INAUGURAL-DISSERTATION

submitted

to the

Combined Faculty of Mathematics, Engineering and Natural Sciences

of

Ruprecht-Karls-University

Heidelberg

for the Degree of Doctor of Natural Sciences

Put forward by

Alireza Shavali, M.Sc.

from Ahvaz, Iran

Day of the oral exam:

On the Image of Automorphic Galois Representations

Supervisor: Prof. Dr. Gebhard Böckle

Abstract

Under the Langlands philosophy, there should be a correspondence between certain automorphic representations of GL_n, certain n-dimensional Galois representations, and motives over number fields. There is a folklore heuristic that the image of the Galois representation should be as big as possible unless there is an automorphic reason for it not to be. In this thesis, we will formulate a precise conjecture in this direction, assuming some standard conjectures in the literature. In the $n = 2$ case, this follows from the work of Ribet, Momose, and Nekovář. We are able to prove this conjecture unconditionally in the $n = 3$ case.

More precisely, Let K be a totally real field and π be a regular algebraic cuspidal automorphic representation of $\mathrm{GL}_3(\mathbb{A}_K)$ that is of general type, i.e. it neither satisfies $\pi \simeq \pi \otimes \chi$ for a non-trivial Hecke character χ, nor $\pi \simeq \pi^\vee \otimes \eta$ for any Hecke character η. Then we show that the $\mathbb{Q}_p$-Zariski closure of the image of the p-adic Galois representation $\rho_{\pi,p}$ associated with π is of the form $(\mathrm{Res}^{F_p}_{\mathbb{Q}_p} H) \cdot \mathbb{G}_{m,\mathbb{Q}_p}$ where $F_p = F \otimes_{\mathbb{Q}} \mathbb{Q}_p$ for a certain subfield F of the Hecke field, and H/F_p is a form of SL_3.

Zusammenfassung

Gemäß der Langlands-Philosophie sollte es eine Korrespondenz zwischen bestimmten automorphen Darstellungen von GL_n, bestimmten n-dimensionalen Galoisdarstellungen und Motiven über Zahlkörpern geben. In der Fachwelt ist eine heuristische Faustregel geläufig, die besagt, dass das Bild der Galoisdarstellung möglichst groß sein sollte, es sei denn, es gibt einen automorphen Grund, der dem entgegensteht. In dieser Arbeit werden wir eine präzise Vermutung in dieser Richtung formulieren, unter Annahme einiger in der Literatur bekannter Standardvermutungen. Im Fall $n = 2$ folgt dies aus den Arbeiten von Ribet, Momose und Nekovář. Wir können diese Vermutung im Fall $n = 3$ uneingeschränkt beweisen.

Konkret beweisen wir insbesondere das folgende Resultat: Sei K ein total reeller Zahlkörper und π eine reguläre, algebraische, kuspidale automorphe Darstellung von $\mathrm{GL}_3(\mathbb{A}_K)$ vom allgemeinen Typ, d. h. es gilt weder $\pi \simeq \pi \otimes \chi$ für einen nicht-trivialen Hecke-Charakter χ, noch $\pi \simeq \pi^\vee \otimes \eta$ für einen beliebigen Hecke-Charakter η. Dann zeigen wir, dass der $\mathbb{Q}_p$-Zariski-Abschluss des Bildes der mit π assoziierten p-adischen Galoisdarstellung $\rho_{\pi,p}$ von der Form $(\mathrm{Res}^{F_p}_{\mathbb{Q}_p} H) \cdot \mathbb{G}_{m,\mathbb{Q}_p}$ ist, wobei $F_p = F \otimes_{\mathbb{Q}} \mathbb{Q}_p$ für einen bestimmten Unterkörper F des Hecke-Körpers ist, und H/F_p eine Form von SL_3 ist.

Acknowledgement

آسایش دو گیتی، تفسیر این دو حرف است
با دوستان مروت، با دشمنان مدارا
حافظ

Peace in this world and the next, lies in these two simple words
Gratitude towards your friends, and with foes try to relate
Hafez [1]

This thesis owes its existence to so many people in my life and there is no way I can fully express in words how grateful I am for all these people or even to mention them all. So let me first express my deepest gratitude to everyone who helped me with my academic journey so far and more broadly, throughout my life.

First and foremost, the mathematical content of this dissertation owes the most to my Ph.D. advisor Gebhard Böckle, with whom I had the pleasure to study for the last four years. I am grateful to him for proposing this research topic, for the time and dedication he devotes to all his students, and for his constant support throughout my doctoral studies in Heidelberg. I could not have asked for a better supervisor.

I feel very lucky that I had the opportunity to do my Ph.D. within the framework of the Collaborative Research Centre (CRC) 326, Geometry and Arithmetic of Uniformized Structures (GAUS). Being part of this network allowed me to meet so many brilliant mathematicians from different universities within the CRC and I am very grateful for all these interactions. In particular, I would like to thank my second advisor within the CRC, Judith Ludwig, and my mentor Marius Leonhardt for many insightful conversations, and the mathematical and professional advice they have provided. I am also grateful to Andrea Conti for numerous discussions about the content of this thesis, as well as for carefully reading my preprints and offering many valuable improvements, and to

[1] translation inspired by Shahriar Shahriari

Theresa Kaiser for helping me improve the English in this thesis. I would especially like to thank everyone in our research group over the past four years — Andrea, Barinder, Giacomo, Jakob, Judith, Julian, Oguz, Paola, Peter, Sriram, and Theresa — with whom I had the pleasure of organizing many seminars and reading groups. I learned a lot from each of them.

I am also grateful to all the mathematicians with whom I discussed my research. In particular, I would like to thank Gaëtan Chenevier for patiently answering my questions, which helped simplify some of the proofs; Chun Yin Hui for reading an early draft of my preprint and suggesting an argument that strengthened the main result; and Ben Moonen for his helpful responses to my questions concerning the final chapter of this thesis.

It goes without saying that I am forever indebted to all my professors at Sharif University of Technology in Tehran, where I obtained both my bachelor's and master's degrees. It was there that my mathematical background was formed. Without a doubt, my deepest gratitude belongs to my master's advisor, Prof. Dr. Arash Rastegar, who deeply influenced my mathematical interests and viewpoint and taught me a great deal of mathematics, from the specifics of the Langlands program to different paradigms of mathematical thinking.

Finally, and most importantly, I want to thank my parents, Ali and Azam, and my siblings, Narges and Mehdi. During the past four years, I could only visit them for a few weeks each year — and being far from these people dearest to my heart was the hardest part of doing my Ph.D. in Germany. Every time I spoke with them about my studies, they were nothing but supportive and encouraging. It was undoubtedly their love and support that carried me through whenever I doubted myself along the way.

Contents

Abstract **v**

Zusammenfassung **vii**

Acknowledgement **ix**

Contents **xi**

1 Introduction **1**

1.1 Background on the Galois Side 1

1.1.1 Compatible Families of Galois Representations 1

1.1.2 p-adic Hodge Theory 3

1.1.3 p-adic Lie Groups 5

1.2 Background on the Automorphic Side 7

1.2.1 Algebraic Groups and Their Forms 7

1.2.2 Automorphic Representations 10

1.2.3 Langlands Functoriality Conjectures 12

1.3 The Problem and the Results 15

2 The Irreducibility Conjecture **21**

2.1 Harish-Chandra's Philosophy of Cusp Forms 21

2.2 The $\mathbf{GL_2}$ Case 24

2.2.1 Modular Galois Representations 24

2.2.2 Irreducibility Using L-Functions 25

2.3 The $\mathbf{GL_3}$ Case 26

2.3.1 Locally Algebraic Representations 26

2.3.2 Irreducibility Using L-Functions 29

3 Inner-Twists and 2-Dimensional Galois Representations 31
3.1 Serre's Open Image Theorem 31
3.2 Inner-Twists of Modular Forms 34
3.2.1 Classical Modular Forms 34
3.2.2 Hilbert Modular Forms 36
3.3 Explicit Formulas 39
3.3.1 Quer's Formula 39
3.3.2 Generalization to Hilbert Modular Forms 40

4 Extra-Twists and Image of Galois Representations 53
4.1 Inner and Outer Twists 54
4.2 The Lie Algebra Computations 56
4.2.1 Extra-Twists and Galois Representations 58
4.2.2 The Lie Algebra of the Image 61
4.3 Application to Automorphic Galois Representations 64
4.3.1 The $\mathbf{GL_2}$ Case 67
4.3.2 The $\mathbf{GL_3}$ Case 68
4.3.3 The $\mathbf{GL}_n$ Case 73

5 Relations to the Mumford-Tate Conjecture 77
5.1 Mumford-Tate Groups 77
5.2 Extra-Twists of a Motive 78

Bibliography 83

Chapter 1

Introduction

1.1 Background on the Galois Side

1.1.1 Compatible Families of Galois Representations

Let K be a number field, $\Gamma_K := \mathrm{Gal}(\overline{K}/K)$ its absolute Galois group and p a prime number. By a p-adic (Galois) representation of Γ_K, we mean a continuous group homomorphism

$$\rho : \Gamma_K \to \mathrm{GL}_n(\overline{\mathbb{Q}}_p),$$

where Γ_K is considered with its topology as a profinite group (Krull topology) and $\mathrm{GL}_n(\overline{\mathbb{Q}}_p)$ with its canonical topology induced from that of $\overline{\mathbb{Q}}_p$. Moreover, for a number field $E \subset \overline{\mathbb{Q}}_p$, the representation ρ is called E-rational if it is unramified outside a finite set of places of K and for each such (finite) place, the characteristic polynomial of the Frobenius element has coefficients in E.

Definition 1.1.1. Let K and E be number fields and S be a finite set of finite places of K. For each prime number p, let S_p be the set of places of K above p. An E-rational n-dimensional compatible family of Galois representations of Γ_K unramified outside S is a collection of Galois representations ρ_λ for each embedding $\lambda : E \hookrightarrow \overline{\mathbb{Q}}_p$ such that

$$\rho_\lambda : \Gamma_K \to \mathrm{GL}_n(\overline{\mathbb{Q}}_p)$$

is a p-adic Galois representation and ρ_λ is unramified outside $S \cup S_p$ and for each finite place v of K outside S there exists a polynomial $f_v(x) \in E[x]$ such that for each λ and each finite place v of K outside $S \cup S_\lambda$, the characteristic polynomial of $\rho_\lambda(\mathrm{Frob}_v)$ is equal to ${}^\lambda f(x) \in \overline{\mathbb{Q}}_p[x]$, i.e. it is independent of λ.

For instance, the ℓ-adic cyclotomic characters $\{\chi_\ell\}_\ell$ form a compatible family of $\mathbb{Q}$-rational Galois representations of $\Gamma_{\mathbb{Q}}$ and the ℓ-adic Tate modules of an elliptic curve over a number field K form a compatible family of $\mathbb{Q}$-rational Galois representations of Γ_K. More generally, if X is a smooth proper variety over a number field K, then it is well-known that its ℓ-adic cohomologies $H^i_{\text{ét}}(X_{\overline{K}}, \mathbb{Q}_p)$ form a compatible family of $\mathbb{Q}$-rational Galois representations of Γ_K.

For a number field K, an abelian variety A over K is a connected projective algebraic group over K. Morphisms and isogenies of abelian varieties are simply morphisms and isogenies of algebraic groups. The set of the complex points $A(\mathbb{C})$ of an abelian variety A of dimension g, is isomorphic to $\mathbb{C}^g/\Lambda$ as a complex Lie group, where Λ is a rank $2g$ lattice in $\mathbb{C}^g$ admitting a Riemann form (polarization). This identifies the subgroup of n-torsion points of $A(\mathbb{C})$ with $\Lambda/n\Lambda$ which is isomorphic to $(\mathbb{Z}/n\mathbb{Z})^{2g}$. All these points are defined over some finite Galois extension L of K and so are endowed with an action of the absolute Galois group Γ_K which factors through $\text{Gal}(L/K)$. In other words, we get a continuous homomorphism

$$\rho_A[n] : \Gamma_K \to \text{GL}_{2g}\left(\mathbb{Z}/n\mathbb{Z}\right).$$

Now let ℓ be a prime number and $[\ell]$ be the multiplication by ℓ map on A. Then we have an inverse system

$$\cdots \xrightarrow{[\ell]} A[\ell^3] \xrightarrow{[\ell]} A[\ell^2] \xrightarrow{[\ell]} A[\ell]$$

of Γ_K-modules. Taking an inverse limit we define the ℓ-adic Tate module $T_\ell(A)$ of A as

$$T_\ell(A) = \varprojlim_n A[\ell^n],$$

which is clearly isomorphic to $\mathbb{Z}_\ell^{2g}$ and is endowed with a continuous Γ_K action. We also define the rational version of the Tate module $V_\ell(A) := T_\ell(A) \bigotimes_{\mathbb{Z}} \mathbb{Q}$, which is a $2g$-dimensional $\mathbb{Q}_\ell$-vector space. Choosing a basis for this vector space, we get an ℓ-adic Galois representation

$$\rho_{A,\ell} : \Gamma_K \to \text{GL}_{2g}(\mathbb{Q}_\ell).$$

It can be shown that the collection of $\rho_{A,\ell}$'s form a compatible family of $\mathbb{Q}$-rational Galois representations of Γ_K.

The ℓ-adic Galois representation attached to an abelian variety carries a lot of information about the abelian variety. However, it does not uniquely determine the abelian variety since one can prove that isogenous abelian varieties have isomorphic Galois representations attached to them. Faltings' isogeny theorem, which is a special case of Tate's conjecture on algebraic cycles, provides an inverse for this; It shows that the Galois representation determines the abelian variety up to isogeny. Slightly more generally, we have:

Theorem 1.1.2 (Faltings). *Let A and B be abelian varieties over the number field K. Then the natural map*

$$\mathrm{Hom}_K(A,B)\otimes_{\mathbb{Z}}\mathbb{Z}_\ell \to \mathrm{Hom}_{\mathbb{Z}_\ell[\Gamma_K]}\left(T_\ell(A),T_\ell(B)\right)$$

is an isomorphism.

It is not hard to see that this theorem implies that if the Galois representations $V_\ell(A)$ and $V_\ell(B)$ are isomorphic, then there is an isogeny $f\in \mathrm{Hom}_K(A,B)\otimes_{\mathbb{Z}}\mathbb{Q}$ inducing this isomorphism.

1.1.2 p-adic Hodge Theory

For a detailed introduction to p-adic Hodge theory, we refer the reader to [15]. Here, we only summarize the main concepts that we need in this thesis.

Let K be a finite extension of $\mathbb{Q}_p$ and let $\mathbb{C}_p$ be the completion of the algebraic closure of K. It is well-known that the action of Γ_K on $\overline{K}$ extends to a well-defined action on $\mathbb{C}_p$ by continuity. Let $\mathrm{Rep}_{\mathbb{C}_p}(\Gamma_K)$ be the category of finite dimensional $\mathbb{C}_p$-semilinear continuous representations of Γ_K. Let $\chi_p:\Gamma_K\to\mathbb{Q}_p^\times$ be the p-adic cyclotomic character and for each $V\in\mathrm{Rep}_{\mathbb{C}_p}(\Gamma_K)$ and $n\in\mathbb{Z}$, define the n'th Tate twist of V as

$$V(n):=V\otimes_{\mathbb{Q}_p}\chi_p^n.$$

A theorem of Sen and Tate tells us that the only elements of $\mathbb{C}_p$ that are fixed by the action of Γ_K are the elements of K, i.e. $\mathbb{C}_p^{\Gamma_K}=K$. Moreover, for any integer $n\neq 0$, there are no non-zero elements fixed in $\mathbb{C}_p(n)$ by Γ_K, i.e. $\mathbb{C}_p(n)^{\Gamma_K}=0$.

Now let $\mathrm{Rep}_{\mathbb{Q}_p}(\Gamma_K)$ be the category of finite dimensional $\mathbb{Q}_p$-linear continuous representations of Γ_K. For any object V of this category, $V\otimes_{\mathbb{Q}_p}\mathbb{C}_p$ is an object of the category $\mathrm{Rep}_{\mathbb{C}_p}(\Gamma_K)$. We are ultimately interested in the objects of $\mathrm{Rep}_{\mathbb{Q}_p}(\Gamma_K)$ which we call p-adic representations of Γ_K. The Tate twists $V(n)$ are defined similarly in this case.

Now we define the graded ring $B_{HT}:=\bigoplus_{n\in\mathbb{Z}}\mathbb{C}_p(n)$ with the evident addition and multiplication. Notice that this is endowed with an action of Γ_K and $B_{HT}^{\Gamma_K}=\mathbb{C}_p^{\Gamma_K}=K$.

As a graded ring we clearly have $B_{HT} \simeq \mathbb{C}_p[T, T^{-1}]$ and the Galois action turns into acting on T^n by χ_p^n.

Definition 1.1.3. A representation $V \in \mathrm{Rep}_{\mathbb{Q}_p}(\Gamma_K)$ is called Hodge-Tate, if there is a decomposition

$$V \otimes_{\mathbb{Q}_p} \mathbb{C}_p \simeq \oplus_q \mathbb{C}_p(-q)^{h_q}$$

in the category $\mathrm{Rep}_{\mathbb{C}_p}(\Gamma_K)$. The values of q for which h_q is non-zero are called the Hodge-Tate weights of V and h_q is called the multiplicity of weight q.

For example, the p-adic cyclotomic character is clearly Hodge-Tate of Hodge-Tate weight -1 which has multiplicity one. By a result of Faltings, the decomposition also holds for the p-adic cohomologies of smooth proper varieties over K. A useful reformulation of this definition is obtained using the ring B_{HT}. Namely, V is Hodge-Tate if and only if the natural map

$$\alpha_V : (V \otimes_{\mathbb{Q}_p} B_{HT})^{\Gamma_K} \otimes_K B_{HT} \to V \otimes_{\mathbb{Q}_p} B_{HT}$$

is an isomorphism. It can be shown that this map is always injective and both sides are of finite rank. Then checking the isomorphism reduces to comparing dimensions:

Proposition 1.1.4 (Tate)**.** *The representation* $V \in \mathrm{Rep}_{\mathbb{Q}_p}(\Gamma_K)$ *is Hodge-Tate if and only if* $\dim_{\mathbb{Q}_p} V = \dim_K (V \otimes_{\mathbb{Q}_p} B_{HT})^{\Gamma_K}$.

Fontaine defined more refined period rings to better understand and classify objects of the category $\mathrm{Rep}_{\mathbb{Q}_p}(\Gamma_K)$. One example of such period rings is the de Rham period ring B_{dR}. This ring is equipped with a Galois action just like B_{HT}, but instead of a grading it comes with a filtration. The semi-simplification of this filtration gives exactly B_{HT}. We omit the construction of this ring in this discussion.

For any representation $V \in \mathrm{Rep}_{\mathbb{Q}_p}(\Gamma_K)$ one can consider the map

$$\beta_V : (V \otimes_{\mathbb{Q}_p} B_{dR})^{\Gamma_K} \otimes_K B_{dR} \to V \otimes_{\mathbb{Q}_p} B_{dR}$$

of filtered B_{dR}-semilinear Galois modules and V is called a de Rham representation if and only if β_V is an isomorphism. Every de Rham representation is clearly Hodge-Tate since one can look at the graded modules of both sides.

We also need to consider more general coefficient fields. When studying automorphic Galois representations, one is naturally led to the situation where representations of Γ_K

have coefficients in an algebraic extension E of $\mathbb{Q}_p$ (where the Γ_K action on E is still considered to be trivial). In this case, we can still define the map

$$\alpha_V : (V \otimes_{\mathbb{Q}_p} B_{HT})^{\Gamma_K} \otimes_K B_{HT} \to V \otimes_{\mathbb{Q}_p} B_{HT},$$

which is now a morphism of $E \otimes B_{HT}$-modules. V is said to be Hodge-Tate if and only if α_V is an isomorphism. This is equivalent to V being Hodge-Tate when it is considered as a $\mathbb{Q}_p$ vector space by restriction of scalars. In particular, if $E = \overline{\mathbb{Q}}_p$ then for each embedding $\tau : K \to \overline{\mathbb{Q}}_p$ we define the τ-Hodge-Tate weights of V to be those integers q for which $(\mathbb{C}_p(q) \otimes_{\overline{\mathbb{Q}}_p, \tau} V)^{\Gamma_K}$ is non-zero and the dimension of this $\overline{\mathbb{Q}}_p$-vector space is the multiplicity of this weight.

We finish our discussion on p-adic Hodge theory by considering global Galois representations. Let K be a number field and consider a p-adic Galois representation

$$\rho : \Gamma_K \to \mathrm{GL}_n(\overline{\mathbb{Q}}_p).$$

We say that ρ is Hodge-Tate (respectively de Rham) if $\rho|_{K_v}$ is Hodge-Tate (respectively de Rham) for every place v of K above p. For each embedding $\tau : K \to \overline{\mathbb{Q}}_p$ we define the set of τ-Hodge-Tate weights to be that of $\rho|_{K_v}$, where v is the place induced by the embedding τ (the restriction of the absolute value of $\overline{\mathbb{Q}}_p$ to K via τ).

1.1.3 p-adic Lie Groups

For a detailed treatment of the topics in this section, we refer the reader to [36]. Here we quickly review the main points.

Throughout this section, let K be a non-Archimedean field of characteristic 0. One can define manifolds over K similarly to the case of Archimedean fields. Let M be a Hausdorff topological space. An n-dimensional (locally analytic) atlas on M is a collection of charts (U, ϕ) where U is an open subset of M and $\phi : U \hookrightarrow K^n$ identifies U with an open subset of K^n, such that these charts give a covering for M and all the transitions maps $\phi(U \cap V) \leftrightarrows \psi(U \cap V)$ for any two charts (U, ϕ) and (V, ψ) are locally analytic. The notion of equivalent atlases and a maximal atlas are defined in the evident way. There exists exactly one maximal atlas in each equivalence class which makes the following definition plausible:

Definition 1.1.5. A (locally analytic) manifold of dimension n over K is a Hausdorff topological space M equipped with an n-dimensional maximal atlas over K.

The notions of a tangent vector and tangent space can be defined similar to the Archimedean situation, using the local identifications $\phi : U \to \phi(U) \subseteq K^n$. Hence, the tangent $T_x M$ space at a point x of an n-dimensional manifold M over K is an n-dimensional K-vector space. For any locally analytic map $f : M \to N$ of manifolds over K the differential of f at $x \in M$ would then be a linear map

$$df_x : T_x M \to T_{f(x)} N.$$

Definition 1.1.6. A Lie group over K is a (locally analytic) manifold over K which carries a group structure such that the group multiplication and the inverse map are locally analytic morphisms.

Let G be a Lie group over K and $e \in G$ be the identity element. Similar to the Archimedean situation, the group action induces a (K-linear) Lie algebra structure on the tangent space $T_e G$, by identifying left-invariant vector fields on G with tangent vectors.

For any positive integer n, the group $\mathrm{GL}_n(K)$ is a Lie group over K since it can be identified with an open subset of $M_n(K) \simeq K^{n^2}$. In fact, one of the main sources of Lie groups over K are algebraic groups. If $\mathbb{G}$ is an algebraic group over K then the set of K-points $\mathbb{G}(K)$ admits a natural structure of a Lie group over K by choosing a faithful representation of $\mathbb{G}$. The Lie group structure is independent of this choice. The (analytically defined) Lie algebra of the manifold $\mathbb{G}(K)$ coincides with the (algebraically defined) Lie algebra of the algebraic group $\mathbb{G}/K$.

The Lie algebra of G encodes the infinitesimal information of neighborhoods of $e \in G$. The naive philosophy is that properties that are expected to hold only up to an open subgroup can be checked on the level of Lie algebras. The following theorem is the main tool that we will need later to apply this philosophy:

Theorem 1.1.7 ([36], Proposition 18.17)**.** *Let G_1 and G_2 be Lie groups over K and let $\sigma : \mathrm{Lie}(G_1) \to \mathrm{Lie}(G_2)$ be a homomorphism of Lie algebras. Then*

1. *There exist open subgroups $H_1 \subseteq G_1$ and $H_2 \subseteq G_2$ and a homomorphism of Lie groups $f : H_1 \to H_2$ such that $df = \sigma$.*

2. *If (H_1', H_2', f') in place of (H_1, H_2, f) also satisfies part 1, then there exists an open subgroup $H \subseteq H_1 \cap H_1'$ such that $f|_H = f'|_H$.*

We will use this theorem very often when we want to compute the image of Galois representations, sometimes without directly mentioning it. It helps us go back and forth between the groups and their Lie algebras. Let us give one example of how we will apply this result later.

Corollary 1.1.8. *Let $H \subseteq G$ be a closed subgroup of the Lie group G over K. Then H is open if and only if* $\mathrm{Lie(H)} = \mathrm{Lie}(G)$*, i.e. the embedding $\iota : H \hookrightarrow G$ induces the identity map on the Lie algebras.*

1.2 Background on the Automorphic Side

1.2.1 Algebraic Groups and Their Forms

For a more detailed overview of the theory of algebraic groups that we might use in this thesis, we refer the reader to [19, §1]. Here, we briefly review the main players and discuss the construction of forms of algebraic groups which is not discussed in [19].

Let K be a field of characteristic zero.

Definition 1.2.1. An affine algebraic group over K is an affine group scheme over K (a group object in the category of affine schemes over K) that is of finite type over K.

It can be shown (since we assumed $\mathrm{char}(K) = 0$) that every affine algebraic group over K is automatically smooth. The notion of (closed) subgroups, normal subgroups and intersection of subgroups are defined pointwise (using the functor of points) in the evident way. The connected component of the identity element G° is a normal subgroup. The kernel of a morphism of affine algebraic groups can also be defined point-wise and is always a normal subgroup of the source. Defining the quotient group is more subtle, see [19, Definition 1.3.8].

For an affine algebraic group G/K, an n-dimensional representation is a homomorphism (of algebraic groups over K) $\rho : G \to \mathrm{GL}_n$. Every algebraic group over K admits a faithful representation $\rho : G \hookrightarrow \mathrm{GL}_n$ identifying G with a closed subgroup of GL_n for some n. An element $g \in G(\overline{K})$ is called unipotent if the matrix $\rho(g) \in \mathrm{GL}_n(\overline{K})$ is unipotent for some faithful representation. This is independent of the choice of this representation. The group G is called unipotent if all the elements of $G(\overline{K})$ are unipotent. The unipotent radical $R_u(G)$ of G is defined to be the maximal connected normal unipotent subgroup of G. For example, the unipotent radical of GL_n is trivial and the unipotent radical of upper-triangular matrices is the subgroup of upper-triangular matrices with ones on the diagonal.

Definition 1.2.2. A connected affine algebraic group G over K is said to be reductive if $R_u(G_{\overline{K}}) = \{1\}$.

For an affine algebraic group G/K, we define the derived subgroup $G^{\text{der}} = G^{(1)}$ to be the intersection of all normal subgroups N of G such that G/N is abelian. Then G^{der} is normal, so it is the maximal normal subgroup with this property. Higher derived subgroups are defined similar to the case of abstract groups by taking derived subgroups consecutively:

$$G^{(n)} := \left(G^{(n-1)}\right)^{\text{der}}.$$

G is called solvable if there exists a positive integer n such that $G^{(n)} = \{1\}$. The solvable radical $R(G)$ of G is defined to be the maximal connected normal solvable subgroup of G. For example, the solvable radical of GL_n is the subgroup of scalar matrices, and the solvable radical of SL_n is trivial. In fact, for any reductive group G, the solvable radical is $Z(G)^\circ$ where $Z(G)$ is the center.

Definition 1.2.3. A connected affine algebraic group G over K is said to be semi-simple if $R(G_{\overline{K}}) = \{1\}$.

Therefore, SL_n is both semi-simple and reductive. GL_n is reductive but not semi-simple. The group of upper-triangular matrices in GL_n is neither semi-simple nor reductive, for $n > 1$. Clearly any semi-simple group is reductive. For any reductive group G, the derived subgroup G^{der} is semi-simple.

We finish this section by describing how to construct forms of algebraic groups from certain cohomology classes of the group.

Let E/F be either a finite Galois extension of fields or the semi-local Galois extension $E_p/F_p = (E/F) \otimes_{\mathbb{Q}} \mathbb{Q}_p$ for a finite Galois extension $E = F(\alpha)/F$ of number fields and let Φ be the minimal polynomial of α. Let G/E be an algebraic group and let $\Gamma = \mathrm{Gal}(E/F)$. Let $f : \Gamma \to \mathrm{Aut}_E(G)$ be a 1-cocycle and write f_σ for the image of σ. We will construct a form of G defined over F using this cocycle.

Let $G' = \mathrm{Res}_F^E G$, so for every F-algebra R we have

$$G'(R) = G(E \otimes_F R).$$

Therefore, $G'(R)$ is equipped with an action of Γ (where it acts on the first component of the tensor product) which is clearly functorial. Hence, the collection of morphisms $\sigma : G'(R) \to G'(R)$ is a natural transformation and so it is induced from a morphism

$\sigma : G' \to G'$ of algebraic groups. Therefore, we can define H to be the subgroup of G' satisfying $f_\sigma({}^\sigma g) = g$ for every $\sigma \in \Gamma$. In other words, $H = (G')^{tw_f(\Gamma)}$ where we define the f-twisted action of σ on G' to be given by ${}^{tw_f(\sigma)}g = f_\sigma({}^\sigma g)$. This gives a closed subgroup of G'.

By base changing $H \subseteq \mathrm{Res}_F^E G$ to E and then projecting to the identity component one gets

$$H_E \hookrightarrow (\mathrm{Res}_F^E G)_E = \prod_\Gamma G_E \xrightarrow{\pi_{\mathrm{id}}} G.$$

We prove that this is an isomorphism by checking this on points. First, we need to give a description of the algebraic action of Γ on $\prod_\Gamma G$ via the identification $(\mathrm{Res}_F^E G)_E = \prod_\Gamma G$. Let R be an E-algebra. Note that $\mathrm{Res}_F^E G(R) = G(E \otimes_F R)$ and the algebraic action of Γ is just the action on the E component. Now

$$G(E \otimes_F R) = G(E \otimes_F E \otimes_E R),$$

and the action is only on the first E component. Then

$$G(E \otimes_F R) = G(E \otimes_F E \otimes_E R) = G\left(E \otimes_F \frac{F[x]}{\Phi(x)} \otimes_E R\right) = G\left(\frac{E[x]}{\Phi(x)} \otimes_E R\right),$$

where the action is only on the coefficients of the first component. So

$$= G\left(\frac{E[x]}{\prod_\Gamma (x - \sigma\alpha)} \otimes_E R\right) = G\left(\left(\prod_\Gamma E\right) \otimes_E R\right) = G\left(\prod_\Gamma R\right) = \prod_\Gamma G(R),$$

where the action of $\gamma \in \Gamma$ is given by

$$(a_\sigma)_\sigma \mapsto (\gamma a_{\gamma^{-1}\sigma})_\sigma.$$

Proposition 1.2.4. *With the notation as above, for any E-algebra R the map*

$$H(R) \hookrightarrow \mathrm{Res}_F^E G(R) = \prod_\Gamma G(R) \to G(R)$$

is an isomorphism of groups. Therefore, the algebraic group H is a form of G.

Proof. By definition, $H(R)$ is the subgroup of the elements invariant under the twisted action of Γ. Let $(g_\sigma)_\sigma \in \prod G(R)$ be invariant under the twisted action:

$${}^{tw_f(\gamma)}(g_\sigma)_\sigma := (f_\gamma({}^\gamma g_{\gamma^{-1}\sigma}))_\sigma = (g_\sigma)_\sigma.$$

By looking at the component $\sigma = \gamma$ we get

$$g_\gamma = f_\gamma({}^\gamma g_1),$$

so the g_1 component determines all other g_γ's. This shows that the map

$$H(R) \hookrightarrow \prod_{\Gamma} G(R) \xrightarrow{\pi_{id}} G(R)$$

is injective.

To prove the surjectivity, we need to show that for every $g_1 \in G(R)$ the element $(f_\sigma({}^\sigma g_1))_\sigma$ is invariant under the twisted action:

$${}^{tw_f(\gamma)}(f_\sigma({}^\sigma g_1))_\sigma = (f_\gamma({}^\gamma(f_{\gamma^{-1}\sigma}({}^{\gamma^{-1}\sigma}g_1))))_\sigma = ((f_\gamma \circ {}^\gamma f_{\gamma^{-1}\sigma})({}^\sigma g_1))_\sigma.$$

Now, by the cocycle condition

$$f_\gamma \circ {}^\gamma f_{\gamma^{-1}\sigma} = f_\sigma$$

hence

$${}^{tw(\gamma)}(f_\sigma({}^\sigma g_1))_\sigma = (f_\sigma({}^\sigma g_1))_\sigma;$$

which is exactly what we needed. Therefore, the map is an isomorphism for the R-points for any R and hence an isomorphism of affine algebraic groups over E. □

We can generalize Proposition 1.2.4 to understand the behavior of H under any Galois base change of F.

Corollary 1.2.5. *Let $F \subset F_0 \subset E$ be an intermediate field that is Galois over F and let $\Gamma_0 = \mathrm{Gal}(E/F_0)$ and $f_0 : \Gamma_0 \to \mathrm{Aut}_E(G)$ be the restriction of f to Γ_0. Then*

$$H \times_F F_0 \simeq (\mathrm{Res}_{F_0}^E G)^{tw_{f_0}(\Gamma_0)}.$$

Proof. Apply Proposition 1.2.4 to the group $(\mathrm{Res}_{F_0}^E G)^{tw_{f_0}(\Gamma_0)}$ for the twisted action of the group Γ/Γ_0. □

1.2.2 Automorphic Representations

Here we summarize some of the results that we need from the theory of automorphic representations. We will mostly focus on the GL_n case which is the main case for us. Since giving the definition of an automorphic representation needs quite a lot of preparation, we will not do this here and just assume that the reader is already familiar with the theory. For further details, we refer the reader to [19].

Let K be a number field throughout this section and let $\pi = \otimes' \pi_v$ be an automorphic representation of $\mathrm{GL}_n(\mathbb{A}_K)$. Let v be a finite place of K at which π is unramified.

Then we use the notation $t_1^v(\pi), \cdots, t_n^v(\pi)$ for the Satake parameters of π_v. Recall that unramified representations are uniquely determined by their Satake parameters.

Recall that the Hecke characters of K are automorphic representations of $\mathrm{GL}_1(\mathbb{A}_K)$. For an automorphic representation π of $\mathrm{GL}_n(\mathbb{A}_K)$ and Hecke character χ of K, one can consider the tensor product $\pi \otimes_{\mathbb{C}} \chi$ of these two complex representations. It can be shown that this is again an automorphic representation which we simply denote by $\pi \otimes \chi$ from now on. If $\alpha(v) \in \mathbb{A}_K^\times$ is the element $\alpha(v) = (1, \cdots, 1, \varpi, 1, \cdots)$ where only the v-factor is not equal to 1 and ϖ is a uniformizer for O_{K_v}, then the Satake parameters of $\pi \otimes \chi$ at v for unramified v are equal to

$$\{t_1^v(\pi)\chi(\alpha(v)), t_2^v(\pi)\chi(\alpha(v)), \cdots, t_n^v(\pi)\chi(\alpha(v))\}.$$

Now assume that π is cuspidal. Strong multiplicity one tells us that the Satake parameters $\{t_1^v(\pi), \cdots, t_n^v(\pi)\}$ for all but finitely many unramified places v determine π uniquely. We will use this result many times throughout this thesis, usually without mentioning it directly. More precisely, we have:

Theorem 1.2.6 (Theorem 11.7.2 in [19])**.** *Let π and π' be cuspidal automorphic representations and S be a finite set of primes. Assume that for each place $v \notin S$ one has $\pi_v \simeq \pi'_v$. Then $\pi \simeq \pi'$.*

We will also need the basic properties of the Rankin-Selberg L-functions at some point. Let π and π' be two irreducible admissible representations of $\mathrm{GL}_n(\mathbb{A}_K)$. Then one can define the Rankin-Selberg L-function $L(s, \pi \times \pi')$ as in [19, §11.7] by taking the product of local Rankin-Selberg L-functions. In general, it is not even clear if this L-function should converge. But, if we assume that π and π' are cuspidal automorphic representations, then many nice analytic properties of these L-functions are known as is summarized in the next statement:

Theorem 1.2.7 (Theorem 11.7.1 in [19])**.** *Let π and π' be unitary cuspidal automorphic representations. Then $L(s, \pi \times \pi')$ admits a meromorphic continuation to the whole complex plane with the only possible poles being simple poles at $s = 0, 1$. There are poles at $s = 0$ and $s = 1$ if and only if $m = n$ and $\pi \simeq \pi'^\vee$.*

One defines the L-function associated with an irreducible admissible representation π as $L(s, \pi) := L(s, \pi \times 1)$. When π is cuspidal automorphic, the last theorem implies that $L(s, \pi)$ has nice analytic properties.

Now assume that π is a regular (C-)algebraic cuspidal automorphic representation. We also need to address questions regarding the rationality of π. For more details and the proofs, we refer the reader to [11]. Let π^f be the finite part of π, i.e. $\pi^f = \otimes'_{v\nmid\infty}\pi_v$ is a complex representation of $\mathrm{GL}_n(\mathbb{A}_K^\infty)$. We can twist the complex structure of π^f with the automorphism σ and denote this representation by ${}^\sigma\pi^f$.

Definition 1.2.8. The field of rationality, or the Hecke field, of π is the fixed field of π^f, i.e. the smallest subfield $\mathbb{Q}(\pi)$ of $\mathbb{C}$ such that for every $\sigma \in \mathrm{Aut}(\mathbb{C}/\mathbb{Q}(\pi))$ one has ${}^\sigma\pi^f \simeq \pi^f$.

Clozel proves that under the regularity assumption, $\mathbb{Q}(\pi)$ is always a number field [11]. One can show that there exists a (unique) algebraic automorphic representation (which would be automatically regular and cuspidal) ${}^\sigma\pi$ whose finite part is ${}^\sigma\pi^f$ (see [11, Theorem 3.13]). Then we would also have ${}^\sigma\pi \simeq \pi$ for all $\sigma \in \mathrm{Aut}(\mathbb{C}/\mathbb{Q}(\pi))$ by strong multiplicity one. Clozel also proves that (assuming regularity) the representation π^f has a model over $\mathbb{Q}(\pi)$, i.e. there exists a representation V of $\mathrm{GL}_n(\mathbb{A}_K^\infty)$ over $\mathbb{Q}(\pi)$ such that $\pi^f \simeq V \otimes_{\mathbb{Q}(\pi)} \mathbb{C}$.

1.2.3 Langlands Functoriality Conjectures

Here we quickly review the Langlands functoriality conjecture. For further details and more precise statements we refer to [19]. Fix a prime number p.

Let G be a reductive group over a field K of characteristic zero. Let $T_{\overline{K}} \subseteq G_{\overline{K}}$ be a maximal torus and let $\Psi := (X^*, X_*, \Phi, \Phi^\vee)$ be the root datum associated with $(G_{\overline{K}}, T_{\overline{K}})$. Then the dual root datum $\Psi^\vee := (X_*, X^*, \Phi^\vee, \Phi)$ gives a split reductive group over $\mathbb{Q}$, which we call the Langlands dual group $\widehat{G}$. Assume that K is a global filed for now. If G is a split group (or an inner form of a split group) then we define the L-group of G as

$$ {}^LG := \widehat{G}(\overline{\mathbb{Q}}_p) \times \Gamma_K, $$

where Γ_K is the absolute Galois group of K. In the non-split case, one fixes a base for the root datum (a Borel subgroup containing T) and then a pinning for this root datum to get a non-trivial action of Γ_K on $\widehat{G}$ and then defines

$$ {}^LG := \widehat{G}(\overline{\mathbb{Q}}_p) \rtimes \Gamma_K $$

with respect to this action. For more details see [19, §7.3].

If K is a local field instead, then we define

$$^{L}G := \widehat{G}(\overline{\mathbb{Q}}_p) \rtimes W_K,$$

where W_K is the Weil group of K.

Let K be a number field and G be a reductive group over K. Then the global Langlands correspondence predicts a relation between certain automorphic representations of $G(\mathbb{A}_K)$ and certain Galois representations (more precisely, L-parameters) with values in the L-group of G. To have a fully satisfactory formulation of this philosophy, one needs the existence of a hypothetical group called the Langlands group which has yet to be constructed. We will not discuss these issues here and restrict ourselves to the representations of the Galois group.

Analogously, in the local setting when K is a p-adic field, one expects a connection between smooth admissible representations of the p-adic group $G(K)$ and certain Galois or Weil-Deligne representations (or rather L-parameters) over K. This is not a one-to-one correspondence but to each L-parameter, one should be able to associate a finite set of smooth admissible representations called an L-packet. Just like in the global setting, these local L-parameters have values in the L-group of G. This suggests that if one has a map from the L-group of one group to another, there should be a "natural" way of transferring the representation theory of one group to the other. This is known as the functoriality principle in the Langlands program.

Now, let H and G be two reductive groups over a number field or a p-adic field K and assume that G is quasi split. An L-map is a continuous group homomorphism

$$r : {}^{L}H \to {}^{L}G$$

commuting with projection to Γ_K or W_K such that its restriction to $\widehat{H}$ comes from a map of algebraic groups $r_0 : \widehat{H} \to \widehat{G}$.

Now, let K be a p-adic field and r be an L-map as above. The local Langlands functoriality predicts that there should be a natural way of transferring irreducible admissible representations of $H(K)$ to irreducible admissible representations of $G(K)$ (or rather L-packets of these representations), compatible with some expected behaviors and constructions (e.g. L-functions). More precisely, to each irreducible admissible representations π of $H(K)$, one can associate an L-packet of irreducible admissible representations of $G(K)$, i.e. a finite set $\{\Pi_i\}$ of irreducible admissible representations of $G(K)$ with the same L-parameter. In particular, if we let $H = \{1\}$ then we (partially) recover the local Langlands correspondence from the local Langlands functoriality.

Now, let K be a number field and r be an L-map as above. Let $\pi = \otimes' \pi_v$ be an automorphic representation of $H(\mathbb{A}_K)$. Each π_v is an irreducible admissible representation of the local group $H(K_v)$ and by local Langlands functoriality it gives a finite set $\{\Pi_{v,i}\}$ of irreducible admissible representations of $G(K_v)$. Then one can form (possibly infinitely many) representations of $G(\mathbb{A}_K)$ by taking the restricted tensor product of these local representations. This gives a global L-packet. The global Langlands functoriality predicts that this L-packet is automorphic, i.e. at least one of these tensor products must be an automorphic representation.

In other words, one should be able to transfer automorphic representations (or rather automorphic L-packets) of H to G in a way compatible with the local Langlands functoriality. Again taking H to be trivial recovers the global Langlands correspondence.

Apart from the general philosophy, there are two specific known cases of the Langlands functoriality that we will use in this thesis. The second symmetric power of the standard representation of GL_2 gives a homomorphism $\mathrm{sym}^2 : \mathrm{GL}_2 \to \mathrm{GL}_3$ and since GL_n splits it clearly gives an L-map

$$\mathrm{sym}^2 : {}^L\mathrm{GL}_2 \to {}^L\mathrm{GL}_3.$$

Over any number field K, the Langlands transfer is constructed for this L-map in the work of Gelbart and Jacquet [18].

Other important cases of functoriality are automorphic base change and automorphic induction. Let L/K be a finite Galois extension of number fields. Then we have

$${}^L\mathrm{Res}_K^L\mathrm{GL}_n = \left(\mathrm{Res}_K^L\mathrm{GL}_n\right)(\overline{\mathbb{Q}}_p) \rtimes \Gamma_K = \mathrm{GL}_n(\overline{\mathbb{Q}}_p)^{\mathrm{Gal}(L/K)} \rtimes \Gamma_K,$$

where Γ_K acts on $\mathrm{GL}_n(\overline{\mathbb{Q}}_p)^{\mathrm{Gal}(L/K)}$ via its action on $\mathrm{Gal}(L/K)$ coming from the surjection $\Gamma_K \to \mathrm{Gal}(L/K)$. Now we can embed GL_n diagonally in this product and we get

$$\mathrm{Res}_K^L : {}^L\mathrm{GL}_n \to {}^L\mathrm{Res}_K^L\mathrm{GL}_n.$$

The (conjectural) functoriality transfer corresponding to this map is known as automorphic base change. Since $\mathrm{Res}_K^L\mathrm{GL}_n(\mathbb{A}_K) = \mathrm{GL}_n(\mathbb{A}_L)$, this transfer (conjecturally) base changes an automorphic representation of $\mathrm{GL}_n(\mathbb{A}_K)$ to an automorphic representation of $\mathrm{GL}_n(\mathbb{A}_L)$. On the Galois side of the Langlands correspondence, this amounts to restricting a Galois representation of Γ_K to a representation of the subgroup Γ_L.

The standard action of $\mathrm{GL}_n(\overline{\mathbb{Q}}_p)$ on $\overline{\mathbb{Q}}_p^n$ and the evident action of $\mathrm{Gal}(L/K)$ on $\left(\overline{\mathbb{Q}}_p^n\right)^{\mathrm{Gal}(L/K)}$ give an action of ${}^L\mathrm{Res}_K^L\mathrm{GL}_n$ on $\left(\overline{\mathbb{Q}}_p^n\right)^{\mathrm{Gal}(L/K)}$. This is an $n[L : K]$ dimensional vector space over $\overline{\mathbb{Q}}_p$ so this action induces an L-map

$$\mathrm{Ind}_K^L : {}^L\mathrm{Res}_K^L\mathrm{GL}_n \to {}^L\mathrm{GL}_{n[L:K]}$$

whose (conjectural) functorial transfer is called automorphic induction. This transfer (conjecturally) takes an automorphic representation of $\mathrm{GL}_n(\mathbb{A}_L)$ and gives an automorphic representation of $\mathrm{GL}_{n[L:K]}(\mathbb{A}_K)$. On the Galois side of the Langlands correspondence, this amounts to induction of a Galois representation from the smaller group Γ_L to the larger group Γ_K.

In the case that $[L : K]$ is a prime number, automorphic base change and induction are known in many cases by the seminal work of Arthur and Clozel [1]. We state the results from [19, §13.4]:

Theorem 1.2.9 (Base change)**.** *Let L/K be a prime degree Galois extension of number fields and θ be a generator for the (cyclic) group $\mathrm{Gal}(L/K)$. Then, for every cuspidal automorphic representation π of $\mathrm{GL}_n(\mathbb{A}_K)$, the base change $\pi_L = \mathrm{Res}_K^L(\pi)$ exists and is an isobaric automorphic representation of $\mathrm{GL}_n(\mathbb{A}_L)$ such that $\pi_L^\theta \simeq \pi_L$. Moreover, π_L is cuspidal if and only if $\pi \not\simeq \pi \otimes \eta$ for all Hecke characters*

$$\eta : K^\times \backslash \mathbb{A}_K^\times / \mathrm{Nr}_{L/K}(\mathbb{A}_L^\times) \to \mathbb{C}^\times.$$

Conversely, if a cuspidal automorphic representation π' of $\mathrm{GL}_n(\mathbb{A}_L)$ satisfies $\pi'^\theta \simeq \pi'$, then $\pi' \simeq \pi_L$ for some cuspidal automorphic representation π of $\mathrm{GL}_n(\mathbb{A}_K)$.

Theorem 1.2.10 (Induction)**.** *Let L/K be a prime degree Galois extension of number fields and θ be a generator for the (cyclic) group $\mathrm{Gal}(L/K)$. If π is a cuspidal automorphic representation of $\mathrm{GL}_n(\mathbb{A}_L)$, then the automorphic induction $I(\pi) = \mathrm{Ind}_K^L(\pi)$ exists and is an isobaric automorphic representation of $\mathrm{GL}_{n[L:K]}(\mathbb{A}_K)$. The representation $I(\pi)$ is cuspidal if and only if $\pi \not\simeq \pi^\theta$. Moreover, a cuspidal automorphic representation π' of $\mathrm{GL}_{n[L:K]}(\mathbb{A}_K)$ is the automorphic induction of a cuspidal automorphic representation of $\mathrm{GL}_n(\mathbb{A}_L)$ if and only if $\pi' \simeq \pi' \otimes \eta$ for some non-trivial character $\eta : K^\times \backslash \mathbb{A}_K^\times / \mathrm{Nr}_{L/K}(\mathbb{A}_L^\times) \to \mathbb{C}^\times$.*

1.3 The Problem and the Results

In this section we will discuss the problem that we are going to study in this thesis. One can argue that the main goal of arithmetic geometry is to understand varieties defined over fields of arithmetic significance, e.g., number fields. This could mean understanding their rational points, reductions, different cohomologies, etc. Many of these arithmetic properties are encoded in the L-functions attached to these (motivic) objects. For example, the famous conjecture of Birch and Swinnerton-Dyer explains how some of the arithmetic properties of rational elliptic curves (for instance the rank of the group of its rational points) are encoded in the analytic behavior of its L-function. These motivic L-functions are usually very hard to understand. We often cannot even prove that they admit analytic continuations. Automorphic representations, as introduced by Langlands, are supposed to provide a rich supply of L-functions that are easier to understand as we saw in Theorem 1.2.7 for instance. The dream would then be to show that any of our motivic L-functions is equal to one of these automorphic L-functions.

As we discussed earlier, the Langlands correspondence predicts that one can associate Galois representations to certain automorphic representations. One can also associate Galois representations to motives (namely their étale realizations). Therefore, Galois representations are supposed to provide the bridge between the motivic and the automorphic world. If one believes in this philosophy, one should expect to be able to translate different features of these worlds from one to another. The feature that is of interest to us in this thesis is the symmetries of the motivic object. It is reasonable to think that the more symmetric the motive is, the smaller the image of the associated Galois representation should be, and these should have implications for the automorphic representations associated with these objects. The precise (conjectural) relation of these objects is discussed in Chapter 5.

Let us briefly summarize what is known in this direction. The more precise statements and proofs will be discussed in Chapter 3. This line of study can be traced back to the work of Serre on Galois representations associated with rational elliptic curves and his famous open image theorem [39]. Weight 2 eigenforms with rational coefficients correspond to rational elliptic curves in the Langlands program. Therefore, Serre's result can also be viewed as determining the image of the Galois representations associated with weight 2 modular forms with rational Fourier coefficients. If the weight is still 2 but the coefficients are not rational, then there exists an abelian variety associated with the form. Ribet realized that for general modular forms of weight 2, the endomorphism ring of the associated abelian variety being big translates into the form having some kinds of symmetries which he called inner-twists [32], and he was able to generalize Serre's result to this context [33]. Momose then worked out the higher weight case [25] and Nekovář generalized their work to Hilbert modular forms [27]. The main goal of this thesis is to better understand the GL_2 case and to further generalize these results to groups other than GL_2.

Let us explain the structure of this thesis and the results that we prove in this direction. A primary question in determining the image of a Galois representation is if the representation is irreducible or not. This is related to the notion of cuspidality on the automorphic side. This is known as the irreducibility conjecture for automorphic Galois representations and will be explained in Chapter 2. Only special cases of this conjecture are known, mostly for small reductive groups. All of the proofs that we know use the analytic properties of automorphic L-functions.We will discuss the GL_2 and GL_3 cases of this conjecture in more detail and give proofs in these cases over totally real fields.

In Chapter 3, we will give a proof of Serre's open image theorem and explain how inner-twists come into play when one wants to compute the image of modular Galois representations. We will state the results of Ribet and Momose for classical modular forms and Nekovář's generalization to Hilbert modular forms. In these cases, there exists a quaternion algebra D over a certain subfield of the Hecke field of the modular form which describes the image of the associated Galois representation. For classical modular forms of weight 2, the work of Quer gives an explicit formula for this quaternion algebra in terms of the Fourier coefficients of the modular form. In Section 3.3.2, we generalize this formula to the case of Hilbert modular forms under mild assumptions. The first step that one needs to do for this generalization is to prove some results of Ribet (that are used by Quer in the classical case) for the case of Hilbert modular forms. The proofs are identical to those of Ribet but they do not seem to be written down in the literature. The second step is to mimic Quer's strategy. Here some of the Galois cohomology computations become more complicated due to the fact that our base field K is not contained in the coefficient field, whereas in the case of classical modular forms, $K = \mathbb{Q}$ is contained in any field of characteristic 0. Therefore, one needs to carefully go up and down between different fields to be able to carry out the computations.

Let us state our main result in this section which is Theorem 3.3.22 in the text. Let K be a totally real number field such that $[K : \mathbb{Q}]$ is odd, and f be a Hilbert modular newform of parallel weight 2, level $\mathcal{N}$, and trivial nebentype. For every non-zero prime ideal $\mathfrak{p}$ of O_K not dividing $\mathcal{N}$, we denote the eigenvalue of the Hecke operator $T_\mathfrak{p}$ acting on f by $a_\mathfrak{p}$. Let D/F be the quaternion algebra (constructed by Nekovář in this case) which describes the image (see Section 3.2.2). We use inner-twists of f to construct a field extension N of K of the form

$$N = K(\sqrt{t_1}, ..., \sqrt{t_m})$$

for $t_i \in K$ with the property that the characters of $\mathrm{Gal}(N/K)$ are exactly the characters appearing in the inner-twists of f. Let $\{\sigma_i\}_{i=1}^m$ be the $\mathbb{F}_2$-basis for $\mathrm{Gal}(N/K)$ satisfying $\sigma_i(\sqrt{t_j}) = (-1)^{\delta_{i,j}}\sqrt{t_j}$. Then we prove:

Theorem 1.3.1. *Let $\mathfrak{p}_1, \cdots, \mathfrak{p}_m$ be a set of prime ideals of O_K not dividing $\mathcal{N}$ and with $a_{\mathfrak{p}_i} \neq 0$ such that $\sigma_i = \mathrm{Frob}_{\mathfrak{p}_i}$ in $\mathrm{Gal}(N/K)$ (such primes exist by Chebotarev's theorem). Then in $\mathrm{Br}(F)$ one has:*

$$[D] = (\mathrm{Nr}_{FK/F}(t_1), a_{\mathfrak{p}_1}^2)(\mathrm{Nr}_{FK/F}(t_2), a_{\mathfrak{p}_2}^2) \cdots (\mathrm{Nr}_{FK/F}(t_m), a_{\mathfrak{p}_m}^2),$$

where $(a, b) = (a, b)_F$ denotes the Hilbert symbol.

The main objective of Chapter 4 is to generalize the results of Ribet, Momose and Nekovář to the case of GL_n. A key observation is that in this case, one should also take into account the so-called "outer-twists" of an automorphic representation. This is due to the fact that not every automorphism of SL_n is inner for $n > 2$. Inner- and outer-twists of a representation together form a group that we call the group of extra-twists. We use these extra-twists to compute the p-adic Lie algebra of certain n-dimensional Galois representations that satisfy a list of natural properties. These properties are expected to hold for the Galois representations associated to "general type" regular algebraic cuspidal automorphic representations (after possibly a twist), but we are only able to prove this in the GL_3 case.

Let us state our main result in the GL_3 case, which is Theorem 4.3.13 in the text. Let K be a totally real field and let π be a regular algebraic cuspidal automorphic representation of $\mathrm{GL}_3(\mathbb{A}_K)$ that is of general type, i.e. it neither satisfies $\pi \simeq \pi \otimes \chi$ for a non-trivial Hecke character χ, nor $\pi \simeq \pi^\vee \otimes \eta$ for any Hecke character η. Let $\mathbb{Q}(\pi)$ be its Hecke field with Galois closure E, fix a prime p and let $\rho_{\pi,p} : \Gamma_K \to \mathrm{GL}_3(\mathbb{Q}(\pi) \otimes_{\mathbb{Q}} \overline{\mathbb{Q}_p})$ be the p-adic Galois representation attached to π and $\Gamma \subseteq \mathrm{Aut}(E)$ the group of extra-twists which will be defined later. Here is our main result:

Theorem 1.3.2. *Let $F = E^\Gamma$ be the field fixed by all extra-twists of π. Then for every prime number p, there exists a finite extension L/K and a semi-simple algebraic group H_p defined over $F_p := F \otimes_{\mathbb{Q}} \mathbb{Q}_p$ which is a form of SL_3 (constructed using the extra-twists), such that $\rho_{\pi,p}(\Gamma_L)$ is contained in $H_p(F_p) \cdot \mathbb{Q}_p^\times \subseteq \mathrm{GL}_3(E \otimes_{\mathbb{Q}} \overline{\mathbb{Q}_p})$ and it is open in the p-adic topology.*

To construct the group H_p, we use the group of extra-twists of π to define a 1-cocycle which gives us a form of SL_3. To prove the openness we need to compare the Lie algebras. We first twist away the determinant and only focus on the semi-simple part of the Lie algebra. To use our earlier computations on Lie algebras, we need some information about the Lie algebras over the algebraic closure. Here, the key ingredients are the classification of semi-simple Lie subalgebras of $\mathfrak{sl}_3$, Langlands functoriality for $\mathrm{sym}^2 : \mathrm{GL}_2 \to \mathrm{GL}_3$, and automorphic induction for degree 3 extensions.

Assuming the functoriality conjectures of Langlands, one can go through the arguments in the proof of the above theorem and see what assumptions are needed on π to prove such a result for GL_n, i.e. when extra-twists are enough to give a precise description of the image. This should intuitively mean that π is not coming from any smaller group via a Langlands transfer. We make this precise in Section 4.3.3 and define automorphic representations of general type and prove a big image theorem for the Galois representations associated to those, assuming Langlands functoriality.

In Chapter 5, we investigate the relation of our results in Chapter 4 and the Mumford-Tate conjecture. The conjectures of Clozel [11] predict the existence of a motive M_π over K (with coefficients in an extension of $\mathbb{Q}(\pi)$) attached to π. The Mumford-Tate conjecture for this motive tells us that the groups H_p from Theorem 1.3.2 should arise from a global object H defined over the field F. Assuming the existence of such a motive, we will use the action of extra-twists on the Hodge structure to construct a group H_∞ over $F \otimes_{\mathbb{Q}} \mathbb{R}$ that should be the Archimedean part of the Mumford-Tate group. We will also use this action on the rational Hodge structure to construct a global group H. This group will contain the (special) Mumford-Tate group and in particular its dimension (which is equal to the dimension of all the groups H_p from the Theorem 4.3.13) bounds the dimension of the Mumford-Tate group from above.

Chapter 2

The Irreducibility Conjecture

2.1 Harish-Chandra's Philosophy of Cusp Forms

Harish-Chandra's idea of reducing the study of automorphic representations to cuspidal ones, by means of parabolic induction, is a precursor of the Langlands philosophy. Let us explain this in a special case, in the language of Langlands functoriality:

Let K be a number field and consider the reductive group $G = \mathrm{GL}_n$ over K. Let P be a parabolic subgroup of GL_n whose associated Levi subgroup M is given by the natural diagonal block embedding

$$\mathrm{GL}_{n_1} \times \cdots \times \mathrm{GL}_{n_k} \hookrightarrow \mathrm{GL}_n$$

for a partition $n = n_1 + \cdots + n_k$. Since the dual group of GL_n is isomorphic to GL_n and all the groups are split here, one can simply take the L-map

$${}^L M \to {}^L G$$

to be the identity on the Galois component and the natural diagonal block embedding on the algebraic groups:

$$\mathrm{GL}_{n_1}(\overline{\mathbb{Q}}_p) \times \cdots \times \mathrm{GL}_{n_k}(\overline{\mathbb{Q}}_p) \times \Gamma_K \to \mathrm{GL}_n(\overline{\mathbb{Q}}_p) \times \Gamma_K.$$

It is proved by Langlands that Langlands functoriality holds for this L-map. In particular, if π_1,..., π_k are automorphic representations of GL_{n_1},..., GL_{n_k} respectively, then the automorphic representation $\pi_1 \otimes \cdots \otimes \pi_k$ of $\mathrm{GL}_{n_1} \times \cdots \times \mathrm{GL}_{n_k}$ can be transferred to an automorphic representation $\pi_1 \boxplus \cdots \boxplus \pi_k$ of GL_n via the above L-map. This is called the isobaric sum of π_1,..., π_k.

Cuspidal automorphic representations are morally the ones that cannot be found as subquotients of the parabolic inductions of smaller groups. They serve as the building blocks upon which other automorphic representations are constructed by parabolic induction. In particular, one can start with cuspidal automorphic representations π_1,..., π_k and construct a new automorphic representation $\pi_1 \boxplus \cdots \boxplus \pi_k$ which is no longer cuspidal.

Definition 2.1.1. An automorphic representation π of GL_n is called isobaric if it is isomorphic to an isobaric sum $\pi_1 \boxplus \cdots \boxplus \pi_k$ of cuspidal automorphic representations.

Not every automorphic representation is isobaric. But these are the simplest ones that can be constructed from cuspidal automorphic representations. If π is isomorphic to an isobaric sum $\pi \simeq \pi_1 \boxplus \cdots \boxplus \pi_k$ of cuspidal automorphic representations, then one easily observes that

$$L(s,\pi) = L(s,\pi_1) \cdots L(s,\pi_k),$$

therefore the analytic properties of $L(s,\pi)$ are also understood by Theorem 1.2.7.

As was discussed in the last chapter, to have the most satisfactory formulation of the Langlands conjectures, one needs the existence of the hypothetical Langlands $\mathcal{L}$-group of the number field K. Nevertheless, Clozel was able to formulate precise conjectures for the group GL_n by restricting himself to algebraic automorphic representations where one expects a correspondence to the representations of the usual Galois group of K rather than the $\mathcal{L}$-group [11]. Let us recall a version of his conjectures (combined with a conjecture of Fontaine and Mazur):

Conjecture 2.1.2 (Langlands, Clozel, Fontaine-Mazur)**.** There is a (unique) bijection between the two sets:

- algebraic isobaric automorphic representations of $\mathrm{GL}_n(\mathbb{A}_K)$, and
- isomorphism classes of semi-simple continuous representations $\Gamma_K \to \mathrm{GL}_n(\overline{\mathbb{Q}}_p)$ that are unramified outside a finite set of places and de Rham at places above p,

such that at unramified primes, the Satake parameters match with the Frobenius eigenvalues.

The famous conjecture of Fontaine and Mazur predicts that these kinds of Galois representations (if one adds an oddness assumption on their irreducible subquotients) are exactly the ones that come from algebraic geometry, i.e. these are subquotients of the

p-adic étale cohomologies of smooth proper varieties over K. By a famous conjecture of Tate, these Galois representations are in fact semi-simple and can be written as a direct sum of irreducible ones. As one can see from the definition of the L-map

$$^LM \to {}^LG,$$

the operator $\boxplus$ corresponds to the direct sum $\oplus$ on the Galois side under the Langlands correspondence.

Given all these, it is then natural to guess that under the above bijection, cuspidal automorphic representations that are the building blocks on the automorphic side should exactly correspond to irreducible Galois representations that are the building blocks on the Galois side. This is known as the Irreducibility Conjecture:

Conjecture 2.1.3. Under the Langlands correspondence for GL_n (Conjecture 2.1.2), cuspidal representations correspond to irreducible Galois representations.

In fact, assuming that we are in the Artin case, i.e. the Galois representation has finite image, one can easily prove this expectation:

Proposition 2.1.4. *Let π be an isobaric automorphic representation of $\mathrm{GL}_n(\mathbb{A}_K)$ and $\rho : \Gamma_K \to \mathrm{GL}_n(\overline{\mathbb{Q}}_p)$ the associated Galois representation via Conjecture 2.1.2. Assume that $\rho(\Gamma_K)$ is finite. Then π is cuspidal if and only if ρ is irreducible.*

Proof. We follow the argument in the introduction of [29]. Let π be cuspidal. Then by Theorem 1.2.7 the Rankin-Selberg L-function $L(s, \pi \times \pi^\vee)$ has a simple pole at $s = 1$. Since the Satake parameters of π match with Frobenius eigenvalues of ρ outside a finite set S of finite places of K, we have the equality of partial L-functions

$$L^S(s, \pi \times \pi^\vee) = L^S(s, \rho \otimes \rho^\vee),$$

where the superscript S indicates that we are removing the Euler factors for places in S. This does not change the analytic properties (analyticity, order of pole, etc) at $s = 1$ since the removed factors are holomorphic and non-zero around $s = 1$. We conclude that $L(s, \rho \otimes \rho^\vee)$ has a simple pole at $s = 1$.

Now from the theory of Artin L-functions, the only irreducible representation whose L-function has a pole at $s = 1$ is the trivial representation. This means that if we decompose $\rho \otimes \rho^\vee \simeq \mathrm{End}(\rho)$ into irreducible representations, there is exactly one trivial factor, which has to be the subrepresentation of the scalar matrices. But if $\rho = \sigma \oplus \eta$ is reducible, then

$$\rho \otimes \rho^\vee \simeq (\sigma \otimes \sigma^\vee) \oplus (\sigma \otimes \eta^\vee) \oplus (\eta \otimes \sigma^\vee) \oplus (\eta \otimes \eta^\vee)$$

contains at least two trivial representations as a direct summand (one inside $\sigma \otimes \sigma^\vee \simeq \mathrm{End}(\sigma)$ and one inside $\eta \otimes \eta^\vee \simeq \mathrm{End}(\eta)$). This implies that ρ is irreducible.

To prove the other direction, one reverses all the arguments and concludes using [19, Lemma 11.8.1] that π must have exactly one cuspidal factor in its isobaric decomposition. □

Little is known about the irreducibility conjecture in general, even though the Galois representation associated with a cuspidal automorphic representation is constructed in many cases when the base field K is totally real or CM. In the rest of this section, we review some of the known instances of this conjecture. In almost all of these cases, the argument uses analytic properties of automorphic L-functions.

2.2 The $\mathbf{GL_2}$ Case

2.2.1 Modular Galois Representations

Let $f = \sum_{n \geqslant 1} a_n q^n$ be a cuspidal newform of weight $k \geqslant 2$, level N, and nebentype ϵ. If $k = 2$, then it is well known that there exists an abelian variety $A_f/\mathbb{Q}$ associated with f whose dimension is the degree of the number field $E = \mathbb{Q}(f)$ over the rational numbers [14, §6.6]. One can then show that the Tate module of A_f is a free $E_p = E \otimes_{\mathbb{Q}} \mathbb{Q}_p$ module [33, §4] which has to be of rank 2 since $\dim A_f = [E : \mathbb{Q}]$. This gives the Galois representation

$$\prod_{\mathfrak{p}|p} \rho_{f,\mathfrak{p}} = \rho_{f,p} : \Gamma_{\mathbb{Q}} \to \mathrm{GL}_2(E_p) = \prod_{\mathfrak{p}|p} \mathrm{GL}_2(E_{\mathfrak{p}}).$$

The Eichler-Shimura relation can be used to show that $\mathrm{tr}\rho_p(\mathrm{Frob}_\ell) = a_\ell$ for all $\ell \nmid pN$. This means that (possibly up to semi-simplification) this is exactly the Galois representation predicted to be associated with f by Conjecture 2.1.2, viewing f as an automorphic representation by the standard procedure. Coming from abelian varieties, the Galois representation ρ_p (or rather each $\rho_{\mathfrak{p}}$) is de Rham with Hodge-Tate weights $(0, -1)$.

The same story should hold for Hilbert modular forms of parallel weight 2, but this is not known in general. We will revisit this in the next chapter.

Now assume that $k \geqslant 2$. Conjecture 2.1.2 still predicts the existence of a Galois representation

$$\prod_{\mathfrak{p}|p} \rho_{f,\mathfrak{p}} = \rho_{f,p} : \Gamma_{\mathbb{Q}} \to \mathrm{GL}_2(E_p) = \prod_{\mathfrak{p}|p} \mathrm{GL}_2(E_{\mathfrak{p}})$$

whose Frobenius traces at primes away from pN are given by the Fourier coefficients of f. This is indeed known by the work of Deligne [13]. These Galois representations could also be found in the p-adic cohomology of certain varieties and are de Rham with Hodge-Tate weights $(0, 1-k)$.

One can hope to be able to improve Deligne's construction to find a more geometric object whose p-adic cohomology is the above Galois representation. In fact, Scholl constructs a motive M_f associated with f whose p-adic realization is exactly the Galois representation constructed by Deligne [37]. When $k = 2$, this motive is exactly the motive induced by (the first cohomology of) the abelian variety A_f.

2.2.2 Irreducibility Using L-Functions

If f is a cuspidal newform as in the last section, the automorphic representation π_f associated to it is algebraic and cuspidal. Therefore one expects each Galois representation $\rho_{f,\mathfrak{p}}$ to be absolutely irreducible by Conjecture 2.1.3. We will prove this in this section. The argument is due to Ribet. The same argument works for Hilbert modular forms.

Theorem 2.2.1. *Let f be a cuspidal newform of weight $k \geqslant 2$. Then, for each finite place $\mathfrak{p}$ of the Hecke field $E = \mathbb{Q}(f)$, the Galois representation*

$$\rho_{f,\mathfrak{p}} : \Gamma_{\mathbb{Q}} \to \mathrm{GL}_2(E_{\mathfrak{p}})$$

is absolutely irreducible.

Proof. Assume that $\rho_{f,\mathfrak{p}}$ is reducible and the ($\overline{\mathbb{Q}}_p$-valued) character η_1 is a subrepresentation and let η_2 be the quotient representation. Then

$$\rho_{f,\mathfrak{p}}^{ss} = \eta_1 \oplus \eta_2.$$

As was mentioned, $\rho_{f,\mathfrak{p}}$ is de Rham. This implies that both η_1 and η_2, being a subrepresentation and a quotient of it, are also de Rham. The Hodge-Tate weights of $\rho_{f,\mathfrak{p}}$ are $(0, 1-k)$ which means that one of η_1 and η_2 has Hodge-Tate weight 0 and the other one $1-k$. Since every Hodge-Tate character of $\Gamma_{\mathbb{Q}}$ is of the form $\chi \cdot \chi_p^n$ for some Dirichlet character χ, we deduce that one of η_1 and η_2 is a Dirichlet character η and the other one is of the form $\chi \cdot \chi_p^{k-1}$, hence

$$\rho_{f,\mathfrak{p}}^{ss} = \eta \oplus \chi \cdot \chi_p^{k-1}.$$

Now note that $f \otimes \chi^{-1}$ is also a cuspform and its associated Galois representation is clearly equal to $\rho_{f,\mathfrak{p}} \otimes \chi^{-1}$ by the Brauer-Nesbitt Theorem. Therefore twisting by χ^{-1} we get:

$$\rho_{f \otimes \chi^{-1},\mathfrak{p}}^{ss} = \eta\chi^{-1} \oplus \chi_p^{k-1}.$$

Taking the L-functions of both sides we have

$$L(s, f \otimes \chi^{-1}) = L(s, \eta\chi^{-1})L(s, \chi_p^{k-1}) = L(s, \eta\chi^{-1})\zeta(s + 1 - k),$$

where ζ is the Riemann zeta function. The left hand side, being the L-function of a cusp form, has analytic continuation to an entire function. Since $\zeta(s + 1 - k)$ has a pole at $s = k$, this should be canceled by a root of the Dirichlet L-function $L(s, \eta\chi^{-1})$ at k. Since Dirichlet L-functions are non-vanishing for $\mathrm{Re}(s) > 0$, we are done. □

More generally, if one assumes the existence of a Galois representation associated with a cuspidal automorphic representation of $\mathrm{GL}_2(\mathbb{A}_K)$ for a number field K, then one uses a sort of similar argument to show the irreducibility of this Galois representation using analytic properties of the L-functions. See [45, Theorem 1.2.6] for more details.

2.3 The $\mathbf{GL_3}$ Case

The irreducibility conjecture is also known for cuspidal automorphic representation of $\mathrm{GL}_3(\mathbb{A}_K)$ for a totally real field K. This is proved in [6] by Böckle and Hui. In this section we give an overview of their proof. For all of the omitted details we refer to [6].

2.3.1 Locally Algebraic Representations

In his work on abelian Galois representations, Serre defined the notion of locally algebraic representations, which can be seen as a precursor to more sophisticated later notions of p-adic Hodge theory [39]. He used this in the study of the Galois representations associated with elliptic curves. See [40] for more details on this.

Assume that K is a general number field. A p-adic representation of Γ_K is called abelian if its image is an abelian group. In other words it factors through Γ_K^{ab}, the abelianization of Γ_K:

$$\phi_p : \Gamma_K^{ab} \to \mathrm{GL}_n(\overline{\mathbb{Q}}_p).$$

Such abelian representations can be related to adelic groups using the Artin reciprocity map:

$$\mathrm{Art_K} : \mathbb{A}_K^\times / K^\times \to \Gamma_K^{ab},$$

and locally algebraic ones are those that could be described by a morphism of algebraic groups around identity, after the above passage via global class field theory. Let $T := \mathrm{Res}_{\mathbb{Q}}^K \mathbb{G}_m$. Here is the precise definition:

Definition 2.3.1. An abelian semi-simple p-adic Galois representation

$$\phi_p : \Gamma_K^{ab} \to \mathrm{GL}_n(\overline{\mathbb{Q}}_p)$$

is called locally algebraic if there exists a morphism of algebraic groups

$$r : T_{\overline{\mathbb{Q}}_p} \to \mathrm{GL}_{n, \overline{\mathbb{Q}}_p}$$

such that the composition

$$T(\mathbb{Q}_p) = \prod_{v|p} K_v^\times \to \mathbb{A}_K^\times / K^\times \xrightarrow{\mathrm{Art_K}} \Gamma_K^{ab} \xrightarrow{\phi_p} \mathrm{GL}_n(\overline{\mathbb{Q}}_p)$$

is equal to $r|_{T(\mathbb{Q}_p)}$ on a small enough neighborhood of the identity element on the p-adic Lie group $T(\mathbb{Q}_p) = (K \otimes_{\mathbb{Q}} \mathbb{Q}_p)^\times$.

The following result of Fontaine shows the relation of this notion to the more familiar notions of p-adic Hodge theory discussed earlier.

Proposition 2.3.2 (Fontaine)**.** *For an abelian semi-simple Galois representation of* Γ_K*, being locally algebraic, being de Rham at places above* p *and being Hodge-Tate at places above* p *are equivalent.*

What makes the definition of locally algebraic representations interesting is the following result of Serre. Recall the definition of an E-rational Galois representation from Section 1.1.1.

Theorem 2.3.3 (Serre)**.** *For an abelian semi-simple Galois representation of* Γ_K*, being locally algebraic is equivalent to being* E*-rational for some number field* E*.*

At the heart of the above results (and of [6]) lies the following surprisingly strong result of Waldschmidt that heavily uses transcendental theory:

Theorem 2.3.4 (Waldschmidt)**.** *Let*

$$\chi : \Gamma_K \to \overline{\mathbb{Q}}_p^\times$$

be a character, unramified outside a finite set of places of K such that at each such place v, $\chi(\mathrm{Frob}_v)$ is an algebraic number (over $\mathbb{Q}$). Then there exists a positive integer N such that χ^N is locally algebraic.

The main new input of the work of Böckle and Hui in [6], is to generalize Theorem 2.3.3 to the so-called weak abelian direct summands of (not necessarily abelian) semi-simple Galois representations.

Definition 2.3.5. Let

$$\rho_p : \Gamma_K \to \mathrm{GL}_n(\overline{\mathbb{Q}}_p)$$

be an arbitrary Galois representation and

$$\psi_p : \Gamma_K \to \mathrm{GL}_m(\overline{\mathbb{Q}}_p)$$

be a semi-simple abelian Galois representation. We say that ψ_p is a weak abelian direct summand of ρ_p if there exists a density one set of (rational) primes $\mathcal{L}$ such that for each $\ell \in \mathcal{L}$ and finite place v of K above ℓ, the representations ρ_p and ψ_p are both unramified at v and the characteristic polynomial of $\psi_p(\mathrm{Frob}_v)$ divides the characteristic polynomial of $\rho_p(\mathrm{Frob}_v)$.

The obvious example of the above situation is when ψ_p is in fact a direct summand of ρ_p but there are examples where this does not hold. Finally, here is the main result of the work of Böckle and Hui:

Theorem 2.3.6 (Theorem 1.1 of [6])**.** *Let $E \subset \overline{\mathbb{Q}}_p$ be a number field and*

$$\rho_p : \Gamma_K \to \mathrm{GL}_n(\overline{\mathbb{Q}}_p)$$

be a semi-simple, E-rational, p-adic Galois representation. Let the representation ψ_p be a weak abelian direct summand of ρ_p. Then ψ_p is locally algebraic and hence de Rham at all places above p.

2.3.2 Irreducibility Using L-Functions

Having Theorem 2.3.6 in hand, one can apply an L-function argument in the same spirit as in the proof of 2.2.1 to prove the irreducibility conjecture in the GL_3 case, at least over a totally real field K. The existence of the Galois representations is known in this case. In fact, more generally, one has:

Theorem 2.3.7 (Harris-Lan-Taylor-Thorne [22], Scholze [38])**.** *Let K be either a totally real or a CM number field and π be a regular algebraic cuspidal automorphic representation of $\mathrm{GL}_n(\mathbb{A}_K)$. Then for every embedding $\lambda : \mathbb{Q}(\pi) \hookrightarrow \overline{\mathbb{Q}}_p$ there exists a semi-simple Galois representation*

$$\rho_{\pi,\lambda} : \Gamma_K \to \mathrm{GL}_n(\overline{\mathbb{Q}}_p),$$

unramified outside a finite set of primes S, such that for any finite place $v \notin S$ of K, the eigenvalues of $\rho_{\pi,\lambda}(\mathrm{Frob}_v)$ match with the Satake parameters of π at v (after applying λ), as predicted by Conjecture 2.1.2.

The Galois representations in the above theorem are expected to be irreducible by Conjecture 2.1.3. It has been known for a long time, that when K is totally real, $n = 3$, and π is essentially self-dual then the Galois representation associated with it is irreducible [5]. The $n = 3$ case was also studied by [8] in the non essentially self-dual setting. We will reduce the general case to the case of essential self-duality.

Theorem 2.3.8 (Böckle-Hui)**.** *Let K be a totally real number field and π be a regular algebraic cuspidal automorphic representation of $\mathrm{GL}_3(\mathbb{A}_K)$. Then for any embedding $\lambda : \mathbb{Q}(\pi) \hookrightarrow \overline{\mathbb{Q}}_p$, the Galois representation $\rho_{\pi,\lambda}$ is irreducible.*

Proof. We write $\rho := \rho_{\pi,\lambda}$ for simplicity. Assume that ρ is reducible:

$$\rho = \sigma \oplus \tau,$$

where σ is a two dimensional and τ is a one dimensional Galois representation. Therefore τ is abelian and hence, a weak abelian direct summand of ρ. Since ρ is clearly $\mathbb{Q}(\pi)$-rational by its definition and $\mathbb{Q}(\pi)$ is a number field by a result of Clozel (see Definition 1.2.8), Theorem 2.3.6 implies that the character τ is de Rham. Since K is totally real, such a character must be a Tate-twist of a finite character and hence comes from an algebraic Hecke character η via class field theory. Now notice that:

$$\bigwedge^2 \rho = \bigwedge^2 \sigma \oplus (\sigma \otimes \tau) \oplus \bigwedge^2 \tau = \det(\sigma) \oplus (\sigma \otimes \tau).$$

Adding τ^2 to both sides we get

$$\tau^2 \oplus \bigwedge^2 \rho = \det(\sigma) \oplus (\rho \otimes \tau).$$

In particular $\det(\sigma)$ is also locally algebraic by 2.3.6 and hence corresponds to an algebraic Hecke character χ. Twisting both sides with τ^{-2} we have

$$1 \oplus (\tau^{-2} \otimes \bigwedge^2 \rho) = (\rho \otimes \tau^{-1}) \oplus (\det(\sigma) \otimes \tau^{-2}).$$

Now we look at the (partial) L-functions of both sides. Since ρ and τ are automorphic (associated with π and η) and Langlands functoriality is known for $\bigwedge^2 : \mathrm{GL}_3 \to \mathrm{GL}_3$, we can replace the associated L-functions with the automorphic ones. All the L-functions are normalized so that the critical strip is between $\mathrm{Re}(s) = 0$ to 1:

$$\zeta(s)L(s, \eta^{-2} \otimes \bigwedge^2 \pi) = L(s, \pi \otimes \eta^{-1})L(s, \chi\eta^{-2}).$$

Now $L(s, \eta^{-2} \otimes \bigwedge^2 \pi)$ is non-vanishing at $s = 1$ by a result of Shahidi and ζ has a pole at $s = 1$. Since $\pi \otimes \eta^{-1}$ is cuspidal, $L(s, \pi \otimes \eta^{-1})$ is entire and therefore the L-function of the character $\chi\eta^{-2}$ has a pole at $s = 1$ and hence this character is trivial and $\chi = \eta^2$, or equivalently, $\det(\sigma) = \tau^2$. This simply means that $\rho \otimes \tau^{-1}$ is self-dual and therefore ρ and hence π are essentially self-dual. As was mentioned, the theorem is known in this case. □

Böckle and Hui use this theorem to investigate the geometric monodromy group of the above Galois representations and then use this to prove some p-adic Hodge theoretic properties of them. We will make use of the following result of their work later:

Theorem 2.3.9. *Keeping the notation of Theorem 2.3.8, there exists a density one set of (rational) primes $\mathcal{P}$, such that for each $p \in \mathcal{P}$ and embedding $\lambda : \mathbb{Q}(\pi) \hookrightarrow \overline{\mathbb{Q}}_p$, the Galois representation $\rho_{\pi,\lambda}$ is de Rham and regular (has distinct τ-HT weights for each embedding $\tau : K \hookrightarrow \mathbb{R}$).*

Chapter 3

Inner-Twists and 2-Dimensional Galois Representations

3.1 Serre's Open Image Theorem

The first case of big image theorems for Galois representations beyond class field theory was worked out by Serre for Galois representations associated with elliptic curves over number fields. Over the field of rational numbers, this can also be seen as computing the image of Galois representations associated with weight 2 cuspidal newforms with rational Fourier coefficients, since one can associate an elliptic curve to such modular forms by a construction of Eichler and Shimura. We will give a proof of Serre's result using Faltings' results on Tate's Conjecture in this section. The case of more general eigenforms will be discussed in the next section.

Let K be a number field and E/K an elliptic curve. It is well known that the rational endomorphism algebra of this elliptic curve, $\mathrm{End}^0_{\overline{\mathbb{Q}}}(E) := \mathrm{End}_{\overline{\mathbb{Q}}}(E) \otimes_{\mathbb{Z}} \mathbb{Q}$, is either isomorphic to $\mathbb{Q}$ or to $\mathbb{Q}(\sqrt{-d})$ for some square-free positive integer $d > 1$. In the second case E is said to have complex multiplication or CM for short. The first case is then called the non-CM case. If K contains $\sqrt{-d}$ then one can see that the Galois representation associated with E decomposes as the sum of two characters. If not, the restriction to the Galois group of $K(\sqrt{-d})$ does, so we get

$$\rho_{E,p}|\Gamma_{K(\sqrt{-d})} \simeq \chi \oplus {}^{\sigma}\chi,$$

where σ is the non-trivial element of $\mathrm{Gal}(K(\sqrt{-d})/K)$. This means that in this case $\rho_{E,p} \simeq \mathrm{Ind}_K^{K(\sqrt{-d})}\chi$. In any case, when E has CM, the image of the Galois representation is small, i.e. it can be described by a character after at most a degree two extension.

If E is non-CM then Serre proves that the restriction of $\rho_{E,p}$ to the Galois group of any finite extension of K is still irreducible [39]. Here, we will give a proof based on Faltings' theorem.

Theorem 3.1.1 (Serre)**.** *Let K be a number field and E over K an elliptic curve without complex multiplication. Then $\rho_{E,p}$ is strongly absolutely irreducible, i.e. for any finite extension L/K one has that $\rho_{E,p}|_{\Gamma_L}$ is absolutely irreducible.*

Proof. Assume that there exists L such that

$$\rho_{E,p}|_{\Gamma_L} \simeq \chi_1 \oplus \chi_2$$

for two characters $\chi_1, \chi_2 : \Gamma_L \to \overline{\mathbb{Q}}_p^{\times}$. This clearly means that $\mathrm{End}_{\overline{\mathbb{Q}}_p[\Gamma_L]}(V_p(E))$ contains $\overline{\mathbb{Q}}_p \times \overline{\mathbb{Q}}_p$, and hence is at least two-dimensional. By Theorem 3.1 we know that

$$\mathrm{End}_L^0(E) \otimes_{\mathbb{Q}} \overline{\mathbb{Q}}_p \simeq \mathrm{End}_{\overline{\mathbb{Q}}_p[\Gamma_L]}(\rho_{E,p}),$$

therefore $\mathrm{End}_L^0(E)$ is at least two-dimensional which means that E has complex multiplication. This contradiction implies the result. □

We will see later that proving a strong irreducibility result such as the one above is always a crucial part of proving an open image theorem for p-adic Galois representations. The next proposition gives another interpretation of strong irreducibility:

Lemma 3.1.2. *Let $\rho : G \hookrightarrow \mathrm{GL}_n(\overline{\mathbb{Q}_p})$ be a closed embedding. Then the representation ρ is strongly irreducible if and only if the induced Lie algebra representation*

$$d\rho : \mathrm{Lie}(G) \hookrightarrow \mathfrak{gl}_n$$

is irreducible.

Proof. Let V be the underlying vector space of ρ. Then G acts on V via ρ and $\mathfrak{g} = \mathrm{Lie}(G)$ acts on V via $d\rho$. Now assume that $V = W_1 \oplus W_2$ as a representation of $\mathfrak{g}$. Then, by Theorem 1.1.7, the representation $\mathfrak{g} \to \mathrm{End}(W_i)$ can be lifted to a group representation $U_i \to \mathrm{Aut}(W_i)$ for an open subgroup U_i of G for $i = 1, 2$. This simply means that $\rho|_{U_i}$ is reducible unless W_i is not proper. This proves that strong irreducibility implies Lie algebra irreducibility. The other direction is obvious. □

Now, we are ready to prove the main theorem of this section:

Theorem 3.1.3 (Serre)**.** *Let K be a number field and E over K an elliptic curve without complex multiplication. Let*

$$\rho_{E,p} : \Gamma_K \to \mathrm{GL}_2(\mathbb{Z}_p)$$

be the Galois representation associated with E (after choosing a basis for the Tate module). Then the image of this representation is open, i.e. $\rho_{E,p}(\Gamma_K) \subseteq \mathrm{GL}_2(\mathbb{Z}_p)$ is open in the p-adic topology.

Proof. First, note that $\mathrm{GL}_2(\mathbb{Z}_p) \subseteq \mathrm{GL}_2(\mathbb{Q}_p)$ is open so it is enough to prove the result for

$$\rho_{E,p} : \Gamma_K \to \mathrm{GL}_2(\mathbb{Q}_p).$$

Also, notice that $\mathrm{GL}_2(\mathbb{Q}_p)$ is a p-adic Lie group, and since Γ_K is compact, $\rho_{E,p}(\Gamma_K)$ is a closed subgroup of this p-adic Lie group and therefore a Lie subgroup. Now by corollary 1.1.8, it is enough to prove that the Lie algebra of the image is $\mathfrak{gl}_2$. Let $G_p = \rho_{E,p}(\Gamma_K)$ and $\rho : G_p \hookrightarrow \mathrm{GL}_2(\mathbb{Q}_p) = \mathrm{Aut}(V_p)$ be the inclusion map where $V_p = V_p(E)$ is the rational p-adic Tate module. Also, let $\mathfrak{g}_p$ be the Lie algebra of G_p. Notice that

$$d\rho : \mathfrak{g}_p \to \mathfrak{gl}_2 = \mathrm{End}(V_p)$$

gives us a Lie algebra representation of $\mathfrak{g}_p$. This representation is irreducible by 3.1.2 and hence the centralizer $C_{\mathfrak{gl}_2}(\mathfrak{g}_p)$ is a $\mathbb{Q}_p$-division algebra. We want to show that it is actually equal to $\mathbb{Q}_p$. Choose a number field L/K such that all the endomorphisms of E are defined over L, i.e. $\mathrm{End}_L(E) = \mathrm{End}_{\overline{\mathbb{Q}}}(E)$. This means that for any normal open subgroup H of Γ_L one has $\mathrm{End}_L(E) \otimes_{\mathbb{Q}} \mathbb{Q}_p = \mathrm{End}_H(V_p)$ and hence it must hold infinitesimally by Theorem 1.1.7:

$$\mathrm{End}_L(E) \otimes_{\mathbb{Q}} \mathbb{Q}_p = \mathrm{End}_{\mathfrak{g}_p}(V_p) = C_{\mathfrak{gl}_2}(\mathfrak{g}_p).$$

Now by the non-CM assumption $\mathrm{End}_{\overline{\mathbb{Q}}}(E) = \mathbb{Q}$ and therefore $C_{\mathfrak{gl}_2}(\mathfrak{g}_p) = \mathbb{Q}_p$. By the classification of irreducible Lie subalgebras of $\mathfrak{gl}_2$, the only Lie subalgebras with this property are $\mathfrak{sl}_2$ and $\mathfrak{gl}_2$. We need to exclude the first case. Assume that $\mathfrak{g}_p = \mathfrak{sl}_2$. Then $G_p \cap \mathrm{SL}_2(\mathbb{Q}_p)$ is open in G_p by Corollary 1.1.8 which means that there exists a finite extension M/K such that

$$\rho_{E,p}(\Gamma_M) \subseteq \mathrm{SL}_2(\mathbb{Q}_p).$$

On the other hand $\det(\rho_{E,p}) = \chi_p$ is the p-adic cyclotomic character which cannot be trivialized by a finite extension. This implies that $\mathfrak{g}_p = \mathfrak{gl}_2$ and we are done. □

3.2 Inner-Twists of Modular Forms

3.2.1 Classical Modular Forms

If $f \in S_2(\Gamma_\circ(N), \mathbb{Q})$ is a cuspidal newform, then there exists an elliptic curve E_f associated with it whose Tate module is the Galois representation associated with f (see Section 2.2.1). Then the above result of Serre describes the image of $\rho_{f,p}$. One can naturally ask what happens in more general cases. If $f \in S_k(N, \epsilon)$ with $\mathbb{Q}(f) \neq \mathbb{Q}$ then even $\det(\rho_{f,p}) = \epsilon\chi_p^{k-1}$ does not have open image in $\mathrm{GL}_1(\mathbb{Q}(f)_p)$, so for one thing, one needs to be careful about the determinant. This is not hard to fix, since one can for instance study the intersection of the image with SL_2 or twist away the determinant after a finite extension as we will see later. A more subtle issue is the existence of *extra symmetries* which can already be seen in weight 2. In this case there is an abelian variety A_f associated with f whose Tate module (endowed with the right $\mathbb{Q}(f)$-structure) gives $\rho_{f,\lambda}$ for a place λ of $\mathbb{Q}(f)$ above p (or an embedding $\lambda : \mathbb{Q}(f) \to \overline{\mathbb{Q}}_p$ if you prefer). This abelian variety could very well have a non-trivial endomorphism ring. This forces the image of the associated Galois representation $V_p(A_f)$ to be smaller than usual since it must commute with this action. To understand how this affects $\rho_{f,\lambda}$ one also needs to consider how these interact with E. Of course, one likes to rewrite this purely in terms of the automorphic data so that it can be generalized to other situations where the associated algebro-geometric objects are less explicit (or not available at all). As Ribet observed, if the endomorphism ring of A_f is big, then f satisfies some sort of symmetries which he called inner-twists. Ribet defines inner-twists as follows [33, §3]:

Definition 3.2.1. Let $f \in S_k(N, \epsilon)$ be a non-CM cuspidal newform of level N, nebentype ϵ, weight $k \geqslant 2$ and Hecke field $E = \mathbb{Q}(f)$. An inner-twist of f is a pair (σ, χ) of an embedding $\sigma : E \hookrightarrow \mathbb{C}$ and a Dirichlet character χ such that for almost all primes p one has:

$$^{\sigma}a_p(f) = a_p(f) \cdot \chi(p).$$

Ribet shows that in fact (even if $E/\mathbb{Q}$ is not Galois) one has $\sigma(E) = E$ for every σ appearing in an inner-twist. This shows that one can choose σ from $\mathrm{Aut}(E)$ instead. This furthermore shows that if π_f is the automorphic representation associated with f, then by strong multiplicity one an inner-twist could be simply thought of as a pair (σ, χ) as above such that

$$^{\sigma}\pi_f = \pi_f \otimes \chi$$

for any extension of σ to an automorphism of $\mathbb{C}$ (see the discussion after Definition 1.2.8 for the action of σ on π_f).

In the next section we will define inner-twists for Hilbert modular forms as well and we will state many properties of inner-twists. Here, let us just mention that inner-twists form a group under the multiplication

$$(\sigma, \chi) \cdot (\tau, \eta) = (\sigma \circ \tau, \chi \cdot {}^{\sigma}\eta).$$

We denote this group by Γ. One can easily see (as we will in the next section) that the projection to the first component identifies Γ with a subgroup of $\mathrm{Aut}(E)$. Let $F = E^{\Gamma}$ be the field fixed by the inner-twists. Then E/F is a finite Galois extension. Now, assume that f is of weight 2, so that there exists a simple abelian variety A_f associated with it. The following result of Ribet demonstrates the relation between the size of the group of inner-twists Γ, and how big the endomorphism algebra of A_f is:

Theorem 3.2.2 (Ribet)**.** *Keeping the notations as above, let $X = \mathrm{End}^0_{\mathbb{Q}}(A_f)$. Then $E \simeq \mathrm{End}_{\mathbb{Q}}(A_f)$ and $E \hookrightarrow X$ is a maximal subfield of the central simple algebra X. Moreover, the center of X is identified with $F \subseteq E$ under the above embedding. In particular,* $\dim_F X = |\Gamma|^2$.

Ribet used the group of inner-twists to construct a quaternion algebra D/F which describes the image of the Galois representation associated to f and also the class of the endomorphism algebra X in the Brauer group. Momose generalized his results to higher weights. We will state the theorem now but do not give the construction of D and the proof of the theorem at this point because this will follow from our results in the next chapter for n-dimensional Galois representations. However, in the next section, we will give Nekovář's construction of D in the case of Hilbert modular forms since it is slightly different from ours.

Theorem 3.2.3 (Ribet, Momose)**.** *Let $f \in S_k(N, \epsilon)$ be a non-CM newform with Hecke field $E = \mathbb{Q}(f)$, group of inner-twists Γ and $F = E^{\Gamma}$. Let $E_p = E \otimes_{\mathbb{Q}} \mathbb{Q}_p$ and also $F_p = F \otimes_{\mathbb{Q}} \mathbb{Q}_p$ as usual. Then there exists a quaternion algebra D over F that can be realized inside $M_2(E)$ (as an F-algebra) and an open normal subgroup $H \subseteq \Gamma_{\mathbb{Q}}$ such that for every prime p one has*

$$\rho_{f,p}(H) \subseteq D^{\times}(F_p) := (D \otimes_F F_p)^{\times} \subseteq \mathrm{GL}_2(E_p).$$

Moreover, $\rho_{f,p}(H)$ is open in the p-adic Lie group

$$\{x \in D^{\times}(F_p) | \mathrm{Nrd}(x) \in \mathbb{Q}_p^{\times}\}.$$

Note that just like the open image theorem of Serre, this theorem is giving a precise description of the image of the Galois representation associated with f, up to p-adic openness. In the $k = 2$ case, Ribet also clarifies the relation of this quaternion algebra with the abelian variety A_f:

Theorem 3.2.4 (Ribet)**.** *Keeping the notations from the last theorem, assume* $k = 2$ *and* $X = \mathrm{End}^0_{\mathbb{Q}}(A_f)$*. Then one has*

$$[X] = [D] \in \mathrm{Br}(F) = H^2(\Gamma_F, \overline{F}^\times).$$

3.2.2 Hilbert Modular Forms

Nekovář generalized the work of Ribet and Momose to the case of Hilbert modular forms in Appendix B of [27]. When everything is done in the right way, the arguments are similar for the most part. One difference is that in the parallel weight 2 case, where one expects the existence of an abelian variety associated with the form, the construction of such an abelian variety is not known in general.

Let $K/\mathbb{Q}$ be a totally real field of degree d, and f be a non-CM Hilbert modular newform over K of weight $(k_1, \cdots, k_d)$ and level $\mathcal{N} \trianglelefteq O_K$ and we assume that all of the k_i's have the same parity. Equivalently, we can consider a non-CM cuspidal automorphic representation π_f of $\mathrm{GL}_2(\mathbb{A}_K)$ of level $K_1(\mathcal{N}) \trianglelefteq \mathrm{GL}_2(O_K \otimes \widehat{\mathbb{Z}})$ that is a discrete series representation of weight k_i at the infinite place v_i, where $\{v_1, \cdots, v_d\}$ are the Archimedean places of K. Assume that $E = \mathbb{Q}(f)$ is the Hecke field and $\omega = |\cdot|^m \varphi$ is the central character of π_f, with φ is a finite character. Following Nekovář [27], we define the inner-twists:

Definition 3.2.5. Keeping the notations as above, an inner-twist of f is a pair (σ, χ) where $\sigma : E \hookrightarrow \mathbb{C}$ is an embedding and χ is a Hecke character such that

$$^{\sigma}\pi_f = \pi_f \otimes \chi.$$

We could equivalently define inner-twists using the q-expansion of f. Let $\mathfrak{d}$ be the different ideal of the number field K, $z = (z_1, \cdots, z_d)$ be a point in $\mathbb{H}^d$, and

$$f(z) = \sum_{\substack{\mathfrak{n} = u \cdot \mathfrak{d}^{-1} \\ u \gg 0}} a_{\mathfrak{n}}(f) q^{\mathrm{tr}(u \cdot z)}$$

be the q-expansion of f. Then a pair (σ, χ) as above is an inner-twist if and only if

$$^{\sigma}a_{\mathfrak{p}}(f) = a_{\mathfrak{p}}(f) \cdot \chi(\mathfrak{p})$$

for almost all prime ideals $\mathfrak{p}$ of O_K.

Proposition 3.2.6. *Let (σ, χ) be an inner-twist of f. Then one has*

1. $\chi^2 = {}^\sigma\omega/\omega = {}^\sigma\varphi/\varphi$ *and hence it is finite.*
2. $\chi = \varphi^r \mu$ *for $r \in \mathbb{Z}$ and a quadratic character μ.*
3. $\mathrm{Im}(\chi) \subseteq E$.
4. $\sigma(E) \subseteq E$, *hence* $\sigma \in \mathrm{Aut}(E)$.

Proof. See [27, B.3.2]. □

If (σ, χ) and (σ, χ') are two inner-twists of f, then one has

$$\pi_f = \pi_f \otimes \chi'\chi^{-1}$$

which implies $\chi = \chi'$ since f is non-CM. This means that for an inner-twist (σ, χ), the character χ is uniquely determined by σ. This motivates us to use the notation (σ, χ_σ) for an inner-twist sometimes. This also implies that forgetting χ_σ embeds the set of inner-twists into the automorphisms of E. We usually use this identification without warning from now on.

Proposition 3.2.7. *Let Γ be the set of inner-twists of f identified with a subset of* $\mathrm{Aut}(E)$. *Then one has:*

1. Γ *is a group under the multiplication*

$$(\sigma, \chi) \cdot (\tau, \eta) = (\sigma \circ \tau, \chi \cdot {}^\sigma\eta).$$

2. $\Gamma \subseteq \mathrm{Aut}(E)$ *is an abelian subgroup.*
3. *Let $F := E^\Gamma$. Then $\Gamma = \mathrm{Gal}(E/F)$ under the above identification.*

Proof. See [27, B.3.3]. □

Nekovář proves the results analogous to that of Ribet and Momose for the case of Hilbert modular forms. Similar to Theorem 3.2.3 one finds a quaternion algebra describing the image of the associated Galois representation. Let us explain the construction of this algebra locally. Let

$$\rho := \rho_{f,p} : \Gamma_K \to \mathrm{GL}_2(E_p)$$

be the p-adic Galois representation associated with f, where $E_p = E \otimes_{\mathbb{Q}} \mathbb{Q}_p$ as usual. Now let (σ, χ_σ) be an inner-twist of f, i.e. ${}^\sigma a_{\mathfrak{p}}(f) = a_{\mathfrak{p}}(f) \otimes \chi_\sigma(\mathfrak{p})$ for almost all prime ideals $\mathfrak{p}$ of O_K. This means that the two Galois representations ${}^\sigma\rho$ and $\rho \otimes \chi_\sigma$ have the same trace at Frobenius elements $\mathrm{Frob}_{\mathfrak{p}}$ for almost all $\mathfrak{p}$. Note that by abuse of notation, we are denoting the Galois character associated with the finite Hecke character χ_σ by the same notation. Now, since the Frobenius elements are dense by Chebotarev's theorem, we deduce that the two Galois representations ${}^\sigma\rho$ and $\rho \otimes \chi_\sigma$ have the same trace. They are both irreducible since f is cuspidal, therefore by the Brauer-Nesbitt theorem they are isomorphic. This means that there exists a matrix $\alpha_\sigma \in \mathrm{GL}_2(E_p)$ such that

$${}^\sigma\rho = \alpha_\sigma(\rho \otimes \chi_\sigma)\alpha_\sigma^{-1}.$$

Now let $H = \bigcap_{\sigma\in\Gamma} \ker(\chi_\sigma) \subseteq \Gamma_K$, which is an open subgroup of Γ_K. Since all χ_σ's are trivial on H, the above equation becomes

$${}^\sigma\rho|_H = \alpha_\sigma \cdot \rho|_H \cdot \alpha_\sigma^{-1}$$

on H. We can E_p-linearly extend $\rho|_H$ to

$$\tilde{\rho} : E_p[H] \to M_2(E_p),$$

which clearly still satisfies

$${}^\sigma\tilde{\rho} = \alpha_\sigma \cdot \tilde{\rho} \cdot \alpha_\sigma^{-1}$$

for any $\sigma \in \Gamma$. We define a twisted action of Γ on the matrix algebra $M_2(E_p)$ as follows. For any inner-twist $\sigma \in \Gamma = \mathrm{Gal}(E/F)$ and $A \in M_2(E_p)$ we define

$${}^{tw(\sigma)}A := \alpha_\sigma^{-1} \cdot {}^\sigma A \cdot \alpha_\sigma.$$

Therefore, every matrix in $\rho(H)$ satisfies ${}^{tw(\sigma)}A = A$. Now, we define the F_p-algebra D_p as

$$D_p := M_2(E_p)^{tw(\Gamma)}$$

whose group of units clearly contains $\rho(H)$.

Theorem 3.2.8 ([27], Appendix B.4)**.** *There exists a quaternion algebra D/F such that for every prime p one has $D_p = D \otimes_F F_p$. Moreover, $\rho_{f,p}(H)$ is contained and open in the p-adic Lie group*

$$\{x \in D^\times(F_p) | \mathrm{Nrd}(x) \in \mathbb{Q}_p^\times\}.$$

3.3 Explicit Formulas

3.3.1 Quer's Formula

Let $f \in S_k(N, \epsilon)$ be a non-CM newform with Hecke field $E = \mathbb{Q}(f)$, group of inner-twists Γ and $F = E^\Gamma$. As we saw in Theorem 3.2.3, there exists a quaternion algebra D/F describing the image of the Galois representation associated with f. In a more computational direction one can ask if one can explicitly compute this quaternion algebra (for example as an element in the Brauer group of F) in terms of the Fourier coefficients of f. Quer was able to find such an explicit formula for $[D] \in \mathrm{Br}(F)$ when $k = 2$ [28]. This was later generalized to arbitrary weights $k \geqslant 2$ in [20]. In this section we review Quer's formula.

From now on we assume that $k = 2$. Let $f = \sum_{n\geqslant 1} a_n q^n$ be the q-expansion of f. Let $H = \bigcap_{\sigma\in\Gamma} \ker(\chi_\sigma)$. H is clearly an open normal subgroup of $\Gamma_\mathbb{Q}$ so there exists a finite Galois extension $N/\mathbb{Q}$ such that $H = \Gamma_N$. One can show that $\mathrm{Gal}(N/\mathbb{Q})$ is a 2-torsion group and hence it is isomorphic to $(\mathbb{Z}/2\mathbb{Z})^m$ for some positive integer m. Therefore, there exist rational numbers $t_1, \cdots, t_m$ such that

$$N = \mathbb{Q}(\sqrt{t_1}, \cdots, \sqrt{t_m}).$$

Let $\sigma_1, \cdots, \sigma_m$ be an $\mathbb{F}_2$ basis for $\mathrm{Gal}(N/\mathbb{Q})$ such that

$$\sigma_i(\sqrt{t_j}) = (-1)^{\delta_{i,j}} \cdot \sqrt{t_j}.$$

Quer also needs the following lemma of Ribet in his proof:

Lemma 3.3.1 (Ribet)**.** *Let $\mathcal{P}$ be the set of all rational prime numbers. Then the field $F = E^\Gamma$ is the field generated by all but finitely many numbers of the form $a_p^2/\epsilon(p)$, i.e. one has*

$$F = \mathbb{Q}\left(\{a_p^2/\epsilon(p)\}_{p\in\mathcal{P}\setminus\mathcal{S}}\right)$$

for any finite set $S \subset \mathcal{P}$.

Now we can state Quer's main result [28, Theorem 3]. Let c_ϵ be the 2-cocycle in $\mathrm{Br}(F) = H^2(\Gamma_F, \overline{F}^\times)$ given by the formula

$$c_\epsilon(\sigma, \tau) = \sqrt{\epsilon(\sigma)}\sqrt{\epsilon(\tau)}\sqrt{\epsilon(\sigma\tau)^{-1}},$$

where we are considering ϵ as a Galois character. One easily sees that this cocycle takes values in $\{\pm 1\}$.

Theorem 3.3.2 (Quer). *Let $p_1, \cdots, p_m$ be prime numbers not dividing the level of f, such that $a_{p_i} \neq 0$ and $\sigma_i = \mathrm{Frob}_{p_i}$ in $\mathrm{Gal}(N/\mathbb{Q})$. Then one has*

$$[D] = [c_\epsilon] \left(t_1, a_{p_1}^2/\epsilon(p_1)\right) \left(t_2, a_{p_2}^2/\epsilon(p_2)\right) \cdots \left(t_m, a_{p_m}^2/\epsilon(p_m)\right)$$

in $\mathrm{Br}(F)$, where $(\cdot, \cdot)$ denotes the Hilbert symbol over F.

It follows from the Sato-Tate conjecture for f that the coefficients a_p are non-zero for a density 1 set of primes p and then the Chebotarev's density theorem guarantees the existence of $p_1, \cdots, p_m$ in the theorem.

3.3.2 Generalization to Hilbert Modular Forms

In this section we want to generalize Quer's result to the case of Hilbert modular forms of parallel weight 2 with trivial nebentype, under the condition that the degree of the base field over $\mathbb{Q}$ is an odd number. The material of this section is identical to the preprint [42] by the author.

Let K be a totally real number field such that $[K : \mathbb{Q}]$ is odd and let f be a non-CM (Hilbert) newform of parallel weight 2, level $\mathcal{N}$ where $\mathcal{N}$ is an ideal of O_K, and finite central character (nebentype) ϵ. It is well-known that in this case, one can use Shimura curves to construct an abelian variety A_f over K associated with f. Let $\mathfrak{p}$ be a non-zero prime ideal of O_K, $T_\mathfrak{p}$ be the Hecke operator at $\mathfrak{p}$, and $a_\mathfrak{p}$ the eigenvalue of $T_\mathfrak{p}$ acting on f. Let the number field $E = \mathbb{Q}(\{a_\mathfrak{p}\}_\mathfrak{p})$ be the Hecke field of f. The abelian variety A_f/K is of dimension $d = [E : \mathbb{Q}]$ and hence its ℓ-adic Tate module (after tensoring with $\mathbb{Q}$) V_ℓ is of dimension $2d$ over $\mathbb{Q}_\ell$. One can define an E-structure on this Tate module by letting $a_\mathfrak{p}$ act via the Hecke operator at $\mathfrak{p}$. This turns V_ℓ into a rank 2 free module over $E \otimes_\mathbb{Q} \mathbb{Q}_\ell$ endowed with a continuous Γ_K-action. This is the Galois representation associated with f which after a choice of basis can be written as

$$\rho_{f,\ell} : \Gamma_K \to \mathrm{Aut}_E(V_\ell) \simeq \mathrm{GL}_2(E \otimes_\mathbb{Q} \mathbb{Q}_\ell).$$

It is also well known that this Galois representation is unramified outside $\ell\mathcal{N}$ and for any unramified prime ideal $\mathfrak{p}$, the Eichler-Shimura relation implies that the characteristic polynomial of $\rho_{f,\ell}(\mathrm{Frob}_\mathfrak{p})$ is equal to $X^2 - a_\mathfrak{p} X + \epsilon(\mathfrak{p})\mathrm{Nm}(\mathfrak{p})$.

As we saw in Theorem 3.2.8, Nekovář constructs a division algebra D over a subfield F of the Hecke field E which describes the image up to p-adic openness. Just like the case of classical modular forms (Theorem 3.2.4), in the special case where one knows there is an abelian variety associated with f (in particular f is of parallel weight 2) F is equal to the center of the algebra $X := \mathrm{End}_{\overline{\mathbb{Q}}}(A_f) \otimes_{\mathbb{Z}} \mathbb{Q}$ [27, B.4.11]. Since A_f is of GL_2-type over K and f is non-CM, it is a Ribet-Pyle abelian variety, i.e. $E \simeq \mathrm{End}_K(A_f) \otimes_{\mathbb{Z}} \mathbb{Q}$ is a maximal subfield of the simple algebra X [21, Propposition 3.1]. Moreover, D and X have the same class in the Brauer group of F [27, B.4.11].

First, we need to generalize a theorem of Ribet [34, Theorem 5.5] to our situation. This is the main arithmetic input in the proof of Quer's formula. Ribet's proof works without many changes but we will repeat the arguments for the convenience of the reader and because this does not seem to be written down in the literature in this case. Then, we will generalize [34, Theorem 5.6] using the work of Chi [10]. Here some of the Galois cohomology computations become more complicated due to the fact that our base field K is not contained in the field F, whereas in the case of classical modular forms $K = \mathbb{Q}$ is contained in every number field. Therefore one needs to carefully go up and down between different fields to be able to carry out the computations.

The first step is to generalize [34, Theorem 5.5.] to the case of Hilbert modular forms. Ribet uses Faltings' theorem on isogenies (Theorem 3.1) to relate the endomorphism algebra X to the Tate module. We will do the same thing and closely follow Ribet's arguments. Choose a prime number ℓ that splits completely in E. Then one has d different embeddings $\sigma : E \to \mathbb{Q}_\ell$. Let M be a finite Galois extension of K such that all of the endomorphisms of A_f are defined over M. Now by Faltings' isogeny theorem one has

$$X \otimes_{\mathbb{Q}} \mathbb{Q}_\ell = \mathrm{End}_{\mathbb{Q}_\ell[\Gamma_M]}(V_\ell). \tag{3.1}$$

Remember that V_ℓ also carries an E-structure through the Hecke action. Every embedding σ of E into $\mathbb{Q}_\ell$ gives a $E \otimes \mathbb{Q}_\ell$ -module structure on $\mathbb{Q}_\ell$ with respect to which we can define

$$V_\sigma = V_\ell \otimes_{E \otimes \mathbb{Q}_\ell, \sigma} \mathbb{Q}_\ell,$$

which is a $\mathbb{Q}_\ell$-subspace of V_ℓ of dimension 2 that is invariant under the action of Γ_K. Now note that $a \in E$ acts on V_σ via multiplication by $\sigma(a) \in E$, hence for two different embeddings σ and τ, V_σ and V_τ have trivial intersection as subspaces of V_ℓ. This (together with obvious dimension reason) gives a decomposition

$$V_\ell = \bigoplus_{\sigma: E \hookrightarrow \mathbb{Q}_\ell} V_\sigma$$

of $\mathbb{Q}_\ell[\Gamma_K]$-modules.

The following lemma will be useful later:

Lemma 3.3.3. *For each embedding σ one has* $\mathrm{End}_{\mathbb{Q}_\ell[\Gamma_M]}(V_\sigma) = \mathbb{Q}_\ell$. *In particular, V_σ is absolutely irreducible as a Γ_M-representation.*

Proof. From (3.1) one has $X \otimes_{\mathbb{Q}} \mathbb{Q}_\ell = \mathrm{End}_{\mathbb{Q}_\ell[\Gamma_M]}(V_\ell)$. Since E is a maximal subfield of X, taking the centralizer of $E \otimes \mathbb{Q}_\ell$ of both sides one gets

$$E \otimes_{\mathbb{Q}} \mathbb{Q}_\ell = \mathrm{End}_{E\otimes\mathbb{Q}_\ell[\Gamma_M]}(V_\ell),$$

which means

$$\oplus_\sigma \mathbb{Q}_\ell = \oplus_\sigma \mathrm{End}_{\mathbb{Q}_\ell[\Gamma_M]}(V_\sigma),$$

which implies the first part. Since V_σ is semi-simple by Faltings' proof of the Tate conjecture, irreducibility follows immediately. □

For every prime $\mathfrak{p}$ of O_K not dividing $\ell\mathcal{N}$, recall that the $\mathrm{Frob}_\mathfrak{p}$ action on V_ℓ has characteristic polynomial

$$X^2 - a_\mathfrak{p} X + \epsilon(\mathfrak{p})\mathrm{Nm}(\mathfrak{p}) \in E[X].$$

Therefore, for every embedding $\sigma : E \to \mathbb{Q}_\ell$ one has

$$\mathrm{tr}(\mathrm{Frob}_\mathfrak{p} \circlearrowright V_\sigma) = \sigma(a_\mathfrak{p}) \in \mathbb{Q}_\ell.$$

Restricting the compatible family of Galois representation to Γ_M, one gets another compatible family, namely for every finite place v of M not dividing $\ell\mathcal{N}$ there is $t_v \in E$ such that

$$\mathrm{tr}(\mathrm{Frob}_v \circlearrowright V_\sigma) = \sigma(t_v) \in \mathbb{Q}_\ell.$$

Let $\Sigma_{\ell\mathcal{N}}$ be the set of finite places of M not dividing $\ell\mathcal{N}$ and $L = \mathbb{Q}(t_v : v \in \Sigma_{\ell\mathcal{N}}) \subset E$. Then one has the following:

Lemma 3.3.4. *The center of the algebra* $\mathrm{End}_{\mathbb{Q}_\ell[\Gamma_M]}(V_\ell)$ *is* $L \otimes_{\mathbb{Q}} \mathbb{Q}_\ell$.

Proof. First note that by Faltings' theorem

$$E \otimes \mathbb{Q}_\ell = \mathrm{End}_{\mathbb{Q}_\ell[\Gamma_K]}(V_\ell) \subset \mathrm{End}_{\mathbb{Q}_\ell[\Gamma_M]}(V_\ell)$$

and since $E \otimes \mathbb{Q}_\ell$ centralizes itself, it should contain the center of $\mathrm{End}_{\mathbb{Q}_\ell[\Gamma_M]}(V_\ell)$.

Semi-simplicity of V_ℓ implies that V_σ and V_τ are isomorphic as Γ_M-representations if and only if they have the same Frob_v traces for all places v of M not dividing $\ell\mathcal{N}$, or equivalently, σ and τ agree on L. Now let $\gamma : L \to \mathbb{Q}_\ell$ be an embedding and define

$$V_\gamma = \oplus_{\sigma|_L = \gamma} V_\sigma.$$

Thus one has the decomposition $V = \oplus V_\gamma$ and also since there is clearly no non-trivial endomorphism from one V_γ to another, one also has the decomposition

$$\mathrm{End}_{\mathbb{Q}_\ell[\Gamma_M]}(V_\ell) = \oplus_\gamma \mathrm{End}_{\mathbb{Q}_\ell[\Gamma_M]}(V_\gamma).$$

Now let $a \in L$. Then a acts on V_σ by $\sigma(a)$, hence it acts on the whole subspace V_γ by the scalar $\gamma(a) \in \mathbb{Q}_\ell$ which means (because of the decomposition above) it's in the center of $\mathrm{End}_{\mathbb{Q}_\ell[\Gamma_M]}(V_\ell)$. So the E-algebra structure on $\mathrm{End}_{\mathbb{Q}_\ell[\Gamma_K]}(V_\ell)$ induces this L-algebra structure on $Z(\mathrm{End}_{\mathbb{Q}_\ell[\Gamma_M]}(V_\ell))$ which means it is enough to prove

$$Z(\mathrm{End}_{\mathbb{Q}_\ell[\Gamma_M]}(V_\ell)) \simeq L \otimes \mathbb{Q}_\ell$$

as L-algebras. This is easy to check:

$$Z(\mathrm{End}_{\mathbb{Q}_\ell[\Gamma_M]}(V_\ell)) = Z(\oplus_\gamma \mathrm{End}_{\mathbb{Q}_\ell[\Gamma_M]}(V_\gamma)) = \oplus_\gamma Z(\mathrm{End}_{\mathbb{Q}_\ell[\Gamma_M]}(V_\gamma)) \simeq \oplus_\gamma \mathbb{Q}_\ell = L \otimes \mathbb{Q}_\ell.$$

□

Corollary 3.3.5. *L is the center of X, i.e. $L = F$.*

Proof. Recall that from Faltings' isogeny theorem we have

$$X \otimes_{\mathbb{Q}} \mathbb{Q}_\ell = \mathrm{End}_{\mathbb{Q}_\ell[\Gamma_M]}(V_\ell).$$

Now from the last lemma it follows that

$$L \otimes \mathbb{Q}_\ell = Z(\mathrm{End}_{\mathbb{Q}_\ell[\Gamma_M]}(V_\ell)) = Z(X \otimes_{\mathbb{Q}} \mathbb{Q}_\ell) = F \otimes \mathbb{Q}_\ell$$

which implies $F = L$. □

Lemma 3.3.6. *If $\sigma, \tau : E \to \mathbb{Q}_\ell$ are embeddings that agree on F then there exists a character $\phi : \Gamma_K \to \mathbb{Q}_\ell^\times$ such that $V_\sigma \simeq V_\tau \otimes \phi$ as representations of Γ_K.*

Proof. From the proof of Lemma 3.3.4 we know that since σ and τ agree on $F = L$, V_σ and V_τ are isomorphic as representations of Γ_M. So we can choose two bases for V_σ and V_τ such that the homomorphisms $\rho_\sigma : \Gamma_K \to GL_2(\mathbb{Q}_\ell)$ and $\rho_\tau : \Gamma_K \to GL_2(\mathbb{Q}_\ell)$ associated with V_σ and V_τ are equal on Γ_M. Now define

$$\phi(g) := \rho_\sigma^{-1}(g)\rho_\tau(g).$$

A priori ϕ is just a map $\phi : \Gamma_K \to GL_2(\mathbb{Q}_\ell)$ which is trivial on Γ_M. We want to prove that it is actually a homomorphism with values in the center (hence actually a character).

Let $g \in \Gamma_K$ and $h \in \Gamma_M$. Note that $\rho_\sigma(h) = \rho_\tau(h)$ and $\rho_\sigma(ghg^{-1}) = \rho_\tau(ghg^{-1})$ since Γ_M is normal in Γ_K. Now the following computation shows that $\phi(g) = \rho_\sigma^{-1}(g)\rho_\tau(g)$ commutes with $\rho_\tau(h)$:

$$\rho_\sigma^{-1}(g)\rho_\tau(g)\rho_\tau(h) = \rho_\sigma^{-1}(g)\rho_\tau(gh) = \rho_\sigma^{-1}(g)\rho_\tau(ghg^{-1})\rho_\tau(g)$$

$$= \rho_\sigma(g^{-1})\rho_\sigma(ghg^{-1})\rho_\tau(g) = \rho_\sigma(h)\rho_\sigma^{-1}(g)\rho_\tau(g) = \rho_\tau(h)\rho_\sigma^{-1}(g)\rho_\tau(g).$$

Now since $\mathrm{End}_{\mathbb{Q}_\ell[\Gamma_M]}(V_\tau) = \mathbb{Q}_\ell$ we are done. □

Corollary 3.3.7. *Using the notation of the last lemma, $\phi^2 = \frac{^\sigma\epsilon}{^\tau\epsilon}$ and for any prime $\mathfrak{p}$ of K of good reduction for A_f, one has*

$$\sigma(a_\mathfrak{p}) = \phi(\mathrm{Frob}_\mathfrak{p})\tau(a_\mathfrak{p}).$$

Proof. Note that $V_\sigma \simeq V_\tau \otimes \phi$. Taking determinants of both sides one gets the first part and taking trace one gets the second part. □

Ribet also proves that in the $K = \mathbb{Q}$ case, the field F is generated by $\{a_p^2/\epsilon(p)\}_{p\nmid\mathcal{N}}$. This is also true in our case. In fact, by [27, B.4.11] F is exactly the field fixed by inner-twists and the above result is known in much more generality in this context by the results of [12].

Proposition 3.3.8 ([12], Corollary 4.12)**.** *The field F is generated over $\mathbb{Q}$ by the numbers $a_\mathfrak{p}^2/\epsilon(\mathfrak{p})$ for $\mathfrak{p} \nmid \mathcal{N}$.*

If $\sigma \in \Gamma_K$, then σ acts on the $\overline{\mathbb{Q}}$-endomorphisms of A_f by $\sigma(\phi)(x) := \sigma(\phi(\sigma^{-1}x))$ and this linearly extends to an action on X. E is clearly invariant under the action of Γ_K on X (we are identifying E with the maximal subfield of X). Since this is an automorphism of F-algebras, By the Skolem-Noether theorem the action of σ is given by conjugation by some element $\alpha(\sigma) \in X$. Since E is invariant under the Galois action, $\alpha(\sigma)$ commutes with E and therefore $\alpha(\sigma) \in E$ because E is a maximal subfield and hence its own centralizer. The next theorem relates the map α which is of geometric (motivic) nature to the (automorphic) data of Hecke eigenvalues.

Theorem 3.3.9. *For every $\sigma \in \Gamma_K$ one has $\alpha(\sigma)^2/\epsilon(\sigma) \in F^\times$. Moreover, for every prime ideal $\mathfrak{p}$ of O_K away from $\ell\mathcal{N}$, if $a_\mathfrak{p} \neq 0$ then $\alpha(\mathrm{Frob}_\mathfrak{p}) \equiv a_\mathfrak{p}$ modulo $F^\times$.*

Proof. As usual, let ℓ be a prime number that splits completely in E. It enough to prove that for every pair of embeddings σ and τ of E in $\mathbb{Q}_\ell$ that agree on F one has $\sigma(\alpha^2/\epsilon) = \tau(\alpha^2/\epsilon)$.

Now if σ and τ agree on F then by Lemma 3.3.6 there exists a character $\phi : \Gamma_K \to \mathbb{Q}_\ell^\times$ such that $V_\sigma \simeq V_\tau \otimes \phi$ as representations of Γ_K. In particular, it implies that as 1-dimensional representations of Γ_K one has

$$\mathrm{Hom}_{\mathbb{Q}_\ell[\Gamma_M]}(V_\sigma, V_\tau) \simeq \phi.$$

Also note that if the embeddings σ and τ do not agree on F then they are not isomorphic as Γ_M-representations and hence

$$\mathrm{Hom}_{\mathbb{Q}_\ell[\Gamma_M]}(V_\sigma, V_\tau) = 0.$$

Therefore we can completely understand $X \otimes \mathbb{Q}_\ell$:

$$\mathrm{End}_M^0(A_f) \otimes \mathbb{Q}_\ell \simeq \mathrm{End}_{\mathbb{Q}_\ell[\Gamma_M]}(\oplus V_\gamma) = \oplus_{\sigma,\tau}\mathrm{Hom}_{\mathbb{Q}_\ell[\Gamma_M]}(V_\sigma, V_\tau).$$

Now remember that on the LHS, $g \in \Gamma_K$ acts via conjugation by $\alpha(g)$. Hence, it acts on V_σ and V_τ by $\sigma(\alpha(g))$ and $\tau(\alpha(g))$ respectively. Now assume that σ and τ agree on F. Then g acts on $\mathrm{Hom}_{\mathbb{Q}_\ell[\Gamma_M]}(V_\sigma, V_\tau)$ by $\sigma(\alpha(g))/\tau(\alpha(g))$. On the other hand as a representation of Γ_K this is just ϕ, so $\sigma(\alpha(g))/\tau(\alpha(g)) = \phi(g)$. Since $\phi^2 = \frac{\sigma\epsilon}{\tau\epsilon}$ one deduces that $\sigma(\alpha^2/\epsilon) = \tau(\alpha^2/\epsilon)$ and the result follows.

For the second part, first notice that

$$\phi(\mathrm{Frob}_\mathfrak{p}) = \sigma(\alpha(\mathrm{Frob}_\mathfrak{p}))/\tau(\alpha(\mathrm{Frob}_\mathfrak{p})) = \sigma(a_\mathfrak{p})/\tau(a_\mathfrak{p}),$$

therefore

$$\sigma(\alpha(\mathrm{Frob}_\mathfrak{p})/a_\mathfrak{p}) = \tau(\alpha(\mathrm{Frob}_\mathfrak{p})/a_\mathfrak{p})$$

which implies the result. □

Our second goal is to prove an analogue of [34, Theorem 5.6] in our setting. Ribet uses a result of Chi to prove this theorem. In [10], Chi studies the twists of a central simple algebra by a 1-cocycle. We need to review some of his results and generalize some of those to our setting.

First note that the endomorphism ring $\mathrm{End}_{\overline{\mathbb{Q}}}(A_f)$ acts on the space of differential 1-forms on $A_f/\overline{\mathbb{Q}}$ (which we denote by $\Omega^1_{\overline{\mathbb{Q}}}$) via pull back. For an endomorphism ϕ and a 1-form ω we use the usual notation $\phi^*\omega$ for this action. This action linearly extends to an action of X on this space and we use the same notation for this action as well. Also, note that for any $\sigma \in \Gamma_K$ and $\phi \in X$ one has

$$\sigma(\phi^*\omega) = (\sigma\phi)^*(\sigma\omega) = (\alpha(\sigma)\cdot\phi\cdot\alpha(\sigma)^{-1})^*(\sigma\omega). \quad (3.2)$$

For σ and τ in Γ_K, define $c_\alpha(\sigma,\tau) := \alpha(\sigma)\alpha(\tau)\alpha(\sigma\tau)^{-1}$. This commutes with every element in X so it lands in F. Therefore α gives a well-defined group homomorphism

$$\alpha_K : \Gamma_K \to \frac{E^\times}{F^\times}.$$

Let α_{FK} be the restriction of α_K to Γ_{FK}. We sometimes use the same notation to denote the composition of this map with the canonical map to $(EK)^\times/(FK)^\times$:

$$\alpha_{FK} : \Gamma_{FK} \to \frac{E^\times}{F^\times} \to \frac{(EK)^\times}{(FK)^\times}.$$

Let $X_{FK} := X \otimes_F FK$. This is an algebra over FK. Note that every element in FK is a sum of the form $\sum_i f_i k_i$ for $f_i \in F$ and $k_i \in K$, so X_{FK} is generated by pure tensors of the form $\sum_i \phi_i \otimes k_i$ for $k_i \in K$.

As in [10] one can look at the twist of this algebra with (the 1-cocycle defined by) α which we denote by $X_{FK}(\alpha_{FK})$ following Chi.

Proposition 3.3.10 ([10], Proposition 1.1)**.** *One has*

$$\dim_{FK} X_{FK}(\alpha_{FK}) = \dim_{FK} X_{FK} = dim_F X.$$

Moreover

$$X_{FK}(\alpha_{FK}) \otimes_{FK} \overline{\mathbb{Q}} \simeq X_{FK} \otimes_{FK} \overline{\mathbb{Q}}.$$

Therefore, $X_{FK}(\alpha_{FK})$ is a central simple FK-algebra.

One can also view X_{FK} as a K-algebra and twist it with α_K instead to get the K-algebra $X_{FK}(\alpha_K)$. Let us recall the definition of this algebra. First for any $\sigma \in \Gamma_K$ we define the twisted action of σ on $X_{FK} \otimes_K \overline{\mathbb{Q}}$ as follows. On pure tensors of the form $\phi \otimes k \otimes \lambda$ for $\phi \in X$, $k \in K$ and $\lambda \in \overline{\mathbb{Q}}$ we define:

$$tw(\sigma) \cdot (\phi \otimes k \otimes \lambda) := \alpha(\sigma)\phi\alpha(\sigma)^{-1} \otimes k \otimes \sigma(\lambda).$$

Note that $k = \sigma(k)$ in the above expression. Now we define

$$X_{FK}(\alpha_K) := (X_{FK} \otimes_K \overline{\mathbb{Q}})^{tw(\Gamma_K)}.$$

This K-algebra also has the structure of an FK-algebra via $a \cdot \sum \psi_i \otimes \lambda_i := \sum a\psi_i \otimes \lambda_i$ for $a \in FK$, $\psi_i \in X_{FK}$ and $\lambda_i \in \overline{\mathbb{Q}}$.

Proposition 3.3.11 ([10], Proposition 1.2)**.** *One has $X_{FK}(\alpha_{FK}) \simeq X_{FK}(\alpha_K)$ as FK-algebras.*

This implies that $X_{FK}(\alpha_K) = (X_{FK} \otimes_K \overline{\mathbb{Q}})^{tw(\Gamma_K)}$ is also a central simple FK-algebra. From now on we simply write $X_{FK}(\alpha)$ for this central simple algebra.

E is a subfield of X_{FK}. Let L be a maximal subfield of X_{FK} containing E. Then L contains EK as well. So one can look at α_{FK} as a group homomorphism

$$\alpha_{FK} : \Gamma_{FK} \to \frac{L^\times}{(FK)^\times}$$

which has values in E. Now one can apply [10, Proposition 2.4] to get

$$X_{FK}(\alpha_{FK}) \otimes_{FK} \text{End}_{FK} L \simeq X_{FK} \otimes_{FK} \text{End}_{FK} L(\alpha_{FK}).$$

We can conclude that in the Brauer group Br(FK) one has

$$[X_{FK}(\alpha)] = [X_{FK}] + [\text{End}_{FK} L(\alpha_{FK})].$$

From this point onward, we assume that the central character ϵ of f is trivial for simplicity. In the general case, one also needs to carry the 2-cocycle $c_\epsilon = [\text{End}_{FK} L(\epsilon)]$ in the computations. Now we can prove:

Lemma 3.3.12. *Assuming ϵ is trivial, the order of $[X_{FK}(\alpha)]$ in* Br(FK) *divides 2.*

Proof. So far we proved

$$[X_{FK}(\alpha)] = [X_{FK}] + [\text{End}_{FK} L(\alpha_{FK})]$$

in Br(FK). By [27, Proposition B.4.12] we know that X and hence X_{FK} have Schur index dividing 2. Also from Theorem 3.3.9 we know that $\alpha^2 = \epsilon$ modulo $F^\times$. Applying [10, Proposition 2.2] we get

$$2 \cdot [\text{End}_{FK} L(\alpha_{FK})] = [\text{End}_{FK} L(\alpha_{FK}^2)] = [\text{End}_{FK} L(\epsilon)].$$

Since $\epsilon = 1$ we are done. □

From Section 2 of [10] we know that the class $[\text{End}_{FK} L(\alpha_{FK})]$ in the Brauer group Br$(FK) = H^2(\Gamma_{FK}, \overline{\mathbb{Q}})$ is the same as the image of the cohomology class defined by α in $H^1(\Gamma_{FK}, \text{PGL}_n(\overline{\mathbb{Q}}))$ under the connecting homomorphism

$$\delta : H^1(\Gamma_{FK}, \text{PGL}_n(\overline{\mathbb{Q}})) \to H^2(\Gamma_{FK}, \overline{\mathbb{Q}})$$

where $n = [L : FK]$. More concretely, one can view every $\ell \in L$ as an FK-linear endomorphism $\ell : L \to L$ given by multiplication by ℓ. So every ℓ can be viewed as an

$n \times n$ matrix with FK-entries. Now viewing every $\alpha(\sigma) \in E$ as such a matrix, conjugation by this matrix gives an element in $\mathrm{PGL}_n(FK) \subset \mathrm{PGL}_n(\overline{\mathbb{Q}})$. This gives a 1-cocycle with $\mathrm{PGL}_n(FK)$ or rather with $\mathrm{PGL}_n(\overline{\mathbb{Q}})$ values that is invariant under the Γ_{FK} action. Since the connecting homomorphism δ sends a 1-cocycle f to $f(\sigma)\sigma(f(\tau))f(\sigma,\tau)^{-1}$, one concludes:

Corollary 3.3.13. *Let $c_\alpha(\sigma,\tau) = \alpha(\sigma)\alpha(\tau)\alpha(\sigma\tau)^{-1}$ be a 2-cocycle for the trivial action of Γ_K on $F^\times$. Then the image of $[c_\alpha]$ under the sequence*

$$H^2(\Gamma_K, F^\times) \xrightarrow{\mathrm{res}} H^2(\Gamma_{FK}, F^\times) \xrightarrow{\iota_*} H^2(\Gamma_{FK}, \overline{\mathbb{Q}}^\times)$$

is exactly the class of $[X_{FK}(\alpha)]$ in $H^2(\Gamma_{FK}, \overline{\mathbb{Q}}) = \mathrm{Br}(FK)$.

Corollary 3.3.14. *In $\mathrm{Br}(FK)$ one has*

$$[X_{FK}(\alpha)] = [X_{FK}] + \iota_*(\mathrm{res}([c_\alpha])).$$

Our next goal is to prove that $X_{FK}(\alpha)$ is trivial in the Brauer group. The main ingredient is the next proposition.

Proposition 3.3.15. *$X_{FK}(\alpha)$ acts (linearly) on Ω^1_K.*

Proof. First, we define an action of $X_{FK} \otimes \overline{\mathbb{Q}}$ on $\Omega^1_{\overline{\mathbb{Q}}}$ by extending the action of X linearly, namely we define

$$(\phi \otimes k \otimes \lambda)^*\omega := k\lambda\phi^*\omega$$

for $\phi \in X$, $k \in K$ and $\lambda \in \overline{\mathbb{Q}}$. Now using (3.2) one easily sees that for any $\sigma \in \Gamma_K$ and $\psi \in X_{FK} \otimes \overline{\mathbb{Q}}$,

$$\sigma(\psi^*\omega) = (tw(\sigma) \cdot \psi)^*\sigma\omega$$

which means that if ψ is invariant under the twisted Galois action and ω is invariant under the usual Galois action, then $\psi^*\omega$ is also invariant. This means that $X_{FK}(\alpha)$ acts on Ω^1_K. □

Proposition 3.3.16. *$X_{FK}(\alpha) \in \mathrm{Br}(FK)$ is trivial.*

Proof. Let $X_{FK}(\alpha) = M_n(D)$ for some division algebra D over FK of dimension s^2. By Corollary 3.3.12 one has $s|2$. Now $\dim X_{FK}(\alpha) = n^2s^2$ which should be equal to the dimension of X over F, therefore $ns = [E : F]$. By Proposition 3.3.15, Ω^1_K is an

$M_n(D)$-module. So there is a D-vector space W such that $\Omega^1_K \simeq W^n$. The dimension of Ω^1_K over K is equal to the dimension of the abelian variety A_f, which is $[E:\mathbb{Q}]$. Hence

$$s^2 = \dim_{FK} D | \dim_{FK} W = \frac{[E:\mathbb{Q}]}{n[FK:K]} = \frac{ns[F:\mathbb{Q}]}{n[F:F\cap K]} = s[F\cap K:\mathbb{Q}].$$

This implies $s | [F\cap K:\mathbb{Q}]$, but since $s|2$ and $[K:\mathbb{Q}]$ is odd, one has $s=1$. □

From Proposition 3.3.16 and Corollary 3.3.14 and the fact that $[X_{FK}] \in \mathrm{Br}(FK)$ has order dividing 2, one deduces:

Corollary 3.3.17. *In* $\mathrm{Br}(FK)$ *one has*

$$[X_{FK}] = \iota_*(\mathrm{res}([c_\alpha])).$$

Now we need to go down from FK to F to compute the class $[X]$ in $\mathrm{Br}(F)$ using α. We can use the corestriction map to do so. First, note that by the last corollary we know that in the following diagram, the image of $[c_\alpha]$ in $H^2(\Gamma_{FK}, \overline{F}^*)$ is $[X_{FK}]$ which is the image of $[X]$ under the restriction.

$$\begin{array}{ccc} [c_\alpha] \in H^2(\Gamma_K, F^*) & \xrightarrow{\text{res}} & H^2(\Gamma_{FK}, F^*) \\ & \overset{\text{cor}}{\swarrow} & \downarrow \iota_* \\ [X] \in H^2(\Gamma_F, \overline{F}^*) & \xrightarrow[\text{res}]{} & H^2(\Gamma_{FK}, \overline{F}^*) \end{array}$$

This means that

$$\iota_*(\mathrm{res}([c_\alpha])) = \mathrm{res}([X]).$$

On the other hand, cor ∘ res $= [FK:F] = [K:F\cap K]$ which is an odd integer. Since X has order dividing 2 in the Brauer group, $\mathrm{cor}(\mathrm{res}([X])) = X$.

Finally, we can conclude the generalization of [34, Theorem 5.6] to the case of Hilbert modular form (with trivial central character):

Corollary 3.3.18. *In* $\mathrm{Br}(F)$ *one has*

$$[X] = \mathrm{cor}(\iota_*(\mathrm{res}([c_\alpha]))).$$

Now we have all the ingredients to generalize [28]. The proof is essentially the same. First notice that from Theorem 3.3.9 (and the assumption $\epsilon = 1$) we know that α^2 is trivial, hence

$$\alpha^2 : \Gamma_K \to F^\times/(F^\times)^2$$

is a homomorphism. Let N be the finite Galois extension of K associated with its kernel, i.e. $\ker(\alpha^2) = \Gamma_N$. Since $\mathrm{Gal}(N/K) \simeq \mathrm{Im}(\alpha^2) \subset F^\times/(F^\times)^2$ is a 2-torsion group, one has $\mathrm{Gal}(N/K) \simeq (\mathbb{Z}/2\mathbb{Z})^m$ for some positive integer m. Therefore, $N = K(\sqrt{t_1}, ..., \sqrt{t_m})$ for some $t_i \in K$ and if one defines $\sigma_i \in \mathrm{Gal}(N/K)$ with the relations

$$\sigma_i(\sqrt{t_j}) = (-1)^{\delta_{i,j}}\sqrt{t_j},$$

then $\sigma_1, ..., \sigma_m$ form an $\mathbb{F}_2$-basis for $\mathrm{Gal}(N/K)$.

Lemma 3.3.19. *In* $\mathrm{Br}(FK)$ *on has*

$$\iota_*(\mathrm{res}([c_\alpha])) = (t_1, \alpha(\sigma_1)^2)(t_2, \alpha(\sigma_2)^2) \cdots (t_m, \alpha(\sigma_m)^2),$$

where $(a, b) = (a, b)_{FK}$ *denotes the Hilbert symbol.*

Proof. First notice that since $\alpha(\sigma)\sigma(\alpha(\tau))\alpha(\sigma\tau)^{-1}$ is a coboundary, the 2-cocycle $[c_\alpha]$ is also given by the formula $(\sigma, \tau) \mapsto \frac{\alpha(\tau)}{\sigma(\alpha(\tau))}$.

For each $\tau \in \Gamma_K$ let

$$\tau(\sqrt{t_i}) = (-1)^{x_i(\tau)}\sqrt{t_i}.$$

Then $x_i : \Gamma_K \to \mathbb{Z}/2\mathbb{Z}$ is clearly a group homomorphism. Similarly let $y_i : \Gamma_{FK} \to \mathbb{Z}/2\mathbb{Z}$ be the homomorphism given by

$$\sigma(\alpha(\sigma_i)) = (-1)^{y_i(\sigma)}\alpha(\sigma_i)$$

for $\sigma \in \Gamma_{FK}$. Now since $\{\sigma_i\}_{i=1}^m$ provides an $\mathbb{F}_2$ basis for $\mathrm{Gal}(N/K)$, every element $\tau \in \Gamma_K$ can be written as $\eta \prod_{i=1}^m \sigma_i^{x_i(\tau)}$ where η is in $\Gamma_N = \ker(\alpha^2)$. Applying α^2 to both sides one gets

$$\alpha^2(\tau) \equiv \prod_{i=1}^m \alpha^2(\sigma_i)^{x_i(\tau)} \pmod{F^{\times 2}}$$

which implies

$$\alpha(\tau) = \lambda \prod_{i=1}^m \alpha(\sigma_i)^{x_i(\tau)}$$

for some $\lambda \in F^\times$. Now one can use this to give a description of $[c_\alpha]$. Applying $\sigma \in \Gamma_{FK}$ to the both sides one has

$$\sigma(\alpha(\tau)) = \lambda \prod_{i=1}^m \sigma(\alpha(\sigma_i))^{x_i(\tau)} = \lambda \prod_{i=1}^m (-1)^{y_i(\sigma)x_i(\tau)}\alpha(\sigma_i)^{x_i(\tau)} = \alpha(\tau) \prod_{i=1}^m (-1)^{y_i(\sigma)x_i(\tau)}$$

which gives the description

$$\prod_{i=1}^m (-1)^{y_i(\sigma)x_i(\tau)}$$

for $\iota_*(\mathrm{res}([c_\alpha]))$ in $\mathrm{Br}(FK)$. Now, it is well-known that the 2-cocycle $(-1)^{y_i(\sigma)x_i(\tau)}$ is represented by the Hilbert symbol $(t_i, \alpha^2(\sigma_i))$ so we are done. □

From [27] we know that $\Gamma \simeq \text{Gal}(E/F)$ is the group of inner-twists of the form f. Namely, for each $\sigma \in \text{Gal}(E/F)$ there exists a unique character $\chi_\sigma : \Gamma_K \to \mathbb{C}^\times$ such that $\chi_\sigma \otimes f = {}^\sigma f$. This is equivalent to saying that for every finite place $\mathfrak{p}$ of K not dividing N one has

$$\chi_\sigma(\text{Frob}_\mathfrak{p}) \cdot a_\mathfrak{p} = \sigma(a_\mathfrak{p}),$$

where $a_\mathfrak{p}$ is the $\mathfrak{p}$'th Fourier coefficient (Hecke eigenvalue) of f.

Lemma 3.3.20. *The characters χ_σ appearing in the inner-twists are exactly characters of Γ_K that factor through* Gal(N/K). *In particular, the number of inner-twists of f is 2^m.*

Proof. First, we prove that all χ_σ's are trivial on $\Gamma_N = \ker(\alpha^2)$. The Sato-Tate conjecture for Hilbert modular forms is known by [2]. This implies that the set of prime ideals $\mathfrak{p}$ of O_K for which $a_\mathfrak{p} \neq 0$ has density 1. Then by Chebotarev's density theorem the Frobenius elements of these primes are dense in Γ_K, therefore it is enough to check that χ_σ is trivial on the elements of the form $\text{Frob}_\mathfrak{p} \in \Gamma_K$ that are in the kernel of α^2 and $a_\mathfrak{p} \neq 0$.

Now if $a_\mathfrak{p} \neq 0$ then by Theorem 3.3.9, $\alpha^2(\text{Frob}_\mathfrak{p}) \equiv a_\mathfrak{p}^2$ modulo $F^{\times 2}$. Hence, if $\text{Frob}_\mathfrak{p} \in \ker(\alpha^2)$ then $a_\mathfrak{p} \in F$. This implies that $\chi_\sigma(\text{Frob}_\mathfrak{p}) = 1$ by the definition of an inner-twist. So we are done.

To prove that these are all such characters, it is enough to prove that the number of character factoring through Gal(N/K) is equal to the number of the inner-twists. The group of character factoring through Gal(N/K) is the dual group of Gal(N/K) and since this is abelian it has exactly $\frac{1}{2^m}$ elements. Then by Chebotarev's density theorem the density of primes $\mathfrak{p}$ (with $a_\mathfrak{p} \neq 0$) for which $\text{Frob}_\mathfrak{p} \in \Gamma_N$ or equivalently $a_\mathfrak{p} \in F^\times$ is $\frac{1}{2^m}$.

Now, notice that if (σ, χ_σ) is an inner-twist then by definition $\chi_\sigma(\text{Frob}_\mathfrak{p}) \cdot a_\mathfrak{p} = \sigma(a_\mathfrak{p})$. So all χ_σ's are trivial on $\text{Frob}_\mathfrak{p}$ if and only if $a_\mathfrak{p} \in F$. Also, since $a_\mathfrak{p}^2 \in F$ for all $\mathfrak{p}$, $\chi_\sigma^2 = 1$. By [27, Proposition B.3.3] Γ is a finite 2-torsion abelian group. Hence, $\Gamma \simeq (\mathbb{Z}/2\mathbb{Z})^n$ for some n. Clearly, $n \leqslant m$ since all χ_σ's factor through Gal(N/K). Now choose an $\mathbb{F}_2$ basis $\sigma^{(1)}, \cdots, \sigma^{(n)}$ for $\Gamma = \text{Gal}(E/F)$. Let Γ_M be the intersection of kernel of all χ_σ's which is equal to the intersection of the kernel of all $\chi_{\sigma^{(i)}}$'s. Now by Chebotarev's density theorem $M = N$ because they contain the same Frobenius elements of Γ_K. Since Γ_N is the intersection of the kernels of $\chi_{\sigma^{(i)}}$'s which are all of order 2, one deduces that $n \geqslant m$. This implies $n = m$ and we are done.

□

By the last lemma, the group form by the characters χ_σ is the dual group of the group $\mathrm{Gal}(N/K) \simeq (\mathbb{Z}/2\mathbb{Z})^m$. Recall that $\{\sigma_i\}_{i=1}^n$ is an $\mathbb{F}_2$ basis for $\mathrm{Gal}(N/K)$ satisfying $\sigma_i(\sqrt{t_j}) = (-1)^{\delta_{i,j}}\sqrt{t_j}$ where $N = K(\sqrt{t_1}, ..., \sqrt{t_m})$. Now let $\sigma^{(1)}, \cdots, \sigma^{(m)}$ be a dual basis for this (so each $\sigma^{(i)}$ appears in an inner-twist), i.e.

$$\sigma^{(j)}(\sigma_i) = (-1)^{\delta_{i,j}}.$$

Notice that the fixed field of $\ker(\sigma^{(j)})$ is just $K(\sqrt{t_j})$.

Recall that we need to apply the corestriction map to get back over F and find a formula for $[X]$ in $\mathrm{Br}(F)$. The following well-known lemma helps us to do that.

Lemma 3.3.21 ([41], Exercise XIV.2.4)**.** *Let L/F be a finite separable extension and let* $\mathrm{cor} : \mathrm{Br}(L) \to \mathrm{Br}(F)$ *be the corestriction map. Then for any $a \in L^\times$ and $b \in F^\times$ one has*

$$\mathrm{cor}(a, b)_L = (\mathrm{Nr}_{L/F}(a), b)_F.$$

Now we can finally state and prove our main theorem of this section. Note that for any finite place $\mathfrak{p}$ away from $\mathcal{N}$ one has $a_\mathfrak{p}^2 \in F$ by Proposition 3.3.8.

Theorem 3.3.22. *Let $\mathfrak{p}_1, \cdots, \mathfrak{p}_m$ be a set of prime ideals of O_K not dividing $\mathcal{N}$ and with $a_{\mathfrak{p}_i} \neq 0$ such that $\sigma_i = \mathrm{Frob}_{\mathfrak{p}_i}$ in $\mathrm{Gal}(N/K)$ (such primes exist by Chebotarev's theorem). Then in* $\mathrm{Br}(F)$ *one has*

$$[X] = (\mathrm{Nr}_{FK/F}(t_1), a_{\mathfrak{p}_1}^2)(\mathrm{Nr}_{FK/F}(t_2), a_{\mathfrak{p}_2}^2) \cdots (\mathrm{Nr}_{FK/F}(t_m), a_{\mathfrak{p}_m}^2),$$

where $(a, b) = (a, b)_F$ denotes the Hilbert symbol.

Proof. Using Lemma 3.3.19 one only needs to notice that $\alpha(\mathrm{Frob}_{\mathfrak{p}_i})^2 \equiv a_{\mathfrak{p}_i}^2$ modulo $F^{\times 2}$, so they only differ by a square which doesn't affect the Hilbert symbol. Therefore

$$\iota_*(\mathrm{res}([c_\alpha])) = (t_1, a_{\mathfrak{p}_1}^2)(t_2, a_{\mathfrak{p}_2}^2) \cdots (t_m, a_{\mathfrak{p}_m}^2).$$

Now one applies the corestriction map to both sides. The left hand sides gives us $[X]$ by Corollary 3.3.18 and the right hand side gives us

$$(\mathrm{Nr}_{FK/F}(t_1), a_{\mathfrak{p}_1}^2)(\mathrm{Nr}_{FK/F}(t_2), a_{\mathfrak{p}_2}^2) \cdots (\mathrm{Nr}_{FK/F}(t_m), a_{\mathfrak{p}_m}^2)$$

by the previous lemma, since $a_{\mathfrak{p}_m}^2 \in F^\times$. This proves the statement of the theorem. □

Chapter 4

Extra-Twists and Image of Galois Representations

In this chapter, we study extra-twists for automorphic representations of GL_n and use them to give a precise description of the image of the Galois representations associated with regular algebraic cuspidal automorphic representations of GL_3 over totally real fields. We also formulate a conjecture for the GL_n case and show how it follows from some standard conjectures in the Langlands program. The main difference to the case of GL_2 from last chapter is the possibility of the existence of outer-twists.

Let K be a number field and π be a cuspidal automorphic representation of $\mathrm{GL}_n(\mathbb{A}_K)$ such that

- π is not self-twist, i.e. there is no Hecke character $\chi \neq 1$ such that $\pi \simeq \pi \otimes \chi$, and
- If $n > 2$, π is not essentially self-dual, i.e. there is no Hecke character η such that $\pi \simeq \pi^\vee \otimes \eta$.

In comparison to the GL_2 case, these two conditions are analogues to a classical modular form being non-CM in the work of Ribet and Momose. It turns out that these conditions are enough in the GL_3 case to have a big image theorem for the associated Galois representation as we will see later, but of course not in the GL_n case.

Let $\mathbb{Q}(\pi) \subset \mathbb{C}$ be the Hecke (number) field of π and let E be a number field containing $\mathbb{Q}(\pi)$. In what follows, we will frequently use strong multiplicity one (Theorem 1.2.6) for cuspidal automorphic representations of GL_n without mentioning it.

4.1 Inner and Outer Twists

In this section we define extra-twists for certain automorphic representations and Galois representations.

Definition 4.1.1. An (E-)extra-twist of the automorphic representation π is either of the following two:

1. (An inner-twist) A pair (σ, χ) where $\sigma \in \mathrm{Aut}(E)$ and $\chi : \mathbb{A}_K^\times / K^\times \to \mathbb{C}^\times$ is a Hecke character, such that ${}^\sigma\pi \cong \pi \otimes \chi$.
2. (An outer-twist) A pair (τ, η) where $\tau \in \mathrm{Aut}(E)$ and $\eta : \mathbb{A}_K^\times / K^\times \to \mathbb{C}^\times$ is a Hecke character, such that ${}^\tau\pi \cong \pi^\vee \otimes \eta$.

Remark 4.1.2. We make three remarks about this definition.

(a) The role of E might seem a bit auxiliary and one might think it should be enough to take $E = \mathbb{Q}(\pi)$. But, making this slightly more general definition will help us on the Galois side when dealing with issues regarding the field of definition of automorphic Galois representations. Also, later it will be more convenient to assume that E is Galois over $\mathbb{Q}$. Since we will always fix E to begin with, we will usually drop it from the notation.

(b) Notice that the Galois action in the above definition is on the coefficients. In particular, do not confuse an outer-twist with an essential conjugate self-dual of an automorphic representation over a CM field (e.g. as in [3]).

(c) For a general reductive group, there should be a class of extra-twists for every automorphism of a (fixed) based root datum. This would also make sense on the Galois side since the automorphism group of the dual root datum is canonically isomorphic to the one for the group.

One can similarly define the notion of extra-twists for Galois representations. Let E be a number field, p a (rational) prime number, and $E_p = E \otimes_{\mathbb{Q}} \mathbb{Q}_p \cong \prod_{\mathfrak{p}|p} E_{\mathfrak{p}}$. Assume that we have a continuous irreducible Galois representation

$$\prod_{\mathfrak{p}|p} \rho_{\mathfrak{p}} = \rho : \Gamma_K \to \mathrm{GL}_n(E_p) = \prod_{\mathfrak{p}|p} \mathrm{GL}_n(E_{\mathfrak{p}}),$$

which is unramified outside a finite set of places. We also assume that ρ is neither self-twist nor essentially self dual in the $n > 2$ case, i.e. it neither satisfies $\rho \simeq \rho \otimes \chi$ for a non-trivial Galois character χ, nor $\rho \simeq \rho^\vee \otimes \eta$ for any Galois character η in the $n > 2$ case.

Definition 4.1.3. An inner-twist of ρ is a pair (σ, χ) where $\sigma \in \mathrm{Aut}(E)$ is an automorphism and $\chi : \Gamma_K \to {E_p}^\times$ is a (continuous) Galois character, such that ${}^\sigma\rho \cong \rho \otimes \chi$. An outer-twist of ρ is a pair (τ, η) where $\tau \in \mathrm{Aut}(E)$ and η a Galois character, such that ${}^\tau\rho \cong \rho^\vee \otimes \eta$. An extra-twist of ρ is either an inner- or an outer-twist.

Remark 4.1.4. Note that $\rho^\vee$ is just isomorphic to the representation ρ^{-T} and hence has coefficients in E_p (and not just $\overline{E} \otimes_{\mathbb{Q}} \mathbb{Q}_p$). This easily implies that the characters χ appearing in the extra-twists are forced to have values in $E_p^\times$ and we do not lose any generality by making this assumption in the definition.

From now on, we assume that K is totally real and π is a regular algebraic cuspidal automorphic representation. Then it is known (see [22], [38]) that there exists a compatible family of Galois representations $\rho_{\pi,p}$ associated with π. We will see in Lemma 4.3.1 that this compatible family can be defined over a coefficient field E of finite degree and Galois over $\mathbb{Q}$. Then we get a bijection between the set of E-extra-twists of π and E-extra-twists of $\rho := \rho_{\pi,p}$. Therefore, we usually identify the two.

The most basic properties of the extra-twists of ρ (or π) are summarized in the next lemma.

Lemma 4.1.5. *Let K be totally real and $\rho : \Gamma_K \to \mathrm{GL}_n(E_p)$ be a p-adic Galois representation that is neither self-twist nor essentially self-dual in the $n > 2$ case. Then extra-twists of ρ satisfy the following properties:*

(i) For an extra-twist (σ, χ), the automorphism σ uniquely determines χ.

(ii) If (σ, χ) is an inner-twist and (τ, η) an outer-twist, then $\sigma \neq \tau$.

(iii) Extra-twists form a group under the operation $(\sigma, \chi) \circ (\tau, \eta) := (\sigma \circ \tau, \chi \cdot {}^\sigma\eta)$.

(iv) Inner-twists form a subgroup of the group of all extra-twists. If at least one outer-twist exists, then this is an index 2 subgroup.

(v) If ρ is associated with an algebraic automorphic representation π, then for any inner-twist (σ, χ), the character χ is finite.

Proof. First assume that (σ, χ_1) and (σ, χ_2) are both inner-twists. Then $\rho \simeq \rho \otimes \chi_1\chi_2^{-1}$, which implies $\chi_1 = \chi_2$ by our assumptions on ρ. A similar argument proves the other cases of *(i)* and also *(ii)*. For part *(iii)*, let us assume that both (σ, χ) and (τ, η) are inner-twists. Then

$$^{\sigma}(^{\tau}\rho) \cong {}^{\sigma}(\rho \otimes \eta) = {}^{\sigma}\rho \otimes {}^{\sigma}\eta \cong \rho \otimes \chi \otimes {}^{\sigma}\eta.$$

The other cases are similar. For part *(iv)*, one just needs to note that the product of two outer-twists is clearly an inner-twist. Finally, to see *(v)*, note that since K is totally real, the central character of π must be of the form $|\cdot|^m\omega$ for some $m \in \mathbb{Z}$ and a finite order character ω. Hence, $\det(\rho) = \epsilon_p^m\omega$ where ϵ_p is the p-adic cyclotomic character and we are viewing ω as a finite Galois character. Now, taking the determinant of both sides of ${}^{\sigma}\rho \cong \rho \otimes \chi$, we get

$$\chi^n = \frac{^{\sigma}\det(\rho)}{\det(\rho)} = \frac{^{\sigma}\omega}{\omega},$$

which implies that χ^n (and hence χ) is a finite order character. □

We will denote the group of all extra-twists of a Galois representation (or an automorphic representation) by Γ, the subgroup of inner-twists by Γ^{inn}, and the set of outer-twists by Γ^{out}. Lemma 4.1.5 shows that we can identify Γ with a subgroup of $\text{Aut}(E)$ by forgetting the character and we will do so from now on. Let $F := E^{\Gamma}$ be the field fixed by all the extra-twists and $F^{\text{inn}} := E^{\Gamma^{\text{inn}}}$ be the field fixed by the inner-twists. In particular, $\Gamma = \text{Gal}(E/F)$ and $\Gamma^{\text{inn}} = \text{Gal}(E/F^{\text{inn}})$. If there is at least one outer-twist, then $[F^{\text{inn}} : F] = 2$. Otherwise, $F = F^{\text{inn}}$.

4.2 The Lie Algebra Computations

In this section we fix a p-adic Galois representation satisfying a list of natural properties (including the property that the $\overline{\mathbb{Q}}_p$-Lie algebra of the image is big) and use the extra-twists to determine the $\mathbb{Q}_p$-Lie algebra of the image of this Galois representation. In the next section, we will apply the results of this section to the Galois representations associated with certain automorphic representations. We will need to use some automorphic input to show that in the 3-dimensional case this list of properties is satisfied.

Let K be a number field as usual and Γ_K its absolute Galois group. Let E be another number field that is assumed to be Galois over $\mathbb{Q}$ and let $E_p = E \otimes_{\mathbb{Q}} \mathbb{Q}_p \cong \prod_{\mathfrak{p}|p} E_{\mathfrak{p}}$. Assume that for each finite place $\mathfrak{p}|p$ of E, we have a continuous semi-simple Galois representation $\rho_{\mathfrak{p}} : \Gamma_K \to \mathrm{GL}_n(E_{\mathfrak{p}})$. It is usually more convenient to work with the product of all these Galois representations

$$\prod_{\mathfrak{p}|p} \rho_{\mathfrak{p}} = \rho_p : \Gamma_K \to \mathrm{GL}_n(E_p) = \prod_{\mathfrak{p}|p} \mathrm{GL}_n(E_{\mathfrak{p}})$$

or equivalently a free E_p-module V_p of rank n, with a continuous Galois action on it. Then, for each $\mathfrak{p}|p$ we have an n-dimensional vector space $V_{\mathfrak{p}} = V_p \otimes_{E_p} E_{\mathfrak{p}}$ over $E_{\mathfrak{p}}$ with a continuous Galois action, such that $V_p = \bigoplus_{\mathfrak{p}|p} V_{\mathfrak{p}}$ as $\mathbb{Q}_p$-vector spaces.

Each embedding $\lambda : E \hookrightarrow \overline{\mathbb{Q}}_p$ induces an absolute value and hence gives a finite place $\mathfrak{p}$ of E above p. Therefore, λ extends to an embedding $\lambda : E_{\mathfrak{p}} \hookrightarrow \overline{\mathbb{Q}}_p$ by continuity. Now we define

$$V_\lambda := V_p \otimes_{E_p, \lambda \otimes id} \overline{\mathbb{Q}}_p = V_{\mathfrak{p}} \otimes_{E_{\mathfrak{p}}, \lambda} \overline{\mathbb{Q}}_p,$$

which is an n-dimensional vector space over $\overline{\mathbb{Q}}_p$ with a continuous Galois action. We denote this representation by ρ_λ. Note that $\rho_\lambda : \Gamma_K \to \mathrm{GL}_n(\overline{\mathbb{Q}}_p)$ is essentially the same object as $\rho_{\mathfrak{p}} : \Gamma_K \to \mathrm{GL}_n(E_{\mathfrak{p}})$, it is just considered with coefficients in $\overline{\mathbb{Q}}_p$ instead of $E_{\mathfrak{p}}$ via λ.

Now, we need to make the following list of natural assumptions on our Galois representations to be able to compute the Lie algebra of the image later. These properties are expected to hold for Galois representations attached to regular cuspidal algebraic automorphic representations of general type, after possibly a finite base change and a twist by a character.

Definition 4.2.1. Keeping the above notations, the representation $\rho_p : \Gamma_K \to \mathrm{GL}_n(E_p)$ is called valid if

- Each ρ_λ is continuous and unramified outside a finite set S of places of K containing the Archimedean places and all places above p.
- $f_v^{(p)}(x) := \mathrm{CharPoly}(\rho_p(\mathrm{Frob}_v))$ has coefficients in E, for each place $v \notin S$.
- Each ρ_λ is neither self-twist nor essentially self-dual for $n > 2$.
- $\det(\rho_p)$ is trivial.
- For each λ, the $\overline{\mathbb{Q}}_p$-Lie algebra of the p-adic Lie group $\rho_\lambda(\Gamma_K)$ is equal to $\mathfrak{sl}_n(\overline{\mathbb{Q}}_p)$.

Remark 4.2.2. We make the following four remarks about this definition:

1. Note that by the last condition, each ρ_λ is strongly irreducible, i.e. the restriction of ρ_λ to Γ_L, the absolute Galois group of L, is irreducible for any finite extension L/K. This is simply because going to a finite extension does not affect the Lie algebra. In practice, we will usually need to prove this first in order to show that a Galois representation is valid.

2. Note that for $n = 2$, we are not excluding essential self-duality, but we are excluding being self-twist.

3. The condition on the determinant is not very restrictive because we can trivialize the determinant by going to a finite extension of K and a twist. Since our first goal is to compute the semi-simple part of the Lie algebra of the image, it doesn't change anything if we restrict to an open subgroup or twist with a character.

4. In the case where the ρ_λ's come from a compatible family of semi-simple Galois representations, it is enough to check the last condition at only one λ. More precisely, by Theorem 3.19 and Remark 3.22 of [23], the semi-simple rank and the formal character of the tautological representation of the algebraic monodromy group are both independent of λ. Then [24, Theorem 4] implies that this pair uniquely determines the Lie algebra in the type A_n case, hence if we have $\mathfrak{sl}_n$ as our Lie algebra at one place λ, we should have $\mathfrak{sl}_n$ at every place.

4.2.1 Extra-Twists and Galois Representations

From now on, we assume that ρ_p is a valid Galois representation. There are two cases that we have to deal with. Namely, ρ_p either has an outer-twist or it does not. Note that in the $n = 2$ case we only have inner-twists since every representation is essentially self-dual. We try to deal with both cases at the same time but at some places it is easier to make a distinction. We need a few lemmas:

Lemma 4.2.3. *For every extra-twist (σ, χ_σ) of ρ_p, one has that χ_σ is a finite character.*

Proof. This is always true for inner-twists as we saw in Lemma 4.1.5. Let (σ, χ_σ) be an outer-twist. Then

$$^\sigma\rho_p \cong \rho_p^\vee \otimes \chi_\sigma.$$

Looking at the determinants of both sides, it follows that $\chi_\sigma^n = 1$, which implies the result. □

Lemma 4.2.4. *Let L be a finite Galois extension of K and let V_1 and V_2 be p-adic finite dimensional Γ_K-representations such that the restriction of V_2 to Γ_L is absolutely irreducible. If $V_1 \simeq V_2$ as representations of Γ_L, then $V_1 \simeq V_2 \otimes \phi$ as representations of Γ_K, for some character ϕ of Γ_K.*

Proof. First of all, we can choose two bases for V_1 and V_2 such that the representations $\rho_1 : \Gamma_K \to \mathrm{GL}_n(\overline{\mathbb{Q}_p})$ and $\rho_2 : \Gamma_K \to \mathrm{GL}_n(\overline{\mathbb{Q}_p})$ associated with V_1 and V_2 are equal when restricted to Γ_L. Now define

$$\phi(g) := \rho_1^{-1}(g)\rho_2(g).$$

A priori ϕ is just a map $\phi : \Gamma_K \to \mathrm{GL}_n(\overline{\mathbb{Q}_p})$ which is trivial on Γ_L. We want to prove that it actually has values in the group of scalar matrices; and this easily implies that ϕ is a homomorphism, hence a character.

Let $g \in \Gamma_K$ and $h \in \Gamma_L$. Note that $\rho_1(h) = \rho_2(h)$ and $\rho_1(ghg^{-1}) = \rho_2(ghg^{-1})$ since Γ_L is normal in Γ_K. Now the following computation shows that $\phi(g) = \rho_1^{-1}(g)\rho_2(g)$ commutes with $\rho_2(h)$:

$$\rho_1^{-1}(g)\rho_2(g)\rho_2(h) = \rho_1^{-1}(g)\rho_2(gh) = \rho_1^{-1}(g)\rho_2(ghg^{-1})\rho_2(g)$$

$$= \rho_1(g^{-1})\rho_1(ghg^{-1})\rho_2(g) = \rho_1(h)\rho_1^{-1}(g)\rho_2(g) = \rho_2(h)\rho_1^{-1}(g)\rho_2(g).$$

Now, since V_2 is absolutely irreducible when restricted to Γ_L, we have $\mathrm{End}_{\Gamma_L}(V_2) = \overline{\mathbb{Q}_p}$ and we are done. □

Now fix a valid Galois representation ρ_p. Note that by the third assumptions in Definition 4.2.1, it makes sense to consider the group $\Gamma \subseteq \mathrm{Aut}(E)$ of all the extra-twists of ρ_p. Let Γ^{inn}, $F = E^{\Gamma}$, and $F^{\mathrm{inn}} = E^{\Gamma^{\mathrm{inn}}}$ be as usual. By Lemma 4.1.5, the character χ in an extra-twist (σ, χ) is uniquely determined by σ, so we use the notation χ_σ for this character.

We assumed in Definition 4.2.1 that we know the $\overline{\mathbb{Q}}_p$-Lie algebra of the image. The next two lemmas are our main tool to compute the $\mathbb{Q}_p$-Lie algebra. The next lemma is the only place where we use the assumption that $E/\mathbb{Q}$ is Galois.

Lemma 4.2.5. *Let $\Gamma_L = \bigcap_{\sigma\in\Gamma} \ker(\chi_\sigma)$ and let L' be a finite extension of L that is Galois over K. For every finite unramified place v of L', let $a_v, b_v \in E$ be such that*

$$f_v(x) := \mathrm{charPoly}(\rho_p(\mathrm{Fr}_v)) = X^n - a_v X^{n-1} + \cdots + (-1)^{n-1} b_v X + (-1)^n.$$

Then, $F^{inn} = \mathbb{Q}(\{a_v\}_v)$ and $F = \mathbb{Q}(\{a_v + b_v\}_v)$.

Proof. We first prove that $f_v(x) \in F^{inn}[x]$. This is because for any inner-twist (σ, χ), the character χ becomes trivial after restricting to $\Gamma_{L'}$, so $\rho_p|_{\Gamma_{L'}} \cong {}^{\sigma}\rho_p|_{\Gamma_{L'}}$, which means that f_v is invariant under the action of Γ_{inn}, which then implies the result. Now, if (τ, η) is an outer-twist, after restriction to $\Gamma_{L'}$ one has $\rho_p|_{\Gamma_{L'}} \cong {}^{\tau}\rho_p^{-T}|_{\Gamma_{L'}}$. Looking at the characteristic polynomials of Fr_v on both sides, one gets

$$X^n - a_v X^{n-1} + \cdots + (-1)^{n-1} b_v X + (-1)^n = X^n - {}^{\tau}b_v X^{n-1} + \cdots + (-1)^{n-1\tau} a_v X + (-1)^n.$$

In particular, $a_v + b_v$ is invariant under any outer-twist τ. This implies that $a_v + b_v \in F$.

Now, let $F' = \mathbb{Q}(\{a_v\}_v) \subseteq F^{inn}$. We want to prove that F' is the field fixed by the inner-twists. This suffices because E was assumed to be Galois over $\mathbb{Q}$ and hence it is Galois over F'. Therefore, it is enough to construct an inner-twist of ρ_p for every $\sigma \in \mathrm{Gal}(\overline{\mathbb{Q}}/F')$. Note that $\rho_p|_{\Gamma_{L'}}$ has traces in F', so $\rho_p|_{\Gamma_{L'}}$ and ${}^{\sigma}\rho_p|_{\Gamma_{L'}}$ have the same traces and since they are semi-simple, they must be conjugate. Now, by Lemma 4.2.4 there exists a character such that $\rho_p \otimes \chi \simeq {}^{\sigma}\rho_p$, so we are done.

At last, let $F'' = \mathbb{Q}(\{a_v + b_v\}_v) \subseteq F^{inn}$. We want to prove that F'' is the field fixed by all extra-twists. It is enough to construct an inner-twist or an outer-twist of ρ_p, for every $\sigma \in \mathrm{Gal}(\overline{\mathbb{Q}}/F'')$. Now, note that $(\rho_p \oplus \rho_p^{-T})|_{\Gamma_{L'}}$ has traces in F'' because the trace of the image of Fr_v would be $a_v + b_v$. Therefore, $(\rho_p \oplus \rho_p^{-T})|_{\Gamma_{L'}}$ and ${}^{\sigma}(\rho_p \oplus \rho_p^{-T})|_{\Gamma_{L'}}$ have the same trace, and since they are semi-simple, they must be conjugate. Now, by the strong irreducibility assumption, we must have either $\rho_p|_{\Gamma_{L'}} \simeq {}^{\sigma}\rho_p|_{\Gamma_{L'}}$ or $\rho_p|_{\Gamma_{L'}} \simeq {}^{\sigma}\rho_p^{-T}|_{\Gamma_{L'}}$. Then by Lemma 4.2.4, we get either an inner-twist or an outer-twist. □

Lemma 4.2.6. *Let $\lambda, \mu : E \hookrightarrow \overline{\mathbb{Q}}_p$ be two places of E above p and let $\Gamma_L = \bigcap_{\sigma \in \Gamma} \ker(\chi_\sigma)$ and L' a finite extension of L that is Galois over K. Then, in the case that there are outer-twists one has*

i. $V_\lambda \simeq V_\mu$ as representations of $\Gamma_{L'}$ if and only if $\lambda|_{F^{inn}} = \mu|_{F^{inn}}$.

ii. $V_\lambda \simeq V_\mu^\vee$ as representations of $\Gamma_{L'}$ if and only if $\lambda|_F = \mu|_F$ but $\lambda|_{F^{inn}} \neq \mu|_{F^{inn}}$.

and in the case that there are no outer-twists one has that $F = F^{inn} = \mathbb{Q}(\{a_v\}_v)$, and part i of the above is true and part ii never occurs.

Proof. Since all our representations are semi-simple, it is enough to check the equality on the characteristic polynomials, and since they are continuous it is enough to check this on a dense subset. We check this on the Frobenius elements of finite places of L' at which ρ_p is unramified.

Keeping the notation of Lemma 4.2.5, the characteristic polynomial of Fr_v acting on V_λ is

$$X^n - \lambda(a_v)X^{n-1} + \cdots + (-1)^{n-1}\lambda(b_v)X + (-1)^n,$$

and on V_μ is

$$X^n - \mu(a_v)X^{n-1} + \cdots + (-1)^{n-1}\mu(b_v)X + (-1)^n.$$

Since the a_v's generate F^{inn} by Lemma 4.2.5, part *i* follows. For part *ii*, notice that the characteristic polynomial of Fr_v acting on $V_\mu^\vee$ is $X^n - \mu(b_v)X^{n-1} + \cdots + (-1)^{n-1}\mu(a_v)X + (-1)^n$. So, if $V_\lambda \simeq V_\mu^\vee$ then $\lambda(a_v) = \mu(b_v)$ and $\lambda(b_v) = \mu(a_v)$, which implies that $\lambda(a_v + b_v) = \mu(a_v + b_v)$ and hence $\lambda|_F = \mu|_F$. On the other hand, V_λ is not essentially self-dual, so by Lemma 4.2.4 its restriction to $\Gamma_{L'}$ cannot be self-dual. Thus, V_λ and V_μ are not allowed to be isomorphic as $\Gamma_{L'}$-representations in this case, which means that $\lambda|_{F^{inn}} \neq \mu|_{F^{inn}}$. The other direction also follows easily since the characteristic polynomials of Frobenius elements at unramified places clearly match. □

4.2.2 The Lie Algebra of the Image

Now, we want to compute the $\mathbb{Q}_p$-Lie algebra of the image of ρ_p. First, we need to use the results of Section 1.2.1 to construct the right algebraic group which contains the image, and then compare the Lie algebra of the image with the (algebraic) Lie algebra of this group.

Recall that we assumed that ρ_p has trivial determinant. Therefore we have

$$\rho_p : \Gamma_K \to \mathrm{SL}_n(E_p).$$

We first define a 1-cocycle $f : \Gamma \to \mathrm{Aut}_{E_p}(\mathrm{SL}_n)$ using extra-twists. For every inner-twist $\sigma \in \Gamma^{\mathrm{inn}}$ one has that $\rho_p|_{\Gamma_L}$ and ${}^\sigma\rho_p|_{\Gamma_L}$ have the same trace. Since each ρ_λ is strongly irreducible, this means that they are isomorphic and there exists $\alpha_\sigma \in \mathrm{SL}_n(E_p)$ (after possibly slightly enlarging E if necessary) such that $\rho_p|_{\Gamma_L} = \alpha_\sigma \cdot {}^\sigma\rho_p|_{\Gamma_L} \cdot \alpha_\sigma^{-1}$. For the inner-twist (σ, χ_σ), we define $f_\sigma = \mathrm{ad}(\alpha_\sigma)$. If $\tau \in \Gamma^{\mathrm{out}}$ is an outer-twist (if there exist any), then $\rho_p|_{\Gamma_L}$ and ${}^\tau\rho_p^{-T}|_{\Gamma_L}$ have the same trace and there exists $\alpha_\tau \in \mathrm{SL}_n(E_p)$ such that $\rho_p|_{\Gamma_L} = \alpha_\tau \cdot {}^\tau\rho_p^{-T}|_{\Gamma_L} \cdot \alpha_\tau^{-1}$. For the outer-twist (τ, χ_τ), we define $f_\tau = \mathrm{ad}(\alpha_\tau) \circ (\cdot)^{-T}$. One can easily check that $f : \Gamma \to \mathrm{Aut}_{E_p}(\mathrm{SL}_n)$ defined above is in fact a 1-cocycle.

Now, as in Section 1.2.1, we can define the twisted action of Γ by this cocycle on SL_n. From the construction of f, it is clear that every matrix in the image of Γ_L is invariant under this twisted action. Let $H = (\mathrm{Res}_{F_p}^{E_p}\mathrm{SL}_n)^{tw_f(\Gamma)}$ which is an algebraic group over F_p. Then it follows:

Corollary 4.2.7. *The representation $\rho_p|_{\Gamma_L}$ factors through $H(F_p) \subseteq \mathrm{SL}_n(E_p)$.*

Note that by Proposition 1.2.4, H is a form of SL_n and in particular is a semi-simple group. Also, note that since Γ_L is open in Γ_K, the Lie algebras of the p-adic Lie groups $\rho_p(\Gamma_K)$ and $\rho_p(\Gamma_L)$ are the same. Let $\mathfrak{g}$ be the Lie algebra (over $\mathbb{Q}_p$) of $\rho_p(\Gamma_L)$ and $\mathfrak{h}$ be the Lie algebra of the algebraic group H/F_p, both viewed as Lie subalgebras of the Lie algebra of $\mathrm{SL}_n(E_p)$. Our next goal is to show that these two Lie algebras are in fact equal.

Proposition 4.2.8. *With the notation as above, $\mathfrak{g} = \mathfrak{h}$.*

Proof. First, note that $\mathfrak{g} \subseteq \mathfrak{h}$ by Corollary 4.2.7. Since $\mathfrak{h}$ is semi-simple, it suffices to prove that $\overline{\mathfrak{g}^{\mathrm{der}}} = \mathfrak{g}^{\mathrm{der}} \otimes_{\mathbb{Q}_p} \overline{\mathbb{Q}_p}$ is equal to $\overline{\mathfrak{h}} = \mathfrak{h} \otimes_{\mathbb{Q}_p} \overline{\mathbb{Q}_p}$.

For every embedding $\sigma : F \hookrightarrow \overline{\mathbb{Q}}_p$, fix an extension $\widetilde{\sigma} : E \hookrightarrow \overline{\mathbb{Q}}_p$ of σ. All of the other extensions of σ can be obtained by composing with different elements of the Galois group $\Gamma = \mathrm{Gal}(E/F)$, i.e. they are all of the form $\widetilde{\sigma}\tau$ for some $\tau \in \Gamma$. Now, we base change our representation $\rho_p|_{\Gamma_L}$ to $\overline{\mathbb{Q}_p}$ to get

$$\begin{array}{ccccccc}
\rho_p : \Gamma_L & \longrightarrow & \mathrm{SL}_n(E_p) & \hookrightarrow & \mathrm{SL}_n(\overline{E_p}) & \xrightarrow{=} & \mathrm{Res}_F^E(\mathrm{SL}_n)(\overline{F_p}) \\
& \searrow & \uparrow & & \uparrow & & \uparrow \\
& & H(F_p) & \hookrightarrow & H(\overline{F_p}) & \xrightarrow{=} & \mathrm{Res}_F^E(\mathrm{SL}_n)^{tw(\Gamma)}(\overline{F_p}),
\end{array}$$

where $\overline{E_p} := E_p \otimes_{\mathbb{Q}_p} \overline{\mathbb{Q}}_p$ and $\overline{F_p} := F_p \otimes_{\mathbb{Q}_p} \overline{\mathbb{Q}}_p = \prod_{\sigma : F \hookrightarrow \overline{\mathbb{Q}}_p} \overline{\mathbb{Q}}_p$. Note that we have

$$\begin{aligned}
\mathrm{Res}_F^E(\mathrm{SL}_n)(\overline{F_p}) &\simeq \prod_{\sigma : F \hookrightarrow \overline{\mathbb{Q}_p}} \mathrm{SL}_n(E \otimes_{F,\sigma} \overline{\mathbb{Q}_p}) \simeq \prod_{\sigma : F \hookrightarrow \overline{\mathbb{Q}_p}} \mathrm{SL}_n(E \otimes_F E \otimes_{E,\widetilde{\sigma}} \overline{\mathbb{Q}_p}) \\
&\simeq \prod_{\sigma : F \hookrightarrow \overline{\mathbb{Q}_p}} \prod_{\Gamma} \mathrm{SL}_n(\overline{\mathbb{Q}_p}) \simeq \prod_{\lambda : E \hookrightarrow \overline{\mathbb{Q}_p}} \mathrm{SL}_n(\overline{\mathbb{Q}_p}),
\end{aligned}$$

where $\lambda = \widetilde{\sigma}\tau$ for $\tau \in \Gamma$. By Proposition 1.2.4, projecting to the identity component of Γ gives the isomorphism of the form $\mathrm{Res}_F^E(\mathrm{SL}_n)^{tw(\Gamma)}$ of SL_n with SL_n over E_p. So we have:

$$\begin{array}{ccccccc}
\rho_p : \Gamma_L & \longrightarrow & \mathrm{SL}_n(E_p) & \hookrightarrow & \prod_{\sigma : F \hookrightarrow \overline{\mathbb{Q}_p}} \prod_{\Gamma} \mathrm{SL}_n(\overline{\mathbb{Q}_p}) & \xrightarrow{=} & \prod_{\lambda : E \hookrightarrow \overline{\mathbb{Q}_p}} \mathrm{SL}_n(\overline{\mathbb{Q}_p}) \\
& \searrow & \uparrow & & \uparrow & & \\
& & H(F_p) & \hookrightarrow & \prod_{\sigma : F \hookrightarrow \overline{\mathbb{Q}_p}} \mathrm{SL}_n(\overline{\mathbb{Q}_p}) & &
\end{array}$$

For each embedding $\sigma : F \hookrightarrow \overline{\mathbb{Q}}_p$, the composition

$$\rho_\sigma : \Gamma_L \to H(F_p) \hookrightarrow \prod_{\sigma : F \hookrightarrow \overline{\mathbb{Q}}_p} \mathrm{SL}_n(\overline{\mathbb{Q}}_p) \xrightarrow{\mathrm{pr}_\sigma} \mathrm{SL}_n(\overline{\mathbb{Q}}_p)$$

corresponds to the action of Γ_L on the vector space V_λ, for some embedding $\lambda : E \hookrightarrow \overline{\mathbb{Q}}_p$ extending σ. Note that by Lemma 4.2.6, these V_σ's are neither isomorphic nor dual to each other after any finite extension. This is the main point of the rest of the argument.

On the level of the Lie algebras, we have the embedding

$$\overline{\mathfrak{g}^{\mathrm{der}}} \subseteq \overline{\mathfrak{h}} \xrightarrow{\simeq} \prod_{\sigma : F \hookrightarrow \overline{\mathbb{Q}}_p} \mathfrak{sl}_n(\overline{\mathbb{Q}}_p).$$

Let $\mathfrak{g}_\sigma^{\mathrm{der}} \subseteq \mathfrak{sl}_n(\overline{\mathbb{Q}}_p)$ be the projection of $\overline{\mathfrak{g}^{\mathrm{der}}}$ to the σ-component of the above map. This is the $\overline{\mathbb{Q}}_p$-Lie algebra of the image of the representation ρ_σ ($= \rho_\lambda$ for some λ extending σ), so by our assumption of ρ_p being valid, we have $\mathfrak{g}_\sigma^{\mathrm{der}} = \mathfrak{sl}_n(\overline{\mathbb{Q}}_p)$.

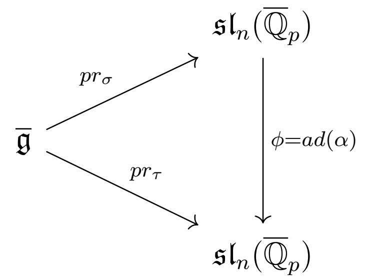
Now we can apply [33, Lemma 4.6] to

$$\overline{\mathfrak{g}^{\mathrm{der}}} \subseteq \overline{\mathfrak{h}} \to \prod_{\sigma : F \hookrightarrow \overline{\mathbb{Q}}_p} \mathfrak{sl}_n(\overline{\mathbb{Q}}_p).$$

We only need to prove that for every $\sigma, \tau : F \hookrightarrow \overline{\mathbb{Q}}_p$ the projections $(\mathrm{pr}_\sigma \times \mathrm{pr}_\tau)(\overline{\mathfrak{g}^{\mathrm{der}}})$ and $(\mathrm{pr}_\sigma \times \mathrm{pr}_\tau)(\overline{\mathfrak{h}})$ are equal. We follow the arguments of [39, §6.2].

Clearly, it is enough to show that $(\mathrm{pr}_\sigma \times \mathrm{pr}_\tau)(\overline{\mathfrak{g}^{\mathrm{der}}}) = \mathfrak{sl}_n(\overline{\mathbb{Q}}_p) \times \mathfrak{sl}_n(\overline{\mathbb{Q}}_p)$. Note that the first factor corresponds to the representation V_σ and the second to V_τ. Now by the Lie algebra version of the Goursat's theorem [31, Lemma 5.2.1], if $(\mathrm{pr}_\sigma \times \mathrm{pr}_\tau)(\overline{\mathfrak{g}^{\mathrm{der}}})$ is not equal to $\mathfrak{sl}_n(\overline{\mathbb{Q}}_p) \times \mathfrak{sl}_n(\overline{\mathbb{Q}}_p)$, then it has to be the graph of an isomorphism. Let us call this automorphism $\phi : \mathfrak{sl}_n(\overline{\mathbb{Q}}_p) \to \mathfrak{sl}_n(\overline{\mathbb{Q}}_p)$. Since $\mathfrak{sl}_n$ is simple, the group of its outer automorphisms is isomorphic to the group of the automorphisms of its Dynkin diagram, which is trivial for $n = 2$ and is isomorphic to $\frac{\mathbb{Z}}{2\mathbb{Z}}$ in the $n > 2$ case. In this case, the class of this non-trivial outer automorphism is given by the map $X \mapsto -X^T$. So ϕ is either an inner automorphism or a conjugate of this outer automorphism.

First, assume ϕ is an inner automorphism and is given by conjugation with some matrix α. In other words, we have the following diagram:



which means that V_σ and V_τ are isomorphic as representations of $\overline{\mathfrak{g}}$. This implies that they are indeed isomorphic as representations of some open subgroup of Γ_L, which contradicts Lemma 4.2.6. Now assume that ϕ is a conjugate of $X \mapsto -X^T$. Similarly, this means that $V_\sigma \cong V_\tau^\vee$ as representations of some small enough open normal subgroups of Γ_L, which again contradicts Lemma 4.2.6. This implies the result. □

Corollary 4.2.9. *The image of Γ_L under the representation ρ_p is an open subgroup of the p-adic Lie group $H(F_p)$.*

If ρ_p has no outer-twists, then the cocycle f is always defined by an inner automorphism and H is an inner-form of SL_n. If there is at least one outer-twist, then this is not true anymore, but H becomes an inner-form after a degree two extension. In fact, the restriction of f to the index two subgroup Γ^{inn} factors through $\mathrm{Inn}_{E_p}(\mathrm{SL}_n) \subset \mathrm{Aut}_{E_p}(\mathrm{SL}_n)$. Then by Corollary 1.2.5, the base change of H to $F_p^{\mathrm{inn}} = F^{\mathrm{inn}} \otimes_{\mathbb{Q}} \mathbb{Q}_p$ is an inner-form. In any case we have:

Corollary 4.2.10. *The group $H_{F_p^{\mathrm{inn}}}$, the base change of H to F_p^{inn}, is an inner-form of the group SL_n which splits over E_p. Moreover, if p splits in the extension F_p^{inn}/F_p (p splits in F^{inn} "more" than it does it F), then H is an inner-form of SL_n.*

Proof. The first part follows from the discussion above. For the second part, if there are no outer-twists, then there is nothing to prove. Otherwise, notice that when p splits, $F_p^{\mathrm{inn}} \simeq F_p \times F_p$ so if H becomes an inner-form over F_p^{inn}, it was already an inner-form over F_p. □

Remark 4.2.11. The arguments of the last two sections can be done for more general reductive groups with the right definition of extra-twists. Although, relating these to the automorphic side would be more difficult as we will see that one needs some cases of Langlands functoriality for this.

4.3 Application to Automorphic Galois Representations

In this section, we will apply the results of the previous section to Galois representations attached to certain automorphic representations. In the case of GL_2, we recover the results of Ribet, Momose, and Nekovář. Throughout this section, we assume that K is a totally real number field and π is a regular cuspidal algebraic automorphic representation of $\mathrm{GL}_n(\mathbb{A}_K)$. It is known by the work of Harris-Lan-Taylor-Thorne [22] or Scholze [38], that there is a compatible family of p-adic Galois representations associated with π. Our goal is to understand the image of these representations.

Let $|.|^m\omega$ be the central character of π where ω is a finite order Hecke character and $m \geqslant 1$ an integer. Then, for each embedding $\lambda : \mathbb{Q}(\pi) \hookrightarrow \overline{\mathbb{Q}_p}$ there exists a continuous semi-simple Galois representation

$$\rho_{\pi,\lambda} : \Gamma_K \to \mathrm{GL}_n(\overline{\mathbb{Q}_p})$$

that is an unramified Galois representation for all unramified places of π not above p, and at each such finite place, say v, the characteristic polynomial of Frob_v is determined by the Satake parameters of π at v. In other words, these representations form a compatible family of Galois representations. Moreover, $\det(\rho_{\pi,\lambda}) = (\lambda \circ \omega) \cdot \epsilon_p^m$ where ϵ_p is the (global p-adic) cyclotomic character and we are regarding the finite order character ω as a Galois character via class field theory.

It is not known wether we can conjugate these Galois representations to have values in the completions of the Hecke field $\mathbb{Q}(\pi)$. Nevertheless, we will show that we can do this for a finite extension E of $\mathbb{Q}(\pi)$. This is exactly the reason why we defined extra-twists for any Galois coefficient field containing the Hecke field in Definition 4.1.1.

The Galois representations $\rho_{\pi,\lambda}$ are expected to be irreducible since π is cuspidal. Let $t_{\pi,\lambda} : \Gamma_K \to \overline{\mathbb{Q}}_p$ be the trace of $\rho_{\pi,\lambda}$. This is an irreducible pseudo-representation and it clearly takes values in $\mathbb{Q}(\pi)_\lambda$. Then by a result of Rouquier [35, Theorem 5.1] (or more generally Chenevier [9, Corollary 2.23]), there exists a central simple algebra D_λ over $\mathbb{Q}(\pi)_\lambda$ of dimension n^2 such that this pseudo-representation can be realized as the reduced trace of a representation $\Gamma_K \to D_\lambda^\times$. The base change of this representation to $\overline{\mathbb{Q}}_p$ clearly gives back $\rho_{\pi,\lambda}$ because of the Brauer-Nesbitt Theorem. In other words, the image of $\rho_{\pi,\lambda}$ is in fact in $D_\lambda^\times$:

$$\rho_{\pi,\lambda} : \Gamma_K \to D_\lambda^\times \subset (D_\lambda \otimes_{\mathbb{Q}(\pi)_\lambda} \overline{\mathbb{Q}}_p)^\times \simeq \mathrm{GL}_n(\overline{\mathbb{Q}}_p).$$

Lemma 4.3.1 (Chenevier)**.** *Assume that $\rho_{\pi,\lambda}$ is irreducible for all λ and regular for at least one λ. Then, there exists a finite extension $E/\mathbb{Q}(\pi)$ that is Galois over $\mathbb{Q}$, such that for all finite places λ of $\mathbb{Q}(\pi)$ and any place μ of E above λ, the central simple algebra D_λ splits over E_μ. In particular, there exists a finite extension $E/\mathbb{Q}(\pi)$ such that all representations $\rho_{\pi,\lambda}$ can be defined over E (can be conjugated to have values in completions of E).*

Proof. First, recall that a central simple algebra $D_\lambda \subset M_n(\overline{\mathbb{Q}}_p)$ splits in an extension M of E_λ if and only if it contains an element with n pairwise distinct eigenvalues in M. Let v be a place of K at which π is unramified, and let $f^{(v)}(x) \in \mathbb{Q}(\pi)[x]$ be the characteristic polynomial of the Frobenius element at v, which is independent of λ. Then, as in the proof of [3, Lemma 5.3.1], choosing a λ for which $\rho_{\pi,\lambda}$ is regular, we get that for infinitely many places v, one has that $f^{(v)}(x)$ has distinct roots. This shows that if E' is the splitting field of $f^{(v)}(x)$, then D_λ splits over the completion of E' at any finite place coprime to v and the level of π. Since it clearly splits over some finite extension of E_λ as well, we can find a number field E which splits all D_λ's at the same time. To conclude, we take the Galois closure over $\mathbb{Q}$.

□

Remark 4.3.2. A natural question that arises after this lemma is if one should expect the D_λ's to come from a global object $D/\mathbb{Q}(\pi)$. We will discuss this more in the next chapter. In particular, in the special case of $n \leqslant 3$, this follows from our results, the existence of a motive associated to π, and the Mumford-Tate conjecture for that motive.

From now on, we take E to be the number field coming from Lemma 4.3.1, and we take our Galois representations to have values in $\mathrm{GL}_n(E_\mathfrak{p})$ for finite primes $\mathfrak{p}$ of E. So, we are in the setting of the pervious sections and we can define $\rho_{\pi,p}$ as follows:

$$\prod_{\mathfrak{p}|p} \rho_{\pi,\mathfrak{p}} =: \rho_{\pi,p} : \Gamma_K \to \mathrm{GL}_n(E_p) = \prod_{\mathfrak{p}|p} \mathrm{GL}_n(E_\mathfrak{p}),$$

where $E_p = E \otimes_\mathbb{Q} \mathbb{Q}_p \cong \prod_{\mathfrak{p}|p} E_\mathfrak{p}$ as usual.

From now on, we assume that π is neither self-twist nor essentially self-dual in the $n > 2$ case. Then it makes sense to talk about E-extra-twists of π. Since E is fixed, we will drop it from the notation from now on. By multiplicity one, inner-twists of π and $\rho_{\pi,p}$ agree (we are using class field theory to identify the characters). So, let Γ, Γ^{inn}, Γ^{out}, F, and F^{inn} be as usual.

The determinant of ρ_p is given by $\omega \cdot \epsilon_p^m$. To apply the results of the last section, we first need to kill the determinant. This is always possible after a finite extension. In fact, after a finite extension, the cyclotomic character will have values in $1 + p\mathbb{Z}_p$, and then we can use the p-adic logarithm on it. Therefore, there exists a finite extension M/K such that $\epsilon_p|_{\Gamma_M}$ has an n'th root. We fix one of those characters and denote it by $\epsilon_p^{1/n}$. Now, we enlarge M to trivialize ω if necessary. Then,

$$\rho'_{\pi,p} := \rho_{\pi,p}|_{\Gamma_M} \otimes \epsilon_p^{-m/n} : \Gamma_M \to \mathrm{GL}_n(E_p)$$

has trivial determinant. This is the Galois representation that we will apply our results from the last section to. Notice that the extra-twists of $\rho'_{\pi,p}$ and $\rho_{\pi,p}$ are the same. More precisely, the characters might have changed after the twist but the group $\Gamma \subseteq \mathrm{Aut}(E)$ has not. It is also not hard to see how the Lie algebra of the image changes.

Lemma 4.3.3. *Assume that $\rho'_{\pi,p}$ is valid. Let $\mathfrak{g}' := \mathrm{Lie}(\rho'_{\pi,p}(\Gamma_M))$ and $\mathfrak{g} = \mathrm{Lie}(\rho_{\pi,p}(\Gamma_L))$. Then $\mathfrak{g}^{\mathrm{der}} = \mathfrak{g}'$.*

Proof. Let $G = \rho_{\pi,p}(\Gamma_M)$ and $G' = \rho'_{\pi,p}(\Gamma_M)$, and let Z be the center of $\mathrm{GL}_n(E_p)$. Then, we clearly have $G \subseteq G' \cdot Z$ and $G' \subseteq G \cdot Z$. Taking the Lie algebras we find that $\mathfrak{g} \subseteq \mathfrak{g}' + \mathfrak{z}$ and $\mathfrak{g}' \subseteq \mathfrak{g} + \mathfrak{z}$ where $\mathfrak{z}$ is the Lie algebra of the center. Since $\mathfrak{g}'$ is semi-simple, by Proposition 4.2.8, this implies that $\mathfrak{g}$ is reductive and $\mathfrak{g}^{\mathrm{der}} = \mathfrak{g}'$. ☐

In the case that $\rho'_{\pi,p}$ is valid, let H_p/F_p be the semi-simple group from Corollary 4.2.7 applied to $\rho'_{\pi,p}$. Then we have:

Proposition 4.3.4. *If $\rho'_{\pi,p}$ is valid, then there exists a finite extension L of K such that $\rho_{\pi,p}(\Gamma_L)$ is contained and p-adically open in $H_p(F_p) \cdot \mathbb{Q}_p^\times \subseteq \mathrm{GL}_n(E_p)$.*

Proof. Since the image of $\rho'_{\pi,p}$ is contained in $H_p(F_p)$ after a finite extension, and the image of $\epsilon_p^{1/n}$ is in $\mathbb{Z}_p^\times$, the image of $\rho_{\pi,p}$ is contained in $H_p(F_p) \cdot \mathbb{Q}_p^\times$ after a finite extension. The image is open in $H_p \subseteq \mathrm{SL}_n(E_p)$ by Lemma 4.3.3 and the image of the determinant is open in $\mathbb{Q}_p^\times$, so we are done. ☐

This in particular implies that the connected component of the $\mathbb{Q}_p$-Zariski closure of the image is the algebraic group $(\mathrm{Res}_{\mathbb{Q}_p}^{F_p} H_p) \cdot \mathbb{G}_{m,\mathbb{Q}_p}$. Now, we only need to check the validity of $\rho'_{\pi,p}$.

4.3.1 The GL_2 Case

As we mentioned in the introduction, essentially everything is known in this case, by the work of Ribet [33], Momose [25], and Nekovar [27]. We repeat the arguments for the sake of completeness. In this case, all representations are essentially self dual, so there are no outer-twists and $\Gamma = \Gamma^{\mathrm{inn}}$. One can in fact take $E = \mathbb{Q}(\pi)$ (then Γ would be abelian), but it is not necessary for our discussion. Recall that we assumed that π is not self-twist, i.e. does not satisfy $\pi \simeq \pi \otimes \chi$ for $\chi \neq 1$. In this case it is more common to say that π does not have complex multiplication (CM).

Proposition 4.3.5. *Assume that $n = 2$ and π does not have CM. Then $\rho'_{\pi,p}$ is valid.*

Proof. Most of the properties are clear from the analogous properties for $\rho_{\pi,p}$. We only need to check strong irreducibility and compute the $\overline{\mathbb{Q}}_p$-Lie algebra.

It is known that $\rho_{\pi,\lambda}$ is irreducible (see Theorem 2.2.1). It is also known that it is de Rham and the Hodge-Tate weights are distinct. Strong irreducibility of $\rho'_{\pi,\lambda}$ and $\rho_{\pi,\lambda}$ are clearly equivalent. Assume that $\rho_{\pi,\lambda}|_{\Gamma_L}$ is reducible for some finite Galois extension L/K. Since $\rho_{\pi,\lambda}$ is semi-simple, so is its restriction to Γ_L and we have $\rho_{\pi,\lambda}|_{\Gamma_L} \simeq \chi_1 \oplus \chi_2$. Since $\rho_{\pi,\lambda}$ is Hodge-Tate with distinct Hodge-Tate weights, so is $\rho_{\pi,\lambda}|_{\Gamma_L}$. This implies that $\chi_1 \neq \chi_2$. Let K' be the fixed field of the stabilizer of χ_1. Then it is clearly a degree 2 extension of K and if $\mathrm{Gal}(K'/K) = \{1, \sigma\}$, then $\chi_2 = \sigma\chi_1$ (otherwise χ_1 would be a direct summand of $\rho_{\pi,\lambda}$). This means that $\rho_{\pi,\lambda} \simeq \mathrm{Ind}_K^{K'}(\chi_1)$, which implies that π is CM. Therefore, $\rho_{\pi,\lambda}$ and hence $\rho'_{\pi,\lambda}$ are strongly irreducible.

Now, let $\mathfrak{g}_\lambda^{\mathrm{der}} \subseteq \mathfrak{sl}_2(\overline{\mathbb{Q}}_p)$ be the derived part of the $\overline{\mathbb{Q}}_p$-Lie algebra of the image of $\rho'_{\pi,\lambda} : \Gamma_M \to \mathrm{SL}_2(\overline{\mathbb{Q}}_p)$. Since $\rho'_{\pi,\lambda}$ is strongly irreducible, the irreducibility holds infinitesimally, i.e. $\mathfrak{g}_\lambda \subseteq \mathfrak{gl}_2(\overline{\mathbb{Q}}_p)$ is an irreducible representation. This means that the centralizer of $\mathfrak{g}$ and hence its center are in the center of $\mathfrak{gl}_2(\overline{\mathbb{Q}}_p)$, which implies that $\mathfrak{g}_\lambda^{\mathrm{der}} \subseteq \mathfrak{sl}_2(\overline{\mathbb{Q}}_p)$ is also irreducible. The only irreducible semi-simple Lie subalgebra of $\mathfrak{sl}_2$ is itself, so we are done.

□

Now, from Proposition 4.3.4 it follows:

Corollary 4.3.6. *Let π be a regular algebraic cuspidal automorphic representation of $\mathrm{GL}_2(\mathbb{A}_K)$ that does not have complex multiplication and let F be the field fixed by the inner-twists. Then, there exists an inner form H_p of SL_2 over F_p and a finite extension L of K such that the image of $\rho_{\pi,p}(\Gamma_L)$ is contained and open in $H_p(F_p) \cdot \mathbb{Q}_p^\times \subseteq \mathrm{GL}_2(E_p)$.*

In the work of Ribet, Momose and Nekovar, they construct an Azumaya algebra D_p/F_p which contains the image. The relation to the last corollary is that if $D_p^\times$ is the algebraic group of units of D_p, then $(D_p^\times)^{\mathrm{der}} = H_p$. In the case that π has parallel weight 2, where we expect an abelian variety to be associated with π, this algebra D_p is closely related to the endomorphism ring of that abelian variety. We will explain the relation of these results to the Mumford-Tate conjecture for that abelian variety in Chapter 5.

4.3.2 The GL_3 Case

In this section we prove our main result. The CM-case for $n = 2$ can be thought of as the π essentially coming from GL_1 by induction. Similarly, in the $n = 3$ case we need to first exclude all the cases where π comes from smaller groups via a Langlands transfer, in which case the image would be easy to describe by previous results. It turns out that we only need to exclude the following two cases to be able to describe the image:

1. π is essentially sym^2, i.e. there exists an automorphic representation θ of $\mathrm{GL}_2(\mathbb{A}_K)$ and a Hecke character η such that $\pi = \mathrm{sym}^2(\theta) \otimes \eta$.

2. π is an induction of a character, i.e. there exist a degree 3 extension L/K and a Hecke character η of $\mathbb{A}_L$ such that $\pi = \mathrm{Ind}_K^L(\eta)$.

Notice that Langlands functoriality is known for $\mathrm{sym}^2 : \mathrm{GL}_2 \to \mathrm{GL}_3$ by [18] and automorphic base change is known for prime degree extensions by [1]. In the first case above, determining the image reduces to the GL_2 case and in the second case to the GL_1 case. The next two lemmas give equivalent classifications for the above cases and show that they follow from our primary assumptions on π (not being essentially self-dual or self-twist) that are needed to define extra-twists to begin with.

Lemma 4.3.7. *π is essentially* sym^2 *if and only if there exist a Hecke character χ such that $\pi = \pi^\vee \otimes \chi$.*

Proof. Since GL_2 representations are essentially self-dual, the "only if" part is clear. Now assume $\pi = \pi^\vee \otimes \chi$ and let ω be the central character of π. Taking the central characters of both sides we have $\chi^3 = \omega^2$. So $\chi = (\omega\chi^{-1})^2$ has a square root and by twisting out this square root we can assume that π is self-dual. Now the result follows from [30]. □

Remark 4.3.8. Since we assumed that K is totally real and π is cuspidal regular algebraic and hence the existence of the associated Galois representations is known, one can equivalently work with the associated Galois representation, by strong multiplicity one. Then, one can give a different proof in the Galois side by investigating the projective image of the representation. We will leave the details to the reader.

We also need the following lemma from [1, Lemma 6.3].

Lemma 4.3.9. *π is an induction of a character if and only if there exists a Hecke character χ such that $\pi = \pi \otimes \chi$.*

Definition 4.3.10. An automorphic representation π of $\mathrm{GL}_3(\mathbb{A}_K)$ is said to be of general type, if π is neither essentially self-dual nor self-twist. Equivalently, π is neither essentially sym^2 nor an induction of a character.

Since we know strong multiplicity one for GL_n, Langlands functoriality for the map $\mathrm{sym}^2 : \mathrm{GL}_2 \to \mathrm{GL}_3$, and automorphic induction for degree 3 extensions, these assumptions are equivalent to similar assumptions on each of the Galois representations $\rho_{\pi,\lambda}$.

Recall that in [6], Böckle and Hui prove that (in the $n = 3$ case), $\rho_{\pi,\lambda}$ is irreducible for all λ. They also prove that for a density 1 set of rational primes $\mathcal{P}$, $\rho_{\pi,\lambda}$ is de Rham with distinct Hodge-Tate weights for all $\lambda : E \hookrightarrow \overline{\mathbb{Q}_p}$ and $p \in \mathcal{P}$. We will use these results to check the validity of $\rho'_{\pi,p}$ (in fact we only need regularity at one place λ by part 4 of Remark 4.2.2).

Proposition 4.3.11. *Assume that π is not self-twist. Then for each $p \in \mathcal{P}$ and $\lambda : E \hookrightarrow \overline{\mathbb{Q}}_p$, one has that $\rho'_{\pi,\lambda}$ is strongly irreducible.*

Proof. Assume that $\rho_{\pi,\lambda}|_{\Gamma_L}$ is reducible for some finite Galois extension L/K. Since $\rho_{\pi,\lambda}$ is semi-simple (in fact irreducible), so is its restriction to Γ_L and we have that $\rho_{\pi,\lambda}|_{\Gamma_L}$ decomposes into the sum of irreducible direct summands. If it decomposes into two irreducible summands, then the action of Γ_K cannot switch the two summands for dimension reasons and hence each summand is actually a subrepresentation of $\rho_{\pi,\lambda}$, which is a contradiction. So we must have $\rho_{\pi,\lambda}|_{\Gamma_L} \simeq \chi_1 \oplus \chi_2 \oplus \chi_3$. Since $\rho_{\pi,\lambda}$ is Hodge-Tate with distinct Hodge-Tate weights, so is $\rho_{\pi,\lambda}|_{\Gamma_L}$. This implies that the three characters are distinct. Let K' be the fixed field of the stabilizer of χ_1. The action of Γ_K on these characters must be transitive, so K' is a degree 3 extension of K and if $\mathrm{Gal}(K'/K) = \{1, \sigma, \sigma^2\}$ then $\chi_2 = \sigma\chi_1$ and $\chi_3 = \sigma^2\chi_1$ (or the other way around). This means that $\rho_{\pi,\lambda} \simeq \mathrm{Ind}_K^{K'}(\chi_1)$, which implies that π is also a degree three automorphic induction of a character, and hence is self-twist, which contradicts the assumption. Therefore, $\rho_{\pi,\lambda}$ and hence $\rho'_{\pi,\lambda}$ are strongly irreducible. □

Proposition 4.3.12. *Assume that π is of general type. Then, for each prime number p and embedding $\lambda : E \hookrightarrow \overline{\mathbb{Q}}_p$ one has that the $\overline{\mathbb{Q}}_p$-Lie algebra of the image of $\rho'_{\pi,\lambda}$ is $\mathfrak{sl}_3(\overline{\mathbb{Q}}_p)$.*

Proof. First assume that $p \in \mathcal{P}$. Let V_λ be the underlying vector space of $\rho'_{\pi,\lambda}$ and let $\mathfrak{g}_\lambda \subseteq \mathfrak{sl}_3(\overline{\mathbb{Q}}_p)$ be the $\overline{\mathbb{Q}}_p$-Lie algebra of the image of $\rho'_{\pi,\lambda} : \Gamma_M \to \mathrm{SL}_3(\overline{\mathbb{Q}}_p)$. Since $\rho'_{\pi,\lambda}$ is strongly irreducible, for every finite extension N of M we have $\mathrm{End}_{\Gamma_N}(V_\lambda) = \overline{\mathbb{Q}}_p$, hence this should be true infinitesimally and we have $\mathrm{End}_{\mathfrak{g}_\lambda}(V_\lambda) = \overline{\mathbb{Q}}_p$. This means that the standard representation $\mathfrak{g}_\lambda \hookrightarrow \mathrm{End}_{\overline{\mathbb{Q}}_p}(V_\lambda) \simeq \mathfrak{gl}_3(\overline{\mathbb{Q}}_p)$ is an irreducible faithful representation, which implies that $\mathfrak{g}_\lambda$ is reductive. Let $\mathfrak{g}_\lambda^{\mathrm{der}}$ be its derived subgroup and hence a semi-simple subgroup of $\mathfrak{sl}_3(\overline{\mathbb{Q}}_p)$. Since the center of $\mathfrak{gl}_3$ acts by scalar multiplication, any $\mathfrak{g}_\lambda^{\text{der}}$-invariant subspace is automatically $\mathfrak{g}_\lambda$-invariant as well, which implies that the centralizer of $\mathfrak{g}$ and hence its center are in the center of $\mathfrak{gl}_3(\overline{\mathbb{Q}}_p)$, which then implies that $\mathfrak{g}_\lambda^{\text{der}} \subseteq \mathfrak{gl}_3(\overline{\mathbb{Q}}_p)$ is also irreducible.

The only irreducible, semi-simple Lie subalgebras of $\mathfrak{sl}_3$ up to conjugation are $\mathfrak{sl}_3$ and $\mathfrak{sl}_2$ embedded into $\mathfrak{sl}_3$ by sym^2. We need to show that the latter does not happen. Assume that $\mathfrak{g}_\lambda^{\text{der}}$ is the image of $\text{sym}^2 : \mathfrak{sl}_2 \to \mathfrak{sl}_3$. Since sym^2 is an irreducible representation, its centralizer in $\mathfrak{gl}_3$ is the center, which means that $\mathfrak{g}_\lambda = \mathfrak{g}_\lambda^{\text{der}} \oplus \mathfrak{z}(\mathfrak{g}_\lambda)$ is in the image of $\text{sym}^2 : \mathfrak{gl}_2 \to \mathfrak{gl}_3$. This means that there is an open subgroup of Γ_M whose image under $\rho'_{\pi,p}$ is in the image of $\text{sym}^2 : \text{GL}_2 \to \text{GL}_3$. So there exists a finite Galois extension M' of M such that $\rho'_{\pi,p}(\Gamma_{M'})$ is in the image of $\text{sym}^2 : \text{GL}_2(E_\lambda) \to \text{GL}_3(E_\lambda)$ and hence is essentially self-dual. Since the determinant of $\rho'_{\pi,p}$ is trivial, there exist a finite extension N of M' (which clearly can be taken to be Galois over M) such that the restriction to Γ_N is in fact self-dual. Now, applying Lemma 4.2.4 to the two representations $\rho'_{\pi,p}|_{\Gamma_N}$ and $\rho'_{\pi,p}|_{\Gamma_N}^\vee$, there exists a character ϕ such that $\rho'_{\pi,p} \simeq (\rho'_{\pi,p})^\vee \otimes \phi$, which contradicts non-essential-self-duality of π. This contradiction implies the result in the case where $p \in \mathcal{P}$.

Now by [6, Theorem 3.2], the irreducible type of $\rho_{\pi,\lambda}$ is independent of λ. This in particular means that if the $\overline{\mathbb{Q}}_p$-Lie algebra of the image of $\rho_{\pi,\lambda}$ contains $\mathfrak{sl}_3$ for one λ (irreducible type A_2), it contains $\mathfrak{sl}_3$ for every λ. This clearly implies the result in general. □

Now we can easily deduce our main result:

Theorem 4.3.13 (Theorem 1.1 in [43]). *Let $F = E^\Gamma$ be the field fixed by all the extra-twists of π. Then, there exists a finite extension L/K and a semi-simple algebraic group H_p defined over $F_p := F \otimes_{\mathbb{Q}} \mathbb{Q}_p$, which is a form of* SL_3 *(constructed using the extra-twists), such that $\rho_{\pi,p}(\Gamma_L)$ is contained in $H_p(F_p) \cdot \mathbb{Q}_p^\times \subseteq \text{GL}_3(E \otimes_{\mathbb{Q}} \overline{\mathbb{Q}_p})$ and it is open in the p-adic topology.*

Proof. The last two propositions imply that for π of general type, $\rho'_{\pi,p}$ is valid for any prime number p. Then Proposition 4.3.4 implies the result. □

In other words, for all p, the algebraic group $(\text{Res}_{\mathbb{Q}_p}^{F_p} H_p) \cdot \mathbb{G}_{m,\mathbb{Q}_p}$ is the connected component of the $\mathbb{Q}_p$-Zariski closure of the image and the image is open in there. Let us apply this theorem to some explicit examples:

Example 4.3.14. In [17], van Geeman and Top construct a 3-dimensional $\mathbb{Q}(i)$-rational compatible family of (motivic) Galois representations of $\Gamma_{\mathbb{Q}}$ which is neither self-twist nor essentially self-dual, and an automorphic representation of $\mathrm{GL}_3(\mathbb{A}_{\mathbb{Q}})$ that should correspond to it. We can apply our results to the Galois representations they construct. For each prime p they construct a Galois representation

$$\rho_p : \Gamma_{\mathbb{Q}} \to \mathrm{GL}_3(\mathbb{Q}(i) \otimes \mathbb{Q}_p)$$

which has the property that $\rho_p \simeq \overline{\rho_p}^{\vee} \otimes \epsilon_p$ where $\overline{(\cdot)}$ indicates complex conjugation. For each unramified p, the characteristic polynomial of Frob_p is of the form

$$X^3 - b_p X^2 + p\overline{b_p} X - p^3$$

and they give a list of values of $b_p \in \mathbb{Q}(i)$ for small primes.

Now, in our notation, the coefficient field is $E = \mathbb{Q}(i)$. There is one outer twist $(\bar{\cdot}, \epsilon_p)$ and there cannot be any more non-trivial extra-twists since $\mathrm{Aut}(E) \simeq \mathbb{Z}/2\mathbb{Z}$. Therefore, $F^{\mathrm{inn}} = \mathbb{Q}(i)$ as well and $F = \mathbb{Q}$. Then, for each prime p we can construct a form H_p of SL_3 over the field $\mathbb{Q}_p$ as before and the image of ρ_p is contained and open in $H_p(\mathbb{Q}_p) \cdot \mathbb{Q}_p^{\times}$. Hence, we get an algebraic group $H_p \cdot \mathbb{G}_m \subseteq \mathrm{Res}_{\mathbb{Q}_p}^{\mathbb{Q}(i)_p} \mathrm{GL}_3$ whose $\mathbb{Q}_p$ points describe the image. We know that H_p is a form of SL_3. Recall that it is constructed as

$$H_p = (\mathrm{Res}_{\mathbb{Q}_p}^{\mathbb{Q}(i)_p} \mathrm{SL}_3)^{tw(\mathrm{Gal}(\mathbb{Q}(i)/\mathbb{Q}))}.$$

Similar to Corollary 4.2.10, if p is a prime that splits in $\mathbb{Q}(i)$, i.e. a prime that is congruent to 1 modulo 4, then H_p is in fact isomorphic to SL_3 over $\mathbb{Q}(i)_p \simeq \mathbb{Q}_p \times \mathbb{Q}_p$. Otherwise, it is not an inner-form and since it splits over $\mathbb{Q}(i)$, it is isomorphic to the the special unitary group SU_3 for the degree two field extension $\mathbb{Q}(i)_p/\mathbb{Q}_p$. So, for half of the primes (primes of the form $p = 4k + 1$) the (Zariski closure of the) image is GL_3 and for the other half (primes of the form $p = 4k + 3$) it is $\mathrm{SU}_3 \cdot \mathbb{G}_m$. Finding the right candidate for the group H_∞ over $\mathbb{R}$, one should be able to prove that all these groups H_p come from a global group $H/\mathbb{Q}$.

Example 4.3.15. In [44], Upton constructs a 3-dimensional $\mathbb{Q}(\zeta_3)$-rational compatible family of (motivic) Galois representations of $\Gamma_{\mathbb{Q}(\zeta_3)}$ which is neither self-twist nor essentially self-dual, and gives a precise description of its image. It is clear from her construction that these Galois representations have an outer-twist. She also observes the same phenomenon as in the last example. Namely, that for half of the primes the image is GL_3

and for the other half it is a unitary group. Although, we believe there is a slight error in her conclusion and the image in the latter case should be $\mathrm{SU}_3 \cdot \mathbb{G}_m$ as above, rather than the general unitary group GU_3 as she claims. In fact since in the split case the image is 9-dimensional (as it is GL_3), if one believes in the Mumford-Tate conjecture, the image cannot be the 10-dimensional group GU_3 in the non-split case.

Even though in her case K is not totally real, we can still directly apply Corollary 4.2.9 to a twist of the Galois representations she constructs (after a finite extension) and then deduce openness. It is easy to check the validity of this twisted Galois representation.

4.3.3 The GL_n Case

In this section we discuss the GL_n case. Everything we say here is conjectural for $n > 3$. We assume the irreducibility conjecture (Galois representations associated with cuspidal automorphic representations are irreducible), Langlands functoriality, and the expected p-adic Hodge theoretic properties (de Rham with distinct Hodge-Tate weights) of our Galois representations. We want to see, assuming all these, when we can apply Proposition 4.3.4 to a regular cuspidal algebraic automorphic representation π of $\mathrm{GL}_n(\mathbb{A}_K)$. First of all, we need to assume that π is neither self-twist nor essentially self-dual (in the $n > 2$ case). Then we only need to check that $\rho_{\pi,\lambda}$ is strongly irreducible for each λ and that the $\overline{\mathbb{Q}}_p$-Lie algebra of the image of $\rho'_{\pi,\lambda}$ is $\mathfrak{sl}_n$.

Assume that $\rho_{\pi,\lambda}$ is reducible after restricting to Γ_L for a finite Galois extension L of K. The irreducible direct summands of $\rho_{\pi,\lambda}|_{\Gamma_L}$ are distinct since the Hodge-Tate weights are distinct and the action of Γ_K on them is transitive since $\rho_{\pi,\lambda}$ is irreducible. This easily implies that $\rho_{\pi,\lambda}$ is an induction of a representation of a proper subgroup. This means that the automorphic representation π is an induction, assuming that the automorphic induction is true. So in order to make sure that $\rho_{\pi,\lambda}$ is strongly irreducible, we only need to assume that it is not an induction.

Now, let $\mathfrak{g}$ be the $\overline{\mathbb{Q}}_p$-Lie algebra of the image of $\rho'_{\pi,\lambda}$. Since $\rho_{\pi,\lambda}$ and hence $\rho'_{\pi,\lambda}$ are strongly irreducible, $\mathfrak{g}$ is an irreducible Lie subalgebra of $\mathfrak{gl}_n$ and hence reductive. We only need to show that $\mathfrak{g}^{\mathrm{der}} = \mathfrak{sl}_n$. It is well known that since $\mathfrak{g}^{\mathrm{der}}$ is semi-simple, there exists a semi-simple (connected) algebraic subgroup G' of GL_n (over $\overline{\mathbb{Q}}_p$) such that $\mathrm{Lie}(G') = \mathfrak{g}^{\mathrm{der}}$. This implies that after a finite Galois extension M/K, the image of Γ_M under $\rho_{\pi,p}$ lies in $G^0(\overline{\mathbb{Q}_p}) \subsetneq \mathrm{GL}_n(\overline{\mathbb{Q}}_p)$ for the (connected) reductive group $G^0 = G' \cdot \mathbb{G}_m \subsetneq \mathrm{GL}_n$. This is exactly the connected component of the $\overline{\mathbb{Q}}_p$-Zariski closure of the image of $\rho_{\pi,\lambda}$. Let the whole image be G. Then the stabilizer of G^0 would give a finite Galois extension L/K and the component group is isomorphic to Gal(L/K). The nicest situation would be if $G(\overline{\mathbb{Q}_p}) = G^0(\overline{\mathbb{Q}_p}) \rtimes \mathrm{Gal}(L/K)$. But it is not clear if this should happen. Nevertheless, one has the following result of Brion [7]:

Lemma 4.3.16 (Brion). *Let G be an algebraic group over a field k and let*

$$1 \to N \to G \to Q \to 1$$

be a short exact sequence (of algebraic groups over k) such that Q is finite. Then there exists a finite subgroup F of G, such that $G = N \cdot F$. In other words, F surjects to Q and G is a quotient of $N \rtimes F$, where F acts on N by conjugation.

Now using this result, one can at least find a finite Galois extension M/K such that G is the quotient of $G^0(\overline{\mathbb{Q}_p}) \rtimes \mathrm{Gal}(M/K)$. Now, we can form the L-group $G^0(\overline{\mathbb{Q}_p}) \rtimes \Gamma_K$ where the Γ_K action factors through the Gal(M/K) action from above. Fix a maximal torus and a Borel subgroup of G^0 containing it. The above action of Gal(M/K) on $G^0(\overline{\mathbb{Q}_p})$ gives an action on the based root datum, which in turn gives an action on the dual root datum. This finally gives a reductive group H over K which splits over M whose Langlands L-group is $G^0(\overline{\mathbb{Q}_p}) \rtimes \Gamma_K$ with the above action. Now, the Langlands functoriality for the L-map ${}^LH \to {}^L\mathrm{GL}_n$ implies that π should come from an automorphic representation of the non-split reductive group H via the Langlands transfer induced by the above L-map. This motivates the following definition:

Definition 4.3.17. A regular cuspidal automorphic representation π of $\mathrm{GL}_n(\mathbb{A}_K)$ is said to be of general type, if it is neither self-twist, nor essentially self-dual (in the $n > 2$ case), and there does not exists any reductive group H over K that is a form of a proper subgroup of GL_n such that π is the image of an automorphic representation of $H(\mathbb{A}_K)$ under the Langlands transfer attached to the L-map

$${}^LH \to {}^L\mathrm{GL}_{n,K}.$$

For instance, in the GL_2 case this just means that π is not CM and in the GL_3 case it agrees with Definition 4.3.10. Note that the condition above also automatically includes that π is not an induction since it would be in the image of the following L-map then:

$${}^L\mathrm{Res}_K^L\mathrm{GL}_d \to {}^L\mathrm{GL}_{d[L:K]}.$$

The discussion above shows that if one believes in Langlands functoriality, the irreducibility conjecture, and that $\rho_{\pi,\lambda}$ is de Rham with distinct Hodge-Tate weights, then $\rho'_{\pi,p}$ is valid. In conclusion, we make the following conjecture:

Conjecture 4.3.18. Let K be totally real and π be a regular cuspidal algebraic automorphic representation of $\mathrm{GL}_n(\mathbb{A}_K)$ of general type. Let $E = \mathbb{Q}(\pi)$ and let F be the field fixed by the E-extra-twists of π. Then, there exists a semi-simple group H over F which is a form of SL_n and a finite extension L/K such that for any prime p, the image of $\rho_{\pi,p}(\Gamma_L)$ is contained and open in $H(F_p) \cdot \mathbb{Q}_p^\times$.

Remark 4.3.19. If π is not of general type, then it comes from a smaller group H. Since the dimension of the group is getting smaller, there should be an optimal choice for H. Loosely speaking, π should be of general type for some group. Then one has to study the image inside this smaller group, via the extra-twists for the Langlands dual of this group. Then it might be possible to give a precise description of the image as above, using the extra-twists.

Chapter 5

Relations to the Mumford-Tate Conjecture

In this chapter we study the relation of our constructions in Chapter 4 with the Mumford-Tate group of the motive associated with an algebraic automorphic representation under the conjectures of Clozel. Almost everything we discuss here is conjectural, but it could give an idea of why one should believe in the conjectures presented here.

5.1 Mumford-Tate Groups

In this section, we give a very quick review of the Mumford-Tate Conjecture for motives. For more details, we refer the reader to [26]. Throughout this chapter, we are thinking of a motive as a collection of realizations whose different structures are compatible through a set of comparison isomorphisms, as in [16, Chapter III].

Let M be a motive over a number field K. Then, the Betti and the de Rham realizations of M, together with their comparison isomorphism, give rise to a rational Hodge structure $V \in \mathbb{Q}$-VS given by

$$h_M : \mathbb{S} \to \mathrm{GL}(V)_{\mathbb{R}},$$

where $\mathbb{S} = \mathrm{Res}_{\mathbb{R}}^{\mathbb{C}}\mathbb{G}_m$ is the Deligne torus.

Definition 5.1.1. The Mumford-Tate group of M is defined to be the smallest algebraic group $\mathrm{MT}(M) \subseteq \mathrm{GL}(V)$ (defined over $\mathbb{Q}$) such that h_M factors through $M_{\mathbb{R}} \subseteq \mathrm{GL}(V)_{\mathbb{R}}$. In other words, $\mathrm{MT}(M)$ is the intersection of all closed subgroups of $\mathrm{GL}(V)/\mathbb{Q}$ with the above property.

One can also realize the Mumford-Tate group from a Tannakian perspective. Namely, the rational Hodge structure (V, h_M) generates a Tannakian subcategory of the category of all rational Hodge structures. One can easily show that the forgetful functor $\omega : (V, h_M) \mapsto V$ is a fiber functor whose automorphism group is the Mumford-Tate group $\mathrm{MT}(M)$. Therefore, the Tannakian category generated by (V, h_M) is equivalent to the category of representations of $\mathrm{MT}(M)$.

Now, let the $\mathbb{Q}_\ell$-vector space V_ℓ be the ℓ-adic realization of M. Then, the absolute Galois group Γ_K acts on V_ℓ and this gives a Galois representation

$$\rho_{M,\ell} : \Gamma_K \to \mathrm{GL}_n(\mathbb{Q}_\ell).$$

Let $\mathrm{G}_{M,\ell}$ be the $\mathbb{Q}_\ell$-Zariski closure of the image of this Galois representation and $\mathrm{G}^\circ_{M,\ell}$ be its connected component of identity.

Conjecture 5.1.2 (Mumford-Tate Conjecture)**.** With the notations as above, there exists an isomporphism

$$\mathrm{G}^\circ_{M,\ell} \simeq \mathrm{MT}(M) \times_{\mathbb{Q}} \mathbb{Q}_\ell$$

for every prime number ℓ.

In particular, this implies that $\mathrm{G}^\circ_{M,\ell}/\mathbb{Q}_\ell$ is independent of ℓ, i.e. they are all base changes of a single global object $\mathrm{MT}(M)/\mathbb{Q}$.

5.2 Extra-Twists of a Motive

Clozel predicts that there should be a correspondence between algebraic automorphic representations of $\mathrm{GL}_n(\mathbb{A}_K)$ and motives over K with coefficients in number fields. Let K be totally real as before and π be a regular cuspidal algebraic automorphic representation of $\mathrm{GL}_n(\mathbb{A}_K)$. Then, Clozel predicts the existence of a motive $M = M_\pi$ over K with coefficients in a number field E containing $\mathbb{Q}(\pi)$ that is associated to π in the way explained in [11]. As mentioned, a motive for us is just a collection of different cohomology theories compatible via a set of comparison isomorphisms. From now on, we assume that such a motive exists. Let $H_B(M)$, $H_{\mathrm{dR}}(M)$ and $H_p(M)$ be the Betti, de Rham and p-adic realizations of M. Note that the first two are E-vector spaces and the last one is an E_p-module.

Let $V = H_B(M)$. The real and complex Betti cohomologies $V \otimes_{\mathbb{Q}} \mathbb{R}$ and $V \otimes_{\mathbb{Q}} \mathbb{C}$ have an $E \otimes_{\mathbb{Q}} \mathbb{R}$ and $E \otimes_{\mathbb{Q}} \mathbb{C} = \prod_{\lambda : E \hookrightarrow \mathbb{C}} \mathbb{C}$ structure, respectively. Similarly, the complex de Rham cohomology $H_{\mathrm{dR}}(M) \otimes_{\mathbb{Q}} \mathbb{C}$ has an $E \otimes_{\mathbb{Q}} \mathbb{C} = \prod_{\lambda: E \hookrightarrow \mathbb{C}} \mathbb{C}$ structure. The ($E \otimes_{\mathbb{Q}} \mathbb{C}$-modules) comparison isomorphism between Betti and de Rham cohomologies, $V_{\mathbb{C}} \simeq H_{\mathrm{dR}}(M) \otimes_{\mathbb{Q}} \mathbb{C}$, equips V with a rational Hodge structure. We denote this Hodge structure with

$$h_\pi : \mathbb{S} \to \mathrm{GL}(V_{\mathbb{R}}),$$

where $\mathbb{S}$ is the Deligne torus. Fixing an E-basis for $V_{\mathbb{Q}}$ which in turn gives an $E \otimes \mathbb{R}$ basis for $V_{\mathbb{R}}$ enables us to write this as

$$h_{\pi,\infty} : \mathbb{S} \to \mathrm{GL}_n(E \otimes_{\mathbb{Q}} \mathbb{R}).$$

This representation should be thought of as the analogue of our p-adic Galois representations $\rho_{\pi,p} : \Gamma_K \to \mathrm{GL}_n(E \otimes_{\mathbb{Q}} \mathbb{Q}_p)$ from pervious chapters, associated to the prime at infinity. Note that this is equipped with an action of $\mathrm{Aut}(E)$ on the coefficients.

Now, let π be of general type and Γ be the group of E-extra-twists of π. Let $|.|^m\omega$ be the central character of π, where ω is a finite order Hecke character. From now on, for simplicity, we assume that m is divisible by n. So, let $m = nd$. Then the outer-twists of π are of the form $(\tau, |\cdot|^{2d}\eta)$ for a finite character η and hence the outer-twists of $\rho_{\pi,p}$ are of the form $(\tau, \epsilon_p^{2d}\eta)$ where ϵ_p is the p-adic cyclotomic character and we think of η as a finite Galois character. The extra-twist of π then induce extra-twists on the motive M_π. An inner-twist (σ, χ) induces an isomorphism ${}^\sigma M_\pi \simeq M_\pi \otimes \chi$ where χ is the Artin motive associated with the finite character χ. The outer-twist $(\tau, |\cdot|^{2d}\eta)$ induces an isomorphism ${}^\tau M_\pi \simeq M_\pi^\vee \otimes \mathbb{Q}(2d) \otimes \eta$.

In particular, the extra-twists also induce symmetries on the Hodge-structure since E acts on the motive M via endomorphisms. Twisting with finite characters does not affect the Hodge structure and twisting with the $2d$'th power of the cyclotomic character amounts to twisting with Tate's Hodge structure $\mathbb{Q}(2d)$. This means that for each inner-twist $\sigma \in \Gamma^{\mathrm{inn}}$, one has ${}^\sigma h_\pi \simeq h_\pi$, and for each outer-twist $\tau \in \Gamma^{\mathrm{out}}$, one has ${}^\tau h_\pi \simeq h_\pi^\vee \otimes_{\mathbb{Q}} \mathbb{Q}(2d)$. Now, if we twist h_π with $\mathbb{Q}(-d)$, we still get ${}^\sigma h_\pi(-d) \simeq h_\pi(-d)$ for each inner-twist and ${}^\tau h_\pi(-d) \simeq h_\pi(-d)^\vee$ for each outer-twist. This is analogous to the representation $\rho'_{\pi,p}$ from the previous chapter.

Since an isomorphism of rational Hodge structures comes from an isomorphism over $\mathbb{Q}$ between the underlying rational vector spaces, and since everything is compatible with the E-structures, we can find matrices $\alpha_\sigma, \alpha_\tau \in \mathrm{GL}_n(E)$ that give the isomorphisms above by conjugation. So we get

$$\begin{cases} h_{\pi,\infty}(-d) = \alpha_\sigma \cdot {}^\sigma h_{\pi,\infty}(-d) \cdot \alpha_\sigma^{-1} \\ h_{\pi,\infty}(-d) = \alpha_\tau \cdot {}^\tau h_{\pi,\infty}^{-T}(-d) \cdot \alpha_\tau^{-1}. \end{cases}$$

Note that the determinant of $h_{\pi,\infty}(-d)$ is trivial, so it has values in SL_n. Define the 1-cocycle $f : \Gamma \to \mathrm{Aut}_E(\mathrm{SL}_n)$ by sending an inner-twist σ to $\mathrm{ad}(\alpha_\sigma)$ and an outer-twist τ to $\mathrm{ad}(\alpha_\tau) \circ (\cdot)^{-T}$. Then, as in Section 1.2.1, we can define the twisted action of Γ on $\mathrm{Res}_F^E \mathrm{SL}_n$ and the matrices in the image of $h_{\pi,\infty}(-d)$ are clearly invariant under this action. We define the groups $H_\infty := (\mathrm{Res}_{F\otimes\mathbb{R}}^{E\otimes\mathbb{R}} \mathrm{SL}_n)^{tw(\Gamma)}$ and $H := (\mathrm{Res}_F^E \mathrm{SL}_n)^{tw(\Gamma)}$. Note that H_∞ is the base change of H to $\mathbb{R}$ and it is the Archimedean analogue of the groups H_p from Chapter 4.

Remark 5.2.1. We could also define the group H_∞ without assuming the existence of the motive, only from the real Hodge structure coming from the Archimedean part of π. But then the connection to the Mumford-Tate group is of course less clear.

Lemma 5.2.2. *The Mumford-Tate group of the motive M_π is contained in the group* $\mathrm{Res}_{\mathbb{Q}}^F(H) \cdot \mathbb{G}_m$.

Proof. The image of $h_{\pi,\infty}(-d)$ lies in H_∞ by the definition of H_∞. This implies that the map $h_{\pi,\infty}$ factors through $H_\infty \cdot \mathbb{G}_{m,\mathbb{R}}$ and therefore the map h_π factors through the base change of the group $\mathrm{Res}_{\mathbb{Q}}^F(H) \cdot \mathbb{G}_{m,\mathbb{Q}}$ to $\mathbb{R}$. This implies the result. □

Note that the dimension of the group H is equal to the dimension of all the groups H_p from the pervious chapter. Our results in Chapter 4 make it reasonable to make the following conjecture.

Conjecture 5.2.3. If π is of general type then the Mumford-Tate group of M_π is equal to $\mathrm{Res}_{\mathbb{Q}}^F(H) \cdot \mathbb{G}_{m,\mathbb{Q}}$.

We showed that the group H (or any of the H_p's from before) gives an upper bound for the Mumford-Tate group. On the other hand, in the special case that M is an abelian motive, Deligne shows that the Mumford-Tate group $\mathrm{MT}(\mathrm{M}_\pi)$, after base changing to $\mathbb{Q}_p$, always contains the connected component of the image of the p-adic Galois representation. Therefore, if we know that the image is open in H_p even for one prime, the conjecture above follows. In particular we have the following result in the $n = 2$ case (which was probably known to Nekovář):

Corollary 5.2.4. *Let f be a non-CM Hilbert modular newform of parallel weight 2 over the totally real field K and assume there is an abelian variety A_f associated to it. Then Conjecture 5.2.3 and the Mumford-Tate conjecture hold for A_f.*

Proof. Note that $\mathrm{MT}(M_\pi) \otimes_{\mathbb{Q}} \mathbb{Q}_p$ is contained in $(\mathrm{Res}_{\mathbb{Q}}^F(H) \cdot \mathbb{G}_m) \otimes \mathbb{Q}_p$ by Lemma 5.2.2 and contains the connected component of the Zariski-closure of the image of $\rho_{f,p}$ which is equal to $\mathrm{Res}_{\mathbb{Q}_p}^{F_p} H_p(\mathbb{Q}_p) \cdot \mathbb{Q}_p^\times$ by 4.3.6. Since the dimensions match, both inclusions must be equality. □

At the end, we want to come back to Remark 4.3.2. Recall that by [9], for each prime p there exist an Azumaya algebra D_p over $\mathbb{Q}(\pi)_p := \mathbb{Q}(\pi) \otimes_{\mathbb{Q}} \mathbb{Q}_p$ such that the Galois representation $\rho_{\pi,p}$ factors through $D_p \subset D_p \otimes_{\mathbb{Q}(\pi)} E \simeq M_n(E_p)$, assuming irreducibility of the Galois representation. We are interested in seeing if the local objects D_p should come from a global object D defined over $\mathbb{Q}(\pi)$. In the $n = 1$ case, this is clear. In the $n = 2$ case, since $\rho_{\pi,p}$ is odd, the residual representation is multiplicity free and a result of Bellaïche and Chenevier [4] shows that in this case every pseudo-representation can be defined over its trace field and hence $D_p \simeq M_2$ for all primes p. So, $D \simeq M_2$ works.

Now assume that π is of general type and assume Conjecture 5.2.3 and the Mumford-Tate Conjecture. First, notice that for every $\sigma \in \mathrm{Gal}(E/\mathbb{Q}(\pi))$ we have an inner-twist $\pi \simeq {}^\sigma\pi$, therefore $F \subseteq \mathbb{Q}(\pi)$. Now, we know by our assumptions that for some finite Galois extension L/K, the image of $\rho_{\pi,p}|_{\Gamma_L}$ is open in $H_p(F_p) \cdot \mathbb{Q}_p^\times$. This is the $\mathbb{Q}_p$-Zariski closure of the image, hence the F_p-Zariski closure of the image is $H_p(F_p) \cdot F_p^\times$. Note that the inner product is happening inside GL_n, so since H_p is a form of SL_n, we deduce that $G_p := H_p \cdot \mathbb{G}_{m,F_p}$ is a form of GL_n. Since the image is contained in $D_p^\times$ which is an algebraic group over $\mathbb{Q}(\pi)_p$ and $F_p \subseteq \mathbb{Q}(\pi)_p$, we have

$$G_p(F_p) \subseteq D_p^\times(\mathbb{Q}(\pi)_p) \subseteq \mathrm{GL}_n(E_p),$$

and the base change of either G_p or $D_p^\times$ to E_p is equal to GL_n. This implies that $D_p^\times$ is the $\mathbb{Q}(\pi)_p$-Zariski closure of the image of the Galois representation. In particular, by the Mumford-Tate conjecture, the groups $D_p^\times$ should come from a global group, namely the $\mathbb{Q}(\pi)$-Zariski closure of the image of the map $h_{\pi,\infty}$. This is an inner-form of GL_n, because all the G_p's become inner-forms over F_p^{inn}. Hence, it is the group of the units of a central simple algebra D over $\mathbb{Q}(\pi)$.

In particular in the $n = 3$ case, π is either an induction of a character from a degree 3 extension, essentially sym^2, or of general type. In the first two cases D_p being global reduces to the $n = 1$ and $n = 2$ case and in the third case it follows from the discussion above (assuming all the above conjectures). This makes it reasonable to make the following conjecture:

Conjecture 5.2.5. Let π be a regular cuspidal algebraic automorphic representation of $\mathrm{GL}_n(\mathbb{A}_K)$ and $\rho_{\pi,p}$ the associated Galois representation and D_p the Azumaya algebra coming from [9, Proposition 2.18]. Then there exists a central simple algebra D over $\mathbb{Q}(\pi)$ such that $D_p \simeq D \otimes_{\mathbb{Q}(\pi)} \mathbb{Q}(\pi)_p$ for every p.

In particular, this conjecture implies that for all but finitely many primes, the representation $\rho_{\pi,p}$ can be defined over its trace field.

Bibliography

[1] James Arthur and Laurent Clozel. *Simple algebras, base change, and the advanced theory of the trace formula*. Number 120 in Ann. of Math. Stud. Princeton Univ. Press, 1989.

[2] Thomas Barnet-Lamb, Toby Gee, and David Geraghty. The Sato-Tate conjecture for Hilbert modular forms. *J. Amer. Math. Soc.*, 24(2):411–469, 2011.

[3] Thomas Barnet-Lamb, Toby Gee, David Geraghty, and Richard Taylor. Potential automorphy and change of weight. *Ann. of Math.*, pages 501–609, 2014.

[4] Joël Bellaıche and Gaëtan Chenevier. Families of Galois representations and Selmer groups. *Astérisque*, 324:1–314, 2009.

[5] Don Blasius and Jonathan Rogawski. Tate classes and arithmetic quotients of the two-ball. *The zeta functions of Picard modular surfaces*, 421:444, 1992.

[6] Gebhard Böckle and Chun-Yin Hui. Weak abelian direct summands and irreducibility of Galois representations. *arXiv:2404.08954*, 2024.

[7] Michel Brion. On extensions of algebraic groups with finite quotient. *Pacific J. Math.*, 279(1-2):135–153, 2015.

[8] Frank Calegari and Toby Gee. Irreducibility of automorphic Galois representations of GL(n), n at most 5. In *Ann. Inst. Fourier*, volume 63, pages 1881–1912, 2013.

[9] Gaëtan Chenevier. *The p-adic analytic space of pseudocharacters of a profinite group and pseudorepresentations over arbitrary rings*, page 221–285. London Math. Soc. Lecture Note Ser. Cambridge Univ. Press, 2014.

[10] Wenchen Chi. Twists of central simple algebras and endomorphism algebras of some abelian varieties. *Math. Ann.*, 276:615–632, 1987.

[11] Laurent Clozel. Motifs et formes automorphes: applications du principe de fonctorialité. In *Automorphic forms, Shimura varieties, and L-functions, Vol. I (Ann Arbor, MI, 1988)*, volume 10 of *Perspect. Math.*, pages 77–159. Academic Press, Boston, MA, 1990.

[12] Andrea Conti, Jaclyn Lang, and Anna Medvedovsky. Big images of two-dimensional pseudorepresentations. *Math. Ann.*, 385(3):1–95, 2023.

[13] Pierre Deligne. Formes modulaires et representations ℓ-adiques. In *Séminaire Bourbaki vol. 1968/69 Exposés 347-363*, pages 139–172. Springer, 2006.

[14] Fred Diamond and Jerry Michael Shurman. *A first course in modular forms*, volume 228. Springer, 2005.

[15] Jean-Marc Fontaine and Yi Ouyang. Theory of p-adic Galois representations. *preprint*, 2008.

[16] Jean-Marc Fontaine and Bernadette Perrin-Riou. Autour des conjectures de Bloch et Kato: cohomologie Galoisienne et valeurs de fonctions l. In *proceedings of Symposia in Pure Mathematics*, pages 599–706. American Mathematical Society, 1994.

[17] Bert van Geemen and Jaap Top. A non-selfdual automorphic representation of GL_3 and a Galois representation. *Invent. Math.*, 117(3):391–401, 1994.

[18] Stephen Gelbart and Hervé Jacquet. A relation between automorphic representations of $\mathrm{GL}(2)$ and $\mathrm{GL}(3)$. *Ann. Sci. École Norm. Sup. (4)*, 11(4):471–542, 1978.

[19] Jayce R Getz and Heekyoung Hahn. *An Introduction to Automorphic Representations: With a View Toward Trace Formulae*, volume 300. Springer Nature, 2024.

[20] Eknath Ghate, Enrique González-Jiménez, and Jordi Quer. On the Brauer class of modular endomorphism algebras. *Int. Math. Res. Not. IMRN*, 2005(12):701–723, 2005.

[21] Xavier Guitart. Abelian varieties with many endomorphisms and their absolutely simple factors. *Rev. Mat. Iberoam.*, 28(2):591–601, 2012.

[22] Michael Harris, Kai-Wen Lan, Richard Taylor, and Jack Thorne. On the rigid cohomology of certain Shimura varieties. *Res. Math. Sci.*, 3:1–308, 2016.

[23] Chun Yin Hui. *Monodromy of Galois representations and equal-rank subalgebra equivalence*. PhD thesis, Indiana University, 2013.

[24] Michael Larsen and Richard Pink. Determining representations from invariant dimensions. *Invent. Math.*, 102(1):377–398, 1990.

[25] Fumiyuki Momose. On the ℓ-adic representations attached to modular forms. *J. Fac. Sci. Univ. Tokyo Sect. IA Math*, 28(1):89–109, 1981.

[26] Ben Moonen. Families of motives and the Mumford–Tate conjecture. *Milan J. Math.*, 85:257–307, 2017.

[27] Jan Nekovář. Level raising and anticyclotomic Selmer groups for Hilbert modular forms of weight two. *Canad. J. Math.*, 64(3):588–668, 2012.

[28] Jordi Quer. La classe de Brauer de l'algèbre d'endomorphismes d'une variété abélienne modulaire. *C. R. Math. Acad. Sci.*, 327(3):227–230, 1998.

[29] Dinakar Ramakrishnan. Irreducibility and cuspidality. *Representation theory and automorphic forms*, pages 1–27, 2008.

[30] Dinakar Ramakrishnan. An exercise concerning the selfdual cusp forms on $\mathrm{GL}(3)$. *Indian J. Pure Appl. Math*, 45(5):777–785, 2014.

[31] Kenneth A Ribet. Galois action on division points of abelian varieties with real multiplications. *Amer. J. Math.*, 98(3):751–804, 1976.

[32] Kenneth A. Ribet. Galois representations attached to eigenforms with nebentypus. In *Modular Functions of one Variable V*, pages 18–52. Springer Berlin Heidelberg, 1977.

[33] Kenneth A Ribet. Twists of modular forms and endomorphisms of abelian varieties. *Math. Ann.*, 253(1):43–62, 1980.

[34] Kenneth A Ribet. Abelian varieties over $\mathbb{Q}$ and modular forms. *Modular curves and abelian varieties*, pages 241–261, 2004.

[35] Raphaël Rouquier. Caractérisation des caracteres et pseudo-caracteres. *J. Algebra*, 180(2):571–586, 1996.

[36] Peter Schneider. *p-adic Lie groups*, volume 344. Springer Science & Business Media, 2011.

[37] Anthony J Scholl. Motives for modular forms. *Invent. Math.*, 100(1):419–430, 1990.

[38] Peter Scholze. On torsion in the cohomology of locally symmetric varieties. *Ann. of Math.*, pages 945–1066, 2015.

[39] Jean-Pierre Serre. Propriétés Galoisiennes des points d'ordre fini des courbes elliptiques. *Invent. Math.*, 15:259–331, 1971/72.

[40] Jean-Pierre Serre. *Abelian l-adic representations and elliptic curves*. AK Peters/CRC Press, 1997.

[41] Jean-Pierre Serre. *Local fields*, volume 67. Springer Science & Business Media, 2013.

[42] Alireza Shavali. On the endomorphism algebra of abelian varieties associated with Hilbert modular forms. *arXiv preprint arXiv:2410.20223*, 2024.

[43] Alireza Shavali. On the image of automorphic Galois representations. *arXiv preprint arXiv:2502.10799*, 2025.

[44] Margaret G Upton. Galois representations attached to Picard curves. *J. Algebra*, 322(4):1038–1059, 2009.

[45] Ariel Weiss. *On Galois representations associated to low weight Hilbert-Siegel modular forms*. PhD thesis, University of Sheffield, 2019.