MASTER'S THESIS

Master of Science in Advanced Mathematics and Mathematical Engineering

Title: Modular curves and complex multiplication

Author:Víctor Hernández Barrios

Advisor: Victor Rotger Cerdà

Department of Mathematics

Academic year: 2017-2018

UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH
Facultat de Matemàtiques i Estadística

Abstract

In this project we explore the connections between elliptic curves, modular curves and complex multiplication (CM). The main theorem of CM shows that the theory of CM for elliptic curves provides an explicit construction of finite abelian extensions of a quadratic imaginary field. The proof we discuss uses many of the properties of the classical modular curve, which is introduced both as a geometrical object and as a moduli space. This theory together with the Modularity theorem is used in the construction of Heegner points, which are in the heart of the proof of Kolyvagin's theorem, a result closely related with the Birch and Swinnerton-Dyer conjecture.

Acknowledgements

I really enjoyed doing this project with Victor Rotger. I thank him for his time and dedication, as well as for his valuable suggestions. He has positively changed my vision on number theory and its current research, and I am sure this work would not be the same without his guidance.
I thank Santiago Molina for clarifying some of the doubts that arose during the realization of the project.
I also express my gratitude to my family for his unconditional support.

Contents

Introduction

This project exposes some of the ideas and connections between elliptic curves, modular curves and complex multiplication. In section 1 we recall some results on number fields which will be required in the following sections.
Complex multiplication is a property some elliptic curves have, discussed in section 2. An elliptic curve E can be defined to be the zero locus of a degree 3 polynomial in two variables having no singularities. Drawing an elliptic curve E on a real plane has its limitations, since the picture we obtain depends on the particular equation E verifies. If we allow complex numbers and instead embed it in the complex projective plane then E is now isomorphic to a complex torus, a genus 1 compact surface. Moreover, one can define through a geometrical construction a law group on E making it an abelian group. In abelian groups one can consider the multiplication-by-n endomorphisms defined by $P \mapsto nP$. In the case of most elliptic curves, there are no other endomorphisms i.e. the endomorphism ring $\mathrm{End}(E)$ is isomorphic to $\mathbb{Z}$. Then E has complex multiplication (CM) if $\mathrm{End}(E)$ is strictly larger than $\mathbb{Z}$. The ring $\mathrm{End}(E)$ is then also an special kind of ring i.e. an order $\mathcal{O}$ of a quadratic imaginary field K. The ring of integers $\mathcal{O}_K$ of K is the maximal order i.e. it contains all the others. In particular, an order $\mathcal{O}$ is a lattice in $\mathbb{C}$ and so it has a fundamental parallelogram $P_{\mathcal{O}}$. The conductor of an order $\mathcal{O}$ is then the ratio $f = f(\mathcal{O}) = A(P_{\mathcal{O}_K})/A(P_{\mathcal{O}})$ where $A(\cdot)$ denotes area. In the case of quadratic imaginary fields, the conductor determines completely the order. There is a quantity attached to every elliptic curve that tells us whether two curves are $\mathbb{C}$-isomorphic or not, the j-invariant. In the case that E has CM its j-invariant is an algebraic number over $\mathbb{Q}$. Every order $\mathcal{O}$ has an associated proper ideal class group $\mathrm{cl}(\mathcal{O})$ and a ring class field K_f, and both are constructed in terms of elliptic curves having CM by $\mathcal{O}$. The main theorem of CM shows that the Galois group of K_f over K is $\mathrm{cl}(\mathcal{O})$. Since $\mathrm{cl}(\mathcal{O})$ is abelian this is an explicit construction of infinitely many finite abelian extensions of K. This solves partially - for the quadratic imaginary case - the main problem in class field theory, which is finding all the finite abelian extensions of a number field.
All of this requires showing that the j-invariants of elliptic curves with CM are algebraic over $\mathbb{Q}$. This last property is proven in section 3 by using the fact that the modular curve $X_0(N)$ can be defined over $\mathbb{Q}$ and an alternative interpretation of $X_0(N)$ as a moduli space i.e. the curve $X_0(N)$ classifies the degree N isogenies between elliptic curves. However, the first definition of $X_0(N)$ is geometrical. In the same way a complex torus $T = \mathbb{C}/\Lambda$ is the quotient space when Λ acts on the plane as a discrete group of isometries, the modular curve is also defined in geometrical terms[1]. It is the quotient space $X_0(N) = \mathfrak{H}^*/\Gamma_0(N)$, where $\Gamma_0(N)$ is a carefully chosen finite index subgroup of $\mathrm{SL}_2(\mathbb{Z})$.
Finally, in section 4 we discuss the construction of Heegner points, which we can think of as an application of the theory of complex multiplication. Together with the Modularity theorem, Kolyvagin, Gross and Zagier gave results closely related to the Birch and Swinnerton-Dyer conjecture, an open problem concerning elliptic curves.

[1] $\mathrm{SL}_2(\mathbb{Z})$ is also a discrete subgroup of isometries of the upper half plane $\mathfrak{H}$ with the Poincaré metric.

1 Number fields

A *number field* is a finite extension $K \subset \mathbb{C}$ of $\mathbb{Q}$. The *algebraic integers* $\mathcal{O}_K$ of K are the elements α of K for which there exists $f \in \mathbb{Z}[x]$ monic with $f(\alpha) = 0$. They form a subring of K and the field of fractions of $\mathcal{O}_K$ is K itself. Furthermore, $\mathcal{O}_K$ is a *Dedekind domain* i.e.

- $\mathcal{O}_K$ is Noetherian
- Every prime ideal $0 \neq \mathfrak{p} \subset \mathcal{O}_K$ is maximal
- $\mathcal{O}_K$ is integrally closed

In a Dedekind domain, every ideal $\mathfrak{a}$ of $\mathcal{O}_K$ factors uniquely up to reordering as a product of prime ideals

$$\mathfrak{a} = \mathfrak{p}_1^{e_1} \cdots \mathfrak{p}_r^{e_r}$$

A *fractional ideal* of K is a finitely generated $\mathcal{O}_K$-submodule of K, and it can be regarded as a generalization of ideal of $\mathcal{O}_K$. The set of fractional ideals $\mathcal{I}(K)$ forms a group under multiplication, and it is freely generated by the prime ideals of K. Note that the subset of principal fractional ideals $\mathcal{P}(K)$ forms a subgroup of $\mathcal{I}(K)$.
The *ideal class group* of K is the abelian group $\mathrm{cl}(K) = \mathcal{I}(K)/\mathcal{P}(K)$. The cardinal of $h_K = |\mathrm{cl}(K)|$ is called the *class number* of K, and it is always finite.
Another result is *Dirichlet's unit theorem* which states that the group of units $\mathcal{O}_K^\times$ of $\mathcal{O}_K$ verifies $\mathcal{O}_K^\times \simeq \mu(K) \times \mathbb{Z}^{r+s-1}$ where r and $2s$ are respectively the number of real and complex embeddings of K in $\mathbb{C}$ and $\mu(K)$ is the *torsion* of $\mathcal{O}_K^\times$, the roots of unity contained in $\mathcal{O}_K$.
Let L/K be a finite extension and let $\alpha \in L$, and let M_α be a matrix of the K-linear map $m_\alpha : L \to L$ in some basis of L/K. The *norm* $\mathrm{N}_{L/K}(\alpha)$ and *trace* $\mathrm{Tr}_{L/K}(\alpha)$ of α is the norm and trace of M_α. An *integral basis* of $\mathcal{O}_L$ is a set $\{\beta_i\}_i$ that generates $\mathcal{O}_L = \langle \beta_i \rangle_\mathbb{Z}$ as a $\mathbb{Z}$-module. The *discriminant* of L/K with respect to β_i is the determinant of the matrix $\{\mathrm{Tr}_{L/K}(\beta_i\beta_j)\}_{i,j}$. As it depends on a choice of basis, it is defined in general up to elements in $(\mathcal{O}_K^\times)^2$.
In this section, we consulted [6, 5, 1].

1.1 Ramification

Let L/K be a finite extension of K and let $\mathfrak{p}$ be a prime ideal of $\mathcal{O}_K$. One can consider the factorization of $\mathfrak{p}\mathcal{O}_L$ in $\mathcal{O}_L$

$$\mathfrak{p}\mathcal{O}_L = \mathfrak{P}_1^{e_1} \cdots \mathfrak{P}_g^{e_g} \tag{1.1}$$

One says that $\mathfrak{P}_i$ *lies* above $\mathfrak{p}$. The coefficients e_i in (1.1) are the *ramification indices*. The *inertial degrees* are $f_i = [\mathbb{F}_{\mathfrak{P}_i}^L : \mathbb{F}_\mathfrak{p}^K]$. These make sense since $\mathbb{F}_\mathfrak{P}^L = \mathcal{O}_L/\mathfrak{P}_i$ can be regarded as a finite Galois extension of $\mathbb{F}_\mathfrak{p}^K = \mathcal{O}_K/\mathfrak{p}$, because these fields are finite and $\mathfrak{p} \subset \mathfrak{P}$.
We say that a prime ideal $\mathfrak{p}$ of K

- *splits* in L if $e_i = 1$ for all i
- *splits completely* in L if it splits and $f_i = 1$ for all i
- *ramifies* in L if it does not split and it is not a prime of K
- is *inert* in L if $\mathfrak{p}\mathcal{O}_L$ is a prime ideal of L

It can be shown that if $\mathfrak{p}$ ramifies then $\mathfrak{p}$ divides the discriminant of L/K. This proves that only finitely many primes of K *ramify* in L. The following relation

$$\sum_{i=1}^{g} e_i f_i = [L : K] \tag{1.2}$$

tells us there is a finite number of possibilities for the e_i and f_i. It is sometimes called the *fundamental identity* for its role in many proofs of the results we discuss.

1.1.1 The Galois case

If L/K is Galois then the situation is simplified, because if we let $G = \mathrm{Gal}(L/K)$ then G acts on the set of prime ideals of $\mathcal{O}_L$ because $\sigma\mathcal{O}_L = \mathcal{O}_L$ for $\sigma \in G$. Moreover, if we fix $\mathfrak{p}$ and choose $\mathfrak{P}$ above $\mathfrak{p}$ the primes $\mathfrak{P}_i$ in the factorization of $\mathfrak{p}\mathcal{O}_L$ all lie in the G-orbit of $\mathfrak{P}$, so G acts transitively on the $\mathfrak{P}_i$. From this fact one can prove that $e = e_i$ and $f = f_i$ for all i, thus

$$efg = [L : K]$$

The inertial and ramification degrees are multiplicative: $e(M/K) = e(M/L) \cdot e(L/K)$ and $f(M/K) = f(M/L)f(L/K)$ for an extension M/L.
Some subgroups of G are studied to derive further properties of the decomposition of $\mathfrak{p}$. The *decomposition group* $G_{\mathfrak{P}} = G(\mathfrak{P}|\mathfrak{p})$ is the stabilizer[2] of $\mathfrak{P}$:

$$G_{\mathfrak{P}} = \{\sigma \in G : \sigma\mathfrak{P} = \mathfrak{P}\}$$

The *inertia group* $I_{\mathfrak{P}} = I(\mathfrak{P}|\mathfrak{p})$ is

$$I_{\mathfrak{P}} = \{\sigma \in G : \sigma x \equiv x \bmod \mathfrak{P} \text{ for all } x \in L\}$$

It follows that $I_{\mathfrak{P}} \subset G_{\mathfrak{P}}$, and each $\sigma \in G_{\mathfrak{P}}$ naturally induces an automorphism $\mathcal{Q}(\sigma) = \overline{\sigma} \in \overline{G} = \mathrm{Gal}(\mathbb{F}_{\mathfrak{P}}^L/\mathbb{F}_{\mathfrak{p}}^K)$ that makes the following diagram[3] commute:

$$\begin{array}{ccc} \mathcal{O}_L & \xrightarrow{\sigma} & \mathcal{O}_L \\ \downarrow{\scriptstyle \pi_{\mathfrak{P}}} & & \downarrow{\scriptstyle \pi_{\mathfrak{P}}} \\ \mathbb{F}_{\mathfrak{P}}^L & \xrightarrow{\overline{\sigma}} & \mathbb{F}_{\mathfrak{P}}^L \end{array}$$

[2] This makes sense since G acts on $\mathcal{O}_L$
[3] Here $\pi_{\mathfrak{P}}$ is the natural quotient by $\mathfrak{P}$ map

The map $\mathcal{Q} : G_{\mathfrak{P}} \to \overline{G}$ is a group homomorphism, and $\ker \mathcal{Q} = I_{\mathfrak{P}}$ because $\mathcal{Q}(\sigma) = 1$ if and only if for any x one has $\overline{\sigma}\pi_{\mathfrak{P}}x = \pi_{\mathfrak{P}}x = \pi_{\mathfrak{P}}\sigma x$, and this is to say $\sigma \in I_{\mathfrak{P}}$. The fixed subfields of L by the decomposition and inertia groups are called the *decomposition field* $Z_{\mathfrak{P}}$ and the *inertia field* $T_{\mathfrak{P}}$. Therefore $T_{\mathfrak{P}}/Z_{\mathfrak{P}}$ is a normal extension because $\ker \mathcal{Q} = I_{\mathfrak{P}} \triangleleft G_{\mathfrak{P}}$ is a normal subgroup. It can be shown that $\mathcal{Q}$ is surjective, so we have an exact sequence

$$1 \longrightarrow I_{\mathfrak{P}} \longrightarrow G_{\mathfrak{P}} \longrightarrow \overline{G} \longrightarrow 1$$

The reason of introducing the decomposition and inertia fields will become clear now in the following diagram

$$\underset{\text{inertial degree}}{\overset{\text{ramification index}}{}} \mathfrak{p} \overset{1}{\underset{1}{-\!\!-}} \mathfrak{P}_Z \overset{1}{\underset{f}{-\!\!-}} \mathfrak{P}_T \overset{e}{\underset{1}{-\!\!-}} \mathfrak{P}$$

Let $\mathfrak{P}$ be a prime of L above $\mathfrak{p}$ and consider the factorization of $\mathfrak{p}$ in the tower of extensions $K \subset Z_{\mathfrak{P}} \subset T_{\mathfrak{P}} \subset L$. Then $\mathfrak{p}$ splits completely in the decomposition field, into primes which remain inert in the inertia field, which finally can ramify in L.
Note that $|G_{\mathfrak{P}}| = ef$ since G acts transitively on the set of primes $\mathfrak{P}$ above $\mathfrak{p}$; the number of factors in the factorization (1.1) is $g = (G : G_{\mathfrak{P}})$, and $|G| = efg = [L : K]$.
In the case L/K is an *abelian extension*, i.e. that G is abelian then the decomposition and inertia groups $G_{\mathfrak{P}}$ and $I_{\mathfrak{P}}$ only depend on $\mathfrak{p}$ and not on $\mathfrak{P}$, because in general

$$\begin{cases} G_{\sigma\mathfrak{P}} = \sigma G_{\mathfrak{P}} \sigma^{-1} \\ I_{\sigma\mathfrak{P}} = \sigma I_{\mathfrak{P}} \sigma^{-1} \end{cases}$$

and all the subgroups of G are normal. Then we write $G_{\mathfrak{P}} = G_{\mathfrak{p}}$ and $I_{\mathfrak{P}} = I_{\mathfrak{p}}$.

1.2 The Artin symbol

We recall that $\overline{G}$ is the Galois group of a finite field, so it is cyclic and generated by a power of the *Frobenius automorphism* φ which maps $\overline{x} \mapsto \overline{x}^p$, namely by $\overline{\phi} = \varphi^{\mathrm{N}\mathfrak{p}}$ where $\mathrm{N}\mathfrak{p} = (\mathcal{O}_K : \mathfrak{p})$ is the *numerical norm* of $\mathfrak{p}$ and $(p) = \mathfrak{p} \cap \mathbb{Z}$. Now consider a prime $\mathfrak{p}$ of K that is *unramified* in L ($e = 1$). It can be shown that $|I_{\mathfrak{P}}| = e$, so in this case the inertia group $I_{\mathfrak{P}}$ is trivial and $\mathcal{Q}$ induces an isomorphism

$$G_{\mathfrak{P}} \simeq \overline{G} = \mathrm{Gal}(\mathbb{F}_{\mathfrak{P}}^L / \mathbb{F}_{\mathfrak{p}}^K) = \langle \overline{\phi} \rangle$$

Taking the preimage $\phi = \mathcal{Q}^{-1}(\overline{\phi})$ of $\overline{\phi}$ by $\mathcal{Q}$ gives the generator of $G_{\mathfrak{P}}$. By the definition of $\mathcal{Q}$, we see that ϕ verifies

$$\phi(x) \equiv x^{\mathrm{N}\mathfrak{p}} \bmod \mathfrak{P} \tag{1.3}$$

because in $\overline{G}$ we had $\overline{\phi}(\overline{x}) = \overline{x}^{\mathrm{N}\mathfrak{p}}$. Note ϕ is the only element in $G_{\mathfrak{P}}$ verifying this condition, since $\mathcal{Q}$ is an isomorphism. In this case, the *Artin symbol* is

$$\left(\frac{L/K}{\mathfrak{P}}\right) := \phi$$

The notation remarks that ϕ depends on L, K and $\mathfrak{P}$. The order of $(\frac{L/K}{\mathfrak{P}})$ is the inertial degree $f = |\overline{G}|$, since $\overline{G}$ is generated by the Frobenius element.
From the uniqueness of an element verifying (1.3) it follows that for any $\sigma \in G$

$$\left(\frac{L/K}{\sigma\mathfrak{P}}\right) = \sigma\left(\frac{L/K}{\mathfrak{P}}\right)\sigma^{-1} \tag{1.4}$$

i.e. the Artin symbol of the conjugate by σ of a prime $\mathfrak{P}$ is the conjugate of the Artin symbol by σ.
If L/K is abelian then conjugation by σ is the identity map in G, so $(\frac{L/K}{\mathfrak{P}})$ does not really depend on what prime $\mathfrak{P}$ above $\mathfrak{p}$ we choose. In this case, given a prime $\mathfrak{p}$ of K we define

$$\left(\frac{L/K}{\mathfrak{p}}\right) := \left(\frac{L/K}{\mathfrak{P}}\right) \tag{1.5}$$

where the right hand side is the Artin symbol and $\mathfrak{P}$ is some prime in L above $\mathfrak{p}$. From our discussion above it follows that is well-defined.
When L/K is not abelian, by taking all the elements in the conjugacy class of $(L/K, \mathfrak{P})$ we define similarly

$$\left(\frac{L/K}{\mathfrak{p}}\right) := \left\{\left(\frac{L/K}{\mathfrak{P}}\right) \text{ with } \mathfrak{P} \text{ above } \mathfrak{p}\right\} \tag{1.6}$$

Remark 1.1. The Artin symbol also characterizes those unramified primes splitting completely; if $\mathfrak{p}$ is unramified and $(L/K, \mathfrak{p}) = 1$ then $\mathfrak{p}$ splits completely, since the hypotheses imply $e = f = 1$ and $efg = g = [L : K]$.

1.3 Artin reciprocity law

We were able to define the Artin symbol for $\mathfrak{p}$ unramified in L, because $\mathcal{Q}$ was an isomorphism. If one wants to define it for all prime ideals, then one can consider the case of an abelian *unramified extension* L/K; that all primes in K are unramified in L. Briefly, the commutativity of G lets one define the Artin symbol for a prime $\mathfrak{p}$ of K as in (1.5) and the condition that L/K is unramified is imposed so that it can be defined for all primes. Let $\mathcal{I}(K)$ be the group of fractional ideals of K. One can define the Artin symbol for any fractional ideal

$$\begin{array}{rcl} \left(\frac{L/K}{\cdot}\right) : \mathcal{I}(K) & \longrightarrow & \mathrm{Gal}(L/K) \\ \mathfrak{a} = \prod_i \mathfrak{p}_i^{r_i} & \longmapsto & \prod_i \left(\frac{L/K}{\mathfrak{p}_i}\right)^{r_i} \end{array}$$

Note that the expression above makes sense because $\mathrm{Gal}(L/K)$ is abelian, hence we may take the product of the $\left(\mathfrak{p}_i, L/K\right)$ in any order. The *Hilbert class field* of K is the maximal abelian unramified extension of K. In the case of $K = \mathbb{Q}$ the Hilbert class field is $\mathbb{Q}$ itself, because every extension of $\mathbb{Q}$ is ramified.

Theorem 1.2 (Artin reciprocity law for the Hilbert class field)**.** *Let L be the Hilbert class field of K. Then the map $\left(\frac{L/K}{\cdot}\right)$ described above is surjective and the kernel is $\mathcal{P}(K)$.*

Hence, to find all the abelian unramified extensions of K one has only to find $\mathrm{Gal}(L/K)$ because of the Galois correspondence i.e. it follows that the Artin symbol induces an isomorphism $\mathrm{cl}(K) \cong \mathrm{Gal}(L/K)$, and by Galois theory we obtain

Corollary 1.3. *There is a one-to-one correspondence between the subgroups of* $\mathrm{cl}(K)$ *and the abelian unramified extensions of K.* [4]

In the case of K being quadratic imaginary, by using some results on elliptic curves the Galois group is completely determined. This is the goal of section 2.
Theorem 1.2 is a particular case of the following result, for which one does not need the extension L/K to be unramified:

Theorem 1.4 (Artin reciprocity)**.** *Let L/K be a finite abelian extension of number fields. Then there exists an integral ideal $\mathfrak{c}$ of K divisible by and only by the primes of K ramifying in L for which the Artin map*

$$\begin{aligned} \left(\tfrac{L/K}{\cdot}\right) : \mathcal{I}_{\mathfrak{c}}(K) &\longrightarrow \mathrm{Gal}(L/K) \\ \mathfrak{a} = \textstyle\prod_i \mathfrak{p}_i^{r_i} &\longmapsto \textstyle\prod_i \left(\frac{L/K}{\mathfrak{p}_i}\right)^{r_i} \end{aligned}$$

is a well-defined morphism from $\mathcal{I}_{\mathfrak{c}}(K) =$ the set of fractional ideals coprime with $\mathfrak{c}$ to the Galois group $\mathrm{Gal}(L/K)$ and so that the kernel of $(\cdot, L/K)$ contains $\mathcal{P}_{\mathfrak{c}}(K) =$ the set of principal ideals (α) with $\alpha \equiv 1$ mod $\mathfrak{c}$.[5]

One can consider the nonempty set of integral ideals $\mathfrak{c}$ of K verifying theorem 1.4. It follows that there must exist a maximal element $\mathfrak{c}_{L/K}$ in this set - because $\mathcal{O}_K$ is a Noetherian ring - and it is called the *conductor of L/K*.

1.4 Completions

We will need some terminology and facts concerning completions of number fields to define the Selmer and Tate-Shafarevich groups in section 4.5.1, and we discuss them below.
An *absolute value* on a field K is a map $|\cdot| : K \to \mathbb{R}$ that verifies the following:

- $|x| \geq 0$ for any $x \in K$, and $|x| = 0$ if and only if $x = 0$.
- $|xy| = |x||y|$
- There exists $C > 0$ such that $|x + y| \leq C \max\{|x|, |y|\}$

[4] In particular, if $\mathcal{O}_K$ is a principal ideal domain (e.g. $K = \mathbb{Q}(i)$ or $K = \mathbb{Q}$) it has no nontrivial abelian unramified extensions.

[5] Here we restrict the Artin map to $\mathcal{I}_{\mathfrak{c}}(K) \subset \mathcal{I}(K)$ to avoid the possible ramifying primes of L/K.

An absolute value is *trivial* if $|x| = 1$ for $x \neq 0$, and two absolute values are *equivalent* if there exists $a \in \mathbb{R}$ nonzero with $|x|_1 = |x|_2^a$ i.e. one is a positive real power of the other. A *valuation* is an absolute value that satisfies the triangle inequality

$$|x + y| \leq |x| + |y|$$

It is a *nonarchimedean valuation* if $|x + y| \leq \max\{|x|, |y|\}$, or an *archimedean valuation* otherwise. Moreover, if the characteristic of K is $\neq 0$ the only valuations on K are nonarchimedean.
If $K \subset \mathbb{C}$ is a number field one can provide K a natural nontrivial archimedean valuation for every embedding of K in $\mathbb{C}$. It is given by

$$|x|_\iota := |\iota(x)|_{\mathbb{C}}$$

where $|\cdot|_{\mathbb{C}}$ is the usual absolute value in $\mathbb{C}$ and $\iota : K \to \mathbb{C}$ is an embedding.
We can also construct nonarchimedean valuations on a number field. For $x \in K - \{0\}$ the fractional ideal (x) admits the following factorization

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

for some $\nu_{\mathfrak{p}}(x) \in \mathbb{Z}$. In this case, one says $\nu_{\mathfrak{p}}(x)$ is an *exponential valuation* on K because the following properties

- $\nu_{\mathfrak{p}}(x) \in \mathbb{Z}$ for any $x \in K - \{0\}$
- $\nu_{\mathfrak{p}}(xy) = \nu_{\mathfrak{p}}(x) + \nu_{\mathfrak{p}}(y)$
- $\nu_{\mathfrak{p}}(x + y) \geq \min\{\nu_{\mathfrak{p}}(x), \nu_{\mathfrak{p}}(y)\}$

make the map $|x|_{\mathfrak{p}} := c^{\nu_{\mathfrak{p}}(x)}$ a nonarchimedean valuation on K for any $c \in (0, 1)$, if we define $\nu_{\mathfrak{p}}(0) = \infty$ formally. It is called the $\mathfrak{p}$-*adic valuation* on K.
A *place* is an equivalence class of valuations. A field together with a valuation is *complete* if Cauchy sequences are convergent. A *completion* of K is an extension L/K that is complete with respect to a valuation ν' that extends ν and for which every element in L is a limit of elements in K.
In the same way $\mathbb{R}$ is obtained from $\mathbb{Q}$ by considering Cauchy sequences with respect to the usual absolute value, one can obtain completions of a number field with respect to any of its places:

Theorem 1.5. *Let K be a number field and ν a place of K. Let C be the set of Cauchy sequences in K with respect to ν and I the ideal of Cauchy sequences tending to zero. Then C has a natural ring structure and the quotient $\hat{K} = C/I$ is a completion of K. Moreover, any two completions of K are isomorphic up to isomorphisms preserving ν.*

We denote by K_ν the completion of K with respect to ν. If ν is nonarchimedean, there are some important associated objects. The *ring of integers* is the set

$$R = \{x \in K_\nu : \nu(x) \leq 1\}$$

It is a *discrete valuation ring* with maximal ideal $\mathfrak{m} = \{x \in K_\nu : \nu(x) < 1\}$, which is also principal. A generator for $\mathfrak{m}$ is called an *uniformizer*. The *residue field* is $k = R/\mathfrak{m}$.
For instance, the p-adic numbers $\mathbb{Q}_p$ are the completion of $\mathbb{Q}$ with respect the p-adic norm, the ring of integers is the ring of p-adic integers $\mathbb{Z}_p$, the maximal ideal is $p\mathbb{Z}_p$ and the uniformizer is just p. The residue field is $\mathbb{Z}/p\mathbb{Z}$, a finite field.
The following theorem characterizes all the completions of $\mathbb{Q}$:

Theorem 1.6 (Ostrowski)**.** *The completions of* $\mathbb{Q}$ *are* $\mathbb{Q}_p$ *and* $\mathbb{R}$*, where* p *is prime.*

There is also a generalization for number fields i.e. it can be shown that all the nontrivial absolute values on a number field K are either the induced by a prime ideal $\mathfrak{p}$ or the euclidean metric induced by an embedding $K \hookrightarrow \mathbb{C}$. Note also that the completion of a number field with respect to the norm induced by a prime is an example of the more general concept of *local field*.

2 Complex multiplication

In the last section we explained what was the Artin map and Artin reciprocity law, which were obtained by patching together local data in unramified primes. In this section, we will expose that in the case that K is a quadratic imaginary field there is a geometric interpretation of the Artin map in terms of lattices: the fact that the j-invariant of an elliptic curve with CM by $\mathcal{O}_K$ is algebraic lets one define a Galois action on the set of $\overline{\mathbb{Q}}$-isomorphism classes of elliptic curves, which turns out to be the same action of the ideal class group on lattices with endomorphism ring $\mathcal{O}_K$ (if the Galois group is that of the Hilbert class field of K).
In other words, complex multiplication will enable us to do class field theory for quadratic imaginary fields.
Here we followed closely [10, 1, 9].

2.1 Lattices

To describe the endomorphism ring of an elliptic curve over $\mathbb{C}$ we introduce lattices, a class of geometric groups whose set of homomorphisms is easily found. In other words, we first study lattices and its endomorphisms and then relate them to elliptic curves over $\mathbb{C}$.
A *lattice* is a discrete abelian group $\Lambda \subset \mathbb{C}$ of the form $\Lambda = \langle w_1, w_2 \rangle_{\mathbb{Z}}$ where $w_1, w_2 \in \mathbb{C}$ are $\mathbb{R}$-linearly independent and $\Im(w_2 w_1^{-1}) > 0$[6]. Denote by Lat the set of lattices $\Lambda \subset \mathbb{C}$. The set of *homomorphisms between two lattices* is

$$\mathrm{Hom}(\Lambda_1, \Lambda_2) = \{\alpha \in \mathbb{C} : \alpha\Lambda_1 \subset \Lambda_2\}$$

and if $\Lambda_1 = \Lambda_2 = \Lambda$ we set $\mathrm{Hom}(\Lambda, \Lambda) = \mathrm{End}(\Lambda)$. Two lattices are *homothetic* if there is an isomorphism between them i.e. if $\alpha\Lambda_1 = \Lambda_2$ for some $\alpha \in \mathbb{C}$.[7] Homothety induces an equivalence relation $\sim$ on Lat. Given an homothety class $[\Lambda] \in \mathrm{Lat}_\sim$ with $\Lambda = \langle w_1, w_2 \rangle$ we can choose a representative $\Lambda_\tau = \langle 1, \tau \rangle_{\mathbb{Z}}$ where $\tau = w_2 w_1^{-1}$ verifies $\tau \in \mathfrak{H}$. However, it is does not follow that $\Lambda_\tau \sim \Lambda_{\tau'}$ if and only if $\tau = \tau'$. To find an unique representative and for simplicity, from now on we work with lattices of the form Λ_τ. Two points $\tau, \tau' \in \mathfrak{H}$ are *equivalent* if $\tau = \gamma\tau'$ for some $\gamma \in \mathrm{SL}_2(\mathbb{Z})$. We denote this by $\tau \sim \tau'$.
From an homothety $\alpha\Lambda_\tau = \Lambda_{\tau'}$ we obtain $\alpha \cdot 1 = a + b\tau'$ and $\alpha \cdot \tau = c + d\tau'$ for some integers $a, b, c, d \in \mathbb{Z}$. By arguing similarly with $\Lambda_\tau = \alpha^{-1}\Lambda_{\tau'}$ it follows that $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{Z})$ and τ, τ' lie in the same $\mathrm{SL}_2(\mathbb{Z})$-orbit of $\mathfrak{H}$. Conversely, for every $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{Z})$ and $\tau' \in \mathfrak{H}$ the lattice Λ_τ with $\tau = (c + d\tau')(a + b\tau')^{-1}$ verifies $\alpha\Lambda_\tau = \Lambda_{\tau'}$ where $\alpha = a + b\tau'$. In other words, $\Lambda_\tau \sim \Lambda_{\tau'}$ if and only if $\tau \sim \tau'$ and so there is a natural bijection

$$\mathrm{Lat}_\sim \leftrightarrow \mathcal{F}$$

between the set of homothety classes of lattices and the fundamental domain[8] $\mathcal{F}$ for the action of $\mathrm{SL}_2(\mathbb{Z})$ in $\mathfrak{H}$.

[6] In other words, we are choosing an orientation.
[7] Geometrically, that Λ_2 can be obtained by rotating and zooming Λ_1.
[8] $\mathcal{F}$ is defined in section 2.4.1

2.1.1 Orders

The endomorphism ring of lattice is a quadratic imaginary order. We introduce them here to know beforehand some of its properties.
Let K be a quadratic imaginary field. An *order* $\mathcal{O}$ of K is a subring of K containing 1 and a $\mathbb{Q}$-basis of K which is a finitely generated $\mathbb{Z}$-module. Then every element of $\mathcal{O}$ must be integral over $\mathbb{Z}$ so $\mathcal{O} \subset \mathcal{O}_K = \langle 1, w_K \rangle_{\mathbb{Z}}$ where

$$w_K = \frac{d_K + \sqrt{d_K}}{2}$$

and d_K is the discriminant of K. Therefore $\mathcal{O}$ is a free $\mathbb{Z}$-module of rank 2 because it contains a $\mathbb{Q}$-basis of K and its rank its bounded by the rank of $\mathcal{O}_K$.
The index $f = (\mathcal{O}_K : \mathcal{O})$ is now finite. It is the *conductor* of $\mathcal{O}$. Then $\mathcal{O}$ can be written as $\mathcal{O} = \langle 1, f w_K \rangle_{\mathbb{Z}}$, and the *discriminant* D of $\mathcal{O}$ is

$$D = \begin{vmatrix} 2 & f d_K \\ f d_K & f^2 \frac{d_K^2 + d_K}{2} \end{vmatrix} = f^2 d_K$$

An order is completely determined by its discriminant:

Theorem 2.1. *For quadratic imaginary fields, there is only one order of discriminant D.*

2.1.2 The endomorphisms of a lattice

Recall that $\mathrm{End}(\Lambda) = \{\alpha \in \mathbb{C} : \alpha\Lambda \subset \Lambda\}$. Note that $\mathrm{End}(\Lambda)$ depends only on the homothety class $[\Lambda] \in \mathrm{Lat}_{\sim}$. Then $\mathcal{O} = \mathrm{End}(\Lambda)$ is a commutative ring and contains $\mathbb{Z}$. To describe it more precisely, let $\alpha \in \mathcal{O} - \mathbb{Z}$. We obtain $\alpha \cdot 1 = a + b\tau$ and $\alpha \cdot \tau = c + d\tau$ for some integers $a, b, c, d \in \mathbb{Z}$. There are two differences between this situation and that of section 2.1; now we cannot deduce that $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{Z})$. On the other hand, by rewriting these equations into a linear system we deduce that α verifies $\alpha^2 - (a+d)\alpha + (ad - bc) = 0$. But $\alpha \notin \mathbb{Z}$ and $\alpha = a + b\tau \in \mathbb{C}$ imply that $\mathbb{Q}(\tau)$ is a quadratic imaginary extension of $\mathbb{Q}$. Then $\mathcal{O}$ is an *order* of the quadratic imaginary field $\mathbb{Q}(\tau)$.

2.2 Tori

To relate lattices with elliptic curves we need another intermediate step.
A *complex torus* for a lattice $\Lambda \subset \mathbb{C}$ is the Riemann surface $\mathbb{C}/\Lambda$, together with structure of abelian group induced by $\mathbb{C}$. Denote by Tor the set of tori $\mathbb{C}/\Lambda$. The set of homomorphisms *between two complex tori* $T_j = \mathbb{C}/\Lambda_j$ is

$$\mathrm{Hom}(T_1, T_2) = \{\phi : T_1 \to T_2 \text{ holomorphic morphism of abelian groups}\}$$

Two tori are *equivalent* if they are isomorphic. The morphisms between tori are easily described:

Proposition 2.2. *If $\phi : T_1 \to T_2$ is an holomorphic morphism then $\phi(z) = \phi_\alpha(z)$ where $\phi_\alpha(z) = \alpha z$ for some $\alpha \in \mathrm{Hom}(\Lambda_1, \Lambda_2)$. Conversely, every $\alpha \in \mathrm{Hom}(\Lambda_1, \Lambda_2)$ induces $\phi_\alpha \in \mathrm{Hom}(T_1, T_2)$ so there is a one-to-one correspondence*

$$\mathrm{Hom}(\Lambda_1, \Lambda_2) \leftrightarrow \mathrm{Hom}(T_1, T_2)$$

This establishes an equivalence of categories between Lat *and* Tor.

In particular, there is a bijection $\mathrm{Tor}_\sim \leftrightarrow \mathrm{Lat}_\sim$ between the sets of equivalence classes.

2.3 Elliptic curves over $\mathbb{C}$

An *elliptic curve* over $\mathbb{C}$ is a nonsingular planar cubic curve $E \subset \mathbb{P}^2(\mathbb{C})$. Denote by $\mathrm{Ell}_\mathbb{C}$ the set of elliptic curves over $\mathbb{C}$. An *isogeny* $\phi : E_1 \to E_2$ of elliptic curves defined over a field k is a surjective morphism of curves that induces a group homomorphism $E_1(\overline{k}) \to E_2(\overline{k})$. The set of isogenies between two elliptic curves is $\mathrm{Hom}(E_1, E_2)$. Two elliptic curves E_1, E_2 are *isomorphic over k* if there exist mutually inverse bijective isogenies between E_1 and E_2. From now on $k = \mathbb{C}$.
Given a lattice Λ, let E_Λ be the following curve

$$E_\Lambda : Y^2 = 4X^3 - g_2(\Lambda)X - g_3(\Lambda) \tag{2.1}$$

where $g_2 = 60G_4(\Lambda)$ and $g_3 = 140G_6(\Lambda)$.[9] Then E_Λ is an elliptic curve. Conversely, for every elliptic curve $E/\mathbb{C}$ there exists Λ unique up to homothety with $E \simeq E_\Lambda$. To show this, we introduce the *j-invariant*. It is defined by

$$j(\tau) = 1728\frac{g_2(\tau)^3}{g_2(\tau)^3 - 27g_3(\tau)^2}$$

It is $\mathrm{SL}_2(\mathbb{Z})$-invariant since it is a quotient of two modular forms[10] of weight 12. It can be shown that $j : \mathcal{F} \to \mathbb{C}$ is injective and surjective. In fact,

Theorem 2.3. *The j-invariant induces an isomorphism between $X(1)$ and $\hat{\mathbb{C}}$.*

Here $X(1)$ can be thought as $\mathcal{F}$ together with a point at infinity (section 3.1). Moreover, two elliptic curves are isomorphic over $\mathbb{C}$ if and only if their j-invariants are the same. Thus, there is a one-to-one correspondence between $\mathrm{Lat}_\sim$ and $(\mathrm{Ell}_\mathbb{C})_\sim$:

$$\begin{array}{ccc} \mathrm{Lat}_\sim & \longleftrightarrow & (\mathrm{Ell}_\mathbb{C})_\sim \\ \Lambda & \longmapsto & E_\Lambda \end{array}$$

It turns out that elliptic curves over $\mathbb{C}$ are tori:

Theorem 2.4. *The following map*

[9] Here $G_k(\Lambda)$ or $G_k(\tau)$ is the *Eisenstein series of weight k*
[10] Modular forms are defined in section 3.5.2.

$$\begin{aligned}\Phi: \quad \mathbb{C}/\Lambda &\longrightarrow E_\Lambda \\ z+\Lambda &\longmapsto (\wp_\Lambda(z) : \wp'_\Lambda(z) : 1) \\ 0+\Lambda &\longmapsto (0:1:0)\end{aligned}$$

is an holomorphic isomorphism of Riemann surfaces, where $\wp_\Lambda$ is the Weierstrass $\wp$ function associated to Λ.

Moreover, $\wp$ and $\wp'$ generate all the functions on a complex torus; the field of the *doubly periodic functions* with respect to Λ is $\mathbb{C}(\wp, \wp')$ where $\wp$ denotes the Weierstrass function with respect to Λ.

Theorem 2.5. *Let $\wp_1$ and $\wp_2$ denote the Weierstrass functions associated to Λ_1 and Λ_2. Then the following are equivalent: $\phi_\alpha \in \mathrm{End}(T_1, T_2)$; $\wp_2(\alpha z) \in \mathbb{C}(\wp_1, \wp_1')$; there is an unique $\alpha \in \mathbb{C}$ making the following diagram commute*

$$\begin{array}{ccc} T_1 & \xrightarrow{\phi_\alpha} & T_2 \\ \downarrow{\scriptstyle \Phi_1} & & \downarrow{\scriptstyle \Phi_2} \\ E_1 & \xrightarrow{\phi} & E_2 \end{array}$$

where $\phi = \Phi \circ \phi_\alpha \circ \Phi^{-1}$. Moreover, there is an isomorphism of abelian groups between $\mathrm{Hom}(E_1, E_2)$ *and* $\mathrm{Hom}(T_1, T_2)$.

The isomorphism of rings exists because if we are given an isogeny $\phi : E_1 \to E_2$ then $\Phi_2^{-1} \circ \phi \circ \Phi_1(z) = f(z)$ is an holomorphic morphism between complex tori, so $f = \phi_\alpha$ for some $\alpha \in \mathbb{C}$, but then $\alpha\Lambda_1 \subset \Lambda_2$ so $\alpha \in \mathrm{End}(\Lambda_1, \Lambda_2)$. Conversely, given $\alpha \in \mathrm{End}(\Lambda_1, \Lambda_2)$ the map $\phi = \Phi_2 \circ \phi_\alpha \circ \Phi_1^{-1}$ is an isogeny $\phi : E_1 \to E_2$.
Therefore, we have established $\mathrm{Lat}, \mathrm{Tor}$ and $\mathrm{Ell}_{\mathbb{C}}$ are **equivalent categories**. We particularize theorem 2.5 for $\Lambda_1 = \Lambda_2 = \Lambda$ to obtain the following:

Corollary 2.6. *Suppose* $\mathrm{End}(E) \supsetneq \mathbb{Z}$. *Then* $\mathrm{End}(E)$ *is naturally isomorphic to an order $\mathcal{O}$ of a quadratic imaginary field $K \subset \mathbb{C}$*

$$\begin{aligned}\Psi: \ \mathrm{End}(E) &\xrightarrow{\cong} \mathcal{O} \subset \mathbb{C} \\ \phi &\longmapsto \alpha = \Psi(\phi)\end{aligned}$$

where $\alpha \in \mathbb{C}$ makes the following diagram commute

$$\begin{array}{ccc} T & \xrightarrow{\phi_\alpha} & T \\ \downarrow{\scriptstyle \Phi} & & \downarrow{\scriptstyle \Phi} \\ E & \xrightarrow{\phi} & E \end{array}$$

Here $T = \mathbb{C}/\Lambda$ and $E_\Lambda \simeq E$.

Remark 2.7. The *dual isogeny* for a degree N isogeny $E \to E'$ is another isogeny $E' \to E$ such that the composition is the multiplication-by-N map, and exists whenever the isogeny is not constant. For the case $E = E'$, it is particularly easy to find. Given $\alpha \in \mathcal{O}$, the dual isogeny is just $\Phi(\overline{\alpha})$ where $\overline{\alpha}$ denotes the conjugate of α.[11]

This motivates the following definition:

Definition 2.8. An elliptic curve $E/\mathbb{C}$ has *complex multiplication* by the quadratic imaginary order $\mathcal{O}$ if $E \simeq E_\Lambda$ and $\mathrm{End}(\Lambda) = \mathcal{O}$. We denote the set of equivalence classes of elliptic curves with CM by $\mathcal{O}$ by

$$\mathrm{Ell}_{\mathbb{C}}(\mathcal{O}) = \{[E] \in \mathrm{Ell}_{\mathbb{C}} \text{ with CM by } \mathcal{O}\}$$

We define $[\alpha] = \Psi^{-1}(\alpha)$ where Ψ is as in corollary 2.6. The *invariant differential* for an elliptic curve is $w = \frac{dx}{y}$. It is invariant under translations in the elliptic curve, hence its name. For instance, the pullback by Φ of w is cdz for some constant c, which is certainly translation invariant. In this case, the pair $(E, [.])$ is *normalized*. The chosen embedding has an additional property:

Corollary 2.9. *For any invariant differential $\omega \in \Omega_E$,*

$$[\alpha]^*\omega = \alpha\omega$$

where $[\alpha]^*$ denotes the pullback of differential forms by $[\alpha]$.

Example 2.10. Consider the lattice $\Lambda = \Lambda_i$. Then $\mathcal{O} = \mathrm{End}(\Lambda_i)$ clearly contains the Gaussian integers $\mathbb{Z}[i]$, so $\mathcal{O} = \mathbb{Z}[i]$ because $\mathcal{O}_K$ is the maximal order of $K = \mathbb{Q}(i)$. Therefore any elliptic curve $\mathbb{C}$-isomorphic to E_Λ has complex multiplication by $\mathbb{Z}[i]$. An expression for the isomorphism $[i] : E_\Lambda \to E_\Lambda$ is easily found because $(X, Y) = (\wp(z), \wp'(z))$ and

$$[i](X, Y) = (\wp(iz), \wp'(iz)) = (-\wp(z), i\wp'(z)) = (-X, iY)$$

because of the properties of the Weierstrass $\wp$ function.

2.4 Binary quadratic forms

Complex multiplication also arises from an algebraic classification problem, namely classifying the set forms of given discriminant. A *form* is a *binary quadratic form*, i.e. an homogeneous polynomial in two variables,

$$f(X, Y) = (a, b, c) = aX^2 + bXY + cY^2$$

satisfying the following conditions: f is *integral, primitive* and positive definite. That is, a, b, c are coprime integers verifying $b^2 - 4ac < 0$ and $a > 0$. These latter conditions follow from Sylvester's criterion of positive definiteness, because one can rewrite f as.

$$f(X, Y) = v^t \begin{pmatrix} a & b/2 \\ b/2 & c \end{pmatrix} v$$

[11] Note that $\overline{\alpha} \in \mathcal{O} \subset \mathcal{O}_K$ because $\mathrm{Tr}(\alpha) = \alpha + \overline{\alpha} \in \mathbb{Z}$ and $1 \in \mathcal{O}$.

with $v = (X, Y)$. The *discriminant* of a form f is $D = b^2 - 4ac$. The set of forms of discriminant D is denoted by $F(D)$. From now on D is a fixed negative integer.
It follows from its definition that the discriminant is invariant under invertible transformations in $\mathbb{Z}^2$, hence by the special linear group $\mathrm{SL}_2(\mathbb{Z})$. Moreover, $\mathrm{SL}_2(\mathbb{Z})$ acts on $F(D)$; for $\gamma = \left(\begin{smallmatrix} u & v \\ s & t \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{Z})$ let $\gamma(X, Y) = (uX + vY, sX + tY)$ and define the action by

$$f^\gamma(X, Y) = f(\gamma(X, Y))$$

Then $f^\gamma \in F(D)$ remains a form. Two forms f, g are *equivalent* if they lie in the same $\mathrm{SL}_2(\mathbb{Z})$-orbit of $F(D)$. We denote this by $f \sim g$. The set of orbits is the *form class group* of discriminant D denoted by $\mathrm{cl}(D)$. That is, $\mathrm{cl}(D)$ is a finite group where the group operation is the composition of equivalence classes of quadratic forms found by Gauss. There is also an isomorphism $\mathrm{cl}(D) \simeq \mathrm{cl}(\mathcal{O}_K)$ when $K = \mathbb{Q}(\sqrt{D})$ has discriminant D.

2.4.1 The form class group is finite

The *principal root* $\tau(f)$ of a form f is the unique solution of the equation $f(\tau, 1) = 0$ lying in $\mathfrak{H}$[12], namely $\tau = \frac{-b+\sqrt{D}}{2a}$. By considering principal roots we will be able to choose an unique representative for each orbit in $\mathrm{cl}(D)$. A *reduced* form verifies $-a < b \leq a \leq c$ or $0 \leq b \leq a = c$. Then f is reduced if and only if $\tau(f) \in \mathcal{F}$ where

$$\mathcal{F} = \{z \in \mathbb{C} : |z| \geq 1 \text{ and } -1/2 \leq \Re z \leq 0\} \cup \{z \in \mathbb{C} : |z| > 1 \text{ and } 0 < \Re z \leq 1/2\}$$

Recall that $\mathrm{SL}_2(\mathbb{Z})$ also acts in $\mathfrak{H}$. Then $\gamma^{-1}\tau$ is the principal root of f^γ. It follows easily that $f \sim g$ if and only if $\tau(f) \sim \tau(g)$.[13] Therefore, every form is equivalent to an unique reduced form because the $\mathrm{SL}_2(\mathbb{Z})$-translates of $\mathcal{F}$ form a disjoint cover of $\mathfrak{H}$; given f compute $\tau(f)$ and find the only $\tau \in \mathcal{F}$ with $\tau \sim \tau(f)$.
Finally, to show that $\mathrm{cl}(D)$ is finite we observe that

$$\min_{\tau \in \mathcal{F}} \Im(\tau) = \frac{\sqrt{3}}{2}$$

Hence $\Im\tau(f) = \frac{\sqrt{|D|}}{2a} \geq \frac{\sqrt{3}}{2}$ implies $1 \leq a \leq \sqrt{|D|/3}$, so there are only finitely many possibilities for a, and $b \in (-a, a]$ for a fixed a. Therefore the form class group is finite

$$h(D) = \Big|\mathrm{cl}(D)\Big| = \text{the number of reduced forms in } F(D) \leq 2\big(\sqrt[2]{|D|/3}\big) \leq |D|/3$$

and finding the reduced forms in $F(D)$ gives an effective procedure to find $h(D)$.

2.4.2 Lattices, proper ideals and forms

Let $\mathcal{O}$ be an order of $K = \mathbb{Q}(\sqrt{D})$. We want to define an ideal class group $\mathrm{cl}(\mathcal{O})$ for $\mathcal{O}$ as we did for number fields to show that $\mathrm{cl}(D) \simeq \mathrm{cl}(\mathcal{O})$. In fact, we only have to show the

[12] Recall that D is negative.
[13] It is also clear that if $f = (a, b, c)$ is reduced then $aX^2 + bX + c$ is the minimal polynomial of $\tau(f)$.

number of equivalence class of ideals is $|\mathrm{cl}(D)|$ because the product in $\mathrm{cl}(D)$ will be the induced by that of $\mathrm{cl}(\mathcal{O})$. For this to be true, we will have to consider only some of the ideals, the *proper* ideals.
A *fractional* $\mathcal{O}$*-ideal* $\mathfrak{b}$ is a finitely generated $\mathcal{O}$-submodule of K. It is *invertible* if $\mathfrak{a}\mathfrak{b} = \mathcal{O}$ for some $\mathfrak{a}$, and *proper* if $\mathcal{O}(\mathfrak{b}) = \mathcal{O}$. It turns out a fractional ideal is proper if and only if it is invertible. Then the set of proper ideals $I(\mathcal{O})$ forms a group, so we can form the ideal class group $\mathrm{cl}(\mathcal{O}) = I(\mathcal{O})/P(\mathcal{O})$ where $P(\mathcal{O}) \subset I(\mathcal{O})$ is the subgroup of principal ideals.
Now let $f = (a, b, c) \in F(D)$ be a form. The *associated lattice* is $\Lambda_f = a\langle 1, \tau(f)\rangle_{\mathbb{Z}}$. Then $f \sim g$ if and only if $\Lambda_f \sim \Lambda_g$. In other words, $[f] \mapsto [\Lambda_f]$ defines an injection $\Lambda_{[\cdot]} : \mathrm{cl}(D) \hookrightarrow \mathrm{Lat}_{\sim}$. Moreover, these lattices are proper $\mathcal{O}$-ideals. To show this, let the *order of a lattice* be its set of endomorphisms $\mathcal{O}(\Lambda) = \mathrm{End}(\Lambda)$. Then

Lemma 2.11. *One has* $\mathcal{O} \circ \Lambda(f) = \mathcal{O}(\Lambda_f) = \langle 1, a\tau\rangle_{\mathbb{Z}}$*, which is the order of discriminant* D *of* $K = \mathbb{Q}(\sqrt{D})$*.*

Proof. One inclusion follows easily by noting $(a\tau)^2 = -b \cdot (a\tau) - ac \cdot 1$. For the other, observe that $\mathcal{O}(\Lambda_f) = \mathcal{O}(\Lambda_{\tau(f)})$ and let $\alpha \in \mathcal{O}(\Lambda_f)$. Then $\alpha \cdot 1 = \alpha = u + v\tau$ for some $u, v \in \mathbb{Z}$ so $\alpha \cdot \tau = (u - \frac{bv}{a})\tau - \frac{vc}{a}$ but a, b, c are coprime so $a|v$ and $\alpha \in \langle 1, a\tau\rangle_{\mathbb{Z}}$. □

In particular, the injection $\Lambda_{[\cdot]} : \mathrm{cl}(D) \to \mathrm{Lat}_{\sim}(\mathcal{O})$ restricts to lattices with $\mathrm{End}(\Lambda) = \mathcal{O}$. Thus $|\mathrm{cl}(D)| \leq |\mathrm{cl}(\mathcal{O})|$. Conversely, it can be shown that every proper $\mathcal{O}$-ideal is the associated lattice of a form in $F(D)$.

2.5 $\mathrm{cl}(\mathcal{O})$ acts on $\mathrm{Ell}_{\mathbb{C}}(\mathcal{O})$

With the results above, one can now find all the elliptic curves over $\mathbb{C}$ with CM. Let $\mathcal{O}$ be a quadratic imaginary order. Then the following map

$$\begin{aligned} \mathrm{cl}(\mathcal{O}) &\longleftrightarrow \mathrm{Ell}_{\mathbb{C}}(\mathcal{O}) \\ [\mathfrak{a}] &\longmapsto E_{\mathfrak{a}} \end{aligned}$$

is a bijection, where $E_{\mathfrak{a}}$ is as in (2.1). It is well-defined, since every proper $\mathcal{O}$-ideal $\mathfrak{a}$ is a lattice with $\mathrm{End}(\mathfrak{a}) = \mathcal{O}$, and so for every $E \in [E_{\mathfrak{a}}]$ one has $\mathrm{End}(E) \simeq \mathrm{End}(\mathfrak{a}) = \mathcal{O}$ since $E \simeq E_{\mathfrak{a}}$. Every lattice Λ with $\mathrm{End}(\Lambda) = \mathcal{O}$ is homothetic to a proper $\mathcal{O}$-ideal, so surjectivity follows. And if $[E_{\mathfrak{a}}] = [E_{\mathfrak{b}}]$ for two fractional proper $\mathcal{O}$-ideals $\mathfrak{a}, \mathfrak{b}$ then $j(E_{\mathfrak{a}}) = j(E_{\mathfrak{b}})$ so $\mathfrak{a}$ and $\mathfrak{b}$ must be $\mathbb{C}$-homothetic, so $[\mathfrak{a}] = [\mathfrak{b}]$ in $\mathrm{cl}(\mathcal{O})$.[14] Let us define

$$\mathrm{Lat}_{\sim}(\mathcal{O}) = \{[\Lambda] \in \mathrm{Lat}_{\sim} \text{ with } \mathrm{End}(\Lambda) = \mathcal{O}\}$$

and

$$\mathcal{F}(\mathcal{O}) = \{\tau(f) \text{ with } f = (a, b, c) \in F(D)\}$$

[14] In this case, $\mathbb{C}$-homothety implies $\mathcal{O}$-homothety.

Then the correspondence of section 2.3 restricts to $\mathrm{Ell}_{\mathbb{C}}(\mathcal{O}) \leftrightarrow \mathrm{Lat}_{\sim}(\mathcal{O})$. From the discussion in section 2.4 and the bijection $\mathrm{Lat}_{\sim} \leftrightarrow \mathcal{F}$ of section 2.1 it follows that we have a correspondence between finite sets $\mathrm{Lat}_{\sim}(\mathcal{O}) \leftrightarrow \mathcal{F}(\mathcal{O})$. Thus

$$|\operatorname{cl}(\mathcal{O})| = |\operatorname{cl}(D)| = h(D) = |\operatorname{Ell}_{\mathbb{C}}(\mathcal{O})| = |\mathcal{F}(\mathcal{O})|$$

Now let us define the following action

$$\begin{aligned} \operatorname{cl}(\mathcal{O}) \times \operatorname{Ell}_{\mathbb{C}}(\mathcal{O}) &\longrightarrow \operatorname{Ell}_{\mathbb{C}}(\mathcal{O}) \\ [\mathfrak{a}], E_{\mathfrak{b}} &\longmapsto \mathfrak{a}E_{\mathfrak{b}} = E_{\mathfrak{a}^{-1}\mathfrak{b}} \end{aligned}$$

It is indeed a well-defined action. The crucial fact here is that it is a *free* action: if $[E_{\mathfrak{a}^{-1}\mathfrak{b}}] = [E_{\mathfrak{b}}]$ then $\mathfrak{a}$ is principal, since $\mathfrak{a}^{-1}\mathfrak{b}$ and $\mathfrak{b}$ must be homothetic and $\mathfrak{b}$ is invertible because it is proper. In general, if we have a free group action $G \circlearrowright X$ with $|G| = |X|$ finite then it is *transitive*, i.e. the orbit of any element is X. Hence, our action above is transitive.

Example 2.12. For instance, there is only one isomorphism class of elliptic curves with CM by $\mathcal{O} = \mathbb{Z}[i]$ or $\mathbb{Z}[\sqrt{-2}]$ because they are both principal ideal domains, and there are two isomorphism classes when $\mathcal{O} = \mathbb{Z}[\frac{1+\sqrt{-5}}{2}]$. These are simple examples because they are all maximal orders.

2.5.1 The action in terms of isogenies

The action $\operatorname{cl}(\mathcal{O}) \circlearrowright \operatorname{Ell}_{\mathbb{C}}(\mathcal{O})$ can be interpreted in terms of isogenies.
Let $\phi : E_1 \to E_2$ be an isogeny with $E_i = E_{\Lambda_i}$ and denote by $T_i = \mathbb{C}/\Lambda_i$. Then by theorem 2.5 there exists $\alpha \in \operatorname{Hom}(\Lambda_1, \Lambda_2)$ with $\alpha\Lambda_1 \subset \Lambda_2$ and $\phi = \Phi_2 \phi_\alpha \Phi_1^{-1}$. Denote by $\Lambda_1' = \alpha\Lambda_1$ and $T_1' = \mathbb{C}/\Lambda_1'$. The lattices Λ_1 and Λ_1' are then $\mathbb{C}$-homothetic, so $\varphi' : T_1 \xrightarrow{z \mapsto \alpha z} T_1'$ is an isomorphism and the natural inclusion $\varphi'' : T_1' \xrightarrow{z \mapsto z} T_2$ has kernel $\ker \varphi'' \simeq \Lambda_1'/\Lambda_2$. Moreover, $\phi_\alpha = \varphi''\varphi'$. In other words, every isogeny is induced by an inclusion of tori up to $\mathbb{C}$-isomorphism.
If $E_1/\mathbb{C}$ has CM by $\mathcal{O}$ then Λ_1 is homothetic to a proper $\mathcal{O}$-ideal $\mathfrak{b}$, and if $\mathfrak{a}$ is a proper $\mathcal{O}$-ideal then so is $\mathfrak{a}^{-1}\mathfrak{b}$. The inclusion of lattices $\mathfrak{b} \subset \mathfrak{a}^{-1}\mathfrak{b}$ induces an isogeny $\phi_{\mathfrak{a}} : E_{\mathfrak{b}} \to E_{\mathfrak{a}^{-1}\mathfrak{b}}$. Therefore, all elliptic curves with CM by $\mathcal{O}$ are isogeneous, because the action is transitive. The kernel of $\phi_{\mathfrak{a}} : E \to \mathfrak{a}E$ can be shown to be

$$E[\mathfrak{a}] = \{P \in E : \alpha P = 0 \text{ for all } \alpha \in \mathfrak{a}\}$$

2.6 The Hilbert class polynomial

Let $\mathcal{O}$ be a imaginary quadratic order of discriminant D. The *Hilbert class polynomial* is

$$H_D(X) = \prod_{[E] \in \operatorname{Ell}_{\mathbb{C}}(\mathcal{O})} \Big(X - j(E)\Big)$$

By using the properties of the modular polynomial Φ_N for N prime in section 3.5.4 we will show that $H_D(X) \in \mathbb{Z}[X]$. We only have to take for granted theorem 3.12 and the following theorem

Theorem 2.13 (Dirichlet,Weber)**.** *There are infinitely many prime ideals $\mathfrak{p}$ of prime norm $(\mathcal{O} : \mathfrak{p})$ in every class of* $\mathrm{cl}(\mathcal{O})$.

We proceed.

Theorem 2.14. *One has $H_D(X) \in \mathbb{Z}[X]$. Thus, the j-invariant of an elliptic curve E with CM is an algebraic integer, and E can be defined over $\overline{\mathbb{Q}}$.*

Proof. Let $E \in \mathrm{Ell}_{\mathbb{C}}(\mathcal{O})$ and let $\mathfrak{p} = \alpha\mathcal{O}$ be a principal ideal of norm p. Then $[\mathfrak{p}]$ acts trivially on $\mathrm{Ell}_{\mathbb{C}}(\mathcal{O})$ and there is an isomorphism $\varphi : \mathfrak{p}E \xrightarrow{\simeq} E$. If $\mathfrak{b}$ is a lattice with $E_{\mathfrak{b}} \simeq E$ then the inclusion of lattices $\mathfrak{b} \subset \mathfrak{p}^{-1}\mathfrak{b}$ induces an isogeny $\phi : E \to \mathfrak{p}E$ with $\ker\phi \simeq \mathfrak{p}^{-1}\mathfrak{b}/\mathfrak{b}$, but $(\mathfrak{p}^{-1}\mathfrak{b} : \mathfrak{b}) = (\mathfrak{b} : \mathfrak{p}\mathfrak{b}) = (\mathcal{O} : \mathfrak{p}) = p$, thus ϕ is cyclic. Hence $\varphi \circ \phi$ is a degree p (cyclic) isogeny of E onto itself, and so by theorem 3.14 one has $\Phi_p(j(E), j(E)) = 0$. Thus $j(E)$ is an algebraic integer because $\Phi_p(X, X)$ is a monic polynomial by theorem 3.15, and E can be defined over $\overline{\mathbb{Q}}$.
Let $\mathrm{Ell}_{\overline{\mathbb{Q}}}(\mathcal{O})$ be the set of equivalence classes of elliptic curves defined over $\overline{\mathbb{Q}}$ modulo $\overline{\mathbb{Q}}$-isomorphism. Then there is a bijection between $\mathrm{Ell}_{\overline{\mathbb{Q}}}(\mathcal{O})$ and $\mathrm{Ell}_{\mathbb{C}}(\mathcal{O})$[15]. The absolute Galois group $G = \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ acts on $\mathrm{Ell}_{\overline{\mathbb{Q}}}(\mathcal{O})$ as follows: for every $\sigma \in G$ and $[E] \in \mathrm{Ell}_{\overline{\mathbb{Q}}}(\mathcal{O})$ with $E : y^2 = 4x^3 - Ax - B$ we define $E^\sigma : y^2 = 4x^3 - A^\sigma x - B^\sigma$. It is well-defined, since given an isogeny $\phi : E \to E$ defined over $\overline{\mathbb{Q}}$, ϕ^σ is obtained by applying σ to the coefficients in an expression for ϕ, so $\mathrm{End}(E) \simeq \mathrm{End}(E^\sigma)$.
The coefficients of $H_D(X)$ are symmetric functions on the j-invariants for $[E] \in \mathrm{Ell}_{\overline{\mathbb{Q}}}(\mathcal{O})$, thus invariant under the action of G so $H_D(X) \in \mathbb{Q}$. But E was arbitrary so $H_D(X)$ divides $\Phi_p(X, X)$ in $\mathbb{Q}[X]$, because $\Phi_p(X, X) \in \mathbb{Q}[X]$ by theorem 3.12. Finally, Φ_p is monic by theorem 3.15, so by Gauss lemma it follows that $H_D(X) \in \mathbb{Z}[X]$. □

2.7 The Galois action

Let $\mathcal{O} \subset K$ be an order of discriminant D and K a quadratic imaginary field. In the proof of theorem 2.14 we considered the action of the absolute Galois group $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$. Note that many elements do not fix K. To simplify the picture, we take an smaller group $G = \mathrm{Gal}(L/K)$ where L is the splitting field of $H_D(X)$. For each $\sigma \in G$ the curve E^σ has CM by $\mathcal{O}$. The action of the class group in section 2.5 is free and transitive so $E^\sigma \simeq \mathfrak{a}E$ for some unique $\mathfrak{a} \in \mathrm{cl}(\mathcal{O})$. Moreover, the Galois action and the class group action are nicely related

Lemma 2.15. *For any $E \in \mathrm{Ell}_{\overline{\mathbb{Q}}}(\mathcal{O})$, $\sigma \in G$ and $\mathfrak{b} \in \mathrm{cl}(\mathcal{O})$ one has $(\mathfrak{b}E)^\sigma = \mathfrak{b}^\sigma E^\sigma$.*

[15] If two elliptic curves are $\overline{\mathbb{Q}}$-isomorphic they are $\mathbb{C}$-isomorphic. Conversely, if E, E' are $\mathbb{C}$-isomorphic with CM by $\mathcal{O}$ then both can be defined over $\overline{\mathbb{Q}}$ because their j-invariant is the same and belongs to $\overline{\mathbb{Q}}$. It can be shown that then they are $\overline{\mathbb{Q}}$-isomorphic.

Taking this for granted, it follows that $\mathfrak{a}$ depends only on σ: for any $E' \in \mathrm{Ell}_{\overline{\mathbb{Q}}}(\mathcal{O})$ one has $E' \simeq \mathfrak{b}E$ for some $\mathfrak{b} \in \mathrm{cl}(\mathcal{O})$ by transitivity. Thus $E'^{\sigma} \simeq (\mathfrak{b}E)^{\sigma} = \mathfrak{b}^{\sigma}E^{\sigma} = \mathfrak{b}E^{\sigma} \simeq \mathfrak{b}\mathfrak{a}E = \mathfrak{a}\mathfrak{b}E \simeq \mathfrak{a}E'$. Here $\mathfrak{b}^{\sigma} = \mathfrak{b}$ because σ fixes K and $\mathfrak{b} \subset \mathcal{O} \subset K$. Therefore we have a well-defined injective[16] map

$$\begin{aligned} \Psi : \mathrm{Gal}(L/K) &\lhook\joinrel\longrightarrow \mathrm{cl}(\mathcal{O}) \\ \sigma &\longmapsto \Psi(\sigma) = \mathfrak{a} \end{aligned}$$

where $E^{\sigma} = \mathfrak{a}E$. It is a group homomorphism since for any E

$$\Psi(\sigma\tau)E = E^{\sigma\tau} = (E^{\sigma})^{\tau} = (\Psi(\sigma)E)^{\tau} = \Psi(\tau)\Psi(\sigma)E = \Psi(\sigma)\Psi(\tau)E$$

The injectivity of Ψ implies that $\mathrm{Gal}(L/K)$ is abelian, and $[L : K] \leq h(D)$. In fact, Ψ surjects.
The splitting field L is known as the *ring class field* of $\mathcal{O}$. In the case where the order $\mathcal{O} = \mathcal{O}_K$ is maximal L is the *Hilbert class field*, which is also the maximal unramified abelian extension of K.

2.8 The First Main theorem of Complex Multiplication

The goal of this section is to prove that Ψ is an isomorphism. The previous results together with the Artin map of section 1.2 help in the proof of this result:

Theorem 2.16. *Let $\mathcal{O}$ be a quadratic imaginary order of discriminant D and L the splitting field of $H_D(X)$ over $K = \mathbb{Q}(\sqrt{D})$. Then there is an isomorphism*

$$\begin{aligned} \Psi : \mathrm{Gal}(L/K) &\xrightarrow{\simeq} \mathrm{cl}(\mathcal{O}) \\ \sigma &\longmapsto \Psi(\sigma) = \mathfrak{a} \end{aligned}$$

where $E^{\sigma} = \mathfrak{a}E$.

Proof. Take $[\mathfrak{a}] \in \mathrm{cl}(\mathcal{O})$ and let $\mathfrak{p}$ be a prime of K satisfying the following conditions

(a) $\mathfrak{q} = \mathfrak{p} \cap \mathcal{O}$ is a proper $\mathcal{O}$-ideal of prime norm with $[\mathfrak{q}] = [\mathfrak{a}]$

(b) p and $\mathfrak{p}$ are unramified in K and L respectively

(c) For every $[E] \in \mathrm{Ell}_L(\mathcal{O})$ there exists an elliptic curve E/L with good reduction modulo $\mathfrak{P}$ for any prime $\mathfrak{P}$ of L above $\mathfrak{p}$

(d) The elements in $\{j(E)\}_{[E] \in \mathrm{Ell}_L(\mathcal{O})}$ are distinct [17] modulo $\mathfrak{P}$ for $\mathfrak{P}|\mathfrak{p}$

[16] Given $E \in \mathrm{Ell}_{\overline{\mathbb{Q}}}(\mathcal{O})$, the only element in G fixing $j(E)$ is the identity.

[17] Condition d lets us detect whether two curves are isomorphic or not by checking if they are isomorphic modulo $\mathfrak{P}$.

There are infinitely many $\mathfrak{p}$ for which these conditions hold, because theorem 2.13 assures the existence of infinitely many $\mathfrak{q}$ verifying condition a and because we can lift the prime $\mathfrak{q}$ to a prime $\mathfrak{p} = \mathfrak{q}\mathcal{O}_K$ of K provided that $\mathfrak{q}$ does not contain (f), where f is the conductor of $\mathcal{O}$ defined in section 2.1.1.
For condition b recall that in a finite extension only finitely many primes ramify[18]. For c take a representative E of $[E]$ i.e. choose a Weierstrass model and avoid primes $\mathfrak{P}$ of L dividing the discriminant of the curve[19], and for condition d avoid primes $\mathfrak{P}$ of L appearing in the factorizations in $\mathcal{O}_L$ of the elements in $\{j(E) - j(E')\}_{[E]\neq[E']}$ where $[E], [E'] \in \mathrm{Ell}_L(\mathcal{O})$.
Fix $\mathfrak{P}$ above $\mathfrak{p}$ and E/L with good reduction modulo $\mathfrak{P}$. Let $\overline{E}/\mathbb{F}_{\mathfrak{P}}$ be the reduction of E modulo $\mathfrak{P}$, obtained by reducing the coefficients in L of the equation of E modulo $\mathfrak{P}$. So if E is given by $E : Y^2 = X^3 + AX + B$ with $A, B \in \mathcal{O}_L$ then $\overline{E} : Y^2 = X^3 + \overline{A}X + \overline{B}$ and $\Delta(\overline{E}) \neq 0$.
Now $\mathfrak{p}$ is unramified so we may form the Artin symbol $\sigma = (\mathfrak{p}, L/K)$ as in section 1.2, and $(\mathcal{O}_K : \mathfrak{p}) = (\mathcal{O} : \mathfrak{q}) = p$. Recall that σ verifies $\sigma(x) \equiv x^p \bmod \mathfrak{P}$, i.e. $\mathcal{Q}(\sigma) = \overline{\sigma}$ is the Frobenius element. Then σ induces an isogeny $\pi : \overline{E} \to \overline{E}^{\overline{\sigma}}$ by applying $\overline{\sigma}$ coordinatewise $\pi(x, y) = (x^p, y^p)$ from $\overline{E}$ to $\overline{E}^{\overline{\sigma}} : Y^2 = X^3 + \overline{A}^p X + \overline{B}^p$.
Recall from section 2.5.1 that $\mathfrak{q}$ induces a degree p isogeny $\phi_{\mathfrak{q}} : E \to \mathfrak{q}E$. Note that $\mathfrak{q}E$ has good reduction modulo $\mathfrak{P}$ by condition c. Then by reducing $\phi_{\mathfrak{q}}$ modulo $\mathfrak{P}$ we obtain $\phi : \overline{E} \to \overline{\mathfrak{q}E}$. It can be shown that ϕ has again degree p. The composition of ϕ with its dual isogeny[20] $\hat{\phi}$ is the multiplication-by-p map m_p, which is *purely inseparable*[21] in characteristic p. Then either ϕ or $\hat{\phi}$ is inseparable, because separability is transitive for extensions, so we may suppose this is the case for ϕ. It turns out that every inseparable isogeny factors as a composition of an n-th power of π and a separable map φ', for some $n \geq 0$, i.e.

$$\varphi = \varphi' \circ \pi^n$$

But here $\varphi = \phi$ and $\deg \varphi = \deg \varphi' \deg \pi^n = p = d \cdot p^n$ with $p \nmid d$ so $d = 1$ and $n = 1$. Thus $\varphi' : \overline{E}^{\overline{\sigma}} \xrightarrow{\simeq} \overline{\mathfrak{q}E}$ is an isomorphism i.e. $\overline{j(E^\sigma)} = j(\overline{E}^{\overline{\sigma}}) = j(\overline{\mathfrak{q}E}) = \overline{j(\mathfrak{q}E)}$ and condition d implies $[\mathfrak{q}E] = [E^\sigma]$. Therefore Ψ surjects. □

2.8.1 Relating $\mathrm{cl}(\mathcal{O})$ with $\mathrm{cl}(\mathcal{O}_K)$

From now on, we shall denote the field L in the proof of theorem 2.16 by K_f or the *ring class field of conductor* f. Note that the result above is an special case of a weakening of theorem 1.4 i.e. the conductor in this case is principal $\mathfrak{c} = (f)$ and it is generated by the conductor of $\mathcal{O}$, hence its name. From theorem 1.4 we deduce that the primes of K ramifying in $L = K_f$ are precisely those dividing the conductor f.

[18] Namely, those dividing the discriminant of the extension defined in section 1.
[19] Because those are the primes of *bad reduction*.
[20] See remark 2.7
[21] A nonconstant morphism $\varphi : C \to C'$ defined over a field K between curves over K has the property $\mathcal{P}$ if so does the field extension $K(C')/\varphi^* K(C)$ where $\varphi^* : K(C') \to K(C)$ is defined by $f \mapsto f \circ \varphi$. $\mathcal{P}$ can be separability, Galois, etc.

Section 4.6 will require an accurate description of some Galois groups involving the ring class fields, and this reduces to relate $\mathrm{cl}(D)$ with $\mathrm{cl}(\mathcal{O}_K)$. For this, we will now relate the proper $\mathcal{O}$-ideals with the ideals of $\mathcal{O}_K$, we introduce another family of ideals, and discuss some results on them that we will not prove[22]. We suppose throughout this section that $\mathcal{O}_K^\times = \{\pm 1\}$.

A nonzero $\mathcal{O}$-ideal $\mathfrak{a}$ is *prime to* f if $\mathfrak{a} + f\mathcal{O} = \mathcal{O}$. This equivalent to the norm $(\mathcal{O} : \mathfrak{a})$ being relatively prime to f, since the injective[23] map $\mathcal{O}/\mathfrak{a} \to \mathcal{O}/\mathfrak{a}$ where $x \mapsto xf$ surjects precisely when $\mathfrak{a}$ is prime to f i.e. when f is coprime to the cardinal $(\mathcal{O} : \mathfrak{a})$ of the finite abelian group $\mathcal{O}/\mathfrak{a}$.

Every $\mathcal{O}$-ideal $\mathfrak{a}$ prime to f is proper since $f\mathcal{O}_K \subset \mathcal{O}$ and if $\alpha \in \mathrm{End}(\mathfrak{a})$ then $\alpha\mathcal{O} = \alpha(\mathfrak{a} + f\mathcal{O}) \subset \mathfrak{a} + f\mathcal{O}_K \subset \mathfrak{a} + \mathcal{O} \subset \mathcal{O}$ as wanted. The $\mathcal{O}$-ideals prime to f generate a subgroup $I(\mathcal{O}, f) \subset I(\mathcal{O})$ since the ideal norm is multiplicative. They also suffice to form the ideal class group $\mathrm{cl}(\mathcal{O})$

$$I(\mathcal{O}, f)/P(\mathcal{O}, f) \simeq \mathrm{cl}(\mathcal{O})$$

Moreover, there is an isomorphism[24]

$$\begin{array}{rcl} \eta : I(\mathcal{O}, f) & \longrightarrow & I(\mathcal{O}_K, f) \\ \mathfrak{a} \cap \mathcal{O} & \longleftarrow\!\shortmid & \mathfrak{a} \\ \mathfrak{b} & \longmapsto & \mathfrak{b}\mathcal{O}_K \end{array} \tag{2.2}$$

Now we are going to see how do the ring class fields and their Galois groups relate to each other.

2.8.2 Some Galois groups of ring class fields

Taking these results for granted, it follows that if $P = \eta(P(\mathcal{O}, f))$ we have a commutative diagram

$$\begin{array}{ccccccc} 1 & \longrightarrow & I(\mathcal{O}_K, f) \cap P(\mathcal{O}_K)/P & \longrightarrow & I(\mathcal{O}_K, f)/P & \longrightarrow & I(\mathcal{O}_K)/P(\mathcal{O}_K) \\ & & & & \downarrow \simeq & & \downarrow \simeq \\ & & & & \mathrm{cl}(\mathcal{O}) & \xrightarrow{\pi} & \mathrm{cl}(\mathcal{O}_K) \end{array}$$

Since the induced map $\pi : \mathrm{cl}(\mathcal{O}) \to \mathrm{cl}(\mathcal{O}_K)$ surjects it follows that

$$\ker \pi \simeq I(\mathcal{O}_K, f) \cap P(\mathcal{O}_K)/P$$

We also have the following exact sequence

$$1 \longrightarrow (\mathbb{Z}/f\mathbb{Z})^\times \longrightarrow (\mathcal{O}_K/f\mathcal{O}_K)^\times \longrightarrow \ker \pi \longrightarrow 1$$

[22] For the proofs see propositions 7.19-7.24 in [1].

[23] The map $\mathcal{O} \to \mathcal{O}$ where $x \mapsto xf$ injects.

[24] It also implies the unique factorization for ideals in $I(\mathcal{O}, f)$, since $\mathcal{O}_K$ is Dedekind.

and if we let $f = \ell$ be a prime ℓ which is inert in K, then the residue field for $\mathfrak{P} = \ell\mathcal{O}_K$ above ℓ is $\mathbb{F}_{\mathfrak{P}} = \mathcal{O}_K/\ell\mathcal{O}_K$ and the inertial degree is $[\mathbb{F}_{\mathfrak{P}} : \mathbb{F}_\ell] = 2$. Thus

$$\ker \pi \simeq (\mathcal{O}_K/f\mathcal{O}_K)^\times/(\mathbb{Z}/f\mathbb{Z})^\times = \mathbb{F}_{\ell O_K}^\times/\mathbb{F}_\ell^\times$$

Since theorem 2.16 implies $\mathrm{Gal}(K_1/K) \simeq \mathrm{cl}(\mathcal{O}_K)$ and $\mathrm{Gal}(K_f/K) \simeq \mathrm{cl}(\mathcal{O})$, by the naturality of these isomorphisms we have that the following diagram commutes

$$\begin{array}{ccccccccc} 1 & \longrightarrow & \mathrm{Gal}(K_f/K_1) & \longrightarrow & \mathrm{Gal}(K_f/K) & \longrightarrow & \mathrm{Gal}(K_1/K) & \longrightarrow & 1 \\ & & \vdots\simeq & & \downarrow\simeq & & \downarrow\simeq & & \\ 1 & \longrightarrow & \ker\pi & \longrightarrow & \mathrm{cl}(\mathcal{O}) & \longrightarrow & \mathrm{cl}(\mathcal{O}_K) & \longrightarrow & 1 \end{array}$$

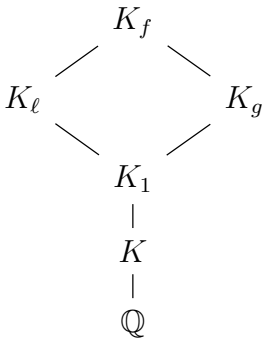
In particular, $\mathrm{Gal}(K_\ell/K_1) \simeq \ker\pi$ is cyclic with $\ell + 1$ elements since

$$|\mathrm{Gal}(K_\ell/K_1)| = |\ker\pi| = |\mathbb{F}_{\ell^2}^\times/\mathbb{F}_\ell^\times| = (\ell^2 - 1)/(\ell - 1) = \ell + 1$$

Moreover, if f is squarefree and we write $f = \prod_j \ell_j$ for distinct primes ℓ_j then there is an isomorphism

$$\mathrm{Gal}(K_f/K_1) \simeq \prod_j \mathrm{Gal}(K_{\ell_j}/K_1)$$

In particular, the extension has degree $[K_f/K_1] = \prod_j(\ell_j + 1)$. The general picture is that for $f = \ell g$ and $\ell \nmid g$ one has the following tower of extensions



3 Modular curves

In this section we define the modular curve in two ways, as a quotient of the upper half-plane modulo a certain subgroup Γ of $\mathrm{SL}_2(\mathbb{Z})$ and as an algebraic curve. The first approach is geometrical since $\mathrm{SL}_2(\mathbb{Z})$ is in fact a discrete subgroup of isometries of $\mathfrak{H}$ with the Poincaré metric, and it provides a formula for its genus in function of the subgroup Γ by applying results on Riemann surfaces. The second approach shows that the modular curve classifies somehow the isogenies between elliptic curves. This requires studying the field of functions that one can define on the curve.
In this section we followed [3, 10].

3.1 The modular curve

The *upper half plane* is $\mathfrak{H} = \{z \in \mathbb{C} : \Im z > 0\}$. The *special linear group* over $\mathbb{Z}$ is

$$\mathrm{SL}_2(\mathbb{Z}) = \left\{ \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) : a, b, c, d \in \mathbb{Z} \text{ and } ad - bc = 1 \right\}$$

and the following map defines an action of $\mathrm{SL}_2(\mathbb{Z})$ on $\mathfrak{H}$.

$$\begin{aligned} \mathfrak{H} \times \mathrm{SL}_2(\mathbb{Z}) &\longrightarrow \mathfrak{H} \\ (z, \gamma) &\longmapsto \gamma(z) := \tfrac{az+b}{cz+d} \end{aligned}$$

In particular, note that for $\gamma \in \mathrm{SL}_2(\mathbb{Z})$ the formula

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

shows that the action is well-defined. However, this action is not faithful. This leads one to consider the *modular group* which identifies elements in $\mathrm{SL}_2(\mathbb{Z})$ according to the action above. This is $\mathrm{SL}_2(\mathbb{Z})/\{\pm I\}$ where $-I = \left(\begin{smallmatrix} -1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$. Each element of the modular group represents a linear fractional transformation, and these can be extended to be the automorphisms of the Riemann sphere. In projective coordinates,

$$\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)(u : v) = (au + bv : cu + dv)$$

It can be shown that the modular group is generated by $\tau \mapsto \tau + 1$ and $\tau \mapsto \tau^{-1}$.
The *principal congruence subgroup* of level $N \geq 1$ is the following subgroup of $\mathrm{SL}_2(\mathbb{Z})$

$$\Gamma(N) = \{\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) : \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \equiv \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right) \bmod n\}$$

A subgroup $\Gamma \subset \mathrm{SL}_2(\mathbb{Z})$ is a *congruence subgroup* if there exists $N \geq 1$ with $\Gamma \supset \Gamma(N)$. It can be shown that the index $[\mathrm{SL}_2(\mathbb{Z}) : \Gamma(N)]$ is finite, hence $[\mathrm{SL}_2(\mathbb{Z}) : \Gamma] < \infty$ holds too. Given a congruence subgroup Γ, the *modular curve* of Γ is the set of orbits for the action of Γ on $\mathfrak{H}$, i.e. $Y(\Gamma) = \Gamma\backslash\mathfrak{H}$ together with the quotient topology induced by the map

$$\begin{aligned} \pi : \mathfrak{H} &\longrightarrow Y(\Gamma) \\ s &\longmapsto \Gamma s \end{aligned}$$

if we endow $\mathfrak{H}$ with the Euclidean topology. Here π is an open map because given an open subset $U \subset \mathfrak{H}$ we have

$$\pi^{-1}(\pi(U)) = \bigcup_{\gamma \in \Gamma} \gamma U = \Gamma U$$

And ΓU is an open set because the elements $\gamma \in \Gamma \subset \mathrm{SL}_2(\mathbb{Z})$ are homeomorphisms. It follows that $Y(\Gamma)$ is connected because π is continuous by construction and $\mathfrak{H}$ is connected.

Example 3.1. One important example is $Y(1) = \mathrm{SL}_2(\mathbb{Z})\backslash\mathfrak{H}$, since every point in $Y(1)$ represents a $\mathbb{C}$-isomorphism class of elliptic curves. Moreover, it turns out that if we let $\mathcal{F}$ be

$$\mathcal{F} = \{z \in \mathbb{C} : |z| \geq 1 \text{ and } -1/2 \leq \Re z \leq 0\} \cup \{z \in \mathbb{C} : |z| > 1 \text{ and } 0 < \Re z < 1/2\}$$

then $Y(1) = \mathcal{F}/\sim$ where $\sim$ here means identifying $z + 1/2 \equiv z - 1/2$ when $\Re z = 0$ and $z \equiv -z^{-1}$ when $|z| = 1$. [25]

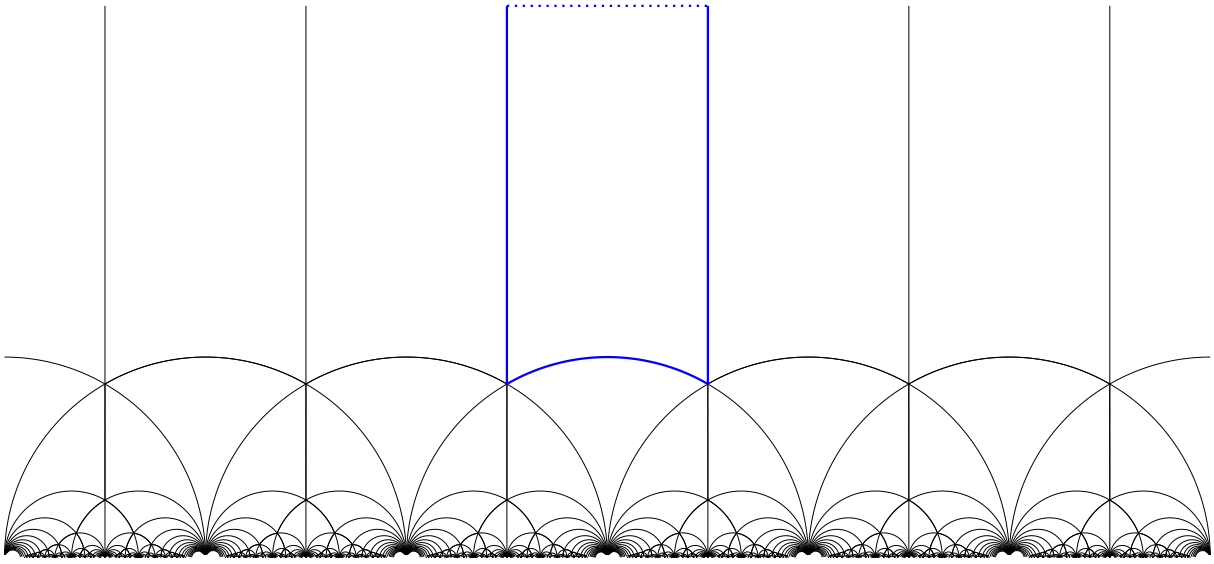


Figure 1: The fundamental domain $\mathcal{F}$, with some of its $\mathrm{SL}_2(\mathbb{Z})$-translates.

This follows from the moduli interpretation of $X_0(1)$ (theorem 3.14): every point in $Y(1)$ represents a pair of elliptic curves together with a degree-1 isogeny. Since degree-1 isogenies are isomorphisms because they are automatically surjective, it follows that every

[25] $\mathcal{F}$ is called a *fundamental domain* for $Y(1)$ because there exists a boundary identification $\sim$ with $Y(1) = \mathcal{F}/\sim$. In particular, the upper-half plane is tiled by the $\mathrm{SL}_2(\mathbb{Z})$-translates of $\mathcal{F}$, $\mathfrak{H} = \mathrm{SL}_2(\mathbb{Z}) \cdot \mathcal{F}$. See figure 1

point represents an isomorphism class of elliptic curves. Moreover, $Y(\Gamma)$ is a Riemann surface. To prove this one has to give local charts in every point of $Y(\Gamma)$. It is natural to consider the map obtained by sending $\Gamma s \mapsto s \in \mathfrak{H}$, and it will work locally on almost every point. However, this map need not to be well-defined for it may happen for instance that $\Gamma s = \Gamma s'$ for $s \neq s'$. The points where this approach does not work are either the *elliptic* points or the *cusps*, and require further study.

3.2 Elliptic points

Given $\tau \in \mathfrak{H}$ we define the *isotropy subgroup* of τ to be the set of elements in Γ that fix τ

$$\Gamma_\tau = \{\gamma \in \Gamma : \gamma(\tau) = \tau\}$$

Then $\tau \in \mathfrak{H}$ is an *elliptic point* for Γ if $\{\pm I\}\Gamma_\tau/\{\pm I\}$ is nontrivial. In other words, if the isotropy subgroup is nontrivial when we identify each element of Γ_τ with a linear fractional transformation. The cardinal $h_\tau^\Gamma = |\{\pm I\}\Gamma_\tau/\{\pm I\}|$ is called the *period of* τ and hence τ is elliptic if $h_\tau^\Gamma > 1$. Moreover, the period depends only on $\Gamma\tau$ because for any $\gamma \in \Gamma$ one has

$$\Gamma_{\gamma\tau} = (\gamma\Gamma\gamma^{-1})_{\gamma\tau} = \gamma\Gamma_\tau\gamma^{-1} \simeq \Gamma_\tau$$

Thus the period is defined for points in $Y(\Gamma)$. In fact, if $\Gamma \lhd \mathrm{SL}_2(\mathbb{Z})$ is normal subgroup the above is valid for any $\gamma \in \mathrm{SL}_2(\mathbb{Z})$ and the period is the same for all points in the orbit of τ when $\mathrm{SL}_2(\mathbb{Z})$ acts on $\mathfrak{H}$. It remains to find the elliptic points, and it is sufficient to find them when $\Gamma = \mathrm{SL}_2(\mathbb{Z})$ because Γ_τ is a subgroup of $\mathrm{SL}_2(\mathbb{Z})_\tau$.

Theorem 3.2. *The elliptic points for $SL_2(\mathbb{Z})$ are the elements in $SL_2(\mathbb{Z}) \cdot i$ or $SL_2(\mathbb{Z}) \cdot \zeta$ (the orbit of i and ζ respectively), where $\zeta^3 = 1, \zeta \neq 1$. Moreover, the isotropy subgroup is finite cyclic of order $1, 2$ or 3.*

There are congruence subgroups Γ for which there are no elliptic points, for instance $\Gamma = \Gamma_0(N)$ with $p|N$ for some prime $p \equiv -1 \bmod 12$ and

$$\Gamma_0(N) = \{\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) : \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \equiv \left(\begin{smallmatrix} * & * \\ 0 & * \end{smallmatrix}\right) \bmod n\}$$

It can be shown that for every point $\Gamma\tau \in Y(\Gamma)$ there exists a sufficiently small neighborhood U of τ containing at most one elliptic point (when τ is elliptic) such that the following map is a chart

$$\begin{aligned} \varphi : \pi(U) &\longrightarrow V \\ \Gamma u &\longmapsto \delta_\tau(u)^{h_\tau} \end{aligned}$$

where $\delta_\tau(u) = \frac{u-\tau}{u-\overline{\tau}}$[26] [27]. These charts define a complex atlas on $Y(\Gamma)$.

[26] The map centers the neighborhood by sending $\tau \mapsto 0$ and $\overline{\tau} \mapsto \infty$.

[27] This suggests that π sends the neighborhood of an elliptic point τ in to a neighborhood of $\pi(\tau)$ in which every point but τ has h_τ preimages.

3.3 Cusps and $X(\Gamma)$

Let $\mathfrak{H}^* = \mathfrak{H} \cup \mathbb{Q} \cup \{\infty\}$. Then $\mathrm{SL}_2(\mathbb{Z})$ acts on $\mathfrak{H}^*$ and the orbit of ∞ is $\mathbb{Q} \cup \{\infty\}$ because if $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ and $c \neq 0$ then $\gamma(\infty) = a/c$ is a rational number.
We define $X(\Gamma)$ as the set of orbits of the action of Γ on $\mathfrak{H}^*$

$$X(\Gamma) = \Gamma \backslash \mathfrak{H}^*$$

The *cusps* of Γ are the equivalence classes of $\mathbb{Q} \cup \{\infty\}$ under Γ. If we endow $\mathfrak{H}^*$ with a convenient topology we will be able to include $Y(\Gamma)$ in $X(\Gamma)$ with $X(\Gamma)$ being a compact space, and this fact will allow us to compute the genus after applying some tools for compact Riemann surfaces.
We define the open sets on $\mathfrak{H}^*$ to be the open sets in $\mathfrak{H}$ together with the sets $\gamma(N_m \cup \{\infty\})$ where γ runs over $\mathrm{SL}_2(\mathbb{Z})$ and

$$N_m = \{z \in \mathbb{C} : \Im z > m\}$$

which are the neighborhoods of ∞. The boundary of these sets $\gamma(N_m \cup \{\infty\})$ are either a line or a circle that is tangent to the real axis, because $\gamma \in \mathrm{SL}_2(\mathbb{Z})$ is a fractional linear transformation. Now $X(\Gamma)$ is $Y(\Gamma)$ together with the cusps of Γ

$$X(\Gamma) = \Gamma \backslash \mathfrak{H}^* = Y(\Gamma) \cup \Gamma \cdot (\mathbb{Q} \cup \{\infty\})$$

and endowed with the quotient topology given by $\pi^* : \mathfrak{H}^* \to X(\Gamma)$. This is adding only finitely many points (the cusps), because the index $[\mathrm{SL}_2(\mathbb{Z}) : \Gamma]$ is finite and

$$\mathrm{SL}_2(\mathbb{Z}) \cdot (\mathbb{Q} \cup \{\infty\}) = \mathbb{Q} \cup \{\infty\} = \left(\bigcup_i \Gamma \gamma_i\right) \cdot (\mathbb{Q} \cup \{\infty\})$$

In this process we have not lost any of the topological properties of $Y(\Gamma)$,

Theorem 3.3. *$X(\Gamma)$ is connected, Hausdorff and compact.*

The compactness of $X(\Gamma)$ follows from the compactness of $\mathcal{F}^* = \mathcal{F} \cup \{\infty\}$ [28] and the finiteness of the index $[\mathrm{SL}_2(\mathbb{Z}) : \Gamma]$, because if $\mathrm{SL}_2(\mathbb{Z}) = \bigcup_i \Gamma \gamma_i$ then

$$X(\Gamma) = \Gamma \cdot \mathfrak{H}^* = \Gamma \cdot (\mathrm{SL}_2(\mathbb{Z}) \cdot \mathcal{F}^*) = \bigcup_i \Gamma \gamma_i \mathcal{F}^*$$

and so $X(\Gamma)$ is compact for being a finite union of compacts.
To prove that $X(\Gamma)$ is still a (compact) Riemann surface, one has to give charts at the cusps. Recall that π has h_τ^Γ preimages in a neighborhood of an elliptic point. It turns out that π has infinitely many preimages in a neighborhood of a cusp [29]. $\mathrm{SL}_2(\mathbb{Z})_\infty$ is generated by $\tau \mapsto \tau + 1$, because if $\gamma(\infty) = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)(\infty) = \infty$ then $c = 0$ and $ad - bc = 1$

[28] Any open cover of $\mathcal{F}^*$ must contain a neighborhood of ∞ of the form N_m. Thus a finite open subcover is N_m together with a finite open subcover of $\mathcal{F}^* - N_m$, which is compact in the Euclidean topology.
[29] This does not enter in contradiction with the theory of Riemann surfaces, because $\mathfrak{H}^*$ is not compact.

implies that $a = d = \pm 1$ and thus $\gamma = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)^{\pm b}$ is a translation. In other words, the period is not defined at a cusp. The *width* of a cusp $s \in \mathbb{Q} \cup \{\infty\}$ is defined as

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

where $\delta(s) = \infty$. Then $h_{s,\Gamma}$ is counting the number of $\mathrm{SL}_2(\mathbb{Z})_\infty$-translates of $\mathcal{F}^*$ which are distinct under Γ-isotropy. At the cusps $\Gamma s \in X(\Gamma)$ the charts are

$$\begin{aligned} \varphi : \pi(U) &\longrightarrow V \\ \Gamma u &\longmapsto \exp(2\pi i \delta_s(u)/h_{s,\Gamma}) \end{aligned}$$

where $\delta_s(s) = \infty$. This transforms $\mathcal{F}^*$ into a $2\pi/h_{s,\Gamma}$ sector of the unit disk, sending $s \mapsto 0$. In short,

Theorem 3.4. *The topology and the charts treated above define a complex atlas for $X(\Gamma)$. Thus $X(\Gamma)$ is a compact Riemann surface.*

3.4 The genus of $X(\Gamma)$

A compact Riemann surface X is in particular a compact topological surface and every such space is homeomorphic to a g-holed torus, where g is the *genus* of X. One can recover the genus g_X of $X = X(\Gamma)$ for Γ a congruence subgroup by counting the cusps and the elliptic points of $X(\Gamma)$, by using some identities that apply to general compact Riemann surfaces.

If $f : X \to Y$ is a nonconstant holomorphic map between compact connected Riemann surfaces, it surjects because $\operatorname{im} f$ is both open (by the open mapping theorem for holomorphic functions) and closed ($\operatorname{im} f$ is compact in a Hausdorff space, thus closed). The *degree* d is defined as follows. For any $y \in Y$ let

$$d = \sum_{x \in f^{-1}(y)} e_x$$

where e_x is the *ramification degree* of f at x defined by the local expression of f given by local charts centered at x. That is

$$\varphi' \circ f \circ \varphi^{-1}(z) = z^{e_x} g(z)$$

for some g holomorphic with $g(0) \neq 0, \infty$. The degree is well-defined by the definition of e_x. Given a nonconstant holomorphic map f, the *Riemann-Hurwitz formula* relates the degree, the ramification degrees and the genus of X and Y

$$2g_X - 2 = d(2g_Y - 2) + \sum_{x \in X} (e_x - 1) \tag{3.1}$$

The map we are going to apply (3.1) is the natural projection

$$\begin{aligned} \Pi : X(\Gamma) &\longrightarrow X(1) \\ \Gamma\tau &\longmapsto \mathrm{SL}_2(\mathbb{Z})\tau \end{aligned}$$

which is an holomorphic map. The ramification index at an elliptic point is

$$|\{\pm I\}\mathrm{SL}_2(\mathbb{Z})_\tau : \{\pm I\}\Gamma_\tau|$$

because a local chart φ in $X(\Gamma)$ sends $\pi(\tau') \mapsto u = \delta_\tau(\tau')^{h_\tau^\Gamma}$, and the local expression of Π sends $u \mapsto \psi \circ \Pi \circ \varphi^{-1}(u) = \delta_\tau(\tau')^{h_\tau^{\mathrm{SL}_2(\mathbb{Z})}}$ and thus

$$\psi \circ \Pi \circ \varphi^{-1}(u) = u^{h_\tau^{\mathrm{SL}_2(\mathbb{Z})}/h_\tau^\Gamma}$$

as wanted. By a similar argument the ramification index at a cusp is

$$|\{\pm I\}\mathrm{SL}_2(\mathbb{Z})_s : \{\pm I\}\Gamma_s|$$

and clearly the ramification index is 1 at everywhere else. Denote by ϵ_h the number of elliptic points τ in $X(\Gamma)$ with $|\Gamma_\tau| = h$ where $h \in \{2, 3\}$, and ϵ_∞ the number of cusps in $X(\Gamma)$. Then by the well-definedness of the degree one has [30]

$$\sum_{x \in X(\Gamma)} (e_x - 1) = (d - \epsilon_\infty) + \frac{1}{2}(d - \epsilon_2) + \frac{2}{3}(d - \epsilon_3)$$

Then by (3.1) the genus of $X(\Gamma)$ is

$$g_X = 1 + \frac{d}{12} - \frac{\epsilon_2}{4} - \frac{\epsilon_3}{3} - \frac{\epsilon_\infty}{2}$$

because $X(1)$ has genus 0. This follows from the fact that the j-invariant

$$j : X(1) \to \hat{\mathbb{C}}$$

is an holomorphic isomorphism of compact Riemann surfaces.
We apply the formula above to $X(1)$, although we already know its genus.

$$g = 1 + \frac{1}{12} - \frac{1}{4} - \frac{1}{3} - \frac{1}{2} = 0$$

because $\Pi = \mathrm{id}_{X(1)}$ has degree 1 and $X(1)$ has only one cusp by the transitivity of $\mathrm{SL}_2(\mathbb{Z})$ on $\mathbb{Q} \cup \infty$ and two elliptic points, one of period 2 and another of period 3 (namely, i and ζ_3) because these are the only elliptic points in the fundamental domain $\mathcal{F}$ for $X(1)$ by theorem 3.2.

[30] Note that Π sends cusps to cusps and elliptic points to elliptic points, although some elliptic points in $X(1)$ may not come from elliptic points in $X(\Gamma)$.

3.5 Modular curves as algebraic curves

There is a general result on Riemann surfaces:

Theorem 3.5. *Every compact connected Riemann surface can be represented as an algebraic curve over* $\mathbb{C}$. [31]

Modular curves are Riemann surfaces, so they must also be represented as algebraic curves over $\mathbb{C}$. They are in fact curves over $\mathbb{Q}$:

Theorem 3.6. *There exists an algebraic curve* X *defined over* $\mathbb{Q}$ *such that the following is an isomorphism of Riemann surfaces*

$$X(\mathbb{C}) \simeq X_0(N)$$

where $X_0(N) = X(\Gamma_0(N))$ *is defined as in 3.3.*

There are at least two ways to show this. The first method shows that the field of functions of $X_0(N)$ is $\mathbb{C}(j, j_N)$ where $j_N = j(N\tau)$, and we devote sections 3.5.2,3.5.3 and 3.5.4 to this approach.
The second approach uses that a moduli space has model over a number field, which in this case can be shown to be $\mathbb{Q}$. All that is left to do is that $X_0(N)$ is a moduli space (section 3.6).
The first approach is however not easily generalizable to other contexts.

3.5.1 Algebraic curves

To begin we must define what is an algebraic curve. Let k be a field of characteristic 0 and $\overline{k}$ its algebraic closure. Let $f_1, \cdots, f_m \in k[X_1, \cdots, X_n]$ be a set of m polynomials and let I be the following ideal $I = \langle f_1, \cdots, f_m \rangle$. Let C be the following set

$$C = \{x \in \overline{k}^n : f(x) = 0 \text{ for all } f \in I\}$$

The *coordinate ring of* C *over* $\overline{k}$ is $\overline{k}[C] = \overline{k}[X_1, \cdots, X_n]/I$ and the *function field of* C *over* $\overline{k}$ denoted by $\overline{k}(C)$ is the field of fractions of $\overline{k}[C]$ when I is a prime ideal in $\overline{k}[X_1, \cdots, X_n]$, because to form a quotient field from a ring requires the ring to be an integral domain. The field of fractions is the set of rational functions on the coordinates of a point in C when we ignore elements in I. This makes sense because C is the vanishing set of I.
C is an *affine algebraic curve over* k if $\overline{k}(C)$ has transcendence degree 1 over $\overline{k}$ and $\overline{k}(C)/\overline{k}(t)$ is a finite algebraic extension, where t is a transcendence base.
C is *nonsingular* if the matrix $(D_j f_i(P))_{i,j}$ has rank $n-1$ for any $P \in C$. For nonsingular curves, one can prove that m_P/m_P^2 is a one dimensional vector space over $\overline{k}$ where $m_P = \{f \in \overline{k}[C] : f(P) = 0\}$ by showing that m_P/m_P^2 is isomorphic to the tangent space to

[31] Here an algebraic curve over $\mathbb{C}$ is loosely speaking a set given by $\varphi(x, y) = 0$ where $\varphi \in \mathbb{C}[x, y]$ is a bivariate polynomial. We will define them more precisely in section 3.5.1.

C at P, which is the kernel of $(D_j f_i(P))_{i,j}$ and this latter has dimension 1 because C is nonsingular. Moreover,

$$\overline{k}[C]_P = \{f/g \in \overline{k}(C) : g(P) \neq 0\}$$

is a local ring with maximal ideal

$$M_P = \{f/g \in \overline{k}[C]_P : f(P) = 0\}$$

and by using Nakayama's lemma (which can be used because $\overline{k}[C]_P$ is local and Noetherian and $M_P = m_P \cdot \overline{k}[C]_P$ implies that M_P is finitely generated) one can show that M_P is a principal ideal, generated by some t. This element t is called an *uniformizer at* P and it induces a valuation in $\overline{k}[C]$ and by extension on $\overline{k}(C)$

$$\begin{aligned} \nu_P : \quad k(C) &\longrightarrow \mathbb{Z} \cup \{\infty\} \\ f = t^e u &\longmapsto e \end{aligned}$$

Here $\nu_P(0) = \infty$ by definition. From now on C is a nonsingular affine algebraic curve over k.
Affine curves arise from *projective* curves when one chooses an affine chart in the projective space $\mathbb{P}^r$. Instead, to define a projective curve the polynomials f_i must be homogeneous in order to make $f_i(p) = 0$ a well-defined expression for $p \in \mathbb{P}^r$. Similarly, a nonsingular projective curve is such that every nonempty affine curve associated to every affine chart is nonsingular.
To define morphisms between curves we include them in an ambient space, which is the projective space $\mathbb{P}^r$. A *morphism* of curves $h : C \to C'$ is the map

$$h(P) = (t^{-\nu} h_0(P) : \cdots : t^{-\nu} h_r(P))$$

(where $\nu = \min_i \nu_P(h_i)$) associated to an element $h = (h_0 : \cdots, h_r) \in \mathbb{P}^r(\overline{k}(C))$ so that $h(C) \subset C'$.
Such a map is an *isomorphism* if there exists another morphism of curves h^{-1} with $h \circ h^{-1}$ and $h^{-1} \circ h$ being the identity maps on C' and C, respectively. This induces an equivalence relation. An stricter criterion is to be *isomorphic over* k, when h and h^{-1} are defined over k. A morphism h is *defined over* k if $h_i \in k(C) =$ the field of fractions of $k[X_j]_j/I_k$ where $I_k = I \cap k[X_j]_j$. Here $k(C)$ can be thought as a subfield of $\overline{k}(C)$ because $I_k \subset I$.[32]
A *function field* K over k is a field with $K \cap \overline{k} = k$ and a finite extension of $k(t)$ with t transcendental over k. We introduce another equivalence relation: two function fields K and K' are *conjugate over* k if there exists $\varphi : K \to K'$ that fixes k pointwise.
The following result from algebraic geometry gives a crucial correspondence:

[32]For instance, if $k = \mathbb{Q}$ the projective curves $C : X^2 + Y^2 = 2Z^2$ and $C' : X^2 + Y^2 = 3Z^2$ are isomorphic when $C, C' \subset \mathbb{P}^2(\overline{\mathbb{Q}})$, but not isomorphic over $\mathbb{Q}$: working modulo 4 or over the Gaussian integers one can easily show that C' has no $\mathbb{Q}$-points in $\mathbb{P}^2(\mathbb{Q})$ while C does (for instance $(1:1:1) \in C$) and if they were isomorphic over $\mathbb{Q}$, $\mathbb{Q}$-points would be mapped to $\mathbb{Q}$-points.

Theorem 3.7. *There is a bijection between the set of isomorphism over k classes of nonsingular projective curves over k and the set of conjugate over k classes of function fields over k, where the class of a curve C is mapped to the class of $k(C)$. Conversely, a function field K is mapped to some desingularization of the curve defined by $F(t,u)=0$ where $F(x,y)$ is obtained by clearing denominators from f, the minimal polynomial of u over $k(t)$ where $K=k(t,u)$.*

Proof. If two curves C, C' are isomorphic $C \overset{\varphi}{\simeq} C'$ over k then so are their function fields; consider the *pullback* $\varphi^* : k(C') \to k(C)$ where $\varphi^*(g) = g \circ \varphi$. It is an isomorphism since $(\varphi^{-1})^* = (\varphi^*)^{-1}$, and it also fixes k so $k(C)$ and $k(C')$ are conjugate over k. Thus the map is well-defined. It is also injective, but we will not show this.
To show it surjects take a function field K and suppose that $K/k(t)$ is separable. Then by the primitive element theorem we may write $K = k(u,t)$. Let $f(Y) = \text{minpoly}(u, k(t), Y)$ and define $F(X,Y)$ by clearing denominators from f. If F is not irreducible over $\overline{k}$ then it factors $F = F_1F_2$ for some $F_i \in \overline{k}[X,Y]$.
The coefficients of F_1, F_2 lie in a finite extension $k(\eta)$. Note that $g(Z) = \text{minpoly}(\eta, k, Z)$ is also irreducible in $k(t)[Z]$, otherwise by unique factorization the coefficients of a factorization of g lie in $k(t) \cap \overline{k} = k$. Now put $L = k(t,\eta)$ so that $[K : k(t)] = \deg(f)$ and $[L : k(t)] = \deg(g)$. Let $M = KL = k(u,t,\eta)$. By hypothesis F factors over $k(\eta)$ so $[M : L] < \deg(f)$ and

$$[M : k(t)] = [M : L][L : k(t)] < \deg(f)\deg(g)$$

Thus $[M : k(t)] = [M : K][K : k(t)]$ implies $[M : K] < \deg(g)$ so g factors $g = g_1g_2$ for $g_i \in K[Z]$. But g factors also in $\overline{k}$ and $g_i \in (K \cap \overline{k})[Z] = k[Z]$, a contradiction.
Thus F is irreducible over $\overline{k}$ and $C : F = 0$ defines a plane curve. If it is nonsingular we are done, otherwise we *desingularize* it. This process gives another curve which is nonsingular and birationally equivalent to C. Note that the function field remains the same, since any birational equivalence induces an automorphism of the function field. □

3.5.2 Automorphic forms

If $\gamma \in \text{SL}_2(\mathbb{Z})$ then we define the *weight-k operator* $[\cdot]_k$ as follows

$$f[\gamma]_k(\tau) = j(\gamma,\tau)^{-k} f(\gamma(\tau))$$

where $j(\gamma,\tau) = c\tau + d$. A function f is *weight-k invariant with respect to* $\Gamma \subset SL_2(\mathbb{Z})$ if $f[\gamma]_k = f$ for any $\gamma \in \Gamma$.
Let Γ be a congruence subgroup. An *automorphic form of weight k with respect to* Γ is a meromorphic weight-k invariant function $f : \mathfrak{H} \to \hat{\mathbb{C}}$ such that $f[\alpha]_k$ is meromorphic at ∞ for any $\alpha \in \text{SL}_2(\mathbb{Z})$ (that is, is meromorphic at the cusps of Γ). We denote the set of automorphic forms of weight k with respect to Γ by $\mathcal{A}_k(\Gamma)$ and define $\mathbb{C}(\Gamma) = \mathcal{A}_0(\Gamma)$.
A *modular form of weight k with respect to* Γ is an automorphic form of weight k with respect to $\text{SL}_2(\mathbb{Z})$ that is holomorphic, an stronger condition than meromorphic, and a

cusp form is a modular form that vanishes at the cusps of Γ. The set of modular forms is $\mathcal{M}_k(\Gamma)$ and the set of cusp forms of weight k is denoted by $\mathcal{S}_k(\Gamma)$.
For instance the j-invariant $j : X(1) \to \hat{\mathbb{C}}$ is an automorphic form of weight 0 with respect to $\mathrm{SL}_2(\mathbb{Z})$, so $\mathbb{C}(j) \subset \mathcal{A}_0(\Gamma(1))$. In fact

Theorem 3.8. *One has* $\mathbb{C}(\Gamma(1)) = \mathbb{C}(j)$

Proof. Let $f : X(1) \to \hat{\mathbb{C}}$ be an automorphic form for $\mathrm{SL}_2(\mathbb{Z})$. Then $f \circ j^{-1} : \hat{\mathbb{C}} \to \hat{\mathbb{C}}$ is an holomorphic map on the Riemann sphere $\hat{\mathbb{C}}$. But every holomorphic map $\hat{\mathbb{C}} \to \hat{\mathbb{C}}$ is a rational function, so $f \circ j^{-1}(z) = r(z)$ for some $r \in \mathbb{C}(z)$. Therefore $f \in \mathbb{C}(j(z))$ as wanted.
It follows also that if f is holomorphic on $Y(1)$ then it is a polynomial on j, because j is exhaustive and any factor in the denominator of r would create a pole. □

3.5.3 The modular polynomial

In other words, the function field of the Riemann sphere is just the rational functions $\mathbb{C}(t)$, and the isomorphism $j : X(1) \to \hat{\mathbb{C}}$ gives $\mathbb{C}(X(1)) = \mathbb{C}(j)$. We will now find $\mathbb{C}(\Gamma_0(N))$, but first a general theorem for $\mathbb{C}(\Gamma)$:

Theorem 3.9. *Let Γ be a congruence subgroup. Then $\mathbb{C}(\Gamma)$ is a finite extension of $\mathbb{C}(\Gamma(1))$, and its degree is at most $[\Gamma(1) : \Gamma]$.*

Proof. Choose n right coset representatives γ_i such that $\Gamma(1) = \bigsqcup_i \Gamma\gamma_i$ with $\gamma_1 = \mathrm{id}$. Let $f \in \mathbb{C}(\Gamma)$. Then the set $\{f_i\}_i$ where $f_i(\tau) = f(\gamma_i(\tau))$ is left invariant by $\tau \mapsto \gamma\tau$ for any $\gamma \in \Gamma(1)$, because f is Γ-invariant and $\gamma_i\gamma = \gamma_{i,j}\gamma_j$ for some j and $\gamma_{i,j} \in \Gamma$. Thus any symmetric function on the f_i is $\Gamma(1)$-invariant, i.e. is in $\mathbb{C}(j)$. Note that $f = f_1$. Then

$$P(Y) = \prod_i (Y - f_i)$$

is a polynomial $P \in \mathbb{C}(j)[Y]$ with $P(f) = 0$ of degree n, and f was arbitrary.[33] □

Theorem 3.10. *One has $\mathbb{C}(\Gamma_0(N)) = \mathbb{C}(j, j_N(t))$ where $j_N(\tau) = j(N\tau)$.*

Proof. j_N is meromorphic because j is meromorphic, but we still have to show that j_N is $\Gamma_0(N)$ invariant. Let $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma_0(N)$. Here is crucial that $N|c$. Then

$$j_N(\gamma\tau) = j(N\gamma\tau) = j\Big(\frac{Na\tau + bN}{c\tau + d}\Big) = j\Big(\frac{a(N\tau) + bN}{\frac{c}{N}(N\tau) + d}\Big)$$

$$= j(\gamma'(N\tau)) = j(N\tau) = j_N(\tau)$$

because $\gamma' = \left(\begin{smallmatrix} a & bN \\ c/N & d \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{Z})$, as wanted.

[33] The degree of any element is bounded by n and the extension is separable.

Again choose $n = [\Gamma(1) : \Gamma_0(N)]$ right coset representatives such that $\Gamma(1) = \bigsqcup_i \Gamma_0(N)\gamma_i$. Let $P \in \mathbb{C}(j)[Y]$ be the minimal polynomial of j_N over $\mathbb{C}(j)$. Then $f(\tau) = P(j(\tau), j_N(\tau))$ is the zero function, and for all i

$$f(\gamma_i\tau) = 0 = P(j(\gamma_i\tau), j_N(\gamma_i\tau)) = P(j(\tau), j_N(\gamma_i\tau))$$

Thus $j_N(\gamma_i\tau)$ is also a root of P. Suppose that $j_N(\gamma_k\tau) = j_N(\gamma_l\tau)$ for some τ. Then both $\alpha(N\gamma_k\tau)$ and $\beta(N\gamma_l\tau)$ are in $\mathcal{F}$ for some $\alpha, \beta \in \mathrm{SL}_2(\mathbb{Z})$ and by the injectivity of j one has $\alpha(N\gamma_k\tau') = \beta(N\gamma_l\tau')$ for all τ' in a neighborhood of τ. If $\alpha^{-1}\beta = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ then

$$\gamma_k\gamma_l^{-1} = \left(\begin{smallmatrix} 1/N & 0 \\ 0 & 1 \end{smallmatrix}\right)\alpha^{-1}\beta\left(\begin{smallmatrix} N & 0 \\ 0 & 1 \end{smallmatrix}\right) = \left(\begin{smallmatrix} a & b/N \\ cN & d \end{smallmatrix}\right) \in \Gamma_0(N)$$

so $k = l$, and $[\mathbb{C}(j, j_N) : \mathbb{C}(j)] = n$ has degree n. □

The *modular polynomial* Φ_N is the minimal polynomial of j_N over $\mathbb{C}(j)$. By the above it has degree $n = [\Gamma(1) : \Gamma_0(N)]$ and it can be written as

$$\Phi_N(Y) = \prod_i (Y - j_N(\gamma_i\tau)) = \sum_i \phi_{N,i}(j)Y^i$$

for some rational functions $\phi_{N,i}(j) \in \mathbb{C}(j)$ in j.

Remark 3.11. Each $\phi_{N,i}$ is a symmetric polynomial in $j_N(\gamma_i\tau)$, hence holomorphic in $\mathfrak{H}$ and $\Gamma(1)$-invariant, so it must be a polynomial in j by theorem 3.8. Thus we may write $\Phi_N(X, Y)$ as the polynomial obtained by replacing $j \mapsto X$ in $\Phi_N(Y) \in \mathbb{C}(j)[Y]$.

3.5.4 Properties of the modular polynomial

Throughout this section N is a prime number. Although the results remain to be true for elliptic curves over any field with characteristic not dividing N we expose only the cases for N prime, because the proofs for the general case are more involved but with the same ideas, i.e. writing down a set of coset representatives explicitly. In other words, we claim that

$$\left\{\gamma_k = ST^k = \left(\begin{smallmatrix} 0 & -1 \\ 1 & k \end{smallmatrix}\right)\right\}_{0 \le k < N}$$

is a set of right coset representatives of $\Gamma(1)/\Gamma_0(N)$, i.e. $\Gamma(1) = \bigsqcup_{0 \le k < N} \Gamma_0(N)\gamma_k$.
By theorem 3.7, $X_0(N)$ is a desingularization of the plane curve $\Phi_N(x, y) = 0$. First we show Φ_N has integer coefficients. This is relevant, for it lets one consider the reduction of the modular curve modulo a prime.

Theorem 3.12. *One has $\phi_{N,i}(X) \in \mathbb{Z}[X]$. Thus $\Phi_N(X, Y) \in \mathbb{Z}[X, Y]$.*

Proof. Consider the following expression for $\Phi_N(Y)$

$$\Phi_N(Y) = \Big(Y - j_N(\tau)\Big) \prod_{k=0}^{N-1} \Big(Y - j_N(\gamma_k\tau)\Big)$$

Note that $\phi_{N,i} \in \mathbb{C}[j]$ by remark 3.11. The q-expansion of the j-invariant has integer coefficients, i.e. $j(\tau) = q^{-1} + 744 + \sum_{n\geq 1} a_n q^n$ with $a_n \in \mathbb{Z}$, because $j(\tau) = 1728 g_2(\tau)^3 (g_2(\tau)^3 - 27 g_3(\tau)^2)^{-1}$ and Eisenstein series have integral q-expansions. Thus if we let $\zeta_N = e^{\frac{2\pi i \tau}{N}}$ one has

$$j_N(\gamma_k \tau) = j\Big(\frac{\tau + k}{N}\Big) = \sum_{n \geq -1} a_n q^{n/N} \zeta_N^{kn}$$

because $e^{2\pi i \frac{\tau+k}{N}} = q^{1/N}\zeta_N^k$. Hence $j_N(\gamma_k\tau) \in \mathbb{Q}(\zeta_N)[[q^{1/N}]]$ i.e. it is a $q^{1/N}$-series with coefficients in $\mathbb{Q}(\zeta_N)$. The Galois group $\mathrm{Gal}(\mathbb{Q}(\zeta_N)/\mathbb{Q})$ acts naturally on $\mathbb{Q}(\zeta_N)[[q^{1/N}]]$ coefficientwise. In particular, the set $S = \{j_N(\gamma_k\tau)\}$ is the union of two orbits because $j_N(t)$ is fixed, so $\phi_{N,i} \in \mathbb{Q}[[q^{1/N}]]$ for it is a symmetric function on S hence invariant by the action. Moreover, since $\phi_{N,i} \in \mathbb{C}[j]$ is a polynomial on j all of its q-powers must be integral i.e. $\phi_{N,i} \in \mathbb{Q}[[q]]$. On the other hand, as the coefficients $a_n\zeta_N^{kn}$ of $j_N(\gamma_k\tau)$ are algebraic integers so are the coefficients of $\phi_{N,i}$. Thus $\phi_{N,i} \in (\mathbb{Q} \cap \overline{\mathbb{Z}})[[q]] = \mathbb{Z}[[q]]$ since $\mathbb{Z}$ is integrally closed. Lemma 3.13 then implies $\phi_{N,i} \in \mathbb{Z}[j]$ is a polynomial with integer coefficients.

□

Lemma 3.13 (Hasse q-expansion principle)**.** *Let $f \in \mathbb{C}(\Gamma(1))$ be holomorphic in $\mathfrak{H}$ with $f \in A[[q]]$ i.e. the coefficients in its q-expansion lie in A for some $\mathbb{Z}$-module $A \subset \mathbb{C}$. Then $f \in A[j]$ is a polynomial in j with coefficients in A.*

Proof. Theorem 3.8 implies $f \in \mathbb{C}[j]$ i.e. $f = P(j)$ for some $P \in \mathbb{C}[X]$ so our goal is to show $P \in A[X]$. We proceed by induction on $d = \deg P$. The case $d = 0$ is clear, since f is constant. Suppose $d > 0$. Since $j(\tau) = q^{-1} + O(1)$ one has $f(\tau) = a_{-d}q^{-d} + O(q^{-d+1})$. Let $Q(X) = P(X) - a_{-d}X^d$ and $g(\tau) = Q(j(\tau))$. Then $g \in A[[q]]$ and $\deg Q < d$ so by induction hypothesis $Q \in A[X]$ and thus $P \in A[X]$. □

The second result says that the modular curve parametrizes somehow the isogenies of degree[34] N. This point of view is used in section 4 to construct Heegner points, which by the following theorem can also be seen to correspond to diagrams $\phi : E \to E'$ where E, E' are two elliptic curves and ϕ is a *cyclic* isogeny. A more precise statement is

Theorem 3.14. *A point $(j, j') \in \mathbb{C}^2$ is in $X_0(N) : \Phi_N(x, y) = 0$ if and only if there exist elliptic curves E, E' over $\mathbb{C}$ with j-invariants j and j' and a cyclic isogeny $E \to E'$ of degree N.*

An isogeny is *cyclic* if it has cyclic kernel. Our isogenies here will have degree N with N prime, hence cyclic.

Proof. With this result, we write down $\Phi_N(j, Y)$

[34] Here the degree is the cardinal of the kernel, because $\mathbb{C}$ is separable.

$$\Phi_N(j(\tau),Y) = (Y-j_N(\tau)) \prod_{0\le k<N} (Y-j_N(\gamma_k\tau)) = (Y-j(N\tau)) \prod_{0\le k<N} (Y-j\Big(\frac{\tau+k}{N}\Big)) \quad (3.2)$$

By the surjectivity of j we may put $j' = j(\tau')$. Thus $\Phi_N(j,j') = 0$ if and only if $\tau' \sim N\tau$ or $\tau' \sim \frac{\tau+k}{N}$, or equivalently if $\Lambda_{\tau'} \sim \Lambda_{N\tau}$ or $\Lambda_{\tau'} \sim \langle N, \tau+k\rangle$ as homothety of lattices. Note that these are sublattices of Λ_τ of index N. Conversely, a sublattice of Λ_τ of index N must be one of these: let $\Lambda' \subset \Lambda$ with $[\Lambda:\Lambda'] = N$ and d the smallest positive integer in Λ'. Then $\Lambda' = \langle d, a\tau+k\rangle$ with $ad = N$ and $0 \le k < N$. But N is prime so either $(a,d) = (1,N)$ or $(a,d) = (N,1)$, as wanted.
The natural inclusion $\Lambda' \subset \Lambda$ with $[\Lambda:\Lambda'] = N$ induces a cyclic isogeny $E_{\Lambda'} \to E_\Lambda$. Conversely, every degree N isogeny is induced by an inclusion of lattices $\Lambda' \subset \Lambda$ with $[\Lambda:\Lambda'] = N$. It is then cyclic because the kernel is isomorphic to Λ/Λ'. □

The existence of the dual isogeny implies that $\Phi_N(j,j') = 0$ if and only if $\Phi_N(j',j) = 0$. Moreover,

Theorem 3.15. *One has $\Phi_N(X,Y) = \Phi_N(Y,X)$. Moreover, $\Phi_N(X,X)$ has leading term $-X^{2N}$ i.e. is monic.*

Proof. Recall that $j_N(\gamma_0\tau) = j\Big(\frac{\tau}{N}\Big)$ is a root of $\Phi_N(j(\tau),Y)$. We also had $\Phi_N(j(\tau),j(N\tau)) = 0$ for all τ. Hence $j\Big(\frac{\tau}{N}\Big)$ is also a root of $\Phi_N(Y,j(\tau))$. $\Phi_N(j,Y)$ is irreducible in $\mathbb{C}(j)[Y]$ so it divides $\Phi_N(Y,j)$. Now for any fixed lattice $\Lambda \subset \mathbb{C}$ the equations $\Phi(j(\Lambda),Y) = 0$ and $\Phi_N(Y,j(\Lambda)) = 0$ have the same number of roots in $\mathbb{C}$. Hence $\Phi_N(j,Y)\cdot f = \Phi_N(Y,j)$ for some $f \in \mathbb{C}(j)$. Now $\mathbb{C}(j)/\mathbb{C}$ is transcendental so $\Phi_N(j,j) \neq 0$ and $f(j) = 1$, as wanted.
The j-invariant has the following q-expansion $j(\tau) = \frac{1}{q} + O(1)$, while $f(\tau) = \Phi_N(j(\tau),j(\tau))$ verifies

$$f(\tau) = \Big(j(\tau) - j(N\tau)\Big) \prod_{0\le k<N} \Big(j(\tau) - j\Big(\frac{\tau+k}{N}\Big)\Big) = -\frac{1}{q^{2N}} + \cdots$$

because $j(\tau) - j(N\tau) = q^{-1} - q^{-N} + \cdots$ and $j(\tau) - j\Big(\frac{\tau+k}{N}\Big) = q^{-1} - \zeta_N^{-k} q^{1/N} + \cdots$ where ζ_N is an N-th primitive root of unity. □

3.6 Modular curves as moduli spaces

The result of this section is quite similar to theorem 3.14; that modular curve parameterize somehow elliptic curves together with some additional data. The point of discussing this is that there is a definition of moduli space in terms of categories given by Mumford in the 1960s, and this categorical definition ensures that a moduli space is unique up to unique isomorphism.
The elements of our moduli space will be *enhanced elliptic curves*. An *enhanced elliptic curve* for $\Gamma_0(N)$ is a pair (E,C) where E is an elliptic curve and $C \subset E$ is a cyclic subgroup of order N (a N-cyclic group). Two enhanced elliptic curves (E,C) and (E',C')

are equivalent if there exists an isomorphism $E \xrightarrow{\simeq} E'$ sending C to C'. The set of equivalence classes of enhanced elliptic curves is denoted by $S_0(N)$.

$$S_0(N) = \{[(E, C)] : E \text{ elliptic curve and } C \subset E \text{ an } N\text{-cyclic subgroup}\}$$

Then $Y_0(N)$ (defined in section 3.1) is a moduli space under the following bijection

$$\begin{aligned} Y_0(N) &\longrightarrow S_0(N) \\ \Gamma_0(N)\tau &\longmapsto [(E_\tau, C_\tau)] \end{aligned}$$

Here $E_\tau = \mathbb{C}/\Lambda_\tau$ is a complex torus where $\Lambda_\tau = \mathbb{Z} \oplus \mathbb{Z}\tau$ and

$$C_\tau = \langle 1/N + \Lambda_\tau \rangle$$

is the N-cyclic subgroup of E_τ generated by $Q_\tau = 1/N + \Lambda_\tau$.
In other words,

Proposition 3.16. *Every enhanced elliptic curve is equivalent to $[(E_\tau, C_\tau)]$ for some $\tau \in \mathfrak{H}$, and the map above is a well-defined bijection.*

Proof. Well-definedness: If $\Gamma_0(N)\tau = \Gamma_0(N)\tau'$ then $\tau = \gamma\tau'$ for some $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \equiv \left(\begin{smallmatrix} * & * \\ 0 & * \end{smallmatrix}\right) \bmod N$. Let $m = c\tau' + d$ so that $m\tau = a\tau' + b$ and $mQ_\tau = (c\tau' + d)/N + \Lambda_{\tau'}$ and $m\Lambda_\tau = \Lambda_{\tau'}$ as before. But $(c, d) \equiv (0, *) \bmod N$ so $mQ_\tau = d/N + \Lambda_{\tau'}$ with $\overline{d} \in \left(\mathbb{Z}/N\mathbb{Z}\right)^\times$. This implies that $\mathbb{C}/\Lambda_\tau \xrightarrow{z \mapsto mz} \mathbb{C}/\Lambda_{\tau'}$ is an isogeny sending

$$\langle Q_\tau \rangle \mapsto \langle d/N + \Lambda_{\tau'} \rangle = \langle Q_{\tau'} \rangle$$

Surjectivity: Let (E, C) be an enhanced elliptic curve. We may suppose $E \simeq \mathbb{C}/\Lambda_{\tau'}$ for some $\tau' \in \mathfrak{H}$, so that C maps to $\langle Q \rangle$ for $Q = (c\tau' + d)/N$ under this isomorphism, but $\gcd(c, d, N) = 1$ so $ad - bc - kN = 1$ for some $a, b, k \in \mathbb{Z}$. Therefore $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})$ and we can suppose $\gamma \in \mathrm{SL}_2(\mathbb{Z})$ because the natural map $\mathrm{SL}_2(\mathbb{Z}) \to \mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})$ surjects and Q depends only on the classes of c and d modulo N. Let $\tau = \gamma\tau'$ and $m = c\tau' + d$ so that $m\tau = a\tau' + b$ and $m\Lambda_\tau = \Lambda_{\tau'}$. Then $\mathbb{C}/\Lambda_\tau \xrightarrow{z \mapsto mz} \mathbb{C}/\Lambda_{\tau'}$ is an isogeny sending $Q_\tau \mapsto Q$ so $(E_\tau, Q_\tau) \sim (E, C)$ as wanted.
Injectivity: Finally, suppose that $[(\mathbb{C}/\Lambda_\tau, \langle Q_\tau \rangle)] = [(\mathbb{C}/\Lambda_{\tau'}, \langle Q_{\tau'} \rangle)]$. Then $\mathbb{C}/\Lambda_\tau \xrightarrow{z \mapsto mz} \mathbb{C}/\Lambda_{\tau'}$ is an isogeny for some $m \in \mathbb{C}$ and $\left(\begin{smallmatrix} m\tau \\ m \end{smallmatrix}\right) = \gamma\left(\begin{smallmatrix} \tau' \\ 1 \end{smallmatrix}\right)$ for some $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{Z})$. Let $m = c\tau' + d$. Then $mC_\tau = C_{\tau'}$ implies $c \equiv 0 \bmod N$ hence $\overline{d} \in \left(\mathbb{Z}/N\mathbb{Z}\right)^\times$ and $\gamma \in \Gamma_0(N)$. This shows $\Gamma_0(N)\tau = \Gamma_0(N)\tau'$. □

4 Heegner points on $X_0(N)$

The last two sections are an ad hoc approach to class field theory for quadratic imaginary extensions, since we relied on the correspondence between elliptic curves and tori, and the fact that the endomorphism ring of an elliptic curve is an order in a quadratic imaginary field. We also constructed the ring class fields K_f analytically through the j-invariant. Moreover, the *First main theorem of CM* (theorem 2.16) is just a weakening of *Artin reciprocity law* (theorem 1.4) and while we proved theorem 2.16 using properties of elliptic curves, theorem 1.4 usually requires Tate cohomology, which we will not discuss here.
In this section, by using the *Modularity theorem* we will show that with this approach one can produce algebraic points on elliptic curves over $\mathbb{Q}$, the *Heegner points*.
To explore all of these results, we will need to take others for granted since discussing them in full detail is out of the scope of this thesis.
We consulted [4, 7, 2, 8].

4.1 The modularity theorem

The following is a very powerful result

Theorem 4.1 (Modularity theorem)**.** *Let $E/\mathbb{Q}$ be an elliptic curve. Then there exists*[35] *N and a surjective morphism over $\mathbb{Q}$ of curves over $\mathbb{Q}$*

$$\varphi : X_0(N)_{alg} \longrightarrow E$$

Here $X_0(N)_{\text{alg}}$ denotes the algebraic curve over $\mathbb{Q}$ underlying the Riemann surface $X_0(N)$ which is given as the desingularization of the projective closure of the affine model

$$\Phi_N(x, y) = 0$$

The morphism φ is called a modular parametrization of E.

In other words, every elliptic curve defined over $\mathbb{Q}$ can be parameterized by some modular curve. Although it was thought to be inaccessible, an special case of theorem 4.1 proven by Taylor and Wiles in 1994 was crucial to prove *Fermat's last theorem*. The full proof was completed in 2001 by Breuil *et al.*
The application of theorem 4.1 that we are interested in is to produce points of infinite order on an elliptic curve, known as *Heegner points*. With these points one can prove some special cases of the *Birch and Swinnerton-Dyer conjecture*.

4.2 The L-function

Given an elliptic curve E over a number field K, the rank of the abelian group $E(K)$ is known to be finite:

[35] The minimal N for which theorem 4.1 holds is now known to be the *conductor* of E.

Theorem 4.2 (Mordell-Weil)**.** *If E and K are as above then one has*

$$E(K) \simeq \mathbb{Z}^r \oplus T$$

where T is a finite abelian group and r is the rank of $E(K)$.

The torsion subgroup $T = E(K)_{\text{tors}}$ for the case $K = \mathbb{Q}$ can be effectively found by using Nagell-Lutz theorem. However, there is no known algorithm for computing the rank r. To be true, the BSD conjecture would solve this problem by giving an effective procedure to determine r. To state this conjecture we introduce the L-function, although discussing more of its properties are out of the scope of this thesis. The L-function encodes all the possible *reductions* of our elliptic curve E, so we have to define *reduction*.

4.2.1 Reduction over local fields

Here L is an arbitrary local field, for instance the p-adic numbers $\mathbb{Q}_p$. To reduce an elliptic curve E over $\mathbb{Q}$ in L is to consider the equation defining E as an equation over L. In order to do this, one first has to find a *good* equation for E that can be reduced over L. Let R be the ring of integers of L, π an uniformizer i.e. a generator of the maximal ideal $\mathfrak{m}$ and $k = R/\mathfrak{m}$ the residue field.
In section 2.3 we claimed that every elliptic curve over $\mathbb{C}$ admits an Weierstrass equation of the form

$$E_\Lambda : Y^2 = 4X^3 - g_2X - g_3$$

If E is over an arbitrary field K, by the Riemann-Roch theorem every elliptic curve can be put as follows

$$\mathcal{W} : Y^2Z + a_1XYZ + a_3YZ^2 = X^3 + a_2X^2Z + a_4XZ^2 + a_6Z^3$$

Then $\mathcal{W}$ is a *minimal Weierstrass model* for E over R if $a_i \in R$ and the discriminant has minimal π valuation i.e. $\nu_\pi(\Delta(\mathcal{W}))$ is minimal. The *reduction* $\tilde{E}$ of E over k is the algebraic subset of $\mathbb{P}^2(k)$ defined by $\mathcal{W} \otimes_R k$ i.e. it is obtained by reducing the a_i modulo $\mathfrak{m}$. The reduction is *good* if the discriminant is nonzero over k, and *bad* otherwise.
There are three cases of bad reduction: additive, split multiplicative or nonsplit multiplicative. These can be described geometrically since an irreducible cubic has at most one singularity, which is either a cusp (additive reduction) or a node (multiplicative reduction). In the split multiplicative reduction case the tangents at the node are defined over k, and in the nonsplit case they are not.
In particular, if we are given $E/\mathbb{Q}$ and we let $L = \mathbb{Q}_p$ then a minimal model always exists, and the *conductor* of E is an integer divisible only by the primes p for which the reduction $\tilde{E}$ of $E/\mathbb{Q}$ modulo $k = \mathbb{Z}_p/p\mathbb{Z}_p \simeq \mathbb{Z}/p\mathbb{Z}$ is bad.

4.2.2 Definition

Let E/K be an elliptic curve and ν a nonarchimedean place of K. Then the completion K_ν of K at ν is a local field[36]. Since $k_\nu = R_\nu/\mathfrak{m}_\nu$ is a finite field we may let $q_\nu = \#k_\nu$

[36]Here $K_\nu, R_\nu, \mathfrak{m}_\nu, k_\nu$ and π_ν are as in section 4.2.1 and depend on the place ν.

and $a_\nu = q_\nu + 1 - \#\tilde{E}_\nu(k_\nu)$ where $\tilde{E}_\nu$ is the reduction of E over k_ν. The local factor is defined as follows, depending on the reduction type of E at ν:

$$L_\nu(T) = \begin{cases} 1 - a_\nu T + q_\nu T^2 & \text{if } E \text{ has good reduction at } \nu, \\ 1 - T & \text{if } E \text{ has split multiplicative reduction at } \nu, \\ 1 + T & \text{if } E \text{ has nonsplit multiplicative reduction at } \nu, \\ 1 & \text{if } E \text{ has additive reduction at } \nu \end{cases}$$

Then the L-function of E/K is

$$L_{E/K}(s) = \prod_\nu L_\nu(q_\nu^{-s})^{-1}$$

where ν runs over the nonarchimedean places of K. The bound $|a_\nu| \leq 2\sqrt{q_\nu}$ implies that the expression defining $L_{E/K}(s)$ converges for $\Re s > 3/2$. Since $L_{E/K}(s)$ can be extended to a meromorphic function in $\mathbb{C}$ one can consider the behavior of $L_{E/K}(s)$ at $s = 1$. This fact was[37] unknown when Birch and Swinnerton-Dyer posed their conjecture:

Conjecture 4.3 (Birch and Swinnerton-Dyer)**.** *Let E/K be an elliptic curve of rank r. Then $L_{E/K}(s)$ has order r at $s = 1$.*

There is also an stronger version of BSD which describes also the coefficient

$$\lim_{s \to 1} L_{E/K}(s)(s-1)^{-r}$$

in terms of additional arithmetical data of E/K, for instance the order of the *Tate-Shafarevich group* $Ш(E/K)$ that we will define in section 4.5.1.

4.3 Modular forms and L-series

We discuss some results on modular forms[38] and their connections with L-functions.
Recall the action of $\mathrm{SL}_2(\mathbb{Z})$ on $\mathfrak{H}$ induced an action on the set of meromorphic functions on $\mathfrak{H}$, and modular forms were invariant up to multiplication by a factor $(c\tau + d)^{-k}$. We now extend this definition as follows. Let

$$\mathrm{GL}_2^+(\mathbb{Q}) = \left\{ \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) : a, b, c, d \in \mathbb{Q} \text{ and } ad - bc > 0 \right\}$$

Then we redefine the *weight-k operator* as follows

$$f[\gamma]_k(\tau) = (\det \gamma)^k (c\tau + d)^{-k} f(\gamma(\tau))$$

This is indeed a generalization since in $\mathrm{SL}_2(\mathbb{Z})$ one has $\det \gamma = 1$.
Throughout this section we put $\mathcal{M}_k(N) = \mathcal{M}_k(\Gamma_0(N))$ and $\mathcal{S}_k(\Gamma_0(N)) = \mathcal{S}_k(N)$ for convenience.

[37] It was only known for E with CM or for $E/\mathbb{Q}$ having a modular parametrization.
[38] We defined them in section 3.5.2.

4.3.1 Hecke operators

Let $M(n,N)$ be the following set

$$M(n,N) = \left\{ \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) : a,b,c,d \in \mathbb{Z}, ad-bc = n, N|c \text{ and } \gcd(a,N) = 1 \right\}$$

Then $\Gamma_0(N)$ acts on $M(n,N)$ and we can consider representatives α_i of the finitely many right cosets $\{\Gamma_0(N)\alpha_i\} = \Gamma_0(N)\backslash M(n,N)$. Given $f \in \mathcal{M}_k(N)$ we define $T_k(n)$ by

$$T_k(n)f = n^{\frac{k}{2}-1} \sum_i f \circ [\alpha_i]_k$$

The operators $T_k(n)$ are the *Hecke operators* for $\Gamma_0(N)$, here defined through their action on modular forms. They are $\mathbb{C}$-linear and preserve modular and cusp forms:

Theorem 4.4. *For any n one has $T_k(n)\mathcal{M}_k(N) \subset \mathcal{M}_k(N)$ and $T_k(n)\mathcal{S}_k(N) \subset \mathcal{S}_k(N)$.*

Moreover, one can endow $\mathcal{S}_k(N)$ with an Hermitian inner product. The *Petersson inner product* on $\mathcal{S}_k(N)$ is Hermitian and defined by[39]

$$\langle f, g \rangle = \int_{\mathcal{F}} f(\tau)\overline{g(\tau)}\sigma^k \frac{d\rho d\sigma}{\sigma^2}$$

where $\mathcal{F}$ is a fundamental domain for $\Gamma_0(N)$ and $\tau = \sigma + i\rho$. An *eigenform* f is a cusp form which is an eigenvector of T_n for all n. They suffice to generate $\mathcal{S}_k(N)$:

Theorem 4.5. *There is a basis $\{f_i\}_i$ of simultaneous eigenforms for $\mathcal{S}_k(N)$ i.e. for all n*

$$T_k(n)f_i \in \mathbb{R} \cdot f_i$$

The eigenvalues are real because the Hecke operators are Hermitian with respect to the Petersson inner product. The algebra generated by the Hecke operators is the *Hecke algebra*.
The Hecke operators for $\Gamma_0(N)$ can also be defined in a natural way on $X_0(N)$. Since it can be shown the T_p generate the full Hecke algebra it suffices to define T_p for p prime. By theorem 3.16 every point $P \in X_0(N)$ is of the form $P = (E, C)$ where E is an elliptic curve and C a cyclic subgroup of order N, or also as (E, ϕ) where $\phi : E \to E'$ is a degree N isogeny by theorem 3.14. We define T_p by

$$T_p((E,C)) = \sum_{\phi} (E', \phi(C))$$

where ϕ runs over the degree p isogenies $\phi : E \to E'$.

[39] It is well-defined since the integral can be taken over a compact set and cusp forms have good behaviour at the cusps of $\Gamma_0(N)$. Note also that $\sigma^{-2}d\rho d\sigma$ is the Poincaré metric for $\mathfrak{H}$ and $\Gamma_0(N) \subset \mathrm{SL}_2(\mathbb{Z})$ preserves it.

4.3.2 Oldforms and newforms

Some forms in $\mathcal{S}_k(N)$ come from forms in $\mathcal{S}_k(M)$, in the following sense. If $M|N$ and $d|\frac{N}{M}$ then for any $f \in \mathcal{M}_k(M)$ the function $g(\tau) = f(d\tau) = \iota_{M,N,d}(f)$ verifies $g \in \mathcal{M}_k(N)$. The map $\iota_{M,N,d}$ restricts to $\iota_{M,N,d} : \mathcal{S}_k(M) \to \mathcal{S}_k(N)$ since the natural map $X_0(M) \to X_0(N)$ sends cusps to cusps. The set of *oldforms* is

$$\mathcal{S}_k(N)^{\text{old}} = \bigcup_{M|N \text{ and } d|\frac{M}{N}} \text{im}\, \iota_{M,N,d}$$

and the set of *newforms* is the orthogonal complement $\mathcal{S}_k(N)^{\text{new}} = \left(\mathcal{S}_k(N)^{\text{old}}\right)^{\perp}$.

4.3.3 The Jacobian of a compact Riemann surface

Suppose X is a compact connected Riemann surface, for instance $X = X_0(N)$. Then by singular homology over $\mathbb{Z}$ we know that $H_0(X) \simeq \mathbb{Z} \simeq H_2(X)$ and $H_1(X) \simeq \mathbb{Z}^{2g}$ where g is the genus of X. It turns out $H_1(X)$ can be thought as a subgroup of the set of meromorphic differentials $\Omega^1(X)$, a family of objects which can be integrated that we define now.
For an open subset $V \subset \mathbb{C}$ a *local meromorphic differential* of degree n on V is an expression of the form $f(z)(dz)^n$ where f is meromorphic on V and z is a local variable on V. The set of local meromorphic differentials on V is denoted by $\Omega^n(V)$.
Let $\varphi_j : U_j \to V_j$ be the coordinate charts of X where $j \in J$. Then a *meromorphic differential* w of degree n on X is an element $w \in \prod_{j\in J} \Omega^n(V_j)$ verifying $\varphi_{k,j}^*(w_k|_{V_{k,j}}) = w_j|_{V_{j,k}}$ for all j, k where $\varphi_{k,j}^*$ is the pullback of the transition map $\varphi_{k,j} : V_{j,k} \to V_{k,j}$ and $V_{j,k} = \varphi_j(U_j \cap U_k), V_{k,j} = \varphi_k(U_j \cap U_k)$. This is essentially restricting w to be locally integrable, regardless of the chosen coordinate chart. We denote by $\Omega^1_{\text{hol}}(X)$ the set of holomorphic differentials on X, in which the local meromorphic differentials are holomorphic.
The reason to introduce meromorphic differentials will become clear now, since the dual space $\Omega^1_{\text{hol}}(X)^\wedge = \text{Hom}(\Omega^1(X), \mathbb{C})$ and $H_1(X)$ are related in the following sense. Recall that $H_1(X) \simeq \mathbb{Z}^{2g}$ so there are $2g$ closed loops $\gamma_1, \cdots, \gamma_{2g}$ in X whose homology class generate $H_1(X)$. If we take a chain $c = \sum_{i=1}^{2g} c_i\gamma_i \in H_1(X)$ and a 1-form $w \in \Omega^1_{\text{hol}}(X)$ then the integral

$$\int_c w := \sum_{i=1}^{2g} c_i \int_{\gamma_i} w \in \mathbb{C}$$

is a well-defined complex number. Thus the operator $\int_c$ which integrates 1-forms over c is an element of $\Omega^1_{\text{hol}}(X)^\wedge$. Moreover, it can be shown that $\Omega^1_{\text{hol}}(X)^\wedge \simeq \bigoplus_{i=1}^{2g} \mathbb{R} \int_{\gamma_i}$ and that $H_1(X)$ naturally identifies under the correspondence $c \mapsto \int_c$ to

$$H_1(X) \simeq \bigoplus_{i=1}^{2g} \mathbb{Z} \int_{\gamma_i}$$

Thus there is a natural inclusion $H_1(X) \subset \Omega^1_{\text{hol}}(X)^\wedge$. The *Jacobian* of X is then defined by

$$\text{Jac}(X) = \Omega^1_{\text{hol}}(X)^\wedge / H_1(X)$$

As an alternate motivation to introduce the Jacobian consider the following. If X is an elliptic curve then we are in the genus $g = 1$ case and X is an abelian group, but for $g > 1$ this does not longer hold. The $\text{Jac}(X)$ is an abelian group obtained from X, which in the $g = 1$ case is isomorphic to X.
Abel's theorem gives an easier description of $\text{Jac}(X)$ in terms of classes of divisors in the Picard group. The *Picard group* is $\text{Pic}(X) = \text{Div}(X)/\text{Pr}(X)$ where

$$\text{Div}(X) = \Big\{ \sum_{x \in X} n_x x \text{ with } n_x \neq 0 \text{ for finitely many } x \Big\}$$

is the set of divisors on X and $\text{Pr}(X) = \{(f) \text{ where } f \in \mathbb{C}(X)\}$ is the set of principal divisors. Here $\mathbb{C}(X)$ is the function field of X and (f) is the principal divisor associated to f. Then

Theorem 4.6 (Abel)**.** *Fix a point $x_0 \in X$ and let $\text{Div}^0(X)$ be the kernel of the degree map i.e. the divisors with $\sum_{x \in X} n_x = 0$. Then the map $\text{Div}^0(X) \to \text{Jac}(X)$ sending $\sum_{x \in X} n_x x \mapsto \sum_{x \in X} n_x \int_{x_0}^{x}$ induces an isomorphism*

$$\text{Pic}^0(X) = \text{Div}^0(X)/\text{Pr}(X) \simeq \text{Jac}(X)$$

The Jacobian is also a complex torus $(S^1)^g$, since $\mathbb{Z}^{2g}$ is a lattice in $\mathbb{R}^{2g}$. It can be shown that since the modular curve $X_0(N)$ is defined over $\mathbb{Q}$ so is $\text{Jac}(X_0(N))$. This is important for the next section.
We choose an inclusion $X_0(N) \hookrightarrow \text{Jac}(X_0(N))$ that sends every point x to the divisor class $(x) - (\infty)$ where ∞ is the cusp of $X_0(N)$ at ∞. This inclusion induces the Hecke operators on $X_0(N)$ to the Jacobian of $X_0(N)$.

4.3.4 Eichler-Shimura theorem

The reason to introduce Hecke operators and the Jacobian is to discuss the Eichler-Shimura theorem. This result is related to the Modularity theorem since it implies the existence of a modular parametrization under certain conditions. To every cusp form f having a q-expansion $f = \sum_{n=1} c_n q^n$ with $c_1 = 1$ we can associate the L-function

$$L(f, s) = \sum_{n=1} \frac{c_n}{n^s}$$

The Eichler-Shimura theorem shows that if $f \in \mathcal{S}_2(N)$ then $L(f, s)$ is the L-function of an elliptic curve E_f over $\mathbb{Q}$ (such L-functions appeared in section 4.2.2):

Theorem 4.7 (Eichler-Shimura). *Let $f(\tau) = \sum_{n=1}^{\infty} c_n q^n$ be a normalized cusp form of weight 2 i.e. with $c_1 = 1$ and $f \in \mathcal{S}_2(N)$. Suppose the $c_n \in \mathbb{Z}$ are integers. Then there exists a pair (E_f, ν) where $E_f/\mathbb{Q}$ is an elliptic curve defined over $\mathbb{Q}$ and a surjective morphism of curves $\nu : \mathrm{Jac}(X_0(N)) \longrightarrow E_f$ defined over $\mathbb{Q}$ such that*

1. *The Hecke operators on $\mathrm{Jac}(X_0(N))$ leave $\ker \nu$ stable and act on E_f as multiplication by c_n.*
2. *The pullback $\nu^*(w)$ of the invariant differential w of E_f is a nonzero multiple of the differential form on $\mathrm{Jac}(X_0(N))$ defined by $f(\tau)d\tau$*
3. *The L-functions of the newform and E_f agree: $L_{E_f/\mathbb{Q}}(s) = L(f, s)$*

If we compose the inclusion $X_0(N) \hookrightarrow \mathrm{Jac}(X_0(N))$ with the map $\nu : \mathrm{Jac}(X_0(N)) \to E_f$ we obtain a modular parametrization $\varphi : X_0(N) \to E_f$ since φ is defined over $\mathbb{Q}$. It can be shown that N is the conductor of E_f. There is also a converse by Wiles:

Theorem 4.8. *If $E/\mathbb{Q}$ has conductor N then there exists a newform $f \in \mathcal{S}_2(N)$ with $L_{E/\mathbb{Q}}(s) = L(f, s)$.*

4.4 Heegner points

Let E be an elliptic curve over $\mathbb{Q}$ without CM, and fix a modular parametrization $\varphi : X_0(N) \to E$. Note that φ exists by theorem 4.1. Let $K = \mathbb{Q}(\sqrt{D})$ an imaginary quadratic field of discriminant D with[40] $D \neq -3, -4$ and such that every prime $p|N$ splits completely in $\mathcal{O}_K$. This condition on N is known as *Heegner's hypothesis*, and it imposed so that the following construction is valid.

There exists $\mathfrak{N}$ with $N\mathcal{O}_K = \mathfrak{N}\overline{\mathfrak{N}}$.[41] Let $\mathcal{O}$ be an order of K with conductor f coprime with N, and let $\mathfrak{M} = \mathcal{O} \cap \mathfrak{N}$. Recall that we can write $\mathcal{O}_K = \langle 1, w_K \rangle$ and $\mathcal{O} = \langle 1, w \rangle$, where w_K is as in section 2.1.1 and $w = f w_K$. Then $\mathfrak{M}$ is proper $\mathcal{O}$-ideal[42] and the lattice inclusion $\mathcal{O} \subset \mathfrak{M}^{-1}$ induces a degree N cyclic[43] isogeny $\phi : E_{\mathcal{O}} \to E_{\mathfrak{M}^{-1}}$. By theorem 3.14 the point $x_f = (j(\mathcal{O}), j(\mathfrak{M}^{-1}))$ belongs to the modular curve $X_0(N)$. We write x_f to remark that x_f depends only on the conductor f. Moreover, both $\mathfrak{M}^{-1}$ and $\mathcal{O}$ have CM by $\mathcal{O}$ because they are proper fractional $\mathcal{O}$-ideals, so by theorem 2.14 the coordinates of x_f are algebraic integers, i.e. $x_f \in \overline{\mathbb{Q}}$. In fact, $x_f \in L$ where $L = K_f$ is the splitting field of the Hilbert class polynomial $H_D(X)$ of section 2.6.

The constructed point $x_f \in X_0(N)$ is called a *Heegner point*. The modular parametrization φ is defined over $\mathbb{Q}$ and x_f has coordinates in K_f, so if we let $y_f = \varphi(x_f)$ then y_f is defined over K_f and $y_f \in E$, i.e. $y_f \in E(K_f)$. Then if we take the trace of y_f we obtain

$$y_{f,K} = \mathrm{Tr}_{K_f/K}(y_f) \tag{4.1}$$

[40] Then $\mathcal{O}_K^{\times} = \{\pm 1\}$.

[41] That is, if $N = \prod_i p_i^{e_i}$ then $p_i\mathcal{O}_K = \mathfrak{p}_i\overline{\mathfrak{p}_i}$ and $N\mathcal{O}_K = \mathfrak{N}\overline{\mathfrak{N}}$ where $\mathfrak{N} = \prod_i \mathfrak{p}_i^{e_i}$. By the Chinese Remainder theorem, $\mathcal{O}_K/\mathfrak{N} \simeq \prod_i \mathcal{O}_K/\mathfrak{p}_i^{e_i} \simeq \prod_i \mathbb{Z}/p^{e_i}\mathbb{Z} \simeq \mathbb{Z}/N\mathbb{Z}$.

[42] Because $\mathfrak{M} = \mathfrak{N} \cap \mathcal{O}$ is proper by the isomorphism in (2.2).

[43] Since $(f, N) = 1$ the natural morphism $\mathcal{O} \to \mathcal{O}_K/\mathfrak{N}$ surjects so $\mathcal{O}/\mathfrak{M} \simeq \mathcal{O}_K/\mathfrak{N}$

But y_f is $\mathrm{Gal}(K_f/K)$-invariant if we consider the natural action of $\mathrm{Gal}(K_f/K)$ on $E(K_f)$, so it is defined over K, i.e. $y_f \in E(K)$.

4.5 Group cohomology

We introduce some key concepts and define cohomology in simple terms, via cochains. For a group G, the *group ring* of G is the ring

$$\mathbb{Z}[G] = \bigoplus_{g \in G} \mathbb{Z}g$$

with the natural sum and the product defined by $(\sum_g a_g g)(\sum_h b_h h) = (\sum_{g,h} a_g b_h gh)$. Then an abelian group A is a *G-module* if it is a $\mathbb{Z}[G]$-module.
The *group of i-cochains of G with coefficients in A* is the abelian group $\mathcal{C}^i(G,A) = \{f : G^i \to A\}$, and the *$i$-th differential* is the map $d^i : \mathcal{C}^i(G,A) \to \mathcal{C}^{i+1}(G,A)$ defined by

$$d^i(f)(g_0, \cdots, g_i) = g_0 f(g_1, \cdots, g_i) + \sum_{j=1}^{i} (-1)^i f(g_0, \cdots, g_{j-1}g_j, \cdots g_i) + (-1)^{i+1} f(g_0, \cdots, g_{i-1})$$

The set of *i-cocycles* is $\mathcal{Z}^i(G,A) = \ker d^i$, and the set of *$i$-coboundaries* is $\mathcal{B}^i(G,A) = \operatorname{im} d^{i-1}$.[44] Moreover, the sequence abelian groups

$$0 \longrightarrow \mathcal{C}^0(G,A) \xrightarrow{d^0} \mathcal{C}^1(G,A) \xrightarrow{d^1} \cdots$$

forms a *cochain complex*, i.e. $d^{i+1} \circ d_i = 0$. That is, every coboundary is a cocycle[45] so one can define the *i-th cohomology group* by $H^i(G,A) = \mathcal{Z}^i(G,A)/\mathcal{B}^i(G,A)$.
Every morphism of G-modules $f : A \to B$ induces a natural morphism of groups $f^i : \mathcal{C}^i(G,A) \to \mathcal{C}^i(G,B)$ that is compatible with the differentials, i.e. $d^i_B f^i = f^{i+1} d^i_A$. Then the natural map $f^\bullet_i : H^i(G,A) \to H^i(G,B)$ induced by f^i is well-defined because of this compatibility. The following result is crucial for our purposes

Proposition 4.9. *The short exact sequence*

$$0 \longrightarrow A \xrightarrow{f} B \xrightarrow{g} C \longrightarrow 0$$

induces the long exact sequence[46]

$$0 \to H^0(G,A) \xrightarrow{f^\bullet_0} H^0(G,B) \xrightarrow{g^\bullet_0} H^0(G,C) \xrightarrow{\delta_0} H^1(G,A) \xrightarrow{f^\bullet_1} H^1(G,B) \xrightarrow{g^\bullet_1} H^1(G,C) \xrightarrow{\delta_1} \cdots$$

[44] Here by definition $\mathcal{B}^0(G,A) = 0$.
[45] Note that $H^i = 0$ only when the converse is true.
[46] The *connecting* morphisms δ_i are given by the Snake lemma and are *natural*.

4.5.1 The Selmer and Tate-Shafarevich groups

Galois cohomology is the study of group cohomology when G is a Galois group, which may be infinite. Because of this we need to modify our definition of cohomology; the groups $C^i(G, A)$ are not convenient for our purposes when G is infinite. This issue can be solved by introducing a topology. Here we endow G with the Krull topology and G becomes a profinite group, and endow A with the discrete topology. Recall that a profinite group is the inverse limit of an inverse system of finite groups,

$$G = \varprojlim_{i \in I} G_i$$

We will not treat in detail this theory, but in the case of Galois groups the G_i ranges over the finite Galois extensions of K. Then the action map $G \times A \to A$ can be shown to be continuous, and if we take $C^i_{\text{cts}}(G, A) = \{f : G^i \to A \text{ continuous}\}$ instead of $C^i(G, A)$ all the standard results remain to be true, in particular proposition 4.9.
Now fix an elliptic curve E over K and a prime number p, and consider the following exact sequence[47]

$$0 \longrightarrow E_p \xrightarrow{\iota} E \xrightarrow{p} E \longrightarrow 0 \tag{4.2}$$

We set $H^i(K, A) = H^i(\mathrm{Gal}(\overline{K}/K), A)$. By proposition 4.9 one has

$$0 \to E(K)_p \xrightarrow{\iota_0} E(K) \xrightarrow{p_0} E(K) \xrightarrow{\delta_0} H^1(K, E_p) \xrightarrow{\iota_1} H^1(K, E) \xrightarrow{p_1} H^1(K, E) \xrightarrow{\delta_1} \cdots$$

because in general $H^0(G, A) \simeq A^G$ is the group of G-invariants of A. By exactness we obtain

$$0 \to \operatorname{coker} p_0^\bullet = E(K)/pE(K) \xrightarrow{\overline{\delta}_0} H^1(K, E_p) \xrightarrow{\iota_1} \ker p_1^\bullet = H^1(K, E)_p \to 0$$

For any prime v of K, the restriction map embeds the Galois group $\mathrm{Gal}(\overline{K}_v/K_v) \hookrightarrow \mathrm{Gal}(\overline{K}/K)$ into the absolute Galois group[48], so that the induced maps in the cohomology make the diagram commute

$$\begin{array}{ccccccccc}
0 & \longrightarrow & E(K)/pE(K) & \xrightarrow{\overline{\delta}_0} & H^1(K, E_p) & \xrightarrow{\iota_1} & H^1(K, E)_p & \longrightarrow & 0 \\
 & & \downarrow & & \downarrow{\scriptstyle \beta} & \searrow^{\alpha=\rho\circ\iota_1} & \downarrow{\scriptstyle \rho} & & \\
0 & \longrightarrow & \prod_v E(K_v)/pE(K_v) & \longrightarrow & \prod_v H^1(K_v, E_p) & \longrightarrow & \prod_v H^1(K_v, E)_p & \longrightarrow & 0
\end{array}$$

Then the *p-Selmer group* is $\mathrm{Sel}_p(E/K) = \ker \alpha$, and the *Tate-Shafarevich group* $Ш(E/K)$ is the kernel of the product of the restriction maps $\hat{\rho}$[49], i.e.

$$\hat{\rho} = \prod_v \Big(\mathrm{Res} : H^1(K, E) \to H^1(K_v, E) \Big)$$

That is, the kernel of ρ is the p-torsion of $Ш(E/K)$. The following expression for $\overline{\delta}_0$ can be obtained by diagram chasing

[47] Here p is the multiplication-by-p map, $E_p = \ker p$ is the *p-torsion* of E and ι the natural inclusion.
[48] Here K_v is the completion of K at v.
[49] $Ш(E/\mathbb{Q})$ is hard to compute in general, but it is conjectured to be finite.

$$\begin{aligned}\overline{\delta}_0 : \quad E(K)/pE(K) &\longrightarrow H^1(K, E_p)\\ P &\longmapsto \overline{\delta}_0(P) : \mathrm{Gal}(\overline{K}/K) \longrightarrow E_p\\ &\qquad \sigma \longmapsto \sigma(\tfrac{1}{p}P) - \tfrac{1}{p}P\end{aligned}$$

This is shown by recalling how the connecting morphism of the Snake lemma was constructed. Given an exact sequence

$$0 \longrightarrow A \xrightarrow{f} B \xrightarrow{g} C \longrightarrow 0$$

let $\mathcal{Q}^k(G, M) = \mathcal{C}^k(G, A)/\mathcal{B}^k(G, A)$ and consider the following diagram of exact rows and columns

$$\begin{array}{ccccccccc}
 & & 0 & & 0 & & 0 & & \\
 & & \downarrow & & \downarrow & & \downarrow & & \\
 & & H^0(G,A) & \xrightarrow{\overline{f_0}} & H^0(G,B) & \xrightarrow{\overline{g_0}} & H^0(G,C) & & \\
 & & \downarrow{\scriptstyle \iota_A} & & \downarrow{\scriptstyle \iota_B} & & \downarrow{\scriptstyle \iota_C} & & \\
 & & \mathcal{Q}^0(G,A) & \xrightarrow{\overline{f_0}} & \mathcal{Q}^0(G,B) & \xrightarrow{\overline{g_0}} & \mathcal{Q}^0(G,C) & \longrightarrow & 0 \\
 & & \downarrow{\scriptstyle \overline{d_A^0}} & & \downarrow{\scriptstyle \overline{d_B^0}} & & \downarrow{\scriptstyle \overline{d_C^0}} & & \\
0 & \longrightarrow & \mathcal{Z}^1(G,A) & \xrightarrow{\overline{f_1}} & \mathcal{Z}^1(G,B) & \xrightarrow{\overline{g_1}} & \mathcal{Z}^1(G,C) & & \\
 & & \downarrow{\scriptstyle \pi_A} & & \downarrow{\scriptstyle \pi_B} & & \downarrow{\scriptstyle \pi_C} & & \\
 & & H^1(G,A) & \xrightarrow{\overline{f_1}} & H^1(G,B) & \xrightarrow{\overline{g_1}} & H^1(G,C) & & \\
 & & \downarrow & & \downarrow & & \downarrow & & \\
 & & 0 & & 0 & & 0 & &
\end{array}$$

By the commutativity of the diagram and the exactness of the second and third rows, for any $c \in H^0(G, C)$ there is $b \in \mathcal{Q}^0(G, B)$ with $\overline{g_0}b = \iota_C c$ and there is $a \in \mathcal{Z}^1(G, A)$ with $\overline{f_1}a = \overline{d_B^0}b$. We[50] define $\delta_0(c) = \pi_A(a)$. In our case (eq 4.2) this is to say that $p \cdot b = c$ and $a = \sigma b - b$ by the definition of d_B^0.
Moreover, the Selmer and Tate-Shafarevich groups fit in the exact sequence

$$0 \longrightarrow E(K)/pE(K) \xrightarrow{\overline{\delta}_0} \mathrm{Sel}_p(E/K) \xrightarrow{\iota'} Ш(E/K)_p \longrightarrow 0$$

where $\iota' = \iota_1|_{\mathrm{Sel}_p(E/K)}$.

4.6 Kolyvagin's theorem

Kolyvagin's theorem relates the Tate-Shafarevich group with the group $E(K)$

Theorem 4.10 (Kolyvagin)**.** *Let $E/\mathbb{Q}$ be an elliptic curve. Let K be a quadratic imaginary field and let $y_{f,K}$ denote the Heegner point of conductor $f = 1$ constructed in section 4.4. If $y_{1,K}$ has infinite order in $E(K)$ then $E(K)$ has rank 1 and $Ш(E/K)$ is finite.*

[50] Proving the Snake lemma is showing that this is a well-defined morphism.

To relate Kolyvagin's result with the BSD conjecture one makes use of the following theorem, proven by Gross and Zagier showed in 1986:

Theorem 4.11 (Gross and Zagier)**.** *Let E and K be as above. Then $y_{1,K}$ has infinite order*[51] *if and only if the analytic rank of E/K is* 1*.*

In other words, under the current hypotheses on E and K the analytic and algebraic rank over K agree.
The proof of theorem 4.10 is quite involved. There is weakening proven by Gross to highlight the main ideas:

Theorem 4.12. *Let p be an odd prime such that $\mathbb{Q}(E_p)/\mathbb{Q}$ has Galois group $\mathrm{GL}_2(\mathbb{F}_p)$ and p does not divide*[52] *$y_{1,K}$ in $E(K)/E(K)_{tors}$. Then $E(K)$ has rank* 1 *and the p-torsion of $Ш(E/K)$ is trivial.*

This is weaker since $Ш(E/K)$ still could have nontrivial infinite subgroups.
From now on E will be an elliptic curve over $\mathbb{Q}$ of *conductor* N. Here the conductor is as in section 4.2.1, a number divisible precisely by those rational primes p for which E has *bad reduction*, i.e. the reduction modulo p is singular, although the precise definition of the conductor is not essential in what follows.

4.6.1 Galois action on torsion points

If $E/\mathbb{Q}$ is an elliptic curve, the Galois group $G_{\mathbb{Q}} = \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ acts naturally on $E(\overline{\mathbb{Q}})$ coordinatewise and this action is compatible with the group law[53] of E, i.e. $\sigma(P+Q) = \sigma(P) + \sigma(Q)$ for any $\sigma \in G_{\mathbb{Q}}$. In particular, a point is in the p-torsion $P \in E_p$ if and only if so does $\sigma(P)$. Here $E_p = E(\overline{\mathbb{Q}})_p$ and the inclusion $E(\overline{\mathbb{Q}})_p \subset E(\mathbb{C})$ and theorem 2.5 imply[54] that $|E_p| = p^2$, because there are precisely p^2 p-torsion points in a lattice $\Lambda \subset \mathbb{C}$. Let $\mathbb{Q}(E_p)$ be the field obtained by adjoining the x and y coordinates of points of E_p and let $G(p) = \mathrm{Gal}(\mathbb{Q}(E_p)/\mathbb{Q})$. Note that the action $G(p)$ on E_p is faithful. Then E_p has a natural structure of $\mathbb{F}_p$-vector space and so there exists a representation $\rho_p : G(p) \hookrightarrow \mathrm{GL}_2(\mathbb{F}_p)$, not surjective in general. However, if E has no CM then by a corollary of *Serre's Uniformity theorem* ρ surjects, i.e. $G(p) \simeq \mathrm{GL}_2(\mathbb{F}_p)$. In other words, we claim

Theorem 4.13. *The extension $\mathbb{Q}(E_p)/\mathbb{Q}$ has Galois group $\mathrm{GL}_2(\mathbb{F}_p)$ for p sufficiently large.*

One may take p sufficiently large so that this is the case. Let $L = K(E_p)$ be the field obtained by adjoining the x and y coordinates of points in E_p to K. Note that the extension L/K is Galois. The ramification of L/K is known,

Proposition 4.14. *The extension L/K is unramified outside the primes that divide pN.*

[51] Here $f = 1$, so $\mathcal{O} = \mathcal{O}_K$ is maximal.

[52] That is, there exists no point $Q \in E(K)$ with $p \cdot Q - y_{1,K} \in E(K)_{\mathrm{tors}}$; since $E(K)/E(K)_{\mathrm{tors}}$ is a finitely generated free $\mathbb{Z}$-module one can rewrite the equation $pQ = y_{1,K} = (c_1, \cdots, c_r)$ in terms of matrices and take $p \nmid c_i$ for all i with $c_i \neq 0$.

[53] Because it can be expressed in terms of $\mathbb{Q}$-rational functions.

[54] The same is true in nonzero characteristic for all but finitely many primes p.

Proof. Let S be the set of primes of K not dividing pN. Take $\mathfrak{p} \in S$ and $\mathfrak{P}$ a prime of L dividing $\mathfrak{p}$. Then it is clear that E has good reduction over $\mathcal{O}_{K_\mathfrak{p}}$. The fundamental identity for number fields of section 1.1 can be applied also to local fields[55], and it reads[56]

$$ef = [L_\mathfrak{P} : K_\mathfrak{p}]$$

Therefore L/K is unramified for $\mathfrak{p} \in S$ if so does $L_\mathfrak{P}/K_\mathfrak{p}$ for $\mathfrak{P}|\mathfrak{p}$ and $\mathfrak{p} \in S$ or equivalently, if $[L_\mathfrak{P} : K_\mathfrak{p}] = f = [\mathbb{F}_\mathfrak{P} : \mathbb{F}_\mathfrak{p}]$. Now the reduction map $E_p \hookrightarrow \tilde{E}(\mathbb{F}_\mathfrak{P})$ injects, and the inertia group

$$I_\mathfrak{P}(L_\mathfrak{P}/K_\mathfrak{p}) = \{\sigma \in \mathrm{Gal}(L_\mathfrak{P}/K_\mathfrak{p}) : \sigma x \equiv x \bmod \hat{\mathfrak{P}} \text{ for all } x \in L_\mathfrak{P}\}$$

is trivial since it fixes pointwise $L_\mathfrak{P}$ because both $\mathbb{F}_\mathfrak{P} \simeq L_\mathfrak{P}/\hat{\mathfrak{P}}$ and $\tilde{E}(\mathbb{F}_\mathfrak{P})$ are fixed. Thus $e = |I(L_\mathfrak{P}/K_\mathfrak{p})| = 1$ as wanted. □

In other words, a rational prime $\ell \neq p$ is unramified in L precisely when it is unramified in K. Let us recall the current hypotheses - and impose some more - on the objects:

1. $E/\mathbb{Q}$ has no CM and has conductor N and every prime $q|N$ splits completely in $\mathcal{O}_K$ (*Heegner hypothesis*).
2. The conductor f of $\mathcal{O}$ is squarefree and the rational prime ℓ considered throughout this section divides f and is coprime with NDp so ℓ is either inert or split in K.
3. The prime p we fixed to consider the short exact sequence in equation 4.2 is sufficiently large so that ρ_p surjects.
4. p does not divide $y_{1,K}$ (equation 4.1) in $E(K)/E(K)_{\mathrm{tors}}$.

Additionally, we impose that the primes ℓ dividing f are *Kolyvagin primes with respect to* (E, K, p) i.e.

5. p divides the coefficient a_ℓ defined in section 4.2.2
6. p^2 divides $\ell + 1$
7. ℓ is inert[57] in K

4.6.2 Constructing cohomology classes in $H^1(K, E_p)$

It can be shown that there are infinitely many Kolyvagin primes. These assumptions on ℓ are used to show the following

Theorem 4.15. *The Heegner points form an Euler system, i.e.*

$$Tr_\ell y_f = \sum_{\sigma \in \mathrm{Gal}(K_\ell/K_1)} \sigma(y_f) = a_\ell y_g$$

[55] Recall that $L_\mathfrak{P}$ and $K_\mathfrak{p}$ are the completions of L and K by the places $|\cdot|_\mathfrak{P}, |\cdot|_\mathfrak{p}$.

[56] Note that in local fields $g = 1$, because there is only one prime ideal.

[57] Note that by Chebotarev density theorem applied to quadratic extensions there is a positive density of inert primes, thus infinitely many.

Each prime factor λ_f of ℓ in K_f divides an unique prime λ_g of K_g and

$$y_f \equiv \sigma y_g \ mod\ \lambda_f$$

for all $\sigma \in (\lambda_g, K_f/K)$. Here $(\lambda_g, K_f/K)$ is as in (1.6) and a_ℓ are the coefficients of the L-function of E.

Proof. We show the first property. Let $G = \mathrm{Gal}(K_f/K_1)$ and $G' = \mathrm{Gal}(K_f/K)$. Recall that for a prime ℓ with $\ell g = f$ the subgroup $G_\ell = \mathrm{Gal}(K_\ell/K_1)$ is cyclic of order $\ell+1$ (see section 2.8.2).
Since our curve E is defined over $\mathbb{Q}$ there exists a newform $f \in \mathcal{S}_2(N)$ with $L_{E/\mathbb{Q}}(s) = L(f,s)$ by theorem 4.8. It can be shown that the trace map

$$\begin{aligned} \mathrm{Tr}_\ell : \quad \mathrm{Jac}(X_0(N))(K_f) &\longrightarrow \mathrm{Jac}(X_0(N))(K_g) \\ x &\longmapsto \textstyle\sum_{\sigma\in G_\ell} \sigma x \end{aligned}$$

verifies $\mathrm{Tr}_\ell(x_f) = T_\ell(x_g)$ where T_ℓ is the Hecke operator on the Jacobian $\mathrm{Jac}(X_0(N))$. Then by theorem 4.7 the Hecke operators act as multiplication by the coefficients a_ℓ of the newform f, the same of the L-function of E

$$\mathrm{Tr}_\ell(y_f) = \mathrm{Tr}_\ell\varphi(x_f) = \varphi(\mathrm{Tr}_\ell(x_f)) = \varphi(T_\ell(x_g)) = a_\ell y_g$$

□

The first property is used below to construct Kolyvagin cohomology classes, that we discuss now. The second property is used to show the local triviality of these classes. This is discussed in section 4.6.3.
Now we define some elements in the group ring $\mathbb{Z}[G]$.
The *augmentation ideal* of the group ring $\mathrm{Z}[G_\ell]$ is the kernel of the *augmentation map*

$$\begin{aligned} \epsilon : \mathbb{Z}[G_\ell] &\longmapsto \mathbb{Z} \\ \textstyle\sum_{\sigma\in G_\ell} n_\sigma\sigma &\longmapsto \textstyle\sum_\sigma n_\sigma \end{aligned}$$

Note that the kernel $\ker\epsilon$ is the $\mathbb{Z}$-submodule generated by elements of the form $\sigma - \sigma'$ for $\sigma, \sigma' \in G_\ell$. Now fix a generator $\sigma_\ell \in G_\ell$ and write

$$\sigma - \sigma' = \sigma_\ell^t - \sigma_\ell^s = \sum_{i=0}^{t-s-1} \sigma_\ell^{t-i} - \sigma_\ell^{t-i-1} = (\sigma_\ell - \mathrm{id})\sigma''$$

for any σ, σ' and some σ''. It follows that $\ker\epsilon = \mathbb{Z}[G_\ell]\cdot(\sigma_\ell - \mathrm{id})$ is principal. Let $\mathrm{Tr}_\ell = \sum_{\sigma\in G_\ell}\sigma \in \mathbb{Z}[G_\ell]$ and let S_ℓ be a solution of the following equation in $\mathbb{Z}[G_\ell]$

$$(\sigma_\ell - \mathrm{id})\cdot S_\ell = (\ell+1)\mathrm{id} - \mathrm{Tr}_\ell \tag{4.3}$$

We must show that at least some S_ℓ exists.

Lemma 4.16 (Kolyvagin)**.** *There exist solutions to equation 4.3.*

Proof. Write $\tau_j = \mathrm{id} + \sigma_\ell + \sigma_\ell^2 + \cdots + \sigma_\ell^{j-1}$ and note that $(\sigma_\ell - \mathrm{id}) \cdot \tau_j = \sigma_\ell^j - \mathrm{id}$. It follows that

$$(\sigma_\ell - \mathrm{id}) \sum_{j=1}^{\ell+1} \tau_j = \sum_{j=1}^{\ell+1} \sigma_\ell^j - (\ell+1) = \mathrm{Tr}_\ell - (\ell+1)\mathrm{id}$$

thus $S_\ell = -\sum_{j=1}^{\ell+1} \tau_j$ is a solution of equation 4.3. □

Note that the class of S_ℓ in $\mathbb{Z}[G_\ell]/\mathbb{Z} \cdot \mathrm{Tr}_\ell$ is well-defined since $(\sigma_\ell - \mathrm{id}) \cdot \mathrm{Tr}_\ell = 0$. Let

$$S_f = \prod_{\ell | f} S_\ell$$

Note that S_f is well-defined since the S_ℓ commute, because the S_ℓ are linear combinations of powers of the same element σ_ℓ. Recall the construction of y_f in section 4.
We claim that the class of the point $S_f y_f \in E(K_f)$ in $E(K_f)/pE(K_f)$ is invariant by G. Since $G \simeq \prod_{\ell|f} G_\ell$ by section 2.8.2 it suffices to show G_ℓ invariance for $\ell \mid f$ i.e. that $S_f y_f \in \ker(\sigma_l - \mathrm{id})$. Since p^2 divides $\ell + 1$ by assumption 6 and equation 4.3 one has

$$(\sigma_\ell - \mathrm{id}) S_f y_f = (\sigma_\ell - \mathrm{id}) S_\ell S_g y_f = \Big((\ell+1)\,\mathrm{id} - \mathrm{Tr}_\ell \Big) S_g y_f \equiv (0 - \mathrm{Tr}_\ell) S_g y_f \mod pE(K_f)$$

But $\mathrm{Tr}_\ell S_g = S_g \mathrm{Tr}_\ell$ so $(\sigma_\ell - \mathrm{id}) S_f y_f \equiv -S_g \mathrm{Tr}_\ell y_f \mod pE(K_f)$. Since p divides a_ℓ by assumption 5 and theorem 4.15 holds we have $\mathrm{Tr}_\ell y_f = a_\ell y_g \in pE(K_g) \subset pE(K_f)$ as wanted. Thus it follows

Proposition 4.17 (Kolyvagin)**.** *The class $[P_f]$ of the point*

$$P_f = \sum_{\sigma \in G'/G} \sigma S_f y_f$$

in $E(K_f)/pE(K_f)$ is invariant under natural the action of G'. Note that the class depends on the choice of the generators σ_ℓ.

Now, the restriction morphism $\mathrm{Res} : H^1(K, E_p) \to H^1(K_f, E_p)$ maps an element $f : \mathrm{Gal}(\overline{K}/K) \to E_p$ to its restriction to $\mathrm{Gal}(\overline{K_f}/K_f) \subset \mathrm{Gal}(\overline{K}/K)$. It can be shown that under the current hypotheses on p (assumptions 3 and 4) and theorem 4.13, the p-torsion $E_p(K_f)$ over K_f is trivial. Then if $\sigma \in G' = \mathrm{Gal}(K_f/K)$ then $\sigma(\mathrm{Res}(f)) = \mathrm{Res}\, f$ since the image of $\mathrm{Res}(f)$ is contained in E_p and the natural action of G' on $E_p(K_f)$ is trivial. Thus the map $\mathrm{Res} : H^1(K, E_p) \to H^1(K_f, E_p)^{G'}$ is well-defined. In fact, Res is an isomorphism

$$H^1(K, E_p) \simeq H^1(K_f, E_p)^{G'}$$

The cohomology classes are built through this isomorphism - together with proposition 4.17 - and it can be proven using the *inflation-restriction sequence*, a general result on group cohomology that we take for granted:

Theorem 4.18. *If G acts on an abelian group A and $N \subset G$ is a normal subgroup then*

$$0 \longrightarrow H^1(G/N, A^N) \xrightarrow{Inf} H^1(G, A) \xrightarrow{Res} H^1(N, A)^{G/N} \longrightarrow H^2(G/N, A^N) \longrightarrow H^2(G, A)$$

is an exact sequence.

The action of G/N on A^N is natural $gN \cdot a := ga$ and well-defined[58]. In our case, $G = \mathrm{Gal}(\overline{K}/K)$ and $N = \mathrm{Gal}(\overline{K_f}/K_f) = \mathrm{Gal}(\overline{K}/K_f)$ satisfy $G/N \simeq G' = \mathrm{Gal}(K_f/K)$ since N is normal by Galois theory, and $A = E_p$. Thus both $H^1(G/N, A)$ and $H^2(G/N, A^N)$ are trivial since $E_p(K_f)$ is trivial.
Consider the diagram from section 4.5.1 for K_f

$$0 \longrightarrow E(K_f)/pE(K_f) \longrightarrow H^1(K_f, E_p) \longrightarrow H^1(K_f, E)_p \longrightarrow 0$$

We can apply the $G' = \mathrm{Gal}(K_f/K)$-invariants functor - which is left-exact - to obtain

$$0 \longrightarrow (E(K_f)/pE(K_f))^{G'} \xrightarrow{\delta_f} (H^1(K_f, E_p))^{G'} \xrightarrow{\iota'} (H^1(K_f, E)_p)^{G'}$$

On the other hand, theorem 4.18 for $G = \mathrm{Gal}(\overline{K}/K), N = \mathrm{Gal}(\overline{K}/K_f)$ and $A = E$ implies[59]

$$0 \longrightarrow H^1(K_f/K, E(K_f)) \longrightarrow H^1(K, E) \longrightarrow H^1(K_f, E)^{G'}$$

Taking torsion is left-exact so we obtain

$$0 \longrightarrow H^1(K_f/K, E(K_f))_p \longrightarrow H^1(K, E)_p \longrightarrow (H^1(K_f, E)^{G'})_p = (H^1(K_f, E)_p)^{G'}$$

Putting everything together gives the following commuting diagram

$$\begin{array}{ccccccccc}
 & & & & & & 0 & & \\
 & & & & & & \downarrow & & \\
 & & & & & & H^1(K_f/K, E(K_f))_p & & \\
 & & & & & & \downarrow{\scriptstyle \mathrm{Inf}} & & \\
0 & \longrightarrow & E(K)/pE(K) & \xrightarrow{\overline{\delta}_0} & H^1(K, E_p) & \xrightarrow{\iota_1} & H^1(K, E)_p & \longrightarrow & 0 \\
 & & \downarrow & & \downarrow{\scriptstyle \mathrm{Res}}\simeq & & \downarrow{\scriptstyle \mathrm{Res}} & & \\
0 & \longrightarrow & (E(K_f)/pE(K_f))^{G'} & \xrightarrow{\delta_f} & (H^1(K_f, E_p))^{G'} & \xrightarrow{\iota'} & (H^1(K_f, E)_p)^{G'} & &
\end{array}$$

Therefore, we can let $c(f) \in H^1(K, E_p)$ be the unique[60] class with

$$\mathrm{Res}\ c(f) = \delta_f[P_f]$$

[58] It is well-defined since $gN = g'N$ implies $g' = gn$ for some $n \in N$ so $g'a = gna = ga$ as wanted.
[59] Here $H^1(K_f/K, E) = H^1(G', E)$.
[60] It is unique because Res is an isomorphism.

where $[P_f]$ is the class of proposition 4.17. Clearly, $P_f \in pE(K_f)$ i.e. $[P_f] = 0$ if and only if $c(f) = 0$. Let $d(n) = \iota_1(c(f))$. And Res $d(n) = 0$ since

$$\text{Res } d(n) = \text{Res } \iota_1 c(f) = \iota' \text{Res } c(f) = \iota' \delta_f [P_f] = 0$$

i.e. by the commutativity of the diagram and exactness of the bottom row. Thus, by the exactness of the rightmost column there exists an unique[61] $\tilde{d}(f) \in H_1(K_f/K, E(K_f))_p$ with

$$\text{Inf } \tilde{d}(f) = d(f)$$

Again, $\tilde{d}(f) = 0$ if and only if $d(f) = 0$ by the injectivity of Inf . But $d(f) = 0 = \iota_1 c(f)$ if and only if $c(f) \in \ker \iota_1 = \text{im } \overline{\delta}_0$ i.e. $P_f \in pE(K_f) + E(K)$.

4.6.3 Idea of the rest of the proof

Recall from section 4.5.1 the following exact sequence

$$0 \longrightarrow E(K)/pE(K) \xrightarrow{\overline{\delta}_0} \text{Sel}_p(E/K) \xrightarrow{\iota'} Ш(E/K)_p \longrightarrow 0$$

and take the following for granted:

Theorem 4.19. *If p satisfies assumptions 3 and 4 then $\text{Sel}_p(E/K)$ is cyclic and generated by $\overline{\delta}_0(y_{1,K})$.*

Assuming theorem 4.19, the proof of 4.12 is as follows:

Proof. E has no p-torsion by assumption 4 the rank $\text{rank}(E/K)$ of E/K verifies

$$\text{rank}(E/K) = \dim_{\mathbb{F}_p} \Big(E(K)/pE(K) \Big)$$

Since $\overline{\delta}_0$ injects and $y_{1,K}$ is not zero in $E(K)/pE(K)$ we must have $\text{rank}(E/K) \neq 0$ and $\text{rank}(E, K) \leq 1$, thus $\text{rank}(E/K) = 1$. Therefore $\overline{\delta}_0$ is an isomorphism and $Ш(E/K)_p$ is trivial by the exact sequence above. □

To prove theorem 4.19 one first shows that the classes $c(f)$ are in the p-Selmer group $\text{Sel}_p(E/K)$ i.e. $c(f) \in \ker \alpha$ where α is as in section 4.5.1 i.e.

$$\begin{array}{ccccccccc}
0 & \longrightarrow & E(K)/pE(K) & \xrightarrow{\overline{\delta}_0} & H^1(K, E_p) & \xrightarrow{\iota_1} & H^1(K, E)_p & \longrightarrow & 0 \\
 & & \downarrow & & \downarrow{\scriptstyle \beta} & \searrow^{\alpha=\rho\circ\iota_1} & \downarrow{\scriptstyle \rho} & & \\
0 & \longrightarrow & \prod_v E(K_v)/pE(K_v) & \longrightarrow & \prod_v H^1(K_v, E_p) & \longrightarrow & \prod_v H^1(K_v, E)_p & \longrightarrow & 0
\end{array}$$

Since β injects, $c(f) \in \text{Sel}_p(E/K)$ if and only if $\beta c(f) = 0$ i.e. $c(f)_\nu$ is trivial at every place ν of K. To show this, the second property of theorem 4.15 and the theory of *Néron models and Tate local duality* is used, but this is out of our scope.

[61] Inf is injective.

Alphabetical index

References

[1] Cox, D. *Primes of the form $x^2 + ny^2$*. Wiley, 1997.

[2] Darmon, H. *Rational points on modular elliptic curves*. 2004.

[3] Diamond, F., and Shurman, J. *A First Course in Modular Forms*. Springer (Graduate Texts in Mathematics), 2005.

[4] Gross, B. *Kolyvagin work on modular elliptic curves. L-functions and Arithmetic*. 1991.

[5] Marcus, D. *Number fields*. Springer, 1977.

[6] Neukirch, J. *Algebraic Number Theory*. Springer, 1999.

[7] Osserman, B., Kedlaya, K., et al. *Kolyvagin's Application of Euler Systems to Elliptic Curves, graduate number theory seminar notes*. 2000.

[8] Silverman, J. *The arithmetic of elliptic curves*. Springer, 1986.

[9] Silverman, J. *Advanced topics in the arithmetic of elliptic curves*. Springer, 1994.

[10] Sutherland, A. *Elliptic curves, course notes*. 2017.