UNIVERSITAT DE BARCELONA

ADVANCED MATHEMATICS

MASTER'S FINAL PROJECT

Large images for Galois representations attached to generic modular forms

---

Author:
Miquel Guiot Cusidó

Supervisor:
Luis Victor Dieulefait

Facultat de Matemàtiques i Informàtica

June, 28, 2023

Abstract

The aim of this project is to study a theorem of Ribet stating that the images of the Galois representations attached to modular forms without Complex Multiplication are large for almost every prime. Firstly, the needed background is introduced in the form of some definitions and basic properties of modular forms and Galois representations. Later, the subgroup classification of general linear groups over finite fields is presented, as well as other useful results from group theory. Finally, Ribet's theorem is stated and proved using all the tools from algebraic number theory and group theory developed in the previous chapters.

2020 Mathematics Subject Classification. 11F11, 11F80

Agraïments

Vull agrair al Dr. Luis Victor Dieulefait no tan sols el seu acompanyament i guiatge al llarg de tot el treball, des de la proposta del tema fins als darrers consells, sinó també l'haver-me introduït en els apassionants mons de les formes modulars i les representacions de Galois. Ha estat un plaer poder desenvolupar aquest treball sota la teva tutorització.

També dono les gràcies a tothom qui, d'alguna manera o altra, m'ha ajudat durant la realització del màster: familiars, amics, companys i professors.

Finalment, m'agradaria fer un agraïment especial als meus pares, a qui agraeixo tota l'estima i suport que sempre m'han mostrat, i al meu germà Quim, de qui he après la importància de la constància i l'esforç. Sense vosaltres no hagués estat possible. Moltes gràcies de tot cor.

Contents

Chapter 1

Introduction

1.1 A bridge between number theory & complex analysis

Number theory can be defined as the branch of mathematics that studies integer numbers, arithmetic functions, and any problem that arise from them. As happens with all mathematical fields, in number theory there are still lots of open problems, questions for which there is no concrete answer yet. Among them, some of the most popular ones are the twin prime conjecture, the Riemann hypothesis, or the study of absolute Galois group.

Each of these problems is interesting in its own right, and the effort that many mathematicians have put into them has led to the development of several new and valuable mathematical constructions. Therefore, it makes sense to take a closer look at at least one of these problems. We opt for the study of the absolute Galois group.

We start by discussing a common question in number theory.

Question 1. *To find all roots of a given polynomial $p(x) \in \mathbb{Q}[x]$.*

To answer Question 1 it suffices to decompose $p(x)$ into linear factors. However, this decomposition can not always be done in $\mathbb{Q}$ since not all rational polynomials have rational roots. For example, $p(x) = x^2 + 1 \in \mathbb{Q}[x]$ factors as $(x+i)(x-i)$ and $\pm i \notin \mathbb{Q}$. Precisely this phenomena motivates the definition of the splitting field of a rational polynomial.

Definition 2. *The splitting field of a polynomial $p(x) \in \mathbb{Q}[x]$ is a field extension L of $\mathbb{Q}$ over which $p(x)$ factors as*

$$p(x) = c \prod_{i=1}^{\deg(p)} (x - a_i),$$

where $c \in \mathbb{Q}$, and the roots a_i generate L over $\mathbb{Q}$.

In general, splitting fields of rational polynomials are well understood and have nice properties, such as being unique and finite extensions. However, we can go a step further and generalize Question 1 as follows.

Question 3. *To determine the smallest field extension of $\mathbb{Q}$ in which the roots of all rational polynomial belong.*

Now, instead of looking for the splitting field of a particular rational polynomial, Question 3 requires to look at all rational polynomials at once and to consider the field extension generated by all the roots. This construction is known as the algebraic closure of $\mathbb{Q}$ and its definition can be synthesised in the following way.

Definition 4. *The algebraic closure of* $\mathbb{Q}$*, denoted by* $\overline{\mathbb{Q}}$*, is the field consisting of those complex numbers which are roots of some non-zero rational polynomial, i.e.*

$$\overline{\mathbb{Q}} = \{\alpha \in \mathbb{C} : p(\alpha) = 0 \text{ for some } p(x) \in \mathbb{Q}\}.$$

Each of the roots α *is defined as an algebraic number over* $\mathbb{Q}$*.*

Contrary to the case of the splitting fields, the algebraic closure of $\mathbb{Q}$ is a mathematical object of which we do not have that much knowledge, and one of the reasons is because it does not have as good properties. For example, the field extension $\overline{\mathbb{Q}}/\mathbb{Q}$ is infinite. Therefore, an strategy consists in using other mathematical constructions to study the infinite extension $\overline{\mathbb{Q}}/\mathbb{Q}$. In particular, since we are dealing with field extensions, it make sense to use Galois theory. This leads to the definition of the absolute Galois group.

Definition 5. *The absolute Galois group corresponds to* $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$*.*

Notice that thanks to Galois theory we have switched from studying the infinite field extension $\overline{\mathbb{Q}}/\mathbb{Q}$ to studying $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$. These are good news for our interests, since now we can use all the machinery from group theory. In this regard, a common and useful technique when dealing with complicated groups is to study them via its actions on finite vector spaces. This is what is known as group representations.

The theory of group representations is studied in detail in Chapter 2, but for the moment it suffices to know that when we restrict group representations to the case of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$, they are called Galois representations.[1] From here, we could simply limit ourselves to studying generic Galois representations, with the disadvantage that they are objects about which we do not have too much information. Therefore, we can mimic the previous strategy and try to combine these representations with some mathematical techniques from other fields to obtain more manageable structures.

It turns out that the additional mathematical tool that we require comes from complex analysis, a branch that is, apparently, far from number theory. In particular, it corresponds to the notion of modular forms, which are holomorphic functions from the upper half plane that satisfy some symmetry conditions with respect to particular subgroups of $\mathrm{SL}(2,\mathbb{Z})$.[2]

The choice of modular forms is by no means accidental and is due to a result known as the Eichler-Shimura relation. Broadly speaking, the Eichler-Shimura relation allows to associate an abelian variety to a particular case of modular form known as an eigenform. Then, by means of another construction known as the Tate module, we can attach a Galois representation to the abelian variety. Hence, since starting from a modular form we end up obtaining a Galois representation, the Eichler-Shimura relation can be seen as a bridge between complex analysis and number theory. Indeed, it is simply the tip of a much

[1] This is a simplification of the actual definition, see Section 2.2.2 for the precise definition.

[2] Again, this is a simplification of the actual definition, see Section 3.1 for the precise definition.

deeper framework of conjectures and results about the connection between these areas known as the Langlands program.

For the case we are interested in, it turns out that Galois representation arising from this construction are endowed with some additional properties that we can take advantage of to study them. From there, one of the most natural questions when dealing with this type of Galois representations is what their image is like. This is precisely what Ribet studied in [Rib85], where he stated and proved a version of the following theorem, which is the main result of study throughout this thesis.

Theorem 6 (Ribet)**.** *The images of Galois representations attached to modular forms without Complex Multiplication are as large as possible for almost every prime.*

1.2 Contents of the thesis

The aim of this thesis is to study the images of Galois representations attached to modular forms without Complex Multiplication, focusing on Theorem 6, which sates that these images are as large as possible for almost every prime. This will involve the understanding of several topics from a priori unrelated mathematical areas. From one side, the notion of Galois representations belongs to the core of this work, both for stating and proving Ribet's theorem. From another side, the theory of modular forms is crucial not only for the theorem itself but also for the construction of all the previous machinery. In addition, several important results arising from group theory also play an important role during the proof of the theorem.

Often, at this mathematical level, topics such as the one treated in this thesis are simply covered in papers in a succinct manner, omitting most of the steps, and providing neither motivation on the ideas nor the necessary mathematical background. Therefore, this work intends to give a self-contained overview of all the notions needed to understand both the statement and the proof of Ribet's theorem, while attempting to give intuition behind all the constructions that appear in it. In that sense, For the sake of simplicity, most of the needed background on related topics is simply introduced by stating their definitions and useful results, motivating their construction, and giving precise references on the proofs. The main theorem is stated and proved with all detail following the work of Ribet in [Rib85], but extending in detail most of his arguments.

To summarize, all the chapters of this thesis are listed below, as well as a brief description of their contents.

First of all, the following chapter introduces the notion of Galois representations, covering from some basic notions of group representations to other advanced topics such as infinite Galois extensions. In addition, it also discusses other related topics from algebraic number theory including the l-adic numbers, ramification of extensions, and group characters.

Chapter 3 is dedicated to the theory of modular forms. It starts by motivating the topic and giving the first definitions and gradually introduces more advanced notions until reaching the construction of newforms, essential for the statement of Ribet's theorem. It

also presents other important concepts of this theory, such as modular forms with complex mutiplication and with inner twists.

Chapter 4 refers to some useful results in group theory. In particular, it is mainly focused on the subgroup classification of $\mathrm{GL}(2, \mathbb{F}_q)$ given by Dickson, including not only the statement of the theorem but also the definitions and some properties of all the groups listed there.

Chapter 5 contains the main goal of the thesis, the statement and proof of Ribet's theorem on the images of Galois representations attached to modular forms without Complex Multiplication. Both the statement and the proof are treated in detail, explaining the intuition behind each step. All the theory developed in the previous chapters is used, as well as some additional results that are introduced in the chapter itself.

Chapter 6 contains all the work done during the thesis in the form of conclusions.

Chapter 2

Galois representations

The history of the origins of Galois theory is as well known as dramatic. It started with the work of Galois in the 19th century, who related the resolubility by radicals of polynomials to the study of the group of permutations of its roots before dying in a duel at the age of 20. Thereafter, the field kept evolving due to the work of several mathematicians such Dedekind and Artin, who popularized it during the last century. In particular, its obvious connections with group theory lead to what is known as Galois representation theory, which consists in the study of Galois groups in terms of its actions over finite vector spaces.

These days, Galois theory, and consequently Galois representations, are present in several research areas inside mathematics, being number theory and algebraic geometry the main ones.

2.1 First notions

The aim of this section is to provide a brief introduction to the field of Galois representations by presenting the main definitions and results needed in the following chapters. We begin by covering the notion of group representations.

2.1.1 Group representations

Definition 7. *Let G be a group, K a field, and V a vector space of dimension n over K. A representation of G on V is a group homomorphism from G to* $\mathrm{GL}(V)$*. It is denoted by*

$$\rho : G \to \mathrm{GL}(V).$$

Remark 8. Since there is an isomorphism between $\mathrm{GL}(V)$ and $\mathrm{GL}(n, K)$, often group representations are expressed as $\rho : G \to \mathrm{GL}(n, K)$.

One of the most straightforward examples of group representations are the ones where G is a finite permutation group.

Example 9. Let $G = S_5$ and consider $V = \mathbb{C}^5$. It is clear that given $\sigma \in G$, it acts on V via the permutation matrix $A = (a_{i,j})$ given by

$$a_{i,j} = \begin{cases} 1 \text{ if } \sigma^{-1}(i) = j \\ 0 \text{ if } \sigma^{-1}(i) \neq j \end{cases}.$$

Therefore, setting $\sigma = (2, 3, 5)$ we have that

$$\rho(\sigma) = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix}$$

As happens with most mathematical constructions, the notion of irreducibility is also present in the context of group representations.

Definition 10. *Let G be a group, V a finite vector space over a field K, and ρ a representation of G on V. A subrepresentation of V is a subspace $W \subset V$ such that for all $w \in W$ and $g \in G$, $\rho(g)w \in W$. In that context, ρ is irreducible if its only subrepresentations are V and $\{0\}$.*

Remark 11. Notice that we can interpret Definition 10 in matricial terms. In particular, a representation ρ of a group G is reducible if and only if for all $g \in G$ the matrices $\rho(g)$ can be expressed as upper triangular matrices up to conjugate by the same invertible matrix.

The following Theorem due to Maschke answers the question of whether group representations can be expressed in terms of irreducible group representations.

Theorem 12. *Let G be a finite group and K a field with characteristic not dividing the size of G. Then any representation $\rho : G \to \mathrm{GL}(n, K)$ can be decomposed into the direct sum of irreducible representations.*

Proof. See [Mas98]. □

2.1.2 *l*-adic numbers

When working with Galois theory, the main object of interest is $\mathrm{Gal}(\overline{\mathbb{Q}}/Q)$ because of the reasons mentioned in Chapter 1. However, often it is useful to also consider other advanced algebraic constructions. In particular, in further sections the l-adic numbers will play a key role for our purposes.

Roughly speaking, the l-adic numbers simply correspond to an extension of the field of rational numbers $\mathbb{Q}$ different from the one leading to the real numbers $\mathbb{R}$. In particular, this construction is based on the idea of establishing the closeness of two given numbers in terms of congruences modulo a prime l. Thus, two numbers will be close if they are congruent modulo a high power of l.

In order to be precise, we first need to introduce the notion of inverse limit.

Definition 13. *Consider a sequence of sets* $(X_n)_{n\in\mathbb{N}}$ *and maps* $f_n : X_{n+1} \to X_n$*. The inverse limit, denoted by* $\varprojlim X_n$*, corresponds to the subset of* $\prod_{n\in\mathbb{N}} X_n$ *given by*

$$\varprojlim X_n := \{(a_n)_{n\in\mathbb{N}} \in \prod_{n\in\mathbb{N}} X_n : f_n(a_{n+1}) = a_n \text{ for all } n \in \mathbb{N}\}.$$

Once the notion of inverse limit is introduced, the l-adic integers simply correspond to a particular case of this construction by considering sequences of integers congruent modulo powers of a prime l.

Definition 14. *Let* l *be a prime,* $X_n = \mathbb{Z}/p^n\mathbb{Z}$*, and* $\pi_n : \mathbb{Z}/p^{n+1}\mathbb{Z} \to \mathbb{Z}/p^n\mathbb{Z}$ *the natural projection for all* $n \in \mathbb{N}$*. Then, the* l*-adic integers, denoted by* $\mathbb{Z}_l$*, correspond to the inverse limit of* $(\mathbb{Z}/p^n\mathbb{Z})_{n\in\mathbb{N}}$*, i.e.*

$$\mathbb{Z}_l := \varprojlim \mathbb{Z}/p^n\mathbb{Z} = \{(a_n)_{n\in\mathbb{N}} \in \prod_{n\in\mathbb{N}} \mathbb{Z}/p^n\mathbb{Z} : \pi_n(a_{n+1}) = a_n \text{ for all } n \in \mathbb{N}\}.$$

Remark 15. Since $\mathbb{Z}/p^n\mathbb{Z}$ is a ring and $\pi_n : \mathbb{Z}/p^{n+1}\mathbb{Z} \to \mathbb{Z}/p^n\mathbb{Z}$ is a ring homomorphism for all $n \in \mathbb{N}$, we can endow $\mathbb{Z}_l$ with a ring structure simply by defining the addition and the product component-wise. Moreover, it can be easily checked that $\mathbb{Z}_l$ is indeed an integral domain.

Remark 16. From Definition 14 one can check that $\mathbb{Z}_l^\times = \mathbb{Z}_l \setminus l\mathbb{Z}_l$.

Remark 17. Notice that the map

$$\begin{aligned} \mathbb{Z} &\to \mathbb{Z}_l \\ a &\mapsto (a + l\mathbb{Z},\ a + l^2\mathbb{Z}, \ldots) \end{aligned}$$

is a ring injection. Thus, we can view $\mathbb{Z}$ as a subring of $\mathbb{Z}_l$.

According to Remark 15, $\mathbb{Z}_l$ is an integral domain. Thus, it makes sense to construct its fraction field, i.e. the smallest field in which it can be embedded. This is precisely the notion of the l-adic rational numbers.

Definition 18. *Let* l *be a prime. The field of* l*-adic rational numbers, denoted by* $\mathbb{Q}_l$*, corresponds to the fraction field of the* l*-adic integers, i.e.*

$$\mathbb{Q}_l = \mathrm{Frac}(\mathbb{Z}_l).$$

Remark 19. There are some alternative ways of constructing both the l-adic rational numbers and the l-adic integers that obviously do not change any of their properties. In particular, Q_l can be defined as the completion of Q with respect to a non-Archimedean norm known as the l-adic absolute value. Then, the construction of $\mathbb{Z}_l$ follows as the closed ball centered at the origin with radius 1 with respect to that norm. See [KKS95], Section 2.4. for details on this construction.

2.1.3 Arithmetic in extensions

Apart from the l-adic numbers, the other important ingredient to take into account when dealing with Galois representations is the theory of arithmetic in extensions, being ramification and the decomposition group its main concepts.

In order to discuss about ramification, we first need to introduce the concept of ring of integers.

Definition 20. *Let $K/\mathbb{Q}$ a finite field extension. Its ring of integers, denoted by $\mathcal{O}_K$, corresponds to the ring of all algebraic integers contained in K, where an algebraic integer corresponds to a root of a monic polynomial with integer coefficients.*

Remark 21. It is easy to check that $\mathcal{O}_K$ is a ring.

In a general sense, the idea of ramification consists in studying how the ideal generated by a prime number decompose in the ring of integers of a given extension. As we will see in the next section, the behaviour of primes in this context is crucial and provides useful information to study Galois representations.

Definition 22. *Let $K/\mathbb{Q}$ a finite extension of degree n and $\mathcal{O}_K$ its ring of integers. A prime p generates a prime ideal of $\mathbb{Z}$ that factors in $\mathcal{O}_K$ as*

$$p\mathcal{O}_K = \mathfrak{p}_1^{e_1}...\mathfrak{p}_r^{e_r},$$

where $\mathfrak{p}_i$ are distinct prime ideals of $\mathcal{O}_K$ and e_i are positive integers. This factorization is unique up to reordering and p is ramified in $K/\mathbb{Q}$ if there is some $e_i > 1$.

Remark 23. In the set up of Definition 22, we can define positive integers $f_1, ..., f_r$ such that $\mathcal{O}_K/\mathfrak{p}_i \cong \mathbb{F}_{p^{f_i}}$. Then, by construction we get $e_1 f_1 + ... + e_r f_r = n$.

Remark 24. Due to a construction known as the discriminant of a number field it can be easily proven that the set of ramified primes in a given finite extension is always finite.

In the case of quadratic extensions over $\mathbb{Q}$, the definition of the discriminant notion referred in the previous remark is straightforward to define and allows us to state a result that will be useful in later sections.

Definition 25. *Let d be a square-free integer different from 0 and 1, and $K = \mathbb{Q}(\sqrt{d})$ the quadratic field associated to d. The discriminant of K, denoted by Δ_K, is given by*

$$\Delta_K := \begin{cases} d \text{ if } d \equiv 1 \pmod 4 \\ 4d \text{ if } d \equiv 2, 3 \pmod 4 \end{cases}.$$

We can characterize the primes ramifying in a finite rational extension in terms of the discriminant by means of the following result due to Dedekind.

Theorem 26 (Dedekind)**.** *Let K be a quadratic extension over $\mathbb{Q}$. Then a prime $p \in \mathbb{Z}$ ramifies in K if and only if p divides Δ_K.*

Proof. See [Ded82]. □

Remark 27. It turns out that Theorem 26 is simply a particular case of a much more general result for finite extensions over any field. However, for our purposes it is enough with the version stated above.

In addition, ramification is closely related to the construction of the Frobenius automoprhism, which in turn is linked to the decomposition and inertia groups. We define all this concepts next.

Definition 28. *Let $L/\mathbb{Q}$ a finite Galois extension of degree n, $\mathcal{O}_L$ its ring of integers, and $G \coloneqq \mathrm{Gal}(L/\mathbb{Q})$. Given a prime p, we take a prime ideal $\mathfrak{p}$ in L lying on the factorization of p. The decomposition group of $\mathfrak{p}$, denoted by $D_\mathfrak{p}$, corresponds to the subgroup of G formed by the elements that fix $\mathfrak{p}$, i.e.*

$$D_\mathfrak{p} \coloneqq \{\sigma \in G : \sigma(\mathfrak{p}) = \mathfrak{p}\}.$$

Remark 29. From the set up of Definition 28, we denote $k_\mathfrak{p} \coloneqq \mathcal{O}_L/\mathfrak{p}$ and $k \coloneqq \mathbb{Z}/p\mathbb{Z}$ to both residue fields. Then, there exists a natural reduction morphism

$$\phi_\mathfrak{p} : D_\mathfrak{p} \to \mathrm{Gal}(k_\mathfrak{p}/k),$$

which is always surjective.

Combining Definition 28 and Remark 29 we can define both the inertia group and the Frobenius automorphism.

Definition 30. *The inertia group of $\mathfrak{p}$, denoted by $I_\mathfrak{p}$, corresponds to the kernel of $\phi_\mathfrak{p}$, i.e.*

$$I_\mathfrak{p} \coloneqq \{\sigma \in G : \sigma(x) \equiv x \pmod{\mathfrak{p}} \text{ for all } x \in \mathcal{O}_L\}.$$

The following proposition illustrates the connection between the inertia group and the ramification of primes. In particular, it states that ramification can be controlled by the structure of the inertia group.

Proposition 31. *Let $L/\mathbb{Q}$ a finite Galois extension, $\mathcal{O}_L$ its ring of integers, and p a prime number. Then p is unramified in K if and only if for any prime ideal $\mathfrak{p}$ in $\mathcal{O}_L$ lying on the factorization of p the inertia group $I_\mathfrak{p}$ is trivial.*

Proof. See [AM18], Theorem 28. □

Remark 32. According to Proposition 31, the ramification of primes characterizes when $\phi_\mathfrak{p}$ is an isomorphism. In particular, when p is unramified we get that $D_\mathfrak{p} \cong \mathrm{Gal}(k_\mathfrak{p}/k)$.

Remark 33. Recall that a classical result in Galois theory states that $\mathrm{Gal}(k_\mathfrak{p}/k)$ is cyclic generated by

$$\begin{aligned} \sigma_p : k_\mathfrak{p} &\to k_\mathfrak{p} \\ x &\mapsto x^p \end{aligned}$$

The union of remarks 32 and 33 is precisely the motivation behind the definition of the Frobenius automorphism.

Definition 34. *Let $L/\mathbb{Q}$ a finite Galois extension, a prime number, $\mathfrak{p}$ a prime ideal in $\mathcal{O}_L$ lying on the factorization of p, and $\phi_{\mathfrak{p}}$ the morphism of Remark 29. A Frobenius automorphism of* $\mathrm{Gal}(L/\mathbb{Q})$*, denoted by* $\mathrm{Frob}_{\mathfrak{p}}$*, corresponds to any $g \in D_{\mathfrak{p}}$ such that $\phi_{\mathfrak{p}}(g) = \sigma_p$, where σ_p is the generator of Gal$(k_{\mathfrak{p}}/k)$ pointed in Remark 33.*

Remark 35. Notice that Proposition 31 implies that when p is unramified there is a unique Frobenius automorphism for any prime ideal $\mathfrak{p}$ in $\mathcal{O}_L$ lying on the factorization of p.

Remark 35 answers the question of whether the Frobenius automorphism is unique given a prime ideal $\mathfrak{p}$. However, another natural question to ask is what happens if instead of $\mathfrak{p}$ we choose another prime ideal $\mathfrak{p}'$ lying on the factorization of p. This is what the following proposition is about.

Proposition 36. *Let $L/\mathbb{Q}$ a finite Galois extension, p an unramified prime, and $\mathfrak{p}$ and $\mathfrak{p}'$ two prime ideals lying on the factorization of p. Then there exists $\sigma \in \mathrm{Gal}(L/\mathbb{Q})$ such that*

$$\mathrm{Frob}_{\mathfrak{p}'} = \sigma^{-1}\mathrm{Frob}_{\mathfrak{p}}\sigma.$$

Proof. See [Die10], p. 13. □

Until now, the Frobenius automorphism seems to be far away from our field of interest. However, it turns out to be important while studying the images of Galois representations derived from modular forms due to the work of Eichler-Shimura. See [Shi71]

2.2 Infinite Galois theory

All Galois theory we have seen so far simply consider finite field extensions. Nevertheless, for our study we need to go a step further and generalize these constructions to also allow infinite extensions. In particular, we are interested in studying Galois representations.

2.2.1 Infinite field extensions

We start this section by recalling an important result in finite field extensions taught at the undergraduate level, the Galois correspondence.

Theorem 37 (Galois correspondence for finite field extensions)**.** *Let L/K be a finite Galois extension. Then there is a bijection between the following sets:*

$$\begin{aligned}\{\text{Subextensions } K \subset N \subset L\} &\overset{1:1}{\longleftrightarrow} \{\text{Subgroups } H \subset \mathrm{Gal}(L/K)\}\\ N &\mapsto \mathrm{Gal}(L/N)\\ L^H &\hookleftarrow H\end{aligned}$$

Theorem 37 is the core of the Galois theory for finite extensions. However, the extension in which we are interested is $\overline{\mathbb{Q}}/\mathbb{Q}$, which is infinite. Therefore, a natural question to ask is whether this correspondence is still valid for infinite extensions, or at least for the particular case of the extension $\overline{\mathbb{Q}}/\mathbb{Q}$.

It turns out that the answer is no, i.e. the correspondence as stated in Theorem 37 is no longer true for the extension $\overline{\mathbb{Q}}/\mathbb{Q}$. Nevertheless, we can sort this out by endowing infinite Galois groups with a topological structure. Broadly speaking, the main idea is to view the Galois group of infinite extensions as a topological groups and from there derive most of the properties that hold on the finite case.

In this regard, we first need to properly define the notion of topological group.

Definition 38. *Let G be a topological space and $\cdot : G \times G \to G$ a binary operation, denoted by product, such that $(G, \cdot)$ is a group. Then, $(G, \cdot)$ is a topological group if the following properties hold:*

1. *The product is continuous.*
2. *The inverse function*

$$\begin{aligned} \iota : G &\to G \\ g &\mapsto g^{-1} \end{aligned}$$

is continuous.

An important property of topological groups is that all open subgroups are also closed.

Proposition 39. *Let G be a topological group and $H \subseteq G$ an open subgroup. Then H is a closed subgroup.*

Proof. Let $g \in G$ and consider the translation map

$$\begin{aligned} m_g : G &\to G \\ h &\mapsto gh \end{aligned}$$

It is straightforward to check that m_g is a homeomorphism. Thus, all cosets $gH = m_g(H)$ are also open. Hence, we get that

$$G = H \cup \bigcup_{\substack{g \in G/H \\ g \notin H}} gH,$$

which implies that H is closed as desired. □

After defining the concept of topological groups, we need to state the particular topology with which Galois groups are endowed. It is known as the Krull topology and it comes from the following construction.

First, notice that given a field $L \subset \overline{\mathbb{Q}}$ such that L/K is a finite Galois extension we can consider the restriction map

$$\begin{aligned} \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) &\to \mathrm{Gal}(L/\mathbb{Q}) \\ \sigma &\mapsto \sigma|_L \end{aligned}$$

Therefore, by taking an infinite product over all these finite Galois extension we obtain the following map:

$$\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \hookrightarrow \prod_{\substack{L/\mathbb{Q} \\ \text{finite Galois} \\ \text{extension}}} \mathrm{Gal}(L/\mathbb{Q}).$$

From there, we define as the Krull topology on $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ to the topology given by considering the discrete topology on each $\mathrm{Gal}(L/\mathbb{Q})$, the product topology on the product, and the subspace topology on $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$.

This construction for the case of the algebraic closure of $\mathbb{Q}$ can be generalized to other field extensions by the following definition of the Krull topology.

Definition 40. *Let L/K be a Galois extension, consider the family of finite Galois extensions given by*

$$\mathcal{N} = \{N : N \text{ is a subfield of } L \text{ sucht that } N/K \text{ is a finite Galois extension}\},$$

and consider the family of subgroups given by

$$\mathcal{H} = \{\mathrm{Gal}(L/N) : N \in \mathcal{N}\}.$$

Then a subset $X \subseteq \mathrm{Gal}(L/K)$ is open in the Krull topology if it is empty or

$$X = \bigcup_i g_i N_i \text{ for some } g_i \in \mathrm{Gal}(L/K) \text{ and } N_i \in \mathcal{N}.$$

Remark 41. There are some equivalent definitions to Definition 40 for the notion in Krull topology. For example, one in terms of the open neighborhood of the identity. See [Die10] for further discussion on this topic.

Once we have endowed Galois groups with the Krull topology, we are able to state the Galois correspondence for infinite Galois extensions.

Theorem 42. *Let L/K be a Galois extension. Then there is a bijection between the following sets:*

$$\begin{aligned} \{\text{Subextensions } K \subset N \subset L\} &\overset{1:1}{\longleftrightarrow} \{\text{Closed subgroups } H \subset \mathrm{Gal}(L/K)\} \\ N &\mapsto \mathrm{Gal}(L/N) \\ L^H &\hookleftarrow H \end{aligned}$$

Proof. See [Mil22], Theorem 7.13. □

Having said that, our main interest along this subsection is to mimic the constructions done in Subsection 2.1.3 to the infinite extension $\overline{\mathbb{Q}}/\mathbb{Q}$. We start by defining its ring of integers.

Definition 43. *The ring of integers of $\overline{\mathbb{Q}}/\mathbb{Q}$, denoted by $\overline{\mathbb{Z}}$, corresponds to the ring formed by the algebraic integers of $\overline{\mathbb{Q}}$, i.e.*

$$\overline{\mathbb{Z}} = \{\alpha \in \overline{\mathbb{Q}} : \ \alpha \textit{ is an algebraic integer}\}.$$

Remark 44. Notice that by construction, for any p prime and $\mathfrak{p} \subset \overline{\mathbb{Z}}$ a maximal ideal containing p it holds that $\overline{\mathbb{Z}}/\mathfrak{p} \cong \overline{\mathbb{F}}_p$.

As expected, the definition of the decomposition group is analogue to the one stated in Definition 28 for finite extensions.

Definition 45. *Let p be a prime and $\mathfrak{p} \subset \overline{\mathbb{Z}}$ a maximal ideal containing p. The decomposition group of $\mathfrak{p}$, denoted by $D_\mathfrak{p}$, corresponds to the subgroup of G formed by the elements that fix $\mathfrak{p}$, i.e.*

$$D_\mathfrak{p} := \{\sigma \in \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) : \sigma(\mathfrak{p}) = \mathfrak{p}\}.$$

Remark 46. Despite that the decomposition group depends on the election of $\mathfrak{p}$ over p, for any other prime ideal $\mathfrak{p}'$ lying over p the decomposition groups are conjugated, i.e. we have that

$$D_{\mathfrak{p}'} = \sigma^{-1} D_\mathfrak{p} \sigma$$

for some $\sigma \in \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$.

In this case, combining the idea of Remark 29 with Remark 44 we obtain a morphism

$$\varphi_\mathfrak{p} : D_\mathfrak{p} \to \mathrm{Gal}(\overline{\mathbb{F}}_p/\mathbb{F}_p).$$

Remark 47. Considering the Krull topology, one can check that φ is indeed a surjective and continuous morphism.

As happened with the decomposition group, the construction of the inertia group follows the same structure to the one from Definition 30 for finite extensions.

Definition 48. *The inertia group of $\mathfrak{p}$, denoted by $I_\mathfrak{p}$, corresponds to the kernel of $\varphi_\mathfrak{p}$, i.e.*

$$I_\mathfrak{p} := \{\sigma \in \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) : \sigma(x) \equiv x \pmod{\mathfrak{p}} \text{ for all } x \in \overline{\mathbb{Z}}\}.$$

Remark 49. As stated in Remark46 for the decomposition group, given two prime ideals $\mathfrak{p}$ and $\mathfrak{p}'$ lying over p the inertia groups are conjugated, i.e. we have that

$$I_{\mathfrak{p}'} = \sigma^{-1} I_\mathfrak{p} \sigma$$

for some $\sigma \in \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$.

At this point, it will not be surprising that the definition of the Frobenius automorphism also sticks to the one presented in Definition 34 for finite extensions. However, in this case we call it the absolute Frobenius over p instead of the Frobenius automorphism in order to underline its importance.

Definition 50. *An absolute Frobenius over p, denoted by $\mathrm{Frob}_\mathfrak{p}$, corresponds to any $g \in D_\mathfrak{p}$ such that $\varphi_\mathfrak{p}(g) = \sigma_p$, where σ_p is given by*

$$\begin{aligned} \sigma_p : \overline{\mathbb{F}}_p &\to \overline{\mathbb{F}}_p \\ x &\mapsto x^p \end{aligned}$$

Remark 51. By definition, the absolute Frobenius is defined only up to the inertia group. So in this set up is also important to keep track of the structure of the inertia group given a prime p. Indeed, this idea will com up later in Subsection 2.2.2 in the definition of the ramification of a Galois extension.

At first sight, it may seem that all the constructions done so far are analogue to the ones done in Subsection 2.1.3 for finite extensions. However, we must take into account that now the Krull topology also plays a role and can be used to prove new results. This is precisely the case for the following theorem due to Chebotarev.

Theorem 52 (Chebotarev density theorem)**.** *For any but a finite set of primes p, and for each maximal ideal $\mathfrak{p} \subset \overline{\mathbb{Z}}$ lying over p, we take an absolute Frobenius* $\mathrm{Frob}_{\mathfrak{p}}$*. Then the set* $\{\mathrm{Frob}_{\mathfrak{p}}\}_{\mathfrak{p}|p}$ *is dense on* $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$*.*

Proof. See [LS94], Appendix. □

Theorem 52 gives us some intuition about the significance of the absolute Frobenius elements. The fact that they form a dense subset will be useful later for our purposes on studying the image of Galois representations.

2.2.2 Galois representations

As mentioned in Chapter 1, there is one Galois group that stands out above all the others: the absolute Galois group. For this reason, when speaking about Galois representations we always restrict ourselves to simply consider representations of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$. Going into details, we move to state its definition.

Definition 53. *Let K be a field and V a K-vector space of finite dimension. A Galois representation of* $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ *is a continuous morphism*

$$\rho : \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{Aut}_K(V).$$

Remark 54. Notice that contrary to the statement from Definition 7, in Definition 53 we require the morphism to be continuous because we are working with $Gal(\overline{\mathbb{Q}}/\mathbb{Q})$ as a topological group.

Remark 55. Despite dealing an infinite Galois extension, the argument used in Remark 8 is still valid. Thus, a galois representation also can be viewed as

$$\rho : Gal(\overline{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{GL}(n, K).$$

Example 56. Let $K = \mathbb{Q}(i)$, and V a K-vector space of dimension 1. By definition, for any $\sigma \in \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ we have that $\sigma(i) = \pm i$ and $\mathrm{Aut}_K(V) \cong K^\times$. Therefore, we can define a representation

$$\begin{aligned} \rho : \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) &\to K^\times \\ \sigma &\mapsto \frac{\sigma(i)}{i} \in \{\pm 1\} \end{aligned}$$

Before presenting some new concepts, we are already in a position to see one of the consequences of Theorem 52 and the importance of the absolute Frobenius.

Proposition 57. *Let ρ_1 and ρ_2 two Galois representations. If for all but finitely many primes p, and for each prime ideal $\mathfrak{p} \subset \overline{\mathbb{Z}}$ containing p, we have that $\rho_1(\mathrm{Frob}_\mathfrak{p}) = \rho_2(\mathrm{Frob}_\mathfrak{p})$, then $\rho_1 = \rho_2$.*

Proof. By Theorem 52, we have that the set $\{Frob_\mathfrak{p}\}$ is dense on $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$. Then, using that both representations are continuous by definition, we conclude that if they are equal on a dense subset they must be equal on the whole set. □

When studying Galois representations, it may be helpful to be able to split them into smaller representations. The representations that allow this phenomena are called semisimple.

Definition 58. *Let $\rho : Gal(\overline{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{Aut}_K(V)$ be a Galois representation. ρ is defined as semisimple if V can be decomposed as*

$$V = \bigoplus_i V_i$$

where each V_i is a subrepresentation of V and the representations $\rho : Gal(\overline{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{Aut}_K(V_i)$ are irreducible for all i.

Therefore, it makes sense to try to characterize the conditions for a Galois representation is semisimple, since it will simplify our task. In that sense, the following theorem states that all Galois representations over the complex numbers are semisimple.

Theorem 59. *Let $\rho : Gal(\overline{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{Aut}_\mathbb{C}(V)$ be a Galois representation. Then ρ is semisimple.*

Proof. The idea of the proof is to reduce the statement of the theorem to a representation from a finite Galois extension via a factorization and then treat it as a particular case of Theorem 12. See [Die10], Theorem 6.8. for further details. □

Remark 60. In general it is not true that any Galois representation is semisimple, since in the previous proof the argument precisely uses the hypotesis of $K = \mathbb{C}$.

Due to Remark 60, we could be tempted to only focus our study to Galois representations with over the complex numbers because they are all semisimple. However, it turns out that complex Galois representations always factor through finite extensions, i.e. by studying them we can only recover information from finite extensions of $\mathbb{Q}$.[1] For this reason, we must also consider the Galois representations on vector spaces over $\mathbb{Q}_l$, called the l-adic Galois representations.

In Subsection 2.2.1 we have generalized the notion of ramification to the absolute Galois group. Thus, another topic that requires our attention is to relate this construction with Galois representations.

Definition 61. *Let $\rho : Gal(\overline{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{Aut}_K(V)$ be a Galois representation. Given a prime p, we say that ρ is unramified at p if $\rho(I_\mathfrak{p}) = \mathrm{Id}$ for any prime ideal $\mathfrak{p} \subset \overline{\mathbb{Z}}$ containing p.*

[1] See [Die10], Theorem 6.5. for further details on this discussion.

Remark 62. It must be checked that Definition 61 does not depend on on the prime ideal $\mathfrak{p}$. In order to do so, we rely on the fact that Remark 49 states that two different inertia groups of the same prime p are conjugated. Therefore, if ρ is trivial in one of them, it must be trivial in the other as well.

At first sight, it may not be clear where this definition of ramification comes from. Recall that Proposition 31 relates the ramification of field extensions with the inertia group not being trivial. With this in mind, the following proposition establishes the link between ramification of a representation and of a field extension.

Proposition 63. *Let p be a prime and $\rho : Gal(\overline{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{Aut}_K(V)$ be a Galois representation. Then ρ is unramified at p if and only if it factors through $\mathrm{Gal}(L/\mathbb{Q})$, where $L/\mathbb{Q}$ is some algebraic extension unramified at p.*

Proof. First, we prove the left to right implication. If ρ factors through some representation $\tilde{\rho} : \mathrm{Gal}(L/\mathbb{Q}) \to \mathrm{Aut}_K(V)$ with $L/\mathbb{Q}$ unramified at p, it is clear that

$$\rho(I_{\mathfrak{p}}) = \tilde{\rho} \circ \pi(I_{\mathfrak{p}}) = \tilde{\rho}(I_{\mathfrak{p}}(L/\mathbb{Q})) = \tilde{\rho}(1) = \mathrm{Id}.$$

On the other hand, if $I_{\mathfrak{p}}$ is trivial the statement follows directly. Moreover, if $I_{\mathfrak{p}}$ is not trivial then ρ factors through $\mathrm{Gal}(L/\mathbb{Q})$, where $L \subset \overline{\mathbb{Q}}$ is the maximal extension in which p is ramified. □

Hence, it makes sense to define a representation as unramified precisely when the image of inertia group is the trivial one. Moreover, as stated in Remark 51, the absolute Frobenius is defined up to the inertia group. Therefore, the fact that ρ is unramified implies that $\rho(I_{\mathfrak{p}}) = \mathrm{Id}$, and consequently, $\rho(\mathrm{Frob}_{\mathfrak{p}})$ only depends on $\mathfrak{p}$.

Moreover, we can go a step further and try to find a relation between $\rho(\mathrm{Frob}_{\mathfrak{p}})$ and $\rho(\mathrm{Frob}_{\mathfrak{p}'})$ for two different prime ideals $\mathfrak{p}$ and $\mathfrak{p}'$ over an unramified prime p. In that sense, notice that since $\mathfrak{p}$ and $\mathfrak{p}'$ are conjugated, $\rho(\mathrm{Frob}_{\mathfrak{p}})$ and $\rho(\mathrm{Frob}_{\mathfrak{p}'})$ are conjugated matrices, which implies that their characteristic polynomials are the same. This leads to the following definition.

Definition 64. *Let ρ be a Galois representation unramified at p. We define the characteristic polynomial at p as*

$$\mathrm{char}(\rho(\mathrm{Frob}_p)) \coloneqq \mathrm{char}(\rho(\mathrm{Frob}_{\mathfrak{p}})).$$

Due to the previous construction, we have obtained a value on the image of a Galois representation that simply depends on the unramified prime p. Moreover, since for any $A \in \mathrm{GL}(2, K)$ we have that $\mathrm{char}(A) = x^2 - \mathrm{trace}(A) + \det(A)$, for any 2-dimensional Galois representation unramified at p we can define

$$\mathrm{trace}(\rho(\mathrm{Frob}_p)) \coloneqq \mathrm{trace}(\rho(\mathrm{Frob}_{\mathfrak{p}})) \text{ and } \det(\rho(\mathrm{Frob}_p)) \coloneqq \det(\rho(\mathrm{Frob}_{\mathfrak{p}})).$$

In addition, the following theorem due to Brauer and Nesbitt shows how the characteristic polynomial is a powerful tool to understand the image of a Galois representation.

Theorem 65. *Let $\rho_i : Gal(\overline{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{Aut}_K(V)$, $i = 1, 2$, two semisimple Galois representations ramified only in a finite set of primes S. If*

$$\mathrm{char}(\rho_1(\mathrm{Frob}_p)) = \mathrm{char}(\rho_2(\mathrm{Frob}_p))$$

for all $p \notin S$, then ρ_1 and ρ_2 are isomorphic. Moreover, if K has characteristic 0 it is enough to check that

$$\mathrm{trace}(\rho_1(\mathrm{Frob}_p)) = \mathrm{trace}(\rho_2(\mathrm{Frob}_p))$$

for all $p \notin S$.

Proof. See [WCR06], Theorem 30.16. □

Notice that the hypothesis of Theorem 65 requires that both Galois representations are unramified at all but finitely many primes p. Therefore, a natural question to ask is when a Galois representation satisfies this requirement. In that line, there is a strong result for the case of $K = \mathbb{C}$.

Theorem 66. *Let $\rho : Gal(\overline{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{Aut}_{\mathbb{C}}(V)$ be a Galois representation. Then ρ is ramified in a finite set of primes.*

Proof. See [Die10], Proposition 6.15. □

Unfortunately, this result does not hold in general. However, the set of ramified primes always has 0 density. See [KR01] for further details.

2.3 Character theory

Character theory is the area of mathematics that deals with the study of certain functions on groups known as characters. One of the key features of character theory is that it provides a powerful tool for studying groups and their properties, without requiring a detailed understanding of the underlying structure of the group itself. Instead, characters allow us to study groups in terms of their algebraic behavior, such as their multiplication and inversion properties, which can often reveal important insights about the group and its properties.

In that sense, character theory is closely related to representation theory and, in particular, with Galois representations. Therefore, in this section we introduce the notions of character theory needed later for our purposes.

2.3.1 Group characters

We start by defining the notion of multiplicative group character.

Definition 67. *Let G be a group and K a field. A multiplicative group character on G is a group homomorphism from G to the multiplicative group of a field, i.e.*

$$\chi : G \to K^{\times}$$

Remark 68. The set of multiplicative characters on G is denoted by $\mathrm{Ch}(G)$ and is an abelian group under pointwise multiplication.

It turns out that in group theory there is a different but related notion in terms of group representations that it is also referred as group character. It corresponds to the following definition.

Definition 69. *Let G be a group, V a finite vector space over a field K, and ρ a representation of G on V. The character of ρ is defined as*

$$\begin{aligned} \chi_\rho : G &\to K \\ g &\mapsto \mathrm{Trace}(\rho(g)) \end{aligned}$$

Remark 70. Notice that the character of a group representation is not necessarily a group homomorphism since the trace may not be one.

Once both notions are defined, it is natural to ask how they relate to each other. In that sense, notice that for 1-dimensional representations the trace coincides with the representation itself. Thus, the concept of multiplicative character of a group can be seen as a kind of particular case of higher-dimensional representation characters.

For our purposes, when dealing with characters we will simply consider 1-dimensional group representations, i.e. the case where both definitions coincide.

2.3.2 Dirichlet characters

Definition 71. *Let N be a positive integer. A Dirichlet character modulo N is a multiplicative character from $(\mathbb{Z}/N\mathbb{Z})^\times$ to $\mathbb{C}^\times$, i.e.*

$$\chi : (\mathbb{Z}/N\mathbb{Z})^\times \to \mathbb{C}^\times$$

Despite being a simple construction, Dirichlet characters show up often in the context of Galois representation due to the existing isomorphism between the Galois group of cyclotomic extensions and $(\mathbb{Z}/N\mathbb{Z})^\times$.

Moreover, for any d positive divisor of N, the existence of the natural projection

$$\pi_{N,\,d} : (\mathbb{Z}/N\mathbb{Z})^\times \to (\mathbb{Z}/d\mathbb{Z})^\times$$

allows to lift every Dirichlet character modulo d χ_d to a Dirichlet character modulo N χ_N via the composition

$$\chi_N = \chi_d \circ \pi_{N,\,d}.$$

From this construction, we define what is known as the conductor of a Dirichlet character.

Definition 72. *Let χ be a Dirichlet character modulo N. The conductor of χ is the smallest positive divisor d of N such that $\chi = \chi_d \circ \pi_{N,\,d}$.*

In a similar way, each Dirichlet character can be extended to a function $\chi : \mathbb{Z}/N\mathbb{Z} \to \mathbb{C}$ by setting $\chi(n) = 0$ for all noninvertible elements n in $\mathbb{Z}/N\mathbb{Z}$. Furthermore, this can be also extended to a function $\chi : \mathbb{Z} \to \mathbb{C}$ simply by defining $\chi(n) = \chi(n \pmod N)$ for all $n \in \mathbb{Z}$. However, this extension is not a group homomorphism.

Remark 73. Beyond their applications in representation theory, which we will see in Section 5.3, Dirichlet characters also appear in the context of modular forms. In particular, they are closely linked to the notion of L-functions. See [Dia05], Chapter 4.3. for a further discussion on this topic.

2.3.3 Cyclotomic characters

Another type of characters that are interesting on their own is the set of cyclotomic characters. Broadly speaking, they consist in characters of a given Galois group based on its action on a multiplicative group of roots of unity.

Going into details, let l be a prime and ζ_{l^n} be a primitive l^n-th root of unity. By the Galois theory of cyclotomic extensions, for any $\sigma \in \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ we have that

$$\sigma(\zeta_{l^n}) = \zeta_{l^n}^{k_{l^n}(\sigma)} \text{ with } k_{l^n}(\sigma) \in (\mathbb{Z}/l^n\mathbb{Z})^\times.$$

From this construction arises a character known as the mod l^n cyclotomic character.

Definition 74. *Let l be a prime and ζ_{l^n} be a primitive l^n-th root of unity. The mod l^n cyclotomic character is given by*

$$\begin{aligned} \chi_{l,\,n} : Gal(\overline{\mathbb{Q}}/\mathbb{Q}) &\to (\mathbb{Z}/l^n\mathbb{Z})^\times \\ \sigma &\mapsto k_{l^n}(\sigma) \end{aligned}$$

Remark 75. As stated in Subsection 2.3.1, the mod l^n cyclotomic character can be viewed both as a character and as a Galois representation.

Remark 76. From the construction of the mod l^n cyclotomic character it is clear that it is surjective.

Starting from Definition 74, if we fix l and σ and we vary the value of n what we get is that the images $k_{l^n}(\sigma)$ for a compatible system. Thus, the combination of this fact with the construction of the l-adic integers as an inverse limit leads us to the definition of the l-adic cyclotomic character.

Definition 77. *Let l be a prime. The l-adic cyclotomic character is given by*

$$\begin{aligned} \chi_l : Gal(\overline{\mathbb{Q}}/\mathbb{Q}) &\to Z_l^\times \\ \sigma &\mapsto \varprojlim k_{l^n}(\sigma) \end{aligned}$$

Remark 78. Combining the construction of the l-adic cyclotomic character as the inverse limit of the mod l^n cyclotomic characters together with Remark 76 we conclude that the l-adic cyclotomic character is surjective.

2.3.4 Fundamental characters

The last set of characters that we need to introduce are the known as fundamental characters, which are defined by means of the l-adic cyclotomic character and the inertia group. In general terms, the idea behind these characters is to define maps from an inertia group I_l to the finite fields $\mathbb{F}_{l^n}$ for all $n \geq 1$.

We start by defining the fundamental character of level 1.

Definition 79. *Let l be a prime, χ_l the l-adic cyclotomic character, and $\overline{\chi}_l$ its reduction modulo l. The fundamental character of level 1 Ψ_1 is the restriction of l-adic cyclotomic character χ_l to the inertia group I_l composed with reduction modulo l, i.e.*

$$\Psi_1 = \overline{\chi}_l|_{I_l}.$$

Remark 80. From this definition, one can easily check that the order of Ψ_1 is $l-1$.

Definition 79 can be generalized to higher levels by the following construction.

Let l be a prime, $K_n/\mathbb{Q}_l$ be the unique unramified field extension of degree n, whose uniqueness is given by local field theory, and consider $\mathbb{F}_{l^n}$ the residue field of K_n. Then, by Hensel's lemma we can assure that K_n contains all (l^n-1)-th roots of unity, so defining $K_n' = K_n((-l)^{\frac{1}{l^n-1}})$ we get that the extension K_n'/K_n is Galois and $\mathrm{Gal}(K_n'/K_n) = \mathbb{F}_{l^n}^{\times}$. Moreover, since $K_n/\mathbb{Q}_l$ is an unramified extension, we can inject the inertia group I_l into $\mathrm{Gal}(\overline{\mathbb{Q}}_l/K_n)$. Then, the fundamental character of level n simply corresponds to the composition of this injection with the restriction of $\mathrm{Gal}(\overline{\mathbb{Q}}_l/K_n)$ to $\mathrm{Gal}(K_n'/K_n) = \mathbb{F}_{l^n}^{\times}$.

Definition 81. *Let l be a prime. The fundamental character of level $n \geq 1$ corresponds to the map $\Psi_n : I_l \to \mathbb{F}_{l^n}^{\times}$ described above.*

Remark 82. By looking at the construction of the fundamental character in detail, it can be observed that by taking the $l+1$-th power of fundamental character of level 2 we obtain the fundamental character of level 1, i.e. $\Psi_2^{l+1} = \Psi_1$. More details about this results as well as other properties of the fundamental character can be found in [Ser72].

Chapter 3

Modular forms

The first appearance of modular forms in mathematics dates back to the beginning of the 19th century in connection with elliptic functions, which were present in the works of Gauss, Abel and Jacobi. Their work on elliptic functions lead Eisenstein to study modular forms arising from the expansions of these functions. However, it was not until the end of the century that the term modular form appeared in Klein's work. Since then, lots of mathematicians, such as Zagier and Langland, contributed to their study. In that sense, in the first half of the 20th century Hecke introduced the modern approach to study modular forms based on $\mathrm{SL}(2,\mathbb{Z})$ and its congruence subgroups.

Nowadays, modular forms are considered a powerful tool from complex analysis with lots of applications not only in number theory but also in other fields such as combinatorics and string theory.

3.1 First notions

Throughout this section, a series of definitions and basic properties of modular forms are presented in order to be used in next sections. We start by defining two fundamental concepts in the theory of modular forms: the modular group and the complex upper half plane.

3.1.1 The modular group

Definition 83. *The modular group is the group of 2-by-2 matrices with integer entries and determinant equal to* 1*. It is denoted by*

$$\mathrm{SL}(2,\mathbb{Z}) := \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} : \ a,b,c,d \in \mathbb{Z},\ ad-bc=1 \right\}.$$

Proposition 84. *The modular group is generated by the matrices*

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

Proof. See [Ser73], Theorem I.7.2. □

Definition 85. *The complex upper half plane is given by*

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

The relation between both concepts is given by the following action of $\mathrm{SL}(2,\mathbb{Z})$ on $\mathcal{H}$.

Proposition 86. $\mathrm{SL}(2,\mathbb{Z})$ *acts on* $\mathcal{H}$ *by the fractional linear transformation*

$$\begin{aligned} \mathrm{SL}(2,\mathbb{Z}) \times \mathcal{H} &\to \mathcal{H} \\ (\gamma, \tau) &\mapsto \gamma\tau := \frac{a\tau + b}{c\tau + d} \end{aligned}$$

Proof. It is well-defined since $c\tau + d \neq 0$ for all $c, d \in \mathbb{Z}, \tau \in \mathcal{H}$ and $\mathrm{Im}(\gamma\tau) = \frac{\mathrm{Im}(\tau)}{|c\tau+d|^2} > 0$. It is also straightforward to check that it is indeed a group action because $\mathrm{Id}\tau = \tau$ and $\gamma_1(\gamma_2\tau) = (\gamma_1\gamma_2)\tau$. □

3.1.2 Modular & Cusp forms

A natural question to ask is which complex functions are invariant under the fractional linear transformation. However, this requirement may seem to strict to be satisfied by any interesting set of functions, so it makes sense to replace the invariance requirement for another slightly less demanding. It is precisely from the relaxation in this condition that modular forms arise.

Definition 87. *Let* $k \in \mathbb{Z}$. *A function* $f : \mathcal{H} \to \mathbb{C}$ *is a modular form of weight* k *if*

1. f *is holomorphic on* $\mathcal{H}$.
2. f *is holomorphic at* ∞, *i.e.* f *is bounded as* $\mathrm{Im}(\tau) \to \infty$.
3. $f(\gamma\tau) = (c\tau + d)^k f(\tau)$ *for all* $\gamma \in \mathrm{SL}(2,\mathbb{Z})$, $\tau \in \mathcal{H}$.

$\mathcal{M}_k(\mathrm{SL}(2,\mathbb{Z}))$ *denotes the set of modular forms of weight* k *and* $\mathcal{M}_*(SL(2,\mathbb{Z})) := \bigoplus_{k\in\mathbb{Z}} \mathcal{M}_k(\mathrm{SL}(2,\mathbb{Z}))$.

Remark 88. In literature, Condition 3 of Definition 87 is usually referred as the weakly modularity condition.

Remark 89. It is straightforward to check that $\mathcal{M}_k(\mathrm{SL}(2,\mathbb{Z}))$ is a $\mathbb{C}$-vector space and $\mathcal{M}_*(\mathrm{SL}(2,\mathbb{Z}))$ is a graded-ring.

Example 90. The most common example corresponds to the Eisenstein series. Given an integer $k > 2$, $G_k(\tau) = \sum_{(0,0)\neq(m,n)\in\mathbb{Z}} \frac{1}{(m+n\tau)^k}$ is a modular form of weight k.

Among all modular forms, there is particular subset of them which play an important role for our purposes. They are called cusp forms.

Definition 91. *A modular form f of weight k such that $\lim_{\mathrm{Im}(\tau)\to\infty} f(\tau) = 0$ is called a cusp form. $\mathcal{S}_k(\mathrm{SL}(2,\mathbb{Z}))$ denotes the set of cusp forms of weight k and $\mathcal{S}_*(SL(2,\mathbb{Z})) \coloneqq \bigoplus_{k\in\mathbb{Z}} \mathcal{S}_k(\mathrm{SL}(2,\mathbb{Z}))$.*

Remark 92. The condition $\lim_{\mathrm{Im}(\tau)\to\infty} f(\tau) = 0$ is equivalent to impose that the leading coefficient of its Fourier expansion is equal to 0.

Remark 93. Again it is straightforward to check that $\mathcal{S}_k(\mathrm{SL}(2,\mathbb{Z}))$ is a subspace of $\mathcal{M}_k(\mathrm{SL}(2,\mathbb{Z}))$ and $\mathcal{S}_*(\mathrm{SL}(2,\mathbb{Z}))$ is an ideal in $\mathcal{M}_*(\mathrm{SL}(2,\mathbb{Z}))$.

3.1.3 Congruence subgroups

Once we have defined with detail the notion of modular forms, the next step is to try to generalize it. In order to do so, the idea is to replace the modular group $\mathrm{SL}(2,\mathbb{Z})$ by some subgroups of it to obtain more functions satisfying Condition 3 from Definition 87. We start by presenting the principal congruence subgroup.

Definition 94. *Let N be a positive integer. The principal congruence subgroup of level N is given by*

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

Remark 95. The index of the principal congruence subgroup is finite for all N. See [Dia05], p. 13 for further details.

The importance of Definition 94 is that it allows to define several different congruence subgroups simply by imposing that they must include $\Gamma(N)$.

Definition 96. *Let N be a positive integer. A subgroup Γ of $\mathrm{SL}(2,\mathbb{Z})$ is a congruence subgroup of level N if $\Gamma(N) \subset \Gamma$.*

Remark 97. From Remark 95 is is immediate that all congruence subgroups have finite index.

Above all congruence subgroups there are two that stand out from the rest in terms of our goals. In particular, they correspond to

$$\Gamma_0(N) \coloneqq \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathrm{SL}(2,\mathbb{Z}) : \begin{bmatrix} a & b \\ c & d \end{bmatrix} \equiv \begin{bmatrix} * & * \\ 0 & * \end{bmatrix} \pmod{N} \right\},$$

$$\Gamma_1(N) \coloneqq \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathrm{SL}(2,\mathbb{Z}) : \begin{bmatrix} a & b \\ c & d \end{bmatrix} \equiv \begin{bmatrix} 1 & * \\ 0 & 1 \end{bmatrix} \pmod{N} \right\}.$$

Remark 98. Notice that $\Gamma(N) \subset \Gamma_1(N) \subset \Gamma_0(N) \subset \mathrm{SL}(2,\mathbb{Z})$.

Given Definition 96, what follows is precisely to define the notion of modular forms with respect congruence subgroups. However, we first need to present two additional concepts that are closely related to the desired definition: the factor of automorphy and the weight k operator.

Definition 99. *Let* $\gamma = \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right] \in \mathrm{SL}(2,\mathbb{Z})$. *The factor of automorphy of* γ *is given by*

$$\begin{aligned} j(\cdot,\cdot) : \mathrm{SL}(2,\mathbb{Z}) \times \mathcal{H} &\to \mathbb{C} \\ (\gamma,\tau) &\mapsto j(\gamma,\tau) \coloneqq c\tau + d \end{aligned}$$

Definition 100. *Let* $\gamma = \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right] \in \mathrm{SL}(2,\mathbb{Z})$. *The weight k operator* $[\gamma]_k$ *on functions* $f : \mathcal{H} \to \mathbb{C}$ *is given by*

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

Remark 101. Notice that now we can restate Condition 3 from Definition 87 as $f[\gamma]_k = f$ for all $\gamma \in \mathrm{SL}(2,\mathbb{Z})$.

With definitions 99 and 100 stated, we are ready to fully define the notion of modular forms with respect to congruence subgroups.

Definition 102. *Let* $k \in \mathbb{Z}$ *and* Γ *a congruence subgroup of* $\mathrm{SL}(2,\mathbb{Z})$. *A function* $f : \mathcal{H} \to \mathbb{C}$ *is a modular form of weight k with respect to* Γ *if*

1. f *is holomorphic on* $\mathcal{H}$.
2. $f[\gamma]_k$ *is holomorphic at* ∞, *i.e.* $f[\gamma]_k$ *is bounded as* $\mathrm{Im}(\tau) \to \infty$ *for all* $\gamma \in \mathrm{SL}(2,\mathbb{Z})$.
3. $f(\gamma\tau) = (cz+d)^k f(\tau)$ *for all* $\gamma =\in \Gamma_0(N)$, $\tau \in \mathcal{H}$.

$\mathcal{M}_k(\Gamma)$ *denotes the set of modular forms of weight k with respect to* Γ *and* $\mathcal{M}_*(\Gamma) \coloneqq \bigoplus_{k\in\mathbb{Z}} \mathcal{M}_k(\Gamma)$.

Remark 103. As in Remark 89, $\mathcal{M}_k(\Gamma)$ is a $\mathbb{C}$-vector space and $\mathcal{M}_*(\Gamma)$ is a graded-ring.

Remark 104. Combining Remark 98 and Definition 87 we get that $\mathcal{M}_k(\Gamma_0(N)) \subset \mathcal{M}_k(\Gamma_1(N))$.

At first sight, Condition 2 from Definition 102 may seem counterintuitive. Broadly speaking, the reason behind the appearance of the weight k operator is to ensure that the function is holomorphic at all cusp points. In particular, when dealing with $\mathrm{SL}(2,\mathbb{Z})$, all cusp points are $\mathrm{SL}(2,\mathbb{Z})$-equivalent to ∞; but when working with some congruence subgroup Γ fewer points are Γ-equivalent, so we must check all other cusps. A fully detailed explanation of this fact can be found in [Dia05], p. 16.

As in the case of $\mathrm{SL}(2,\mathbb{Z})$, when taking any congruence subgroup Γ the cusp forms with respect to Γ are also an important subset to consider.

Definition 105. *Let* $k \in \mathbb{Z}$ *and* Γ *a congruence subgroup of* $\mathrm{SL}(2,\mathbb{Z})$. *A modular form of weight k with respect to* Γ *whose Fourier expansion has leading coefficient* a_0 *is called a cusp form with respect to* Γ. $\mathcal{S}_k(\Gamma)$ *denotes the set of cusp forms of weight k and* $\mathcal{S}_*(\Gamma) \coloneqq \bigoplus_{k\in\mathbb{Z}} \mathcal{S}_k(\Gamma)$.

Remark 106. As in Remark 93, $\mathcal{S}_k(\Gamma)$ is a subspace of $\mathcal{M}_k(\Gamma)$ and $\mathcal{S}_*(\Gamma)$ is an ideal in $\mathcal{M}_*(\Gamma)$.

3.2 Hecke theory

This section introduces the notion of Hecke operators and discusses some of its properties and results. In particular, we will show that the space of cusp forms of weight k $\mathcal{S}_k(\Gamma)$ has an orthogonal basis of modular forms which are simultaneously eigenfunctions for almost all the Hecke operators. We start by defining the core of the construction of the Hecke operators: the double coset operators.

3.2.1 Diamond & Hecke operators

Definition 107. *Let Γ_1 and Γ_2 be congruence subgroups of $\mathrm{SL}(2,\mathbb{Z})$ and $\alpha \in \mathrm{GL}^+(2,\mathbb{Q})$. A double coset of $\mathrm{GL}^+(2,\mathbb{Q})$ is given by*

$$\Gamma_1\alpha\Gamma_2 \coloneqq \{\gamma_1\alpha\gamma_2 \mid \gamma_1 \in \Gamma_1,\ \gamma_2 \in \Gamma_2\}.$$

Remark 108. Since Γ_1 acts on the double coset $\Gamma_1\alpha\Gamma_2$ by left multiplication, we can express it as a disjoint union of orbit spaces of this group action, i.e.

$$\Gamma_1\alpha\Gamma_2 = \cup\Gamma_1\beta_j,$$

where β_j are the orbit representatives.

Our intention is to define the double coset operator as the sum of the weight k operator for each right coset representative β_j. Therefore, in order to be well-defined, we first need to ensure that the set of right coset representatives β_j is finite.

Proposition 109. *Let Γ_1 and Γ_2 be congruence subgroups of $\mathrm{SL}(2,\mathbb{Z})$, $\alpha \in \mathrm{GL}^+(2,\mathbb{Q})$ and $\Gamma_1\alpha\Gamma_2 = \cup\Gamma_1\beta_j$, where β_j are the orbit representatives. The set of right coset representatives β_j is finite.*

Proof. See [Dia05], Lemma 5.1.1 & Lemma 5.1.2. □

Thanks to Proposition 109 we are ready to define the double coset operator.

Definition 110. *Let Γ_1 and Γ_2 be congruence subgroups of $\mathrm{SL}(2,\mathbb{Z})$, $\alpha \in \mathrm{GL}^+(2,\mathbb{Q})$, $\Gamma_1\alpha\Gamma_2 = \cup\Gamma_1\beta_j$, where β_j are the orbit representatives, and $f \in \mathcal{M}_k(\Gamma_1)$. The weight k double coset operator $[\Gamma_1\alpha\Gamma_2]_k$ is given by*

$$f[\Gamma_1\alpha\Gamma_2]_k \coloneqq \sum_j f[\beta_j]_k.$$

Remark 111. Notice that given two orbit representatives β_1, β_2, we have that $\beta_1\beta_2^{-1} \in \Gamma_1$. Therefore, $f[\beta_1]_k = f[\beta_1\beta_2^{-1}]_k f[\beta_2]_k = f[\beta_2]_k$. So combining this with Proposition 109 it is clear that the double coset operator is well-defined.

Remark 112. The weight k double coset operator $[\Gamma_1\alpha\Gamma_2]_k$ sends $\mathcal{M}_k(\Gamma_1(N))$ to $\mathcal{M}_k(\Gamma_2(N))$, and $\mathcal{S}_k(\Gamma_1(N))$ to $\mathcal{S}_k(\Gamma_2(N))$.

Remark 113. When $\Gamma_1 = \Gamma_2$ we have that $f[\Gamma_1\alpha\Gamma_1]_k = f[\alpha]_k$ for all $\alpha \in \mathrm{GL}^+(2,\mathbb{Q})$.

The importance of the double coset operator lies in the fact that the definitions of the Hecke operators arise from it. However, it must be combined with another operator: the diamond operator.

Definition 114. *The diamond operator is given by*

$$\begin{aligned} \langle d \rangle : \mathcal{M}_k(\Gamma_1(N)) &\to \mathcal{M}_k(\Gamma_1(N)) \\ f &\mapsto \langle d \rangle f \coloneqq f[\alpha]_k \end{aligned}$$

for any $\alpha = \left[\begin{smallmatrix} a & b \\ c & \delta \end{smallmatrix}\right] \in \Gamma_0(N)$ *with* $\delta \equiv d \pmod{N}$.

Remark 115. At first sight it may not be clear that the diamond operator is well-defined. However, combining Remark 113 and the fact that $\Gamma_1(N)$ is a normal subgroup of $\Gamma_0(N)$ one can easily verify this. See [Dia05], p. 168 for further details.

Once we have defined the double coset and the diamond operators, we are ready to properly define the Hecke operators.

Definition 116. *Let p be a prime. The Hecke operator is given by*

$$\begin{aligned} T_p : \mathcal{M}_k(\Gamma_1(N)) &\to \mathcal{M}_k(\Gamma_1(N)) \\ f &\mapsto T_p f \coloneqq f[\Gamma_1(N) \left[\begin{smallmatrix} 1 & 0 \\ 0 & p \end{smallmatrix}\right] \Gamma_1(N)]_k \end{aligned}$$

Since the weight k double coset operator is defined in terms of the weight k operators of the right coset representatives β_j, it makes sense to try to find an explicit representation of the Hecke operators by means of the β_j. This is precisely what the following proposition is about.

Proposition 117. *Let p be a prime. The Hecke operator T_p is given by*

$$T_p = \begin{cases} \sum_{j=0}^{p-1} f[\left[\begin{smallmatrix} 1 & 0 \\ 0 & p \end{smallmatrix}\right]]_k & \text{if } p \mid N \\ \\ \sum_{j=0}^{p-1} f[\left[\begin{smallmatrix} 1 & 0 \\ 0 & p \end{smallmatrix}\right]]_k + f[\left[\begin{smallmatrix} m & n \\ N & p \end{smallmatrix}\right] \left[\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\right]]_k & \text{if } p \nmid N, \text{ where } mp - nM = 1. \end{cases}$$

Proof. See [Dia05], Proposition 5.2.1. □

Although Proposition 117 gives an explicit representation for the Hecke operators, we could still complain about the fact that they are not defined in full generality, since in the statement we restricted to the case of p prime. In the same way, until now the diamond operator is only defined for values in $(\mathbb{Z}/N\mathbb{Z})^\times$. Thus, the next step is precisely to generalize the current definitions of the diamond and the Hecke operators.

The generalization fo the diamond operator is straightforward, we simply need to set $\langle n \rangle$ to zero if $\gcd(n, N) \geq 1$.

Definition 118. *Let $n \in \mathbb{Z}$. The diamond operator is given by*

$$\langle n \rangle := \begin{cases} f[\alpha]_k \text{ for any } \alpha = \left[\begin{smallmatrix} a & b \\ c & n \end{smallmatrix}\right] \in \Gamma_0(N) & \text{if } \gcd(n, N) = 1 \\ \\ 0 & \text{if } \gcd(n, N) \geq 1. \end{cases}$$

In the case of the Heche operators, we generalize them by combining Definition 116 and Definition 118 with the fundamental theorem of arithmetic.

Definition 119. *Let $n \in \mathbb{Z}$. The Hecke operator is given by*

$$T_n := \begin{cases} f[\Gamma_1(N)\left[\begin{smallmatrix} 1 & 0 \\ 0 & n \end{smallmatrix}\right]\Gamma_1(N)]_k & \text{if } n \text{ prime} \\ \\ T_p T_{p^{r-1}} - p^{k-1}\langle p \rangle T_{p^{r-2}} & \text{if } n = p^r \text{ with } p \text{ prime} \\ \\ \prod T_{p^{e_i}} & \text{if } n = \prod p^{e_i}. \end{cases}$$

3.2.2 The Petersson inner product

As briefly mentioned earlier, the key aspect of the Hecke operators is that they characterize an orthogonal basis of the space of cusp forms of weight k. However, we first need to check that $\mathcal{S}_k(\Gamma_1(N))$ is a inner product space. To do so, we must develop the known as Petersson inner product.

We start by defining the measure we consider on the upper half plane.

Definition 120. *Let $\tau = x + iy \in \mathcal{H}$. The hyperbolic measure on the upper half plane is given by*

$$d\mu(\tau) = \frac{dxdy}{y^2}.$$

Remark 121. The hyperbolic measure is $\mathrm{SL}(2, \mathbb{Z})$-invariant, i.e. $d\mu(\gamma\tau) = d\mu(\tau)$ for any $\gamma \in \mathrm{SL}(2, \mathbb{Z}),\ \tau \in \mathcal{H}$.

The Petersson inner product is defined by means of the hyperbolic measure and integrating as follows.

Definition 122. *Let Γ be a congruence subgroup, $X(\Gamma)$ its associated modular curve, and $V_\Gamma = \int_{X(\Gamma)} d\mu(\tau)$ the volume of $X(\Gamma)$. The Petersson inner product is given by*

$$\begin{aligned} \langle\ ,\ \rangle_\Gamma : \mathcal{S}_k(\Gamma) \times \mathcal{S}_k(\Gamma) &\to \mathcal{S}_k(\Gamma) \\ (f, g) &\mapsto \frac{1}{V_\Gamma} \int_{X(\Gamma)} f(\tau)\overline{g(\tau)}(\mathrm{Im}(\tau))^k d\mu(\tau) \end{aligned}$$

Remark 123. The integral from Definition 122 is well-defined, convergent, and it satisfies all the requirements for being an inner product on $\mathcal{S}_k(\Gamma)$, i.e. it is linear in f, conjugate linear in g. Hermitian-symmetric, and positive definite. See [Dia05], pp. 181-183.

The following theorem links the notions of the Hecke operators and the Petersson inner product and is the key point for giving an orthogonal basis of the space of the cusp forms of weight k.

Theorem 124. *Let p be an integer such that $p \nmid N$, and consider the inner product space $\mathcal{S}_k(\Gamma_1(N))$. Then the Hecke operators $\langle p \rangle$ and T_p have adjoints*

$$\langle p \rangle^* = \langle p \rangle^{-1} \text{ and } T_p^* = \langle p \rangle^{-1} T_p.$$

Therefore, for any $n \in \mathbb{Z}$ such that $\gcd(n, N) = 1$ the Hecke operators $\langle n \rangle$ and T_n are normal.

Proof. See [Dia05], Theorem 5.5.3. □

Now, the idea is to combine all the development done so far with some linear algebra. In particular, recall that the Spectral Theorem states that given a family of normal operators from a finite-dimensional inner product space, there always exists an orthogonal basis of simultaneous eigenvectors for these operators. Therefore, combining this with Theorem 124 we obtain the following result.

Theorem 125. *There is a set of modular forms that form an orthogonal basis of $\mathcal{S}_k(\Gamma_1(N)$ and are simultaneously eigenvectors for the Hecke operators $\{\langle n \rangle,\ T_n : \gcd(n, N) = 1\}$. These modular forms are called eigenforms.*

Proof. See [Dia05], Theorem 5.5.4. □

3.2.3 Oldforms & Newforms

Theorem 125 gives us an orthogonal basis of $\mathcal{S}_k(\Gamma_1(N)$ formed by eigenforms. However, these modular forms are not eigenvectors for all Hecke operators, but only for the ones coprime with N. So a natural question to ask is whether there is a way to get rid of this restriction. To anwser this question we need to develop the theory of the oldforms and newforms.

Definition 126. *Let $\alpha_d = \left[\begin{smallmatrix} 1 & 0 \\ 0 & d \end{smallmatrix}\right] \in \mathrm{GL}^+(2, \mathbb{Q})$, d a divisor of N, and*

$$\begin{aligned} i_d : \mathcal{S}_k(\Gamma_1(Nd^{-1})) \times \mathcal{S}_k(\Gamma_1(Nd^{-1})) &\to \mathcal{S}_k(\Gamma_1(N)) \\ (f, g) &\mapsto f + g[\alpha_d]_k \end{aligned}$$

The subspace of oldforms at level N is given by

$$\mathcal{S}_k(\Gamma_1(N))^{\text{old}} := \sum_{\substack{p \mid N \\ \text{prime}}} i_p(\mathcal{S}_k(\Gamma_1(Np^{-1})) \times \mathcal{S}_k(\Gamma_1(Np^{-1}))).$$

Remark 127. $\mathcal{S}_k(\Gamma_1(N))^{\text{old}}$ can be interpreted as the subspace of $\mathcal{S}_k(\Gamma_1(N))$ that comes from lower levels, i.e. from any $\mathcal{S}_k(\Gamma_1(M))$ with $M \mid N$.

Definition 128. *The subspace of newforms corresponds to the orthogonal complement of the subspace of oldforms with respect to the Petersson inner product, i.e.*

$$\mathcal{S}_k(\Gamma_1(N))^{\text{new}} := (\mathcal{S}_k(\Gamma_1(N))^{\text{old}})^{\perp}.$$

As we could expect, there is an analogue of Theorem 125 for both the subspaces of oldforms and newforms. In particular, we have the following result.

Theorem 129. *The spaces $\mathcal{S}_k(\Gamma_1(N))^{\text{old}}$ and $\mathcal{S}_k(\Gamma_1(N))^{\text{new}}$ have orthogonal basis of eigenforms for the Hecke operators $\{\langle n\rangle,\ T_n : \gcd(n, N) = 1\}$.*

Proof. See [Dia05], Corollary 5.6.3. □

As we mentioned before, we are interested in removing the restriction regarding the coprimality between the Hecke operators and the level, which is a condition still required in Theorem 129. The reason behind that is that this is not true for the case of the space of oldforms. However, if we restrict ourselves to $\mathcal{S}_k(\Gamma_1(N))^{\text{new}}$ we obtain the desired result.

In order to properly state the theorem, we need to introduce one last notion: the concept of newform.

Definition 130. *Let $f \in \mathcal{S}_k(\Gamma_1(N))^{\text{new}}$ be an eigenform for the Hecke operators $\{\langle n\rangle,\ T_n : n \in \mathbb{N}\}$ with Fourier expansion $f(\tau) = \sum_{n=0}^{\infty} a_n q^n$. f is a newform if $a_1(f) = 1$.*

At first sight, it may seem that newforms are simply a particular and rare case of modular forms. However, they have a number of properties that make them an interesting object of study. In particular, the following theorem is an example of them.

Theorem 131. *Let $f \in \mathcal{S}_k(\Gamma_1(N))^{\text{new}}$ be an eigenform for the Hecke operators $\{\langle n\rangle,\ T_n : \gcd(n, N) = 1\}$. Then f is a eigenform for the Hecke operators $\{\langle n\rangle,\ T_n : n \in \mathbb{N}\}$ and a suitable scalar multiple of f is a newform. Moreover, the set of newforms is an orthogonal basis of $f \in \mathcal{S}_k(\Gamma_1(N))^{\text{new}}$ and satisfies $T_n f = a_n f$, i.e. the Fourier coefficients coincide with the eigenvalues.*

Proof. See [Dia05], Theorem 5.8.2. □

Another interesting fact about newforms is that there exists a bound for its Fourier coefficients of prime order, which is known as the Petersson conjecture (despite being proved by Deligne in [Del71]). Indeed, this result will be useful later when proving the main theorem of this work.

Theorem 132. *Let $f \in \mathcal{S}_k(\Gamma_1(N))^{\text{new}}$ be an newform with Fourier expansion $f = \sum_{n\geq 0} a_n q^n$. Then we have that $|a_p| \leq 2\sqrt{p}$ for all prime p.*

Proof. See [Del71]. □

Remark 133. The bound from Theorem 132 reminds to the Hasse bound for the number of points on an elliptic curve over a finite field. In fact, it turns out that this similarity is not only a coincidence but a consequence of a deep result known as the modularity Theorem. See [Dia05], Chapters 7, 8 & 9 for further details on this topic.

Finally, there is still another interesting property regarding the coefficients of newforms, which can be seen as kind of first connection between modular forms and algebraic number theory.

Proposition 134. *Let $f \in \mathcal{S}_k(\Gamma_1(N))^{\text{new}}$ be an newform with Fourier expansion $f = \sum_{n\geq 0} a_n q^n$. Then we have that $\{a_n\}_{\geq 1}$ are algebraic numbers and $K_f = \mathbb{Q}(\{a_n\}_{\geq 1})$ is a number field.*

Proof. See [Dia05], Theorem 6.5.1. □

3.3 Twisting on modular forms

As has been seen throughout this chapter, the theory of modular forms is very broad and has many constructions with interesting properties in their own. However, there are a couple of cases with special relevance for our study: those with complex multiplication and inner twists. Thus, the aim of this section is to introduce these notions and some of their properties for use in the next chapter.

3.3.1 Modular forms with complex multiplication

At first, it is important to point out that there are several equivalent definitions of modular forms with complex multiplication (from now on denoted by CM modular forms) in literature. Each of them arises from a different perspective and usually one or another is chosen depending on which suits better in that specific context. In our case, we follow the construction proposed by Ribet in [Rib77]. In order to do so, we first need to introduce a slightly generalization of the notion of modular forms.

Recall that, apart from the holomorphic conditions, modular forms must satisfy the weakly modularity condition. Thus, an option to generalize them is to replace this condition for a more flexible one. This is precisely the idea behind modular forms with nebentypus.

Definition 135. *Let χ be a Dirichlet character mod N. A function $f : \mathcal{H} \to \mathbb{C}$ is a modular form of weight k and Nebentypus χ if*

1. *f is holomorphic on $\mathcal{H}$.*
2. *$f[\gamma]_k$ is holomorphic at ∞, i.e. $f[\gamma]_k$ is bounded as $\text{Im}(\tau) \to \infty$ for all $\gamma \in \text{SL}(2,\mathbb{Z})$.*
3. *$f(\gamma\tau) = \chi(d)(cz+d)^k f(\tau)$ for all $\gamma \in \Gamma$, $\tau \in \mathcal{H}$.*

$\mathcal{M}_k(N,\chi)$ denotes the set of modular forms of weight k and Nebentypus χ.

Remark 136. Notice that when χ is the trivial character, we get that $\mathcal{M}_k(N,\chi) = \mathcal{M}_k(\Gamma_0(N))$.

Almost all the theory we have been developed so far for modular forms also holds for modular forms with Nebentypus. However, our interest behind them lies in the following construction.

Definition 137. *Let $f \in \mathcal{S}_2(\Gamma_0(N))$ be a newform with Fourier expansion $f = \sum_{n\geq 0} a_n q^n$, and let χ be a Dirichlet character mod D. The twisting of f by χ, denoted by $f \otimes \chi$, is given by*

$$f \otimes \chi = \sum_{n\geq 0} \chi(n) a_n q^n.$$

At first sight, it may not seem clear if the twisting of a newform by a Dirichlet character is again a modular form. It turns out that it is not necessarily the case, but this can be corrected if we consider the notion of modular forms with Nebentypus.

Proposition 138. *Let $f \in \mathcal{S}_2(\Gamma_0(N))$ be a newform, and let χ be a Dirichlet character mod D. Then the twisting of f by χ belongs to $\mathcal{M}_k(ND^2, \chi^2)$.*

Proof. See [Shi71], Proposition 3.64. □

With this in mind, we are ready to define the notion of CM modular forms.

Definition 139. *Let $f \in \mathcal{S}_2(\Gamma_0(N))$ be a newform with Fourier expansion $f = \sum_{n\geq 0} a_n q^n$, and let χ be a non-trivial Dirichlet character. The modular form f is a CM modular form by χ if*

$$\chi(p) a_p = a_p$$

for all primes p in a set of primes of density 1.

Remark 140. As previously mentioned, there are several equivalent definitions for CM modular forms. In that sense, one of the most natural ones is defining them as the modular forms arising from abelian varieties with complex multiplication. See [Shi97], Chapter VI for a detailed treatment of this topic.

3.3.2 Modular forms with inner twists

Apart from CM modular forms, there is another particular case of modular forms that can be characterized in terms of the twisting by a Dirichlet character. These are known as modular forms with inner twists. Again, despite existing several alternatives that lead to equivalent definitions of modular forms with inner twists, we follow the construction proposed by Ribet in [Rib77].

Definition 141. *Let $f \in \mathcal{S}_2(\Gamma_0(N))$ be a newform without CM, and let $\mathbb{Q}_f$ be the number field generated by its Fourier coefficients a_n. The modular form f has an inner twist if there exist $\tau \in \mathrm{Aut}(\mathbb{Q})$ not pointwise fixing $\mathbb{Q}_f$, and a Dirichlet character χ unramified outside N such that*

$$a_p = \chi(p)\tau(a_p)$$

for all primes p in a set of primes of density 1.

Based on Definition 141, it makes sense to give a name to the set of automorphisms of the field $\mathbb{Q}_f$ giving an inner twist.

Definition 142. *We define by Γ to be the set of automorphisms of the field $\mathbb{Q}_f$ giving an inner twist together with the identity, i.e.*

$$\Gamma = \{\tau \in \mathrm{Aut}(\mathbb{Q}) : \text{ there exists a character } \chi \text{ such that } a_p = \chi(p)\tau(a_p) \text{ for almost all } p\}.$$

The set Γ has some nice properties. In particular, it has an abelian group structure.

Proposition 143. *Γ is an abelian group.*

Proof. See [Rib80], Proposition 3.3. □

The theory of modular forms with inner twists is considerably broad, with several interesting results based on definitions 141 and 142 and in Proposition 143. However, it turns out that among all of them there is one that is really useful for our later purposes. It corresponds to the following proposition.

Proposition 144. *Let $f \in \mathcal{S}_2(\Gamma_0(N))$ be a newform without CM, let $\mathbb{Q}_f$ be the number field generated by its Fourier coefficients a_n, and let $F_f = \mathbb{Q}_f^\Gamma$ be the fixed field of Γ. Then, F_f is the field generated by $\{a_p^2\}$, with $p \nmid N$ ranging over a set of primes of density 1.*

Proof. See [Rib80], p. 49. □

Moreover, there are some particular cases in which the statement of Proposition 143 can be even refined assuring that F_f is not only generated by $\{a_p^2\}$ with $p \nmid N$ ranging over a set of primes of density 1, but also just by one of these values. Precisely this is what states the corollary that follows.

Corollary 145. *Let $f \in \mathcal{S}_2(\Gamma_0(N))$ be a newform without CM, let $\mathbb{Q}_f$ be the number field generated by its Fourier coefficients a_n, and let $F_f = \mathbb{Q}_f^\Gamma$ be the fixed field of Γ. If $|\Gamma| \leq 2$, then there exists a prime v not dividing N such that a_v^2 generates F_f over $\mathbb{Q}$.*

Proof. See [Die01], p. 397. □

Chapter 4

Subgroup classification of $\mathbf{GL}(2, \mathbb{F}_q)$

In chapters 2 and 3 we have introduced the notions of Galois representations and modular forms, which are the essential objects needed to understand the statement of Ribet's theorem. However, as mathematicians we are not only interested in the statement but also in the proof. In this regard, it turns out that this proof is mainly based in another interesting result: the subgroup classification of $\mathrm{GL}(2, \mathbb{F}_q)$.

The subgroup classification of $\mathrm{GL}(2, \mathbb{F}_q)$ can be seen as a particular case of a much more broader program for describing the finite sugbroups of $\mathrm{GL}(2, K)$ for an arbitrary field K. The first steps in this direction were made by Klein, who did this classification for the case of the complex numbers at the end of the 19th century. A couple of decades later, at the beginning of the 20th century, Dickson extended this result to deal with the case over finite fields, which is the one in which we are interested. Thereafter, many mathematicians, among which Serre, Beauville or Suzuki, have worked in this classification over other particular fields using advanced mathematical techniques such as Galois cohomology.

The aim of this section is to introduce the subgroup classification of $\mathrm{GL}(2, \mathbb{F}_q)$ as well as all the notions needed to understand its statement. Moreover, some particular group properties needed for the proof of the Ribet's theorem are also presented and proved.

4.1 Groups appearing in Dickson's theorem

The first thing to point out is that despite being a subgroup classification of $\mathrm{GL}(2, \mathbb{F}_q)$, Dickson's result also relies on the projective general linear group over finite fields. Thus, we start by defining it.

Definition 146. *Let q be a power of a prime, $\mathbb{F}_q$ the finite field of q elements and $Z(2, \mathbb{F}_q) \subset \mathrm{GL}(2, \mathbb{F}_q)$ the subgroup of non-zero scalar matrices. Then the projective general linear group over $\mathbb{F}_q$, denoted by $\mathrm{PGL}(2, \mathbb{F}_q)$, is defined as*

$$\mathrm{PGL}(2, \mathbb{F}_q) := \mathrm{GL}(2, \mathbb{F}_q)/\mathrm{Z}(2, \mathbb{F}_q).$$

Remark 147. Definition 146 states that the projective general linear group simply corresponds to the general linear group except that it also identifies all matrices that are equal up to a scalar constant. The analogy with the construction of the projective space is therefore clear and hence the name.

There is a subgroup of $\mathrm{PGL}(2, \mathbb{F}_q)$ that will appear later when discussing the applications of the Ribet's theorem, which is known as the projective special linear group.

Definition 148. *Let q be a power of a prime, $\mathbb{F}_q$ the finite field of q elements and $SZ(2, \mathbb{F}_q) \subset \mathrm{SL}(2, \mathbb{F}_q)$ the subgroup of non-zero scalar matrices with unit determinant. Then the projective special linear group over $\mathbb{F}_q$, denoted by $\mathrm{PSL}(2, \mathbb{F}_q)$, is defined as*

$$\mathrm{PSL}(2, \mathbb{F}_q) := \mathrm{SL}(2, \mathbb{F}_q)/\mathrm{SZ}(2, \mathbb{F}_q).$$

An important property of $\mathrm{PSL}(2, \mathbb{F}_q)$ is that it corresponds to a maximal subgroup of $\mathrm{PGL}(2, \mathbb{F}_q)$ whenever q is odd. This is an immediate consequence of the following proposition.

Proposition 149. *Let q be a power of an odd prime. Then $[\mathrm{PGL}(2, \mathbb{F}_q) : \mathrm{PSL}(2, \mathbb{F}_q)] = 2$.*

Proof. We start by computing the order of $\mathrm{GL}(2, \mathbb{F}_q)$. For the first column, we have $q^2 - 1$ different possibilities that correspond to all combinations except the zero column. Then, since the determinant must be different from 0, for the second column we have $q^2 - q$ possibilities, which correspond to all combinations except to the zero column and the $q - 1$ multiples of the first column. Hence, $|\mathrm{GL}(2, \mathbb{F}_q)| = (q^2 - 1)(q^2 - q) = q(q + 1)(q - 1)^2$.

Next, it is clear that $|\mathrm{Z}(2, \mathbb{F}_q)| = q - 1$, so combining this with the previous computation we get that we get that $|\mathrm{PGL}(2, \mathbb{F}_q)| = q(q + 1)(q - 1)$.

To compute the order of $\mathrm{SL}(2, \mathbb{F}_q)$ we consider the exact sequence

$$1 \to \mathrm{SL}(2, \mathbb{F}_q) \to \mathrm{GL}(2, \mathbb{F}_q) \xrightarrow{\det} \mathbb{F}_q^\times \to 1.$$

From there, we get that $|\mathrm{GL}(2, \mathbb{F}_q)/\mathrm{SL}(2, \mathbb{F}_q)| = q - 1$, which together with the previous computation implies that $|\mathrm{SL}(2, \mathbb{F}_q)| = q(q + 1)(q - 1)$.

Then, since q is odd, we have that $\mathrm{SZ}(2, \mathbb{F}_q) = \{\pm \mathrm{Id}\}$. Therefore, using the previous result we get that $|\mathrm{PSL}(2, \mathbb{F}_q)| = \frac{q(q+1)(q-1)}{2}$.

From all this computations is immediate to conclude that $[\mathrm{PGL}(2, \mathbb{F}_q) : \mathrm{PSL}(2, \mathbb{F}_q)] = 2$ as desired.

□

4.1.1 Cyclic, dihedral, symmetric, and alternate groups

Once we have introduced the projective linear group, we move to define all group classes appearing in Dickson's classification. Some of them are familiar for almost all mathematicians, but for the sake of completeness we also list them. In fact, we start by the best known ones.

Definition 150. *Let G be a group. G is a cyclic group if it is generated by a single element, i.e. there exists $g \in G$ such that*

$$G = \langle g \rangle = \{g^k : k \in \mathbb{Z}\}.$$

Definition 151. *Let $n \in \mathbb{N}$. The dihedral group of order $2n$, denoted by D_n, is the group of symmetries of the regular n-gon. In particular,*

$$D_n = \{\alpha, \beta : \alpha^n = \beta^2 = e,\ \beta\alpha\beta = \alpha^{-1}\}.$$

α is known as a rotation and β as a reflection.

Definition 152. *Let $n \in N$. The n-th symmetric group, denoted by S_n corresponds to the permutations that can be performed on n elements.*

Definition 153. *Let $n \in N$. The n-th alternating group, denoted by A_n corresponds to the even permutations that can be performed on n elements.*

4.1.2 Classical subgroups of $\mathbf{GL}(2, \mathbb{F}_q)$

So far, the classes of groups we have defined are generic in the sense that they definition does not depend on the general linear group $\mathrm{GL}(2, \mathbb{F}_q)$. This is not the case for the remaining groups that we must introduce, which are known as classical subgroups of $\mathrm{GL}(2, \mathbb{F}_q)$.

The first in this list correspond to the known as Borel subgroups.

Definition 154. *A Borel subgroup of $\mathrm{GL}(2, \mathbb{F}_q)$ is any subgroup conjugate to the subgroup of upper triangular matrices in $\mathrm{GL}(2, \mathbb{F}_q)$.*

Remark 155. From Definition 154 it is clear that any Borel subgroup has order $q(q-1)^2$.

Next, we move to define the Cartan subgroups, in which is necessary to distinguish between the split and non-split cases.

Definition 156. *A split Cartan subgroup of $\mathrm{GL}(2, \mathbb{F}_q)$ is any subgroup conjugate to the subgroup of diagonal matrices in $\mathrm{GL}(2, \mathbb{F}_q)$.*

Remark 157. From Definition 156 it is straightforward to notice that any split Cartan subgroup is isomorphic to $(\mathbb{F}_q^\times)^2$.

Definition 158. *A non-split Cartan subgroup of $\mathrm{GL}(2, \mathbb{F}_q)$ is any subgroup conjugate to the following subgroup:*

$$N_{sp} := \left\{ \begin{bmatrix} a & \delta b \\ b & a \end{bmatrix} \right\} \subset \mathrm{GL}(2, \mathbb{F}_q),$$

where δ is any fixed quadratic non-residue modulo q.

Remark 159. There are several equivalent definitions for Cartan subgroups, both split and non-split. For example, they can be defined in terms of the centralizer of a maximal torus or as a particular subalgebra of $\mathrm{End}((\mathbb{F}_q)^2)$ for the non-split case. A further discussion on this subject can be found in [Bor91], Chapter IV, Section 11.

Finally, the last type of groups in the list are the normalizers of Cartan subgroups. Again, we must distinguish between the split and non-split cases. In this regard, we could leave it like that and define them as being the normalizers of Cartan subgroups, but we have opted to give an explicit description of them.

Definition 160. *A normalizer of a split Cartan subgroup of* $\mathrm{GL}(2, \mathbb{F}_q)$ *is any subgroup conjugate to the following subgroup:*

$$N_{sp} \coloneqq \left\{ \begin{bmatrix} a & 0 \\ 0 & d \end{bmatrix}, \begin{bmatrix} 0 & c \\ d & 0 \end{bmatrix} \right\} \subset \mathrm{GL}(2, \mathbb{F}_q).$$

Definition 161. *A normalizer of a non-split Cartan subgroup of* $\mathrm{GL}(2, \mathbb{F}_q)$ *is any subgroup conjugate to the following subgroup:*

$$N_{ns} \coloneqq \left\{ \begin{bmatrix} a & \delta b \\ b & a \end{bmatrix}, \begin{bmatrix} a & -\delta b \\ b & -a \end{bmatrix} \right\} \subset \mathrm{GL}(2, \mathbb{F}_q),$$

where δ is any fixed quadratic non-residue modulo q.

Remark 162. From definitions 160 and 161 it is straightforward to notice that the normalizer of a split (resp. non-split) Cartan subgroup contains the split (resp. non-split) Cartan subgroup as a subgroup with index 2. Hence, a Cartan subgroup is normal in its normalizer.

4.2 Dickson's theorem & other group results

Once we have defined all groups appearing in Dickson's classification of subgroups of $\mathrm{GL}(2, \mathbb{F}_q)$, we are ready to present Dickson's theorem as well as other group theoretical results that will be useful for proving Ribet's theorem.

4.2.1 Dickson's theorem

We start by stating Dickson's classification of subgroups of $\mathrm{GL}(2, \mathbb{F}_q)$.

Theorem 163 (Dickson)**.** *Let G be a subgroup of* $\mathrm{GL}(2, \mathbb{F}_q)$ *and let H be its image in* $\mathrm{PGL}(2, \mathbb{F}_q)$*. If the order of G is divisible by q, one of the following holds:*

1. *G is contained in a Borel subgroup.*
2. *G contains* $\mathrm{SL}(2, \mathbb{F}_q)$*.*

If the order of G is not divisible by q, one of the following holds

1. *H is cyclic and G is contained in a Cartan subgroup.*
2. *H is dihedral and G is contained in the normalizer of a Cartan subgroup but not in a Cartan subgroup.*
3. *H is isomorphic to one of the following special groups: A_4, S_4, A_5.*

Proof. Despite being out of the scope of this work, we give a brief summary of the techniques used to prove this theorem, since they correspond to classical results in group theory. In that sense, in his proof Dickson mainly uses the Sylow theorems together with some theory of group actions. In particular, it deals with the action of $\mathrm{PGL}(2, \mathbb{F}_q)$ on $\mathbb{P}^1(\mathbb{F}_q)$.

See [Dic58], Chapter XII, Section 260 for a detailed proof. □

4.2.2 Other group results

As previously mentioned, apart from Dickson's result on the classification of subgroups of $\mathrm{GL}(2, \mathbb{F}_q)$ we also need to use some other properties of the groups listed above for the proof of Ribet's theorem.

First, we must recall the order of some particular groups.

Proposition 164. *Let G be a group. Then the following hold:*

1. *If G is a cyclic group generated by g, then $|G| = \mathrm{ord}(g)$.*
2. *If $G \cong D_n$, then $|G| = 2n$.*
3. *If $G \cong S_n$, then $|G| = n!$.*
4. *If $G \cong A_n$, then $|G| = \frac{n!}{2}$.*

Proof. We prove each of the statements separately.

1. If G is a cyclic group generated by g, then it is clear that $|G| = \mathrm{ord}(g)$ simply by looking at Definition 150.
2. If $G \cong D_n$, we consider G as the group of symmetries of the regular n-gon. Then, it is clear that $|G| \geq 2n$ since we have n rotations and also the same n rotations reflected, which are clearly pairwise distinct. Then, since these symmetries are distance preserving and map vertices to vertices, given two adjacent vertices A and B we get that we have n possibilities for the first vertex A (any of the n vertices), while for the second there are only 2 possibilities (the 2 adjacent vertices). Hence, it is clear that $|G| = 2n$.
3. If $G \cong S_n$, we have that G corresponds to the set of permutations of n elements. Then, since there are exactly $n!$ different permutations of this type, we get that $|G| = n!$.
4. If $G \cong A_n$, we have that G corresponds to the set of even permutations of n elements. Then, since there are exactly $\frac{n!}{2}$ different permutations of this type, we get that $|G| = \frac{n!}{2}$.

□

Next, we present some results regarding the structure of cyclic subgroups of the dihedral groups.

Proposition 165. *Let $n \geq 3$ and D_n a dihedral group. Then there exists a unique cyclic subgroup of D_n of order n.*

Proof. Let $D_n = \{\alpha, \beta : \alpha^n = \beta^2 = e,\ \beta\alpha\beta = \alpha^{-1}\}$. By definition, it is clear that $\langle\alpha\rangle$ is a cyclic subgroup of order n of D_n. Hence, the existence is proved.

Let's move to check its uniqueness. By construction, any other cyclic subgroup of order n of D_n must be generated by an element either of the form $\beta\alpha^i$ or α^i for $0 \leq i \leq n$. Therefore, it suffices to check that $\text{ord}(\beta\alpha^i) < n$ for $0 \leq i \leq n$. In that sense, since $\alpha\beta = \beta\alpha^{-1}$ we have that

$$(\beta\alpha^i) = \beta\alpha^i\beta\alpha^i = \beta\alpha^{i-1}\beta\alpha^{-1}\alpha^i = \beta\alpha^{i-1}\beta\alpha^{i-1} = ... = s^2 = e.$$

Thus, $\text{ord}(\beta\alpha^i) = 2 < 3 \leq n$ and uniqueness is proved as desired. □

Proposition 165 motivates the following definition, which is crucial for the next result.

Definition 166. *Let $n \geq 3$ and D_n a dihedral group. The cyclic subgroup of D_n of order n is defined as the core of D_n.*

Remark 167. Often in literature, the notion of core introduced in Definition 166 is referred as center. However, this leads to ambiguity with the usual notion of the center of a group, Therefore, we have opted to denoted as core.

Once the core of a dihedral group is defined, we state and proof the following proposition, that will be key later during the proof of Ribet's theorem.

Proposition 168. *Let $n, k \geq 3$, D_n a dihedral group, and $G \subset D_n$ a cyclic subgroup of order k. Then G is contained in the core C of D_n.*

Proof. Let g be the generator of G and α the generator of the core C. Notice that to prove the proposition it is enough to prove that $g = \alpha^i$ for some $0 \leq i \leq n$. In that sense, by Proposition 165 we know that for any $h \in D_n \setminus C$ the order of h is equal to 2. Therefore, since $\text{ord}(g) = k \geq 3$, we conclude that $g \in C$ and $G \subset C$ as desired. □

After studying in detail the cyclic subgroups of the dihedral group, we must also take a look at the cyclic subgroups of the particular cases of the symmetric and alternating groups appearing in Theorem 163.

Proposition 169. *Neither of the groups A_4, S_4, and A_5 have a cyclic subgroup of order ≥ 6.*

Proof. First, notice that since $A_4 \subset S_4$, it suffices to prove the statement only for S_4 and A_5. In order to do so, we simply check that there is not an element of order ≥ 6 in none of both groups.

On one hand, S_4 consists by definition in the group of permutations that can be performed on 4 elements. Thus, taking into account that each permutation can be decomposed as a product of disjoint cycles and that its order corresponds to the least common multiple of the orders of these cycles, we conclude that the maximum order for an element of S_4 is precisely 4. Hence, neither A_4 nor S_4 have a cyclic subgroup of order ≥ 6.

On the other hand, A_5 corresponds by definition to the group of even permutations that can be performed on 5 elements. Therefore, using again the previous reasoning in the order of these permutations we conclude that the maximum order for an element of A_5 is 5. So A_5 does not have a cyclic subgroup of order ≥ 6.

□

Finally, the last results we need to introduce are related to Cartan subgroups and their normalizers. In particular, they are about the eigenvalues of the elements in a non-split Cartan subgroup and the traces of the normalizers.

Proposition 170. *Any non-split Cartan subgroup does not have any element with distinct rational eigenvalues.*

Proof. Let G be a non-split Cartan subgroup. From Definition158 we know that any $g \in G$ is of the form $g = \begin{bmatrix} a & \delta b \\ b & a \end{bmatrix}$, with $a, b, \delta \in \mathbb{F}_q$, δ a non-residue module q, and $\det(g) \neq 0$. Now, in order to obtain the eigenvalues of g, we compute its characteristic polynomial. In particular, it corresponds to

$$f(x) = x^2 - 2ax + a^2 - \delta b^2.$$

From here, notice that $a^2 - \delta b^2 \neq 0$. This is because if $b = 0$, then $a \neq 0$ due to the fact that $\det(g) \neq 0$; and if $b \neq 0$, then $a^2 - \delta b^2$ can not equal to 0 since this would imply that δ is a residue modulo q.

Once we have done this observation, in order to compute eigenvalues, i.e. the roots of $f(x)$, we split by cases depending on the characteristic of $\mathbb{F}_q$.

If $\text{char}(\mathbb{F}_q) \neq 2$, by the formula of the quadratic equation we get that $x = a \pm \sqrt{\delta} b$, and since δ a non-residue module q we get that either both eigenvalues are non-rational or they are both equal to a.

If $\text{char}(\mathbb{F}_q) = 2$, we get that $f(x) = x^2 + a^2 - \delta b^2$, which implies that $x = a - \sqrt{\delta} b$ is a double root. Again, we see that since δ a non-residue module q g can not have distinct rational eigenvalues. □

Proposition 171. *Let C be a Cartan subgroup and N its normalizer. Then any element $n \in N \setminus C$ has trace 0.*

Proof. It suffices to take a look at the definitions of Cartan subgroups and their normalizers.

In particular, if C is a split Cartan subgroup then

$$N \setminus C = \left\{ \begin{bmatrix} 0 & b \\ c & 0 \end{bmatrix} \right\} \subset GL(2, \mathbb{F}_q),$$

while if C is a non-split Cartan subgroup then

$$N \setminus C = \left\{ \begin{bmatrix} a & -\delta b \\ b & -a \end{bmatrix} \right\} \subset GL(2, \mathbb{F}_q).$$

Hence, in both cases it is clear that for any $n \in N \setminus C$ we have that trace$(n) = 0$ as desired. □

Chapter 5

Images of Galois representations attached to newforms

The image of Galois representations attached to newforms without Complex Multiplication of weight 2 are well-understood due to the work of Ribet [Rib85]. In this paper, Ribet states that for almost all prime numbers l, the image of the Galois representation associated to a modular form of weight 2 and level N is *as large as possible*, the meaning of which we will precise later on this Chapter.

This theorem was a major step in understanding the relationship between modular forms and Galois representations. In particular, it helped to realize projective linear groups over finite fields as Galois groups over $\mathbb{Q}$. Moreover, this theorem is also part of the framework that lead to Andrew Wiles' proof of Fermat's Last Theorem.

The aim of this Chapter is to state and prove Ribet's theorem, using a combination of all the theory developed in previous chapters together with other advanced algebraic and geometric techniques.

5.1 The statement of Ribet's theorem

Along this section we will develop all the framework needed for stating Ribet's theorem as well as a brief outline of the strategy we will follow to prove it.

5.1.1 The Eichler-Shimura relation

Before stating Ribet's theorem about the images of the Galois representations attached to modular forms, it is worth spending a few lines explaining how to obtain Galois representations from modular forms since it corresponds to a deep result in number theory known as the Eichler-Shimura relation.

As one may expect, try to summarize it in a couple of paragraphs is an unfeasible task due to the huge amount of theory, constructions, and subtleties hidden behind its statements. Thus, we simply outline the main results that are useful for our purposes and give references where the subject is dealt with in greater depth.

Broadly speaking, the Eichler-Shimura construction provides a way of obtaining a Galois representation from a modular form of weight 2. This connection is made precise through the language of abelian varieties and the theory of Hecke operators that we have seen so far. It starts by constructing the Jacobian $J_0(N)$ of the modular curve $X_0(N)$, which decomposes as a product of abelian varieties A_f arising from newforms f of weight 2. In particular, the dimension of each abelian variety A_f is equal to $[K_f : \mathbb{Q}]$, where K_f is the field extension generated by the Fourier coefficients of f. Then, a two dimensional Galois representation is attached to each of these abelian varieties A_f by means of their Tate modules. Therefore, through the construction of the abelian varieties A_f we are able to obtain Galois representations from modular forms. A complete treatment of this construction can be found in [Dia05], Chapter 8 & 9.

In our case, the interesting result derived from the Eichler-Shimura theory is the following.

Theorem 172. *Let $f \in \mathcal{S}_2(\Gamma_0(N))$ be a newform without inner twists nor CM. Let $\mathbb{Q}_f$ be the number field generated by its Fourier coefficients a_n and $\mathcal{O}$ its ring of integers. For every prime l, let $\mathcal{O}_l = \mathcal{O} \otimes_{\mathbb{Z}} \mathbb{Z}_l$ and $\mathbb{Q}_{f,l} = \mathbb{Q}_f \otimes_{\mathbb{Q}} \mathbb{Q}_l$. Then there exists a Galois representation*

$$\rho_l : \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{GL}(2, \mathcal{O}_l) \subset \mathrm{GL}(2, \mathbb{Q}_{f,l})$$

unramified outside lN such that for every prime $p \nmid lN$

$$\mathrm{trace}(\rho_l(\mathrm{Frob}_{\mathfrak{p}})) = a_p \text{ and } \det(\rho_l(\mathrm{Frob}_{\mathfrak{p}})) = p.$$

Remark 173. Theorem 172 can be seen a particular case of a much more general statement due to Deligne. In that sense, in [Del71] Deligne proved a version of this theorem for newforms of any weight.

Remark 174. Being precise, the condition of requiring that the newform does not have inner twists is not necessary for Theorem 172 to hold. However, as we will see later, adding this assumption simplifies our work.

The Galois representations that we will study correspond precisely to those stated in Theorem 172. The reason for this restriction lies in the fact that this theorem not only allows us to construct a Galois representation attached to a newform of weight 2, but also gives us additional information of the representation by establishing its images on the absolute Frobenius. Therefore, it offers an interesting set up for studying the images of Galois representations by combining both the tools from Galois representation theory and the ones from modular forms.

5.1.2 The statement

Once enclosed the framework in which we will work, we are ready to state Ribet's theorem about the images of the Galois representations attached to modular forms.

Theorem 175 (Ribet's theorem)**.** *Let ρ_l be a Galois representation as in Theorem 172, $G_{\mathbb{Q}} = Gal(\overline{\mathbb{Q}}/\mathbb{Q})$, and $G_l = \rho_l(G_{\mathbb{Q}})$. Let $A_l = \{x \in \mathrm{GL}(2, \mathcal{O}_l) : \det(x) \in \mathbb{Z}_l^{\times}\}$. Then the equality $G_l = A_l$ holds for almost every prime.*

At first sight, it is not clear at all why the statement of Theorem 175 implies that the image of the Galois representations attached to newforms of weight 2 is *as large as possible*. Indeed, we have still not defined rigorously what does this notion mean.

In that sense, by *as large as possible* we mean that the image is the biggest group it can be taking into account the restrictions imposed by the Galois representation given by Theorem 172. Therefore, since ρ goes from $G_{\mathbb{Q}}$ to $\mathrm{GL}(2, \mathcal{O}_l)$ it is obvious that $G_l \subset \mathrm{GL}(2, \mathcal{O}_l)$. Moreover, Theorem 172 also states that the determinant map coincides on the image of the absolute Frobenius with the image of the l-adic cyclotomic character. Then, since ρ_l is continuous and the set of absolute Frobenius is dense by Theorem 52, we conclude that the determinant map coincides with the l-adic cyclotomic character in G_l. Thus, since the output of the l-adic cyclotomic character takes values in $\mathbb{Z}_l^{\times}$, we get that $\det(x) \in \mathbb{Z}_l^{\times}$ for all $x \in G_l$. Hence, for any l it holds that $G_l \subset A_l$, and it makes sense to say that the image is *as large as possible* when $G_l = A_l$.

5.2 Outline of the proof of Ribet's theorem

In order to prove Ribet's theorem, we combine the strategy followed by Ribet in [Rib85] and the one appearing in [Die01], Chapter 2. In particular, we develop some of their arguments with the aim of stating the proof as self-contained as possible.

Having said that, the first step to be taken refers to the decomposition of the Galois representation ρ_l.

Using the decompositions

$$\mathbb{Q}_{f,l} = \prod_{\lambda | l} \mathbb{Q}_{f,\lambda} \text{ and } \mathcal{O}_l = \prod_{\lambda | l} \mathcal{O}_\lambda$$

we can decompose ρ_l as a direct sum of Galois representations

$$\rho_\lambda : \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{GL}(2, \mathcal{O}_\lambda) \subset \mathrm{GL}(2, \mathbb{Q}_{f,\lambda}),$$

where λ refers to the primes of $\mathcal{O}$ over l.

Then, let $\mathbb{F}_\lambda$ be the residue field of λ and consider the reduction $\overline{\rho}_\lambda$ of ρ_λ consisting in composing ρ_λ with the reduction map $\mathrm{GL}(2, \mathcal{O}_\lambda) \to \mathrm{GL}(2, \mathbb{F}_\lambda)$. We define $\overline{G}_\lambda$ as $\overline{\rho}_\lambda(G_{\mathbb{Q}})$.

At first sight, it may not seem clear why do we care about this decomposition of the Galois representation. However, it turns out that in his previous work [Rib75], Ribet proved the following theorem, which is strongly related with our purposes and where this decomposition plays an important role.

Theorem 176. *The equality $G_l = A_l$ holds whenever all the following conditions are satisfied:*

1. *$l \geq 5$.*
2. *l does not ramify in $\mathbb{Q}_f/\mathbb{Q}$.*

3. *The determinant map* $G_l \to \mathbb{Z}_l^\times$ *is surjective.*
4. G_l *contains an element* x_l *such that* $(\text{trace}(x_l))^2$ *generates* $\mathcal{O}_l$ *as a* $\mathbb{Z}_l$*-algebra.*
5. *For each* $\lambda|l$*, the group* $\overline{G}_\lambda$ *is an irreducible subgroup of* $\text{GL}(2, \mathbb{F}_\lambda)$ *whose order is divisible by* l*.*

Remark 177. Notice that condition 5 of Theorem 176 refers explicitly to the decompositions previously mentioned.

The importance of Theorem 176 relies in the fact that it enables us to turn the proof of Ribet's theorem into a matter of verifying these 5 conditions. Thus, after proving it, this is precisely the strategy we will follow in order to complete the Ribet's theorem.

5.2.1 Proof of Theorem 176

The proof of Theorem 176 is not trivial at all and it requires several steps. Indeed, to present it in full detail is out of the scope of this work since it would require to many auxiliary results. Hence, we have opted to outline the general strategy followed by Ribet, to prove the main results, and to give some comments on the remaining steps. The fully detailed proof can be found in [Rib75].

Having said that, the main idea of the proof consists in using condition 5 to show that

$$\prod_{\lambda|l} \text{SL}(2, \mathcal{O}_\lambda) \subset G_l.$$

To do so, notice that condition 5 together with Theorem 163 assures that $\text{SL}(2, \mathbb{F}_\lambda) \subset \overline{G}_\lambda$. In particular, the order of $\overline{G}_\lambda$ being multiple of l discards the cyclic, the dihedral, and the special groups cases; while the irreducibility discards the possibility of the Borel subgroup case. With this in mind, the following theorem theorem allows us to pass from the image on the residue field to the ring of integers.

Theorem 178. *Let* $l \geq 5$ *be a prime,* $K_1, ..., K_t$ *finite extensions of* $\mathbb{Q}_l$*, and* $\mathcal{O}_1, ..., \mathcal{O}_t$ *their rings of integers. If* H *is a closed subgroup of*

$$\text{GL}(2, \mathcal{O}_1) \times ... \times \text{GL}(2, \mathcal{O}_t)$$

such that its image mod l *contains*

$$\text{GL}(2, \mathcal{O}_1/l\mathcal{O}_1) \times ... \times \text{GL}(2, \mathcal{O}_t/l\mathcal{O}_t),$$

then H *contains*

$$\text{SL}(2, \mathcal{O}_1) \times ... \times \text{SL}(2, \mathcal{O}_t).$$

The proof of Theorem 178 mainly relies on adapting a famous lemma by Serre, which is the one that follows.

Lemma 179. *Let* $l \geq 5$ *be a prime and* X *be a closed subgroup of* $\text{SL}(2, \mathbb{Z}_l)$ *whose image in* $\text{SL}(2, \mathbb{F}_l)$ *is* $\text{SL}(2, \mathbb{F}_l)$*. Then* $X = \text{SL}(2, \mathbb{Z}_l)$*.*

Proof. We argue by induction on n that X maps onto $\mathrm{SL}(2, \mathbb{Z}/l^n\mathbb{Z})$.

The base case $n = 1$ is clear, so we assume that it holds for n and we prove it for $n+1$. In this regard, it suffices to show that for any $s = \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right] \in \mathrm{SL}(2, \mathbb{Z}_l)$ congruent to $\mathrm{Id} \pmod{l^n}$ there exists $x \in X$ such that $x \equiv s \pmod{l^{n+1}}$.

At this point, we write s as $S = \mathrm{Id} + l^n u$, and since $\det(s) = 1$ we get that $\mathrm{trace}(u) = 0$. From there, it can be seen that u is congruent mod l to a sum of matrices u_i with $u_i^2 = 0$, so we can also assume that $u^2 = 0$.

Then, by induction hypothesis there exists $y \in X$ such that $y = \mathrm{Id} + l^{n-1}u + l^n v$, where v has coefficients in $\mathbb{Z}_l$. Therefore, by setting $x = y^l$ we get that

$$x = \mathrm{Id} + l(l^{n-1}u + l^n v) + \binom{l}{2}(l^{n-1}u + l^n v)^2 + ... + (l^{n-1}u + l^n v)^l.$$

Now, if $n \geq 2$ it is clear that $x \equiv \mathrm{Id} + l^n u \pmod{l^{n+1}}$. Moreover, if $n = 1$ we can use that $u^2 = 0$ and $u + lv \equiv u \pmod{l}$ to get that $x \equiv \mathrm{Id} + lu + (u + lv)^l \pmod{l^2}$. Then, since $(u + lv)^2 \equiv l(uv + vu) \pmod{l^2}$, we get that

$$(u + lv)^l \equiv l(uv + vu)u^{l-2} \equiv 0 \pmod{l^2}$$

because $l \geq 5$. Hence, we conclude that $x \equiv 1 + l^n u \pmod{l^{n+1}}$ in all cases as desired. □

From there, by means of an argument regarding the closure of the commutator subgroup of $\mathrm{GL}(2, \mathcal{O}_1) \times ... \times \mathrm{GL}(2, \mathcal{O}_t)$, Lemma 179 is applied to prove Theorem 178. See [Rib75], Theorem 2.1 for the details on this argument.

After that, the next step consists precisely in using Theorem 178 to prove the following corollary.

Corollary 180. *Let G be a closed subgroup of*

$$A = \{(x_1, ..., x_t) \in \prod_{i=1}^{t} \mathrm{GL}(2, \mathcal{O}_i) : \det(x_1) = ... = \det(x_t) \in \mathbb{Z}_l^\times\}.$$

The equality $G = A$ holds whenever all the following conditions are satisfied:

1. *$l \geq 5$.*
2. *The determinant map $G \to \mathbb{Z}_l^\times$ is surjective.*
3. *The image mod l contains $\prod_{i=1}^{t} \mathrm{SL}(2, \mathcal{O}_i/l\mathcal{O}_i)$.*

Proof. By all the previous development we have that conditions 1 and 3 assure that $\prod_{i=1}^{t} \mathrm{SL}(2, \mathcal{O}_i) \subset G$. Then, using condition 2 we get the following exact sequence.

$$1 \to \prod_{i=1}^{t} \mathrm{SL}(2, \mathcal{O}_i) \to G \xrightarrow{\det} \mathbb{Z}_l^\times \to 1.$$

However, by definition of the group A we also have the exact sequence that follows.

$$1 \to \prod_{i=1}^{t} \mathrm{SL}(2, \mathcal{O}_i) \to A \xrightarrow{\det} \mathbb{Z}_l^{\times} \to 1.$$

Thus, from here we can conclude that $A = G$ as desired.

□

We can view Corollary 180 as a kind of halfway result since it almost assures the desired equality between groups. In that sense, the only remaining step is to solve the difference between the original representation and the λ-adic representations. In particular, one must prove that if each of the ρ_λ has large image, then their product also has large image. Despite not seeming difficult, it turns out that one must be careful when proving this because there could be some overlapping and intersection between components. In this regard, in order to do so the main idea is to combine condition 4 of Theorem 176 together with some technical results in group theory, such as for example the Goursat's lemma. As before, all these details can be found in [Rib75], Theorem 3.1.

5.2.2 Verification of the conditions

Once the proof of Theorem 176 is outlined, we move to verify that all the 5 required conditions are satisfied in the framework we are working in.

First of all, notice not all conditions require the same amount of effort to be proven, since some of them are almost trivial while others involve much more mathematical development. In particular, conditions 1 and 2 are straightforward to check.

Conditions 1 **&** 2

Recall that the statement of Theorem 175 refers to all but a finite set of primes. Thus, condition 1 is for free simply by not considering the primes $l = 2, 3$. Moreover, Remark 24 assures that the amount of primes not satisfying condition 2 is also finite, so condition 2 also holds for almost every prime as desired.

Condition 3

Condition 3 is also quite straightforward. However, it requires to combine the theory of characters seen in Section 2.3 together with Theorem 172. In particular, following the reasoning of Subsection 5.1.2, we get that the determinant map coincides with the l-adic cyclotomic character in G_l. Thus, using Remark 78 we automatically get condition 3.

Condition 4

In the checking of condition 4 is where the hypothesis on the newform not having inner twists is applied. In particular, by Corollary 145 we have that there exists a prime v not dividing N such that a_v^2 generates F_f over $\mathbb{Q}$. In this case, notice that

$$F_f = \mathbb{Q}_f^{\Gamma} = \mathbb{Q}_f^{\mathrm{Id}} = \mathbb{Q}_f,$$

so combining this together with the equality on the trace of the representation from Theorem 172 we get that $(\text{trace}(\rho_\lambda(\text{Frob}_\mathfrak{v})))^2 = a_v^2$ generates Q_f over $\mathbb{Q}$. From there, since Proposition 134 guarantees that a_v is an algebraic integer, it is clear that $a_v \in \mathcal{O}_l$ and a_v^2 generates $\mathcal{O}_l$ as a $\mathbb{Z}_l$-algebra.

Condition 5

In order to prove Theorem 175 we are left to simply verify condition 5. However, it turns out that the proof of this condition requires much more machinery, so we rephrase it as a theorem on its own.

Before stating the theorem, we must point out that from now on we slightly switch the approach used so far. In particular, instead of picking a rational prime l and then consider all primes $\lambda \in \mathcal{O}$ over l, we consider all primes $\lambda \in \mathcal{O}$ and then we set $l = l(\lambda)$ to be the rational prime such that λ is over l.

In this set up, we can state condition 5 as follows.

Theorem 181. *For all but finitely many λ the following conditions are satisfied:*

(a) The representation $\overline{\rho}_\lambda$ is an irreducible 2-dimensional representation over $\mathbb{F}_\lambda$.

(b) The order of the group $\overline{G}_\lambda$ is divisible by l.

As already mentioned, the proof of Theorem 181 is highly nontrivial and requires some other results such as Dickson's classification of all subgroups of the general linear group. In that sense, the next section is fully devoted to the proof of Theorem 181.

5.3 Proof of Theorem 181

The aim of this section is to prove the statement of Theorem 181. In order to do so, we split the proof in several steps. First, we introduce a couple of results from Raynaud about the characterization of 2-dimensional representations coming from a newform $f \in \mathcal{S}_k(\Gamma_0(N))$ that play a key role in proving the theorem. Later, we prove part (a) by using one of these statements. Finally, the proof of part (b) is based on the Dickson's result about the classification of subgroups of $\text{GL}(\mathbb{F}_\lambda)$ stated in Theorem 163, as well as on the use of Raynaud's other result.

5.3.1 On the characterization of 2-dimensional representations

Since our final goal is to prove that the image of given Galois representations attached to a newform is as large as possible, it makes sense to try to study how this kind of representations look like under some particular assumptions.

In this regard, it turns out that this was one of the approaches that Serre considered when trying to answer the same question for the case of Galois representations attached to elliptic curves in [Ser72]. A couple of years later, Raynaud extended this work in [Ray74] by giving some results on the characterization of 2-dimensional Galois representations. Among all of them, we state the one that is useful for the proof of Theorem 181.

The result of our interest is about the characterization of these Galois representations when restricted to the inertia group. As commented in Chapter 2, the inertia group is closely related to the notion of ramification of field extensions, which in turn has some finiteness properties. Hence, since Theorem 181 is stated for almost every prime, knowing the structure of the representation restricted to the inertia groups can be useful information.

Theorem 182. *Let l be a prime, $\lambda \in \mathcal{O}$ be a prime over l, and I_λ the inertia group. If $l \geq 3$, and $l \nmid N$, then the semisimplification of $\overline{\rho}_\lambda$ in I_λ, denoted by $\overline{\rho}_\lambda|_{I_\lambda}$, is one of the following:*

$$\overline{\rho}_\lambda|_{I_\lambda} = \begin{bmatrix} \Psi_1 & 0 \\ 0 & 1 \end{bmatrix} \text{ or } \overline{\rho}_\lambda|_{I_\lambda} = \begin{bmatrix} \Psi_2 & 0 \\ 0 & \Psi_2^l \end{bmatrix},$$

where Ψ_1 and Ψ_2 are the fundamental characters of level 1 and 2 respectively.

From there, there is an interesting result that can be derived from Theorem 182 simply by assuming that the λ-adic representation is reducible.

Theorem 183. *Let l be a prime and $\lambda \in \mathcal{O}$ be a prime over l. If $\overline{\rho}_\lambda$ is reducible, $l \geq 3$, and $l \nmid N$, then*

$$\overline{\rho}_\lambda = \epsilon_1 \oplus \epsilon_1^{-1}\overline{\chi}_l,$$

where ϵ_1 is a Dirichlet character unramified outside N and $\overline{\chi}_l$ is the mod l cyclotomic character.

Remark 184. Despite not providing the proof of these statements, which are a consequence of the works from Serre in [Ser72] and Raynaud in [Ray74], it is interesting at least to give some intuition on the results. Mainly, these characterizations are based on the constraint that the determinant map coincides with the mod l cyclotomic character. In that sense, it is clear that

$$\epsilon_1\epsilon_1^{-1}\overline{\chi}_l = \overline{\chi}_l,$$

and by Remark 82 we get that

$$\Psi_2\Psi_2^l = \Psi_2^{l+1} = \Psi_1 = \overline{\chi}_l|_{I_\lambda}.$$

Remark 185. Both theorems 183 and 182 are particular cases of much more general results. Indeed, in [FJ95] Faltings and Jordan stated these results for newforms of any weight. However, we have restricted ourselves to newforms of weight 2 because is the set up in which we are working.

A useful consequence derived from Theorem 182 is the following proposition.

Proposition 186. *Let l be a prime, $\lambda \in \mathcal{O}$ be a prime over l, and I_λ the inertia group. If $l \geq 3$, and $l \nmid N$, then the projective image of the semisimplification of $\overline{\rho}_\lambda$ in I_λ is cyclic of order $l \pm 1$.*

Proof. First, by Theorem 182 we have that the semisimplification of $\overline{\rho}_\lambda$ in I_λ is given by

$$\overline{\rho}_\lambda|_{I_\lambda} = \begin{bmatrix} \Psi_1 & 0 \\ 0 & 1 \end{bmatrix} \text{ or } \overline{\rho}_\lambda|_{I_\lambda} = \begin{bmatrix} \Psi_2 & 0 \\ 0 & \Psi_2^l \end{bmatrix}.$$

From there, to prove the statement we simply need to check the two cases independently taking into account that its projective image simply corresponds to identifying all matrices that are equal up to a scalar constant.

In the first case, by Remark 80 we immediately have that the projective image is cyclic of order $l-1$.

On the second case, by Remark 82 we get that $\Psi_2^{l+1} = \Psi_1$, so combining this with the first case we get that the projective image is cyclic of order dividing $(l+1)(l-1) = l^2-1$. Moreover, again by Remark 80 we get that $(\Psi_2^l)^{l+1} = (\Psi_2^{l+1})^l = (\Psi_1)^l = \Psi_1$. Thus, when taking the $l+1$-th power in the second case we obtain the diagonal matrix given by

$$\begin{bmatrix} \Psi_1 & 0 \\ 0 & \Psi_1 \end{bmatrix},$$

which is identified with the identity matrix in $\mathrm{PGL}(2, \mathbb{F}_\lambda)$. Hence, we get that in this case the projective image is cyclic of order $l+1$.

Therefore, we conclude that he projective image of the semisimplification of $\overline{\rho}_\lambda$ in I_λ is cyclic of order $l \pm 1$ as desired. □

5.3.2 On the reducibility of representations

In order to prove part (a), our strategy consists in verifying that there are only finitely many primes $\lambda \in \mathcal{O}$ over l for which $\overline{\rho}_\lambda$ is reducible. In other words, our goal is to show that

$$\Lambda = \{\lambda \in \mathcal{O} : \lambda \text{ is a prime and } \overline{\rho}_\lambda \text{ is reducible}\}$$

is finite.

First, notice that we can apply Theorem 183 to characterize the representation because by hypothesis $\overline{\rho}_\lambda$ is reducible. In particular, combining it with the additional information we have about $\overline{\rho}_\lambda$ due to Theorem 172 we get that

$$\begin{aligned} a_p = \mathrm{trace}(\rho_\lambda(\mathrm{Frob}_\mathfrak{p})) &\equiv \epsilon_1(p) + \epsilon_1^{-1}(p)p \pmod{\lambda} \\ p = \det(\rho_\lambda(\mathrm{Frob}_\mathfrak{p})) &\equiv \epsilon_1(p)\epsilon_1^{-1}(p)p \pmod{\lambda}. \end{aligned} \tag{5.1}$$

Thus, merging congruences 5.1 with the results on the bounds for the conductors done by Carayol in [Car89] and Livné in [Liv89], we obtain the following statement.

Proposition 187. *Let l be a prime such that $l \geq 3$ and $l \nmid N$, and let $\lambda \in \Lambda$. For any prime $p \nmid lN$, we have that*

$$a_p \equiv \epsilon(p) + \epsilon^{-1}(p)p \pmod{\lambda},$$

where ϵ is a Dirichlet character unramified outside N whose conductor c satisfies $c^2|N$.

From there, if Λ is infinite, using the bound on the conductors we can pick an infinite subset Λ_c of Λ such that c is the conductor for all $\lambda \in \Lambda_c$. Then, we consider a prime p such that

$$p \equiv 1 \pmod{c},$$

which automatically gives us the relation

$$a_p \equiv 1 + p \pmod{\lambda}$$

for all $\lambda \in \Lambda_c$.

Hence, since Λ_c is infinite, the congruences 5.3.2 hold for infinitely many values of λ, so they become equalities. However, this contradicts the bounds on the absolute values of the a_p from Theorem 132.

Therefore, Λ must be finite, which automatically proves part (a) as desired.

5.3.3 On the groups of order not divisible by l

To prove part (b) we follow the same philosophy as in part (a), i.e. we simply check that there are only finitely many primes $\lambda \in \mathcal{O}$ for which the order of $\overline{G}_\lambda$ is not divisible by l. In that sense, since part (a) is already proven, we can also assume that $\overline{\rho}_\lambda$ is irreducible. In other words, our goal is to prove that

$$\Lambda' = \{\lambda \in \mathcal{O} : \lambda \text{ is a prime, } \overline{\rho}_\lambda \text{ is irreducible, and the order of } \overline{G}_\lambda \text{ is not divisible by } l\}$$

is finite.

In this set up, as mentioned earlier, the strategy of the proof mainly relies in applying Dickson's result from Theorem 163. In particular, combining it with our current hypothesis we obtain the following theorem:

Theorem 188. *Let $G \in \mathrm{GL}(2, \mathbb{F}_\lambda)$ be an irreducible subgroup whose order is not divisible by l, and H its image in $\in \mathrm{PGL}(2, \mathbb{F}_\lambda)$. Then, one of the following holds:*

1. *H is cyclic and G is contained in a Cartan subgroup.*
2. *H is dihedral and G is contained in the normalizer of a Cartan subgroup but not in a Cartan subgroup.*
3. *H is isomorphic to one of the following special groups: $A_4,\ S_4,\ A_5$.*

In this framework, we first denote P_λ to be the projective image of $\overline{G}_\lambda$, i.e. its image in $\mathrm{PGL}(2, \mathbb{F}_\lambda)$. Then, to prove that Λ' is finite it suffices to check that $\overline{G}_\lambda$ and P_λ satisfy each of the cases from Theorem 188 only for finitely many λ.

The cyclic case

We start by treating the case where P_λ is cyclic and $\overline{G}_\lambda$ is contained in a Cartan subgroup denoted by C_λ.

Let $c \in G_\mathbb{Q}$ be a complex conjugation. By definition, c has order 2 and we recall that $\det(\overline{\rho}_\lambda)$ is the mod l cyclotomic character. This implies that $\det(\overline{\rho}_\lambda(c)) = -1$ because the complex conjugate of any root of unity is its reciprocal. From there, we get that $\overline{\rho}_\lambda(c)$ has ± 1 as eigenvalues, i.e. it has different rational eigenvalues. Then, Proposition 170 tells us that C_λ can not be a non-split Cartan subgroup, so it must be a split Cartan. However, split Cartan groups correspond, up to conjugacy, to diagonal matrices. Hence, by Remark 11 we get that ρ_λ is a reducible representation contradicting the initial hypothesis.

Therefore, we conclude that the cyclic case does not hold for any $\lambda \in \Lambda'$.

The dihedral case

In this case, we suppose that there is an infinite subset $\Lambda'_D \subseteq \Lambda'$ such that for all $\lambda \in \Lambda'_D$ we have that P_λ is a dihedral group and $\overline{G}_\lambda$ is contained in the normalizer N_λ of a Cartan subgroup C_λ but not in C_λ.

First, by Remark 162 we can construct the character

$$\alpha_\lambda : G_\mathbb{Q} \to \{\pm 1\},$$

which consists in the composition

$$\alpha_\lambda : G_\mathbb{Q} \xrightarrow{\overline{\rho}_\lambda} \overline{G}_\lambda \subset N_\lambda \to N_\lambda / C_\lambda \cong \{\pm 1\}. \tag{5.2}$$

From there, notice that kernel of α_λ is an open subgroup of $G_\mathbb{Q}$ of order 2, so by the Galois correspondence from Theorem 42 we can construct the associated fixed quadratic field K_λ. In addition, K_λ is unramified at lN because $\overline{\rho}_\lambda$ is by hypothesis.

At this point, the next step is to prove the following theorem.

Theorem 189. *There exists an infinite subset $\Lambda'_K \subseteq \Lambda'_D$ such that the quadratic field K_λ is the same for all $\lambda \in \Lambda'_K$.*

Proof. First, notice that to prove this statement it suffices to prove that for almost any $\lambda \in \Lambda'_D$ the subextension K_λ does not ramify at l. This is because by construction K_λ must ramify only in a subset of the finite set of primes in which the original representation ramifies. Hence, if l is not in this subset, the subset of candidate primes in which K_λ may ramify is finite and does not depend on λ, so by Theorem 26 there are only finitely many possible K_λ and since Λ'_D is infinite by hypothesis there must exist $\Lambda'_K \subseteq \Lambda'_D$ such that K_λ is the same for all $\lambda \in \Lambda'_K$.

Due to the previous argument, we are left to prove that K_λ does not ramify at l. To do so, we must take a look at the image of $\overline{\rho}_\lambda$ in I_λ. In that sense, notice that since by hypothesis we are assuming that the order of $\overline{G}_\lambda$ is not divisible by l, the image of the inertia does not have elements of order l, which implies that it cannot contain matrices of the form

$$\begin{bmatrix} 1 & * \\ 0 & 1 \end{bmatrix} \text{ with } * \text{ different from } 0.$$

Thus, this means that it corresponds to a semisimple representation and by Proposition 186 we have that the projective image of $\overline{\rho}_\lambda$ in I_λ, which we denote by $\overline{I}_\lambda$, is cyclic of order $l \pm 1$. Then, by taking $l \geq 5$ we get that the order is strictly larger than 2, so we can apply Proposition 168 to get that $\overline{I}_\lambda$ is contained in the core of the dihedral group.

Finally, this implies that the inertia is trivial in K_λ for almost every $\lambda \in \Lambda'_D$, i.e. K_λ does not ramify at l if $l \geq 5$, which completes the proof. □

Now, for all $\lambda \in \Lambda'_K$ we denote K to the fixed quadratic field K_λ, and α to the quadratic character α_λ associated to K defined in (5.2).

Having said that, we simply need to combine the definition of α in (5.2) with Proposition 171 to get that for all $\lambda \in \Lambda'_K$ we have that given any $g \in G_{\mathbb{Q}}$

1. If $\overline{\rho}_\lambda(g) \in C_\lambda$, then $\alpha(g) = 0$.
2. If $\overline{\rho}_\lambda(g) \in N_\lambda \setminus C_\lambda$, then $\alpha(g) = -1$ and $\operatorname{trace}(\overline{\rho}_\lambda(g)) = 0$.

Then, notice that these two conditions together with the description of the initial representation from Theorem 172 imply the following statement.

Theorem 190. *For every $\lambda \in \Lambda'_K$ and every prime $p \nmid lN$ we have that*

$$\alpha(\mathrm{Frob}_\mathfrak{p})a_p \equiv a_p \pmod{\lambda}.$$

Finally, the infinity of Λ'_K in Theorem 190 implies that the congruence of the theorem becomes an equality. However, according to Definition 139 this means that the original newform has complex multiplication, which contradicts our initial hypothesis.

Hence, Λ'_D must be finite as desired.

The special groups case

We now suppose that P_λ is isomorphic to either A_4, S_4, or A_5.

As in the previous case, the key point to prove that this situation can only happen for finitely many values of λ is to use the characterization of the projective image of the semisimplification of $\overline{\rho}_\lambda$ in I_λ, which we denote again by $\overline{I}_\lambda$.

In particular, Proposition 186 tells us that $\overline{I}_\lambda$ corresponds to a cyclic subgroup of order $l \pm 1$. Hence, arguing on the order of the groups A_4, S_4, and A_5 we can apply Proposition 164 to conclude that the maximum order among them corresponds to $|A_5| = 60$, which implies that for $l \geq 67$ the group $\overline{I}_\lambda$ can not be included in neither of these groups contradicting the initial hypothesis.

However, we can go a step further to lower even more the bound on l. To do so, we simply need to apply Proposition 169 to conclude that neither A_4, S_4, nor A_5 have cyclic subgroups of order ≥ 6. Hence, this case can not hold for any $l \geq 7$.

Grouping the cases

By taking a detailed look at each of the three cases, we have concluded that each of them hold at most for a finite number of λ. Thus, the set Λ' for which Theorem 188 holds is finite, which immediately proves part (b) of Theorem 181 as desired.

5.4 Applications & consequences of Ribet's theorem

Once we have proved Ribet's theorem, a natural question to ask is what consequences follow from its statement. In this regard, it turns out that we can use it to realize projective linear groups over finite fields as Galois groups over $\mathbb{Q}$. Hence, the aim of this section is to outline this application of Ribet's theorem.

Recall that we denote by P_λ the image if $\overline{G}_\lambda$ in $\mathrm{PGL}(2,\mathbb{F}_\lambda)$. With this in mind, the following theorem directly realises $\mathrm{PGL}(2,\mathbb{F}_{q^n})$ and $\mathrm{PSL}(2,\mathbb{F}_{q^n})$ as Galois groups over $\mathbb{Q}$.

Theorem 191. *Let l be a prime satisfying Ribet's theorem and $\lambda \in \mathcal{O}$ be a prime over l. Then*

$$P_\lambda = \begin{cases} \mathrm{PGL}(2,\mathbb{F}_\lambda) \text{ if } [\mathbb{F}_\lambda : \mathbb{F}_l] \text{ is odd} \\ \mathrm{PSL}(2,\mathbb{F}_\lambda) \text{ if } [\mathbb{F}_\lambda : \mathbb{F}_l] \text{ is even} \end{cases}.$$

Proof. First of all, notice that the equality $G_l = A_l$ from Theorem 175 implies that

$$\overline{G}_\lambda = \{x \in \mathrm{GL}(2,\mathbb{F}_\lambda) : \det(x) \in \mathbb{F}_\lambda^\times\},$$

so it is clear that $\mathrm{PSL}(2,\mathbb{F}_\lambda) \subset P_\lambda$. In addition, since l is odd, by Proposition 149 we get that $[\mathrm{PGL}(2,\mathbb{F}_\lambda) : \mathrm{PSL}(2,\mathbb{F}_\lambda)] = 2$. Hence, P_λ is either $\mathrm{PGL}(2,\mathbb{F}_\lambda)$ or $\mathrm{PSL}(2,\mathbb{F}_\lambda)$.

Now, we simply need to take a look at the determinant in order to distinguish between both cases. In this regard, by construction the determinant of $\mathrm{PGL}(2,\mathbb{F}_\lambda)$ corresponds to $\mathbb{F}_\lambda^*/(\mathbb{F}_\lambda^*)^2$, so the projective image of any $x \in \mathrm{GL}(2,\mathbb{F}_\lambda)$ lies in $\mathrm{PSL}(2,\mathbb{F}_\lambda)$ if and only if $\det(x)$ is a square in $\mathbb{F}_\lambda$.

With this in mind, we move to split between cases depending on the parity of the extension.

On the one hand, if $[\mathbb{F}_\lambda : \mathbb{F}_l]$ is even the quadratic extension of $\mathbb{F}_l$ is indeed a subextension of $\mathbb{F}_\lambda/\mathbb{F}_l$. Then by the uniqueness of this quadratic extension we can assure that all elements in $\mathbb{F}_l$ are squares on it, so they are also squares in $\mathbb{F}_\lambda$. Hence, by the previous argument on the determinant we are done in this case.

On the other hand, if $[\mathbb{F}_\lambda : \mathbb{F}_l]$ is odd the quadratic extensions $\mathbb{F}_l$ can not be a subextension of $\mathbb{F}_\lambda/\mathbb{F}_l$. Therefore, not all the elements of $\mathbb{F}_l$ are squares in $\mathbb{F}_\lambda$, so in this case it is clear that $P_\lambda = \mathrm{PSL}(2,\mathbb{F}_\lambda)$. □

Once Theorem 191 is proved, simply by applying it we are able to realize multiple projective linear groups over finite fields as Galois groups over $\mathbb{Q}$. Nevertheless, we could argue that this result only holds for the primes l satisfying Ribet's theorem, the statement of which does not explicitly specify them.

In this regard, there are two main considerations to be taken into account. First, notice that Ribet's theorem holds for almost every prime, so the chances of picking a prime for which it does not hold are low. Moreover, by looking at the conditions required in Theorem 176 we can give a criteria for determining the desired primes. Indeed, in [Die10], Section 2.2. Dieulefait presented an algorithm for computing the exceptional primes in Ribet's theorem. Broadly speaking, the algorithm simply consists in mimicking the steps during the proof of Theorem 181 in order to detect these primes.

Chapter 6

Conclusions

Throughout this thesis we have studied the images of Galois representations attached to modular forms without Complex Multiplication. In particular, we have focused our attention into develop all the theory needed to understand Theorem 175 from Ribet, which sates that the images of this kind of representations are as large as possible for almost every prime.

Going into details, we have devoted chapters 2 and 3 to introduce the topics of Galois representations and modular forms respectively, which are the main characters of Ribet's theorem. From one hand, when studying Galois representations we have seen several notions as well as some important results such as the Chebotarev density theorem. In addition, we have treated other notions from algebraic number theory such as the l-adic numbers, ramification of extensions, and group characters. On the other hand, in the modular forms chapter we have presented and motivated multiple concepts as cusp forms and congruence subgroups. Moreover, we have introduced all the theory and intuition behind Hecke operators and also the notions of newforms, complex multiplication, and inner twists.

Once the needed background has been introduced, we have targeted Ribet's theorem about the images of Galois representations attached to newforms without complex multiplication. To do so, we first have discussed about multiple results on group theory, specially regarding general linear groups. In that sense, we have put emphasis on Dickson's classification of subgroups of $\mathrm{GL}(2, \mathbb{F}_q)$, which is an essential tool for the proof of Ribet's theorem. After that, we have focused on the main theorem on its own by explaining the Eichler-Shimura relation, the statement itself, a detailed proof developing all steps left by Ribet in [Rib85], and commenting some of its main applications.

To conclude, we propose a series of research lines to continue this work. The first one corresponds to generalize the argument from Ribet to newforms of any weight instead of simply assuming weight 2. This is something that could be done using similar strategies based on results as the one in [FJ95]. Moreover, the generalization can also be done in another direction, which corresponds to take a look at the case where inner twists are considered. In particular, this leads to a slightly more complex characterization of G_l that can be found in [Die01]. Finally, the other alternative is to study in detail the algorithm for computing the exceptional primes in Ribet's theorem mentioned at the end of Chapter 5.

Bibliography

[AM18] Daniel A. Marcus, *Number fields*, Universitext, Springer, New York, NY, 2018.

[Bor91] Armand Borel, *Linear algebraic groups*, Springer, 1991.

[Car89] Henri Carayol, *Sur les représentations galoisiennes modulo l attachées aux formes modulaires*, Duke Math. J. **59** (1989), no. 3, 785–801.

[Ded82] R. Dedekind, *Ueber die discriminanten endlicher körper*, Abhandlungen der Königlichen Gesellschaft der Wissenschaften in Göttingen **29** (1882), 1–56.

[Del71] Pierre Deligne, *Formes modulaires et représentations ℓ-adiques*, Séminaire Bourbaki **11** (1971), 139–172, Talk no. 355.

[Dia05] Fred Diamond, *A first course in modular forms*, Graduate texts in mathematics; 228, Springer, New York, NY, 2005.

[Dic58] Leonard Eugene Dickson, *Linear groups with an exposition of the galois field theory*, Dover Publications, New York, 1958.

[Die01] Luis Victor Dieulefait, *Modular galois realizations of linear groups*, Ph.D. thesis, Universitat de Barcelona, 2001.

[Die10] ———, *Representaciones de galois*, Available at `sweet.ua.pt/apacetti/cursos/Final.pdf`, 2010.

[FJ95] Gerd Faltings and Bruce Jordan, *Crystalline cohomology and* $\mathrm{GL}(2,\mathbb{Q})$, Israel J. Math. **90** (1995), no. 1-3, 1–66.

[KKS95] Kazuya Kato, Nobushige Kurokawa, and Takeshi Saito, *Number theory 1: Fermat's dream*, first ed., Translations of Mathematical Monographs, vol. 1, American Mathematical Society, 1995.

[KR01] Chandrashekhar Khare and C. S. Rajan, *The density of ramified primes in semisimple p-adic galois representations.*, Internat. Math. Res. Notices **12** (2001), 601–607.

[Liv89] Ron Livné, *On the conductors of mod l galois representations coming from modular forms*, J. Number Theory **31** (1989), no. 2, 133–141.

[LS94] H.W. Lenstra and P. Stevenhagen, *Chebotarev and his density theorem*, Available at `https://pub.math.leidenuniv.nl/~lenstrahw/PUBLICATIONS/1994c/art.pdf`, 1994.

[Mas98] Heinrich Maschke, *Ueber den arithmetischen charakter der coefficienten der substitutionen endlicher linearer substitutionsgruppen*, Math. Ann. **50** (1898), no. 4, 492–498.

[Mil22] James S. Milne, *Fields and galois theory (v5.10)*, 2022, Available at www.jmilne.org/math/, p. 144.

[Ray74] Michel Raynaud, *Schémas en groupes de type* $(p, \dots, p)$, Bull. Soc. Math. France **102** (1974), 241–280.

[Rib75] Kenneth Ribet, *On l-adic representations attached to modular forms*, Invent. Math. **28** (1975), 245–275.

[Rib77] Kenneth A. Ribet, *Galois representations attached to eigenforms with nebentypus*, 1977.

[Rib80] ———, *Twists of modular forms and endomorphisms of abelian varieties*, Math. Ann. **253** (1980), 43–62.

[Rib85] K.A. Ribet, *On l-adic representations attached to modular forms ii*, Glasgow Math. J. **27** (1985), 185–194.

[Ser72] Jean-Pierre Serre, *Propriétés galoisiennes des points d'ordre fini des courbes elliptiques*, Invent. Math. **15** (1972), 259–331.

[Ser73] ———, *A course in arithmetic*, Graduate texts in mathematics; 7, Springer, New York, NY, 1973.

[Shi71] Goro Shimura, *Introduction to the arithmetic theory of automorphic functions*, Publications of the Mathematical Society of Japan, no. 11, Mathematical Society of Japan, Princeton University Press, Tokyo, Japan and Princeton, NJ, 1971.

[Shi97] ———, *Abelian varieties with complex multiplication and modular functions*, Princeton University Press, Princeton, NJ, 1997.

[WCR06] Charles W. Curtis and Irving Reiner, *Representation theory of finite groups and associative algebras*, Providence, AMS Chelsea Publishins, 2006.