ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

EPFL - SMA Bachelor Semestre 6
Under the supervision of : Prof. Dr. Eva Bayer-Fluckiger
Assisted by : Dr. Peter Jossen

Elliptic curves, modular forms and the Taniyama-Shimura Conjecture

Semester project presented by

David Ter-Borch Gram Schjoldager Lilienfeldt

December 11, 2018

Abstract

We study elliptic curves and modular forms with the aim to state and understand the Taniyama-Shimura Conjecture, now known as the Modularity Theorem. We illustrate the theorem on two examples using an approach with modular symbols.

Contents

Introduction

The aim of this work is to state and understand the Taniyama-Shimura Conjecture. This conjecture, known since 2001 as the Modularity Theorem, states that:

All rational elliptic curves arise from modular forms.

Understanding this statement requires the study of elliptic curves and modular forms. The link provided by the Modularity Theorem between these two completely unrelated areas in mathematics is far from being trivial.

In the first chapter we study elliptic curves and their fundamental properties. We follow an approach similar to the one in [He02]. After having defined elliptic curves over a field K and looked at birational equivalence between algebraic curves, we introduce Weierstrass equations and show that every rational elliptic curve may be represented by an equation of this type. We study the reduction modulo a prime p of elliptic curves. As we will see, the characterization of these reductions depends on the chosen Weierstrass equation of the curve and this motivates the introduction of minimal Weierstrass equations and proving Néron's theorem which states that every rational elliptic curve is isomorphic to a curve given by a minimal Weierstrass equation. We then declare that the type of reduction modulo p of an elliptic curve is the type of reduction of a minimal Weierstrass equation representing the curve. This enables us to define the conductor N of an elliptic curve. Up to isomorphism, there is only a finite number of elliptic curves for a given conductor N and thus elliptic curves may be ordered using this conductor. We end the first chapter by defining the Hasse-Weil L-function of a rational elliptic curve and expressing it as a Dirichlet series. This series, called the L-series of the elliptic curve, contains information about the reduction modulo p prime of the curve. We will see that its coefficients are related to the conductor of the curve and the number of points of the same curve viewed over $\mathbb{F}_p$ for p prime. The conductor and the L-series are the two major concepts related to rational elliptic curves that are needed to state the Conjecture.

The second chapter is dedicated to the study of modular forms. In the first part of this chapter we mainly follow the approaches of [DS05] and [He02]. We introduce the modular group $\mathrm{SL}_2(\mathbb{Z})$ and congruence subgroups of $\mathrm{SL}_2(\mathbb{Z})$ with special emphasis on $\Gamma_0(N)$ which is the subgroup considered in the Conjecture. Then we define several actions of the modular group on different sets and finish this preliminary part of the chapter with the notion of fundamental domains for congruence subgroups. We characterize this first part as a preliminary part because the concepts introduced are needed in order to understand and define modular forms, which are the real object of study of this chapter. Following the approach of [DS05] we define modular forms of weight k for any congruence subgroup of $\mathrm{SL}_2(\mathbb{Z})$, keeping in mind that our main focus is on $\Gamma_0(N)$ and the weight 2. We follow up with some examples of modular forms on $\mathrm{SL}_2(\mathbb{Z})$, which are not needed for the Conjecture but still interesting as they enable us to get a better grasp of the rather abstract definition of modular forms. Furthermore, some interesting properties related to Eisenstein series are exposed here. Mainly following [Se70] we introduce Hecke operators for $\mathrm{SL}_2(\mathbb{Z})$ and then for $\Gamma_0(N)$. This results in the definition of Hecke forms that is needed in the Conjecture.

After the two first chapters we are ready to state the Taniyama-Shimura Conjecture and this is what we do in the third chapter after a brief history of the Conjecture. In a short paragraph we relate the Modularity Theorem with Fermat's Last Theorem.

The fourth and final chapter is dedicated to the computational illustration of the Conjecture. Following [St07], we choose the modular symbol approach to do these computations. After a brief definition of these symbols, we expose a trick due to Manin that enables us to express every modular symbol as a $\mathbb{Q}$-linear combination of a finite number of specific symbols. We state a theorem due to Manin that allows us to compute a basis for the space of modular symbols for $\Gamma_0(N)$. We then define Hecke operators on this same space and introduce a pairing of modular symbols for $\Gamma_0(N)$ with modular forms for $\Gamma_0(N)$. We show that this pairing is compatible with the Hecke operators and thus the eigenvalues of the Hecke operators on modular forms may be computed through the operators on modular symbols. Putting all this together we finally compute two explicit examples verifying the Conjecture, the first one computed by hand and the second using SAGE.

The prerequisites for a good understanding of this work is knowledge of complex analysis, linear algebra and basic abstract algebra with special emphasis on group actions. Familiarity with differentiable manifolds may be helpful but not crucial since this part of the theory is not really treated in this work.

1 Elliptic curves

In this chapter, we define elliptic curves and study some of their important properties. These properties will be needed later for understanding the Taniyama-Shimura Conjecture. We start with the definition of elliptic curves.

1.1 Definition

We need to define elliptic curves over a given field K. In order to do this, we need a couple of definitions. First, recall that a polynomial $f \in K[X, Y]$ is said to be homogeneous if all of its monomials have the same total degree, which we call then the degree of the polynomial. For example the polynomial $Y^2 - YX$ is homogeneous, but the polynomial $Y^2 - X^3 + X$ is not.

Now, it comes naturally that if we are given a non-homogeneous polynomial we would like to homogenize it. The following definition gives the answer to this question.

Definition 1.1. Let K be a field and $f \in K[X, Y]$ be a non-zero polynomial of positive degree d. The *homogenization* of f is the unique homogeneous polynomial $F \in K[X, Y, Z]$ of degree d satisfying $F(X, Y, 1) = f(X, Y)$.

For example, the homogenization of the polynomial $f(X, Y) = Y^2 - X^3 + X$ is the polynomial $F(X, Y, Z) = Y^2Z - X^3 + XZ^2$. We now define plane projective curves.

Definition 1.2. Let $d \in \mathbb{N} \setminus \{0\}$. A *plane projective curve of degree* d defined over a field K is an element of $\mathbb{P}(K_h[X, Y, Z]_d)$, where $K_h[X, Y, Z]_d$ is the K-vector space of homogeneous polynomials of degree d defined in the variables X, Y and Z and $\mathbb{P}(K_h[X, Y, Z]_d)$ is the associated projective space.

So a plane projective curve C is given by a non-zero polynomial $F \in K_h[X, Y, Z]_d$ up to non-zero scalar multiples. We speak of the curve C given by the equation $F(X, Y, Z) = 0$. For any field K' that contains K, we set

$$C(K') = \{(a, b, c) \in \mathbb{P}^2(K') : F(a, b, c) = 0\},$$

and this makes sense since F is homogeneous. In fact, recall that the projective plane of a field K can be viewed as $K^3 \setminus \{(0, 0, 0)\}$ quotiented by the homothetic equivalence relation:

$$(a, b, c) \sim (a', b', c') \iff \exists \lambda \in K^* \text{ such that } (a', b', c') = (\lambda a, \lambda b, \lambda c).$$

So a point in $\mathbb{P}^2(K)$ is a homothetic equivalence class. Furthermore, it is easy to see that

$$F(\lambda X, \lambda Y, \lambda Z) = 0 \iff F(X, Y, Z) = 0,$$

since F is homogeneous. Note that if F is homogeneous of degree d, then its partial derivatives are also homogeneous polynomials but of degree $d-1$. We now introduce the concept of singular points.

Definition 1.3. Let $\overline{K}$ be an algebraic closure of K and $P=(a,b,c)\in C(\overline{K})$. The point P is said to be a *singular point* of C if:

$$\operatorname{grad}(F)(a,b,c)=(0,0,0).$$

Otherwise, P is said to be *non-singular.* A curve that has no singular points is said to be *smooth.*

Example 1.4. Consider the curve C defined over $\mathbb{Q}$ by $F(X,Y,Z)=Y^2Z-X^3=0$. We have

$$\operatorname{grad}(F)(X,Y,Z)=(-3X^2,2YZ,Y^2)$$

and thus the point $(0,0,1)$ is clearly a singular point of C and we conclude that C is singular.

Consider the curve C' defined over any field K with characteristic different from 2 by the equation $F(X,Y,Z)=X^2+Y^2+Z^2=0$. We compute that

$$\operatorname{grad}(F)(X,Y,Z)=(2X,2Y,2Z)$$

and notice that it only vanishes at the point $(X,Y,Z)=(0,0,0)$. But this point is not in $\mathbb{P}^2(K)$, thus C is smooth. Note that if K has characteristic 2, then every point in C is singular.

Finally, we need to speak about the genus of a projective algebraic curve. When C is a smooth algebraic curve defined over $\mathbb{C}$, its genus is equal to the number of "holes" of the corresponding Riemann surface. We will not need this definition. There is a simple formula to compute the genus g of a smooth plane projective curve. It depends only on the degree d of the polynomial that defines the curve and it is given by:

$$g=\frac{1}{2}(d-1)(d-2).$$

Thus, a smooth plane projective curve of genus 1 is defined by a polynomial of degree 3. We have now defined all the concepts needed to introduce elliptic curves defined over any field K.

Definition 1.5. Let K be any field. An *elliptic curve over* K is a pair (E,O) consisting of a plane projective smooth curve E of genus 1 and a point $O\in E(K)$.

For example, the curve given by the polynomial $X^3+Y^3+Z^3$ together with $O=(1,-1,0)$ is an elliptic curve over any field with characteristic different from three, but the curve given by X^2+Y^2 is not an elliptic curve since it is not smooth (and the genus of this curve is 0).

For simplicity, we will most of the time speak of an elliptic curve E instead of (E,O). The point O will be implicit.

1.6. In practice, when we speak about elliptic curves, we always have in mind a specific equation $f(X,Y,Z)=0$ for it. But there is no such thing as a unique equation that defines the curve. An elliptic curve E is a geometrical object for which we may choose a projective embedding such that $E \subset \mathbb{P}^2$. Then E is given by an equation up to non-zero scalar multiples and from there on we may choose an actual $f \in K[X,Y,Z]$ to "represent" E. It follows that properties of the elliptic curve E must be distinguished from those of the chosen equation $f=0$.

1.2 Birational maps between algebraic curves

Now that we have defined elliptic curves we will see that some of them are isomorphic. It is in fact more appropriate to speak about isomorphism classes of elliptic curves rather than just elliptic curves. Elliptic curves are particular cases of algebraic curves and this section is about this larger class of curves. We start with the notion of rational maps.

Definition 1.7. Let F and G be two plane projective curves defined over a field K and let $A,B,C \in K_h[X,Y,Z]_d$. The polynomials A,B and C are said to define a *rational map* ϕ from F to G defined over K if for all points (x,y,z) in $F(\overline{K})$, except for a finite number,

$$\phi(x,y,z) = (A(x,y,z), B(x,y,z), C(x,y,z))$$

is well defined, in other words not equal to zero, and lies in $G(\overline{K})$.

For example, let $F(X,Y,Z)=Z$ and $G(X,Y,Z)=Y^2-XZ$. Then $\phi(x,y,z)=(x^2,xy,y^2)$ is a rational map from F to G defined over any field.

Definition 1.8. Let F and G be two plane projective curves and $\phi : F \to G$ be a rational map defined over K. Then ϕ is said to be *birational* if there exists a rational map $\psi : G \to F$ such that for all points of $F(\overline{K})$ and $G(\overline{K})$, except for a finite number, the composition maps

$$\psi \circ \phi \quad \text{and} \quad \phi \circ \psi$$

are well defined and equal to the identity map. The curves F and G are said to be *birationally equivalent* over K if there exists a birational map between them.

Definition 1.9. Two plane projective curves are said to be *isomorphic* over K if there exists a birational map between them defined over K.

Let C be a smooth plane projective curve over a field K given by $f(X,Y,Z)=0$. We denote by $K(C)$ the *function field* of the curve C that we will now construct. Since K is a field, $K[X,Y]$ is an integral ring. Consider the ring $A_C = K[X,Y]/\langle f(X,Y,1)\rangle$. We need the following lemma and proposition in our construction.

Lemma 1.10. *If A is a unique factorization domain, then every irreducible element of A is prime.*

Proof. Let $a \in A$ be an irreducible element of A. Then a is non-zero and not a unit. Since A is a unique factorization domain, we may write a as a finite product of prime elements of A. Thus, there exists a positive integer n and prime elements $p_i \in A$ for $1 \leq i \leq n$ such that $a = p_1 p_2 \dots p_n$. If $n > 1$, this contradicts the fact that a is irreducible. Thus $n = 1$ and a is prime. □

Proposition 1.11. *Let C be a plane projective curve over K given by the equation $f(X,Y,Z) = 0$. If $f(X,Y,1)$ is irreducible in $K[X,Y]$, then $K[X,Y]/\langle f(X,Y,1)\rangle$ is an integral domain.*

Proof. By basic abstract algebra, we know that $K[X,Y]$ is a unique factorization domain. Since $f(X,Y,1)$ is irreducible in $K[X,Y]$, we know by Lemma 1.10 that $f(X,Y,1)$ is also prime. Thus the ideal generated by $f(X,Y,1)$ is a prime ideal and this is equivalent to $K[X,Y]/\langle f(X,Y,1)\rangle$ being an integral domain. □

1.12. Let D and E be two plane projective curves defined over K and given respectively by the polynomials f and g. Suppose $D \neq E$, $\gcd(f,g) = c \in K$ and let $\deg(f) = d_D$ and $\deg(g) = d_E$. Then $D \cup E$ is a plane projective curve given by the polynomial fg. Bézout 's Theorem concerning the number of intersection points of two plane projective curves, tells us that $\#D(\overline{K}) \cap E(\overline{K}) = d_D d_E$. Also, all points of this intersection are singular points. In fact, an intersection between the two curves is either a "cross" intersection or a "tangent" intersection. In the first case, there are two distinct tangent lines to $D \cup E$ at the intersection point. In the second, there is one tangent line, but with multiplicity two.

Suppose that f is reducible. Then we may write $C = D \cup E$ where D and E are two plane projective curves. Using the above remark, the intersection between D and E is non-empty and contains only singular points. This contradicts our assumption of C being smooth. Hence f is irreducible and Proposition 1.11 applies. Thus A_C is an integral ring and we may consider the fraction field of A_C. We set

$$K(E) = \mathrm{Frac}(A_C),$$

where $\mathrm{Frac}(A_C)$ denotes the fraction field of A_C.

1.13. The function field is a birational invariant.

1.3 Weierstrass equations

We define elliptic curves via Weierstrass equations. This definition is less restrictive than one would think as we will see. We then define the discriminant of these equations and look at singular cubic curves. We start with the definition of Weierstrass equations.

1.3.1 Long and short Weierstrass equations

Definition 1.14. Let K be a field. A *long Weierstrass equation* defined over K is an equation of the form:

$$Y^2 + a_1XY + a_3Y = X^3 + a_2X^2 + a_4X + a_6, \tag{1.14.1}$$

where $a_1, a_2, a_3, a_4, a_6 \in K$.

Theorem 1.15. *Let (E, O) be an elliptic curve defined over a field K.*

(i) There exist functions x, y in $K(E)$ and elements a_1, a_2, a_3, a_4, a_6 of K such that the map $\psi : E \to \mathbb{P}^2(K)$ sending $P \in E$ to $(x(P), y(P), 1)$ is a birational map from (E, O) to the projective curve defined over K given by

$$Y^2Z + a_1XYZ + a_3YZ^2 = X^3 + a_2X^2Z + a_4XZ^2 + a_6Z^3 \tag{1.15.1}$$

and such that $\psi(O) = (0, 1, 0)$.

(ii) Every other couple (x', y') that suits the same conditions as above can be written:

$$x' = u^2x + r \qquad y' = u^3y + sx + t,$$

where $u, r, s, t \in K$ and u is non-zero. We say that (u, r, s, t) is a feasible change of coordinates.

Proof. We refer to [He02] Chapter 4, Section 4.12 p. $211 - 212$. □

Let E be an elliptic curve. Using Theorem 1.15, it may be assumed that the projective equation for E is of the form:

$$Y^2Z + a_1XYZ + a_3YZ^2 = X^3 + a_2X^2Z + a_4XZ^2 + a_6Z^3. \tag{1.15.2}$$

We search for points at infinity, in other words points of the form $(a, b, 0)$ with $(a, b) \neq (0, 0)$, by setting $Z = 0$. This yields the equation $X^3 = 0$, which is solved by $X = 0$. We see that E has exactly one point at infinity and its projective coordinates are $\mathcal{O} = (0, 1, 0)$. In other words, E only intersects the projective line $Z = 0$ in one point. Thus, we can set $Z = 1$ and work with the equation $Y^2 + a_1XY + a_3Y = X^3 + a_2X^2 + a_4X + a_6$, which happens to be a Weierstrass equation. It follows that

$$E(K) = \{(x, y) \in K^2 : y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6\} \cup \{\mathcal{O}\}.$$

If we start with a long Weierstrass equation of type 1.14.1 over a field K with $\mathrm{char}(K) \neq 2, 3$ we can transform it into what we call a short Weierstrass equation.

Definition 1.16. A *short Weierstrass equation* defined over a field K is an equation of the form:

$$Y^2 = X^3 + AX + B, \tag{1.16.1}$$

with $A, B \in K$.

Consider the long Weierstrass equation $y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6$. We set

$$\begin{aligned} b_2 &= a_1^2 + 4a_2 \\ b_4 &= a_1a_3 + 2a_4 \\ b_6 &= a_3^2 + 4a_6 \\ b_8 &= a_1^2a_6 - a_1a_3a_4 + 4a_2a_6 + a_2a_3^2 - a_4^2 \\ c_4 &= b_2^2 - 24b_4 \\ c_6 &= -b_2^3 + 36b_2b_4 - 216b_6. \end{aligned}$$

Since $\mathrm{char}(K) \neq 2$, we get

$$y^2 = x^3 + \frac{b_2}{4}x^2 + \frac{b_4}{2}x + \frac{b_6}{4}.$$

Finally, since $\mathrm{char}(K) \neq 2, 3$, we get

$$y^2 = x^3 - \frac{c_4}{48}x - \frac{c_6}{864},$$

which is the short Weierstrass equation we were looking for.

1.3.2 The discriminant

Definition 1.17. The quantity:

$$\Delta = -b_2^2b_8 - 8b_4^3 - 27b_6^2 + 9b_2b_4b_6 = \frac{c_4^3 - c_6^2}{12^3}$$

is called the *discriminant* of the cubic 1.14.1.

It is easy to see that the discriminant of the cubic 1.16.1 is $\Delta = -16(4A^3 + 27B^2)$. In fact, if we consider this equation and compare it with the cubic 1.14.1, we see that $a_1 = a_3 = a_2 = 0, a_4 = A, a_6 = B$. Then $b_2 = 0, b_4 = 2A, b_6 = 4B, b_8 = -A^2$ and it follows that

$$\Delta = -8.8A^3 - 27.16B^2 = -16(4A^3 + 27B^2). \tag{1.17.1}$$

A calculation shows that the curve defined by the cubic 1.14.1 is singular (or non-smooth) if and only if $\Delta = 0$. Hence, a Weierstrass equation defines an elliptic curve if and only if the discriminant is non-zero.

It is important to note that the discriminant is not attached to the isomorphism class of an elliptic curve. In fact, if we make a feasible change of coordinates using Theorem 1.15, we get $u^4 c_4' = c_4$, $u^6 c_6' = c_6$ and thus $\Delta = u^{12}\Delta'$. We now define a quantity that has the property of being invariant under feasible change of coordinates.

Definition 1.18. The quantity:

$$j = \frac{c_4^3}{\Delta}$$

is called the modular invariant of the elliptic curve E. We shall denote it $j(E)$.

Theorem 1.19. *Let E and E' be two elliptic curves defined over an algebraically closed field K. We then have the following assertion:*

$$E \text{ and } E' \text{ are isomorphic} \iff j(E) = j(E').$$

Proof. The direct implication is easy since E and E' are isomorphic and thus have the same Weierstrass equation up to a feasible change of coordinates, which does not affect j.

For the converse, suppose that the characteristic of K is different form 2 and 3. Then E and E' have short Weierstrass equations of the form:

$$\begin{cases} E: & y^2 &=& x^3 + Ax + B \\ E': & y'^2 &=& x'^3 + A'x' + B'. \end{cases}$$

It is easy to see that $c_4 = -48A, c_4' = -48A'$ and thus

$$j(E) = j(E') \iff 1728\frac{4A^3}{4A^3 + 27B^2} = 1728\frac{4A'^3}{4A'^3 + 27B'^2} \iff A'^3B^2 = A^3B'^2.$$

We are confronted with several cases.

If $j = 0$, then $A = A' = 0$ and $B, B' \neq 0$. Let $u \in K$ such that $u^6 = (B'/B)$. The existence of such a u is guaranteed since K is supposed algebraically closed. Then $(u, 0, 0, 0)$ is a feasible change of coordinates from E' to E.

If $j = 1728$, then $B = B' = 0$ and $A, A' \neq 0$. Let $u \in K$ such that $u^4 = (A'/A)$. Then $(u, 0, 0, 0)$ is a feasible change of coordinates from E' to E.

If $j \neq 0, 1728$, then $(B'/B)^2 = (A'/A)^3$. Thus, $(B'/B)^4 = (A'/A)^6$. Let $u \in K$ such that $u^4 = (A/A')$ and $u^6 = (B/B')$. Then $(u, 0, 0, 0)$ is a feasible change of coordinates from E' to E. In fact,

$$\begin{aligned} y'^2 = x'^3 + A'x' + B' &\implies u^6y^2 = u^6x^3 + u^2A'x + B' \implies u^6u^4y^3 = u^6u^4x^3 + u^6A'x + u^4B' \\ &\implies \frac{B'}{B}\frac{A'}{A}y^3 = \frac{B'}{B}\frac{A'}{A}x^3 + A'\frac{B'}{B}x + B'\frac{A'}{A} \implies y^3 = x^3 + Ax + B, \end{aligned}$$

which is the equation of E.

If $\mathrm{char}(K) = 2$ or 3, we have to do the computations with long Weierstrass equations, a thing that we will not do here. □

Example 1.20. This example will serve as an illustration of the concepts covered in the last two sections. Provided two equivalent elliptic curves, we will expose a feasible change of coordinates and compute both the discriminant and the modular invariant of these curves.

Let E and E' be two elliptic curves over $\mathbb{Q}$ given by the following equations:

$$\begin{cases} E: & y^2 & = & x^3+2x+2 \\ E': & y'^2+\frac{1}{2}y'x' & = & x'^3+\frac{11}{16}x'^2+\frac{5}{16}x'+\frac{5}{64}. \end{cases}$$

A calculation shows that $(2,1,2,0)$ is a feasible change of coordinates from E to E'. Thus the two curves are isomorphic. We now compute the discriminant of the two curves.

(E) We have defined the elliptic curve E using a short Weierstrass equation. Identifying with Equation 1.16.1 we get: $A=B=2$. Using Equation 1.17.1, we compute:

$$\Delta_E = -16(4.2^3+27.2^2) = -2240.$$

(E') We have defined the elliptic curve E' using a long Weierstrass equation. Identifying with Equation 1.14.1 we get: $a'_1 = 1/2, a'_2 = 11/16, a'_3 = 0, a'_4 = 5/16, a'_5 = 5/64$. We now compute that $b'_2 = 3, b'_4 = 5/8, b'_6 = 5/16$. Further computations give $c'_4 = -6, c'_6 = -27$. Finally, by using Definition 1.17, we compute the discriminant $\Delta_{E'} = -\frac{35}{64}$. Note that $u^{12}\Delta_{E'} = -2^{12}.\frac{35}{64} = -2240 = \Delta_E$, just as expected. We now show that the modular invariants of E and E' are equal.

(E) By comparing the equation of E with the long Weierstrass Equation 1.14.1 we get: $a_4 = a_6 = 2, a_i = 0, i = 1,2,3$. Thus $b_2 = 0, b_4 = 4, b_6 = 8$. Using this, we find that $c_4 = -96$. Using Definition 1.18, we compute:

$$j(E) = \frac{c_4^3}{\Delta_E} = \frac{96^3}{2240} = \frac{13824}{35}.$$

(E') Recall that $c'_4 = -6$. Using Definition 1.18, we compute:

$$j(E') = \frac{c_4'^3}{\Delta_{E'}} = 6^3.\frac{64}{35} = \frac{13824}{35}.$$

We see that $j(E) = j(E')$, just as expected.

1.3.3 Singular cubic curves

Let $C: f(x,y) = 0$ be a cubic curve defined over a field K by a long Weierstrass equation with

$$f(x,y) = y^2 + a_1xy + a_3y - x^3 - a_2x^2 - a_4x - a_6, \tag{1.20.1}$$

where $a_i \in K,\ i = 1,2,3,4,6$. Suppose E is singular. Then there exists a point $P = (x_0,y_0)$ such that $\mathrm{grad}(f)(x_0,y_0) = 0$. We compute the Taylor expansion of f in a neighborhood of P up till the third order:

$$\begin{aligned} f(x,y) &= -(3x_0+a_2)(x-x_0)^2 + a_1(x-x_0)(y-y_0) + (y-y_0)^2 - (x-x_0)^3 \\ &= [(y-y_0)-\alpha(x-x_0)][(y-y_0)-\beta(x-x_0)] - (x-x_0)^3, \end{aligned}$$

where α and β lie in an algebraic closure of K.

Definition 1.21. If $\alpha \neq \beta$, then P is said to be a *node* (or double point). In this case there are two distinct tangent lines to C at P, namely

$$y - y_0 = \alpha(x - x_0) \text{ and } y - y_0 = \beta(x - x_0).$$

If $\alpha = \beta$, then P is said to be a *cusp*. In this case there is a unique tangent line to C at P with multiplicity two.

There is an easy way to identify cusps and nodes of an singular curve given by a Weierstrass equation as we see in the following proposition that we give without proof since this one is a simple computation.

Proposition 1.22. *Let C be a curve defined over a field K by a long Weierstrass Equation 1.20.1.*

(i) The curve C admits a cusp if and only if $\Delta = c_4 = 0$.

(ii) The curve C admits a node if and only if $\Delta = 0$ and $c_4 \neq 0$.

1.23. Note that the condition $\Delta = 0$ is equivalent to the curve C being singular as we have already seen.

1.4 Weierstrass equations over $\mathbb{Z}$

We study the reduction modulo a prime p of elliptic curves. We start by characterizing the reduction of Weierstrass equations and then introduce minimal Weierstrass equations in order to characterize the reduction of elliptic curves which is the aim of this section.

1.4.1 Reduction modulo p

Now that we have got a better grasp of elliptic curves and some of their properties, we will turn our attention towards elliptic curves defined over $\mathbb{Q}$ and the reduction modulo a prime number p of the defining equations.

Consider an elliptic curve E defined over $\mathbb{Q}$ by the short Weierstrass equation:

$$y^2 = x^3 + Ax + B, \tag{1.23.1}$$

where $A, B \in \mathbb{Q}$. If we define a feasible change of coordinates as follows $x' = \lambda^2 x, y' = \lambda^3 y$, where $\lambda \in \mathbb{Q}^*$, then we obtain the equation $y'^2 = x'^3 + \lambda^4 Ax' + \lambda^6 B$. We see that choosing λ wisely yields an equation with integer coefficients. Thus, up to a feasible change of coordinates, we may assume that $A, B \in \mathbb{Z}$.

Let p be a prime and consider the equation

$$E_p : y^2 = x^3 + A_p x + B_p,$$

where A_p and B_p are respectively the classes of A and B modulo p.

Definition 1.24. Let $E : y^2 = x^3 + Ax + B$ be a short Weierstrass equation with $A, B \in \mathbb{Z}$ and let p be a prime. We say that

(i) E has *good reduction* at p if the equation E_p defines a smooth curve.

(ii) E admits *multiplicative reduction* at p if the curve defined by the equation E_p admits a node in $\mathbb{P}^2(\overline{\mathbb{F}}_p)$. There are two types of multiplicative reduction. If the slopes of the tangent lines (α and β in Section 1.3.3) are in $\mathbb{F}_p$, we say that the reduction is *split* multiplicative. Otherwise, we say that the reduction is *non-split* multiplicative.

(iii) E admits *additive reduction* at p if the curve defined by the equation E_p admits a cusp in $\mathbb{P}^2(\overline{\mathbb{F}}_p)$.

We see that if E has good reduction at p, then $\Delta \not\equiv 0 \pmod{p}$. In the opposite case, we say that E has bad reduction at p. So the bigger the discriminant, the more places of bad reduction there exist for E.

Example 1.25. Let $p > 3$ be a prime and consider the two elliptic curves E and E' defined over $\mathbb{Q}$ by:

$$E : y^2 = x^3 + 1 \quad \text{and} \quad E' : y'^2 = x'^3 + p^6.$$

It is easy to see that $(\frac{1}{p}, 0, 0, 0)$ is a feasible change of coordinates form E to E', thus the two elliptic curves are isomorphic over $\mathbb{Q}$. If we compute their discriminants, we get:

$$\Delta_E = -432 = -2^4.3^3 \quad \text{and} \quad \Delta_{E'} = -432p^{12} = -2^4.3^3.p^{12}.$$

Since $p > 3$, E has good reduction at p while E' has bad reduction at p. Hence Definition 1.24 concerns the equation of the curve and not the curve as a geometrical object.

We want to be able to choose one particular representative Weierstrass equation from the isomorphism class of an elliptic curve and determine the reduction modulo p of the curve using this equation. The goal is for the curve to have a minimal number of places of bad reduction and this motivates the introduction of minimal Weierstrass equations. We state and prove a theorem due to Néron that states that every elliptic curve may be represented by such a minimal equation and we deduce the reduction of the curve using the reduction of such a representative equation. In order to introduce minimal Weierstrass equations, we first need the definition of the p-adic absolute value on $\mathbb{Q}$ of a prime number p.

1.4.2 The p-adic absolute value on $\mathbb{Q}$

An *absolute value* on an integral ring A is a map $A \longrightarrow \mathbb{R}$ defined by $x \longmapsto |x|$, that satisfies the following conditions:

(i) $|x| \geq 0$ for all $x \in A$ and $|x| = 0$ if and only if $x = 0$.

(ii) $|xy| = |x|.|y|$ for all $x, y \in A$.

(iii) $|x + y| \leq |x| + |y|$, for all $x, y \in A$ (triangle inequality).

For example, the usual absolute value $|\cdot|$ on $\mathbb{R}$ is an absolute value in this sense, but also $x \mapsto 2|x|$.

Definition 1.26. The *p-adic absolute value* on $\mathbb{Q}$ is the map from $\mathbb{Q}$ to $\mathbb{R}$ defined by $x \longmapsto |x|_p$, where $|0|_p = 0$ and $|x|_p = p^{-n}$ if $x = p^n \frac{a}{b}$ with $n \in \mathbb{Z}$ and $a, b \in \mathbb{Z} \setminus p\mathbb{Z}$.

The function $x \mapsto |x|_p$ is indeed an absolute value. It satisfies the inequality $|x + y|_p \leq \sup\{|x|_p, |y|_p\}$, for all $x, y \in \mathbb{Q}$, which is stronger than the triangle inequality. Because of this, we say it is an *ultrametric* or *non-archimedean* absolute value on $\mathbb{Q}$.

Now that we have defined the p-adic absolute value we may consider the following set:

$$\mathbb{Z}_{(p)} = \{x \in \mathbb{Q} : |x|_p \leq 1\} = \left\{ \frac{r}{q} : r \in \mathbb{Z},\ q \in \mathbb{Z} \setminus p\mathbb{Z} \right\}.$$

This is a subring of $\mathbb{Q}$ and is referred to as the *valuation ring* of $|\cdot|_p$. If $x \in \mathbb{Q}$ is an element of $\mathbb{Z}_{(p)}$, then we say that x is a *p-integer*, or that x is *p-integral.*

1.4.3 Minimal Weierstrass equations

Let E be an elliptic curve defined over $\mathbb{Q}$. In Section 1.3.1 we saw that such a curve can be defined by a long Weierstrass equation with integer coefficients:

$$y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6. \tag{1.26.1}$$

Let p be a prime. If the coefficients of this equation are all p-integers, then it is said to be *p-integral.*

Definition 1.27. Equation 1.26.1 is said to be *p-minimal* if the following conditions hold:

(i) The equation is p-integral.

(ii) $|\Delta|_p$ cannot increase if we perform a feasible change of coordinates in $\mathbb{Q}$ resulting in a new p-integral equation.

The last condition tells us that if we perform a feasible change of coordinates and the new equation is p-integral, then we must have $|\Delta|_p \geq |\Delta'|_p$. In other words, the discriminant of a p-minimal Weierstrass equation is maximal in terms of the p-adic absolute value.

From now on we shall write a feasible change of coordinates

$$x = u^2x' + r \qquad y = u^3y' + su^2x' + t \tag{1.27.1}$$

with $u, r, s, t \in \mathbb{Q}$. We then have:

$$\begin{array}{lcl} ua'_1 & = & a_1 + 2s \\ u^2a'_2 & = & a_3 - sa_1 + 3r - s^2 \\ u^3a'_3 & = & a_3 + ra_1 + 2t \\ u^4a'_4 & = & a_4 - sa_3 + 2ra_2 - (t + rs)a_1 + 3r^2 - 2st \\ u^6a'_6 & = & a_6 + ra_4 + r^2a_2 + r^3 - ta_3 - t^2 - rta_1 \\ u^4c'_4 & = & c_4 \\ u^6c'_6 & = & c_6 \\ u^{12}\Delta' & = & \Delta. \end{array} \tag{1.27.2}$$

We now state and prove two lemmas that we will use in the proof of Néron's Theorem on the existence of minimal Weierstrass equations..

Lemma 1.28. *Let p be a prime and suppose that the coefficients of Equation 1.26.1 are p-integers. Then:*

(i) If $|\Delta|_p > p^{-12}$ or $|c_4|_p > p^{-4}$ or $|c_6|_p > p^{-6}$, then the equation is p-minimal.

(ii) If $p > 3$ and if $|\Delta|_p \leq p^{-12}$ and $|c_4|_p \leq p^{-4}$, then the equation is not p-minimal.

Proof. To prove (i), assume that $|\Delta|_p > p^{-12}$. Suppose, by contradiction, that the equation is not p-minimal. Then there exists a feasible change of coordinates (u, r, s, t) such that the new equation is p-integral and

$$|\Delta'|_p > |\Delta|_p. \tag{1.28.1}$$

Let $|u|_p = p^{-n}$, with n an integer. Since $u^{12}\Delta' = \Delta$ we see that $|\Delta|_p = |\Delta'|_p.|u|_p^{12}$. Using Inequality 1.28.1, we see that $|\Delta'|_p > |\Delta'|_p.|u|_p^{12}$, and thus $|u|_p < 1$. Furthermore, $\Delta' \in \mathbb{Z}_{(p)}$ since $\mathbb{Z}_{(p)}$ is a ring and $a'_i \in \mathbb{Z}_{(p)}, i = 1, 2, 3, 4, 6$ by hypothesis. This implies that $|\Delta'|_p \leq 1$. From this it follows that

$$|u|_p^{12} = \frac{|\Delta|_p}{|\Delta'|_p} > \frac{p^{-12}}{|\Delta'|_p} \geq p^{-12} \implies p^{-12n} > p^{-12} \implies p^{12(1-n)} > 1 \implies n < 1.$$

But $|u|_p < 1$, and therefore we must have $n > 0$. Finally this gives us $0 < n < 1$ and this is a contradiction. The cases where $|c_4|_p > p^{-4}$ and $|c_6|_p > p^{-6}$ can be proved in a similar way.

In order to prove (ii) we let $p > 3$ be a prime number and suppose that $|\Delta|_p \leq p^{-12}$ and $|c_4|_p \leq p^{-4}$. Using Definition 1.17, we see that $1728\Delta = c_4^3 - c_6^2$. Note that the prime factor decomposition of 1728 is 2^63^3. Since $p > 3$, p does not enter this decomposition and therefore $|1728|_p = 1$. Using the fact that $|\cdot|_p$ is ultrametric, we compute:

$$|c_6|_p^2 = |c_4^3 - 1728\Delta|_p \leq \sup\{|c_4^3|_p, |-1728\Delta|_p\} = \sup\{|c_4^3|_p, |1728\Delta|_p\}.$$

But $|c_4^3|_p = |c_4|_p^3 \leq (p^{-4})^3 = p^{-12}$ and $|1728\Delta|_p = |1728|_p.|\Delta|_p = |\Delta|_p \leq p^{-12}$. This shows that $\sup\{|c_4^3|_p, |1728\Delta|_p\} \leq p^{-12}$, which implies that $|c_6|_p \leq p^{-6}$. Similarly,

$$\begin{array}{lcl} |\Delta|_p \leq p^{-12}, |c_6|_p \leq p^{-6} & \Longrightarrow & |c_4|_p \leq p^{-4}, \\ |c_4|_p \leq p^{-4}, |c_6|_p \leq p^{-6} & \Longrightarrow & |\Delta|_p \leq p^{-12}. \end{array}$$

We have seen before that transforming, via a feasible change of coordinates, Equation 1.26.1 into a short Weierstrass equation yields the equation:

$$y^2 = x^3 - \frac{c_4}{48}x - \frac{c_6}{864}.$$

On this equation we now perform the feasible change of coordinates $(p, 0, 0, 0)$. This yields the equation:

$$y'^2 = x'^3 - \frac{c_4}{48}p^{-4}x' - \frac{c_6}{864}p^{-6}. \tag{1.28.2}$$

The coefficients of this new equation are p-integers. In fact,

$$|\frac{c_4}{48}p^{-4}|_p = |c_4|_p.|p^{-4}|_p = |c_4|_p.p^4 \leq p^{-4}.p^4 = 1.$$

This computation can only be done because $p > 3$ and therefore $|48|_p = 1$. A similar calculation shows that $|\frac{c_6}{864}p^{-6}|_p \leq 1$. Thus Equation 1.28.2 is p-integral. From System 1.27.2, we know that $\Delta' = p^{-12}\Delta$ and thus

$$|\Delta'|_p = |p^{-12}|_p.|\Delta|_p = p^{12}.|\Delta|_p > |\Delta|_p.$$

So the initial Equation 1.26.1 is not minimal. □

Lemma 1.29. *Let E be an elliptic curve defined over $\mathbb{Q}$ and p be a prime.*

(i) Any equation for E can be made p-minimal via a feasible change of coordinates with coefficients in $\mathbb{Q}$.

(ii) If an equation for E is already p-integral, then the resulting equation of this change of coordinates is also p-integral.

(iii) Two minimal equations at p for E are related by a feasible change of coordinates for which $|u|_p = 1$ and r, s, t are p-integers.

Proof. For this proof, we will suppose that $p > 3$. We start by proving (i). We may suppose that E is given by a Weierstrass equation with integer coefficients. Thus $|\Delta|_p \leq 1$. Since $\Delta \neq 0$ (because otherwise E would not be an elliptic curve), we have $|\Delta|_p > 0$. So there is only a finite number of $|\Delta'|_p$ between $|\Delta|_p$ and 1, and this proves the existence of a p-minimal equation.

We prove statement (ii). If the resulting equation is p-minimal, then the eighth equation of System 1.27.2 tells us that $|u|_p \leq 1$. Since $a_i, a_i' \in \mathbb{Z}_{(p)}$, $i = 1, 2, 3, 4, 6$, the first three equations of System 1.27.2 tell us that $r, s, t \in \mathbb{Z}_{(p)}$.

Finally we prove the last statement. If a given equation for E is p-minimal, then we have already seen in (ii) that $|u|_p \leq 1$. The u-parameter of the inverse transformation is u^{-1}. So, by (ii), we also have $|u^{-1}|_p \leq 1$. But $|u^{-1}|_p = |u|_p^{-1}$, so $|u|_p \geq 1$ and therefore $|u|_p = 1$. We already know from (ii) that $r, s, t \in \mathbb{Z}_{(p)}$ and thus the proof is complete. □

Definition 1.30. Equation 1.26.1 is said to be *globally minimal* if:

(i) Its coefficients are integers.

(ii) The equation is p-minimal for all prime numbers p.

Example 1.31. Let E be an elliptic curve defined over $\mathbb{Q}$ and given by the equation $E : y^2 = x^3+1$. Clearly, the coefficients of this equation are integers. We saw in Example 1.25 that $\Delta = -432 = -2^4 3^3$. If $p > 3$ is a prime number, then $|\Delta|_p = 1$ and the equation is p-minimal. Otherwise, we compute that $|\Delta|_2 = 2^{-4} > 2^{-12}$ and $|\Delta|_3 = 3^{-3} > 3^{-12}$. By Lemma 1.28, the equation is 2-minimal and 3-minimal. Thus the equation is globally minimal.

In order to prove that all elliptic curves possess a globally minimal equation we need this next lemma.

Lemma 1.32. *Let p_1, ..., p_n be a finite set of prime numbers and ϵ_1, ..., ϵ_n be positive real numbers. For all $i \in \{1, ..., n\}$ take p_i-integers $x_i \in \mathbb{Z}_{(p_i)}$. Then there exists $x \in \mathbb{Z}$ such that, for all $1 \leq i \leq n$, we have $|x - x_i|_{p_i} \leq \epsilon_i$.*

Proof. Let p be prime and notice that $\mathbb{Z}_{(p)} \subset \mathbb{Z}_p = \{x \in \mathbb{Q}_p : |x|_p \leq 1\}$, where $\mathbb{Q}_p$ is the p-adic field, that is, the completion of $\mathbb{Q}$ with respect to $|\cdot|_p$. Actually, one can show that $\mathbb{Z}_p$ is the completion of $\mathbb{Z}$ with respect to $|\cdot|_p$. By construction of this completion, $\mathbb{Z}$ is dense in $\mathbb{Z}_p$.

Considering the above remark, we may choose integers $x'_i \in \mathbb{Z}$ such that $|x_i - x'_i|_{p_i} < \frac{\epsilon_i}{2}$ for all $1 \leq i \leq n$. Let m be an arbitrary positive integer. By the classic Chinese Remainder Theorem, the system of linear congruences $x \equiv x'_i (\text{mod } p_i^m)$ for $1 \leq i \leq n$ has a unique solution x (mod $(p_1 p_2 \dots p_n)^m$). But then, for all i, there exists $c_i \in \mathbb{Z} \setminus p_i\mathbb{Z}$ and a positive integer k_i such that $x - x'_i = c_i p_i^{k_i m}$. Thus $|x - x'_i|_{p_i} = p_i^{-k_i m} \to 0$, when $m \to \infty$. Hence, there exists a positive integer N_i such that $|x - x'_i|_{p_i} < \frac{\epsilon_i}{2}$ whenever $m \geq N_i$. Let N be the maximum of all N_i. If we choose $m \geq N$, then we obtain

$$|x - x_i|_{p_i} \leq |x - x'_i|_{p_i} + |x'_i - x_i|_{p_i} < \epsilon_i$$

for all $1 \leq i \leq n$, as desired. □

1.33. Lemma 1.32 is actually equivalent to the Chinese Remainder Theorem. We do not prove the other implication since we do not need it but it is worth noting. This lemma is therefore often directly referred to as the Chinese Remainder Theorem.

Theorem 1.34 (Néron)**.** *Let E be an elliptic curve defined over $\mathbb{Q}$ and given by a Weierstrass equation.*

(i) There exists a change of coordinates 1.27.1, with coefficients in $\mathbb{Q}$, such that the resulting equation is globally minimal.

(ii) Two globally minimal equations for the same curve E are related by a change of coordinates 1.27.1 such that $u = \pm 1$ and r, s, t are integers.

Proof. We do not need to prove (ii) since it is a direct consequence of Lemma 1.29. So we only need to prove part (i), that is, the existence of a globally minimal equation.

Let E be an elliptic curve defined over $\mathbb{Q}$ by a Weierstrass equation. As seen already, up to feasible change of coordinates we may suppose that the coefficients of this equation are all integers. This implies that Δ is also an integer.

Let p be a prime that divides Δ. Then, by Lemma 1.29, there exists a feasible change of coordinates (u_p, r_p, s_p, t_p) such that the resulting equation is p-minimal. This same lemma also tells us that $u_p, r_p, s_p, t_p \in \mathbb{Z}_{(p)}$. From System 1.27.2 we know that $|u_p|_p^{12}|\Delta_p|_p = |\Delta|_p$, where Δ_p is the discriminant of the new equation. Let us write $u_p = p^{\lambda_p}\theta_p$, where $\theta_p \in \mathbb{Z}_{(p)}, |\theta_p|_p = 1$ and $\lambda_p \geq 0$ (point (ii) of Lemma 1.29). Then $|u_p|_p = p^{-\lambda_p} \leq 1$. Now, we perform a feasible change of coordinates with

$$u = \prod_{p|\Delta} p^{\lambda_p} \in \mathbb{N}.$$

Then

$$|\Delta'|_p = |u|_p^{-12}|\Delta|_p = \left(\prod_{\ell|\Delta} |\ell^{\lambda_\ell}|_p\right)^{-12} |\Delta|_p = |p^{\lambda_p}|_p^{-12}|\Delta|_p = p^{12\lambda_p}|\Delta|_p = |u_p|_p^{-12}|\Delta|_p = |\Delta_p|_p.$$

So this new equation is minimal at all primes that divide Δ.

Since $|\Delta'|_\ell = 1$ if ℓ is a prime that does not divide Δ, the equation is p-minimal for all p, where p is prime. The only thing we need to do now is to show that the coefficients of this equation are integers. Then by definition it is globally minimal.

We need to choose r, s, t wisely and we will do this using Lemma 1.32. We shall take $r, s, t \in \mathbb{Z}$ such that, for all primes p that divide Δ, we have

$$|r - r_p|_p \leq p^{-6\lambda_p} \qquad |s - s_p|_p \leq p^{-6\lambda_p} \qquad |t - t_p|_p \leq p^{-6\lambda_p}.$$

This is possible since there is only a finite number of primes dividing the discriminant. To prove that a'_i, $i = 1, 2, 3, 4, 6$ are integers we need to show that $|a'_i|_\ell \leq 1$ for all $i = 1, 2, 3, 4, 6$ and all p prime. If ℓ does not divide Δ, then there is no problem since a_i, r, s, t are integers.

For $i = 1$, if $\ell = p$ does divide Δ, then write

$$ua'_1 = a_1 + 2s = a_1 + 2(s - s_p) + 2s_p = u_p a_{1,p} + 2(s - s_p).$$

It follows that

$$|ua_1'|_p \leq \max\{|u_p a_{1,p}|_p, |2(s-s_p)|_p\} \leq \max\{|u_p|_p, p^{-6\lambda_p}\} = \max\{p^{-\lambda_p}, p^{-6\lambda_p}\} = p^{-\lambda_p} = |u_p|_p.$$

But $|u_p|_p = |u|_p$, so $|ua_1'|_p = |u|_p|a_1'|_p \leq |u|_p$, thus $|a_1'|_p \leq 1$.

For $i=2$, if $\ell = p$ does divide Δ, then write

$$\begin{aligned} u^2a_2' = a_3 - sa_1 + 3r - s^2 &= a_3 - s_pa_1 + 3r_p - s_p^2 + (s_p - s)a_1 + 3(r - r_p) + (s_p^2 - s^2) \\ &= u_p^2a_{2,p} + (s_p - s)(a_1 + s + s_p) + 3(r - r_p)|u^2a_2'|_p \leq \max\{|u_p|_p^2, p^{-6\lambda_p}\} = |u_p|_p^2, \end{aligned}$$

thus $|a_2'|_p \leq 1$.

The other cases for i can be proved in a similar way using System 1.27.2. □

We may now characterize the reduction modulo a prime p of an elliptic curve. The following definition does not depend on the choice of the minimal Weierstrass equation.

Definition 1.35. Let p be a prime, E be an elliptic curve defined over $\mathbb{Q}$ and let W be a minimal Weierstrass equation for E. We say that

(i) E admits good reduction at p if W admits good reduction at p.

(ii) E admits split (respectively non-split) multiplicative reduction at p if W admits split (respectively non-split) multiplicative reduction at p.

(iii) E admits additive reduction at p if W admits additive reduction at p.

We now introduce the conductor of an elliptic curve E. It is a positive integer N_E, which contains information about the reduction of E. The conductor will be of major importance in what will follow. In the case of additive reduction at 2 or 3, we refer to [Og67].

Definition 1.36. Let E be an elliptic curve defined over $\mathbb{Q}$. The *conductor* N_E is the product

$$N_E = \prod_{p \text{ prime}} p^{f_p},$$

where

$$f_p = \begin{cases} 0 & \text{if } E \text{ has good reduction at } p, \\ 1 & \text{if } E \text{ admits multiplicative reduction at } p, \\ 2 + \delta_p & \text{if } E \text{ admits additive reduction at } p. \end{cases}$$

where $\delta_p \geq 0$ is a certain integer whose precise definition we omit, except for saying that $\delta_p = 0$ if $p > 3$.

Example 1.37. In this example we compute the conductor of a well-known elliptic curve. Let E be an elliptic curve defined over $\mathbb{Q}$ and given by the equation $E : y^2 - y = x^3 - x^2$. Comparing with Definition 1.14.1, we see that $a_1 = a_4 = a_6 = 0$ as well as $a_2 = a_3 = -1$. A computation shows that $c_4 = 16$ and $c_6 = -152$, thus

$$\Delta = \frac{16^3 - 152^2}{12^3} = -11.$$

The coefficients of the equation are integers and if $p \neq 11$ is prime, then $|\Delta|_p = 1$ and thus the equation is p-minimal. Note that $|\Delta|_{11} = 11^{-1} > 11^{-12}$ and by Lemma 1.28, the equation is 11-minimal. Thus the equation is globally minimal.

If $p \neq 11$ is a prime number, then $f_p = 0$. If $p = 11$, then $c_4 \not\equiv 0 \bmod 11$. So E admits a double point and the reduction modulo 11 is multiplicative, hence $f_{11} = 1$. This leads to $N_E = 11$.

We will show that E has split multiplicative reduction at 11. We shall use the notation $\overline{E}$ for the curve E modulo 11. We have

$$E_{11} : f(x, y) = 0,$$

where $f(x, y) = y^2 + 10y + 10x^3 + x^2$. We look for the singular point $P = (x_0, y_0)$ of E_{11} by forcing the equality between the gradient of f at P and 0:

$$\operatorname{grad}(f)(x_0, y_0) = (8x_0^2 + 2x_0, 2y_0 + 10) = (0, 0).$$

A calculation shows that $P = (8, 6)$. We take the third order Taylor expansion of f in a neighborhood of P:

$$f(x, y) = (y - 6)^2 - 9(x - 8)^2 - (x - 8)^3 = [(y - 6) - 3(x - 8)][(y - 6) - 8(x - 8)] - (x - 8)^3.$$

Thus, the tangent lines of E_{11} at P are $y = 3x + 7$ and $y = 8x + 8$. Both of these lines have slopes in $\mathbb{F}_{11}$, hence E admits split multiplicative reduction at 11.

1.5 The L-series of elliptic curves over $\mathbb{Q}$

We define and study the L-series of elliptic curves over $\mathbb{Q}$. These series will constitute the link that we will establish between elliptic curves and modular forms. We start with some important theorems and results on elliptic curves over finite fields, mostly without proof.

1.5.1 Results concerning elliptic curves over finite fields

Let p be a prime and n a positive integer. The following theorem was first conjectured by E. Artin in 1924 in his thesis and proved in 1934 by the German mathematician H. Hasse.

Theorem 1.38 (Hasse)**.** *Let E be an elliptic curve over $\mathbb{F}_{p^n}$. Then the order of $E(\mathbb{F}_{p^n})$ satisfies*

$$|\#E(\mathbb{F}_{p^n}) - (p^n + 1)| \leq 2\sqrt{p^n}.$$

Proof. For the original article by Hasse we refer to [Ha36]. For a modern proof, see Theorem 1.1 of [Si86] Section V.I. □

This result gives an approximation of the order of $E(\mathbb{F}_{p^n})$ and tells us that it does not differ very much from $p^n + 1$. We state the following theorem without proof.

Theorem 1.39. *Let E be an elliptic curve over $\mathbb{F}_p$ and define $a = p + 1 - \#E(\mathbb{F}_p)$. Write $X^2 - aX + p = (X - \alpha)(X - \beta)$. Then, the equality*

$$\#E(\mathbb{F}_{p^n}) = 1 - \alpha^n - \beta^n + p^n$$

holds for all $n \geq 1$.

1.40. The numbers α, β in Theorem 1.39 are two conjugate complex numbers that satisfy $|\alpha| = |\beta| = \sqrt{p}$. In fact, we have

$$\alpha = \frac{a - \sqrt{a^2 - 4p}}{2} \qquad \beta = \frac{a + \sqrt{a^2 - 4p}}{2}.$$

Since $|a| \leq 2\sqrt{p}$, we have $a^2 - 4p \leq 0$. Hence α and β are complex conjugates. Since $2\alpha = a - i\sqrt{4p - a^2}$, we have $2|\alpha| = \sqrt{a^2 + 4p - a^2} = 2\sqrt{p}$. Since α and β are conjugates, we know that $|\alpha| = |\beta| = \sqrt{p}$.

Example 1.41. In this example, we compute $\#E(\mathbb{F}_9)$ for an elliptic curve E over $\mathbb{F}_3$ in two different ways. We numerically illustrate the above theorems. Let E be the elliptic curve of Example 1.37 viewed over $\mathbb{F}_3$:

$$E : y^2 - y = x^3 - x^2.$$

Since $\Delta = -11$, we know that E has good reduction at 3. A short computation shows that $E(\mathbb{F}_3) = \{\mathcal{O}, (0,0), (0,1), (1,0), (1,1)\}$, thus $\#E(\mathbb{F}_3) = 5$. By definition, we then have $a = 3 + 1 - 5 = -1$. Then

$$X^2 - aX + p = X^2 + X + 3 = \left(X - \frac{-1 - i\sqrt{11}}{2}\right)\left(X - \frac{-1 + i\sqrt{11}}{2}\right) = (X - \alpha)(X - \beta).$$

Using Theorem 1.39, $\#E(\mathbb{F}_9) = 1 + 5 + 9 = 13$.

We now compute $\#E(\mathbb{F}_9)$ by hand. Consider the irreducible polynomial $X^2 - 2 \in \mathbb{F}_3[X]$. Let ξ be a root of this polynomial. Then we may identify $\mathbb{F}_9$ with $\mathbb{F}_3(\xi) = \{0, 1, 2, \xi, \xi + 1, \xi + 2, 2\xi, 2\xi + 1, 2\xi + 2\}$.

Using Table 1, we can count the order of $E(\mathbb{F}_9)$. Remember to add the point $\mathcal{O}$ to the list, since it is not represented in this table. Doing this, we find that $\#E(\mathbb{F}_9) = 13$.

x	x^2	x^3	$x^3 - x^2$	$x^2 - x$	$\#y \mid y^2 - y = x^3 - x^2$
0	0	0	0	0	2
1	1	1	0	0	2
2	1	2	1	2	1
ξ	2	2ξ	$2\xi + 1$	$2\xi + 2$	0
$\xi + 1$	2ξ	$2\xi + 1$	1	$\xi + 2$	1
$\xi + 2$	ξ	$2\xi + 2$	$\xi + 2$	$\xi + 1$	2
2ξ	2	ξ	$\xi + 1$	$\xi + 2$	1
$2\xi + 1$	ξ	$\xi + 1$	1	$2\xi + 2$	1
$2\xi + 2$	2ξ	$\xi + 2$	$2\xi + 2$	1	2

Table 1: Computations in $\mathbb{F}_9$

1.5.2 The zeta function of elliptic curves over finite fields

We define the zeta function of an elliptic curve and study some of its properties.

Definition 1.42. Let p be a prime and let E be an elliptic curve over $\mathbb{F}_p$. Let $N_m = \#E(\mathbb{F}_{p^m})$, for all $m \geq 1$ integer. The *Z-function* of E and the *local zeta-function* ζ_p of E are respectively the formal power series and the function defined by:

$$Z(E,T) = \exp\left(\sum_{m=1}^{\infty} \frac{N_m}{m} T^m\right) \qquad \zeta_p(E,s) = Z(E, p^{-s}),$$

where $s \in \mathbb{C}$.

Proposition 1.43. *Let E be an elliptic curve over $\mathbb{F}_p$ and let $a = p + 1 - \#E(\mathbb{F}_p)$. Then*

$$Z(E,T) = \frac{pT^2 - aT + 1}{(1-T)(1-pT)} \qquad \zeta_p(E,s) = \frac{1 - ap^{-s} + p^{1-2s}}{(1-p^{-s})(1-p^{1-s})}.$$

Proof. In this proof we use the fact that,

$$\sum_{m=1}^{\infty} \frac{(\xi x)^m}{m} = -\log(1 - \xi x),$$

where the term on the left is a formal series. Since $\#E(\mathbb{F}_p) = p + 1 - a$, we know by Theorem 1.39 that $\#E(\mathbb{F}_{p^m}) = 1 - \alpha^m - \beta^m + p^m$, where α and β are the complex roots of the polynomial

$X^2 - aX + p$. Knowing this we compute:

$$\begin{aligned}
Z(E,T) &= \exp\left(\sum_{m=1}^{\infty} \frac{1-\alpha^m-\beta^m+p^m}{m} T^m\right) \\
&= \exp\left(\sum_{m=1}^{\infty} \frac{T^m}{m} + \frac{(pT)^m}{m} - \frac{(\alpha T)^m}{m} - \frac{(\beta T)^m}{m}\right) \\
&= \exp\left(-\log(1-T) - \log(1-pT) + \log(1-\alpha T) + \log(1-\beta T)\right) \\
&= \frac{(1-\alpha T)(1-\beta T)}{(1-T)(1-pT)} \\
&= \frac{pT^2 - aT + 1}{(1-T)(1-pT)}.
\end{aligned}$$

To prove the equality concerning the function ζ_p, one only needs to substitute T with p^{-s}. □

1.44. Notice that Proposition 1.43 implies Theorem 1.39. Hence, the two are equivalent.

We can experiment a little with this Z-function. Let p be a prime, $m \geq 1$ and C be the curve defined over $\mathbb{F}_{p^m}$ by the equation $xy = 0$. The curve C is obviously not an elliptic curve. However, let us define the Z-function of C as in Definition 1.42 by

$$Z(C,T) = \exp\left(\sum_{m=1}^{\infty} \frac{N_m}{m} T^m\right),$$

where $N_m = \#C(\mathbb{F}_{p^m})$. It is easy to see that $N_m = 2p^m - 1$. We now compute that

$$\begin{aligned}
Z(C,T) &= \exp\left(\sum_{m=1}^{\infty} \frac{2p^m-1}{m} T^m\right) \\
&= \exp\left(\sum_{m=1}^{\infty} \frac{(pT)^m}{m} + \frac{(pT)^m}{m} - \frac{T^m}{m}\right) \\
&= \exp\left(-2\log(1-pT) + \log(1-T)\right) \\
&= \frac{(1-T)}{(1-pT)^2}.
\end{aligned}$$

So the Z-function of C is again rational. This is not a coincidence. In fact, a result by Dwork and Grothendieck states that we can define a Z-function for any finite system of equations over $\mathbb{F}_p$ by counting solutions and this function will always be rational.

1.5.3 The L-series

We start by defining the L-function of an elliptic curve and then we show that this function takes the form of a Dirichlet series.

Let E be an elliptic curve over $\mathbb{Q}$ and let p be a prime.

Definition 1.45. The *L-function* of E at p is defined, for $s \in \mathbb{C}$, by

$$L_p(E,s) = \begin{cases} (1 - a_p p^{-s} + p^{1-2s})^{-1} & \text{if } E \text{ has good reduction at } p, \\ (1 - a_p p^{-s})^{-1} & \text{if } E \text{ admits bad reduction at } p, \end{cases}$$

where

$$a_p = \begin{cases} p+1-\#E(\mathbb{F}_p) & \text{if } E \text{ has good reduction at } p, \\ 1 & \text{if } E \text{ admits split multiplicative reduction at } p, \\ -1 & \text{if } E \text{ admits non-split multiplicative reduction at } p, \\ 0 & \text{if } E \text{ admits additive reduction at } p. \end{cases}$$

1.46. If E admits good reduction at p, then $L_p(E,s)$ is exactly the numerator of the local zeta function of E.

Definition 1.47. The *Hasse-Weil L-function* of E is defined by

$$L(E,s) = \prod_{p \text{ prime}} L_p(E,s),$$

where $s \in \mathbb{C}$.

The Hasse-Weil L-function of E has analogous properties to the ones of the Riemann zeta function

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

We recall that the Riemann zeta function converges for $\mathrm{Re}(s) > 1$ and that it can be expressed as an Euler product:

$$\zeta(s) = \prod_{p \text{ prime}} \frac{1}{1-p^{-s}}.$$

The local L_p-functions of E correspond by analogy to the terms $\frac{1}{1-p^{-s}}$.

1.48. We may write the Hasse-Weil L-function under the form

$$L(E,s) = \prod_{\text{bad } p} (1-a_p p^{-s})^{-1} \prod_{\text{good } p} (1-a_p p^{-s}+p^{1-2s})^{-1},$$

where the first product is finite since it only concerns primes p that induce bad reduction and the second product is infinite.

Proposition 1.49. *The Hasse-Weil L-function of E converges for* $\mathrm{Re}(s) > \frac{3}{2}$.

Proof. Suppose that $\mathrm{Re}(s) > 3/2$. By Remark 1.48, we only need to show the result for

$$\prod_{\text{good } p} (1-a_p p^{-s}+p^{1-2s})^{-1}, \tag{1.49.1}$$

since the product over the bad primes is finite.

For any good prime p, consider the polynomial $1 - a_pT + pT^2$ and let θ_p be any one of its two complex roots. Then product 1.49.1 becomes

$$\prod_{\text{good } p} (1 - \theta_p p^{-s})^{-1}(1 - \overline{\theta}_p p^{-s})^{-1}.$$

Since $|\theta_p| = p^{\frac{1}{2}}$, we see that $|\theta_p p^{-s}| = p^{\frac{1}{2}-\mathrm{Re}(s)}$. Thus, $|1 - \theta_p p^{-s}| \geq |1 - |\theta_p p^{-s}|| = |1 - p^{\frac{1}{2}-\mathrm{Re}(s)}|$. The same arguments hold for the conjugate of θ_p. Now,

$$\prod_{\text{good } p} |(1 - a_p p^{-s} + p^{1-2s})^{-1}| \leq \prod_{\text{good } p} |\frac{1}{1 - p^{\frac{1}{2}-\mathrm{Re}(s)}}|^2.$$

We state that $|1 - p^{\frac{1}{2}-\mathrm{Re}(s)}|^{-1} \geq 1$. In fact, since $\mathrm{Re}(s) > 3/2$, we have

$$|1 - p^{\frac{1}{2}-\mathrm{Re}(s)}| < |1 - p^{\frac{1}{2}-\frac{3}{2}}| = |1 - \frac{1}{p}| \leq 1.$$

This implies that

$$\prod_{\text{good } p} (1 - a_p p^{-s} + p^{1-2s})^{-1} \leq \prod_{p \text{ prime}} |\frac{1}{1 - p^{\frac{1}{2}-\mathrm{Re}(s)}}|^2 = |\zeta(\mathrm{Re}(s) - \frac{1}{2})|^2.$$

We have $\mathrm{Re}(s) > \frac{3}{2} \iff \mathrm{Re}(s) - \frac{1}{2} > 1 \iff \mathrm{Re}(\mathrm{Re}(s) - \frac{1}{2}) > 1$. Thus $|\zeta(\mathrm{Re}(s) - \frac{1}{2})|^2$ is finite. Furthermore, the product 1.49.1 is non-zero and this proves the absolute convergence of 1.49.1 and finishes the proof.

An alternative way to prove this is to use the following result of real analysis: for real numbers $\alpha_n \geq 0$,

$$\sum_{n=1}^{\infty} \alpha_n \text{ converges} \iff \prod_{n=1}^{\infty} (1 + \alpha_n) \text{ converges}.$$

Thus, we need only to verify that $\sum_p | - a_p p^{-s} + p^{1-2s}| < +\infty$. But

$$\sum_p | - a_p p^{-s} + p^{1-2s}| \leq \sum_p |a_p p^{-s}| + \sum_p |p^{1-2s}|.$$

We start by dealing with the second term:

$$\sum_p |p^{1-2s}| = \sum_p p^{1-2\,\mathrm{Re}(s)} < \sum_p p^{-2} \leq \sum_{n=1}^{\infty} n^{-2} < +\infty.$$

For the other term, we notice that $|a_p| \leq 2p^{\frac{1}{2}}$ by Theorem 1.38. We then get:

$$\sum_p |a_p p^{-s}| = \sum_p |a_p| p^{\mathrm{Re}(s)} \leq 2\sum_p p^{\frac{1}{2}-\mathrm{Re}(s)} \leq 2\sum_p p^{\frac{1}{2}-(\frac{3}{2}-\epsilon)} \leq 2\sum_{n=1}^{\infty} n^{-\delta} < +\infty,$$

where $\epsilon > 0$ and $\delta = 1 - \epsilon < 1$. This ends the alternative proof. □

Definition 1.50. We can expand the Hasse-Weil function of an elliptic curve E and transform it into a Dirichlet series:

$$L(E,s) = \sum_{m=1}^{\infty} \frac{a_n}{n^s},$$

where $s \in \mathbb{C}$. This is called the L-*series* of E.

This definition does not give us the coefficients of the L-series. We now show how to obtain these. Recall that

$$L(E,s) = \prod_{\text{bad } p} (1 - a_p p^{-s})^{-1} \prod_{\text{good } p} (1 - a_p p^{-s} + p^{1-2s})^{-1}$$

and recall the formula for the infinite geometric series:

$$\frac{1}{1-x} = \sum_{m=0}^{\infty} x^m. \tag{1.50.1}$$

Let p be a good prime, in other words a prime such that E has good reduction at p. Then, by using Equation 1.50.1, we see that

$$\frac{1}{1 - a_p p^{-s} + p^{1-2s}} = \sum_{m=0}^{\infty} (a_p p^{-s} - p^{1-2s})^m.$$

Using the Newton binomial, we compute that

$$\begin{aligned}
(a_p p^{-s} - p^{1-2s})^2 &= a_p^2 p^{-2s} - 2a_p p p^{-3s} + p^2 p^{-4s}, \\
(a_p p^{-s} - p^{1-2s})^3 &= a_p^3 p^{-3s} - 3a_p^2 p p^{-4s} + 3a_p p^2 p^{-5s} - p^3 p^{-6s}, \\
(a_p p^{-s} - p^{1-2s})^4 &= a_p^4 p^{-4s} - 4a_p^3 p p^{-5s} + 6a_p^2 p^2 p^{-6s} - 4a_p p^3 p^{-7s} + p^4 p^{-8s}.
\end{aligned}$$

Thus,

$$\begin{aligned}
\frac{1}{1 - a_p p^{-s} + p^{1-2s}} = 1 + \frac{a_p}{p^s} + \frac{a_p^2 - p}{p^{2s}} + \frac{a_p^3 - 2a_p p}{p^{3s}} + \frac{a_p^4 - 3a_p^2 p + p^2}{p^{4s}} + \frac{3a_p p^2 - 4a_p^3 p}{p^{5s}} \\
+ \frac{6a_p^2 p^2 - p^3}{p^{6s}} - \frac{4a_p p^3}{p^{7s}} + \frac{p^4}{p^{8s}} + \sum_{m=5}^{\infty} (a_p p^{-s} - p^{1-2s})^m.
\end{aligned}$$

1.51. The definition of a_{p^k} is that it is the coefficient of p^{-ks}. So we can set

$$a_{p^2} = a_p^2 - p,\ a_{p^3} = a_p^3 - 2a_p p,\ a_{p^4} = a_p^4 - 3a_p^2 p + p^2.$$

Generalizing this, we find that $a_p a_{p^n} = a_{p^{n+1}} + p a_{p^{n-1}}$ if p is prime, $n \geq 1$.

Let p be a bad prime, that is, a prime such that E has bad reduction at p. Then, by using Equation 1.50.1, we get

$$\frac{1}{1 - a_p p^{-s}} = \sum_{m=0}^{\infty} (a_p p^{-s})^m = 1 + \frac{a_p}{p^s} + \frac{a_p^2}{p^{2s}} + \sum_{m=3}^{\infty} (a_p p^{-s})^m.$$

1.52. By multiplying out these infinite series for various p to get terms in the Dirichlet series we see that:

(i) If $n = \prod_j p_j^{e_j}$, then $a_n = \prod_j a_{p_j^{e_j}}$.

(ii) If $n = p$ with p prime, then a_n is the same as a_p of Definition 1.45.

Example 1.53. We will compute the first terms of the L-series of the elliptic curve E defined over $\mathbb{Q}$ by

$$E : y^2 - y = x^3 - x^2.$$

We have already seen in Example 1.37 that $\Delta = -11$, that this equation is globally minimal and that E admits split multiplicative reduction at 11. Now, since E admits split multiplicative reduction at 11, we use Definition 1.45 to set $a_{11} = 1$. A quick calculation shows that $\#E(\mathbb{F}_2) = 5$. Thus, by definition, $a_2 = 2 + 1 - 5 = -2$. We already saw in Example 1.41 that $\#E(\mathbb{F}_3) = 5$ and thus $a_3 = 3 + 1 - 5 = -1$. Similar calculations show that $a_5 = 1$ and $a_7 = -2$. With this information, we compute

$$\begin{cases} a_4 & = & a_{2^2} = a_2^2 - 2 = 4 - 2 = 2, \\ a_6 & = & a_{2\times 3} = a_2.a_3 = 2, \\ a_8 & = & a_{2^3} = a_2^3 - 2a_2.2 = -8 + 8 = 0, \\ a_9 & = & a_{3^2} = a_3^2 - 3 = -2, \\ a_{10} & = & a_{2\times 5} = a_2.a_5 = -2. \end{cases}$$

Hence the Hasse-Weil L-series of E up till the eleventh term is:

$$L(E, s) = 1 - \frac{2}{2^s} - \frac{1}{3^s} + \frac{2}{4^s} + \frac{1}{5^s} + \frac{2}{6^s} - \frac{2}{7^s} - \frac{2}{9^s} - \frac{2}{10^s} + \frac{1}{11^s} + \cdots$$

2 Modular forms

In this chapter we radically change the subject from elliptic curves to modular forms. As we will see, modular forms are functions of the complex upper half plane that behave in a certain manner under the action of the matrix group $\mathrm{SL}_2(\mathbb{Z})$, which we call *the modular group*. Elliptic curves and modular forms emerge from two completely unrelated areas of mathematics but by the end of this chapter we shall have constructed a link between the two notions.

2.1 The modular group

A general theory of modular forms can be developed around $\mathrm{SL}_2(\mathbb{Z})$ and some of its particular subgroups called congruence subgroups. We shall not do this here since, as we will see, we will only work with the congruence subgroup $\Gamma_0(N)$. For a more general approach we refer to [DS05] Chapters 1 and 2.

2.1.1 Congruence subgroups

Let N be a positive integer and consider the group homomorphism

$$\begin{array}{rccc} \Pi_N : & \mathrm{SL}_2(\mathbb{Z}) & \longrightarrow & \mathrm{SL}_2(\mathbb{Z}/\mathrm{N}\mathbb{Z}) \\ & \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) & \longmapsto & \left(\begin{smallmatrix} [a]_N & [b]_N \\ [c]_N & [d]_N \end{smallmatrix}\right). \end{array}$$

Definition 2.1. The kernel of Π_N is called the *principal congruence subgroup of level* N and is denoted $\Gamma(N)$. Explicitly, that is

$$\Gamma(N) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}) : \begin{pmatrix} a & b \\ c & d \end{pmatrix} \equiv \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \pmod{N} \right\}.$$

Definition 2.2. A subgroup Γ of $\mathrm{SL}_2(\mathbb{Z})$ is said to be a *congruence subgroup* if there exists N a positive integer such that $\Gamma \supset \Gamma(N)$. The smallest such N is called the *level* of the congruence subgroup.

We may now introduce the particular congruence subgroup of level N mentioned above

$$\Gamma_0(N) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}) : c \equiv 0 \pmod{N} \right\}.$$

Notice that $\Gamma_0(1) = \mathrm{SL}_2(\mathbb{Z})$.

2.3. We verify that $\Gamma_0(N)$ is indeed a subgroup of $\mathrm{SL}_2(\mathbb{Z})$. Let $M = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma_0(N)$ and $M' = \left(\begin{smallmatrix} a' & b' \\ c' & d' \end{smallmatrix}\right) \in \Gamma_0(N)$. Clearly $\det(MM') = 1$ since $\det(M) = 1 = \det(M')$. Furthermore $(MM')_{21} = ca' + dc' \equiv 0 \pmod{N}$ since both c and c' are zero modulo N. So $MM' \in \Gamma_0(N)$. By definition, M is invertible and $M^{-1} = \left(\begin{smallmatrix} d & -b \\ -c & a \end{smallmatrix}\right)$. Thus $\det(M^{-1}) = 1$ and $(M^{-1})_{21} = -c \equiv 0 \pmod{N}$. So $M^{-1} \in \Gamma_0(N)$.

2.1.2 Results on indices

We state some results concerning the indices of $\Gamma_0(N)$ and $\Gamma(N)$ in $\mathrm{SL}_2(\mathbb{Z})$.

Proposition 2.4. *Let N be a positive integer. Then $\Gamma(N)$ is a normal subgroup of $\mathrm{SL}_2(\mathbb{Z})$ and the quotient group $\mathrm{SL}_2(\mathbb{Z})\,/\Gamma(N)$ is isomorphic to $\mathrm{SL}_2(\mathbb{Z}/\mathrm{N}\mathbb{Z})$.*

Proof. Since, by definition, $\Gamma(N)$ is the kernel of Π_N we know by basic algebra that it is a normal subgroup of $\mathrm{SL}_2(\mathbb{Z})$, and we only need to show that Π_N is surjective to prove the isomorphism.

Let $\gamma \in \mathrm{SL}_2(\mathbb{Z}/\mathrm{N}\mathbb{Z})$ and lift it to obtain a matrix $M = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in M_2(\mathbb{Z})$ such that $M \equiv \gamma \pmod{N}$. Then $\det(M) \equiv 1 \pmod{N}$. There exist matrices U, $V \in \mathrm{SL}_2(\mathbb{Z})$ such that UMV is diagonal. We write $UMV = \left(\begin{smallmatrix} a_1 & 0 \\ 0 & a_2 \end{smallmatrix}\right)$. Then $\det(UMV) = \det(U)\det(M)\det(V) = \det(M)$. Thus $a_1a_2 \equiv 1 \pmod{N}$. Define $W = \left(\begin{smallmatrix} a_2 & 1 \\ a_2-1 & 1 \end{smallmatrix}\right)$ and $X = \left(\begin{smallmatrix} 1 & -a2 \\ 0 & 1 \end{smallmatrix}\right)$. Note that $\det(W) = 1 = \det(X)$, thus both these matrices are in $\mathrm{SL}_2(\mathbb{Z})$. We have

$$WUMVX = \left(\begin{smallmatrix} a_2 & 1 \\ a_2-1 & 1 \end{smallmatrix}\right)\left(\begin{smallmatrix} a_1 & 0 \\ 0 & a_2 \end{smallmatrix}\right)\left(\begin{smallmatrix} 1 & -a2 \\ 0 & 1 \end{smallmatrix}\right) = \left(\begin{smallmatrix} a_1a_2 & a_2(1-a_1a_2) \\ a_1a_2-a_1 & a_2(1-a_1a_2)+a_1a_2 \end{smallmatrix}\right) \equiv \left(\begin{smallmatrix} 1 & 0 \\ 1-a_1 & 1 \end{smallmatrix}\right) \pmod{N},$$

since $a_1a_2 \equiv 1 \pmod{N}$. Write $A = \left(\begin{smallmatrix} 1 & 0 \\ 1-a_1 & 1 \end{smallmatrix}\right)$. The determinant of A is 1 and therefore $A \in \mathrm{SL}_2(\mathbb{Z})$. Finally set $M' = (WU)^{-1}A(VX)^{-1}$. Then $M' \equiv M \pmod{N} \equiv \gamma \pmod{N}$. Since $M' \in \mathrm{SL}_2(\mathbb{Z})$, we have proved that Π_N is surjective. □

2.5. A calculation that we do not expose here shows that

$$\#\,\mathrm{SL}_2(\mathbb{Z}/\mathrm{N}\mathbb{Z}) = N^3 \prod_{p|N} \left(1 - \frac{1}{p^2}\right).$$

Note that if $N = p$ is prime this is easy. Start by noticing that the sequence

$$1 \to \mathrm{SL_n}(\mathbb{F}_\mathrm{p}) \hookrightarrow \mathrm{GL_n}(\mathbb{F}_\mathrm{p}) \overset{\det}{\twoheadrightarrow} \mathbb{F}_p^* \to 1$$

is exact. The group $\mathrm{GL_n}(\mathbb{F}_\mathrm{p})$ is the set of all n by n matrices with coefficients in $\mathbb{F}_p$ for which the columns are linearly independent. For the first column, there are $p^n - 1$ possibilities (the all zero column is not valid). For the second column, there are $p^n - p$ possibilities since there are p multiples of the first column. Proceeding like this, we find that there are $p^n - p^{n-1}$ possibilities for the n^{th} column. Combining this, we have

$$\#\,\mathrm{GL_n}(\mathbb{F}_\mathrm{p}) = (p^n - 1)(p^n - p)\cdots(p^n - p^{n-1}) = \prod_{k=0}^{n-1}(p^n - p^k)$$

Thus $(p-1)\#\,\mathrm{SL_n}(\mathbb{F}_\mathrm{p}) = \#\,\mathrm{GL_n}(\mathbb{F}_\mathrm{p}) = \prod_{k=0}^{n-1}(p^n - p^k)$. Setting $n=2$, we get

$$\#\,\mathrm{SL}_2(\mathbb{Z}) = \frac{(p^2-1)(p^2-p)}{p-1} = p(p^2-1) = p^3(1-\frac{1}{p^2}).$$

Then one needs to show the result for p^n and that $\#\,\mathrm{SL}_2(\mathbb{Z}/n_1n_2\mathbb{Z}) = \#\,\mathrm{SL}_2(\mathbb{Z}/n_1\mathbb{Z})\#\,\mathrm{SL}_2(\mathbb{Z}/n_2\mathbb{Z})$ for n_1, n_2 relatively prime. Then the general result follows.

By Proposition 2.4, we know that $\mathrm{SL}_2(\mathbb{Z})\,/\Gamma(N) \xrightarrow{\sim} \mathrm{SL}_2(\mathbb{Z}/\mathrm{N}\mathbb{Z})$ and thus $[\mathrm{SL}_2(\mathbb{Z}) : \Gamma(N)] = \#\,\mathrm{SL}_2(\mathbb{Z}/\mathrm{N}\mathbb{Z})$.

Proposition 2.6. *Let N be a positive integer. The index of $\Gamma_0(N)$ in $\mathrm{SL}_2(\mathbb{Z})$ is*

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

Proof. In order to compute the index of $\Gamma_0(N)$ in $\mathrm{SL}_2(\mathbb{Z})$ we introduce the congruence subgroup

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

where $*$ means unspecified. We have the inclusions

$$\Gamma(N) \subset \Gamma_1(N) \subset \Gamma_0(N) \subset \mathrm{SL}_2(\mathbb{Z})\,.$$

Consider the group homomorphism $\Gamma_1(N) \longrightarrow \mathbb{Z}/N\mathbb{Z}$ given by $\left(\begin{smallmatrix} a & b \\ c & d\end{smallmatrix}\right) \longmapsto [b]_N$. This map is surjective. In fact, let $n \in \mathbb{Z}/N\mathbb{Z}$ and consider the matrix $\left(\begin{smallmatrix} 1 & n \\ 0 & 1\end{smallmatrix}\right)$. This matrix is an element of $\Gamma_1(N)$ and its image by the homomorphism is $[n]_N = n$. The kernel of this map is

$$\left\{\begin{pmatrix} a & b \\ c & d\end{pmatrix} \in \Gamma_1(N) : [b]_N = 0\right\} = \left\{\begin{pmatrix} a & b \\ c & d\end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}) : \begin{pmatrix} a & b \\ c & d\end{pmatrix} \equiv \begin{pmatrix} 1 & 0 \\ 0 & 1\end{pmatrix} \pmod{N}\right\} = \Gamma(N).$$

Hence, we have an isomorphism $\Gamma_1(N)/\Gamma(N) \xrightarrow{\sim} \mathbb{Z}/N\mathbb{Z}$ and thus $[\Gamma_1(N) : \Gamma(N)] = N$.

Consider the group homomorphism $\Gamma_0(N) \longrightarrow (\mathbb{Z}/N\mathbb{Z})^*$ given by $\left(\begin{smallmatrix} a & b \\ c & d\end{smallmatrix}\right) \longmapsto [d]_N$. This map is surjective. In fact, since $[d]_N \in (\mathbb{Z}/N\mathbb{Z})^*$, there exists an element $[d']_N$ of $(\mathbb{Z}/N\mathbb{Z})^*$ such that $[d]_N[d']_N = 1$. Consider the matrix $\left(\begin{smallmatrix} [d']_N & 0 \\ 0 & [d]_N\end{smallmatrix}\right)$. It is an element of $\Gamma_0(N)$ and its image by the map is $[d]_N$. The kernel of this map is

$$\left\{\begin{pmatrix} a & b \\ c & d\end{pmatrix} \in \Gamma_0(N) : [d]_N = 1\right\} = \left\{\begin{pmatrix} a & b \\ c & d\end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}) : \begin{pmatrix} a & b \\ c & d\end{pmatrix} \equiv \begin{pmatrix} 1 & * \\ 0 & 1\end{pmatrix} \pmod{N}\right\} = \Gamma_1(N).$$

So we have an isomorphism $\Gamma_0(N)/\Gamma_1(N) \xrightarrow{\sim} (\mathbb{Z}/N\mathbb{Z})^*$ and $[\Gamma_0(N) : \Gamma_1(N)] = \phi(N)$, where ϕ is the Euler totient function.

We know that $[\mathrm{SL}_2(\mathbb{Z}) : \Gamma(N)] = [\mathrm{SL}_2(\mathbb{Z}) : \Gamma_0(N)][\Gamma_0(N) : \Gamma_1(N)][\Gamma_1(N) : \Gamma(N)]$. Thus

$$[\mathrm{SL}_2(\mathbb{Z}) : \Gamma_0(N)] = \frac{N^3}{N\phi(N)}\prod_{p|N}\left(1-\frac{1}{p^2}\right).$$

Recall that the Euler totient function can be expressed as an Euler product:

$$\phi(N) = N \prod_{p|N} \left(1 - \frac{1}{p}\right).$$

Using this we finally get

$$[\mathrm{SL}_2(\mathbb{Z}) : \Gamma_0(N)] = \frac{N^3}{N^2} \prod_{p|N} \left(\frac{1 - \frac{1}{p^2}}{1 - \frac{1}{p}}\right) = N \prod_{p|N} \left(1 + \frac{1}{p}\right).$$

□

Proposition 2.7. *Let $N = p$ be a prime. Then the matrices*

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}, \ldots, \begin{pmatrix} 1 & 0 \\ N-1 & 1 \end{pmatrix}, \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

form a system of right coset representatives of $\Gamma_0(N)$ *in* $\mathrm{SL}_2(\mathbb{Z})$.

Proof. We prove that the union

$$\Gamma_0(N) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \cup \Gamma_0(N) \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \cup \ldots \cup \Gamma_0(N) \begin{pmatrix} 1 & 0 \\ N-1 & 1 \end{pmatrix} \cup \Gamma_0(N) \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

is disjoint. This union is contained in $\mathrm{SL}_2(\mathbb{Z})$ and by Proposition 2.6, $[\mathrm{SL}_2(\mathbb{Z}), \Gamma_0(N)] = N + 1$. Thus proving that the union is disjoint will end the proof.

Suppose there exists $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma_0(N)$ and $0 \leq k, k' < N$ such that

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 1 & 0 \\ k & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ k' & 1 \end{pmatrix}.$$

Then $b = 0, d = 1$ and $a = 1$. This implies that $c + k = k'$. But since $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma_0(N)$ we have $c \equiv 0 \pmod{N}$. Thus $k \equiv k' \pmod{N}$ and since $0 \leq k, k' < N$ this implies that $k = k'$.

Suppose there exists $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma_0(N)$ and $0 \leq k < N$ such that

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ k & 1 \end{pmatrix}.$$

Then $b = 1, a = 0, d = k$ and $c = -1$. But then $c \not\equiv 0 \pmod{N}$ and this contradicts the assumption that $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma_0(N)$.

This proves that the union is disjoint and hereby ends the proof. □

2.1.3 Generators of $\mathrm{SL}_2(\mathbb{Z})$

We state and prove an important proposition concerning the structure of $\mathrm{SL}_2(\mathbb{Z})$.

Proposition 2.8. *The group* $\mathrm{SL}_2(\mathbb{Z})$ *is generated by the matrices*

$$S = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \text{ and } T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

Proof. Let Γ be the subgroup of $\mathrm{SL}_2(\mathbb{Z})$ generated by S and T. Note that $T^n = \left(\begin{smallmatrix} 1 & n \\ 0 & 1 \end{smallmatrix}\right) \in \Gamma$ for all $n \in \mathbb{Z}$ and $S^2 = -I_2 \in \Gamma$. Let $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ be a matrix in $\mathrm{SL}_2(\mathbb{Z})$. We perform an induction on $|c|$. Note the identity

$$AT^n = \begin{pmatrix} a & b' \\ c & nc+d \end{pmatrix}, \tag{2.8.1}$$

where $b' = na + b$.

If $c = 0$, then $ad = 1$ which implies that $a = d = \pm 1$. If $a = d = 1$, then $A = \left(\begin{smallmatrix} 1 & b \\ 0 & 1 \end{smallmatrix}\right) = T^b \in \Gamma$. Otherwise, $A = S^2 \left(\begin{smallmatrix} 1 & -b \\ 0 & 1 \end{smallmatrix}\right) = S^2 T^{-b} \in \Gamma$.

Now, suppose $c \neq 0$ and that every $\left(\begin{smallmatrix} a' & b' \\ c' & d' \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{Z})$ with $|c'| < |c|$ belongs to Γ. Choose $N \in \mathbb{Z}$ such that $-1/2 - d/c \leq N \leq 1/2 - d/c$. Then, using Equation 2.8.1, we see that $AT^N = \left(\begin{smallmatrix} a & b' \\ c & d' \end{smallmatrix}\right)$ with $d' = Nc + d$ and $b' = Na + b$. Hence

$$-\frac{1}{2} - \frac{d}{c} \leq N \leq \frac{1}{2} - \frac{d}{c} \implies -\frac{c}{2} \leq d' \leq \frac{c}{2} \implies |d'| \leq \frac{|c|}{2} < |c|.$$

Note that $AT^N S = \left(\begin{smallmatrix} b' & -a \\ d' & -c \end{smallmatrix}\right)$. Since $|d'| < |c|$ we can apply the induction hypothesis and state that $AT^N S \in \Gamma$. Let $B \in \Gamma$ such that $AT^N S = B$. Then $A = BS^{-1}T^{-N} \in \Gamma$ and we have proved the proposition. □

2.2 Actions of the modular group

In this section we define and study several actions of the modular group. We will understand their purpose in the next section when we introduce modular forms. Note that we define the actions of $\mathrm{SL}_2(\mathbb{Z})$ and thus $\Gamma_0(N)$, being a subgroup of $\mathrm{SL}_2(\mathbb{Z})$, naturally inherits these actions by restriction.

2.2.1 Action on the Poincaré half plane

We define the complex upper half plane $\mathcal{H}$, also known as the Poincaré half plane, as follows

$$\mathcal{H} = \{z \in \mathbb{C} : \mathrm{Im}(z) > 0\}.$$

We now consider the following group action

$$\begin{array}{ccc} \mathrm{SL}_2(\mathbb{Z}) \times \mathcal{H} & \longrightarrow & \mathcal{H} \\ (\gamma, z) & \longmapsto & \gamma \cdot z = \frac{az+b}{cz+d}, \end{array}$$

where $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{Z})$.

2.9. The above map is indeed a group action. To see this we first show that $\gamma \cdot z$ is in fact in the upper half plane. Let $z \in \mathcal{H}$ and write $z = x + iy$, with x a real number and y a positive real number. Then

$$\gamma \cdot z = \frac{a(x+iy)+b}{c(x+iy)+d} = \frac{(ax+b)+iay}{(cx+d)+icy} = \frac{((ax+b)(cx+d)+acy^2)+iy(ad-cb)}{(cx+d)^2+(cy)^2}.$$

So the imaginary part of $\gamma \cdot z$ is $y/(cx+d)^2+(cy)^2$, since $\gamma \in \mathrm{SL}_2(\mathbb{Z})$ and therefore $ad - bc = 1$. The denominator is positive and we already know that $y > 0$. Thus $\gamma \cdot z \in \mathcal{H}$.

We now verify that it is a group action. It is easy to see that $I_2 \cdot z = z$ for all z in the upper half plane so we only need to show that $\gamma \cdot (\gamma' \cdot z) = (\gamma\gamma') \cdot z$ for all $\gamma, \gamma' \in \mathrm{SL}_2(\mathbb{Z})$ and $z \in \mathcal{H}$. We compute that

$$(\gamma\gamma') \cdot z = \left(\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix} \right) \cdot z = \begin{pmatrix} aa'+bc' & ab'+bd' \\ ca'+dc' & cb'+dd' \end{pmatrix} \cdot z = \frac{(aa'+bc')z+(ab'+bd')}{(ca'+dc')z+(cb'+dd')}.$$

On the other hand,

$$\begin{aligned} \gamma \cdot (\gamma' \cdot z) = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot \left(\begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix} \cdot z \right) &= \begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot \frac{a'z+b'}{c'z+d'} \\ &= \frac{a(\frac{a'z+b'}{c'z+d'})+b}{c(\frac{a'z+b'}{c'z+d'})+d} = \frac{(aa'+bc')z+(ab'+bd')}{(ca'+dc')z+(cb'+dd')}, \end{aligned}$$

as desired.

2.2.2 Weight k right action

Let k be an integer. We define the *weight k right action* of $\mathrm{SL}_2(\mathbb{Z})$ on the set of function $f : \mathcal{H} \longrightarrow \mathbb{C}$ as follows. For all $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{Z})$,

$$(f^{[\gamma]_k})(z) = (cz+d)^{-k} f(\gamma \cdot z).$$

2.10. The above map is indeed a group action. Note that $(f^{[I_2]_k})(z) = f(I_2 \cdot z) = f(z)$ for all $z \in \mathcal{H}$. We need to show that if γ and γ' are matrices in $\mathrm{SL}_2(\mathbb{Z})$, then $(f^{[\gamma\gamma']_k})(z) = ((f^{[\gamma]_k})^{[\gamma']_k})(z)$. We compute

$$\begin{aligned} (f^{[\gamma\gamma']_k})(z) &= ((ca'+dc')z+(cb'+dd'))^{-k} f((\gamma\gamma') \cdot z) \\ &= ((ca'+dc')z+(cb'+dd'))^{-k} f(\gamma \cdot (\gamma' \cdot z)). \end{aligned}$$

On the other hand,

$$\begin{aligned} ((f^{[\gamma]_k})^{[\gamma']_k})(z) &= (c'z+d')^{-k} f^{[\gamma]_k}(\gamma' \cdot z) \\ &= (c'z+d')^{-k}(c(\gamma' \cdot z)+d)^{-k} f(\gamma \cdot (\gamma' \cdot z)) \\ &= ((ca'+dc')z+(cb'+dd'))^{-k} f(\gamma \cdot (\gamma' \cdot z)) = (f^{[\gamma\gamma']_k})(z). \end{aligned}$$

In this paragraph we shall consider an action of $\mathrm{SL}_2(\mathbb{Z})$ on the set of holomorphic differential forms of degree 1. For the reader who is unfamiliar with differential forms, we refer to [Ca94] Chapters 1 and 2. What follows will be crucial when we introduce modular symbols later.

Consider the left action of $\mathrm{SL}_2(\mathbb{Z})$ on the set of holomorphic differential forms on $\mathcal{H}$ defined by

$$\gamma \cdot (f(z)dz) = f(\gamma \cdot z)d(\gamma \cdot z),$$

where $f : \mathcal{H} \to \mathbb{C}$ is holomorphic and $\gamma \in \mathrm{SL}_2(\mathbb{Z})$.

To see that this is indeed an action, note that $I_2 \cdot (f(z)dz) = fdz$ and let $\gamma, \gamma' \in \mathrm{SL}_2(\mathbb{Z})$. Then compute

$$\begin{aligned}\gamma' \cdot (\gamma \cdot (f(z)dz)) = \gamma' \cdot (f(\gamma \cdot z)d(\gamma \cdot z)) &= f(\gamma' \cdot (\gamma \cdot z))d(\gamma' \cdot (\gamma \cdot z)) \\ &= f((\gamma'\gamma) \cdot z)d((\gamma'\gamma) \cdot z) = (\gamma'\gamma) \cdot (f(z)dz).\end{aligned}$$

2.11. As we will see, modular forms satisfy a modularity condition which means that modular forms are *invariant* under the weight k right action. Our main object of study are modular forms of weight 2 for $\Gamma_0(N)$ and this motivates the next observation that will prove to be of great importance.

A function $f : \mathcal{H} \to \mathbb{C}$ is a modular form of weight 2 for a congruence subgroup Γ if and only if the differential form $f(z)dz$ is invariant under the left action of Γ.

Let f be a modular form of weight 2. The above observation only relies on the modularity condition that f must satisfy in order to be a modular form. This condition is exactly that $f^{[\gamma]_2} = f$ for all $\gamma \in \Gamma$. Let $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma$. We compute

$$d(\gamma \cdot z) = \frac{\partial}{\partial z}(\gamma \cdot z)dz = \frac{\partial}{\partial z}\left(\frac{az+b}{cz+d}\right)dz = \frac{ad-bc}{(cz+d)^2}dz = (cz+d)^{-2}dz.$$

Thus, $\gamma \cdot (f(z)dz) = f(\gamma \cdot z)d(\gamma \cdot z) = f(\gamma \cdot z)(cz+d)^{-2}dz = f^{[\gamma]_2}(z)dz = f(z)dz$. The last equality holds because of the modularity condition on f.

2.2.3 Action on $\mathbb{P}^1(\mathbb{Q})$

As we will see, modular forms are complex functions of the upper half plane $\mathcal{H}$ that satisfy a group action condition and a holomorphicity condition. In order to define this last condition we introduce the notion of cusps.

We consider $\mathbb{P}^1(\mathbb{Q})$ that we have already defined. We may identify elements $(s : t)$ of $\mathbb{P}^1(\mathbb{Q})$ with $\frac{s}{t}$ if $t \neq 0$ and with ∞ if $t = 0$. It is therefore appropriate to write

$$\mathbb{P}^1(\mathbb{Q}) = \mathbb{Q} \cup \{\infty\}.$$

The modular group acts on this set in the following way: let $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{Z})$ and $r \in \mathbb{P}^1(\mathbb{Q})$. Then

$$\gamma \cdot r = \begin{cases} \frac{ar+b}{cr+d} & \text{if } r \neq \infty, \\ \frac{a}{c} & \text{if } r = \infty. \end{cases}$$

Note that if the denominator c or $cr + d$ equal zero, then $\gamma \cdot r = \infty \in \mathbb{P}^1(\mathbb{Q})$.

Definition 2.12. Let Γ be a congruence subgroup of $\mathrm{SL}_2(\mathbb{Z})$. We define

$$C(\Gamma) = \Gamma \setminus \mathbb{P}^1(\mathbb{Q})$$

to be the *set of cusps of* Γ.

We prove that $C(\Gamma)$ is of finite cardinality for all congruence subgroups of $\mathrm{SL}_2(\mathbb{Z})$.

Proposition 2.13. *Let G be a group, $H \leq G$ a subgroup of G with finite index in G. Furthermore, suppose $X \neq \emptyset$ is a set for which we define a G-action and suppose G acts transitively on X. Then the cardinal of the quotient $H \setminus X$ satisfies*

$$\#\left(H \setminus X\right) \leq [G : H].$$

In particular, the quotient $H \setminus X$ is finite.

Proof. Let $x_0 \in X$. We define the map

$$\begin{aligned} \alpha_{x_0} : H \setminus G &\longrightarrow H \setminus X \\ Hg &\longmapsto H(g \cdot x_0). \end{aligned}$$

We show that this is a well defined map. Let $g, g' \in G$ such that $Hg = Hg'$. Then there exists $h' \in H$ such that $g' = h'g$ and thus

$$\begin{aligned} H(g' \cdot x_0) = \{h \cdot (g' \cdot x_0) \mid h \in H\} = \{h \cdot ((h'g) \cdot x_0) \mid h \in H\} = \{(hh') \cdot (g \cdot x_0) \mid h \in H\} \\ = \{h \cdot (g \cdot x_0) \mid h \in H\} = H(g \cdot x_0). \end{aligned}$$

Hence $\alpha_{x_0}(Hg) = \alpha_{x_0}(Hg')$.

The map α_{x_0} is surjective. In fact, consider $Hx \in H \setminus X$. Since G acts transitively on X, there exists $g \in G$ such that $x = g \cdot x_0$. Thus $Hx = H(g \cdot x_0) = \alpha_{x_0}(Hg)$. It follows from the surjectivity of α_{x_0} that $\#\left(H \setminus X\right) \leq [G : H]$ and since $[G : H]$ is finite by assumption, $H \setminus X$ is of finite cardinality. □

Lemma 2.14. *The group $\mathrm{SL}_2(\mathbb{Z})$ acts transitively on $\mathbb{P}^1(\mathbb{Q})$. In other words, for all α and β in $\mathbb{P}^1(\mathbb{Q})$ there exists $\gamma \in \mathrm{SL}_2(\mathbb{Z})$ such that $\gamma \cdot \alpha = \beta$.*

Proof. Let $\alpha \in \mathbb{P}^1(\mathbb{Q})$. Either $\alpha = \infty$ or $\alpha \in \mathbb{Q}$. Suppose $\alpha = p/q \in \mathbb{Q}$, with p, q relatively prime. Since $\gcd(p, q) = 1$, by Bézout's Identity, there exist integers r, s such that $rp + sq = 1$. Consider the matrix $\gamma = \left(\begin{smallmatrix} p & -s \\ q & r \end{smallmatrix}\right)$. It is an element of $\mathrm{SL}_2(\mathbb{Z})$ since $\det(\gamma) = rp + sq = 1$, and $\gamma \cdot \infty = p/q = \alpha$.

Let α, β be two arbitrary elements of $\mathbb{P}^1(\mathbb{Q})$. Then there exist $\gamma, \delta \in \mathrm{SL}_2(\mathbb{Z})$ such that $\alpha = \gamma \cdot \infty$ and $\beta = \delta \cdot \infty$. Then $\beta = (\delta\gamma^{-1}) \cdot \alpha$ and $\delta\gamma^{-1} \in \mathrm{SL}_2(\mathbb{Z})$ since $\mathrm{SL}_2(\mathbb{Z})$ is a group. □

Corollary 2.15. *Let Γ be a congruence subgroup of $\mathrm{SL}_2(\mathbb{Z})$. Then the set of cusps $C(\Gamma)$ is finite.*

Proof. By the previous lemma, the group $\mathrm{SL}_2(\mathbb{Z})$ acts transitively on $\mathbb{P}^1(\mathbb{Q})$ and by Remark 2.5, Γ has finite index in $\mathrm{SL}_2(\mathbb{Z})$. Thus, applying Proposition 2.13 with $G = \mathrm{SL}_2(\mathbb{Z})$, $H = \Gamma$ and $X = \mathbb{P}^1(\mathbb{Q})$ yields $\#\left(\Gamma \setminus \mathbb{P}^1(\mathbb{Q})\right) \leq [\mathrm{SL}_2(\mathbb{Z}) : \Gamma]$ as desired. □

2.16. Without proof we give a formula for the exact number of cusps of $\Gamma_0(N)$:

$$\#C(\Gamma_0(N)) = \sum_{d|N} \phi(\gcd(d, \frac{N}{d})),$$

where ϕ is the Euler totient function. This enables us to see that $\Gamma_0(11)$ has exactly 2 cusps and that $\Gamma_0(39)$ for example has 4 cusps.

2.3 Fundamental domains

We describe a fundamental domain for the action of $\mathrm{SL}_2(\mathbb{Z})$ on $\mathcal{H}$. Fundamental domains are most important if we want to consider quotients $\Gamma \setminus \mathcal{H}$, where Γ is a congruence subgroup of $\mathrm{SL}_2(\mathbb{Z})$.

Definition 2.17. A *fundamental domain* for the action of a subgroup Γ of $\mathrm{SL}_2(\mathbb{Z})$ on $\mathcal{H}$ is an open subset D of $\mathcal{H}$ such that

(i) If z_1 and z_2 are two equivalent points (under the action of Γ) contained in D, then $z_1 = z_2$.

(ii) The closure $\overline{D}$ of D contains at least one point of each Γ-orbit.

2.18. From this definition we can observe:

(i) The map $\overline{D} \longrightarrow \mathcal{H}/\,\mathrm{SL}_2(\mathbb{Z})$ is surjective and its restriction to D is injective.

(ii) When γ runs through Γ, the sets $\{\gamma\overline{D}\}$ cover $\mathcal{H}$.

Lemma 2.19. *Let $z \in \mathcal{H}$ and fix $a, b \in \mathbb{Z}$. Then there is only a finite number of integers c, d such that $\mathrm{Im}(\gamma \cdot z) \geq \mathrm{Im}(z)$, where $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{Z})$.*

Proof. Let $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{Z})$ and $z' = \gamma \cdot z = (az+b)/(cz+d)$ such that $\mathrm{Im}(z') \geq \mathrm{Im}(z)$. A calculation performed in Section 2.2.1 shows that

$$\mathrm{Im}(z') = \frac{\mathrm{Im}(z)}{(c\,\mathrm{Re}(z)+d)^2 + (c\,\mathrm{Im}(z))^2} = \frac{\mathrm{Im}(z)}{|cz+d|^2}.$$

The condition $\mathrm{Im}(z') \geq \mathrm{Im}(z)$ translates as $|cz+d|^2 \leq 1$. But $(c\,\mathrm{Im}(z))^2 \leq |cz+d|^2 \leq 1$ which implies $|c| \leq 1/\,\mathrm{Im}(z)$ and thus there is only a finite number of $c \in \mathbb{Z}$. Then the condition $(c\,\mathrm{Re}(z) + d)^2 + (c\,\mathrm{Im}(z))^2 \leq 1$ gives a finite number of d. □

Theorem 2.20. *The subset D of $\mathcal{H}$ defined by*

$$D = \left\{ z \in \mathcal{H} : |z| > 1,\ |\,\mathrm{Re}(z)| < \frac{1}{2} \right\}$$

is a fundamental domain for the action of $\mathrm{SL}_2(\mathbb{Z})$ on $\mathcal{H}$. Furthermore, two distinct elements of $\overline{D}$ can only be equivalent if they both lie on the boundary of D. We shall refer to this set as the "standard" fundamental domain.

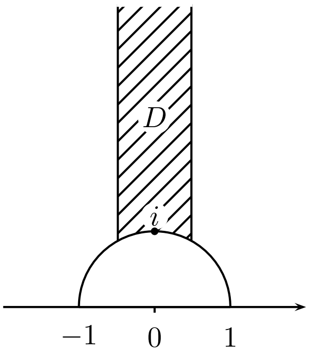

Figure 1: The standard fundamental domain

Proof. Define D_1 to be the following subset of $\mathcal{H}$:

$$D_1 = \left\{ z \in \mathcal{H} : |\,\mathrm{Re}(z)| < \frac{1}{2},\ |cz+d| > 1 \text{ if } (c,d) \neq (0,0) \text{ and } (c,d) \neq (0,1) \right\}.$$

The strategy of the proof is the following: (1) $D = D_1$; (2) $\overline{D}$ contains at least one element of each orbit; (3) two distinct elements of $\overline{D}$ can only be equivalent if they both lie on the boundary of D.

(1) Clearly we have $D_1 \subset D$. In fact, one only needs to set $(c,d) = (1,0)$. For the other inclusion, let $z = x+iy \in D$, where $x, y \in \mathbb{R}$ with $|x| < 1/2$, $x^2+y^2 > 1$. Then

$$|cz+d|^2 = (cx+d)^2 + (cy)^2 = c^2(x^2+y^2) + 2cdx + d^2 > c^2 - |cd| + d^2 \geq 1.$$

Thus $D \subset D_1$ and we have proved (1).

(2) Consider $S = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$. By Lemma 2.19, there exists $z \in \mathcal{H}$ such that

$$\mathrm{Im}(z) = \max\left\{ \mathrm{Im}\left(\left(\begin{smallmatrix} 0 & -1 \\ c & d \end{smallmatrix}\right) \cdot z\right) \right\},$$

where the maximum is taken over all $(c,d) \in \mathbb{Z}^2$ such that $\left(\begin{smallmatrix} 0 & -1 \\ c & d \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{Z})$. We may translate z into $\{z \in \mathcal{H} : |\operatorname{Re}(z)| \leq 1/2\}$. Then we must have $|z| \geq 1$. In fact, suppose that $|z| < 1$. Then $\operatorname{Im}(S \cdot z) = \operatorname{Im}(z)/|z|^2 > \operatorname{Im}(z)$ and this contradicts the maximality of $\operatorname{Im}(z)$. Thus $z \in \overline{D}$ and $\overline{D}$ contains at least one element of each orbit.

(3) Let $z, z' \in \overline{D}$ such that $z' = \frac{az+b}{cz+d}$. Since $D = D_1$, we have that $z, z' \in \overline{D_1}$. Thus

$$\operatorname{Im}(z') = \frac{\operatorname{Im}(z)}{|cz+d|^2} \leq \operatorname{Im}(z)$$

because $|cz + d| \geq 1$. But we also have

$$\operatorname{Im}(z) = \frac{\operatorname{Im}(z')}{|-cz'+a|^2} \leq \operatorname{Im}(z')$$

because $|-cz' + a| \geq 1$. So $\operatorname{Im}(z) = \operatorname{Im}(z')$. Thus

$$1 = |cz+d|^2 \geq c^2 - |cd| + d^2 \geq 1 \implies (c,d) = (0, \pm 1) \text{ or } (c,d) = (\pm 1, 0).$$

This implies that $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) = \pm \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)^n$ or $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) = \pm \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$.

In the first case, if $n \neq 0$, $z' = \pm(z + n)$. But then $n = 1$ since the width of D is 1. Thus z, z' lie on the vertical boundaries of D.

In the second case, $z' = S \cdot z = -1/z$ and we know that $|z| \geq 1$ and $|z'| \geq 1$. But

$$|z'| = \frac{1}{|z|} \implies |z'| \leq 1 \implies |z'| = |z| = 1,$$

hence z, z' lie on the unit circle.

Otherwise, $z' = T^0 \cdot z = z$. □

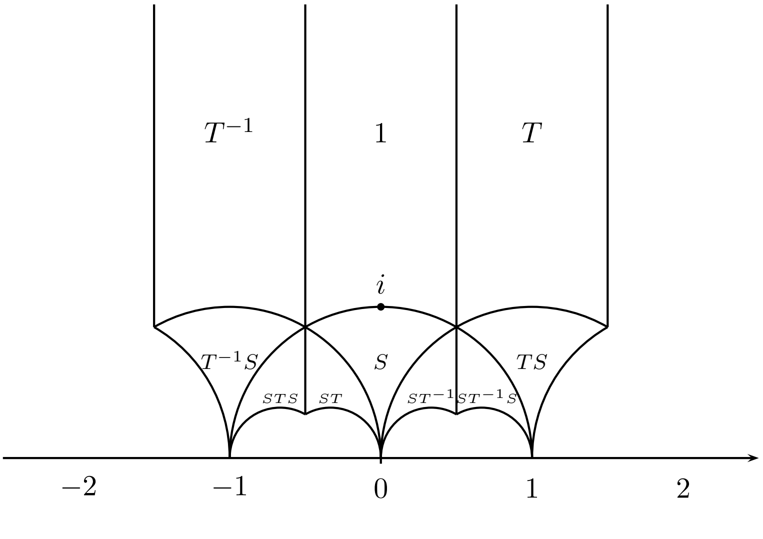

Figure 2: Transformations of the standard fundamental domain

Corollary 2.21. *Let Γ be a subgroup of finite index n in $\mathrm{SL}_2(\mathbb{Z})$ and let $\Gamma g_1, \cdots, \Gamma g_n$ be a system of left classes of $\mathrm{SL}_2(\mathbb{Z})$ modulo Γ such that $\mathrm{SL}_2(\mathbb{Z}) = \Gamma g_1 \cup \cdots \cup \Gamma g_n$. Then $\Delta := g_1 D \cup \cdots \cup g_n D$ is a fundamental domain for Γ.*

Proof. Basic topological properties show that Δ is open. Let $z \in \mathcal{H}$, then there exists $\sigma \in \mathrm{SL}_2(\mathbb{Z})$ such that $\sigma \cdot z \in \overline{D}$. Since $\sigma \in \mathrm{SL}_2(\mathbb{Z})$, there exists $i \in \{1, \cdots, n\}$ such that $\sigma \in \Gamma g_i$. Thus there exists $\gamma \in \Gamma$ such that $\sigma = \gamma g_i$. Then $\sigma \cdot z = \gamma \cdot (g_i \cdot z) \subset \gamma(g_i \overline{D}) \subset \gamma(\overline{\Delta})$. So the closure of Δ contains at least one point of each Γ-orbit.

Let z, z' be two Γ-equivalent points of Δ. Then there exist $i, j \in \{1, \cdots, n\}$ such that $z \in g_i D$ and $z' \in g_j D$. Then $g_i^{-1} \cdot z \in D$ and $g_j^{-1} \cdot z' \in D$. Clearly $g_i^{-1} \cdot z$ and $g_j^{-1} \cdot z'$ are Γ-equivalent and since D is a fundamental domain for $\mathrm{SL}_2(\mathbb{Z})$ we must have $g_i^{-1} \cdot z = g_j^{-1} \cdot z'$. Hence $i = j$ and $z = z'$. So Δ contains at most one point of each Γ-orbit. Thus Δ is a fundamental domain for Γ. □

2.22. Let Γ be a congruence subgroup of $\mathrm{SL}_2(\mathbb{Z})$. Using the previous corollary we can choose representatives for the classes of $\mathrm{SL}_2(\mathbb{Z})$ modulo Γ in a certain way and visualize the cusps of Γ as being the intersection points of $\overline{\Delta}$ with the real axis. Since we have chosen the cusps to be equivalence classes, this determination does not depend on the choice of the representatives. This illustrates the result of Corollary 2.15 on the cardinality of $C(\Gamma)$.

Example 2.23. We use the previous remark to visualize the cusps of $\Gamma_0(11)$.

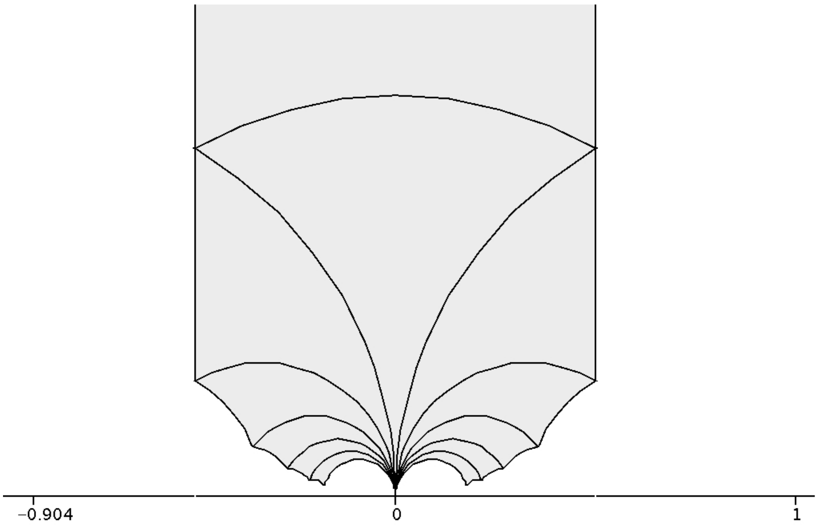

Figure 3: A fundamental domain for $\Gamma_0(11)$[1]

Using Figure 3 we see that $\Gamma_0(11)$ only has two cusps, namely the classes of 0 and ∞.

2.4 Modular forms

We are now almost ready to define modular forms for congruence subgroups of $\mathrm{SL}_2(\mathbb{Z})$. We start by defining weakly modular functions and then, by using the actions introduced in Section 2.2, we

[1] Drawn with *Fundamental Domain Drawer* program developed by H.A. Verrill. http://www.math.lsu.edu/~verrill/fundomain/index2.html accessed on 02/12/13.

define modular forms. Again our main concern is the definition of modular forms for $\Gamma_0(N)$ but in this section we construct a more general theory around congruence subgroups and thus we view $\Gamma_0(N)$ as a particular case to which the theory applies.

2.4.1 Definition

Definition 2.24. Let k be an integer and Γ be a congruence subgroup of level N. A *weakly modular function* of weight k and level N for Γ is a meromorphic function $f : \mathcal{H} \longrightarrow \mathbb{C}$ that is Γ-invariant under the weight k right action. In other words,

$$f^{[\gamma]_k} = f$$

for all $\gamma \in \Gamma$.

Lemma 2.25. *If f is a weakly modular function of weight k for a congruence subgroup Γ and if $\delta \in \mathrm{SL}_2(\mathbb{Z})$, then $f^{[\delta]_k}$ is a weakly modular function of weight k for the congruence subgroup $\delta^{-1}\Gamma\delta$.*

Proof. Note that $f^{[\delta]_k}$ is meromorphic and has the same poles as f. Let $\gamma = \delta^{-1}\alpha\delta \in \delta^{-1}\Gamma\delta$. Then

$$(f^{[\delta]_k})^{[\gamma]_k} = f^{[\delta\gamma]_k} = f^{[\delta\delta^{-1}\alpha\delta]_k} = f^{[\alpha\delta]_k} = f^{[\delta]_k},$$

since $\alpha \in \Gamma$. □

We define the extended upper half plane $\mathcal{H}^*$ to be

$$\mathcal{H}^* = \mathcal{H} \cup \mathbb{P}^1(\mathbb{Q}).$$

Modular forms of weight k for a congruence subgroup Γ are weakly modular functions that are holomorphic on the extended upper half plane. In order to clarify the difference between the different functions satisfying modularity conditions we introduce a table summarizing modularity of a function $f : \mathcal{H} \to \mathbb{C}$ for Γ and weight k.

Weakly modular function	Modular function	Modular form	Cusp form
$f^{[\gamma]_k} = f$ and f meromorphic on $\mathcal{H}$	$f^{[\gamma]_k} = f$ and f meromorphic on $\mathcal{H}$ and at cusps	$f^{[\gamma]_k} = f$ and f holomorphic on $\mathcal{H}$ and at cusps	$f^{[\gamma]_k} = f$ and f holomorphic on $\mathcal{H}$ and at cusps and f is zero at all cusps

Table 2: Summary (modularity for Γ, weight k)

In order to properly define modular forms we need to clarify the meaning of this holomorphicity condition.

Fix a weakly modular function f of weight k for a congruence subgroup Γ of level N. The principal congruence subgroup of level N, $\Gamma(N)$, contains the matrix $T^N = \left(\begin{smallmatrix} 1 & N \\ 0 & 1 \end{smallmatrix}\right)$. Since Γ is a congruence subgroup of level N it also contains T^N and thus there exists a positive integer h such that $f(z+h) = f(z)$ for all z in $\mathcal{H}$. Note that $h = N$ always works but there might exist a smaller such h. Then f is $h\mathbb{Z}$-periodic for some minimal $h > 0$. Let D be the open unit disc and let $D' := D \setminus \{0\}$ be the punctured open unit disc. We consider the holomorphic map $q_h : \mathcal{H} \longrightarrow D'$ defined by $q_h(z) = e^{\frac{2i\pi z}{h}}$. We define the function $g_h : D' \longrightarrow \mathbb{C}$ such that $g_h(q_h) = f(\log(q)/\frac{2i\pi}{h})$. This function is well defined and $f(z) = g_h(q_h(z))$. If f is holomorphic on $\mathcal{H}$, then g_h is holomorphic on the punctured disc D' by composition of holomorphic functions. Thus g_h has a Laurent expansion in a neighborhood of 0:

$$g_h(q_h) = \sum_{n=-\infty}^{+\infty} a_n q_h^n.$$

We say that f is *meromorphic at* ∞ if g_h is meromorphic at 0, that is, if $a_n = 0$ for all sufficiently small n. Then there exists $m \in \mathbb{Z}$ such that $g_h(q) = \sum_{n=m}^{\infty} a_n q^n$. If $m \geq 0$, then g_h extends holomorphically to the puncture point 0 and we say that f is *holomorphic at* ∞. In this case, f has a Fourier expansion:

$$f(z) = \sum_{n=0}^{\infty} a_n q_h^n(z) = \sum_{n=0}^{\infty} a_n e^{\frac{2i\pi n z}{h}}.$$

But we also have to deal with the rest of the cusps. Fix $\alpha \in \mathbb{P}^1(\mathbb{Q})$. It follows from Lemma 2.14 that there exists $\delta \in \mathrm{SL}_2(\mathbb{Z})$ such that $\alpha = \delta \cdot \infty$. By Lemma 2.25, $f^{[\delta]_k}$ is a weakly modular function of weight k for the congruence subgroup $\delta^{-1}\Gamma\delta$. We define f to be holomorphic at α if $f^{[\delta]_k}$ is holomorphic at ∞. We say that f is holomorphic on the extended upper half plane $\mathcal{H}^*$ if f is holomorphic on $\mathcal{H}$ and at all $\alpha \in \mathbb{P}^1(\mathbb{Q})$.

2.26. If f is holomorphic at $\alpha \in \mathbb{P}^1(\mathbb{Q})$, then f is holomorphic at $\Gamma\alpha =: \overline{\alpha} \in C(\Gamma)$. In fact, let $\gamma \in \Gamma$ and consider $\gamma \cdot \alpha$. Then $\gamma \cdot \alpha = (\gamma\delta) \cdot \infty$ and $f^{[\gamma\delta]_k} = f^{[\delta]_k}$ is a weakly modular function of weight k for the congruence subgroup $(\gamma\delta)^{-1}\Gamma(\gamma\delta) = \delta^{-1}\Gamma\delta$. Thus, in order for f to be holomorphic on $\mathcal{H}^*$, f only needs to be holomorphic at all $\overline{\alpha} \in C(\Gamma)$. By Corollary 2.15, the number of cusps is finite so f only needs to be holomorphic at a finite number of $\alpha \in \mathbb{P}^1(\mathbb{Q})$.

Definition 2.27. Let Γ be a congruence subgroup of level N and let k be an integer. A function $f : \mathcal{H} \longrightarrow \mathbb{C}$ is said to be a *modular form (respectively a modular function) of weight k and level N for Γ* if

(i) The function f is a weakly modular function of weight k for Γ.

(ii) The function f is holomorphic (respectively meromorphic) on $\mathcal{H}^*$.

We denote by $\mathcal{M}_k(\Gamma)$ the set of modular forms of weight k for Γ.

2.28. Note that $T = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right) \in \Gamma_0(N)$. Thus every modular form in $\Gamma_0(N)$ is $\mathbb{Z}$-periodic and

$$f(z) = \sum_{n=0}^{\infty} a_n e^{2i\pi nz}.$$

Proposition 2.29. *Let Γ be a congruence subgroup of level N. The product of two modular forms of weights k and l for Γ is a modular form of weight $k+l$ for Γ.*

Proof. Let $f \in \mathcal{M}_k(\Gamma)$ and $g \in \mathcal{M}_l(\Gamma)$. It follows from basic complex analysis that fg is holomorphic on $\mathcal{H}$ since both f and g are holomorphic on $\mathcal{H}$ by definition. We show that fg is Γ-invariant under the weight $k+l$ right action. Let $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma$. Then, for all $z \in \mathcal{H}$,

$$(fg)^{[\gamma]_{k+l}}(z) = (cz+d)^{-(k+l)}(fg)(\gamma \cdot z) = f^{[\gamma]_k}(z)g^{[\gamma]_l}(z) = (fg)(z),$$

since f and g are modular forms of weight k and l respectively.

Finally, we check that fg is holomorphic on $\mathcal{H}^*$. Let $\alpha \in \mathbb{P}^1(\mathbb{Q})$. There exists $\delta \in \mathrm{SL}_2(\mathbb{Z})$ such that $\delta \cdot \infty = \alpha$. Since both f and g are holomorphic at α, we know that $f^{[\delta]_k}$ and $g^{[\delta]_l}$ are holomorphic at ∞. Thus there exist sequences $\{a_n\}, \{b_n\}$ such that

$$(fg)^{[\delta]_{k+l}}(z) = (f^{[\delta]_k}g^{[\delta]_l})(z) = \left(\sum_{n=0}^{\infty} a_n e^{\frac{2i\pi nz}{h}}\right)\left(\sum_{n=0}^{\infty} b_n e^{\frac{2i\pi nz}{h}}\right) = \sum_{n=0}^{\infty}\left(\sum_{i+j=n} a_i b_j\right) e^{\frac{2i\pi nz}{h}}.$$

Thus fg is holomorphic at α. Since α is arbitrary, we have shown that fg is holomorphic on $\mathcal{H}^*$ and this ends the proof. □

2.30. It can easily be checked that $\mathcal{M}_k(\Gamma)$ is a $\mathbb{C}$-vector space. It follows from Proposition 2.29 that $\mathcal{M}_k(\Gamma)\mathcal{M}_l(\Gamma) \subset \mathcal{M}_{k+l}(\Gamma)$. Thus

$$\mathcal{M}(\Gamma) := \bigoplus_{k \in \mathbb{Z}} \mathcal{M}_k(\Gamma)$$

is a graded $\mathbb{C}$-algebra.

Proposition 2.31. *Let k be an integer. If N' divides N, then every modular form for $\Gamma_0(N')$ is a modular form of the same weight for $\Gamma_0(N)$. In other words, we have the inclusion*

$$\mathcal{M}_k(\Gamma_0(N')) \hookrightarrow \mathcal{M}_k(\Gamma_0(N)).$$

The image of this inclusion is called the set of "Oldforms".

Proof. Let n be an integer. Since $n \equiv 0 \pmod{N}$ implies $n \equiv 0 \pmod{N'}$, we easily see that $\Gamma_0(N) \subset \Gamma_0(N')$. Let $f \in \mathcal{M}_k(\Gamma_0(N'))$. We need to show that f is holomorphic on $C(\Gamma_0(N))$ and that f is invariant under the weight k right action restricted to $\Gamma_0(N)$.

Let $[\alpha]_{\Gamma_0(N)} \in C(\Gamma_0(N))$. Then

$$[\alpha]_{\Gamma_0(N)} = \{\gamma \cdot \alpha | \gamma \in \Gamma_0(N)\} \subset \{\gamma \cdot \alpha | \gamma \in \Gamma_0(N')\} = [\alpha]_{\Gamma_0(N')},$$

since $\Gamma_0(N) \subset \Gamma_0(N')$. Thus $C(\Gamma_0(N)) \subset C(\Gamma_0(N'))$ and it is then clear that f is holomorphic on $C(\Gamma_0(N))$.

The fact that f is invariant under the weight k right action restricted to $\Gamma_0(N)$ follows immediately from the fact that $\Gamma_0(N) \subset \Gamma_0(N')$. □

2.4.2 Examples of modular forms of level 1

We are later interested in modular forms of weight 2 for $\Gamma_0(N)$. In this section we give examples of modular forms for $\mathrm{SL}_2(\mathbb{Z})$ in order to get in touch with a concrete modular form. We will not need this later. We start by making a few remarks on modular forms of level 1.

First, recall that by Proposition 2.8, $\mathrm{SL}_2(\mathbb{Z})$ is generated by the matrices S and T. Hence, in order to show that a meromorphic function $f : \mathcal{H} \longrightarrow \mathbb{C}$ is weakly modular of integer weight k it suffices to show that

$$f(z+1) = f(z) \text{ and } f(-\frac{1}{z}) = z^k f(z).$$

Next, recall that by Lemma 2.14, $\mathrm{SL}_2(\mathbb{Z})$ only has one cusp and we may choose ∞ as a representative. Thus, if $f \in \mathcal{M}(\mathrm{SL}_2(\mathbb{Z}))$, then f is holomorphic at ∞ and therefore has a Fourier expansion

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

where $q : z \mapsto e^{2i\pi z}$. Since $q(z) \to 0$ when $\mathrm{Im}(z) \to \infty$, we see that a weakly modular holomorphic function f is holomorphic at ∞ if the limit of $f(z)$ exists when $\mathrm{Im}(z) \to \infty$.

We start with some trivial examples. The all-zero function is a modular form of level 1 of any integer weight k. Note that the weight of a non-zero form is unique. Constant functions are modular forms of level 1 and weight 0.

2.32. Note that this last observation can be extended. In fact, it is true that for any N, *a function f is a modular form of level N and weight 0 if and only if f is a constant function.*

2.33. Let f be a weakly modular function of weight k for $\mathrm{SL}_2(\mathbb{Z})$. Letting $\gamma = -I_2 \in \mathrm{SL}_2(\mathbb{Z})$ yields the following equation

$$f(z) = (-1)^k f(z), \; \forall z \in \mathcal{H}.$$

We see that if k is odd, then the only such function is the all-zero function.

To see non-trivial examples of modular forms we introduce Eisenstein series.

Definition 2.34. Let $k > 2$ be an integer. The weight k Eisenstein series $G_k : \mathcal{H} \longrightarrow \mathbb{C}$ is defined by

$$G_k(z) = \sum_{(c,d)\in\mathbb{Z}^2}{}' \frac{1}{(cz+d)^k},$$

for $z \in \mathcal{H}$. Here, $'$ means that the sum is over $\mathbb{Z}^2 \setminus \{(0,0)\}$.

Notice that if $\tau \in \mathcal{H}$ is fixed, then the sum is taken over all non-zero elements of the lattice $\Lambda = \Lambda(1,\tau) = \mathbb{Z} + \mathbb{Z}\tau$. To see that the Eisenstein series defined above is indeed well-defined, we need to show that the series is absolutely convergent and this is a consequence of this next lemma as we will see.

Lemma 2.35. *Let α be a real number. Then the series*

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

converges absolutely if and only if $\alpha > 2$.

Proof. Consider the parallelogram P centered at the origin with vertices $-1-\tau, 1-\tau, 1+\tau$ and $-1+\tau$. Let r and R respectively denote the minimal and the maximal distances from the origin to P.

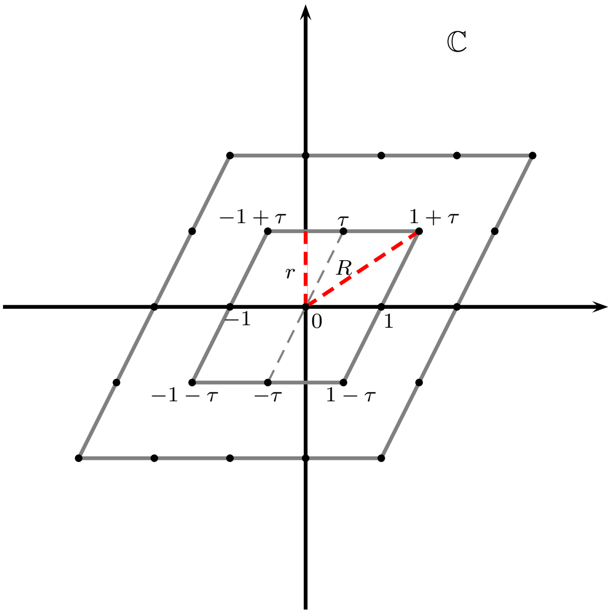

Figure 4: The first two layers of the lattice Λ

Consider the first layer of Λ. If ω is any of the 8 non-zero lattice points of this layer (see Figure 4), then it holds that $r \leq |\omega| \leq R$ by definition of r and R.

Consider the second layer of Λ. There are 16 new lattice points in this layer (see Figure 4). If ω is any of these 16 points, then the inequality $2r \leq |\omega| \leq 2R$ holds.

In general, in the k^{th} layer of Λ, there are $8k$ new lattice points and for each of these it holds that $kr \leq |\omega| \leq kR$. Therefore, we have

$$\frac{1}{(kR)^\alpha} \leq \frac{1}{|\omega|^\alpha} \leq \frac{1}{(kr)^\alpha}$$

for all non-zero lattice points ω in the first k layers of Λ.

Let $S(n)$ denote the sum over the $8\sum_{k=1}^n k$ distinct non-zero lattice points ω in the first n layers of Λ of the terms $\frac{1}{|\omega|^\alpha}$. In other words, $S(n) = \sum \frac{1}{|c\tau+d|^\alpha}$ where the sum is taken over $c, d = -n, \ldots, -1, 0, 1, \ldots, n$ with $(c, d) \neq (0, 0)$. We then have

$$\sum_{k=1}^{n} \frac{8k}{(kR)^\alpha} = \frac{8}{R^\alpha} \sum_{k=1}^{n} \frac{1}{k^{\alpha-1}} \leq S(n) \leq \sum_{k=1}^{n} \frac{8k}{(kr)^\alpha} = \frac{8}{r^\alpha} \sum_{k=1}^{n} \frac{1}{k^{\alpha-1}}.$$

Recall that the Riemann zeta-function ζ converges for complex numbers with real part greater than 1 and thus $S(n)$ is bounded from above by $\frac{8}{r^\alpha}\zeta(\alpha-1)$ if $\alpha > 2$. This bound is independent of n and therefore we may write

$$\lim_{n\to\infty} S(n) \leq \frac{8}{r^\alpha}\zeta(\alpha-1)$$

and this proves that the series $\sum'_{\omega\in\Lambda} \frac{1}{\omega^\alpha}$ converges absolutely if $\alpha > 2$.

On the other hand, if $\alpha \leq 2$, then $\zeta(\alpha-1)$ does not converge. Combined with the inequality

$$S(n) \geq \frac{8}{R^\alpha}\zeta(\alpha-1)$$

this shows that the series $\sum'_{\omega\in\Lambda} \frac{1}{\omega^\alpha}$ does not converge if $\alpha \leq 2$. □

Proposition 2.36. *Let $k > 1$ be an integer. The Eisenstein series G_{2k} is a non-constant modular form of weight $2k$ and level 1, that is, $G_{2k} \in \mathcal{M}_{2k}(\mathrm{SL}_2(\mathbb{Z}))$. Furthermore, we have*

$$G_{2k}(\infty) = 2\zeta(2k),$$

where ζ is the Riemann zeta-function.

Proof. We start by showing that G_{2k} is well defined and holomorphic on $\mathcal{H}$ and at ∞. Let $z \in \overline{D}$ where $D = \{z \in \mathcal{H} : |z| > 1 \text{ and } |\operatorname{Re}(z)| < 1/2\}$ is the standard fundamental domain of $\mathrm{SL}_2(\mathbb{Z})$ described in Theorem 2.20. Let $(c, d) \in \mathbb{Z}^2 \setminus \{(0, 0)\}$. Then

$$|cz + d|^2 = (c\operatorname{Re}(z) + d)^2 + (c\operatorname{Im}(z))^2 = c^2 z\overline{z} + 2cd\operatorname{Re}(z) + d^2 \geq c^2 - cd + d^2 = |c\rho - d|^2,$$

where $\rho = e^{\frac{i\pi}{3}}$. Since $2k > 2$, we may apply Lemma 2.35 to see that the series $\sum'_{(c,d)\in\mathbb{Z}^2} \frac{1}{|c\rho-d|^{2k}}$ converges for $k > 1$. Hence, G_{2k} converges normally on $\overline{D}$. This implies that G_{2k} also converges uniformly on $\overline{D}$ and thus G_{2k} is holomorphic on $\overline{D}$ since it is the uniform limit of a sequence of holomorphic functions. This is also true for all $\gamma\overline{D}$ with $\gamma \in \mathrm{SL}_2(\mathbb{Z})$. In fact, it suffices to apply the previous result to $G_{2k}(\gamma^{-1} \cdot z)$. But Remark 2.18 (*ii*) states that $\mathcal{H}$ is covered by these subsets, thus G_{2k} is well defined and holomorphic on $\mathcal{H}$.

We show that G_{2k} is holomorphic at ∞. It is sufficient to show that the limit of $G_{2k}(z)$ exists when $\mathrm{Im}(z) \to \infty$ and we may suppose that z stays in $\overline{D}$. Since the convergence is uniform on $\overline{D}$ we can take the limit term by term. The terms $1/(cz+d)^{2k}$ for which $c \neq 0$, vanish. Those for which $c = 0$, become $1/d^{2k}$. Thus

$$\lim_{\mathrm{Im}(z)\to\infty} G_{2k}(z) = \sum_{d\in\mathbb{Z}}{}' \frac{1}{d^{2k}} = 2\sum_{d=1}^{\infty} \frac{1}{d^{2k}} = 2\zeta(2k).$$

It remains to show that G_{2k} is modular. By the remarks above, we only need to show this for S and T. Let $z \in \mathcal{H}$. Since G_{2k} converges absolutely for $k > 1$ it is permitted to change the order of summation of the terms and this will not affect the convergence of the series. Thus, we may compute that

$$G_{2k}(z+1) = \sum_{(c,d)\in\mathbb{Z}^2}{}' \frac{1}{(cz+(c+d))^{2k}} = \sum_{(c,d)\in\mathbb{Z}^2}{}' \frac{1}{(cz+d)^{2k}} = G_{2k}(z).$$

The second equality holds because the map on $\mathbb{Z}^2$ defined by $(c, c+d) \mapsto (c, d)$ is invertible. Thus as $(c, c+d)$ runs through $\mathbb{Z}^2 \setminus \{(0,0)\}$, so does (c, d).

$$G_{2k}(-\frac{1}{z}) = \sum_{(c,d)\in\mathbb{Z}^2}{}' \frac{1}{(-c\frac{1}{z}+d)^{2k}} = z^{2k} \sum_{(c,d)\in\mathbb{Z}^2}{}' \frac{1}{(dz-c)^{2k}} = z^{2k} \sum_{(c,d)\in\mathbb{Z}^2}{}' \frac{1}{(cz+d)^{2k}} = z^{2k} G_{2k}(z).$$

The third equality holds because the map on $\mathbb{Z}^2$ defined by $(d, -c) \mapsto (c, d)$ is invertible. This ends the proof. □

2.37. It can be proved (see [Se70] Chapter VII Section 3.2) that G_4 and G_6 are algebraically independent over $\mathbb{C}$ and

$$\{G_4^a G_6^b : a, b \in \mathbb{N} \text{ and } 4a + 6b = k\}$$

is a basis for the $\mathcal{C}$-vector space $\mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z}))$. This is equivalent to saying that $\epsilon : \mathbb{C}[X, Y] \to \mathcal{M}(\mathrm{SL}_2(\mathbb{Z}))$ is an isomorphism, where ϵ assigns G_4 to X and G_6 to Y. Thus, we can identify $\mathcal{M}(\mathrm{SL}_2(\mathbb{Z}))$ with $\mathbb{C}[G_4, G_6]$. That means that every element of the graded $\mathbb{C}$-algebra $\mathcal{M}(\mathrm{SL}_2(\mathbb{Z}))$ can be expressed as a polynomial in G_4 and G_6 with complex coefficients.

For example, every modular form of weight 30 can be expressed as

$$aG_4^6 G_6 + bG_4^3 G_6^3 + cG_6^3,$$

for some unique $a, b, c \in \mathbb{C}$. It follows that the dimension of $\mathcal{M}_{30}(\mathrm{SL}_2(\mathbb{Z}))$ is 3. More generally the dimension of $\mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z}))$ is

$$\dim(\mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z}))) = \#\{(a,b) \in \mathbb{N}^2 : 4a + 6b = k\}.$$

Hence we see that if $k = 0, 4, 6, 8, 10$, then $\mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z}))$ has dimension 1. Also, the dimension of $\mathcal{M}_2(\mathrm{SL}_2(\mathbb{Z}))$ is 0 and since our object of study is modular forms of weight 2 for $\Gamma_0(N)$ we see that the case $N = 1$ is not relevant from our point of view.

We proceed to computing the Taylor series of the Eisenstein series G_{2k} with respect to $q = q(z) = e^{2\pi iz}$. We shall make the convention for the next part that

$$\sum_{n\in\mathbb{Z}} = \lim_{N\to\infty} \sum_{n=-N}^{N}.$$

This gives sense to the series we are about to consider, since it does not converge absolutely.

Proposition 2.38. *Let $z \in \mathbb{C} \setminus \mathbb{Z}$. We have the following formula*

$$\pi \cot(\pi z) = \sum_{n\in\mathbb{Z}} \frac{1}{z+n}.$$

Proof. The function $x \mapsto \cos(zx)$ is even and thus its Fourier series only contains terms in cosine:

$$\cos(zx) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos(nx) \quad \text{with} \quad a_n = \frac{2}{\pi}\int_0^{\pi} \cos(zx)\cos(nx)dx.$$

We compute that

$$a_0 = \frac{2}{\pi}\int_0^{\pi} \cos(zx)dx = \frac{2\sin(\pi z)}{\pi z}$$

and for n positive,

$$\begin{aligned}
a_n &= \frac{2}{\pi}\int_0^{\pi} \cos(zx)\cos(nx)dx \\
&= \frac{1}{2\pi}\int_0^{\pi} e^{i(z+n)x} + e^{-i(z+n)x} + e^{i(z-n)x} + e^{-i(z-n)x}dx \\
&= \frac{1}{\pi}\int_0^{\pi} \cos((z+n)x) + \cos((z-n)x)dx \\
&= \frac{1}{\pi}\left(\frac{\sin((z+n)x)}{z+n} + \frac{\sin((z-n)x)}{z-n}\right) \\
&= \frac{(-1)^n 2z\sin(\pi z)}{\pi(z^2-n^2)}.
\end{aligned}$$

Thus, for $x \in [-\pi; \pi]$,

$$\cos(zx) = \frac{2z\sin(\pi z)}{\pi}\left(\frac{1}{2z^2} + \sum_{n=1}^{\infty} \frac{(-1)^n\cos(nx)}{z^2-n^2}\right).$$

Setting $x = \pi$ yields

$$\frac{\cos(\pi z)}{\sin(\pi z)} = \cot(\pi z) = \frac{1}{\pi}\left(\frac{1}{z} + \sum_{n=1}^{\infty} \frac{2z}{z^2-n^2}\right),$$

and thus

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

□

Notice that

$$\pi\cot(\pi z) = \pi\frac{\cos(\pi z)}{\sin(\pi z)} = i\pi\frac{e^{i\pi z}+e^{-i\pi z}}{e^{i\pi z}-e^{-i\pi z}} = i\pi\frac{q(z)+1}{q(z)-1} = i\pi - \frac{2i\pi}{1-q(z)} = i\pi - 2i\pi\sum_{n=1}^{\infty} q(z)^n.$$

Combining this with Proposition 2.38 yields the equality

$$\sum_{n\in\mathbb{Z}} \frac{1}{z+n} = i\pi - 2i\pi\sum_{n=1}^{\infty} q(z)^n$$

and differentiating $k-1$ times ($k \geq 2$), we see that

$$\sum_{n\in\mathbb{Z}} \frac{1}{(z+n)^k} = \frac{(-2i\pi)^k}{(k-1)!}\sum_{n=1}^{\infty} n^{k-1}q(z)^n. \tag{2.38.1}$$

We introduce the following notation: $\sigma_m(n) = \sum_{d|n} d^m$.

Proposition 2.39. *Let $k \geq 2$. The Taylor series of the Eisenstein series G_{2k} with respect to q is*

$$G_{2k} = 2\zeta(2k) + 2\frac{(2i\pi)^{2k}}{(2k-1)!}\sum_{n=1}^{\infty}\sigma_{2k-1}(n)q^n.$$

Proof. We already know from Proposition 2.36 that $G_{2k}(\infty) = 2\zeta(2k)$ and thus the constant term of the Taylor series with respect to q is $2\zeta(2k)$. We compute

$$G_{2k}(z) = \sum_{(c,d)\in\mathbb{Z}^2}{}' \frac{1}{(cz+d)^{2k}} = 2\zeta(2k) + 2\sum_{c=1}^{\infty}\sum_{d\in\mathbb{Z}}\frac{1}{(cz+d)^{2k}}.$$

Using Equation 2.38.1 with z replaced by cz, we see that

$$\sum_{d\in\mathbb{Z}}\frac{1}{(cz+d)^{2k}} = \frac{(-2i\pi)^{2k}}{(2k-1)!}\sum_{d=1}^{\infty} d^{2k-1}q(z)^{dc}.$$

Hence

$$G_{2k}(z) = 2\zeta(2k) + 2\frac{(2i\pi)^{2k}}{(2k-1)!}\sum_{c=1}^{\infty}\sum_{d=1}^{\infty} d^{2k-1}q(z)^{dc}.$$

We perform the change of variables $n = dc$. Then c divides n and

$$\sum_{c=1}^{\infty}\sum_{d=1}^{\infty} d^{2k-1}q^{dc} = \sum_{n=1}^{\infty}\sum_{c|n}\left(\frac{n}{c}\right)^{2k-1}q^n = \sum_{n=1}^{\infty}\sum_{d|n} d^{2k-1}q^n = \sum_{n=1}^{\infty}\sigma_{2k-1}(n)q^n.$$

We then have

$$G_{2k}(z) = 2\zeta(2k) + 2\frac{(2i\pi)^{2k}}{(2k-1)!}\sum_{n=1}^{\infty}\sigma_{2k-1}(n)q(z)^n.$$

□

We want to give an explicit formula for $\zeta(2k)$. In order to find this formula we consider the Bernoulli numbers introduced by Jakob Bernoulli in his 1713 *Ars Conjectandi* (posthum). These are denoted B_k and defined by the equality

$$\frac{x}{e^x-1}=\sum_{k=0}^{\infty}B_k\frac{x^k}{k!}. \tag{2.39.1}$$

We give some of the first values of the Bernoulli numbers in the following table:

k	0	1	2	3	4	5	6	7	8	9	10
B_k	1	$-\frac{1}{2}$	$\frac{1}{6}$	0	$-\frac{1}{30}$	0	$\frac{1}{42}$	0	$-\frac{1}{30}$	0	$\frac{5}{66}$

Table 3: First values of the Bernoulli numbers

It can be shown that $B_n=0$ for all odd $n>1$.

Proposition 2.40. *If $k\geq 1$ is integer, then*

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

Proof. We start by noticing that

$$z\cot(z)=1+\sum_{k=2}^{\infty}B_k\frac{(2iz)^k}{k!}. \tag{2.40.1}$$

In fact, we compute

$$z\cot(z)=z\frac{\cos(z)}{\sin(z)}=iz\frac{e^{iz}+e^{-iz}}{e^{iz}-e^{-iz}}=iz+\frac{2iz}{e^{2iz}-1}.$$

Using Equation 2.39.1, with x is replaced by $2iz$, yields

$$z\cot(z)=iz+\sum_{k=0}^{\infty}B_k\frac{(2iz)^k}{k!}.$$

Using Table 3, we see that

$$z\cot(z)=1+\sum_{k=2}^{\infty}B_k\frac{(2iz)^k}{k!}.$$

Since all Bernoulli numbers of odd index greater than 1 is zero we may write this as

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

since $i^{2k}=(-1)^k$.

Recall that the sine function can be written as an Euler product in the following way

$$\sin(z)=z\prod_{n=1}^{\infty}\left(1-\frac{z^2}{n^2\pi^2}\right).$$

We take the logarithm of both sides of this equation and differentiate. We see that

$$\frac{d}{dz}\log(\sin(z)) = \cot(z)$$

and

$$\frac{d}{dz}\log\left(z\prod_{n=1}^{\infty}\left(1-\frac{z^2}{n^2\pi^2}\right)\right) = \frac{d}{dz}\left(\log(z)+\sum_{n=1}^{\infty}\log\left(1-\frac{z^2}{n^2\pi^2}\right)\right) = \frac{1}{z}+2\sum_{n=1}^{\infty}\frac{z}{z^2-n^2\pi^2}.$$

Combining these facts yields

$$z\cot(z) = 1+2\sum_{n=1}^{\infty}\frac{z^2}{z^2-n^2\pi^2} = 1-2\sum_{n=1}^{\infty}\frac{z^2}{n^2\pi^2-z^2}.$$

Notice that

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

Using this, we find that

$$z\cot(z) = 1-2\sum_{n=1}^{\infty}\sum_{k=1}^{\infty}\frac{z^{2k}}{n^{2k}\pi^{2k}} = 1-2\sum_{k=1}^{\infty}\zeta(2k)\frac{z^{2k}}{\pi^{2k}}.$$

Finally, if we compare this last equation with Equation 2.40.1, we get

$$\sum_{k=1}^{\infty}2\zeta(2k)\frac{z^{2k}}{\pi^{2k}} = \sum_{k=1}^{\infty}(-1)^{k+1}B_{2k}\frac{2^{2k}z^{2k}}{(2k)!}$$

and thus

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

□

In the following table we give some values of $\zeta(2k)$:

k	1	2	3	4	5
$\zeta(2k)$	$\frac{\pi^2}{6}$	$\frac{\pi^4}{90}$	$\frac{\pi^6}{945}$	$\frac{\pi^8}{9450}$	$\frac{\pi^{10}}{93555}$

Table 4: Some values of the Riemann zeta function

Definition 2.41. Let $k > 2$. We define the normalized Eisenstein series E_k as

$$E_k = \frac{(k-1)!}{2(2i\pi)^k}G_k.$$

Using this definition wee see that the Taylor series of the normalized Eisenstein series E_{2k}, for $k > 1$ is

$$E_{2k} = -\frac{B_{2k}}{4}+\sum_{n=1}^{\infty}\sigma_{2k-1}(n)q^n.$$

We do the computations for E_4 and E_6:

$$E_4 = \frac{1}{240} + q + 9q^2 + 28q^3 + 73q^4 + 126q^5 + 252q^6 + \cdots$$

and

$$E_6 = -\frac{1}{504} + q + 33q^2 + 244q^3 + 1057q^4 + 3126q^5 + 8052q^6 + \cdots$$

2.4.3 Cusp forms

Definition 2.42. Let Γ be a congruence subgroup of level N and let k be an integer. A function $f : \mathcal{H} \longrightarrow \mathbb{C}$ is said to be a *cusp form of weight k and level N for* Γ if it is zero at all cusps.
We denote by $\mathcal{S}_k(\Gamma)$ the set of cusp forms of weight k for Γ. It is a $\mathbb{C}$-linear subspace of $\mathcal{M}_k(\Gamma)$.

Example 2.43. Consider the modular forms G_4 and G_6. We set $g_2 = 60G_4$ and $g_3 = 140G_6$. By Proposition 2.36,

$$g_2(\infty) = 120\zeta(4) = \frac{4}{3}\pi^4 \quad \text{and} \quad g_3(\infty) = 280\zeta(6) = \frac{8}{27}\pi^6.$$

Consider the modular form $\Delta = g_2^3 - 27g_3^2$. We compute that

$$\Delta(\infty) = \frac{64}{27}\pi^{12} - \frac{64}{27}\pi^{12} = 0.$$

Thus Δ is a cusp form of weight 12 for $\mathrm{SL}_2(\mathbb{Z})$.

2.5 Hecke operators on modular forms

In this section we introduce Hecke operators. We start by defining these on the space $\mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z}))$ and then we extend this definition to $\mathcal{M}_k(\Gamma_0(N))$. We will see the action of Hecke operators on the set $\mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z}))$ and we will describe it via the Fourier series of the modular forms. We follow an approach similar to the ones in both [Se70] Chapter VII Section 5 and [He02] Chapter 5 Section 5.7. A more geometrical approach can be found in [St07] Chapter 3 Section 3.1 but we will leave this aside.

2.5.1 Hecke operators for $\mathrm{SL}_2(\mathbb{Z})$

Let n be a positive integer. Consider the set of matrices

$$\Delta_n = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \,\middle|\, a, b, c, d \in \mathbb{Z},\ ad - bc = n \right\}.$$

The group $\mathrm{SL}_2(\mathbb{Z})$ acts on Δ_n by left multiplication:

$$\begin{array}{rcl} \mathrm{SL}_2(\mathbb{Z}) \times \Delta_n & \longrightarrow & \Delta_n \\ (\gamma, \alpha) & \longmapsto & \gamma\alpha. \end{array}$$

We decompose Δ_n in orbits of this action and choose a system of representatives R for $\mathrm{SL}_2(\mathbb{Z}) \backslash \Delta_n$. The finiteness of this system is a consequence of this next proposition.

Proposition 2.44. *For any positive integer n, we have the disjoint union*

$$\Delta_n = \bigcup \left\{ \mathrm{SL}_2(\mathbb{Z}) \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} \middle| \, a > 0, ad = n, 0 \leq b < d \right\}.$$

Proof. Let $\alpha = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Delta_n$. We proceed to finding $\gamma = \left(\begin{smallmatrix} x & y \\ z & w \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{Z})$ such that $\gamma\alpha = \left(\begin{smallmatrix} a' & b' \\ 0 & d' \end{smallmatrix}\right)$. We need to have $za + wc = 0$. We set $z = c/(a,c)$ and $w = -a/(a,c)$. To simplify notations we shall write $(\cdot,\cdot)$ instead of $\gcd(\cdot,\cdot)$. Then the condition is satisfied and $(z,w) = 1$. By Bézout's Identity, there exists integers x and y such that $xw - yz = 1$ and thus $\gamma = \left(\begin{smallmatrix} x & y \\ z & w \end{smallmatrix}\right)$ is in $\mathrm{SL}_2(\mathbb{Z})$. Note that $a'd' = \det(\gamma\alpha) = \det(\gamma)\det(\alpha) = n$. Without loss of generality, we may assume that $a', d' > 0$ (otherwise just take $-\gamma$). Finally, multiplying by $\left(\begin{smallmatrix} 1 & h \\ 0 & 1 \end{smallmatrix}\right)$ on the left with a suitable h gives the condition $0 \leq b' < d'$.

To see that the union is disjoint, suppose there exists $\gamma = \left(\begin{smallmatrix} x & y \\ z & w \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{Z})$ and positive integers a, d, a', d', $0 \leq b < d$, $0 \leq b' < d'$ such that

$$\begin{pmatrix} x & y \\ z & w \end{pmatrix} \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} = \begin{pmatrix} a' & b' \\ 0 & d' \end{pmatrix}.$$

Since the set of all invertible upper triangular matrices form a group, we must have $z = 0$ and thus $xw = 1$. Then $ax = a'$ and $dw = d'$. Since a, a', d, d' are all positive by assumption, we must have $x = w = 1$. Then $a = a'$ and $d = d'$. Finally, $b + yd = b'$ and thus $b \equiv b' \pmod{d}$. But since $0 \leq b, b' < d$, we conclude that $b = b'$. This shows that the union is disjoint. □

Recall that we defined the weight k right action of $\mathrm{SL}_2(\mathbb{Z})$ on the set of functions $f : \mathcal{H} \to \mathbb{C}$. One can extend this action, for k even, to $\mathrm{GL}_2(\mathbb{Z})$ in the following way. For all $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{GL}_2(\mathbb{Z})$,

$$(f^{[\gamma]_k})(z) = \det(\gamma)^{\frac{k}{2}}(cz+d)^{-k} f(\gamma \cdot z).$$

Let k be an even integer and let $f : \mathcal{H} \to \mathbb{C}$ be a function which is $\mathrm{SL}_2(\mathbb{Z})$-invariant under the weight k right action. Then the sum $\sum f^{[\mu]_k}$ taken over all $\mu \in R$ is independent of the choice of the representatives. In fact, if $\gamma \in \mathrm{SL}_2(\mathbb{Z})$ and $\mu \in R$, then

$$f^{[\gamma\mu]_k} = (f^{[\gamma]_k})^{[\mu]_k} = f^{[\mu]}.$$

Definition 2.45. Let n be a positive integer. The n^{th} Hecke operator for $\mathrm{SL}_2(\mathbb{Z})$ is a linear map from the set of functions $f : \mathcal{H} \to \mathbb{C}$ to the same set. It is denoted T_n and defined as

$$T_n f = n^{\frac{k}{2}-1} \sum_{\mu \in R} f^{[\mu]_k},$$

where R is a system of representatives for $\mathrm{SL}_2(\mathbb{Z}) \backslash \Delta_n$.

Using Proposition 2.44, we may express the n^{th} Hecke operator in a more explicit way:

$$T_n f(z) = n^{k-1} \sum_{\substack{0 \leq b < d \\ a \geq 1, ad=n}} d^{-k} f\left(\frac{az+b}{d}\right). \tag{2.45.1}$$

Proposition 2.46. *Let n be a positive integer. If f is a weakly modular function of weight k, then so is $T_n f$.*

Proof. We start by verifying the modularity condition for $T_n f$. If $\gamma \in \mathrm{SL}_2(\mathbb{Z})$, then

$$(T_n f)^{[\gamma]_k} = n^{\frac{k}{2}-1} \sum_{\mu \in R} (f^{[\mu]_k})^{[\gamma]_k} = n^{\frac{k}{2}-1} \sum_{\mu \in R} (f^{[\gamma]_k})^{[\mu]_k} = n^{\frac{k}{2}-1} \sum_{\mu \in R} (f^{[\mu]_k})^{[\gamma]_k} = T_n f,$$

where the third equality follows from the fact that f is $\mathrm{SL}_2(\mathbb{Z})$-invariant under the weight k right action.

Equation 2.45.1 shows that $T_n f$ is meromorphic on $\mathcal{H}$ if f is meromorphic on $\mathcal{H}$. In fact, a finite sum of meromorphic functions is meromorphic. The same argument applies if f is holomorphic and thus if f is holomorphic on $\mathcal{H}$, then so is $T_n f$. □

In the rest of this section we shall write $c(m)$ instead of c_m for the Fourier coefficients of f. This is only done to make the proofs easier to read.

Proposition 2.47. *Let n be a positive integer and let $f = \sum_{m \in \mathbb{Z}} c(m) q^m$ be a modular function of weight k. Then $T_n f$ is also a modular function of weight k. We have*

$$T_n f = \sum_{m \in \mathbb{Z}} \gamma(m) q^m \quad \textit{with} \quad \gamma(m) = \sum_{\substack{a \mid (n,m) \\ a \geq 1}} a^{k-1} c\left(\frac{mn}{a^2}\right).$$

Proof. Using Equation 2.45.1, we see that

$$T_n f(z) = n^{k-1} \sum_{\substack{0 \leq b < d \\ a \geq 1, ad=n}} d^{-k} \sum_{m \in \mathbb{Z}} c(m) e^{2\pi i \frac{az+b}{d} m}.$$

Suppose $d|m$, and set $m' = m/d$. Then

$$\sum_{b=0}^{d-1} e^{2\pi i \frac{mb}{d}} = \sum_{b=0}^{d-1} e^{2\pi i m' b} = \sum_{b=0}^{d-1} 1 = d.$$

Otherwise,

$$\sum_{b=0}^{d-1} e^{2\pi i \frac{mb}{d}} = \frac{1-(e^{2\pi i \frac{m}{d}})^d}{1-e^{2\pi i \frac{m}{d}}} = \frac{1-e^{2\pi i m}}{1-e^{2\pi i \frac{m}{d}}} = 0.$$

Thus

$$T_n f(z) = n^{k-1} \sum_{\substack{a\geq 1\\ ad=n}} d^{-k+1} \sum_{m'\in\mathbb{Z}} c(m'd)e^{2\pi i z a m'}.$$

Write $\mu = am'$. Then

$$T_n f = \sum_{\mu\in\mathbb{Z}} q^{\mu} \sum_{\substack{a\geq 1\\ a|(n,\mu)}} \left(\frac{n}{d}\right)^{k-1} c\left(\frac{\mu d}{a}\right).$$

Since f is meromorphic at ∞, there exists a positive constant N such that $c(m) = 0$, $\forall m \leq -N$. Thus $c(\mu d/a) = 0$, $\forall \mu \leq -nN$. From this, we see that $T_n f$ is meromorphic at ∞ and this shows that $T_n f$ is a modular function of weight k. To finish the proof, remember that $ad = n$ and therefore $a = n/d$ and $d/a = n/a^2$. Thus

$$T_n f = \sum_{\mu\in\mathbb{Z}} \left(\sum_{\substack{a\geq 1\\ a|(n,\mu)}} a^{k-1} c\left(\frac{\mu n}{a^2}\right) \right) q^{\mu}.$$

□

Corollary 2.48. *Let n and m be positive integers. Then*

$$T_n T_m = \sum_{d|(n,m)} d^{k-1} T_{\frac{nm}{d^2}} = T_m T_n.$$

In particular, $T_n T_m = T_{nm}$ if n and m are relatively prime.

Proof. Let $f = \sum_{m\in\mathbb{Z}} c(m)q^m$. Using the previous proposition, we compute that

$$\sum_{d|(n,m)} d^{k-1} T_{\frac{nm}{d^2}} f = \sum_{d|(n,m)} d^{k-1} \sum_{\alpha\in\mathbb{Z}} \left(\sum_{a|(\frac{nm}{d^2},\alpha)} a^{k-1} c\left(\frac{mn\alpha}{d^2a^2}\right) \right) q^{\alpha}.$$

Similarly, we compute that

$$T_n f = \sum_{\alpha\in\mathbb{Z}} \left(\sum_{a|(n,\alpha)} a^{k-1} c\left(\frac{n\alpha}{a^2}\right) \right) q^{\alpha}$$

and

$$\begin{aligned} T_m T_n f &= \sum_{\alpha\in\mathbb{Z}} \left(\sum_{d|(m,\alpha)} d^{k-1} \sum_{a|(n,\frac{m\alpha}{d^2})} a^{k-1} c\left(\frac{mn\alpha}{d^2a^2}\right) \right) q^{\alpha} \\ &= \sum_{d|(n,m)} d^{k-1} \sum_{\alpha\in\mathbb{Z}} \left(\sum_{a|(\frac{nm}{d^2},\alpha)} a^{k-1} c\left(\frac{mn\alpha}{d^2a^2}\right) \right) q^{\alpha} \end{aligned}$$

as desired. □

Let $f = \sum_{m=0}^{\infty} a(m)q^m \in \mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z}))$. Consider now the particular case when $n = p$ is prime. Then

$$T_p f = \sum_{m=0}^{\infty} a(pm)q^m + p^{k-1} \sum_{m=0}^{\infty} a(m)q^{pm}.$$

Furthermore consider two particular cases of Corollary 2.48:

(i) If p divides n once, then $\gcd(\frac{n}{p}, p) = 1$ and $T_n = T_{\frac{n}{p}} T_p$.

(ii) If p^2 divides n, then $T_n = T_{\frac{n}{p}} T_p - p^{k-1} T_{\frac{n}{p^2}}$.

Corollary 2.49. *If f is a modular form (resp. a cusp form), then $T_n f$ is also a modular form (resp. a cusp form).*

Corollary 2.50. *Consider the same notations as in Proposition 2.47. Then $\gamma(0) = \sigma_{k-1}(n)c(0)$ and $\gamma(1) = c(n)$.*

Corollary 2.51. *If $n = p$ is prime, then*

$$\begin{cases} \gamma(m) = c(pm) & \text{if } p \nmid m, \\ \gamma(m) = c(pm) + p^{k-1} c\left(\frac{m}{p}\right) & \text{if } p | m. \end{cases}$$

Example 2.52. We compute the first terms of $T_2 E_4$. Recall that we already computed the Taylor series of E_4 with respect to q in the end of Section 2.4.2. We let γ_m denote the coefficients of the series of $T_2 E_4$. Using the results of Corollary 2.50, we see that $\gamma_0 = 3/80$ and $\gamma_1 = 9$. Using Corollary 2.51, we compute γ_m for $m = 2, \cdots, 6$. We can then write

$$T_2 E_4 = \frac{3}{80} + 9q + 81q^2 + 252q^3 + 657q^4 + 1134q^5 + 2268q^6 + \cdots$$

2.5.2 Hecke forms for $\mathrm{SL}_2(\mathbb{Z})$

Let $f = \sum_{m=0}^{\infty} a_m q^m \in \mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z}))$, not equal to the all-zero function, and suppose f is an eigenfunction for all Hecke operators T_n. In other words, there exists a sequence $\{\lambda_n\}_{n=1}^{\infty} \subset \mathbb{C}$ such that $T_n f = \lambda_n f$ for all positive integer n.

Theorem 2.53. *The coefficient a_1 of f is non-zero. Furthermore, if f is normalized such that $a_1 = 1$, then $a_n = \lambda_n$ for all positive integer n.*

Proof. Using Corollary 2.50, we see that the q coefficient of T_nf is a_n. But since $T_nf = \lambda_n f$, this coefficient is also $\lambda_n a_1$. Thus $a_n = \lambda_n a_1$ for all n. So if $a_1 = 0$, then $a_n = 0$ for all n. Hence, f is the all-zero function and this contradicts our assumption. Thus $a_1 \neq 0$. If $a_1 = 1$, then $a_n = \lambda_n$ for all n. □

Definition 2.54. Let $f \in \mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z}))$. We say that f is a *Hecke form* (for $\mathrm{SL}_2(\mathbb{Z})$) if f is normalized in the sense that $a_1 = 1$ and if f is an eigenfunction for all Hecke operators T_n.

Corollary 2.55. *Two Hecke forms of weight k which have the same eigenvalues coincide.*

Corollary 2.56. *Let $f = \sum_{m=0}^{\infty} a_m q^m$ be a Hecke form of weight k. Then*

$$\begin{cases} a_m a_n = a_{mn} & \text{if } \gcd(m,n) = 1, \\ a_p a_{p^n} = a_{p^{n+1}} + p^{k-1} a_{p^{n-1}} & \text{if } p \text{ prime, } n \geq 1. \end{cases}$$

Proof. Using Theorem 2.53 and Corollary 2.48, we get

$$T_n T_m f(z) = T_n a_m f(z) = a_m a_n f(z) = \sum_{d|(m,n)} d^{k-1} T_{\frac{mn}{d^2}} f(z) = \sum_{d|(m,n)} d^{k-1} a_{\frac{mn}{d^2}} f(z).$$

Thus, if m and n are relatively prime, $a_m a_n = a_{mn}$. From the above computation we see that

$$a_p a_{p^n} = \sum_{d|(m,n)} d^{k-1} a_{\frac{p^{n+1}}{d^2}} = a_{p^{n+1}} + p^{k-1} a_{p^{n-1}}.$$

□

2.57. Observe that the function $n \mapsto a_n$ is multiplicative. That is, if m, n are relatively prime, then $a_{mn} = a_m a_n$.

Theorem 2.58 (Hecke)**.** *Let $k \geq 4$ be even and $f = \sum_{m=0}^{\infty} a_m q^m \in \mathcal{S}_k(\mathrm{SL}_2(\mathbb{Z}))$. Then*

$$a_n = \mathcal{O}(n^{\frac{k}{2}}) \quad \text{when} \quad n \to +\infty.$$

In other words, there exists a positive real constant C such that $|a_n| \leq C n^{\frac{k}{2}}$ for n sufficiently large.

Proof. Since f is a cusp form, $a_0 = 0$. We can then factorize f by q and write

$$f = \left(\sum_{m=1}^{\infty} a_m q^{m-1} \right) q.$$

If we write $z = x + iy$, then $|f(z)| = \mathcal{O}(|q(z)|) = \mathcal{O}(e^{-2\pi y})$ when $q(z) \to 0$, that is, when $y \to +\infty$.

Let $\phi(z) = |f(z)|y^{\frac{k}{2}}$. Then ϕ is invariant under the action of $\mathrm{SL}_2(\mathbb{Z})$. In fact, let $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{Z})$ and recall that $\mathrm{Im}(\gamma \cdot z) = y/|cz+d|^2$. Thus

$$\phi(\gamma \cdot z) = |f(\gamma \cdot z)| \, \mathrm{Im}(\gamma \cdot z)^{\frac{k}{2}} = |cz+d|^k |f(z)| \frac{y^{\frac{k}{2}}}{|cz+d|^k} = \phi(z).$$

Furthermore, ϕ is continuous on the standard fundamental domain D. Since $|f(z)| = \mathcal{O}(e^{-2\pi y})$, we see that $\phi(z) \to 0$ when $y \to +\infty$. Thus ϕ is bounded and there exists a constant M such that

$$|f(z)| \leq My^{-\frac{k}{2}} \quad \text{when} \quad y \to +\infty. \tag{2.58.1}$$

Fix y and let x run through $[0,1]$. Then $q(z) = e^{2\pi i(x+iy)}$ describes a circle of center 0 and radius $e^{-2\pi y}$. We shall denote this circle C_y. Now consider the functions $g_n = fq^{-(n+1)}$, with n positive. Observe that

$$g_n = \sum_{m=1}^{\infty} a_m q^{m-n-1} = \sum_{k=-n}^{\infty} a_{k+n+1} q^k = \frac{a_1}{q^n} + \cdots + \frac{a_n}{q} + \sum_{k=0}^{\infty} a_{k+n+1} q^k.$$

Thus g_n has one pole at 0 of order n and $\mathrm{Res}_0(g_n) = a_n$. Applying the Residue Theorem, we get

$$a_n = \frac{1}{2\pi i} \int_{C_y} f(z) q(z)^{-n-1} dz = \int_0^1 f(x+iy) q(x+iy)^{-n} dx.$$

Using Equation 2.58.1 we see that $|a_n| \leq My^{-\frac{k}{2}} e^{2\pi n y}$. This inequality holds for all positive y. In particular, if $y = \frac{1}{n}$, then $|a_n| \leq Me^{2\pi} n^{\frac{k}{2}}$ and thus $a_n = \mathcal{O}(n^{\frac{k}{2}})$. □

2.5.3 Hecke operators for $\Gamma_0(N)$

We define Hecke operators for $\Gamma_0(N)$ and expose some of their main properties. These properties and their proofs are analogous to those of Section 2.5.1 and 2.5.2.

Let N and n be positive integers and consider the set of matrices

$$\Delta_n^N = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Delta_n \;\middle|\; a,b,c,d \in \mathbb{Z},\ (a,N) = 1,\ c \equiv 0 \pmod{N} \right\}.$$

The group $\Gamma_0(N)$ acts on Δ_n^N by left multiplication

$$\begin{array}{rcl} \Gamma_0(N) \times \Delta_n^N & \longrightarrow & \Delta_n^N \\ (\gamma, \alpha) & \longmapsto & \gamma\alpha. \end{array}$$

We decompose Δ_n^N in orbits of this action and choose a system of representatives R^N for $\Gamma_0(N)\backslash\Delta_n^N$. The finiteness of this system is a consequence of this next proposition.

Proposition 2.59. *For any positive integer n, we have the disjoint union*

$$\Delta_n^N = \bigcup \left\{ \Gamma_0(N) \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} \,\middle|\, a > 0,\ ad = n,\ 0 \leq b < d,\ (a, N) = 1 \right\}.$$

Proof. The inclusion "$\supseteq$" is straightforward. For the other inclusion, consider $\alpha = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Delta_n^N$. We know by Proposition 2.44 that there exists $\gamma = \left(\begin{smallmatrix} x & y \\ z & w \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{Z})$ such that

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} x & y \\ z & w \end{pmatrix} \begin{pmatrix} a' & b' \\ 0 & d' \end{pmatrix},$$

where $a'd' = n$, $a', d' > 0$ and $0 \leq b' < d'$. Since $a = a'x$ and $(a, N) = 1$, we have that $(a', N) = 1$. Then, since $c \equiv 0 \pmod N$ and $c = a'z$, we find that $z \equiv 0 \pmod N$ and thus $\gamma \in \Gamma_0(N)$. □

By analogy with Hecke operators for $\mathrm{SL}_2(\mathbb{Z})$, we define Hecke operators for $\Gamma_0(N)$, denoted T_n^N, as follows:

$$T_n^N f = n^{\frac{k}{2}-1} \sum_{\mu \in R^N} f^{[\mu]_k},$$

for $f \in \mathcal{M}_k(\Gamma_0(N))$.

2.60. If N is clear from the context, we shall write T_n instead of T_n^N.

Similarly, we get a more explicit formula for T_n if we choose the system that follows from Proposition 2.59. This yields,

$$T_n f(z) = n^{k-1} \sum_{\substack{0 \leq b < d, (a,N)=1 \\ a \geq 1, ad = n}} d^{-k} f(\frac{az+b}{d}).$$

Note that since $\left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right) \in \Gamma_0(N)$, every $f \in \mathcal{M}_k(\Gamma_0(N))$ can be written as

$$f = \sum_{m=0}^{\infty} a_m q^m.$$

Thus Proposition 2.47 can be fitted for $\Gamma_0(N)$. All other properties from Section 2.5.1 can be derived for $\Gamma_0(N)$. We end this section by stating an important theorem.

Theorem 2.61. *The Hecke forms form a basis for $\mathcal{M}_2(\Gamma_0(N))$.*

Proof. This is Theorem 5.7.2 of [He02]. The proof is not technical but requires the Petersson scalar product that we do not define. □

3 The Taniyama-Shimura Conjecture

We now have all the elements in hand needed to state the famous conjecture, which is the objective of our study. The statement was originally conjectured in 1955 by Yutaka Taniyama. His colleague Goro Shimura gave a precise statement of the conjecture in the following year. In 1967, a paper of André Weil provided strong evidence in favor of the conjecture. On September 19^{nth} 1994, Andrew Wiles, with help from Richard Taylor, obtained the full proof of the Taniyama-Shimura Conjecture for a special class of elliptic curves defined over $\mathbb{Q}$. This version of the conjecture sufficed to prove Fermat's Last Theorem. Extending the methods developed by Wiles in his proof, Christophe Breuil, Brian Conrad, Fred Diamond, and Richard Taylor obtained the proof of the full conjecture in 2001, which then became known as the *Modularity Theorem.* We now state the theorem.

Theorem 3.1 (Breuil, Conrad, Diamond, Taylor, Wiles)**.** *Let E be an elliptic curve defined over $\mathbb{Q}$ with conductor N and let $L(E,s)$ be the L-series of E:*

$$L(E,s) = \sum_{n=1}^{\infty} \frac{a(n)}{n^s}.$$

(See Section 1.5.3 for the definition of the coefficients of this L-series). Then the function $f : \mathcal{H} \to \mathbb{C}$ defined by

$$f = \sum_{n=1}^{\infty} a(n) q^n,$$

with $q : z \mapsto e^{2i\pi z}$, is a cusp form and a Hecke form for $\Gamma_0(N)$.

In 1637, Pierre de Fermat famously claimed to have proved, in the margin of a copy of the *Arithmetica*, what would become known as Fermat's Last Theorem. He claimed to have found a quite marvelous proof of this statement but that the margin was too small to contain it. Unfortunately, no sign of this proof was ever found and this lead to a 358-year long search for the proof. The statement is as follows.

Conjecture 3.2 (Fermat)**.** *Let n be a positive integer. The equation:*

$$\begin{cases} a^n + b^n = c^n \\ abc \neq 0 \\ a, b, c \in \mathbb{Z} \end{cases}$$

has no solutions for $n \geq 3$.

After the death of Fermat, mathematicians undertook the work of proving his many assertions one by one. The famous conjecture was the last one that was attempted and that is why it became known as *Fermat's Last Theorem.* It eventually became one of the most famous unsolved problems in mathematics and its resolution stimulated the development of several areas in mathematics during the many years of search. The conjecture was proved for several special cases of n and many advances were made by Euler, Goldbach, Germain, Legendre, Dirichlet, Lamé, Liouville, Cauchy, Kummer and many more famous mathematicians. But no one succeeded in proving the general case.

Around 1955, Taniyama and Shimura noticed a possible link between elliptic curves and modular forms. This was surprising since the two areas were completely unrelated. This resulted in the Taniyama-Shimura Conjecture which was precisely formulated around 1956. It is sometimes referred to as the Taniyama-Shimura-Weil Conjecture because of the contribution of Weil in 1967. In general, the conjecture was believed inaccessible. In 1983, Gerhard Frey noticed a link between the conjecture and Fermat's Last Theorem. He asserted that if a solution existed to Fermat's problem for $n \geq 3$, then from this solution, one would be able to extract an elliptic curve with such bizarre properties that its existence would contradict the Taniyama-Shimura Conjecture. It became clear that if this statement was correct, then a proof of Taniyama-Shimura would imply Fermat's Last Theorem. But he did not succeed in proving his statement. The missing part was identified by Jean-Pierre Serre and became known as the *epsilon-conjecture*. In 1986, Ken Ribet obtained a proof for this conjecture, which became known as Ribet's Theorem. The ground was then laid for Andrew Wiles who embarked on a 7-year long lonely search which resulted in him obtaining the proof of the Modularity Theorem in 1993 for elliptic curves that only admit good or multiplicative reduction modulo a prime p. This special case of the theorem was enough to imply Fermat's Last Theorem. But a flaw was discovered in his proof. With help from Richard Taylor, he corrected this flaw and the correct proof was submitted in 1995, making Wiles' childhood dream of proving Fermat come true. The methods used in his proof were extended by Breuil, Conrad, Diamond, and Taylor between 1996 and 2001 and led to the full proof of the Modularity Theorem.

A nice description of the historical context surrounding Fermat's Last Theorem and the Modularity Theorem can be found throughout [He02].

4 Computations

The aim of this section is to numerically illustrate the Modularity Theorem for some selected examples. We choose an approach using modular symbols and follow [St07] Chapter 3.

4.1 Modular symbols

We introduce modular symbols with the aim of constructing a basis for the vector space $\mathcal{S}_2(\Gamma_0(N))$. We then expose a trick due to Manin that gives a finite list of generators and relations for the space of modular symbols. We then define Hecke operators on modular symbols and study the relation between these and the Hecke operators for weight 2 modular forms for $\Gamma_0(N)$.

4.1.1 Definition

We consider homotopy classes of oriented paths in $\mathcal{H}^*$ with endpoints belonging to $\mathbb{P}^1(\mathbb{Q})$ and denote them $\{\alpha, \beta\}$, where $\alpha, \beta \in \mathbb{P}^1(\mathbb{Q})$. Note that the order of the pair matters. Consider the following homology relations:

(1) $\{\alpha, \beta\} + \{\beta, \gamma\} + \{\gamma, \alpha\} = 0$

(2) $\{\alpha, \alpha\} = 0$

(3) $\{\alpha, \beta\} = -\{\beta, \alpha\}$

for $\alpha, \beta, \gamma \in \mathbb{P}^1(\mathbb{Q})$.

Definition 4.1. We define $\mathbb{M}_2$ to be the $\mathbb{Q}$-vector space generated by symbols $\{\alpha, \beta\}$ quotiented by the three homology relations. An element of this space is called a *modular symbol*. Modular symbols of the form $\{\alpha, \beta\}$ are called *elementary symbols*.

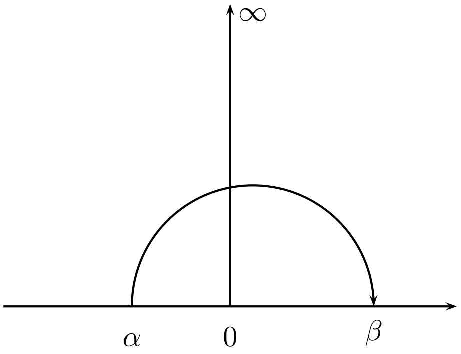

Figure 5: The elementary symbols $\{\alpha, \beta\}$ and $\{0, \infty\}$

4.2. Recall that an element s of a $\mathbb{Q}$-vector space V with basis B may be expressed in one and only one way as a $\mathbb{Q}$-linear combination of elements of B

$$s = \sum_{i=1}^{\infty} \lambda_i b_i,$$

where each $\lambda_i \in \mathbb{Q}$ and all but at most finitely many λ_i are 0. Elements of V are formal sums over B.

The group $\mathrm{GL}_2(\mathbb{Q})$ acts from the right on the space of modular symbols $\mathbb{M}_2$ in the following way: let $\gamma \in \mathrm{GL}_2(\mathbb{Q})$ and $s = \sum_{i=0}^{\infty} \lambda_i \{\alpha_i, \beta_i\} \in \mathbb{M}_2$. Then

$$\gamma \cdot s = \sum_{i=0}^{\infty} \lambda_i \gamma \cdot \{\alpha_i, \beta_i\} = \sum_{i=0}^{\infty} \lambda_i \{\gamma \cdot \alpha_i, \gamma \cdot \beta_i\}.$$

Here, $\mathrm{GL}_2(\mathbb{Q})$ acts on $\mathbb{P}^1(\mathbb{Q})$ by linear fractional transformation. That means that this action is the same as the one in Section 2.2.3 but extended to $\mathrm{GL}_2(\mathbb{Q})$. Explicitly this corresponds to the action

$$\gamma \cdot r = \begin{cases} \frac{ar+b}{cr+d} & \text{if } r \neq \infty \\ \frac{a}{c} & \text{if } r = \infty \end{cases}$$

where $r \in \mathbb{P}^1(\mathbb{Q})$ and $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{GL}_2(\mathbb{Q})$. If the denominator c or $cr + d$ equal zero, then we declare that $\gamma \cdot r = \infty \in \mathbb{P}^1(\mathbb{Q})$. Seeing that the action of $\mathrm{GL}_2(\mathbb{Q})$ on $\mathbb{M}_2$ is indeed an action and that it is well defined is immediate.

Definition 4.3. We define the space of modular symbols for $\Gamma_0(N)$ to be the space $\mathbb{M}_2(N)$ defined as $\mathbb{M}_2$ quotiented by the sub-vector space of $\mathbb{M}_2$ that is generated by modular symbols of the form $\{\alpha, \beta\} - \gamma\{\alpha, \beta\}$, where $\{\alpha, \beta\} \in \mathbb{M}_2$ and $\gamma \in \Gamma_0(N)$. An element of this space is called a modular symbol for $\Gamma_0(N)$.

4.1.2 Cuspidal modular symbols

Let $\mathbb{B}_2(N)$ denote the $\mathbb{Q}$-vector space with basis the finite set of cusps $C(\Gamma_0(N))$. We define the *boundary map*

$$\delta_N \ : \ \mathbb{M}_2(N) \longrightarrow \mathbb{B}_2(N)$$

to be the linear map that sends $\{\alpha,\, \beta\} \in \mathbb{M}_2(N)$ to $\{\beta\} - \{\alpha\}$, where $\{\beta\}$ denotes the basis element of $\mathbb{B}_2(N)$ corresponding to $\beta \in \mathbb{P}^1(\mathbb{Q})$. When the context is clear we shall write δ instead of δ_N.

4.4. This is a well defined map. Let $\alpha, \beta, \gamma \in \mathbb{P}^1(\mathbb{Q})$. We verify that the four relations we declared above hold in the image of δ. For the first one, we simply compute that $\delta(\{\alpha, \beta\} + \{\beta, \gamma\} + \{\gamma, \alpha\}) = (\{\beta\} - \{\alpha\}) + (\{\gamma\} - \{\beta\}) + (\{\alpha\} - \{\gamma\}) = 0$. For the second, $\delta(\{\alpha, \beta\}) = \{\beta\} - \{\alpha\} = -(\{\alpha\} - \{\beta\}) =$

$-\delta(\{\beta,\alpha\})$. Finally, it is clear that $\delta(\{\alpha,\alpha\}) = \{\alpha\} - \{\alpha\} = 0$. For the last relation, let $\gamma \in \Gamma_0(N)$. Then for all $\alpha \in \mathbb{P}^1(\mathbb{Q})$, $\{\alpha\} = \{\gamma \cdot \alpha\}$ since α and $\gamma \cdot \alpha$ represent the same cusp. Thus

$$\delta(\{\alpha,\beta\} - \gamma \cdot \{\alpha,\beta\}) = \{\beta\} - \{\alpha\} - \{\gamma \cdot \beta\} + \{\gamma \cdot \alpha\} = 0.$$

Definition 4.5. The kernel of δ_N is called the space of *cuspidal modular symbols* and is denoted $\mathbb{S}_2(N)$.

Thus, an element of $\mathbb{S}_2(N)$ can be thought of as a linear combination of paths in $\mathcal{H}^*$ whose endpoints are cusps.

4.1.3 Manin's trick

We know by Proposition 2.6 that $\Gamma_0(N)$ has finite index in $\mathrm{SL}_2(\mathbb{Z})$. Let $r_0, \ldots, r_m$ be a system of right coset representatives for $\Gamma_0(N)$ in $\mathrm{SL}_2(\mathbb{Z})$. Then

$$\mathrm{SL}_2(\mathbb{Z}) = \bigcup_{i=0}^{m} \Gamma_0(N) r_i,$$

where the union is disjoint. Manin's trick enables us to write every modular symbol as a $\mathbb{Q}$-linear combination of symbols of the form $r_i\{0,\infty\}$. It shows that $\{r_i\{0,\infty\} \mid i = 0, \cdots, m\}$ generates $\mathbb{M}_2(N)$.

Theorem 4.6 (Manin)**.** *Let $r_0, \ldots, r_m$ be a system of right coset representatives for $\Gamma_0(N)$ in $\mathrm{SL}_2(\mathbb{Z})$. The linear map*

$$\begin{array}{rcl} \varphi : \mathbb{Q}[r_0, \ldots, r_m] & \longrightarrow & \mathbb{M}_2(N) \\ \sum_{i=0}^{m} \lambda_i r_i & \longmapsto & \sum_{i=0}^{m} \lambda_i r_i \{0,\infty\} \end{array}$$

is surjective.

Proof. Let $\{\alpha,\beta\} \in \mathbb{M}_2(N)$. Using the first and the third homology relation, we notice that $\{\alpha,\beta\} = \{0,\beta\} - \{0,\alpha\}$ for all $\alpha, \beta \in \mathbb{P}^1(\mathbb{Q})$. Thus it suffices to show that there exists $\{\lambda_i\}_{i=0}^{m} \subset \mathbb{Z}$ such that $\sum_{i=0}^{m} \lambda_i r_i\{0,\infty\} = \{0, \frac{b}{a}\}$, where $(a,b) = 1$.

The first step is to write $\frac{b}{a}$ as a *simple continued fraction*. That is

$$\frac{b}{a} = a_0 + \cfrac{1}{a_1 + \frac{1}{a_2 + \frac{1}{a_3 + \ldots}}},$$

where a_i for $i = 0, 1, \ldots$ are integers. We shall write this in the following way, which is more compact:

$$\frac{b}{a} = a_0 + \frac{1}{a_1+} \frac{1}{a_2+} \cdots$$

We have written the continued fraction as if it is infinite but it is actually finite in our case since we only consider rationals. Thus, there exists a positive integer N such that

$$\frac{b}{a} = a_0 + \frac{1}{a_1+}\frac{1}{a_2+}\cdots\frac{1}{a_N}.$$

The second step of the trick consists in considering the successive *convergents* of $\frac{b}{a}$. The n^{th} convergent, for $0 \leq n \leq N$, is denoted $\frac{p_n}{q_n}$ and defined as

$$\frac{p_n}{q_n} = a_0 + \frac{1}{a_1+}\frac{1}{a_2+}\cdots\frac{1}{a_n}.$$

In addition we also consider two formal convergents, namely the ones of order -2 and -1 defined formally and respectively as $\frac{p_{-2}}{q_{-2}} = \frac{0}{1}$ and $\frac{p_{-1}}{q_{-1}} = \frac{1}{0}$.

We notice that $p_k q_{k-1} - p_{k-1} q_k = (-1)^{k-1}$ and thus the matrix defined by

$$\gamma_k = \begin{pmatrix} p_k & (-1)^{k-1} p_{k-1} \\ q_k & (-1)^{k-1} q_{k-1} \end{pmatrix}$$

is in $\mathrm{SL}_2(\mathbb{Z})$. Furthermore, we see that

$$\gamma_k\{0,\infty\} = \{\gamma_k \cdot 0, \gamma_k \cdot \infty\} = \left\{\frac{p_{k-1}}{q_{k-1}}, \frac{p_k}{q_k}\right\} = r_i\{0,\infty\},$$

for a certain $i \in \{0, \cdots, m\}$. In fact, there exists $\delta \in \Gamma_0(N)$ and $0 \leq i \leq m$ such that $\delta r_i = \gamma_k$. Thus $\gamma_k\{0,\infty\} = \delta r_i\{0,\infty\} = r_i\{0,\infty\}$ since we quotiented by the relation $\{\alpha,\beta\} - \gamma\{\alpha,\beta\} = 0$.

Finally, note that

$$\left\{0, \frac{b}{a}\right\} = \left\{\frac{p_{-2}}{q_{-2}}, \frac{p_{-1}}{q_{-1}}\right\} + \left\{\frac{p_{-1}}{q_{-1}}, \frac{p_0}{q_0}\right\} + \ldots + \left\{\frac{p_{N-1}}{q_{N-1}}, \frac{p_N}{q_N}\right\}.$$

This is a consequence of the homology relations we stated and the fact that $p_{-2}/q_{-2} = 0$ and $p_N/q_N = b/a$ by definition. Using our observation above, we know that for each $-1 \leq k \leq N$, there exists $0 \leq i \leq m$ such that

$$\left\{\frac{p_{k-1}}{q_{k-1}}, \frac{p_k}{q_k}\right\} = r_i\{0,\infty\}$$

and this completes the proof. □

Example 4.7. Let $N = 11$ and $\frac{b}{a} = \frac{4}{7}$. We use Manin's trick to express the modular symbol $\{0, 4/7\}$ as a $\mathbb{Q}$-linear combination of $r_i\{0,\infty\}$. We know from Proposition 2.6 that the index of $\Gamma_0(11)$ in $\mathrm{SL}_2(\mathbb{Z})$ is 12. Using Proposition 2.7 we choose the following system of right coset representatives for $\Gamma_0(11) \setminus \mathrm{SL}_2(\mathbb{Z})$:

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}, \ldots, \begin{pmatrix} 1 & 0 \\ 10 & 1 \end{pmatrix}, \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

We start by expanding $\frac{4}{7}$ as a continued fraction. The algorithm is simple. Let a_0 be the integer part of $\frac{4}{7}$, that is 0 and write

$$\frac{4}{7} = a_0 + \frac{1}{\frac{7}{4}}$$

Let a_1 be the integer part of $\frac{7}{4}$, that is 1. Repeat this process until you obtain the desired form. Doing this yields

$$\frac{4}{7} = 0 + \cfrac{1}{1 + \frac{1}{1+\frac{1}{3}}}$$

and we may write

$$\frac{4}{7} = 0 + \frac{1}{1+}\frac{1}{1+}\frac{1}{3}.$$

The convergents of $\frac{4}{7}$ are

$$\frac{p_{-2}}{q_{-2}} = \frac{0}{1}, \quad \frac{p_{-1}}{q_{-1}} = \frac{1}{0}, \quad \frac{p_0}{q_0} = \frac{0}{1}, \quad \frac{p_1}{q_1} = \frac{1}{1}, \quad \frac{p_2}{q_2} = \frac{1}{2}, \quad \frac{p_3}{q_3} = \frac{4}{7}.$$

Thus $\{0, 4/7\} = \{0, \infty\} + \{\infty, 0\} + \{0, 1\} + \{1, 1/2\} + \{1/2, 4/7\} = \{0, 1\} + \left(\begin{smallmatrix} 1 & -1 \\ 2 & -1 \end{smallmatrix}\right)\{0, \infty\} + \left(\begin{smallmatrix} 4 & 1 \\ 7 & 2 \end{smallmatrix}\right)\{0, \infty\}$. We state that $\{0, 1\} = 0$ for all N. To see why, note that $T = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right) \in \Gamma_0(N)$ for all N. Thus $\{\infty, 0\} = \{T \cdot \infty, T \cdot 0\} = \{\infty, 1\}$. So

$$0 = \{\infty, 1\} - \{\infty, 0\} = \{\infty, 1\} + \{0, \infty\} = \{0, \infty\} + \{\infty, 1\} = \{0, 1\}.$$

Also, notice that

$$\begin{pmatrix} 1 & -1 \\ 2 & -1 \end{pmatrix} = \begin{pmatrix} 10 & -1 \\ 11 & -1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 9 & 1 \end{pmatrix} \text{ and } \begin{pmatrix} 4 & 1 \\ 7 & 2 \end{pmatrix} = \begin{pmatrix} -5 & 1 \\ -11 & 2 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 9 & 1 \end{pmatrix}.$$

and $\left(\begin{smallmatrix} 10 & -1 \\ 11 & -1 \end{smallmatrix}\right), \left(\begin{smallmatrix} -5 & 1 \\ -11 & 2 \end{smallmatrix}\right) \in \Gamma_0(11)$. Thus we may write

$$\{0, 4/7\} = 2\Big[\begin{pmatrix} 1 & 0 \\ 9 & 1 \end{pmatrix}\{0, \infty\}\Big].$$

4.1.4 Manin symbols

Fix right coset representatives $r_0, \ldots, r_m$ for $\Gamma_0(N)$ in $\mathrm{SL}_2(\mathbb{Z})$. We introduce the following notation $r_i\{0, \infty\} = [r_i]$ for $i = 0, \ldots, m$ and we consider the formal symbols $[r_i]'$, for $i = 0, \ldots, m$. We define the right action

$$\begin{array}{ccc} \mathrm{SL}_2(\mathbb{Z}) \times \{[r_0]', \ldots, [r_m]'\} & \longrightarrow & \{[r_0]', \ldots, [r_m]'\} \\ (\gamma,\ [r_i]') & \longmapsto & [r_i]' \cdot \gamma = [r_j]', \end{array}$$

where $\Gamma_0(N) r_j = \Gamma_0(N) r_i \gamma$. We extend the notion by setting

$$[\alpha]' = [\Gamma_0(N)\alpha]' = [r_i]',$$

where r_i is such that $\alpha \in \Gamma_0(N) r_i$. Then the right action of $\Gamma_0(N)$ is $[\alpha]' \cdot \gamma = [\alpha\gamma]'$.

Definition 4.8. The symbols $[r_0]', \ldots, [r_m]'$ are called *Manin symbols.*

Note that this definition does not depend on the choice of the representatives r_i. In fact, let $\{g_0, \dots, g_m\}$ be another system of representatives. Then for each $0 \leq i \leq m$, there exists $0 \leq k = k_i \leq m$ such that $\Gamma_0(N)g_i = \Gamma_0(N)r_{k_i}$ and $k_i \neq k_j$ for $i \neq j$. Thus, $[g_i]' = [\Gamma_0(N)g_i]' = [r_{k_i}]'$. We may reorder the g_i's in a way such that $[g_i]' = [r_i]'$ for all i.

Recall from Proposition 2.8 that the modular group $\mathrm{SL}_2(\mathbb{Z})$ is generated by the matrices $S = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$ and $T = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$.

Theorem 4.9 (Manin)**.** *Let M denote the quotient of the $\mathbb{Q}$-vector space generated by the Manin symbols $[r_0]', \dots, [r_m]'$ by the sub-vector space generated by the relations (for all $i = 0, \dots, m$)*

$$[r_i]' + [r_i]' \cdot S = 0 \quad \textit{and} \quad [r_i]' + [r_i]' \cdot (TS) + [r_i]' \cdot (TS)^2 = 0.$$

Then there is an isomorphism

$$\Psi : M \xrightarrow{\sim} \mathbb{M}_2(N)$$

given by $[r_i]' \mapsto [r_i]$ for all i.

Proof. The injectivity of Ψ requires a certain amount of work and therefore we refer to [Ma72] Chapter 1 Section 7 for the proof of this part. The surjectivity of Ψ is a consequence of Theorem 4.6. We show that Ψ is well defined. We need to verify that the relations considered in the theorem hold in the image. To prove that the first relation holds, we observe that

$$\begin{aligned} \Psi([r_i]' + [r_i]' \cdot S) &= \Psi([r_i]' + [r_iS]') = [r_i] + [r_iS] = r_i\{0, \infty\} + r_iS\{0, \infty\} \\ &= \{r_i \cdot 0, r_i \cdot \infty\} + \{r_i \cdot (S \cdot 0), r_i \cdot (S \cdot \infty)\} = \{r_i \cdot 0, r_i \cdot \infty\} + \{r_i \cdot \infty, r_i \cdot 0\} = 0, \end{aligned}$$

where we used the third homology relation for the last equality.

For the second relation, start by noticing that $TS = \left(\begin{smallmatrix} 1 & -1 \\ 1 & 0 \end{smallmatrix}\right)$ and $(TS)^2 = \left(\begin{smallmatrix} 0 & -1 \\ 1 & -1 \end{smallmatrix}\right)$. Then

$$\begin{aligned} \Psi([r_i]' + [r_i]' \cdot (TS) + [r_i]' \cdot (TS)^2) &= \Psi([r_i]' + [r_i(TS)]' + [r_i(TS)^2]') = [r_i] + [r_i(TS)] + [r_i(TS)^2] \\ &= \{r_i \cdot 0, r_i \cdot \infty\} + \{r_i \cdot ((TS) \cdot 0), r_i \cdot ((TS) \cdot \infty)\} + \{r_i \cdot ((TS)^2 \cdot 0), r_i \cdot ((TS)^2 \cdot \infty)\} \\ &= \{r_i \cdot 0, r_i \cdot \infty\} + \{r_i \cdot \infty, r_i \cdot 1\} + \{r_i \cdot 1, r_i \cdot 0\} = 0, \end{aligned}$$

where we used the first homology relation for the last equality. □

4.10. This theorem provides a finite presentation for the space of modular symbols. In particular

$$\dim(\mathbb{M}_2(N)) = \dim(M) \leq [\mathrm{SL}_2(\mathbb{Z}) \, : \, \Gamma_0(N)].$$

4.2 Hecke operators on modular symbols

We define Hecke operators on modular symbols and then introduce a perfect pairing between these and the Hecke operators for modular forms for $\Gamma_0(N)$ and weight 2.

4.2.1 Definition

Let N and n be positive integers and consider the set of matrices

$$\Delta_n^N = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Delta_n \,\middle|\, a, b, c, d \in \mathbb{Z},\ (a, N) = 1,\ c \equiv 0 \pmod N \right\}$$

just as we did when we defined Hecke operators for modular forms for $\Gamma_0(N)$. Recall from Proposition 2.59, that

$$R_n = \left\{ \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} \,\middle|\, ad = n, a \geq 1, (a, N) = 1, 0 \leq b < d \right\}$$

is a system of right coset representatives of $\Gamma_0(N) \setminus \Delta_n^N$.

Definition 4.11. Let N and n be positive integers. We define the n^{th} Hecke operator on $\mathbb{M}_2(N)$ to be the map from $\mathbb{M}_2(N)$ to $\mathbb{M}_2(N)$ defined by

$$T_n s = \sum_{g \in R_n} g \cdot s.$$

It is easy to verify that this is a well defined map. We now consider the case where $n = p$ is prime. If $p \nmid N$ and $g = \left(\begin{smallmatrix} a & b \\ 0 & d \end{smallmatrix}\right) \in R_p$, then since $ad = p$ and a is positive we must have $a = 1, d = p$ or $a = p, d = 1$. In the first case, $g = \left(\begin{smallmatrix} 1 & b \\ 0 & p \end{smallmatrix}\right)$ with $0 \leq b < p$. In the second case, $g = \left(\begin{smallmatrix} p & b \\ 0 & 1 \end{smallmatrix}\right)$, with $0 \leq b < 1$ and therefore $g = \left(\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\right)$. If $p|N$, then $a \neq p$ because otherwise $(a, N) \neq 1$. Thus $g = \left(\begin{smallmatrix} 1 & b \\ 0 & p \end{smallmatrix}\right)$, with $0 \leq b < p$. Hence, if $p \nmid N$

$$R_p = \left\{ \begin{pmatrix} 1 & r \\ 0 & p \end{pmatrix} \,\middle|\, 0 \leq r < p \right\} \cup \left\{ \begin{pmatrix} p & 0 \\ 0 & 1 \end{pmatrix} \right\}.$$

Otherwise,

$$R_p = \left\{ \begin{pmatrix} 1 & r \\ 0 & p \end{pmatrix} \,\middle|\, 0 \leq r < p \right\}.$$

It follows from these observations and our definition of Hecke operators that if N is a positive integer, p is prime and $\{\alpha, \beta\} \in \mathbb{M}_2(N)$, then the p^{th} Hecke operator on the elementary symbol $\{\alpha, \beta\}$ is

$$T_p\{\alpha, \beta\} = \begin{cases} \left(\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\right) \cdot \{\alpha, \beta\} + \displaystyle\sum_{r=0}^{p-1} \left(\begin{smallmatrix} 1 & r \\ 0 & p \end{smallmatrix}\right) \cdot \{\alpha, \beta\} & \text{if } p \nmid N \\ \displaystyle\sum_{r=0}^{p-1} \left(\begin{smallmatrix} 1 & r \\ 0 & p \end{smallmatrix}\right) \cdot \{\alpha, \beta\} & \text{if } p \mid N. \end{cases}$$

Thus if $s = \sum_{i=0}^{\infty} \lambda_i \{\alpha_i,\ \beta_i\} \in \mathbb{M}_2(N)$, then

$$T_p s = \sum_{g \in R_p} g \cdot s = \sum_{i=0}^{\infty} \lambda_i \sum_{g \in R_p} g \cdot \{\alpha_i,\ \beta_i\}.$$

4.2.2 Pairing of modular symbols with modular forms

Consider the pairing $\langle \cdot, \cdot \rangle \ : \ \mathbb{M}_2(N) \times \mathcal{M}_2(\Gamma_0(N)) \longrightarrow \mathbb{C}$ that sends $\{\alpha, \beta\} \in \mathbb{M}_2(N)$ and $f \in \mathcal{M}_2(\Gamma_0(N))$ to

$$\int\limits_{\alpha\to\beta} f(z)dz$$

that we extend $\mathbb{Q}$-linearly in the first variable in the sense that if $s = \sum_{i=0}^{\infty} \lambda_i\{\alpha_i,\ \beta_i\} \in \mathbb{M}_2(N)$ with $\{\lambda_i\}_{i=0}^{\infty} \subset \mathbb{Q}$, then

$$\langle s, f\rangle = \sum_{i=0}^{\infty} \lambda_i \int\limits_{\alpha_i\to\beta_i} f(z)dz.$$

We immediately notice that $\langle \cdot, \cdot \rangle$ is $\mathbb{Q}$-bilinear and $\mathbb{C}$-linear in the second variable. The integral is taken over a path $\alpha \to \beta$ that represents the elementary symbol $\{\alpha, \beta\}$. Note that this pairing makes sense since the integral is independent of the choice of the path $\alpha \to \beta$. To see this, we adopt the notation $\delta_1 := \alpha \to \beta$ and let δ_2 be another path that represents the symbol $\{\alpha, \beta\}$. Since f is holomorphic on $\mathcal{H}$, Cauchy's theorem tells us that the integral of f on any closed path in $\mathcal{H}$ is equal to zero. But this does not suffice to prove the independence of the choice of paths since the closed path $C := \delta_1 - \delta_2$ contains two points that belong to $\mathbb{P}^1(\mathbb{Q})$, namely the points α and β. But there is a way to avoid this problem. Let $\epsilon > 0$ and let C_ϵ^α and C_ϵ^β be the circles with radius ϵ and respective centers α and β.

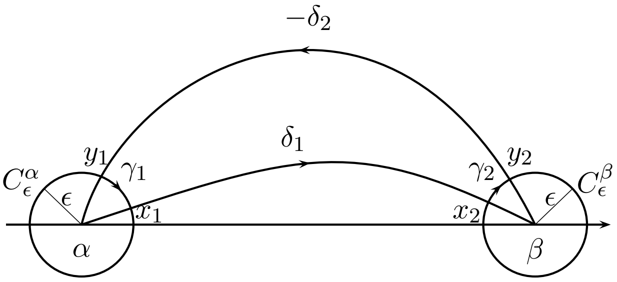

Figure 6: The closed path Γ in $\mathcal{H}$

Let $x_1 = \delta_1 \cap C_\epsilon^\alpha, y_1 = \delta_2 \cap C_\epsilon^\alpha, x_2 = \delta_1 \cap C_\epsilon^\beta$ and $y_2 = \delta_2 \cap C_\epsilon^\beta$. Define $\gamma_1 = (y_1 \to x_1) \subset (C_\epsilon^\alpha \cap \overset{\circ}{C})$ and $\gamma_2 = (x_2 \to y_2) \subset (C_\epsilon^\beta \cap \overset{\circ}{C})$. Observe that the path $\Gamma := \delta_1 + \gamma_2 - \delta_2 + \gamma_1$ is closed in $\mathcal{H}$ (see Figure 6). Thus by our previous discussion, the integral of f on Γ is zero. Thus

$$\int_{\delta_2} f(z)dz - \int_{\delta_1} f(z)dz = \int_{\gamma_1} f(z)dz + \int_{\gamma_2} f(z)dz.$$

As $\epsilon \to 0$, the path γ_i "shrinks" towards zero. But does the integral of f of γ_i also tend to zero? The answer is yes. Since f is holomorphic at the cusps and holomorphicity at a cusp implies that f is bounded in a close neighborhood of the cusp, we see that there exists $M > 0$ such that

$$\int_{\gamma_i} |f(z)|dz \leq M \int_{\gamma_i} dz \to 0 \text{ as } \epsilon \to 0,$$

for $i = 1, 2$. Hence

$$\int_{\delta_1} f(z)dz = \int_{\delta_2} f(z)dz$$

as we set out to prove.

4.12. The above pairing is well defined. We need to verify that the pairing satisfies the three homology relations from Section 4.1.1 and the relation defined in Definition 4.3. Let $\alpha, \beta, \gamma \in \mathbb{P}^1(\mathbb{Q})$ and $f \in \mathcal{M}_2(\Gamma_0(N))$. In order to verify the three relations of Section 4.1.1, we need to show that

$$\int_{(\alpha\to\beta)+(\beta\to\gamma)+(\gamma\to\alpha)} f(z)dz = 0 \qquad \int_{\alpha\to\alpha} f(z)dz = 0 \qquad \int_{(\alpha\to\beta)+(\beta\to\alpha)} f(z)dz = 0.$$

Note that all three paths $(\alpha \to \beta) + (\beta \to \gamma) + (\gamma \to \alpha), \alpha \to \alpha$ and $(\alpha \to \beta) + (\beta \to \alpha)$ form closed loops in $\mathcal{H}^*$. Thus, the desired results follow from arguments similar to the ones already developed above.

In order to show the result for the relation of Definition 4.3, let $\gamma \in \Gamma_0(N)$. We need to show that

$$\langle \gamma \cdot \{\alpha, \beta\}, f\rangle = \langle \{\alpha, \beta\}, f\rangle.$$

By definition,

$$\langle \gamma \cdot \{\alpha, \beta\}, f\rangle = \int_{\gamma\cdot\alpha\to\gamma\cdot\beta} f(\gamma \cdot z)d(\gamma \cdot z).$$

Since $f \in \mathcal{M}_2(\Gamma_0(N))$, we know by Remark 2.11 that $f(\gamma \cdot z)d(\gamma \cdot z) = f(z)dz$. Furthermore, $\gamma \cdot \{\alpha, \beta\} = \{\alpha, \beta\}$. Thus

$$\int_{\gamma\cdot\alpha\to\gamma\cdot\beta} f(\gamma \cdot z)d(\gamma \cdot z) = \int_{\alpha\to\beta} f(z)dz,$$

as desired.

Proposition 4.13. *The Hecke operators are compatible with the integration pairing $\langle\cdot, \cdot\rangle$ in the sense that for all $s \in \mathbb{M}_2(N)$, $f \in \mathcal{M}_2(\Gamma_0(N))$ and p prime, $\langle T_p s, f\rangle = \langle s, T_p f\rangle$.*

Proof. Since the pairing is $\mathbb{Q}$-linear in the first variable it suffices to prove the equality for an elementary symbol $\{\alpha, \beta\} \in \mathbb{M}_2(N)$. Let p be a prime and $f \in \mathcal{M}_2(\Gamma_0(N))$. Then

$$\langle\{\alpha, \beta\}, T_p f\rangle = \langle\{\alpha, \beta\}, p^{\frac{2}{2}-1} \sum_{g\in R_p} f^{[g]_2}\rangle = \langle\{\alpha, \beta\}, \sum_{g\in R_p} f^{[g]_2}\rangle$$

$$= \int_{\alpha\to\beta} \sum_{g\in R_p} f^{[g]_2}(z)dz = \sum_{g\in R_p} \int_{\alpha\to\beta} pd^{-2} f(\frac{az + b}{d})dz.$$

We perform the change of variables $w = \frac{az+b}{d}$. Then

$$\langle\{\alpha, \beta\}, T_p f\rangle = \sum_{g\in R_p} \int_{g\cdot\alpha\to g\cdot\beta} \frac{d}{a} pd^{-2} f(w)dw = \sum_{g\in R_p} \int_{g\cdot\alpha\to g\cdot\beta} \frac{p}{ad} f(w)dw = \sum_{g\in R_p} \int_{g\cdot\alpha\to g\cdot\beta} f(w)dw.$$

On the other hand,

$$\langle T_p\{\alpha, \beta\}, f\rangle = \langle \sum_{g\in R_p} g\cdot\{\alpha,\ \beta\}, f\rangle = \sum_{g\in R_p}\int_{g\cdot\alpha\to g\cdot\beta} f(z)dz.$$

Thus $\langle T_p\{\alpha, \beta\}, f\rangle = \langle\{\alpha, \beta\}, T_p f\rangle$. □

Restricting the pairing as follows $\langle\cdot,\cdot\rangle\ :\ \mathbb{S}_2(N)\times\mathcal{S}_2(\Gamma_0(N)) \longrightarrow \mathbb{C}$, it remains $\mathbb{Q}$-bilinear and $\mathbb{C}$-linear in the second variable. Hence, the map

$$\alpha\ :\ \mathbb{S}_2(N) \longrightarrow \mathrm{Hom}_{\mathbb{C}}(\mathcal{S}_2(\Gamma_0), \mathbb{C}) = \mathcal{S}_2(\Gamma_0(N))^*$$

that sends $s \in \mathbb{S}_2(N)$ to $\langle s, \cdot\rangle$ is well defined. Here, $\mathcal{S}_2(\Gamma_0(N))^*$ denotes the dual of $\mathcal{S}_2(\Gamma_0(N))$. We state the following theorem without proof.

Theorem 4.14. *The map $\alpha_{\mathbb{R}} : \mathbb{S}_2(N)\otimes_{\mathbb{Q}}\mathbb{R} \longrightarrow \mathcal{S}_2(\Gamma_0(N))^*$ that sends s to $\langle s, \cdot\rangle$ is an isomorphism. In other words, our pairing is perfect.*

Proof. See Theorem 3.4 of [St07]. □

Corollary 4.15.

$$\dim_{\mathbb{Q}} \mathbb{S}_2(N) = 2\dim_{\mathbb{C}} \mathcal{S}_2(\Gamma_0(N))$$

Proof. This is Proposition 3.8 of [St07]. □

Corollary 4.16. *The Hecke operators for $\mathcal{S}_2(\Gamma_0(N))$ and those for $\mathbb{S}_2(N)$ have the same eigenvalues.*

Proposition 4.17. *The Hecke operators on $\mathbb{S}_2(N)$ commute. Moreover, the function $n \mapsto T_n$ is multiplicative. Thus, T_n for n composite is determined by the operators T_p with p prime.*

Proof. This proposition is a consequence of Corollary 2.48 which states the same result for Hecke operator on $\mathcal{S}_2(\Gamma_0(N))$. Let $s \in \mathbb{S}_2(N)$ and $f \in \mathcal{S}_2(\Gamma_0(N))$. Then $\langle s, T_nT_mf\rangle = \langle s, T_mT_nf\rangle$ and using Proposition 4.13 we see that $\langle T_mT_ns, f\rangle = \langle T_nT_ms, f\rangle$ which implies that $\alpha_{\mathbb{R}}(T_mT_ns) = \alpha_{\mathbb{R}}(T_nT_ms)$ since f is arbitrary. Since $\alpha_{\mathbb{R}}$ is an injection, we must then have $T_mT_ns = T_nT_ms$.

If n and m are relatively prime, then $T_nT_mf = T_{nm}f$. Thus

$$\langle s, T_nT_mf\rangle = \langle s, T_{nm}f\rangle = \langle T_{nm}s, f\rangle.$$

But we also have $\langle s, T_nT_mf\rangle = \langle T_mT_ns, f\rangle$ and thus $\alpha_{\mathbb{R}}(T_{nm}s) = \alpha_{\mathbb{R}}(T_mT_ns)$ and since $\alpha_{\mathbb{R}}$ is an injection, we see that $T_mT_ns = T_{nm}s$. □

4.18. If $2d = \dim(\mathbb{S}_2(N))$ and $\{e_1, \dots, e_{2d}\}$ is a basis for this space, then the series

$$f_i = \sum_{n=0}^{\infty} \langle T_n e_i, e_i \rangle q^n$$

are Hecke forms for $\Gamma_0(N)$ for $i = 1, \dots, 2d$. Here, $\langle \cdot, \cdot \rangle$ denotes the standard scalar product on $\mathbb{C}$. Thus, we may compute a basis for $\mathcal{S}_2(\Gamma_0(N))$ by looking only at the eigenvalues of the Hecke operators T_p on $\mathbb{S}_2(N)$. It suffices to list the series f_i until we have enough to form a basis for $\mathcal{S}_2(\Gamma_0(N))$ and this is possible since $\dim(\mathcal{S}_2(\Gamma_0(N))) = d$ by Corollary 4.15.

4.3 Computations

We have now covered the theory that is necessary and sufficient for us to numerically illustrate the Modularity Theorem. We give a table of the dimension g of $\mathcal{S}_2(\Gamma_0(N))$.

N	1	2	3	4	5	6	7	8	9	10	11	12	13
g	0	0	0	0	0	0	0	0	0	0	1	0	0
N	14	15	16	17	18	19	20	21	22	23	24	25	26
g	1	1	0	1	0	1	1	1	2	2	1	0	2
N	27	28	29	30	31	32	33	34	35	36	37	38	39
g	1	2	2	3	2	1	3	3	3	1	2	4	3

Table 5: Dimension g of $\mathcal{S}_2(\Gamma_0(N))$.

We shall proceed in the following way.

(i) We compute a basis for the space M of Manin symbols. By Theorem 4.9, we already know that the Manin symbols generate M. We compute the relations defined on M in this theorem using the action of $\mathrm{GL}_2(\mathbb{Z})$ on M defined in Section 4.1.4. Then, using linear algebra we study the independence between these symbols by looking at the relations. This will give us a basis for M. Once we have a basis for M, we also have one for $\mathbb{M}_2(N)$ by Theorem 4.9.

(ii) We compute the image of each basis element of this basis by the Hecke operator T_n. This allows us to express T_n as a matrix. We compute its eigenvalues.

(iii) We compute the boundary map and find a basis for its kernel, which is the space $\mathbb{S}_2(N)$. We compute T_n on $\mathbb{S}_2(N)$ and find its eigenvalues.

(iii) We construct a basis for $\mathcal{S}_2(\Gamma_0(N))$ using Remark 4.18.

(iv) We consider an elliptic curve with conductor N and compute the coefficients a_n of its L-series. We then express the function $f = \sum a_n q^n$ in the basis we constructed for $\mathcal{S}_2(\Gamma_0(N))$.

4.3.1 Illustration for $N = 3$

We start by computing T_2 on the space $\mathbb{M}_2(3)$. The strategy is to compute a basis for the space M on Manin symbols. Then, using Theorem 4.9, we know that $M \cong \mathbb{M}_2(3)$, and thus we will also have computed a basis for $\mathbb{M}_2(N)$. We then apply T_2 on this basis and express the operator as a matrix.

Recall from Proposition 2.7 that the matrices

$$r_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},\ r_1 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix},\ r_2 = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix},\ r_3 = \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}$$

form a system of right coset representatives of $\Gamma_0(3)$ in $\mathrm{SL}_2(\mathbb{Z})$.

We list the relations defined in Theorem 4.9. We do the computations of the first relations. The first relation is $[r_0]' + [r_0]' \cdot S = 0$. Referring to the action defined in Section 4.1.4, we see that $[r_0]' + [r_0]' \cdot S = [r_0]' + [r_0 S]' = [r_0]' + [r_i]'$, where $\Gamma_0(3)(r_0 S) = \Gamma_0(3) r_i$. Notice that $r_0 S = S = r_1$ and thus $[r_0 S]' = [r_1]'$. Hence, the relation is $[r_0]' + [r_1]' = 0$.

The second relation is $[r_0]' + [r_0]' \cdot (TS) + [r_0]' \cdot (TS)^2 = 0$. First, compute that $r_0 TS = TS = \left(\begin{smallmatrix} 1 & -1 \\ 1 & 0 \end{smallmatrix}\right)$ and notice that $\left(\begin{smallmatrix} 1 & -1 \\ 1 & 0 \end{smallmatrix}\right) = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)\left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right) = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right) r_1$ and that $\left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right) \in \Gamma_0(3)$. Thus, $\Gamma_0(3)(r_0 TS) = \Gamma_0(3) r_1$ and we may conclude that $[r_0 TS]' = [r_1]'$. Similarly, we compute that $r_0 (TS)^2 = (TS)^2 = \left(\begin{smallmatrix} 0 & -1 \\ 1 & -1 \end{smallmatrix}\right)$. Notice that $\left(\begin{smallmatrix} 0 & -1 \\ 1 & -1 \end{smallmatrix}\right) = \left(\begin{smallmatrix} 2 & -1 \\ 3 & -1 \end{smallmatrix}\right)\left(\begin{smallmatrix} 1 & 0 \\ 2 & 1 \end{smallmatrix}\right) = \left(\begin{smallmatrix} 2 & -1 \\ 3 & -1 \end{smallmatrix}\right) r_3$ and that $\left(\begin{smallmatrix} 2 & -1 \\ 3 & -1 \end{smallmatrix}\right) \in \Gamma_0(3)$. Thus, $\Gamma_0(3)(r_0 TS) = \Gamma_0(3) r_3$ and we may conclude that $[r_0 (TS)^2]' = [r_3]'$. Thus, the second relation is $[r_0]' + [r_1]' + [r_3]' = 0$. By iterating this procedure we find the six remaining relations:

$$\begin{array}{llll}
[r_0]' + [r_1]' = 0 & [r_0]' + [r_1]' = 0 & [r_2]' + [r_3]' = 0 & [r_2]' + [r_3]' = 0 \\
[r_0]' + [r_1]' + [r_3]' = 0 & [r_0]' + [r_1]' + [r_3]' = 0 & 3[r_2]' = 0 & [r_0]' + [r_1]' + [r_3]' = 0.
\end{array}$$

Recall that the aim is to find a basis for the space M. We already know that $[r_0]', [r_1]', [r_2]'$ and $[r_3]'$ generate M but since we quotient by the above relations, we need to look at independence between the Manin symbols modulo these relations. We express the above relations as a linear system:

$$\begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 3 & 0 \end{pmatrix} \begin{pmatrix} [r_0]' \\ [r_1]' \\ [r_2]' \\ [r_3]' \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Notice that we omit the recurrent relations. Taking the reduced row echelon form of this matrix yields the equivalent system

$$\begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} [r_0]' \\ [r_1]' \\ [r_2]' \\ [r_3]' \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Thus, our complicated system of relations above simplifies to

$$[r_0]' + [r_1]' = 0 \qquad [r_2]' = 0 \qquad [r_3]' = 0.$$

Hence, M has dimension 1 and its basis is the class of $[r_1]'$ (or $[r_2]'$). We choose $[r_1]'$. Notice that by Theorem 4.9, these relations hold in the image. Thus

$$[r_0] + [r_1] = 0 \qquad [r_2] = 0 \qquad [r_3] = 0. \tag{4.18.1}$$

This implies that $\dim(\mathbb{M}_2(N)) = 1$ and we may choose as basis of $\mathbb{M}_2(N)$ the class of the modular symbol $[r_1] = r_1 \cdot \{0, \infty\} = \{\infty, 0\}$. Now we compute $T_2[r_1]$:

$$T_2[r_1] = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} \{\infty, 0\} + \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} \{\infty, 0\} + \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix} \{\infty, 0\}.$$

Note that $\left(\begin{smallmatrix} 2 & 0 \\ 0 & 1 \end{smallmatrix}\right) \{\infty, 0\} = \{\infty, 0\} = [r_1]$, $\left(\begin{smallmatrix} 1 & 0 \\ 0 & 2 \end{smallmatrix}\right) \{\infty, 0\} = \{\infty, 0\} = [r_1]$ and $\left(\begin{smallmatrix} 1 & 1 \\ 0 & 2 \end{smallmatrix}\right) \{\infty, 0\} = \{\infty, \frac{1}{2}\} = \{\infty, 0\} + \{0, \frac{1}{2}\} = [r_1] + [r_3]$, since $[r_3] = r_3 \cdot \{0, \infty\} = \{0, \frac{1}{2}\}$. Using the relations 4.18.1, we see that $[r_3] = 0$ and thus

$$T_2[r_1] = 3[r_1].$$

Notice that this example is very easy and the computation of $T_2[r_1]$ is almost immediate but in more complicated cases one might need to use the continued fraction method described in the proof of Theorem 4.6 in order to compute the Hecke operator. We will see this in the next sections.

In conclusion, T_2 is the 1×1 matrix with coefficient 3. Its characteristic polynomial is $x - 3$ and its only eigenvalue is 3.

Computing the boundary map only involves looking at the cusps of $\Gamma_0(3)$. Using the formula of Remark 2.16 we see that there are only two cusps.

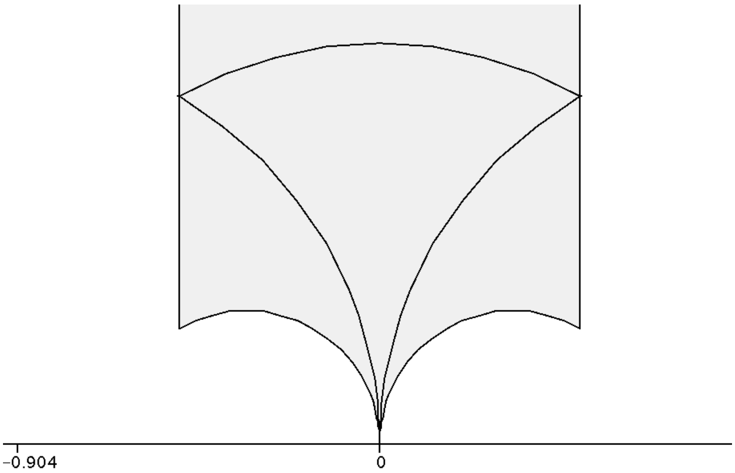

Figure 7: A fundamental domain for $\Gamma_0(3)$[2]

[2] Drawn with *Fundamental Domain Drawer* program developed by H.A. Verrill.
http://www.math.lsu.edu/~verrill/fundomain/index2.html accessed on 19/12/13.

This is confirmed by Figure 7. We may choose 0 and ∞ as a basis for $\mathbb{B}_2(3)$. We compute that $\delta([r_1]) = \{0\} - \{\infty\}$ and thus

$$\delta = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

The kernel of the boundary is the null space and thus $\mathbb{S}_2(3) = \emptyset$. Using Corollary 4.15, we see that $\dim(\mathcal{S}_2(\Gamma_0(3))) = 0$ as indicated in Table 5 and thus the Modularity Theorem is automatically verified for $N = 3$.

4.3.2 Illustration for $N = 11$

Recall from Proposition 2.7 that the matrices

$$r_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, r_1 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, r_2 = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}, r_3 = \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}, r_4 = \begin{pmatrix} 1 & 0 \\ 3 & 1 \end{pmatrix}, \ldots, r_{11} = \begin{pmatrix} 1 & 0 \\ 10 & 1 \end{pmatrix}$$

form a system of right coset representatives of $\Gamma_0(11)$ in $\mathrm{SL}_2(\mathbb{Z})$. Using the same technique as in the previous example we compute the relations on the corresponding Manin symbols. Expressing each relation as a vector in the basis $[r_0]', \ldots, [r_{11}]'$, this gives us the following relation matrix:

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

where again we omit the recurrent relations. Then

$$\begin{pmatrix}
1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & -1 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{pmatrix}$$

is the reduced echelon form of this matrix. Thus we obtain the equivalent relations:

$$\begin{array}{lllll} [r_0]' = -[r_1]' & [r_2]' = 0 & [r_3]' = [r_{10}]' & [r_4]' = -[r_9]' + [r_{10}]' & \\ [r_5]' = -[r_9]' & [r_6]' = -[r_{10}]' & [r_7] = -[r_{10}]' & [r_8]' = [r_9]' - [r_{10}]' & [r_{11}]' = 0. \end{array}$$

Hence, we see that every Manin symbol can be expressed as a linear combination of the symbols $[r_1]', [r_9]', [r_{10}]'$ and therefore M has dimension 3 and basis $\{[r_1]', [r_9]', [r_{10}]'\}$. Using Theorem 4.9 as before, $\mathbb{M}_2(11)$ has dimension 3 and basis $\{[r_1], [r_9], [r_{10}]\} = \{\{\infty, 0\}, \{0, \frac{1}{8}\}, \{0, \frac{1}{9}\}\}$.

We compute some Hecke operators on $\mathbb{M}_2(11)$. In order to compute T_p on $\mathbb{M}_2(11)$ we need to compute the image of each basis element and then we may express T_p as a matrix.

We start by computing the Hecke operator T_2 on $\mathbb{M}_2(11)$. Thus, we need to compute $T_2[r_1], T_2[r_9]$ and $T_2[r_{10}]$. The first calculation is similar to the one we did in Section 4.3.1 and we get

$$T_2[r_1] = 3[r_1] + [r_3] = 3[r_1] + [r_{10}]$$

.

We compute $T_2[r_9]$:

$$T_2[r_9] = T_2\left\{0, \frac{1}{8}\right\} = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}\left\{0, \frac{1}{8}\right\} + \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}\left\{0, \frac{1}{8}\right\} + \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}\left\{0, \frac{1}{8}\right\}.$$

We do the calculation step by step. For the first term of the sum, $\left(\begin{smallmatrix} 2 & 0 \\ 0 & 1 \end{smallmatrix}\right)\{0, 1/8\} = \{0, 1/4\} = \left(\begin{smallmatrix} 1 & 0 \\ 4 & 1 \end{smallmatrix}\right)\{0, \infty\} = [r_5] = -[r_9]$. The second term is $\left(\begin{smallmatrix} 1 & 0 \\ 0 & 2 \end{smallmatrix}\right)\{0, 1/8\} = \{0, 1/16\} = \left(\begin{smallmatrix} 1 & 0 \\ 16 & 1 \end{smallmatrix}\right)\{0, \infty\}$. We notice that $\left(\begin{smallmatrix} 1 & 0 \\ 16 & 1 \end{smallmatrix}\right) = \left(\begin{smallmatrix} 1 & 0 \\ 11 & 1 \end{smallmatrix}\right)\left(\begin{smallmatrix} 1 & 0 \\ 5 & 1 \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} 1 & 0 \\ 11 & 1 \end{smallmatrix}\right) \in \Gamma_0(11)$. Hence, $\left(\begin{smallmatrix} 1 & 0 \\ 0 & 2 \end{smallmatrix}\right)\{0, 1/8\} = [r_6] = -[r_{10}]$. For the third term, $\left(\begin{smallmatrix} 1 & 1 \\ 0 & 2 \end{smallmatrix}\right)\{0, 1/8\} = \{1/2, 9/16\} = \{0, 9/16\} - \{0, 1/2\} = \{0, 9/16\} - [r_3] = \{0, 9/16\} - [r_{10}]$. We use the continued fraction method illustrated in Example 4.7 to find that $\{0, 9/16\} = 2[r_{10}] - [r_9]$. Thus $\left(\begin{smallmatrix} 1 & 1 \\ 0 & 2 \end{smallmatrix}\right)\{0, 1/8\} = [r_{10}] - [r_9]$. Combining all these results, we finally get that

$$T_2[r_9] = -[r_9] - [r_{10}] + [r_{10}] - [r_9] = -2[r_9].$$

We compute $T_2[r_{10}]$ step by step like we did for $T_2[r_9]$. Computing the first term of the sum, we get $\left(\begin{smallmatrix} 2 & 0 \\ 0 & 1 \end{smallmatrix}\right)\{0, 1/9\} = \{0, 2/9\}$. Using the continued fraction method, we see that $\{0, 2/9\} = -[r_9] - [r_{10}]$. For the second term of the sum, $\left(\begin{smallmatrix} 1 & 0 \\ 0 & 2 \end{smallmatrix}\right)\{0, 1/9\} = \{0, 1/18\} = \left(\begin{smallmatrix} 1 & 0 \\ 18 & 1 \end{smallmatrix}\right)\{0, \infty\} = \left(\begin{smallmatrix} 1 & 0 \\ 11 & 1 \end{smallmatrix}\right)\left(\begin{smallmatrix} 1 & 0 \\ 7 & 1 \end{smallmatrix}\right)\{0, \infty\} = [r_8] = [r_9] - [r_{10}]$, since $\left(\begin{smallmatrix} 1 & 0 \\ 11 & 1 \end{smallmatrix}\right) \in \Gamma_0(11)$. Finally, for the third term, $\left(\begin{smallmatrix} 1 & 1 \\ 0 & 2 \end{smallmatrix}\right)\{0, 1/9\} = \{1/2, 5/9\} = \{0, 5/9\} - [r_3] = \{0, 5/9\} - [r_{10}]$. Using the continued fraction method, we see that $\{0, 5/9\} = [r_{10}]$ and thus the third term of the sum is 0. Combining the three terms gives us

$$T_2[r_{10}] = -[r_9] - [r_{10}] + [r_9] - [r_{10}] = -2[r_{10}].$$

Having computed the Hecke operator T_2 for all basis elements of $\mathbb{M}_2(N)$ we may now explicit T_2:

$$T_2 = \begin{pmatrix} 3 & 0 & 0 \\ 0 & -2 & 0 \\ 1 & 0 & -2 \end{pmatrix}.$$

The characteristic polynomial of T_2 is $-(x-3)(x+2)^2$ and its eigenvalues are 3 and -2.

Using the same method as above we compute T_3 on $\mathbb{M}_2(11)$:

$$T_3\{\alpha,\beta\} = \begin{pmatrix}3 & 0\\ 0 & 1\end{pmatrix}\{\alpha,\beta\} + \begin{pmatrix}1 & 0\\ 0 & 3\end{pmatrix}\{\alpha,\beta\} + \begin{pmatrix}1 & 1\\ 0 & 3\end{pmatrix}\{\alpha,\beta\} + \begin{pmatrix}1 & 2\\ 0 & 3\end{pmatrix}\{\alpha,\beta\}.$$

Clearly, $\left(\begin{smallmatrix}3 & 0\\ 0 & 1\end{smallmatrix}\right)[r_1] = \left(\begin{smallmatrix}1 & 0\\ 0 & 3\end{smallmatrix}\right)[r_1] = [r_1]$ and $\left(\begin{smallmatrix}1 & 1\\ 0 & 3\end{smallmatrix}\right)[r_1] = [r_1]+[r_4] = [r_1]-[r_9]+[r_{10}]$. For the last term, we obtain $\left(\begin{smallmatrix}1 & 2\\ 0 & 3\end{smallmatrix}\right)[r_1] = \{\infty, 2/3\} = [r_1] + \{0, 2/3\}$. Using the continued fraction method, we find that $\{0,2/3\} = [r_9]$ and thus $\left(\begin{smallmatrix}1 & 2\\ 0 & 3\end{smallmatrix}\right)[r_1] = [r_1]+[r_9]$. Combining the four terms, we get

$$T_3[r_1] = 4[r_1] + [r_{10}].$$

For the second basis element, compute that $\left(\begin{smallmatrix}3 & 0\\ 0 & 1\end{smallmatrix}\right)[r_9] = \{0, 3/8\}$. Using continued fractions, we see that $\{0,3/8\} = [r_0]+[r_1]+[r_3]+[r_5]+[r_{11}] = -[r_9]+[r_{10}]$. The second term is easily computed: $\left(\begin{smallmatrix}1 & 0\\ 0 & 3\end{smallmatrix}\right)[r_9] = \{0, 1/24\} = [r_3] = [r_{10}]$ by continued fractions. Next, $\left(\begin{smallmatrix}1 & 1\\ 0 & 3\end{smallmatrix}\right)[r_9] = \{1/3, 3/8\} = \{0,3/8\} - \{0,1/3\} = -[r_9]+[r_{10}]-[r_4] = 0$. Finally, $\left(\begin{smallmatrix}1 & 2\\ 0 & 3\end{smallmatrix}\right)[r_9] = \{2/3, 17/24\} = \{0, 17/24\} - [r_9]$ and by continued fractions we find that $\{0,17/24\} = [r_9]+2[r_7] = [r_9]-2[r_{10}]$. Thus $\left(\begin{smallmatrix}1 & 2\\ 0 & 3\end{smallmatrix}\right)[r_9] = -2[r_{10}]$. Combining the four terms, we obtain

$$T_3[r_9] = -[r_9].$$

For the last basis element, we immediately see that $\left(\begin{smallmatrix}3 & 0\\ 0 & 1\end{smallmatrix}\right)[r_{10}] = [r_4] = -[r_9]+[r_{10}]$. It is also easy to see that $\left(\begin{smallmatrix}1 & 0\\ 0 & 3\end{smallmatrix}\right)[r_{10}] = \{0, 1/27\} = \left(\begin{smallmatrix}1 & 0\\ 27 & 1\end{smallmatrix}\right)\{0,\infty\} = [r_6] = -[r_{10}]$. The third term becomes $\left(\begin{smallmatrix}1 & 1\\ 0 & 3\end{smallmatrix}\right)[r_{10}] = \{1/3, 10/27\} = \{0, 10/27\} - [r_4]$ and by the continued fraction method, $\{0, 10/27\} = [r_0]+[r_1]+[r_3]+[r_4]-[r_3]+[r_{11}]+[r_{10}] = [r_4]+[r_{10}]$. Thus $\left(\begin{smallmatrix}1 & 1\\ 0 & 3\end{smallmatrix}\right)[r_{10}] = [r_{10}]$. The last terms is $\left(\begin{smallmatrix}1 & 2\\ 0 & 3\end{smallmatrix}\right)[r_{10}] = \{2/3, 19/27\} = \{0, 19/27\} - [r_9]$. By the continued fraction method, $\{0, 19/27\} = [r_9]-[r_{10}]+[r_9]-[r_{10}]$. Thus $\left(\begin{smallmatrix}1 & 2\\ 0 & 3\end{smallmatrix}\right)[r_{10}] = [r_9]-2[r_{10}]$. Combining the four terms, we obtain

$$T_3[r_{10}] = -[r_{10}].$$

We may finally write

$$T_3 = \begin{pmatrix}4 & 0 & 0\\ 0 & -1 & 0\\ 1 & 0 & -1\end{pmatrix}.$$

The characteristic polynomial of T_3 is $-(x-4)(x+1)^2$ and its eigenvalues are 4 and -1.

We compute a basis for $\mathbb{S}_2(11)$. In order to do this we first need to compute the boundary map defined in Section 4.1.2 and then compute its kernel. Recall that $\Gamma_0(11)$ only has two cusps (see Figure 3). We choose as representatives for $C(\Gamma_0(11))$ the elements 0 and ∞ of $\mathbb{P}^1(\mathbb{Q})$. Thus the classes $\{0\}$ and $\{\infty\}$ form a basis of $\mathbb{B}_2(11)$. Recall that $\{[r_1],[r_9],[r_{10}]\}$ is a basis for the space

$\mathbb{M}_2(11)$ of modular symbols. We compute that $\delta([r_1]) = \delta(\{\infty, 0\}) = \{0\} - \{\infty\}$. If we apply the boundary map to $[r_9]$, we get $\delta([r_9]) = \delta(\{0, 1/8\}) = \{1/8\} - \{0\} = 0$. In fact, $\left(\begin{smallmatrix} -4 & 1 \\ -33 & 8 \end{smallmatrix}\right) \cdot 0 = 1/8$ and $\left(\begin{smallmatrix} -4 & 1 \\ -33 & 8 \end{smallmatrix}\right) \in \Gamma_0(11)$ and thus $\{1/8\} = \{0\}$. Similarly, $\delta([r_{10}]) = \delta(\{0, 1/9\}) = \{1/9\} - \{0\} = 0$, since $\left(\begin{smallmatrix} 5 & 1 \\ 44 & 9 \end{smallmatrix}\right) \cdot 0 = 1/9$ and $\left(\begin{smallmatrix} 5 & 1 \\ 44 & 9 \end{smallmatrix}\right) \in \Gamma_0(11)$. Hence the boundary map, with respect to our chosen bases is

$$\delta = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}.$$

Now that we have expressed δ as a matrix we can compute the kernel of δ, which is the space $\mathbb{S}_2(11)$ of cuspidal modular symbols, by solving the linear system

$$\begin{pmatrix} 1 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

where (a, b, c) with $a, b, c \in \mathbb{Q}$ is an element of $\mathbb{M}_2(11)$ expressed in the basis $\{[r_1], [r_9], [r_{10}]\}$. This equation is solved by all elements of the form $(0, b, c)$ and thus $\mathbb{S}_2(11)$ is a space of dimension 2 and with basis $\{[r_9], [r_{10}]\}$.

We compute some Hecke operators on $\mathbb{S}_2(11)$. We start by computing T_2. We already computed that $T_2[r_9] = -2[r_9]$ and $T_2[r_{10}] = -2[r_{10}]$ and thus T_2 on the space $\mathbb{S}_2(11)$ is the matrix

$$T_2 = \begin{pmatrix} -2 & 0 \\ 0 & -2 \end{pmatrix}.$$

We compute T_3 on $\mathbb{S}_2(11)$. We already computed that $T_3[r_9] = -[r_9]$ and $T_3[r_{10}] = -[r_{10}]$ and thus T_3 on $\mathbb{S}_2(11)$ is

$$T_3 = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Since $\dim(\mathbb{S}_2(11)) = 2$ we know from Corollary 4.15 that $\dim(\mathcal{S}_2(\Gamma_0(11))) = 1$ as indicated in Table 5. Remark **??** tells us that the Hecke forms f_1, f_2 defined by

$$f_1 = \sum_{n=1}^{\infty} \langle T_n[r_9], [r_9] \rangle q^n \qquad f_2 = \sum_{n=1}^{\infty} \langle T_n[r_{10}], [r_{10}] \rangle q^n$$

are candidates for a basis of $\mathcal{S}_2(\Gamma_0(11))$. In this case $\langle T_n[r_9], [r_9] \rangle = \langle T_n[r_{10}], [r_{10}] \rangle$ and thus a basis for $\mathcal{S}_2(\Gamma_0(11))$ is given by the form

$$f = \sum_{n=1}^{\infty} b_n q^n,$$

where $b_n = \langle T_n[r_9], [r_9] \rangle$. We compute the first coefficients. We must have $b_1 = 1$ since f is a Hecke form. The coefficient b_2 is the eigenvalue of T_2 and thus $b_2 = -2$. Similarly, $b_3 = -1$. It is clear that since the eigenvalues of T_n are the same on $\mathbb{S}_2(11)$ and $\mathcal{S}_2(\Gamma_0(11))$, then the relations in Corollary

2.56 holds for the eigenvalues of T_n on $\mathbb{S}_2(11)$. In particular, $b_4 = b_2^2 - 2b_1 = 4 - 2 = 2$. We may then write

$$f(z) = q - 2q^2 - q^3 + 2q^4 + \cdots$$

In Example 1.53, we considered the elliptic curve of conductor 11 given by

$$E : y^2 - y = x^3 - x^2$$

and computed its L-series up to the eleventh term :

$$L(E, s) = \sum_{n=1}^{\infty} \frac{a_n}{n^s} = 1 - \frac{2}{2^s} - \frac{1}{3^s} + \frac{2}{4^s} + \frac{1}{5^s} + \frac{2}{6^s} - \frac{2}{7^s} - \frac{2}{9^s} - \frac{2}{10^s} + \frac{1}{11^s} + \cdots .$$

The Modularity Theorem asserts that $g := \sum_{n=1}^{\infty} a_n q^n \in \mathcal{S}_2(11)$. But this is a space of one dimension and with basis the Hecke form f. Thus we must have $g = \lambda f$ for some scalar λ. But noticing that $b_1 = a_1$ we may identify $\lambda = 1$. We have verified the theorem for $N = 11$.

Notice that for $p \neq 11$ prime we have $b_p = p+1-\#E(\mathbb{F}_p)$. We can therefore use Hecke operators on modular symbols to compute the number of points of the reduced elliptic curve E modulo p for any p different from 11. We illustrate this for $p = 101$. Using SAGE, we compute T_{101} on $\mathbb{S}_2(11)$:

$$T_{101} = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}.$$

Thus $2 = 101 + 1 - \#E(\mathbb{F}_{101})$ and therefore $\#E(\mathbb{F}_{101}) = 100$.

4.3.3 Illustration for $N = 35$

Consider the elliptic curve E defined over $\mathbb{Q}$ by $E : y^2+y = x^3+x^2-x$. Using the notations of Section 1.3, we identify $a_1 = 0, a_2 = 1, a_3 = 1, a_4 = -1$ and $a_6 = 0$. We compute that $b_2 = 4, b_4 = -2$ and $b_6 = 1$. Thus

$$c_4 = 64 \text{ and } c_6 = -568.$$

Using Definition 1.17, the discriminant of E is $\Delta = -35$. Therefore, if $p \neq 5, 7$ is prime, then $|\Delta|_p = 1$ and thus the equation of E is p-minimal. Moreover, note that $|\Delta|_5 = 5^{-1} > 5^{-12}$ and $|\Delta|_7 = 7^{-1} > 7^{-12}$ hence, by Lemma 1.28, the equation of E is 5-minimal and 7-minimal. Since the coefficients of the equation are all integers, we may state that the equation of E is globally minimal. We can therefore use this equation to study the reduction modulo p of E.

We compute the conductor of E. If $p \neq 5, 7$ is prime, then E admits good reduction at p and thus f_p from Definition 1.36 is equal to zero. If $p = 5$, then $\Delta = 0$ and $c_4 \neq 0$ modulo 5. Hence, the reduction modulo 5 of the elliptic curve E admits a node and by Definition 1.24 the reduction of E at 5 is multiplicative. Therefore $f_5 = 1$. If $p = 7$, then $\Delta = 0$ and $c_4 \neq 0$ modulo 7. Hence, the reduction modulo 7 of the elliptic curve E admits a node and again the reduction of E at 7 is multiplicative. Therefore $f_7 = 1$. Hence the conductor of E is $N_E = 35$.

We now study the type of multiplicative reduction of the curve E at $p = 5$. Let E_5 denote the reduced curve. We have $E_5 : f(x, y) = 0$ with $f(x, y) = y^2 + y - x^3 - x^2 + x$. We look for the singular point $P = (x_0, y_0)$ of E_5 by solving the equation $\text{grad}(f)(x_0, y_0) = (0, 0)$. Computing the gradient in $\mathbb{F}_5$, we get

$$\text{grad}(f)(x, y) = (-3x^2 - 2x + 1, 2y + 1).$$

Thus finding the point P is equivalent to solving the system

$$\begin{cases} 2x_0^2 + 3x_0 + 1 = 0 \\ 2y_0 + 1 = 0. \end{cases}$$

We see that $y_0 = -2^{-1}$, where this is the inverse of 2 in $\mathbb{F}_5^*$. Thus $y_0 = 2$. Solving the equation in x_0 we get $x_0 = 4$ or $x_0 = 2$. Thus $P = (2, 2)$ or $P = (4, 2)$. We verify that $f(2, 2) = 1$ and $f(4, 2) = 0$ and therefore the only possibility is $P = (4, 2)$. Following Section 1.3.3, the third order Taylor series of f at a neighborhood of P is

$$f(x, y) = -3(x - 4)^2 + (y - 2)^2 - (x - 4)^3 = [(y - 2) - \sqrt{3}(x - 4)][(y - 2) + \sqrt{3}(x - 4)] - (x - 4)^3.$$

Thus the slopes of the tangent lines at P are $\sqrt{3}$ and $-\sqrt{3}$. Since $\sqrt{3} \notin \mathbb{F}_5$, Definition 1.24 tells us that the reduction of E at 5 is non-split multiplicative.

We study the type of multiplicative reduction of the curve E at $p = 7$. Let E_7 denote this time the reduced curve modulo 7. We have $E_7 : f(x, y) = 0$ with $f(x, y) = y^2 + y - x^3 - x^2 + x$. We look for the singular point $P = (x_0, y_0)$ and therefore compute the gradient of f in $\mathbb{F}_7$:

$$\text{grad}(f)(x, y) = (4x^2 + 5x + 1, 2y + 1).$$

Finding the point P requires solving the system

$$\begin{cases} 4x_0^2 + 5x_0 + 1 = 0 \\ 2y_0 + 1 = 0. \end{cases}$$

We see that $y_0 = -2^{-1} = 3$ and solving the equation in x_0, we find that $x_0 = 6$ or $x_0 = 5$. Thus $P = (6, 3)$, or $P = (5, 3)$. But $f(6, 3) = 4$ and $f(5, 3) = 0$ and therefore the only possibility is $P = (5, 3)$. Following Section 1.3.3, the third order Taylor series of f at a neighborhood of P is

$$f(x, y) = -2(x - 5)^2 + (y - 3)^2 - (x - 5)^3 = [(y - 3) - \sqrt{2}(x - 5)][(y - 3) + \sqrt{2}(x - 5)] - (x - 5)^3.$$

Hence the slopes of the tangent lines at P are $\sqrt{2} = 3$ and $-\sqrt{2} = 4$, which are both elements of $\mathbb{F}_7$. We may then conclude that the reduction of E at 7 is split multiplicative.

We now compute the L-series of E. A quick calculation shows that $\#E(\mathbb{F}_2) = 3$ and $\#E(\mathbb{F}_3) = 3$. Thus, referring to Definition 1.45, the coefficients 2 and 3 of the L-series are $a_2 = 0$ and $a_3 = 1$. Thus the L-series of E is given by

$$L(E, s) = 1 + \frac{0}{2^s} + \frac{1}{3^s} + \cdots$$

The Modularity Theorem states that the function

$$f = \sum_{n=1}^{\infty} a_n q^n = q + 0q^2 + q^3 + ???$$

is a Hecke form for $\Gamma_0(N_E) = \Gamma_0(35)$.

We now compute a basis for $\mathcal{S}_2(\Gamma_0(35))$ constituted by Hecke forms using modular symbols. The goal is to express the function f in this basis and we will then have verified the statement of the theorem.

Using SAGE we compute a basis for $\mathcal{S}_2(\Gamma_0(35))$ to precision $\mathcal{O}(q^{45})$. See [St07] Section 3.6 for the commands to use. We find that

$$\begin{aligned}
f_1 = q &- 2q^6 - 2q^8 - q^{11} + 4q^{13} - q^{15} + 2q^{16} + 2q^{18} + 2q^{20} + q^{21} + 2q^{22} - 4q^{23} + q^{25} + 2q^{26} - 4q^{27} \\
&- 2q^{28} + q^{29} - 2q^{30} - 2q^{31} - 2q^{32} - 4q^{33} - 2q^{34} - q^{35} + 2q^{36} + 4q^{37} + 4q^{38} - q^{39} - 2q^{40} - 6q^{41} \\
&+ 2q^{42} - 2q^{43} + 2q^{44} + \mathcal{O}(q^{45}) \\
f_2 = q^2 &- 3q^4 - q^5 + 2q^6 + q^7 + 3q^8 - q^9 + q^{10} - q^{11} - 4q^{12} + 2q^{13} - q^{14} - q^{15} - q^{16} + 2q^{17} - q^{18} \\
&+ 4q^{19} - q^{20} + q^{21} - 2q^{22} - 4q^{23} + 4q^{24} + q^{28} - q^{29} + 2q^{30} - 2q^{31} + 3q^{32} + 3q^{36} - 2q^{37} - 8q^{38} \\
&+ 3q^{39} + 3q^{40} - 8q^{41} - 2q^{42} - 6q^{43} + 6q^{44} + \mathcal{O}(q^{45}) \\
f_3 = q^3 &- 2q^4 - q^5 + 2q^6 + q^7 + 2q^8 - 2q^9 - 2q^{11} - 2q^{12} + q^{13} + 2q^{16} + 3q^{17} - 2q^{18} + 2q^{19} - 2q^{22} \\
&- 2q^{23} - 2q^{26} - q^{27} + 2q^{29} + 2q^{30} - 2q^{31} + 2q^{32} + q^{33} + 2q^{34} + 2q^{36} - 2q^{37} - 4q^{38} + 6q^{39} \\
&+ 2q^{40} - 6q^{41} - 2q^{42} - 8q^{43} + 4q^{44} + \mathcal{O}(q^{45})
\end{aligned}$$

form a basis of $\mathcal{S}_2(\Gamma_0(35))$. Using Table 5, we see that indeed $\dim \mathcal{S}_2(\Gamma_0(35)) = 3$.

We want to express f in this basis. We write $f = \lambda_1 f_1 + \lambda_2 f_2 + \lambda_3 f_3$. Looking at the coefficients of f, we see that

$$1 = 1\lambda_1 + 0\lambda_2 + 0\lambda_3 \qquad 0 = 0\lambda_1 + 1\lambda_2 + 0\lambda_3 \qquad 1 = 0\lambda_1 + 0\lambda_2 + 1\lambda_3.$$

Hence $\lambda_1 = 1, \lambda_2 = 0$ and $\lambda_3 = 1$ and $f = f_1 + f_3$. We conclude that f is indeed a Hecke form for $\Gamma_0(35)$.

With this information we can compute the unknown coefficients of f. For example, $a_5 = 1.0 + 1.(-1) = -1$. Our previous computations for $p = 5$ show that the reduction of E at 5 is non-split multiplicative and Definition 1.45 tells us that indeed $a_5 = -1$. We may now express f to precision $\mathcal{O}(q^{45})$:

$$\begin{aligned}
f = q &+ q^3 - 2q^4 - q^5 + q^7 - 2q^9 - 3q^{11} - 2q^{12} + 5q^{13} - q^{15} + 4q^{16} + 3q^{17} + 2q^{19} + 2q^{20} + q^{21} \\
&- 6q^{23} + q^{25} - 5q^{27} - 2q^{28} + 3q^{29} - 4q^{31} - 3q^{33} - q^{35} + 4q36 + 2q^{37} + 5q^{39} - 12q^{41} \\
&- 10q^{43} + 6q^{44} + \mathcal{O}(q^{45}).
\end{aligned}$$

For example, $a_{43} = -10$. Definition 1.45 tells us that $a_{43} = 43 + 1 - \#E(\mathbb{F}_{43})$ and thus

$$\#E(\mathbb{F}_{43}) = 54.$$

References

[Ca94] M.P. do Carmo, *Differential forms and applications.* Springer-Verlag, 1994.

[DS05] F. Diamond, J. Shurman, *A first course in modular forms.* Springer, 2005.

[Ha36] H. Hasse, *Zur Theorie der abstrakten elliptischen Funktionenkörper I, II & III.* Crelle's Journal **175**, 1936.

[He02] Y. Hellegouarch, *Invitation au mathématiques de Fermat-Wiles.* Dunod, 2002.

[Ko93] N. Koblitz, *Introduction to elliptic curves and modular forms.* Springer, 1993.

[Ma72] J.I. Manin, *Parabolic points and zeta functions of modular curves.* Ivz. Akad. Nauk SSSR Ser. Mat. **36**, 1972.

[Og67] A-P. Ogg, *Abelian curves of small conductor.* J. Reine Angew. Math. **226** (1967) 204 – 215.

[Se70] J-P. Serre, *Cours d'arithmétique.* Presses universitaires de France, 1970.

[Si86] J.H. Silverman, *The arithmetic of elliptic curves.* Springer, 1986.

[St07] W. Stein, *Modular forms, a computational approach.* Graduate studies in Mathematics, 2007.