

**Exploring Neural Networks for  
computing the Hilbert Class Field of  
Quadratic Extensions of  $\mathbb{Q}$**

Mateo David Céspedes Gil

Supervised by: Jorge Plazas and Andrés Vargas

Pontificia Universidad Javeriana - Bogotá

August 2023


# Exploring Neural Networks for computing the Hilbert Class Field of Quadratic Extensions of $\mathbb{Q}$

**Mateo David Céspedes Gil**


Nota: Aprobado


## Directores:

Jorge Andrés Plazas Vargas, Ph.D.  \_\_\_\_\_

Andrés Vargas Domínguez, Ph.D.  \_\_\_\_\_

## Jurados:

Daniel Cabarcas Jaramillo, Ph.D.  \_\_\_\_\_

Guillermo Arturo Mantilla Soler, Ph.D.  \_\_\_\_\_

Exploring Neural Networks for computing the Hilbert Class Field of  
Quadratic Extensions of  $\mathbb{Q}$

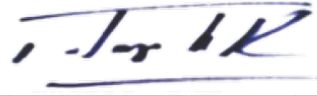
Mateo David Céspedes Gil

APROBADO:



---

Jhon Jairo Sutachan Rubio, Ph.D.  
Director de Posgrado  
Facultad de Ciencias



---

Alba Alicia Trespalacios Rangel, Ph.D.  
Decana  
Facultad de Ciencias

Bogotá, septiembre de 2023

# Introduction

Inspired by DeepMind's article "Advancing Mathematics by Guiding Human Intuition with AI" [8], where connections between mathematical objects were discovered leading to the formulation and subsequent proof of conjectures in knot theory and representation theory, we embark on an approach utilizing machine learning and deep learning techniques, with the primary objective to investigate the Hilbert class field of a real quadratic field. By leveraging the capabilities of modern AI, we aim to guide our intuition and uncover patterns in number fields that may facilitate explicit constructions of abelian extensions.

The explicit class field theory problem consists on finding a way to explicitly construct all abelian extensions of a given number field. This is a difficult problem, and it has only been solved in a few special cases.

The first case that was solved is the case of the rational number field,  $\mathbb{Q}$ . In this case, the maximal abelian extension is the cyclotomic field, obtained by adjoining to the rational numbers all roots of unity. In the resulting field we can then write every element as a finite sum of products with rational coefficients of primitive roots of unity.

The second case that was solved was the case of imaginary quadratic fields. An imaginary quadratic field is a number field that is generated by the square root of a negative integer. The theory of elliptic curves with complex multiplication settles the problem by generating first the maximal abelian unramified extension using the  $j$ -invariant and then adjoining values of the Weber function at the torsion points on elliptic curves.

Other cases of explicit class field theory have been solved as well, but the problem is still unsolved in general. Following results in machine learning as stated above together with tools from noncommutative geometry our goal is to tackle this problem for the case of real quadratic fields.

# Contents

<b>Preface</b>	<b>1</b>
<b>1 Preliminaries and Background</b>	<b>3</b>
1.1 Basic definitions about quadratic fields . . . . .	3
1.2 Class numbers and real quadratic fields . . . . .	9
1.2.1 Continued Fractions . . . . .	10
1.3 Abelian Extensions and the Hilbert Class Field . . . . .	12
1.4 Kronecker-Weber Theorem . . . . .	13
1.5 Elliptic Curves . . . . .	14
1.5.1 Elliptic curves over the rationals . . . . .	15
1.5.2 Elliptic curves over the complex numbers . . . . .	15
1.6 Modular curves . . . . .	17
1.7 The $j$ -invariant . . . . .	18
1.8 Morphisms and Complex Multiplication . . . . .	18
1.9 Theta Functions . . . . .	20
1.10 Concepts on functional analysis . . . . .	22
1.11 Deep Learning and Terminology . . . . .	24
<b>2 Real Multiplication</b>	<b>27</b>
2.1 Noncommutative spaces . . . . .	27
2.2 Real Multiplication . . . . .	31
<b>3 AI and mathematical intuition</b>	<b>32</b>
3.1 Framework . . . . .	32
3.2 Current computations of $H_K$ for real quadratic fields . . . . .	34
<b>4 Some Computations</b>	<b>36</b>
4.1 Initial setting and generating Data . . . . .	36
4.2 Method Used . . . . .	38
4.3 Preliminary Results and Next Steps . . . . .	39
<b>Bibliography</b>	<b>43</b>

# Chapter 1

## Preliminaries and Background

### 1.1 Basic definitions about quadratic fields

In this section we will introduce the basic definitions and results about fields and some of the rudiments of algebraic number theory for the quadratic case, which will be used in the following sections to explain the present problem and its developments. Proofs of the results can be found in Trifkovic's [24]; we also follow his notation.

As usual an **extension** of a base field  $F$ , is a field  $K$  such that  $F \subseteq K$ . By a **number field** we mean a finite degree extension of the rational numbers.

A cyclotomic field is an extension obtained by adjoining a root of unity to  $\mathbb{Q}$ .

Let  $K$  be an extension of the field  $F$ . The set of automorphisms  $\sigma \in \text{Aut}(K)$  such that  $\sigma(a) = a$  for all  $a \in F$ , is a group under composition and it is called the **Galois Group** of  $K$  over  $F$ . We use the usual notation  $\text{Gal}(K/F)$ .

Every finite group  $G$  appears as a Galois group of certain field extension. To study such groups is reasonable to start with the simplest case, when  $G$  is abelian. An extension  $K$  of  $F$  is said to be abelian if  $\text{Gal}(K/F)$  is a commutative group.

Here will concentrate in a particular type of number field.

**Definition 1.1.1.** A **quadratic field** is an extension of degree two of the rational numbers. Any such field has the form

$$\mathbb{Q}[\sqrt{D}] = \{a + b\sqrt{D} : a, b \in \mathbb{Q}\}$$

with  $D$  square-free integer. The adjectives *real* or *imaginary* are added to  $\mathbb{Q}[\sqrt{D}]$  depending on whether  $D$  is positive or negative.

Of utmost importance are the analogs to the integers in these more general fields which also happen to be a subring of the number field under consideration.

**Definition 1.1.2.** A **quadratic integer** is a number  $\alpha \in \mathbb{Q}[\sqrt{D}]$  that satisfies a monic polynomial with coefficients in  $\mathbb{Z}$ .

The set of all the quadratic integers in a given field forms a ring<sup>1</sup>. Moreover there is a nice characterization for such ring of integers in a quadratic field:

**Proposition 1.1.1.** The set of quadratic integers of  $K = \mathbb{Q}[\sqrt{D}]$  is the ring  $\mathcal{O}_K = \mathbb{Z} + \mathbb{Z}\delta = \mathbb{Z}[\delta]$ , where

$$\delta = \begin{cases} \sqrt{D}, & \text{if } D \equiv 2, 3 \pmod{4} \\ \frac{1+\sqrt{D}}{2}, & \text{if } D \equiv 1 \pmod{4} \end{cases}$$

An useful numerical invariant is the **discriminant** of the field which in some way measures the size of its ring of integers. In the case of quadratic fields it is defined as

$$\Delta_K = \begin{cases} 4D, & \text{if } D \equiv 2, 3 \pmod{4} \\ D, & \text{if } D \equiv 1 \pmod{4} \end{cases}$$

In any quadratic field we can define trace and norm as  $Tr(\alpha) = \alpha + \bar{\alpha}$  and  $N(\alpha) = \alpha\bar{\alpha}$ , where the bar means complex conjugation. In the case of  $\alpha = a + b\sqrt{D} \in \mathbb{R}$  is just  $\bar{\alpha} = a - b\sqrt{D}$ .

Trace and norm are group homomorphisms:

$$Tr : \mathbb{Q}[\sqrt{D}] \rightarrow \mathbb{Q}$$

$$N : \mathbb{Q}[\sqrt{D}]^\times \rightarrow \mathbb{Q}^\times$$

For example the norm can be used to find the units in the ring of quadratic integers.

**Proposition 1.1.2.** An element  $\alpha \in \mathcal{O}_K$  is an unit if and only if  $N(\alpha) = \pm 1$

---

<sup>1</sup>This definition makes sense for any number field: An **integer** in the number field  $K$  is a number  $\alpha \in K$  that satisfies a monic polynomial with coefficients in  $\mathbb{Z}$ . The set of such numbers is a subring of  $K$  denoted by  $\mathcal{O}_K$

Also these operations appear when working with polynomials satisfied by elements in the field. If  $p(\alpha) = 0$  also  $p(\bar{\alpha}) = 0$  then  $p(x)$  is divisible by

$$(x - \alpha)(x - \bar{\alpha}) = x^2 - \text{Tr}(\alpha) + N(\alpha)$$

Following proposition 1.1.1, the ring of integers is of the form  $\mathbb{Z}\alpha + \mathbb{Z}\beta$ . In the complex quadratic case this is a nice subset of the complex plane.

**Definition 1.1.3.** A *lattice*  $\Lambda \subset \mathbb{C}$  is a subgroup (under addition) of the complex numbers that has the form  $\Lambda = \mathbb{Z}v_1 + \mathbb{Z}v_2$ , with  $v_1$  and  $v_2$  linearly independent over  $\mathbb{R}$ .

**Definition 1.1.4.** A *fundamental parallelogram* is a subset of the complex numbers of the form

$$\Pi_{(a,b)} = \{at_1 + bt_2 : t_1, t_2 \in [0, 1)\}$$

Given a lattice  $\mathbb{Z}a + \mathbb{Z}b$ , it is possible to tile the complex plane with translations of a fundamental parallelogram.

Notation as above and identifying  $\mathbb{C}$  with  $\mathbb{R}^2$ , the prototype for all lattices is  $\Lambda_0$ , the set of all column vectors with entries in  $\mathbb{Z}$  or equivalently to  $\mathbb{Z} \times \mathbb{Z}$ . Then the map  $\begin{bmatrix} 1 \\ 0 \end{bmatrix} \mapsto a$  and  $\begin{bmatrix} 0 \\ 1 \end{bmatrix} \mapsto b$  is an isomorphism between  $\Lambda_0$  and  $\mathbb{Z}a + \mathbb{Z}b$ .

**Proposition 1.1.3.** Given an imaginary quadratic field  $K$ , its ring of integers  $\mathcal{O}_K \subseteq \mathbb{C}$  is a lattice. Moreover, the inverse of the above isomorphism maps nonzero ideals of  $\mathcal{O}_K$  to sublattices of  $\Lambda_0$ .

This correspondence is used with the objective of studying the ring of quadratic integers. The following proposition will be a step in this direction.

**Proposition 1.1.4.** A subgroup  $\Lambda \subseteq \Lambda_0$  is a sublattice if and only if there exist a  $2 \times 2$  matrix  $\gamma$  with integer coefficients such that  $\Lambda = \gamma\Lambda_0$  and  $\det\gamma \neq 0$ .

**Proposition 1.1.5.** Let  $\Lambda = \gamma\Lambda_0$  be a lattice. Then  $\Lambda_0/\Lambda$  is finite and  $|\Lambda_0/\Lambda| = |\det\gamma|$ .

This provides us with two important properties of the quadratic integers of a given field.

**Corollary 1.1.1.** Let  $I$  be a nonzero ideal of  $\mathcal{O}_K$ . Then  $\mathcal{O}_K/I$  is finite and any strictly ascending chain of ideals of  $\mathcal{O}_K$  is finite.



**Corollary 1.1.2.** *Every prime ideal in  $\mathcal{O}_K$  is maximal.*

We can extend the notion of norm to ideals, for  $I$  an ideal of  $\mathcal{O}_K$  its norm is the natural number  $N I = |\mathcal{O}_K/I|$ .

If  $I, J$  are ideals of the ring  $R$ , we say that  $I$  divides  $J$ , written  $I|J$  if  $J \subseteq I$ .

**Proposition 1.1.6.** *Let,  $I, J$  be non-zero ideals of  $\mathcal{O}_K$ . If  $I|J$ , then  $N I|N J$ .*

The proposition below will have as consequences two results that led to a generalization of unique factorization to the realm of ideals.

**Proposition 1.1.7.** *If  $I$  is a nonzero ideal of  $\mathcal{O}_K$  then  $I\bar{I} = \mathcal{O}_K N I$ . The principal ideal in  $\mathcal{O}_K$  generated by  $N I$ .*

**Corollary 1.1.3.** *For all  $I, J$  ideals of  $\mathcal{O}_K$ ,  $N(IJ) = (N I)(N J)$ .*

**Corollary 1.1.4.** *Let  $I, J, L$  be ideals of  $\mathcal{O}_K$  with  $I \neq 0$ . If  $IJ = IL$  then  $J = L$ .*

To make this last corollary meaningful we have to add inverses since within all nonzero ideals of  $\mathcal{O}_K$  the full ring is the only element with an inverse.

**Definition 1.1.5.** *Let  $K = \mathbb{Q}[\sqrt{D}]$  and  $\mathcal{O}_K = \mathbb{Z} + \mathbb{Z}\delta$  its ring of integers. A fractional ideal  $\mathfrak{F}$  of  $\mathcal{O}_K$  is a subgroup of  $K$  such that the following holds:*

- $\mathfrak{F} = \mathbb{Z}\alpha + \mathbb{Z}\beta$ , for some  $\alpha, \beta$  linearly independent over  $\mathbb{Z}$ .
- $\delta\mathfrak{F} \subseteq \mathfrak{F}$ .

With  $\mathcal{O}_K = \mathbb{Z} + \mathbb{Z}\delta$ . The set of all non zero fractional ideals in a given field  $K$  is denoted by  $\mathbb{I}_K$ .

It is possible to consider fractional ideals in an arbitrary number field  $K$ , these are just the finitely generated  $\mathcal{O}_K$ -submodules of  $K$ , see for instance [2].

In the case of quadratic fields we can determine easily which are the fractional ideals. All fractional ideals are of the form  $\mathfrak{I} = r(\mathbb{Z}a + \mathbb{Z}(-b + \delta))$  with  $a, b \in \mathbb{Z}$  and  $r \in \mathbb{Q}^\times$ .

Also fractional ideals have a well defined group structure given by the operation  $\mathfrak{F} \cdot \mathfrak{G} = \frac{1}{kl}(k\mathfrak{F})(l\mathfrak{G})$ , where  $k$  and  $l$  are integers such that  $k\mathfrak{F}, l\mathfrak{G} \subseteq \mathcal{O}_K$ .

The first main result is the well known unique factorization of ideals for quadratic numbers.

**Theorem 1.1.1.** *Let  $K = \mathbb{Q}[\sqrt{D}]$  and  $\mathcal{O}_K$  its ring of integers. For any non-trivial ideal  $I \subseteq \mathcal{O}_K$ , there exists prime ideals  $P_1, \dots, P_n$  of  $\mathcal{O}_K$  such that  $I = P_1 \cdot P_2 \cdot \dots \cdot P_n$ , this factorization is unique up to a permutation of the  $P_i$ .*

This factorization is never too bad due to  $\mathcal{O}_K$  fulfilling the ascending chain condition of corollary 1.1.1 and the following lemma.

**Lemma 1.1.1.** *Let  $P$  be a prime ideal in  $\mathcal{O}_K$ . Then exist a unique prime  $p \in \mathbb{N}$  such that  $P | \langle p \rangle = p\mathcal{O}_K$ .*

To find the prime factorization of  $P$  we only then need to factor principal ideals generated by primes in  $\mathbb{N}$ .

**Proposition 1.1.8.** *Let  $p \in \mathbb{N}$  be prime. The prime factorization of the ideal  $\langle p \rangle$  is one of the following:*

- $\langle p \rangle = P$  with  $NP = p^2$
- $\langle p \rangle = P^2$  with  $NP = p$
- $\langle p \rangle = P\bar{P}$  with  $NP = p$  and  $P$  not equal to its conjugate.

*We call  $p$  and  $P$ ; inert, ramified or split respectively.*

An integer prime  $p$  is ramified if and only if  $p | \Delta_K$  so only finitely many primes ramify.

A similar theory of ramification applies to more general fields including extensions of a quadratic fields  $K$ , see for example [3].

Finally all non zero ideals in a quadratic field have a standard form  $\mathbb{Z}a + \mathbb{Z}(-b + \delta)$  and can be factored this way

$$I = \prod_{i=1}^r (\mathbb{Z}p_i + \mathbb{Z}(-b + \delta))^{e_i}$$

where  $a = p_1^{e_1} \cdot \dots \cdot p_r^{e_r}$ .

Another invariant that we won't discuss but will be used in our computational example in the last chapter is the regulator.

Let  $\sigma_1, \dots, \sigma_{r_1}$  and  $\sigma_{r_1+1}, \dots, \sigma_{r_1+r_2}$  be the real and complex embeddings (respectively) of a number field  $K$  into  $\mathbb{C}$ .

If  $u_1, \dots, u_r$  a full set of fundamental units of  $K$ . Then  $r = r_1 + r_2 - 1$ .

If we denote by  $M$  the  $(r_1 + r_2 - 1) \times (r_1 + r_2)$  matrix with coefficients  $(d_i \log \sigma_j(u_i))$ , where  $d_i = 1$  if  $i \leq r_1$  or  $d_i = 2$  otherwise.

**Definition 1.1.6.** *The regulator of  $K$  is the absolute value of the determinant of the matrix  $M$  minus one column.*

To end this section we mention explicitly how are the subrings of  $\mathcal{O}_K$  this will be useful to relate quadratic fields with elliptic curves.

Subrings of  $\mathcal{O}_K = \mathbb{Z} + \delta\mathbb{Z}$  are either  $\mathbb{Z}$  the usual integers or a ring of the form  $\mathcal{O}_f = \mathbb{Z} + \mathbb{Z}f\delta$  where  $f$  is a positive integer called the conductor and equals the size of  $\mathcal{O}_K/\mathcal{O}_f$ . These are called **orders** in  $K$ .

## 1.2 Class numbers and real quadratic fields

**Definition 1.2.1.** Let  $K$  be a quadratic field and  $\mathbb{P}_K$  be the set of all non zero fractional principal ideals of  $\mathcal{O}_K$ . The ideal class group of  $K$  is the quotient

$$Cl(K) = \mathbb{I}_K / \mathbb{P}_K$$

The class number of  $K$  is  $h(K) = |Cl(K)|$ .

The inverses in the group are given by conjugation and, using unique factorization of ideals, classes of prime elements generate the group. A complete explanation for the following two results is out of the scope of this work but its proofs can be found in most algebraic number theory books (see for example [3]).

**Theorem 1.2.1.** The class number of a number field is finite.

**Proposition 1.2.1.** Each ideal class in  $Cl(K)$  contains an ideal with norm at most

$$\mathfrak{M}_K = \sqrt{|\Delta_K|} \cdot \frac{2}{\pi}$$

for  $K$  quadratic imaginary, or

$$\mathfrak{M}_K = \sqrt{|\Delta_K|} \cdot \frac{1}{2}$$

for  $K$  real quadratic.

Combining the two big results up to now; unique factorization and finiteness of the class number we obtain

**Theorem 1.2.2.** The ideal class group of  $K$  is generated by a finite number of elements  $P_i$  with the  $P_i$  having prime norm bounded by  $\mathfrak{M}_K$ .

We want to have as much information about a quadratic field  $K$  as possible, this includes its ideal class group structure and class number. The difference between imaginary and real quadratic fields which will be of importance later is that the former has nicer properties to start with.

Firstly, the ring of integers of  $K$  imaginary quadratic has a lattice structure. Second, its norm homomorphism involves solving an equation of the form  $N(x + y\delta) = (x - \frac{t}{2}y)^2 - \frac{\Delta_K}{4}y^2 = n$  which has a finite number of solutions due to  $\Delta_K$  being negative.

On the other hand the norm homomorphism in the real case leads to solving the equation  $x^2 - Dy^2 = \pm n$  which is Pell's equation and has in general infinitely many solutions which are not always easy to find.

### 1.2.1 Continued Fractions

The proper setting for working with Pell's equation and real quadratic fields is by using continued fractions. We now give a short introduction to its theory just enough to aid our quadratic field studies. Its usefulness lies in the fact that continued fractions approximate real numbers in efficient way. Furthermore, even though not thoroughly covered in this work, convergents of a continued fraction appear in the theory of generalized theta functions and also in the theory of noncommutative tori with real multiplication.

**Definition 1.2.2.** *A finite continued fraction is an expression of the form:*

$$[a_0; a_1, a_2, \dots, a_n] = a_0 + \frac{1}{a_1 + \frac{1}{a_2 + \frac{1}{\ddots + \frac{1}{a_n}}}}$$

where the  $a_i$  are integers.

The  $i$ th convergent  $p_i/q_i$  is obtained by truncating this expansion at the  $i$ -th element. Convergents are given by the recursive formulas:

$$p_i = a_i p_{i-1} + p_{i-2}, \quad q_i = a_i q_{i-1} + q_{i-2}$$

where  $p_{-1} = 1, p_{-2} = 0, q_{-1} = 0, q_{-2} = 1$ .

These recursions can be written in matrix form:

$$\begin{bmatrix} p_i \\ p_{i-1} \end{bmatrix} = \begin{bmatrix} a_i & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} p_{i-1} \\ p_{i-2} \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} q_i \\ q_{i-1} \end{bmatrix} = \begin{bmatrix} a_i & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} q_{i-1} \\ q_{i-2} \end{bmatrix}$$

An infinite continued fraction is denoted as:

$$[a_0; a_1, a_2, a_3, \dots] = a_0 + \frac{1}{a_1 + \frac{1}{a_2 + \frac{1}{a_3 + \frac{1}{\ddots}}}}$$

A periodic continued fraction is of the form:

$$[a_0; a_1, a_2, \dots, a_m, \overline{a_{m+1}, a_{m+2}, \dots, a_{m+k}}]$$

where  $m$  is the index where the periodic part starts, and  $k$  is the period length.

A purely periodic continued fraction is of the form:

$$[a_0; \overline{a_1, a_2, \dots, a_k}]$$

where  $k$  is the period length.

Any real number can be expressed as a continued fraction, with rational numbers corresponding to finite continued fractions.

The relationship between continued fractions and the theory of quadratic numbers comes from the following theorem.

**Theorem 1.2.3.** *A continued fraction for an irrational number is periodic if and only if that number is a quadratic number.*

In our study we will want to obtain fast approximations to quadratic numbers. There are algorithms that calculate periods of a continued fractions efficiently, in this work we used the current implementations included in SAGE and PARI/GP.

Continued fractions are also used for obtaining the units in the ring of integers of a real quadratic field  $\mathcal{O}_K$ .

**Proposition 1.2.2.** *If  $K$  is a real quadratic field, there exist a unique  $\delta$  such that  $\mathcal{O}_K = \mathbb{Z}[\delta]$  and its continued fraction is purely periodic.*

**Definition 1.2.3.** *Given a real quadratic field  $K$ , a fundamental unit of  $\mathcal{O}_K$  is an element  $\epsilon_K \in \mathcal{O}_K^\times$  satisfying the following:*

- For any  $\epsilon \in \mathcal{O}_K^\times$ ,  $\epsilon = \pm \epsilon_K^n$  for some  $n \in \mathbb{Z}$ ,
- $\epsilon_K > 1$ .

The way to construct the group of units  $\mathcal{O}_K^\times$  using continued fractions is by means of

**Theorem 1.2.4.** *Let  $\mathcal{O}_K = \mathbb{Z}[\delta]$  with  $\delta > \bar{\delta}$  be the ring of integers of  $K = \mathbb{Q}[\sqrt{D}]$ . Let  $p'_i/q'_i$  be the convergents of  $\delta$ , and  $l$  the period length of  $D$ .*

*Set  $\epsilon_1 = p'_l - q'_l \delta$ . Then any unit  $\epsilon \in \mathcal{O}_K^\times$  is given by  $\epsilon = \pm \epsilon_1^n$  for some integer  $n$ . The fundamental unit  $\epsilon_K$  is then  $\pm \epsilon_1$  or  $\pm \bar{\epsilon}_1$ , whichever fulfills the second condition above.*

## 1.3 Abelian Extensions and the Hilbert Class Field

The goal of the theory as stated in the introduction is to find the maximal abelian extension in an explicit way. This is no easy task and far from being solved, even the quadratic case which is the most simple and nontrivial is incomplete.

**Definition 1.3.1.** *Let  $K$  be a field. The maximal abelian extension of  $K$ , denoted by  $K^{\text{ab}}$ , is the largest field extension of  $K$  that is abelian over  $K$ .*

The maximal abelian extension  $K^{\text{ab}}$  has several important properties:

1.  $K^{\text{ab}}$  is a Galois extension of  $K$ , meaning it is a normal and separable extension.
2. Every abelian extension of  $K$  is contained in  $K^{\text{ab}}$
3.  $K^{\text{ab}}$  is the union of all finite abelian extensions of  $K$ .

**Proposition 1.3.1.** *Let  $K$  be a number field and  $K^{\text{ab}}$  be its maximal abelian extension. Then the following relationship between Galois groups holds*

$$\text{Gal}(K^{\text{ab}}/K) \cong \text{Gal}(\bar{K}/K)^{\text{ab}}$$

Where  $\bar{K}$  is the algebraic closure of  $K$  and the right hand side is the abelianization of the absolute Galois group.

The maximal abelian extension  $K^{\text{ab}}$  plays a fundamental role in algebraic number theory. It provides a way to study the abelian extensions of a field and the behavior of Galois groups. Even though there is a description of what  $K^{\text{ab}}$  is <sup>2</sup>, in general it is hard or not know its explicit generators and structure.

**Definition 1.3.2.** *Let  $K$  be a number field. The Hilbert class field of  $K$ , denoted by  $H_K$ , is the maximal abelian extension of  $K$  which is unramified at all primes.*

The Hilbert class field  $H_K$  has several important properties:

1.  $H_K$  is a finite extension of  $K$ .
2.  $H_K$  is Galois over  $K$ .

---

<sup>2</sup>using class field theory

3.  $H_K$  is an abelian extension of  $K$ .
4.  $H_K$  is the smallest field containing  $K$  such that all prime ideals in  $\mathcal{O}_K$  become principal in  $H_K$ .

David Hilbert conjectured the existence of this field sometimes also called the absolute class field and then Philipp Furtwängler prove its existence in 1907 ([3] Chapter XI).

**Theorem 1.3.1** (Furtwängler). *For any number field  $K$ , the Hilbert class field  $H_K$  exists and is unique up to isomorphism.*

**Theorem 1.3.2.** *Let  $H_K$  be the Hilbert class field of  $K$ . Then  $\text{Gal}(H_K/K)$  is isomorphic to the ideal class group  $Cl(K)$ .*

The Hilbert class field  $H_K$  has deep connections to the arithmetic properties of the number field  $K$ . It provides a way to study the behavior of prime ideals in  $K$  and is related to important topics such as class field theory and the study of quadratic forms.

Understanding the Hilbert Class field is an important first step to understand the maximal abelian extension of a number field. This applies in particular to the problem of finding explicit generators.

## 1.4 Kronecker-Weber Theorem

We are looking for descriptions of abelian extensions of fields, explicit generators of these extensions, and the action of the Galois group for such elements. One of the first steps in this direction is the Kronecker-Weber theorem.

We begin with a well known result of field theory [10]:

**Proposition 1.4.1.** *Let  $K$  be a cyclotomic extension obtained from  $\mathbb{Q}$  by adjoining  $\zeta$ , a  $n$ -th root of unity. Then  $K$  is abelian with Galois group isomorphic to  $(\mathbb{Z}/n\mathbb{Z})^\times$ .*

A partial converse is

**Theorem 1.4.1.** (Kronecker-Weber) *Let  $K$  be a number field that is Galois over  $\mathbb{Q}$ , suppose that  $\text{Gal}(K/\mathbb{Q})$  is an abelian group. Then there exists a cyclotomic extension  $\mathbb{Q}(\zeta)$  that contains  $K$ , for  $\zeta$  an  $n$ -th root of unity for some positive integer  $n$ .*



In other words,  $K^{\text{ab}} = \mathbb{Q}^{\text{ab}} = \mathbb{Q}^{\text{cycl}}$ . the field resulting by adjoining all roots of unity to  $\mathbb{Q}$ .

This result is special not only because of the description given for  $\mathbb{Q}^{\text{ab}}$ . Theorem 1.4.1 says that there is an specific function that produces generators for the extension.

Consider the Taylor series of the complex exponential function which converges in all  $\mathbb{C}$ :

$$f(z) = e^{2\pi iz} = \sum_{k=0}^{\infty} \frac{(2\pi iz)^k}{k!},$$

By evaluating in  $1/n$  for some  $n$  we obtain a transcendental function with the property that  $\mathbb{Q}[f(\frac{1}{n})]$  is abelian and that every other finite abelian extension of  $\mathbb{Q}$  is contained in one of these for some value of  $n$ .

The point is, we can describe all finite abelian extensions of the rationals in terms of the special values of an analytic function. Moreover, Galois theory gives us an isomorphism  $\phi : Gal(\mathbb{Q}(\zeta)/\mathbb{Q}) \rightarrow (\mathbb{Z}/n\mathbb{Z})^\times$ . And we can capture the action of each element in the Galois group on the elements of the field:

$$\sigma \left( f \left( \frac{1}{n} \right) \right) = f \left( \frac{\phi(\sigma)}{n} \right)$$

For every  $\sigma$  en  $Gal(\mathbb{Q}(f(1/n))/\mathbb{Q})$ .

**Kronecker Jugendtraum:** refers to the discovery of a similar theory to that described above to fields more general than  $\mathbb{Q}$ . Given a field  $F$ , we want to find a function  $f$  with the following property: for every abelian extension  $K$  of  $F$ , there exist some values  $\alpha_i$  that satisfy  $F(f(\alpha_1), \dots, f(\alpha_n))$  and  $K$  is contained in the resulting field.

The next step in the development of this problem was the case where  $F$  is an imaginary quadratic extension of  $\mathbb{Q}$ . This was solved by using the theory of elliptic curves with complex multiplication that we briefly outline in the following section.

## 1.5 Elliptic Curves

As a motivation and starting point we are going to skim over the theory of elliptic curves, their basic properties for completeness and finally review the theory of such curves with complex multiplication. All the following results are

standard and their proofs can be found in most books on the subject (see for instance [21]).

### 1.5.1 Elliptic curves over the rationals

An elliptic curve over  $\mathbb{Q}$  is a smooth cubic projective curve defined over the field of rational numbers together with a rational point  $O$ .

A general equation for such curve is of the form

$$aX^3 + bX^2Y + cXY^2 + dY^3 + eX^2Z + fXYZ + gY^2Z + hXZ^2 + iYZ^2 + jZ^3 = 0$$

As a consequence of the Riemann-Roch theorem, if the elliptic curve is defined over a field of characteristic different to 2 or 3, as is the case of  $\mathbb{Q}$ , number fields and  $\mathbb{C}$ , then its equation can be simplified to the so-called Weierstrass form:

$$ZY^2 = X^3 + aXZ^2 + bZ^3$$

with  $4a^3 + 27b^2 \neq 0$ . In affine coordinates,  $y^2 = x^3 + ax + b$ .

In this context two elliptic curves are isomorphic if there is a bijective change of variables that converts the Weierstrass form of one into the corresponding Weierstrass form of the other.

Let  $E/\mathbb{Q}$  be an elliptic curve and  $E(\mathbb{Q})$  its set of rational points.  $E(\mathbb{Q})$  can be equipped with a (geometric) group structure. The result of adding two points  $P + Q$  will be the reflection over the x-axis of the third point of intersection of  $\overline{PQ}$  with  $E$ .

This group is abelian and finitely generated according to Mordell-Weil theorem. Its structure is given by

$$E(\mathbb{Q}) \cong E(\mathbb{Q})_{tors} \oplus \mathbb{Z}^r$$

For some  $r \in \mathbb{N}$ .

### 1.5.2 Elliptic curves over the complex numbers

Considering an elliptic curve  $E$  defined by a cubic with rational coefficients as being defined over the complex numbers yields to some additional structure. If  $\Lambda$  is a lattice, then  $\mathbb{C}/\Lambda$  is topologically a torus since every fundamental parallelogram is a complete set of representatives in  $\mathbb{C}/\Lambda$ .

Following the propositions about lattices in section 1.1 we have

**Proposition 1.5.1.** *Let  $\Lambda = \langle \omega_1, \omega_2 \rangle$  and  $\Lambda' = \langle \omega'_1, \omega'_2 \rangle$  be oriented lattices. Then*

- $\Lambda = \Lambda'$  if and only if there is a matrix  $\gamma \in SL(2, \mathbb{Z})$  such that  $\begin{bmatrix} \omega'_1 \\ \omega'_2 \end{bmatrix} = \gamma \begin{bmatrix} \omega_1 \\ \omega_2 \end{bmatrix}$
- There is a holomorphic isomorphism  $\mathbb{C}/\Lambda \xrightarrow{\sim} \mathbb{C}/\Lambda'$  if and only if  $\Lambda' = \alpha\Lambda$ , for some  $\alpha \in \mathbb{C}$ .

**Corollary 1.5.1.** *Let  $\Lambda = \langle \omega_1, \omega_2 \rangle$  and  $\Lambda' = \langle \omega'_1, \omega'_2 \rangle$  be oriented bases of lattices, such that there is an analytic isomorphism  $\mathbb{C}/\Lambda \cong \mathbb{C}/\Lambda'$  of abelian groups. Then, there exist  $\alpha \in \mathbb{C}^*$  and  $\gamma \in SL(2, \mathbb{Z})$  such that  $(\omega'_1, \omega'_2) = \alpha\gamma \begin{bmatrix} \omega_1 \\ \omega_2 \end{bmatrix}$ .*

If  $\Lambda = \langle \omega_1, \omega_2 \rangle$  we can always simplify using proposition above to obtain a isomorphism  $\mathbb{C}/\Lambda \cong \mathbb{C}/\Lambda'$  where  $\Lambda' = \langle \tau, 1 \rangle$  and  $\tau = \omega_1/\omega_2$ .

With the above notation, and using the previous corollary we have that given two lattices  $\Lambda, \Lambda'$ , given by fundamental periods  $\tau, \tau'$ . Then  $\mathbb{C}/\Lambda \cong \mathbb{C}/\Lambda'$  if and only if there is a  $\gamma = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in SL(2, \mathbb{Z})$  with

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

**Definition 1.5.1.** *Let  $k \geq 2$  and  $\Lambda$  a lattice. The Eisenstein series of  $\Lambda$  is*

$$G_{2n}(\Lambda) = \sum_{\omega \in \Lambda - \{0\}} \frac{1}{\omega^{2n}}$$

The main result is the following

**Theorem 1.5.1.** *Uniformization theorem: If  $y^2 = 4x^3 + ax + b$  is an elliptic curve  $E$ . Then there exists a lattice  $\Lambda$  such that  $a = -60G_4(\Lambda)$ ,  $b = -140G_6(\Lambda)$  and that  $\mathbb{C}/\Lambda \cong E(\mathbb{C})$ .*

In other words, there is a correspondence between elliptic curves defined over  $\mathbb{C}$  and quotients of the complex plane modulo lattices. Also, since we determined that arbitrary (oriented) lattices can be obtained by those of the form  $\langle \tau, 1 \rangle$  with  $\tau$  in the upper half plane.

Another important function defined on lattices is the Weierstrass  $\wp$ -function.

**Definition 1.5.2.** *The Weierstrass  $\wp$  function is defined as follows:*

$$\wp(z, \Lambda) = \frac{1}{z^2} + \sum_{\omega \in \Lambda \setminus \{0\}} \left( \frac{1}{(z - \omega)^2} - \frac{1}{\omega^2} \right),$$

where  $\Lambda$  is the lattice associated with the elliptic curve.

## 1.6 Modular curves

The modular group is the quotient  $\Gamma(1) = SL(2, \mathbb{Z}) / \{\pm Id\}$ , its fundamental domain is the subset of the upper-half plane of complex numbers such that  $|z| \geq 1$  and  $Re(z) \geq 1/2$  and its generators are the matrices

$$S = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, T = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

And their action in  $\mathbb{H}$  is given by  $T\tau = \tau + 1$  and  $S\tau = -1/\tau$ .

Of importance for us are the so called modular curves. Define  $Y(1) = \mathbb{H}/\Gamma(1)$ , this is homeomorphic to a sphere without a point so we compactify it by adding a point to infinity  $X(1) = Y(1) \cup \{\infty\}$ . More formally:

**Definition 1.6.1.** *Let  $\mathbb{H}^* = \mathbb{H} \cup \mathbb{P}^1(\mathbb{Q})$ . The modular curve  $X(1)$  is defined as*

$$X(1) = \mathbb{H}^*/\Gamma(1)$$

Taking subgroups of the modular group is possible to extend this definition by considering the quotients  $X(n) = \mathbb{H}^*/\Gamma(n)$ , where

$$\Gamma(n) = \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in SL(2, \mathbb{Z}) : a, d \equiv 1 \pmod{n}, b, c \equiv 0 \pmod{n} \right\}$$

A congruence subgroup is a subgroup  $G$  of  $SL(2, \mathbb{Z})$  such that  $\Gamma(n) \leq G$  for some positive integer  $n$ . The groups  $X(n) = \mathbb{H}^*/\Gamma(n)$  are other examples of modular curves.

Cusps of a modular curves are elements in the quotient with representatives in  $\mathbb{P}^1(\mathbb{Q})$ .

Using the uniformization theorem we see that every class  $[\tau]$  (no in the cusp) of the modular curve corresponds to an elliptic curve so  $X(1)$  separates curves into isomorphism classes.

## 1.7 The $j$ -invariant

Let  $a, b$  be as in theorem 1.5.1 with  $y^2 = x^3 + ax + b$  the corresponding elliptic curve  $E$ . The discriminant for  $E$  is the quantity

$$\Delta = a^3 - 27b^2$$

**Definition 1.7.1.** *The  $j$ -invariant for an elliptic curve  $E$  is*

$$j(\tau) = 1728 \frac{a^3}{\Delta}$$

with  $\tau \in \mathbb{H}$  and  $a$  as in 1.5.1. The name comes from the fact that this function is an invariant for isomorphism classes of elliptic curves, See chapter III.1 proposition 1.4 in [21]:

**Theorem 1.7.1.** *Let  $K$  be a field and  $E$  an elliptic with Weierstrass form. Then*

- *Two elliptic curves are isomorphic over  $\bar{K}$  if and only if they have the same  $j$ -invariant.*
- *Let  $j_0 \in \bar{K}$ . There exists an elliptic curve defined over  $K(j_0)$  whose  $j$ -invariant is equal to  $j_0$ .*

## 1.8 Morphisms and Complex Multiplication

For the group law described in the previous section, the identity is assumed to be the point at infinity  $O$  of the curve seen as a projective curve. Group morphisms in this context are called isogenies. For elliptic curves  $E, E'$ , with point-wise addition  $End(E, E')$  forms a group and  $End(E, E) = End(E)$  is a ring with composition of isogenies as product.

The simplest example of isogeny is the multiplication by  $m \in \mathbb{C}$  map and if  $\Lambda$  is the associated lattice of  $E$ , then  $m\Lambda \subseteq \Lambda$ .

If the base field  $K \subseteq \mathbb{C}$ , then  $End(E)$  is either  $\mathbb{Z}$  or an order in an imaginary quadratic field. An analogue to this in the real case will be introduced in the following chapter.

Culmination of this elliptic curve overview is the analogous to the Kronecker-Weber theorem for the case of imaginary quadratic fields, see [21] C.11.2 or [3] XIII for the following two culminating results:

**Theorem 1.8.1.** *Weber-Fueter: Let  $K = \mathbb{Q}[\sqrt{D}]$ ,  $D < 0$  and  $E = \mathbb{C}/\Lambda$  an elliptic curve with complex multiplication.*

1.  $j(E)$  is an algebraic integer.
2.  $[K(j(E)) : K] = [\mathbb{Q}(j(E)) : \mathbb{Q}]$ .
3. The field  $H_K = K(j(E))$  is the maximal unramified abelian extension of  $K$ , i.e.,  $H_K$  is the Hilbert class field of  $K$ .
4.  $\text{Gal}(K(j(E))/K)$  permutes the  $j(E)$ s associated to  $\Lambda$  transitively.

This can be generalized to other subrings  $O_f$  of  $\mathcal{O}_K$ . Let  $O_f = \mathbb{Z} + \mathbb{Z}f\delta$ . When  $E$  has complex multiplication we have that  $\text{End}(E) \cong O_f$ . This goes both ways and for each possible  $O_f$  there is an elliptic curve with endomorphism ring isomorphic to it.

**Theorem 1.8.2.** *Up to isomorphism, there are finitely many curves  $E$  with  $\text{End}(E) \cong O_f$ , exactly  $|\text{Cl}(O_f)|$ .*

And correspondingly the  $j(E)$ s are algebraic integers.

Using class field theory, the maximal abelian extension can be obtained from  $H_K$  by adjoining certain elements obtained from the points of finite order of  $E$ . We are not going any further in this direction but in short terms we have that

$$K^{\text{ab}} = K(j(E), w(E_{\text{tors}}))$$

Where  $w$  is the Weber function from the elliptic curve to  $\mathbb{P}^1$  defined as

$$w(f(z)) = \begin{cases} \frac{ab}{\Delta} \wp(z, \Lambda), & \text{if } j(E) \neq 0, 1728, \\ \frac{a^2}{\Delta} \wp(z, \Lambda)^2, & \text{if } j(E) = 0, \\ \frac{b^2}{\Delta} \wp(z, \Lambda)^3, & \text{if } j(e) = 1728. \end{cases}$$

$f : \mathbb{C}/\Lambda \rightarrow E$  is an isomorphism and  $a, b$  are as in theorem 1.5.1.

It is also possible to find a curve  $E = \mathbb{C}/\Lambda$  such that  $\Lambda$  is the same as the lattice  $\mathcal{O}_K$ .

Ideally we would like to have a similar construction for real quadratic fields. That is a generating function for  $H_K$ . Unfortunately at this time there is no such description and further investigation is required.

## 1.9 Theta Functions

The theta function is denoted as  $\vartheta(z, \tau)$  and is defined as follows:

$$\vartheta(z, \tau) = \sum_{n \in \mathbb{Z}} e^{\pi i n^2 \tau + 2\pi i n z}$$

where  $z \in \mathbb{C}$  and  $\tau$  is in the upper half complex plane.

By setting  $a_n(\tau) = e^{\pi i n^2 \tau}$  we can see  $\vartheta$  as a Fourier series of a function of  $z$ .

$$\vartheta(z, \tau) = \sum_{n \in \mathbb{Z}} a_n(\tau) e^{2\pi i n z}$$

If  $\Lambda$  is a lattice, a short calculation shows that  $\vartheta(z + \tau, \tau) = e^{-\pi i \tau - 2\pi i z} \vartheta(z, \tau)$  and  $\vartheta(z + 1, \tau) = \vartheta(z, \tau)$  which shows that this functions is quasi-periodic in the variable  $z$  with respect to lattices.

We will work with a set of variations of the general theta function above called theta functions with characteristics.

Take  $a, b \in \mathbb{R}$  and  $f(z)$  holomorphic. Set

$$(S_b f)(z) = f(z + b)$$

and

$$(T_a f)(z) = e^{\pi i a^2 \tau + 2\pi i a z} f(z + a\tau)$$

This way we obtain

$$S_b \cdot T_a = e^{2\pi i a b} T_a \cdot S_b$$

Let  $\mathcal{G} = S^1 \times \mathbb{R} \times \mathbb{R}$ . Then

$$(U_{(\lambda, a, b)} f)(z) = \lambda (T_a \cdot S_b f)(z) = \lambda e^{\pi i a^2 \tau + 2\pi i a z} f(z + a\tau + b)$$

And we can define a group operation in  $\mathcal{G}$  as

$$(\lambda, a, b)(\lambda', a', b') = (\lambda\lambda' e^{2\pi i b a'}, a + a', b + b')$$

This group is a function space representation of the continuous Heisenberg group in quantum mechanics.

Let  $\mathcal{F} = \{(1, a, b) : a, b \in \mathbb{Z}\}$ , it is a subgroup of  $\mathcal{G}$  and by quasi-periodicity of  $\vartheta$ , this function is invariant under the action of  $\mathcal{F}$ .

Consider the sets

$$l\mathcal{F} = \{(1, la, lb)\} \subseteq \mathcal{F}$$

and

$$V_l = \{f : f \text{ is entire and invariant under } l\mathcal{F}\}$$

This is a vector space with basis

$$\vartheta_{a,b} = S_b T_b \vartheta = e^{2\pi i a b} T_a S_b \vartheta$$

Where  $a, b \in 1/l\mathbb{Z}$

In explicit form writing the action of the transformations  $S$  and  $T$ :

$$\vartheta_{a,b} = \sum_{n \in \mathbb{Z}} e^{\pi i (a+n)^2 \tau + 2\pi i (n+a)(z+b)}$$

The following properties hold

- $\vartheta_{0,0} = \vartheta$
- If  $a, b, c \in 1/l\mathbb{Z}$ , then  $S_c(\vartheta_{a,b}) = \vartheta_{a,b+c}$
- For  $a, d, b \in 1/l \in \mathbb{Z}$ , then  $T_d(\vartheta_{a,b}) = e^{-2\pi i d b} \vartheta_{a+d,b}$
- $\vartheta_{a+p,b+q} = e^{2\pi i a q} \vartheta_{a,b}$ , with  $p, q \in \mathbb{Z}$

There are four auxiliary theta functions that are just translates of  $\vartheta_{a,b}$  with  $a, b = 0, 1$ , denoted as  $\vartheta_i(z, \tau)$  for  $i = 1, 2, 3, 4$ . If  $q = e^{\pi i \tau}$ , they are defined as:

$$\theta_1(z, \tau) = 2 \sum_{n=0}^{\infty} (-1)^n q^{(n+1/2)^2} \sin((2n+1)z)$$

$$\theta_2(z, \tau) = 2 \sum_{n=0}^{\infty} q^{(n+1/2)^2} \cos((2n+1)z)$$

$$\theta_3(z, \tau) = 1 + 2 \sum_{n=1}^{\infty} q^{n^2} \cos(2nz)$$

$$\theta_4(z, \tau) = 1 + 2 \sum_{n=1}^{\infty} (-1)^n q^{n^2} \cos(2nz)$$

where  $z \in \mathbb{C}$  and  $q = e^{\pi i \tau}$ .

These are just examples of theta functions with rational (integer) characteristics described above.



We immediately obtain

$$\begin{aligned}\theta_1(z, \tau) &= -\vartheta_{11}(z, \tau) \\ \theta_2(z, \tau) &= \vartheta_{10}(z, \tau) \\ \theta_3(z, \tau) &= \vartheta_{00}(z, \tau) = \theta(z, \tau) \\ \theta_4(z, \tau) &= \vartheta_{01}(z, \tau)\end{aligned}$$

In the special case where  $z = 0$ , the auxiliary theta functions are usually called theta constants.

There are several reasons why theta functions are important. A great deal of the theory of elliptic curves can be derived in terms of theta functions, see for example the projective embedding of  $\mathbb{C}/\Lambda$  by means of theta functions in [16].

In the context of Manin's real multiplication program (to be described later) they are important because the absolute  $j$ -invariant of an elliptic curve can be described in terms of theta functions with characteristics.

If  $\lambda = \frac{\vartheta_{1,0}^4}{\vartheta_{0,0}^4}$ , then

$$j(\lambda) = \frac{4(1 - \lambda + \lambda^2)^3}{\lambda^2(1 - \lambda)^2}$$

This  $\lambda$  function is known as the modular lambda function and can be obtained via the Weierstrass  $\mathcal{P}$  function too.

There's a large and useful formulae for theta functions and  $j$ -invariants, see for example [9] and [16].

## 1.10 Concepts on functional analysis

As opposed to the customary approach of using class field theory, Manin proposed tackling the problem using non-commutative geometric tools. Since functional analysis plays an important role here, we will summarize the necessary definitions and concepts in this section.

**Definition 1.10.1.** *An algebra over a field  $K$  is a vector space  $A$  over  $K$  with a product such that for all  $a, b, c \in A$  and  $\alpha \in K$ :*

- $(ab)c = a(bc)$
- $a(b + c) = ab + ac$
- $(a + b)c = ac + bc$
- $\alpha(ab) = (\alpha a)b = b(\alpha y)$

*We say that  $A$  is commutative if the product commutes.*

All our algebras in the following will be defined over the complex numbers.

**Definition 1.10.2.** A Banach algebra  $\mathcal{A}$  is a normed algebra which is complete and satisfies  $\|ab\| \leq \|a\|\|b\|$ , for all  $a, b \in \mathcal{A}$ .

**Definition 1.10.3.** An involution on an algebra  $\mathcal{A}$  is a unary operation denoted by  $*$  that satisfies the following properties:

1.  $(a^*)^* = a$  for all  $a \in \mathcal{A}$  (involution property).
2.  $(ab)^* = b^*a^*$  for all  $a, b \in \mathcal{A}$  (anti-multiplicative property).
3.  $(\alpha a + \beta b)^* = \bar{\alpha}a^* + \bar{\beta}b^*$  for all  $a, b \in \mathcal{A}$  and  $\alpha, \beta \in \mathbb{C}$  (linearity property).

**Definition 1.10.4.** A Banach  $*$ -algebra  $\mathcal{A}$  is a Banach algebra equipped with an involution  $*$  that is compatible with the norm, i.e.,  $\|a^*\| = \|a\|$  for all  $a \in \mathcal{A}$ .

**Definition 1.10.5.** A  $C^*$ -algebra  $\mathcal{A}$  is a Banach  $*$ -algebra that satisfies the additional property,  $\|a^*a\| = \|a\|^2$  for all  $a \in \mathcal{A}$ .

In addition, a simple algebra is an algebra with no nontrivial two-sided ideals.

The most important example of a  $C^*$ -algebra is the set of bounded operators  $B(H)$  on a Hilbert space  $H$  with product given by composition and with adjoint operation as its involution. Also an example of the ubiquity of  $C^*$ -algebras is the following result.

**Proposition 1.10.1.** Let  $X$  be a compact Hausdorff space and  $C(X)$  its algebra of continuous functions, then  $C(X)$  is a commutative  $C^*$ -algebra.

The crucial theorem that simplifies the study of  $C^*$ -algebras is the following result similar to Cayley's theorem in group theory

**Theorem 1.10.1** (Gelfand-Naimark). Every  $C^*$ -algebra is isomorphic to a  $C^*$ -subalgebra of the algebra of bounded operators on certain Hilbert space.

## 1.11 Deep Learning and Terminology

Following the introduction, we would like to use a framework similar to [8] in order to exploit the capabilities of machine learning in our problem. Our goal is to use neural networks as a predictive tool to find properties of the Hilbert class field of a given quadratic extension and assess our intuition in the right direction. Below is a high-level introduction of how neural networks work and chapter 4 explains our implementation.

Deep learning is a subfield of machine learning. What distinguishes deep learning from traditional machine learning approaches is its ability to automatically learn and extract hierarchical representations of data. Deep neural networks are made up of multiple layers, each consisting of many interconnected neurons.

In a neural network, a *neuron* is a computational unit that receives inputs, performs a weighted sum of these inputs, and applies an activation function to produce an output.

*Layers* in a neural network transform data into increasingly different representations. This provides additional information about our desired output. As stated by Chollet in [4] Chapter 1, we can think of a neural network as a *multi-stage information distillation process*.

A neural network starts with an input layer of raw data/features. Each connection between neurons has weights, and each neuron has a bias, initially set to small random values. The layers then transform the input by weighted sums and application of activation functions.

An *activation function* is a function applied to the output of each neuron in a layer, it introduces non-linearity into the model since the usual transformation operations in layers are linear (a dot product and a sum) allowing neural networks to learn complex, non-linear relationships in the data.

After this step, an output layer produces predictions based on the transformed data. The error is measured by a loss/cost function that we want to minimise. The approach is to use the output of this loss function to adjust the weights in a direction that will reduce the loss in the next iteration. An *optimiser* is the algorithm used for this task.

To update the parameters of a neural network during training to minimize the loss function. Optimisers use gradient descent methods to adjust the weights and biases of the network's neurons. The central method for doing this in deep learning is the use of *backpropagation* which is an application of the chain rule of differentiation.

This process is iterative and so the weights are updated several times to improve the predictions.

Other concepts to consider when building a neural network are learning rate, batches and batch size and epochs:

The *learning rate* is a parameter that determines the step size or rate at which a model's parameters (weights and biases) are updated during training. It affects the convergence and stability of the training.

A batch refers to a subset of the training data that is processed together during a single iteration of training. Batch training is more efficient than training on individual data points and allows for parallel processing.

An *epoch* is a complete pass through the entire training dataset during the training of a neural network. In each epoch, the model is trained on all available training data, typically in multiple batches. Multiple epochs are used to improve the performance of the model over time.

In order to begin the training process described above, further work is required. A general outline could be described as follows.

- **Data Collection:**
  - Gather and compile the dataset containing input features and corresponding target labels.
- **Data Splitting:**
  - Training Set: Used to train the neural network.
  - Validation Set: Used for parameter tuning and model evaluation during training.
  - Test Set: Reserved for the final evaluation of the trained model's generalization performance. That is, data that has not been used before to train or tune the model.
- **Data Preprocessing:**
  - Feature Scaling/Normalization: Scale input features to ensure they have similar magnitudes, typically between 0 and 1, to aid in convergence during training.
  - Handling Missing Data: Address any missing or incomplete data by imputation or removal.
  - Encoding Categorical Variables (if applicable): Convert categorical variables into numerical representations.
- **Data Batching:**
  - Divide the training and validation data into mini-batches to facilitate efficient gradient computation and model training.

After the initial training, it is necessary to experiment with different hyperparameters (e.g., learning rate, batch size, number of layers) on the validation set to find optimal settings that reduce loss and improve the model's ability to predict and generalise.

## Chapter 2

# Real Multiplication

As opposed to the case  $K = \mathbb{Q}$  or  $K$  imaginary quadratic, most other cases where we want an explicit description of  $K^{ab}$  are far from being solved. Those that are already understood or partially solved are:

- $K = \mathbb{F}_p(t)$ , using Drinfeld modules [19]
- $K$  a finite extension of  $\mathbb{Q}_p$  the  $p$ -adic rationals, due to Lubin and Tate. They used a similar method to that of elliptic curves but over  $p$ -adic numbers [13].

In addition there's work due to Dasgupta and Kadke [7] that deals with totally real fields but is not published yet.

Before, we have shown the usefulness of the theory of elliptic curves with complex multiplication to address the explicit class field theory problem for the case of imaginary quadratic fields. In this chapter we are going to explain an analogous theoretical framework for the case of real quadratic fields due to Manin. For this purpose we will use the tools of non-commutative geometry.

### 2.1 Noncommutative spaces

Proposition 1.10.1 and theorem 1.10.1 provide a correspondence between categories:

$$\{\text{Locally compact Hausdorff spaces}\} \cong \{\text{commutative } C^*\text{-algebras}\}^{op}$$

And we can regard non-commutative spaces as a dual category of the category of non-commutative  $C^*$ -algebras (see chapter 1 of [11]).

$$\{\text{NC locally compact spaces}\} := \{\text{Non-commutative } C^*\text{-algebras}\}^{op}$$

Some objects of study in noncommutative geometry and mainly in Manin proposal are noncommutative tori, these are examples of  $C^*$ -algebras that we will study next following the treatment of [18].

**Definition 2.1.1.** *The rotation algebra  $A_\theta$  is a  $C^*$ -algebra generated by two unitary operators  $U$  and  $V$  subject to the relations  $UV = e^{2\pi i\theta}VU$  and  $U^*U = V^*V = I$ , where  $\theta$  is a real number. This algebra can be realized as a  $C^*$ -subalgebra of  $B(L^2(\mathbb{R}/\mathbb{Z}))$ , the algebra of bounded linear operators of square integrable functions over  $S^1$ .*

A list of basic facts

1. If  $\theta \in \mathbb{Z}$  the algebra  $A_\theta$  is isomorphic to  $C(\mathbb{T}^2)$ .
2. If  $\theta \in \mathbb{Q}$  the algebra  $A_\theta$  is isomorphic to the algebra of global sections of the endomorphism bundle of a complex vector bundle over  $\mathbb{T}^2$ .
3. If  $\theta \in \mathbb{R} \setminus \mathbb{Q}$  the algebra  $A_\theta$  is a simple  $C^*$ -algebra.

For irrational values of  $\theta$ ,  $A_\theta$  is also called *the algebra of continuous functions on the noncommutative torus  $\mathbb{T}_\theta^2$* . Following the correspondences in the prior section, as a topological space the noncommutative torus  $\mathbb{T}_\theta^2$  is defined as the dual object of  $C(\mathbb{T}_\theta^2) := A_\theta$ .

Smooth elements in this algebra are defined by  $C^\infty(\mathbb{T}_\theta^2) := \mathcal{A}_\theta$  and

$$\mathcal{A}_\theta = \left\{ \sum_{n,m \in \mathbb{Z}} a_{n,m} U^n V^m \in A_\theta \mid \{a_{n,m}\} \in \mathcal{S}(\mathbb{Z}^2) \right\}$$

where  $\mathcal{S}(\mathbb{Z}^2)$  denotes the space of sequences of rapid decay in  $\mathbb{Z}^2$ .

Noncommutative tori are defined in terms of function algebras. Morphisms  $\mathbb{T}_\theta^2 \rightarrow \mathbb{T}_{\theta'}^2$  between two noncommutative tori  $\mathbb{T}_\theta^2$  and  $\mathbb{T}_{\theta'}^2$ , and the corresponding  $C^*$ -algebra morphisms  $\mathcal{A}_{\theta'} \rightarrow \mathcal{A}_\theta$  are not enough in the context of noncommutative geometry. A more suitable notion of morphisms here is that of Morita equivalences.

Recalling the definition of projective modules and bimodules

**Definition 2.1.2.** A module  $P$  over a ring  $R$  is called *projective* if, for every surjective module homomorphism  $f : M \rightarrow N$  and every module homomorphism  $g : P \rightarrow N$ , there exists a module homomorphism  $h : P \rightarrow M$  such that  $f \circ h = g$ .

**Definition 2.1.3.** A  $R - S$ -bimodule is an abelian group  $M$  such that  $M$  is a left  $R$ -module and a right  $S$ -module. And such that, for all  $r \in R$ ,  $s \in S$ , and  $m \in M$ , we have:

$$r \cdot (m \cdot s) = (r \cdot m) \cdot s$$

**Definition 2.1.4.** Let  $R$  and  $S$  be rings. We say that  $R$  and  $S$  are *Morita equivalent* if there exists a bimodule  $P$  such that:

- $P$  is a projective left  $R$ -module and a projective right  $S$ -module.
- The functors
  - $\text{Hom}_R(P, -) : R\text{-Mod} \rightarrow S\text{-Mod}$
  - $\text{Hom}_S(P, -) : S\text{-Mod} \rightarrow R\text{-Mod}$

provide an equivalences of categories.

For real multiplication program we are looking for Morita equivalences between  $\mathcal{A}_{\theta'}$  and  $\mathcal{A}_\theta$ . These are given by the isomorphism class of a  $\mathcal{A}_{\theta'}$ - $\mathcal{A}_\theta$ -bimodule  $E$  which is projective and of finite-type both as a left  $\mathcal{A}_{\theta'}$ -module and as a right  $\mathcal{A}_\theta$ -module.

If such bimodule exists we say that  $\mathcal{A}_{\theta'}$  and  $\mathcal{A}_\theta$  are Morita equivalent. We can consider a Morita equivalence between  $\mathcal{A}_{\theta'}$  and  $\mathcal{A}_\theta$  as a morphism between  $\mathcal{A}_{\theta'}$  and  $\mathcal{A}_\theta$  inducing a morphism between  $\mathbb{T}_\theta^2$  and  $\mathbb{T}_{\theta'}^2$ . Composition of morphisms is provided by tensor product of modules.

Let  $SL_2(\mathbb{Z})$  act on  $\mathbb{R} \setminus \mathbb{Q}$  by fractional linear transformations, i.e. given

$$g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL_2(\mathbb{Z}), \quad \theta \in \mathbb{R} \setminus \mathbb{Q}$$

we take

$$g\theta = \frac{a\theta + b}{c\theta + d}.$$

Morita equivalences between noncommutative tori are characterized by the following result:



**Theorem 2.1.1.** (Rieffel [20]) *Let  $\theta', \theta \in \mathbb{R} \setminus \mathbb{Q}$ . Then the algebras  $\mathcal{A}_{\theta'}$  and  $\mathcal{A}_{\theta}$  are Morita equivalent if and only if there exist a matrix  $g \in SL_2(\mathbb{Z})$  such that  $\theta' = g\theta$ .*

In the context of noncommutative geometry we'll call our irrationalities  $\theta = \sqrt{D}$ .

Let  $\theta \in \mathbb{R}$  be quadratic irrational and  $\theta'$  its image under the nontrivial element of  $Gal(K/\mathbb{Q})$ . If  $K = \mathbb{Q}[\sqrt{\theta}]$ , given a matrix

$$\gamma = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \Gamma(1)$$

such that  $\epsilon = c\theta + d$  is a fundamental unit of  $\mathcal{O}_K$ . Denote by  $\gamma^n$  the powers of  $\gamma$ :

$$\gamma^n = \begin{bmatrix} a_n & b_n \\ c_n & d_n \end{bmatrix}$$

These matrices have fixed points  $\theta$  and  $\theta'$  and the fundamental unit still satisfies  $\epsilon^n = c_n\theta + d_n$ .

A profinite group is an inverse limit of finite groups. This means that it can be constructed by taking the limit of a sequence of finite groups, where each group is a quotient of the previous group.

From the above statements we can define a profinite group as follows. Since  $m|n$  implies  $c_m|c_n$

$$\mathcal{G}_{\epsilon} = \varprojlim \mathbb{Z}/c_n\mathbb{Z}$$

And we would like to think of this group as an analogous to a Galois group and see if it is of use in the study of the field  $K$ .

## 2.2 Real Multiplication

Inspired by Alain Connes work, Manin's real multiplication program is described in detail in his paper [14]. An informal summary is described below.

Stark conjectures suggest that an analogous to the Kronecker-Weber theorem or Complex Multiplication theory should exist for real quadratic fields. That is, a description of the Hilbert Class Field (and eventually  $K^{ab}$ ) given by a generating transcendental function.

This function say  $\kappa$  is  $z \mapsto e^{2\pi iz}$  in KW case and the  $j$ -invariant together with modular functions in Complex Multiplication.

Let  $K$  be a imaginary quadratic field. An elliptic curve with complex multiplication has a description of the form  $\mathbb{C}/\Lambda$  with  $\Lambda \cong \mathcal{O}_K$ . One may try to find an analogue in the real quadratic case, this by assuming the existence of a space  $\mathbb{R}/\mathcal{O}_K$ . We don't have anymore the nice discrete lattice structure but instead a dense subset of the real line, also this quotient happen to be non-Hausdorff.

Since part of the motivation for Alain Connes to develop noncommutative geometry was to aid in the study of "ill-behaved" quotients as  $\mathbb{R}/\mathcal{O}_K$  is reasonable to use his tools here.

Manin's proposal is to replace elliptic curves by noncommutative tori as defined in 2.1.1. The following theorem and its similarity with the analogous in complex multiplication gives the reason why this approach might be fruitful

**Theorem 2.2.1.** *Let  $\theta \in \mathbb{R} - \mathbb{Q}$ . The following are equivalent*

- $\mathcal{A}_\theta$  has nontrivial Morita autoequivalences (in the context of categories)
- There exists a matrix  $g \in SL_2(\mathbb{Z})$  such that  $\theta = g\theta$
- $\theta$  is a real quadratic so  $[\mathbb{Q}(\theta) : \mathbb{Q}] = 2$

A noncommutative torus that satisfies any of these conditions is said to have Real Multiplication (see [18]).

We denote by  $(\text{End}_{\text{Morita}}(\mathcal{A}_\theta))$  the set of isomorphism classes of  $\mathcal{A}_\theta - \mathcal{A}_\theta$  bimodules.

There is a well defined ring homomorphism  $\phi : \text{End}_{\text{Morita}}(\mathcal{A}_\theta) \rightarrow \mathbb{R}$

If  $\mathbb{T}_\theta^2$  is a real multiplication noncommutative torus then

$$\begin{aligned} \phi(\text{End}_{\text{Mor}}(\mathcal{A}_\theta)) &= \{\alpha \in \mathbb{R} \mid \alpha\Gamma_\theta \subset \Gamma_\theta\} \\ &= \mathbb{Z} + f\mathcal{O}_k \end{aligned}$$

where  $f \geq 1$  is an integer and  $\mathcal{O}_k$  is the ring of integers of the real quadratic field  $K = \mathbb{Q}(\theta)$ .

## Chapter 3

# AI and mathematical intuition

### 3.1 Framework

As its work title states, inspiration draws from the paper Advancing mathematics by guiding human intuition with AI [8].

The authors of the paper used machine learning (deep learning specifically) to find relations between mathematical objects and prove two conjectures, one in topology (knot theory) and the other in representation theory. They already have some conjectures about the possible behavior of some invariants in each case and used the method to reject or refine them and prove a proper result.

Some previous usages of machine learning in mathematics included: finding counterexamples, accelerate computations, generating conjectures.

The focus being the use of supervised learning to inquire about an hypothesized function, the paper shows a framework that we expect to follow in this work.

A summary of the above method is as follows, see figure 3.1 below.

There is a conjecture (maybe not necessarily be a strongly supported one) about the relationship between two features  $X(z)$  and  $Y(z)$  associated with an object  $z$ . This connection is thought to be a function  $\hat{f}$  such that  $\hat{f}(X(z)) \approx Y(z)$  and analysing could allow the mathematician to understand properties of the relationship.

In the supervised learning stage, the researcher proposes a hypothesis that there exists a relationship between  $X(z)$  and  $Y(z)$ . Now generate data for  $X(z)$

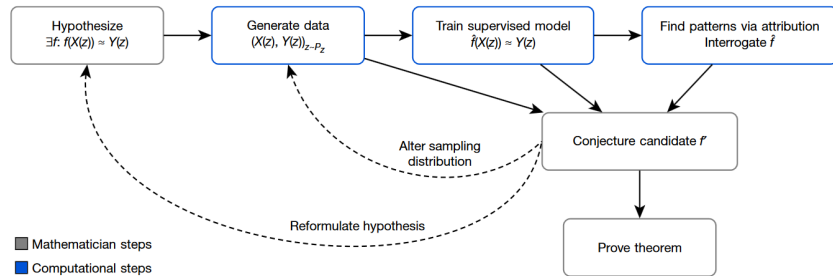


Figure 3.1: How to include AI in mathematical research, Chart from [8]

and  $Y(z)$  pairs and train a function  $\hat{f}$  that predicts  $Y(z)$ , using only  $X(z)$  as input.

The fundamental contribution here are all the possible functions that can be learned (linear, nonlinear, continuous or not, etc) given an enough amount of data. If the measures of performance and loss for  $\hat{f}$  are more accurate than what would be expected by chance, it indicates that there may be a relationship worth studying deeper.

In that case, attribution techniques like gradient saliency used in [8] can aid to understand  $\hat{f}$  and help conjecture a candidate  $f'$ . Though it may be out of the actual scope of this work, the use of these techniques could aid to identify which components of  $\hat{f}$  are more relevant for prediction the outputs  $Y(z)$  and therefore for  $f'$ .

By iterating the process just described by generating data with our candidate  $f'$  and reformulating our hypotheses given the attribution analysis when necessary, we obtain what the paper calls a ‘test bed for intuition’ by finding what relationships and features weight more so we can adjust our experiment.

[8] makes a simple example. Let  $z$  run between the set of convex regular polyhedra,  $X(z) \in \mathbb{Z}^2 \times \mathbb{R}^2$  be a tuple representing its vertices, edges, volume and surface area. Then  $Y(z) \in \mathbb{Z}$  would represent the number of faces of  $z$  given by Euler’s characteristic:  $X(z) \cdot (-1, 1, 0, 0) + 2 = Y(z)$ .

One of the two results proved in the paper using the framework was a relation between an algebraic invariant of hyperbolic knots  $z$  in topology, their signature  $\sigma$  and some of its geometric invariants slope, volume and ratio.

It was initially conjectured that there exist constants  $c_1$  and  $c_2$  such that, for every hyperbolic knot  $z$ ,

$$|2\sigma(z) - slope(z)| < c_1 vol(z) + c_2$$

This was supported by large amounts of data but they were able to find counterexamples to it. A framework as described in this section lead to a new conjecture that could be proved at the end with the support of this method:

**Theorem 3.1.1.** *There exists a constant  $c$  such that, for any hyperbolic knot  $z$ ,*

$$|2\sigma(z) - \text{slope}(z)| \leq \text{cvol}(z)\text{inj}(z)^{-3}$$

Where  $\text{slope}(z)$ ,  $\text{inj}(z)$  the injectivity radius,  $\text{vol}(z)$  are geometric invariants of the know and  $\sigma(z)$  is an algebraic invariant. For more information about know invariants see [1].

The conclusion being that instead of generating data-driven conjectures directly, we can iteratively guide our thinking in a direction with the aid of AI.

## 3.2 Current computations of $H_K$ for real quadratic fields

Where does the previous section fit into our study of quadratic fields and their abelian extensions? An example of how we can use the framework from Nature's paper is to set our objects  $z$  as a given quadratic field, say  $\mathbb{Q}[\sqrt{D}] = K$  with  $D < 0$ . The tuple  $X(z)$  storing invariants of  $K$  and finally  $Y(z)$  representing its corresponding Hilbert Class Field (or a description of it). With these pieces of information we could try to recreate in part, the theory of complex multiplication of elliptic curves and obtain an approximation of  $j(E)$ ,  $E$  an elliptic curve given by the orders of a lattice  $\Lambda$  in the ring of integers  $\mathcal{O}_K$ .

As complex multiplication for imaginary quadratic fields is already proven we can readily check results with existing algorithms see [5] for example.

Now for the real part of the story  $\mathbb{Q}[\sqrt{D}] = K$ ,  $D > 0$ , no proven method yields a description of a generating function for  $H_K$ , similar to the complex exponential for  $\mathbb{Q}$  or the  $j$ -invariant for an imaginary quadratic field. According to [6] and [14], there is evidence that a similar relation/generating function may exists.

Concretely in [6], Stark units are described as a special set of units in a number field that appear in Stark's conjectures on the values of L-functions at  $s = 0$ . They are used in the algorithm presented in the paper to compute the Hilbert class field of a real quadratic field. Algorithm involves computing the Stark units of the real quadratic field and using them to generate the Hilbert class field and then as output provide the minimal polynomial which defines  $H_K$ .

Even though Stark units come from the still unproven Stark conjectures in [22], Cohen's and Roblot's paper verifies independently of them that the generated field is in fact the Hilbert Class Field. This algorithm together with a complex multiplication version for imaginary fields is implemented in the computer algebra system PARI/GP. Created by Henri Cohen and designed for computations in number theory it is available at <https://pari.math.u-bordeaux.fr/>.

# Chapter 4

## Some Computations

### 4.1 Initial setting and generating Data

Our approach will deviate from the classical techniques of class field theory and instead try to find a way through different methods, here machine learning and possibly non commutative geometric tools. Following the framework in 3.1.

The objects of study are real quadratic fields and our target values are Hilbert Class Fields.

Let  $z = \mathbb{Q}[\sqrt{D}]$  with  $D > 0$  squarefree integer, be a real quadratic field. Mimicking [8] we'll work with a set of known data of  $z$ , and divide it into vectors  $X(z)$  and  $Y(z)$ .

For each  $z$ ,  $X(z)$  contains the following data

- $X_1 = D$
- $X_2 = \Delta_z$
- $X_3, X_4, X_5$ : the coefficients of the minimal polynomial of definition for  $\mathbb{Q}[\sqrt{D}]$
- $X_6$ : Class number for  $z$
- $X_7$ : Period length of the expression of  $\sqrt{D}$  as a continued fraction
- $X_8, \dots, X_{206}$ : numbers  $a_i$  in  $[a_0; a_1, a_2, \dots, a_n]$ . Depending on the period length most of the time just a few of these  $X_i$  apply for each field  $z$
- $X_{207}, X_{210}, X_{209}$ : Three numbers representing the decomposition of the class group of  $z$  according to the structure theorem for finitely generated abelian groups. In our data no real quadratic field had a Class Group where such a decomposition consisted of more than three groups

- $X_{210}$ : Regulator of  $z$
- $X_{211} = p_n/q_n$ , where  $n = X_7$  the period length of the continued fraction expansion of  $\sqrt{D}$
- $X_{212}, X_{213}$ :  $p_n$  and  $q_n$  as above respectively.

Similarly,  $Y(z)$  is made of the following components:

- $Y_1$ : Degree of the polynomial  $p$  that generates the Hilbert Class Field
- $Y_2, \dots, Y_{28}$ : Possible roots of  $p$
- $Y_{29}, \dots, Y_{40}$ : Coefficients of  $p$

We would expect the Hilbert Class Field of a given  $z$ , identified with the values  $Y(z)$  to be related to some if not all of these components  $X_i(z)$  through an unknown function  $\hat{f}(X(z)) \approx Y(z)$ .

The non-standard addition here is based mainly in non-commutative geometry and corresponding to the relationship between the expression of  $\sqrt{D}$  as a continued fraction and real quadratic fields. For example, Manin proposes in [14] the use of pseudolattices (an analog to lattices within  $\mathbb{R}$ ) as tools to find the Hilbert Class Field of  $\mathbb{Q}[\sqrt{D}]$ . Two such pseudolattices are generated by irrationals  $\theta, \theta'$  and they are isomorphic if and only if their continued fraction expansions coincide starting from some place.

This connection is supported by several sources, see for example [17] and [15], or [12].

The software used to generate the above data for each field was PARI/GP and SageMath. As described in 3.2 the  $Y(z)$  values we obtained via the algorithm in [6]. Fortunately for quadratic extensions, several invariants from Galois theory are trivial and the remaining features in  $X(z)$  had fast algorithms implemented in PARI/GP to compute them.

Initially, computations were performed on the same dataset as in Cohen and Roblot's paper [6] that consisted in all quadratic real fields of discriminant less than 2000 (around 700 fields) but it was later enlarged to include all fields  $z = \mathbb{Q}[\sqrt{D}]$  for  $0 < D < 10000$ ,  $D$  squarefree, that is, 6082 fields with  $\Delta_z < 40000$ .

The routine `quadhilbert(d)` which calculates the Hilbert class field of the quadratic field with discriminant  $d$ , running in a intel i5 at 3.60GHz and 16Gb of RAM took around 20 minutes to compute the Hilbert Class Field polynomial of 6082 fields without taking into account its roots.

Our purpose is to train a function  $\hat{f}$  that predicts  $Y(z)$  (or a part of it) using only  $X(z)$  values that do not require the use of Stark conjectures. If not



possible, then a posterior study and implementation of attribution techniques could lead us in the right direction.

## 4.2 Method Used

For simplicity the machine learning model applied is a feedforward neural network constructed with the `sequential` API, part of the Keras library in Python. As such, we focused on obtaining an individual target value, a root of the generating polynomial of the Hilbert Class Field, specifically  $Y_2$  the first found root since it was the only available for all fields in the dataset. A future approach could consist on finding the degree of the polynomial or using a more flexible method as the Functional API.

An original attempt consisted in trying to find the coefficients of the polynomial instead of its roots. Apparently the large nature of some coefficients for several fields made that computation too large for a model of this type even after removing outliers or applying different normalization techniques in the data.

The `pandas` data analysis library and `scikit learn` were used to manipulate the data. `TensorFlow` and `Keras` for models

Standard methods of data normalization and regularization were applied. `StandScaler` standarizes features removing mean and scaling to unit variance.

This normalization was applied firstly to make the data conform to what is expected by the method in the Keras library, and secondly to standardize our data due to the nature of the features. For example, discriminants and coefficients of the Hilbert class field tend to increase and are far from being uniformly distributed. Also, some others like the coefficient of  $x^2$  in the minimal defining polynomial of  $z$  have no deviation since this polynomial is by definition monic.

- First a feature selection was implemented to simplify the model so it would take into account only the first 50 and then the first 20 numbers in  $[a_0; a_1, a_2, \dots, a_n]$ .
- Another method was the removal of outliers, that included some fields with an abnormal/bigger than usual polynomial for the Hilbert class field.
- Finally, due to the high validation loss, simultaneous L1 and L2 regularization was applied at different penalization values.
- L1 regularization consists of a cost added proportional to the absolute value of the nets weight coefficients (L1 norm)

- L2 regularization consists of an added cost proportional to the square of the value of the weight coefficients (L2 norm)

See [4] for more in relation to L1 and L2 regularization.

In this work, each layer is a fully connected or dense layer. Data was tested with 2 up to 4 layers with different number of neurons 32, 64, 200, 300, 400, and 500.

For each of this possible configurations batch and epochs sizes of 32, 64, 128, 256 and 50, 100, 150, 200 respectively were tested with similar results.

Learning rates of 0.001, 0.01 and 0.5 were used at different stages of the experiment.

Model was implemented with the following functions

- Activation: Rectified linear activation function ReLU
- Loss function: mean squared error
- Optimizer: Adam optimization algorithm which is an extension of the stochastic gradient descent.

Since our data was small for the task at hand, we performed k-fold cross-validation at  $K = 4$  and  $K = 5$ . This consisted in splitting the data in  $K$  partitions and iterating training and evaluation steps.

### 4.3 Preliminary Results and Next Steps

The results of the experiment were fairly consistent between different runs using different parameters as mentioned in the previous section.

Across all different tests, training set loss was in the interval  $[5, 36]$ , validation set loss within  $[19.4336, 27]$  and Test loss inside  $[20, 32]$ . The best results obtained lead to a training set loss of 5.2557, a validation set loss of 19.4336 and a Test/total loss of 20.1540.

A relative and improving low training set loss and a high, constant or increasing validation test loss indicate that even though the algorithm is learning for the data, it is not good at generalizing and consequently inappropriate to make predictions for unseen data.

The training loss of 5.2557 indicates that, on average, the squared prediction errors (the squared differences between predicted and actual values) on the training data are approximately 5.2557. The target values in the full dataset were in the interval  $[-50, 0]$  and we considered this to be low for our case.

However, further investigation is required as this could be due to the somewhat small amount of data retrieved.

Similarly the relatively high validation loss compared to the training loss indicated to us that the model might not generalise well. A test/total loss was expected to be higher than the validation loss since the test set is totally new to the model and not tuned for the problem. In summary, this indicates an issue with the model's performance.

An attribution score plot for the most accurate iteration of the experiment using the absolute value of the gradients of the root with respect to the inputs is given below

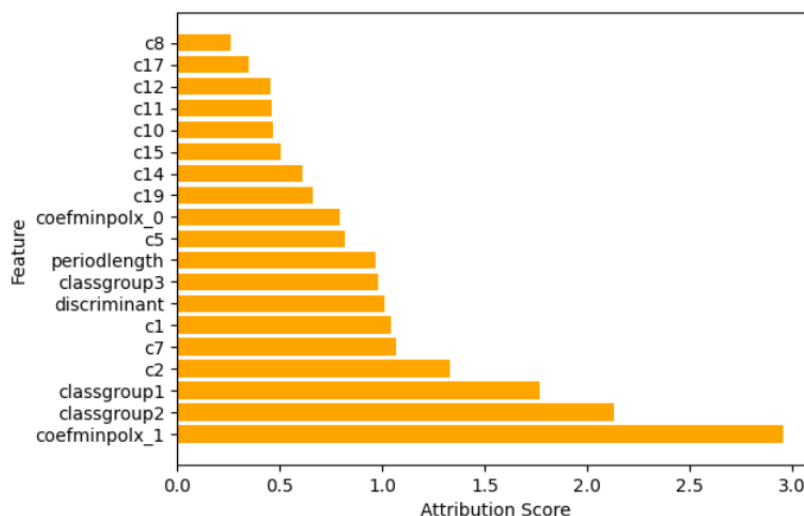


Figure 4.1: Attribution score

Where `coefminpolx1` is the coefficient of  $x$  in the polynomial generator of  $z = \mathbb{Q}[\sqrt{D}]$ , `classgroupi` corresponds to a component in the finite abelian decomposition of  $Cl(z)$ . Similarly the  $c_i$ 's are the numbers in the continuous fractional expansion of  $\sqrt{D}$ .

One of the purposes of [8] was to introduce a framework that uses attribution techniques in machine learning to guide mathematical thought. In our case figure 4.3 shows that `coefminpolx1` and the class group having bigger attribution scores might be more important to our study than other features. This however is hypothetical and is subject to the results given by a further improved model with better generalisation.

As mentioned in [8]. The fact that the model cannot predict a relationship between  $X(z)$  and  $Y(z)$  better to what is expected by chance doesn't mean that such relationship or pattern doesn't exist.

Even if the results are inconclusive, we can still gain some insights. First, possible ways to improve our experiment including generating additional data including new features/invariants of  $\mathbb{Q}[\sqrt{D}]$ , choosing different and more appropriate target values as the degree of the generating polynomial. Furthermore we are not led to abandon the deep learning approach as the training set results and attribution do not rule out a relationship between our features and target values.

Currently we are working in further implementation of Manin's real multiplication program into the algorithm.

A possible approach is to compute special values of theta functions with characteristics  $(a, b) = (p_n, q_n)$ , where  $p_n, q_n$  are the numerator and denominator of the  $n$ -th convergent of  $\sqrt{D}$  and  $n$  is the period length. This is outlined in Manin's paper [14] with ideas that come from Hecke.

As of our current knowledge there is not a built-in library to work with theta functions with arbitrary rational characteristics in PARI/GP, Sage or Mathematica, the computer algebra systems inquired for this work. However a custom library for SageMath has been written and published in [23]. We expect this library to help us in further developments.

# Bibliography

- [1] Knotinfo - the knot atlas. <https://knotinfo.math.indiana.edu/>. Accessed: 2023.
- [2] M. F. Atiyah and I. G. Macdonald. *Introduction to commutative algebra*. Addison-Wesley Series in Mathematics. Westview Press, Boulder, CO, economy edition, 2016. For the 1969 original see [MR0242802].
- [3] J. W. S. Cassels and A. Fröhlich, editors. *Algebraic number theory*. Academic Press, London; Thompson Book Co., Inc., Washington, D.C., 1967.
- [4] François Chollet. *Deep Learning with Python*. Manning Publications, Shelter Island, NY, 2021.
- [5] Henri Cohen. *A course in computational algebraic number theory*, volume 138 of *Graduate Texts in Mathematics*. Springer-Verlag, Berlin, 1993.
- [6] Henri Cohen and Xavier-François Roblot. Computing the hilbert class field of real quadratic fields. *Math. Comp.*, 69:1229–1244, 2000.
- [7] Samit Dasgupta and Sayan Kadke. Brumer-stark units and explicit class field theory, 2023. <https://arxiv.org/abs/2103.02516>.
- [8] Alex Davies, Petar Veličković, and Lars et al. Buesing. Advancing mathematics by guiding human intuition with ai. *Nature*, 600:70–74, 2021.
- [9] Hershel M. Farkas and Irwin Kra. *Theta constants, Riemann surfaces and the modular group*, volume 37 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2001. An introduction with applications to uniformization theorems, partition identities and combinatorial number theory.
- [10] Eknath Ghate. The Kronecker-Weber theorem. In *Cyclotomic fields and related topics (Pune, 1999)*, pages 135–146. Bhaskaracharya Pratishthana, Pune, 2000.
- [11] Masoud Khalkhali. *Basic noncommutative geometry*. EMS Series of Lectures in Mathematics. European Mathematical Society (EMS), Zürich, second edition, 2013.

- [12] J. Lewis and D. Zagier. Period functions and the Selberg zeta function for the modular group. In *The mathematical beauty of physics (Saclay, 1996)*, volume 24 of *Adv. Ser. Math. Phys.*, pages 83–97. World Sci. Publ., River Edge, NJ, 1997.
- [13] Jonathan Lubin and John Tate. Formal complex multiplication in local fields. *Ann. of Math. (2)*, 81:380–387, 1965.
- [14] Yu. I. Manin. Real multiplication and noncommutative geometry (ein alterstrraum). In *The legacy of Niels Henrik Abel*, pages 685–727. Springer-Verlag, 2004.
- [15] Yuri I. Manin and Matilde Marcolli. Continued fractions, modular symbols, and noncommutative geometry. *Selecta Math. (N.S.)*, 8(3):475–521, 2002.
- [16] David Mumford. *Tata lectures on theta. I*. Modern Birkhäuser Classics. Birkhäuser Boston, Inc., 2007. Reprint of the 1983 edition.
- [17] Jorge Plazas. Arithmetic structures on noncommutative tori with real multiplication. *Int. Math. Res. Not. IMRN*, (2):Art. ID rnm147, 41, 2008.
- [18] Jorge Plazas. Heisenberg modules over real multiplication noncommutative tori and related algebraic structures. In *Geometric and topological methods for quantum field theory*, pages 405–421. Cambridge Univ. Press, Cambridge, 2010.
- [19] Bjorn Poonen. Introduction to Drinfeld modules. In *Arithmetic, geometry, cryptography, and coding theory 2021*, volume 779 of *Contemp. Math.*, pages 167–186. Amer. Math. Soc., [Providence], RI, [2022] ©2022.
- [20] Marc A. Rieffel.  $C^*$ -algebras associated with irrational rotations. *Pacific J. Math.*, 93(2):415–429, 1981.
- [21] Joseph H. Silverman. *The arithmetic of elliptic curves*, volume 106 of *Graduate Texts in Mathematics*. Springer, Dordrecht, second edition, 2009.
- [22] H. M. Stark.  $L$ -functions at  $s = 1$ . III. Totally real fields and Hilbert’s twelfth problem. *Advances in Math.*, 22(1):64–84, 1976.
- [23] Christopher Swierczewski and Bernard Deconinck. Computing Riemann theta functions in Sage with applications. *Math. Comput. Simulation*, 127:263–272, 2016.
- [24] Mak Trifković. *Algebraic theory of quadratic numbers*. Universitext. Springer, 2013.