractional linear transformations $T : \widehat{\mathbb{C}} \to \widehat{\mathbb{C}}$,

$$T(z) = \frac{az+b}{cz+d}, \quad \text{with } ad - bc \neq 0, \text{ are bijective morphisms.}$$

3. Let $\mathbb{C}/\Lambda_1, \mathbb{C}/\Lambda_2$ be complex tori and α a complex number such that $\alpha\Lambda_1 \subset \Lambda_2$. For each $\beta \in \mathbb{C}$, the map $F_{\alpha,\beta} : \mathbb{C}/\Lambda_1 \to \mathbb{C}/\Lambda_2$,
$$F_{\alpha,\beta}(z + \Lambda_1) = (\alpha z + \beta) + \Lambda_2, \quad \forall\, z + \Lambda \in \mathbb{C}/\Lambda_1,$$
is a morphism.

4. Let Γ_1, Γ_2 be congruence subgroups of $\mathrm{SL}_2(\mathbb{Z})$ such that $\Gamma_1 \subset \Gamma_2$. The natural projection of the corresponding compact modular curves
$$F : X(\Gamma_1) \to X(\Gamma_2), \quad \Gamma_1\tau \mapsto \Gamma_2\tau,$$
is a surjective morphism.

In the special case $Y = \mathbb{C}$, it is important to note that a function $f : U \to Y$ is holomorphic at $p \in U$ if and only if there exists a chart $\phi_1 : U_1 \to V_1$ on X, with $p \in U_1$, such that the composition $f \circ \phi_1^{-1}$ is holomorphic at $\phi_1(p)$.

The set of holomorphic functions on U is denoted by $\mathcal{O}_X(U)$,
$$\mathcal{O}_X(U) = \{f : U \to \mathbb{C} \mid f \text{ is holomorphic on } U\}.$$
It is a $\mathbb{C}$-algebra, since the constant functions are holomorhic functions and the sum and product of holomorhic functions are also holomorphic functions. Moreover, if U is connected, then $\mathcal{O}(U)$ is an integral domain.

The following proposition states the main results concerning holomorphic maps between Riemann surfaces.

Proposition A.11. *Let U be an open subset of X, $F : U \to Y$ a map and $p \in U$.*

1. *If F is holomorphic at p, then F is continuous at p.*

2. *If F is holomorphic at p, then F is holomorphic on an open subset $U_p \subset U$ with $p \in U_p$.*

3. *F is holomorphic at p if and only if for any pair of charts $\phi_1 : U_1 \to V_1$ on X, with $p \in U_1$, and $\phi_2 : U_2 \to V_2$ on Y, with $F(p) \in U_2$, the composition*
$$\phi_2 \circ F \circ \phi_1^{-1}$$
is holomorphic at $\phi_1(p)$.

4. *If F is continuous at p and there exists an open subset $U_p \subset U$, with $p \in U_p$, such that F is holomorphic on $U_p \backslash \{p\}$, then F is holomorphic at p.*

5. *Let V be an open subset of Y and $G : V \to Z$ a map such that $F(U) \subset V$. If F is holomorphic at p and G is holomorphic at $F(p)$, then $G \circ F$ is holomorphic at p.*

It is not difficult to prove that if $F : X \to Y$ is a bijective morphism, then $F^{-1} : Y \to X$ is also a morphism. A bijective morphism $F : X \to Y$ is called isomorphism. The Riemann surfaces X and Y are isomorphic if there exists an isomorphism $F : X \to Y$. Note that the relation of being isomorphic is an equivalence relation on Riemann surfaces, since the composition of morphisms is a morphism.

As usual, a self-isomorphism $F : X \to X$ is also called automorphism. The automorphisms of X form a group under the operation of composition.

A.2.1 Theorems on morphisms

Now we present several results about morphisms between Riemann surfaces. In the majority of cases these results are immediate consequences of the corresponding results of complex analysis about holomorphic functions.

Theorem A.12 (Identity Theorem)**.** *Let $F, G : X \to Y$ be morphisms. If there exists a subset $S \subset X$ such that $S' \neq \emptyset$ and $F(x) = G(x)$, $\forall\, x \in S$, then $F = G$.*

Theorem A.13 (Open Mapping Theorem)**.** *Let $F : X \to Y$ be a nonconstant morphism. Then F is an open map.*

Corollary A.14. *Let $F : X \to Y$ be a nonconstant morphism. If X is compact, then F is surjective and Y is compact.*

As a consequence, observe that if X is compact, then $\mathcal{O}(X) = \mathbb{C}$.

Corollary A.15 (Discreteness of Preimages)**.** *Let $F : X \to Y$ be a nonconstant morphism. Then $F^{-1}(q)$ is a discrete subset of X for all $q \in Y$. In particular, if X is compact, then $F^{-1}(q)$ is a nonempty finite subset of X for all $q \in Y$.*

Theorem A.16. *Let $F : X \to Y$ be an injective morphism. Then its restriction on its image $F : X \to F(X)$ is an isomorphism.*

Theorem A.17 (Local normal form)**.** *Let $F : X \to Y$ be a nonconstant morphism and $p \in X$. There exists a unique integer $m \geq 1$ which satisfies the following property: For each chart $\phi_2 : U_2 \to V_2$ on Y, centered at $F(p)$, there exists a chart $\phi_1 : U_1 \to V_1$ on X, centered at p, with $F(U_1) \subset U_2$, such that*

$$\phi_2 \circ F \circ \phi_1^{-1}(z) = z^m, \quad \forall\, z \in V_1.$$

Proof: [Mir95, p.44] □

Definition A.18. *Let $F : X \to Y$ be a nonconstant morphism. The ramification index of F at $p \in X$ is the unique integer $m \geq 1$ which satisfies the property cited in the previous theorem. It is denoted by $e_p(F)$.*

Examples A.19.

1. Let $f : \mathbb{C} \to \mathbb{C}$ be the holomorphic function $f(z) = z^m$, with $m \geq 1$. The index ramification of f at $z \in \mathbb{C}^*$ is $e_z(f) = 1$ and at 0 is $e_0(f) = m$.

2. Let $f : \widehat{\mathbb{C}} \to \widehat{\mathbb{C}}$ be the holomorphic function

$$f(z) = z^m + c_1 z^{m-1} \cdots + c_m \quad \text{(where } f(\infty) := \infty\text{)},$$

with $m \geq 1$. The index ramification of f at ∞ is $e_\infty(f) = m$.

There exists an easy way to compute the ramification index without having to find centered charts which put the morphism into local normal form. Let us fix two charts $\phi_1 : U_1 \to V_1$ on X, with $p \in U_1$, and $\phi_2 : U_2 \to V_2$ on Y, with $F(p) \in U_2$. As the composition

$$\phi_2 \circ F \circ \phi_1^{-1}$$

is holomorphic at $z_0 = \phi_1(p)$,

$$\phi_2 \circ F \circ \phi_1^{-1} = \sum_{n=1}^{\infty} a_n (z - z_0)^n.$$

Then

$$e_p(F) = \min\{n \geq 1 \,|\, a_n \neq 0\}.$$

Proposition A.20. *Let $F : X \to Y$ be a nonconstant morphism and $p \in X$. The following conditions are equivalent:*

1. *$e_p(F) = 1$*

2. *F is a local isomorphism at p, i.e., there exist connected open subsets $U_p \subset X$, with $p \in U_p$, and $V_{F(p)} \subset Y$, with $F(p) \in V_{F(p)}$, such that $F_{|U_p} : U_p \to V_{F(p)}$ is an isomorphism.*

Definition A.21. *Let $F : X \to Y$ be a nonconstant morphism. A point $p \in X$ is a ramification point of F if $e_p(F) > 1$. A point $y \in Y$ is a branch point of F if it is the image of a ramification point of F.*

As a consequence of the above, the ramification points of a nonconstant morphism $F : X \to Y$ form a discrete closed subset of X.

Definition A.22. *A morphism $F : X \to Y$ is a branched covering if for each $q \in Y$ holds that*

- ▷ *$F^{-1}(q)$ is a nonempty finite subset of X, $F^{-1}(q) = \{p_1, \dots, p_r\}$, and*
- ▷ *there exist charts $\phi_i : U_i \to V_i$ on X, centered at p_i, $\phi : U \to V$ on Y, centered at q, with*

$$F^{-1}(U) = \bigsqcup_{i=1}^{r} U_i,$$

and integers $e_i \geq 1$, such that for all $i = 1, \dots, r$,

$$\phi \circ F \circ \phi_i^{-1}(z) = z^{e_i}, \qquad \forall\, z \in V_i.$$

Note that the integers e_i are the ramification indexes of the points p_i.

Examples A.23.

1. The holomorphic function $f : \mathbb{C} \to \mathbb{C}$, $f(z) = z^m$, with $m \geq 1$, is a branched covering.
2. The holomorphic function $f : \widehat{\mathbb{C}} \to \widehat{\mathbb{C}}$, $f(z) = z^m + c_1 z^{m-1} \cdots + c_m$, with $m \geq 1$, is a branched covering.

The following theorem is really useful to prove that a morphism is a branched covering since it characterizes the branched coverings in a simple way.

Theorem A.24 (Characterization of Branched Coverings)**.** *Let $F : X \to Y$ be a nonconstant morphism. The following conditions are equivalent:*

1. F is a branched covering.

2. For each $q \in Y$, $F^{-1}(q)$ is a nonempty finite subset of X, and in addition, the map

$$q \in Y \mapsto \sum_{p \in F^{-1}(q)} e_p(F) \in \mathbb{Z}$$

is constant.

3. F is a proper map.

In particular, observe that if X is compact, then F is a branched covering.

Proof: [Gir70, p.7]

□

Since any proper map between locally compact topological spaces is closed, we deduce from this result that any branched covering is a closed map.

Definition A.25. *Let $F : X \to Y$ be a branched covering. The number*

$$\deg(F) = \sum_{p \in F^{-1}(q)} e_p(F)$$

is called the degree of the branched covering F.

Proposition A.26. *Let $F : X \to Y$, $G : Y \to Z$ be branched coverings. Then the composition $G \circ F$ is a branched covering and*

$$\deg(G \circ F) = \deg(G)\deg(F).$$

Proposition A.27. *Let $F : X \to Y$ be a branched covering and*

$$R = \{p \in X \,|\, e_p(F) > 1\}.$$

Then $F(R)$ is a closed discrete subset of Y. Moreover, R is empty if and only if F is an étale covering.

Finally, we present the Riemann-Hurwitz formula which relates the genus of two compact Riemann surfaces through a nonconstant morphism between each other.

Theorem A.28 (Riemann-Hurwitz formula)**.** *Let $F : X \to Y$ be a nonconstant morphism. If X is compact, then*

$$2g(X) - 2 = \deg(F)(2g(Y) - 2) + \sum_{p \in X}(e_p(F) - 1).$$

Proof: [Mir95, p.52]

□

A.2.2 Meromorphic functions

Definition A.29. *Let U be a nonempty open subset of X.*

- ▷ *A function $f : U \to \widehat{\mathbb{C}}$ is meromorphic at $p \in U$ if there exists a chart $\phi_1 : U_1 \to V_1$ on X, with $p \in U_1$, such that the composition $f \circ \phi_1^{-1}$ is meromorphic at $\phi_1(p)$. In this case, for any chart $\psi_1 : U_1' \to V_1'$ on X, with $p \in U_1'$, the composition $f \circ \psi_1^{-1}$ is meromorphic at $\psi_1(p)$.*
- ▷ *A function $f : U \to \widehat{\mathbb{C}}$ is meromorphic on a nonempty open subset $V \subset U$ if f is meromorphic at each point of V.*

Note that a function $f : U \to \widehat{\mathbb{C}}$ is meromorphic at $p \in U$ if and only if f is holomorphic at p (as Riemann surfaces) and there does not exist a neighbourhood U_p of p, with $U_p \subset U$, such that f is identically infinity on U_p.

Examples A.30.

1. The meromorphic functions on a nonempty open subset of $\mathbb{C}$.
2. Let $\mathbb{C}/\Lambda$ be a complex torus. For each elliptic function f with periods Λ, the induced function $F : \mathbb{C}/\Lambda \to \widehat{\mathbb{C}}$,
$$F(z+\Lambda) = f(z), \quad \forall\, z+\Lambda \in \mathbb{C}/\Lambda,$$
is a meromorphic function on $\mathbb{C}/\Lambda$.
3. Let Γ be a congruence subgroup of $\mathrm{SL}_2(\mathbb{Z})$. For each automorphic form f of weigh 0 with respect to Γ, the induced function $F : X(\Gamma) \to \widehat{\mathbb{C}}$,
$$F(x) = \begin{cases} f(\tau) & \text{if } x = \Gamma\tau, \\ f(s) & \text{if } x = \Gamma s, \end{cases} \qquad \forall\, x \in X(\Gamma),$$
is a meromorphic function on $X(\Gamma)$.

The set of meromorphic functions on U is denoted by $\mathcal{M}_X(U)$,
$$\mathcal{M}_X(U) = \{f : U \to \mathbb{C} \,|\, f \text{ is meromorphic on } U\}.$$
Note that it is a $\mathbb{C}$-algebra that contains $\mathcal{O}_X(U)$ as subalgebra. Furthermore, if U is connected, then $\mathcal{M}_X(U)$ is a field, since if f is a meromorphic function which is not identically zero, then $1/f$ is also a meromorphic function.

Definition A.31. *Let $f : X \to \widehat{\mathbb{C}}$ be a nonzero meromorphic function on X. The order of f at $p \in X$ is defined as*
$$\mathrm{ord}_p(f) = \min\{n \in \mathbb{Z} \,|\, a_n \neq 0\},$$
where $\{a_n\}_{n\in\mathbb{Z}}$ is the sequence of Laurent coefficients of f around p with respect to a chart $\phi_1 : U_1 \to V_1$ on X, with $p \in U_1$, i.e.,
$$f \circ \phi_1^{-1}(z) = \sum_{n\in\mathbb{Z}} a_n (z-z_0)^n, \quad z_0 = \phi_1(p).$$

One can easily check that this definition is independent of the chosen chart to define the coefficients of the Laurent series.

We say that the function f has a zero of order n at $p \in X$ if $\mathrm{ord}_p(f) = n \geq 1$ and has a pole of order n at $p \in X$ if $\mathrm{ord}_p(f) = -n \leq -1$.

Examples A.32.

1. The function induced by the Weierstrass $\wp$-function for a lattice Λ in $\mathbb{C}$
$$\wp_\Lambda : \mathbb{C} \to \widehat{\mathbb{C}}, \quad \wp_\Lambda(z) = \frac{1}{z^2} + {\sum_{w\in\Lambda}}' \left(\frac{1}{(z-w)^2} - \frac{1}{w^2} \right), \quad \forall\, z \in \mathbb{C}, z \notin \Lambda,$$
has a double pole at $0 + \Lambda$.

2. The function induced by the modular function

$$j : \mathbb{H} \to \mathbb{C}, \quad j(\tau) = 1728\frac{g_2(\tau)^3}{\Delta(\tau)}, \quad \forall\, \tau \in \mathbb{H},$$

has a simple pole at $\mathrm{SL}_2(\mathbb{Z})\infty$.

Convention. If $f : X \to \widehat{\mathbb{C}}$ is identically zero, then $\mathrm{ord}_p(f) = +\infty, \forall\, p \in X$.

Remarks A.33.

▷ Let f be a nonzero meromorphic function on X:

- If f has a zero at $p \in X$, then $e_p(f) = \mathrm{ord}_p(f)$.
- If f has a pole at $p \in X$, then $e_p(f) = -\mathrm{ord}_p(f)$.
- If f has neither a zero nor a pole at $p \in X$, then

$$e_p(f) = \mathrm{ord}_p(f - f(p)).$$

▷ Let f, g be nonzero meromorphic functions on X. Then

$$\begin{array}{ll} \mathrm{ord}_p(f \pm g) \geq \min\{\mathrm{ord}_p(f), \mathrm{ord}_p(g)\}, & \mathrm{ord}_p(1/f) = -\mathrm{ord}_p(f), \\ \mathrm{ord}_p(fg) = \mathrm{ord}_p(f) + \mathrm{ord}_p(g), & \mathrm{ord}_p(f/g) = \mathrm{ord}_p(f) - \mathrm{ord}_p(g). \end{array}$$

The following theorems are immediate consequences of the characterization of branched coverings, since if $f : X \to \widehat{\mathbb{C}}$ is a branched covering, then the map

$$q \in \widehat{\mathbb{C}} \mapsto \sum_{p \in f^{-1}(q)} e_p(f) \in \mathbb{Z}$$

is constant.

Theorem A.34. *Let $f : X \to \widehat{\mathbb{C}}$ be a branched covering. If f has a unique simple pole, then f is an isomorphism. Therefore, X is isomorphic to $\widehat{\mathbb{C}}$.*

Theorem A.35. *Let $f : X \to \widehat{\mathbb{C}}$ be a branched covering. Then*

$$\sum_{p \in X} \mathrm{ord}_p(f) = (\textit{number of zeros of } f) - (\textit{number of poles of } f) = 0.$$

A.3 Meromorphic differentials

Let X be a Riemann surface and n an integer.

Definition A.36. *A local representation of a meromorphic differential on X of degree n is a collection of pairs*

$$(\phi_\alpha, f_\alpha),$$

where $\phi_\alpha : U_\alpha \to V_\alpha$ is a chart on X and $f_\alpha : V_\alpha \to \widehat{\mathbb{C}}$ is a meromorphic function, that satisfies the following conditions:

▷ *$\{\phi_\alpha : U_\alpha \to V_\alpha\}$ is a complex atlas on X,*

▷ *and in addition, if $U_{\alpha_1} \cap U_{\alpha_2} \neq \emptyset$, then*

$$f_{\alpha_1}(z) = f_{\alpha_2}(\phi_{\alpha_2,\alpha_1}(z))(\phi'_{\alpha_2,\alpha_1}(z))^n, \quad \forall\, z \in \phi_{\alpha_1}(U_{\alpha_1} \cap U_{\alpha_2}),$$

where ϕ_{α_2,α_1} is the transition function from the chart ϕ_{α_1} to the chart ϕ_{α_2}.

Two local representations of meromorphic differentials on X of degree n,

$$\{(\phi_\alpha, f_\alpha)\} \quad \text{and} \quad \{(\phi_{\alpha'}, f_{\alpha'})\}$$

are equivalent (or represent the same meromorphic differential of degree n) if its union

$$\{(\phi_\alpha, f_\alpha)\} \cup \{(\phi_{\alpha'}, f_{\alpha'})\}$$

also satisfies the second condition of the previous definition.

Let $\{(\phi_\alpha, f_\alpha)\}$ be a local representation of a meromorphic differential of degree n. If $\{\phi_{\alpha'} : U_{\alpha'} \to V_{\alpha'}\}$ is another complex atlas, then we can define

$$f_{\alpha'}(z) = f_\alpha(\phi_{\alpha,\alpha'}(z))(\phi'_{\alpha,\alpha'}(z))^n, \quad \forall\, z \in \phi_{\alpha'}(U_\alpha \cap U_{\alpha'}),$$

where $\phi_{\alpha,\alpha'}$ is the transition function from the chart $\phi_{\alpha'}$ to the chart ϕ_α. As a consequence, observe that the representations

$$\{(\phi_\alpha, f_\alpha)\} \quad \text{and} \quad \{(\phi_{\alpha'}, f_{\alpha'})\}$$

are equivalent.

Definition A.37. *A meromorphic differential w on X of degree n is a equivalence class of local representations of meromorphic differentials of degree n.*

Although a meromorphic differential w of degree n is a equivalence class, we keep the same notation

$$w = \{(\phi_\alpha, f_\alpha)\}$$

if there is no risk of confusion.

The set of meromorphic differentials on X of degree n is denoted $\mathcal{M}^{(n)}(X)$. It is a $\mathbb{C}$-vector space together with the operations

$$w + w' = \{(\phi_\alpha, f_\alpha + g_\alpha)\}, \quad w = \{(\phi_\alpha, f_\alpha)\},\ w' = \{(\phi_\alpha, g_\alpha)\}$$

and

$$\lambda w = \{(\phi_\alpha, \lambda f_\alpha)\}, \quad w = \{(\phi_\alpha, f_\alpha)\},\ \lambda \in \mathbb{C}.$$

Observe that these operations are independent of the local representations.

If $w = \{(\phi_\alpha, f_\alpha)\} \in \mathcal{M}^{(s)}(X)$ and $w' = \{(\phi_\alpha, g_\alpha)\} \in \mathcal{M}^{(j)}(X)$, then the product

$$ww' = \{(\phi_\alpha, f_\alpha g_\alpha)\} \in \mathcal{M}^{(s+j)}$$

is well-defined. Thus the direct sum over all the degrees

$$\bigoplus_{n\in\mathbb{Z}} \mathcal{M}^{(n)}(X)$$

forms naturally a graded ring.

Remark A.38. The space $\mathcal{M}^{(0)}$ is a $\mathbb{C}$-algebra together with this product. Furthermore, the map

$$f \in \mathcal{M}(X) \mapsto \{(\phi_\alpha, f \circ \phi_\alpha^{-1})\} \in \mathcal{M}^{(0)}, \quad \text{where } \{\phi_\alpha\} \text{ is an atlas on } X,$$

defines an isomorphism of $\mathbb{C}$-algebras.

Examples A.39.

1. The collection
$$\{(\phi_{\mathbb{R}^2}, f)\}, \quad \text{with } f \in \mathcal{M}(\mathbb{C}),$$
is a meromorphic differential on the complex plane of any degree $n \in \mathbb{Z}$.
2. The collection
$$\{(\phi_1, f_1 = 1), (\phi_2, f_2(z) = (-1/z^2)^n)\}$$
is a meromorphic differential on the Riemann sphere of any degree $n \in \mathbb{Z}$.
3. Let $\mathbb{C}/\Lambda$ be a complex torus. The collection
$$\{(\phi_z : \pi(D_z) \to D_z, f_z = 1) \,|\, z \in \mathbb{C}\}$$
is a meromorphic differential on the complex torus of any degree $n \in \mathbb{Z}$.

Definition A.40. *Let $w = \{(\phi_\alpha, f_\alpha)\}$ be a meromorphic differential on X of degree n. The order of w at $p \in X$ is defined as*

$$\mathrm{ord}_p(w) = \nu_{z_p}(f_\alpha),$$

where $(\phi_\alpha : U_\alpha \to V_\alpha, f_\alpha)$ is a pair in w, with $p \in U_\alpha$, and $z_p = \phi_\alpha(p)$.

One can easily check that this definition is independent of the chosen pair, i.e, if $(\phi_{\alpha'}\!:\!U_{\alpha'}\!\rightarrow\!V_{\alpha'}, f_{\alpha'})$ is another pair in w, with $p \in U_{\alpha'}$, then

$$\nu_{z_p}(f_\alpha) = \nu_{z'_p}(f_{\alpha'}), \quad \text{where } z'_p = \phi_{\alpha'}(p).$$

A.4 Divisors and Riemann-Roch Theorem

Let X be a compact Riemann surface.

Definition A.41. *A divisor D on X is a map from X to $\mathbb{Z}$ whose support*

$$\{p \in X \mid D(p) \neq 0\}$$

is a finite subset of X.

A divisor D on X is denoted

$$D = \sum n_p p, \quad \text{where } n_p = D(p), \ \forall\, p \in X.$$

The set $\operatorname{Div}(X)$ of divisors on X forms a Abelian group, the free abelian group generated by the set X, under the addition

$$D + D' = \sum (n_p + n'_p)p, \quad D = \sum n_p p, \ D' = \sum n'_p p.$$

The degree of a divisor $D = \sum n_p p$ on X is defined as $\deg(D) = \sum n_p$. Observe that the map $\deg : \operatorname{Div}(X) \to \mathbb{Z}$ is a homomorphism of groups. Its kernel is the subgroup

$$\operatorname{Div}_0(X) = \{D \in \operatorname{Div}(X) \mid \deg(D) = 0\}.$$

Each meromorphic function $f : X \to \widehat{\mathbb{C}}$ defines a divisor on X,

$$\operatorname{div}(f) = \sum \operatorname{ord}_p(f)p \quad (f \neq 0).$$

The divisors D on X of the form $D = \operatorname{div}(f)$ are called principal divisors. Let $\mathcal{M}(X)^*$ be the multiplicative group of nonzero meromorphic functions. Observe that the map $\operatorname{div} : \mathcal{M}(X)^* \to \operatorname{Div}(X)$ is also a homomorphism of groups, since

$$\operatorname{div}(fg) = \operatorname{div}(f) + \operatorname{div}(g), \quad \forall\, f, g \in \mathcal{M}(X)^*.$$

By Theorem A.35 its image

$$\operatorname{PDiv}(X) = \{\operatorname{div}(f) \mid f \in \mathcal{M}(X)^*\}$$

is a subgroup of the group $\mathrm{Div}_0(X)$. Abel's Theorem [Mir95, p.250-263] states that the quotient $\mathrm{Div}_0(X)/\mathrm{PDiv}(X)$ is isomorphic to a g-dimensional complex torus $\mathbb{C}^g/\Lambda_g$, where g is the genus of X and Λ_g is a lattice in $\mathbb{C}^g$.

Let $D = \sum n_p p$, $D' = \sum n'_p p$ be divisors on X. If $n'_p \leq n_p$, $\forall\, p \in X$, then we write

$$D' \leq D \quad \text{or} \quad D \geq D'.$$

The linear space $L(D)$ associated to a divisor D on X is defined as

$$L(D) = \{f \in \mathcal{M}(X) \,|\, f = 0 \ \text{ or } \ \mathrm{div}(f) + D \geq 0\}.$$

The inequality

$$\mathrm{ord}_p(f+g) \geq \min\{\mathrm{ord}_p(f), \mathrm{ord}_p(g)\}, \quad \forall\, p \in X,\ \forall\, f, g \in \mathcal{M}(X),$$

shows that $L(D)$ is a vector subspace. It is well-known that this subspace turns out to be finite-dimensional [Mir95, p.151]. Its dimension is denoted $\ell(D)$.

Remarks A.42.

▷ If D is the zero divisor, then

$$L(D) = \mathbb{C},$$

since the meromorphic functions without poles are the constants functions.

▷ If D is a divisor with degree $\deg(D) < 0$, then $L(D) = \{0\}$, since

$$\deg(\mathrm{div}(f) + D) = \deg(D), \quad \forall\, f \in \mathcal{M}(X)^*.$$

Each meromorphic differential $w \in \mathcal{M}^{(n)}(X)$ also defines a divisor on X,

$$\mathrm{div}(w) = \sum \mathrm{ord}_p(w) p \quad (w \neq 0).$$

Observe that

$$\mathrm{div}(w_1 w_2) = \mathrm{div}(w_1) + \mathrm{div}(w_2)$$

for any nonzero meromorphic differentials $w_1 \in \mathcal{M}^{(s)}(X)$ and $w_2 \in \mathcal{M}^{(j)}(X)$. The divisors D on X of the form $D = \mathrm{div}(\lambda)$, with $\lambda \in \mathcal{M}^{(1)}(X)$, $\lambda \neq 0$, are called canonical divisors.

Lemma A.43. *Let $\lambda_1, \lambda_2 \in \mathcal{M}^{(1)}(X)$ be nonzero meromorphic differentials. Then there exists a unique nonzero meromorphic function $f : X \to \widehat{\mathbb{C}}$ such that $\lambda_2 = f\lambda_1$. As a consequence, observe that*

$$\mathrm{div}(\lambda_2) = \mathrm{div}(f) + \mathrm{div}(\lambda_1).$$

Proof: Let $\{\phi_\alpha : U_\alpha \to V_\alpha\}$ be a complex atlas on X. There exist local representations

$$\lambda_1 = \{(\phi_\alpha, f_\alpha^1)\} \quad \text{and} \quad \lambda_2 = \{(\phi_\alpha, f_\alpha^2)\}.$$

Then the meromorphic function $f : X \to \widehat{\mathbb{C}}$,

$$f(p) = \frac{f_\alpha^2(\phi_\alpha(p))}{f_\alpha^1(\phi_\alpha(p))}, \quad \forall\, p \in U_\alpha,$$

is well-defined and is the unique nonzero meromorphic function on X that satisfies the desired property.

□

Theorem A.44. *[*Riemann-Roch*] Let X be a compact Riemann surface of genus g. If* $\operatorname{div}(\lambda)$ *is a canonical divisor on X, then for any divisor* $D \in \operatorname{Div}(X)$*,*

$$\ell(D) = \deg(D) - g + 1 + \ell(\operatorname{div}(\lambda) - D).$$

Proof: [Mir95, p.192]

□

Observe that if $\operatorname{div}(\lambda')$ is another canonical divisor on X, then the vector spaces $L_\lambda = L(\operatorname{div}(\lambda) - D)$ and $L_{\lambda'} = L(\operatorname{div}(\lambda') - D)$ are isomorphic, since the map

$$g \in L_\lambda \mapsto g/f \in L_{\lambda'}, \quad \text{where } f \in \mathcal{M}(X)^*,\ \lambda' = f\lambda,$$

is an isomorphism of $\mathbb{C}$-vector spaces. As a consequence,

$$\ell(\operatorname{div}(\lambda) - D) = \ell(\operatorname{div}(\lambda') - D).$$

Corollary A.45. *Let X, g,* $\operatorname{div}(\lambda)$ *and D be as above. Then*

1. $\ell(\operatorname{div}(\lambda)) = g$

2. $\deg(\operatorname{div}(\lambda)) = 2g - 2$

3. If $\deg(D) < 0$*, then* $\ell(D) = 0$

4. If $\deg(D) > 2g - 2$*, then* $\ell(D) = \deg(D) - g + 1$

Proof: To prove the first equality let $D = 0$ in the Riemann-Roch Theorem and recall that $\ell(0) = 1$. To prove the second equality it suffices to apply the Riemann-Roch Theorem with $D = \operatorname{div}(\lambda)$, since

$$g \overset{(1)}{=} \ell(\operatorname{div}(\lambda)) = \deg(\operatorname{div}(\lambda)) - g + 2.$$

We now argue the third assertion by contradiction. Let us suppose that there exists a nonzero meromorphic function $f \in L(D)$. Then $\mathrm{div}(f) \geq -D$, and taking degrees follows that $\deg(D) \geq 0$, which is a contradiction. Finally, the fourth assertion is consequence of *(2)* and *(3)* since if $\deg(D) > 2g-2$, then

$$\deg(\mathrm{div}(\lambda) - D) \overset{(2)}{=} 2g - 2 - \deg(D) < 0.$$

□

Appendix B

Complex elliptic curves

The compact Riemann surfaces of genus equal to 1 are called complex elliptic curves for reasons to be explained in this chapter. It is possible to prove that any complex elliptic curve is isomorphic to a complex torus (as Riemann surfaces). As a consequence, these curves can be endowed with an analytic group structure which is uniquely determined by the choice of identity element. In this chapter we detail all these results over complex elliptic curves.

B.1 Complex Tori

Let us begin by defining the Riemann surfaces called complex torus.

It is well-known that the discrete subgroups of $\mathbb{C}$ are

- $\{0\}$,
- $\mathbb{Z}w$, with $w \in \mathbb{C}^*$, and
- $\mathbb{Z}w_1 + \mathbb{Z}w_2$, with $w_1, w_2 \in \mathbb{C}$ linearly independent over $\mathbb{R}$.

A lattice Λ in $\mathbb{C}$ is a discrete subgroup of $\mathbb{C}$ of the third kind, that is, $\Lambda = \mathbb{Z}w_1 + \mathbb{Z}w_2$, with $w_1, w_2 \in \mathbb{C}$ linearly independent over $\mathbb{R}$. Note that the pair (w_1, w_2) has the property that any $w \in \Lambda$ has a unique representation

$$w = m_1 w_1 + m_2 w_2, \quad \text{with } m_1, m_2 \in \mathbb{Z}.$$

Any pair with this property is called a basis of the lattice Λ.

It is usual to make the normalizing convention $w_1/w_2 \in \mathbb{H}$, but this still does not determine a basis given a lattice. In fact, two pairs (w_1, w_2) and (w_1', w_2') are bases of a same lattice if and only if

$$\begin{bmatrix} w_1' \\ w_2' \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}, \quad \text{for some } \alpha = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathrm{GL}_2(\mathbb{Z}).$$

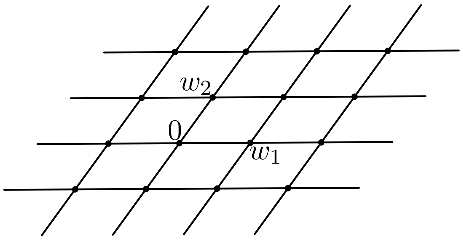


Figure B.1: Lattice with basic (w_1, w_2)

If we make the normalizing convention $w_1/w_2, w_1'/w_2' \in \mathbb{H}$, then $\alpha \in \mathrm{SL}_2(\mathbb{Z})$.

Let Λ be a lattice in $\mathbb{C}$, with basis (w_1, w_2),

$$\Lambda = \mathbb{Z}w_1 + \mathbb{Z}w_2.$$

Let us consider the quotient group $\mathbb{C}/\Lambda$ endowed with the topology induced by the natural projection $\pi : \mathbb{C} \to \mathbb{C}/\Lambda$, that is, a subset $U \subset \mathbb{C}/\Lambda$ is open if and only if $\pi^{-1}(U)$ is open in $\mathbb{C}$. Note that $\mathbb{C}/\Lambda$ is a compact connected topological space, since π is a continuous map with respect to this topology and $\pi(\overline{P_a}) = \mathbb{C}/\Lambda$, for any period parallelogram

$$P_a = P_a(w_1, w_2) = \{a + \lambda w_1 + \beta w_2 \,|\, 0 \leq \lambda, \beta < 1\}, \quad a \in \mathbb{C}.$$

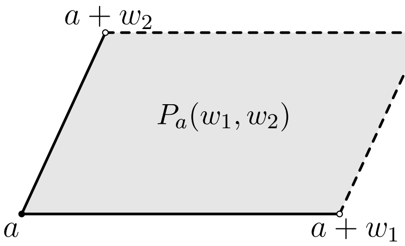


Figure B.2: Period parallelogram

Furthermore, π is an open map, since

$$\pi^{-1}(\pi(V)) = \bigcup_{w \in \Lambda} w + V, \quad \forall\, V \subset \mathbb{C}.$$

Let us now prove that $\mathbb{C}/\Lambda$ is a Hausdorff topological space. Let $z_1 + \Lambda$, $z_2 + \Lambda$ be points in $\mathbb{C}/\Lambda$, with $z_2 - z_1 \notin \Lambda$. Then

$$\eta = \min_{w \in \Lambda} |(z_2 - z_1) - w| > 0.$$

As a consequence,

$$\pi(D(z_1, \eta/2)) \quad \text{and} \quad \pi(D(z_2, \eta/2))$$

are disjoint open subsets of $\mathbb{C}/\Lambda$ containing $z_1 + \Lambda$ and $z_2 + \Lambda$, respectively.

Finally, let us define an atlas on $\mathbb{C}/\Lambda$. For each $z \in \mathbb{C}$, consider the disc $D_z = D(z, \delta/2)$, where

$$\delta = \min_{w \in \Lambda \setminus \{0\}} \{|w|\}.$$

Note that $\pi_{|D_z} : D_z \to \pi(D_z)$ is a homeomorphism. We denote by ϕ_z its inverse homeomorphism. Below, we show that the collection of charts

$$\mathcal{A} = \{\phi_z : \pi(D_z) \to D_z \,|\, z \in \mathbb{C}\}$$

is an atlas on $\mathbb{C}/\Lambda$:

- Evidently, $\mathbb{C}/\Lambda = \bigcup_{z \in \mathbb{C}} \pi(D_z)$.
- Furthermore, if $\pi(D_{z_1}) \cap \pi(D_{z_2}) \neq \emptyset$, with $z_1, z_2 \in \mathbb{C}$, then there exists $w \in \Lambda$ such that

  $$\phi_{z_2} \circ \phi_{z_1}^{-1}(z) = z + w, \quad \forall\, z \in V_{1,2},$$

  where $V_{1,2} = \phi_{z_1}(\pi(D_{z_1}) \cap \pi(D_{z_2}))$. Indeed,

  $$\varphi(z) := \phi_{z_2} \circ \phi_{z_1}^{-1}(z) - z \in \Lambda, \quad \forall\, z \in V_{1,2},$$

  and

  $$|\varphi(s_1) - \varphi(s_2)| < \delta, \quad \forall\, s_1, s_2 \in V_{1,2}.$$

The topological space $\mathbb{C}/\Lambda$ endowed with the complex structure determined by the atlas $\mathcal{A}$ is a compact Riemann surface called Complex Torus. It is important to note that the group structure on $\mathbb{C}/\Lambda$ is analytic, i.e., in terms of local charts about any two given points in the complex torus, addition is a holomorphic function of two complex variables.

Let $z_1 + \Lambda, z_2 + \Lambda \in \mathbb{C}/\Lambda$. Consider the charts

$$\begin{aligned} \phi_{z_1} &: \pi(D_{z_1}) \to D_{z_1}, \\ \phi_{z_2} &: \pi(D_{z_2}) \to D_{z_2}, \\ \phi_{z_3} &: \pi(D_{z_3}) \to D_{z_3}, \end{aligned}$$

where $z_3 = z_1 + z_2$, and fix an open neighbourhood V of (z_1, z_2) in $D_{z_1} \times D_{z_2}$ such that $s_1 + s_2 \in D_{z_3}$, $\forall\, (s_1, s_2) \in V$. Then

$$\phi_{z_3}\left(\phi_{z_1}^{-1}(s_1) + \phi_{z_2}^{-1}(s_2)\right) = s_1 + s_2, \quad \forall\, (s_1, s_2) \in V.$$

B.1.1 Analytic group structure

As mentioned at the beginning of this chapter, if E is a complex elliptic curve, then there exists a lattice Λ in $\mathbb{C}$ such that E is isomorphic to $\mathbb{C}/\Lambda$ as Riemann surfaces [Mir95, p.265]. Therefore, there exists an isomorphism

$$F : \mathbb{C}/\Lambda \to E.$$

Now, through this isomorphism the complex elliptic curve E inherits the analytic group structure on $\mathbb{C}/\Lambda$. To see this, it suffices to define

$$F(z_1 + \Lambda) + F(z_2 + \Lambda) := F((z_1 + z_2) + \Lambda), \quad \forall\, z_1 + \Lambda, z_2 + \Lambda \in \mathbb{C}/\Lambda.$$

Note that the isomorphism becomes a group isomorphism from $\mathbb{C}/\Lambda$ to E with respect to this group structure.

The following theorem states that an analytic group structure on a complex elliptic curve is uniquely determined by the identity element.

Theorem B.1. *Let E be a complex elliptic curve and $p_0 \in E$. Then there exists a unique analytic group structure on E such that p_0 is the identity element.*

PROOF: Let Λ be a lattice in $\mathbb{C}$ such that E is isomorphic to $\mathbb{C}/\Lambda$ and $F : E \to \mathbb{C}/\Lambda$ an isomorphism. The map $G : E \to \mathbb{C}/\Lambda$,

$$G(p) = F(p) - F(p_0), \quad \forall\, p \in E,$$

is an isomorphism, taking p_0 to $0 + \Lambda$. So the inverse isomorphism defines an analytic group structure on E as above, such that

$$p_0 = G^{-1}(0 + \Lambda)$$

is the identity element.

Let us now prove the uniqueness.
First, we consider the case $E = \mathbb{C}/\Lambda$, with Λ a lattice in $\mathbb{C}$, and $p_0 = 0 + \Lambda$. Let $\oplus : E \times E \to E$ be another addition which defines an analytic group structure with identity element $0 + \Lambda$. Below, we show that

$$\widetilde{z_1} + \widetilde{z_2} = \widetilde{z_1} \oplus \widetilde{z_2}, \quad \forall\, \widetilde{z_1} = z_1 + \Lambda, \widetilde{z_2} = z_2 + \Lambda \in E.$$

- For $\widetilde{z_2} = 0 + \Lambda$, it is obvious.
- For $\widetilde{z_2} \neq 0 + \Lambda$, the map $F_{\widetilde{z_2}} : E \to E$,

  $$F_{\widetilde{z_2}}(\widetilde{z}) = \widetilde{z} \oplus \widetilde{z_2}, \quad \forall\, \widetilde{z} = z + \Lambda \in E,$$

  is an automorphism without fixed points. Then

  $$F_{\widetilde{z_2}}(\widetilde{z}) = \widetilde{z} + \widetilde{z_3}, \quad \forall\, \widetilde{z} = z + \Lambda \in E,$$

for some $\widetilde{z_3} = z_3 + \Lambda \in E$, with $z_3 \notin \Lambda$ (Corollary B.5). As a consequence,

$$\widetilde{z_2} = F_{\widetilde{z_2}}(0 + \Lambda) = \widetilde{z_3}.$$

The arbitrary case of an elliptic curve E, with identity element $p_0 \in E$, follows from the case just considered since if there exists another analytic group structure on E, with identity element p_0, then the isomorphism

$$G : E \to \mathbb{C}/\Lambda$$

defined above carries the two analytic group structures on E to two distinct analytic group structures on $\mathbb{C}/\Lambda$, with identity element $0 + \Lambda$.

□

From the uniqueness stated in the previous theorem, we obtain the following corollary.

Corollary B.2. *Let E_1, E_2 be two complex elliptic curves with identity elements p_0^1 and p_0^2, respectively. If $F : E_1 \to E_2$ is an isomorphism of Riemann surfaces such that $F(p_0^1) = p_0^2$, then*

$$F(p) + F(q) = F(p + q), \quad \forall\, p, q \in E_1.$$

Therefore, the isomorphism F is also a group isomorphism from E_1 to E_2.

B.2 Morphisms between complex tori

Let $\mathbb{C}/\Lambda_1, \mathbb{C}/\Lambda_2$ be two complex tori.

Theorem B.3. *If $F : \mathbb{C}/\Lambda_1 \to \mathbb{C}/\Lambda_2$ is a morphism of Riemann surfaces, then there exist $\alpha, \beta \in \mathbb{C}$, with $\alpha\Lambda_1 \subset \Lambda_2$, such that*

$$F(z + \Lambda_1) = (\alpha z + \beta) + \Lambda_2, \quad \forall\, z + \Lambda_1 \in \mathbb{C}/\Lambda_1.$$

Reciprocally, if there exists $\alpha \in \mathbb{C}$, with $\alpha\Lambda_1 \subset \Lambda_2$, then for each $\beta \in \mathbb{C}$ the map $F_{\alpha,\beta} : \mathbb{C}/\Lambda_1 \to \mathbb{C}/\Lambda_2$,

$$F_{\alpha,\beta}(z + \Lambda_1) = (\alpha z + \beta) + \Lambda_2, \quad \forall\, z + \Lambda_1 \in \mathbb{C}/\Lambda_1.$$

is a morphism of Riemann surfaces. Its degree is the index of $\alpha\Lambda_1$ in Λ_2,

$$\deg(F_{\alpha,\beta}) = [\Lambda_2 : \alpha\Lambda_1].$$

Proof: [Mir95, p.63]

□

Note that the map $F_{\alpha,\beta}$ is bijective if and only if $\alpha\Lambda_1 = \Lambda_2$. In such case, its inverse is

$$F^{-1}_{\alpha,\beta}(z + \Lambda_2) = (1/\alpha)z - \beta/\alpha + \Lambda_1, \quad \forall\, z + \Lambda_2 \in \mathbb{C}/\Lambda_2.$$

Corollary B.4. *The isomorphisms from $\mathbb{C}/\Lambda_1$ to $\mathbb{C}/\Lambda_2$ are*

$$F(z+\Lambda_1) = (\alpha z+\beta)+\Lambda_2, \quad \forall\, z+\Lambda_1 \in \mathbb{C}/\Lambda_1,$$

where $\alpha,\beta \in \mathbb{C}$, $\alpha\Lambda_1 = \Lambda_2$.

Therefore, the complex tori $\mathbb{C}/\Lambda_1$ and $\mathbb{C}/\Lambda_2$ are isomorphic (as Riemann surfaces) if and only if there exists $\alpha \in \mathbb{C}$ such that $\alpha\Lambda_1 = \Lambda_2$, i.e., if the lattices Λ_1 and Λ_2 are homothetic.

Corollary B.5. *Let $\mathbb{C}/\Lambda$ be a complex torus.*

1. *The morphisms from $\mathbb{C}/\Lambda$ to itself are*

$$F(z+\Lambda) = (\alpha z+\beta)+\Lambda, \quad \forall\, z+\Lambda \in \mathbb{C}/\Lambda,$$

where $\alpha,\beta \in \mathbb{C}$, $\alpha\Lambda \subset \Lambda$.

2. *The automorphisms of $\mathbb{C}/\Lambda$ are*

$$F(z+\Lambda) = (\alpha z+\beta)+\Lambda, \quad \forall\, z+\Lambda \in \mathbb{C}/\Lambda,$$

where $\alpha,\beta \in \mathbb{C}$, $\alpha\Lambda = \Lambda$.

3. *The automorphisms of $\mathbb{C}/\Lambda$ without fixed points are*

$$F(z+\Lambda) = (z+\beta)+\Lambda, \quad \forall\, z+\Lambda \in \mathbb{C}/\Lambda,$$

where $\beta \in \mathbb{C}\setminus\Lambda$.

Remark B.6. The maps $F_{\alpha,\beta} : \mathbb{C}/\Lambda_1 \to \mathbb{C}/\Lambda_2$,

$$F_{\alpha,\beta}(z+\Lambda_1) = (\alpha z+\beta)+\Lambda_2, \quad \forall\, z+\Lambda_1 \in \mathbb{C}/\Lambda_1.$$

are group homomorphisms if and only if $F_{\alpha,\beta}(0+\Lambda_1) = 0+\Lambda_2$, i.e., $\beta \in \Lambda_2$.

B.3 Weierstrass $\wp$-function

The Weierstrass $\wp$-function for a lattice Λ in $\mathbb{C}$ is defined as

$$\wp_\Lambda(z) = \frac{1}{z^2} + \sum_{w\in\Lambda}{}' \left(\frac{1}{(z-w)^2} - \frac{1}{w^2}\right), \quad \forall\, z \in \mathbb{C},\ z \notin \Lambda,$$

where the primed summation means to sum over the points $w \in \Lambda\setminus\{0\}$. It is the most important specific example of elliptic function with periods Λ,

$$\wp_\Lambda(z) = \wp_\Lambda(z+w), \quad \forall\, z \in \mathbb{C},\ \forall\, w \in \Lambda.$$

As the series converges uniformly on compact subsets which do not meet Λ, the function is holomorphic on $\mathbb{C}/\Lambda$ and has double poles at lattice points. So its order is 2, i.e, the number of poles in any period parallelogram

$$P_a(w_1, w_2) = \{a + \lambda w_1 + \beta w_2 \,|\, 0 \leq \lambda, \beta < 1\}, \quad a \in \mathbb{C},$$

where (w_1, w_2) is a basic of Λ, is equal to 2.

Its derivative is also an elliptic function with periods Λ,

$$\wp_\Lambda'(z) = -2 \sum_{w \in \Lambda}{}' \frac{1}{(z-w)^3}, \quad \forall\, z \in \mathbb{C}, z \notin \Lambda.$$

In fact, it turns out that $\wp_\Lambda$ and $\wp_\Lambda'$ are the only specific examples we need since the field of elliptic functions with periods Λ is generated by $\wp_\Lambda$ and $\wp_\Lambda'$. It is important to note that $\wp_\Lambda$ is an even function while $\wp_\Lambda'$ is an odd function.

The following theorem states the main results of the Weierstrass $\wp$-function.

Theorem B.7. *Let $\wp_\Lambda$ be the Weierstrass function with respect to a lattice Λ.*

1. *The Laurent series of $\wp_\Lambda$ is*

$$\wp_\Lambda(z) = \frac{1}{z^2} + \sum_{\substack{n=2 \\ n \text{ even}}}^{\infty} (n+1) G_{n+2}(\Lambda) z^n, \quad \forall\, z \in \dot{D}(0, \delta),$$

where

$$G_k(\Lambda) = \sum_{w \in \Lambda}{}' \frac{1}{w^k}, \quad k > 2 \text{ even},$$

and

$$\delta = \min_{w \in \Lambda \setminus \{0\}} \{|w|\}.$$

2. *The functions $\wp_\Lambda$ and $\wp_\Lambda'$ satisfy the cubic equation*

$$(\wp_\Lambda')^2 = 4\wp_\Lambda^3 - g_2(\Lambda)\wp_\Lambda - g_3(\Lambda),$$

where $g_2(\Lambda) = 60G_4(\Lambda)$ and $g_3(\Lambda) = 140G_6(\Lambda)$.

3. *Let (w_1, w_2) be a basis of Λ and $w_3 = w_1 + w_2$. Then the cubic equation satisfied by $\wp_\Lambda$ and $\wp_\Lambda'$ is*

$$y^2 = 4(x - e_1)(x - e_2)(x - e_3), \quad \text{where } e_i = \wp_\Lambda(w_i/2), \ \forall\, i = 1, 2, 3.$$

This equation is nonsingular, meaning its right side has distinct roots. In particular, its discriminant up to constant multiple

$$\Delta(\Lambda) = g_2(\Lambda)^3 - 27g_3(\Lambda)^2 \neq 0.$$

Proof: [Apo90, p.9-14]

□

As a consequence of this result, the following question arises:

> Let c_2, c_3 be complex numbers satisfying $c_2^3 - 27c_3^2 \neq 0$. Does there exist a lattice Λ in $\mathbb{C}$ such that
>
> $$g_k(\Lambda) = c_k, \quad k = 2, 3\,?$$

Using properties of the modular function

$$j : \mathbb{H} \to \mathbb{C}, \quad j(\tau) = 1728\frac{g_2(\tau)^3}{\Delta(\tau)}, \quad \forall\, \tau \in \mathbb{H},$$

it is possible to demonstrate that the answer to this question is affirmative.

Theorem B.8. *Let c_2, c_3 be complex numbers satisfying $c_2^3 - 27c_3^2 \neq 0$. Then there exists a unique lattice Λ in $\mathbb{C}$ such that*

$$g_k(\Lambda) = c_k, \quad k = 2, 3.$$

Proof: [Apo90, p.42]

□

Therefore, we can conclude that there exists a natural bijection between

$$\begin{array}{ccc} \text{Complex} & & \text{Nonsingular} \\ \text{tori} & \longrightarrow & \text{cubic equations} \\ \mathbb{C}/\Lambda & & y^2 = 4x^2 - c_2x - c_3 \end{array}$$

Definition B.9. *A elliptic curve over $\mathbb{C}$ is any nonsingular cubic equation of this form,*

$$E: \, y^2 = 4x^3 - c_2x - c_3, \quad c_2, c_3 \in \mathbb{C}, \, c_2^3 - 27c_3^2 \neq 0.$$

A good reference book about elliptic curves is for example [Sil09].

B.3.1 Algebraic models

Let Λ be a lattice in $\mathbb{C}$, with basis (w_1, w_2). Consider the nonsingular cubic equation satisfied by $\wp_\Lambda$ and $\wp'_\Lambda$,

$$y^2 = 4x^3 - g_2(\Lambda)x - g_3(\Lambda).$$

Observe that it determines a noncompact topological subspace X_Λ of $\mathbb{C}^2$,

$$X_\Lambda = \{(x, y) \in \mathbb{C}^2 \,|\, y^2 - 4x^3 + g_2(\Lambda)x + g_3(\Lambda) = 0\}.$$

To compactify it we add an additional point that will be denoted by (∞, ∞),

$$E_\Lambda = X_\Lambda \cup \{(\infty, \infty)\},$$

and we define the following topology: The open subsets of E_Λ are

- the open subsets of X_Λ, and
- the complements of compact subsets in X_Λ.

Theorem B.10. *The complex elliptic curve*

$$E_\Lambda = X_\Lambda \cup \{(\infty, \infty)\}$$

is a compact connected Hausdorff topological space with respect to this topology.

Let us now define an atlas on E_Λ. Consider the natural projections

$$\pi_x : \mathbb{C}^2 \to \mathbb{C},\ \pi_x(x, y) = x, \quad \text{and} \quad \pi_y : \mathbb{C}^2 \to \mathbb{C},\ \pi_y(x, y) = y.$$

We need a version of the implicit function theorem for polynomials in $\mathbb{C}[x, y]$.

Theorem B.11. *Let $p \in \mathbb{C}[x, y]$ and $(a, b) \in \mathbb{C}^2$. Suppose that*

$$p(a, b) = 0 \quad \textit{and} \quad \partial_y p(a, b) \neq 0.$$

Then there exist open subsets $V_a, V_b \subset \mathbb{C}$, with $a \in V_a$, $b \in V_b$, and a holomorphic function $g : V_a \to V_b$ such that for all $(x, y) \in V_a \times V_b$,

$$p(x, y) = 0 \quad \textit{if and only if} \quad y = g(x).$$

This result allows us to invert locally the natural projections π_x and π_y on the curve E_Λ. Define

$$f = y^2 - 4x^3 + g_2(\Lambda)x + g_3(\Lambda) \in \mathbb{C}[x, y]$$

and

$$f_\infty = y^2 - x^4\big(\frac{4}{x^3} - g_2(\Lambda)\frac{1}{x} - g_3(\Lambda)\big) \in \mathbb{C}[x, y]$$

The map $\psi : U \to U_\infty$ defined as

$$\psi(x, y) = (\frac{1}{x}, \frac{y}{x^2}), \quad \forall\, (x, y) \in U,$$

where

$$U = \{(x, y) \in \mathbb{C}^2 \,|\, x \neq 0, f(x, y) = 0\}$$

and

$$U_\infty = \{(x, y) \in \mathbb{C}^2 \,|\, x \neq 0, f_\infty(x, y) = 0\},$$

is a homeomorphism. Its inverse homeomorphism is defined in the same way,

$$\psi^{-1} : U_\infty \to U, \quad \psi^{-1}(x, y) = (\frac{1}{x}, \frac{y}{x^2}), \quad \forall\, (x, y) \in U_\infty.$$

Remark B.12. Let $(x, y) \in U$. Observe that

$$\left(\frac{y}{x^2}\right)^2 = \frac{1}{x^4}(4x^3 - g_2(\Lambda)x - g_3(\Lambda)) \sim \frac{1}{x}, \quad \text{as } x \to \infty,$$

Making the change of variables $z = 1/x$ and $w = y/x^2$, we obtain that

$$w^2 = z^4(\frac{4}{z^3} - g_2(\Lambda)\frac{1}{z} - g_3(\Lambda))$$

This argument justifies the previous definition of the polynomial $f_\infty \in \mathbb{C}[x, y]$.

Let $(a, b) \in X_\Lambda$. Suppose that $\partial_y f(a, b) \neq 0$. Then there exist open subsets $V_a, V_b \subset \mathbb{C}$, with $a \in V_a$, $b \in V_b$, and a holomorphic function $g : V_a \to V_b$ such that for all $(x, y) \in V_a \times V_b$, $f(x, y) = 0$ if and only if $y = g(x)$. Therefore, the projection $\pi_x : U_{(a,b)} \to V_a$, where $U_{(a,b)} = (V_a \times V_b) \cap X_\Lambda$, is a chart on E_Λ. Its inverse homeomorphism is $\pi_x^{-1}(x) = (x, g(x)), \forall\, x \in V_a$.

We can define a chart $\pi_y : U_{(a,b)} \to V_b$ in the same way as above if $\partial_x f(a, b) \neq 0$. Since the partial derivatives of f are not both zero,

$$\partial_x f(a, b) \neq 0 \quad \text{or} \quad \partial_y f(a, b) \neq 0,$$

we have defined a complex chart on the curve E_Λ for each point $(a, b) \in X_\Lambda$.

Observe now that $\partial_x f_\infty(0, 0) = -4$, since

$$\partial_x f_\infty = -4 + 3g_2(\Lambda)x^2 + 4g_3(\Lambda)x^3.$$

Then there exist open subsets $V_1, V_2 \subset \mathbb{C}$, with $0 \in V_1 \cap V_2$, and a holomorphic function $h : V_2 \to V_1$ such that for all $(x, y) \in V_1 \times V_2$, $f_\infty(x, y) = 0$ if and only if $x = h(y)$. As above, the projection $\pi_y : U_{1,2} \to V_2$, where $U_{1,2} = \{(x, y) \in V_1 \times V_2 \,|\, f_\infty(x, y) = 0\}$, is a homeomorphism. Therefore, the composition $\pi_\infty = \pi_y \circ \psi : \psi^{-1}(U_{1,2}) \cup \{(\infty, \infty)\} \to V_2$,

$$\pi_\infty(x, y) = y/x^2, \quad \forall\, (x, y) \in \psi^{-1}(U_{1,2}) \quad [\psi(\infty, \infty) := (0, 0)],$$

is a chart on E_Λ about the point (∞, ∞).

Proving the compatibility of these charts is left as an exercise to the lector. The elliptic curve E_Λ endowed with the complex structure determined by this collection of pairwise compatible charts is a compact Riemann surface.

Let $F_\Lambda : \mathbb{C}/\Lambda \to E_\Lambda$ be the map

$$F_\Lambda(z + \Lambda) = (\wp_\Lambda(z), \wp'_\Lambda(z)), \quad \forall\, z + \Lambda \in \mathbb{C}/\Lambda.$$

Observe that it is well-defined since $\wp_\Lambda$ and $\wp'_\Lambda$ are both elliptic functions.

Theorem B.13. *The map F_Λ is an isomorphism of Riemann surfaces.*

Proof: Let us begin by proving that F_Λ is bijective. Let $(x_0, y_0) \in E_\Lambda$. If $(x_0, y_0) = (\infty, \infty)$, then the only preimage of this point is $0+\Lambda$. Otherwise, the equation

$$\wp_\Lambda(\widetilde{z}) = x_0, \quad \widetilde{z} = z + \Lambda \in \mathbb{C}/\Lambda,$$

has two solutions $\widetilde{z_1}, \widetilde{z_2} \in \mathbb{C}/\Lambda$, since $\wp_\Lambda$ is an elliptic function of order 2. If x_0 is a root of the polynomial

$$-4x^3 + g_2(\Lambda)x + g_3(\Lambda) \in \mathbb{C}[x],$$

then these two solutions are the same, since

$$(\wp'_\Lambda(z_i))^2 = -4(\wp_\Lambda(z_i))^3 + g_2(\Lambda)\wp_\Lambda(z_i) + g_3(\Lambda) = 0, \quad i = 1, 2.$$

And if x_0 is not a root of this polynomial, then these two solutions must be distinct, since otherwise we can conclude that $\wp_\Lambda$ does not have order 2. Therefore, it suffices to see that

$$\wp'_\Lambda(z_1) \neq \wp'_\Lambda(z_2),$$

to conclude that (x_0, y_0) has only a preimagen in $\mathbb{C}/\Lambda$. As $\wp_\Lambda$ is even and $\widetilde{z_1} \neq -\widetilde{z_1}$ (since the derivative $\wp'_\Lambda$ only vanishes at the points of $\mathbb{C}/\Lambda$ that have order 2, Theorem B.7), we deduce that $\widetilde{z_2} = -\widetilde{z_1}$. Using now that $\wp'_\Lambda$ is odd, we obtain that

$$\wp'_\Lambda(z_2) = \wp'_\Lambda(-z_1) = -\wp'_\Lambda(z_1).$$

To prove that F_Λ is holomorphic at $\widetilde{z_0} = z_0 + \Lambda \in \mathbb{C}/\Lambda$ we distinguish the following three cases:

- If $z_0 \notin \Lambda$ and $\wp_\Lambda(z_0) \neq 0$, then a complex chart on E_Λ about the point $(x_0, y_0) = F_\Lambda(\widetilde{z_0})$ is π_x, since $\partial_y f(x_0, y_0) \neq 0$. So

$$\begin{aligned}\pi_x \circ F_\Lambda \circ \phi_{z_0}^{-1}(z) &= \pi_{x_0}(F_\Lambda(z + \Lambda)) \\ &= \pi_{x_0}(\wp_\Lambda(z), \wp'_\Lambda(z)) = \wp_\Lambda(z)\end{aligned}$$

- If $z_0 \notin \Lambda$ and $\wp_\Lambda(z_0) = 0$, then a complex chart on E_Λ about the point $(0, y_0) = F_\Lambda(\widetilde{z_0})$ is π_y, since $\partial_x f(0, y_0) \neq 0$. So

$$\begin{aligned}\pi_y \circ F_\Lambda \circ \phi_{z_0}^{-1}(z) &= \pi_y(F_\Lambda(z + \Lambda)) \\ &= \pi_y(\wp_\Lambda(z), \wp'_\Lambda(z)) = \wp'_\Lambda(z).\end{aligned}$$

- If $z_0 \in \Lambda$ $(z_0 + \Lambda = 0 + \Lambda)$, then a complex chart on E_Λ about the point $(\infty, \infty) = F_\Lambda(\widetilde{z_0})$ is π_∞. Therefore,

$$\begin{aligned}\pi_\infty \circ F_\Lambda \circ \phi_{z_0}^{-1}(z) &= \pi_\infty(F_\Lambda(z + \Lambda)) \\ &= \pi_\infty(\wp_\Lambda(z), \wp'_\Lambda(z)) = \frac{\wp'_\Lambda(z)}{(\wp_\Lambda(z))^2}\end{aligned}$$

and

$$\lim_{z\to z_0} \frac{\wp'_\Lambda(z)}{(\wp_\Lambda(z))^2} = \lim_{z\to z_0} -2(z-z_0) + (\text{higher order terms}) = 0.$$

Observe that the three compositions are well-defined in neighbourhoods of z_0. □

Final conclusion: Let E be a complex elliptic curve (compact Riemann surface of genus 1). Then there exist a lattice Λ in $\mathbb{C}$ and an isomorphism $F : E \to \mathbb{C}/\Lambda$. Therefore,

$$E \xrightarrow{F} \mathbb{C}/\Lambda \xrightarrow{F_\Lambda} E_\Lambda$$

is an isomorphism from E to E_Λ.

Bibliography

[Ahl78] Lars V. Ahlfors. *Complex analysis.* McGraw-Hill Book Co., New York, third edition, 1978. An introduction to the theory of analytic functions of one complex variable, International Series in Pure and Applied Mathematics.

[Apo90] Tom M. Apostol. *Modular functions and Dirichlet series in number theory*, volume 41 of *Graduate Texts in Mathematics.* Springer-Verlag, New York, second edition, 1990.

[DS05] Fred Diamond and Jerry Shurman. *A first course in modular forms*, volume 228 of *Graduate Texts in Mathematics.* Springer-Verlag, New York, 2005.

[Gir70] Jean Giraud. *Surfaces de Riemann Compactes.* Notes d'un cours professé à Orsay au second semestre de l'année scolaire 1969–70 dans le cadre du troisième cycle de Géométrie Algébrique. Faculté des Sciences d'Orsay, 1970. `http://sites.mathdoc.fr/PMO/PDF/J_GIRAUD_1969-70.pdf`.

[IR90] Kenneth Ireland and Michael Rosen. *A classical introduction to modern number theory*, volume 84 of *Graduate Texts in Mathematics.* Springer-Verlag, New York, second edition, 1990.

[Mir95] Rick Miranda. *Algebraic curves and Riemann surfaces*, volume 5 of *Graduate Studies in Mathematics.* American Mathematical Society, Providence, RI, 1995.

[Miy06] Toshitsune Miyake. *Modular forms.* Springer Monographs in Mathematics. Springer-Verlag, Berlin, english edition, 2006. Translated from the 1976 Japanese original by Yoshitaka Maeda.

[NR11] Terrence Napier and Mohan Ramachandran. *An introduction to Riemann surfaces.* Cornerstones. Birkhäuser/Springer, New York, 2011.

[Ser73] J.-P. Serre. *A course in arithmetic.* Springer-Verlag, New York-Heidelberg, 1973. Translated from the French, Graduate Texts in Mathematics, No. 7.

[Shi71] Goro Shimura. *Introduction to the arithmetic theory of automorphic functions.* Publications of the Mathematical Society of Japan, No. 11. Iwanami Shoten, Publishers, Tokyo; Princeton University Press, Princeton, N.J., 1971. Kanô Memorial Lectures, No. 1.

[Sil09] Joseph H. Silverman. *The arithmetic of elliptic curves*, volume 106 of *Graduate Texts in Mathematics.* Springer, Dordrecht, second edition, 2009.

[SS03] Elias M. Stein and Rami Shakarchi. *Complex analysis*, volume 2 of *Princeton Lectures in Analysis.* Princeton University Press, Princeton, NJ, 2003.

[TW95] Richard Taylor and Andrew Wiles. Ring-theoretic properties of certain Hecke algebras. *Ann. of Math. (2)*, 141(3):553–572, 1995.

[vdP96] Alf van der Poorten. *Notes on Fermat's last theorem.* Canadian Mathematical Society Series of Monographs and Advanced Texts. John Wiley & Sons, Inc., New York, 1996. A Wiley-Interscience Publication.

[Wil95] Andrew Wiles. Modular elliptic curves and Fermat's last theorem. *Ann. of Math. (2)*, 141(3):443–551, 1995.