BACHELOR'S DEGREE THESIS

Degree in Mathematics

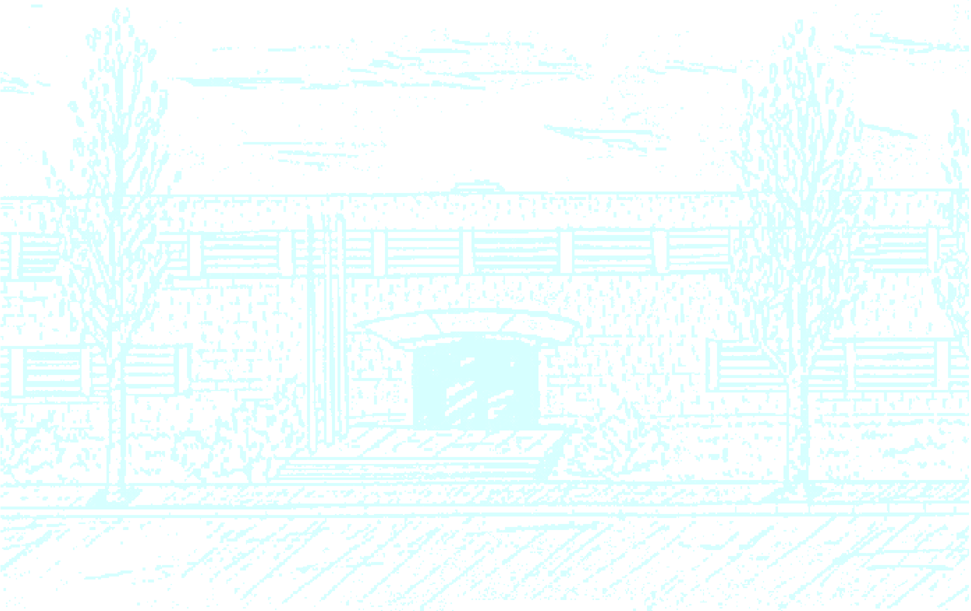
Title: Birch and Swinertonn-Dyer conjecture

Author: Óscar Rivero Salgado

Advisor: Víctor Rotger Cerdà

Department: Departament de Matemàtica Aplicada

Academic year: 2015-2016



UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH
Facultat de Matemàtiques i Estadística

Para Marta, por estar siempre ahí en los malos momentos y recordarme que a veces lo mejor es soñar despierto para vencer el cansancio.

I could just remember how my father used to say that the reason for living was to get ready to stay dead for a long time.
William Faulkner

porque la soledad le había seleccionado los recuerdos, y había incinerado los entorpecedores montones de basura nostálgica que la vida había acumulado en su corazón, y había purificado, magnificado y eternizado los otros, los más amargos
Gabriel García Márquez

Contents

Introduction

The objective of this thesis is to introduce the topics involved in the formulation of the conjecture of Birch and Swinertonn-Dyer, one of the six unsolved problems of the Millenium Problem list of the Clay Institute of Mathematics, together with the Riemann hypothesis, the existence of solutions for the Navier-Stokes equation, the P vs NP problem, the Hodge conjecture and the Yang-Mills equation (the existence of a mass gap in the solution to the quantum version of this equation). Roughly speaking, this conjecture relates the number of points of an elliptic curve over finite fields with the rank of the elliptic curve over the rationals. This brief statement needs some previous considerations: the first one is that for speaking of rank we need a group structure, but we have it in any elliptic curve and explained in geometric terms in any undergraduate course of algebraic geometry. The second fact is that the rank of this group is finite (in fact, the rank of an elliptic curve over a number field is finite). This is a non-trivial theorem proved by Mordell and generalized by Weil around 1930. The relation with the number of points over finite fields comes through the L-function associated to the curve, in a natural generalization of the classical zeta functions for number fields, or even in an easier way, the one that motivates the Riemann hypothesis. The first remarkable results around Birch and Swinertonn-Dyer conjecture (BSD) were done in 1977 by Coates and Wiles and in the last years there has been a great progress, being an active topic in research in number theory nowadays. Some of the ideas around the proof of Wiles and Taylor of the Last Fermat's Theorem (and then the Modularity theorem of Breuil, Conrad, Diamond and Taylor) will also play a crucial role.
However, for a deep understanding of all these concepts it is necessary a wide background, and this is what we try to do in this bachelor thesis, organized in twelve chapters that try to explain the several topics we will need to understand the statement and the most elementary ideas in the proofs of the main results that have been made:

- The first chapter begins by stating the Hasse-Minkowski theorem for quadratic forms. We have to bear in mind that BSD is an example of local-global principle (from the behavior in local fields we infer the behavior in a global field), so we begin by reviewing the most classical and simplest example of this. The last section of this chapter explains why Hasse's principle fails for cubic polynomials, by giving an explicit example of an equation that has solution both in the real field and in the p-adic fields but not over the rationals. For a proper explanation of the proof, we need to introduce the main results in algebraic number theory, that are also needed in several parts of the work. Of special importance are those concerning finiteness (of the class number and Dirichlet's unit theorem) and those about the behavior of primes in extensions of $\mathbb{Q}$ (number fields).

- The second chapter is a quick review of some fundamental results in algebraic geometry and Riemann surfaces needed along the thesis. We do not give proofs of most of them, since in several cases are long and technical. At the end of the chapter, there should be a brief introduction to the theory of schemes, an angular stone in further developments of algebraic geometry and consequently number theory (it is omitted due to a lack of time). We try to point out the great importance geometry plays in arithmetic. We explain Abel's theorem and the role it plays to study maps between jacobians.

- The third chapter is one of the cores (together with chapter seven) of the thesis: it introduces the concept of elliptic curves, its basic properties, the geometrical aspects that will be around us all over the time and its most remarkable properties in both local and global fields. The last sections are not so introductory and they deal with concepts like the Tate module, the endomorphism ring or the Weil pairing.

- The forth chapter introduces ζ and L-functions of an elliptic curves, starting with a proof of Hasse's theorem that gives a bound for the number of points that an elliptic curve has over a finite field. We explain why this is also called the Riemann hypothesis for elliptic curves, and state the celebrated Weil conjectures, one of the highlights in the mathematics of the twentieth century (proved by Deligne).

- In the five chapter we just give a brief presentation of a topic that has interest by its own, but that we use here in several moments and for different purposes: group cohomology. It appears in the proof of Mordell's theorem, in the definition of homogeneous spaces and it is the natural framework to develop class field theory. We try to give an explicit interpretation of the meaning of H^0 and H^1 and state some basic results in algebraic topology and category theory.

- The sixth chapter contains the proof of one of the most elementary and important results in the theory of elliptic curves: Mordell-Weil theorem about the finiteness of the rank of an elliptic curve over the rationals (and more generally over a number field). It consists basically of two steps, being the first one related with establishing the finiteness of the Selmer group (that will be properly defined arrived that point), and that can be generalized to number field, and then, defining a height over the elliptic curve, use an appropriate descent procedure to conclude. The extension of this part to number fields is complicated and we omit it.

- Together with chapter three, this seventh chapter is one of the most important in the thesis. It develops all the theory of modular forms, beginning by the basic topological and analytic issues, and defining all the key concepts: from the more geometrical ones to those more delicate, like Hecke operators or eigenforms. The first sections are very elementary, and use concepts of complex analysis and general topology, but when we go further we start to see some similitudes with the theory of elliptic curves, about all when we define the L-series attached to a modular form.

- Begin we move to more complicated topics, we include another one that can still be considered introductory: in the eighth chapter we properly define quaternion algebras, that had already appeared in chapter three as one of the three possibilities for the endomorphism ring of an elliptic curve. Furthermore, they will be necessary to generalize several concepts about modular forms and to define Shimura curves, one of the main objects in number theory.

- Chapter nine is about the theory of complex multiplication, that Hilbert considered the most beautiful not only of mathematics, but of all science. However, although the statements are simple, most of the proofs are complicated and they require the introduction of a powerful tool: class field theory. We use a few sections to explain the main results of this field, omitting most of the proofs, and then we move to the main results in complex multiplication. Roughly speaking, it is just the study of elliptic curves (here over the rational) whose endomorphism ring is an order in an imaginary quadratic field.

- Chapter ten is extremely related with chapter nine, and in fact we begin with the proof of a result that had been already stated: that there is a curve defined over the rationals (i.e., a polynomial $F(X,Y)$ with coefficients in $\mathbb{Q}$) birationally equivalent to $X_0(N)$). We continue by stating (and trying to give an idea of the proof) important results like the Eichler-Shimura correspondence, explaining the concept of Heegner points and quoting some recent results of great importance.

- In chapter eleven we finally explain the conjecture of Birch and Swinertonn-Dyer: we give several formulations of it, its generalizations and the consequences it would have. Furthermore, we state the main results in that direction, trying to indicate the most relevant ideas in the proofs. In particular, the results of Gross-Zagier are quite natural after the theory developed in the previous chapter. In the last sections, we state of the last known results about BSD due to Darmon and Rotger, that gives a positive answer beyond the most frequent cases of order of vanishing zero or one.

- The last chapter is devoted to briefly comment some topics that did not fit in the thesis but that are of great importance and have a deep relation with BSD and at this level cannot be stated in a proper way. We introduce

> and give some first results about Galois representations, and state its importance in the proof of Fermat's last theorem. We also give a brief insight of how can we adapt Mordell-Weil for the case of function fields, and try to introduce in an extremely vague way the theory of Neron models, that is useful for proving results like Mazur's theorem about the torsion of an elliptic curve. Finally, we reinterpret some results about modular forms introducing Hilbert modular forms.

There are several books which play a prominent role in the writing of this dissertation. In a first walk through number theory, I highly believe that the books of Milne are a very good choice to understand in a proper way the first non-trivial facts about elliptic curves or modular forms; furthermore, his books in algebraic number theory and algebraic geometry were also of great aid. The chapters devoted to elliptic curves are based (in most of its sections) in the two books of Silverman, whose last chapters were not included here and would be a good start for an improvement version of this thesis. To understand modular forms, I do not conceive a better book than the one from Diamond and Shurman, focused on the most relevant aspects which lead to the proof of the modularity theorem. Last, but not least important, Darmon's notes about rational points on modular elliptic curves were the base for the writing of chapters ten and eleven, maybe the most difficult ones in a first reading.

I apologize beforehand for the mistakes this thesis contain, most of them due to a lack of time for a more exhaustive revision; I also request your benevolence in which concerns some inconsistencies with the notation: depending on the context, I use $\sigma(P)$ or P^σ for the action of the Galois group, or the Frobenius is sometimes referred as ϕ and others as Π. Anyway, all of these notations are highly used and the meaning, from my point of view, is clear at any moment.

I would like to thank the support and guidance of Víctor Rotger along these months, and I hope that this experience were only the beginning of a fruitful cooperation. His preclair vision of mathematics has helped me to learn many new things and to understand all those concepts that seemed dark and unclear in many books. I am also indebted with Jordi Quer, not only for introducing me Víctor but for being an excellent teacher that taught us along the subjects I shared with him to love algebra and number theory. This first approach to mathematics would not have possible without the support and guidance of people from both the CFIS and the School of Mathematics FME. I would like to express my gratitude to all teachers, head-masters and personal stuff that helped me during the last three years and a half. In particular to the chairmen of the institutions, Miguel Ángel Barja and Jaume Franch, for always being ready to share a conversation about any topic. Last, but not least, thanks to the people (both teachers and mates) from the Mathematical Olympiad, which showed me to love mathematics from an early age. And of course, thanks to my family and friends who were always close to me; without them, life would be less interesting.

Chapter 1

Algebraic number theory

The aim of this chapter is to provide the necessary background from basic algebraic number theory to understand the proofs that we will be presenting along the thesis. Most of these results are very classical (from the nineteenth century) but provide a lot of information about number fields and their structure. We begin with a motivation from the theory of classical quadratic forms, and then we move to present all the machinery from algebraic number theory, starting from basic definitions (with some easy propositions to show the type of proofs we usually have) and then passing to state and prove some of the most relevant theorems concerning finiteness and factorization in extension. We finish given an example of how the local-global principle fails for the case of cubics.

1.1 Hasse-Minkowski theorem for quadratic forms

One of the key aspects so as to understand BSD conjecture is the local-global principle, or the possibility of using properties of local fields (locally compact topological fields with respect to a non-discrete topology, here the reals and the p-adics) to say what will happen in global fields (here the rationals, or in general, a number field). As we already pointed out, this is not always possible, but a first naive example is the Hasse-Minkowski theorem for quadratic forms:

Theorem 1.1. *Let f be a quadratic form and let V be the set of places of $\mathbb{Q}$. In order that f represents 0 in $\mathbb{Q}$, it is necessary and sufficient that, for all $v \in V$, the form f_v represents 0.*
Alternatively, a quadratic form has solutions over the rationals if and only if it has solutions both at the real numbers and at the p-adics (for all p).

To develop this theory (quite elementary and that does not require many technical tools), we will need some lemmas that are not difficult to prove and that we will state along the thesis. We will give proofs or not depending on the ideas they provide and on its length. We start with the famous Chevalley's theorem (or Chevalley-Warning).

Theorem 1.2. *Let K be a finite field with q elements and characteristic p. Let $f_\alpha \in K[X_1, \cdots, X_n]$ be such that $\sum_\alpha \deg f_\alpha < n$ and let V the set of their common zeros in K^n. Then, the cardinal of V is a multiple of p.*

Proof. Let $P = \prod_\alpha (1 - f_\alpha^{q-1})$ and let $x \in K^n$. By Fermat's little theorem, P is the characteristic function of V and if we define for a generic polynomial f, $S(f) = \sum_{x \in K^n} f(x)$, we clearly have that the cardinal of V is equal to $S(p)$ modulo p. Recall now the following almost obvious lemma:

Lemma 1.1. *Let u be a non negative integer; then, $S(X^u) = \sum_{x \in K} x^u$ is -1 when $u \geq 1$ and divisible by $q-1$ and is 0 otherwise.*

The proof of the lemma is an exercise where you only have to consider, for the case of not divisible by $q-1$, the fact that K^* is cyclic; multiplying by y such that $y^u \neq 1$, $S(X^u) = y^u S(X^u)$, and we are done.
With the lemma in mind, we have that the degree of P is less than $n(q-1)$, so P is a linear combination of monomials $X^u = X_1^{u_1} \cdots X_n^{u_n}$ where at least one of the exponents must be smaller than $q-1$. □

Another crucial result that will be used several times along this thesis is Hensel's lemma. It is a result that allows us to lift solutions from congruence groups to p-adic solutions (in some way it is similar to Newton's method).

Theorem 1.3. *Let $f \in \mathbb{Z}_p[X_1, \cdots, X_m]$, $x = (x_i) \in (\mathbb{Z}_p)^m$, $n, k \in \mathbb{Z}$ and let j be an integer such that $0 \leq j \leq m$. Suppose that $0 < 2k < n$ and that $f(x) \equiv 0 \mod p^n$ and that $v_p\left(\frac{\partial f}{\partial X_j}(x)\right) = k$. Then, there exists a zero y of f in $(\mathbb{Z}_p)^m$ congruent to x modulo p^{n-k}*

We continue by introducing the Hilbert symbol (a particular case of a more general theory). In this section, k will denote either $\mathbb{R}$ or $\mathbb{Q}_p$. Let $a, b \in k^*$. We define:

$$(a,b) = \begin{cases} 1 & \text{if } z^2 - ax^2 - by^2 = 0 \text{ has a solution in } k^3 \\ -1 & \text{otherwise} \end{cases}$$

This number is called the Hilbert symbol of a and b relative to k and clearly defines a map from $k^*/k^{*^2} \times k^*/k^{*^2}$ to $\{+1, -1\}$. This will appear again when we study quaternion algebras. We mention some properties that almost follow directly from the definition:

a) $(a,b) = (b,a)$ and $(a, c^2) = 1$.

b) $(a, -a) = 1$ and $(a, 1-a) = 1$.

c) $(a,b) = 1 \implies (aa', b) = (a', b)$.

d) $(a,b) = (a, -ab) = (a, (1-a)b)$.

If $k = \mathbb{R}$ the computation of the Hilbert symbol is straightforward: $(a,b) = -1$ if and only if both a and b are strictly greather than 0. We now state one of the main theorems of this section:

Theorem 1.4. *If $k = \mathbb{Q}_p$ and we write $a = p^\alpha u, b = p^\beta v$, where u, v are p-adic units, we have:*

$$(a,b) = \begin{cases} (-1)^{\alpha\beta\epsilon(p)}\left(\frac{u}{p}\right)^\beta\left(\frac{v}{p}\right)^\alpha & \text{if } p \neq 2 \\ (a,b) = (-1)^{\epsilon(u)\epsilon(v)+\alpha\omega(v)+\beta\omega(u)} & \text{if } p = 2 \end{cases}$$

where as usual $\epsilon(p) = \frac{p-1}{2}$.

Proof. Since the Hilbert symbol is defined modulo squares (in k^*/k^{*^2}), to prove the result we only have to consider the different possibilities for α and β modulo 2. We start by the case in which $p \neq 2$:

- $\alpha = 0, \beta = 0$. In this case we only have to prove that $(a,b) = 1$, i.e., that the equation $z^2 - ux^2 - vy^2$ has a nontrivial solution modulo p. By virtue of Chevalley principle, we have a solution modulo p and by Hensel lemma we can lift it to a p-adic solution.

- $\alpha = 1, \beta = 0$; here we have to prove that $(pu, v) = \left(\frac{v}{p}\right)$; since $(u,v) = 1$, using the third of the properties stated below, it will be enough to show $(p, v) = \left(\frac{v}{p}\right)$. We mention now two lemmas that are required to finish:

  Lemma 1.2. *Let v be a p-adic unit. If the equation $z^2 - px^2 - vy^2 = 0$ has a nontrivial solution in $\mathbb{Q}_p$ it has a solution (z,x,y) such that z, y are p-adic units and $x \in \mathbb{Z}_p$.*

  Lemma 1.3. *Let $p \neq 2$ and let $x = p^n u$ an element of $\mathbb{Q}_p^*$ such that $n \in \mathbb{Z}$ and u is a p-adic unit. x is a square if and only if n is even and the image of u in $U/U_1 \cong \mathbb{F}_p^*$ is a square. Here, we use the typical notation $U_n = 1 + p^n\mathbb{Z}_p$.*

  If v is a square, no explanation is required (just put $x = 0$); otherwise, $\left(\frac{v}{p}\right) = -1$ and the equation $z^2 - px^2 - vy^2$ does not have a nontrivial zero (reduce for instance modulo p, and you get $z^2 = vy^2$ where v is not a square so it is not possible).

- $\alpha = 1, \beta = 1$; $(pu, pv) = (pu, -p^2uv) = (pu, -uv)$. We have now a situation where we can use the previous case, i.e., $(pu, pv) = (pu, -uv) = \left(\frac{-uv}{p}\right) = \left(\frac{-1}{p}\right)\left(\frac{u}{p}\right)\left(\frac{v}{p}\right) = (-1)^{\epsilon(p)}\left(\frac{u}{p}\right)\left(\frac{v}{p}\right)$.

□

The proof of the case 2 uses these same ideas and can be read in Serre's book.

Corollary 1.1. *The Hilbert symbol is multiplicative, i.e., $(aa', b) = (a,b)(a',b)$ and $(a, bb') = (a,b)(a,b')$.*

We finish our digression of Hilbert symbol (it will maybe reappear in some place of the thesis) with the following result:

Corollary 1.2. *If $a, b \in \mathbb{Q}^*$, then $(a,b)_v = 1$ for almost all $v \in V$ and*

$$\prod_{v \in V} (a,b)_v = 1$$

We sketch now the way we can prove Hasse-Minkowski theorem, quoting along the explanation several facts about quadratic forms that we do not prove.
The necessity of representing 0 in $\mathbb{R}$ and $\mathbb{Q}_v$ (v a prime) is clear. To see the converse, we write (using the fact that there exists an orthogonal basis)

$$f = a_1 x_1^2 + \cdots a_n x_n^2$$

where the a_i are defined modulo squares and replacing f by $a_1 \cdot f$ we can assume $a_1 = 1$. We have to distinguish now different cases depending on the value of n:

- Case $n = 2$. We consider $x^2 - ay^2$. Seeing the equation in f_∞ we have that $a \geq 0$. Write the equation in the form $x^2 = \prod_{i=1}^r p_i^{v_{p_i}(a)} y^2$. Since this equation has a solution modulo p_i for each i, we conclude that all the $v_{p_i}(a)$ must be even, and therefore the equation has a non trivial solution in $\mathbb{Q}$.

- Case $n = 3$. We have to deal with the equation $x^2 - ay^2 - bz^2$, where we can consider that a and b are defined modulo square and so the p-adic valuations of a and b are 0 or 1. We assume that $|a| \leq |b|$ and use induction in the value of $m = |a| + |b|$. If $m = 2$ we have that in the case $x^2 + y^2 + z^2$ there are no solutions neither in $\mathbb{R}$ (f_∞) neither in $\mathbb{Q}$; for the case $x^2 - y^2 - z^2$ just consider the solution $(1,1,0)$ and for $x^2 + y^2 - z^2$ take $(1,0,1)$.
  Consider now the general case, so $|b| = \pm p_1 \cdots p_r \geq 2$ (all the primes are distinct). Take any of the p_i; then, we have that a is a square modulo p_i. To show that, take a primitive solution of the equation (it is easy to prove that it must exist) and see the equation modulo p. In that case, if a is 0 modulo p we are done, and elsewhere, we have that $x^2 - ay^2 \equiv 0$ and from here, we know, first, that neither x nor y are multiples of p (if one of them is, so is the other, and we look the equation modulo p^2 to conclude that y is also multiple of p and this is a contradiction). From all this, we conclude that a is a square (a is a quadratic residue modulo b if and only if it is a quadratic residue modulo each prime factor of b). So a is a square modulo b and we have integers m, n such that $m^2 = a + bn$ and in particular, since we have two option for m that add up 0 (modulo b) we can take $|m| \leq |b|/2$. We can also write $bn = m^2 - a$, showing that bn is a norm in the quadratic extension $k(\sqrt{a})$ where $k = \mathbb{Q}$ or $k = \mathbb{Q}_v$. From here, $(a, bb') = 1$ (since $(a,b) = 1$ if and only if a belongs to the group of norms of elements of k_b^*, note how lengthened is the shadow of class field theory). We now have that f represents 0 if and only if is represented by $f' = x^2 - ay^2 - b'z^2$ (multiplicativity of the Hilbert symbol) and the result follows from the inductive hypothesis since $|b'| < |b|$.

- Case $n = 4$. Write $f = ax^2 + by^2 - (cz^2 + dt^2)$. We will see that there exists a unity in $\mathbb{Q}_v$ that is represented both by $ax^2 + by^2$ and by $cz^2 + dt^2$, and that the same occurs for $\mathbb{R}$.

  Lemma 1.4. *Let g, h be two non-degenerate forms over k of rank ≥ 1 and let $f = g - h$. Then, the following properties are equivalent:*

  a) f represents 0.

  b) There exists $a \in k^*$ represented both by g and by h.

  c) There exists $a \in k^*$ such that $g - aZ^2$ and $h - aZ^2$ represent 0.

  We call x_v that element represented both by $ax^2 + by^2$ and by $cz^2 + dt^2$. This is the same as saying $(x_v, -ab)_v = (a, b)_v$, $(x_v, -cd)_v = (c, d)_v$ for all v. We need here another result, the approximation theorem:

  Lemma 1.5. *Let S be a finite subset of V. The image of $\mathbb{Q}$ in $\prod_{v \in S} \mathbb{Q}_v$ is dense in this product (when endowed with the product topology).*

  Combining the approximation theorem with the Chinese remainder, we get the following:

  Lemma 1.6. *Let $(a_i)_{i \in I}$ be a finite family of elements in $\mathbb{Q}^*$, and let now $(\epsilon_{i,v})_{i \in I, v \in V}$ be a family of numbers equal to ± 1. In order that there exists $x \in \mathbb{Q}^*$ such that $(a_i, x)_v = \epsilon_{i,v}$ for all $i \in I, v \in V$, it is necessary and sufficient that almost all the $\epsilon_{i,v}$ are 1, that $\prod_{v \in V} \epsilon_{i,v} = 1$ and that for all $v \in V$ there exists $x_v \in \mathbb{Q}_v^*$ such that $(a_i, x_v)_v = \epsilon_{i,v}$ for all $i \in I$.*

  Coming back to our problem and using the product formula, $\prod_{v \in V}(a, b)_v = \prod_{v \in V}(c, d)_v = 1$, we conclude that there exists $\alpha \in \mathbb{Q}$ such that $(\alpha, -ab)_v = (a, b)_v$ and $(\alpha, -cd)_v = (c, d)_v$ for all $v \in V$. So the form $ax^2 + by^2 - \alpha r^2$ represents 0 in each of the $\mathbb{Q}_v$ so in $\mathbb{Q}$; therefore α is represented in $\mathbb{Q}$ both by $ax^2 + by^2$ and by $cz^2 + dt^2$.

- For the case $n \geq 5$ we should use induction, writing $f = h - g$, where $h = ax^2 + by^2$ and $g = -(c_3 z_3^2 + \cdots c_n z_n^2)$. We take the subset of V, S, that contains the place of infinity, 2 and the primes such that $v_p(a_i) \neq 0$ for some $i \geq 3$ (note that is finite). Take $v \in S$, and since f_v represents 0, we have $\alpha_v \in \mathbb{Q}_v^*$ represented both by g and h. What we have said so is that there exist $x_i^v \in \mathbb{Q}_v$ such that $h(x_1^v, x_2^v) = \alpha_v = g(x_3^v, \cdots, x_n^v)$. But we know that the squares of $\mathbb{Q}_v^*$ form an open set, and using the approximation theorem we can affirm the existence of $x_1, x_2 \in \mathbb{Q}$ such that if $\alpha = h(x_1, x_2)$, then $\alpha/\alpha_v \in \mathbb{Q}_v^{*2}$ for all $v \in S$. Here we need a technical result:

  Definition 1.1. *Let $f = a_1 X_1^2 + \cdots + a_n X_n^2$ be a quadratic form in n variables. Then, we define the following two invariants:*

$$\epsilon(f) = \prod_{i<j}(a_i, a_j)$$

$$d(f) = a_1 \cdots a_n \textit{ modulo squares}$$

Proposition 1.1. *In $\mathbb{Q}_v, v \neq 2$, a quadratic form of rank 4 represents 0 if and only either $d \neq 1$ or $d = 1$ and $(-1,-1) = \epsilon$.*

Define $F = \alpha z^2 - g$: if $v \in S$, then g represents α_v in $\mathbb{Q}_v^{*2}$, and it also represents α since $\alpha/\alpha_v \in \mathbb{Q}_v^{*2}$. Hence F represents 0 in $\mathbb{Q}_v$. If $v \notin S$, then $-c_3, \cdots, -c_n$ are v-adic units. The same occurs for $d_v(g)$ and since $v \neq 2$, we have that $\epsilon_v(g) = 1$. In any case F represents 0 in $\mathbb{Q}_v$ so g represents a in $\mathbb{Q}$, and the proof is complete by induction (since F is of rank $n-1$, it represents 0, or alternatively g represents a; therefore h represents a and f represents 0, as wanted).

1.2 Basic definitions in algebraic number theory

Our aim in this section (and the following ones of this chapter) is to introduce some of the most basic facts in algebraic number theory, and some theorems that appeared during the nineteenth century. Our main objectives will be: the theorem of unique factorization in ideals in Dedekind domains, the finiteness of the class number, the unit or Dirichlet theorem and then some results concerning absolute values in $\mathbb{Q}$ and in number fields.
We start with some definitions and easy proposition basically to illustrate the way things work, and as it can be seen most of the ideas are just introductory ring theory. We introduce some names that will be appearing constantly along the thesis:

Definition 1.2. *A ring A is noetherian if it fulfills some of the following three equivalent conditions:*

a) Every ideal in A is finitely generated.

b) Every ascending chain of ideals eventually becomes constant.

c) Every nonempty set S of ideals in A has a maximal element (there exists an ideal in S not properly contained in any other ideal in S)

Definition 1.3. *A ring A is integrally closed if it is its own integral closure in its field of fractions K.*

In particular, a unique factorization domain is integrally closed.

Definition 1.4. *A discrete valuation ring (DVR) is a principal ideal domain A satisfying one of this three equivalent conditions:*

a) A is local and is not a field.

b) A has exactly one nonzero prime ideal.

c) A has exactly one prime element (up to associates)

Proposition 1.2. *An integral domain A is a DVR if and only if A is (at the same time) noetherian, integrally closed and has exactly one nonzero prime ideal.*

Proof. The necessity of the three conditions is clear, so let as assume that all three things hold and let us prove that it is a principal ideal domain. We begin by showing that the nonzero prime ideal is principal (the third condition already implies that A is a local ring).
Choose now an element $c \in A$ such that c is neither 0 nor a unit, and define $M = A/(c)$. For a nonzero element m of M, $\mathrm{Ann}(m) = \{a \in A \mid am = 0\}$ is a proper ideal in A. For the third equivalent condition of being noetherian, choose m such that $\mathrm{Ann}(m)$ is maximal among these ideals, $m = b+(c), p =$Ann$(b+(c))$. Since $c \in p$, we have that p is non empty and that can be characterized by $p = \{a \in A \mid c \text{ divides } ab\}$.
We prove that p is prime: suppose that $x, y \notin p$ such that $xy \in p$. Then $yb + (c)$ is not zero in M, and $\mathrm{Ann}(yb + (c))$ clearly contains p and x, contradicting the maximality of p.
If $b/c \in A$, then $b = (b/c) \cdot c \in (c)$ and $m = b + (c) = 0$ in M (that is not possible for hypothesis). However, $c/b \in A$ and furthermore $p = (c/b)$; to see this, note that $pb \subset (c)$ so $p \cdot (b/c) \subset A$ and it is an ideal in A. If $p \cdot b/c \subset p$, b/c is integral over A since p is finitely generated (this is one of the typical characterizations of integrality), and $b/c \in A$, that is not possible. Consequently, $p \cdot b/c = A$ (where we have used that there is only one prime element) and so $p = (c/b)$.
Let now $\pi = c/b$, so $p = (\pi)$. If a is a proper ideal of A, consider $a \subset a\pi^{-1} \subset a\pi^{-2} \subset \cdots$. If in some moment $a\pi^{-i} = a\pi^{-i-1}$, then $\pi^{-1}(a\pi^{-i}) = a\pi^{-i}$, so π^{-1} is integral over A and so it is in A (and again this was not possible since π is not a unit). Therefore the sequence is strictly increasing, but the ring is noetherian. so it cannot be contained in A, being an integer n such that $a\pi^{-n} \subset A$ but $a\pi^{-n-1}$ is not contained in A. Then, $a\pi^{-m}$ is not in p and $a\pi^{-m} = A$, concluding that $a = (\pi^m)$. □

Definition 1.5. *An integral domain A is a Dedekind domain if these three conditions simultaneously hold:*

a) A is noetherian.

b) A is integrally closed.

c) A nonzero prime ideal is maximal.

A first remarkable fact about these type of rings is the following one:

Proposition 1.3. *A noetherian integral domain A is Dedekind if and only if for every nonzero prime ideal p in A, the localization A_p is a discrete valuation ring (in this context, with A_p we mean the localization in the complement of the prime ideal p, that is, $A_p = S_p^{-1}A$, where $S_p = A \backslash p$).*

The main interest of Dedekind domains is that although we cannot factor an element in general as a product of primes, we have a factorization in ideals that in some sense it is very similar to working over a factorial ring.

Theorem 1.5. *Let A be a Dedekind domain. Every proper nonzero ideal I of A factors uniquely as a product of prime ideals.*

The proof will follow after we proof several lemmas which are rather intuitive and in fact of easy verification.

Lemma 1.7. *Let A be a noetherian ring, then every ideal I in A contains a product of nonzero prime ideals.*

Proof. We proceed by contradiction, by choosing a maximal counterexample I, that obviously cannot be prime. The fact of not being prime means that we have two elements a, b such that $ab \in I$ but neither a nor b are in I. But take now the ideal $a + I$, that properly contains I. The same occurs with $b + I$, and furthermore the product of these two ideals is contained in I. But for hypothesis, since I is maximal, $a+I$ and $b+I$ are a product of prime ideals, and consequently I contains a product of prime ideals. □

Lemma 1.8. *Let A be a ring and I and J relatively prime ideals. Then, for $m, n \in \mathbb{N}$, I^m and J^n are relatively prime.*

Proof. If they were not relatively prime they are contained in some prime ideal, but if a prime contains the m-th power on an element, it contains the element. Therefore, if I^m is in the ideal, so is I, and the same for J. We have therefore that I, J are relatively prime ideals contained in a prime ideal, and that's not possible. □

Lemma 1.9. *Let p be a maximal ideal of an integral domain A, and let q be the ideal it generates in A_p, $q = pA_p$. Then the map*

$$A/p^m \to A_p/q^m, a + p^m \mapsto a + q^m$$

is an isomorphism.

The proof is just an easy verification, where the key fact for injectivity is that $q^m \cap A = p^m$.

We can now prove the theorem: according to the first lemma, a contains a product of nonzero prime ideals, $b = p_1^{r_1} \cdots p_n^{r_n}$, where the p_i are distinct. Then, using the Chinese remainder theorem,

$$A/b \simeq \prod A/p_i^{r_i} \simeq \prod A_{p_i}/q_i^{r_i}$$

where $q_i = p_i A_{p_i}$ is the maximal ideal of A_{p_i}. That way, a/b is $\prod q_i^{s_i}/q_i^{r_i}$, where $s_i \leq r_i$ (we are using that it is a DVR). We conclude that $a = \prod p_i^{s_i}$ in A/b, and since these ideals both contain b, we conclude that $a = p_i^{s_i}$ in A (for the usual correspondence between ideals in A/b and ideals in A containing b).

We only have to check uniqueness. If $\prod p_i^{s_i} = a = \prod p_i^{t_i}$, then $q_i^{s_i} = aA_{p_i} = q_i^{t_i}$, where q_i is the maximal ideal in A_{p_i}. Therefore, $s_i = t_i$.

Factorization in extensions

We consider a Dedekind domain A whose field of fractions is K, and we consider the integral closure of A, B, in a finite separable extension L of K. A prime ideal p of A will factor in B as $pB = P_1^{e_1} \dots P_g^{e_g}$ (here we will use capital letter for primes of B instead of the most conventional notation of gothic characters), where the e_i are positive integers. If some of the e_i is strictly greater than 1, p is said to be ramified in B. e_i will be called the ramification index. As it would be expected, we say that P_1 divides p if it appears in its factorization in B. $e(P_i|p)$ or simply e_i will denote the ramification index and $f(P_i|p)$ or f_i will be the degree of the field extension $[B/P_i : A/p]$. A prime p splits in L if $e_i = f_i = 1$ for all i and is inert when pB is a prime ideal ($e = g = 1$).

This should be illustrated with an example for the sake of a better clarity: we take $A = \mathbb{Z}$, $B = \mathbb{Z}[i]$. In $\mathbb{Z}[i]$ we know from any undergraduate course in algebra that there are three kinds of primes:

- The prime $1 + i$, associated to the rational prime 2, that is the only prime of $\mathbb{Z}$ that ramifies, since $2 = (1 + i)^2$. Here, $e = 2, g = 1$ and f is also 1, since $\mathbb{Z}[i]/(1 + i) \simeq \mathbb{F}_2$.

- The primes $a + bi$ and $a - bi$, where a, b are chosen in such a way that $a^2 + b^2 = p$ where p is any rational prime of the form $4k + 1$. In this case $p = (a + bi)(a - bi)$, so $g = 2, e_i = 1$ and $f_i = 1$. This primes split in $\mathbb{Z}[i]$ and so $\mathbb{Z}[i]/(a + bi) = \mathbb{F}_p$

- The primes of $\mathbb{Z}$ of the form $4k+3$ are inert. In this case, $e = g = 1, f = 2$, since $\mathbb{Z}[i]/(p) \simeq \mathbb{F}_{p^2}$.

We present now an easy lemma and an important theorem.

Lemma 1.10. *A prime ideal P of B divides p if and only if $p = P \cap K$.*

Proof. Assume first that P divides p, from which we know that $p \subset P \cap K \subset A$. But p is a maximal ideal of A different from the total, so either $P \cap K = A$, which is not possible, or $p = P \cap K$.

For the other implication, from $p \subset P$, $pB \subset P$, and this implies that P appears in the factorization of pB. □

Theorem 1.6. *Let $m = [L : K]$, and let $P_1, \cdots, P_g$ be the prime ideals appearing in the factorization of p. Then $\sum_{i=1}^{g} e_i f_i = m$ and if L is Galois over K all the ramification numbers and residue class degrees are equal, verifying $efg = m$.*

Proof. We prove that both sides of the first equality are equal to $[B/pB : A/p]$. From the Chinese remainder theorem, $B/pB = \prod B/P_i^{e_i}$, so we have to show just that $e_i f_i = [B/p_i^{e_i} : A/p]$. But we know (from the definition of f_i) that B/p_i is a field of degree f_i over A/p. P_i^r/P_i^{r+1} is a B/P_i module and since there are no other ideals between P_i^r and P_i^{r+1}, its dimension as vector spaces is 1. From here, it follows that the dimension of $B/P_i^{e_i}$ is $e_i f_i$.

For the other equality, $[B/pB : A/p] = m$, note the following: if B is a free A module the result is almost immediate, since an isomorphism from A^n to B when tensored with K gives an isomorphism from K^n to L, showing that $m = n$, and when tensored with A/p gives an isomorphism $(A/p)^n \to B/pB$ showing that $n = [B/pB : A/p]$.

Consider now $S = A\backslash p$. Consider $B' = S^{-1}B, A' = S^{-1}A$. It is not difficult to show that B' is the integral closure of A' in L and that $pB' = \prod(P_iB')^{e_i}$. Consequently, $\sum e_if_i = [B'/pB' : A'/pA']$, and since A' is principal the result follows.
The second part of the statement follows directly form the fact that $\mathrm{Gal}(L/K)$ acts transitively on the prime ideals of B dividing p □

Theorem 1.7. *With the same notations that before, if B is a free A-module, a prime p ramifies in L if and only if it divides the discriminant of the extension.*

The ideal class group

Definition 1.6. *Let A be a Dedekind domain. A fractional ideal of A is a nonzero A-submodule I of K such that $dI = \{di \mid i \in I\}$ is contained in A for some nonzero d.*

Every nonzero element b of K defines a fractional ideal

$$(b) = bA = \{ba \mid a \in A\}$$

One such fractional ideal is called principal.
The following theorem affirms that the set of fractional ideals is a group:

Theorem 1.8. *Let A be a Dedekind domain. The set $\mathrm{Id}(A)$ of fractional ideals is a free abelian group on the set of nonzero prime ideals.*

We finish the section with a definition that will be crucial in several moments, for instance when explaining the theory of complex multiplication of elliptic curves:

Definition 1.7. *The ideal class group $\mathrm{CL}(A)$ is the quotient $\mathrm{Id}(A)/P(A)$ where $P(A)$ are the principal fractional ideals. Its order is the class number of K.*

1.3 The finiteness of the class number

In the following two sections we introduce some key theorems in algebraic number theory related with finiteness. The first one is that the group of ideal classes of a number field is finite. Related with this fact, we also have another remarkable result due to Hermite that asserts that the set of number fields with a given discriminant is finite. The third one is the unit theorem, that states that the set of units of a finite field is a finitely generated group and furthermore provides an effective way of knowing the rank.
For proving these results, the key observation is Minkowski's theorem. Before introducing it, we do a basic definition:

Definition 1.8. *Let V be a vector space of dimension n over $\mathbb{R}$. A lattice Λ in V is a subgroup of the form $\Lambda = \mathbb{Z}e_1 + \cdots + \mathbb{Z}e_r$, where the e_i are linearly independent. When $r = n$ the lattice is said to be full.*

When we have a full lattice $\Lambda = \sum \mathbb{Z}e_i$, we define for any $\lambda_0 \in \Lambda$ the fundamental parallelopied as

$$D = \{\lambda_0 + \sum a_i e_i \mid 0 \leq a_i < 1\}$$

It is not difficult to verify that its volume is given by the determinant of the e_i

Minkowski's theorem

Theorem 1.9. *Let Λ be a lattice in $\mathbb{R}^n$ whose fundamental parallelepiped is D_0. Let A be a compact, convex and symmetric in the origin subset of $\mathbb{R}^n$. If $\mathrm{Vol}(B) \geq 2^n \mathrm{Vol}(D_0)$, then B has a nontrivial point of Λ.*

We begin by showing an easy example of application:
Let $M \in \mathrm{SL}_3(\mathbb{Z})$ being positive definite. Prove that there is a vector $v \in \mathbb{Z}^3$ such that $v^T M v = 1$.
The solution is in fact quite simple, taking as our lattice the points of integer coordinates (so the fundamental parallelepied has volume 1) and considering B the set of vectors in $\mathbb{R}^3$ such that $0 \leq v^T M v \leq (2-\epsilon)$. Clearly, B is compact, convex and symmetric, and we just have to prove that the volume is smaller or equal than 8; but for the condition of being positive definite, it is an ellipsoid, so the volume will be $\frac{4\pi(2-\epsilon)^{3/2}}{3} > 8$ (we are using that the determinant is 1 since in the denominator it appears that factor) and we are done. We show another application of Minkowski that will be useful in our work.

Lemma 1.11. *Let $q : V \to \mathbb{R}$ be a quadratic form on a real vector space (finite dimensional). If there is a lattice $\Lambda \in V$ such that for every constant k the set $\{P \in \Lambda \mid q(P) \leq k\}$ is finite and the only $P \in \Lambda$ with $q(P) = 0$ is $P = 0$, then q is positive definite on V.*

Proof. We begin by writing q in diagonal form as $q = x_1^2 + \cdots x_r^2 - x_{r+1}^2 - \cdots x_{r+s}^2$. We proceed by contradiction assuming that r is smaller than the dimension of V. Take now the shortest vector of Λ, and call it λ (the smallest value of $q(P)$ where P runs over the nonzero points of Λ, and for our hypothesis $\lambda > 0$). Take now the set

$$B(\delta) = \{(x_i) \in \mathbb{R}^n \mid x_1^2 \cdots x_r^2 \leq \lambda/2, x_{r+1}^2 + \cdots x_{r+s}^2 \leq \delta\}$$

All the vectors of $B(\delta)$ have length smaller than $\lambda/2$ so there are no elements of the lattice, but when increasing the value of δ we would violate Minkowski, reaching that way a contradiction. □

We give now a proof of Minkowski's theorem (quite elementary in fact):

Proof. We will begin by showing that any measurable set of $\mathbb{R}^n$ whose volume is greater than that of D_0 has two points whose difference is in Λ; then we will apply that result to $S = 1/2B = \{x/2 \mid x \in B\}$, whose volume is $\frac{\mathrm{Vol}(B)}{2^n} > \mathrm{Vol}(D_0)$, so there exist two points in B, α, β such that $\alpha/2 - \beta/2 \in \Lambda$. For the symmetry, $-\beta \in B$ and for the convexity $\frac{\alpha+(-\beta)}{2} \in B$.
To prove the first claim let S be our set; we observe that $\mathrm{Vol}(S) = \sum \mathrm{Vol}(S \cap D)$,

where we are summing over all the translated of the fundamental parallelepiped. Since $\mathrm{Vol}(S) > \mathrm{Vol}(D_0)$, at least two of the translated will overlap, so there exist $\alpha, \beta \in S$ such that $\alpha - \lambda = \beta - \lambda'$ $(\lambda \neq \lambda')$, and then $\alpha - \beta \in \Lambda \backslash \{0\}$. □

We go now to the proof of the finiteness of the class number. There are some quite technical proofs that we omit, and that can be found on Milne's book on Algebraic Number theory (chapter 4).

As usual, we will take A to be a Dedekind domain with field of fractions K, L a finite separable extension of K and B the integral closure of A in L. We will define a homomorphims $N : \mathrm{Id}(B) \to \mathrm{Id}(A)$ compatible with taking norm of elements. If P is a prime ideal in B, we define $N(P) = p^{f(P/p)}$, where $p = P \cap A$ (we will sometimes write N for the norm of an ideal). We state some of the main properties concerning norms (all of them are immediate):

Proposition 1.4. *a) For any nonzero ideal $a \subset A, N_{L/K}(aB) = a^m$.*

b) If L is Galois and if P is a nonzero prime ideal of B, $p = P \cap A$, we already know that $pB = (P_1 \cdots P_g)^e$. Then $N(pB) = (P_1 \cdots P_g)^{ef} = \prod_{\sigma \in \mathrm{Gal}(L/K)} \sigma P$.

c) Let $\beta \in B$ be a nonzero element. Then, $\mathrm{Nm}(\beta) \cdot A = \mathrm{Nm}(\beta \cdot B)$.

Let now a be a nonzero ideal in the ring of integers O_K of a number field K; since a is of finite index in O_K, we can define the numerical norm of the ideal a, $\mathbb{N}a$ to be that index.

Proposition 1.5. *The numerical norm is multiplicative and if $b \subset a$ are fractional ideals of K, then $(a : b) = \mathbb{N}(a^{-1}b)$.*

The main result toward the proof of the finiteness of the class number is the following:

Theorem 1.10. *Let K be an extension of $\mathbb{Q}$ of degree n, Δ_K its discriminant and $2s$ the number of complex embeddings. Then, there exists a set of representatives for the ideal class group of K consisting of integral ideals a with*

$$N(a) \leq \frac{n!}{n^n}\left(\frac{4}{\pi}\right)^s |\Delta_K|^{1/2}$$

The quantity of the right is called the Minkowski bound.

From this, we have a theorem that, to be more precise, is a corollary of the previous result:

Theorem 1.11. *The class number of K is finite.*

Proof. We will be done if we show that there are only finitely many integral ideals a in O_K with $N(a) < M$. If $a = \prod P_i^{r_i}$, $N(a) = \prod p_i^{r_i f_i}$, where $(p_i) = P_i \cap \mathbb{Z}$. For the boundedness condition, we have only finitely many possibilities for the p_i (and so for the P_i) and obviously for the exponents r_i. □

We will consider an example for the sake of clarity: we are going to compute the class group of $K = \mathbb{Q}(\sqrt{82})$. $\Delta_K = 4 \cdot 82$, $s = 0$, so the Minkowski bound is $\sqrt{82}$. We have to look at the primes of $2, 3, 5, 7$. For that, we observe the factorization of $T^2 - 82$ over the different $\mathbb{F}_p$, since a well-known result is the following one:

Proposition 1.6. *Let $B = A[\alpha]$, where A is a Dedekind domain and B its integral closure in a finite separable extension L of its field of fractions K. Let $f(X)$ be the minimum polynomial of α over K, and let p be a prime ideal in A. Write $f(X) \equiv \prod g_i(X)^{e_i}$ modulo p, where the g_i are distinct and irreducible in $\mathbb{F}_p$. Then,*

$$pB = \prod (p, g_i(\alpha))^{e_i}$$

is the factorization of pB into a product of powers of distinct prime ideals.

For $p = 5, 7$, note that 3 is not a square, so the polynomial is irreducible and p is inert. For $p = 2$ it factors as T^2, so $(2) = P_2^2$ and for $p = 3$ it is $(T-1)(T+1)$ so $(3) = P_3 P_3'$. Thus, the class group is generated by $[P_2]$ (that has order 2) and by $[P_3]$. Since the norm of $10 + \sqrt{82}$ is 18 and that number is not divisible by 3, only P_3 or P_3' is a divisor (assume for instance P_3). Therefore, $(10 + \sqrt{82}) = P_2 P_3^2$, so $p_2 \equiv p_3^{-2}$ so the class group is generated by $[P_3]$ and has order dividing 4. If we show that P_2 is non-principal, we will be done and the class group will be $\mathbb{Z}/4\mathbb{Z}$. If $P_2 = (a + b\sqrt{82})$, then $(2) = P_2^2 = ((a + b\sqrt{82})^2)$, so $2 = (a + b\sqrt{82})^2 u$, where u is a unit. Taking norms, $N(u) = 1$, but the unit group is $\pm(9 + \sqrt{82})^n, n \in \mathbb{Z}$ (it is non-immediate at all; the next section clarifies the unit structure of a number field). Therefore, the units of norm 1 are the even power of $9 + \sqrt{82}$, that are all squares, so 2 is also a square, $2 = (a + b\sqrt{82})^2$ and that is not possible.

Recall that in $V = \mathbb{R}^r \times \mathbb{C}^s$ we have a norm defined by

$$||x|| = \sum_{i=1}^{r} |x_i| + 2 \sum_{i=r+1}^{r+s} |z_i|$$

and using the usual tools of calculus (changes of coordinates, manipulations of the Gamma function) it is possible to prove

Lemma 1.12. *Let $t > 0$ be a real number, and let*

$$X(t) = \{x \in V \text{ such that } ||x|| \leq t\}$$

Then,

$$\mu(X(t)) = 2^r (\pi/2)^s t^n / n!$$

The proof of the main theorem (the finiteness of the class number) relies now on the following statement:

Proposition 1.7. *Let K a number field of degree n and let a be a nonzero ideal in O_K. Then $\sigma(a)$ is a full lattice in V and the volume of a fundamental parallelepiped of $\sigma(a)$ is $2^{-s}\mathbb{N}(a)|\Delta_K|^{1/2}$.*

Proof. We take a basis for a as a $\mathbb{Z}$-module: $\alpha_1, \cdots, \alpha_n$. To prove this, we consider the corresponding images under the following morphism (that will appear in this section and in the following one):

$$\sigma : K \to \mathbb{R}^r \times \mathbb{C}^s : \alpha \mapsto (\sigma_1\alpha, \cdots, \sigma_{r+s}\alpha)$$

and prove that they are also linearly independent. Here, σ_i for $i = 1, \cdots, r$ denotes the set of real embeddings, and $\sigma_{r+1}, \bar{\sigma}_{r+1}, \cdots$ the set of complex embeddings.
To prove the independence, consider two different matrices. First of all, matrix A will be that whose i-th row is

$$(\sigma_1(\alpha_i), \cdots, \sigma_r(\alpha_i), \Re(\sigma_{r+1}(\alpha_i)), \Im(\sigma_{r+1}(\alpha_i)), \cdots)$$

and let B that whose i-th row is

$$(\sigma_1(\alpha_i), \cdots, \sigma_r(\alpha_i), \sigma_{r+1}(\alpha_i), \bar{\sigma}_{r+1}(\alpha_i), \cdots)$$

Clearly one can be obtained just by a linear combination of the other, so we will show that the determinant of A is nonzero seeing that the determinant of B is nonzero. But then

$$\det(A) = (-2i)^{-s}\det(B)^2 = \pm(-2i)^{-s}D(\alpha_1, \cdots, \alpha_n)^{1/2} \neq 0$$

where D is the discriminant of the α_i. Then $\sigma(a)$ is a lattice whose fundamental parallelepiped has volume $|\det(A)|$. Furthermore, from the identity

$$|D(\alpha_1, \cdots \alpha_n)| = (O_K : a)^2 \cdot |\operatorname{disc}(O_K/\mathbb{Z})|$$

we get that the measure of D is what we wanted. □

This proposition is the final step towards the proof of the theorem:

Proposition 1.8. *Let a be a nonzero ideal in O_K. Then a contains a nonzero element α such that*

$$|\operatorname{Nm}(\alpha)| \leq B_K Na = \left(\frac{4}{\pi}\right)^s \frac{n!}{n^n}\mathbb{N}a|\Delta_K|^{1/2}$$

Proof. Let $X(t)$ be the set of elements in V whose norm is smaller than t and let D be a fundamental domain for the lattice $\sigma(a)$. Then $X(t)$ is compact, convex and symmetric so choosing a sufficiently large t verifying the hypothesis of Minkowski, we have that $X(t)$ contains a point $\sigma(\alpha) \neq 0$. For this $\alpha \in a$, using the mean inequalities, we have that is norm is smaller or equal than $t^n/n!$.
In particular, if we want $\mu(X(t)) \geq 2^n\mu(D)$, we need that $2^r(\pi/2)^s t^n/n! \geq 2^n 2^{-s}\mathbb{N}a|\Delta_K|^{1/2}$ We can take now a t such that equality holds, and in that case

$$|\operatorname{Nm}(\alpha)| \leq t^n/n! = \frac{n!2^{2s}}{n^n\pi^s}\mathbb{N}a|\Delta_K|^{1/2}$$

that is the desired formula. □

We can now finally prove the finiteness of the class number:

Proof. Let c be a fractional ideal in K, and let us prove that its class is represented by an integral ideal a whose norm is smaller than our bound $B_K = \left(\frac{4}{\pi}\right)^s \frac{n!}{n^n}|\Delta_K|^{1/2}$. Take $d \in K^*$ such that dc^{-1} is an integral ideal, $(d)c^{-1} = b$, so there is a $\beta \in b$ such that its norm is smaller than $B_k{\cdot}\mathbb{N}b$ (previous result). We now have that $\beta O_K \subset b$, and from here $\beta O_K = ab$, where a is integral and equivalent to b^{-1} and to c in the class group. Furthermore, $\mathbb{N}a \cdot \mathbb{N}b = |\operatorname{Nm}_{K/\mathbb{Q}} \beta| \leq B_K \cdot \mathbb{N}b$ and the theorem is proved. □

1.4 Dirichlet's unit theorem

In this section we prove one of the main theorems in algebraic number theory, with several applications in which will follow. In a basic undergraduate course in algebra, we easily prove that $\mathbb{Z}[i]^* \simeq \mathbb{Z}/4\mathbb{Z}$ or that $\mathbb{Z}[\sqrt{2}]^* \simeq \mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$. In fact, it is not difficult to prove that the set of units in an imaginary quadratic field is finite (using the norm). Here we prove a stronger result:

Theorem 1.12. *The group of units in a number field K is finitely generated with rank equal to $r + s - 1$, where r is the number of real embeddings of K and $2s$ the number of non-real complex embeddings.*

Lemma 1.13. *An element $\alpha \in K$ is a unit if and only if $\alpha \in O_K$ and $\operatorname{Nm}_{K/\mathbb{Q}} \alpha = \pm 1$*

Proof. If α is a unit, we have $\beta \in O_K$ such that $\alpha\beta = 1$; the norms of α, β are integers, so $1 = \operatorname{Nm}(\alpha\beta) = \operatorname{Nm}(\alpha)\operatorname{Nm}(\beta)$. It is clear that is value is ± 1.
The converse is also easy. Fix an embedding σ_0 of K into $\mathbb{C}$ and recall that $\operatorname{Nm}(\alpha) = \alpha \prod_{\sigma \neq \sigma_0} \sigma\alpha$. If $\beta = \prod_{\sigma \neq \sigma_0} \sigma\alpha$, then it is an algebraic integer, since each of the factors is (they are roots of the same monic polynomial) and so β is an integer. If $\operatorname{Nm}(\alpha) = \pm 1$, then $\alpha\beta = \pm 1$ and so $\beta \in O_K$. We conclude that whenever $\alpha \in O_K$ has norm ± 1, it has an inverse $\pm\beta \in O_K$, so it is a unit. □

Lemma 1.14. *For any integers m, M, the set of algebraic integers α such that the degree of α is $\leq m$ and $|\alpha'| < M$ for all conjugates α' of α is finite.*

Proof. We are looking for irreducible polynomials of degree m whose coefficients are bounded, since they are symmetric polynomials evaluated at points whose norm is also bounded: therefore, each coefficient is bounded (using the triangular inequality) by a binomial coefficient (the number of summands of the symmetric polynomials) multiplied by M raised to the degree of this homogeneous polynomial. □

An immediate consequence of this is that an algebraic integer, each of whose conjugates in $\mathbb{C}$ has norm 1, is a root of 1. This is due to the fact that $\{1, \alpha, \alpha^2, \cdots\}$ is finite (according to the proposition).

Let's now consider the same map than in the previous section:

$$\sigma : K \to \mathbb{R}^r \times \mathbb{C}^s : \alpha \mapsto (\sigma_1\alpha, \cdots \sigma_r\alpha, \sigma_{r+1}\alpha, \cdots \sigma_{r+s}\alpha)$$

where the σ_i are the corresponding embeddings (the first ones are real, and the second ones are supposed to be taken in such a way that for σ_{r+i} we also have the corresponding conjugate).
In a similar way, we define

$$L : K^* \to \mathbb{R}^r \times \mathbb{C}^s : \alpha \mapsto (\log|\sigma_1\alpha_1|, \cdots \log|\sigma_r\alpha|, \log|\sigma_{r+1}\alpha|, \cdots \log|\sigma_{r+s}\alpha|)$$

that is a homomorphism. When u is a unit in O_K, its norm is ± 1 and each of the logarithms is 0, so the set of units is contained in the hyperplane

$$H : x_1 + \cdots + x_r + 2x_{r+1} + \cdots + 2x_{r+s} = 0$$

Note that there is an isomorphism between H and $\mathbb{R}^{r+s-1}$.

Lemma 1.15. *The image of $L : U \to H$ is a lattice in H, whose kernel is a finite group (here U denotes the group of units of O_K).*

Proof. Consider the bounded subset C of H formed by those elements of norm smaller or equal than M. If $L(u) \in C$, we know that $|\sigma_j u| \leq e^M$ for each j, and that only happens for finitely many u (by the previous lemma). Since $L(U) \cap C$ is finite, $L(U)$ is a lattice in H. For the kernel, the same reasoning applies, since if an element is there, $|\sigma_i\alpha| = 1$ for all i, and we use the same lemma. Because the kernel is finite, rank(U) =rank$(L(U)) \leq \dim H = r + s - 1$. □

We only need to prove now that the image $L(U)$ in H is a full lattice and this would complete the proof of the main theorem.

Proof. For that, consider again $\sigma : K \to \mathbb{R}^r \times \mathbb{C}^s$ and for $x = (x_1, \cdots, x_r, \cdots) \in \mathbb{R}^r \times \mathbb{C}^s$ define $\mathrm{Nm}(x) = x_1 \cdots x_r x_{r+1}\bar{x}_{r+1} \cdots x_{r+s}\bar{x}_{r+s}$. Clearly, $\mathrm{Nm}(\sigma(\alpha)) = \mathrm{Nm}(\alpha)$ and $|\mathrm{Nm}(x)| = |x_1| \cdots |x_r||x_{r+1}|^2 |x_{r+s}|^2$. We have already seen that $\sigma(O_K)$ is a full lattice and the volume of its fundamental parallelepiped is $2^{-s}|\Delta|^{1/2}$. We will take now an element $x \in \mathbb{R}^r \times \mathbb{C}^s$ with the norm between 1/2 and 1. Define then $x\sigma(O_K) = \{x\sigma(\alpha) \mid \alpha \in O_K\}$. This is again a lattice whose fundamental parallelepiped has volume the determinant of the matrix whose i-th row is

$$(x_1\sigma_1(\alpha_i), \cdots, \Re(x_{r+1}\sigma_{r+1}(\alpha_i)), \Im(x_{r+1}\sigma_{r+1}(\alpha_i)), \cdots)$$

In the same way than before, manipulating the expression, its absolute value is 2^{-s} times the absolute value of the determinant whose i-th row is

$$(x_1\sigma_1(\alpha_i), \cdots, x_{r+1}\sigma_{r+1}(\alpha_i)), \bar{x}_{r+1}\bar{\sigma}_{r+1}(\alpha_i)), \cdots)$$

Therefore, $x\sigma(O_K)$ is a lattice of volume $2^{-s}|\Delta|^{1/2}|\mathrm{Nm}(x)|$ (so when x varies this quantity is bounded).
Let now T be a compact convex subset, symmetric in the origin and whose volume

is so large that, for every x with the norm between $1/2$ and 1, the hypothesis of Minkowski's theorem hold and there is a nonzero point $\gamma \in O_K$ such that $x\sigma(\gamma) \in T$. The norm of the points of T is bounded, so there is an M such that $|\operatorname{Nm}(\gamma)| \leq 2M$.
Consider now the set of ideals γO_K, where γ runs through the γ's in O_K such that $x\sigma(\gamma) \in T$ for some x in our set. The norm of such an ideal is $\leq 2M$, so there are only finitely many ideals $\gamma_1 O_K, \cdots, \gamma_t O_K$. Note that if $\gamma \in O_K$ is such that $x\sigma(\gamma) \in T$, then $\gamma O_K = \gamma_i O_K$ for some i, and consequently $\gamma = \gamma_i \epsilon$. We have so that $x\sigma(\epsilon) \in \sigma(\gamma_i^{-1})T$. Since the set $T' = \sigma(\gamma_1^{-1})T \cup \cdots \cup \sigma(\gamma_t^{-1})T$ is bounded, we have shown that for each x there exists a unit ϵ such that the coordinates of $x\sigma(\epsilon)$ are bounded uniformly in x, since the set T' does not depend on the choice of x.

We can prove now that $L(U)$ is a full lattice. When $r + s - 1 = 0$, it is trivial, so assume is ≥ 1. For an i such that $1 \leq i \leq r + s$ take an x such that all the coordinates of x but x_i are very large, in such a way that $|\operatorname{Nm}(x)| = 1$. Since there exists ϵ_i such that $x\sigma(\epsilon_i)$ has bounded coordinates, $|\sigma_j \epsilon_i| < 1$ for $j \neq i$ and so $\log|\sigma_j \epsilon_i| < 0$. We will be done by proving that $L(\epsilon_1), \cdots, L(\epsilon_{r+s-1})$ are linearly independent in $L(U)$, so take the matrix whose i-th row is

$$(l_1(\epsilon_i), \cdots, 2l_r(\epsilon_i), 2l_{r+1}(\epsilon_i), \cdots, 2l_{r+s-1}(\epsilon_i))$$

where $l_i(\epsilon) = \log|\sigma_i \epsilon|$. All the elements not lying in the diagonal are negative, but the sum of each row is positive. In can be easily proved that any real matrix with negatives entries in the diagonal and positive sum of each row is invertible, and so we are done. □

1.5 Absolute values and local fields

Definition 1.9. *An absolute value or valuation on a field K is a function $x \mapsto |x| : K \to \mathbb{R}$ such that:*

a) $|x| \geq 0$ with equality if and only if $x = 0$.

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

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

A classical result in number theory is a theorem due to Ostrowski that says which are all the absolute values on $\mathbb{Q}$ (up to equivalence). We will denote $|\cdot|_\infty$ the usual absolute value on $\mathbb{R}$ and say that is normalized.

Theorem 1.13. *Let $|\cdot|$ be a nontrivial absolute value on $\mathbb{Q}$. If it is archimedean, then it is equivalent to $|\cdot|_\infty$ and if it is not, it is equivalent to $|\cdot|_p$ for exactly one prime p.*

Proof. We consider the expansion of a certain number m in base n (both m and n greater than one): $m = a_0 + \cdots + a_r n^r$. We take $N = \max\{1, |n|\}$, so by the

triangle equality $|m| \leq \sum |a_i||n|^i \leq \sum |a_i||N|^i \leq N^r \sum |a_i|$.
It is clear that $r \leq \log_n(m) = \log(m)/\log(n)$ and again the triangular inequality (applied a_i times) gives us $|a_i| \leq n$. Combining all these things in the previous inequality (and recalling that the sum has $r+1$ summands, we get

$$|m| \leq N^r(1+r)n \leq (1+\log(m)/\log(n))nN^{\log(m)/\log(n)}$$

Now we will substitute m by a power, m^q, take roots and we will see what happens when t goes to infinity:

$$|m| \leq (1+t\log(m)/\log(n))^{1/t}n^{1/t}N^{\log(m)/\log(n)}$$

Note that $n^{1/t}$ goes to 1 and $(1+tk)^{1/t}$ also goes to 1 for any constant k (for l'Hopital, for instance). We conclude that $|m| \leq N^{\log(m)/\log(n)}$. We now have two cases:

- $|n| > 1$ for all integers $n > 1$. Taking $\log(m)$-roots in the previous inequality, it yields that $|m|^{1/\log(m)} \leq |n|^{1/\log(n)}$ and for the symmetry we must have the reverse inequality, so there is a constant $c = |m|^{1/\log(m)} > 1$. Consequently,

  $$|n| = c^{\log(n)} = n^{\log(c)}$$

  for all integers $n > 1$. We conclude so that $|n| = |n|_\infty^a$, where $|\cdot|$ is the usual absolute value. But clearly, $|\cdot|$ and $|\cdot|_\infty^a$ are homomorphisms from $\mathbb{Q}^*$ to $\mathbb{R}_{>0}$ that agree on a set of generators (the primes and -1), so they agree on all of $\mathbb{Q}^*$. The value of a does not affect the topology we obtain, so they are all equivalent.

- For some $n > 1$, $|n| \leq 1$. In this case, take $N = 1$ in the inequality

  $$|m| \leq N^{\log(m)/\log(n)}$$

  and so $|m| \leq 1$ for all integers m, so it is non-archimedean. In that case, it makes sense to define A as the numbers with absolute value smaller than or equal to 1 and m as those with absolute value strictly smaller than 1. It is not difficult to see that m is maximal (and so it is prime) and that here $m \cap \mathbb{Z}$ is a prime ideal of $\mathbb{Z}$ (and nonzero). We conclude that $|m| = 1$ when m is not divisible by p, and so $|np^r| = |p|^r$, when n is a rational whose numerator and denominator are not divisible by p. Therefore, any such absolute value is equivalent to $|\cdot|_p$.

□

Corollary 1.3. *Let $a \neq 0, a \in \mathbb{Q}$. Then, $\prod |a|_p = 1$, where $|\cdot|_p$ denotes the normalized absolute value ($|p|_p = 1/p$).*

In a general number field K we have a similar picture. An equivalence class of absolute values will be called a prime or a place of K. The number of different places in an arbitrary number field is:

a) One for each prime ideal p.

b) One for each real embedding.

c) One for each pair of complex embeddings.

We will do a few remarks about this in next chapters.

Completions

It is common that we want to work with complete fields, and so extend our absolute value to have that every Cauchy sequence is convergent. In that sense, we have the following result:

Proposition 1.9. *Let K be a field with an absolute value $|\cdot|$. Then, there exists a complete valued field $(\hat{K}, |\cdot|)$ and a homomorphism from K to $\hat{K}$ preserving the absolute value and universal in the following way: a homomorphims from K into a complete valued field $(L, |\cdot|)$ preserving the absolute value extends uniquely to a homomorphism $\hat{K} \to L$.*

We begin by discussing the non-archimedean case. Take $|\cdot|$ a discrete non-archimedean absolute value on K and let π be an element whose valuation is the greatest one among those that are smaller than 1 (π is a generator of the maximal ideal m). π will be called a local uniformizing parameter, and the set of values taken by the absolute value is $\{|\pi|^m \mid m \in \mathbb{Z}\} \cup \{0\}$. When $a \in \hat{K}^*$, we can take a sequence a_n converging to a, and so $|a_n|$ also converges to $|a|$ (because the valuation is a continuous map) and consequently $|a|$ is a limit point for $|K^*|$, that is closed (for being discrete) and so $|a| \in |K^*|$. We conclude that $|\hat{K}| = |K|$, so $|\cdot|$ is a discrete absolute value on $\hat{K}$.
In $\hat{K}$, we have analogous constructions to those in K (such as the maximal ideal), and in particular if π is a generator of the maximal ideal of K also generates the maximal ideal of $\hat{K}$. In particular,

Lemma 1.16. *For every $n \in \mathbb{N}$, $A/m^n \to \hat{A}/\hat{m}^n$ is an isomorphism (where $\hat{m}$ is the maximal ideal of $\hat{K}$).*

The main result in that sense is the following one, that can be proved using the preceding lemma:

Proposition 1.10. *Let S be a set of representatives of A/m, and let π a generator of m. Then, the series*

$$\cdots + a_{-n}\pi^{-n} + \cdots + a_0 + a_1\pi + \cdots + a_m\pi^m + \cdots$$

is a Cauchy series, and conversely every Cauchy series is equivalent to exactly one of this form. That way, each element of $\hat{K}$ has a unique representative of this form.

An obvious example is $\mathbb{Q}_p$, the completion of the rational numbers with the p-adic valuation.
We begin with a proposition about how to extend a absolute value to a larger field (for instance, how to extend an absolute value of $\mathbb{Q}$ to a number field):

Proposition 1.11. *Let K be complete with respect to a discrete absolute value $|\cdot|_K$ and let L be a finite separable extension of K of degree n. Then, $|\cdot|$ extends uniquely to a discrete absolute value $|\cdot|_L$ on L, and L is complete for the extended absolute value. Furthermore, if $\beta \in L$, $|\beta|_L = |\operatorname{Nm}_{L/K}\beta|_K^{1/n}$.*

1.6 Global fields

A global field is an algebraic number field (finite extension of $\mathbb{Q}$ or a function field in one variable over a finite field. Our main interest will be in the first case. We already know that when K is a field with an absolute value (archimedean or discrete non-archimedean), and L is a finite separable extension of K, then there is a unique extension of $|\cdot|$ to L. We are interested now in the case when K is not complete. We will state here the main results in that sense, and since they will not be specially relevant in our work, we do not provide proofs (that are quite technical):

Proposition 1.12. *Let $L = K(\alpha)$ be a finite separable extension of K and let $f(X)$ be the minimum polynomial of α over K. Then, there is a natural one to one correspondence between the extensions of $|\cdot|$ to L and the irreducible factors of $f(X)$ in $\hat{K}(X)$.*

Consider so a finite Galois extension L of a number field K and let $G = \mathrm{Gal}(L/K)$. If w is an absolute value of L, we write σw for the absolute value such that $|\sigma\alpha|_{\sigma w} = |\alpha|_w$. For instance, if w is the valuation defined by a prime ideal P, σw is the valuation defined by the prime ideal σP. An important remark is that G acts on the set of primes of L lying over a prime v of K, so it is natural to define the decomposition group (or splitting group) G_w of w to be the stabilizer of w in G. A $\sigma \in G_w$ extends uniquely to a continuous automorphism of L_w and also $G_{\tau w} = \tau G_w \tau^{-1}$.

Proposition 1.13. *The homomorphism $G_w \to \mathrm{Gal}(L_w/K_v)$ is an isomorphism.*

We are going to make a fictional drawing of our situation, fixing once for all $\mathbb{Q}$ as the base field. For a maximal ideal P of L (L Galois extension) over a rational prime p, we have that the decomposition group is

$$D_p = \{\sigma \in \mathrm{Gal}(L/\mathbb{Q}) \mid \sigma P = P\}$$

This group has order ef so its index in the Galois group $\mathrm{Gal}(L/\mathbb{Q})$ is g. It clearly acts on the residue field $f_P = O_L/P$ by

$$(x+P)^\sigma = x^\sigma + P$$

The inertia group of P is the kernel of the action, that has order e. Since $\mathbb{F}_p$ is a subfield of f_P, there is an injection

$$D_P/I_P \to \mathrm{Gal}(f_p/\mathbb{F}_p) = \langle \sigma_p \rangle$$

where σ is the Frobenius. But both groups have order f, and so the injection is an isomorphism. Recall that when $L/\mathbb{Q}$ is Galois, the Galois group acts transitively on the maximal ideals lying over p, so the decomposition and the inertia group corresponding to different primes over p are conjugate.

The main object in this picture is the Frobenius element, of a crucial importance in our future study of class field theory. In the same conditions than before, consider an ideal P of L unramified in L/K. The Frobenius element $\sigma = (P, L/K)$ is the element of $G(P)$ that acts as the Frobenius automorphism on the residue field, that is:

a) $\sigma \in G(P)$, i.e., $\sigma P = P$.

b) Let $\alpha \in O_L$; then, $\sigma\alpha \equiv \alpha^q$ modulo P, where q is the number of elements in the residue field O_K/p, where $p = P \cap K$.

We state without proof an important theorem of which Dirichlet's Theorem of primes in arithmetic progression is a special case. This is due to Chebotarev:

Theorem 1.14. *Let F be a Galois number field. Then, every element of* $\mathrm{Gal}(F/\mathbb{Q})$ *takes the form* Frob_P *for infinitely many maximal ideals P of O_F.*

1.7 Selmer example

Let's now put an example of how this local-global principle does not work in general, by using the same example shown by Selmer.

Theorem 1.15. $3x^3 + 4y^3 + 5y^3$ *has solutions both in $\mathbb{R}$ and in $\mathbb{Q}_p$, $p \geq 2$, but not in $\mathbb{Q}$.*

The proof of the first fact is not complicated and it uses Hensel's lemma several times. First of all, the existence of a real solution is straightforward. Let's now distinguish several cases. First of all, to look for a 3-adic solution, we put $x = 0, z = 1$, and we have the equation $4y^3 + 5 = 0$. In a first attempt, we see that this has a solution mod 3 ($a = 1$), but that is not enough, since defining $f(y) = 4y^3 + 5$, we would have that $v_p(f(a)) = 2$ and this is not strictly smaller than $v_p(f'(a))^2 = 2$. So we take the equation mod 27 and put $a = 7$, and in that case $v_p(f(a)) = 4$ and $v_p(f'(a)) = 1$, so by Hensel's lemma we can lift the solution to $\mathbb{Z}_3$.
We continue just by stating a trivial result in the theory of finite fields: in $\mathbb{F}_q^*$, the map sending x to x^3 is surjective if and only if q is of the form $3k + 2$ and elsewhere the image has index 3. In the first case, just putting $x = 1, z = 0$, we will have to find a zero of the function $f = 3 + 4y^3$, but that's trivial by Hensel again, since now we have that 3 does not divide p. If q has de form $3k + 1$, we can have that 3 is a cube (and in that case take $y = 1, z = -1$ and the equation $3x^3 = 1$ has a solution that we can lift with Hensel) or that 3 is not a cube. In that case, the subgroup of cubes in $\mathbb{F}_p^*$ has index 3 and $\{1, 3, 9\}$ is a set of representatives. We have then three possibilities:

- If $5 = 1$, 5 is a cube and we can take $x = 1, y = -1$ just like above for having $5z^3 = 1$ and since 5 is a cube so is its inverse.
- If $5 = 3$, then $5/3$ is a cube so we take $y = 0, z = -1$ and we are done.
- If $5 = 9$, we have to proceed in a slightly different way; until now we were taking the solutions mod p and then lifting them; now we work directly in $\mathbb{Z}_p$, where 15 is a cube (since it is in $\mathbb{F}_p^*$ for our assumption). We take a such that $a^3 = 15$ and put $(3a/7, 5/7, -1)$ that works in $\mathbb{Q}_p$.

We now proof the non-existence of solutions in $\mathbb{Q}$ proceeding by contradiction. Multiplying by 2 and rearranging the variables, we can rewrite the equation like $x^3+6y^3 = 10z^3$. For being the equation homogeneous, we can assume that we are working in $\mathbb{Z}$, and in particular, if one of them is zero so are the other two (since neither 6 or 10 are cubes). Again for being homogeneous, we can assume that we have a primitive solution and furthermore, if a prime p divides two of x, y, z, it should divide the other; therefore, the three numbers are pairwise coprime. We also have that x, z are not divisible by 3 and x, y are not divisible by 5. We factor the LHS using $\alpha = \sqrt[3]{6}$ and we now have

$$(x + \alpha y)(x^2 - \alpha xy + \alpha^2 y^2) = 10z^3$$

It is an easy-to-prove result that the discriminant of $\mathbb{Z}[\sqrt[3]{d}]$ is $-27d^2$ (it is the matrix where we have all the entries 0 but $a_{11} = 3, a_{23} = a_{32} = -3d$), so in this case is $-4 \cdot 243$.

Lemma 1.17. *Let $K = \mathbb{Q}(\alpha)$, being α the root of a p-Eisenstein polynomial with degree n. Then*

a) p does not divide $[O_k : \mathbb{Z}[\alpha]]$.

b) p divides with multiplicity $n - 1$ the discriminant of K when p does not divide n and when $p|n$, we can only say that the multiplicity is at least n.

The key for proving the lemma is the observation that if $K/\mathbb{Q}$ is a number field of degree n of the form $K = \mathbb{Q}(\alpha)$ (where $\alpha \in O_K$), and its minimal polynomial over $\mathbb{Q}$ is p-Eisenstein, then for integer numbers a_i such that

$$a_0 + a_1\alpha + \cdots a_{n-1}\alpha^{n-1} \equiv 0 \mod pO_K$$

then a_i is a multiple of p for all i.
Since $t^3 - 6$ is both 2 and 3 Einsenstein the index of $\mathbb{Z}[\sqrt[3]{6}]$ in the ring of integers of $\mathbb{Q}(\sqrt[3]{6})$ is 1 (the index should divide the discriminant and we have seen that neither 2 nor 3 divides the index).

Our aim is to rewrite the previous equation as a product of principal ideals. We pass to an equation of ideals in $\mathbb{Z}[\alpha]$:

$$(x + \alpha y)(x^2 - \alpha xy + \alpha^2 y^2) = (10)(Z)^3$$

We will see how 10 factors. For that, consider again the polynomial $T^3 - 6$, that in $\mathbb{F}_2$ factors as p_2^3 and in $\mathbb{F}_5$ factors as $p_5 \cdot p_{25}$, where p_5 has degree 1 and p_{25} has degree 2. Therefore $(10) = p_2^3 p_5 p_{25}$.
Let now n be an integer. We know that $\mathrm{Nm}(n + \alpha) = n^3 + 6$ (here Nm denotes the norm), in such a way that $\mathrm{Nm}(-1 + \alpha) = 5, \mathrm{Nm}(-2 + \alpha) - 2$. That way, it must be $p_5 = -1 + \alpha, p_2 = -2 + \alpha$.

Lemma 1.18. $\mathbb{Z}[\alpha]$ *has class number* 1

Proof. The Minkowski bound is between 8 and 9, so we have to look at the primes smaller that 9. We have already detected which are the primes of norm 2 and 5, that are $p_2 = (-2+\alpha), p_5 = (-1+\alpha)$. From the factorization of $T^3 - 6$ we see that $(3) = p_3^3$; since $\mathrm{Nm}(\alpha) = 6, (\alpha) = p_2 p_3$, we conclude that p_3 is principal. For which concerns 7, $T^3 - 6 = (T-3)(T-5)(T-6)$, so $(7) = p_7 q_7 r_7$. We observe that $\mathrm{Nm}(1+\alpha) = 7$, so $(\alpha+1) = p_7$, $\mathrm{Nm}(2+\alpha) = 14, (\alpha+2) = p_2 q_2, \mathrm{Nm}(4+\alpha) = 70, (\alpha+4) = p_2 q_5 r_7$. From our knowledge that p_2 and p_5 are principal, we conclude that the others too. □

Lemma 1.19. *The units of $\mathbb{Z}[\alpha]$ modulo unit cubes are represented by*

$$(1 - 6\alpha + 3\alpha^2)^k$$

where $k = 0, 1, 2$.

Proof. We know that the unit group has rank 1, and trivially our units modulo cubes is a cyclic group of order 3; we just have to find a unit that is not a cube. Since $(2) = (-\alpha+2)^3$, the ratio $\frac{(2-\alpha)^3}{2} = 1 - 6\alpha + 3\alpha^2$ is therefore a unit; to see that is not a cube look at the residue field $Z[\alpha]/p_7$; there $1 - 6\alpha + 3\alpha^2 \equiv 3$, that is not a cube (the only cubes there are $0, 1, -1$). □

Note now that if a prime ideal p divides both $x + \alpha y$ and $x^2 - \alpha xy + y^2$, then $3xy\alpha \equiv 0$ in the residue field of p, so $p|(3)(x)(y)(\alpha)$. If p divides (y), it also divides (x), but we had assumed that x and y were relatively prime. Also, if $p|(x)$, then $p|(y)(\alpha)$, so $p|(\alpha)$. If $p|(3)$, then $\mathrm{Nm}(p)$ is a power of 3 and since p divides $(z)^3$, this would force z to be divisible by 3 and that is not possible. Therefore p divides (α) but not (3), so it is a factor of (2). We conclude that $p = p_2$. Consequently, $(x + y\alpha) = p_2 c, (x^2 - \alpha xy + \alpha^2 y^2) = p_2 c'$ (c, c' are coprime and p_2 does not divide c). Putting this factorization in the equation, we have $p_2^2 cc' = p_2^3 p_5 p_{25} (z)^3$, so p_2 is a factor of c'.
Again, from $x^3 + 6y^3 = 10z^3$, $x \equiv -y$ modulo 5, so $X + Y \equiv 0$ modulo p_5, what forces $p_5 | x + y\alpha$ (since $p_5 | (\alpha - 1)$). If p_{25} also divides it, then 5 would divide $x + y\alpha$ and x, y would be both divisible by 5 in $\mathbb{Z}$ and that is not true. Therefore p_{25} is a factor of $(x^2 - \alpha xy + \alpha^2 y^2)$. If $c = p_5 m$, $c' = p_2 p_{25} m'$, then

$$(x + \alpha y) = p_2 p_{25} m; \quad (x^2 - \alpha xy + \alpha^2 y^2) = p_2^2 p_{25} m'$$

and in addition $mm' = (z)^3$ so they are both cubes. In conclusion, we have the following equation

$$(x + \alpha y) = (\alpha - 2)(\alpha - 1) b^3$$

Use now that the class number is one, so the ideal b is principal, and also the previous lemma about units:

$$x + \alpha y = (\alpha - 2)(\alpha - 1)\beta^3 u$$

$$u = ((2-\alpha)^3/2)^k v^3 = ((2-\alpha)^k v)^3 / 2^k; v \in \mathbb{Z}[\alpha]^*; k \in \{0, 1, 2\}$$

$$2^k x + 2^k y\alpha = (\alpha - 2)(\alpha - 1)\gamma^3$$

Writing now $\gamma = a + b\alpha + c\alpha^2$, where a, b, c are integers not all 0, and plugging all these in the previous equation, we can equate the coefficients of α^2 to get

$$0 = a^3 + 6b^3 + 36c^3 + 36abc - 9(a^2b + 6ac^2 + 6b^2c) + 6(ab^2 + a^2c + 6bc^2)$$

Obviously $3|a$, and from kindergarten arguments, then $3|b$ and again $3|c$ (a classical Fermat descent in a homogeneous equation that lead us to conclude that all three are 0). We have consequently reached a contradiction and finished the proof.

Chapter 2

Algebraic curves and their jacobians

When we work with elliptic curves, we are frequently interested in arithmetic properties, but we cannot forget that we are dealing with objects defined also from a geometric point of view, so sometimes it is important to consider this approximation. This chapter is a brief summary of some of the algebraic geometry we need, and if this thesis would deal with deeper facts, it would be necessary to enlarge this part, since the theory of schemes introduced by Grothendieck and other developments along the last (say) seventy years are of great relevance in number theory.

2.1 Algebraic curves

A curve will be a projective variety of dimension one. We will denote by $\bar{K}[C]_P$ the local ring of C at P, and by M_P the maximal ideal of $\bar{K}[C]_P$. Let P be a non-singular point; using that M_P/M_P^2 is a one dimensional vector space over $\bar{K} = \bar{K}[C]_P/M_P$, we have the following useful result:

Proposition 2.1. *Let C be a curve and let $P \in C$ be a smooth point. Then, $\bar{K}[C]_P$ is a discrete valuation ring.*

Definition 2.1. *Let C and P as in the previous proposition, and let $f \in \bar{K}(C)$. The order of f at P is $\operatorname{ord}_P(f)$. When it is greater than 0, f has a zero at P, and when it is smaller it has a pole. When it is ≥ 0, f is regular or defined at P and we can evaluate $f(P)$.*

A uniformizer for C at P is any function $t \in \bar{K}(C)$ with order 1 (a generator of M_P). A first important result is the following:

Proposition 2.2. *Let C/K be a curve and let $t \in K(C)$ be a uniformizer at some non-singular point $P \in C(K)$. Then $K(C)$ is a finite separable extension of $K(t)$.*

We continue with a result that will appear for instance as our first step in the proof of Mordell's theorem:

Proposition 2.3. *Let $\phi : C_1 \to C_2$ be a morphism of curves (C_2 connected). Then ϕ is either constant or surjective.*

Consider two curves over K, C_1 and C_2 and a non-constant rational map (over K), $\phi : C_1 \to C_2$. Then composition with ϕ induces an injection of function fields that fixes K

$$\phi^* : K(C_2) \to K(C_1); \quad \phi^* f = f \circ \phi$$

Proposition 2.4. *Let C_1/K and C_2/K be two algebraic curves. With the previous notations,*

a) $K(C_1)$ is a finite extension of $\phi^*(K(C_2))$.

b) Let $\iota : K(C_2) \to K(C_1)$ an injection of function fields that fixes K. Then there exists a unique non-constant map $\phi : C_1 \to C_2$ such that $\phi^* = \iota$.

c) Let $\mathbb{K} \subset K(C_1)$ be a subfield of finite index that contains K. Then, there exists a smooth curve C'/K and a non-constant map $\phi : C_1 \to C'$ such that $\phi^* K(C') = \mathbb{K}$. Furthermore, C' is unique up to K -isomorphism.

When we have a map of curves defined over K, $\phi : C_1 \to C_2$, we can define its degree: it will be 0 when ϕ is constant, and otherwise we will say that is a finite map of degree $\deg \phi = [K(C_1) : \phi^* K(C_2)]$. We say that ϕ is separable, inseparable or purely inseparable when the corresponding field extension has that property.

These results show a strong connection between curves and their functions field. We can say that there is an equivalence of categories: on one side, smooth curves defined over K, where the maps are non-constant rational maps over K; on the other, finitely generated extension of $\mathbb{K}/K$ of transcendence degree one with $\mathbb{K} \cap \bar{K} = K$, being the maps field injections fixing K. The following two theorems (that could be called curves-fields correspondence) explain that fact:

Theorem 2.1. *The map $C \to K(C)$ induces a bijection from the set of isomorphim classes over K of non-singular projective algebraic curves over K to the set of conjugacy classes over K of function fields over K.*

That theorem gives the map from curve-classes to field-classes explicitly. To describe the map from field-classes to curve-classes, let $\mathbb{K}$ be a function field over K. Since the extension $\mathbb{K}/K(t)$ is finite, in characteristic zero the primitive element theorem says that $\mathbb{K} = K(t, u)$, where as we have already pointed out u satisfies an irreducible polynomial over $K(t)$. We will have therefore, clearing denominators, a relation of the form $\phi(t, u) = 0, \phi \in K[x, y]$. Since $\mathbb{K} \cap \bar{K} = K$, it can be seen that the polynomial $\phi(x, y)$ will be irreducible over $\bar{K}$ and the set of points $(x, y) \in \bar{K}^2$ such that $\phi(x, y) = 0$ gives a plane curve C'. The typical process of desingularizing (chapter 7 of Fulton) will produce a non-singular curve C with function field $\mathbb{K}$. The problem is that in general C' and C will not be isomorphic, just birationally equivalent. The point is that maybe the non-singular curve corresponding to the field $\mathbb{K}$ is not plane, and in fact in the theorem we are

not saying anything about that.
Let now $h : C \to C'$ be a non-constant morphism over K, and consider $h^* : K(C') \to K(C)$ defined as usual.

Theorem 2.2. *Let C, C' be non-singular projective algebraic curves over K. Then, the map $h \to h^*$ is a bijection from the set of non-constant morphisms over K from C to C' to the set of K-injections of $K(C')$ in $K(C)$.*

Definition 2.2. *Let $\phi : C_1 \to C_2$ be a non-constant map of smooth curves and let $P \in C_1$. The ramification index of ϕ at P, $e_\phi(P)$ is $e_\phi(P) = \mathrm{ord}_P(\phi^* t_{\phi(P)})$.*

Proposition 2.5. *Let $\phi : C_1 \to C_2$ be a non-constant map of smooth curves.*

a) For every $Q \in C_2$,

$$\sum_{P\in\phi^{-1}(Q)} e_\phi(P) = \deg(\phi)$$

b) For all but finitely may $Q \in C_2$, $\#\phi^{-1}(Q) = \deg_s(\phi)$

c) Let $\psi : C_2 \to C_3$ be another non-constant map of smooth curves. Then, $e_{\psi\circ\phi} = e_\phi(P)e_\psi(\phi P)$.

We introduce now the Frobenius map. For that, assume that $\mathrm{char}(K) = p > 0$ and let $q = p^r$. We define $f^{(q)}$ to be the polynomial obtained by raising each coefficient of the polynomial f to the q-th power. So when we have a curve C/K we can define a new curve $C^{(q)}/K$ (as the one whose homogeneous ideal is given by the one generated by $\{f^{(q)} : f \in I(C)\}$).

Proposition 2.6. *Let K be a field of characteristic $p > 0$, let $q = p^r$, let C/K be a curve and let $\phi : C \to C^{(q)}$ be the q-th power Frobenius morphism.*

a) $\phi^*(K(C^{(q)}) = K(C)^q$.

b) ϕ is purely inseparable.

c) $\deg \phi = q$.

We continue our walk through basic algebraic geometry recalling some basic definitions.

Definition 2.3. *The divisor group of a curve C is the free abelian group generated by the points of C. A divisor $D \in \mathrm{Div}(C)$ is principal if it has the form $D = \mathrm{div}(f)$ (where $\mathrm{div}(f) = \sum_{P\in C} \mathrm{ord}_P(f)(P)$) for some $f \in \bar{K}(C)^*$. Two divisors are equivalent when their difference is principal. The divisor class group or Picard group of C is the quotient of $\mathrm{Div}(C)$ by its subgroup of principal divisors. Finally, $\mathrm{Pic}_K(C)$ is the subgroup of $\mathrm{Pic}(C)$ fixed by $G_{\bar{K}/K}$.*

The degree of a divisor n_iP_i is $\sum n_i$; a super-index Pic^0 will mean that we are considering only those divisors of zero degree.

Definition 2.4. *Let C be a curve. The space of meromorphic differential forms of C, Ω_C, is the $\bar{K}$-vector space generated by symbols of the form dx, for $x \in K(C)$ subjec to the relations of linearity, Leibnitz rule and that $da = 0$ for all $a \in \bar{K}$.*

We now state Riemann-Roch theorem, one of the central results of algebraic geometry:

Theorem 2.3. *Let C be a smooth curve and let K_C be a canonical divisor on C. Let g be the genus of C. Then, for every divisor $D \in \mathrm{Div}(C)$,*

$$l(D) - l(K_C - D) = \deg D - g + 1$$

We conclude this first section with a classical relationship connecting the genera of curves linked by a non-constant map (Riemann-Hurwitz):

Theorem 2.4. *Let $\phi : C_1 \to C_2$ be a non-constant separable map of smooth curves of genera g_1 and g_2 respectively. Then,*

$$2g_1 - 2 \geq (\deg \phi)(2g_2 - 2) + \sum_{P \in C_1} (e_\phi(P) - 1)$$

with equality if and only if K is of characteristic zero or the characteristic p does not divide $e_\phi(P)$ for all $P \in C_1$.

2.2 Algebraic varieties

This section is not strictly necessary for this thesis, but in the last chapter we will try to explain some concepts that require the notion of scheme. We will not properly explain what does it mean, but we consider appropriate to state here the situation we have in the framework of classical algebraic geometry, to point out after how this can be generalized to give rise to the modern theory of schemes.

Ringed spaces

Let V be a topological space and k a field.

Definition 2.5. *Suppose that for every open subset U of V we have a set $O_V(U)$ of functions $U \to k$. Then O_V is a sheaf of k-algebras if it satisfies the following conditions:*

a) $O_V(U)$ is a k-subalgebra of the algebra of all k-valued functions on U (equivalently, $O_V(U)$ contains the constant functions and if f, g are in $O_V(U)$ then so also do $f + g$ and fg).

b) If U' is an open subset of U and $f \in O_V(U)$, then $f|U' \in O_V(U')$.

c) A function $f : U \to k$ on an open subset U of V is in $O_V(U)$ if $f|U_i \in O_V(U_i)$ for all U_i in some open covering of U.

When only the two first conditions are satisfied, we say that it is a presheaf.

Definition 2.6. *A pair (V, O_V) consisting on a topological space V and a sheaf of k-algebras will be called a ringed space. $O_V(U)$ is frequently written as $\Gamma(U, O_V)$. Its elements are called sections of O_V over U.*

A morphism of ringed spaces $(V, O_V) \to (W, O_W)$ is a continuous map $\phi : V \to W$ such that $f \in \Gamma(U, O_W)$ implies that $f \circ \phi \in \Gamma(\phi^{-1}U, O_V)$.
Every ringed space isomorphic to an algebraic set $V \in k^n$ is an affine algebraic variety over k. A map $f : V \to W$ of affine varieties is regular if it is a morphism of ringed spaces. An affine k-algebra is a reduced (without nilpotent elements) finitely generated k-algebra. We can attach a ringed space (V, O_V) to such an algebra A by letting V be the set of maximal ideals in A. For $f \in A$ let $D(f) = \{m \mid f(m) \neq 0\}$. We consider the topology for which the $D(f)$ form a base. It is immediate to see that the pair (V, O_V) is an affine (algebraic) variety with $\Gamma(V, O_V) = A$. We write $\mathrm{spm}(A)$ for the topological space V and $\mathrm{Spm}(A)$ for the ringed space (V, O_V).

Proposition 2.7. *A ringed space (V, O_V) is an affine variety if and only if $\Gamma(V, O_V)$ is an affine k-algebra and the canonical map $V \to \mathrm{spm}(\Gamma(V, O_V))$ is an isomorphism of ringed spaces.*

With the usual notations, we have the following elementary result:

Proposition 2.8. *Let $V = V(a) \subset k^m$, $W = V(b) \subset k^n$. The following conditions on a continuous map $\phi : V \to W$ are equivalent:*

a) ϕ is regular.

b) The components $\phi_1, \cdots, \phi_m$ of ϕ are all regular.

c) $f \in k[W]$ implies $f \circ \phi \in K[V]$.

Definition 2.7. *An algebraic prevariety over k is a ringed space (V, O_V) such that V is compact and every point of V has an open neighborhood U for which $(V, O_V|U)$ is an affine algebraic variety over k. When (V, O_V) and (W, O_W) are algebraic varieties, a map $\phi : V \to W$ is regular if it is a morphism of ringed spaces.*

Definition 2.8. *An algebraic variety is an algebraic prevariety that is separated, that is, for every pair of regular maps $\phi_1, \phi_2 : Z \to V$ with Z an affine algebraic variety, the set where $\phi_1(z) = \phi_2(z)$ is closed in Z.*

Definition 2.9. *An algebraic group is a variety G with regular maps, multiplication, inverse and identity $e : \mathbb{A}^0 \to G$, that make G into a group in the usual sense.*

Complete varieties

The introduction of this concrete type of varieties is necessary for the proper definition of abelian varieties that are the natural generalization of elliptic curves. In the category of algebraic varieties, complete varieties are the analogue to compact spaces in the category of Hausdorff topological spaces. The basic fact (that will be used several times, for instance at the beginning of the proof of Mordell-Weil) is that the image of a complete variety is complete, and we will look for analogies: in general topology, the image of a compact is compact, and if the space is Hausdorff, then it is closed; furthermore, a Hausdorff space is compact if and only if for all topological spaces W the projection map $\pi : V \times W \to W$ is closed.

Definition 2.10. *An algebraic variety V is complete if for all algebraic varieties W the projection map $\pi : V \times W \to W$ is closed.*

We now state the so called rigidity theorem:

Proposition 2.9. *Let $\phi : V \times W \to Z$ be a regular map, and assume that V is complete, that V and W are irreducible and that Z is separated. If there exists points $v_0 \in V$, $w_0 \in W$ and $z_0 \in Z$ such that*

$$\phi(V \times \{w_0\}) = \{z_0\} = \phi(\{v_0\} \times W)$$

then $\phi(V \times W) = \{z_0\}$.

2.3 Abel's theorem

Let ω be a closed C^∞ 1-form on X. If D is a triangulated subset of X, Stoke's theorem told us that $\int_{\partial D} w = \int \int_D dw = 0$. We know, therefore, that the integrals of ω around any closed chain only depend on the homology class of the chain. Hence for every homology class $[c]$ we obtain a well defined functional on the space $\Omega^1(X)$ of holomorphic 1-forms, which is integration around c.

Definition 2.11. *A linear functional $\lambda : \Omega^1(X) \to \mathbb{C}$ is a period if it is $\int_{[c]}$ for some homology class $[c]$.*

Note that the set of periods Λ is a subgroup of the dual space $\Omega^1(X)^*$.

Definition 2.12. *Let X be a compact Riemann surface. The jacobian of X, $\mathrm{Jac}(X)$ is the quotient group*

$$\mathrm{Jac}(X) = \frac{\Omega^1(X)^*}{\Lambda} = \frac{\Omega^1(X)^*}{H_1(X, Z)}$$

i.e., the space of holomorphic 1-forms modulo periods.

The Abel-Jacobi map

Choose a base point p_0 on the compact Riemann surface X and for any other point $p \in X$, we choose a path γ_p from p_0 to p contained in X. So it is possible to define the map $A : X \to \Omega^1(X)^* : A(p)(\omega) \mapsto \int_{\gamma_p} \omega$. Of course this is not well defined, and what we have to do is define it modulo periods (since the value of $A(p)$, when choosing a different path, changes by integration around a closed chain), so A goes from X to $\mathrm{Jac}(X)$. This is the so called Abel-Jacobi map for X, which depends on the base point.

Theorem 2.5. *Let X be a compact Riemann surface of genus g and let D be a divisor of degree* 0 *on X. Then D is principal if and only if $A_0(D) = 0$ in* $\mathrm{Jac}(X)$*.*

Another way of stating the theorem is the following: if we denote by $\mathrm{Pic}^0(X)$ the subgroup of $\mathrm{Pic}(X)$ given by classes of divisors of degree 0, what we are saying is that

$$\mathrm{Pic}^0(X) \cong \mathrm{Jac}(X)$$

For example, when $X = \hat{\mathbb{C}}$ the jacobian is trivial because $g = 0$ and Abel's Theorem states that every degree zero divisor on the Riemann sphere is principal. One of the main results around this thesis will be, as announced, modularity theorem: in one of its formulations, what it says is that there is a surjective holomorphic homomorphism between the jacobian of a certain curve $X_0(N)$ and a complex elliptic curve whose j-invariant is rational.

Maps between jacobians

Let $h : X \to Y$ be a non-constant holomorphic map of compact Riemann surfaces. We are going do define now forward and reverse holomorphic homomorphisms of jacobians, $h_J : \mathrm{Jac}(X) \to \mathrm{Jac}(Y), h^J : \mathrm{Jac}(Y) \to \mathrm{Jac}(X)$ and consequently, due to the isomorphism above mentioned, between the Picard groups h_P, h^P. Begin by considering the pullback map induced by h,

$$h^* : \mathbb{C}(Y) \to \mathbb{C}(X)$$

defined as usual as $h^*g = g \circ h$. From the theory of algebraic curves, each point $x \in X$ has a ramification degree $e_x \in \mathbb{Z}^+$ such that h is locally e_x-to-1 at x The following relation between the orders of vanishing ν of a function and its pullback will be useful

$$\nu_x(h^*g) = e_x \nu_{h(x)}(g)$$

where $g \in \mathbb{C}(Y)^*$. The important fact now is that the pullback can be extended to a linear map of holomorphic differential

$$h^* : \Omega^1_{\mathrm{hol}}(Y) \to \Omega^1_{\mathrm{hol}}(X)$$

We omit the explanation but is the rather typical procedure of geometry of working in local coordinates.

The pullback dualizes to a linear map of dual spaces, that we will denote by h_*.

Definition 2.13. *The forward map of jacobians is the holomorphic homomorphism induced by composition with the pullback,*

$$h_J : \mathrm{Jac}(X) \to \mathrm{Jac}(Y) \quad h_J[\phi] = [h_*\phi] = [\phi \circ h^*]$$

If we return to the general situation, a map $h : X \to Y$, we also have a norm map between $\mathbb{C}(X)$ and $\mathbb{C}(Y)$, denoted norm_h

$$(\mathrm{norm}_h\, f)(y) = \prod_{x \in h^{-1}(y)} f(x)^{e_x}$$

The orders of vanishing of a nonzero function and its norm are related by

$$\nu_y(\mathrm{norm}_h\, f) = \sum_{x \in h^{-1}(y)} \nu_x(f)$$

Consequently, we will have that

$$\mathrm{div}(\mathrm{norm}_h\, f) = \sum_y \Big(\sum_{x \in h^{-1}(y)} \nu_x(f) \Big) y = \sum_x \nu_x(f) h(x)$$

The map on general divisor that extends this will be

$$h_D : \mathrm{Div}(X) \to \mathrm{Div}(Y) \quad h_D(\sum_x n_x x) = \sum_x \nu_x(f) h(x)$$

Note that it takes zero-degree divisors to zero-degree divisors, and also principal divisors to principal divisors, since $h_D(\mathrm{div}(f)) = \mathrm{div}(\mathrm{norm}_h\, f)$.

Definition 2.14. *The forward map of Picard groups is the homomorphism*

$$h_P : \mathrm{Pic}^0(X) \to \mathrm{Pic}^0(Y) \quad h_P(\sum_x n_x x) = (\sum_x n_x h(x))$$

Defining the reverse map h^J is more delicate. Let now h be a surjection of finite degree d, locally e_x-to-1 at each $x \in X$, and let $\epsilon = \{x \in X : e_x > 1\}$ be the finite set of points where h is ramified. If $Y' = Y - h(\epsilon)$ and $X' = h^{-1}(Y')$ are the Riemann surfaces obtained by removing the exceptional points and their preimages, we get a restriction map $h : X' \to Y'$ that is a d-fold covering map. That means that every point $y \in Y'$ has a neighborhood $\tilde{U}$ whose inverse image is a disjoint union of neighborhoods $U_1, \dots, U_d$ in X' such that each restriction $h_i : U_i \to \tilde{U}$ of h is invertible.

We recall now some facts in general topology: given a path δ in Y' and a preimage $x \in h^{-1}(\delta(0))$ in X', there is a unique lift γ of δ to X' starting at x. When δ is a path in Y and only its endpoints might lie in $h(\epsilon)$ the local mapping theorem shows that for every $x \in h^{-1}(\delta(0))$ there exist e_x lifts of γ starting at x. When we have a loop β in Y' the map taking the initial point of each of its lifts to the terminal point is a permutation of the d-element set $h^{-1}(\beta(0))$. Summing up, since any path can be perturbed to avoid $h(\epsilon)$ without changing integration of holomorphic differential, any path integral of holomorphic differentials on Y can

be taken over a path δ such that only its endpoints might lie in $h(\epsilon)$.
Define the trace map induced by h as

$$\operatorname{tr}_h : \Omega^1_{\text{hol}}(X) \to \Omega^1_{\text{hol}}(Y)$$

To do this, we have to consider local inverse $h_i^{-1} : \tilde{U} \to U_i$ (in local coordinates) and the trace will be

$$(\operatorname{tr}_h \omega)|_{\tilde{U}} = \sum_{i=1}^{d} (h_i^{-1})^* (w|_{U_i})$$

As usual, this dualizes to a linear map of dual spaces

$$\operatorname{tr}_h^* : \Omega^1_{\text{hol}}(Y)^* \to \Omega^1_{\text{hol}}(X)^*$$

defined as $\operatorname{tr}_h^* \psi = \psi \circ \operatorname{tr}_h$.

Definition 2.15. *The reverse map of jacobians is the holomorphic homomorphism induced by composition with the trace*

$$h^J : \operatorname{Jac}(Y) \to \operatorname{Jac}(X) \quad h^J(\psi) = [\psi \circ \operatorname{tr}_h]$$

Proposition 2.10. *Let $h : X \to Y$ be a non-constant holomoprhic map of compact Riemann surfaces. Take h_* to be the dual map of h^* restricted to $H_1(X, \mathbb{Z})$*

$$h_* : H_1(X, \mathbb{Z}) \to H_1(Y, \mathbb{Z}) \quad h_*\Big(\sum_\alpha n_\alpha \int_\alpha\Big) = \sum_\alpha n_\alpha \int_{h\circ\alpha}$$

Then, $h_*(H_1(X, \mathbb{Z}))$ is a subgroup of finite index in $H_1(Y, \mathbb{Z})$. If V is a subspace of $\Omega^1_{\text{hol}}(Y)$ then the restriction $h_*(H_1(X, \mathbb{Z}))|V$ is a subgroup of finite index in $H_1(Y, \mathbb{Z})|V$.

To finish, we just establish the reverse map of Picard groups. Observe that

$$\operatorname{div}(h^* g) = \sum_x e_x \nu_{h(x)}(g) x = \sum_y \nu_y(g) \sum_{x \in h^{-1}(y)} e_x x$$

Definition 2.16. *The reverse map of Picard groups is the homomorphism*

$$h^P : \operatorname{Pic}^0(Y) \to \operatorname{Pic}^0(X) \quad h^P(\sum_y n_y y) = (\sum_y n_y \sum_{x \in h^{-1}(y)} e_x x)$$

2.4 Further results in Algebraic Geometry

Definition 2.17. *Let S be a set of meromorphic functions on a compact Riemann surface X. We say that S separates points of X if for every pair of distinct points p and q in X there is a meromorphic function $f \in S$ such that $f(p) \neq f(q)$. We say that S separates tangents of X if for every point $p \in X$ there is a meromorphic function $f \in S$ which has multiplicity one at p. A compact Riemann surface X is an algebraic curve if the field $M(X)$ of global meromorphic functions separates the points and tangents of X.*

This will be important for us, but we will be looking for something stronger, canonical models over the rationals for certain curves, and this will be more difficult A rather deep theorem is the following one:

Theorem 2.6. *Every compact Riemann surface is an algebraic curve.*

This is a deep theorem that requires tools of functional analysis for its proof. Indeed, it is not trivial that a compact Riemann surface has any nonconstant meromorphic functions at all. This result is used in the proof of Riemann Roch and in some sense it can be considered equivalent to it.
Another important result that will be around us in some moments is this:

Proposition 2.11. *Every algebraic curve of genus one is isomorphic to a smooth projective plane cubic curve.*

Abelian varieties

As we mentioned in the introduction, the main object of this thesis will be elliptic curves (together maybe with modular curves). A tentative definition of elliptic curve could be: a non-singular projective curve together with a group structure defined by regular maps (in fact, we will give four different definitions and just sketch how this can be taken as equivalent to them). But the important fact is that is the definition that can be generalized. We have then the concept of an abelian variety, that fulfills most of the good properties of elliptic curves.

Definition 2.18. *An abelian variety is a complete connected group variety.*

Proposition 2.12. *Every regular map $\alpha : A \to B$ of abelian varieties is the composite of a homomorphism with a translation; in particular, a regular map $\alpha : A \to B$ such that $\alpha(0) = 0$ is a homomorphism.*

Proof. After composing α with a translation we may suppose that $\alpha(0) = 0$. Consider the map $\phi : A \times A \to B$ given by $\phi(a, a') = \alpha(a + a') - \alpha(a) - \alpha(a')$. Then, $\phi(A \times 0) = 0 = \phi(0 \times A)$ and so $\phi = 0$ and α is a homomorphism. □

Proposition 2.13. *The group law on an abelian variety is commutative.*

Proof. A property that characterizes commutative groups is that the map that takes an element to its inverse is a homomorphism. Since the map that takes a to $-a$ takes the identity element to itself, by the preceding result we know that it is a homomorphism. □

Chapter 3

Elliptic curves: definitions and first properties

As we pointed out in the introduction, elliptic curves are one of the main topics of this essay, and one of the most well-known objects in number theory. They are introduced in any course in basic algebraic geometry when studying algebraic curves over algebraically closed fields. From that approach, it is convenient to keep in mind that they have a group structure, that can be defined with a particular geometric construction (the classical proof of associativity is in some sense weird). The important fact here is that they are group algebraic varieties, so we can take advantage both of its algebraic and geometric properties. We begin by giving four different (and equivalent) definitions of an elliptic curve over a field k that, for convenience will be assumed to be perfect (this is only necessary in some moments).

Definition 3.1. *An elliptic curve over k can be defined as:*

a) A non-singular projective plane curve E over k of degree 3 together with a point $0 \in E(k)$.

b) A non-singular projective plane curve E over k of degree 3 together with an **inflection** *point $0 \in E(k)$.*

c) A non-singular projective plane curve over k of the form

$$Y^2Z + a_1XYZ + a_3YZ^2 = X^3 + a_2X^2Z + a_4XZ^2 + a_6Z^3$$

d) A non-singular projective plane curve E of genus 1 together with a point $O \in E(k)$.

Proof. That *a)* implies *b)* follows from the fact that, in a projective cubic plane curve over k, we can transform a point to a point of inflection: this is maybe the hardest of the implications and is based in a change of variables suggested around 1928 by Nagell. Let $O \in C(k)$ and suppose that it is not an inflection point, so the tangent line to C at O meets C in another point P. We make a change of variables in such a way that the tangent line at O is the Y-axis and

$P = (0 : 0 : 1)$. That way, C does not have term in Z^3 and the corresponding affine curve (defined as the intersection of C with $Z = 1$) does not have constant term. Write $C^{\text{aff}} = F_1(X,Y) + F_2(X,Y) + F_3(X,Y)$, where F_i is homogeneous of degree i. Since $O = (0,y)$ is a double point, we have that y is a double root of $F_1(0,1)Y + F_2(0,1)Y^2 + F_3(0,1)Y^3$ (since when putting $F_i(0,1)$ we obtain the coefficient of Y^i). Extracting Y as a common factor, we have a second degree polynomial with a double root, so its discriminant is 0, i.e.,

$$F_2(0,1)^2 = 4F_1(0,1)F_3(0,1)$$

Consider now a line of the form $Y = tX$, and look at their intersection with C^{aff}; this happens at those points whose x-coordinate verifies

$$x(F_1(1,t) + xF_2(1,t) + x^2F_3(1,t)) = 0$$

We have by one side the origin and we must work with the other factor. The relation $F_1(1,t) + xF_2(1,t) + x^2F_3(1,t)$ can be written completing squares in the same spirit that when we solve a quadratic equation as

$$(2F_3(1,t)x + F_2(1,t))^2 = F_2(1,t)^2 - 4F_1(1,t)F_3(1,t)$$

Now we can define a change of variables in a natural way:

$$s \mapsto 2F_3(1,y/x)x + F_2(1,y/x) \text{ and } t \mapsto y/x$$

This defines a homomorphism $k[s,t] \to k[x,y][x^{-1}]$ where

$$s^2 = G(t) \text{ being } G(t) = F_2(1,t)^2 - 4F_1(1,t)F_3(1,t)$$

That way, we have a regular map from $C^{\text{aff}}\backslash\{0\} \to E$, where E is the affine curve defined by $s^2 = G(t)$. The polynomial $G(t)$ can be interpreted as a fourth-degree polynomial that we can homogenize. Previously, we saw that it has a root in $(0,1)$, so when substituting $(1,t)$ it can only have degree three (the homogeneous polynomial of degree four has already the root $(0,1)$). We can extend this application to a morphism $C^{\text{aff}} \to E$ by sending O to $(0 : 1 : 0)$ (the point with s-coordinate 1, since it does not appear term in s^3) and we get that the origin is now an inflection point.

We continue with the chain of implications: it is clear from basic algebraic geometry that after an invertible linear change of variables with coefficients in k, the point O (that was an inflection point) can have coordinates $(0 : 1 : 0)$ and the tangent line to a curve C at O can be $L_\infty : Z = 0$. The general form of the cubic would be

$$c_1X^3 + c_2X^2Y + c_3X^2Z + c_4XY^2 + c_5XYZ + c_6XZ^2 + c_7Y^3 + c_8Y^2Z + c_9YZ^2 + c_{10}Z^3$$

Since $(0 : 1 : 0) \in C(k)$, we must have $c_7 = 0$. We consider now

$$U_1 = \{(x : y : z) \mid y = 1\}$$

and we identify as usual U_1 with $\mathbb{A}^2$ through $(x : 1 : z) \mapsto (x, z)$. We also know that $C \cap U_1$ is the affine curve

$$c_1X^3 + c_2X^2 + c_3X^2Z + c_4X + c_5XZ + c_6XZ^2 + c_8Z + c_9Z^2 + c_{10}Z^3$$

but the tangent line at $(0, 0)$ is $c_4X + c_8Z = 0$ and by hypothesis it was $Z = 0$, so $c_4 = 0$ (and $c_8 \neq 0$ to be non-singular). The intersection number is $I(Z, c_1X^3 + c_2X^2) = I(Z, X^2) + I(Z, c_1X + c_2) = 2 + I(Z, c_1X + c_2)$, so we must have (since this intersection number is ≥ 3 for the condition of being a point of inflection) that $c_2 = 0$. In addition, we must have that $c_1 \neq 0$, otherwise the polynomial would be divisible by Z. We divide by c_1 and replace Z by $-c_1Z/c_8$, to finally obtain the desired equation.

The fact that *c)* implies *d)* is the easiest one: clearly, the equation defined in *c)* is non-singular and by the genus formula it has genus 1; the point $(0 : 1 : 0)$ also belongs to the curve.

Let now E be a complete non-singular curve of genus 1 over k and let $O \in E(k)$. By virtue of Riemann-Roch, the rational functions on E having no poles except at O (where they are allowed to have at most a pole of order m), form a k-vector space of dimension m (for $m \geq 1$). In $L([O])$ we have only constant functions ($\{1\}$ is a basis). Choose now x such that $\{1, x\}$ is a basis for $L(2[O])$ and y such that $\{1, x, y\}$ is a basis for $L(3[O])$. In $L(4[O])$ the basis will be $\{1, x, y, x^2\}$ and in $L(5[O])$, $\{1, x, y, x^2, xy\}$. But in $L(6[O])$ we already know about the existence of seven elements, $\{1, x, y, x^2, xy, y^2, x^3\}$, so we must have a relation between them of the form

$$a_0y^2 + a_1xy + a_3y = a_0'x^3 + a_2x^2 + a_4x + a_6$$

where both a_0 and a_0' must be nonzero; we can replace y with a_0y/a_0' and x with a_0x/a_0'. Multiplying by $a_0'^2/a_0^3$, the coefficients in y^2 and x^3 can be assumed to be both equal to 1. To sum up, the map $P \to (x(P), y(P))$ sends $E\backslash\{0\}$ onto the plane affine curve

$$C : Y^2 + a_1XY + a_2Y = X^3 + a_5X^2 + a_6X + a_7$$

x has a double pole at O so it has only two zeros, and so $x + c$ has two zeros, too (counting multiplicities). That way, the map

$$E\backslash\{0\} \to \mathbb{A}^1, P \mapsto x(P)$$

has degree two, and the map sending P to $y(P)$ has degree 3. Therefore, the degree of $E\backslash\{O\} \to C$ must divide both 2 and 3 so it is 1. If C were singular, its genus would be 0, and this cannot be. For being C non-singular, the map is an isomorphism and it extends to an isomorphism onto $\bar{C}$. □

3.1 The Weierstrass equation for an elliptic curve

Definition 3.2. *Let E an elliptic curve over k. An equation of the form*

$$Y^2Z + a_1XYZ + a_3YZ^2 = X^3 + a_2X^2Z + a_4XZ^2 + a_6Z^3$$

is called a Weierstrass equation for the elliptic curve. The quantity $\Delta = -16(4a^3 + 27b^2)$ *is the discriminant of the elliptic curve (we sometimes omit the factor* -16*).*

If $\mathrm{char}(k) \neq 2, 3$ we perform first the change

$$X' = X, Y' = Y + a_1/2X, Z' = Z$$

to eliminate the term XYZ, and then

$$X' = X + a_2/3, Y' = Y + a_3/2, Z' = Z$$

to eliminate the terms in X^2 and Y. We finally get an equation of the form

$$E(a, b) : Y^2Z = X^3 + aXZ^2 + bZ^3$$

which is the way kids know elliptic curves. Note that a curve in this form is non-singular if and only if $4a^3 + 27b^2 \neq 0$. From now on, we will write the curve sometimes as a projective curve but in some moments we will consider its intersection with one of the affine planes.
The following two propositions are our first approach to study isomorphisms of elliptic curves.

Proposition 3.1. *Let* $\phi : E(a', b') \to E(a, b)$ *be an isomorphism sending* $O = (0 : 1 : 0)$ *to* $O' = (0 : 1 : 0)$*, then there exists a* $c \in k^*$ *such that* $a' = c^4a, b' = c^6b$ *and* ϕ *is is the map* $(x : y : z) \mapsto (c^2x : c^3y : z)$*. Conversely, if* $a' = c^4a, b' = c^6b$ *for some* $c \in k^*$*, then* $(x : y : z) \mapsto (c^2x : c^3y : z)$ *is an isomorphsim sending* O *to* O'*.*

Proof. Consider the function $x \circ \phi$ on $E(a', b')$ that has a double pole at O', like x'; therefore $x \circ \phi = u_1x' + r$ $(u_1 \in k^*, r \in k)$. For the same reason, and using now that $y \circ \phi$ has order three and so is an affine combination of the base functions of less order, $y \circ \phi = u_2y' + sx' + t, u_2 \in k^*, s, t \in k$. Use now that the map $f \to f \circ \phi$ is a homomorphism between $k[x, y]$ and $k[x', y']$; since we know $Y^2 = X^3 + aX + b$, then the same equation holds for $x \circ \phi$ and $y \circ \phi$

$$(u_2y' + sx' + t)^2 = (u_1x' + r)^3 + a(u_1x' + r) + b$$

and any polynomial satisfied by x', y' should be a multiple of $Y^2 - X^3 - aX - b$. From here, we deduce that $u_2^2 = u_1^3, r = s = t = 0, a' = (u_2/u_1)^4a, b' = (u_2/u_1)^6b$, as we wanted. The reciprocal is obvious. □

Proposition 3.2. *Let* (E, O) *be an elliptic curve with a point* O *isomorphic to* $(E(a, b), O)$*. We define*

$$j(E) = \frac{1728(4a^3)}{4a^3 + 27b^2}$$

Then (j, E) *depends only on* (E, O) *and two elliptic curves* E, E' *are isomorphic over the algebraic closure of the field* k *if and only if* $j(E) = j'(E)$*.*

Proof. If two curves are isomorphic, there exists a c like in the preceding proposition, so the numerator and the denominators of $j(E)$ are multiplied by c^6 and give the same result. To proof the converse, suppose that $j(E) = j(E')$. A first trivial observation is that $a = 0$ if and only if $a' = 0$. Over k^{al}, two elliptic curves of the form $Y^2Z = X^3 + bZ^3$ are isomorphic (the isomorphisms here are of the form $b' = c^6 b$ and for the condition of being algebraically closed we can put any b'). Assume now that $a, a' \neq 0$ and replace now (a, b) with $(c^4 a, c^6 b)$, where $c = \sqrt[4]{a'/a}$. That way we already have $a = a'$ and from $j(E) = j(E')$, $b = \pm b$. A change like before where $c = \sqrt{-1}$ shows the final isomorphim between the two curves with opposite b. □

Note that the condition of being algebraically closed is necessary and two curves can have the same j-invariant without being isomorphic. Two curves like that, which become isomorphic over the algebraic closure, are called twists and will be studied in detail later.

A similar study of the isomorphisms between elliptic curves can be done in general, without assuming that the characteristic is different from 2 or 3, but the ideas are the same and we just get formulas and proofs that are more tedious.

3.2 Reduction of an elliptic curve modulo p

Let's consider now an elliptic curve

$$E : Y^2Z = X^3 + aXZ^2 + bZ^3, \quad a, b \in \mathbb{Q}, \quad \Delta = -16(4a^3 + 27b^2) \neq 0$$

After a change of variables $X' \mapsto X/c^2, Y' \mapsto Y/c^3$ we can assume that the coefficients are in $\mathbb{Z}$ so we can look at them modulo p to obtain a curve $\bar{E}$ over $\mathbb{F}_p$. But this curve is not necessary an elliptic curve, since the discriminant may vanish.

To introduce the topic, we will take a look to different irreducible algebraic curves over k (that is assumed to be perfect) having group structures defined by regular maps.

- Elliptic curves: they will result to be the only irreducible projective curves that have a group structure defined by polynomial maps.
- Clearly, we also have that the affine line $\mathbb{A}^1$ is a group under addition. We will write it as G_a.
- The same occurs for the affine line with the origin removed, which is a group under multiplication, G_m.
- We have the so called twisted multiplicative groups: to define them, take a a non-square in k^* and let $L = k[\sqrt{a}]$ the corresponding quadratic extension, and consider the elements of L^* of norm 1 (or equivalently, the affine plane curve $X^2 - aY^2 = 1$). The group operation is defined as

$$(x, y) \cdot (x', y') = (xx' + ayy', xy' + x'y)$$

We write for this group $G_m[a]$ and since it can be transformed by an invertible change of coordinates in $G_m[ac^2]$, it will only depend on the field $k(\sqrt{a})$. In particular $G_m[a]$ is isomorphic to G_m over $k(\sqrt{a})$. An important fact is that in $\mathbb{F}_q$, $G_m[a]$ has $q+1$ elements. This is because there exists an exact sequence where the map from $\mathbb{F}_{q^2}^*$ to $\mathbb{F}_q^*$ corresponding to taking norms is surjective, since a quadratic form in three variables always has a nontrivial zero over a finite field (Chevalley).

Lefschetz fixed point theorem

We are going to give an intuition of why the previous examples are the only ones (a clear and rigurous statement of this would lead too far). For our purposes, we will enunciate a theorem with reminiscences from algebraic topology, the Lefschetz fixed point theorem.

Theorem 3.1. *Let M be a compact oriented manifold, and let $\alpha : M \to M$ be a continuous function. Let Δ denotes the diagonal in $M \times M$ and Γ_α the graph of α ($\Delta \cdot \Gamma_\alpha$ is therefore the number of fixed points taking into account multiplicities).*

$$(\Delta \cdot \Gamma_\alpha) = \sum (-1)^i \operatorname{Tr}(\alpha | H^i(M, \mathbb{Q}))$$

Let $L(\alpha)$ be the integer on the right. Assume that M has a group structure; we take the translation map, that has no fixed points: $\tau_a = (x \mapsto x + a)$. But the map $a \mapsto L(\tau_a) : M \to \mathbb{Z}$ is continuous, then constant on each connected component. If we let a tend to 0, $L(\tau_0) = 0$ and we get

$$0 = L(\tau_0) = \sum \dim_{\mathbb{Q}} H^i(M, \mathbb{Q})$$

Consequently, if the manifold has group structure, its Euler characteristic is 0 so the genus is 1.
We will still have to prove that G_a and G_m are the only affine algebraic groups of dimension one over an algebraically closed field.

Singular cubic curves

Consider so an elliptic curve expressed in reduced form

$$E : Y^2 Z = X^3 + aXZ^2 + bZ^3, a, b \in \mathbb{Q} \text{ with } \Delta = -16(4a^3 + 27b^2) \neq 0$$

where (maybe performing a change of the form $X \mapsto X/c^2, Y \mapsto Y/c^3$) a, b are integers chosen in such a way that $|\Delta|$ is minimal. Consider now the reduction $\bar{E}$ of the curve to $\mathbb{F}_p$.
We will have three options for a generic curve: good reduction, cuspidal reduction and nodal reduction, depending on the value of $-2ab$. Since $(0 : 1 : 0)$ is always non-singular, we study the affine curve $Y^2 = X^3 + aX + b$ and we want to see if it is possible to find a t such that $Y^2 = (X - t)^2 (X + 2t) = X^3 - 3t^2 X + 2t^3$, i.e., $t^2 = -a/3, t^3 = b/2$, from where $t = -\frac{3b}{2a}$ (and in this case $\Delta = 0$ as

expected). Expressing $X + 2t = (X - t) + 3t$ we can write the affine curve as $Y^2 = (X-t)^3 + 3t(X-t)^2$, that is singular at $(t, 0)$ and it has a cusp when $3t = 0$, a node with rational tangents if $3t$ is a square in k^* and a node with non-rational tangents elsewhere.
Since $-2ab = -2(-3t^2)(2t^3) = (2t^2)^2(3t)$, the value of $-2ab$ will determine the type of singularity.

a) Good reduction: it occurs when $p \neq 2$ and p does not divide Δ. In that case, $\bar{E}$ is an elliptic curve in $\mathbb{F}_p$. Any point in E, $P = (x : y : z)$ can be reduced by taking a primitive representative (x, y, z) (integer coordinates without common factors), obtaining a point $(\bar{x} : \bar{y} : \bar{z})$. $(0 : 1 : 0)$ reduces to $(0 : 1 : 0)$ so the reduction map is a homeomorphism. In the next chapter we will see that the number of points in this case is $p + 1$ more or less a small factor $(2\sqrt{p})$.

b) Additive reduction (or cuspidal): when the reduced curve has a cusp. The name comes from the fact that the set of non-singular points (that still has a group structure) is isomorphic to $\mathbb{G}_a$. Note that the set of non-singular is a group since when we add up two of them, the result cannot be the singular point since, for having multiplicity strictly greater than 1, we would have a line cutting the elliptic curve at more that three points. When $p \neq 2, 3$, this case occurs when p divides Δ and at the same time divides $-2ab$.
We will justify the different affirmations we have made, about all the isomorphism with $\mathbb{G}_a$. We are considering a projective curve $E : Y^2Z = X^3$, that has a cusp in $P = (0 : 0 : 1)$ since the affine curve $y^2 = x^3$ has a cusp at $(0, 0)$ (two equal tangents). P is the only point of the curve whose Y coordinate is zero. This is a way to state that the set $E(k)\backslash\{S\}$ (i.e., our non-singular points) are the points of the curve $E \cap \{Y = 0\}$ $(F : Z = X^3)$. A generic line $Z = aX + b$ cuts F in three points, satisfying that $X^3 - aX - b$ (the quadratic term is 0, so using Cardano-Viéte formulae the sum of the x-coordinates is 0). It is natural now to consider the map $F \to k : P \mapsto x(P)$ that has the property that if $P_1 + P_2 + P_3 = 0$ (the points are collinear), then $x(P_1) + x(P_2) + x(P_3) = 0$. From here, we deduce that $x(-P) = -x(P)$ (we can also argue that since P, Q and $-(P+Q)$ add up to 0, then $x(P) + x(Q) = x(P+Q)$). We conclude from this that the map $P \mapsto x(P)$ is a homomorphism and also an isomorphism of algebraic varieties (so an isomorphism of algebraic groups). We finally have that the map $P \mapsto \frac{x(P)}{y(P)}$ is an isomorphism of algebraic group from the set of non-singular points to $\mathbb{G}_a$ (p points).

c) Multiplicative reduction (or nodal): when the reduced curve has a node. When $p \neq 2, 3$, this occurs when p divides Δ and does not divide $-2ab$. The tangents at the node can be rational over $\mathbb{F}_p$: the rational point (once we are in the affine chart corresponding to $Z = 1$) is $(-\frac{3b}{2a}, 0)$, so looking for the tangents in the usual way we obtain that they are the two lines corresponding to the quadratic part $y^2 + \frac{9b}{2a}x^2$, so we require $-2ab$ to be a square. In this case the set of non-singular points is isomorphic to $\mathbb{G}_m$ and E has split multiplicative reduction (and $p - 1$ points). Elsewhere, when $-2ab$ is not a square, $E^{\text{ns}} \simeq \mathbb{G}_m[-2ab]$

and E has split multiplicative reduction ($p+1$ points).
We have to justify again the stated isomorphisms, and we will follow the approach taken in Cassels' book. Considering the projective curve $Y^2Z = X^3 + cX^2Z$, where $c \neq 0$, we see that at $(0:0:1)$ it has a node because the corresponding affine curve is $y^2 - cx^2 - x^3 = 0$ and the tangents at $(0,0)$ are different. When $c \neq 0$, the tangent cone factors as $(Y - \sqrt{c}X)(Y + \sqrt{c}X) = 0$. We will say that the tangent lines are defined over k (or are rational over k). We shall now exhibit an isomorphism of E^{ns} and $\mathbb{G}_m$. We define $c = \gamma^2$ and $U = Y + \gamma X, V = Y - \gamma X$ so the curve is $8\gamma^3 UVZ = (U-V)^3$. A line that not passes through the origin can be written as $Z = aU + bV$ and it meets the curve when

$$(U-V)^3 - 8\gamma^3 UV(aU + bV) = 0$$

Since everything is homogeneous, we can assume that $V = 1$ and in that case the product of the U-th coordinates is 1; in general, it would be equal to the product of the V-th coordinates, resulting that $\frac{u_1}{v_1}\frac{u_2}{v_2}\frac{u_3}{v_3} = 1$. The same reasoning than above implies that this is a homomorphism of groups (three points that add up zero map to three points whose product is 1) and also an algebraic isomorphism.
We finish with the case when γ is not a square. We adjoin γ to the ground field, where $\gamma = c^2$. For a point (x, y, z) of the curve, write $\frac{y+\gamma x}{y-\gamma x} = r + s\gamma$, where $r^2 - cs^2 = 1$ (since the norm is multiplicative). So we have an isomorphism with the curve $R^2 - cS^2$, that has the twisted multiplication law explicated before.

Definition 3.3. *An elliptic curve is semistable if it has bad reduction only of multiplicative type.*

From here, it is interesting to keep in mind the number of points we have in the reduced elliptic curve: p for the case of additive reduction, $p-1$ for the split multiplicative and $p+1$ for the non-split multiplicative. In the next chapter, we will do the counting for the case of good reduction.

3.3 Elliptic curves over $\mathbb{Q}_p$

After having studied some basic facts of elliptic curves over $\mathbb{Q}$, we notice that we can replace $\mathbb{Q}$ or $\mathbb{Z}$ for $\mathbb{Q}_p$ and $\mathbb{Z}_p$ and in general for any finite extension of $\mathbb{Q}_p$ and its ring of integers. Recall that by Hensel's lemma the image of the reduction map $E(\mathbb{Q}_p) \to E(\mathbb{F}_p)$ includes every non.singular point (we can lift a solution in $\mathbb{F}_p$, in the same way that we did in the proof of Selmer example). Obviously, we can also replace $\mathbb{Q}$ with a number field, but it may happen that the ring of integers is not a PID, and in that case it can occur that there is not an equation for the elliptic curve that is minimal for all primes at the same time. We will focus our attention now in $\mathbb{Q}_p$.
Consider as usual our standard model for an elliptic curve

$$E : Y^2Z = X^3 + aXZ^2 + bZ^3, \quad a, b \in \mathbb{Q}_p, \quad 4a^3 + 27b^2 \neq 0$$

where with an admissible change of variables we can assume that $a, b \in \mathbb{Z}_p$. As we did in the previous section, we have a reduction map $E(\mathbb{Q}_p) \to \bar{E}(\mathbb{F}_p), P \mapsto \bar{P}$. Along these lines, we will define a particular filtration identifying the quotients. We begin by defining

$$E^0(\mathbb{Q}_p) = \{P \mid \bar{P} \text{ is nonsingular}\}$$

As we have already pointed out, this has group structure since O is always non singular and a line through two non-singular points will meet the curve again in a non-singular point, since elsewhere the intersection number of that point and the curve would be at least two, and the cubic will meet a line in more than three points. Define now $E^1(\mathbb{Q}_p)$ to be the kernel of the reduction map $E^0(\mathbb{Q}_p) \to \bar{E}^{\text{ns}}(\mathbb{F}_p), P \mapsto \bar{P}$, that is a homomorphism. $E^1(\mathbb{Q}_p)$ is the set of points whose x and z coordinates are 0 modulo p and y is not. In general, we can define

$$E^n(\mathbb{Q}_p) = \{P \in E^1(\mathbb{Q}_p) \mid \frac{x(P)}{y(P)} \in p^n\mathbb{Z}_p\}$$

Note that we have the filtration

$$E(\mathbb{Q}_p) \supset E^0(\mathbb{Q}_p) \supset \cdots \supset E^n(\mathbb{Q}_p) \supset \cdots$$

Their properties are summarized in the following proposition:

Proposition 3.3. *The previous filtration fulfills:*

a) $E(\mathbb{Q}_p)/E^0(\mathbb{Q}_p)$ *is finite.*

b) The reduction map defines an isomorphism $E^0(\mathbb{Q}_p)/E^1(\mathbb{Q}_p) \to \bar{E}^{\text{ns}}(\mathbb{F}_p)$

c) $E^n(\mathbb{Q}_p)$ *is a subgroup of* $E(\mathbb{Q}_p)$ *when* $n \geq 1$ *and the map* $P \mapsto p^{-n}\frac{x(P)}{y(P)} \mod p$ *defines an isomorphism of groups between* $\mathbb{F}_p$ *and* $E^n(\mathbb{Q}_p)/E^{n+1}(\mathbb{Q}_p)$.

d) The filtration is exhaustive ($\cap_n E^n(\mathbb{Q}_p) = \{0\}$).

Proof. For the first claim, we will analyze the topological aspects: in $\mathbb{Q}_p^3$ we put the product topology and then in $\mathbb{P}^2(\mathbb{Q}_p)$ the corresponding quotient topology; note that the equivalence relation that defines $\mathbb{P}^2(\mathbb{Q}_p)$ allows us to clean denominators and consider the coordinates in $\mathbb{Z}_p$ (one of them, at least, being a unity, since elsewhere we divide by p). Therefore, $\mathbb{P}^2(\mathbb{Q}_p)$ is the union of the sets $\mathbb{Z}_p^* \times \mathbb{Z}_p \times \mathbb{Z}_p$, $\mathbb{Z}_p \times \mathbb{Z}_p^* \times \mathbb{Z}_p$, $\mathbb{Z}_p \times \mathbb{Z}_p \times \mathbb{Z}_p^*$. Each of them is compact and open (just write the corresponding definitions), so $\mathbb{P}^2(\mathbb{Q}_p)$ is compact and $E(\mathbb{Q}_p)$ is closed since it is the set of zeros of a polynomial there. We have a closed set inside a compact, so it is also compact. Note now that $E^0(\mathbb{Q}_p)$ is an open subgroup (we already have seen that it is a subgroup and it is open because we are considering a topology where a point is close to another when they have the same reduction). To sum up, $E(\mathbb{Q}_p)$ is compact, and it is the union of the cosets of its subgroup $E^0(\mathbb{Q}_p)$, and to stablish the compactness, note that there is only a finite number.

The second claim is the first isomorphism theorem applied to the map

$$E^0(\mathbb{Q}_p) \to \bar{E}^{\text{ns}}(\mathbb{F}_p)$$

already defined, where the surjectivity is given by Hensel.

For the third claim, we proceed by induction, assuming that $E^n(\mathbb{Q}_p)$ is a subgroup of $E(\mathbb{Q}_p)$. Let $P = (x : y : 1)$ a point of $E^1(\mathbb{Q}_p)$; note that $y \notin \mathbb{Z}_p$ since as we commented before in $\mathbb{E}^1(\mathbb{Z}_p)$, x and z are divisible by p and y is not, i.e., the multiplicity of p in z is greater than the multiplicity in y, so now $y \notin \mathbb{Z}_p$. Write $x = p^{-m}x_0, y = p^{-m'}y_0$ where x_0, y_0 are units. Putting this in the equation results

$$p^{-2m'} = p^{-3m}x_0^3 + ap^{-m} + b$$

and since a, b were chosen to be in $\mathbb{Z}_p$ we can take orders and conclude that $2m' = 3m$. From this, it follows that there is a positive integer such that $m = 2n, m' = 3n$. Furthermore $n = m' - m$. These considerations allow us to conclude that if $P = (x : y : z)$ is in $E^n(\mathbb{Q}_p)$ but not in $E^{n+1}(\mathbb{Q}_p)$ then $P = (p^n x_0, y_0, p^{3n}z_0)$, where x_0, y_0, z_0 are units of $\mathbb{Z}_p$. A general point of $E^n(\mathbb{Q}^p)$ has the same form, but we only require y_0 to be a unit.
Substituting again in the equation, dividing by p^{3n} and reducing modulo p we get that $P_0 = (\bar{x}_0 : \bar{y}_0 : \bar{z}_0)$ verifies $E_0 : Y^2Z = X^3$. This is a homomorphism between $E^n(\mathbb{Q}_p) \to E_0(\mathbb{F}_p)$ where nobody goes to the singular point (y_0 is a unit) and whose kernel will be $E^{n+1}(\mathbb{Q}_p)$ (the points such that x_0 is a multiple of p are those in $E^{n+1}(\mathbb{Q}_p)$). From Hensel again, we see that any non-singular point of $E_0(\mathbb{F}_p)$ is in the image, and since we already now that $Q \mapsto \frac{x(Q)}{y(Q)}$ is an isomorphim, the composition of the two is also an isomorphism.

For the last item, just consider a point in the intersection, so $x(P) = 0$ and $y(P) \neq 0$, which implies that either $z(P) = 0$ or $y(P)^2 = bz(P)^3$; the second option is not possible, since P would not be in $E^1(\mathbb{Q}_p)$, so $z(P) = 0$ and $P = (0 : 1 : 0)$. □

Corollary 3.1. *Let m be an integer not divisible by p. Then,*

$$E^1(\mathbb{Q}_p) \to E^1(\mathbb{Q}_p), P \mapsto mP$$

is a bijection.

Proof. The injectivity is due to the fact that $E^n(\mathbb{Q}_p)/E^{n+1}(\mathbb{Q}_p) \simeq \mathbb{Z}/p\mathbb{Z}$; if $P \in E^n(\mathbb{Q}_p)\backslash E^{n+1}(\mathbb{Q}_p)$ and at the same time $mP = 0$ (assuming that of course $P \neq 0$), then the image of P in $\mathbb{Z}/p\mathbb{Z}$ is not zero but the image of mP is 0, and that is a contradiction.
For the surjectivity, take $P \in E^1(\mathbb{Q}_p)$. Multiplication by m is an isomorphism in $E^1(\mathbb{Q}_p)/E^2(\mathbb{Q}_p)$, so there exists $Q_1 \in E^1(\mathbb{Q}_p)$ such that $P = mQ_1$ modulo $E^2(\mathbb{Q}_p)$. Iterating the process, we will have a point $Q_2 \in E^2(\mathbb{Q}_p)$ such that $P - mP_1 = mQ_2$ modulo $E^3(\mathbb{Q}_p)$ and in general we have a sequence of points

$\{Q_i\}$, where $Q_i \in E^i(\mathbb{Q}_p)$ and such that $P - m\sum_{i=1}^n Q_i \in E^{n+1}(\mathbb{Q}_p)$. Bearing in mind that $E(\mathbb{Q}_p)$ is compact, we have that $\sum Q_i$ converges to a point in $E(\mathbb{Q}_p)$ and from what we said the limit verifies $P = mQ$. □

3.4 Torsion points

In Birch and Swinertonn-Dyer conjecture, we are concerned about the rank of an elliptic curve, for which relatively little is known. For instance, there is not an easy way to compute it. However, the torsion part admits an easy treatment, and the following results will clearly establish an algorithmic procedure to determine it. For this section, we consider again an elliptic curve E given by

$$E : Y^2Z = X^3 + aXZ^2 + bZ^3 a, b \in \mathbb{Z}, \Delta = 4a^3 + 27b^2 \neq 0$$

We will denote as $E(\mathbb{Q})_{\text{tors}}$ the torsion group of $E(\mathbb{Q})$.

Theorem 3.2. (Lutz-Nagell) *If $P = (x : y : 1) \in E(\mathbb{Q})_{tors}$, then $x, y \in \mathbb{Z}$ and $y = 0$ or $y|\Delta$*

We will prove first two proposition and the theorem will directly follow.

Proposition 3.4. *Let $P = (x : y : 1) \in E(\mathbb{Q})$. If P and $2P$ have integer coordinates, then either $y_1 = 0$ or $y_1|\Delta$.*

Proof. We assume that $y_1 \neq 0$. To find $-2P = (x_2 : y_2 : 1)$, we do the intersection of the tangent at P with the affine curve $Y^2 = X^3 + aX + b$. For the formula to get the inverse of a point, we know that if Q has integer coordinates, also $-Q$ has. Let $Y = \alpha X + \beta$ be the tangent line at P; the X-coordinates of the intersection satisfy

$$0 = (\alpha X + \beta)^2 - X^3 - aX - b$$

The term in X^2 is α^2, that when divided by the leading term remains $-\alpha^2$. Therefore, the X-coordinates of the intersection verify $x_0 + x_1 + x_2 = \alpha^2$, so α^2 is an integer. We have now that $\frac{f'(x_1)}{2y_1}$ is an integer, where $f(X) = X^3 + aX + b$. We already knew from $y_1^2 = f(x_1)$ that $y_1|f(x_1)$, and we now have that $y_1|f'(x_1)$. For the properties of the resultant, there exist $r(X), s(X) \in \mathbb{Z}[X]$ such that

$$\Delta = r(X)f(X) + s(X)f'(X)$$

and so $y_1|\Delta$. □

Proposition 3.5. *The group $E^1(\mathbb{Q}_p)$ is torsion-free*

Proof. Since we already now that multiplication by an integer coprime with p is a bijection, we are just concerned with pP. In particular, we want to show that $E^1(\mathbb{Q})$ does not contain any point such that $pP = 0$. When $P \in E^1(\mathbb{Q}_p)$, $y(P) \neq 0$, so we are considering the intersection with the affine plane corresponding to $y \neq 0$, $E_1 : Z = X^3 + axZ^2 + bZ^3$. The new coordinates of a point $P = (x : y : z)$ are $x'(P) = \frac{x(P)}{y(P)}, z'(P) = \frac{z(P)}{y(P)}$. With this map, the property that three points are collinear if and only if they add up to 0 still holds. Now, we can reformulate the definition of $E^n(\mathbb{Q}_p)$ as the points $P \in E^1(\mathbb{Q}_p)$ such that $x'(P) \in p^n\mathbb{Z}_p$. A tedious algebraic manipulation (that we omit) shows the following:

Lemma 3.1. *Let $P_1, P_2, P_3 \in E(\mathbb{Q}_p)$ collinear points. If $P_1, P_2 \in E^n(\mathbb{Q}_p)$, then $P_3 \in (\mathbb{Q}_p)$ and $x'(P_1) + x'(P_2) + x'(P_3) \in p^{5n}\mathbb{Z}_p$.*

We define now a projection map given by

$$E^n(\mathbb{Q}_p) \to p^n\mathbb{Z}_p/p^{5n}\mathbb{Z}_p : P \mapsto \bar{x}(P)$$

It is a homomorphism of abelian groups because $\bar{x}(-P) = -\bar{x}(P)$ and $P_1 + P_2 + P_3 = 0$ implies $\bar{x}(P_1) + \bar{x}(P_2) + \bar{x}(P_3) = 0$ (this suffices, just taking now $0 = \bar{x}(P+Q) + \bar{x}(-P) + \bar{x}(-Q) = \bar{x}(P+Q) - \bar{x}(P) - \bar{x}(Q)$). If now $P \in E^1(\mathbb{Q}_p)$ has order p, it lies in $E^n(\mathbb{Q}_p)\backslash E^{n+1}(\mathbb{Q}_p)$ for some n and so $\bar{x}(P)$ is in $p^n\mathbb{Z}_p\backslash p^{n+1}\mathbb{Z}_p$, and so $\bar{x}(pP) \in p^{n+1}\mathbb{Z}_p\backslash p^{n+2}\mathbb{Z}_p$, contradicting the fact that $pP = 0$. □

We are now about to finish the proof of the main theorem. We state the following result as a proposition but it is quite direct using the previous result:

Proposition 3.6. *Let $P = (x : y : 1) \in E(\mathbb{Q}_p)_{\text{tors}}$, then $x, y \in \mathbb{Z}_p$*

Proof. We write P with primitive coordinates $(\bar{x} : \bar{y} : \bar{z})$ (all in $\mathbb{Z}_p$ but not all three in $p\mathbb{Z}_p$). If $P = (x : y : 1)$ with x or y not in $\mathbb{Z}_p$, any primitive coordinates will have $\bar{z} \in p\mathbb{Z}_p$. But therefore reducing mod p, the z-coordinate will be 0 and so the point will be $\bar{P} = (0 : 1 : 0)$ that is $E^1(\mathbb{Q}_p)$. Thus, if $P = (x : y : 1) \notin E^1(\mathbb{Q}_p)$, then $x, y \in \mathbb{Z}_p$. □

Corollary 3.2. *If $P = (x : y : 1) \in E(\mathbb{Q})_{\text{tors}}$, then $x, y \in \mathbb{Z}$.*

Corollary 3.3. *If E has good reduction at p, then the reduction map*

$$E(\mathbb{Q})_{\text{tors}} \to \bar{E}(\mathbb{F}_p)$$

is injective.

We give now some examples of how to calculate the torsion. For instance, let

$$E_1 : Y^2 = X^3 + 1$$

Its discriminant is 27. We see that the candidates that work as torsion points are O, $(-1, 0)$ (unique point of order 2), $(0, 1), (0, -1)$ (order 3) and $(2, 3), (2, -3)$ (order 6). We have so that the torsion group is isomorphic to $\mathbb{Z}/6\mathbb{Z}$. We give another example

$$E_2 : Y^2 = X^3 - X$$

Here, the only possibilities (that work) apart from O are $(1, 0), (-1, 0), (0, 0)$ that all have order 2; the torsion group is therefore isomorphic to $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$. Mazur proved in the seventies that $E(\mathbb{Q})$ is isomorphic to one of the groups

$$\mathbb{Z}/m\mathbb{Z} \text{ for } m = 1, 2, \cdots, 10, 12$$

$$\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/m\mathbb{Z} \text{ for } m = 2, 4, 6, 8$$

It would be nice to give a proof of this result, but we do not have time to develop of the algebraic geometry required to prove it in this essay. In a number field this is much more difficult. The main result is due to Loic Merel:

Theorem 3.3. *For all $d \in \mathbb{Z}, d \geq 1$ there exists a constant $B(d) \geq 0$ such that for all elliptic curves E over a number field K with $[K : \mathbb{Q}] = d$, then*

$$|E(K)_{tors}| \leq B(d)$$

3.5 The invariant differential

We begin by recalling some concepts from algebraic geometry. Differentials are an important topic in this area, allowing the use of some classical geometric arguments in the context of varieties over any field. They can be used to define the genus of a curve and to analyze the ramification of morphisms between curves. The following definitions mimic those given in any undergraduate course in differential manifolds:

Definition 3.4. *Let X be a nonsingular variety, and $U \subset X$ an open subset. A differential form on U associates to each point $P \in U$ an element of the Zariski cotangent space $T^*_P(X)$.*

As it can be seen, they play a role analogous to that of arbitrary functions, and we need to restrict to a much smaller collection of them to obtain a useful concept:

Definition 3.5. *Given $U \subset X$ an open subset of a nonsingular variety, and $f \in O(U)$, the differential form df associated to f is defined as follows: for $P \in U$, let $df(P) \in m_P/m_P^2$ be the equivalence class of $(U, f - f(P))$. A differential form ω on U is regular if for every $P \in U$, there exists an open neighborhood $V \in U$ of P and regular functions $f_1, \dots, f_m, g_1, \dots, g_m \in O(V)$ such that $\omega|V = \sum_i f_i dg_i$.*

In this section, we consider the following Weierstrass equation for our elliptic curve, writing in non homogeneous coordinates as

$$y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6$$

where the $a_i \in K$.

Definition 3.6. *The invariant differential associated to an elliptic curve (also referred as Néron diferential) is*

$$\omega = \frac{dx}{2y + a_1x + a_3} = \frac{dy}{3x^2 + 2a_2x + a_4x - a_1y}$$

The name of invariant will be soon explained. First, we prove that it is holomorphic and non-vanishing.

Proposition 3.7. *The invariant differential associated to the Weierstrass equation of an elliptic curve is regular and non-vanishing, i.e.,* $\operatorname{div}(\omega) = 0$.

Proof. Take $P = (x_0, y_0) \in E$ and let $E : F(x, y) = y^2 + a_xy + a_3y - x^3 - a_2x^2 - a_4x - a_6 = 0$ in such a way that

$$w = \frac{d(x - x_0)}{F_y(x, y)} = \frac{-d(y - y_0)}{F_x(x, y)}$$

Clearly, if P were a pole, $F_x(x, y) = 0, F_y(x, y) = 0$ and this cannot be true for our definition of elliptic curve (it cannot have singular points). To see that P cannot be a zero, define the map $E \to \mathbb{P}^1 : (x : y : 1) \mapsto (x : 1)$; the map has degree 2, so $\mathrm{ord}_P(x-x_0) \leq 2$ and equality holds if and only if the polynomial in one single variable $F(x_0, y)$ has a double root and in that case $F_y(x,y)$ would have a single zero at P. In any case $\mathrm{ord}_P(\omega) = \mathrm{ord}_P(x-x_0) - \mathrm{ord}_P(F_y(x,y)) - 1 = 0$. This is enough for concluding that no affine point will be a zero or a pole, but we have to check what happens at infinity. We consider an uniformizer t and since the order of O at x is -2 and at y is -3, then $x = t^{-2}f, y = t^{-3}g$ for functions f and g that have neither zeros nor poles at O. Making this substitutions in the definition of ω, we get $\omega = \frac{-2f+tf'}{2g+a_1tf+a_3t^3}dt$. We see therefore that ω also behaves well at O, assuming that the characteristic of the field is not 2. In that case, it remains true but we should perform more careful changes of variables. □

We will explain now the reason for the name of invariant, basically that it does not change under translations.

Proposition 3.8. *Let Q be a point in $E(K)$, and let τ_Q be the translation by Q map. Then, $\tau_Q^*\omega = \omega$.*

Proof. Of course there is a straightforward way to do this. Use the addition formulas and compute $x(P+Q), y(P+Q)$ in terms of $x(P), y(P), x(Q), y(Q)$, and then check that $\frac{dx(P+Q)}{2y(P+Q)+a_1x(P+Q)+a_3}$ remains invariant. Alternatively, we can use a more elegant approach using properties of divisors: let Ω_E the space or holomorphic 1-forms, that has dimension one over $\bar{K}(E)$. We have so a function $a_Q \in K(E)^*$ depending on the point and such that $\tau_Q^*\omega = a_Q\omega$. But since $\mathrm{div}(\omega) = 0$, we can conclude that $\mathrm{div}(a_Q) = 0$, so it is constant.
We now consider the map $f : E \to \mathbb{P}^1$ that sends Q to $(a_Q : 1)$. It is clear that if we compute explicitly a_Q, it would be a rational function of $x(Q)$ and $y(Q)$, so f is a rational map that is not surjective (since it omits at least $(0,1)$ and $(1,0)$). Consequently f is constant and evaluating at O we find that its value is 1. We conclude that $\tau_Q^*\omega = \omega$. □

3.6 Isogenies

Definition 3.7. *Let E_1 and E_2 be elliptic curves. An isogeny from E_1 to E_2 is a non-zero morpshim of curves $\phi : E_1 \to E_2$ such that $\phi(O) = O$. Two elliptic curves E_1 and E_2 are isogenous if there is an isogeny from E_1 to E_2.*

The fact of excluding the zero morphism is a convention to preclude the possibility that any two elliptic curves were isogenous. Because of that, we will say that $\mathrm{Hom}(E_1, E_2)$ are the morphisms of curves sending 0 to 0 (or the set of isogenies together with the zero-map). In fact, some authors do not exclude this possibility. The most natural example is the multiplication-by-m isogeny, that we will denote as $[m]$.

Proposition 3.9. *Let E/K be an elliptic curve and let $m \in \mathbb{Z}, m \neq 0$.*

a) The multiplication-by-m map $[m]$ is not constant.

b) Let E_1 and E_2 be elliptic curves. Then the group $\mathrm{Hom}(E_1, E_2)$ is a torsion free $\mathbb{Z}$-module.

c) *Let E be an elliptic curve. Then the endomorphism ring* $\operatorname{End}(E)$ *is a ring of characteristic* 0 *that is an integral domain.*

Proof. For the first item, recall that $[2] \neq [0]$. For that, use the formula that gives you the x-coordinate of $2P$ and put it equal to 0 (we have not provided in any moment explicit formulas for the double of a point but there are relatively straightforward to obtain): $4x^3 + b_2x^2 + 2b_4x + b_6 = 0$. When the characteristic is not two, this shows that there are only finitely many solutions. In the case of characteristic two, we need $b_2 = b_6 = 0$, and this forces $\Delta = 0$. Now, taking into account that $[mn] = [m] \circ [n]$ we just have to deal with the case m odd. This requires some tedious computations that we only sketch: the first observation is that (in characteristic different from two) $x^4 - b_4x^2 - 2b_6x - b_8$ is not a multiple of $4x^3 + b_2x^2 + 2b_4x + b_6$ (since $\Delta \neq 0$). Therefore, there exists x_0 such that the second polynomial vanishes at higher order than does the first. We choose one of the possible $y_0 \in \bar{K}$ so that $P = (x_0, y_0) \in E$ and we have that $[2]P = O$. This proves that E has a nontrivial point of order 2 and so $[m]P_0 = P_0 \neq O$. In characteristic two, we need the triplication formula and everything is the same but with longer computations.
The second statement is direct: if $\phi \in \operatorname{Hom}(E_1, E_2)$ and $m \in \mathbb{Z}$ satisfy $[m] \circ \phi = [0]$ we can use the multiplicativity of the degrees to conclude that either $m = 0$ or $\phi = [0]$.
Once we know that the endomorphism ring has characteristic 0, take ϕ, ψ such that $\phi \circ \psi = [0]$. Then, using again the multiplicativity of the degrees, either $\phi = [0]$ or $\psi = [0]$. □

An important result is that although we have defined an isogeny as a morphism of curves, it behaves well a morphism of groups. Note that in the definition we only impose that 0 is sent to 0 but we can prove the following result:

Theorem 3.4. *Let $\phi : E_1 \to E_2$ be an isogeny. Then, for all $P, Q \in E_1$,*

$$\phi(P + Q) = \phi(P) + \phi(Q)$$

Proof. When we have the 0-map there is nothing to prove. Elsewhere, ϕ is a finite map so it induces a homomorphism $\phi_* : \operatorname{Pic}^0(E_1) \to \operatorname{Pic}^0(E_2)$ sending the class of $\sum n_i(P_i)$ to the class of $\sum n_i(\phi P_i)$. Furthermore, we have a group isomorphisms $k_i : E_i \to \operatorname{Pic}^0(E_i)$ that sends P to the class of $(P) - (O)$ (for $i = 1, 2$). Recall that

$$\phi = k_2^{-1} \circ \phi_* \circ k_1$$

and since all k_1, k_2^{-1}, ϕ_* are group homomorphisms, so is ϕ. □

Corollary 3.4. *Let $\phi : E_1 \to E_2$ be an isogeny. Then,* $\ker \phi$ *is a finite group.*

We state now one of the main theorems, that says that the number of points in the kernel coincides with the degree when the map is separable. We do not provide the proof, just say that everything reduces to take into account that for a morphism between curves ϕ, $\phi^{-1}(Q)$ is the degree of separability of ϕ but for a finite number of points.

Theorem 3.5. *Let $\phi : E_1 \to E_2$ be an isogeny.*

a) Let $Q \in E_2$. Then, $\#\phi^{-1}(Q) = \deg_s \phi$. Furthermore, if $P \in E_1$, $e_\phi(P) = \deg_i \phi$.

b) The map $\ker \phi \to \operatorname{Aut}(\bar{K}(E_1)/\phi^* \bar{K}(E_2)) : T \mapsto \tau_T^*$ is an isomorphism.

c) Suppose that ϕ is separable. Then, $\# \ker \phi = \deg \phi$ and $\bar{K}(E_1)$ is a Galois extension or $\phi^* \bar{K}(E_2)$.

An important result (that will be the key for proving Hasse's theorem), is the following:

Proposition 3.10. *Let E be an elliptic curve over a finite field $\mathbb{F}_q$ of characteristic p and let ϕ the q-th power Frobenius morphism. Let $m, n \in \mathbb{Z}$. Then $m + n\phi$ is separable if and only if p does not divide m (and $1 - \phi$ is therefore separable).*

Proof. Let ω be an invariant differential on E. We know that the map $\psi : E \to E$ is inseparable if and only if $\psi^* \omega = 0$. Applying this to $m + n\phi$, we get that $(m + n\phi)^* \omega = m\omega + n\phi^* \omega$.

$$\phi^* \Big(\frac{dx}{2y + a_1 x + a_3} \Big) = \frac{d(x^q)}{2y^q + a_1 x^q + a_3} = \frac{q x^{q-1}}{2y^q + a_1 x^q + a_3} = 0$$

Hence, $(m + n\phi)^* \omega = m\omega$, and this last one expression is 0 if and only if $p|m$. □

The dual isogeny

When we have an isogeny ϕ from E_1 to E_2 this induces a map $\phi^* : \operatorname{Pic}^0(E_2) \to \operatorname{Pic}^0(E_1)$. Furthermore, we also have group isomorphisms $k_i : E_i \to \operatorname{Pic}^0(E_i)$ sending P to the class of $(P) - (O)$. This allows us to construct a homomorphism in the opposite direction:

$$E_2 \to \operatorname{Pic}^0(E_2) \to \operatorname{Pic}^0(E_1) \to E_1$$

It is basically $k_1^{-1} \circ \phi^* \circ k_2(Q)$ and we will see that if $Q \in E_2$, $P \in E_1$ satisfy $\phi(P) = Q$ then $k_1^{-1} \circ \phi^* \circ k_2(Q) = [\deg \phi](P)$. But previously we have to check several things: when we take a preimage P of Q we are taking roots of polynomials, so we have to check that applying $[\deg \phi]$ to P make the roots to appear symmetrically when ϕ is separable. The non-separable case will be more technical.

Theorem 3.6. *Let $E_1 \to E_2$ be a non-constant isogeny of degree m.*

a) There exists a unique isogeny $\hat{\phi} : E_2 \to E_1$ such that $\hat{\phi} \circ \phi = [m]$.

b) As a group homomorphism, $\hat{\phi}$ is the composition

$$E_2 \to \operatorname{Div}^0(E_2) \to \operatorname{Div}^0(E_1) \to E_1$$

where we go from $\operatorname{Div}^0(E_2)$ to $\operatorname{Div}^0(E_1)$ through ϕ^*.

Proof. Uniqueness is clear: in case of having $\hat{\phi}$ and $\hat{\phi}'$ with the desired properties, then $(\hat{\phi} - \hat{\phi}') \circ \phi = [0]$. ϕ is surjective (since it is non-constant), so necessarily $\hat{\phi} - \hat{\phi}'$ must be constant, and therefore $\hat{\phi} = \hat{\phi}'$.
Note also that if we have $\psi : E_2 \to E_3$ another non-constant isogeny, now of degree n, and we assume the existence of $\hat{\phi}$ and $\hat{\psi}$, then $(\hat{\phi} \circ \hat{\psi}) \circ (\psi \circ \phi) = \hat{\phi} \circ [n] \circ \phi = [n] \circ \hat{\phi} \circ \phi = [mn]$. We have proven that $\hat{\phi} \circ \hat{\psi} = \widehat{\psi \circ \phi}$. We recall here a very important lemma in the theory of algebraic curves:

Lemma 3.2. *Every map $\psi : C_1 \to C_2$ of smooth curves over a field of characteristic $p > 0$ factors as $C_1 \to C_1^{(q)} \to C_2$, where $q = \deg_i(\psi)$, the first map is the q-th power Frobenius map and the second one is a separable map.*

From this lemma, and for the property of compositions, it is enough with proving the existence of the isogeny when ϕ is separable and when ϕ is the Frobenius morphism.
For the first case, note that $\# \ker \phi = m$, so every element of the kernel has order a divisor of m and $\ker[\phi] \subset [m]$. To prove this we need another lemma:

Lemma 3.3. *Let $\phi : E_1 \to E_2$ and $\psi : E_1 \to E_3$ be non-constant isogenies such that ϕ is separable and $\ker \phi \subset \ker \psi$. Under these hypothesis, there is a unique isogeny $\lambda : E_2 \to E_3$ such that $\psi = \lambda \circ \phi$.*

The idea for the proof of the lemma is that the separability condition gave us the inclusions

$$\psi^* \bar{K}(E_3) \subset \phi^* \bar{K}(E_2) \subset \bar{K}(E_1)$$

Now, from our general theory of algebraic curves, we have a map $\lambda : E_2 \to E_3$ such that $\phi^*(\lambda^* \bar{K}(E_3)) = \psi^* \bar{K}(E_3)$ and so $\lambda \circ \phi = \psi$. It is trivial to check that O goes to O, so it is an isogeny.
This finishes the proof for the separable case. For the case of the Frobenius, due to the composition property, it is enough to do the proof when $\deg \phi = p$. Look at the multiplication-by-p map. If ω is an invariant differential, then $[p]^* \omega = p\omega = 0$. We can affirm now that $[p]$ is not separable and can be decomposed as a Frobenius morphism followed by a separable map: $[p] = \psi \circ \phi^e$. We take now $\hat{\phi} = \psi \circ \phi^{e-1}$ and this works. □

We state now some of the most remarkable properties of the dual isogeny:

Theorem 3.7. *Let $\phi : E_1 \to E_2$ be an isogeny and $\hat{\phi}$ the corresponding dual isogeny.*

a) Let $m = \deg \phi$. Then, $\hat{\phi} \circ \phi = [m]$ and $\phi \circ \hat{\phi} = [m]$.

b) Let $\lambda : E_2 \to E_3$ be another isogeny. Then, $\widehat{\lambda \circ \phi} = \hat{\phi} \circ \hat{\lambda}$.

c) Let $\psi : E_1 \to E_2$ be another isogeny. Then $\widehat{\phi + \psi} = \hat{\phi} + \hat{\psi}$.

d) The dual isogeny of the multiplication-by-m map is again the multiplication by m-map and its degree is m^2.

e) $\deg \hat{\phi} = \deg \phi$.

f) $\hat{\hat{\phi}} = \phi$.

We recall now that in an abelian group A we can define the concept of quadratic form $d : A \to \mathbb{R}$ as a function satisfying that $d(\alpha) = d(-\alpha)$ for all α and that the pairing $A \times A \to \mathbb{R}$ sending $(\alpha, \beta) \mapsto d(\alpha + \beta) - d(\alpha) - d(\beta)$ is bilinear.
Using the properties of the dual isogeny, we can prove that the degree map is a quadratic form (positive definite):

Proposition 3.11. *Let E_1, E_2 be elliptic curves. The degree map*

$$\deg : \mathrm{Hom}(E_1, E_2) \to \mathbb{Z}$$

is a positive definite quadratic form.

Proof. The unique part that is not clear is that $\langle \phi, \psi \rangle$ is bilinear. For that, use that there is an injection from $\mathbb{Z}$ to $\mathrm{End}(E_1)$ (multiplication-by-m) and so

$$[\langle \phi, \psi \rangle] = [\deg(\phi + \psi)] - [\deg(\phi)] - [\deg(\psi)] = \hat{\phi} \circ \psi + \hat{\psi} \circ \phi$$

Using now that the dual isogeny is linear (the third item of the previous theorem), the result follows. □

We finish this section with an important fact about the torsion of an elliptic curve, that is a direct consequence of the properties of the Frobenius map.

Proposition 3.12. *Let E be an elliptic curve and let $m \in \mathbb{Z}$ with $m \neq 0$. Then,*

a) $\deg[m] = m^2$.

b) If $m \neq 0$ in K, then

$$E[m] = \mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/m\mathbb{Z}$$

c) If $\mathrm{char}(K) = p > 0$, then either $E[p^e] = \{O\}$ for all $e > 0$ or $E[p^e] = \mathbb{Z}/p^e\mathbb{Z}$ for all $e > 0$.

3.7 The Tate Module

Consider as usual an elliptic curve and an integer $m \geq 2$ that is coprime with the characteristic of the field when this is not zero. We already know that $E[m] \cong \mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/m\mathbb{Z}$. But this group isomorphism can be viewed as something deeper, since $E[m]$ has more structure. For instance, an element of the absolute Galois group σ acts on $E[m]$ since $[m](P^\sigma) = ([m]P)^\sigma = O^\sigma = O$. That way, we have obtained a representation

$$G_{\bar{K}/K} \to \mathrm{Aut}(E[m]) \cong \mathrm{GL}_2(\mathbb{Z}/m\mathbb{Z})$$

But this is not very interesting, since we tend to prefer representations of groups in a ring of characteristic 0. We copy the idea of the construction of the p-adic integers $\mathbb{Z}_p$ as a projective limit, and now we define the l-adic Tate module of E:

Definition 3.8. *Let E be an elliptic curve and let $l \in \mathbb{Z}$ be a prime. The l-adic Tate module of E is the group $T_l(E) = \varprojlim E[l^n]$, where we have the natural maps to pass from $E[l^{n+1}]$ to $E[l^n]$ consisting on multiplication by l.*

Note that $E[l^n]$ is a $\mathbb{Z}/l^n\mathbb{Z}$ module, so the Tate module has a structure of $\mathbb{Z}_l$-module. The next proposition is a direct corollary of the previous section:

Proposition 3.13. *As a $\mathbb{Z}_l$-module, the Tate module has the following structure:*

a) When $l \neq \operatorname{char}(K)$, then $T_l(E) = \mathbb{Z}_l \times \mathbb{Z}_l$.

b) When $p = \operatorname{char}(K) > 0$, then $T_p(E) \cong \{0\}$ or $\mathbb{Z}_p$.

Definition 3.9. *The l-adic representation of $G_{\bar{K}/K}$ associated to E is the homomorphism*

$$\rho_l : G_{\bar{K}/K} \to \operatorname{Aut}(T_l(E))$$

induced by the action of the absolute Galois group on the l^n-torsion points of E.

A very important fact is that there is a natural way in which the m-torsion subgroup $E[m]$ can be identified with the homology group $H_1(E, \mathbb{Z}/m\mathbb{Z})$ and similarly $T_l(E)$ with $H_1(E, \mathbb{Z}_l)$ While the homology does not admit a Galois action (generally), the torsion subgroup and the Tate module do admit such action. This idea has been generalized in the theory of etale cohomology.

When we have an isogeny of elliptic curves $\phi : E_1 \to E_2$, it induces a map between the l^n-torsion points, and therefore a $\mathbb{Z}_l$-linear map: $\phi_l : T_l(E_1) \to T_l(E_2)$. We thus obtain a natural homomorphism

$$\operatorname{Hom}(E_1, E_2) \to \operatorname{Hom}(T_l(E_1), T_l(E_2))$$

that when $E_1 = E_2$ is a homomorphism of rings. This map is injective, and furthermore we have the following result:

Theorem 3.8. *Let E_1, E_2 be elliptic curves, and let $l \neq \operatorname{char}(K)$ be a prime. Then, the natural map*

$$\operatorname{Hom}(E_1, E_2) \otimes \mathbb{Z}_l \to \operatorname{Hom}(T_l(E_1), T_l(E_2))$$

sending ϕ to ϕ_l is injective.

We omit the proof, and content ourselves with give a very important corollary:

Corollary 3.5. *Let E_1, E_2 be elliptic curves. Then,* $\operatorname{Hom}(E_1, E_2)$ *is a free $\mathbb{Z}$-module of rank at most* 4.

Proof. We already know that $\operatorname{Hom}(E_1, E_2)$ is torsion free, and from the previous injectivity

$$\operatorname{rk}_{\mathbb{Z}_l} \operatorname{Hom}(E_1, E_2) \otimes \mathbb{Z}_l \leq \operatorname{rk}_{\mathbb{Z}_l} \operatorname{Hom}(T_l(E_1), T_l(E_2))$$

Choosing now a $\mathbb{Z}_l$-basis for $T_l(E_1)$ and $T_l(E_2)$ we see that $\operatorname{Hom}(T_l(E_1), T_l(E_2)) = M_2(\mathbb{Z}_l)$, whose rank is four. □

A natural question now is when $\operatorname{Hom}_K(E_1, E_2) \otimes \mathbb{Z}_l \to \operatorname{Hom}_K(T_l(E_1), T_l(E_2))$ is an isomorphism (since we already know that is injective). The answer is that this happens if and only if K is a finite field or a number field. The proof of this result is beyond the scope of this thesis.

3.8 The Weil pairing

In this section, we are going to consider an elliptic curve E/K and an integer $m \geq 2$ which we assume to be prime with the characteristic p of K, when $p > 0$. Recall that, seen as a group, $E[m] \simeq \mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/m\mathbb{Z}$. We have that $E[m]$ is a free $\mathbb{Z}/m\mathbb{Z}$-module or rank 2, and we have so a natural non-degenerate alternating multilinear map, the determinant, that has the drawback that is not Galois invariant, i.e., $\det(P^\sigma, Q^\sigma)$ is not necessarily equal to $\det(P,Q)^\sigma$. To have this Galois invariance, we will do a small modification and consider a pairing of the form $\zeta^{\det(P,Q)}$.
We are going to develop the definition to do it from an intrinsic point of view, recalling for that the result that a divisor $\sum n_i(P_i)$ is the divisor of a function (over an elliptic curve) if and only if $n_i = 0$ and $\sum [n_i]P_i = 0$. Therefore, if $T \in E[m]$ there is a function $f \in \bar{K}(E)$ such that

$$\operatorname{div}(f) = m(T) - m(O)$$

From our hypothesis that m is coprime with the characteristic, we can take $T' \in E$ such that $[m]T' = T$. Consequently, we have a function $g \in \bar{K}(E)$ such that

$$\operatorname{div}(g) = [m]^*(T) - [m]^*(O) = \sum_{R \in E[m]} (T' + R) - (R)$$

The sum of the divisors is clearly O since there are m^2 points of M-torsion and now we use that $[m^2]T' = O$. The functions $f \circ [m]$ and g^m have the same divisors, so multiplying f by a constant in K^*, we can assume that $f \circ [m] = g^m$.
Consider now another point $S \in E[m]$. Then, for any $X \in E$,

$$g(X+S)^m = f([m]X + [m]S) = f([m]X) = g(X)^m$$

and therefore the function $g(X+S)/g(X)$ takes finitely many values (m-th roots of the unity). We can define in these circumstances a map from E to $\mathbb{P}^1$ given by $S \mapsto g(X+S)/g(X)$, that clearly is not surjective, so it must be constant (a map of curves over an algebraically closed field is either constant or surjective). We have defined therefore the map

$$e_m : E[m] \times E[m] \to \mu_m \quad e_m(S,T) = \frac{g(X+S)}{g(X)}$$

(note that the dependence of T is seen when constructing the function g); it is well-defined where $X \in E$ is any point such that $g(X+S)$ and $g(X)$ are both defines and nonzero. We call this pairing Weil e_m-pairing.

Proposition 3.14. *The Weil e_m-pairing has the following properties:*

a) It is bilinear.

b) It is alternating.

c) It is non-degenerate: if $e_m(S,T) = 1$ for all $S \in E[m]$ then $T = O$.

d) It is Galois invariant.

e) It is compatible: $e_{mm'}(S,T) = e_m([m']S,T)$ for all $S \in E[mm']$ and $T \in E[m]$.

Proof. The linearity in the first factor is obvious. For the second one, let f_1, f_2, f_3, g_1, g_2, g_3 be the functions for the points $T_1, T_2, T_3 = T_1 + T_2$. Choose also a function $h \in \bar{K}(E)$ with divisor $\operatorname{div}(h) = (T_1 + T_2) - (T_1) - (T_2) + (O)$. Then,

$$\operatorname{div}\left(\frac{f_3}{f_1 f_2}\right) = m \cdot \operatorname{div}(h)$$

and we conclude that $f_3 = cf_1 f_2 h^m$. Compose now with the multiplication-by-m map, and then take m-th roots. That way, we have that $g_3 = d \cdot g_1 g_2 (h \circ [m])$, where $d \in \bar{K}^*$. Now,

$$e_m(S, T_1 + T_2) = \frac{g_1(X+S) g_2(X+S) h([m]X + [m]S)}{g_1(X) g_2(X) h([m]X)} = e_m(S,T_1) e_m(S,T_2)$$

The other properties follow also from the same game of algebraic manipulations. □

Corollary 3.6. *There exist points $S, T \in E[m]$ such that $e_m(S,T)$ is a primitive m-th root of unity. In particular, if $E[m] \subset E(K)$, then $\mu_m \subset K^*$.*

Proof. The image of $e_m(S,T)$ is a subgroup of μ_m, say μ_d. We have so that $1 = e_m(S,T)^d = e_m([d]S,T)$ for all S,T. Since the pairing is non-degenerate, $[d]S = O$, but S is arbitrary, so $d = m$. On the other side, if $E[m] \subset E(K)$, using the Galois invariance of the e_m-pairing, we have that $e_m(S,T) \in K^*$ for all S,T, and hence $\mu_m \subset K^*$. □

Proposition 3.15. *Let $\phi : E_1 \to E_2$ be an isogeny of elliptic curves. Then for all m-torsion points $S \in E_1[m], T \in E_2[m]$,*

$$e_m(S, \hat{\phi}(T))) = e_m(\phi(S), T)$$

Proposition 3.16. *There exists a bilinear, alternating, non-degenerate, Galois invariant pairing:*

$$e : T_l(E) \times T_l(E) \to T_l(\mu)$$

Furthermore, if $\phi : E_1 \to E_2$ is an isogeny, then ϕ and $\hat{\phi}$ are adjoints for the pairing: $e(\phi S, T) = e(S, \hat{\phi} T)$. More generally, if $\phi : E_1 \to E_2$ is a non-constant isogeny, there is a Weil pairing

$$e_\phi : \ker \phi \times \ker \hat{\phi} \to \mu_m$$

3.9 The endomorphism ring

Until now, we know that the endomorphism ring of an elliptic curve $\operatorname{End}(E)$ has characteristic 0, no zero divisors and rank at most 4 as a $\mathbb{Z}$-module. Furthermore, it possesses an anti-involution $\phi \to \hat{\phi}$ and for $\phi \in \operatorname{End}(E)$ the product $\phi\hat{\phi}$ is a nonnegative integer, being 0 if and only if $\phi = 0$. We define a concept that will reappear several times along the thesis:

Definition 3.10. *Let K be a $\mathbb{Q}$-algebra finitely generated over $\mathbb{Q}$. An order R of K is a subring of K finitely generated as a $\mathbb{Z}$-module satisfying $R \otimes \mathbb{Q} = K$. A definite quaternion algebra (we will then dedicate a whole chapter to them) is an algebra of the form $K = \mathbb{Q} + \mathbb{Q}\alpha + \mathbb{Q}\beta + \mathbb{Q}\alpha\beta$ satisfying $\alpha^2, \beta^2 \in \mathbb{Q}, \alpha^2, \beta^2 < 0, \alpha\beta = -\beta\alpha$.*

In this section we prove the following result:

Theorem 3.9. *Let R be a ring of characteristic 0, without 0 divisors, of rank at most four as a $\mathbb{Z}$-module, with an anti-involution satisfying: $\widehat{\alpha + \beta} = \alpha + \hat{\beta}, \hat{\alpha\beta} = \hat{\beta}\hat{\alpha}, \hat{\hat{\alpha}} = \alpha, \hat{\alpha} = \alpha$ for $\alpha \in \mathbb{Z} \subset R$. Assume also that $\alpha\hat{\alpha} = 0$ if and only if $\alpha = 0$. Then, R can only be one of the following type of rings:*

a) $R \cong \mathbb{Z}$.

b) R is an order in an imaginary quadratic extension of $\mathbb{Q}$.

c) R is an order in a definite quaternion algebra.

Proof. Let $K = R \otimes \mathbb{Q}$. Since R is finitely generated as a $\mathbb{Z}$-module, it suffices to prove that K is either $\mathbb{Q}$, an imaginary quadratic field or a quaternion algebra. For that, we extend the anti-involution to K and define the norm and trace from K to $\mathbb{Q}$ by $\operatorname{Nm}\alpha = \alpha\hat{\alpha}$ and $\operatorname{Tr}(\alpha) = \alpha + \hat{\alpha}$. We begin by observing that

$$\operatorname{Tr}(\alpha) = 1 + \operatorname{Nm}\alpha - \operatorname{Nm}(\alpha - 1)$$

(just routine). From here, we observe that the trace belongs to $\mathbb{Q}$ and that is $\mathbb{Q}$-linear. Furthermore, $\operatorname{Tr}(\alpha) = 2\alpha$ when $\alpha \in \mathbb{Q}$ and if $\alpha \in K$ satisfies $\operatorname{Tr}(\alpha) = 0$, then $0 = \alpha^2 + \operatorname{Nm}\alpha$. That way, $\alpha^2 \in \mathbb{Q}$ and $\alpha^2 < 0$.
If $K = \mathbb{Q}$, everything is clear. Otherwise, take $\alpha \in K$ but not in $\mathbb{Q}$ and replacing α by $\alpha - 1/2\operatorname{Tr}(\alpha)$ we can assume that the trace of α is 0. Then $\mathbb{Q}(\alpha)$ is a quadratic imaginary field, and if $K = \mathbb{Q}(\alpha)$ we are done. Elsewhere, take $\beta \in K$ not in $\mathbb{Q}(\alpha)$. Replace β by $\beta - 1/2\operatorname{Tr}(\beta) - \frac{\operatorname{Tr}(\alpha\beta)}{2\alpha^2} \cdot \alpha$. In that case, $\operatorname{Tr}(\beta) = \operatorname{Tr}(\alpha\beta) = 0$ and consequently $\beta^2 \in \mathbb{Q}, \beta^2 < 0$. We also have that $\operatorname{Tr}(\alpha) = \operatorname{Tr}(\beta) = \operatorname{Tr}(\alpha\beta) = 0$ and that $\alpha = -\hat{\alpha}, \beta = -\hat{\beta}, \alpha\beta = -\beta\alpha$. From here, $\alpha\beta = -\beta\alpha$ and we can conclude that

$$\mathbb{Q}(\alpha, \beta) = \mathbb{Q} + \mathbb{Q}\alpha + \mathbb{Q}\beta + \mathbb{Q}\alpha\beta$$

It only remains to prove that $\mathbb{Q}(\alpha, \beta) = K$, and it will be done if we see that $1, \alpha, \beta, \alpha\beta$ are linearly independent over $\mathbb{Q}$. If $w + x\alpha + y\beta + z\alpha\beta = 0$, then taking traces $w = 0$ and multiplying on the left by α and β, it yields

$$(x\alpha^2)\beta + (y\beta^2)\alpha + z\alpha^2\beta^2 = 0$$

What we know is that $1, \alpha, \beta$ are linearly independent, so $x\alpha^2 = y\beta^2 = z\alpha^2\beta^2 = 0$, which implies that $x = y = z = 0$. □

The next corollary is a consequence of the previous theorem together with the fact that when the characteristic of K is 0, then $\operatorname{End}(E)$ is commutative and so we cannot have the case of a quaternion algebra (this is because when the characteristic is 0, every endomorphism is separable and so $\operatorname{End}(E)$ injects into $\bar{K}^*$).

Corollary 3.7. *The endomorphism ring of an elliptic curve E/K is either $\mathbb{Z}$, an order in an imaginary quadratic field or an order in a quaternion algebra. If the characteristic of K is 0, the last option is not possible.*

Determine the endomorphism ring is not usually easy. For the automorphism group the situation is more direct and we have the following result, that is a consequence of a little bit work with the elliptic curve in its Weierstrass form.

Proposition 3.17. *Let E/K be an elliptic curve. Then, its automorphism group $\mathrm{Aut}(E)$ is a finite group. More precisely,*

a) It has order 2 when the j-invariant is neither 0 nor 1728.

b) It has order 4 when $j(E) = 1728$ and the characteristic of K is nor 2 neither 3.

c) It has order 6 when $j(E) = 0$ and the characteristic of K is nor 2 neither 3.

d) It has order 12 when $j(E) = 0$ and the characteristic is 3.

e) It has order 24 when $j(E) = 0$ and the characteristic is 2.

Corollary 3.8. *Let E/K be a curve over a field of characteristic nor equal to 2 or 3. Let $n = 2$ when $j(E) \neq 0, 1728$, $n = 4$ when $j(E) = 1728$ and $n = 6$ if $j(E) = 0$. Then, $\mathrm{Aut}(E) \cong \mu_n$.*

Chapter 4

Elliptic curves over finite fields

BSD conjecture relates the rank of an elliptic curve with the order of vanishing of a certain L-function obtained from taking information about the number of points the curve has over finite fields. In the previous chapter, we introduce elliptic curves from a very general point of view, and now we study its behavior in finite fields, with special attention to Hasse's theorem, that provides a bound for the number of points, and to the L-function associated to an elliptic curve. In chapter six, we will prove that over a number field, the rank of an elliptic curve is finite, and so we will have all the ingredients we need to understand the statement of BSD.

4.1 Number of rational points over a finite field

Take the usual Weierstrass equation of the affine curve

$$y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6$$

A first insight shows that for each value of x we can only have 2 possible values of y, so the number of points is bounded by $2q$. But, being more realistic, what we expect is that for half of the values of x, we will have a pair of solutions (half of the values are squares), and for the other half we will not have any solution. This would give us a value of q points ($q + 1$ taking into account the infinity). The main result of this section is:

Theorem 4.1. *Let $E/\mathbb{F}_q$ an elliptic curve over a finite field. Then,*

$$|\#E(\mathbb{F}_q) - q - 1| \leq 2\sqrt{q}$$

Proof. We take a Weierstrass equation with coefficients in $\mathbb{F}_q$ and we take the Frobenius morphism $(x, y) \mapsto (x^q, y^q)$. In the infinite Galois group $G_{\bar{\mathbb{F}_q}/\mathbb{F}_q}$, we take the Frobenius morphism acting on $\bar{\mathbb{F}_q}$, and for any point in the algebraic closure we know that $P \in E(\mathbb{F}_q)$ if and only if $\phi(P) = P$. We conclude that $E(\mathbb{F}_q) = \ker(1 - \phi)$. Since the map $1 - \phi$ is separable (as it was seen in the previous chapter) we know that $\# \ker(1-\phi) = \deg(1-\phi)$. Therefore, $\#E(\mathbb{F}_q) = \# \ker(1 - \phi) = \deg(1 - \phi)$. But the degree map on $\mathrm{End}(E)$ is a positive definite quadratic form (seen in the previous chapter) and $\deg \phi = q$. We will have the result from the following Cauchy-Schwarz type inequality. □

Proposition 4.1. *Let G be an abelian group and let d be a positive definite quadratic form taking values on the integers. Then*

$$|d(\psi - \phi) - d(\phi) - d(\psi)| \leq 2\sqrt{d(\phi)d(\psi)} \text{ for all } \psi, \phi \in G$$

Proof. It is basically the typical Cauchy-Schwarz inequality. Consider the associated bilinear form $\langle x, y\rangle = d(x+y) - d(x) - d(y)$, that takes values in $\mathbb{Q}$. In that case $|d(\psi - \phi) - d(\phi) - d(\psi)| = |\langle \phi, \psi\rangle|$ and what we obtain now is directly the Cauchy inequality (note that here we define the bilinear form without the usual 2 dividing, so we obtain that 2 in the RHS). □

Corollary 4.1. *Let now $f(x) = ax^3 + bx^2 + cx + d$ be a cubic polynomial with distinct roots in $\bar{\mathbb{F}_q}$ and let χ be the unique nontrivial character of order 2 (the one that takes the value 1 if it is a square and -1 otherwise). Take $E : y^2 = f(x)$. We will have that*

$$E(\mathbb{F}_q) = 1 + \sum_{x \in \mathbb{F}_q} (1 + \chi(f(x))) = q + 1 + \sum_{x \in \mathbb{F}_q} \chi(f(x))$$

so the sum must be smaller or equal than $2\sqrt{q}$.

4.2 Zeta functions of affine plane curves over finite fields

Zeta functions over number fields

Recall that the usual zeta function is defined as

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s} = \prod_p \Big(1 - \frac{1}{p^s}\Big)^{-1}$$

An elementary result in complex analysis is that the zeta function can be extended to the whole complex plane, being holomorphic except for a pole at $s = 1$:

Theorem 4.2. *The function $\zeta(s)$ can be extended to a function over all the complex plane, analytic except for a pole at $s = 1$ with residue 1. Further, it satisfies a functional equation: let $\xi(s) = \pi^{-s/2}\Gamma(s/2)\zeta(s)$ (where $\Gamma(s)$ is the usual Gamma function), so except at 0 and 1, $\xi(s)$ is bounded in all the vertical strip and satisfies*

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

Corollary 4.2. *The function $\xi(s)$ does not vanish in the half-plane $\Re(s) > 1$; in the half-plane $\Re(s) < 0$ only vanishes at the even negative integers $-2, -4, -6, \ldots$. All the other zeros are in the critical strip $0 \leq \Re(s) \leq 1$.*

We will not reproduce here the proof, that is a classical result. It uses harmonic analysis, defining the Fourier transform of a function f as $\hat{f} = \int_{\mathbb{R}} f(x)\exp(2\pi ixy)dx$. Then, we have Poisson's summation formula:

$$\sum_{n\in\mathbb{Z}} f(n) = \sum_{n\in\mathbb{Z}} \hat{f}(n)$$

Defining $\theta(u) = \sum_{n\in\mathbb{Z}} \exp(-\pi un^2)$ we can also prove that $\theta(1/u) = \sqrt{u}\theta(u)$ and combining this with a careful manipulation of integrals we arrive to the desired symmetry.

Over a number field we can also define a zeta function, usually called Dedekind zeta function. It is defined both by a sum or by a product that are convergent for $\Re(s) > 1$

$$\zeta_K(s) = \sum_I \frac{1}{\mathbb{N}(I)^s} = \prod_p \Big(1 - \frac{1}{\mathbb{N}(p)^s}\Big)^{-1}$$

where I runs over all nonzero ideals of the ring O_K and p runs over the nonzero prime ideals. Note that behind the equality of the two expressions we have the idea that the decomposition of an ideal as a product of prime ideals is unique in a Dedekind domain (and particularly in a number field).
The functional equation will involve the discriminant Δ_K, the number r_1 of real embeddings and the number r_2 of pairs of complex embeddings (we use here r_2 instead s to avoid a confusion with the complex variable s).

Theorem 4.3. *The function $\zeta_K(s)$ can be extended to a function over all the complex plane, holomorphic except for a simple pole at $s = 1$ with residue $\lambda(K)$. Furthermore, it satisfies a functional equation. Let*

$$\xi_K(s) = \Delta_K^{s/2}\Gamma_{\mathbb{R}}(s)^{r_1}\Gamma_{\mathbb{C}}(s)^{r_2}\zeta_K(s)$$

where $\Gamma_{\mathbb{R}}(s) = \pi^{-s/2}\Gamma(s/2)$ and $\Gamma_{\mathbb{C}}(s) = (2\pi)^{-s}\Gamma(s)$. Then, outside 0 and 1, $\xi_K(s)$ is bounded over all the vertical strip and verifies

$$\xi_K(s) = \xi_K(1-s)$$

There is an expression for the residue

$$\lambda(K) = \frac{h_K R_K}{\sqrt{\Delta_K}} \cdot \frac{2^{r_1}(2\pi)^{r_2}}{\omega_K}$$

where h_K is the number of classes; R_K is the regulator of units (we have not defined it, is a measure of how dense units are, for instance it is 1 in an imaginary quadratic field where units are a finite group and it is the logarithm of the fundamental unit in a real quadratic extension); finally ω_K is the number of roots of the unity.

Zeta functions of curves over finite fields

In the same spirit that we have defined zeta functions over number fields, we can do the same for curves over finite fields. For that, we start by considering an affine plane curve $C : f(X,Y) = 0$ over $\mathbb{F}_p$. As it could be expected, we just write, for complex numbers s with $\Re(s) > 1$,

$$\zeta(C,s) = \prod_p \frac{1}{1 - \mathbb{N}p^{-s}}$$

where the product is over the prime ideals of the ring of functions of the curve $\mathbb{F}_p[x,y] = \mathbb{F}_p[X,Y]/(f(X,Y)$. Note that when doing the quotient by a prime ideal, it is finite, and of course it is an integral domain, and it is a basic result in ring theory that under this hypothesis is also a field. We will denote by $\deg(p)$ its degree over $\mathbb{F}_p$ (since it would be a finite extension of this base field). Writing $\mathbb{N}p = p^{\deg(p)}$ we can rewrite our definition in the following way:

$$Z(C,T) = \prod_p \frac{1}{1 - T^{\deg(p)}}$$

and we trivially have the relation $\zeta(C,s) = Z(C,p^{-s})$. We will not discuss for the moment convergence issues and we will work with it as a formal series. The easiest way to understand this zeta functions is through an example.

Take for instance the affine curve $X = 0$. The quotient $F_p[X,Y]/(X)$ is just the polynomial ring in one variable, and there prime ideals are the same than irreducible polynomials. Our product is therefore a product over all the irreducible polynomials, and $\deg(p)$ is just the degree of the polynomial. But we do not know how many polynomials of degree n there is (in a closed form, I mean), what we know is that $X^{p^m} - X$ is the product of all the irreducible polynomials whose degree divides m. Taking logarithms, we put

$$\log Z(\mathbb{A}^1,T) = -\sum_f \log(1 - T^{\deg(f)})$$

We take now derivatives and finally get something that we can compute explicitly:

$$\frac{Z'(\mathbb{A}^1,T)}{Z(\mathbb{A}^1,T)} = \sum_f \frac{\deg(f)T^{\deg(f)-1}}{1 - T^{\deg(f)}} = \sum_f \sum_{n\geq 0} \deg(f)T^{(n+1)\deg(f)-1}$$

For instance, a prime polynomial of degree 3 will contribute with a coefficient of 3 to all the powers of T of the form T^{3k-1}. Using now the preceding observation

$$\frac{Z'(\mathbb{A}^1,T)}{Z(\mathbb{A}^1,T)} = \sum p^m T^{m-1}$$

If we finally integrate, it results that

$$\log Z(\mathbb{A}^1,T) = \log \frac{1}{1 - pT}$$

and finally

$$Z(\mathbb{A}^1, T) = \frac{1}{1 - pT}$$

We observe that in this example occurs a particular coincidence that is our first proposition of this section:

Proposition 4.2. *If N_m is the number of points of an affine curve C in $\mathbb{F}_{p^m}$, and let $Z(C, T)$ be the zeta function of C over $\mathbb{F}_p$. Then*

$$Z(C, T) = \exp\Big(\sum_{m \geq 1} \frac{N_m T^m}{m}\Big)$$

Proof. We just reproduce the steps of the example to get

$$\frac{Z'(C, T)}{Z(C, T)} = \sum_f \frac{\deg(p) T^{\deg(p)-1}}{1 - T^{\deg(p)}} = \sum_p \sum_{n \geq 0} \deg(p) T^{(n+1)\deg(p)} / T$$

Note that the coefficient of T^{m-1} is $\sum \deg(p)$ where the sum is over all the nonzero prime ideals such that $\deg(p)$ divides m. It also holds that

$$\deg(p) = [\mathbb{F}_p[x, y]/p : \mathbb{F}_p]$$

Recall also that, from field theory $\mathbb{F}_{p^r}$ is a subfield of $\mathbb{F}_{p^s}$ if and only if r divides s, so the condition $\deg(p)|m$ can be reformulated by saying that there is a homomorphism $\mathbb{F}_p[C]/p \to \mathbb{F}_{p^m}$. In particular, since the extension is separable, there will be $\deg(p)$ such homomorphisms (take a primitive element of the multiplicative field and send it to one of the m roots of its irreducible polynomial). Conversely a homomorphism between $\mathbb{F}_p[C]$ and $\mathbb{F}_{p^m}$ factors through $\mathbb{F}_p[C]/p$ for a prime ideal p, with $\deg(p)|m$ (the preimage of $\{0\}$ will be a prime ideal). We conclude that the coefficient of T^{m-1} is the number of homomorphisms of $\mathbb{F}_p$-algebras

$$\mathbb{F}_p[x, y] \to \mathbb{F}_{p^m}$$

But one such homomorphism is determined by the images a, b of x, y and conversely the homomorphism $P(X, Y) \mapsto P(a, b)$ factors through $\mathbb{F}_p[X, Y]/(f(X, Y))$ if and only if $f(a, b) = 0$. We conclude that there is a correspondence between homomorphisms from $\mathbb{F}_p[C]$ to $\mathbb{F}_{p^m}$ and the points of $C(\mathbb{F}_{p^m})$. Now, the result follows. □

We now extend this definition to the case of plane projective curves:

Definition 4.1. *For a projective plane curve C over $\mathbb{F}_p$ we define*

$$Z(C, T) = \exp\Big(\sum_{m \geq 1} \frac{N_m T^m}{m}\Big)$$

where as usual N_m denotes the number of points in $C(\mathbb{F}_{p^m})$.

For example, a trivial substitution gives us that

$$Z(\mathbb{P}^1, T) = \frac{1}{(1-T)(1-pT)}$$

and for an elliptic curve since we have for every finite field a point in the infinity,

$$Z(E, T) = \frac{Z(E^{\mathrm{aff}}, T)}{1-T}$$

To continue with this theory, we need a lemma to know the number of divisors of a fixed degree on an elliptic curve:

Lemma 4.1. *On an elliptic curve E over $\mathbb{F}_p$ the number of positive divisors of degree $m \geq 1$ is $N\frac{p^m-1}{p-1}$, where N is the number of points of $E(\mathbb{F}_p)$.*

Proof. Start by taking a divisor D_0 and consider the set of divisors that are equivalent to D_0, $P(D_0)$ (all those divisors D such that $D = D_0 + (f)$, where $f \in \mathbb{F}_p(E)^*$. We have a bijection between $(L(D_0)\backslash\{0\})/\mathbb{F}_p^*$ and $P(D_0)$ since the map $L(D_0)\backslash\{0\} \to P(D_0) : f \mapsto D_0 + (f)$ is clearly surjective (by definition) and two functions have the same image if and only if one is multiple of the other. We recall that using Riemann-Roch, if $m \geq 1$ and $\deg(D_0) = m$, we know that $\dim(L(D_0)) = m$. From here, the cardinal of $P(D_0)$ will be the quotient of the cardinal of $L(D_0)\backslash\{0\}$ (that is, $p^m - 1$) divided by $p - 1$. As we pointed out in the introduction to Riemann surfaces, the degree of a principal divisor factors through $\mathrm{Pic}(E)$. The map deg is surjective since we can take for instance p_∞, that is mapped to 1. If we consider $\mathrm{Pic}^m E$ as the elements in $\mathrm{Pic}(E)$ of degree m, we have these two trivial affirmations:

- The map $\mathrm{Pic}^0(E) \to \mathrm{Pic}^m(E) : D \to D + mp_\infty$ is a bijection.
- There is a bijection between $E(k)$ and $\mathrm{Pic}^0(E)$ (as we pointed out in chapter two).

Now the conclusion follows, since the number of elements in $\mathrm{Pic}^m(E)$ is the same as the number of points in $E(\mathbb{F}_p)$. □

We are now in conditions to state the most important theorem of this section:

Theorem 4.4. *Let E be an elliptic curve over $\mathbb{F}_p$. Then,*

$$Z(E, T) = \frac{1 + (N_1 - p - 1)T + pT^2}{(1-T)(1-pT)}$$

Proof. For what we say before, and passing to the affine form of the curve

$$Z(E, T) = \frac{1}{1-T}\prod_{p}\frac{1}{1-T^{\deg p}}$$

(the product again over the prime ideals of $\mathbb{F}_p[C] = \mathbb{F}_p[X, Y]/(Y^2 - X^3 - aX - b)$. That expression can also be written as $Z(E, T) = \sum d_m T^m$, where now d_m is the number of divisors of degree m, that have already been counted in the previous lemma. Therefore,

$$Z(E,T) = 1 + \sum_{m\geq 1} N_1 \frac{p^m-1}{p-1} T^m = 1 + \frac{N_1}{(p-1)(1-pT)} - \frac{N_1}{(p-1)(1-T)} =$$

$$= \frac{1+(N_1-p-1)T+pT^2}{(1-T)(1-pT)}$$

□

Beyond the simplicity of this proof, there are some deep remarks: take α, β to be the zeros of $Z(E,T)$. After some algebra,

$$\log Z(E,T) = \sum (1+p^m-\alpha^m-\beta^m)T^m/m$$

and we get that $N_m(E) = 1+p^m-\alpha^m-\beta^m$. There is also a relation with the Riemann hypothesis (the zeros of the ζ function have real part equal to 1/2). In fact

$$\zeta(E,s) = \frac{(1-\alpha p^{-s})(1-\beta p^{-s})}{(1-p^{-s})(1-p^{1-s})}$$

The poles are at $s=0,1$ and the zeros are the roots of $p^s=\alpha, p^s=\beta$, so the Riemann hypothesis is the same as saying that α, β have absolute value $\sqrt{p}$. But α, β are the inverse of the roots of a quadratic polynomials and by Cardano-Viete, their product is 1. So, if we see that the roots are complex conjugates, we will be done. But that is the same as saying that $|N_1-p-1| \leq 2\sqrt{p}$, and that is the content of Hasse's theorem.

4.3 Weil conjectures

Weil conjectures are one of the most well-known problems of the twentieth century. They were one of the excuses to produce a further development of algebraic geometry and as frequently happened in mathematics, they are elementary in its formulation. Furthermore, it has deep implications, as its relations with the Ramanujan conjecture about the coefficients of $\Delta(q)$, that will be introduced later as one of the most classical examples of modular forms.

$$\Delta(q) = q\prod_{n=1}^{\infty}(1-q^n)^2 4 = \sum_{n=1}^{\infty}\tau(n)q^n$$

Weil conjectures state that $|\tau(p)| \leq 2p^{11/2}$ for any prime number p and this is related to the problem of the number of ways to write a number as a sum of 24 squares. Until now, we have proved Riemann hypothesis for elliptic curves, but our aim would be to extend it to general curves over finite fields. In 1941, Andre Weil produce one such proof, but it relied on some facts in algebraic geometry that were only proved for varieties over the complex numbers. To address this deficiency, Weil wrote Foundations of Algebraic Geometry (1948), where it was introduced for the first time the notion of an abstract algebraic variety. In the next years, people made some deep conjectures about projective varieties in higher dimensions, the celebrated Weil conjectures, announced now in form of a deep theorem:

Theorem 4.5. *Let $V/\mathbb{F}_q$ be a smooth projective variety of dimension N. Then, we have:*

a) Rationality: $Z(V,T) \in \mathbb{Q}(T)$.

b) Functional equation: there is an integer ϵ (Euler characteristic of V), such that

$$Z(V, 1/q^n T) = \pm q^{N\epsilon/2} T^{\epsilon} Z(V,T)$$

c) Riemann Hypothesis: The zeta function factors as

$$Z(V,T) = \frac{P_1(T)\cdots P_{2N-1}(T)}{P_0(T)\cdots P_{2N}(T)}$$

where $P_i(T) \in \mathbb{Z}[T]$ and $P_0(T) = 1 - T$, $P_{2N}(T) = 1 - q^N T$. Furthermore $P_i(T) = \prod_{j=1}^{b_i}(1-\alpha_{ij}T)$, with $|\alpha_{ij}| = q^{1/2}$. We will call b_i the i-th Betti number of V.

Weil proved the conjecture for curves and abelian varieties, and Dwork established the rationality in 1960 using p-adic functional analysis. The development of l-adic cohomology by Artin, let Grothendieck to give another proof of rationality. It was in 1973 when Deligne, a student of Grothendieck, gave a proof of the Riemann hypothesis.

We have proved basically the Weil Conjectures for the case of elliptic curves using arguments of counting divisors. Another approach is also possible, which reveals some properties of the Frobenius morphism in characteristic p. For that, take a prime l different from the characteristic p of $\mathbb{F}_q$. We know that there is a representation

$$\mathrm{End}(E) \to \mathrm{End}(T_l(E)); \psi \mapsto \psi_l$$

just choosing a $\mathbb{Z}_l$-basis of $T_l(E)$. It makes sense to talk about the determinant and the trace of ψ_l. We have the following result:

Proposition 4.3. *Let $\psi \in \mathrm{End}(E)$. Then,*

$$\det(\psi_l) = \deg(\psi)$$

and

$$\mathrm{Tr}(\psi_l) = 1 + \deg(\psi) - \deg(1-\psi)$$

Proof. We will use some of the properties concerning the Weil pairing states in the previous chapter. Take a $\mathbb{Z}_l$-basis for $T_l(E)$ and write $\psi_l(v_1) = av_1 + bv_2$,

$\psi_l(v_2) = cv_1 + dv_2$.
Then,

$$e(v_1, v_2)^{\deg \phi} = e([\deg \phi]v_1, v_2) = e(\hat{\phi}_l \phi_l v_1, v_2) = e(\phi_l v_1, \phi_l v_2) =$$

$$= e(av_1 + cv_2, bv_1 + dv_2) = e(v_1, v_2)^{ad-bc} = e(v_1, v_2)^{\det \phi_l}$$

Since e is non-degenerate, $\deg \phi = \det \phi_l$. Furthermore, it is a trivial check that for any 2×2 matrix A,

$$\mathrm{Tr}(A) = 1 + \det(A) - \det(1 - A)$$

□

Using this, we can prove the following theorem, that has interest by itself and that can be eventually used to derive the proof of the Weil conjectures for elliptic curves in a similar way than before:

Theorem 4.6. *Let $E/\mathbb{F}_q$ be an elliptic curve and let $\phi : E \to E$ be the q-th power Frobenius endomorphism. Define $a = q + 1 - \#E(\mathbb{F}_q)$. Then,*

a) Let $\alpha, \beta \in \mathbb{C}$ be the roots of $T^2 - aT + q$. Then, α, β are complex conjugates of modulo $\sqrt{q}$ and for every $n \geq 1$,

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

b) The Frobenius endomorphism satisfies the following equation in $\mathrm{End}(E)$*:*

$$\phi^2 - a\phi + q = 0$$

Proof. The characteristic polynomial of ϕ_l can be factored as $\det(T - \phi_l) = T^2 - aT + q = (T - \alpha)(T - \beta)$. Note also that for every rational number $\det(m/n - \phi_l) = \deg(m - n\phi_l)/n^2 \geq 0$. A quadratic polynomial that is non-negative for all rational numbers, has either two complex conjugate roots or a double root. In either case, the modulo of both of them is the same and equal to the square root of q. Similarly, when we consider the q^n-th power Frobenius endomorphism, its roots will be α^n, β^n and so

$$\#E(\mathbb{F}_{q^n}) = \deg(1 - \phi^n) = \det(1 - \phi_l^n) = 1 - \alpha^n - \beta^n + q^n$$

The second part of the proposition follows from Cayley-Hamilton: ϕ_l satisfies its characteristic polynomial, and so $\phi_l^2 - a\phi_l + q = 0$. Consequently,

$$\deg(\phi^2 - a\phi + q) = 0$$

and so $\phi^2 - a\phi + q$ is the zero map. □

4.4 The endomorphism ring (revisited)

The following theorem relates the values of $E[p]$ and $End(E)$.

Theorem 4.7. *Let K be a field of characteristic p, and let E/K be an elliptic curve. For each integer $r \geq 1$, let*

$$\phi_r : E \to E^{(p^r)} \text{ and } \hat{\phi}_r : E^{(p^r)} \to E$$

The following conditions are equivalent:

a) $E[p^r] = 0$ for one (all) $r \geq 1$.

b) $\hat{\phi}_r$ is (purely) inseparable for one (all) $r \geq 1$.

c) The map $[p] : E \to E$ is purely inseparable and $j(E) \in \mathbb{F}_{p^2}$.

d) $\mathrm{End}(E)$ *is an order in a quaternion algebra.*

Proof. That a is equivalent to b is almost direct: since the Frobenius map is purely inseparable,

$$\deg_s(\hat{\phi}_r) = \deg_s[p^r] = (\deg_s[p])^r = (\deg_s \hat{\phi})^r$$

and now we recall that the cardinal of $E[p^r]$ is then

$$\deg_s(\hat{\phi}_r) = \deg(\hat{\phi})^r$$

The result now follows.

We prove now that b implies c. From the second condition, $[p] = \hat{\phi} \circ \phi$ is purely inseparable. We consider now the map $\hat{\phi} : E^{(p)} \to E$; by hypothesis $\hat{\phi}$ is purely inseparable. We can then factor it through a morphism ϕ' from $E^{(p)}$ to $E^{(p^2)}$ (the p-th power Frobenius map) and then a morphism ψ from $E^{(p^2)}$ to E; comparing degrees we see that it has degree one so it is an isomorphism. Then,

$$j(E) = j(E^{p^2}) = j(E)^{p^2}$$

To see that c implies d we proceed by contradiction, assuming that $K = \mathrm{End}(E) \otimes \mathbb{Q}$ is either $\mathbb{Q}$ or an imaginary quadratic extension of $\mathbb{Q}$. We consider E' an elliptic curve isogenous to E, $\psi : E \to E'$. Since $\psi \circ [p] = [p] \circ \psi$ and $[p] : E \to E$ is purely inseparable (hypothesis), taking the inseparability degrees we have that $[p] : E' \to E'$ is purely inseparable. Then, the number of points in $E'[p] = \deg_s[p] = 1$, and since we already know that a implies c, $j(E') \in \mathbb{F}_{p^2}$. We have seen that up to isomorphim there are finitely many elliptic curves isogenous to E.
Take now a prime $l \neq p$ such that l is still prime in $\mathrm{End}(E')$ for every elliptic curve E' isogenous to E. We know that

$$E[l^i] = \mathbb{Z}/l^i\mathbb{Z} \times \mathbb{Z}/l^i\mathbb{Z}$$

so we can take a sequence of subgroups

$$\Phi_1 \subset \Phi_2 \subset \ldots \subset E$$

with $\Phi_i \cong \mathbb{Z}/l^i\mathbb{Z}$. If $E_i = E/\Phi_i$, there is an isogeny $E \to E_i$ with kernel Φ_i. But up to isomorphism, there are only finitely many distinct E_i, so there are integers m, n such that E_{m+n} and E_m are isomoprhic. This yields an endomorphism of E_m,

$$\lambda : E_m \to E_{m+n} \cong E_m$$

The kernel of λ is cyclic of order l^n. But l is prime in the ring $\mathrm{End}(E_m)$ so by comparing degree $\lambda = u \circ [l^{n/2}]$ for some $u \in \mathrm{Aut}(E_m)$ (n even). But the kernel of $[l^{n/2}]$ is not cyclic for any n, so K is not a number field. We omit the proof that d implies b. □

4.5 The zeta function of a variety over $\mathbb{Q}$

Let V be a non-singular projective variety over $\mathbb{Q}$, so it is the zero set of a collection of homogeneous polynomials $F(X_0, \cdots, X_n) \in \mathbb{Q}([X_0, \cdots, X_n])$. We can assume that the polynomials are in $\mathbb{Z}$ and have no common factor, and consider its reductions $\bar{F}(X_0, \cdots, X_n)$ to $\mathbb{F}_p[X_0, \cdots, X_n]$. As it occurred with elliptic curves, when $\bar{F}$ defines a non-singular variety V_p we say that p is good for V. All but finitely many primes are good for a given variety. For each good prime, we have the usual zeta function

$$\zeta(V_p, s) = Z(V_p, p^{-s}), \text{ with } \log Z(V_p, T) = \sum \#V_p(\mathbb{F}_{p^m})\frac{T^m}{m}$$

It is natural to define

$$\zeta(V, s) = \prod_p \zeta(V_p, s)$$

where the product is over the good primes. Since the Riemann hypothesis holds for V_p, the product is convergent for $\Re(s) > d + 1$, where d is the dimension of V. Take for instance the variety $\mathbb{A}^0$ consisting on one single point, good at any prime. Then,

$$\zeta(\mathbb{A}^0, s) = \prod_p \frac{1}{1 - p^{-s}}$$

which is just the Riemann zeta function. If $V = \mathbb{P}^n$, then all primes are good and now

$$\zeta(\mathbb{P}^n, s) = \zeta(s)\zeta(s-1)\cdots\zeta(s-n)$$

where $\zeta(s)$ is the usual Riemann zeta function.
The following statement is known as the Hasse-Weil conjecture.

Conjecture 4.1. *For any non-singular projective variety V over $\mathbb{Q}$, $\zeta(V, s)$ can be analytically continued to a meromorphic function on the whole complex plane, and satisfies a function equation relating $\zeta(V, s)$ with $\zeta(V, d + 1 - s)$, where d is the dimension of V.*

The case of elliptic curves

Let E be an elliptic curve over $\mathbb{Q}$ and let S be the set of primes where E has bad reduction. Then

$$\zeta(E,s)=\prod_{p\notin S}\frac{1+(N_p-p-1)p^{-s}+p^{1-2s}}{(1-p^{-s})(1-p^{1-s})}=\frac{\zeta_S(s)\zeta_S(s-1)}{L_S(s)}$$

where $\zeta_S(s)$ is the Riemann's zeta function omitting the factors corresponding to the primes in S and L_S is the celebrated L-function of an elliptic curve. In a first approximation, we can define it as

$$L_S(E,s)=\prod_{p\notin S}\frac{1}{1+(N_p-p-1)p^{-s}+p^{1-2s}}$$

and if we write the denominator as $(1-\alpha_pT)(1-\beta_pT)$, then

$$L_S(E,s)=\prod_{p\notin S}\frac{1}{1-\alpha_pp^{-s}}\frac{1}{1-\beta_pp^{-s}}$$

Taking into account that $\prod_p \frac{1}{1-p^{-s}}$ converges for $\Re(s)>1$, then $\prod_p \frac{1}{1-p^{1/2-s}}$ will converge for $\Re(s)>3/2$. Since precisely $|\alpha_p|=|\beta_p|=\sqrt{p}$, it follows (the formal justification would require more rigor but the idea is clear enough) that $L_S(E,s)$ converges for $\Re(s)>3/2$.

The good definition of the L function will also take into account the bad primes:

Definition 4.2. *For any prime p, define $L_p(T)$ as:*

- $1-a_pT+pT^2$ *where* $a_p=p+1-N_p$ *if p is good.*
- $1-T$ *if E has split multiplicative reduction at p.*
- $1+T$ *if E has non-split multiplicative reduction at p.*
- 1 *if E has additive reduction.*

Note (and this will be important then in the understanding of BSD) that

$$L_p(p^{-1})=\frac{N_p}{p},\frac{p-1}{p},\frac{p+1}{p},\frac{p}{p}$$

In any case $L_p(p^{-1})=\#E^{ns}(\mathbb{F}_p)/p$ where E^{ns} is the non-singular part of the reduction of the elliptic curve. We finally define

$$L(E,s)=\prod\frac{1}{L_p(p^{-s})}$$

Another concept that should be properly understood is the conductor $N_{E/\mathbb{Q}}$ (or simply N) of an elliptic curve. It is defined to be an expression of the form $\prod_p p^{f_p}$ where f_p is zero for the good primes, 1 for those where p has multiplicative reduction and ≥ 2 for those where the reduction is additive, being equal to 2 if $p\neq 2,3$. For those values, the definition is more complicated and we do not need it here. The work of Wiles and others around modularity has lead to the following theorem:

Theorem 4.8. *The L-function $L(E,s)$ extends to an entire function on $\mathbb{C}$ and has a functional equation relating its value at s and $2-s$ of the form*

$$\Lambda(E,s) = \pm\omega_E \Lambda(E, 2-s)$$

where

$$\Lambda(E,s) = (2\pi)^{-s}\Gamma(s)N^{s/2}L(E,s)$$

Wiles (together with Taylor) proved the theorem for the case where E is a particular type of elliptic curve (semistable) and it was finally established in the modularity theorem.

The concept of L-function can be of course extended to the case of number fields: in general, take v a prime ideal and let $|v|$ be its norm. Then,

$$L(E/K,s) = \prod_v L_v(E/K,s)$$

where the local factors are given by

$$L_v(E/K,s) = (1 - a_{|v|}|v|^{-s} + |v|^{1-2s})^{-1}$$

when v does not divide N and

$$L_v(E/K,s) = (1 - a_{|v|}|v|^{-s})^{-1}$$

when $v|N$.

We prove here a theorem relating the L-function of an elliptic curve over a quadratic field with the L-function of the elliptic curve (over the rationals):

Theorem 4.9. *Let K be a quadratic extension of $\mathbb{Q}$. Then, the following formula holds:*

$$L(E/K,s) = L(E,s)L(E^D,s)$$

where E^D is the quadratic twist of E over K (this will be explained later; basically, it is a curve isomorphic to the given one over a quadratic extension of discriminant D).

Proof. Let

$$E : y^2 = x^3 + ax^2 + bx + c = f(x)$$

and consider now the quadratic twist, obtained by considering

$$E^D : Dy^2 = x^3 + ax^2 + bx + c$$

where D is the discriminant of the extension. Let a_p the corresponding coefficients of the first L-series, and let b_p the coefficients attached to E^D. Recall that $a_p = p + 1 - N_p$. We will denote by N_p the number of points of E and by M_p the number of points of E^D. If N is the conductor of the curve and p is a prime not dividing neither D nor N, both curves have good reduction in p. We distinguish now two cases:

If p splits in K (D is a square), clearly $y^2 = f(x)$ have the same number of

solutions than $Dy^2 = f(x)$ (it will depend if $f(x)$ is zero, a non-zero square or a non-square). Then, $M_p = N_p$ and so $a_p = b_p$ and since p splits, each of the primes contributes to the L-function of the curve over K with the same factor, that will be present once both in the L-function of E and E^D (seen now as curves in $\mathbb{Q}$).
We move now to the case in which p is inert. Then, we have that if $f(x)$ is zero modulo p then both $y^2 = f(x)$ and $Dy^2 = f(x)$ have one solution. Elsewhere, one of the equations will have two solutions and the other zero. Taking into account the infinite point

$$N_p + M_p = 2 + 2p$$

and so

$$a_p + b_p = p + 1 - N_p + p + 1 - M_p = 0$$

We conclude that

$$(1 - a_p p^{-s} + p^{1-2s})(1 - b_p p^{-s} p^{1-2s}) = 1 + 2p^{1-2s} + p^{2-4s} + (a_p b_p)p^{-2s}$$

If we see now the curve over K, we will have that $|p| = p^2$, and the inverse of the local factor will be

$$1 - a_{p^2} p^{-2s} + p^{2-4s}$$

and consequently all we need is

$$a_{p^2} = a_p^2 - 2p$$

that is true for the theory developed in the previous sections.
The case in which $p|ND$ requires a more careful manipulation, but the same conclusion holds. □

Chapter 5

An introduction to cohomology

In this chapter we present the main definitions about group cohomology, without giving many examples and proofs, focusing on the main theorems we will need in subsequent chapters. Cohomology is the natural language to state most of the theorems of number theory, so a good understanding of this seems to be necessary to be able to formulate many results not only around BSD, but in the whole area. However, in this first and naive approach in many of the topics, we will skip many technical aspects due to a lack of time.

5.1 Definition of the cohomology groups

Definition 5.1. *Let G be a group (abelian or not). A G-module is an abelian group M together with a map $G \times M \to M : (g, m) \mapsto gm$ such that for all $g, g' \in G, m, m' \in M$ satisfies:*

a) $g(m + m') = gm + gm'$.

b) $(gg')(m) = g(g'm), 1m = m$

Definition 5.2. *A homomorphism of G-modules is a map $\alpha : M \to N$ such that:*

a) $\alpha(m + m') = \alpha(m) + \alpha(m')$.

b) $\alpha(gm) = g\alpha(n)$.

An important remark is that if M and N are G- modules, then the set $\mathrm{Hom}(M, N)$ is a G-module with the structures

$$(\phi + \phi')(m) = \phi(m) + \phi'(m)$$

$$(g\phi)(m) = g(\phi(g^{-1}m))$$

Definition 5.3. *Let H be a subgroup of G, and let M be an H- module. We define $\mathrm{Ind}_H^G(M)$ to be the set of maps $\phi : G \to M$ such that $\phi(hg) = h\phi(g)$ for all $h \in H$.*

Again, $\mathrm{Ind}_H^G(M)$ is a G-module with the operations

$$(\phi+\phi')(x)=\phi(x)+\phi'(x)$$
$$(g\phi)(x)=\phi(xg)$$

Also note that when we have a homomorphism $\alpha : M \to M'$ of H-modules, this induces a homomorphism between $\mathrm{Ind}_H^G(M)$ and $\mathrm{Ind}_H^G(M')$ by sending ϕ to $\alpha\circ\phi$.

Proposition 5.1. *For every G-module M and H-module N,*

$$\mathrm{Hom}_G(M, \mathrm{Ind}_H^G(N)) \simeq \mathrm{Hom}_H(M,N)$$

Further, the functor

$$\mathrm{Ind}_H^G : \mathrm{Mod}_H \to \mathrm{Mod}_G$$

is exact.

In the particular case in which $H=\{1\}$, we use the notation $\mathrm{Ind}^G(M_0)$, that will denote the set of maps ϕ from G to M_0. A G-module will be called induced if it is isomorphic to $\mathrm{Ind}^G(M_0)$ for some abelian group M_0. We continue with more definitions:

Definition 5.4. *A G-module I is injective if every G-homomorphism from a submodule of a G-module extends to the whole module.*

Proposition 5.2. *The category Mod_G has enough injectivities, that is, every G-module M can be embedded into an injective G-module I.*

We are now ready to define the cohomology groups. Let M^G be the set of points of M fixed by G. The functor $M \mapsto M^G$ is left exact, and since the category of G-modules has enough injectivities we can use what is called the theory of derived functors (imagine that here goes an unexisting subsection about that). Summing up, if M is an object in an abelian category C, a resolution of M is a long exact sequence

$$0 \to M \to I^0 \to I^1 \to \cdots \to I^r \to \cdots$$

and when the I^r are injective objects of C, the resolution is said to be injective. One of the first results is that an injective resolution $M \to I^\bullet$ of M exists and it is unique in the sense that if $M \to J^\bullet$ is a second injective resolution there exists a homomorphism from $M \to I^\bullet$ to $M \to J^\bullet$. Basically, what we have is an injective resolution

$$0 \to M \to I^0 \to I^1 \to \dots$$

where the morphism that goes from I^i to I^{i+1} is d^i. Take now fixed points, and so the new complex need no longer be exact, and define now

$$H^r(G,M)=\frac{\mathrm{Ker}(d^r)}{\mathrm{Im}(d^{r-1})}$$

A first result is Shapiro's lemma. When M is a G-module, we can also regard $\mathbb{Z}$ as a G-module. Since any homomorphism from $\mathbb{Z}$ to M is uniquely determined by $\alpha(1)$, we have that

$$\mathrm{Hom}_G(\mathbb{Z}, M) \simeq M^G$$

Proposition 5.3. *Let H be a subgroup of G. For every H-module N, there is a canonical isomorphism*

$$H^r(G, \mathrm{Ind}_H^G(N)) \to H^r(H, N)$$

It is possible to define in a more intuitive way (at least, more similar to already known things like homology groups in algebraic topology) the cohommology groups. If P_r is the free $\mathbb{Z}$ module with basis the $(r+1)$-tuples $(g_0, \ldots, g_r)$ of elements of G, with G acting by multiplication, define a homomorphism $d_r : P_r \to P_{r-1}$ given by

$$d_r(g_0, \ldots, g_r) = \sum_{i=0}^{r} (-1)^i (g_0, \ldots, \hat{g}_i, \ldots, g_r)$$

Let ϵ be the map $P_0 \to \mathbb{Z}$ sending each basis element to 1. It is immediate to see that $d_{r-1} \circ d_r$ it is a complex.
An element of $\mathrm{Hom}(P_r, M)$ can be identified with a function $\phi : G^{r+1} \to M$, and ϕ is fixed by G if and only if

$$\phi(gg_0, \ldots, gg_r) = g(\phi(g_0, \ldots, g_r))$$

$\tilde{C}^r(G, M)$ is the set of ϕ satisfying the condition. It is natural to define a boundary map $\tilde{d^r} : \tilde{C}^r(G, M) \to \tilde{C}^{r+1}(G, M)$ as

$$(\tilde{d^r}\phi)(g_0, \ldots, g_{r+1}) = \sum (-1)^i \phi(g_0, \ldots, \hat{g}_i, \ldots, g_{r+1})$$

We will have that

$$H^r(G, M) \simeq \frac{\mathrm{Ker}(\tilde{d^r})}{\mathrm{Im}(\tilde{d^{r-1}})}$$

A homogeneous cochain $\phi : G^{r+1} \to M$ is determined by its values on the elements $(1, g_1, g_1 g_2, \ldots, g_1 \ldots g_r)$. We can introduce the group $C^r(G, M)$ of inhomogenous r-cochains of G with values in M, that are all maps $\phi : G^r \to M$. Define now a map $d^r : C^r(G, M) \to C^{r+1}(G, M)$ by

$$(d^r\phi)(g_1, \ldots, g_{r+1}) =$$

$$= g_1\phi(g_2, \ldots, g_{r+1}) + \sum_{j=1}^{r} (-1)^j \phi(g_1, \ldots, g_j g_{j+1}, g_{r+1}) + (-1)^{r+1}\phi(g_1, \ldots, g_r)$$

Now, letting $Z^r(G, M) = \mathrm{Ker}(d^r); B^r(G, M) = \mathrm{Im}(d^{r-1})$ the group of r-cocycles and r-coboundaries respectively, we have the following:

Proposition 5.4. *The sequence of groups $C^i(G, M)$ joined by the morphisms d^i is a complex ($d^r \circ d^{r-1} = 0$) and there is a canonical isomorphism*

$$H^r(G, M) \simeq \frac{Z^r(G, M)}{B^r(G, M)}$$

5.2 A more down to earth vision

Since this may seem a very abstract framework, we explain first the meaning of the 0-th and 1-st cohomology groups:

Proposition 5.5. *The* 0*-th cohomology group of the G-module M, denoted by M^G or $H^0(G,M)$ is the submodule of M consisting of all G-invariant elements:*

$$H^0(G,M) = \{m \in M \mid \sigma(m) = m, \forall \sigma \in G\}$$

A very typical picture is when we have a short exact sequence of G-modules. A natural question is what happens when we take G-invariants. It is straightforward to see that the first two exactness are preserved, but not the third one (there is a lack of surjectivity). This is measured through the first cohomology group. We particularize some of the previous definitions:

Definition 5.5. *Let M be a G-module. The group of* 1*-cochains from G to M is*

$$C^1(G,M) = \{\textit{maps } \xi : G \to M\}$$

The group of 1*-cocycles from G to M is given by*

$$Z^1(G,M) = \{\xi \in C^1(G,M) \mid \xi(\tau\sigma) = \tau(\xi(\sigma)) + \xi(\tau)\}$$

Finally, the group of 1*-coboundaries from G to M is*

$$B^1(G,M) = \{\xi \in C^1(G,M) \mid \textit{ there exists an } m \in M \textit{ such that } \xi(\sigma) = \sigma(m) - m\}$$

With these definitions,

$$H^1(G,M) = \frac{Z^1(G,M)}{B^1(G,M)}$$

The existence of a long exact sequence between the cohomology groups when we have a short exact sequence applied to

$$0 \to P \to M \to N$$

(where we will denote by ϕ and ψ the maps between P and M and M and N respectively) gives us a map $\delta : H^0(G,N) \to H^1(G,P)$ defined as follows: let $n \in H^0(G,N)$. Choose $m \in M$ such that $\psi(m) = n$ and define the cochain ξ by $\xi(\sigma) = \sigma(m) - m$. Since the value of ξ are in P, $\xi \in Z^1(G,P)$ and we define $\delta(n)$ to be the cohomology class of ξ.

Take now a subgroup H of G. As we pointed out, a G-module is also an H-module and a cochain from G to M can be also seen (by restriction) as a cochain from H to M. This process takes cocycles to cocycles and coboundaries to coboundaries, so we have a restriction homomorphism

$$\text{Res} : H^1(G,M) \to H^1(H,M)$$

In the same way, if H is a normal subgroup of G, the submodule M^H has a natural structure of G/H-module Let $\xi : G/H \to M^H$ be a 1-cochain. The projection map, followed by ξ and then by the inclusion $M^H \subset M$ gives a cochain from G to M that is well behaved, so we can define

$$\mathrm{Inf} : H^1(G/H, M^H) \to H^1(G, M)$$

We state now a result relating the inflation and restriction maps:

Proposition 5.6. *Let M be a G-module and let H be a normal subgroup of G. Then, the following sequence (when the maps are the inflation and restriction homomorphisms already defined) is exact:*

$$0 \to H^1(G/H, M^H) \to H^1(G, M) \to H^1(H, M)$$

5.3 Galois cohomology

Let K be a perfect field, $\bar{K}$ an algebraic closure and $G_{\bar{K}/K}$ the absolute Galois group (inverse limit of $G_{L/K}$ when L varies over all finite Galois extensions of K. $G_{\bar{K}/K}$ is a profinite group, equipped with a topology (Krull topology) in which a basis of open sets around the identity is that formed by the normal subgroups having finite index in $G_{\bar{K}/K}$.

Definition 5.6. *A discrete $G_{\bar{K}/K}$-module is an abelian group on which the Galois group acts such that the action is continuous for the profinite topology on $G_{\bar{K}/K}$ and the discrete topology on M. Equivalently, for all $m \in M$ the stabilizer of M is of finite index in the whole group.*

With this language, we can remember the following results studied in a standard course in Galois theory:

Proposition 5.7. *Let K be a field.*

a) $H^1(G_{\bar{K}/K}, \bar{K}^+) = 0.$

b) $H^1(G_{\bar{K}/K}, \bar{K}^*) = 0$ *(Hilbert's Theorem 90).*

c) If $\mathrm{char}(K)$ *does not divide m or it is 0, then* $H^1(G_{\bar{K}/K}, \mu_m) \simeq K^*/(K^*)^m$.

5.4 Twisting: general theory

In the next chapter we will talk about the computation of the rank of an elliptic curve and the possibility of finding a set of generators. When doing this, we have the problem of the existence or nonexistence of a single rational point on various other curves. These other curves are twists of E that are called homogeneous spaces.

Definition 5.7. *Let C/K be a smooth projective curve. The isomorphism group of C, $\mathrm{Isom}(C)$, is the group of $\bar{K}$-isomorphisms from C to itself. We denote the subgroup of $\mathrm{Isom}(C)$ consisting of isomorphisms defined over K by $\mathrm{Isom}_K(C)$. This group we have defined is sometimes called the automorphism group of C, but we have already defined $\mathrm{Aut}(E)$ to be the group of isomorphisms from E to E taking O to O. For instance, the group $\mathrm{Isom}(E)$ will contain translation maps.*

Definition 5.8. *A twist of C/K is a smooth curve C'/K that is isomorphic to C over $\bar{K}$. Two twists are equivalent if they are isomorphic over K. The set of twists of C/K modulo K-isomorphisms, is denoted by $\mathrm{Twist}(C/K)$.*

Let C'/K be a twist of C/K. There is an isomorphism $\phi : C' \to C$ defined over $\bar{K}$. A measure of how ϕ fails to be defined over $\bar{K}$ is given by the map

$$\xi : G_{\bar{K}/K} \to \mathrm{Isom}(C),\, \xi(\sigma) = \phi^\sigma \phi^{-1}$$

We note that ξ is a 1-cocycle:

$$\xi(\tau\sigma) = \phi^{\tau\sigma}\phi^{-1} = (\phi^\sigma\phi^{-1})^\tau(\phi^\tau\phi^{-1}) = (\xi(\sigma))^\tau\xi(\tau)$$

Theorem 5.1. *The cohomology class $\{\xi\}$ is determined by the K-isomorphism class of C' and is independent of the choice of ϕ. We thus obtain a natural map from $\mathrm{Twist}(C/K) \to H^1(G_{\bar{K}/K}, \mathrm{Isom}(C))$. Not only this: the map is a bijection.*

Proof. Let C''/K be another twist of C/K K-isomorphic to C'. Choose a $\bar{K}$-isomorphism $\psi : C'' \to C$. We have to show that the cocycles $\phi^\sigma\phi^{-1}$ and $\psi^\sigma\psi^{-1}$ are cohomologous. We know that there is a K-isomorphism $\theta : C'' \to C'$ and we can consider $\alpha = \phi\theta\psi^{-1} \in \mathrm{Isom}(C)$. Note that

$$\alpha^\sigma(\psi^\sigma\psi^{-1}) = (\phi\theta\psi^{-1})^\sigma(\psi^\sigma\psi^{-1}) = \phi^\sigma\theta\psi^{-1} = (\phi^\sigma\phi^{-1})(\phi\theta\psi^{-1}) = (\phi^\sigma\phi^{-1})\alpha$$

So the difference of the two is a principal crossed homomorphism and both are cohomologous.
In order to establish the bijection, recall that if C'/K and C''/K give the same cohomology class there is a map $\alpha \in \mathrm{Isom}(C)$ such that $\alpha^\sigma(\psi^\sigma\psi^{-1}) = (\phi^\sigma\phi^{-1})\alpha$. It is a simple calculation to prove that $\theta = \phi^{-1}\alpha\psi$ is defined over K by showing that $\theta^\sigma = \theta$. Therefore C'' and C' are isomorphic and give the same element of $\mathrm{Twist}(C/K)$. Surjectivity is not so direct and would require a longer discussion. □

We put an example: let E/K be an elliptic curve and let $K(\sqrt{d})$ be a quadratic extension of K. Consider the character $\chi : G_{\bar{K}/K} \to \{\pm 1\}$ such that $\chi(\sigma) = \sqrt{d}^\sigma/\sqrt{d}$. χ defines a 1-cocycle $\phi : G_{\bar{K}/K} \to \mathrm{Isom}(E)$, $\chi(\sigma) = [\xi(\sigma)]$. Let now C/K be the corresponding twist of E/K and let us derive an equation for C/K. We begin with an equation for E/K of the form $y^2 = f(x)$. Note that $\bar{K}(E) = \bar{K}(x, y)$ and $\bar{K}(C) = \xi(\bar{K}(x, y))$. Note that $[-1](x, y) = (x, -y)$ and this implies that the action of σ on $\xi(\bar{K}(x, y)$ is determined by $\sqrt{d}^\sigma = \xi(\sigma)\sqrt{d}$, $x^\sigma = x, y^\sigma = \xi(\sigma)y$. If we define $x' = x, y' = y/\sqrt{d}$ they are fixed by $G_{\bar{K}/K}$ and satisfy

$$dy'^2 = f(x')$$

the equation of an elliptic curve over K. The isomorphism of the curves is over the quadratic extension $K(\sqrt{d})$. It is an easy verification to check that we have a cocycle.

Chapter 6

Mordell-Weil theorem for elliptic curves

In this chapter we prove one of the main results in the theory of elliptic curves: Mordell-Weil theorem. The proof has two clear steps: the first one, known as the Weak Mordell Theorem, in which what we prove is that $E(K)/mE(K)$ is a finite group; this can be done for any number field and we present two approaches that are similar in the ideas but that have some differences. The final part needs to introduce what we call height functions to do a descendence procedure. This is quite technical and we just explain it for the case of $\mathbb{Q}$. For a number field, the details can be found in the second book of Silverman. In the last part of the chapter, we explore the notion of homogeneous spaces.

6.1 The Selmer and Tate-Shafarevich groups

We begin by recalling this result already stated in chapter two:

Lemma 6.1. *For every elliptic curve E over an algebraically closed field k and integer n, the map $E(k) \to E(k) : P \mapsto nP$ is surjective (assume n does not divide the characteristic).*

Proof. We can identify $E(k)$ with an algebraic variety (since k is algebraically closed); since this map is regular and E is connected and complete, the image is connected and closed, i.e., it is either a point or the whole of E, but the first is impossible. □

From the lemma, we have the following exact sequence:

$$\begin{tikzcd} 0 \arrow[r] & E_n(\mathbb{Q})^{\mathrm{al}} \arrow[r] & E(\mathbb{Q})^{\mathrm{al}} \arrow[r, "n"] & E(\mathbb{Q})^{\mathrm{al}} \arrow[r] & 0 \end{tikzcd}$$

and now we will take cohomologies; note that H^0 correspond to the points of the base field $\mathbb{Q}$ that are the ones fixed by the Galois group.

$$0 \longrightarrow E_n(\mathbb{Q}) \longrightarrow E(\mathbb{Q}) \xrightarrow{n} E(\mathbb{Q}) \longrightarrow H^1(\mathbb{Q}, E_n) \longrightarrow H^1(\mathbb{Q}, E) \xrightarrow{n} H^1(\mathbb{Q}, E)$$

and from this it is possible to extract another exact sequence, where $H^1(\mathbb{Q}, E)_n$ will denote the subgroup of elements in $H^1(\mathbb{Q}, E)$ killed by n:

$$0 \longrightarrow E(\mathbb{Q})/nE(\mathbb{Q}) \longrightarrow H^1(\mathbb{Q}, E_n) \longrightarrow H^1(\mathbb{Q}, E)_n \longrightarrow 0$$

It would be enough to prove that $H^1(\mathbb{Q}, E_n)$ is finite, but that does not necessarily holds. So we will proceed in a different way: consider E as an elliptic curve over $\mathbb{Q}_p$, obtaining a similar exact sequence and the following commutative diagram.

$$\begin{tikzcd}
0 \arrow[r] & E(\mathbb{Q})/nE(\mathbb{Q}) \arrow[r] \arrow[d] & H^1(\mathbb{Q}, E_n) \arrow[r] \arrow[d] & H^1(\mathbb{Q}, E)_n \arrow[r] \arrow[d] & 0 \\
0 \arrow[r] & E(\mathbb{Q}_p)/nE(\mathbb{Q}_p) \arrow[r] & H^1(\mathbb{Q}_p, E_n) \arrow[r] & H^1(\mathbb{Q}_p, E)_n \arrow[r] & 0
\end{tikzcd}$$

We cannot affirm the finiteness of $H^1(\mathbb{Q}, E_n)$, but we would like to have in its place a set containing the image of $E(\mathbb{Q})/nE(\mathbb{Q})$ and of course, being well-behaved in which concerns finiteness. If $\gamma \in H^1(\mathbb{Q}, E_n)$ comes from an element of $E(\mathbb{Q})$, its image γ_p in $H^1(\mathbb{Q}_p, E_n)$ comes from an element of $E(\mathbb{Q}_p)$. Once we are at this point, it is natural to define both the Selmer and Tate-Shafarevich group.

$$\begin{aligned} S^{(n)}(E/\mathbb{Q}) &= \{\gamma \in H^1(\mathbb{Q}, E_n) \mid \forall p, \gamma_p \text{ comes from } E(\mathbb{Q}_p)\} \\ &= \operatorname{Ker}\Big(H^1(\mathbb{Q}, E_n) \to \prod_{p=2,3,\cdots\infty} H^1(\mathbb{Q}_p, E)\Big) \\ W(E/\mathbb{Q}) &= \operatorname{Ker}\Big(H^1(\mathbb{Q}, E) \to \prod_{p=2,3,\cdots\infty} H^1(\mathbb{Q}_p, E)\Big) \end{aligned}$$

We need now some very basic homological algebra, in concrete a lemma that allow us to establish a connection between the Selmer and the Tate-Shafarevich groups:

Lemma 6.2. *Let α, β two maps of abelian groups (or modules) of the form:*

$$A \xrightarrow{\alpha} B \xrightarrow{\beta} C$$

In this case, we have the following exact sequence:

$$0 \longrightarrow \operatorname{Ker}(\alpha) \longrightarrow \operatorname{Ker}(\beta \circ \alpha) \xrightarrow{\alpha} \operatorname{Ker}(\beta)$$

$$\operatorname{Ker}(\beta) \longrightarrow \operatorname{Coker}(\alpha) \longrightarrow \operatorname{Coker}(\beta \circ \alpha) \xrightarrow{\alpha} \operatorname{Coker}(\beta) \longrightarrow 0$$

Most of the exactness are trivial and just a matter of writing down the definitions. For instance, note that when we go from $\text{Ker}(\alpha)$ to $\text{Ker}(\beta \circ \alpha)$ the image is precisely $\text{Ker}(\alpha)$, and the kernel of the next application is also $\text{Ker}(\alpha)$. For the next exactness, it is trivial that both the image and the kernel of the corresponding morphisms are $\text{Ker}(\beta) \cap \text{Im}(\alpha)$. The other ones follow in a similar way.
We will now apply this lemma to the following maps:

$$\begin{tikzcd} H^1(\mathbb{Q}, E_n) \arrow[r] & H^1(\mathbb{Q}, E)_n \arrow[r] & \prod_{p=2,3,\cdots,\infty} H^1(\mathbb{Q}_p, E)_n \end{tikzcd}$$

From here we will extract an exact sequence:

$$\begin{tikzcd} 0 \arrow[r] & E(\mathbb{Q})/nE(\mathbb{Q}) \arrow[r] & S^{(n)}(E/\mathbb{Q}) \arrow[r] & W(E/\mathbb{Q})_n \arrow[r] & 0 \end{tikzcd}$$

Note that by virtue of the first diagram of this section, the first map is surjective, so $\text{Coker}(\alpha) = 0$. Then we only have to identify the other sets; for instance, $\text{Ker}(\alpha) = E(\mathbb{Q})/nE(\mathbb{Q})$ also because of the same exact sequence. The other identifications directly follow from the corresponding definitions.
We go no to one of the keys aspects, the proof of the finiteness of the Selmer group. We begin with a lemma that tells us information about the behavior in $\mathbb{Q}_p$ once we know about $\mathbb{F}_p$.

Lemma 6.3. *Let E be an elliptic curve over $\mathbb{Q}_p$ with good reduction, and let n be an integer not divisible by p. A point $P \in (Q_p)$ is of the form nQ if and only if its image $\bar{P}$ in $E(\mathbb{F}_p)$ is of the form $n\bar{Q}$ for some $Q \in E(\mathbb{F}_p)$.*

$$\begin{tikzcd} 0 \arrow[r] & E^1(\mathbb{Q}_p) \arrow[r] \arrow[d, "n"] & E(\mathbb{Q}_p) \arrow[r] \arrow[d, "n"] & \bar{E}(\mathbb{F}_p) \arrow[r] \arrow[d, "n"] & 0 \\ 0 \arrow[r] & E^1(\mathbb{Q}_p) \arrow[r] & E(\mathbb{Q}_p) \arrow[r] & \bar{E}(\mathbb{F}_p) \arrow[r] & 0 \end{tikzcd}$$

Proof. First of all, necessity is obvious. In previous chapters we saw that the first vertical arrow is an isomorphism. So, let $P \in E(\mathbb{Q}_p)$ be such that $\bar{P} = n\bar{Q}$; then $P - nQ$ is 0 in $\bar{E}(\mathbb{F}_p)$ and is therefore 0 in $E^1(\mathbb{Q}_p)$. We conclude that $P - nQ = nQ'$ for $Q' \in E^1(\mathbb{Q}_p)$, so $P = n(Q + Q')$. □

We prove now another lemma in the context of algebraic number theory, but that will be very useful here, that also relates in some way extensions of $\mathbb{F}_p$ with those of $\mathbb{Q}_p$.

Lemma 6.4. *For any finite extension k of $\mathbb{F}_p$, there is an unramified extension K of $\mathbb{Q}_p$ of the same degree than $[k : \mathbb{F}_p]$ such that $O_k/pO_k = k$.*

Proof. Take a primitive element for k over $\mathbb{F}_p$ and let $f_0(X)$ be its minimum polynomial over $\mathbb{F}_p$. Take now a monic polynomial $f(X) \in \mathbb{Z}_p[X]$ such that $f_0(X) \equiv f(X)$ modulo p. It is a simple verification to see that $K = \mathbb{Q}_p[X]/(f(X))$ has the required properties. □

Lemma 6.5. *Let E be an elliptic curve over $\mathbb{Q}_p$ with good reduction, and let n be an integer not divisible by p. For $P \in E(\mathbb{Q}_p)$, there is a finite unramified extension K of $\mathbb{Q}_p$ such that $P \in nE(K)$.*

Proof. Take the same K than before. O_K is a principal ideal with p as the only prime, so every $\alpha \in K^*$ is of the form up^m, where $u \in O_K^*$ and $m \in \mathbb{Z}$. We can define m to be the order of α, and that way we have a homomorphism between K^* and $\mathbb{Z}$ that extends the usual homomorphism between $\mathbb{Q}_p^*$ and $\mathbb{Z}$. The key fact is that we can now translate our study of elliptic curves over $\mathbb{Q}_p$ to its unramified extensions, and in particular $E^0(K)/E^1(K) \simeq \bar{E}^{\mathrm{ns}}(k), E^n(K)/E^{n+1}(K) \simeq k$. Take so $P \in E(\mathbb{Q}_p)$ and $\bar{P} \in n\bar{E}(k)$ for a finite extension k of $\mathbb{Q}_p$. We conclude that $P \in nE(K)$ for any unramified extension K of $\mathbb{Q}_p$ where k is the residue field. □

We can now state the most important proposition of this part of the proof:

Proposition 6.1. *Let E be an elliptic curve over $\mathbb{Q}$ (let Δ be its discriminant) and let T the set of primes which do not divide Δ. Take $p \notin T, \gamma \in S^{(n)}(\mathbb{Q})$. There exists a finite unramified extension K of $\mathbb{Q}_p$ such that γ is 0 in $H^1(K, E_n)$.*

Proof. From the definition, we have $P \in E(\mathbb{Q}_p)$ mapping to the image γ_p of γ in $H^1(\mathbb{Q}_p, E_n)$. Since p does not divide 2Δ, E has good reduction at p and so there exists an unramified extension K of $\mathbb{Q}_p$ such that $P \in nE(K)$ and so γ_p maps to zero in $H^1(K, E_n)$.

$$\begin{tikzcd}
E(\mathbb{Q}) \arrow[r, "n"] \arrow[d] & E(\mathbb{Q}) \arrow[r] \arrow[d] & H^1(\mathbb{Q}, E_n) \arrow[d] \\
E(\mathbb{Q}_p) \arrow[r, "n"] \arrow[d] & E(\mathbb{Q}_p) \arrow[r] \arrow[d] & H^1(\mathbb{Q}_p, E_n) \arrow[d] \\
E(K) \arrow[r, "n"] & E(K) \arrow[r] & H^1(K, E_n)
\end{tikzcd}$$

□

Consider now L a number field, where as we already commented we have one valuation for each prime ideal of O_L, one for each embedding of L into $\mathbb{R}$ and another one for each pair of embeddings of L into $\mathbb{C}$. Write $P(p)$ for the set of valuations that extends $|\cdot|_p$. We have that

$$L \otimes_{\mathbb{Q}} \mathbb{Q}_p \simeq \prod_{v \in P(p)} L_v$$

For an elliptic curve over L, we can generalize the definition of the Selmer group and define

$$S^{(n)}(E/L) = \mathrm{Ker}\left(H^1(L/E_n) \to \prod_{v \in P} H^1(L_v, E)\right)$$

For our purposes, it will be easier to prove the finiteness of $S^{(n)}$ rather than that of $S^{(n)}(E/\mathbb{Q})$. This suffices because of the following lemma:

Lemma 6.6. *For any finite Galois extension L of $\mathbb{Q}$, the kernel of $S^{(n)}(E/\mathbb{Q}) \to S^{(n)}(E/L)$ is finite.*

Proof. We are dealing with the finiteness of an application between subgroups of $H^1(\mathbb{Q}, E_n)$ and $H^1(L, E_n)$, so we will prove that the kernel of this application is finite. But this coincides with the kernel of $H^1(\mathrm{Gal}(L/\mathbb{Q}), E_n(L))$, which is finite because both the Galois group and $E_n(L)$ are finite. □

We will call now C to the ideal class group C of O_L. Note also that a is a unit in O_L if $\mathrm{ord}_p(a) = 0$ for all prime ideals p. We clearly have the following exact sequence:

$$0 \to U \to L^* \to \bigoplus_p \mathbb{Z} \to C \to 0$$

where from algebraic number theory U is finitely generated (of rank $r + s - 1$ and C is finite).
Recall that if T is a finite set of prime ideals in L, we have the following pair of maps of abelian groups

$$L^* \to \bigoplus_p \mathbb{Z} \to \bigoplus_{p \notin T} \mathbb{Z}$$

where the second arrow is the natural projection. Use here the kernel-cokernel exact sequence and extract the following exact sequence

$$0 \to U \to U_T \to \bigoplus_{p \in T} \mathbb{Z} \to C \to C_T \to 0$$

We just need a final lemma to finish the proof of the finiteness of the Selmer group:

Lemma 6.7. *For any finite subset T of P (the set of valuations of the number field) that contains $P(\infty)$, let N be the kernel of*

$$L^*/L^{*n} \to \bigoplus_{p \notin T} \mathbb{Z}/n\mathbb{Z}$$

Therefore, we have an exact sequence

$$0 \to U_T/U_T^n \to N \to (C_T)_n$$

Proof. Recall that we have the following exact sequence:

$$0 \to U_T \to L^* \to \bigoplus_{p \notin T} \mathbb{Z} \to C_T \to 0$$

Let $\alpha \in L^*$ an element of N; then $n | \mathrm{ord}_p(\alpha)$ for all $p \notin T$ so we can map α to the class c of $\mathrm{ord}_p(\alpha)/n$ in C_T (with $nc = 0$). If $c = 0$, then there exists $\beta \in L^*$ such that $\mathrm{ord}_p(\beta) = \mathrm{ord}_p(\alpha)/n$ for all $p \notin T$. We conclude that α/β^n is in U_T, so it is well defined up to an element of U_T^n. □

6.2 Heights

The title of this section may seem unclear, and it is just to point out that heights, in general, are a key concept in number theory; different types of heights can be defined and its theory is crucial for the proof of many theorems, like this of Mordell-Weil or Bilu's equidistribution theorem. Here, let $P = (a_0 : \cdots : a_n) \in \mathbb{P}^n(\mathbb{Q})$; we will call it primitive representative for P if all the components are coprime integers. In that case, we define $H(P) = \max_i |a_i|$ and $h(P) = \log H(P)$.

We need now a technical proposition that tells us about then case when we have two homogeneous polynomials of degree m, $F(X,Y), G(X,Y) \in \mathbb{Q}[X,Y]$, that clearly define a map from $\mathbb{P}^1(\mathbb{Q})\backslash V(\mathbb{Q})$ to $\mathbb{P}^1(\mathbb{Q})$ by sending $(x,y) \mapsto (F(x,y) : G(x,y)$. Intuitively, our proposition will state that the height of $\phi(P)$ is approximately $mh(P)$ (note that this would be the case when taking for instance $(x^m : y^m)$).

Proposition 6.2. *If $F(X,Y)$ and $G(X,Y)$ do not have commons zeros in $\mathbb{P}(\mathbb{Q}^{\text{al}})$ there is a constant B such that*

$$|h(\phi(P)) - mh(P)| \leq B \textit{ for all } P \in \mathbb{P}^1(\mathbb{Q})$$

Proof. We may assume that F and G have integer coefficients. If $(a : b)$ is a primitive representative of P, then $H(\phi(P)) \leq CH(P)^m$, for some constant C, just by considering that the maximum of $|F(a,b)|, |G(a,b)|$ is a sum of $m+1$ monomials and considering the greater coefficient. That way we have, taking logarithms, that $h(\phi(P)) \leq mh(P) + \log C$, that is the easy inequality. For the remaining part, we will have to use some properties of the resultant of two polynomials:
F and G do not have common zeros, so its resultant is nonzero. We are going to consider the homogeneization of the two polynomials, $F(X/Y,1), G(X/Y,1)$, that can be viewed as polynomials in one variable X/Y (and its resultant R is the same than that of $F(X,Y), G(X,Y)$). Therefore, we can assure the existence of polynomials $U(X/Y,1), V(X/Y,1)$ of degree $m-1$ verifying

$$U(X/Y)F(X/Y,1) + V(X/Y)G(X/Y,1) = R$$

We now homogeneize, and obtain

$$U(X,Y)F(X,Y) + V(X,Y)G(X,Y) = RY^{2m-1}$$

and we perform the same trick with Y/X to obtain

$$U'(X,Y)F(X,Y) + V'(X,Y)G(X,Y) = RX^{2m-1}$$

Putting in the equation (a,b), we observe (Bezout) that $\gcd(F(a,b), G(a,b))$ divides $R \gcd(a^{2m-1}, b^{2m-1}) = R$. As we did before we have another constant D such that $U(a,b), U'(a,b), V(a,b), V'(a,b) \leq C(\max |a|, |b|)^{m-1}$. Combining this last inequality with the previous equations, we observe that

$$|R||a|^{2m-1} \leq 2C(\max(|a|,|b|)^{m-1} \max(|F(a,b)|, |G(a,b)|)$$

and the same for b. From here, and taking into account that $\gcd(F(a,b), G(a,b))$ divides R

$$H(\phi(P)) \geq \frac{1}{|R|} \max(|F(a,b)|, |G(a,b)|) \geq \frac{H(P)^m}{2C}$$

and taking logarithms we get the desired inequality. □

But the height we have defined could be called naive and it is not very interesting, so we will try to give a better definition more adjusted to our goals. Define therefore

$$\hat{h}(P) = \lim_{n\to\infty} \frac{h(2^n P)}{4^n}$$

Lemma 6.8. *For any $P \in E(\mathbb{Q})$, the sequence $h(2^n P)/4^n$ is Cauchy in $\mathbb{R}$ (and so it converges in $\mathbb{R}$).*

Proof. As we have seen before, there exists a constant C such that $|h(2P) - 4h(P)| \leq C$ for all P. Consider now positive integers $n \geq m$ and a point $P \in E(\mathbb{Q})$. We have the following chain of inequalities:

$$\Big|\frac{h(2^n P)}{4^n} - \frac{h(2^m P)}{4^m}\Big| = \Big|\sum_{i=m}^{n-1} \Big(\frac{h(2^{i+1}P)}{4^{i+1}} - \frac{h(2^i P)}{4^i}\Big)\Big| \leq$$

$$\leq \sum_{i=m}^{n-1} \frac{1}{4^{i+1}} |h(2^{i+1}P) - 4h(2^i P)| \leq \sum_{i=m}^{n-1} \frac{C}{4^{n+1}} \leq \frac{C}{3 \cdot 4^M}$$

□

The importance of this new height is that it behaves as a quadratic function, and in some way, it could be said that starting from h, is the unique function that, being close to h, behaves as a quadratic function. All these things are summarized in the following propositions:

Lemma 6.9. *There exists at most a function $g : E(\mathbb{Q}) \to \mathbb{R}$ that satisfies that $g(P) - h(P)$ is bounded and $g(2P) = 4g(P)$*

Proof. From the boundedness condition $|g(2^n P) - h(2^n P)| < C$. Impose now the second fact and we get that $|g(P) - \frac{h(2^n P)}{4^n}| \leq \frac{C}{4^n}$. From this, we get that $g(P)$ should be the above defined function $\hat{h}$. □

Our next aim will be the proof that g is a quadratic form. It is a relatively well-known result that it suffices to show to verify the parallelogram law:

Lemma 6.10. *A function $f : M \to K$ from an abelian group into a field whose characteristic is not 2 is a quadratic form if and only if it satisfies the parallelogram law:*

$$f(x+y) + f(x-y) = 2f(x) + 2f(y)$$

Proof. The necessity is straightforward. Recall that for proving that a function $f : M \to K$ is a quadratic form when $f(2x) = 4f(x)$ and the associated form defined as $\langle x, y\rangle = f(x+y) - f(x) - f(y)$ (we sometimes normalize this value dividing by 2) is biadditive, i.e., $\langle x+z, y\rangle = \langle x, y\rangle + \langle z, y\rangle$. Putting the corresponding values, we see that our objective is proving that

$$f(x+y+z) + f(x) + f(y) + f(z) = f(x+y) + f(y+z) + f(z+x)$$

We have to perform some tedious calculations using several times the parallelogram law:

$$f(x+y+z) = 2f(x+y) + 2f(z) - f(x+y-z) =$$
$$= 2f(x+y) + 2f(z) - 2f(y-z) - 2f(x) + f(x+z-y) =$$
$$= 2f(x+y) + 2f(z) - 2f(y-z) - 2f(x) + 2f(x+z) + 2f(y) - f(x+y+z)$$

We divide now by two and get

$$f(x+y+z) = f(x+y) + f(x+z) - f(x) + f(y) + f(z) - f(y-z) =$$

$$= f(x+y) + f(x+z) + f(y+z) - f(x) - f(y) - f(z)$$

as we wanted. To prove that $f(2x) = 4f(x)$ just put $x = y = 0$ and get $f(x) = 0$ (here we are using once more that the characteristic is not 2) and then $x = y$ and the result follows. □

Proposition 6.3. *The height function $\hat{h} : E(\mathbb{Q}) \to \mathbb{R}$ satisfies the parallelogram law, so it is a quadratic form.*

Proof. The proof is not immediate at all. We will begin by proving the existence of a constant C such that

$$h(P+Q) + h(P-Q) \leq 2h(P) + 2h(Q) + C$$

This would imply the desired fact in an easy way, since substituting P for 2^nP and the same for Q, then dividing for 4^n and taking limits, we get

$$\hat{h}(P+Q) + \hat{h}(P-Q) \leq 2\hat{h}(P) + 2\hat{h}(Q)$$

For the reverse inequality, perform the change $P = (P'+Q')/2, Q = (P'-Q')/2$ and using that we already know that $\hat{h}(2P) = 4\hat{(}P)$, we get

$$\hat{h}(P') + \hat{h}(Q') \leq \frac{\hat{h}(P'+Q') + \hat{h}(P'-Q')}{2}$$

All we have to prove therefore is that

$$H(P_1+P_2)H(P_1-P_2) \leq KH(P_1)^2H(P_2)^2$$

for a certain constant K. Call $P_1+P_2 = P_3, P_1-P_2 = P_4$ and for the coordinates of the points put $P_i = (x_i : y_i : z_i)$.
Consider now the point of coordinates $(x_3x_4 : x_3z_4 + x_4z_3 : z_3z_4) = (i : j : k)$.

Now, after tedious manipulations that we will not reproduce (see either Milne or Silverman for a more detailed computation) we obtain expressions for i, j, k and conclude that

$$H(i:j:k) \leq KH(P_1)^2H(P_2)^2$$

Now it only remains to prove that

$$H(i:j:k) \geq 1/2H(P_3)H(P_4)$$

and this is again a matter of being careful with algebraic manipulations. □

Now we are almost done and just need one more result:

Proposition 6.4. *Let $C > 0$ be a real number such that*

$$S = \{P \in E(\mathbb{Q}) \mid \hat{h}(P) \leq C\}$$

contains a set of coset representatives for $2E(\mathbb{Q})$ in $E(\mathbb{Q})$. Then, S generates $E(\mathbb{Q})$.

Proof. By contradiction. Take $Q \in E(\mathbb{Q})$ not in the subgroup generated by S and such that $\hat{h}(Q)$ takes the smallest value among these points. We know that there is a P such that $Q = P + 2R$, where $P \in S$ and $R \in E(\mathbb{Q})$. Since R is not in the subgroup generated by S (elsewhere Q would be), $\hat{h}(R) \geq \hat{h}(Q)$. That way

$$2\hat{h}(P) = \hat{h}(P+Q) + \hat{h}(P-Q) - 2\hat{h}(Q) \geq \hat{(}2R) - 2\hat{(}Q) = 4\hat{h}(R) - 2\hat{h}(Q) \geq 2\hat{h}(Q)$$

but $\hat{h}(P) \leq C$ and $\hat{h}(Q) > C$, and that is a contradiction. □

A few remarks about the canonical height

Proposition 6.5. *The canonical height extends to a positive definite quadratic form on the real vector space $E(K) \otimes \mathbb{R}$.*

The picture we have now is one contains an elliptic curve E/K, a finite dimensional vector space $E(K) \otimes \mathbb{R}$, a positive definite quadratic form in that space $\hat{h}$ and a lattice in $E(K) \otimes \mathbb{R}$, $E(K)/E_{\text{tors}}(K)$. In such a situation, an important invariant is the volume of a fundamental domain for the lattice (computed with the metric induced by the quadratic form):

Definition 6.1. *The canonical height pairing on E/K is the bilinear form*

$$\langle , \rangle : E(\bar{K}) \times E(\bar{K}) \to \mathbb{R}$$

defined by

$$\langle P, Q \rangle = \hat{h}(P+Q) - \hat{h}(P) - \hat{h}(Q)$$

The elliptic regulator of E/K, $R_{E/K}$ is the volume of a fundamental domain for $E(K)/E_{\text{tors}}(K)$ computed with the quadratic form $\hat{h}$. In other words, given a set of generators of the lattice $P_1, \ldots, P_r$,

$$R_{E/K} = \det(\langle P_i, P_j \rangle)$$

with the convention that if $r = 0$ the regulator is 1.

Extension to number fields

If K is a number field and O_K is not a PID there may be a problem: the non-existence of a primitive representative for a point $P \in \mathbb{P}^n(K)$. We define therefore a new heigth:

$$H(P) = \prod_{p=2,\cdots,\infty} \max_i(|a_i|_p)$$

For the product formula this does not depend of the representative chosen. For a number field, we copy this definition:

$$H(P) = \prod_v \max_i(|a_i|_v)$$

where the product runs over all the valuations.
Roughly speaking, we can say that the canonical height functions we would like to define

$$\hat{h} : E(\bar{K}) \to [0, \infty)$$

is a quadratic form that measures the arithmetic complexity at P, relating the geometrically defined group law to the arithmetic properties of the points on E. In the last chapter of the second book of Silverman there is a detailed treatment of local height functions in number fields.

6.3 The weak-Mordell-Weil revisited

Recall that the weak Mordell-Weil theorem states that if K is a number field and E/K is an elliptic curve, then, for $m \geq 2$

$$E(K)/mE(K)$$

is finitely generated. It is possible to proceed in a slightly different way simplifying some of the technicalities involved in the first proof. We begin by observing the following:

Lemma 6.11. *Let L/K be a finite Galois extension. If $E(L)/mE(L)$ is finite, then $E(K)/mE(K)$ is also finite.*

Proof. The inclusion of $E(K)$ in $E(L)$ induces a natural map

$$E(K)/mE(K) \to E(L)/mE(L)$$

The kernel of the map Φ will be $\Phi = \frac{E(K)\cap mE(L)}{mE(K)}$. For each P modulo $mE(K)$ we can take a point $Q_P \in E(L)$ such that $[m]Q_P = P$. Once we have done this, we define the map

$$\lambda_P : G_{L/K} \to E[m]; \quad \lambda_P(\sigma) = Q_P^\sigma - Q_P$$

It is straightforward that $Q_P^\sigma - Q_P \in E[m]$ and that if $P, P' \in E(K) \cap mE[L]$ satisfy $\lambda_P = \lambda'_P$, then they are equal modulo $mE(K)$. Hence, there is a one to one correspondence between the elements of Φ and the maps from $G_{L/K}$ to $E[m]$. But since both sets are finite, the set of maps is also finite. Therefore Φ is finite and we are done. □

So we can prove now the weak Mordell-Weil theorem under the assumption that $E[m] \subset E[K]$.

Definition 6.2. *The Kummer pairing*

$$\kappa : E(K) \times G_{\bar{K}/K} \to E[m]$$

is defined in the following way. Let $P \in E(K)$ and take $Q \in E(\bar{K})$ such that $[m]Q = P$. Then

$$\kappa(P, \sigma) = Q^\sigma - Q$$

Proposition 6.6. *The Kummer pairing is well defined and it is bilinear. Furthermore, it satisfies:*

1. *The kernel of the Kummer pairing on the left is $mE(K)$.*
2. *The kernel of the Kummer pairing on the right is $G_{\bar{K}/L}$, where*

   $$L = K([m]^{-1}E(K))$$

   is the compositum of all fields $K(Q)$ as Q runs over the points in $E(\bar{K})$ satisfying $[m]Q \in E(K)$.

Hence the Kummer pairing induces a perfect bilinear pairing

$$E(K)/mE(K) \times G_{L/K} \to E[m]$$

If we were interested now in proving this proposition, we would go through some of the facts we have seen in the first section of this chapter, but without explicitly introducing the Selmer and Shafarevich groups.
We have not still studied in many detail elliptic curve over local fields (we will go back when talking about complex multiplication), but for the moment let M_K^0, M_K^∞ be the non-archimedean and archimedean absolute values of K, respectively. Let $v \in M_K^0$ be a discrete valuation. We say that E has good reduction at v if E has good reduction over the completion K_v. Taking a minimal Weierstrass equation for E over K_v, denote by $\tilde{E}_v/k_v$ the reduced curve over the residue field (we have not done all these definition but they can be understood as a generalization of the p-adic case). It is not always possible to choose a single Weierstrass equation for E over K simultaneously minimal for all K_v, but it can be done if $K = \mathbb{Q}$.

Proposition 6.7. *Let $v \in M_K^0$ be a discrete valuation ring such that $v(m) = 0$ and such that E has good reduction at v. Then, the reduction map*

$$E(K)[m] \to \tilde{E}_v(k_v)$$

is injective.

We proved this for the case of $K = \mathbb{Q}$ when the residue field is simply $\mathbb{F}_p$. The proof of the general case is similar.
Let us analyze the extension L/K:

Proposition 6.8. *Let*

$$L = K([m]^{-1}E(K))$$

defined as before. Then, L/K is abelian with exponent m (every element of $G_{L/K}$ has order dividing m) and if

$$S = \{v \in M_K^0 : E \text{ has bad reduction at } v\} \cup \{v \in M_K^0 : v(m) \neq 0\} \cup M_K^\infty$$

then L/K is unramified outside S (that is, if $v \in M_K$ and $v \notin S$, L/K is unramified at v).

Assuming these results, there is just one thing remaining: prove that a field extension satisfying the conditions of the previous proposition is finite. Here is where we use both the finiteness of the ideal class group and the finite generation of the group of units:

Proposition 6.9. *Let K be a number field, $S \subset M_K$ a finite set of places containing M_K^∞ and $m \geq 2$ an integer. Let L/K be the maximal abelian extension of K with exponent m unramified outside S. Then, L/K is a finite extension.*

Proof. Imagine that we know that the proposition is true for some finite extension K' of K, where S' is the set of places of K' lying over S. Then, LK'/K' is abelian of exponent m unramified outside S' and so is finite, and hence L/K is also finite. So it is enough to prove the result under the assumption that K contains the m-th roots of unity μ_m. Similarly, we can increase the size of S, just making L larger. Using this fact, we can adjoin a finite number of elements to S so that the ring of S-integers, R_S (those that have non-negative valuation for all the valuations not in S) is a principal ideal domain, since it is more or less clear that there is an extension in which every element of the ideal class group becomes principal (for instance, the Hilbert class field). We also enlarge S to ensure that $v(m) = 0$ for all $v \notin S$.

We recall now the main theorem of Kummer theory:

Proposition 6.10. *If a field of characteristic 0 contains μ_m, its maximal abelian extension of exponent m is obtained by adjoining the m-th roots of all of its elements.*

Thus, L is the largest subfield of $K(\sqrt[m]{a} : a \in K)$ that is unramified outside S. Let now $v \in M_K$ with $v \notin S$, and consider $X^m - a = 0$ over K_v. Since $v(m) = 0$ and since the discriminant of the polynomial is $\pm p^m a^{m-1}$, we see that $K_v(\sqrt[m]{a})/K_v$ is unramified if and only if $\mathrm{ord}_v(a)$ is a multiple of m. Recalling that when adjoining m-th roots it is necessary to take only one representative for each class in $K^*/(K^*)^m$, if we let

$$T_s = \{a \in K^*/(K^*)^m : \mathrm{ord}_v(a) \equiv 0 \mod m \text{ for all } v \in M_K \text{ with } v \notin S\}$$

then

$$L = K(\sqrt[m]{a} : a \in T_S)$$

It will be enough to show that T_S is finite and this comes from the Dirichlet unit theorem applied to prove this claim:

Lemma 6.12. *The natural map*

$$R_S^* \to T_S$$

is surjective.

□

Summing up, if $L = K([m]^{-1}E(K))$, since $E[m]$ is finite, the perfect pairing induced by the Kummer pairing shows that $E(K)/mE(K)$ is finite if and only if $G_{L/K}$ is finite. Now, establishing that L has certain properties, we showed that any extension with these properties is finite.

6.4 Homogeneous spaces

Associated to an elliptic curve E/K we have a Kummer sequence like the following one:

$$0 \to \frac{E(K)}{mE(K)} \to H^1(G_{\bar{K}/K}, E[m]) \to H^1(G_{\bar{K}/K}, E)[m] \to 0$$

The key in the proof of Mordell's theorem was that the image of the first term in the second consists of elements unramified outside of a certain finite set of primes. Now, we analyze the third term, associating to each element of $H^1(G_{\bar{K}/K}, E)$ a twist of E called a homogeneous space.

Definition 6.3. *Let E/K be an elliptic curve. A principal homogeneous space for E/K is a smooth curve C/K with a simply transitive algebraic group action of E on C defined over K. Equivalently, a homogeneous space for E/K is a pair (C, μ) where C/K is a smooth curve and $\mu : C \times E \to C$ is a morphism over K satisfying:*

a) $\mu(p, O) = p$ for all $p \in C$.

b) $\mu(\mu(p, P), Q) = \mu(p, P + Q)$ for all $p \in C, P, Q \in E$.

c) For all $p, q \in C$ there is a unique $P \in E$ such that $\mu(p, P) = q$.

We frequently write $p + P$ instead $\mu(p, P)$.

With these definitions in mind, we can define a subtraction map from $C \times C$ to E, ν, characterized by $\nu(p, q)$ is the unique $P \in E$ satisfying $\mu(p, P) = q$. We now here state the most basic properties of addition and subtraction maps, just to be sure that they behave well:

Proposition 6.11. *Let C/K be a homogeneous space for E/K. Then, for all $p, q \in C$ and $P, Q \in E$,*

a) $p + O = p, p - p = O$.

b) $p + (q - p) = q, (p + P) - p = P$.

c) $(q+Q)-(p+P)=(q-p)+Q-P$.

Next, we show that a homogeneous space C/K for E/K is a twist of E/K and describe addition and subtraction on C in terms of a given $\bar{K}$-isomorphism $E \to C$.

Proposition 6.12. *Let E/K be an elliptic curve, and let C/K be a homogeneous space for E/K. Fix a point $p_0 \in C$ and define a map $\theta : E \to C$ by $\theta(P)=p_0+P$. Then,*

a) θ is an isomorphism over $K(p_0)$. In particular C/K is a twist of E/K.

b) For all $p \in C$ and all $P \in E$, $p+P=\theta(\theta^{-1}(p)+P)$.

c) For all $p,q \in C$, $p-p=\theta^{-1}(q)-\theta^{-1}(p)$.

d) The subtraction map is a morphism and is defined over K.

Proof. Everything is routine. For the first item, take $\sigma \in G_{\bar{K}/K}$ fixing p_0. Then, $\theta(P)^\sigma=\theta(P^\sigma)$, so θ is defined over $K(p_0)$. Further, since the action is simple and transitive, θ has degree one and so is an isomorphism.
For the second, $\theta(\theta^{-1}(p)+P)=p_0+\theta^{-1}(p)+P=p+P$, using only that $\theta^{-1}(p)$ is the unique point of E that gives p when added to p_0. The other claims follow in a similar way. □

Definition 6.4. *Two homogeneous spaces C/K and C'/K for E/K are equivalent if there is an isomorphism $\theta : C \to C'$ over K compatible with the action of E on C and C', or what is the same*

$$\theta(p+P)=\theta(p)+P$$

for all $p \in C$ and all $P \in E$. The equivalence class containing E/K is the trivial class, and the collection of equivalence class of homogeneous spaces for E/K is called the Weil-Chatelet group for E/K, $WC(E/K)$ (it is not still immediate why it is a group).

Proposition 6.13. *Let C/K be a homogeneous space for E/K. Then, C/K is in the trivial class if and only if $C(K)$ is not the empty set.*

Proof. If it is in the trivial class, there is a K-isomorphism $\theta : E \to C$ and so $\theta(O) \in C(K)$. For the converse, suppose that $p_0 \in C(K)$. Then, the map $\theta : E \to C$ defined by $\theta(P)=p_0+P$ is an isomorphism over $K(p_0)=K$. The compatibility condition on θ is $p_0+(P+Q)=(p_0+P)+Q$, which is part of the definition of homogeneous space. □

We state now the two main theorems of this section:

Theorem 6.1. *Let E/K be an elliptic curve. There is a natural bijection from $WC(E/K)$ to $H^1(G_{\bar{K}/K},E)$ defined as follows: when C/K is a homogeneous space for E/K, choose a point $p_0 \in C$ and do*

$$\{C/K\} \mapsto \{\sigma \mapsto p_0^\sigma - p_0\}$$

The following result states that if C/K is a homogeneous space for E/K, then $\mathrm{Pic}^0(C)$ can be canonically identified with E. This means that E is the jacobian variety of C/K. It can be proved that every curve C/K of genus one is a homogeneous space for some elliptic curve E/K, this shows that the group $\mathrm{Pic}^0(C)$ is always the group of points of an elliptic curve. This remains true in higher dimension, but the proof is clearly much more complicated.

Theorem 6.2. *Let C/K be a homogeneous space for an elliptic curve E/K. Choose a point $p_0 \in C$ and consider the map sum from $\mathrm{Div}^0(C)$ to E that sends $\sum n_i(p_i)$ to $\sum[n_i](p_i - p_0)$. Then,*

a) There is an exact sequence

$$1 \to \bar{K}^* \to \bar{K}(C)^* \to \mathrm{Div}^0(C) \to E \to 0$$

b) The summation map is independent of the choice of p_0.

c) The summation map commutes with the natural action of the Galois group on $\mathrm{Div}^0(C)$ and on E. Hence, it induces an isomorphism of $G_{\bar{K}/K}$ modules between $\mathrm{Pic}^0(C)$ and E.

Chapter 7

Modular functions and modular forms

The aim of this chapter is to introduce a basic tool not only in the study of elliptic curves, but in the whole area of number theory: modular forms. A priori, it may seem that this is unrelated with our previous work but as we have already pointed out in some moments, the deep connection comes from the modularity theorem, that assures that we can attach to any elliptic curve over $\mathbb{Q}$ a modular curve where things may seem easy. This is a recent theorem, proved in the last twenty years and with a lot of implications. The most important results around elliptic curves and Birch and Swinertonn-Dyer conjectures arise from this concept of modularity, so the study of this (at first sight) strange objects is highly advisable for a better understanding of elliptic curves. We will begin with some basic definitions, the analytic theory and we will move at the end of the chapter to the study of certain moduli interpretations that will play a preponderant role then in chapter ten. Throughout the pages, the connections with elliptic curves will be clear, for instance when proving that we have an L-function attached to the modular curve satisfying identical properties than those we saw in chapter four. Here, we will denote $\mathbb{H}$ as the complex upper plane ($\mathbb{H} = \{z \in \mathbb{C} | \Im(z) > 0\}$), that will be the natural place to develop our concepts. We will then see it as the Poincare upper half place, with its corresponding Haar measure given by the form $\frac{dxdy}{y^2}$.

7.1 Elliptic modular curves as Riemann surfaces

In this section we present some facts about $\mathbb{H}$.
There is a natural action of $\mathrm{SL}_2(\mathbb{R})$ on $\mathbb{H}$, given by

$$\mathrm{SL}_2 \times \mathbb{H} \to \mathbb{H}, \ (\alpha, z) \mapsto \alpha(z) = \frac{az+b}{cz+d}, \text{ where } \alpha = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

This is clear since $\Im(\alpha z) = \frac{\Im(z)}{|cz+d|^2}$; when we give SL_2 and $\mathbb{H}$ their natural topologies, the action is continuous. We can see this as a group acting on a topological space, so a natural question is to determine whether or not the action is transitive, and try to obtain some information from this fact. We have some easy-to-see properties:

Proposition 7.1. *The group* SL_2 *acts transitively on* $\mathbb{H}$

Proof. It is enough to show that, given $a \in \mathbb{H}$ we can map i to it; if $a = x + iy$, we consider the matrix $\begin{pmatrix} \sqrt{y} & \frac{x}{\sqrt{y}} \\ 0 & \frac{1}{\sqrt{y}} \end{pmatrix}$ which is clearly in our group and maps i to a. □

We can write now the orbit-stabilizer theorem; for that, take point i and impose $\frac{ai+b}{ci+d} = i$; from that, it results that $a = d$, $b = -c$ and from the group condition $a^2 + b^2 = 1$; this clearly correspond to the special orthogonal group, so we have that $\mathrm{SL}_2(\mathbb{R})/\mathrm{SO}_2(\mathbb{R}) \simeq \mathbb{H}$ is a group isomorphism. But we can give a stronger result: this is also a topological homeomorphism, just using the following result:

Proposition 7.2. *Suppose that* G *is a group acting continuously and transitively on* X. *If* G *and* X *are locally compact and Hausdorff, and* G *is second countable, then the map*

$$[g] \mapsto gx : G/\operatorname{Stab}(x) \to X$$

is a homeomorphism.

Proof. We saw that the map is a bijection (just orbit-stabilizer theorem) and continuity is also clear; we will show that it is open. For that, we take an open set U of G and let $g \in U$: we will show that gx is an interior point of Ux. For that, consider the map $\phi : G \times G \to G, (h, h') \mapsto ghh'$ (g is fixed); it is continuous and maps (e, e) to g (which is in U), so we have a neighborhood V of e (which can be assumed to be compact for being in a locally compact space) such that $V \times V$ is mapped into U (in fact, we can replace V with $V \cap V^{-1}$), and we have $gV^2 \subset U$.
Recall that $e \in V$, and we can write $G = \bigcup gV$, but each of the sets is a union of open sets of the basis, and for the assumption that G is 2AN, we only need a countable number of g's. Therefore $G = \bigcup g_n V$. We know that $g_n V$ is compact, so $g_n V x$ is also compact, and for being in a Hausdorff space, it is closed. At this point of the proof, we recall a well-known fact from general topology (Baire's theorem): if a nonempty locally compact and Hausdorff space X is a countable union $X = \bigcup V_n$ of closed subsets, then at least one of them has an interior point. Assume then that $g_n V x$ has an interior point, and consider the homeomorphism between Vx and $g_n V x$ $\psi : X \to X, y \mapsto g_n y$; we conclude that Vx has an interior point hx, for which we can take an open subset W of X such that $hx \in W \subset Vx$ and consequently $gx = gh^{-1}hx \in gh^{-1}W \subset gV^2x \subset Ux$. □

We continue our study recalling what is the group of automorphisms of $\mathbb{H}$.

Proposition 7.3. *The action of* $\mathrm{SL}_2(\mathbb{R})$ *on* $\mathbb{H}$ *induces an isomorphism*

$$\mathrm{SL}_2(\mathbb{R})/\{\pm \mathrm{Id}\} \to \mathrm{Aut}(\mathbb{H})$$

where $\mathrm{Aut}(\mathbb{H})$ *means here biholomorphic automorphisms (*$\mathbb{H}$ *seen as a Riemann surface)*

Proof. If an element of $\mathrm{SL}_2(\mathbb{R})$ fixes every z, we have that $cz^2+(d-a)z-b=0$, which forces $b=c=0$, $a=d$, but $ad=1$, so the only possibilities are $\pm\,\mathrm{Id}$. Let α be an automorphism of $\mathbb{H}$, and consider $\beta\in\mathrm{SL}_2(\mathbb{R})$ such that $\beta(i)=\alpha(i)$. If we change α by $\beta^{-1}\circ\alpha$ we have that i is a fixed point. Take now the usual isomorphism from $\mathbb{H}$ onto the open disk D, $\rho : z\mapsto\frac{z-i}{z+i}$, that sends i to 0. But it is an easy corollary of the Schwarz's lemma that an automorphism of the disk fixing the origin is of the form $f(z)=\lambda z$ (with $|\lambda|=1$), and therefore $\rho\circ\beta^{-1}\circ\alpha\circ\rho^{-1}=e^{2\theta i}z$. After an algebraic manipulation, we conclude that $\beta^{-1}\circ\alpha$ is an element of $\mathrm{SO}_2(\mathbb{R})$, and therefore also of $\mathrm{SL}_2(\mathbb{R})$; consequently $\alpha\in\mathrm{SL}_2(\mathbb{R})$. □

We move now to the study of congruence subgroups of $\mathrm{SL}_2(\mathbb{R})$.

Definition 7.1. *Let* $M=\begin{pmatrix} a & b\\ c & d\end{pmatrix}\in\mathrm{SL}_2(\mathbb{Z})$*. We say that one such matrix* M *is in* $\Gamma(N)$ *if it is congruent with the identity modulo* N*.* $\Gamma_0(N)$ *is the set of those satisfying* $c\equiv 0$ *modulo* N*, and* $\Gamma_1(N)$ *are the ones such that* $c\equiv 0$*,* $a,d\equiv 1$ *modulo* N*.*
A congruence subgroup of $\mathrm{SL}_2(\mathbb{Z})$ *is a subgroup containing* $\Gamma(N)$ *for some* N*.*
A Fuschian group is a discrete subgroup of $\mathrm{PSL}_2(\mathbb{R})$*.*

Proposition 7.4. *The sequence*

$$1\to\Gamma(N)\to\mathrm{SL}_2(\mathbb{Z})\to\mathrm{SL}_2(\mathbb{Z}/n\mathbb{Z})\to 1$$

is exact.

Proof. The only non-obvious fact is the surjectivity of the last one: take one element of $\mathrm{SL}_2(\mathbb{Z}/n\mathbb{Z})$, i.e., a matrix A whose determinant is congruent with 1 modulo n and with integral entries. We want to prove the existence of another matrix B (with integral coefficients) such that $A\equiv B$ and $\det(B)=1$. Let $A=\begin{pmatrix} a & b\\ c & d\end{pmatrix}$; we know, from the condition, that $ad-bc-Nk=1$. We can replace d with $d+nN$, where n is chosen in such a way that $\gcd(c,d+nN)=1$ (that can be done just by using the Chinese Remainder theorem, taking n such that $d+nN\equiv 1 \mod p$, for every prime dividing c but not N, and $n\equiv 0 \mod p$ for every prime dividing both c and N). We assume now that $(c,d)=1$ and we take $B=\begin{pmatrix} a+eN & b+fN\\ c & d\end{pmatrix}$, whose determinant is $ad-bc+N(ed-fc)=1+(m+ed-fc)N$; but since $(c,d)=1$, we can select e,f such that $ed-fc=-m$ (Bézout). □

We are now concerned with the classification of this linear fractional transformations in $\mathrm{SL}_2(\mathbb{R})$ (that we take as acting on $\mathbb{P}^1(\mathbb{C})$), identified with $\mathbb{C}\cup\infty$. We say that one such transformation α is:

a) Parabolic: if it has a unique fixed point (which will be real or infinite); it is the case when the Jordan form is not diagonal or equivalently, when $\mathrm{Tr}(\alpha) = \pm 2$ and the matrix is not $\pm\,\mathrm{Id}$.

b) Elliptic: two fixed points that are complex conjugates (one in the upper half plane and the other in the lower half plane). It admits diagonal form and $|\mathrm{Tr}(\alpha)| < 2$.

c) Hyperbolic: two real fixed points in $\mathbb{R} \cup \infty$. It admits diagonal form and $|\mathrm{Tr}(\alpha)| > 2$.

This can also be extended to $\mathrm{SL}_2(\mathbb{C})$ and the definitions are similar: α is parabolic when the trace is ± 2, elliptic when the trace is real and smaller than 2 in absolute value, hyperbolic when it is real and greater than 2 in absolute value and loxodromic when the trace is not real.
We introduce now some useful terminology.

Definition 7.2. *Let Γ be a discrete subgroup of $\mathrm{SL}_2(\mathbb{R})$. A point $z \in \mathbb{H}$ will be called elliptic point if it is fixed by some elliptic element of Γ; in the same way, it will be called a cusp if there is a parabolic element of Γ that has it as a fixed point.*

We start by the following proposition:

Proposition 7.5. *If z is an elliptic point of Γ, we have that $C = \{\gamma \in \Gamma \mid \gamma z = z\}$ is a finite cyclic group.*

Proof. We take an element β in $\mathrm{SL}_2(\mathbb{R})$ such that $\beta(i) = z$, and we then have an isomorphism given by $\gamma \mapsto \beta^{-1}\gamma\beta$ between C and $\mathrm{SO}_2(\mathbb{R}) \cap (\beta^{-1}\Gamma\beta)$ (recall that $\mathrm{SO}_2(\mathbb{R})$ are the ones fixing i); this last group is compact and discrete, so it is finite. Recall also that $\mathrm{SO}_2(\mathbb{R})_{\mathrm{tors}} \equiv \mathbb{Q}/\mathbb{Z}$ (since $\mathrm{SO}_2(\mathbb{R}) \equiv \mathbb{R}/\mathbb{Z}$) and every finite subgroup of $\mathbb{Q}/\mathbb{Z}$ is clearly cyclic. □

Take as a first example $\mathrm{SL}_2(\mathbb{Z})$; here the cusps must be between the points of $\mathbb{Q} \cup \infty$ and all of them will be equivalent to ∞. If we take $T = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$, ∞ is the only fixed point, and generally if $m/n \in \mathbb{Q}, (m,n) = 1$, we take r, s such that $mr - sn = 1$ and $R = \begin{pmatrix} m & s \\ n & r \end{pmatrix}$, for which $R(\infty) = m/n$ and RTR^{-1} is a parabolic element fixing m/n.

Definition 7.3. *Let Γ be a discrete subgroup of $\mathrm{SL}_2(\mathbb{R})$. A fundamental domain for Γ is a connected open subset D of $\mathbb{H}$ such that no two points of D are equivalent under Γ and $\mathbb{H} = \bigcup \gamma\bar{D}$ (equivalently, $D \to \Gamma\backslash\mathbb{H}$ is injective, or $\bar{D} \to \Gamma\backslash\mathbb{H}$ is surjective).*

It is a theorem that every Γ has a fundamental domain, but we we do not prove this here. In the literature, there are many pages written about how to calculate these fundamental domains, how to count their vertexes, how to find its area (it is not difficult to prove that they are polygons with finite sides in the Poincare plane with its usual metric $\frac{dx \cdot dy}{y^2}$). For instance, the fundamental domain for $\Gamma(1)$ is $D = \{z \in \mathbb{H}$ such that $|z| > 1, |\Re(z)| < 1/2\}$. The key step in the proof is to observe that $\Gamma(1)/\{\pm \mathrm{Id}\}$ is generated by $Tz = z + 1$ and $Sz = -1/z$.

Now we have some topological work to define our Riemann surfaces. Take first $\Gamma(1)\backslash\mathbb{H}$: if P is any non-elliptic point there, we take $Q \in \mathbb{H}$ mapping to it, and we can choose a neighborhood U of Q such that there is a homeomorphism between U and $p(U)$, where p is the projection map. The only elliptic points in $\Gamma(1)$ are i, ρ, ρ^2, and so for instance if Q is equivalent to i, take it directly equal to i and consider the map $z \mapsto \frac{z-i}{z+i}$ that is an isomorphism of an open neighborhood D of i stable under the inversion $S = -1/z$ onto an open disk D' centered at the origin, so the action of S on D is now the automorphism $\sigma : z \mapsto -z$ of D'. Note that $\langle S\rangle\backslash D$ is homeomorphic to $\langle\sigma\rangle\backslash D'$; we would like also a biholomorphic isomorphism between them. Summing up: $\frac{z-i}{z+i}$ is a holomorphic function defined in a neighborhood of i and S maps it to $-\frac{z-i}{z+i}$. Thus, $z \mapsto (\frac{z-i}{z+i})^2$ is a holomorphic function defined in a neighborhood of i invariant under S, so it is σ-invariant in a neighborhood of $p(i)$. With ρ the treatment is similar.
The problem is that this orbit space is not compact so we need to compactify it to have a new space $\mathbb{H}^*$. We consider so $\mathbb{H}^* = \mathbb{H} \cup \mathbb{P}^1(\mathbb{Q})$, the union of $\mathbb{H}$ with the set of cusps; here, the cusps are ∞ and the rational points of the real axis, that are of the form $\sigma(\infty)$ for some $\sigma \in \Gamma(1)$. We give now to $\sigma(\infty)$ the fundamental system of neighborhoods for which σ is a homeomorphism, and that way $\Gamma(1)$ acts continuously on $\mathbb{H}^*$ making sense to consider the quotient $\Gamma(1)\backslash\mathbb{H}^*$.
The next theorem is very direct from our construction (to prove it, just consider a triangulation or observe that $\Gamma(1)\backslash\mathbb{H}$ is simply connected and the only Riemann surface that is simply connected is the sphere.

Proposition 7.6. *The Riemann surface* $\Gamma(1)\backslash\mathbb{H}^*$ *is compact and of genus zero (and therefore isomorphic to the Riemann sphere).*

All these things can be done in any subgroup Γ of $\Gamma(1)$ of finite index, putting a complex structure just in the same way (remember also that we have to check the Hausdorff condition). We will write $X(\Gamma) = \Gamma\backslash H^*$ and $Y(\Gamma) = \Gamma\backslash\mathbb{H}$. $X(\Gamma)$ will be called a modular curve. In chapter two of [1] it is possible to find a detailed explanation.

To compute the genus of $X(\Gamma)$, consider it as a covering of $X(\Gamma(1))$ (of degree m). Riemann-Hurwitz's formula gives us

$$g = 1 - m + \sum(e_p - 1)/2$$

where e_p is the ramification index at a point P. It is not difficult using the multiplicity of the ramification indexes to obtain the following fundamental theorem:

Theorem 7.1. *Let* Γ *be a subgroup of* $\Gamma(1)$ *of finite index, let* ν_2 *be the number of inequivalent elliptic points of order* 2, ν_3 *the number of inequivalent points of* *order* 3 *and* ν_∞ *the number of inequivalent cusps. Then the genus of* $X(\Gamma)$ *is*

$$g = 1 + m/12 - \nu_2/4 - \nu_3/3 - \nu_\infty/2$$

The announced Taniyama-Weil conjecture states that, for an elliptic curve $E/\mathbb{Q}$, there exists a surjective function $X_0(N) \to E$, where N is the conductor of E and $X_0(N)$ the compatification of $\Gamma_0(N)\backslash\mathbb{H}$. This conjecture, that implies Fermat's last theorem, was proved by Breuil, Conrad, Diamond and Taylor, and previously the work of Wiles and Taylor made possible to deduce Fermat theorem just by proving the conjecture for semistable elliptic curves.
An elliptic curve for which there is a nonconstant map $X_0(N) \to E$ for some N is a modular elliptic curve, that is not the same than elliptic modular curves, that are the ones of the form $\Gamma\backslash\mathbb{H}^*$.

7.2 Elliptic functions

We take two complex numbers ω_1, ω_2 and maybe interchanging their roles we can assume that $\tau = \omega_1/\omega_2$ is in the upper half plane (we do not consider, for its lack of interest, the case where the number are not algebraically independent over the reals). We consider $\Lambda = \mathbb{Z}\omega_1 + \mathbb{Z}\omega_2$ the lattice generated by ω_1 and ω_2.
If Λ is a lattice in $\mathbb{C}$, we can make the quotient $\mathbb{C}/\Lambda$ into a Riemann surface as follows: let Q be a point in $\mathbb{C}$ and let P be its image in $\mathbb{C}/\Lambda$; for our construction, we have neighborhoods V of Q and U of P such that the quotient map p defines a homeomorphism; we take every such pair $(U, p^{-1} : U \to V)$ to be a coordinate neighborhood, so we have a complex structure that verifies that the map p is holomorphic and for every open subset U of the lattice, $f : U \to \mathbb{C}$ is holomorphic if and only if $f \circ p$ is holomorphic on $p^{-1}(U)$. Obviously, all these surfaces are topologically homeomorphic (torus of genus 1), but they will not be isomorphic as Riemann surfaces. In fact, two lattices are isomorphic if and only we can pass from one to the other through a homothety, and that every Riemann surface of genus one is in fact isomorphic to one such torus.
We make some remarks: algebraically, a complex torus is an abelian group under the addition it inherits from $\mathbb{C}$; geometrically, it is a parallelogram with its sides identified in opposing pairs (from this points of view it is clear that it is a Riemann surface).
From now on, we will deal with doubly periodic functions (also called elliptic functions) in a lattice generated by 1 and τ (after the normalization which is nothing but a homothety of the complex plane). Therefore we are in a situation of a function such that

$$f(z+1) = f(z), f(z+\tau) = f(z)$$

We have two different objects: first of all, the lattice, defined as

$$\Lambda = \{m + n\tau, m, n \in \mathbb{Z}\}$$

on the other, the fundamental parallelogram,

$$P_0 = \{z \in \mathbb{C} : z = a + b\tau, 0 \leq a, b < 1\}$$

An equivalence relation is defined, where two points are related if their difference is a point of the lattice. We state without proof some really obvious properties:

Proposition 7.7. *Let $\mathbb{C}/\Lambda$ be a lattice and let P_0 be its fundamental parallelogram. Let f be a complex function in the lattice, then:*

- *Every point in $\mathbb{C}$ is congruent to a unique point in the fundamental parallelogram.*
- *The function f is completely determined by its values in any period parallelogram.*
- *An entire doubly periodic function is constant (by Liouville's theorem).*
- *The number of poles at P_0 is at least 2 (by the residue theorem for Riemann surfaces, the sum of the residues is 0, so if there is one single pole its residue would be zero and that makes no sense).*

We give now our first example of elliptic function, the Weierstrass $\wp$ function:

Proposition 7.8. *Let $\Lambda^* = \Lambda\backslash\{(0,0)\}$ (the lattice minus the origin), and consider the function*

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

Then, $\wp(z)$ is an elliptic function with double poles at the points of the lattice.

Proof. We begin by analyzing the convergence of the series. We will see that the following two series converge when $r > 2$:

$$\sum_{(m,n)\neq(0,0)} \frac{1}{(|m|+|n|)^r}, \sum_{(m,n)\neq(0,0)} \frac{1}{(|n+m\tau|)^r}$$

To do this, note that if $n \neq 0$, then

$$\sum_{m\in\mathbb{Z}} \frac{1}{(|n|+|m|)^r} = \frac{1}{|n|^r} + \sum_{k\geq|n|+1} \frac{2}{|k|^r} \leq \frac{1}{|n|^r} + 2\int_{|n|}^{\infty} \frac{dx}{x^r} \leq \frac{1}{|n|^r} + \frac{C}{|n|^{r-1}}$$

From this, it is immediate (using that the series $\sum_{n\geq1}\frac{1}{n^\alpha}$ converges when $\alpha > 1$) that

$$\sum_{(m,n)\neq(0,0)} \frac{1}{(|m|+|n|)^r} \leq \sum_{m\neq0} \frac{1}{|n|^r} + \sum_{n\neq0} \left(\frac{1}{|n|^r} + \frac{C}{|n|^{r-1}}\right) < \infty$$

To prove the convergence of the second series, we will be done if we find a constant D such that $|m| + |n| \leq D|n + m\tau|$. But if $\tau = a + bi$ (where $b > 0$), we know that $|n + m\tau| = |(n + ma) + mbi| \geq |n + mbi| = \sqrt{n^2 + m^2b^2}$. If $b \geq 1$, then this last quantity is at least $\sqrt{n^2 + m^2} \geq \frac{|m|+|n|}{\sqrt{2}}$. If $b < 1$, we have that $\sqrt{n^2 + m^2b^2} > b\sqrt{n^2 + m^2} \geq b\frac{|m|+|n|}{\sqrt{2}}$. Now, the convergence of the $\wp$ is clear, since the function is

$$\wp(z) = \frac{1}{z^2} + \sum_{|\omega|\leq 2R} \frac{1}{(z+\omega)^2} - \frac{1}{w^2} + \sum_{|\omega|>2R} \frac{1}{(z+\omega)^2} - \frac{1}{w^2}$$

The second sum is $O(1/|\omega|^3)$ uniformly when $|z| < R$ and by the previous remarks we are done.
The existence of double poles in the lattice points is also clear, and for the periodicity we consider an important function, the derivative of $\wp(z)$

$$\wp'(z) = \sum_{(n,m)\in\mathbb{Z}} \frac{2}{(z+n+m\tau)^3}$$

This series converges absolutely when z is not in the lattice and it is clearly periodic with periods 1 and τ and therefore $\wp(z+1) = \wp(z)+a, \wp(z+\tau) = \wp(z)+b$ but since $\wp(z)$ is even, substituting $z = -1/2$ and $z = -\tau/2$, we conclude that $a = b = 0$. □

Define now $\wp(1/2) = a, \wp(\tau/2) = b, \wp((1+\tau)/2) = c$. The following equality holds:

Proposition 7.9.

$$(\wp')^2 = 4(\wp - a)(\wp - b)(\wp - c)$$

Proof. Note that $\wp'$ is elliptic of order 3 so it has three zeros; but for being odd, $\wp'(1/2) = -\wp'(-1/2) = -\wp'(-1/2+1)$, so the function has a zero in $1/2$ (and similarly in $\tau/2$ and $(1+\tau)/2$), and these three are the only ones. On the other side, since $\wp'$ has a single zero at $1/2$ and $\wp - a$ has also zero at $1/2$, it must have a double zero there (same reasoning for the other factors). But the function is of order 2, so it has no more zeros. Therefore, both the LHS and RHS have the same zeros and the same poles (all of order six in the points of the lattice). We conclude that $\frac{(\wp')^2}{(\wp-a)(\wp-b)(\wp-c)}$ is constant, and near zero $\wp(z) = \frac{1}{z^2} + \cdots, \wp'(z) = \frac{2}{z^3} + \cdots$. Consequently, the constant must be $2^2/1 = 4$. □

We finish this digression with a final proposition that states that every elliptic function is a simple combination of $\wp$ and $\wp'$.

Proposition 7.10. *Every elliptic function with periods* 1 *and* τ *is a rational function of* $\wp$ *and* $\wp'$.

Proof. We begin by showing that every elliptic function F with these periods and the additional property of being even, is a rational function of $\wp$. If F has a pole or a zero at the origin, it must have even order (F is even); we have so an integer m such that $F\wp^m$ has neither zeros nor poles at the lattice points. We assume so that F has neither zeros nor poles in the lattice points (elsewhere multiply by $\wp^m$). We consider $\wp(z) - \wp(a)$, and we already know that it has a double zero in a if a is a half-period and in any other case, it has two zeros in a and $-a$.
We now count the zeros and poles of F: if a is a zero so is $-a$, and the same with the poles: we consider that the zeros are $z_1, \ldots, z_m$ and the poles $p_1, \ldots, p_n$ and we define now

$$G(z) = \frac{(\wp(z) - \wp(z_1)) \ldots (\wp(z) - \wp(z_m))}{(\wp(z) - \wp(p_1)) \ldots (\wp(z) - \wp(p_n))}$$

We now have that F/G is holomorphic and doubly periodic, hence constant and we are done.

To finish the proof, having in mind that $\wp$ is even and $\wp'$ is odd, we write $f = f_{\text{even}} + f_{\text{even}}$, where $f_{\text{even}} = \frac{f(z)+f(-z)}{2}$ and $f_{\text{odd}} = \frac{f(z)-f(-z)}{2}$. But $f_{\text{odd}}/\wp'$ is even, so applying the previous result about even functions to $f_{\text{odd}}/\wp'$ and f_{even} we reach the desired result about $f(z)$. □

We state now the following remarkable result that is a trivial consequence of Riemann-Roch theorem, and that characterizes all the functions we have in a complex torus:

Proposition 7.11. *Let $P_1, \cdots, P_n$ and $Q_1, \cdots, Q_n$ be two sets of points ($n \geq 2$) in the complex plane such that P_i is not Q_j modulo the lattice Λ. If $\sum P_i \equiv \sum Q_i$ modulo Λ, there exists a periodic function $f(z)$ whose poles are the P_i and whose zeros are the Q_j, and $f(z)$ is unique up to multiplication by a nonzero constant.*

Eisenstein series

Define

$$G_{2k}(\Lambda) = \sum_{\omega \in \Lambda, \omega \neq 0} \omega^{-2k}$$

and define also $G_{2k}(z) = G_{2k}(z\mathbb{Z} + \mathbb{Z})$. Then, we have the following result:

Proposition 7.12. *$G_{2k}(z), k \geq 1$ converges to a holomorphic function on $\mathbb{H}$ taking the value $2\zeta(2k)$ at infinity.*

Proof. The convergence is clear from the previous results. To see the value at infinity, taking into account that converges uniformly and absolutely on compact sets,

$$\lim_{z \to i\infty} G_{2k}(z) = \sum \lim_{z \to i\infty} \frac{1}{(mz+n)^{2k}}$$

and the limit of each summand is 0 unless $m = 0$. Therefore,

$$\lim_{z \to i\infty} G_{2k}(z) = 2 \sum_{n \geq 1} \frac{1}{n^{2k}} = 2\zeta(2k)$$

□

The elliptic curve $E(\Lambda)$

We observe now a very remarkable fact. If Λ is a lattice in $\mathbb{C}$, we know that

$$\wp'(z)^2 = 4\wp(z)^3 - g_2\wp(z) - g_3$$

Let $E(\Lambda)$ be the projective curve defined by

$$Y^2Z = 4X^3 - g_2XZ^2 - g_3Z^3$$

It is natural to state the following proposition:

Proposition 7.13. *The curve $E(\Lambda)$ is an elliptic curve and the map $\mathbb{C}/\Lambda \to E(\Lambda)$ sending 0 to $(0:1:0)$ and any other point z to $(\wp(z):\wp'(z):1)$ is an isomorphism of Riemann surfaces, and reciprocally, every elliptic curve over $\mathbb{C}$ is isomorphic to $E(\Lambda)$ for some Λ.*

While the first statement is obvious, the second one is the content of the uniformization theorem. Its proof begins by observing that the universal cover of a Riemann surface of genus one is $\mathbb{C}$ and the proof that if G is a group of automorphisms of $\mathbb{C}$ without fixed points such that every orbit of G is discrete, then G is either the trivial group, the group of all translations of the form $z \mapsto z+n\gamma$ or the groups of translations $z \mapsto z+m\gamma_1+n\gamma_2$, where γ_1, γ_2 are linearly independent over $\mathbb{R}$.
Although it has no relation with modular forms, we state here a remarkable consequence of the uniformization theorem that can be useful to bear in mind:

Proposition 7.14. *Let $E/\mathbb{C}$ be an elliptic curve with Weierstrass coordinate functions x and y. Then,*

a) Let α, β closed paths on $E(\mathbb{C})$ giving a basis for $H_1(E,\mathbb{Z})$. Then, the periods

$$\omega_1 = \int_\alpha \frac{dx}{y}, \quad \omega_2 = \int_\beta \frac{dx}{y}$$

are linearly independent over $\mathbb{R}$.

b) Let Λ be the lattice generated by ω_1 and ω_2. Then the map $F : E(\mathbb{C}) \to \mathbb{C}/\Lambda$

$$F(P) \mapsto \int_O^P \frac{dx}{y} \mod \Lambda$$

is a complex analytic isomorphism of Lie groups.

Endomorphisms of $\mathbb{C}/\Lambda$

Proposition 7.15. *Let Λ, Λ' two lattices in $\mathbb{C}$. An element $\alpha \in \mathbb{C}$ such that $\alpha\Lambda \subset \Lambda'$ defines a holomorphic map ϕ_α between $\mathbb{C}/\Lambda$ and $\mathbb{C}/\Lambda', [z] \mapsto [\alpha z]$ that sends $[0]$ to $[0]$ and any such map is of this form.*

Proof. That α defines a holomorphic map of this form is trivial. Consider a holomorphic map $\phi : \mathbb{C}/\Lambda \to \mathbb{C}/\Lambda'$ such that $\phi([0]) = [0]$. $\mathbb{C}$ is the universal covering of both spaces, so by general topology ϕ lifts to a continuous map $\bar{\phi} : \mathbb{C} \to \mathbb{C}$ such that $\bar{\phi}(0) = 0$. The projection maps are local isomorphisms, then $\bar{\phi}$ is holomorphic (for being the composition of holomorphic maps). Take now $\omega \in \Lambda$; then $z \mapsto \bar{\phi}(z+\omega) - \bar{\phi}(z)$ takes values in Λ', but $\bar{\phi}$ is a continuous map from $\mathbb{C}$ to a discrete set, so for connection properties it must be constant. From here, we deduce that the derivative of $\bar{\phi}$ is a double periodic function, but it is also holomorphic, so it must be constant. We conclude that $\bar{\phi}(z) = \alpha z + \beta$, and if we impose that $\phi(0) = 0$, then $\beta = 0$. □

Note also that $\mathbb{Z} \subset \operatorname{End}(\mathbb{C}/\Lambda)$. From our previous proposition is already clear that two torus $\mathbb{C}/\Lambda$, $\mathbb{C}/\Lambda'$ are isomorphic if and only if $\Lambda' = \alpha\Lambda$. We can recover from here the structure of the ring of endomorphisms of $\mathbb{C}/\Lambda$. The main result here was that is either $\mathbb{Z}$ or a certain subring of O_K, for K an imaginary quadratic field:

Corollary 7.1. *$R = \operatorname{End}(\mathbb{C}/\Lambda)$ is either $\mathbb{Z}$ or there is imaginary quadratic field K such that R is a subring of O_K of rank 2 over $\mathbb{Z}$.*

Proof. We put $\Lambda = \mathbb{Z}\omega_1 + \mathbb{Z}\omega_2, \tau = \omega_1/\omega_2 \in \mathbb{H}$. For the previous result, we consider the α associated to a particular endomorphism. In concrete, assume that we have $\alpha \notin \mathbb{Z}$ such that $\alpha\Lambda \subset \Lambda$. Then, we must have $a, b, c, d, c \neq 0$ such that $\alpha\tau = a\tau + b, \alpha = c\tau + d$. Dividing both equations by α, it remains $c\tau^2 + (d - a)\tau - b = 0$. From here $[\mathbb{Q}(\tau) : \mathbb{Q}] = 2$ and to see that α is integral over $\mathbb{Z}$ (from any of the two equations we already have that is in $\mathbb{Q}(\tau)$) we just eliminate τ and it remains the equation $\alpha^2 - (a + d)\alpha + (ad - bc) = 0$. α is so integral over $\mathbb{Z}$ and is contained in the ring of integers of $\mathbb{Q}(\tau)$. □

7.3 Modular Functions and Modular Forms

Once we have constructed our subgroups Γ of finite index in $\Gamma(1)$ we are interested in studying holomorphic and meromorphic functions there.

Definition 7.4. *Let k be an integer and let Γ be a subgroup of finite index in $\Gamma(1)$. A meromorphic function $f : \mathbb{H} \to \mathbb{C}$ is weakly modular of weight k if*

$$f(\gamma(\tau)) = (c\tau + d)^k f(\tau) \text{ for } \gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma, \tau \in \mathbb{H}$$

Definition 7.5. *Let Γ be a subgroup of finite index in $\Gamma(1)$. A modular function for Γ is a meromorphic function on the compact Riemann surface $\Gamma\backslash\mathbb{H}^*$ (or a meromorphic function on $\mathbb{H}^*$ invariant under Γ). Generally, we can define a modular function f for Γ as a function on $\mathbb{H}$ that satisfies these three conditions:*

a) $f(\gamma z) = f(z)$ for all $\gamma \in \Gamma$ and all $z \in \mathbb{H}$.

b) f(z) is meromorphic in $\mathbb{H}$.

c) f(z) is meromorphic at the cusps.

We will do a few comments about the definition and that condition, that at first sight may seem subtle, of meromorphicity (maybe an invented word) at the cusps: for the cusp $i\infty$, we know that the group that fixes it in $\Gamma(1)$ is the free abelian group of rank 1 generated by T, and therefore the subgroup of Γ fixing $i\infty$ is of finite index in $\langle T\rangle$, so generated by $\begin{pmatrix} 1 & h \\ 0 & 1 \end{pmatrix}$ for some $h \in \mathbb{N}$. Therefore, $f(z + h) = f(z)$ and $f(z)$ can be expressed as a function $f^*(q)$ in the variable $q = \exp(2\pi iz/h)$. $f^*(q)$ is defined in a punctured disk around the origin, and the condition for f to be meromorphic at $i\infty$ means that is meromorphic at $q = 0$. This condition is extended to other cusps τ just by considering an element $\sigma \in \Gamma(1)$ taking $i\infty$ to τ; then the function $z \mapsto f(\sigma z)$ is invariant under $\sigma\Gamma\sigma^{-1}$ and $f(\sigma z)$ is required to be meromorphic at $i\infty$ like before.

Definition 7.6. *Let Γ be a subgroup of finite index in $\Gamma(1)$. A modular form for Γ of weight $2k$ is a function on $\mathbb{H}$ verifying these three properties:*

a) $f(\gamma z) = (cz + d)^{2k} f(z)$

b) $f(z)$ *is holomorphic in* $\mathbb{H}$

c) $f(z)$ *is holomorphic at the cusps of* Γ

With the notations above, the condition of holomorphicity means that the expansion of $f^*(q)$ does not have negative terms (when the first term vanishes we call it a cusp form). Note that condition a) is what we have called before weakly modularity. The easiest example of modular form are the Einsenstein series.
This definition may seem strange but is not. A modular form of weight 0 is nothing but a holomorphic modular function; a form of weight 2 will correspond to a differential one-form on the Riemann surface. That is, consider $\omega = f(z) \cdot dz$, where f is meromorphic. If $\gamma(z) = \frac{az+b}{cz+d}$, then $\gamma^*\omega = f(\gamma z)(cz + d)^{-2}dz$, so it is clear that modular forms of weight $2k$ correspond to Γ-invariant differential forms on $\Gamma\backslash\mathbb{H}^*$ (and in general, modular forms of weight $2k$ correspond to k-fold differential forms on $\Gamma\backslash\mathbb{H}^*$). It also makes sense to consider modular forms of odd weight, but the interpretation is not so clear.

The dimension of the space of modular forms

In these next sections we restrict our attention to modular forms of even weight. Our next objective is to compute the dimension of $M_{2k}(\Gamma)$, the space of modular forms of weight $2k$ for a subgroup of finite index of $\Gamma(1)$, and also of $S_{2k}(\Gamma)$, the subspace of cusp forms of weight $2k$. Note that when multiplying a modular form of weight k with another of weight l we get a modular form of weight $k + l$. Therefore we have that

$$M(\Gamma) = \bigoplus_{k\geq 0} M_{2k}(\Gamma)$$

is a graded ring (it would have been also possible to allow forms of odd weight).

Theorem 7.2. *The dimension of $M_{2k}(\Gamma)$ is:*

$$\dim(M_{2k}(\Gamma) = \begin{cases} 0 & \text{if } k < 0 \\ 1 & \text{if } k = 0 \\ (2k-1)(g-1) + \nu_\infty k + \sum_P \lfloor k(1 - 1/e_p\rfloor & \text{if } k > 0 \end{cases}$$

where g is the genus of $X(\Gamma)$, ν_∞ is the number of inequivalent cusps and the sum is over a set of representatives for the elliptic points, and e_P is the order or the stabilizer of P in the image of Γ in $\Gamma(1)/\{\pm \mathrm{Id}\}$

The proof of the theorem is almost direct combining the following lemma with the Riemann-Roch theorem.

Lemma 7.1. *Let f be a meromorphic modular form of weight $2k$ and let ω be the associated k-fold differential on $\Gamma\backslash\mathbb{H}^*$. Let $Q \in \mathbb{H}^*$ maps to $P \in \Gamma\backslash\mathbb{H}^*$.*

a) If Q is an elliptic point with multiplicity e, then $\mathrm{ord}_Q(f) = e\cdot\mathrm{ord}_p(\omega)+k(e-1)$

b) If Q is a cusp, then $\mathrm{ord}_Q(f) = \mathrm{ord}_p(\omega) + k$

c) If Q is any other point, then $\mathrm{ord}_Q(f) = \mathrm{ord}_p(\omega)$

Using the preceding lemma, we can also count the number of zeros and poles of a meromorphic differential form and get the following theorem:

Theorem 7.3. *Let f be a meromorphic modular form of weight $2k$, then*

$$\sum(\mathrm{ord}_Q(f)/e_Q - k(1 - 1/e_Q)) = k(2g - 2) + k\nu_\infty$$

where the sum is over a set of representatives for the points in $\Gamma\backslash\mathbb{H}^*$, ν_∞ is the number of inequivalent cusps and e_Q is the ramification index of Q over $p(Q)$ when $Q \in \mathbb{H}$ and 1 when Q is a cusp.

Proof. We just have to sum the equalities of the previous lemma over the three types of points we have (elliptic, cusps and the remaining ones). □

We are going to state some specific properties for the case of $\Gamma(1)$, where we can give a precise description of all the modular forms.

Proposition 7.16. *In $\Gamma(1)$, the following facts hold:*

a) For $k < 0$ and $k = 1$, $M_{2k} = 0$.

b) For $k = 0, 2, 3, 4, 5$, M_{2k} is a space of dimension 1 where a possible basis is $1, G_4, G_6, G_8, G_{10}$, respectively.

c) Multiplication by $\Delta = g_4^3 - 27g_6^2$ (where $g_4 = 60G_4, g_6 = 140G_6$) defines an isomorphism of M_{2k-12} onto S_{2k}.

d) $\oplus M_{2k} = \mathbb{C}[G_4, G_6]$.

We quote here a proposition that will be useful:

Proposition 7.17. *The assignment which to a cuspidal form $f \in S_2(\Gamma)$ of weight two associates the expression*

$$\omega_f = 2\pi i f(\tau)\tau$$

identifies $S_2(\Gamma)$ with the space of holomorphic differential forms on $X_0(N)(\mathbb{C})$.

Petersson inner product

Before the proper beginning of this part, we quote a theorem from linear algebra that will be our motivation (the spectral theorem):

Theorem 7.4. *Let V be a finite dimensional complex vector space with a positive definite hermitian form $\langle , \rangle$. Then,*

1. *Any self-adjoint linear map $\alpha : V \to V$ is diagonalizable, that is, V is a direct sum of eigenspaces for α.*

2. *Let $\alpha_1, \alpha_2, \dots$ be a sequence of commuting self-adjoint linear maps $V \to V$. Then V has a basis consisting of vectors that are eigenvectors for all α_i.*

Let f and g be two modular forms of weight $2k > 0$ for a subgroup Γ of finite index in $\Gamma(1)$.

Lemma 7.2. *Let $w(z)$ be a holomorphic function. Then the map $z \mapsto w(z)$ multiplies areas by $|\frac{dw}{dz}|^2$*

Proof. Write $w(z) = u(x,y) + iv(x,y)$, and so the jacobian of our map is $u_x v_y - v_x u_y$ and using now the equations of Cauchy-Riemann this is equal to $u_x^2 + v_x^2 = |w'(z)|^2$. □

The next lemma is now almost straightforward just using that a k-fold differential form on a Riemann surface can be written, locally, as $\omega = f(z)(dz)^k$ and if $g = g(z)$, then $g^*\omega = f(g(z))(dg(z))^k = f(g(z))g'(z)^k(dz)^k$ (this is a classical result for differential forms).

Lemma 7.3. *The differential $f(z)\overline{g(z)}y^{2k-2}dxdy$ is invariant under the action of* $\mathrm{SL}_2(\mathbb{R})$

Proof. Note that $f(\gamma z)\overline{g(\gamma z)} = (cz+d)^{2k}\overline{(cz+d)}^{2k} f(z)\overline{g(z)}$ and $\gamma^*(dx \cdot dy) = \frac{dx \cdot dy}{|cz+d|^4}$ (for the previous lemma). Multiplying all this, we get the result. □

For which concerns convergence, we have the following:

Lemma 7.4. *Let D be a fundamental domain for Γ. If f or g is a cusp form then the integral*

$$\int\int_D f(z) \cdot \overline{g(z)} y^{2k-2} dxdy$$

converges.

If we recapitulate, what we have is that, if f and g are modular forms of weight $2k$ for some group $\Gamma \subset \Gamma(1)$, at least one of them being a cusp form, the Petersson inner product satisfies that is a positive definite hermitian form on $S_{2k}(\Gamma)$, so this is a finite dimensional Hilbert space. For the Hilbert basis theorem, we should have that every cusp form is a linear combination (not necessarily finite) of a certain basis, called in this case the Poincare basis.

7.4 Hecke operators

This is one of the most important points in the theory of modular points. We will define two types of Hecke operators: first of all we will define them on L, the set of full lattices in $\mathbb{C}$. Then we will define them on modular forms. They are like some kind of endomorphisms of the modular curve, but this is not exactly true and they are only correspondences (in fact, they are endomorphisms of the jacobian).
Let now D be the free abelian group generated by the elements of L.

Definition 7.7. *We define $T(n) : D \to D$ as the sum of all sublattices of Λ of index n.*

$$T(n)[\Lambda] = \sum_{(\Lambda:\Lambda')=n} [\Lambda']$$

We also define the operator $R(n)$.

$$R(n)[\Lambda] = [n\Lambda]$$

Proposition 7.18. *a) Let m, n be coprime positive integers. Then $T(m) \circ T(n) = T(n) \circ T(m) = T(mn)$.*

b) Let p be a prime number and $n \geq 1$. Then

$$T(p^n) \circ T(p) = T(p^{n+1}) + pR(p) \circ T(p^{n-1})$$

Proof.

$$T(m) \circ T(n)[\Lambda] = T(m)\big(\sum_{\Lambda:\Lambda'=n} [\Lambda']\big) = \sum_{\Lambda:\Lambda'=n, \Lambda':\Lambda''=m} [\Lambda'']$$

But each lattice Λ'' of index mn admits a unique chain of sublattices $\Lambda \supset \Lambda' \supset \Lambda''$ where Λ' is of index n in Λ.
For the second part, note that:

a) $T(p^n) \circ T(p)[\Lambda]$ is the sum over all lattices Λ'' such that there exists another lattice Λ' such that $(\Lambda : \Lambda') = p, (\Lambda' : \Lambda'') = p^n$.

b) On the other side, $T(p^{n+1})[\Lambda]$ is the sum of all the lattices Λ'' with $(\Lambda : \Lambda') = p^{n+1}$.

c) Note also that $pR(p) \circ T(p^{n-1})$ is p times the sum over all lattices Λ' with $(\Lambda : \Lambda') = p^{n-1}$ of $R(p)[\Lambda']$.

Fix a lattice Λ'' of index p^{n+1} in Λ and count how many times occurs in the first sum (call it a) and b the number of times it appears in the last expression. Our objective is to prove that $a = 1 + pb$. If Λ'' is not in $p\Lambda$, then $b = 0$ and a is the number of lattices Λ' that contain Λ'' of index p in Λ. It is not difficult to check that there is only one such lattice. When $\Lambda'' \subset p\Lambda$, $b = 1$ and every lattice Λ' of index p contains $p\Lambda$; the problem is to count the number of subgroups of $\Lambda/p\Lambda$ of index p, i.e., the number of lines through te origin in the $\mathbb{F}_p$ plane, that is, $p + 1$. □

Corollary 7.2. *For any m, n,*

$$T(m) \circ T(n) = \sum_{d|\,\gcd(m,n), d>0} dR(d) \circ T(mn/d^2)$$

The importance of this operators is that if $F : L \to \mathbb{C}$ is a function of weight $2k$, then $T(n) \cdot F$ is again of weight $2k$.

We can now define Hecke operators for $\Gamma(1)$. For that, we note that there is a correspondence between functions F on L of weight $2k$ and functions f on $\mathbb{H}$ that are weakly modular of weight $2k$:

$$F(\Lambda(\omega_1, \omega_2)) = \omega_2^{-2k} f(\omega_1/\omega_2)$$

$$f(z) = F(\Lambda(z, 1))$$

We need first an easy lemma concerning 2×2 matrices.

Lemma 7.5. *Let $A \in M_2(\mathbb{Z})$ and determinant n. There exists an invertible matrix U in $M_2(\mathbb{Z})$ such that $U \cdot A = \begin{pmatrix} a & b \\ 0 & d \end{pmatrix}$ with $ad = n, a \geq 1, 0 \leq b < d$ (and moreover the integers a, b, d are uniquely determined).*

Proof. We put A into upper triangular form using just elementary operations: add a multiple of one row to another and swap two rows (the proof of this is just routine). This operations are invertible in $M_2(\mathbb{Z})$ since the corresponding matrices have determinant 1. These operations do not change the gcd of one column, so a is uniquely determined (it is the gcd of the first column), and d is the number such that $ad = n$. Once we have this it is clear that b is also fixed modulo n. □

With the correspondence between functions on L of weight $2k$ and weakly modular function of weight $2k$ in $\mathbb{H}$ in mind, we will define $T(n) \cdot f(z)$ to be the function associated with $T(n) \cdot F$, but multiplying by a factor n^{2k-1} to have integer coefficients. Therefore,

$$T(n) \cdot f(z) = n^{2k-1} \cdot (T(n)F)(\Lambda(z, 1))$$

If we want a more explicit formula, we can put the following sum over all the a, b, d satisfying the condition of the previous lemma

$$T(n) \cdot f(z) = n^{2k-1} \sum d^{-2k} f\Big(\frac{az + b}{d}\Big)$$

Proposition 7.19. *If f is a weakly modular form of weight $2k$ for $\Gamma(1)$, then $T(n) \cdot f$ is also weakly modular of the same weight $2k$, and it verifies these properties (note the analogy with the previously defined Hecke operators for lattices):*

a) $T(m) \cdot T(n) \cdot f = T(mn) \cdot f$, when $(m, n) = 1$.

b) $T(p) \cdot T(p^n) \cdot f = T(p^{n+1}) \cdot f + p^{2k-1} T(p^{n-1}) \cdot f$, when p prime and $n > 0$.

c) Let f be a modular form of weight $2k$ for $\Gamma(1)$ with Fourier expansion $f = \sum_{m\geq 0} c(m)q^m$, where as usual $q = e^{2\pi iz}$. Then $T(n)\cdot f$ is a modular form satisfying $T(n)\cdot f(z) = \sum_{m\geq 0}\gamma(m)q^m$, where $\gamma(m) = \sum_{a|\,\gcd(m,n), a\geq 1} a^{2k-1}c(\frac{mn}{a^2})$ (the Fourier coefficients will be denoted by $c(m)$ when it is clear the function associated to them and by $c_m(f)$ elsewhere).

Proof. We begin with the last item.

$$T(n)\cdot f(z) = n^{2k-1}\sum_{ad=n,a>0}\sum_{b=0}^{d-1} d^{-2k} f\Big(\frac{az+b}{d}\Big) =$$

$$= n^{2k-1}\sum_{ad=n,a>0}\sum_{b=0}^{d-1} d^{-2k}\Big(\sum_{m\geq 0} c(m)e^{2\pi im(\frac{az+b}{d})}\Big) =$$

$$= n^{2k-1}\sum_{ad=n,a>0} d^{-2k+1}\Big(\sum_{m'\geq 0} c(m'd)e^{2\pi im'az}\Big) =$$

$$= \sum_{m''\geq 0}\Big(\sum_{a|(n,m'')} a^{2k-1}c\Big(\frac{m''n}{a^2}\Big)\Big)q^{m''}$$

We prove now that T is weakly multiplicative, i.e., that $T(m)\cdot T(n) = T(mn)$ when $(m,n)=1$. It is enough with showing that the Fourier coefficients coincide. The r-th Fourier coefficient of $T(mn)\cdot f(z)$ is

$$\sum_{a|(mn,r)} a^{2k-1}c\Big(\frac{mnr}{a^2}\Big)$$

On the other hand, the r-th Fourier coefficient of $T(m)\cdot T(n)\cdot f(z)$ is

$$\sum_{e|(r,m)} e^{2k-1}c_{rm/e^2}(T(m)\cdot f(z)) = \sum_{e|(r,m)}\sum_{d|(m,rn/e^2)} d^{2k-1}e^{2k-1}c\Big(\frac{rmn}{e^2d^2}\Big) =$$

$$= \sum_{h|(mn,r)} h^{2k-1}c\Big(\frac{rmn}{h^2}\Big)$$

where we have written $c_i(g)$ to denote the i-th Fourier coefficient of g.
We finally prove b, that is quite technical but there are no great ideas, just algebraic manipulations (we will sum over a set of representatives, as it would be expected):

$$T(p^n)\cdot f(z) = p^{n(k-1)}\sum_{0\leq i\leq n} p^{-2ik}\sum_{0\leq b<p} f\Big(\frac{p^{n-i}z+b}{p^i}\Big)$$

$$T(p)\cdot g(z) = p^{2k-1}g(pz) + p^{-1}\sum_{0\leq b<p} g\Big(\frac{z+b'}{p}\Big)$$

Combining these two equalities,

$$T(p)\cdot T(p^n)\cdot f(z) = p^{(n+1)(2k-1)}\sum_{0\leq i\leq n} p^{-2ik}\sum_{0\leq b<p} f\Big(\frac{p^{n+1-i}z+b}{p^i}\Big)+$$

$$+p^{-1}p^{n(2k-1)}\sum_{0\leq b'<p}\sum_{0\leq i\leq n}p^{-2ik}\sum_{0\leq b<p^i}f\Big(\frac{p^{n-i}(z+b')+pb}{p^{i+1}}\Big)$$

Consider now, in the second summand, the case $i=n$; there, what we have is

$$p^{-1-n}\sum_{0\leq b'<p}\sum_{0\leq b<p^n}f\Big(\frac{z+b'+pb}{p^{n+1}}\Big)=p^{-1-n}\sum_{0\leq b\leq p^{n+1}}f\Big(\frac{z+b}{p^{n+1}}\Big)$$

This term should be now added to the expression obtained for $T(p^n)\cdot f(z)$ and we obtain that way $T(p^{n+1})\cdot f(z)$. For the rest of the terms, we will see now. For each i, the set $\{b+p^{n-1-i}b' : 0\leq b<p^i, 0\leq b'<p\}$ has p^{i+1} numbers, having representatives of all classes modulo p^i, p times each one. Since $f(z+1)=f(z)$, we always obtain the same value, and therefore that is equal to

$$p^{n(2k-1)}\sum_{0\leq i\leq n-1}p^{-2ki}\sum_{0\leq b<p}f\Big(\frac{p^{n-1-i}z+b}{p^i}\Big)=p^{k-1}T(p^{n-1})\cdot f(z)$$

as we wanted. □

For instance, it is clear now that $\gamma(0)=\sigma_{2k-1}(n)\cdot c(0)$ (where as usual $\sigma_i(n)=\sum_{d|n}d^i$), and $\gamma(1)=c(n)$. That is, the Fourier coefficient gave us information about arithmetic properties. Another virtue of this operator is that it preserves cusp forms, or said in another way, the $T(n)'s$ act on the spaces $M_k(\Gamma(1))$ and $S_k(\Gamma(1))$.
Let us now look further: beyond this long computations, we are now in a situation where is natural to define the Hecke algebra H as the algebra generated by the Hecke operators T_n for all $n\geq 1$. The two previous results show that H is a commutative algebra, generated by the operators T_p, where p is a prime. The key fact will be that all these operators would be self adjoint with respect to the Petersonn inner product. We need a lemma before proving that:

Lemma 7.6. *Let $f\in S_{2k}(\Gamma)$. Then, f satisfies the bound $|f(z)|\leq C|\Im(z)|^{-k}$ for all $z\in\mathbb{H}$. Furthermore, $|c_n(f)|=O(n^k)$*

Theorem 7.5. *$T(n)$ is a self-adjoint operator in $(S_{2k}(\Gamma),\langle\cdot,\cdot\rangle)$.*

Proof. We start by introducing some notation. If $\alpha\in\mathrm{GL}_2(\mathbb{R})^+$, and f is a function on $\mathbb{H}$, then $f|_k\alpha=(\det\alpha)^k(cz+d)^{-2k}f(\frac{az+b}{cz+d})$. It is easy to check that f is weakly modular of weight $2k$ for Γ if and only if $f|_k\alpha=f$ when $\alpha\in\Gamma$. Writing with this new notation the definition of the Hecke operators,

$$T(n)\cdot f(z)=\sum n^{k-1}f|_k\alpha$$

A first claim here is that $\langle f|_k\alpha,g|_k\alpha\rangle=\langle f,g\rangle$.
To see that, consider $\omega(f,g)=f(z)\bar{g}(z)y^{k-2}dxdy$ and prove that $\omega(f|_k\alpha,g|_k\alpha)=\alpha^*\omega(f,g)$. Since multiplication of α by an scalar does not change any of the sides, assume that $\det\alpha=1$ and therefore

$$f|_k\alpha=(cz+d)^{-2k}f(\alpha z)\text{ and also }\bar{g}|_k\alpha=(c\bar{z}+d)^{-2k}\overline{g(\alpha z)}$$

We know that $\alpha^*(dx \cdot dy) = dx \cdot dy/|cz+d|^4$ and combining all these facts

$$\alpha^*(\omega(f,g)) = f(\alpha z) \cdot \overline{g(\alpha z)} \cdot |cz+d|^{4-4k} y^{2k-2} \cdot |cz+d|^{-4} dxdy = \omega(f|_k\alpha, g|_k\alpha)$$

This tells us that

$$\int\int_D \omega(f|_k\alpha, g|_k\alpha) = \int\int_{\alpha D} \omega(f,g)$$

which is equal to $\langle f, g\rangle$ if αD is also a fundamental domain for $\Gamma(1)$, but that is not true. However, we can take a sufficiently small congruence subgroup Γ such that $\alpha\Gamma\alpha^{-1} \subset \Gamma(1)$; if D is a fundamental domain for Γ, what we have is that αD is a fundamental domain for $\alpha\Gamma\alpha^{-1}$ that has the same volume as D and so by our choice of Γ, both f, g are modular with respect to $\alpha\Gamma\alpha^{-1}$.

This lemma implies that

$$\langle f|_k\alpha, g\rangle = \langle f, g|_k\alpha^{-1}\rangle$$

for all $\alpha \in \mathrm{GL}_2(\mathbb{R})^+$. Since the Hecke algebra is generated by $T(p)$ it is enough with proving that

$$\langle T(p)f, g\rangle = \langle f, T(p)g\rangle$$

But this is easier if we assume the following straightforward lemma:

Lemma 7.7. *Let $M(n)$ be, as usual, the set of integer matrices with determinant n. There exists a common set of representatives $\{\alpha_i\}$ for the set of left orbits $\Gamma(1)\backslash M(p)$ and for the set of right orbits $M(p)/\Gamma(1)$.*

Then,

$$\langle T(p)f, g\rangle = p^{k-1}\sum_i \langle f|_k\alpha_i, g\rangle = p^{k-1}\sum_i \langle f, g|_k\alpha_i^{-1} =$$

$$= p^{k-1}\sum_i \langle f, g|_k\alpha_i'\rangle = \langle f, T(p)g\rangle$$

□

Using the results of linear algebra we quote, together with the commutativity of the Hecke operators, we will have the following:

Corollary 7.3. *For each $k \geq 1$ there is an orthonormal basis of $S_{2k}(\Gamma)$ of eigenfunctions of all Hecke operators, $T(n) \cdot f = \lambda(n)f$, where $\lambda(n) \in \mathbb{R}$. These eigenfunctions are called Hecke autoforms (or eigenforms).*

Proposition 7.20. *Let $f \in M_{2k}(\Gamma)$ be a Hecke autoform such that $T(n) \cdot f = \lambda(n)f$ for all $n \in \mathbb{N}$. If $k > 0$ we have that $c_1(f) \neq 0$ and if $k = 0$ then $c_k(f) = 0$ and f is constant. If $k > 0$ and $c_1(f) = 1$ (we will say that f is normalized) then we have:*

a) $c_n(f)c_m(f) = c_{nm}(f)$ if $(m,n) = 1$.

b) $c_p(f)c_{p^n}(f) = c_{p^{n+1}}(f) + p^{2k-1}c_{p^{n-1}}(f)$

Proof. The coefficient of q in $T(n)\cdot f$ is $c(n)$ but at the same time is $\lambda(n)c(1)$, so $c(n)=\lambda(n)\cdot c(1)$, so if $c(1)$ were 0 all the other coefficients would be zero, so f would be constant. From the relations between the coefficients, it is straightforward knowing that these relations hold for the eigenvalues, and since $c(n)=\lambda(n)$ we are done. □

Corollary 7.4. *Two cuspidal normalized eigenforms are either orthogonal or equal.*

Proof. Let f,g such that $T(n)\cdot f=\lambda(n)\cdot f, T(n)\cdot g=\mu(n)\cdot g$ for all n. Then

$$\lambda(n)\langle f,g\rangle=\langle T(n)\cdot f,g\rangle=\langle f,T(n)\cdot g\rangle=\mu(n)\langle f,g\rangle$$

We only have two options: or $\langle f,g\rangle=0$ (and in that case are orthogonal) or $\lambda(n)=\mu(n)$ for each n, so $c_n(f)=c_n(g)$. □

Integral structure on the space of modular functions

We have already introduced the Einsenstein series

$$G_{2k}(z)=\sum_{(m,n)\neq(0,0}\frac{1}{(mz+n)^{2k}}$$

Its values at $q=0$ are $2\zeta(2k)$, so we will divide for this factor to obtain the so called normalized Eisenstein series, $E_{2k}(z)=G_{2k}(z)/2\zeta(2k)$.

Proposition 7.21. *The Eisenstein series G_k ($k\geq 2$) is an eigenform of $T(n)$ with eigenvalue $\sigma_{2k-1}(n)$.*

Proof. Note that $M_{2k}=S_{2k}\oplus\langle G_k\rangle$ and since $T(n)$ is Hermitian and preserves S_{2k}, $T(n)\cdot G_k$ should be a multiple of G_k. To find the eigenvalue some manipulations are required. □

A useful fact is that $\oplus_k M_{2k}(\mathbb{Z})$ is a $\mathbb{Z}$-structure on $M_k(\Gamma(1))$ (in a vector space V over $\mathbb{C}$ a $\mathbb{Z}$-structure is a $\mathbb{Z}$-submodule V_0 free of rank equal the dimension of V): since we already know that $\oplus_k M_{2k}(\mathbb{C})=\mathbb{C}[E_4,E_6]$ it will be enough with showing that $\oplus_k M_{2k}(\mathbb{Z})=\mathbb{Z}[E_4,E_6]$. $E_4(z),E_6(z)$ and $\Delta'=q\prod(1-q^n)^{24}$ have integer coefficients, and an easy induction shows that $M_{2k}(\mathbb{Z})$ is the $2k$-th graded piece of $\mathbb{Z}[E_4,E_6]$. Then, given $f(z)=\sum a_nq^n$, with a_n integers,

$$f=a_0E_4^a\cdot E_6^b+\Delta\cdot g$$

where $4a+6b=2k$ and $g\in M_{k-12}$. Then $a_0\in\mathbb{Z}$ and it can be seen that $g\in M_{2k-12}(\mathbb{Z})$.

We finish with an important result. Since $M_k(\mathbb{Z})$ is stabilized by $T(n)$, the matrix of $T(n)$ with respect to a basis for $M_k(\mathbb{Z})$ has integer coefficients, and so the eigenvalues of $T(n)$ are algebraic integers.

Proposition 7.22. *The eigenvalues of the Hecke operators are algebraic integers.*

7.5 Interpretation of Hecke Operators

In the previous section we have introduced a particular type of operators, we checked that they have some good properties but, however, everything we did was quite restrictive (only of $\Gamma(1)$) and there was not a clear idea of the intuition behind that. In this section we try to give a more accurate interpretation of the Hecke operators.
We begin by introducing a notation that is very useful and that we have still not used. For any matrix $\gamma \in \mathrm{SL}_2(\mathbb{Z})$, define the factor of automorphy $j(\gamma, \tau) \in \mathbb{C}$, for any $\tau \in \mathbb{H}$ as

$$j(\gamma, \tau) = c\tau + d$$

For an integer k, define now the weight k-operator $[\gamma]_k$ on functions $f : \mathbb{H} \to \mathbb{C}$ as

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

With these notations, a weakly modular function of weight k with respect to Γ is just a meromorphic functions such that $f[\gamma]_k = f$ for all $\gamma \in \Gamma$. Some straightforward properties are the following:

Proposition 7.23. *Let $\gamma, \gamma' \in \mathrm{SL}_2(\mathbb{Z})$ and $\tau \in \mathbb{H}$, we have that:*

a) $j(\gamma\gamma', \tau) = j(\gamma, \gamma'(\tau)) j(\gamma', \tau)$.

b) $[\gamma\gamma']_k = [\gamma]_k [\gamma']_k$.

c) $\Im(\gamma(\tau)) = \frac{\Im(\tau)}{|j(\gamma,\tau)|^2}$.

d) $\frac{d\gamma(\tau)}{d\tau} = \frac{1}{j(\gamma,\tau)^2}$.

Fix now a group of finite index in $\Gamma(1)$, and let $\alpha \in \mathrm{GL}_2(\mathbb{R})^+$, so α defines a map $x \mapsto \alpha x$ from $\mathbb{H}$ to itself, and we would like to define $\alpha : \Gamma\backslash\mathbb{H} \to \Gamma\backslash\mathbb{H}, \Gamma z \mapsto \alpha\Gamma z$, and this cannot obviously be done, since Γ is not normal in $\mathrm{GL}_2(\mathbb{R})$. In fact, $\alpha\Gamma z$ is not even a Γ-orbit, and what will interest us are orbits of the form $\Gamma\alpha\Gamma z$. Every left or right coset of Γ in $\mathrm{GL}_2(\mathbb{R})^+$ that meets $\Gamma\alpha\Gamma$ is contained in it, so

$$\Gamma\alpha\Gamma = \bigcup \Gamma\alpha_i$$

where the $\cup$ is a disjoint union. That way, α can define a many valued map from $\Gamma\backslash\mathbb{H}$ to itself sending Γz to $\{\Gamma\alpha_i z\}$. But this is totally illicit, since talking about many-valued maps sounds like an ad-hoc trick. We first justify that the maps are finite:

Lemma 7.8. *Let $\alpha \in \mathrm{GL}_2(\mathbb{R})^+$. Then $\Gamma\alpha\Gamma$ is a finite union of right (left) cosets if and only if α is a scalar multiple of a matrix with integer coefficients.*

Lemma 7.9. *Let $\alpha \in \mathrm{GL}_2(\mathbb{R})^+$. If we write*

$$\Gamma = \bigcup (\Gamma \cap \alpha^{-1}\Gamma\alpha)\beta_i$$

where the $\cup$ is a disjoint union, then

$$\Gamma\alpha\Gamma = \bigcup \Gamma\alpha_i$$

where $\alpha_i = \alpha\beta_i$ and the $\cup$ is again a disjoint union.

In a more general way, we can consider operators between $M_k(\Gamma_1)$ and $M_k(\Gamma_2)$. We omit some of the proofs of the most technical results.

Lemma 7.10. *Let Γ be a congruence subgroup of $\mathrm{SL}_2(\mathbb{Z})$ and let α be an element of $\mathrm{GL}_2^+(\mathbb{Q})$. Then, $\alpha^{-1}\Gamma\alpha \cap \mathrm{SL}_2(\mathbb{Z})$ is a congruence subgroup of $\mathrm{SL}_2(\mathbb{Z})$.*

Lemma 7.11. *Let Γ_1, Γ_2 be congruence subgroups of $\mathrm{SL}_2(\mathbb{Z})$ and let α be an element of $\mathrm{GL}_2^+(\mathbb{Q})$. Let $\Gamma_3 = \alpha^{-1}\Gamma_1\alpha \cap \Gamma_2$, a subgroup of Γ_2. Then, left multiplication by α*

$$\Gamma_2 \to \Gamma_1\alpha\Gamma_2 \text{ given by } \gamma_2 \mapsto \alpha\gamma_2$$

induces a natural bijection from the coset space $\Gamma_3\backslash\Gamma_2$ and the orbit space $\Gamma_1\backslash\Gamma_1\alpha\Gamma_2$. Furthermore, $\{\gamma_{2,j}\}$ is a set of coset representatives for $\Gamma_3\backslash\Gamma_2$ if and only if $\{\beta_j\} = \{\alpha\gamma_{2,j}\}$ is a set of orbit representatives for $\Gamma_1\backslash\Gamma_1\alpha\Gamma_2$.

Definition 7.8. *Let Γ_1, Γ_2 be congruence subgroups of $\mathrm{SL}_2(\mathbb{Z})$ and let $\alpha \in \mathrm{GL}_2^+(\mathbb{Q})$. The weight-$k$ $\Gamma_1\alpha\Gamma_2$ operators takes functions $f \in M_k(\Gamma_1)$ to*

$$f[\Gamma_1\alpha\Gamma_2]_k = \sum_j f[\beta_j]_k$$

where the $\{\beta_j\}$ are orbit representatives: $\Gamma_1\alpha\Gamma_2 = \cup_j \Gamma_1\beta_j$.

We would like to check that this double coset operator is independent of how the β_j are chosen and that for each $f \in M_k(\Gamma_1)$, $f[\Gamma_1\alpha\Gamma_2]_k$ is Γ_2-invariant and holomorphic at the cusps. We can also see that this operators carry $S_k(\Gamma_1)$ to $S_k(\Gamma_2)$.
We first note that if $\beta = \gamma_1\alpha\gamma_2$ and $\beta' = \gamma_1'\alpha\gamma_2'$ represent the same orbit, then $\alpha\gamma_2 \in \Gamma_1\alpha\gamma_2'$ and using now that f is weight-k invariant under Γ_1, we have that $f[\beta]_k = f[\beta']_k$.
For the invariance, note that γ_2 permutes the orbit space $\Gamma_1\backslash\Gamma_1\alpha\Gamma_2$ by right multiplication. Thus,

$$(f[\Gamma_1\alpha\Gamma_2]_k)[\gamma_2]_k = \sum_j f[\beta_j\gamma_2]_k = f[\Gamma_1\alpha\Gamma_2]_k$$

We show now holomorphy at the cusps: note first that if $f \in M_k(\Gamma_1)$, for any $\gamma \in \mathrm{GL}_2^+(\mathbb{Q})$, the function $g = f[\gamma]_k$ is holomorphic at infinity, so it has a Fourier expansion

$$g(\tau) = \sum_{n\geq 0} a_n(g)e^{2\pi i n\tau/h}$$

for some $h \in \mathbb{Z}^+$. Also note that the sum of a finite number of these holomorphic functions is holomorphic, and for any $\delta \in \mathrm{SL}_2(\mathbb{Z})$, the function $(f[\Gamma_1\alpha\Gamma_2]_k)[\delta]_k$ is a sum of functions $g_j = f[\gamma_j]_k$ (where $\gamma_j = \beta_j\delta$), so it is holomorphic at infinity. A similar reasoning yields to the fact that it preserves the space of cusp forms.
We comment three special case of the double coset operator $[\Gamma_1\alpha\Gamma_2]_k$:

a) $\Gamma_1 \supset \Gamma_2$. In that case $M_k(\Gamma_1) \subset M_k(\Gamma_2)$ and taking $\alpha = I$ the double coset operator is simply $f[\Gamma_1\alpha\Gamma_2]_k = f$ and what we have is simply the inclusion between the two spaces.

b) Γ_1 and Γ_2 are conjugates, and take α such that $\alpha^{-1}\Gamma_1\alpha = \Gamma_2$. Here the double coset operator is $f[\alpha]_k$, the natural translation (isomorphism) between $M_k(\Gamma_1)$ and $M_k(\Gamma_2)$.

c) $\Gamma_1 \subset \Gamma_2$. Taking $\alpha = I$ and letting $\{\gamma_{2,j}\}$ be a set of coset representatives for $\Gamma_1\backslash\Gamma_2$ makes the double coset operator $f[\Gamma_1\alpha\Gamma_2]_k = \sum_j f[\gamma_{2,j}]_k$ the natural trace map that projects $M_k(\Gamma_1)$ onto its subspace $M_k(\Gamma_2)$ (it is a surjection).

It comes as no surprise that any double coset operator is a composition of these (as a linear application is a composition of an injection, a bijection and a surjection). Given $\Gamma_1, \Gamma_2, \alpha$, set $\Gamma_3 = \alpha^{-1}\Gamma_1\alpha \cap \Gamma_2$ and let $\Gamma_3' = \alpha\Gamma_3\alpha^{-1} = \Gamma_1 \cap \alpha\Gamma_2\alpha^{-1}$. Then, $\Gamma_1 \supset \Gamma_3'$, $\alpha^{-1}\Gamma_3'\alpha = \Gamma_3$ and $\Gamma_3 \subset \Gamma_2$, giving the three cases. The corresponding composition is

$$f \mapsto f \mapsto f[\alpha]_k \mapsto \sum_j f[\alpha\gamma_{2,j}]_k$$

which is the general $[\Gamma_1\alpha\Gamma_2]_k$.

This also admits a **geometric interpretation**: we are transferring points back between the corresponding modular curves. To see this in a more precise way, recall that every congruence subgroup Γ has a modular curve $X(\Gamma) = \Gamma\backslash\mathbb{H}^*$ consisting of orbits $\Gamma\tau$. In this sense, note that what we will have is a map between $\mathrm{Div}(X_2)$ and $\mathrm{Div}(X_1)$ and since it behaves well with respect to principal divisor what we really have with Hecke operators is a map between the jacobians of the corresponding modular curves.

A few remarks: the map $\alpha : X_3 \to X_3'$ given by $\Gamma_3\tau \mapsto \Gamma_3'\alpha(\tau)$ is well defined. As usual, we put $\Gamma_3\backslash\Gamma_2 = \cup_j\Gamma_3\gamma_{2,j}$ and $\beta_j = \alpha\gamma_{2,j}$ so that $\Gamma_1\alpha\Gamma_2 = \cup_j\Gamma_1\beta_j$. Call π_2, π_1 the projections of X_3 in X_2 and X_3' in X_1 respectively. Then, each point of X_2 is taken back by $\pi_1 \circ \alpha \circ \pi_2^{-1}$ to a set of points of X_1. π_2^{-1} takes a point $x \in X_2$ to the multiset of overlying points $y \in X_3$: $\pi_2^{-1}(x) = \{e_y y : y \in X_3, \pi_2(y) = x\}$. Summing up, what we have is a map between $\mathrm{Div}(X_2)$ and $\mathrm{Div}(X_1)$, as we previously announced.

We are going to analyze now in this framework the T_p operators and also the so called diamond operators.

To define this first type of operators, take any $\alpha \in \Gamma_0(N)$ and set $\Gamma_1 = \Gamma_2 = \Gamma_1(N)$. Take now $[\Gamma_1\alpha\Gamma_2]_k$. Since $\Gamma_1(N)$ is a normal subgroup of $\Gamma_0(N)$ we have an operator of the second case in the list above (conjugation). We are translating each function $f \in M_k(\Gamma_1(N))$ to $f[\alpha]_k$. We have so that $\Gamma_0(N)$ acts on $M_k(\Gamma_1(N))$ and since its subgroup $\Gamma_1(N)$ acts trivially, this is really an action of the quotient, that can be identified with $(\mathbb{Z}/n\mathbb{Z})^*$. The action of a generic matrix α is determined by the d element (modulo N):

$$\langle d \rangle : M_k(\Gamma_1(N)) \to M_k(\Gamma_1(N))$$

and is given by $\langle d\rangle f = f[\alpha]_k$ for any α with its $(2,2)$ position congruent with d modulo N. This operator is also called diamond operator. It is important to note that for any character $\chi : (\mathbb{Z}/N\mathbb{Z})^* \to \mathbb{C}$ the space

$$M(N,\chi) = \{f \in M_k(\Gamma_1(N)) : f[\gamma]_k = \chi(d_\gamma)f \text{ for all } \gamma \in \Gamma_0(N)\}$$

(d_γ the $(2,2)$-position) is just the χ-eigenspace of the diamond operator, that is, those functions such that $\langle d\rangle f = \chi(d)f$. Observe that

$$M_k(\Gamma_1(N)) = \bigoplus_\chi M_k(N,\chi)$$

The second kind of Hecke operators, already commented for the particular case of $\mathrm{SL}_2(\mathbb{Z})$ occurs also when $\Gamma_1 = \Gamma_2 = \Gamma_1(N)$, and now α is the diagonal matrix with 1 and p along the diagonal (p prime). The double coset is here

$$\Gamma_1(N)\begin{pmatrix} 1 & 0 \\ 0 & p \end{pmatrix}\Gamma_1(N) = \{\gamma \in M_2(\mathbb{Z}) \mid \gamma \equiv \begin{pmatrix} 1 & * \\ 0 & p \end{pmatrix} \pmod N, \det\gamma = p\}$$

It is not difficult to verify that $\langle d\rangle T_p f = T_p\langle d\rangle f$.
Working carefully with the expressions we also obtain expression for T_p:

Proposition 7.24. *Let $N \in \mathbb{Z}^+$, and let $\Gamma_1 = \Gamma_2 = \Gamma_1(N)$. Let $\alpha = \begin{pmatrix} 1 & 0 \\ 0 & p \end{pmatrix}$ where p is a prime. The operator $T_p = [\Gamma_1\alpha\Gamma_2]_k$ is given by:*

$$T_p f = \sum_{j=0}^{p-1} f\Big[\begin{pmatrix} 1 & j \\ 0 & p \end{pmatrix}\Big]_k \quad \textit{if } p|N$$

$$T_p f = \sum_{j=0}^{p-1} f\Big[\begin{pmatrix} 1 & j \\ 0 & p \end{pmatrix}\Big]_k + f\Big[\begin{pmatrix} m & n \\ N & p \end{pmatrix}\begin{pmatrix} p & 0 \\ 0 & 1 \end{pmatrix}\Big]_k \quad \textit{if } p \nmid N; mp - nN = 1$$

As it occurred in the simplest case, we now have $T_pT_q = T_qT_p$ so we can define T_n, when n is square free, as the product of the T_p, where p are the prime factors of n. For a prime power, define

$$T_{p^r} = T_pT_{p^{r-1}} - p^{k-1}\langle p\rangle T_{p^{r-2}}$$

and again extend by multiplicativity to all n.

7.6 *L*-series attached to modular forms

Consider $\sum c(n)q^n$, a cusp form that is a normalized eigenfunction for all Hecke operators $T_{2k}(n)$. We have already pointed out the multiplicative relations between the coefficients

$$c(m)c(n) = c(mn), (m,n) = 1$$

$$c(p^e)c(p) = c(p^{e+1}) + p^{2k-1}c(p^{e-1})$$

We will prove now that this is equivalent to an Euler product decomposition for the Dirichlet series attached to f; we introduce here again a mysterious ad-hoc definition for the L-series, that will seem natural after a few chapters:

Definition 7.9. *For any power series $f = \sum_{n\geq 1} c(n)q^n$, the L-series attached to f is the Dirichlet series*

$$L(f,s) = \sum_{n\geq 1} c(n)n^{-s}$$

For the moment, we will see that L-series as a formal series, without dealing with convergence issues.

Proposition 7.25. *Let $f = \sum_{n\geq 1} c(n)q^n$ be a power series where $c(1) = 1$. Then, the coefficients of f satisfy*

$$c(m)c(n) = c(mn), (m,n) = 1$$

$$c(p^e)c(p) = c(p^{e+1}) + p^{2k-1}c(p^{e-1})$$

if and only if the associated $L-$ series has the Euler product expansion

$$L(f,s) = \prod_p \frac{1}{1 - c(p)p^{-s} + p^{2k-1-2s}}$$

Proof. We will begin by proving that the relation between the coefficients imply the Euler product. Note that, as with the usual Dirichlet function,

$$L(f,s)\prod_p \sum_{e\geq 0} c(p^e)p^{-es}$$

We will do now some algebraic manipulations bearing in mind that $c(1) = 1$:

$$(1 - c(p)p^{-s} + p^{2k-1-2s})(\sum_{e\geq 0} c(p^e)p^{-es}) =$$

$$= \sum_{e\geq 0} c(p^e)p^{-es} - \sum_{e\geq 0} c(p^e)c(p)p^{-e(s+1)} + c(p^e)p^{2k-1-(2+e)s} =$$

$$= (c(1) + c(p)p^{-s}) - (c(p)c(1)p^{-s}) + \sum_{e\geq 2}(c(p^e) - c(p)c(p^{e-1}) + c(p^{e-2}p^{2k-1})p^{-es} = 1$$

The result now follows (for the converse, just go back taking care of little modifications). □

For which concerns convergence, we recall here a few facts that are quite direct:

Proposition 7.26. *Let $f(\tau)$ be a cusp form of weight $2k$ with Fourier expansion $\sum c(n)q^n$. There exists a constant C that only depends on f such that*

$$|c(n)| \leq Cn^k \textit{ for all } n \geq 1$$

Corollary 7.5. *Let f be a cusp form of weight $2k$. Then the associated L-series converges to a holomorphic function in the upper half plane for $\Re(s) > k+1$. If f is not a cusp form, the L-series converges for all s with $\Re(s) > k$.*

We have now the following theorem, also due to Hecke, that asserts that is possible to do the analytic continuation of $L(f, s)$ to all of $\mathbb{C}$. For the case of elliptic curves, this will not be always obvious, so it would be a great advance to justify that the L-function of an elliptic curve is the same as the one of a certain modular form.

Proposition 7.27. *Let $f(\tau)$ be a cusp form of weight $2k$. Then:*

a) $L(f, s)$ has an analytic continuation to all of $\mathbb{C}$.

b) Let $R(s) = (2\pi)^{-s}\Gamma(s)L(f, s)$. Then

$$R(f, 2k - s) = (-1)^k R(f, s) \ s \in \mathbb{C}$$

7.7 Oldforms and newforms

From the beginning of this chapter, we have worked in a particular level N. Now we try to explain what happens when we move between levels, more precisely the relation between levels M and N when $M|N$. We begin by observing that when $M|N$, then $S_k(\Gamma_1(M)) \subset S_k(\Gamma_1(N))$. Another way to embed $S_k(\Gamma_1(M))$ in $S_k(\Gamma_1(N))$ is composing with the map consisting in multiplication by d, where d is a factor of N/M; to do this, let

$$\alpha_d = \begin{pmatrix} d & 0 \\ 0 & 1 \end{pmatrix}$$

and consider $(f[\alpha_d]_k)(\tau) = d^{k-1}f(d\tau)$ for any $f : \mathbb{H} \to \mathbb{C}$. It can be checked that this carries the level M to level N. We distinguish therefore the part of $S_k(\Gamma_1(N))$ that comes from lower levels.

Definition 7.10. *For each divisor d of N, define*

$$i_d : (S_k(\Gamma_1(Nd^{-1})))^2 \to S_k(\Gamma_1(N))$$

such that

$$(f, g) \mapsto f + g[\alpha_d]_k$$

The subspace of old forms is so

$$S_k(\Gamma_1(N))^{\text{old}} = \sum_{p|N} i_p((S_k(\Gamma_1(Np^{-1})))^2)$$

and the subspace of newforms (at level N) is the orthogonal complement with respect to the Petersson inner product (it is possible to change the sum and do it over all the divisors of N, not only primes, and nothing changes).

Proposition 7.28. *The subspaces $S_k(\Gamma_1(N))^{\text{old}}$ and $S_k(\Gamma_1(N))^{\text{new}}$ are stable under the Hecke operators T_n and $\langle n\rangle$ for all $n \in \mathbb{Z}^+$.*

Corollary 7.6. *The spaces $S_k(\Gamma_1(N))^{\text{old}}$ and $S_k(\Gamma_1(N))^{\text{new}}$ have orthogonal bases of eigenforms for the Hecke operators away from the level, $\{T_n, \langle n\rangle : (n, N) = 1\}$.*

Let now $M|N$ and let $d|(N/M)$, with $d > 1$. We have two important maps from $S_k(\Gamma_1(M))$ to $S_k(\Gamma_1(N))$. Inclusion and the weight-k operators $[\alpha_d]_k$ operator. Define now a variant of the map i_d, ι_d, again from $S_k(\Gamma_1(M))$ to $S_k(\Gamma_1(N))$ as

$$\iota_d = d^{1-k}[\alpha_d]_k, \quad (\iota_d f)(\tau) = f(d\tau)$$

If $f \in S_k(\Gamma_1(N))$ takes the form $f = \sum_{p|N} \iota_p f_p$ with $f_p \in S_k(\Gamma_1(N/p))$, and if the Fourier expansion of f is $f(\tau) = \sum a_n(f)q^n$, then $a_n(f) = 0$ for all n such that $(n, N) = 1$. The main lemma in the theory of new forms is that the converse holds. In fact, this result is referred as the Main Lemma.

Theorem 7.6. *If $f \in S_k(\Gamma_1(N))$ has Fourier expansion $f(\tau) = \sum a_n(f)q^n$ with $a_n(f) = 0$ when $(n, N) = 1$, then f is of the form $f = \sum_{p|N} \iota_p f_p$ with $f_p \in S_k(\Gamma_1(N/p))$.*

Write $S_2(N)$ for $S_2(\Gamma_0(N))$. We are going to define now $\mathbb{T}$ as the commutative subalgebra of $\mathrm{End}_{\mathbb{C}}(S_2(N))$ generated over $\mathbb{Z}$ by the Hecke operators T_n and $\mathbb{T}^0$ as the subalgebra generated by the operators T_n, where $(n, N) = 1$. We have the following result:

Proposition 7.29. *The Hecke algebras $\mathbb{T}$ and $\mathbb{T}^0$ are finitely generated as $\mathbb{Z}$-modules and its rank is g, the genus of $X_0(N)$.*

Proof. Let V be the vector space dual of $S_2(N)$, that is, the homomorphisms from $S_2(N)$ to $\mathbb{C}$. By the theory of Abel-Jacobi, $H^1(X_0(N)(\mathbb{C}), \mathbb{Z})$ is a sublattice Λ of V, just by associating to a closed cycle c on $X_0(N)$ the functional $\eta_c \in V$ that will be the integral of ω_f around c. The action of $\mathbb{T}$ on $S_2(N)$ induces an action on V which leaves stable Λ (we will turn over this on chapter ten). Hence, $\mathbb{T}$ is a subalgebra of the endomorphisms of Λ, and since this latter ring is finitely generated as a $\mathbb{Z}$-module the same will be true for $\mathbb{T}$ and for $\mathbb{T}^0$.

To see that the rank is at most g, we must consider the action of complex conjugation τ on $X_0(N)(\mathbb{C})$. It induces an action on Λ commuting with that of $\mathbb{T}$. Hence, $\mathbb{T}$ preserves the submodules Λ^+, Λ^- on which τ acts as multiplication by 1 and -1 respectively. Both of them are free of rank g and so $\mathbb{T}$ is identified with a commutative subalgebra of $M_g(\mathbb{Z})$. It is an algebraic fact that there exists $T \in \mathbb{T}$ such that $\mathbb{T}$ contains $\mathbb{Z}[T]$ with finite index, and hence these two rings have the same rank as $\mathbb{Z}$-modules; but $\mathbb{Z}[T]$ is generated by $1, T, \ldots, T^{g-1}$ (Cayley-Hamilton), so the result follows.

It remains to prove that the rank is exactly g. Take $T_{\mathbb{C}} = \mathbb{T} \otimes \mathbb{C}$, and so we have a $\mathbb{C}$-bilinear pairing

$$\langle , \rangle : T_{\mathbb{C}} \times S_2(N) \to \mathbb{C}$$

given by $\langle T, f \rangle = a_1(Tf)$ (a_i are now the Fourier coefficients). We see that $\langle T_n, f \rangle = a_n(f)$ and so the pairing is non-degenerate on the right and the natural map

$$S_2(N) \to \mathrm{Hom}(T_{\mathbb{C}}, \mathbb{C})$$

induced by $\langle , \rangle$ is injective. Hence, the complex dimension of $T_{\mathbb{C}}$ is greater or equal than g, and hence the rank of $\mathbb{T}$ is $\geq g$, as we wanted. □

We quote now some results that follow almost directly from our previous work.

Corollary 7.7. *$S_2(N)$ has a basis consisting of modular forms with integer Fourier coefficients.*

Proposition 7.30. *If T is in $\mathbb{T}^0$, then it is self-adjoint with respect to the Petersson scalar product.*

(this last fact is not immediate at all).
All this, combined with the spectral theorem, assures that

$$S_2(N) = \bigoplus_{\lambda} S^0_{\lambda}$$

taken over all $\mathbb{C}$-algebra homomorphisms $\lambda : \mathbb{T}^0 \to \mathbb{C}$, where S^0_λ denotes the corresponding eigenspace in $S_2(N)$. These spaces need not be of dimension one, but if $\lambda : \mathbb{T} \to \mathbb{C}$ is a ring homomorphism defined over the full Hecke algebra, and S_λ is the associated eigenspace, then:

Proposition 7.31. *The eigenspace S_λ attached to $\lambda : \mathbb{T} \to \mathbb{C}$ is one-dimensional*

But $S_2(N)$ does not decompose in general into direct sum of one dimensional eigenspace S_λ, since the operators in $\mathbb{T}$ need not act semisimply on $S_2(N)$. But the space of newforms does decompose as a direct sum of one dimensional eigenspaces under both the action of $\mathbb{T}$ and $\mathbb{T}^0$. We finish this section with a remarkable theorem of Atkin-Lehner.

Theorem 7.7. *Let $f \in S_2^{\text{new}}$ be a simultaneous eigenform for the action of $\mathbb{T}^0$. Let S be any finite set of prime numbers and $g \in S_2(N)$ an eigenform for T_p for all $p \notin S$. If $a_p(f) = a_p(g)$ for all $p \notin S$, then $g = \lambda f$ for some $\lambda \in \mathbb{C}$.*

An immediate corollary of this is the following:

Corollary 7.8. *The full Hecke algebra $\mathbb{T}$ acts semisimply on $S_2^{\text{new}}(N)$ with one dimensional eigenspaces. We therefore have an orthogonal decomposition*

$$S_2(N) = S_2^{\text{old}}(N) \bigoplus_{\lambda} \mathbb{C} f_\lambda$$

where the sum is over all algebra homomorphisms $\lambda : \mathbb{T} \to \mathbb{C}$ corresponding to eigenvectors in $S_2^{\text{new}}(N)$ and $f_\lambda(\tau) = \sum_{n=1}^{\infty} \lambda(T_n) e^{2\pi i n \tau}$.
The simultaneous eigenvector f_λ is called a normalized eigenform or a newform of level N (it satisfies $a_1(f) = 1$).

7.8 Eigenforms

Let us summarize what we have done until the moment: after having presented our place of work, $\mathbb{H}$, we have defined modular forms and study Hecke operators for the case of $\mathrm{SL}_2(\mathbb{Z})$. Then, we have done an interpretation of these operators in terms of jacobians and in that context it was natural to generalize the operators to other groups of finite index in $\mathrm{SL}_2(\mathbb{Z})$. There, we had to distinguish between two types of forms: newforms and oldforms. In $\Gamma(1)$, the Hecke operators were self-adjoint and they diagonalize simultaneously. In other congruence subgroups, we can repeat those procedure but with some changes; for instance we have the following:

Theorem 7.8. *In the inner product space $S_{2k}(\Gamma_1(N))$, the Hecke operators $\langle p\rangle$ and T_p for $p \nmid N$ have adjoints $\langle p\rangle^* = \langle p\rangle^{-1}$ and $T_p^* = \langle p\rangle^{-1}T_p$. Thus, the Hecke operators $\langle n\rangle$ and T_n for n relatively prime to N are normal.*

Note that the Petersson inner product is defined in the same way, with a normalizing factor to take into account the measure of the fundamental region, that will satisfy

$$\mathrm{Vol}(\Gamma) = 2\pi[\mathrm{SL}_2(\mathbb{Z}) : \{\pm\,\mathrm{Id}\}\Gamma]$$

Recall also that the spaces $S_{2k}(\Gamma_1(N))^{\mathrm{old}}$ and $S_{2k}(\Gamma_1(N))^{\mathrm{new}}$ have orthogonal bases of eigenforms for the Hecke operators $\{T_n, \langle n\rangle : (n, N) = 1\}$. Let f be such an eigenform; the Main Lemma can be used to deduce that if $f \in S_{2k}(\Gamma_1(N))^{\mathrm{new}}$ then f is an eigenform for all T_n and $\langle n\rangle$. We explore these ideas:

Definition 7.11. *A nonzero modular form $f \in M_{2k}(\Gamma_1(N))$ that is an eigenform for the Hecke operators T_n and $\langle n\rangle$ for all $n \in \mathbb{Z}^+$ is a Hecke eigenform (or simply eigenform). The eigenform $f(\tau) = \sum_{n=0}^{\infty} a_n(f)q^n$ is normalized when $a_1(f) = 1$. A newform is a normalized eigenform in $S_{2k}(\Gamma_1(N))^{\mathrm{new}}$.*

If $f \in S_{2k}(\Gamma_1(N))$ is an eigenform for the Hecke operators T_n and $\langle n\rangle$ with $(n, N) = 1$, then take the corresponding eigenvalues for each n, c_n and d_n in such a way that $T_n f = c_n f, \langle n\rangle = d_n f$. The map $n \mapsto d_n$ defines a Dirichlet character $\chi : (\mathbb{Z}/N\mathbb{Z})^* \to \mathbb{C}$ and so $f \in S_{2k}(N, \chi)$. Consequently, $a_n(f) = c_n a_1(f)$ when $(n, N) = 1$, and so if $a_1(f) = 0$ then $a_n(f) = 0$ for all n coprime with N and so $f \in S_{2k}(\Gamma_1(N))^{\mathrm{old}}$. Recall that $M_{2k}(N, \chi)$ denotes those $f \in M_{2k}(\Gamma_1(N))$ such that $f[\gamma]_k = \chi(d_\gamma)f$ for all $\gamma \in \Gamma_0(N)$ (where χ is a Dirichlet character modulo N).

Theorem 7.9. *Let $f \in S_{2k}(\Gamma_1(N))^{\mathrm{new}}$ be a nonzero eigenform for the Hecke operators T_n and $\langle n\rangle$ for all n with $(n, N) = 1$. Then,*

a) f is a Hecke eigenform, that is, an eigenform for T_n and $\langle n\rangle$ for all $n \in \mathbb{Z}^+$. A suitable scalar multiple of f is a newform.

b) If $\tilde{f}$ satisfies the same conditions as f and has the same T_n-eigenvalues, then $\tilde{f} = cf$ for some constant c.

The set of newforms in $S_{2k}(\Gamma_1(N))^{\mathrm{new}}$ is an orthogonal basis of the space. Each such newform lies in an eigenspace $S_{2k}(N, \chi)$ and satisfies $T_n f = a_n(f)f$ for all $n \in \mathbb{Z}^+$. That is, its Fourier coefficients are its T_n-eigenvalues.

We finish the section with a proposition where we see again the important role played by L-functions:

Proposition 7.32. *Let $f \in M_{2k}(N, \chi)$. Then the following conditions are equivalent:*

a) f is a normalized eigenform.

b) Its Fourier coefficients satisfy:

1. $c(1) = 1$.

2. $c(mn) = c(m)c(n)$ when $(m, n) = 1$.

3. $c(p)c(p^r) = c(p^{r+1}) + \chi(p)p^{2k-1}c(p^{r-1})$ for all prime p and $r \geq 1$.

c) $L(s, f)$ has an Euler product expansion

$$L(s,f) = \prod_p (1 - c(p)p^{-s} + \chi(p)p^{2k-1-2s})^{-1}$$

7.9 Modular curves as algebraic curves

We have seen that a complex elliptic curve could be described as an algebraic curve via the Weierstrass $\wp$-function. Let N be a positive integer. We have sketched how $X_0(N) = \Gamma_0(N)\backslash\mathbb{H}$, $X_1(N) = \Gamma_1(N)\backslash\mathbb{H}$ and $X(N) = \Gamma(N)\backslash\mathbb{H}$ can also be described as algebraic curves. The existence can be assured by a general theorem on Riemann surfaces, that states that any Riemann surface is isomorphic (as a Riemann surface) to an algebraic curve with complex coefficients. But here we want something stronger, since $X_0(N)$ and $X_1(N)$ are curves over the rational numbers. This will be the content of some of the theorems of chapter ten.
From the theory we have developed until now, we know that two elliptic curves over the complex numbers, $\mathbb{C}/\Lambda, \mathbb{C}/\Lambda'$ are holomorphically group isomorphic if and only if $m\Lambda = \Lambda', m \in \mathbb{C}$. It is natural to consider therefore an equivalence relation consisting in viewing two elliptic curves as the same if they are isomorphic. Similarly, we take in $\mathbb{H}$ the equivalence relation given by $\Gamma(1)$ and consider the fundamental domain. We will show in this section that there is a bijection between the two sets. We have to introduce first some terminology:

Definition 7.12. *A moduli problem over k is a contravariant functor F from the category of algebraic varieties over k to the category of sets. In particular, for each variety V over k we are given a set $F(V)$ and for each regular map $\phi : W \to V$, we are given a map $F(\phi) : F(V) \to F(W)$ (typically, $F(V)$ will be the set of isomorphism classes of certain objects over V.*

For instnce, the j-invariant would be a solution to the moduli problem of classifying elliptic curves over an algebraically closed field, since two curves are isomorphic if and only they have the same j-invariant.
Let N be now a positive integer. An enhanced elliptic curve for $\Gamma_0(N)$ is an ordered pair (E, C), where E is a complex elliptic curve and C is a cyclic group of E of order N. Two pairs (E, C) and (E', C') are equivalent if some isomorphims between E and E' takes C to C'. We will write $S_0(N)$ for the set of equivalence

classes and $[E, C]$.
In the same way, we define an enhanced elliptic curve for $\Gamma_1(N)$ to be a pair (E, Q), where E is a complex elliptic curve and Q is a point of E of order N. Here, the equivalence relation is defined by an isomorphism taking Q to Q'. Call $S_1(N)$ to the set of equivalence classes.
Finally, an enhanced elliptic curve for $\Gamma(N)$ is a pair $(E, (P, Q))$ where here (P, Q) is a pair of points of E that generates the N-torsion subgroup $E[N]$ with Weil pairing $e_N(P, Q) = e^{2\pi i/N}$. Two pairs are equivalent when we have an isomorphism taking P to P' and Q to Q'. The set of equivalences is called here $S(N)$.

Theorem 7.10. *The moduli space of* $\Gamma_1(N)$ *is* $S_1(N)$*, and two points* $[E_\tau, 1/N + \Lambda_\tau]$*,* $[E_{\tau'}, 1/N + \Lambda_{\tau'}]$ *are equal if and only* $\Gamma_1(N)\tau = \Gamma_1(N)\tau'$*. So there is a bijection* ψ_1 *between* $S_1(N)$ *and* $Y_1(N)$*.*

Proof. Take a point $[E, Q] \in S_1(N)$ and recall that E is isomorphic to $\mathbb{C}/\Lambda_{\tau'}$, with $\tau' \in \mathbb{H}$. That way, $Q = (c\tau' + d)/N$, for some $c, d \in \mathbb{Z}$ and $(c, d, N) = 1$ for the condition that the order of Q is exactly N. For that reason, we have that there exist $a, b, k \in \mathbb{Z}$ such that $ad - bc - kN = 1$ and the matrix $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_2(\mathbb{Z})$ can be seen as a matrix in $\mathrm{SL}_2(\mathbb{Z}/n\mathbb{Z})$ (reducing modulo N). We can furthermore modify the entries of γ modulo N, and this does not affect Q. Since we know that $\mathrm{SL}_2(\mathbb{Z})$ surjects to $\mathrm{SL}_2(\mathbb{Z}/n\mathbb{Z})$ we take the matrix γ to be directly in $\mathrm{SL}_2(\mathbb{Z})$. Writing $m = c\tau' + b$ and $\tau = \gamma(\tau')$ we have that $m\tau = a\tau' + b$ and so

$$m\Lambda_\tau = m(\tau\mathbb{Z} \oplus \mathbb{Z}) = (a\tau' + b)\mathbb{Z} \oplus (c\tau' + d)\mathbb{Z} = \tau'\mathbb{Z} \oplus \mathbb{Z} = \Lambda_{\tau'}$$

On the other side

$$m(1/N + \Lambda_\tau) = (c\tau' + d)/N + \Lambda_{\tau'} = Q$$

From these two observations we have that $[E, Q] = [C/\Lambda_\tau, 1/N + \Lambda_\tau]$. Similarly, when we have two points $\tau, \tau' \in \mathbb{H}$ such that $\Gamma_1(N)\tau = \Gamma_1(N)\tau'$, we have $\gamma \in \Gamma_1(N)$ such that $\tau = \gamma(\tau')$ and it can be easily verified that $[\mathbb{C}/\Lambda_\tau, 1/N + \Lambda_\tau] = [\mathbb{C}/\Lambda_{\tau'}, 1/N + \Lambda_{\tau'}]$.
Conversely, when $[\mathbb{C}/\Lambda_\tau, 1/N + \Lambda_\tau] = [\mathbb{C}/\Lambda_{\tau'}, 1/N + \Lambda_{\tau'}]$, it is clear that we have $m \in \mathbb{H}$ such that $m\Lambda_\tau = \Lambda_{\tau'}$ and a straightforward manipulation shows that it suffices to assure that $\Gamma_1(N)\tau = \Gamma_1(N)\tau'$. □

In the same way, it is possible to prove that there is a bijection between $S_0(N)$ and $Y(N)$ and another between $S(N)$ and $Y(N)$.

Chapter 8

Quaternion algebras

This chapter may seem unrelated with our previous work, but that would be an erroneous thinking. There are at least three basic reasons why quaternion algebras play a prominent role in the study of elliptic curves, modular forms and consequently BSD:

- They arise in a natural way as one of the three possibilities for the endomorphism ring of an elliptic curve (chapter 3).
- There is a close relationship (that we will explain in this chapter), between the second cohomology group and the Brauer group, that classifies quaternion algebras over a given field.
- We will see that under certain circumstances, we have a map $X_0(N) \to E$, where N is the conductor of the elliptic curve. This kind of modular parametrizations are in some sense restrictive and to enlarge this we need to introduce Shimura curves, that provide a plentyful supply of constructions.

8.1 First definitions

Let F be a field of characteristic different than 2. Given $a, b \in F^*$, let us define an algebra over F with basis $\{1, i, j, k\}$ where multiplication is given by

$$i^2 = a, \quad j^2 = b \quad ij = -ji = k$$

This algebra will be written as $\left(\frac{a,b}{F}\right)$ and will be called quaternion algebra. A particular case corresponds to the celebrated Hamilton's quaternions, namely $\left(\frac{-1,-1}{\mathbb{R}}\right)$. The basis $\{1, i, j, ij\}$ is called a standard basis, and obviously the standard basis is not unique; for instance $\left(\frac{a,b}{F}\right) = \left(\frac{ax^2,by^2}{F}\right) = \left(\frac{a,-ab}{F}\right)$. We have to make now some straightforward verifications:

Proposition 8.1. *Let $a, b \in F^*$. Then $\left(\frac{a,b}{F}\right)$ exists.*

Proof. Take α, β in an algebraic closure E of F such that $\alpha^2 = a, \beta^2 = -b$ and consider the matrices

$$i = \begin{pmatrix} \alpha & 0 \\ 0 & -\alpha \end{pmatrix}, \quad j = \begin{pmatrix} 0 & \beta \\ -\beta & 0 \end{pmatrix}, \quad k = \begin{pmatrix} 0 & \alpha\beta \\ \alpha\beta & 0 \end{pmatrix}$$

We clearly have that $k = ij = -ji$ and that $\{Id_2, i, j, k\}$ are independent over E (so also over F); therefore, they generate a four dimensional algebra H over F, what we call $H = \left(\frac{a,b}{F}\right)$. □

Theorem 8.1. *A quaternion algebra over F is central and simple (where central means that the center is F and simple that it does not have any nonzero proper two-sided ideal)*

Proof. Take an element of the form $x = \alpha + \beta i + \gamma j + \delta k$ and suppose that is in the center of H (where $\alpha, \beta, \gamma, \delta \in F$). Then, $0 = xj - jx = 2k(\beta + \delta j)$, so $\beta = \delta = 0$. Doing the same but multiplying now by i, we reach $\gamma = 0$. We conclude that $x \in F$.
The next step is showing that a nonzero two-sided ideal A must be equal to H. We will be done by showing that A contains an element of F; for that, take a nonzero element $y = a + bi + cj + dk$ in A, where $a, b, c, d \in F$ and one of b, c, d is nonzero; assume also that $a \neq 0$ (if not, replace it with iy, jy or ky). Using now that $yj - jy = 2k(b + dj)$ and $2k$ is a unity in H, we have that $b + dj, bi + dk \in A$. Subtracting this from y, we also have that $a + cj \in A$. The same reasoning allows us to say that $a + bi, a + dk$ are in A. We have now another element in a $(a + bi) + (a + cj) + (a + dk) - (a + bi + cj + dk) = 2a \in F$. We have therefore a contradiction. □

We quote now without proof some classical theorems about the classification of central simple algebras that will help us in studying quaternion algebras. The first one is the Wedderburn's structure theorem and the second one, the Skolem-Noether theorem.

Theorem 8.2. *Let A be a finite dimensional simple algebra. Then A is isomorphic to $M_n(D)$, where $D \simeq \mathrm{End}_A(N)$ is a division algebra over F with N a nonzero minimal right ideal of A. The integer n and the isomorphism class of the division algebra D are uniquely determined by A.*

Theorem 8.3. *Let A be a finite dimensional central simple algebra over F and let B be a finite dimensional simple algebra over F. If ϕ, ψ are algebra homomorphism from B to A, then there exists an invertible element $c \in A$ such that $\phi(b) = c^{-1}\psi(b)c$ for all $b \in B$ (ϕ, ψ are conjugate). In particular, all nonzero endomorphism of A are inner automorphisms.*

Let us apply Wedderburn theorem to a quaternion algebra. $H \simeq M(n, D)$, so $4 = \dim_K H = n^2 \dim_K D$ which gives only two possibilities: either $n = 1$ and so $H \simeq D$ is a division algebra of $n = 2$ and $D = K$, so $H \simeq M(2, K)$ is a matrix algebra (split).
From the construction of the quaternion algebra, we can see that if F is algebraically closed, we only obtain matrix algebras (in general, when we put $H = \left(\frac{a,b}{F}\right)$ we can construct a morphism ϕ of quaternion algebras between H and $M(2, F(\sqrt{a}))$ given by

$$\phi(x + yi + zj + tij) = \begin{pmatrix} x + y\sqrt{a} & x + t\sqrt{a} \\ b(z - t\sqrt{a}) & x - y\sqrt{a} \end{pmatrix}$$

This proves that if $\sqrt{a}$ is a square, we obtain matrix algebras.
We state now the first main theorem about quaternion algebras:

Theorem 8.4. *Let H be a quaternion algebra over F. Then,*

a) H is a division algebra or $H \simeq M_2(F)$.

b) Let E be a subfield of H which is a quadratic extension of F, and let τ be a nontrivial automorphism of E/F. Then there exists $j \in H^*$ such that $j^2 \in F^*, H = E + Ej$ and $jx = \tau(x)j$ for all $x \in E$.

We will discuss now the case when a quaternionic algebra H is isomorphic to $M_2(F)$. The treatment of this problem is analogous to the case of quadratic forms discussed at the beginning, so we omit most of the proofs.

Theorem 8.5. *Let $H = \left(\frac{a,b}{F}\right)$ be a quaternion algebra. Then, the following conditions are equivalent.*

a) $H \simeq \left(\frac{1,1}{F}\right) \simeq M_2(F)$.

b) H is not a division algebra.

c) H has an element of norm zero (isotropic).

d) H_0 (the pure quaternions) has an isotropic element.

e) The equation $ax^2 + by^2 = 1$ as a solution in $F \times F$ (note the presence here of the Hilbert symbol).

f) If $E = F(\sqrt{b})$, then $a \in N_{E/F}(E)$

We give a first easy proposition:

Proposition 8.2. *If $\mathbb{F}_q$ is a finite field, any quaternion algebra is isomorphic to $M_2(\mathbb{F}_q)$.*

Proof. It is enough with proving that $ax^2 + by^2 = 1$ has a solution in $\mathbb{F}_q$. But the image of ax^2 has $(q + 1)/2$ elements and the same for the image of $1 - by^2$. We conclude that they must have at least a common value, i.e., there is a solution of $ax^2 = 1 - by^2$. □

Proposition 8.3. *If K is a local field (different from $\mathbb{C}$) there exists a unique division quaternion F-algebra up to isomorphisms. When $F = \mathbb{R}$ we get the Hamilton quaternions.*

We can recover here the notations of the Hilbert symbol: when $(a, b)_v = 1$, H_v is a matrix algebra and H is non-ramified at v; when $(a, b) = -1$, H_v is a division algebra and H is ramified at v.

Definition 8.1. *The reduced discriminant D_H of a quaternion $\mathbb{Q}$-algebra H is the integral ideal of $\mathbb{Z}$ equal to the product of prime ideals of $\mathbb{Z}$ that ramify in H. It can be identified with an integer number.*

The main theorem about ramification is the following. Note that is not new, is a direct consequence of the theory of quadratic forms.

Theorem 8.6. *Consider a quaternion $\mathbb{Q}$-algebra H. Then,*

a) H is ramified at a finite even number of places.

b) Given an even number of non complex places of $\mathbb{Q}$, there exists a quaternion $\mathbb{Q}$-algebra ramifying exactly at these places.

c) Two quaternion $\mathbb{Q}$-algebras are isomorphic if and only if they are ramified at the same places (they have the same reduced discriminant). In particular, H is a matrix $\mathbb{Q}$-algebra if and only if $D_H = 1$.

Let us finish this section with some notation. A quaternion algebra over $\mathbb{Q}$ is called definite if when tensored with $\mathbb{R}$ is isomorphic to the Hamilton quaternions. Elsewhere (isomorphic to $M_2(\mathbb{R})$, it is called indefinite.
For the case of a number field F, we typically use the word indefinite to mean that there is at least an embedding $i : F \to \mathbb{R}$ such that the tensor product of the algebra H with $\mathbb{R}$ is isomorphic to $M_2(\mathbb{R})$ seeing $\mathbb{R}$ as an F-algebra via the inclusion i. When this works for any embedding i we call it totally indefinite.

8.2 Orders in Quaternions Algebras

We take F to be a number field or a p-adic field. Its rings of integers, O_F, is a Dedekind domain, and F is the field of fractions of O_F. O_F is, furthermore, an integrally closed noetherian ring in which every nonzero prime ideal is maximal. We call I_F the set of nonzero finitely generated o-submodules of F (fractional ideals). As we already commented, it makes sense to define the product of two fractional ideals and also the inverse of a, a^{-1} as that formed by elements $x \in F$ such that $xa \subset O_F$. We already know from basic algebraic number theory that I_F is the free abelian group on the set of nonzero prime ideals of O_F. Let P_F be here the set of principal fractional ideals, and consider as usual I_F/P_F, the ideal class group (finite). In the case of p-adic fields, the class number is trivially one. Remember also that an O_F-lattice L over a finite dimensional F-vector space, V, is complete when $FL = V$.

Definition 8.2. *Let H be a quaternion algebra over F. An O_F-ideal in H is a complete O_F-lattice in H. An order in H is an O_F-ideal which is also a ring.*

Definition 8.3. *Let O be an O_F-order in H. The discriminant of O, $d(O)$, is the fractional ideal of O_F generated by the elements* $\det(\mathrm{Tr}(x_i x_j))$*, where $x_i \in O$.*

We will work from now on over $\mathbb{Q}$. We can reformulate the definition and say that a $\mathbb{Z}$-order O of H is a subring of H whose elements are integral (or what is the same here, the trace and the norm are integers), that contains $\mathbb{Z}$ and that $\mathbb{Q} \otimes O = H$.
Further, we can affirm that if $\{v_1, v_2, v_3, v_4\}$ is a basis of the order O, then $d(O)$

is the principal ideal generated by $\det(\mathrm{Tr}(x_ix_j))$. An important fact is that if $O \subset O'$ are orders, then $d(O')|d(O)$.
We introduce now maximal orders:

Proposition 8.4. *A maximal order is an order that is not properly contained in any other order. Then,*

a) Each order is contained in a maximal order.

b) O is maximal if and only if O_v is maximal for every finite place v.

c) O is a maximal order if and only if $d(O) = D_H$. In particular, all the maximal orders have the same discriminant.

Definition 8.4. *An Eichler order in a quaternion algebra is the intersection of two maximal orders. The level of an Eichler order is the index of $O_1 \cap O_2$ either in O_1 or in O_2 (it gives the same number).*

There are two alternative characterizations of Eichler orders that are very useful:

Proposition 8.5. *Let O be an order in a quaternion $\mathbb{Q}$-algebra H of discriminant D. Let $N \subset \mathbb{Z}$ be coprime to D. Then, the following conditions are equivalent:*

a) O is an Eichler order of level N.

b) For every prime p, O satisfies that if p does not divide N, the local $\mathbb{Z}_p$ order O_p is maximal, and if $p|N$, O_p is isomorphic to the order $\begin{pmatrix} \mathbb{Z}_p & \mathbb{Z}_p \\ N\mathbb{Z}_p & \mathbb{Z}_p \end{pmatrix}$.

c) For every prime p, O satisfies that if p divides D, the local $\mathbb{Z}_p$ is maximal and if p does not divide D, O_p is isomorphic to the order $\begin{pmatrix} \mathbb{Z}_p & \mathbb{Z}_p \\ N\mathbb{Z}_p & \mathbb{Z}_p \end{pmatrix}$.

There is no an explicit characterization of Eichler orders in terms of their discriminant, but we can state some of their properties:

Proposition 8.6. *Let H be a quaternion $\mathbb{Q}$-algebra of discriminant D.. Then,*

a) For each integer N such that $(D, N) = 1$ there exists an Eichler order of level N.

b) Let $O(D, N) \subset O(D, 1)$. Then, the index as $\mathbb{Z}$-modules is $[O(D, 1) : O(D, N)] = N$.

c) For $O = O(D, N)$, $d(O) = DN$.

d) If $d(O) = DN$ is a square free, then O is an Eichler order of level N.

We finish this section with a remarkable result in number fields:

Theorem 8.7. *Let F be a totally real number field and let H be an indefinite quaternion F-algebra. If the ideal class number of F is odd, there is only one conjugacy class of Eichler orders having the same level. In particular, for indefinite quaternion rational algebras, all Eichler orders having the same level are conjugated.*

8.3 Shimura curves: an introduction

In this section we define Shimura curves $X(D,N)$ attached to Fuschian groups defined from Eichler orders $O(D,N)$ in a quaternion $\mathbb{Q}$-algebra. Define first $O(D,N)^*_+$ as the elements in $O(D,N)$ with norm one. Assume that D is the product of an even number of different primes and take an isomorphism $\Phi : \mathbb{R}\otimes H \to M(2,\mathbb{R})$. By the last theorem of the previous section, $O(D,N)^*_+$ will only depend on D and N up to conjugation. Given $H = \left(\frac{a,b}{\mathbb{Q}}\right)$, with $a > 0$, we will take

$$\Phi(x+yi+zj+tij) = \begin{pmatrix} x+y\sqrt{a} & x+t\sqrt{a} \\ b(z-t\sqrt{a}) & x-y\sqrt{a} \end{pmatrix}$$

Let $\Gamma(D,N) = \Phi(O(D,N)^*_+)$. It is a subgroup of $\mathrm{SL}(2,\mathbb{R})$ whose elements will be called quaternion transformations.
We note that

$$\Gamma(D,N) \subset \{\begin{pmatrix} \alpha & \beta \\ b\beta' & \alpha' \end{pmatrix} \mid \alpha,\beta\in\mathbb{Q}(\sqrt{a})\} \subset SL(2,\mathbb{Q}(\sqrt{\alpha}))$$

where α' denotes the usual Galois conjugation.
We are now in conditions to define a Shimura curve. Let D, N be natural numbers, where D is the product of an even number of different primes and $(D,N) = 1$.. Fix an indefinite quaternion $\mathbb{Q}$-algebra H of discriminant $D_H = D$, an Eichler order $O(D,N)$ of level N in H and a monomorphism $\Phi : H \to M(2,\mathbb{R})$. Consider also the group of quaternion transformations $\Gamma(D,N)$ associated with the order $O(D,N)$ and Φ. The group $\Gamma(D,N)$ is a Fuschian group that acts on the upper half plane, and the quotient $\Gamma(D,N)\backslash\mathbb{H}$ is a Riemann surface. We define for this Riemann surface a canonical model with the following properties:

a) $X(D,N)$ is a projective curve defined over $\mathbb{Q}$.

b) There exists a map $j_{D,N} : \mathbb{H} \to X(D,N)(\mathbb{C})$ that factorizes in an isomoprhism between the analytic space $\Gamma(D,N)\backslash\mathbb{H}$ and a Zariski open set in $X(D,N)(\mathbb{C})$.

c) Let $F = \mathbb{Q}(\sqrt{d})$ be an imaginary quadratic field splitting the algebra H. Let ϕ be an embedding of F into H, and let $z \in \mathbb{H}$ the unique common fixed point of all the elements in $\Phi(\phi(F^*))$. Then, the coordinates of the point $j_{D,N}(z)$ are algebraic, more specifically $j_{D,N}(z) \in X(D,N)(F^{\mathrm{ab}})$ where F^{ab} is the maximal abelian extension of F.

$X(D,N)$ is called the Shimura curve associated to $\Gamma(D,N)$. The case $D = 1$ corresponds to a non-ramified quaternion algebra. In this case, $\Gamma(1,N)\backslash\mathbb{H}$ is a non-compact Riemann surface with finite volume. It is clear that the corresponding compact Shimura curve $X(1,N)$ is the modular curve $X_0(N)$. If $D > 1$ the quaternion algebra H is ramified and the Riemann surface $\Gamma(D,N)\backslash\mathbb{H}$ is already compact.
We can give the following moduli interpretation: a point in $X(D,N)(\mathbb{C})$ corresponds to an isomorphism class of triples (A,i,G), where A is a certain kind of abelian surface, $i : H \to \mathbb{Q}\otimes_{\mathbb{Z}}\mathrm{End}(A)$ is a monomorphism such that $i(O(D,1)) \subset$

$\mathrm{End}(A)$ and G is a subgroup of the group of N-torsion points of A which is a cyclic $O(D,N)$-module.

8.4 Brauer group

Let A, B be k-algebras, and let $A \otimes_k B$ be the tensor product of A and B as k-vector spaces. There is a unique k-bilinear multiplication on $A \otimes_k B$ such that

$$(a \otimes b)(a' \otimes b') = aa' \otimes bb' \text{ for all } a, a' \in A, b, b' \in B$$

Identifying k with $k \times (1 \otimes 1) \subset A \otimes_k B$ we give it structure of k-algebra.
We state now some propositions about the structure of these tensor products.

Proposition 8.7. *Let A, A' be k-algebras with subalgebras B, B' and let $C(B)$ and $C(B')$ be the centralizers of B, B' in A, A' respectively. Then, the centralizer of $B \otimes_k B'$ in $A \otimes_k A'$ is $C(B) \otimes C(B')$. In addition, the center of a simple k-algebra is a field.*

Proposition 8.8. *The tensor product of two simple k-algebras, at least one of which is central, is again simple. Furthermore, the tensor product of two central simple k-algebra is again central simple.*

Recall that from Noether-Skolem theorem we can say a lot about the homomorphisms of a k-algebra. For instance, it is immediate that when A is a central simple algebra over k and B_1, B_2 are simple k-subalgebras, any isomorphism $f : B_1 \to B_2$ is induced by an inner automorphism of A, that is, there exists an invertible $a \in A$ such that $f(b) = aba^{-1}$ for all $b \in B_1$. This implies that all automorphisms of a central simple k-algebra are inner (the classical example: for $M_n(k)$ the automorphism group is $\mathrm{PGL}_n(k)$).

We can now move to the definition of the Brauer group. Let A and B central simple algebras over k. We say that they are similar ($A \sim B$) if $A \otimes_k M_n(k) \approx B \otimes_k M_m(k)$ for some m, n. It is direct to check that it is an equivalence relation. We define the Brauer group of k, $\mathrm{Br}(k)$, to be the set of equivalence classes of central simple algebras over k, and write $[A]$ for one such element. In $\mathrm{Br}(k)$ we have the following operation

$$[A][B] = [A \otimes_k B]$$

that is well defined and is also associative and commutative. Since for every n, $[M_n(k)]$ is an identity element, the fact that $A \otimes_k A^{\mathrm{opp}} \approx M_n(k)$ implies that $[A]$ has $[A^{\mathrm{opp}}]$ as inverse. We conclude that $\mathrm{Br}(k)$ is an abelian group.
Wedderburn's theorem states that every central simple algebra over k is isomorphic to $M_n(D)$ for some central division algebra D and that D is uniquely determined by A up to isomorphism. Therefore each similarity class is represented by a central division algebra and two central division algebras represent the same similarity class if and only if they are isomorphic.

We give now some examples: when k is algebraically closed, $\mathrm{Br}(k) = 0$ since if $\alpha \in D$ (where D is a central division algebra), we can take $k[\alpha]$, the subalgebra generated by k and α that is a commutative field of finite degree over k because it is an integral domain of finite degree over k. Hence, $k[\alpha] = k$ and this shows that $D = k$. Frobenius showed that Hamilton's quaternion algebra is the only central division algebra over $\mathbb{R}$. Therefore, $\mathrm{Br}(\mathbb{R})$ is cyclic of order 2. In any undergraduate course of algebra it is studied the little Wedderburn's theorem: a finite division algebra is commutative. This says that the Brauer group of a finite field is zero.
We move now to more interesting examples. The first one refers to the Brauer group of a non-archimedean local field and it is due to Hasse. It says that the Brauer group is canonically isomorphic to $\mathbb{Q}/\mathbb{Z}$. The proof of this fact is very related with cohomology. Recall that for a Galois extension L/k of fields, we write $H^2(L/k) = H^2(\mathrm{Gal}(L/k), L^*)$.

Theorem 8.8. *There is a natural isomorphism between $H^2(L/k)$ and $\mathrm{Br}(L/k)$. In other word, the second cohomology group classifies the central simple algebras over k split by L.*

To understand this theorem we have to do some previous considerations. The first one is the following proposition:

Proposition 8.9. *Let A be a central simple algebra over k, and let K be a field containing k. Then, $A \otimes_k K$ is a central simple algebra over K.*

As we have already pointed out in our discussion of quaternion algebras, a central simple algebra A is said to be split by L (and L is called a splitting field for A) if $A \otimes_k L$ is a matrix algebra over L. Thus, we can define $\mathrm{Br}(L/k)$ are the elements of $\mathrm{Br}(k)$ siplit by L or alternatively, the kernel of the homomorphism $\mathrm{Br}(k) \to \mathrm{Br}(L)$ defined by sending A to $A \otimes_k L$.

Proposition 8.10. *For every field k, $\mathrm{Br}(k) = \cup \mathrm{Br}(K/k)$ where K runs over the finite extensions of k contained in some fixed algebraic closure k^{al} of k.*

The proof of the theorem relating the second cohomology group with the Brauer group is quite long and we skip it. However, we explain a consequence of this fact. With class field theory it is possible to prove that for non-archimedean local fields, $H^2(K^{\mathrm{al}}/K) \simeq \mathbb{Q}/\mathbb{Z}$ and so $\mathrm{Br}(K) \simeq \mathbb{Q}/\mathbb{Z}$. From here, we see that the Brauer group is a torsion group.

The study of the Brauer group of a number field is considerably much more difficult. Albert, Brauer, Hasse and Noether showed that for any number field K there is an exact sequence

$$0 \to \mathrm{Br}(K) \to \bigoplus_v \mathrm{Br}(K_v) \to \mathbb{Q}/\mathbb{Z} \to 0$$

where the sum is over all primes of K (including the infinite primes) and the morphism that goes from $\oplus \mathrm{Br}(K_v)$ to $\mathbb{Q}/\mathbb{Z}$ is the sum of the respective images

of each of the isomorphisms we have between $\mathrm{Br}(K_v)$ and $\mathbb{Q}/\mathbb{Z}$.
We finish the section with a theorem due to Tsen:

Theorem 8.9. *If k is a field of transcendence degree one over an algebraically closed field, k has trivial Brauer group.*

Chapter 9

Complex multiplication

The study of curves whose endomorphism ring is greater than $\mathbb{Z}$ is of special interest for us, since these curves have certain special properties that we will study in the next chapter and that are of special interest to understand the most basic results around BSD. One of the main motivations for the study of complex multiplication comes from the Kronecker-Weber theorem, that says that the maximal abelian extesion of $\mathbb{Q}$ equals the maximal cyclotomic extension; the theory of complex multiplication tries to study abelian extensions of quadratic imaginary fields. There is no a natural generalization to the case of real quadratic, and it is not still well understood how to know the maximal abelian extension.

9.1 Complex multiplication over $\mathbb{C}$

Recall that over $\mathbb{C}$ the ring of endomorphisms of an elliptic curve can be either $\mathbb{Z}$ either an order of a quadratic imaginary field. In this last case we say that E has complex multiplication by R (where R is the order) or by K, defining $K = R \otimes \mathbb{Q}$. It is necessary to bear in mind the usual isomorphism given by the Uniformization Theorem:

$$f : \mathbb{C}/\Lambda \to E(\mathbb{C}) : z \mapsto (\wp(z, \Lambda), \wp'(z, \Lambda))$$

When E has complex multiplication, there are two ways to embed the order $\operatorname{End}(E)$ into $\mathbb{C}$, but it is important to pin down one of these embeddings.

Proposition 9.1. *Let $E/\mathbb{C}$ be an elliptic curve with complex multiplication by the ring $R \subset \mathbb{C}$. There is a unique isomorphism*

$$[\cdot] : R \to \operatorname{End}(E)$$

such that for any invariant differential $\omega \in \Omega_E$ on E, $[\alpha]^*\omega = \alpha\omega$ for all $\alpha \in R$. In this case, we say that the pair $(E, [\cdot])$ is normalized.

In order to study a particular elliptic curve with complex multiplication, we should look at the set of all elliptic curves with the same endomorphism ring. We define $\mathrm{ELL}(R)$ to be the set of elliptic curves whose endomorphism ring is R

modulo isomorphism over $\mathbb{C}$. The most natural question in this moment is how can we construct an elliptic curve with complex multiplication by O_K. If a is a nonzero fractional ideal of K, we know that a is a lattice in $\mathbb{C}$ (the definition of fractional ideals for quadratic imaginary fields implies that a is a $\mathbb{Z}$-module of rank two not contained in $\mathbb{R}$). That way, any nonzero fractional ideal a of K will give an elliptic curve with complex multiplication by O_K, the ring of integers of a quadratic number field. Since we are taking the relation modulo isomorphisms, we see that a and ca give the same elliptic curve, so we just have to look at the ideal class group of O_K. These observations are contained in the following proposition:

Proposition 9.2. *Let Λ be a lattice with $E_\Lambda \in \mathrm{ELL}(O_K)$ and let a, b be nonzero fractional ideals of K. Then,*

1. *$a\Lambda$ is a lattice in $\mathbb{C}$.*
2. *$E_{a\Lambda}$ satisfies that its endomorphism ring is O_K.*
3. *$E_{a\Lambda} \simeq E_{b\Lambda}$ if and only if they are equal in $\mathrm{CL}(O_K)$.*

Proof. For the first part, choose a nonzero integer d such that $da \in O_K$; then, $a\Lambda \subset 1/d\Lambda$, so $a\Lambda$ is a discrete subgroup of $\mathbb{C}$. Similarly, take another nonzero integer d such that $dO_K \subset a$ and so $d\Lambda \subset a\Lambda$. We have that $a\Lambda$ spans $\mathbb{C}$ and so is a lattice.

Let now $\alpha \in \mathbb{C}$, and let $a \neq 0$ be a fractional ideal. We have that $\alpha a\Lambda \subset a\Lambda$ if and only if $\alpha\Lambda \subset \Lambda$. Hence, $\mathrm{End}(E_{a\Lambda}) = \mathrm{End}(E_\Lambda) = O_K$.

For the last part, in previous chapters we have seen that the isomorphism class of $E_{a\Lambda}$ depends only on the homothety class of $a\Lambda$. Alternatively, we want to know if there is a $c \in \mathbb{C}^*$ such that $a\Lambda = cb\Lambda$. Manipulating now the expression, $E_{a\Lambda} \cong E_{b\Lambda}$ if and only if both $ca^{-1}b$ and $c^{-1}ab^{-1}$ take Λ to itself (it is the same than saying that they are in O_K). Therefore, $a = cb$ (or they are the same in the class group). □

Note that it is natural to consider now an action of $\mathrm{CL}(O_K)$ over $\mathrm{ELL}(O_K)$ given by $\bar{a} * E_\Lambda = E_{a^{-1}\Lambda}$. The choice of a^{-1} is arbitrary and it is done in order to ease expressions in the future.

Proposition 9.3. *The previous action is simply transitive. In particular, the number of classes $\mathrm{CL}(O_K)$ coincides with the number of elliptic curves with complex multiplication by O_K modulo isomorphisms.*

Orders in K

Let K be a quadratic imaginary number field. An order there is a subring that contains $\mathbb{Z}$ and free of rank 2 over $\mathbb{Z}$. For the definition (an order is also a ring), every element of R is integral over $\mathbb{Z}$, so $R \subset O_K$ and O_K is the unique maximal order.

Proposition 9.4. *Let R be an order in K. Then, there is a unique integer $f > 0$ such that $R = \mathbb{Z} + fO_K$. Conversely, $\mathbb{Z} + f \cdot O_K$ is an order in K for every integer $f > 0$. f is called the conductor of R.*

Proof. Take $\{1, \alpha\}$ a $\mathbb{Z}$-basis for O_K, in such a way that $O_K = \mathbb{Z} + \mathbb{Z}\alpha$. Then $R \cap \mathbb{Z}\alpha$ is a subgroup of $\mathbb{Z}\alpha$, so it is equal to $\mathbb{Z}\alpha f$ (the only subgroups of $\mathbb{Z}$ are the multiples of f, for $f > 0$). We have so that $\mathbb{Z} + fO_K \subset \mathbb{Z} + \mathbb{Z}\alpha f \subset R$. Now, if $m + n\alpha \in R$ ($m, n \in \mathbb{Z}$), then $n\alpha \in R$ and therefore $n \in f\mathbb{Z}$. We have proved that $m + n\alpha \in \mathbb{Z} + f\alpha\mathbb{Z} \subset \mathbb{Z} + fO_K$. □

The easy case, and that will be the one for which we develop most of our results, is when $f = 1$. But we will be interested also in the cases where the endomorphism ring is not maximal. We now give a characterization about R-submodules of K that can be useful in some moments.

Proposition 9.5. *Let R be an order in K. The following conditions on an R-submodule a of K are equivalent:*

a) a is a projective R-module (a module a is projective if there is a free module f and another module b such that the direct sum of a and b is f).

b) $R = \{c \in K \mid ca \subset a\}$.

c) $a = xO_K$ for some $x \in I$ (here $\mathbb{I}$ is the group of ideles of $\mathbb{Q}$, to be introduced in a few sections).

9.2 Rationality questions

In this section we will talk about the field of definition for complex multiplication elliptic curves and their endomorphisms. The key result is that when an elliptic curve has complex multiplication, its j invariant is an algebraic number. We begin with the following proposition:

Proposition 9.6. *Let $E/\mathbb{C}$ be an elliptic curve. Then,*

a) Let $\sigma : \mathbb{C} \to \mathbb{C}$ be a field automorphism of $\mathbb{C}$. It holds that $\mathrm{End}(E^\sigma) = \mathrm{End}(E)$.

b) If E has complex multiplication by O_K, then $j(E) \in \bar{\mathbb{Q}}$ (in fact, it is an algebraic integer).

c) $\mathrm{ELL}(O_K)$ is isomorphic to the elliptic curves over $\bar{\mathbb{Q}}$ with O_K as the endomorphism ring modulo isomorphism over $\bar{\mathbb{Q}}$.

Proof. The first item is clear, since if ϕ is an endomorphism of E, ϕ^σ is an endomorphism of E^σ.

For the second one, let σ be an automorphism of $\mathbb{C}$. E^σ is obtained by letting σ act on the coefficients of the Weierstrass equation for E, and $j(E)$ is a rational combination of those coefficients, and $j(E^\sigma) = j(E)^\sigma$. But $\mathrm{End}(E^\sigma) \cong O_K$, and there are only finitely many classes of elliptic curves with the same endomorphism ring (modulo isomorphism). But j determines the isomorphism class of an elliptic curve, so $j(E)^\sigma$ takes only finitely many values when σ is an automorphism of $\mathbb{C}$. Therefore, $[\mathbb{Q}(j(E)) : \mathbb{Q}]$ is finite.

For the last part of the proposition, consider a subfield F of $\mathbb{C}$ and define $\mathrm{ELL}_F(R_K)$ to be the set of elliptic curves over F with O_K as the endomorphism ring modulo isomorphisms of F. When we fix an embedding of $\bar{\mathbb{Q}}$ in $\mathbb{C}$ there is a natural map $\epsilon : \mathrm{ELL}_{\bar{\mathbb{Q}}}(O_K) \to \mathrm{ELL}_{\mathbb{C}}(O_K)$. The proposition says that this map is a bijection.
For any element of $\mathrm{ELL}_{\mathbb{C}}(O_K)$ we already know that $j(E)$ is an algebraic number, that there is an elliptic curve E' over $\mathbb{Q}(j(E))$ with $j(E) = j(E')$ and that E' is isomorphic to E over $\mathbb{C}$. Consequently $\epsilon(E') = E$. Surjectivity is so established. On the other hand, if two curves E_1, E_2 over $\bar{\mathbb{Q}}$ fulfills that $\epsilon(E_1) = \epsilon(E_2)$, then $j(E_1) = j(E_2)$ and since $\bar{\mathbb{Q}}$ is algebraically closed and the j invariant coincide, they must be isomorphic. We have so that ϵ is also injective. □

Recall the analogy with cyclotomic fields: we know that if ζ is a primitive N-th root of unity and $\sigma \in \mathrm{Gal}(\mathbb{Q}(\zeta)/\mathbb{Q})$, then ζ^σ s another primitive N-th root, say $\zeta^\sigma = \zeta^{\rho(\sigma)}$ and it is easy to check that $\rho : \mathrm{Gal}(\mathbb{Q}(\zeta)/\mathbb{Q}) \to (\mathbb{Z}/N\mathbb{Z})$ is an injective homomorphism (also surjective in the case of $\mathbb{Q}$) and that the extension is abelian. We state now a similar fact for elliptic curves:

Theorem 9.1. *Let $E/\mathbb{C}$ be an elliptic curve with complex multiplication by O_K. Let $L = K(j(E), E_{\mathrm{tors}})$ be the field generated by the j-invariant and the coordinates of all of the torsion points of E. Then L is an abelian extension of $K(j(E))$*

Proof. Let $H = K(j(E))$ and $L_m = H(E[m])$, the extension of H generated by the m-torsion points of E. It will be enough to show that L_m is an abelian extension of H, since L is the compositum of all the L_m. We have the usual representation

$$\rho : \mathrm{Gal}(\bar{K}/H) \to \mathrm{Aut}(E[m])$$

determined by the condition $\rho(\sigma)(T) = T^\sigma$, where $T \in E[m]$. For a general elliptic curve we can deduce that $\mathrm{Gal}(L_m/H)$ injects into the automorphism group of the abelian group $E[m]$ so $\mathrm{Gal}(L_m/H)$ is isomorphic to a subgroup of $\mathrm{GL}_2(\mathbb{Z}/m\mathbb{Z})$. But once we know that the curve has complex multiplication we can take a model for E over H and for what we have seen, every endomorphism of E is also defined over H. Consequently, elements of $\mathrm{Gal}(L_m/H)$ commute with elements of O_K when acting on $E[m]$.

$$([\alpha]T)^\sigma = [\alpha](T^\sigma)$$

Alternatively, ρ is a homomorphism from $\mathrm{Gal}(\bar{K}/H)$ to the group of O_K/mO_K-module automorphisms of $E[m]$. Hence, we have an injection of $\mathrm{Gal}(L_m/H)$ in $\mathrm{Aut}\, O_K/mO_K(E[m])$. But using that $E[m]$ is a free O_K/mO_K-module of rank one, this last group is isomorphic to $(O_K/mO_K)^*$ and so $\mathrm{Gal}(L_m/H)$ is abelian. □

Until now, we have identified $\mathrm{ELL}(O_K)$ with the elliptic curves with complex multiplication by O_K modulo isomorphisms of $\bar{\mathbb{Q}}$. That way, we have a natural action of the absolute Galois group sending the isomorphism class of E to that of E^σ. We also know that the action of the class group is simply transitive, so there is a unique $a \in \mathrm{CL}(O_K)$ such that $a * E = E^\sigma$. We conclude that we have a well defined map $F : \mathrm{Gal}(\bar{K}/K) \to \mathrm{CL}(O_K)$ characterized by that property. Furthermore, F is a homomorphism and is independent of the choice of the curve $E \in \mathrm{ELL}(O_K)$. Note also the analytic component of F, that can also be characterized by $j(\Lambda)^\sigma = j(F(\sigma)^{-1}\Lambda)$.

Proposition 9.7. *Let K be a quadratic imaginary field. There exists a homomorphism*

$$F : \mathrm{Gal}(\bar{K}/K) \to \mathrm{CL}(O_K)$$

uniquely characterized by $E^\sigma = F(\sigma) * E$ for all $\sigma \in \mathrm{Gal}(\bar{K}/K)$ and all $E \in \mathrm{ELL}(O_K)$.

Proposition 9.8. *Let $E/\bar{\mathbb{Q}}$ be an elliptic curve that represents an element of $\mathrm{ELL}(O_K)$, and let $a \in \mathrm{CL}(O_K), \sigma \in \mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q})$. Then*

$$(a * E)^\sigma = a^\sigma * E^\sigma$$

9.3 Two words about class field theory

The goal of class field theory is to describe the abelian Galois extensions of a local or a global field in terms of the arithmetic of the field (the case of non-abelian extensions is much more complicated and it appears as one of the central topics in the Langlands' programme). It is an essential tool for a full understanding of arithmetic and extensions of number fields so we use the following lines to give a brief insight that will be necessary to continue our exposition of complex multiplication. When we have a Galois extension L/K, a prime in K factors as $pO_L = (p_1 \cdots p_g)^e$, where $n = efg$ and f is the degree of the residue field extension (seen in the first chapter). Let $\mathrm{Spl}(L/K)$ be the set of primes of K that split in L. Frobenius proved that this set has density $1/[L : K]$ in the set of all primes. Several results from class field will explain that the Galois extensions of K are classified by the sets $\mathrm{Spl}(L/K)$.

Let us consider the example of quadratic extensions of $\mathbb{Q}$ for the sake of clarity. Let p be an odd prime, and let $p^* = (-1)^{\frac{p-1}{2}}$ (that way $p^* \equiv 1$ modulo 4). This implies that $\mathbb{Q}(\sqrt{p^*})/\mathbb{Q}$ is ramified only at p (the discriminant is p, this is precisely the quadratic extension contained in $\mathbb{Q}(\zeta_p)/\mathbb{Q}$), and a prime number $q \neq p$ will split in the extension if and only if p^* is a square modulo q, and when q is odd for the quadratic reciprocity law this is the same than saying that q is a square modulo p. We have that $\mathrm{Spl}(\mathbb{Q}(\sqrt{p^*}/\mathbb{Q})$ are the primes q such that q mod p lies in the unique subgroup of index 2 of $(\mathbb{Z}/p\mathbb{Z})$.

Consider now an unramified abelian extension (for instance, in $\mathbb{Q}$ this is not possible). Let I be the group of fractional ideals of K, and let $i : K^* \to I$ be the map sending $a \in K^*$ to the principal ideal (a). Clearly $I/i(K^*)$ is the class group C and there is a bijection between subgroups H of C and subgroups $\tilde{H}$ of I containing $i(K^*)$. As usual, we call primes of K to the equivalence classes of nontrivial valuations on K. A real prime of K is said to split in an extension L/K if every prime lying over it is real, and otherwise is said to ramify. For example, $\mathbb{Q}(\sqrt{-5})$ is ramified over $\mathbb{Q}$ in $(2), (5)$ and ∞. Let H be a subgroup of C; a finite unramified extension L of K is a class field of H if the prime ideals of K splitting in L are exactly those in $\tilde{H}$.

Theorem 9.2. *A class field exists for each subgroup of C, it is unique and every finite unramified abelian extension of K arises as the class group of some subgroup of C. If L is the class field of H, then $\mathrm{Gal}(L/K) \simeq C/H$ and $f(p)$ is the order of the image of p in C/H for all prime ideals p of K.*

Two brief remarks: the subgroup H of C corresponding to a finite unramified abelian extension L of K is that generated by the primes splitting in L. The class field of the trivial subgroup of C is called the Hilbert class field of K, and it is the largest abelian extension L of K unramified at all primes of K, the prime ideals that split are the principal ones and the Galois group is isomorphic to C.

Dealing with ramified abelian extensions is not so easy since we need a generalization of the notion of ideal class group. Take for instance a cyclotomic extension, $\mathbb{Q}(\zeta_m)$, and clearly the primes that ramify are the ideals (p) where $p|m$ and also ∞. We want to identify $(\mathbb{Z}/m\mathbb{Z})^*$ with an ideal class group: to do so, let S be the set of prime ideals (p) such that $p|m$ and let I^S be the group of fractional ideals of $\mathbb{Q}$ generated by the prime ideals not in S. An element of I^S can be written as (r/s), where r, s are positive integers coprime with m, and we map (r/s) to $[r][s]^{-1}$ in $(\mathbb{Z}/m\mathbb{Z})^*$. We have so a homomorphism $I^S \to (\mathbb{Z}/m\mathbb{Z})^*$ whose kernel can be thought as those elements such that r, s have the same sign and they map to the same element in $(\mathbb{Z}/p^{\mathrm{ord}_p(m)}\mathbb{Z})^*$ for all primes p dividing m. We introduce now an important concept: a modulus of a number field K is the formal product $m = m_0 m_\infty$, where m_0 is an integral ideal and m_∞ is the product of some real primes. $I^{S(m)}$ is the group of fractional ideals generated by the primes not dividing m_0. C_m is the quotient of $I^{S(m)}$ by the subgroup of principal ideals in I^S and generated by an element a such that $a > 0$ for all real primes dividing m_∞ and $\mathrm{ord}_p(a-1) \geq \mathrm{ord}_p(m_0)$ for all prime ideals dividing m_0.
Let now H be a subgroup of C_m for some modulus m and let $\tilde{H}$ be its inverse image in $I^{S(m)}$. An abelian extension L of K is a class field for H if the prime ideals of K not dividing m_0 that split in L are those in $\tilde{H}$.

Theorem 9.3. *A class field exists for each subgroup of a class group C_m, it is unique and every finite abelian extension of K is the class field of some subgroup of a class group. If L is the class field of $H \subset C_m$, then $\mathrm{Gal}(L/K) \simeq C_m/H$ and the prime ideals p of K not dividing m are unramified in L are unramified in L and have $f(p)$ equal to the order of the image of p in the group C_m/H.*

We are going to state now one of the main theorems in class field theory, the local reciprocity law:

Theorem 9.4. *Let K be a nonarchimedean local field, then there exists a unique homomorphism*

$$\phi_K : K^* \to \mathrm{Gal}(K^{\mathrm{ab}}/K)$$

with the properties that:

a) For every prime π of K and every finite unramified extension L, $\phi_K(\pi)$ acts on L as $\mathrm{Frob}_{L/K}$.

b) For every finite abelian extension L of K, $Nm_{L/K}(L^*)$ is contained in the kernel of $a \mapsto \phi_k(a)|L$, and ϕ_K induces an isomorphism

$$\phi_{L/K} : K^*/\,\mathrm{Nm}_{L/K}(L^*) \to \mathrm{Gal}(L/K)$$

For the moment, we do not need all the theory, so we will explain some particular facts that will be useful in our context. Assume from the moment that we have K, a totally imaginary number field, and L is a finite abelian extension. Let p be a prime of K that does not ramify in L and let P be a prime of L lying over p. We consider as usual the Galois group of R_L/P over R_K/p, that is cyclic and generated by the Frobenius element σ_p, that is under these circumstances uniquely determined by the condition $\sigma_p(x) = x^q$, being q the cardinal of the residue field R_K/p.
Let now c be an integral ideal of K divisible by all primes that ramify, and let $I(c)$ the group of fractional ideals of K relatively prime to c. Then we can define the Artin map using the Frobenius σ_p

$$I(c) \to \mathrm{Gal}(L/K); \quad (\prod_p p^{n_p}, L/K) \mapsto \prod_p \sigma_p^{n_p}$$

Proposition 9.9. *Let L/K be a finite abelian extension of number fields. There exists an integral ideal $c \subset O_K$ divisible by the primes of K that ramify in L such that*

$$\big((\alpha), L/K\big) = 1$$

for all $\alpha \in K^*$ such that $\alpha \equiv 1$ modulo c.

If the proposition is true for two ideals c_1 and c_2 so it is for the sum, so there is a larger ideal for which the statement is true. We call it the conductor of the extension and write $c_{L/K}$. Artin reciprocity states that the kernel of the Artin map contains $P(c)$ for an appropriate choice of c. If p is an unramified prime of K, then p splits completely in L if and only if the extensions of residue fields has degree one and that way, the unramified prime ideals in the kernel of the Artin map are those of K that split completely.

Definition 9.1. *Let c be an integral ideal of K. A ray class field of K (modulo c) is a finite abelian extension K_c/K with the property that for any finite abelian extension L/K, if $c_{L/K}|c$, then $L \subset K_c$.*

More generally, take a module m and put $I_k(M)$ as the group of fractional ideals in O_K whose norm is coprime with m. We can define $P_{K,1}(m)$ as the subgroup of $I_K(m)$ generated by the principal ideals αO_K where $\alpha \equiv 1 \mod m_0$ and $\sigma(\alpha) > 0$ for all the infinite primes dividing m_∞. We say that a subgroup H of $I_K(m)$ is a congruence subgroup for the module m if it satisfies $P_{K,1}(m) \subset H \subset I_K(m)$. In this case, $I_K(m)/H$ is a generalized class group for m.
When $K \subset L$ is an abelian extension, we have seen that there exists a module $f = f(L/K)$ such that a prime of K ramifies if and only if it divides f, and if m is a module divisible by all the primes of K ramifying in L, then the kernel of the Artin map that goes from $I_K(m)$ to $\mathrm{Gal}(L/K)$ is a congruence subgroup for m if and only if $f|m$.
Our focus of interest will be in the case when O is an order in a quadratic imaginary field of conductor c. We define $P_{K,\mathbb{Z}}(c)$ as the congruence subgroup generated by the principal ideals $\alpha \equiv a \mod cO_K$, where a is an integer coprime with c. The quotient $I_K(c)/P_{K,\mathbb{Z}}(c)$ is isomorphic to $\mathrm{Pic}(O)$. We have so that there must exist an abelian extension L/K (the class field of O) satisfying the previous hypothesis. When $c = 1$, the extension we obtain is the Hilbert class field, that is, the maximal non-ramified abelian extension of K.

9.4 Idelic formulation of class field theory

It was a matter of time to define the concept of adeles and ideles, since it is a key definition in number theory and it will reappear in the next chapters when studying deeper facts of modular forms (over quaternion algebras, for instance).

Definition 9.2. *Let S be a finite set of places (in $\mathbb{Q}$) containing the infinity. We define the set of S-adeles, denoted by $A_{\mathbb{Q}}^S$, to be*

$$A_{\mathbb{Q}}^S = \prod_{p \notin S} \mathbb{Z}_p \times \prod_{p \in S} \mathbb{Q}_p \times \mathbb{R}$$

This ring is a topological ring endowed with the product topology. The ring of adeles in the union over all the finite sets S of $A_{\mathbb{Q}}^S$. Note that $\mathbb{Q}$ is a subring of $A_{\mathbb{Q}}$.
The group of invertible elements of the adele ring is the idele group. It can be realized as the restricted product of the unit group of the different places with respect to the subgroup of local integral units.

This can be generalized in the obvious way to the case of number fields, just letting the product be over all the primes. For the case we will need here, let K be an arbitrary number field and for each absolute value v on K, let K_v be the completion of K at v. Let R_v be the ring of integers of K_v when v is non archimedean and let $R_v = K_v$ otherwise. The idele group is nothing but

$$I_K^* = \prod_v{}' K_v^*$$

where the $'$ indicates that the product is restricted relative to the R_v's.
Let $s \in A_K^*$ be an idele. The ideal of s is the fractional ideal of K given by

$$(s) = \prod_p p^{\mathrm{ord}_p s_p}$$

where the product is over all prime ideals of K and this is well defined since s_p is a p-adic unit for all but finitely many p. For an integral ideal, define U_c as the subgroup of I_K^* given by

$$U_c = \{s \in A_K^* \mid s_p \in R_p^* \text{ and } s_p \equiv 1 \mod cR_p \text{ for all primes } p\}$$

Proposition 9.10. *Let K be a number field and K^{ab} the maximal abelian extension of K. There exists a unique continuous homomorphism*

$$A_K^* \to \mathrm{Gal}(K^{\mathrm{ab}}/K) \mid s \mapsto [s, K]$$

with the property that if L/K is a finite abelian extension and $s \in I_K^*$ is an idele whose ideal (s) is not divisible by any prime ramifying in L, then

$$[s, K]|_L = ((s), L/K)$$

where $(\cdot, L/K)$ is the Artin map. This homomorphism, called the reciprocity map, verifies that is surjective, with K^* in the kernel, that is compatible with the norm map

$$[x, L]|_L = [N_{L/K}(x), K] \text{ for all } x \in A_L^*$$

9.5 Applications of class field to complex multiplication

The main theorem of this part will be the following one:

Theorem 9.5. *Let $K/\mathbb{Q}$ be a quadratic imaginary field with ring of integers O_K, and let $E/\mathbb{C}$ be an elliptic curve with O_K as the endomorphism ring. Then $K(j(E))$ is the Hilbert class field H of K.*

For proving this result, we strongly need the following proposition:

Proposition 9.11. *There is a finite set of rational primes $S \subset \mathbb{Z}$ such that if $p \notin S$ is a prime splitting in K ($pO_K = pp'$), then*

$$F(\sigma_p) = p \in \mathrm{CL}(O_K)$$

We show how using this result, we can prove that $K(j(E))$ is the Hilbert class field H of K:

Proof. Let L/K be the finite extension corresponding to the homomorphism $F : \mathrm{Gal}(\bar{K}/K) \to \mathrm{CL}(O_K)$, that is, L is the fixed field of the kernel of F. Then,

$$\mathrm{Gal}(\bar{K}/L) = \ker F = \{\sigma \in \mathrm{Gal}(\bar{K}/K) : F(\sigma) = 1\} =$$
$$= \{\sigma \in \mathrm{Gal}(\bar{K}/K) : F(\sigma) * E = E\} = \{\sigma \in \mathrm{Gal}(\bar{K}/K) : E^\sigma = E\} =$$
$$= \{\sigma \in \mathrm{Gal}(\bar{K}/K) : j(E)^\sigma = j(E)\} = \mathrm{Gal}(\bar{K}/K(j(E)))$$

We conclude that $L = K(j(E))$, and since F maps injectively $\mathrm{Gal}(L/K)$ into $\mathrm{CL}(O_K)$ we see that L/K is an abelian extension and so L is an abelian extension. Let now $c_{L/K}$ be the conductor of L/K and consider the composition of the Artin map with F

$$I(c_{L/K}) \to G_{L/K} \to \mathrm{CL}(O_K)$$

We will prove that $F((a, L/K)) = \bar{a}$ for all $a \in I(c_{L/K})$. To see this, take $a \in I(c_{L/K})$ and let S the finite set described in the proposition. From the class field theory version of Dirichlet theorem, there exists a degree one prime $p \in I(c_{L/K})$ in the same $P(c_{L/K})$ ideal class as a and not lying over a prime in S. Equivalently, we have $\alpha \equiv 1$ modulo $c_{L/K}$ and $a = (\alpha)p$. Now, just observe

$$F((a, L/K)) = F(((\alpha)p, L/K)) = F((p, L/K)) = \bar{p} = \bar{a}$$

□

The following proposition is almost immediate using the previous results:

Proposition 9.12. *Let E be an elliptic curve representing an isomorphism class in $\mathrm{ELL}(O_K)$. Then:*

a) $[\mathbb{Q}(j(E)) : \mathbb{Q}] = [K(j(E)) : K] = h_K$, where h_K is the class number of K (and it is also equal to the cardinal of $\mathrm{Gal}(H/K)$.

b) If $E_1, \ldots, E_h$ is a complete set of representatives for $\mathrm{ELL}(O_K)$, then $j(E_1), \ldots, j(E_h)$ is a complete set of Galois conjugates for $j(E)$.

c) For every prime ideal p of K,

$$j(E)^{\sigma_p} = j(\bar{p} * E)$$

and more generally for every nonzero fractional ideal a of K, then

$$j(E)^{(a,H/K)} = j(\bar{a} * E)$$

We do not prove the technical result that lead us to state the main theorems, but just quote the lemma required to show it. It says basically that isogenies behave nicely under reduction.

Proposition 9.13. *Let L be a number field, P a maximal ideal of L, and $E_1/L, E_2/L$ two elliptic curves with good reduction at P. Let $\tilde{E}_1, \tilde{E}_2$ the corresponding reductions. Then, the natural map*

$$\mathrm{Hom}(E_1, E_2) \to \mathrm{Hom}(\tilde{E}_1, \tilde{E}_2)$$

sending ϕ to $\tilde{\phi}$ is injective and preserves degrees.

9.6 Integrality of j

Here we state one of the most celebrated results in this theory: that $j(E)$ is an algebraic integer for every elliptic curve with complex multiplication. There are several approaches to this, and the ideas will be more clear after the proof of some results in chapter ten. We sketch now the idea: let Λ_1, Λ_2 the lattices corresponding to two isogenous elliptic curves $E_1/\mathbb{C}, E_2/\mathbb{C}$. We will have that $j(E_1)$ and $j(E_2)$ are algebraically dependent over $\mathbb{Q}$ by explicitly constructing a polynomial in two variables such that $F(j(E_1), j(E_2))$. Note the following: two isomorphic elliptic curves have the same j-invariant, and when they are isogenous, this is weakened but they still have some relation. When E has complex multiplication, taking $E_1 = E_2 = E$ we will have a monic polynomial with $j(E)$ as a root.

Proof. Let D_n the set of integral matrices with determinant n and S_n those that are upper-triangular, with determinant n and with $d > 0, 0 \leq b < d$.

Lemma 9.1. *Let*

$$F_n(X) = \prod_{\alpha \in S_n} (X - j(\alpha)) = \sum_m s_m X^m$$

(the coefficients are clearly the m-th elementary symmetric function in $j \circ \alpha$). Then,

a) $s_m(\gamma\tau) = s_m(\tau)$ for all $\gamma \in \mathrm{SL}_2(\mathbb{Z})$ and $\tau \in \mathbb{H}$.

b) $s_m \in \mathbb{C}[j]$.

c) The Fourier expansion of s_m has coefficients in $\mathbb{Z}$.

d) $s_m(\tau) \in \mathbb{Z}[j]$.

This told us that there is a polynomial $F_n(Y, X) \in \mathbb{Z}[Y, X]$ such that

$$\prod_{\alpha \in S_n} (X - j(\alpha) = F_n(j, X)$$

But that is not all: when $\beta \in M_2(\mathbb{Z})$ is a matrix with integer coefficients and β has positive determinant, the function $j \circ \beta$ is integral over $\mathbb{Z}[j]$ and when n is not a perfect squre $H_n(X) = F_n(X, X)$ is non constant with leading coefficient ±1.

This suffices to prove the main theorem. We begin with the case in which $R = O_K$; take an element of the ring ρ whose norm is not a perfect square, and so $[\rho]$ is an isogeny of degree n. Fixing τ such that $j(\tau) = j(E)$ multiplication by ρ send the lattice $\mathbb{Z}\tau + \mathbb{Z}$ to a sublattice of inde n. Let α be the matrix of the endomorphism in our basis, and since by definition $j \circ \alpha$ is a root of $F_n(j, X)$, substituting $X = j \circ \alpha$ and evaluating at τ, we get that

$$0 = F_n(j(\tau), j(\alpha\tau)) = F_n(j(E), j(E)) = H_n(j(E))$$

From the previous observations, we see that $j(E)$ is integral over $\mathbb{Z}$. The case when the endomorphism ring is not the whole O_K is done in a similar way but taking care of small modifications. □

This proof has a clear analytic character, and the problem can be tackled with more algebraic tools, in at least two other ways, and the other approaches have the virtue that can be generalized to abelian varieties. This idea is due to Serre and Tate and is based on the criterion of Néron, Ogg and Shafarevich.
To introduce it, consider a local field with a discrete valuation v.

Definition 9.3. *Let Σ be a set on which $G_{\bar{K}/K}$ acts. We say that Σ is unramified at v if the action of I_v, the inertia group of $G_{\bar{K}/K}$, on Σ is trivial.*

Lemma 9.2. *Let E/K be an elliptic curve such that the reduced curve $\tilde{E}/k$ is non-singular. Let $m \geq 1$ an integer prime with the characteristic ($v(m) = 0$). Then, $E[m]$ is unramified at v. Further, if l is a prime different from the characteristic, $T_l(E)$ is also unramified at v.*

We state now Néron-Ogg-Shafarevich theorem.

Theorem 9.6. *Let E/K be an elliptic curve. Then, the following are equivalent:*

1. *E has good reduction at K.*
2. *$E[m]$ is unramified at v for all integers $m \geq 1$ relatively prime to the characteristic of k.*
3. *The Tate module $T_l(E)$ is unramified at v for some (all) primes l satisfying $l \neq \mathrm{char}(k)$.*
4. *$E[m]$ is unramified at v for infinitely many integers $m \geq 1$ relatively prime to $\mathrm{char}(k)$.*

Using this, we can prove that E has potential good reduction at all primes, and from this, $j(E)$ is integral at all primes. Thus, if L is a local field and E/L an elliptic curve with complex multiplication, using that the action of $\mathrm{Gal}(\bar{L}/L)$ on the Tate module is abelian we can conclude that this action factors through a finite quotient of $\mathrm{Gal}(L^{\mathrm{ab}}/L)$. We will give a rough presentation of the ideas behind Néron models in chapter twelve and we will comment this again.

9.7 The Main Theorem of Complex Multiplication

Let K be a quadratic imaginary field with ring of integers O_K. Let E be an elliptic curve with endomorphism ring O_K, σ an automorphism of the complex numbers, σ an idele satisfying $[s, K] = \sigma|_{K^{\mathrm{ab}}}$. Further, fix

$$f : \mathbb{C}/a \to E(\mathbb{C})$$

a complex analytic isomorphism, where a is a fractional ideals. Then, there exists a unique complex isomorphism

$$f' : \mathbb{C}/s^{-1}a \to E^{\sigma}(\mathbb{C})$$

such that $f' \circ s^{-1} = \sigma \circ f$, where s^{-1} is the morphism sending K/a to $K/s^{-1}a$. This statement can be adapted for elliptic curve whose endomorphism ring is not O_K, but then some modifications are required.

Notice for instance how this theorem transforms the algebraic action of σ on the torsion subgroup $f(K/a) = E_{\text{tors}}$ into the analytic action of multiplication by s^{-1}, that is

$$f(t)^{[s,K]} = f'(s^{-1}t) \text{ for } t \in K/a \text{ and } s \in A_K^*$$

The maximal abelian extension

Forget for one moment about elliptic curves and take the multiplicative group $\mathbb{G}_m(\mathbb{C})^*$. Consider the morphism given by $z \mapsto z^N$ and take

$$\mu_N = \ker\Big(\mathbb{G}_m(\mathbb{C}^*) \to \mathbb{G}_m(\mathbb{C}^*)\Big)$$

the group of N-torsion points of $\mathbb{G}_m$. The extension $\mathbb{Q}(\mu_N)/\mathbb{Q}$ is a cyclotomic extension ramified at primes dividing N. It is easy to see that $\mathbb{Q}(\zeta) = \mathbb{Q}(\mu_N)$ is the ray class field of $\mathbb{Q}$ of conductor N.

Let now $L/\mathbb{Q}$ be an abelian extension and let N be the conductor of L. Then, for class field theory, L will be contained in the ray class field of conductor N and this is basically the content of the Kronecker-Weber Theorem. Thus, the ray class field of $\mathbb{Q}$ is generated by the value of the analytic function

$$e^{2\pi iz} = \sum_{n\geq 0} \frac{(2\pi iz)^n}{n!}$$

evaluated at points of finite order in the group $\mathbb{R}/\mathbb{Z}$. It can be seen that the action of a Frobenius element σ_p on $e^{2\pi ia/N}$ is given by

$$(e^{2\pi ia/N})^{\sigma_p} = e^{2\pi iap/N} \text{ assuming } p \nmid N$$

That way, the Galois action of σ_p is transformed into a multiplication action on the circle group.

The importance of complex multiplication is that the torsion points of an elliptic curve E with complex multiplication by O_K can be used to generate abelian extensions of K. The problem is that the torsion points themselves do not generate abelian extensions of K, but of the Hilbert class field H of K. Take so a model for E defined over H and fix a finite map

$$h : E \to E/\operatorname{Aut}(E) \cong \mathbb{P}^1$$

also defined over H. This map H will be called a Weber function. To generated abelian extensions of K, we will use the values of a Weber function on torsion points, which roughly speaking means that we will take the x-coordinates of the torsion points. Note the analogy of the following result with the cyclotomic case:

Theorem 9.7. *Let K be a quadratic imaginary field, and let E be an elliptic curve with complex multiplication by O_K. Let $h : E \to \mathbb{P}^1$ be a Weber function for E/H. Let c be an integral ideal of O_K. Then, the field*

$$K(j(E), h(E[c]))$$

is the ray class field of K modulo c. In particular,

$$K^{\mathrm{ab}} = K(j(E), h(E_{\mathrm{tors}}))$$

(if $j(E) \neq 0, 1728$ and E is an elliptic curve with coefficients in $K(j(E))$, the maximal abelian extension of K is generated by $j(E)$ and the x-coordinates of the torsion points of E).

Chapter 10

Connections between elliptic curves and modular forms

Until now, we have described elliptic curves and modular forms separately, seeing in some moments relationships between them (when dealing with L-functions, in the moduli interpretation of certain modular groups that classify elliptic curves, ...). We will now apply the preceding theory first to obtain elliptic modular curves over number fields and also to study the zeta functions of modular curves and of elliptic curves. We prove, for instance, the Eichler-Shimura relation, that leads us to prove that when we have a normalized eigenform f whose Fourier coefficients are integer numbers, then we can get an elliptic curve over $\mathbb{Q}$ with the same L-function.

10.1 $X_0(N)$ as an algebraic curve over $\mathbb{Q}$

For a congruence subgroup Γ of $\Gamma(1)$, we have seen that $\Gamma \backslash \mathbb{H}^*$ is an algebraic curve that will be defined over a certain number field. In this section we will find a canonical polynomial $F(X, Y)$ with coefficients in $\mathbb{Q}$ such that the curve $F(X, Y) = 0$ is birrationally equivalent to $X_0(N) = \Gamma_0(N) \backslash \mathbb{H}^*$. It is possible to derive explicit formulas for the genus of $X_0(N)$, or for the number of inequivalent cusps. But our interest now is to study the field of functions of $X_0(N)$.
The simplest case will be $X_0(1)$. Note that for the Riemann sphere, the meromoprhic functions are the rational functions of z and the automorphisms of $\mathbb{S}^2$ are the Mobius transformations.

Proposition 10.1. *There exists a unique meromorphic function J on $X_0(1)$ that is holomorphic except at ∞, where it has a simple pole, and that takes the value $J(i) = 1, J(\rho) = 0$. Moreover, the meromorphic functions on $X_0(1)$ are the rational functions of J.*

Proof. $X_0(1)$ has genus zero, so it is isomorphic to the Riemann sphere as a Riemann surface $\mathbb{S}^2$. If we take a map $f : X_0(1) \to \mathbb{S}^2$ that is an isomorphism and P, Q, R are the images of ρ, i, ∞, there is a unique fractional transform carrying g them to $0, 1, \infty$, so $g \circ f$ has the desired properties. In case of having another such function J' then $J' \circ J^{-1}$ is an automorphism of S fixing $0, 1, \infty$. But from the study of isomorphisms of the Riemann sphere, we have that the only ones are the Mobius transformations, and if ones fixes these three points, it must be the identity map. □

Recall that for a general lattice Λ, we have $G_{2k}(\Lambda) = \sum_{\omega \in \Lambda, \omega \neq 0} 1/\omega^{2k}$, $G_{2k}(\mathbb{Z}z + \mathbb{Z}), g_4(z) = 60G_4(z), g_6 = 140G_6(z)$. Then we had a map sending $(\wp, \wp')$ onto the curve $Y^2Z = 4X^3 - g_4(z)XZ^2 - g_6(z)Z^3$ whose discriminant was not 0 and its j-invariant was $j(z) = \frac{1728 g_4(z)^3}{\Delta}$. We can find a Fourier development of j in terms of $q = e^{2\pi i z}$,

$$j = 1/q + 744 + 196884q + 21493760q^2 + \cdots$$

We observe these three facts:

a) j is invariant under $\mathrm{SL}_2(\mathbb{Z})$ since $j(z)$ depends only on the lattice $\mathbb{Z}z + \mathbb{Z}$.

b) $j(\rho) = 0$ since $\mathbb{Z}/\rho + \mathbb{Z}$ has complex multiplication by $\rho^2 = \sqrt[3]{1}$ and so is of the form $Y^2 = X^3 + b$.

c) $j(i) = 1728$ since $\mathbb{C}/\mathbb{Z}i + \mathbb{Z}$ has complex multiplication by i and therefore is of the form $Y^2 = X^3 + aX$.

Our previous proposition can be written also in this way:

Proposition 10.2. *The function $j = 1728J$ is the unique meromorphic function on $X_0(1)$ that is holomorphic except at ∞, where it has a simple pole, and takes the values $j(i) = 1728, j(\rho) = 0$. In particular j defines an isomorphism from $X_0(1)$ onto the Riemann sphere, so the field of meromorphic functions on $X_0(1)$ is $\mathbb{C}(j)$.*

In $X_0(N)$, we define $j_N(z) = j(Nz)$. It is an easy matter to check that $j_N(\gamma z) = j_N(z)$ if $\gamma \in \Gamma_0(N)$. With these notations, we state the following theorem:

Theorem 10.1. *The field $\mathbb{C}(X_0(N))$ of modular functions for $\Gamma_0(N)$ is generated by $j(z)$ and $j(Nz)$. Further, the minimum polynomial $F(j, Y) \in \mathbb{C}(j)[Y]$ of $j(Nz)$ over $\mathbb{C}(j)$ has degree $\mu = (\Gamma(1) : \Gamma_0(N))$. $F(j, Y)$ is a polynomial in j with coefficients in $\mathbb{Z}$. When $N > 1$, $F(X, Y)$ is symmetric in X, Y, and when N is a prime p, then*

$$F(X, Y) \equiv X^{p+1} + Y^{p+1} - X^p - Y^p - XY \mod p$$

Proof. Let $\gamma \in \Gamma_0(N)$. Then, $j(N\gamma z) = j(Nz)$. Therefore, $\mathbb{C}(j(z), j(Nz))$ is contained in the field of modular functions for $\Gamma_0(N)$. Now there is an easy way to finish, that is observing that $X_0(N)$ is a covering of $X(1)$ of degree μ. From algebraic geometry, we can say that the field of meromorphic functions $\mathbb{C}(X_0(N))$ on $X_0(N)$ has degree μ over $\mathbb{C}(j)$. But let us do it explicitly:
Let $\{\gamma_1, \ldots, \gamma_\mu\}$ be a set of representatives for the right cosets of $\Gamma_0(N)$ in $\Gamma(1)$. Note that if $\gamma \in \Gamma(1)$, then $\{\gamma_1\gamma, \ldots, \gamma_\mu\gamma\}$ is also a set of right representatives. When $f(z)$ is a modular functions for $\Gamma_0(N)$, then $f(\gamma_i z)$ depends only on the

coset $\Gamma_0(N)\gamma_i$. Hence, $\{f(\gamma_i\gamma z)\}$ is a permutation of $\{f(\gamma_i z)\}$ and every symmetric polynomial in the $f(\gamma_i z)$ is invariant under $\Gamma(1)$. Such a polynomial will be a modular function for $\Gamma(1)$ and hence a rational function of j. Note also that $f(z)$ satisfies a polynomial of degree μ with coefficient in $\mathbb{C}(j)$

$$\prod (Y - f(\gamma_i z))$$

This holds for every $f \in \mathbb{C}(X_0(N))$, so $\mathbb{C}(X_0(N))$ has degree at most μ over $\mathbb{C}(j)$. Our next claim is that $f(\gamma_i z)$ are conjugate to $f(z)$ over $\mathbb{C}(j)$. Take $F(j, Y)$ to be the minimum polynomial of $f(z)$ (in particular, it will be monic and irreducible as a polynomial in Y with coefficients in $\mathbb{C}(j)$). Replace z with $\gamma_i z$ and since $j(\gamma_i z) = j(z)$, we find that $F(j(z), f(\gamma_i z)) = 0$, as we wanted. We will be done if we prove that the μ functions $j(N\gamma_i z)$ are all distinct. If $j(N\gamma_i z) = j(N\gamma_{i'} z)$ where $i \neq i'$ there will exist $\gamma \in \Gamma(1)$ such that $N\gamma_i z = \gamma N \gamma_{i'} z$ for all z. But this will force $\gamma_i \gamma_{i'}^{-1}$ to be in $\Gamma_0(N)$, which contradicts our hypothesis.
The minimum polynomial of $j(Nz)$ is

$$F(j, Y) = \prod (Y - j(N\gamma_i z))$$

The symmetric polynomials in $j(N\gamma_i z)$ are holomorphic on $\mathbb{H}$, so they must be polynomials in $j(z)$ and $F(X, Y) \in \mathbb{C}[X, Y]$. But recall that $j(z) = q^{-1} + \sum_{n=0}^{\infty} c_n q^n$, with $c_n \in \mathbb{Z}$. We also know that taking an appropriate representative, $j(N\gamma z) = j\Big(\frac{az+b}{d}\Big)$ for integers $a, b, d, ad = N$. We conclude from all this that $j(N\gamma z)$ has a Fourier expansion in powers of $q^{1/N}$ whose coefficients are in $\mathbb{Z}[e^{2\pi i/N}]$ and hence are algebraic integers. Obviously the same is true for the symmetric polynomials in $j(N\gamma_i z)$. These polynomials, that are in $\mathbb{C}[j(z)]$, will result to be polynomials in j with coefficients that are algebraic integers.
If $P = \sum c_n j^n$ and the coefficients are not algebraic integers, take the one of smallest index that is not an algebraic integer (c_m). Then, the coefficient of q^{-m} in the q-expansion is not an algebraic integer and P cannot be equal to a symmetric polynomial in the $j(N\gamma_i z)$. Thus, $F(X, Y) = \sum c_{m,n} X^m Y^n$, with $c_{m,n}$ algebraic integers and $c_{0,\mu} = 1$. Using now the q-expansion, $F(j(z), j(Nz)) = 0$ and equating coefficients we obtain a set of linear equations for $c_{m,n}$. We can see that the system is compatible with a unique solution in $\mathbb{C}$, and so also in $\mathbb{Q}$. But we already knew that they are algebraic integers, so they are in $\mathbb{Z}$.
Replacing z with $-1/Nz$ and using the invariance of j, we see that $F(Y, X)$ is a multiple of $F(X, Y)$. So, $F(Y, X) = cF(X, Y)$ and equating coefficients, $c^2 = 1$ and since $c = -1$ would imply that $F(X, X) = 0$ (and $X - Y$ would be a factor of $F(X, Y)$), $c = 1$.
Finally, if $N = p$ is a prime, the functions $j(p\gamma_i z)$ are

$$j\Big(\frac{z+m}{p}\Big)$$

where $0 \leq m \leq p-1$. Let ζ_p a p-th root of unity and m the maximal ideal $1 - \zeta_p$ in $\mathbb{Z}[\zeta_p]$. Then, $m^{p-1} = (p)$. Regarding the previous functions as power series in

q, they are all congruent modulo m, and so

$$F(j(z),Y) = (Y - j(pz)) \prod \Big(Y - j\Big(\frac{z+m}{p}\Big)\Big) =$$

$$= (Y - j(pz))(Y - j(z/p))^p = (Y - j(z)^p)(Y^p - j(z))$$

always working modulo p. We have so the last equality. □

Now, recall from algebraic geometry that every compact Riemann surface X has a unique structure of a complete non-singular algebraic curve, in this case $X_0(N)_{\mathbb{C}}$. This is the unique complete non-singular curve over $\mathbb{C}$ having the field $\mathbb{C}(j(z), j(Nz))$ of modular functions as its field of rational functions (curves-field correspondence). If we compare it with the polynomial that we have constructed, $F_N(X,Y)$, we see that this is also a curve C, but with singularities that we can remove to obtain a non-singular curve over $\mathbb{Q}$, say C'. We can embed C' into a complete regular curve $\bar{C}$; the coordinate functions x, y are rational functions on $\bar{C}$, they generate the field of rational functions on $\bar{C}$ and satisfy $F_N(x,y) = 0$. Seeing this $\bar{C}$ inside the complex field, $\bar{C}_{\mathbb{C}}$ there is an isomorphism between $\bar{C}_{\mathbb{C}} \to X_0(N)_{\mathbb{C}}$ making the rational functions x, y correspond to $j(z)$ and $j(Nz)$. Summing up, $\bar{C}$ can be seen as a model of $X_0(N)$ over $\mathbb{Q}$. The curve $X_0(N)_{\mathbb{Q}}$ is what we call the canonical model of $X_0(N)$ over $\mathbb{Q}$.

10.2 *L*-series revisited

L-series and isogeny classes

Let us recall some facts about L-series that will be useful now. In chapter seven we already pointed out the close connection between the L-series of modular forms and elliptic curves, and now we will study some deeper facts.
An isogeny $E \to E'$ defines a group homomorphism $E(\mathbb{Q}) \to E'(\mathbb{Q})$ that has a finite kernel and cokernel. Therefore, $E(\mathbb{Q})$ and $E'(\mathbb{Q})$ have the same rank (but of course, not necessarily the same torsion), but not only this: there is a result that states they also have the same number of points over a finite field.

Theorem 10.2. *Let E, E' be isogenous elliptic curves over $\mathbb{Q}$. Then $N_p(E) = N_p(E')$ and conversely, if those numbers are equal for sufficiently many good p, then E is isogenous to E'.*

Proof. We already know that $N_p(E) = \deg(1 - \phi)$, where ϕ is the Frobenius map. An isogeny $\alpha : E \to E'$ induces an isogeny $\alpha_p : E_p \to E'_p$ on the reductions modulo p that commutes with the Frobenius map. We have so that

$$\deg(\alpha) N_p(E) = N_p(E') \deg(\alpha)$$

and we are done. The converse is beyond the scope of this thesis (it was conjectured by Tate and proved by Faltings when he proved Mordell's conjecture). □

Working around the ideas of Faltings, it is possible to prove also the following important theorem:

Theorem 10.3. *Two elliptic curves E, E' are isogenous if and only if $L(E,s) = L(E',s)$.*

This is a rough way of saying that there is a one to one correspondence between isogeny classes of elliptic curves over $\mathbb{Q}$ and certain L-series. In many cases we are interested in classifying elliptic curves up to isogeny, and for that is important a theorem of Shafarevich that assures that there are only finitely many isomorphism classes of elliptic curves over $\mathbb{Q}$ with a given conductor (and hence only finitely many in each isogeny class).

L-series of modular forms revisited

Recall that we can attach to a cusp form of weight $2k$

$$f(q) = \sum_{n\geq 1} c(n)q^n$$

an L-series defined as

$$L(f,s) = \sum_{n\geq 1} c(n)n^{-s}$$

We saw that $|c(n)| \leq Cn^k$ for some constant C, and so the Dirichlet series converges for $\Re(s) > k+1$.
We introduce now an object whose importance may not be clear for the moment, but it is a rather important tool in the study of modular forms. Let $a_1, a_2, \ldots$ be a sequence of complex numbers such that $a_n = O(n^M)$ for some M. For the same reason than before, $\phi(s) = \sum a_n n^{-s}$ is absolutely convergent for $\Re(s) > M+1$, and $f(q) = \sum a_n q^n$ is absolutely convergent for $|q| < 1$. There is a clear correspondence between f and ϕ, in a similar way than in harmonic analysis we see a correspondence between a functions and its Fourier transform.
A suggesting example is the following (sometimes referred as Mellin inversion formula):

Lemma 10.1. *For every real $c > 0$,*

$$e^{-x} = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} \Gamma(s)x^{-s}ds \text{ for } x > 0$$

(where the integral is taken upwards on a vertical line).

The proof is a direct application of the residue theorem.

Definition 10.1. *The Mellin transform of $f = c(n)q^n$, or more accurately of the function $y \mapsto f(iy)$ that goes from $\mathbb{R}_{>0}$ to $\mathbb{C}$ is defined as*

$$g(s) = \int_0^\infty f(iy)y^s\frac{dy}{y}$$

The Mellin transform can be viewed as a version of the Fourier transform appropriate for the multiplicative group $\mathbb{R}_{>0}$ with invariant measure dx/x.

Lemma 10.2.

$$g(s) = (2\pi)^{-s}\Gamma(s)L(f,s)$$

Proof. Ignoring for the moment convergence issues,

$$g(s) = \int_0^\infty \sum_{n=1}^\infty c(n)e^{-2\pi ny}y^s\frac{dy}{y} = \sum_{n=1}^\infty c_n \int_0^\infty e^{-t}(2\pi n)^{-s}t^s\frac{dt}{t} =$$

$$= (2\pi)^{-s}\Gamma(s)\sum_{n=1}^\infty c(n)n^{-s} = (2\pi)^{-s}\Gamma(s)L(f,s)$$

□

This can be written in a slightly different way: if we consider $a_1, a_2, \dots$ (with $a_n = O(n^M)$ for convergence reasons), and if then $f(x) = \sum a_n e^{-nx}$, $\phi(s) = \sum a_n n^{-s}$ we say that $f(x)$ and $\phi(s)$ are the Mellin transforms of each other.
The following result is due to Hecke and was established around 1936 and it is a straightforward application of what we say about the Mellin transform.

Theorem 10.4. *Let $a_0, a_1, a_2, \dots$ be a sequence of complex numbers such that $a_n = O(n^M)$ for some M. Given $\lambda > 0, k > 0, C = \pm 1$, consider:*

a) $\phi(s) = \sum a_n n^{-s}$ ($\phi(s)$ converges for $\Re(s) > M+1$).

b) $\Phi(s) = (\frac{2\pi}{\lambda})^{-s}\Gamma(s)\phi(s)$.

c) $f(z) = \sum a_n e^{2\pi i n z/\lambda}$ (converges for $\Im(z) > 0$).

Then, these conditions are equivalent:

a) The function $\Phi(s) + \frac{a_0}{s} + \frac{Ca_0}{k-s}$ can be analytically continued to a holomorphic function on the entire complex plane, bounded on vertical strips, satisfying

$$\Phi(k-s) = C\Phi(s)$$

b) In the upper half plane, f satisfies the functional equation

$$f(-1/z) = C(z/i)^k f(z)$$

These kind of tools are very important when trying to establish convergence: for instance, in BSD we have an L-function that we know is convergent when $\Re(s) > 3/2$ but we want to determine its vanishing order at $s = 1$, so the first step must be to study its analytic continuation. We will return to this in the next chapter, since the results of Wiles around Fermat's last theorem quarantee that these functions can be continued.

We continue with some definitions: w_N is defined to be the operator acting on $\Gamma_0(N)$ such that

$$(w_N f)(z) = (\sqrt{N}z)^{2k} f(-1/z)$$

w_N preserves $S_{2k}(\Gamma_0(N))$ and is an involution. Therefore its eigenvalues must be ± 1 and so $S_{2k}(\Gamma_0(N))$ is a direct sum of the eigenspaces $S_{2k} = S_{2k}^{+1} \oplus S_{2k}^{-1}$.

Theorem 10.5. *Let $f \in S_{2k}(\Gamma_0(N))$ be a cusp form in the ϵ-eigenspace (where $\epsilon = \pm 1$). Then, f extends analytically to a holomorphic function on the whole complex plane and satisfies the functions equation*

$$\Lambda(f, s) = \epsilon(-1)^k \Lambda(f, k - s)$$

where as usual

$$\Lambda(f, s) = N^{s/2}(2\pi)^{-s}\Gamma(s)L(f, s)$$

From our previous results, we already knew that

$$S_{2k}(\Gamma_0(N)) = \bigoplus V_i$$

where the V_i are orthogonal subspaces each of which is a simultaneous eigenspace for all $T(n)$ with $(n, N) = 1$. The $T(p)$ for $p|N$ stabilize each V_i and commute, so there does exist at least one f in each V_i that is also an eigenform for the $T(p)$ with $p|N$. Scaling f such that $f = q + \sum_{n\geq 2} c(n)q^n$, then

$$L(f, s) = \prod_{p\nmid N} \frac{1}{1 - c(p)p^{-s} + p^{2k-1-2s}} \prod_{p|N} \frac{1}{1 - c(p)^{-s}}$$

Note that w_N is self adjoint for the Petersson inner product and commute with the $T(n)$ when $(n, N) = 1$, so $V_i = V_i^{+1} \oplus V_i^{-1}$, where this is a decomposition into orthogonal subspaces for w_N.
But we have a problem: w_N does not commute with the $T(p)$ when $p|N$ and so the decomposition in two subspaces is not stable under all the $T(p)$. Thus, we do not know if there is a single f that is simultaneously an eigenvector for w_n and all the $T(p)$.
Note that we have already commented similar results when at the end of chapter seven we talk about oldforms and newforms and they appear now in a natural way. If we are in $S_2(N)$ (the common notation for $S_2(\Gamma_0(N))$, there we consider two distinguished subspaces. It was a result of Atkin and Lehner that S_2^{new} decompose in a direct sum of orthogonal subspaces of dimension one, old and new (in general this holds for S_{2k}). Since the $T(p)$ for $p|N$ and w_N commute with the $T(n)$ for $(n, N) = 1$, each stabilizes each W_i. So, in that case, the functional equation is written as

$$\Lambda(f, s) = \epsilon\Lambda(f, 2 - s)$$

where $\epsilon = \pm 1$ is the eignevalues of w_n acting on W_i.

We will make a picture of our situation: begin with an elliptic curve and its L-series, $L(E, s) = \sum a_n n^{-s}$, whose coefficients are integers; it can be expressed as an Euler product and it is expected to satisfy a certain functional equation. We will have therefore a map $E \mapsto L(E, s)$, from the set of elliptic curves over $\mathbb{Q}$ under the equivalence relation given by isogeny to Dirichlet series. Falting's theorem implies that this map is injective. But on the other hand, from the theory of Atkin and Lehner we know that the subspace $S_2^{\text{new}}(N)$ decomposes into a direct sum $\oplus W_i$ of one-dimensional subspaces that are simultaneously eigenspaces for all the $T(n)$ with $(n, N) = 1$. These elements that are simultaneously eigenforms for $S_2^{\text{new}}(N)$ are called newforms (and are said to be normalized when $c(1) = 1$). In the next sections we prove that a Dirichlet L-series $c(n)n^{-s}$ is the L-series of an elliptic curve over $\mathbb{Q}$ with conductor N if it is the L-series of a normalized newform for $\Gamma_0(N)$ (under the assumption that the $c(n)$ are rational numbers). This is done in two steps: given f, construct E_f; then check that $L(E_f, s) = L(f, s)$.

Review of zeta functions

Until now, we have studied in chapter four the q-th power of the Frobenius endomorphism, ϕ, for an elliptic curve $E/\mathbb{F}_q$. There, we prove that defining $a = q + 1 - \#E(\mathbb{F}_q)$, we had

$$\phi^2 - a\phi + q = 0$$

where the equality is in the endomorphism ring. We need to develop further the theory to prove some other results:
Let Λ be a free module over a ring R and $\alpha : R \to R$ a R-linear map. It makes to sense to consider the determinant of the trace relative to some basis (and it would be independent of that choice). We quote an almost immediate lemma:

Lemma 10.3. *Let Λ be a free $\mathbb{Z}$-module of finite rank, and let $\alpha : \Lambda \to \Lambda$ be a $\mathbb{Z}$-linear map with nonzero determinant. Then, the cokernel of α is finite with order equal to $|\det(\alpha)|$ and the kernel of the map*

$$\tilde{\alpha} : (\Lambda \otimes \mathbb{Q})/\Lambda \to (\Lambda \otimes \mathbb{Q})/\Lambda$$

is finite with order $|\det(\alpha)|$.

Proof. For the general theory of modules, we can take a basis $e'_1 \cdots e'_m$ such that the matrix with respect to the bases $e_1, \cdots, e_m$ is $e'_1, \cdots, e'_m$ is $\operatorname{diag}(n_1, \cdots, n_m)$, and now is clear that $|\det(\alpha)| = n_1 \cdots n_m$ and that the cokernel is finite of that same order.
For the kernel of $\tilde{\alpha}$, just consider $\Lambda, \Lambda \otimes \mathbb{Q}, (\Lambda \otimes \mathbb{Q})/\Lambda$ and consider the respective endomorphisms $\alpha, \alpha \otimes 1, \tilde{\alpha}$. Note that the second is an isomorphisms for being nonzero the determinant, and from the snake lemma we obtain an isomorphisms between $\ker(\tilde{\alpha})$ and $\operatorname{coker}(\alpha)$. □

Using this, it is not difficult to see that the following propositions hold:

Proposition 10.3. *The degree of a nonzero endomorphism α of an elliptic curve E with $E(\mathbb{C}) = \mathbb{C}/\Lambda$ is the determinant of α acting on Λ.*

Recall that we have already introduced the Tate module, that can be thought as $T_l E = \Lambda \oplus \mathbb{Z}_l$.

Proposition 10.4. *For any nonzero endomorphism α of E, $\det(\alpha | T_l E) = \deg \alpha$.*

Further, let α be an endomorphism of an elliptic curve over a field k, and let

$$\mathrm{Tr}(\alpha) = 1 + \deg(\alpha) - \deg(1-\alpha) \in \mathbb{Z}$$

Define the characteristic polynomial of α to be

$$f_\alpha(X) = X^2 - \mathrm{Tr}(\alpha)X + \deg(\alpha) \in \mathbb{Z}[X]$$

Proposition 10.5. *The endomorphism $f_\alpha(\alpha)$ of E is zero and for $l \neq \mathrm{char}(k)$, $f_\alpha(X)$ is the characteristic polynomial of α acting on $V_l(E)$.*

Recall that for an elliptic curve $E = \mathbb{C}/\Lambda$ over $\mathbb{C}$ the degree of a nonzero endomofphism of E is the determinant of α acting on Λ. More generally, if k is algebraically closed and l is a prime different from the characteristic of k, then

$$\deg \alpha = \det(\alpha | T_l E)$$

Proposition 10.6. *Let E be an elliptic curve over $\mathbb{F}_p$. Then the trace of the Frobenius endomorphism ϕ_p on $T_l E$ is*

$$\mathrm{Tr}(\phi_p | T_l E) = a_p = p + 1 - N_p$$

Similarly, let E be an elliptic curve over $\mathbb{F}_p$. Then,

$$\mathrm{Tr}(\phi_p^{\mathrm{tr}} | T_l E) = \mathrm{Tr}(\phi_p | T_l E)$$

Proof. For any 2×2 matrix A,

$$\det(A - I_2) = \det A - \mathrm{Tr}\, A + 1$$

We apply this to the matrix of ϕ_p acting on $T_l E$ to find that

$$\deg(\phi_p - 1) = \deg(\phi_p) - \mathrm{Tr}(\phi_p | T_l E) + 1$$

But we already proved that $\deg(1 - \phi_p) = N_p$ and that $\deg(\phi_p) = p$.
For the second part, note that ϕ_p has degree p and so $\phi_p \circ \phi_p^{\mathrm{tr}} = p$. Consequently, if α, β are the eigenvalues of ϕ_p, $\alpha\beta = \deg \phi = p$. Then,

$$\mathrm{Tr}(\phi_p^{\mathrm{tr}} | T_l E) = p/\alpha + p\beta = \beta + \alpha$$

□

10.3 The ring of correspondences of a curve

Let X, X' be projective nonsingular curves over an algebraically closed field k. A correspondence T between X and X' is a pair of finite surjective regular maps $\alpha : Y \to X, \beta : Y \to X'$, also pictured as

$$X \leftarrow Y \to X'$$

that can be thought of as a many valued map $X \to X'$ that sends $P \in X(k)$ to $\{\beta(Q_i)\}$ where the Q_i are the preimages of P. That way, we have a natural map between $\mathrm{Div}(X)$ and $\mathrm{Div}(X')$, $[P] \mapsto \sum_i [\beta(Q_i)]$. Through this map the degree is multiplied by $\deg(\alpha)$ and so sends divisors of degree zero on X to divisors of degree zero on X', and also principal divisors to principal divisors, defining a map $T : J(X) \to J(X')$. The ring of correspondences $A(X)$ will be the subring of $\mathrm{End}(J(X))$ generated by maps defined by correspondences. We can consider the correspondence $X' \leftarrow Y \to X$; it will be called the transpose of the correspondence.

The Hecke correspondence

As usual, let Γ be a subgroup of $\Gamma(1)$ of finite index, and let α be a matrix with integer coefficients and positive determinant. Writing $\Gamma\alpha\Gamma = \cup\Gamma\alpha_i$, we get a map

$$T(\alpha) : J(X(\Gamma) \to J(X(\Gamma)), [z] \mapsto \sum [\alpha_i z]$$

As we already pointed out, this is the map defined by a correspondence

$$X(\Gamma) \leftarrow X(\Gamma_\alpha) \to X(\Gamma)$$

where the second arrow means multiplication by α and the first one is just an inclusion.
We are going to consider now a particular case, when $\Gamma = \Gamma_0(N)$ and $T = T(p)$. Here, the Hecke correspondence is defined by the double coset

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

If p does not divide N we can give two other characterizations of $T(p)$.

- $Y_0(N)$ over $\mathbb{C}$ can be identified with an isomorphism class (E, C), where E is an elliptic curve and C is a cyclic group of order N (or alternatively a homomorphim $\alpha : E \to E'$ of elliptic curves with kernel a cyclic group of order N). If E_p is the subgroup of points of E of order p, is isomorphic to $(\mathbb{Z}/p\mathbb{Z}) \times (\mathbb{Z}/p\mathbb{Z})$ so there are $p+1$ subgroups of order p, $S_0, S_1, \cdots, S_p$. Thus, $T(p)$ sends $\alpha : E \to E'$ to

  $$\{E_i \to E'_i \mid i = 0, 1, \cdots, p\}$$

  where $E_i = E/S_i, E'_i = E'/\alpha(S_i)$ (the kernel still has order N).

- $Y_0(N)$ is the curve C given by a polynomial $F_N(X, Y)$; take a point (j, j'). There are elliptic curves E, E' (defined up to isomorphism) such that $j = j(E), j' = j(E')$. Since $F_N(j, j') = 0$, there is a homomorphism $\alpha : E \to E'$ with kernel a cyclic subgroup of order N, so $T(p)$ maps (j, j') to $\{(j_i, j'_i) \mid i = 0, 1, \cdots, p\}$ where $j_i = j(E/S_i), j'_i = j(E'/\alpha(S_i))$.

The Eichler-Shimura relation

Recall that the curve $X_0(N)$ is defined over $\mathbb{Q}$ and $T(p)$ is defined over a number field K. It is a result that for almost every prime p that does not divide N, $X_0(N)$ still reduces to a non-singular curve $\tilde{X}_0(N)$. For one such prime, $T(p)$ defines a correspondence $\tilde{T}(p)$ on $\tilde{X}_0(N)$. We will denote by Π_q the Frobenius map from C to $C^{(q)}$ and by Π'_q its transpose. It was a result of the third chapter that the Frobenius map is either purely inseparable of degree q, and from here, we deduce that multiplication by p (when $\text{char}(k) = p$) is either purely inseparable (and so E has no points of order p) or its separable and inseparable degrees are p (and so E has p points of order dividing p). If E has no points of order p, then $j(E) \in \mathbb{F}_{p^2}$.

Theorem 10.6. *Let p be a prime where $X_0(N)$ has good reduction. Then, we have the following equality in the ring $A(\tilde{X}_0(N))$ of correspondences of $\tilde{X}_0(N)$ over the algebraic closure $\mathbb{F}$ of $\mathbb{F}_p$:*

$$\bar{T}(p) = \Pi_p + \Pi'_p$$

Proof. We proof that they coincide as many valued maps on an open subset of $\tilde{X}_0(N)$. Recall that over $\mathbb{Q}_p^{\text{al}}$ we have that $T_p(j(E), j(E')) = \{(j(E_i), j(E'_i))\}$ where $E_i = E/S_i, E'_i = E'/\alpha(S_i)$.
Take now a point $\tilde{P} \in \tilde{X}_0(N)$ with coordinates in $\mathbb{F}$; ignoring a finite number of points, assume that $\tilde{E} \in \tilde{Y}_0(N)$ and hence is of the form $(j(\tilde{E}), j(\tilde{E}'))$ for some map $\tilde{\alpha} : \tilde{E} \to \tilde{E}'$. We deal with the case where $\tilde{E}$ has p points of order dividing p. We consider a lifting α of $\tilde{\alpha}$ to $\mathbb{Q}_p^{\text{al}}$. The reduction map has kernel of order p, and let us number the subgroups of order p in E in such a way that S_0 is the kernel of this map. Each $S_i, i \neq 0$, maps to a subgroup of order p in $\tilde{E}$. So we have that the map $p : \tilde{E} \to \tilde{E}$ factors through $\tilde{E}/S_i$ as the composition $\psi \circ \phi$. When $i = 0$, ϕ is purely inseparable of degree p so ψ is separable of degree p (assuming that $\tilde{E}$ has p points of order dividing p). But recall that under these circumstances we have an isomorphism $\tilde{E}^{(p)} \to \tilde{E}/S_0$ and $\tilde{E}'^{(p)} \to \tilde{E}'/S_0$ (we have seen this when working with isogenies of elliptic curves). Therefore,

$$(j(\tilde{E}_0), j(\tilde{E}'_0)) = (j(\tilde{E}^{(p)}), j(\tilde{E}'^{(p)})) = (j(\tilde{E})^p, j(\tilde{E}')^p) = \Pi_p(j(\tilde{E}), j(\tilde{E}'))$$

When $i \neq 0$, ϕ is separable since its kernel is the reduction of S_i so ψ is purely inseparable and so $\tilde{E}$ is isomorphic to $\tilde{E}_i^{(p)}$ and $\tilde{E}'$ to $\tilde{E}_i^{\prime(p)}$. Therefore

$$(j(\tilde{E}_i)^{(p)}, j(\tilde{E}'_i)^{(p)}) = (j(\tilde{E}), j(\tilde{E}'))$$

and so $\{j(\tilde{E}_i), j(\tilde{E}'_i)) \mid i = 1, 2, \ldots, p\}$ is the inverse image of Π_p, that is, $\Pi'(j(\tilde{E}), j(\tilde{E}'))$, as we wanted to prove. □

A consequence of this is the following remarkable theorem that we have commented several times:

Theorem 10.7. *Let f be a normalized eigenform whose Fourier coefficients $a_n(f)$ are integers. Then, there is an elliptic curve E_f over $\mathbb{Q}$ such that*

$$L(E_f, s) = L(f, s)$$

The proof of this fact is basically as follows: we consider the canonical model $X_0(N)$ defined over $\mathbb{Q}$. Let $J_0(N)$ be its jacobian, that has dimension g. The Hecke correspondences give rise to endomorphisms of $J_0(N)$ defined over $\mathbb{Q}$. If I_f is the kernel of the homomorphism $\lambda : \mathbb{T} \to \mathbb{Z}$ attached to f, the quotient $J_0(N)/I_f J_0(N)$ will be the desired elliptic curve. Now, we would have to see that the L-functions are in fact equal: this is done with the Eichler-Shimura relation and will be commented after some brief digression, trying to understand a quite deep result that was pending, why Hecke operators act on the jacobian.
As it can be seen this is very related with the ideas behind the work of Wiles. Recall that the Shimura-Taniyama-Weil conjecture states:

Theorem 10.8. *Let E be an elliptic curve over $\mathbb{Q}$ with conductor N. Then, there exists a newform $f \in S_2(N)$ such that*

$$L(E, s) = L(f, s)$$

and furthermore, E is isogenous to the elliptic curve E_f obtained from f via the Eichler-Shimura construction.

In the following subsections we give some ideas that help to a better understanding of what is happening here:

Modular jacobians and Hecke operators

Recall that when we explained double cosets, denoting by $J_1(N) = \mathrm{Jac}(X_1(N))$ we have seen that the Hecke operators act naturally on $J_1(N)$ and we have a natural map

$$[\Gamma_1 \alpha \Gamma_2]_2 : \mathrm{Div}(X_2) \to \mathrm{Div}(X_1)$$

Recalling the concepts of chapter two, we see this as a composition of forward and reverse induced maps, $[\Gamma_1 \alpha \Gamma_2]_2 = (\pi_1)_D \circ \alpha_D \circ \pi_2^D$. It therefore descends to the corresponding map of Picard groups

$$[\Gamma_1 \alpha \Gamma_2]_2 = (\pi_1)_P \circ \alpha_P \circ \pi_2^P : \mathrm{Pic}^0(X_2) \to \mathrm{Pic}^0(X_1)$$

Since the holomorphic differentials $\Omega^1(X(\Gamma))$ and the weight two cusp forms $S_2(\Gamma)$ are naturally identified, $\omega : S_2(\Gamma) \to \Omega^1(X(\Gamma))$ is a linear isomorphism, whose dual spaces are naturally identified under ω^*.
Take $H_1(X(\Gamma), \mathbb{Z})$ as a subgroup of $S_2(\Gamma)^*$, and consider its corresponding image under ω^* (the jacobian is nothing but the quotient of $S_2(\Gamma)^*$ and $H_1(X(\Gamma), \mathbb{Z})$).
So, consider now X, Y, the modular curves associated to the congruence subgroups Γ_X, Γ_Y. Let $\alpha \in \mathrm{GL}_2^+(\mathbb{Q})$ be such that $\alpha \Gamma_X \alpha^{-1} \subset \Gamma_Y$ and consider the corresponding holomorphic map $h : X \to Y$ given by $h(\Gamma_X \tau) = \Gamma_Y \alpha(\tau)$. This weight-two operator on functions is compatible with the pullback on differentials in the sense that $\omega_X \circ [\alpha]_2 = h^* \circ \omega_Y$. The induced forward map on dual spaces is nothing but $h_* : S_2(\Gamma_X)^* \to S_2(\Gamma_Y)^*$ given by $h_* \phi = \phi \circ [\alpha]_2$. Similarly, if $\alpha \Gamma_X \alpha^{-1} \backslash \Gamma_Y = \cup_j \alpha \Gamma_X \alpha^{-1} \gamma_{Y,j}$, we have again that $\omega_Y \circ \sum_j [\gamma_{Y_j}]_2 = \mathrm{tr}_h \circ \omega_X$. The induced reverse map $\mathrm{tr}_h^* : S_2(\Gamma_Y)^* \to S_2(\Gamma_X)^*$ is given by $\mathrm{tr}_h^* \psi = \psi \circ \sum_j [\gamma_{Y,j}]_2$

By the compatibilities we have commented, h_*, tr_h^* descend to maps of Jacobians, that we will denote h_J and h^J. Recall that the double coset operator on divisor groups correspond to the double coset operator on modular forms (here weight two cusp forms)

$$[\Gamma_1\alpha\Gamma_2]_2 : S_2(\Gamma_1) \to S_2(\Gamma_2) \text{ given by } f[\Gamma_1\alpha\Gamma_2]_2 = \sum_J f[\beta_j]_2$$

Its corresponding pullback is

$$[\Gamma_1\alpha\Gamma_2]_2 : S_2(\Gamma_2)^* \to S_2(\Gamma_1)^* \text{ given by } [\Gamma_1\alpha\Gamma_2]_2 f = f[\Gamma_1\alpha\Gamma_2]_2$$

Now, as it occurred with the divisor map, we clearly have

$$[\Gamma_1\alpha\Gamma_2]_2 = (\pi_1)_* \circ \alpha_* \circ \mathrm{tr}_{\pi_2}^*$$

and this operators descends to a composition of maps of jacobians, given by $(\pi_1)_J \circ \alpha_J \circ \pi_2^J$.
Let us summarize: the double coset operators acts on jacobians as composition with its action on modular forms in the other direction. For the special case of Hecke operators,

Proposition 10.7. *The Hecke operators $T = T_p$ and $T = \langle d\rangle$ act by composition on the jacobian associated to $\Gamma_1(N)$.*

Action of the Hecke operators on $H_1(E, \mathbb{Z})$

We begin by recalling some basic facts from linear algebra. When V is a real vector space and we are given the structure of a complex vector space on V, that means we are given an $\mathbb{R}$ linear map $J : V \to V$ such that $J^2 = -1$. J extends by linearity to $V \otimes_{\mathbb{R}} \mathbb{C}$ and so

$$V \otimes_{\mathbb{R}} \mathbb{C} = V^+ \oplus V^-$$

with $V^\pm$ the ±1 eigenspaces of J. Then, the map

$$V \to V \otimes_{\mathbb{R}} \mathbb{C} \to V^+$$

(where the maps are $v \mapsto v \otimes 1$ and the projection) is an isomorphism of complex vector spaces and then the map $v \otimes z \mapsto v \otimes \bar{z} : V \otimes_{\mathbb{R}} \mathbb{C} \to V \otimes_{\mathbb{R}} \mathbb{C}$ is an $\mathbb{R}$-linear involution of $V \otimes_{\mathbb{R}} \mathbb{C}$ interchanging V^+ and V^-.
This can be stated in the following terms:

Proposition 10.8. *Let α be an endomorphism of V which is $\mathbb{C}$-linear. Write A for the matrix of α regarded as an $\mathbb{R}$-linear endomorphism of V, and A_1 for the matrix of α as a $\mathbb{C}$-linear endomorphism of V. Then,*

$$A \sim A_1 \oplus \bar{A}_1$$

Using that $H_1(X_0(N),\mathbb{Z})$ is a lattice in $\Omega^1(X_0(N))^*$, it is clear that

$$H_1(X_0(N),\mathbb{Z}) \oplus_{\mathbb{Z}} \mathbb{R} = \Omega^1(X_0(N))^*$$

Clearly,

$$\mathrm{Tr}(T(p)|H_1(X_0(N),\mathbb{Z})) = \mathrm{Tr}(T(p)|H_1(X_0(N),\mathbb{Z}) \otimes_{\mathbb{Z}} \mathbb{R})$$

so we can apply the preceding result to get:

Proposition 10.9. *For any p not dividing N,*

$$\mathrm{Tr}(T(p)|H_1(X_0(N),\mathbb{Z})) = \mathrm{Tr}(T(p)|\Omega^1(X_0(N))) + \overline{\mathrm{Tr}(T(p)|\Omega^1(X_0(N)))}$$

The Eichler-Shimura construction

Let us state again our previous problem with slightly different words, taking advantage of the facts we have underlined now (here, we sketch a construction of E_f that may seem different but it is not):

Theorem 10.9. *Let $f = \sum c(n)q^n$ be a normalized newform in $S_2(\Gamma_0(N))$. Assume that $c(n) \in \mathbb{Z}$ for all n. Then there exists an elliptic curve E_f and a map $\alpha : X_0(N) \to E_f$ with the following properties:*

a) α factors uniquely through $\mathrm{Jac}(X_0(N))$, that is

$$X_0(N) \to \mathrm{Jac}(X_0(N)) \to E_f$$

where the second map realizes E_f as the largest quotient of $\mathrm{Jac}(X_0(N))$ on which the endomorphism $T(n)$ and $c(n)$ of $\mathrm{Jac}(X_0(N))$ agree.

b) The inverse image of an invariant differential ω on E_f under $\mathbb{H} \to X_0(N) \to E_f$ is a nonzero rational multiple of $f dz$.

Consider now $f = c(n)q^n$ and a map $X_0(N) \to E$ as in the theorem.

Theorem 10.10. *We have that the L-function associated to the elliptic curve we construct coincides with the L-function of the given newform:*

$$c(p) = a_p = p + 1 - N_p(E)$$

To begin, assume that $X_0(N)$ has genus one. Then, $X_0(N) \to E$ is an isogeny and we can take $E = X_0(N)$. Let p be a prime not dividing N. Then E has good reduction at p for any $l \neq p$ and the reduction map $T_l E \to T_l \tilde{E}$ is an isomorphism. The Eichler-Shimura relation states that

$$\tilde{T}(p) = \Pi_p + \Pi_p^{\mathrm{tr}}$$

Taking traces on $T_l \tilde{E}$, we get

$$2c(p) = a_p + a_p$$

As we have already pointed out this works for the general case, but at some points we just have to consider the jacobian variety of the curve. Anyway, the key is that $T_l E$ is the largest quotient of $T_l \operatorname{Jac}(X_0(N))$ on which $T(p)$ acts as multiplication by $c(p)$ for all p not dividing N.

An alternative (and maybe more direct approach) would be to define

$$E_f = \cap_p \ker(T_p - a_p)$$

whose dimension, at first sight, is unknown (maybe it is trivial). But what we know is that this dimension coincides with the one of the tangent space, that is obviously the intersection of the kernels of $T_p - a_p$, when we see the T_p as endomorphisms of the tangent space to the jacobian (that is naturally identified with $S_2(N)$). By the theory of newforms, there exists exactly one modular form (and its multiples) in the kernels of all the $T_p - a_p$. This modular form is obvioulsy f. Therefore, the tangent space has dimension one and this implies that E_f is an elliptic curve.

In general, if we do not put the conditions that the $c(n)$ are integer numbers, they will generate a number field K_f of degree $d \geq 1$ and in that case the Eichler-Shimura construction produces an abelian variety A_f of dimension d with the property that $\operatorname{End}(A_f)$ contains an order of the field K_f.

10.4 Heegner points

Definition 10.2. *A Heegner point is a point in $Y_0(N)$ classifying pairs of N-isogenous elliptic curves with the same ring of endomorphisms O (modulo isomorphisms).*

If $y = (E, E')$ is a Heegner point with complex multiplication by O, then it has two associated lattices M, M' that are O-projective modules of rank one. After a homothety, we can assume that $M = a, M' = b$, with a, b O-invertible submodules of K with $a \subset b$. The ideal $n = ab^{-1}$ is an O-proper ideal (invertible) of cyclic quotient O/n of order N. Alternatively, we can think that $E = C/M, E' = C/M', M \subset M'$ and that there is an N-isogeny $\phi : C/M \to C/M'$ that is the identity of the covering spaces.
Conversely, when we have such an ideal (a proper ideal of cyclic quotient of order N), we can construct a Heegner point in a similar way: let a be an O-invertible submodule and let $[a]$ be its class in $\operatorname{Pic}(O)$. Let n be the ideal with cyclic quotient of order N and let $E = C/a, E' = C/an^{-1}$. These curves are related through the obvious isogeny whose kernel is $an^{-1}/a \simeq a/an \simeq \mathbb{Z}/N\mathbb{Z}$. Since curves E, E' only depend on the class of a in the Picard group, so we have proven the following:

Proposition 10.10. *Let O be an order in a quadratic imaginary field, and let n be an O-proper invertible ideal of cyclic quotient of order N. Then, the Heegner points with endomorphism ring O are in correspondence with* $\operatorname{Pic}(O)$.

We can then denote a Heegner point through a triplet of the form $y = (O, n, [a])$. The following proposition provides alternative characterizations of the existence of a Heegner point:

Proposition 10.11. *Let O be an order of discriminant D and let $N \in \mathbb{N}$. The following statements are equivalent:*

a) There is a Heegner point in $X_0(N)$ with endomorphism ring O.

b) There is an ideal $n \in O$ of norm N and such that O/n is cyclic.

c) There exist integers B, C with $\gcd(B, C, N) = 1$ such that $D = B^2 - 4CN$.

We have already sketched the proof of the equivalence between the first two items. The equivalence with the third one is a consequence of the theory of binary quadratic forms.
From now on, we assume that $(c, N) = 1$, where c is the conductor of the order. Write now $D = dc^2$, where d is the discriminant of the field K. The third condition is equivalent now to the fact that D is a square modulo $4N$.

The theory we have developed of complex multiplication assumed in some points that the ring we where working with was O_K; in general, we have seen that an order in a quadratic imaginary field is of the form $O = \mathbb{Z} + \mathbb{Z}c\omega_D$, where c is called the conductor. If A has complex multiplication by O, the corresponding period lattice of A is a projective O-module of rank one, and its isomorphism class depends only on the isomorphism type of A. We have, as in the other case, a bijection between elliptic curves with complex multiplication by O and rank one projective O-modules (always up to isomorphism), also called the Picard group. Since it is finite, as before, we have finitely many isomorphism classes of elliptic curves with complex multiplication by O, $A_1, \cdots, A_h$, whose j-invariants are algebraic numbers.
Again, we have a natural action of $\mathrm{Pic}(O)$ over $\mathrm{ELL}(O)$ given by

$$[\Lambda] * [A] = \mathrm{Hom}(\Lambda, A)$$

When p is a prime ideal of K of norm prime to c the inclusion $p \to O$ yields an isogeny of kernel $A[p]$ (elements in A annihilated by all elements of p). Since everything is essentially the same than before, we consider the action of the absolute Galois group G_K given by

$$\eta : G_K \to \mathrm{Pic}(O) \text{ satisfying } A^\sigma = \eta(\sigma) * A$$

Since the Picard group is commutative, the definition of η does not depend on the choice of the base curve A. The j-invariants are defined over the abelian extension $H = \bar{K}^{\ker \eta}$. From class field theory, we also know the following:

Theorem 10.11. *There exists an abelian extension H_c of K unramified outside the primes dividing c whose Galois group is identified via the Artin map, with $\mathrm{Pic}(O)$.*

Theorem 10.12. *The abelian extension H is equal to the ring class field H_c; in fact, for all primes p of K which do not divide c,*

$$\eta(\sigma_p) = [p] \in \operatorname{Pic}(O)$$

We need now some definitions:

Definition 10.3. *Consider $M_2(\mathbb{Z})$ and let $\tau \in \mathbb{H}$. We define the order associated to τ by*

$$O_\tau = \{\gamma \in M_2(\mathbb{Z}) \mid \det\gamma \neq 0, \gamma\tau = \tau\} \cup 0_{2\times 2}$$

This order O_τ is isomorphic to the endomorphism ring of the elliptic curve A_τ. In general, when O is an order in a quadratic imaginary field, we can write

$$CM(O) = \{\tau \in \mathbb{H}/\operatorname{SL}_2(\mathbb{Z}) \mid O_\tau = O\}$$

$\operatorname{Pic}(O)$ acts on $CM(O)$ as follows: a class $\alpha \in O$ can be represented by an integral ideal I such that the quotient O/I is cyclic. Choosing an I like that, the lattice $(1,\tau)I^{-1}$ is a projective O-module containing 1 as an indivisible element, so

$$\langle 1,\tau\rangle I^{-1} = \langle 1,\tau'\rangle$$

where the generator τ' is defined modulo the action of $\operatorname{SL}_2(\mathbb{Z})$ and $\alpha * \tau = \tau'$ (by definition). We have endowed $CM(O)$ with an action of $\operatorname{Pic}(O)$ compatible with the action of this group on $\operatorname{ELL}(O)$. We can reformulate so the following theorem about complex multiplication:

Theorem 10.13. *Let $K \subset \mathbb{C}$ be a quadratic imaginary field and let $\tau \in \mathbb{H} \cap K$ be an element of $\mathbb{H}$ quadratic over $\mathbb{Q}$. Then, $j(\tau)$ belongs to H, where H is the ring class field attached to the order $O = O_\tau$. More precisely, for all $\alpha \in \operatorname{Pic}(O)$ and $\tau \in CM(O)$,*

$$j(\alpha * \tau) = \operatorname{rec}(\alpha)^{-1} j(\tau)$$

We continue introducing terminology:

Definition 10.4. *The associated order of τ relative to the level N is $O_\tau^{(N)} = O_\tau \cap O_{N\tau}$ or equivalently the matrices with determinant N satisfying $\gamma\tau = \tau$ (together with the zero matrix).*

The map from $\mathbb{H}$ to $E(\mathbb{C})$ induced by the parametrization Φ_N is transcendental for being of infinite degree, so it will not take (generally) algebraic values when evaluated on algebraic arguments. But there is an exception, that is one of the most important results of this chapter:

Theorem 10.14. *Let τ be any element in $\mathbb{H}\cap K$ and let $O = O_\tau^{(N)}$ be its associated order in $M_0(N)$. Let H/K be the ring class field attached to O. Then $\Phi_N(\tau)$ belongs to $E(H)$.*

We focus our attention again on Heegner points.

Proposition 10.12. *Let O be an order of discriminant prime to N. Then the set $CM(O)$ is non empty if and only if the primes dividing N split in $K/\mathbb{Q}$.*

Proof. If $CM(O)$ is nonempty, O can be realized as a subring of $M_0(N)$. We have then a ring homomorphism from O to $\mathbb{Z}/N\mathbb{Z}$. Since the conductor of O is prime to N, all l dividing N are split in K. Alternatively, and in a more tangible way, writing $O = \langle 1, \omega \rangle$, what we have to do is to find a matrix M in $\Gamma_0(N)$ such that its trace is the trace of ω and its determinant is the norm of ω. We distinguish two cases, depending on the value of d ($K = Q(\sqrt{d})$)

a) If $d \equiv 2, 3 \mod 4$, then $\omega = c\sqrt{d}$ and $D = 4dc^2$. In this case, the trace of ω is 0 and its norm $-dc^2$. Then, if we write

$$M = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$$

we have that $\alpha + \delta = 0, \alpha\delta - \beta\gamma = -dc^2$, or what is the same

$$D = (2\alpha)^2 + 4\beta N \gamma'$$

where we have written $N\gamma' = \gamma$. From the equivalence stated above, this will have a solution.

b) When $d \equiv 1 \mod 4$, $\omega = c\left(\frac{1+\sqrt{d}}{2}\right)$ and $D = dc^2$. In that case, the trace is 1 and the norm $(1-d)c^2/4$. Imposing these conditions, we will have that

$$D = (2\alpha - 1)^2 + 4\beta\gamma$$

and again we conclude that it is possible.

Note that once we have the matrix M, it is enough with taking $\tau \in \mathbb{H}$ such that $M\tau = \tau$. □

We will call Heegner hypothesis to the assumption that all primes dividing N are split in $K/\mathbb{Q}$. Let n be any integer prime to N and let O_n be the order of K of conductor n. A point of the form $\Phi_n(\tau)$ with $\tau \in CM(O_n)$ is a Heegner point of conductor n. Let $HP(n) \in E(H_n)$ the set of all Heegner points of conductor n in $E(H_n)$, where H_n is the ring class field of K of conductor n. There is a set of norm-compatibility relation between these points:

Proposition 10.13. *Let n be an integer, and let l be a prime number, both prime to N. Let P_{nl} be any point in $HP(nl)$. Then, there exist points $P_n \in HP(n)$ and (if $l|n$) $P_{n/l} \in HP(n/l)$ such that the trace of P_{nl} relative to the extension H_{nl}/H_n is (if we call σ_λ the Frobenius element):*

a) $a_l P_n$ if l does not divide n and is inert in K.

b) $(a_l - \sigma_\lambda - \sigma_\lambda^{-1})P_n$ if $l = \lambda\bar{\lambda}$ does not divide n and splits in K.

c) $(a_l - \sigma_\lambda)P_n$ if $l = \lambda^2$ is ramified in K.

d) $a_l P_n - P_{n/l}$ *when* $l|n$.

Proof. We prove, for instance, the second relation: consider the point P_n, that is associated to a pair $A \to A'$ of N-isogenous elliptic curves. If l is a prime that splits in K, the action of $\mathrm{Gal}(\bar{K}/H_n)$ on $A[l]$ leaves invariant two cyclic subgroups of order l: $C_0 = A[\lambda]$ and $C_l = A[\bar{\lambda}]$, permuting the other $l-1$ subgroups transitively. This action factors through a simply transitive action of $\mathrm{Gal}(H_{nl}/H_n)$ on $\{C_1, \cdots, C_{l-1}\}$. Let P_{nl}^j be the point in $E(H_{nl})$ corresponding to the pair $A/C_j \to A'/\phi(C_j)$ and put $P_{nl} = P_{nl}^1$. But by the description of the Hecke operator, we clearly have that

$$a_l P_n = P_{nl}^0 + P_{nl}^l + P_{nl}^1 + \cdots + P_{nl}^{l-1}$$

Furthermore, since $P_{nl}^0 = \sigma_\lambda P_n$, $P_{nl}^l = \sigma_\lambda^{-1} P_n$ and $P_{nl}^1 + \cdots + P_{nl}^{l-1} = \mathrm{Tr}_{H_{nl}/H_n}(P_{nl})$ the result follows. □

An element $\tau \in \mathrm{Gal}(H/\mathbb{Q})$ is a reflection if its restriction to K is not the identity. It can be proved that any reflection is of order 2 and that any two reflections differ by multiplication by an element of $\mathrm{Gal}(H/K)$. We also have the following:

Proposition 10.14. *Let* $\tau \in \mathrm{Gal}(H/\mathbb{Q})$ *be a reflection. Then there is a* $\sigma \in \mathrm{Gal}(H/K)$ *such that*

$$\tau P_n = -\operatorname{sign}(E, \mathbb{Q}) \sigma P_n$$

where the equality is modulo torsion and $\operatorname{sign}(E, \mathbb{Q})$ *is the sign attached to* $E/\mathbb{Q}$ *(the one appearing in the functional equation of* $L(E, s)$*.*

Definition 10.5. *A Heegner system attached to* (E, K) *is a collection of points* $P_n \in E(H_n)$ *indexed by integers n prime to N satisfying the norm compatibilities and the behavior under the action of reflections.*

We have an important theorem in that direction:

Theorem 10.15. *If* (E, K) *satisfies the Heegner hypothesis, then there is a nontrivial Heegner system attached to* (E, K)*.*

Denote by H_∞ the union of all the class fields with conductor coprime with N. The proof of the theorem relies on the following lemma.

Lemma 10.4. *The torsion subgroup of* $E(H_\infty)$ *is finite.*

Proof. An inert prime in K splits completely or ramifies in its class fields. Then, the residual field in H_∞ of one such prime q is $\mathbb{F}_{q^2}$. Since the torsion is coprime with q it can be injectively set in $E(\mathbb{F}_{q^2})$, the whole torsion group can be set inside $E(\mathbb{F}_{q^2}) \oplus E(\mathbb{F}_{p^2})$, where p, q are different inert primes in K. □

10.5 Modular forms on quaternion algebras

Let B be an indefinite quaternion algebra over $\mathbb{Q}$. Fix an identification

$$\iota : B \otimes_{\mathbb{Q}} \mathbb{R} \simeq M_2(\mathbb{R})$$

Let R be an order in B and denote by R_1^* the group of elements of R^* of reduced norm one. Let $\Gamma = \iota(R_1^*)$.

Lemma 10.5. *Γ acts discretely on $\mathbb{H}$ with compact quotient.*

Definition 10.6. *Let N be a positive integer. The factorization $N = N^+N^-$ is an admissible factorization if:*

a) $(N^+, N^-) = 1$.

b) N^- is square-free and the product of an even number of primes.

Let B denote the quaternion algebra ramified at the primes dividing N^- (it will be unique, up to isomorphism). Choose a maximal order R_0 in B, that will be also unique up to conjugation by B^*. Since B is split at the primes dividing N^+, we can fix an identification

$$\eta : R_0 \otimes (\mathbb{Z}/N^+\mathbb{Z}) \to M_2(\mathbb{Z}/N^+\mathbb{Z})$$

Let R denote the subring of R_0 consisting of all elements x with $\eta(x)$ upper triangular. The subring R is an Eichler order of level N^+ in B, and is unique up to conjugation. With the same identification ι as before, we will have

$$\Gamma_{N^+,N^-} = \iota(R_1^*)$$

We want to translate all the properties of the classical case: Hecke operators, Petersson inner product, Atkin-Lehner.... But the problem is that when $N^{-1} \neq 1$ one does no have the notion of Fourier expansion in the cusps since the quotient of the upper half plane by these groups is compacts. It can be proved that through a similar construction of that of Eichler and Shimura, one can construct a modular parametrization

$$\Phi_{N^+,N^-} : \mathrm{Div}^0_{\mathbb{H}\backslash\Gamma_{N^+,N^-}} \to E(\mathbb{C})$$

Let us sketch briefly how to do it, explaining first the analogies of the space $S_2(\Gamma_{N^+,N^-}) = S_2(N^+, N^-)$ with the case of $S_2(N)$.

a) $S_2(N^+, N^-)$ is a Hilbert space, where the duality is given by the wedge product of differential one forms.

b) It is endowed with a natural action of Hecke operators. As usual, write

$$\Gamma\alpha\Gamma = \bigcup_{i=0}^{p} \alpha_i\Gamma$$

and define T_p by summing the translates of f by the left coset representatives α_i.

c) The Hecke operators T_n for $(n, N) = 1$ commute and are self-adjoint, so $S_2(\Gamma_{N^+,N^-})$ is diagonalisable under the action of these operators.

d) When f is an eigenform for the Hecke operators, its associated L-function can be defined as the product of the local factors (at least for those primes l not dividing N)
$$(1 - a_l(f)l^{-s} + l^{1-2s})^{-1}$$
where $T_l f = a_l f$.

As in the classical case, let f be an eigenform in $S_2(\Gamma_{N^+,N^-})$ with integer Hecke eigenvalues $a_n(f)$. One can then associate to such an eigenform an elliptic curve over $\mathbb{Q}$.

Theorem 10.16. *There exists an elliptic curve E over $\mathbb{Q}$ such that $a_n(E) = a_n(f)$ for all integers n such that $(n, N) = 1$.*

This result, combined with the modularity theorem, leads to the conclusion that for every admissible factorization N^+N^- of N and for every newform g on Γ_{N^+,N^-} with integer Hecke eigenvalues there is an associated newform f on $\Gamma_0(N)$ with the same Hecke eigenvalues as those of g at the primes l not dividing N. Not only this: it is not necessary to assume rationality of the Fourier coefficients, thanks to the following theorem of Jacquet-Langlands whose proof uses non-abelian harmonic analysis.

Theorem 10.17. *Let f be a newform on $\Gamma_0(N)$ and let $N = N^+N^-$ be an admissible factorisation of N. Then, there is a newform $g \in S_2(\Gamma_{N^+,N^-})$ with*

$$L(f, s) = L(g, s) \text{ (up to finitely many Euler factors)}$$

Combining all the previous results, we can rewrite the Shimura-Taniyama-Weil conjecture in terms of modular forms on Γ_{N^+,N^-}

Theorem 10.18. *Let $E/\mathbb{Q}$ be an elliptic curve of conductor N and let $N = N^+N^-$ be an admissible factorization of N. Then, there exists a unique eigenform $f \in S_2(\Gamma_{N^+,N^-})$ such that*
$$T_l(f) = a_l(E)f$$
for all l not dividing N.

10.6 An explicit example

Consider the maximal order in the quadratic extension $\mathbb{Q}(\sqrt{-7})$ whose discriminant is -7 and whose class number is one. Then, $O_K = \mathbb{Z} + \mathbb{Z}\left(\frac{1+\sqrt{-7}}{2}\right)$. Since -7 is a square modulo 11, 11 splits in K and verifies Heegner hypothesis. Repeating the same procedure explained in the proof, we find in $M_0(11)$ the matrix

$$\begin{pmatrix} -4 & -2 \\ 11 & 5 \end{pmatrix}$$

whose fixed point is $\tau = \frac{-9+\sqrt{-7}}{22}$. Writing $q = e^{2\pi i \tau}$ it is possible to compute its image through the modular parametrization (and with appropriate software, this example is taken from Darmon's book), and we obtain the point

$$P = (x, y) = \Big(\frac{1-\sqrt{-7}}{2}, -2-2\sqrt{-7}\Big)$$

Taking traces over $\mathbb{Q}$ we obtain the point

$$P + \bar{P} = (16, -61)$$

and since the rank of the curve is 0 this must be a torsion point (in fact, its order is five).
When the class number is different from one thins become more delicate. Take $Q(\sqrt{-6})$ and there the maximal order $O_K = \mathbb{Z}[\sqrt{-6}]$ whose discriminant is -24. Its class number is 2 and it verifies Heegner hypothesis since $-24 =\equiv 9$ is a quadratic residue modulo 11. What we have to do here is to consider the group of primitive quadratic forms of discriminant -24, and since they will represent 11 for satisfying Heegner hypothesis we can turn them into equivalent forms with the coefficient in x^2 multiple of 11. Two non-equivalent forms are

$$11x^2 + 8xy + 2y^2 \text{ and } 22x^2 + 8xy + y^2$$

A quadratic form of the form $ax^2+bxy+cy^2$ corresponds to the point $\frac{-b+\sqrt{D}}{2a} \in \mathbb{H}$. In this case, $\tau_1 = \frac{-4+\sqrt{-6}}{11}, \tau_2 = \frac{-4+\sqrt{-6}}{22}$ and calculating

$$P = \Phi_{11}(\tau_1) + \Phi_{11}(\tau_2) = (-2-\sqrt{-6}, 5) \in E(K)$$

Taking traces we obtain the point $(5, -6)$ which will be a torsion point.

Chapter 11

The Birch and Swinertonn-Dyer conjecture

11.1 Motivation

The classical statement of BSD is that the rank of an elliptic curve over the rationals (over a number field in general) equals the order of vanishing of the associated L-function at $s = 1$, where we do not even know that the function can be analytically continued.

We recall that the problem that we cannot solve, at least a priori, is the computation of the rank of $E(\mathbb{Q})$ (or more generally over a number field). Our approach when proving Mordell-Weil provides an upper bound in terms of the Selmer group, and the difference between the real rank and the upper bound is measured by the Tate-Shafarevich group, and there is no easy way to decide if an element from $S^{(2)}(E/\mathbb{Q})$ comes from an element of infinite order or gives a nontrivial element of the Tate-Shafarevich group. Call N_p the number of points over $\bar{E}(\mathbb{F}_p)$, where $\bar{E}$ is the reduction to $\mathbb{F}_p$.

Recall that if a prime p is good there is a reduction map from $E(\mathbb{Q})$ to $\bar{E}(\mathbb{F}_p)$, that in general will not be injective (for instance if $E(\mathbb{Q})$ is infinite) neither surjective (if $E(\mathbb{Q})$ is of rank zero it has at most 16 points, and by Hasse's theorem we know that in $\bar{E}(\mathbb{F}_p)$ there is at least $p + 1 - 2\sqrt{p}$ points). In the fifties, Birch and Swinertonn-Dyer suggested that if $E(\mathbb{Q})$ is large, then the same should occur with the N_p. For P a large number, define

$$f(P) = \prod_{p \leq P} \frac{N_p}{p}$$

(each quotient is approximately one). They formulate the following conjecture:

Conjecture 11.1. *For each elliptic curve E over $\mathbb{Q}$ there exists a constant C such that*

$$\lim_{P \to \infty} f(P) = C \log(P)^r$$

where r is the rank of $E(\mathbb{Q})$.

Before going to the conjecture let us look for some analogy. Take the case of the zeta function for a number field: there, if we take into account the functional equation given in chapter 4, it is immediate that the order of vanishing of the zeta function at $s = 0$ is $r_1 + r_2 - 1$, where r_1, r_2 are the number of real and complex embeddings, respectively. This is precisely the rank of the group of units, as it states Dirichlet's theorem.
Here it is very difficult to formalize this to present it as a good analogy, since there are important difference: the first one is that in the case of BSD we are evaluating the L-function at $s = 1$, that is the line of symmetry, while in the other case the line of symmetry would be $s = 1/2$. It is a general thinking that is more difficult to deal with values of L-functions over the line of symmetry, as it occurred for instance with the Riemann hypothesis. The other difference between the case of number fields is that we are taking two different objects: in BSD the rank of a certain curve, in the case of number fields the rank of the unit group. We will need to go into a deep interpretation of the cohomology groups to explain these phenomena.

Before going direct into the subject, let us try to do something similar for the case of a quadratic equation. Take the classical

$$x^2 + y^2 = 1$$

and look at the solutions it has modulo a certain prime p. The trick here is that solutions are given in a certain parametric form:

$$(x, y) = \Big(\frac{t^2 - 1}{t^2 + 1}, \frac{2t}{t^2 + 1}\Big)$$

so the number of solutions is $p - 1$ or $p + 1$ according to the fact that p is 1 or 3 modulo 4 (this determines if -1 is a square or not). Then, by Wallis' formula

$$\prod_p \frac{p}{N_p} = \frac{\pi}{4}$$

and bearing in mind that this equation has eight integer solutions we have the following:

$$\frac{N_p}{p} N_{\mathbb{R}} = 2N_{\mathbb{Z}}$$

This computation, that is ridicously easy in the quadratic case, is still unsolvable for the cubic case.

11.2 Some known results about BSD: a first insight

The first thing we must say is that since Birch and Swiertonn-Dyer formulated the conjecture, a great computational work has been made, and all the terms

appearing in the conjecture have been computed for many curves. For a pair of isogenous elliptic curves over $\mathbb{Q}$, most of the terms will differ for the two curves, but Cassels showed in 1965 that if the conjecture is true for one curve, then it holds for all those that are isogenous to it. The first positive results were over some function fields, where it is known to be true.
Tate described the problem as a relation between the behavior of an L-function at a point where it is not known to be defined to the order of a group (the Tate-Shafarevich group W) which is not known to be finite. One first result is due to Coates and Wiles in 1977:

Theorem 11.1. *Let E be a elliptic curve with complex multiplication such that $E(\mathbb{Q})$ is infinite. Then $L(E,1)=0$.*

What this says is that if the rank is greater than 0, then so is the order of vanishing of the L-function at $s=1$. The modularity theorem supposes a very remarkable tool: we can affirm that $L(E/\mathbb{Q},s)$ extends to the whole complex plane and satisfies a functional equation of the form

$$\Lambda(E,s)=\omega_E\Lambda(E,2-s)$$

where $\omega_E=\pm 1$. Recall that $\omega_E=1$ if and only if $L(E/\mathbb{Q},s)$ has a zero of even order at $s=1$.
Gross and Zagier (1983 and 1986) proved that if E is modular over $\mathbb{Q}$ (a fact that we now know), then

$$L'(E/K,1)=C\hat{h}(P_K)$$

where $C\neq 0$. That way, P_K has infinite order if and only if $L'(E/K,1)\neq 0$.
In 1988 Kolyvagin showed that if $\omega_E=1$ and P_K has infinite order for some complex quadratic extension K of $\mathbb{Q}$, then $E(\mathbb{Q})$ and $W(E/\mathbb{Q})$ are both finite. In 1989, Bump, Friedberg and Hoffstein proved that if $\omega_E=1$, there exists a complex quadratic field such that $L(E^K/\mathbb{Q},s)$ (where E^K denotes the usual quadratic twist) has a zero of order one at $s=1$, and so $L'(E/K,1)\neq 0$ if $L(E/\mathbb{Q},1)\neq 0$, taking into account the formula $L(E/K,s)=L(E/\mathbb{Q},s)\cdot L(E^K/\mathbb{Q},s)$. Combining these results, we arrive to the following result:

Theorem 11.2. *$L(E/\mathbb{Q},1)\neq 0$ implies that $E(\mathbb{Q})$ and $W(E/\mathbb{Q})$ are finite.*

Let now $E/\mathbb{Q}$ be an elliptic curve and let K be a quadratic imaginary field satisfying Heegner hypothesis respect to E. Let $\{P_n\}=\{\Phi_n(\tau_n)\}$ be a Heegner system. Consider also $P_K=\mathrm{Tr}_{H_1/K}(P_1)\in E(K)$, the trace of a Heegner point of conductor one over the Hilbert class field of K. More generally, let χ be a character in the class field of conductor n, i.e., $\chi:\mathrm{Gal}(H_n/K)\to\mathbb{C}^*$ is a primitive character in the ring class field extension of K of conductor n. We define

$$P_n^\chi=\sum_{\sigma\in\mathrm{Gal}(H_n/K)}\bar{\chi}(\sigma)P_n^\sigma\in E(H_n)\otimes\mathbb{C}$$

The following result is due to Gross, Zagier and Zhang:

Theorem 11.3. *Let $\langle , \rangle_n$ be the canonical Neron-Tate height in $E(H_n)$ extended by linearity to the pairing in $E(H_n) \otimes \mathbb{C}$. Then,*

1. $\langle P_K, P_K \rangle \simeq L'(E/K, 1)$.
2. $\langle P_n^{\chi}, P_n^{\bar{\chi}} \rangle \simeq L'(E/K, \chi, 1)$

where we have used the symbol $\simeq$ to indicate equality up to a factor.

The main consequence of this result is that the Heegner vector P_n^{χ} is nonzero if and only if $L'(E/K, \chi, 1)$ does not vanish.
We are going to give now a more precise statement of BSD:

Conjecture 11.2. *Let r be the rank of $E(\mathbb{Q})$ and let $P_1, \dots, P_r$ be linearly independent elements of $E(\mathbb{Q})$. Then,*

$$L(E, s) \sim \Big(\Omega \prod_{p \ bad} c_p\Big) \frac{[W(E/\mathbb{Q})] \det(\langle P_i, P_j \rangle)}{(E(\mathbb{Q}) : \sum \mathbb{Z}P_i)^2} (s-1)^r \ \textit{as } s \to 1$$

where $[*]$ is the order or $*$, $\Omega = \int_{E(\mathbb{R})} |\omega|$ and $c_p = (E(\mathbb{Q}_p) : E^0(\mathbb{Q}_p))$.

Let us do a few remarks. The first one is that the quotient

$$\frac{\det(\langle P_i, P_j \rangle)}{E(\mathbb{Q}) : \sum \mathbb{Z}P_i)^2}$$

is independent of the choice of $P_1, \dots P_r$ and equals

$$\frac{\operatorname{disc}\langle , \rangle}{[E(\mathbb{Q})_{\text{tors}}]^2}$$

when they form a basis for $E(\mathbb{Q})$ modulo torsion.
Note also that the following integral, $\int_{E(\mathbb{Q}_p)} |\omega|$ makes sense and is equal to $(E(\mathbb{Q}_p) : E^1(\mathbb{Q}_p))/p$. The explanation of this fact is that there is bijection between $E^1(\mathbb{Q}_p)$ and $p\mathbb{Q}_p$ under which ω corresponds to the Haar measure on $\mathbb{Z}_p$ for which $\mathbb{Z}_p$ has measure 1 and therefore $p\mathbb{Z}_p$ has measure $1/p$. Consequently

$$\int_{E(\mathbb{Q}_p)} |\omega| = (E(\mathbb{Q}_p) : E^1(\mathbb{Q}_p)) \int_{E^1(\mathbb{Q}_p)} |\omega| = \frac{c_p N_p}{p}$$

Note now that for any finite set S of prime numbers including those of bad reduction of E,

$$L_S^*(s) = \Big(\prod_{p \in S \cup \{\infty\}} \int_{E(\mathbb{Q}_p)} |\omega| \Big)^{-1} \prod_{p \notin S} \frac{1}{L_p(p^{-s})}$$

When p is good,

$$L_p(p^{-1}) = N_p/p = \Big(\int_{E(\mathbb{Q}_p)} |\omega| \Big)$$

and so the behavior of $L_S^*(s)$ near s is independent of S satisfying the condition and BSD can be stated as

$$L_S^*(E, s) \sim \frac{[W(E/\mathbb{Q})] \operatorname{disc}\langle , \rangle}{[E(\mathbb{Q})_{\text{tors}}]^2} (s-1)^r \text{ as } s \to 1$$

11.3 Relation of Heegner points with BSD

Let K be a number field and D_K its discriminant. If v is a fractional ideal, let $|v|$ its norm. When E is an elliptic curve over $\mathbb{Q}$ we can consider the L-function $L(E/K, s)$ of E over K:

$$L(E/K, s) = \prod_{v} L_v(E/K, s)$$

where $L_v(E/K, s)^{-1}$ is a polynomial of degree 1 or 2, that is $(1 - a_{|v|}|v|^{-s} + |v|^{1-2s})^{-1}$ when v does not divide N and $(1 - a_{|v|}|v|^{-s})^{-1}$ if $v|N$. If K is a quadratic field, it is easy to see that

$$L(E/K, s) = L(E, s)L(E', s)$$

where E' is any quadratic twist of E over K.
Let $\chi : \mathrm{Gal}(H/K) \to \mathbb{C}^*$ a character of the ring class field H of conductor c with $(c, N) = 1$ and put $D = D_K c^2$. The twisted L-series is defined as

$$L(E/K, \chi, s) = \prod_{v} L_v(E/K, \chi, s)$$

where $L_v(E/K, \chi, s)$ is given, when v does not divide ND, by

$$L_v(E/K, \chi, s) = (1 - \chi(\sigma_v)a_{|v|}|v|^{-s} + \chi(\sigma_v)^2|v|^{1-2s})^{-1}$$

At the infinite primes, set $L_\infty(E/K, \chi, s) = (2\pi)^{-2s}\Gamma(s)^2$.
Now, let $A = (ND)^2/\gcd(N, D)$.

Theorem 11.4. *Let*

$$\Lambda(E/K, \chi, s) = A^{s/2}L_\infty(E/K, \chi, s)L(E/K, \chi, s)$$

The L-function $L(E/K, \chi, s)$ has an analytic continuation to the entire complex plane, satisfying the functional equation

$$\Lambda(E/K, \chi, s) = \mathrm{sign}(E/K)\Lambda(E/K, \chi, 2 - s)$$

where $\mathrm{sign}(E, K) = \pm 1$ *depends only on E and K.*

Kolyvagin's theorem

The idea of the result we present in this section is that a non-trivial Heegner system yields certain lower bounds on the size of the Mordell-Weil group of E over ring class field of K. But what is more surprising is that also lead to upper bounds on the Mordell-Weil group and Tate-Shafarevich group of E/K. This result is the celebrated Kolyvagin's theorem:

Theorem 11.5. *Let $\{P_n\}_n$ be a Heegner system attached to (E, K). If P_K is non-torsion, the following facts hold:*

a) The Mordell-Weil group $E(K)$ is of rank one, and so P_K generates a subgroup of $E(K)$ of finite index.

b) The Tate-Shavarevich group is finite.

The proof of this theorem would be a good excuse to introduce a great amount of concepts from cohomology, since it requires several technical results from that area.

11.4 Sketch of the proof of Gross-Zagier-Kolyvagin theorem

Theorem 11.6. *If E is an elliptic curve over $\mathbb{Q}$ and $\operatorname{ord}_{s=1} L(E,s) \leq 1$. Then,*

$$\operatorname{rank}(E(\mathbb{Q})) = \operatorname{ord}_{s=1} L(E,s) \text{ and } W(E,\mathbb{Q}) \leq \infty$$

Proof. Denote by $\operatorname{sign}(E,\mathbb{Q})$ the sign in the function equation for $L(E,s) = L(E/\mathbb{Q},s)$. Suppose first that this sign is equal to -1:

Lemma 11.1. *There exist infinitely many quadratic Dirichlet characters ϵ such that*

a) $\epsilon(l) = 1$ for all $l|N$.

b) $\epsilon(-1) = -1$.

c) $L(E,\epsilon,1) \neq 0$.

If a characters satisfies condition a and b, then it vanishes to even order at $s = 1$ because the quadratic imaginary field satisfies the Heegner hypothesis with respect to E, and so $L(E/K,s) = L(E,s)L(E,\epsilon,s)$ vanishes to odd order at $s = 1$. If the sign is 1, for parity reason $L(E,\epsilon,1) = 0$ for all quadratic Dirichlet characters satisfying the first two conditions. Several complicated results guarantees that in this circumstances we have a character ϵ such that

$$L'(E,\epsilon,1) \neq 0$$

In any case, if K is the quadratic imaginary field associated to ϵ, by construction K satisfies the Heegner hypothesis relative to E and since $\operatorname{ord}_{s=1} L(E/K,s) = 1$ then $L'(E/K,1) \neq 0$. If $\{P_n\}$ is the Heegner system arising from the CM points on $X_0(N)$ attached to K. For previous results, this Heegner system is nontrivial in the sense that P_K is non-torsion. Using Kolyvagin's theorem, $E(K)$ has rank one and so the quotient of $E(K)$ by $\langle P_K \rangle$ is finite. Then, P_K belongs to $E(\mathbb{Q})$ up to torsion if and only if the sign is -1 and it follows that the rank of $E(\mathbb{Q})$ is equal to the order of vanishing of $L(E,s)$ at $s = 1$. Finally, the finiteness of that Shafarevich group $W(E/K)$ implies the finiteness of $W(E/\mathbb{Q})$ since the natural restriction map has finite kernel. □

11.5 Generalizations of BSD

In the last chapter we have included a section explaining the main ideas in Galois representations, that are a key tool in number theory. Here, we content with an intuitive view to explain the Galois equivariant version of BSD. The usual result is simply that when we have an elliptic curve E/F (F a number field), then the rank of $E(F)$ is $\text{ord}_{s=1} L(E/F, s)$. In this section, we will consider finite Galois extension F'/F. Some examples to bear in mind are H/K, where K is a quadratic imaginary field and H is the Hilbert class field (in this case we know that the Galois group is abelian and isomorphic to $\text{CL}(O_K)$); recall that $H/\mathbb{Q}$ has generalized dihedral Galois group.

Definition 11.1. *Let G be a group and M a module over a ring R. A representation of G in M is a morphism of groups*

$$G \to \text{Aut}_R(M)$$

The most common case is when $R = K$ is a field and $M = K^d$ is a vector space.

Consider for the sake of clarity the following example: $M = E(F')$ seen as a $\mathbb{Z}$-module, and

$$M_{\mathbb{C}} = E(F') \otimes_{\mathbb{Z}} \mathbb{C} \cong \mathbb{C}^r$$

where the last isomorphism is due to Mordell-Weil. Another thing we know is that $G = \text{Gal}(F'/F)$ acts naturally in $M_{\mathbb{C}}$ and gives a morphism

$$\rho : G \to \text{GL}_r(\mathbb{C}) = \text{GL}(M_{\mathbb{C}})$$

From the classical results of representation theory, we know that ρ decomposes uniquely as a direct sum of irreducible representations:

$$E(F') \otimes \mathbb{C} \cong \oplus V_i^{r_i}$$

We cleary have that

$$r = \sum r_i \dim(V_i)$$

Let us return to our examples: when we just have $K = \mathbb{Q}(\sqrt{-D})$ over $\mathbb{Q}$, the Galois group has two elements, the identity and another one (call it χ). Then

$$E(K) \otimes \mathbb{C} = V_1^{r_1} \oplus V_{\chi}^{r_{\chi}} = (E(\mathbb{Q}) \otimes \mathbb{C}) \oplus (E(K)^{\chi} \otimes \mathbb{C})$$

where the last summand is the set of vectors in $E(K) \otimes \mathbb{C}$ such that $\bar{v} = -v$. Observe that if $P \in E(K)$, then $P + \bar{P} \in E(\mathbb{Q})$ and that $P - \bar{P} \in E(K)^{\chi}$.
When $F = K, F' = H$, $G = \text{CL}(O_K)$. Since G is abelian, every V_i has dimension one and so what we have are the characters

$$\phi_i : G \to \mathbb{C}^* = \text{Aut}(\mathbb{C})$$

It is easy to see that there are $|G|$. Therefore,

$$E(H) \otimes_{\mathbb{Z}} \mathbb{C} \cong \oplus \mathbb{C}(\psi)^{r_i} \cong \oplus E(H)^{\psi_i}$$

where $E(H)^{\psi}$ is the set of vectors in $E(H) \otimes \mathbb{C}$ such that for all $\sigma \in G, \sigma(v) = \psi(\sigma)v$; r_i is the dimension of $E(H)^{\psi}$.
Given $P \in E(H)$, we can construct a vector $v \in E(H)^{\psi}$ as follows

$$v = \sum_{\sigma \in G} \psi^{-1}(\sigma) \otimes \sigma(P)$$

The next observation is almost trivial:

Lemma 11.2. *Let $\tau \in G$. Then, $\tau(v) = \psi(\tau)v$*

For proving it, just observe that

$$\tau(v) = \sum_{\sigma \in G} \psi^{-1}(\sigma) \otimes \tau\sigma(P) = \psi(\tau)v$$

We can do an analogous treatment for the L-functions

$$L(E/F', s) = \prod_i L(E/F, V_i, s)^{\dim(V_i)}$$

and so

$$\operatorname{ord}_{s=1} L(E/F', s) = \sum \operatorname{ord}_{s=1} L(E/F', V_i, s) \dim(V_i)$$

The Galois equivariant version of BSD states that if E/F is an elliptic curve and we consider

$$\rho : \mathrm{Gal}(F'/F) \to \mathrm{GL}(V_\rho)$$

then the order at $s = 1$ of $L(E/F, \rho, s)$ equal the multiplicity of V_ρ in $E(F') \otimes \mathbb{C}$.

Complex L-functions revisited

The definition of L-functions we made can seem a very ad-hoc construction to deal with our problems, but is something deeper, as we pointed out when we introduced it in chapter four. The appropriate framework to deal with L-functions, which will be explored later, is when working with Galois representations. Take $G_{\mathbb{Q}} = \mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q})$ and consider

$$\rho : \mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{GL}(V) = \mathrm{GL}_n(F)$$

(F a field). We will say that a prime p is unramified for ρ when $\rho(I_p) = \{\mathrm{Id}\}$ (I_p is the inertia group). In that case, we have a distinguished element $\rho(\mathrm{Frob}_p)$, good defined up to conjugation (here we are using that the image of the inertia is trivial). Let $P_{\rho,p}(T) \in F(T)$ the characteristic polynomial of $\rho(\mathrm{Frob}_p)$. Then, we define

$$L(\rho, s) = \prod_p \frac{1}{P_{\rho,p}(p^{-s})}$$

This is a very general definition and from that we can recover some of the things we did.
As a first example, let

$$\rho : G_{\mathbb{Q}} \to \mathbb{Q}^*$$

the trivial representation sending everyone to 1. It is clear that $P_{\rho,p}(T) = 1 - T$ and so

$$L(\rho, s) = \prod_p \frac{1}{1 - p^{-s}}$$

which is the classical Riemann's zeta function.

In the same way, consider now a Dirichlet character modulo N, that is, a homomorphism

$$\chi : (\mathbb{Z}/N\mathbb{Z})^* \to \mathbb{C}^* = \mathrm{GL}_1(\mathbb{C})$$

Recall that $(\mathbb{Z}/N\mathbb{Z})^* \simeq \mathrm{Gal}(\mathbb{Q}(e^{\frac{2\pi i}{N}})/\mathbb{Q})$ (so there is a canonical surjection of the absolute Galois group there, we are again very close to more general facts from the perspective of class field theory). It is almost trivial to check that here

$$L(\chi, s) = \prod_p \frac{1}{1 - \chi(p)p^{-s}} = \sum_{n \geq 1} \frac{\chi(n)}{n^{-s}}$$

which is the usual L-function used for instance in the proof of the Dirichlet's theorem of primes in arithmetic progression.

Let us now move to a more interesting example, the case of elliptic curves. Let $E/\mathbb{Q}$ be an elliptic curve, l a prime number and let

$$V = T_l(E) \otimes \mathbb{Q}_l \cong \mathbb{Q}_l \oplus \mathbb{Q}_l$$

We are interested in a representation of the form

$$\rho_{E,l} : G_{\mathbb{Q}} \to \mathrm{Aut}(T_l(E) \otimes \mathbb{Q}_l) = \mathrm{GL}_2(\mathbb{Q}_l)$$

Consider a prime not dividing $N_E \cdot l$ (N_E the conductor of the curve). When considering $\rho_{E,l}(\mathrm{Frob}_p)$, it will have a second degree characteristic polynomial in $\mathbb{Q}_l[T]$ but we have a more general result: this polynomial does not depend on l (has coefficients in $\mathbb{Q}$) and we saw that it was

$$P_{E,l}(T) = 1 - a_p(E)T + pT^2$$

Consequently, the L-series will be

$$L(E, s) = \prod_p \frac{1}{1 - a_p p^{-s} + p^{1-2s}}$$

(we sometimes add extra factors corresponding to those primes dividing the conductor, but for being only a finite number they do not modify the essential properties of the functions).

We can incorporate now the character χ of the previous example to have

$$L(E, \chi, s) = L(\rho, s) \text{ where } \rho = V_l(E) \otimes_{\mathbb{C}_l} V_\chi$$

where we are using what at first sight may seem a strange object, $\mathbb{C}_p^*$, but is just the completion of an algebraic closure of $\mathbb{Q}_p$ (of course the algebraic closure of the rationals is inside). Since V_χ has dimension one, ρ will have dimension two and we simply have

$$P_{\rho,p}(T) = \frac{1}{1-\chi(p)a_p(E)T+\chi^2(p)pT^2}$$

Consequently,

$$L(E,\chi,s) = \prod_p \frac{1}{1-\chi(p)a_pp^{-s}+\chi^2(p)p^{1-2s}}$$

An easy to prove result is the following

Lemma 11.3. *Let $\rho = \rho_1 \oplus \rho_2$. Then,*

$$L(\rho,s) = L(\rho_1,s)\cdot L(\rho_2,s)$$

The case of the tensor product is more interesting, and when we have two elliptic curves E_1, E_2 over $\mathbb{Q}$ it is natural to consider $L(E_1,E_2,s) = L(\rho,s)$, where

$$\rho = V_l(E_1)\otimes V_l(E_2)$$

It will be an interesting topic to study the analytic properties of this new L-series (Rankin L-function), that lead to some interesting results.
We are going to consider a last example to finish. Now, H will be a number field and $G_H = \mathrm{Gal}(\mathbb{Q}/H)$. Take

$$\rho : G_H \to \mathrm{GL}(V) = \mathrm{GL}_n(F)$$

Now we require a small modification in the definition of L-function over a general number field, that will be

$$L(\rho,s) = \prod_{P\subset O_H} \frac{1}{P_{\rho,\mathrm{Frob}_P}(\mathrm{Nm}_{H/\mathbb{Q}}(p)^{-s})}$$

where the product is over the nonzero ideals of O_H.
It is natural to consider for instance an elliptic curve E/K (K imaginary quadratic field) and $\rho = T_l(E)\otimes\mathbb{Q}_l$ that can be seen as a representation of G_K or G_H (now H is the Hilbert class field). A very nice property is the following

$$L(E/H,s) = \prod_\psi L(E/K,\psi,s)$$

where the product is over all the characters that go from $\mathrm{Gal}(H/K)$ to $\mathbb{C}^*$ (recall that the extension H/K is abelian with Galois group isomorphic to $\mathrm{CL}(O_K)$.
It is important the fact that $\rho = T_l(E)\otimes\mathbb{Q}_l$ is nothing but a map

$$\rho : G_H \to \mathrm{Aut}(T_l(E)\otimes\mathbb{Q}_l)$$

$T_l(E)$ is frequently called a G_H-module since it is a module over the non-commutative ring

$$\mathbb{Z}[G_H] = \oplus_{\sigma\in G_H}\mathbb{Z}\sigma$$

where the product is given by the composition.

We finish this section reviewing some ideas that where introduced when we exposed the first notions about cohomology. We introduce the concept of representation induced by a subgroup. Let $H \subset G$ be a subgroup of finite index, that is

$$G = \cup_{i=1}^{n} g_i H \text{ (disjoint union)}$$

Assume we are given a representation $\rho_H : H \to \mathrm{GL}(V)$ of the subgroup H (alternatively, a $\mathbb{Z}[H]$-module V). From ρ_H we can induce a representation of G in a natural way.
For this, consider $W = \oplus_{i=1}^{n} g_i V = \mathrm{Ind}_H^G(V)$ (that is, n copies of the vector space V). The induced representation ρ_G will map an element $g \in G$ to an element from $\mathrm{GL}(W)$, that is, to an endomorphism of $\oplus g_i V$. To describe how it acts, consider an element of the vector space, $\sum g_i v_i$ and note that $gg_i = g_{j(i)}h_i$ ($j(i)$ is nothing but an element of the symmetric group). Then, we define

$$\rho_G(g)(\sum_{i=1}^{n} g_i v_i) = \sum_{i=1}^{n} g_{j(i)} \rho_H(h_i) v_i$$

As a $\mathbb{Z}$-module, we can write

$$\mathrm{Ind}_H^G(V) \cong \mathbb{Z}[G] \otimes_{\mathbb{Z}[H]} V$$

We can relate this with the relation we had between the L-function of an elliptic curve over $\mathbb{Q}$ and a quadratic twist of it: when we have $K = \mathbb{Q}(\sqrt{-D})$, we consider D^D to be the twist of E with respect to K and we prove that

$$L(E/K, s) = L(E/\mathbb{Q}, s)L(E^D/\mathbb{Q}, s) = L(E/\mathbb{Q}, s)L(E/\mathbb{Q}, \chi, s)$$

where χ is here the order two character of $(\mathbb{Z}/D\mathbb{Z})^*$ (in the case of D prime, the Legendre symbol). In this framework of representation theory, this is equivalent to

$$L(E/K, \psi, s) = L(E/\mathbb{Q}, \mathrm{Ind}_{G_K}^{G_\mathbb{Q}}(\psi), s)$$

In the left we have $\rho = T_l(E) \otimes V_\psi$ as a G_K-representation, and in the right $\rho = T_l(E) \otimes \mathrm{Ind}(\psi)$ as a $G_\mathbb{Q}$ representation, and all we need is that $\mathrm{Ind}_{G_K}^{G_\mathbb{Q}}$ is the direct sum of ψ and $\sigma\psi$, where σ is an element in $G_\mathbb{Q}$ not in G_K. Using that in this case the corresponding characteristic polynomial is the product of the polynomial associated to each representation, we see that the L-function in the right factors as the product of $L(E/\mathbb{Q}, \psi, s)$ and $L(E^D/\mathbb{Q}, \psi, s)$.

Galois representations and BSD. A result of Darmon and Rotger.

Let $E/\mathbb{Q}$ be as usual an elliptic curve, and

$$\rho : G_\mathbb{Q} \to \mathrm{GL}(V_\rho) = \mathrm{GL}_n(\mathbb{C})$$

what we will define as an Artin representation, that is, a continuous representation of $G_{\mathbb{Q}}$. The hypothesis of continuity here is a very strong request: this forces immediately that the kernel is Galois (for being the kernel of an homomorphism) and of the form G_M, where M is a finite extension of $\mathbb{Q}$. Therefore, ρ factors through the kernel and can be seen as an injection of $\mathrm{Gal}(M/\mathbb{Q})$ in $\mathrm{GL}_n(\mathbb{C})$ (a natural question would be, for instance, which Galois groups may appear here: for instance, if $n = 1$ it must be a cyclic group).

We already mentioned the Galois equivariant version of BSD, that basically says that $L(E, \rho, s)$ vanishes at $s = 1$ with order $r(E, \rho)$, the multiplicity of the representation V_ρ in the $\mathrm{Gal}(M/\mathbb{Q})$-module $E(M) \oplus \mathbb{C}$ that is inside $\mathrm{Sel}_p(E/M)$. Obviously, they also require the same hypothesis than in their proof of BSD, that $\mathrm{ord}_{s=1} L(E, \rho, s) \leq 1$.

In 1987, Gross, Zagier and Kolyvagin proved this conjecture when $\rho = \mathrm{Ind}_{G_K}^{G_{\mathbb{Q}}}(\psi)$ for all the anti-cyclotomic (also referred as dihedral or ring class) characters.

Definition 11.2. *We say that a character ψ is anti-cyclotomic when one of the following equivalent properties hold:*

a) $H_\psi/\mathbb{Q}$ is Galois and the Galois group is dihedral (generalized). With dihedral generalized we mean that the usual semidirect product is done with an abelian (not necessarily cyclic) group.

b) $\mathrm{Ind}_{G_K}^{G_{\mathbb{Q}}}$ is an self-dual representation of $G_{\mathbb{Q}}$.

c) Let σ_0 be any element of $G_{\mathbb{Q}}$ not in G_K. We can define the character ψ' : $\mathrm{Gal}(H_\psi/K) \to \mathbb{C}^*$ by the formula $\psi'(\sigma) = \psi(\sigma_0 \sigma \sigma_0^{-1})$ and the definition does not depend of σ_0. Then, the anti-cyclotomic property is that

$$\psi' = \psi^{-1}$$

(the name cyclotomic is used when $\psi' = \psi$).

One of the main recent theorems towards BSD is the following result, from 2012, of Darmon and Rotger. One of the relevant aspects of this result is that it says something when the order of vanishing is even and non-zero:

Theorem 11.7. *Let $\rho_1, \rho_2 : G_{\mathbb{Q}} \to \mathrm{GL}_2(\mathbb{C})$ 2-dimensional odd Artin representations (odd means that the eigenvalues are 1 and -1). Let ρ be an irreducible component of the $G_{\mathbb{Q}}$-representation $\rho_1 \otimes \rho_2$. Then:*

a) If $\mathrm{ord}_{s=1} L(E, \rho, s) = 0$, then the conjecture is true for (E, ρ).

b) If $\mathrm{ord}_{s=1} L(E, \rho, s) = 1$ (and under some additional p-adic hypothesis), it is possible to build a copy of V_ρ in $\mathrm{Sel}_p(E)$.

c) If $\mathrm{ord}_{s=2} L(E, \rho, s) = 2$, then it is possible to build an injection of $V_\rho \oplus V_\rho$ in $\mathrm{Sel}_p(E)$ (again under some hypothesis).

This additional requirements say, in a rough way, that the p-adic analogous of $L''(E, \rho, 1)$ is not zero.

The proof is far beyond the scope of this thesis: the key ingredients are p-adic L-functions and the construction of a new Euler system using diagonal cycles in the cube of a modular curve.

Chapter 12

Miscellaneous topics

12.1 Galois representations

Along this thesis, we have seen that through reducing modulo p we can obtain information about a modular or an elliptic curve. For instance, we prove Eichler-Shimura relation for $X_0(N)$, that could be also stated as

$$T_p = \Pi_p + \Pi_p'$$

(where Π is the Frobenius) as an endormophism of $\mathrm{Pic}^0(\tilde{X}_0(N))$. But we have the same for elliptic curves

$$a_p(E) = \Pi_p + \Pi_p'$$

as an endomorphism of $\mathrm{Pic}^0(\tilde{E})$. The proof in that case is quite simple: just note that $x \in \bar{\mathbb{F}_p}$ satisfies $x^p = x$ if and only if $x \in \mathbb{F}_p$. Thus,

$$\tilde{E}(\mathbb{F}_p) = \ker(\sigma_p - 1)$$

and consequently

$$|\tilde{E}(\mathbb{F}_p)| = \deg(\sigma_p - 1) = (\sigma_p - 1)_* \circ (\sigma_p - 1)^* = p + 1 - \sigma_{p,*} - \sigma_p^*$$

These relations hold for all but finitely many p, and each involves different geometric objects as p varies. What we try to emphasize now is that we can lift these two relations from characteristic p to characteristic 0 (as with Hensel's lemma we lift a solution in $\mathbb{F}_p$ to a p-adic field).

For any prime l, the l-power torsion groups of an elliptic curve give rise to vector spaces $V_l(E)$ over the l-adic number field $\mathbb{Q}_l$ (we already studied this). But similarly, the l-power torsion groups of the Picard group of a modular curve give an l-adic vector space $V_l(X)$. $V_l(E)$ and $V_l(X)$ are acted on by the absolute Galois group of $\mathbb{Q}$, that subsumes the Galois groups of all number fields and contains absolute Frobenius elements Frob_P for maximal ideals P of the algebraic closure $\mathbb{Z}$ lying over rational primes p. The previous relations lead to

$$\mathrm{Frob}_P^2 - a_p(E)\,\mathrm{Frob}_P + p = 0$$

as an endomorphism of $V_l(E)$ and

$$\mathrm{Frob}_P^2 - T_p\,\mathrm{Frob}_P + p = 0$$

as an endomorphism of $V_l(X_0(N))$.

Representations of $\mathrm{End}(E)$

Let $A = \mathrm{End}(E)$. Since E has genus one, the map $\sum n_i[P_i] \mapsto \sum n_i P_i$ is a map from $\mathrm{Div}^0(E)$ to $E(k)$ that defines an isomorphism between $J(k)$ and $E(k)$. A is nothing but the full ring of correspondences of E, since there is a one to one correspondence between these two objects. There are three natural representations for A:

a) When l is a prime different from the characteristic of k, the Tate module $T_l E$ is a free $\mathbb{Z}_l$ module of rank two, so we have a homomorphism $\rho_l : A \to \mathrm{End}(T_l E)$.

b) Let $W = \mathrm{Tgt}_0(E)$ the tangent space to E at O. This is a one dimensional vector space over k, and since every element α of A fixes 0, α defines an endomorphism $d\alpha$ of W. We have another homomorphism $\rho : A \to \mathrm{End}(W)$.

c) When $k = \mathbb{C}$, $H_1(E, \mathbb{Z})$ is a free $\mathbb{Z}$-module of rank 2, and now the homomorphism is $\rho : A \to \mathrm{End}(H_1(E, \mathbb{Z}))$.

Proposition 12.1. *If $k = \mathbb{C}$, we have that $\rho \otimes \mathbb{Z}_l \ \simeq \rho_l$ and $\rho_B \otimes \mathbb{C} \simeq \rho \oplus \bar{\rho}$.*

We can even consider a forth representation, taking $\Omega^1(E)$, the space of holomorphic differentials on E. There is a canonical non-degenerate pairing $\Omega^1(E) \times \mathrm{Tgt}_0(E) \to k$. The representation of A on $\Omega^1(E)$ is the transpose of the representation on $\mathrm{Tg}\, t_0(E)$ and since both representations are one-dimensional, they are equal. We recover here some results that already appeared in some moment of the thesis and that now seem very natural in this context:

Proposition 12.2. *For every nonzero endomorphism of an elliptic curve E, the degree of α is equal to $\det(\rho_l \alpha)$. Further, let E now be an elliptic curve over $\mathbb{F}_p$. Then the numbers α_1, α_2 satisfying $N_p = 1 + p - \alpha_1 - \alpha_2$ in Hasse's theorem are the eigenvalues of Π_p acting on $T_l E$ (for $l \neq p$).*

We define now properly what we understand by a Galois representation. The philosophy will be that $\mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q})$ is not easy to understand and we cannot tell many things about it, so a natural way to study it is from its representations. In a course in Galois theory, for instance, we learn that the absolute Galois group only contains, up to conjugation, one element of finite order, the complex conjugation. With the following tool we will be able to go further in our understanding of this group.

Definition 12.1. *A Galois representation is a continuous morphism*

$$\rho : \mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{GL}_n(k)$$

where k is a field (typically, $k = \mathbb{C}, \overline{\mathbb{Q}_p}, \overline{\mathbb{F}_p}$). The first ones will be called Artin representations, the second one p-adic representations and the third ones modulo p representations. In this section G_K denotes the absolute Galois group of K.

We state some results concerning these representations:

Proposition 12.3. *The image of any representation ρ modulo p is contained in $\mathrm{GL}_n(\mathbb{F}_q)$ where $\mathbb{F}_q$ is a finite extension of $\mathbb{F}_p$. In particular, it is finite.*

Proof. Observe that the topology of $\overline{\mathbb{F}_p}$ is discrete, and so the topology of $\mathrm{GL}_n(\overline{\mathbb{F}_p})$ is also discrete. On the other hand, $\mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q})$ is compact, and since ρ is continuous, the image of ρ is a compact subset in a discrete space, so it must be finite. □

Proposition 12.4. *Every Artin representation has finite image.*

Proof. Note that in $\mathrm{GL}_n(\mathbb{C})$ we have the usual operator norm. Since ρ is continuous (and therefore bounded), we can take an open set $U \subset G_{\mathbb{Q}}$ in such a way that $\rho(U)$ is in the ball centered at Id and radius 1/2. Suppose now that there exists $u \in U$ such that $\rho(u) \neq \mathrm{Id}$, say for instance $T = \rho(u)$. If T has an eigenvalue $\lambda \neq 1$, we can take v an eigenvector of norm 1 and observe that

$$||T^n - \mathrm{Id}\,|| \geq ||(T^n - \mathrm{Id})(v)|| = |\lambda^n - 1| \geq 1/2$$

if n is big enough. This is a contradiction since T^n is in $\rho(U)$. On the other hand, if 1 is the only eigenvalue, the Jordan form J will be a matrix of ones along the diagonal and maybe some other ones. Then, the matrix $J - \mathrm{Id}$ has norm not smaller than one, and so

$$||T - \mathrm{Id}\,|| = ||J - \mathrm{Id}\,|| \geq 1/2$$

which is a contradiction again, and so $\rho(U) = \{\mathrm{Id}\}$.
Finally, since U is an open subgroup, it has finite index in $G_{\mathbb{Q}}$, from where we have that $\rho(U) = \{0\}$ has finite index in $\mathrm{Im}(\rho)$, and so the image of ρ is finite. □

Proposition 12.5. *The image of any p-adic representation ρ is contained in $\mathrm{GL}_n(F)$, where F is a finite extension of $\mathbb{Q}_p$ (in general, this image will be infinite)*

Proof. The first step is writing $\overline{\mathbb{Q}_p}$ as the union of a countable number of finite extensions of $\mathbb{Q}_p$. The maximal non-ramified extension of $\mathbb{Q}_p$ ($\mathbb{Q}_p^{\mathrm{nr}}$) is the union of $\mathbb{Q}_p(\omega_n)$, where ω_n is an n-th root of the unity and n is relatively prime with p. Then, $\overline{\mathbb{Q}_p}$ can be seen as the union of the extensions $\mathbb{Q}_p^{nr}(\sqrt[n]{p})$ where n runs over the set of natural numbers (this follows from class field theory).
Now, $\overline{\mathbb{Q}_p} = \cup F_i$ and so $\mathrm{GL}_n(\overline{\mathbb{Q}_p})$ is also $\cup\,\mathrm{GL}_n(F_i)$. Further, since $\mathrm{GL}_n(F_i)$ is closed and $\mathrm{Im}\,\rho$ is a compact written as the countable union of closed subspaces, applying Baire's theorem, at least one of the subspaces must have a nonempty interior, for instance $\mathrm{GL}_n(F_r) \cap \mathrm{Im}(\rho)$. Since it is also a subgroup, it is also open and has finite index in $\mathrm{Im}\,\rho$. We deduce from here that $\mathrm{Im}(\rho)$ is generated by $\mathrm{GL}_n(F_r) \cap \mathrm{Im}(\rho)$ and finite elements $T_1, \ldots, T_s$ and so it is included in the group generated by $\{\mathrm{GL}_n(F_r), T_1, \ldots, T_s\} \subset \mathrm{GL}_n(\mathbb{F})$ where F is the field generated by F_r and the coefficients of the T_i. □

Definition 12.2. *Given a prime $p \in \mathbb{Z}$ and a Galois representation $\rho : G_{\mathbb{Q}} \to \mathrm{GL}_n(K)$, we say that ρ is unramified in p when $\rho(I_P) = \{\mathrm{Id}\}$ for all prime P of $\bar{\mathbb{Q}}$ over p (here I_P denotes the inertia group).*

We state without proof this interesting result, relating the notion of ramification in a Galois representation with the traditional notion of ramification of a prime in an extension:

Proposition 12.6. *Let ρ be a Galois representation and $L/\mathbb{Q}$ the subextension of $\overline{\mathbb{Q}}$ fixed by* $\operatorname{Ker}(\rho)$*. Then, ρ is unramified in a prime p if and only if p does not ramify in L.*

Galois representations and modular forms

In chapter 3, we saw how to associate a Galois representation to an elliptic curve via the Tate module, that gives a map

$$\rho_{E_l} : G_{\mathbb{Q}} \to \mathrm{GL}_2(\mathbb{Z}_l) \subset \mathrm{GL}_2(\mathbb{Q}_l)$$

For that kind of representations, we have the following result whose proof is not quite complicated:

Theorem 12.1. *Let l be a prime and let E be an elliptic curve over $\mathbb{Q}$ with conductor N. The Galois representation $\rho_{E,l}$ is unramified at every prime $p \nmid lN$. For any such p, let $P \subset \bar{\mathbb{Z}}$ be any maximal ideal over p. Then, the characteristic equation of* $\rho_{E,l}(\mathrm{Frob}_P)$ *is*

$$x^2 - a_p(E)x + p = 0$$

Further, the Galois representation $\rho_{E,l}$ is irreducible.

A much more difficult result, due to Serre, is that if we let F be the field generated by the torsion points of an elliptic curve, that is,

$$F = \bigcup \mathbb{Q}\Big(\{x, y\}_{(x,y)\in E[n]}\Big)$$

then, $\mathrm{Gal}(F/\mathbb{Q})$ has finite index in $\mathrm{GL}_2(\hat{\mathbb{Z}})$ if E does not have complex multiplication.

Our objective now will be to associate Galois representations to modular curves and then decompose them into two dimensional representations associated to modular forms. If N is a positive integer and l is a prime, $X_1(N)$ is a projective non-singular algebraic curve over $\mathbb{Q}$ of genus g. Seeing the curve $X_1(N)_{\mathbb{C}}$ as a compact Riemann surface, we know that $J_1(N) \cong \mathbb{C}^g/\Lambda_g$. The Picard group of the modular curve is the abelian group of divisor classes on the points of $X_1(N)$, and it can be proved that $\mathrm{Pic}^0(X_1(N))$ can be identified with a subgroup of $\mathrm{Pic}^0(X_1(N)_{\mathbb{C}})$. Thus, there is an inclusion of the l^n-torsion

$$i_n : \mathrm{Pic}^0(X_1(N))[l^n] \to \mathrm{Pic}^0(X_1(N)_{\mathbb{C}})[l^n] \cong (\mathbb{Z}/l^n\mathbb{Z})^{2g}$$

There is a result called Igusa's Theorem that states that $X_1(N)$ has good reduction at the primes not dividing N, so there is also a natural surjective reduction map from $\mathrm{Pic}^0(X_1(N))$ to $\mathrm{Pic}^0(\tilde{X}_1(N))$ restricting to the l^n-torsion (call this map π_n). The following result relies in techniques of algebraic geometry.

Theorem 12.2. *The inclusion i_n is an isomorphism and the surjection π_n is also an isomorphism if p does not divide lN.*

We define the l-adic Tate module of $X_1(N)$, $\mathrm{Ta}_l(\mathrm{Pic}^0(X_1(N))$ as the projective limit of $\mathrm{Pic}^0(X_1(N))[l^n]$. It is clear that any automorphism σ of the absolute Galois group defines an automorphism of $\mathrm{Div}^0(X_1(N))$ that descends to $\mathrm{Pic}^0(X_1(N))$. Again, this will lead to a continuous homomorphism

$$\rho_{X_1(N),l} : G_{\mathbb{Q}} \to \mathrm{GL}_{2g}(\mathbb{Z}_l) \subset \mathrm{GL}_{2g}(\mathbb{Q}_l)$$

(this is the $2g$-dimensional representation associated to $X_1(N)$.
We had previously defined the Hecke algebra over $\mathbb{Z}$ as the algebra of endomorphisms of $S_2(\Gamma_1(N))$ generated over $\mathbb{Z}$ by the Hecke and diamond operators. This algebra acts on $\mathrm{Pic}^0(X_1(N))$ and since the action is linear it restricts to l-power torsion and it extends to $\mathrm{Ta}_l(\mathrm{Pic}^0(X_1(N)))$. Not only that: the Hecke action is defined over $\mathbb{Q}$ and so the Galois action and the Hecke action on $\mathrm{Pic}^0(X_1(N))$ commute and therefore also the two actions on $\mathrm{Ta}_l(\mathrm{Pic}^0(X_1(N)))$.

Theorem 12.3. *Let l be a prime number and N be a positive integer. The Galois representation $\rho_{X_1(N),l}$ is unramified at every prime p not dividing lN. For any such p let $P \subset \bar{\mathbb{Z}}$ (the bar denoting the integral closure of $\mathbb{Z}$ in $\bar{\mathbb{Q}}$) be a maximal ideal over p. Then, $\rho_{X_1(N),l}(\mathrm{Frob}_P)$ satisfies the equation*

$$x^2 - T_p x + \langle p \rangle p = 0$$

Recall also that the Hecke algebra contains an ideal associated to f, the kernel of the eigenvalue map I_f; the abelian variety of f is defined as $A_f = J_1(N)/I_f J_1(N)$. It can be seen that if $\mathbb{T}_{\mathbb{Z}}$ is the Hecke algebra, then $\mathbb{T}_{\mathbb{Z}}/I_f$ is isomorphic to O_f, where

$$O_f = \mathbb{Z}[\{a_n(f) : n \in \mathbb{Z}^+\}]$$

Let $\mathrm{Ta}_l(A_f)$ be the projective limit of $A_f[l^n]$. The action of O_f on A_f is defined on l-power torsion and the extends to $\mathrm{Ta}_l(A_f)$.

Proposition 12.7. *The map $\mathrm{Pic}^0(X_1(N))[l^n] \to A_f[l^n]$ is a surjection whose kernel is stable under $G_{\mathbb{Q}}$.*

So $G_{\mathbb{Q}}$ acts on $A_f[l^n]$ and so on $\mathrm{Ta}_l(A_f)$. The action commutes with the action of O_f since the $G_{\mathbb{Q}}$ action and the $\mathbb{T}_{\mathbb{Z}}$ action commute on $\mathrm{Ta}_l(\mathrm{Pic}^0(X_1(N)))$. Choosing coordinates what we have is a Galois representation

$$\rho_{A_f,l} : G_{\mathbb{Q}} \to \mathrm{GL}_{2d}(\mathbb{Q}_l)$$

The representation is continuous since $\rho_{X_1(N),l}$ is continuous and

$$\rho^{-1}_{X_1(N),l}(U(n,g)) \subset \rho^{-1}_{A_f,l}(U(n,d))$$

where $U(n,g)$ is the kernel of $\mathrm{GL}_{2g}(\mathbb{Z}_l) \to \mathrm{GL}_{2g}(\mathbb{Z}/l^n\mathbb{Z}))$ and similarly for $U(n,d)$. The representation is unramified at the primes p not dividing lN since the kernel contains $\ker \rho_{X_1(N),l}$. For those p, let $P \subset \bar{\mathbb{Z}}$ be any maximal ideal over p. Then, since T_p acts as $a_p(f)$ and $\langle p \rangle$ acts as $\chi(p)$, at the level of abelian varieties we have that $\rho_{A_f,l}(\mathrm{Frob}_P)$ satisfies the equation

$$x^2 - a_p(f)x + \chi(p)p = 0$$

The Tate module $\mathrm{Ta}_l(A_f)$ has rank $2d$ over $\mathbb{Z}_l$, and since it is an O_f-module, then $V_l(A_f) = \mathrm{Ta}_l(A_f) \otimes \mathbb{Q}$ is a module over $O_f \otimes \mathbb{Q}_l = K_f \otimes_{\mathbb{Q}} \mathbb{Q}_l$.

Lemma 12.1. *$V_l(A_f)$ is a free module of rank 2 over $K_f \otimes_{\mathbb{Q}} \mathbb{Q}_l$.*

In the next section we will describe a special type of representations, those called modular representations.

12.2 A brief insight into Fermat's last theorem

Let S denote the usual sphere $\mathbb{S}^2$. Then, $\pi = \pi_1(S \backslash \{P_1, \dots, P_s\}, O)$ is generated by $\gamma_1, \dots, \gamma_s$ loops around each of the points P_i. π classifies the coverings of S unramified except over $P_1, \dots, P_s$. The idea is that we want to do an analogy of this fact over $\mathbb{Q}$. When K is a number field, take O_K the ring of integers and there any ideal factors as a product of prime ideals $pO_K = \prod P^{e_P}$; p is said to be unramified when $e_P = 1$ for each p. Let now the extension be Galois with Galois group G. Recall from basic algebraic number theory that when P is a prime ideal diving pO_K, we can consider $G(P)$, the subgroup of G formed by those σ such that $\sigma P = P$. The action of $G(P)$ on the residue field $O_K/P = K(P)$ defines a surjection from $G(P)$ to $\mathrm{Gal}(k(P)/\mathbb{F}_p)$, which is an isomorphism if and only if p is unramified in K. Recall also that in $G(P)$ we have a distinguished element, the Frobenius element at P, F_P. Finally, note that if P' also divides pO_K, there exists a $\sigma \in G$ such that $\sigma P = P'$ and so the Frobenius of P and P' are conjugated by σ and so the conjugacy class of F_P depends on p (and we will write it as F_p). This can be generalized to infinite extensions. Let S be a finite nonempty set of prime numbers and let K_S be the union of all $K \subset \mathbb{C}$ that are of finite degree over $\mathbb{Q}$ and unramified outside S. For each $p \in S$ there is an element $F_p \in \mathrm{Gal}(K_S/\mathbb{Q})$ well defined up to conjugation called the Frobenius element at p.

Proposition 12.8. *Let E be an elliptic curve over $\mathbb{Q}$. Let l be a prime and let*

$$S = \{p \mid E \text{ has bad reduction at } p\} \cup \{l\}$$

Then all points of order l^n on E have coordinates in K_S, that is $E(K_S)_{l^n} = E(\mathbb{Q}^{\mathrm{al}})_{l^n}$ for all n.

Let now T_lE be the Tate module of $E_{k^{\mathrm{al}}}$. Thus, for E over $\mathbb{Q}$ and S as in the proposition T_lE is a free $\mathbb{Z}_l$-module of rank two such that

$$T_lE/l^nT_lE = E(K_S)_{l^n} = E(\mathbb{Q}^{\mathrm{al}})_{l^n}$$

for all n. The action of G_S on the quotients defines a continuous action of G_S on T_lE, that is, a representation

$$\rho_l : G_S \to \mathrm{Aut}_{\mathbb{Z}_l}(T_lE) \approx \mathrm{GL}_2(\mathbb{Z}_l)$$

Proposition 12.9. *Let E, l, S as above. For all $p \notin S$*

$$\mathrm{Tr}(\rho_l(F_p)|T_lE) = p + 1 - N_p(E) = a_p$$

Definition 12.3. *A continuous homomorphism $\rho : G_S \to \mathrm{GL}_2(\mathbb{Z}_l)$ is modular if $\mathrm{Tr}(\rho(F_p)) \in \mathbb{Z}$ for all $p \notin S$ and there exists a cusp form $f = \sum c(n)q^n \in S_{2k}(\Gamma_0(N))$ for some k and N such that*

$$\mathrm{Tr}(\rho(F_p)) = c(p)$$

for all $p \notin S$.

To prove that E is modular what we must prove is that $\rho_l : G_s \to \mathrm{Aut}(T_lE)$ is modular for some l (and in that case ρ_l will be modular for all l). Similarly, a continuous homomorphism $\rho : G_S \to \mathrm{GL}_2(\mathbb{F}_l)$ is modular if there exists a cusp form $f = \sum c(n)q^n$ in $S_{2k}(\Gamma_0(N))$ for some k and N such that $\mathrm{Tr}(\rho(F_p)) \equiv c(p)$ modulo l for all $p \notin S$. A representation is odd if $\det \rho(c) = -1$, where c is complex conjugation. A conjecture due to Serre states that every odd irreducible representation $\rho : G_S \to \mathrm{GL}_2(\mathbb{F}_l)$ is modular (here irreducible means that there is no one dimensional subspace of $\mathbb{F}_l^2$ stable under the action of G_S). One of the classical results is the following, due to Langlands and Tunnell:

Theorem 12.4. *If $\rho : G_S \to \mathrm{GL}_2(\mathbb{F}_3)$ is odd and irreducible, then it is modular.*

Let now R be a complete local noetherian ring with residue field $\mathbb{F}_l$. Two homomorphisms $\rho_1, \rho_2 \to \mathrm{GL}_2(R)$ are strictly equivalent if $\rho_1 = M\rho_2M^{-1}$, where $M \in \ker(\mathrm{GL}_2(R) \to \mathrm{GL}_2(k))$. A deformation of ρ_0 is a strict equivalence class of homomorphisms $\rho : G_s \to \mathrm{GL}_2(R)$ whose composite with $\mathrm{GL}_2(R) \to \mathrm{GL}_2(\mathbb{F}_p)$ is ρ_0. Now, if we put a set of conditions $*$ on representations ρ, under suitable hypothesis there is a universal $*$-deformation of ρ_0, that is, a ring $\tilde{R}$ and a deformation $\tilde{\rho} : G_s \to \mathrm{GL}_2(\tilde{R})$ satisfying $*$ and such that any other representation with that property factors through $\tilde{\rho}$.
Roughly speaking, one of the strategies of Wiles was, first, state conditions $*$ as strong as possible but satisfied by the representation of G_S on T_lE for E a semistable elliptic curve over $\mathbb{Q}$. Fixing a modular representation ρ_0, we get a homomorphism $\delta : \tilde{R} \to \mathbb{T}$, which is in fact an isomorphism. We cannot go further since a proper explanation of this ideas would be extremely complicated. The philosophy is that both Wiles' proof as some of the main theorems around BSD needs, in some way or another, the concept of Galois representations and modular representations.

12.3 Elliptic surfaces

There are several ways to introduce elliptic surfaces. They can be viewed as one-parameter algebraic families of elliptic curves, algebraic surfaces containing a pencil of elliptic curves or elliptic curves over one-dimensional function fields. We will begin with this last approach and we will assume that we are working all the time over a field k of characteristic zero. We will consider an elliptic curve of the form

$$y^2 = x^3 + A(T)x + B(T)$$

where $A(T), B(T) \in k(T)$ are rational functions of the parameter T. For most values of $t \in \bar{k}$ we can substitute $T = t$ and get an elliptic curve

$$E_t : y^2 = x^3 + A(t)x + B(t)$$

(this will happen when $\Delta(t) = -16(4a^3 + 27b^2) \neq 0$).
Our interest will be now in proving the weak Mordell-Weil theorem for elliptic curves defined over function fields in characteristic zero.

Theorem 12.5. *Let k be an algebraically closed field, $K = k(C)$ the function field of a curve and E/K an elliptic curve. Then, $E(K)/2E(K)$ is finite.*

We give just an idea of the proof. Recall the steps in the proof of this theorem for number fields: we begin by saying that the extension field $L = K([m]^{-1}E(K))$ is an abelian extension of K of exponent m unramified outside a finite set of primes S. This will work in the same way for function fields by developing the theory of valuations. In the second part we will prove with Kummer theory that the maximal abelian extension of K of exponent m unramified outside of S is a finite extension. We need in this part to translate some concepts from algebraic number theory involved in results of finiteness: for instance, the unit group is nothing but the set of elements satisfying $v(\alpha) = 0$ for all discrete valuations on K^*. But the discrete valuations on a function field $K = k(C)$ correspond to the points of $C(k)$, since k is algebraically closed. Thus if a function f has valuation zero for all valuations, it has no zeros or poles so it is constant. Hence the unit group of K will be k^* (that is not finitely generated). Furthermore, the ideal class group will be now the Picard group, that is neither finitely generated.
But we need not such a strong statement. We only used the facts that the ideal class group has only finitely many elements of order m and that the unit group R^* has the property that the quotient R^*/R^{*m} is finite. These results remain true for function fields under certain assumptions on the constant field k of K. The next proposition is one of the key facts:

Proposition 12.10. *Let C be a non-singular projective curve over an algebraically closed field k. Then, for any integer $m \geq 1$, the Picard group has only finitely many elements of order m.*

Another important observation is that if E/K is the elliptic curve, that admits a Weierstrass equation of the form

$$y^2 = (x - e_1)(x - e_2)(x - e_3)$$

where $e_1, e_2, e_3 \in K$, and if $S \subset C$ is a set of points where any one of e_1, e_2, e_3 has a pole, together with those points where $\Delta = (e_1 - e_2)^2(e_2 - e_3)^2(e_3 - e_1)^2$ vanishes. Then, for any point $P = (x, y) \in E(K)$ with $x \neq e_1$,

$$\text{ord}_t(x - e_1) \equiv 0 \pmod 2 \text{ for all } t \in C \text{ with } t \notin S$$

where ord_t is the normalized valuation on $k(C)$ which measures the order of vanishing of a function at t.
The following lemma, combined with this technical observation we have made, finishes the proof of the weak Mordell-Weil theorem:

Lemma 12.2. *Let k be an algebraically closed field, $K = k(C)$ the function field of a curve and $S \subset C$ a finite set of points. Let $m \geq 1$ be an integer. Then, the group*

$$K(S, m) = \{f \in K^*/K^{*m} \mid \mathrm{ord}_t(f) \equiv 0 \textit{ (mod } m\textit{) for all } t \notin S\}$$

is a finite subgroup of K^*/K^{*m}.

As we have said, it is interesting to consider elliptic surfaces as a one-parameter family of elliptic curves. We might consider a family

$$E_T : y^2 = x^3 + A(T)x + B(T)$$

with $A(T), B(T) \in k(T)$. Or more generally we can fix a projective curve C/k and take $A, B \in k(C)$ with $4A^3 + 27B^2 \neq 0$. Then, for almost all points $t \in C(\bar{k})$ we can evaluate A and B at t to get an elliptic curve E_t. Consider now

$$\epsilon = \{([X, Y, Z], t) \in \mathbb{P}^2 \times C \mid Y^2Z = X^3 + A(t)XZ^2 + B(t)Z^3\}$$

ϵ is a subvariety of $\mathbb{P}^2 \times C$ of dimension two (a surface formed from a family of elliptic curves).
Since ϵ is a subvariety of $\mathbb{P}^2 \times C$, projection onto the second factor defines a morphism $\pi : \epsilon \to C$ sending $([X, Y, Z], t) \mapsto t$. Further, for almost every point $t \in C$, the fiber $\epsilon_t = \pi^{-1}(t)$ is the curve E_t we considered earlier.
But note that our family of elliptic curves has other important property. An elliptic curve is nothing but a pair (E, O) where E is a curve of genus one and O is a point of E. The equation that defines ϵ gives a one-parameter family of elliptic curves, that is, for almost all values of t we get an elliptic curve ϵ_t, which means a pair (ϵ_t, O_t), that is, each one is equipped with a zero element. The interesting property here is that the collection of zero elements O_t is an algebraic family of points, and so, since each fiber ϵ_t is an elliptic curve with zero element O_t we get a map $\sigma_0 : C \to \epsilon$ mapping $t \mapsto O_t$. This map satisfies that $\pi(\sigma_0(t)) = t$ for all $t \in C(\bar{k})$. Since O_t is an algebraic family, σ_0 is a rational map of varieties and σ is a section.

Definition 12.4. *Let C be a non-singular projective curve. An elliptic surface over C consists of the following data:*

a) A surface ϵ (a two dimensional projective variety).

b) A morphism $\pi : \epsilon \to C$ such that for all but finitely many points $t \in C(\bar{k})$ and the fiber $\epsilon_t = \pi^{-1}(t)$ is a non-singular curve of genus one.

c) A section to π, $\sigma_0 : C \to E$.

12.4 Néron models

This is a deep topic that needs a good knowledge of algebraic geometry to be properly understood. We content ourselves with giving an intuitive introduction to it. Recall that when we have an elliptic curve $E/\mathbb{Q}$, we say that a prime is of good reduction if there exists $f(x, y) \in \mathbb{Z}[x, y]$ such that the curve E over $\mathbb{Q}$ is isomorphic to $\{f(x, y) = 0\}$ and at the same time the curve $\{\bar{f}(x, y) = 0\}$ is non-singular, where the bar denotes reduction modulo p. Recall that there may be a model where the reduction is singular and the curve can still be of good reduction.
If p is a prime of good reduction, we can see E as a curve over $\mathbb{Q}_p$ and consider $f(x, y) \in \mathbb{Z}[x, y]$ isomorphic to E over $\mathbb{Q}_p$. It is perfectly possible to have two polynomials over $\mathbb{Z}_p$, f_1, f_2 such that E is isomorphic to $\{f_1 = 0\}$ and to $\{f_2 = 0\}$ over $\mathbb{Q}_p$ but the curves associated to f_1 and to f_2 are not isomorphic over $\mathbb{Z}_p$ (in the same way that for instance two curves could be isomorphic over a number field and not over $\mathbb{Q}$). We will say here that f_1 and f_2 are distinct p-adic integer models for $E/\mathbb{Q}_p$.

Definition 12.5. *The Néron model of $E/\mathbb{Q}_p$ is a scheme $\epsilon/\mathbb{Z}_p$ such that $\epsilon/\mathbb{Q}_p \cong E/\mathbb{Q}_p$ and such that is optimal and canonical in a certain sense.*

Before continuing with this introduction, we recall the following:

Proposition 12.11. *If p is of good reduction, there is a unique model for E over $\mathbb{Z}_p$ (but for isomorphism over $\mathbb{Z}_p$) and will be the Néron model $\epsilon/\mathbb{Z}_p$.*

Now suppose that p is of bad reduction. The Néron model is optimal in the following sense, characterizing ϵ in a unique way but for $\mathbb{Z}_p$-isomorphism:

a) The reduction modulo p is a curve whose singularities are all of double ordinary type, i.e., locally all the singularities are nodes.

b) The natural morphism $\epsilon(\mathbb{Z}_p) \to E(\mathbb{Q}_p)$ is an isomorphism.

c) The scheme ϵ is regular (its local rings are regular everywhere).

> **Definition 12.6.** *A regular local ring is a noetherian local ring with the property that the minimal number of generators of its maximal ideal is equal to its Krull dimension.*

Note that this is very related with some of the topics we commented in several moments. For instance, recall the criterion of Néron-Ogg-Shafarevich, that characterizes if p is of good or bad reduction, and in this last case measures how bad the reduction is. Recall also that we have seen that there is a natural action of $G_{\mathbb{Q}}$ in $T_l(E)$ that gives us a Galois representation $\rho_{E,l}$ that we can restrict to the decomposition group D_p or to the inertia I_p.

Theorem 12.6. *Let $l \neq p$. Then, $\rho_{E,l}(I_p) \subset \mathrm{GL}_2(\mathbb{Z}_l)$ is a finite group and E has good reduction in p if and only if $\rho_{E,l}(I_p) = \{\mathrm{Id}\}$ (and this is independent of the choice of $l \neq p$).*

In general, for an elliptic curve E/K over a number field, and for a prime P of O_K, we have three options: either P is of good reduction or it has bad reduction, that can be multiplicative (or stable) or additive (non-stable). The following theorem is due to Grothendieck:

Theorem 12.7. *Let E/K be an elliptic curve over a number field, and let P be an ideal of O_K. Then, there exists a finite extension L/K and an ideal P_L over P such that the reduction of E/L in P_L is either good or multiplicative. Furthermore, if L'/L is an extension of L, then the character (good or multiplicative) of the reduction of E/L' over the primes above P_L does not change.*

We can also relate this with the concepts we see about complex multiplication, thanks to the following theorem due to Tate:

Theorem 12.8. *Let E/K be an elliptic curve over a number field and let P be a prime of O_K where E has stable reduction (good or multiplicative). Then, the reduction is good if and only if $v_P(j(E)) \geq 0$.*

And a direct corollary of this, taking advantage of the fact that we already know that the j-invariant of a curve with complex multiplication is an algebraic integer, is that if the elliptic curve has CM, then for all prime P of H (the Hilbert class field), then $v_P(j(E)) \geq 0$ and so they have potential good reduction in all the primes.

We now return to the schemes and reinterpret what is the Néron model in terms of algebraic geometry. Let E/K be an elliptic curve and let W/R be its minimal Weierstrass model. Since W is proper over R, we have that $W(R) = W(K) = E(K)$. However, W is typically singular and its smooth locus W_{sm} is a group scheme over R. Typically, it is not proper, and not all K-points of E extend to W_{sm}. Those that do are the subgroup $E_0(K)$, of finite index in $E(K)$.

The Néron model can be seen as an extension ϵ of E over R which combines the desirable properties of W and W_{sm}, since it is a smooth scheme and all K-points extend to R-points. The identity component of ϵ is W_{sm} while the component of ϵ_k (at least for k algebraically closed) is $E(K)/E_0(K)$. So all the points of $E(K)$ extend to points of $\epsilon(R)$ and $E_0(K)$ is the subgroup of points which extend to the identity component of ϵ.

Let now C/K be a curve. A regular model for C is a proper flat scheme C over R which is regular and whose generic fiber is C. A regular model C is minimal if for any other regular C' there exists a map of schemes $C' \to C$ extending the identity on the generic fiber. The main theorem is that minimal regular models exist and are canonically unique. One can find a regular model for C by starting with any model and applying blowing-up and normalization. From there, we can find a minimal regular model. Let E/K be an elliptic curve and let C/R be its minimal regular model. The Néron model of E is then the smooth locus in C.

We present now the theorem that guarantees the existence of Néron models for elliptic curves: the development of the proof will require all a course in scheme theory:

Theorem 12.9. *Let R be a Dedekind domain with fraction field K, let E/K be an elliptic curve, C/R a minimal proper regular model for E/K (this should be properly define) and let ϵ/R be the larger subscheme of C/R which is smooth over R. Then, ϵ/R is a Néron model for E/K.*

One of the first applications of all this theory in the context of elliptic curves is the proof of Mazur's theorem, that characterizes which groups arises as torsion groups of elliptic curves.

12.5 Reinterpretation of modular forms

A natural generalization of the theory of modular forms comes from Hilbert forms, that is a natural extension of the same concept to several variables. Summing up, when F is a totally real number field of degree m over the rationals, consider $\sigma_1, \dots, \sigma_m$ the real embeddings of F. That way, we have a natural map $\mathrm{GL}_2(F) \to \mathrm{GL}_2(\mathbb{R})^m$. If O_F is the ring of integers of F, the group $\mathrm{GL}_2^+(O_F)$ is what we call the full Hilbert modular group, and for every element $z = (z_1, \dots, z_m) \in \mathbb{H}^m$ there is a group action of $\mathrm{GL}_2^+(O_F)$ defined as one can expect. A Hilbert modular form of weight $(k_1, \dots, k_m)$ will be an analytic function on $\mathbb{H}^m$ such that for every $\gamma \in \mathrm{GL}_2^+(O_F)$,

$$f(\gamma z) = \prod_{i=1}^{m} j(\sigma_i(\gamma), z_i)^{k_i} f(z)$$

In the development of this theory, we find some interesting analogies with the theory of certain function spaces, that is only sketched in the following lines. Recall that when we have an element g of $\mathrm{GL}_2(\mathbb{R})^+$, we write $j(g,z) = \det(g)^{-1/2}(cz+d)$. To a modular form $f \in M_k(\Gamma)$ we can attach a function $\phi_f : \mathrm{SL}_2(\mathbb{R}) \to \mathbb{C}$ defined as follows

$$\phi_f(g) = f(g \cdot i) j(g, i)^{-k}$$

Proposition 12.12. *The function ϕ_f satisfies the following properties:*

a) $\phi_f(\gamma g) = \phi_f(g)$ for all $\gamma \in \Gamma$.

b) Let $r(\theta)$ denotes the element of $\mathrm{SO}_2(\mathbb{R})$ corresponding to a rotation of angle θ. Then, $\phi_f(gr(\theta)) = \exp(ik\theta)\phi_f(g)$.

c) If f is a cusp form, then ϕ_f is a bounded function satisfying

$$\int_{\Gamma \backslash \mathrm{SL}_2(\mathbb{R})} |\phi_f(g)|^2 dg < \infty$$

d) If f is a cusp form, then ϕ_f is what we will call cuspidal, that is, for each $g \in \mathrm{SL}_2(\mathbb{R})$, for each $g \in \mathrm{SL}_2(\mathbb{R})$ and for each $\gamma \in \mathrm{SL}_2(\mathbb{Z})$,

$$\int_0^1 \phi_f\Big(\gamma \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} g\Big) dx = 0$$

The proof of the properties is just easy manipulations.
We also have, associated to any modular form $f \in S_k(\Gamma)$, a function ϕ_f in $L^2(\Gamma\backslash \mathrm{SL}_2(\mathbb{R}))$. We would like to study the image of this association.
Recall that the C^∞ functions of $\Gamma\backslash \mathrm{SL}_2(\mathbb{R})$ are dense in $L^2(\Gamma\backslash \mathrm{SL}_2(\mathbb{R})))$; for those functions, we can consider the following operator, called Casimir's operator, and that in coordinates (x, y, θ) can be written as

$$\Delta = -y^2\Big(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\Big) - y\frac{\partial^2}{\partial x \partial \theta}$$

Proposition 12.13. *The function $f \mapsto \phi_f$ gives a bijection between $S_k(\Gamma)$ and the functions $\phi \in \mathrm{SL}_2(\mathbb{R})$ satisfying:*

a) $\phi(\gamma g) = \phi_f(g)$ for all $\gamma \in \Gamma$.

b) $\phi(gr(\theta)) = \exp(ik\theta)\phi_f(g)$.

c) $\Delta\phi = -\frac{k}{2}(\frac{k}{2} - 1)\phi$.

d) ϕ is bounded and cuspidal.

It would be interesting now to relate this with the theory of group representations of Lie groups to obtain some remarkable results, but again this will lead us too far. What we will do is to define what is an automorphic form in $\mathrm{GL}_2(\mathbb{A}_\mathbb{Q})$, the ring of adeles. If we decompose $g \in \mathrm{GL}_2(\mathbb{A}_\mathbb{Q})$ as $g = \gamma g_\infty k_0$, where $\gamma \in \mathrm{GL}_2(\mathbb{Q}), g_\infty \in \mathrm{GL}_2(\mathbb{R})^+$ and $k_0 \in \prod_p K_p$, given $f \in S_k(\Gamma_0(N))$ (K_p are subgroups of $\mathrm{SL}_2(\mathbb{Z}_p)$), then we can define $\phi_f : \mathrm{GL}_2(\mathbb{A}_\mathbb{Q}) \to \mathbb{C}$ by

$$\phi_f(g) = f(g_\infty i)j(g_\infty, i)^{-k}$$

This is a well-defined function, and the function $f(z) \to \phi_f(g)$ gives an isomorphism between $S_k(\Gamma_0(N))$ and the space of functions ϕ of $\mathrm{GL}_2(\mathbb{A}_\mathbb{Q})$ satisfying certain conditions about growth and invariance. We can now define Hecke operators and like before, establish a correspondence between modular forms and adelic representations.
In this context it is a must to introduce Hilbert forms:
Given an ideal B, we can define the group

$$\mathrm{GL}_2^+(O_K, B) = \Big\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mid a, d \in O_K, b \in B^{-1}, c \in B \Big\}$$

These groups are all maximal (and not necessarily conjugated if B_1, B_2 are in different classes of ideals), and so as to study modular forms we must look at all them. In the same way, given an ideal N, we define

$$\Gamma_0(N, B) = \Big\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{GL}_2^+(O_K, B) \mid c \in NB \Big\}$$

Let now $X_0(N, B)$ the quotient $\Gamma_0(N, B)\backslash \mathbb{H}^2 \cup \mathbb{P}^1(K)$.

Definition 12.7. *A holomorphic function* $f : \mathbb{H}^2 \to \mathbb{C}$ *is a modular form of weight* $k = (k_1, k_2)$ *(where* k_1, k_2 *are non-negative integers) for* $\Gamma_0(N, B)$ *if for all* $\gamma \in \Gamma_0(N, B)$ *we have*

$$f(\gamma \cdot (z_1, z_2)) = \tau_1(\det(\gamma))^{-k_1/2} \tau_2(\det(\gamma))^{-k_2/2}$$

$$(\tau_1(c)z_1 + \tau_1(d))^{k_1} (\tau_2(c)z_2 + \tau_2(d))^{k_2} f(z_1, z_2)$$

We will denote by $M_k(\Gamma_0(N, B))$ *the space of these functions.*

Note that we do not request holomorphy in the cusps; the explanation of this comes from the famous Koecher's principle, that states that the condition of holomorphy at the cusps is immediate from the definition. This same theory, as was pointed out at the beginning of this section, generalizes to the case of m variables and everything remains equal.

12.6 What is missing in this thesis?

When one starts the writing of a project like this, there are many topics that one hopes to have time to develop and finally, due to the lack of time, it is not always possible. We try to point out here several things that we projected to include in the thesis and that would have complemented some of the chapters and give a better global understanding of the topic:

a) A clear and concise vision of class field theory, since many of the results we gave for instance when dealing with complex multiplication are particular cases of a more general theory. It is also missing a better understanding of Galois representations. For that, it would be necessary to properly explain the proof and to develop more cohomology of groups and homological algebra.

b) An introduction to the theory of schemes and group schemes (and in general, this thesis has a lack of algebraic geometry, that is crucial to understand deeper results in number theory). This would lead us to a better introduction of Néron models and to be able to develop some advanced results that rely on this. The same applies for the theory of arithmetic geometry.

c) Explain more results around BSD, Stak-Heegner points, p-adic uniformization, extensions to totally real fields, sketch the proof of Coates-Wiles . . . Formulate BSD in a more general framework: this would have lead us to the Block-Kato conjecture about motives.

d) Some results were quoted without proofs. I would like to have included them to have a more self-contained thesis.

Bibliography

[1] Diamond, F and Shurman, J. *A first course in modular forms*, 2005.

[2] Silverman, J. *The arithmetic of elliptic curves*, 1992.

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

[4] Milne, J. *Elliptic curves*, 2006.

[5] James Milne's online notes

[6] Darmon, H. *Rational points on modular elliptic curves*, 2001.

[7] Serre, J.P. *A course in arithmetics*, 1973.

[8] Neukirch, J. *Algebraic number theory*, 1991.

[9] Neukirch, J. *Class Field Theory*, 1980.

[10] Keith Conrad's online blurbs

[11] Hindry, M. and Silverman, J. *Diophantine geometry. An introduction*, 1991.

[12] Miranda, R. *Algebraic Curves and Riemann Surfaces*, 1953.

[13] Hartshorne, R. *Algebraic Geometry*, 2000.

[14] Cassels, J. *Lectures on Elliptic Curves*, 1991.

[15] Online notes from the AGRA school (Cuzco, August 2015)

[16] Stein, E. and Shakarchi, R. *Complex analysis*, 2011.

[17] Alsina, M. *Quaternion algebras and Quadratic forms towards Shimura curves*, 2013.

[18] Silverman, J. and Tate J. *Rational points on elliptic curves*, 2015.

[19] Cornell, G. et at. *Modular forms and Fermat's last theorem*, 2000.

[20] Lang, S. *Algebra*, 2004.