Automorphic representations with local constraints

by

Jared Seth Weinstein

A.B. (Harvard College) 2002

A Dissertation submitted in partial satisfaction of the requirements for the degree

Doctor of Philosophy

in

Mathematics

at the

University of California, Berkeley

Committee in charge:

Professor Kenneth A. Ribet, Chair
Professor Bjorn Poonen
Professor Mary Gaillard

Spring 2007

The dissertation of Jared Seth Weinstein is approved.

Professor Bjorn Poonen

Professor Mary Gaillard

Professor Kenneth A. Ribet, Committee Chair

Spring 2007

Automorphic representations with local constraints

by

Jared Seth Weinstein

To everyone who loves mathematics,

everywhere.

Acknowledgements

I cannot stress enough how wonderful it is to study mathematics at Berkeley. I had the great privilege to share ideas there with many fantastic mathematicians. My advisor, Ken Ribet, granted me the freedom to pursue this project while masterfully answering the lion's share of my questions. The ideas for the final chapter grew out of interesting conversations with Robert Coleman. I give special thanks to Bjorn Poonen: his generosity with time led to many, *many* improvements in the final draft.

Some more words of gratitude: Stefan Wewers taught me almost everything I know about the local Langlands correspondence. Without having had a conversation with Richard Taylor and Teruyoshi Yoshida, I couldn't possibly have written Section 4.2. I have also profited from communication with Akshay Venkatesh, Darren Semmen, and Chandrashekhar Khare.

Yet more thanks are due to my student colleagues Lauren Deason, Janak Ramakrishnan, Jomy Alappattu, Andrew Presley, and Aaron Greicius for their help and friendship, and to everyone who encouraged me to pursue a career in mathematics. I made the right decision after all!

Part of this research was carried out under an NSF Graduate Fellowship.

Contents

Chapter 1

Introduction

Let f be a modular cusp form for $\Gamma_1(N)$ of weight $k \geq 2$ and character ε which is an eigenvalue for all the Hecke operators. Let ℓ be a prime number, and choose an isomorphism $\iota\colon \mathbf{C} \tilde{\to} \overline{\mathbf{Q}}_\ell$. By a construction of Deligne, there is a corresponding Galois representation

$$\rho_f\colon \operatorname{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathrm{GL}_2(\overline{\mathbf{Q}}_\ell),$$

unramified outside ℓN. That is, the restriction $\rho_{f,p} = \rho_f|_{\operatorname{Gal}(\overline{\mathbf{Q}}_p/\mathbf{Q}_p)}$ vanishes on the inertia group $I_{\mathbf{Q}_p} \subset \operatorname{Gal}(\overline{\mathbf{Q}}_p/\mathbf{Q}_p)$ unless $p|\ell N$. Note that $\det \rho_{f,p}\colon \operatorname{Gal}(\overline{\mathbf{Q}}_p/\mathbf{Q}_p)^{\mathrm{ab}} \to \mathbf{C}^*$ is identified via local class field theory with the local component of ε at p. In particular $\prod_p \det \rho_{f,p}(-1) = \varepsilon(-1) = (-1)^k$.

Our main theorem establishes the existence of cuspidal eigenforms f for which ρ_f has prescribed restriction to $I_{\mathbf{Q}_p}$ at all $p \neq \ell$. Specifically, it is

Theorem 1.0.1. *For $p \neq \ell$, let $\{\rho_p\}$ be a family of representations $\rho_p\colon I_{\mathbf{Q}_p} \to \mathrm{GL}_2(\overline{\mathbf{Q}})$ which extend to semisimple representations of $\operatorname{Gal}(\overline{\mathbf{Q}}_p/\mathbf{Q}_p)$. Assume that ρ_p is trivial for almost all p, including $p = 2, 3$, and that $\prod_p \det \rho_p(-1) = 1$. Let $k \geq 2$ be even. Then there exists a cuspidal eigenform f of weight k and prime-to-ℓ level for which*

$$\rho_{f,p}|I_{\mathbf{Q}_p} = \rho_p,$$

unless the pair $(\{\rho_p\}, k)$ belongs to a finite set of exceptional families or their twists.

A more precise form of the theorem appears in Chapter 3, where we work instead with complex Weil-Deligne representations rather than the ℓ-adic representations $\rho_{f,p}$, in order to remove from the discussion the auxiliary prime ℓ. Theorem 1.0.1 is a Grunwald-Wang type theorem for GL_2 which generalizes a theorem in [KP96], which establishes our main result when $k = 2$ and when one ρ_p is irreducible and tame and the others are trivial.

The proof of Theorem 1.0.1 combines two methods, one local and one global. The first method is a detailed study of the *inertial types* of two-dimensional representations ρ_0 of the absolute inertia group I_F of a p-adic field F which extend to the Weil group W_F. We first recall the Local Langlands Correspondence for GL_2: this is a correspondence $\rho \mapsto \mathrm{rec}(\rho)$ between two-dimensional representations of W_F and irreducible admissible representations of $\mathrm{GL}_2(F)$. Now if ρ_0 is a representation of I_F extending to W_F, one can associate to ρ_0 a certain finite-dimensional representation $\tau(\rho_0)$ of the compact group $K = \mathrm{GL}_2(\mathcal{O}_F)$. We call $\tau(\rho_0)$ the *inertial type* of ρ_0; it is characterized by the property that it appears in $\mathrm{rec}(\rho)|_K$ for every lift ρ of ρ_0 to W_F. In Chapter 2 we determine the character of $\tau(\rho_0)$ on a large portion of the conjugacy classes of K.

The second method involves the relationship between (global) automorphic representations and Galois representations. To a newform f as in the first paragraph, one has an associated representation π_f of the adelic group $\mathrm{GL}_2(\mathbf{A})$. By a theorem of Deligne-Carayol ([Car83]), the local component $\pi_{f,p}$ determines $\rho_{f,p}$ for all $p \neq \ell$, and Theorem 1.0.1 becomes a question of constructing a π_f for which $\pi_{f,p}|_{\mathrm{GL}_2(\mathbf{Z}_p)}$ contains $\tau(\rho_p)$ for all p *and* for which $\pi_{f,\infty}$ is discrete series of lowest weight k. This is a matter of producing automorphic representations subject to local constraints, and yet it admits a solution using surprisingly classical technology, namely a study of the geometry of the classical modular curves $X(N)$.

To wit, let $(\{\rho_p\}, k)$ be as in the hypotheses of Theorem 1.0.1, let $\lambda = \bigotimes_p \tau(\rho_p)$; this is a representation of the profinite group $\mathrm{GL}_2(\hat{\mathbf{Z}})$ which factors through a finite quotient $G = \mathrm{GL}_2(\mathbf{Z}/N\mathbf{Z})/\{\pm 1\}$ for some N. Let S_k denote the space of cusp forms of weight k for the full congruence subgroup $\Gamma(N)$. Then G acts on S_k. We show that

Theorem 1.0.2. *The number of newforms f of weight k for which $\pi_{f,p}|_{\mathrm{GL}_2(\mathbf{Z}_p)}$ contains $\tau(\rho_p)$ equals the multiplicity of λ in S_k as a G-module.*

We then compute the G-module structure of S_k completely, using an "equivariant" version of the Riemann-Roch formula applied to the Galois cover of curves $X(N) \to X(1)$ with group G. We prove: that

Theorem 1.0.3. *In the Grothendieck ring of G we have $[S_k] = \lfloor k/12 \rfloor \mathbf{C}G + \varepsilon_k$, where ε_k depends only on $k \pmod{12}$.*

For the exact statement, see Section 3.4. It implies that λ appears in S_k at least once if $k \geq 14$. If $k \leq 12$ we show that there exists a bound M for which λ appears in S_k unless some twist $\lambda \times \chi \circ \det$ of λ factors through some $\mathrm{GL}_2(\mathbf{Z}/m\mathbf{Z})$ for $m \leq M$. Up to twisting there are of course only finitely many such exceptional λ; this establishes Theorem 1.0.1.

We use the calculations behind Theorem 1.0.1 to prove a theorem about abelian varieties with everywhere good reduction:

Theorem 1.0.4. *Assume $F/\mathbf{Q}$ is Galois, and that there is a ramified prime $\mathfrak{p}|p$ of F for which one of the following holds:*

1. *$p \geq 29$, $p \equiv 1 \mod 4$, and $F_{\mathfrak{p}}/\mathbf{Q}_p$ is ramified quadratic,*
2. *$p \geq 23$ and $F_{\mathfrak{p}}/\mathbf{Q}_p$ is cyclic with ramification degree at least 3, or*
3. *$p \geq 17$, $p \equiv 1 \mod 4$ and $\mathrm{Gal}(F_{\mathfrak{p}}/\mathbf{Q}_p)$ is a dihedral group of order at least 6.*

Then there exists a modular abelian variety $A/\mathbf{Q}$ for which A_F has everywhere good reduction.

This is a strengthening of a theorem of Langlands (Prop. 2 on p. 263 in [MW84]), which implies that for $p \geq 17$, the kernel of $J_1(p) \to J_0(p)$ is a nonzero abelian variety with good reduction everywhere over $\mathbf{Q}(\zeta_p)$. According to our theorem, if F is of the type specified in the hypotheses, then for a suitable n there exists a nonzero abelian variety quotient of $J_1(p^n)$ which attains everywhere good reduction over F.

In the last chapter, we calculate the minimal field extension of $\mathbf{Q}_p^{\mathrm{nr}}$ over which the modular abelian variety $J_1(p^n)$ becomes semi-stable, thus proving a converse to a result of Krir ([Kri96]). This leads us into a discussion of stable models for modular curves. We remark that the structure of the special fiber of a stable model has been calculated through various sorts of methods for $X_0(p)$ ([DR73]), for $X_0(p^2)$ ([Edi90]), for $X(p)$ ([BW04]), for $X_0(p^3)$ ([CM06]), and for $X_1(p^2)$ ([Joy06]). The general case of $X_1(p^n)$ remains open for $n \geq 4$. Chapter 4 gives a conjectural form for the stable reduction of $X(p^2)$.

In Section 4.3 we review the theory of Deligne-Carayol regarding the cohomology of modular curves. Let p be prime and let $N \geq 4$ be prime to p. For $n \geq 1$, let $X_n/\mathbf{Q}_p$ be the curve parametrizing elliptic curves with full level p^n structure, along with a point of order N. The ℓ-adic ($\ell \neq p$) cohomology of the tower of modular curves X_n inherits an action of the product group $\mathrm{GL}_2(\mathbf{Q}_p) \times W_{\mathbf{Q}_p}$ and decomposes as a sum of terms of the form $\pi_{f,p} \otimes \rho_{f,p}$. Here f runs over newforms whose prime-to-p conductor divides N, $\pi_{f,p}$ is the local component of the automorphic representation attached to f, and $\rho_{f,p}$ is the local ℓ-adic representation attached to f at p. This theory involves a delicate study of a regular model for X_n over $\mathbf{Z}_p^{\mathrm{nr}}$ and the action of a quaternion group on the vanishing cycles on that model.

On the other hand, the analysis of the cohomology of a curve X over a p-adic field becomes dramatically simpler once one can compute a stable model for X, rather than just a regular one, via Grothendieck's theory of semi-stable reduction. The cohomology of the modular curves X_n already being known, it seems possible to reverse-engineer the stable reduction of X_n in a manner consistent both with the Deligne-Carayol theory and with the cases already established for n small. In the last section, we conjecture a form of the stable reduction of X_2 together with an action of $\mathrm{GL}_2(\mathbf{Z}/p^2\mathbf{Z}) \times I_{\mathbf{Q}_p}$, and prove that this form is consistent with Deligne-Carayol. The conjecture seems not too far off in terms of a solution, and one hopes that this line of investigation may lead to a purely local approach to the local Langlands correspondence and to the theory of Deligne-Carayol.

1.1 Notation

A complex-valued *character* of a group is any homomorphism from that group into $\mathbf{C}^*$, not necessarily a unitary one.

When F is a finite extension of $\mathbf{Q}_p$, we will always use the symbols $\mathcal{O}_F$ and $\mathfrak{p}_F$ for the ring of integers of F and its maximal ideal, respectively. When χ is a character of $\mathcal{O}_F^*$, we let $c(\chi)$ denote the smallest ideal $\mathfrak{p}_F^n$ for which χ vanishes on $U_F^n := 1 + \mathfrak{p}_F^n$.

When G is a compact group and ψ and ϕ are two finite-dimensional representations of G over a field of characteristic zero, we use $\langle \psi, \phi \rangle_G$ to mean $\dim \operatorname{Hom}_G(\psi, \phi)$. In the case that G is finite, this reduces to the usual formula

$$\langle \psi, \phi \rangle_G = \frac{1}{\#G} \sum_{g \in G} \operatorname{Tr} \psi(g) \operatorname{Tr} \phi(g^{-1}).$$

Chapter 2

Study of inertial types for GL_2

2.1 Number theoretic background: Weil-Deligne representations

The title of this section, as well as its content, is adapted from [Tat79]. We are going to develop the tools to discuss representations of local Galois groups in the context of the local Langlands correspondence.

Let $F/\mathbf{Q}_p$ be a finite extension, with ring of integers $\mathcal{O}_F$, uniformizer ϖ and residue field k. Let $q = \#k$ and let $|\ |_F$ be the absolute value on F for which $|\varpi|_F = 1/q$.

Let G_F and I_F denote the absolute Galois group and inertia group of F, respectively. An attempt to investigate the structure of G_F naturally leads to the study of linear representations of this group. However, the category of continuous representations $G_F \to \mathrm{GL}_n(\mathbf{C})$ is not quite rich enough for many purposes. For instance, the image of such a representation must be compact and totally disconnected, hence finite. For arithmetic applications we pass to the study of Weil-Deligne representations.

Definition 2.1.1. The *Weil group* W_F is the preimage of $\mathbf{Z}$ in the map $G_F \to \mathrm{Gal}(\overline{k}/k) \cong \hat{\mathbf{Z}}$. Its topology is such that $I_F \subset W_F$ is open and I_F is given the topology induced from G_F.

The Weil group fits into the diagram

$$\begin{tikzcd}
1 \arrow[r] & I_F \arrow[r] \arrow[d, equal] & W_F \arrow[r] \arrow[d, hook] & \mathbf{Z} \arrow[r] \arrow[d, hook] & 0 \\
1 \arrow[r] & I_F \arrow[r] & G_F \arrow[r] & \hat{\mathbf{Z}} \arrow[r] & 0.
\end{tikzcd}$$

Let W_F^{ab} be the largest abelian Hausdorff quotient of W_F. Local class field theory furnishes an isomorphism $\mathrm{Art}\colon F^* \tilde{\to} W_F^{\mathrm{ab}}$, which sends a uniformizer of F^* to a geometric Frobenius element. We define an absolute value on W_F by

$$|\sigma| = \left|\mathrm{Art}^{-1}(\sigma)\right|_F .$$

Let $\ell \neq p$ be prime.

Definition 2.1.2. An *ℓ-adic representation* of F is a continuous homomorphism $\rho_\ell\colon W_F \to \mathrm{GL}_n(E)$, where $E/\mathbf{Q}_\ell$ is algebraic and $\rho_\ell(\Phi)$ is semisimple for every Frobenius element $\Phi \in G_F$.

From the point of view of arithmetic geometry, these are the sort of local Galois representations that occur "in the wild," by which we mean inside the ℓ-adic cohomology $H^*(X \otimes \overline{F}, E)$ of a variety defined over F. The philosophy developed in [Tat79] suggests that the choice of prime ℓ is irrelevant to the discussion so long as $\ell \neq p$, and that one might as well work over the field of complex numbers. To that end, we define

Definition 2.1.3. A *Weil-Deligne (WD) representation* of F over a field E of characteristic 0 is a pair (ρ, N), where

$$\rho\colon W_F \to \mathrm{GL}_n(E)$$

is a homomorphism, $\mathrm{N} \in \mathrm{End}\, E^n$, and the pair (ρ, N) satisfies the conditions

1. ρ is continuous for the discrete topology on E,

2. $\rho(\Phi) \in \mathrm{GL}_n(E)$ is semisimple for every Frobenius element $\Phi \in W_K$, and

3. $\rho(\sigma)\,\mathrm{N}\,\rho(\sigma)^{-1} = |\sigma|\,\mathrm{N}$ for all $\sigma \in W_F$.

The matrix N is called the *monodromy operator*. The relationship 3 above implies that it is nilpotent, for if λ is an eigenvalue of N, then $\lambda = |\sigma|\,\lambda$ for every $\sigma \in W_F$, whence $\lambda = 0$. When referring to a WD representation of F over E, we will often omit the monodromy operator from the notation and instead refer to the pair (ρ, N) simply as ρ. Let $\mathcal{G}_n(F)_E$ denote the set of WD representations of F over E up to isomorphism. Let $\mathcal{G}_n(F)_E^{\mathrm{ss}} \subset \mathcal{G}_n(F)_E$ be the set of isomorphism classes of WD representations with $\mathrm{N} = 0$. Finally let $\mathcal{G}_n(F)_E^0 \subset \mathcal{G}_n(F)_E^{\mathrm{ss}}$ denote the set of isomorphism classes of WD representations $(\rho, 0)$ of F over E for which ρ is irreducible.

Theorem 2.1.4 (Grothendieck, see [Tat79])**.** *Let E be a nonarchimedean local field containing $\mathbf{Q}_\ell$. There is a bijection*

$$\mathcal{G}_n(F)_E \xrightarrow{\sim} \left\{ \begin{matrix} \textit{isomorphism classes of} \\ \ell\textit{-adic representations} \\ \rho_\ell \colon W_F \to \mathrm{GL}_n(E) \end{matrix} \right\}. \tag{2.1}$$

The relationship between a WD representation (ρ, N) and its image ρ_ℓ in the above bijection is given as follows: Let $\Phi \in W_F$ be a Frobenius element and let $t_\ell \colon I_K \to E$ be an ℓ-adic tame character. Then for all $\sigma \in I_F$ and $a \in \mathbf{Z}$:

$$\rho_\ell(\Phi^a \sigma) = \rho(\Phi^a \sigma) \exp(t_\ell(\sigma)\,\mathrm{N}).$$

The bijection in the theorem does not depend on the choices of Φ and t_ℓ.

The advantage of working with WD representations is that it allows us to sidestep the choice of the prime ℓ. Indeed, since the topology on the field of scalars is irrelevant in Definition 2.1.3, any abstract isomorphism of fields $E \cong E'$ gives a bijection $\mathcal{G}_n(F)_E \to \mathcal{G}_n(F)_{E'}$. When the field E is omitted from the notation, it is assumed that the field of scalars is $\mathbf{C}$.

WD representations admit direct sums by the formula

$$(\rho_1, \mathrm{N}_1) \oplus (\rho_2, \mathrm{N}_2) = (\rho_1 \oplus \rho_2, \mathrm{N}_1 \oplus \mathrm{N}_2)$$

and tensor products by the formula

$$(\rho_1, \mathrm{N}_1) \otimes (\rho_2, \mathrm{N}_2) = (\rho_1 \otimes \rho_2, (\mathrm{Id}_1 \otimes \mathrm{N}_2) \oplus (\mathrm{N}_1 \otimes \mathrm{Id}_2)),$$

where Id_i is the identity endomorphism of the representation space of ρ_i for $i = 1, 2$. These definitions are compatible via the dictionary (2.1) with the usual notions of direct sum and tensor product of ℓ-adic representations of W_F.

If L/F is a finite extension of degree d, we have the usual notion of induced representation $\mathrm{Ind}_{W_L}^{W_F} : \mathcal{G}_n(L)^{\mathrm{ss}} \to \mathcal{G}_{nd}(F)^{\mathrm{ss}}$. We abbreviate this as $\mathrm{Ind}_{L/F}$.

Example 2.1.5. Let $n \geq 1$. We define the special representation $\mathrm{Sp}(n)$ as the pair (ρ, N), where $\rho \colon W_F \to \mathrm{GL}_n(\mathbf{Q})$ is defined by

$$g \mapsto \begin{pmatrix} 1 & & & \\ & |g|_F & & \\ & & \ddots & \\ & & & |g|_F^{n-1} \end{pmatrix}$$

and N is the matrix

$$\mathrm{N} = \begin{pmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & 0 & 1 \\ & & & & 0 \end{pmatrix}.$$

The next proposition shows that any WD representation is a sum of terms of the form $\rho \otimes \mathrm{Sp}(n)$:

Proposition 2.1.6 ([Tat79], 4.1.5)**.** *Let $\rho \in \mathcal{G}_n(F)$ be indecomposable, i.e. not the sum of two nonzero WD representations. Then ρ is of the form $\rho' \otimes \mathrm{Sp}(m)$ for a divisor m of n and an irreducible $\rho' \in \mathcal{G}_{n/m}(F)$.*

Let $G_i \subset G_F$ be the filtration by ramification subgroups.

Definition 2.1.7. Let $\rho \in \mathcal{G}_n(F)^{\mathrm{ss}}$. The *Artin conductor* of ρ is

$$c(\rho) = \sum_{i=0}^{\infty} [G_F : G_i] \left(n - \langle \rho|_{G_i}, 1 \rangle_{G_i} \right).$$

If $\rho = \sigma \otimes \mathrm{Sp}(d)$ for $d > 1$, then $c(\rho) := \min \{c(\sigma), 1\}$.

For proofs of Propositions 2.1.8 and 2.1.10 below, we refer to [Ser67].

Proposition 2.1.8. *We have the following facts concerning the Artin conductor:*

1. $c(\rho) \in \mathbf{Z}$.

2. *If* $\chi \in \mathcal{G}_1(F) = \mathrm{Hom}(F^*, \mathbf{C}^*)$, *then* $c(\chi)$ *is the conductor of* χ *in the usual sense,*

3. *If* L/F *is a finite extension and* $\theta \in \mathcal{G}_n(L)^{ss}$, *then*

$$c(\mathrm{Ind}_{L/F}\,\theta) = nv_F(D_{L/F}) + f_{L/F}c(\theta),$$

   where $f_{L/F}$ *is the residue degree of* L/F *and* $D_{L/F}$ *is the discriminant.*

4. $c(\rho)$ *only depends on* $\rho|_{I_F}$.

Definition 2.1.9. An irreducible WD representation $\rho \in \mathcal{G}_n^0(F)$ is *imprimitive* if there exists a nontrivial extension K/F and a representation $\theta \in \mathcal{G}_m^0(F)$ for which $\rho = \mathrm{Ind}_{K/F}\,\theta$. If ρ is not imprimitive then it is *primitive*.

Proposition 2.1.10. *If* $p > n$ *then every* $\rho \in \mathcal{G}_n^0(F)$ *is imprimitive.*

Example 2.1.11. The category $\mathcal{G}_1(F)$ of one-dimensional WD representations of F is easy to study, thanks to the isomorphism $\mathrm{Art} : F^* \to W_F^{\mathrm{ab}}$ of local class field theory. If $\rho \in \mathcal{G}_1(F)$, its monodromy operator, being a 1×1 nilpotent matrix, must vanish. Thus ρ is simply a homomorphism $W_F^{\mathrm{ab}} \to \mathbf{C}^*$; composition with Art gives us an identification

$$\mathrm{rec} : \mathcal{G}_1(F) \xrightarrow{\sim} \mathrm{Hom}(F^*, \mathbf{C}).$$

We will have such frequent occasion to use this identification that if $\chi \in \mathcal{G}_1(F)$ we will abbreviate the character $\chi \circ \mathrm{Art}$ as simply χ.

Definition 2.1.12. A WD representation $\rho \in \mathcal{G}_2(F)$ is *minimal* if $c(\chi \otimes \rho) \geq c(\rho)$ for each $\chi \in \mathcal{G}_1(F)$.

Example 2.1.13. A WD representation $\rho \in \mathcal{G}_2(F)$ is a twist by a character of F^* of a WD representation of one of the following types:

1. $\rho = \varepsilon \oplus 1$ for a character ε of F^*. Then $c(\rho) = c(\varepsilon)$.

2. $\rho = \mathrm{Sp}(2)$. Then $c(\rho) = 1$.

3. $\rho = \mathrm{Ind}_{L/F}\,\theta$, where θ is a minimal character of L^*, meaning that $c(\theta \times \chi \circ \mathrm{N}_{L/F}) \geq c(\theta)$ for all characters χ of F^*: then by part (3) of Proposition 2.1.8, $c(\rho) = 2c(\theta)$ if L/F is unramified and $c(\rho) = c(\theta) + 1$ if L/F is ramified

4. ρ is a primitive representation: occurs only when $p = 2$, by Proposition 2.1.10.

2.2 The Local Langlands correspondence

Example 2.1.11 shows that $\mathcal{G}_1(F)$ is in bijection with irreducible representations of $F^* = \mathrm{GL}_1(F)$. In higher dimensions, the *local Langlands correspondence* (LLC) puts $\mathcal{G}_n(F)$ into bijection with the set of isomorphism classes of representations of $\mathrm{GL}_n(F)$ of a certain sort.

Definition 2.2.1. Let G be a locally profinite group and $\pi\colon G \to \mathrm{Aut}\, V$ be a representation of G on a vector space over a field of characteristic 0. We say π is *admissible* if

1. for every vector $v \in V$, the stabilizer $\mathrm{Stab}_G(v)$ is an open subgroup of G, and

2. the subspace of V fixed by any compact open subgroup of G is finite-dimensional.

Definition 2.2.2. For $n \geq 1$ and E a field of characteristic 0, let $\mathcal{A}_n(F)_E$ denote the set of isomorphism classes of irreducible admissible representations of $\mathrm{GL}_n(F)$ over E. When E is omitted from the notation, it is assumed that the field of scalars is $\mathbf{C}$.

Let $\psi\colon F \to E^*$ be a nontrivial additive character. When $\pi \in \mathcal{A}_n$, there are notions of *L-factor* $L(\pi, s)$ and *ε-factor* $\varepsilon(\pi, \psi, s)$. If $\chi \in \mathcal{A}_1$, we define the twist $\chi\pi$ as $g \mapsto \chi(\det g)\pi(g)$. We are now ready to state the

Theorem 2.2.3 (Local Langlands Correspondence for GL_n)**.** *Let $\psi\colon F \to \mathbf{C}^*$ be a nontrivial additive character. There is a bijection*

$$\mathrm{rec} : \mathcal{G}_n(F) \xrightarrow{\sim} \mathcal{A}_n(F)$$

such that for all $\rho \in \mathcal{G}_n(F)$ and $\chi \in \mathrm{Hom}(F^*, \mathbf{C}^*)$, we have

$$\begin{aligned} L(\chi\,\mathrm{rec}(\rho), s) &= L(\chi \otimes \rho, s) \\ \varepsilon(\chi\,\mathrm{rec}(\rho), \psi, s) &= \varepsilon(\chi \otimes \rho, \psi, s) \end{aligned}$$

The theorem first made an appearance as a conjecture in [Lan70]. A proof for $n = 2$ was given by Kutzko in [Kut80] and [Kut84]. The full statement of LLC was proved by Harris and Taylor in [HT01] by examining the cohomology of deformation spaces of formal groups with Drinfeld level structure.

There is also a notion of conductor for admissible representations which is compatible under the LLC with the same notion for WD representations. We define it only for $n = 2$. Let $K = \mathrm{GL}_2(\mathcal{O}_F)$ and define a filtration $K_0(N) \subset K$ for $N \geq 1$ by

$$K_0(N) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in K\colon\ c \in \mathfrak{p}_F^N \right\}.$$

When ε is a character of $\mathcal{O}_F^*$, and $N \geq c(\varepsilon)$, extend ε to a character of $K_0(N)$ by $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto \varepsilon(a)$.

Theorem 2.2.4 ([Cas73a])**.** *Let $\pi \in \mathcal{A}_2(F)$ have central character ε. There is an integer N for which there exists a nonzero vector in the space of π on which $\pi(g)$ acts as $\varepsilon(g)$ for all $g \in K_0(N)$. If c is the smallest such integer, then there is exactly one such vector up to scaling.*

Definition 2.2.5. The integer c is the *conductor* of π, and the vector v is the *new vector*.

We note without proof that $c(\rho) = c(\mathrm{rec}(\rho))$ for $\rho \in \mathcal{G}_2(F)$. If ι is an isomorphism from $\mathbf{C}$ onto another field E, we also use the symbol rec to mean the $\mathcal{G}_n(F)_E \xrightarrow{\sim} \mathcal{A}_n(F)_E$ for which rec and ι commute.

2.3 Inertial WD representations

Definition 2.3.1. Let E be a field of characteristic 0. An *inertial WD representation* of F over E is a pair (ρ_0, N) is a continuous finite-dimensional representation $\rho_0 \colon I_F \to \mathrm{GL}_n(E)$, together with an endomorphism $\mathrm{N} \in \operatorname{End} E^n$, for which there exists an extension ρ of ρ_0 to W_K making (ρ, N) into a WD representation of F.

Remark 2.3.2. An equivalent formulation of Definition 2.3.1 runs as follows: Let $\Phi \in W_F$ be a Frobenius element. Then Φ normalizes I_F in W_F. If $\rho_0 \colon I_F \to \mathrm{GL}_n(E)$ is a continuous homomorphism and $\mathrm{N} \in \operatorname{End} E^n$ an endormorphism commuting with the image of ρ, let $(\rho, \mathrm{N})^\Phi$ denote the pair $((\sigma \mapsto \rho(\Phi^{-1}\sigma\Phi)), q^{-1}\,\mathrm{N})$. Then (ρ_0, N) is an inertial WD representation of F if and only if $(\rho, \mathrm{N})^\Phi \cong (\rho, \mathrm{N})$ in the sense that there is a change of basis matrix $M \in \operatorname{End} E^n$ for which $\rho_0^\Phi(\sigma) = M^{-1}\rho_0(\sigma)M$ and $q^{-1}\,\mathrm{N} = M^{-1}\,\mathrm{N}\,M$. Indeed, the required extension of ρ_0 to W_F is given by $\rho(\sigma\Phi^a) = \rho_0(\sigma)M^a$.

We may refer to the pair (ρ_0, N) simply as ρ_0. There is an evident notion of morphism between inertial WD representations, as well as a restriction map $(\rho, \mathrm{N}) \mapsto (\rho|_{I_F}, \mathrm{N})$ carrying WD representations to inertial WD representations. Finally, the Artin conductor $c(\rho_0)$ is defined as the Artin conductor $c(\rho)$ for any $\rho \in \mathcal{G}_n(F)$ for which $\rho|_{I_F} = \rho_0$.

Let $\mathcal{G}_n^I(F)_E$ denote the set of isomorphism classes of n-dimensional inertial WD representations over E. This is the quotient of $\mathcal{G}_n(F)_E$ by the equivalence relation $\rho|_{I_F} \cong \rho'|_{I_F}$. Let $\mathcal{G}_n^I(F)_E^{\mathrm{ss}} \subset \mathcal{G}_n^I(F)_E$ be the set of isomorphism classes of WD representations with $\mathrm{N} = 0$, and let $\mathcal{G}_n^I(F)_E^0 \subset \mathcal{G}_n^I(F)_E^{\mathrm{ss}}$ be the set of isomorphism classes of WD representations with $\mathrm{N} = 0$ which are not the sum of two nonzero inertial WD representations. Note that an irreducible ρ_0 may be the sum of two nonzero representations of I_F, so long as those representations do not admit extensions to W_F. In fact, this is always the case when ρ_0 is the restriction to I_F of $\operatorname{Ind}_{L/F}\theta$, where L/F is unramified and θ is a character of L^* which does not factor through the norm map $N_{L/F} \colon L^* \to K^*$. As usual, when the notion E is suppressed, we assume the field of scalars is $\mathbf{C}$.

Proposition 2.3.3. *There is a natural identification*

$$\mathrm{rec} : \mathcal{G}_1^I(F) \xrightarrow{\sim} \mathrm{Hom}(\mathcal{O}_F^*, \mathbf{C}^*). \tag{2.2}$$

Proof. If $(\rho_0, \mathrm{N}) \in \mathcal{G}_1^I(F)$, then N, being a nilpotent 1×1 matrix, must vanish. Choose a lifting ρ of ρ_0 to W_F; then ρ factors through a character of W_F^{ab}. Since in the class field theory isomorphism $\mathrm{Art} : F^* \to W_F^{\mathrm{ab}}$, the image of $\mathcal{O}_F^*$ equals the image of I_F in $W_F \to W_F^{\mathrm{ab}}$, the character $\chi := \rho \circ \mathrm{Art}\,|_{\mathcal{O}_F^*}$ depends on ρ_0 and not on the lifting ρ.

Conversely if χ is a character of $\mathcal{O}_F^*$, choose an extension $\tilde{\chi}$ of χ to F^* and let $\rho = \tilde{\chi} \circ \mathrm{Art}^{-1} : W_F^{\mathrm{ab}} \to \mathbf{C}^*$; the inverse image of χ in (2.2) is given by $\rho_0 = \rho|_{I_F}$. $\square$

2.4 The Inertial Langlands Correspondence

Proposition 2.3.3 suggests that there ought to be an analogue of the LLC relating the set $\mathcal{G}_n^I(F)$ to the set of isomorphism classes of irreducible representations of the profinite group $K = \mathrm{GL}_n(\mathcal{O}_F)$. Indeed, a map between these sets exists, but is not quite a bijection.

We will restrict our attention to the case of two-dimensional inertial WD representations, although we expect a similar statement to hold in general. Let $\hat{K}$ be the set of isomorphism classes of continuous irreducible complex representations of $K = \mathrm{GL}_2(\mathcal{O}_F)$. These all factor through a finite quotient $K^{(n)} = \mathrm{GL}_2(\mathcal{O}_F/\mathfrak{p}_F^n)$ for some n. For $\lambda \in \hat{K}$, define the *level* $\ell(\lambda)$ to be the least such n.

Theorem 2.4.1 (Inertial LLC)**.** *There exists a map*

$$\tau \colon \mathcal{G}_2^I(F) \to \hat{K}$$

satisfying the following properties:

1. *If $\rho \in \mathcal{G}_2(F)$, then $\mathrm{rec}(\rho)|_K$ contains $\tau(\rho|_{I_F})$ with multiplicity one.*

2. *Conversely, if* $\rho \in \mathcal{G}_2(F)$ *and* $\eta \in \mathcal{G}_2^I(F)$ *are such that* $\operatorname{rec}(\rho)|_K$ *contains* $\tau(\eta)$*, then* $\rho|_{I_F} \cong \eta$ ***unless*** $\rho = \varepsilon_1 \oplus \varepsilon_2$ *for characters* $\varepsilon_1, \varepsilon_2$ *of* F^* *such that* $\varepsilon_1\varepsilon_2^{-1} = |\ |_F^{\pm 1}$. *In that case, either* $\rho|_{I_F} \cong \eta$ *or else* $\eta \cong \varepsilon_1|_{\mathcal{O}_F^*} \otimes \mathrm{Sp}(2)$.

3. *If* $\rho, \rho' \in \mathcal{G}_2(F)$ *have* $\rho|_{I_F} \cong \rho'|_{I_F}$ *then* $\operatorname{rec}(\rho)|_K \cong \operatorname{rec}(\rho')|_K$.

Definition 2.4.2. When $\rho_0 \in \mathcal{G}_2^I(F)$, we refer to $\tau(\rho_0)$ as the *inertial type* of ρ_0.

Remark 2.4.3. The theorem is the content of [Hen02]. In the next section we give the construction of τ presented therein. We remark that if $p = \operatorname{char} k \neq 2$ then τ is unique for the properties (1) and (2) above. And in any residue characteristic, a map $\tau\colon \mathcal{G}_2^I(F)^0 \to \hat{K}$ is uniquely determined by (1) and (2) without the "unless" clause in (2).

Remark 2.4.4. For all dimensions n, a construction of the analogue of τ for irreducible ρ satisfying (1) and (2) is given in [Pas05].

2.5 Characters of the inertial types

Recall that $K = \mathrm{GL}_2(\mathcal{O}_F)$.

As we describe the inertial types $\tau(\rho_0)$ for $\mathrm{GL}_2(F)$ we will determine the characters of $\tau(\rho_0)$ on particular sorts of elements of K, namely

1. $g = \begin{pmatrix} 1 & a \\ & 1 \end{pmatrix}$: the "parabolic case"
2. $g = \begin{pmatrix} x & \\ & 1 \end{pmatrix}$: the "hyperbolic case"
3. $g = \iota(\alpha)$, where $\iota\colon \mathcal{O}_{F'} \to M_2(\mathcal{O}_F)$ is an embedding of an unramified quadratic extension F'/F and $\overline{\alpha} \notin k^*$: the "elliptic case."

g	$\varepsilon_0 \oplus 1$	$\mathrm{Sp}(2)$	$\mathrm{Ind}_{L/F}\,\theta$, L/F unram.	$\mathrm{Ind}_{L/F}\,\theta$, L/F ram.
Id_2	$q^{n-1}(q+1)$	q	$q^{n-1}(q-1)$	$q^{n-2}(q^2-1)$
$u(\varpi^{n-1})$	q^{n-1}	0	$-q^{n-1}$	$-q^{n-2}$
$u(\varpi^{n-2})$	0	N/A	0	$-q^{n-2}$
$u(\varpi^r)$, $r < n-2$	0	N/A	0	0
$\delta(x)$	$\varepsilon_0(x)+1$	1	0	0
$\iota(\alpha)$	0	-1	$(-1)^n(\theta(\alpha)+\theta(\alpha^\sigma))$	0

Figure 2.1. The characters of $\tau(\rho)$ on certain elements g of G. It is assumed that $g \not\equiv \mathrm{Id}_2 \pmod{\mathfrak{p}_F^n}$.

We will find it useful to put

$$\begin{aligned} u(x) &= \begin{pmatrix} 1 & x \\ & 1 \end{pmatrix} \\ \delta(x) &= \begin{pmatrix} x & \\ & 1 \end{pmatrix} \end{aligned}$$

The next theorem shows that when $\rho_0 \in \mathcal{G}_2^I(F)$, the inertial type $\tau(\rho_0)$ has small trace on elliptic and hyperbolic elements. As we will see in Chapter 3, this fact enables us to give a sharp estimate of the multiplicity of the inertial types inside spaces of cusp forms.

Theorem 2.5.1. *Up to twist, the characters of $\tau(\rho_0)$ belonging to each nontrivial $\rho_0 \in \mathcal{G}_2^I(F)$ which is decomposable or imprimitive are given in Figure 2.1. In particular, let $\lambda = \tau(\rho_0)$ be a type of level n, and let $g \in K$ be elliptic or hyperbolic with $g \not\equiv \mathrm{Id}_2 \pmod{\mathfrak{p}_F^n}$. Then*

$$|\mathrm{Tr}\,\lambda(g)| \leq 2. \tag{2.3}$$

In the proof of Theorem 2.5.1, the following form of Mackey's Theorem will be useful:

Lemma 2.5.2. *Let G be a profinite group, $H \subset G$ a subgroup of finite index, and η a finite-dimensional representation of H. Then if $\tau = \mathrm{Ind}_H^G\,\eta$, we have*

$$\mathrm{Tr}\,\tau(g) = \sum_{[x]\in(G/H)^g} \mathrm{Tr}\,\eta(x^{-1}gx),$$

where the sum ranges over cosets $[x] = xH \in G/H$ which are fixed under the left action of g.

Note that $[x] \in (G/H)^g$ is equivalent to $x^{-1}gx \in H$; we use the former formulation because in our situation the coset space G/H can be given a geometric interpretation.

The proof of the theorem proceeds case by case.

2.5.1 The principal series case

Suppose $\rho_0 \in \mathcal{G}_2^I(F)$ is decomposable. After twisting by a character of $\mathcal{O}_F^*$ we may assume $\rho_0 = \varepsilon_0 \oplus 1$ for a character ε_0 of $\mathcal{O}_F^*$. If $\varepsilon_0 = 1$ then $\tau(\rho_0) = 1_K$. Assume $\varepsilon_0 \neq 1$ has conductor $\mathfrak{p}^n$ for $n > 0$; in this case $\operatorname{rec}(\rho)$ is infinite-dimensional principal series for any lift ρ of ρ_0 to $\mathcal{G}_2(F)$. The inertial type of ρ_0 is

$$\tau(\rho_0) = \operatorname{Ind}_{K_0(n)}^K \varepsilon, \text{ where } \varepsilon \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \varepsilon(a).$$

We can identify the coset space $K/K_0(n)$ with $\mathbf{P}^1(\mathcal{O}_F/\mathfrak{p}^n)$ via $g \mapsto g\infty$. In particular $\dim \tau(\rho_0) = q^{n-1}(q+1)$.

In the parabolic case, we have up to conjugacy $g = u(\varpi^r)$ for some $0 \leq r < n$. The fixed points of g on $\mathbf{P}^1(\mathcal{O}_F/\mathfrak{p}^n)$ are precisely the points $[a : \varpi^i]$ with $i \geq (n-r)/2$, with a running through a set of representatives for $(\mathcal{O}_F/\mathfrak{p}^{n-i})^*$. We find

$$\operatorname{Tr} \tau(g) = \sum_{i=\lceil (n-r)/2 \rceil}^{n} \sum_{a \in (\mathcal{O}_F/\mathfrak{p}^{n-i})^*} \varepsilon_0(1 - a^{-1}\varpi^{i+k}).$$

This is simply the sum of $\varepsilon_0(1+t)$ for t running through a set of representatives of $\mathfrak{p}^{\lceil (n+r)/2 \rceil}/\mathfrak{p}^{n+r}$. Because $c(\varepsilon_0) = \mathfrak{p}^n$, the sum vanishes unless $\lceil (n+r)/2 \rceil = n$, that is $r = n-1$, in which case the sum is q^{n-1}.

In g is hyperbolic, then g has exactly the fixed points 0 and ∞ in $\mathbf{P}^1(\mathcal{O}_F/\mathfrak{p}^n)$, corresponding to the left cosets $hK_0(n)$ with $h = \mathrm{Id}_2$ and $h = \begin{pmatrix} & 1 \\ 1 & \end{pmatrix}$, respectively. Lemma 2.5.2 implies that $\operatorname{Tr} \tau(\rho_0)(g) = \varepsilon_0(g) + 1$.

If g is elliptic, it has no fixed points in $\mathbf{P}^1(\mathcal{O}_F/\mathfrak{p}^n)$, so that $\operatorname{Tr} \tau(\rho_0)(g) = 0$.

2.5.2 The Steinberg case

Now suppose $\rho_0 = \mathrm{Sp}(2)|_{I_F}$. Then $\tau(\rho_0) = \mathrm{St}_K$ is the Steinberg representation of K. This is the pullback through $K \mapsto \mathrm{GL}_2(k)$ of the representation of $\mathrm{GL}_2(k)$ on $\mathrm{Ind}_B^{\mathrm{GL}_2(k)}[1] - [1]$, where $B \subset \mathrm{GL}_2(k)$ is the subgroup of upper triangular matrices. It is simple to check that St_K has character as shown in Figure 2.1.

2.5.3 The cuspidal case

In this section, ρ_0 is irreducible and imprimitive. This means it is the restriction to I_F of $\mathrm{Ind}_{L/F}\,\theta$ for a quadratic extension L/F and a character θ of L^*. We note that ρ_0 is determined by $\theta|_{\mathcal{O}_L^*}$. By twisting, assume that θ is minimal of conductor c. Let e be the ramification degree of L/F. We remark that c is even if $e = 2$, and that the level n of $\tau(\rho_0)$ is c if $e = 1$ and $c/2 + 1$ if $e = 2$.

Let $\mathfrak{A} \subset M_2(F)$ be the algebra $M_2(\mathcal{O}_F)$ if $e = 1$ and

$$\left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_2(\mathcal{O}_F) | c \in \mathfrak{p} \right\} \text{ if } e = 2.$$

Then $\mathfrak{A}$ is a hereditary chain order with $e_{\mathfrak{A}} = e$ (see [HB06]). Choose an embedding of $\mathcal{O}_F$-algebras $\mathcal{O}_L \hookrightarrow \mathfrak{A}$ and identify $\mathfrak{A}$ with $\mathcal{O}_L \oplus \mathcal{O}_L\sigma \subset \mathrm{End}_{\mathcal{O}_F}\,\mathcal{O}_L$.

Let $\mathfrak{P}$ be the Jacobson radical of $\mathfrak{A}$, and let $U_{\mathfrak{A}}^j \subset \mathfrak{A}^* \subset K$ be the subgroup $1 + \mathfrak{P}^j$ for $j > 0$. Also define $U_0 = \mathcal{A}^*$.

Explicitly, $\mathfrak{P}^j = \mathfrak{p}^j M_2(\mathcal{O}_F)$ in the case of $e = 1$ and when $e = 2$ we have

$$\mathfrak{P}^j = \begin{pmatrix} \mathfrak{p}^{\lceil j/2 \rceil} & \mathfrak{p}^{\lceil (j-1)/2 \rceil} \\ \mathfrak{p}^{\lceil (j+1)/2 \rceil} & \mathfrak{p}^{\lceil j/2 \rceil} \end{pmatrix} = \mathfrak{p}_L^{\lceil j/2 \rceil} + \mathfrak{p}_L^{\lceil (j+1)/2 \rceil}\sigma.$$

Let m be defined by $m = \lfloor c/2 \rfloor$ if $e = 1$ and $m = c - 1$ if $e = 2$. Let $J = L^* U_{\mathfrak{A}}^m$ and $J^0 = \mathcal{O}_L^* U_{\mathfrak{A}}^m$. We are going to define a certain finite-dimensional representation η of J^0 depending on θ for which $\tau(\rho_0) = \mathrm{Ind}_{J^0}^K\,\lambda$. We define the representation η in cases.

If L/F is unramified and $c = 1$, then θ arises by pullback from a character of the unique quadratic extension k_2 of k. Then η is the character of $J = K$ arising by pullback from the cuspidal representation of $\mathrm{GL}_2(k)$ associated with θ. The construction of η is described in many places, see for instance Section 3.1 of [CDT99].

If L/F is unramified and $c > 1$ is even, then $\eta\colon J^0 \to \mathbf{C}^*$ is the character $\alpha + \beta\sigma \mapsto \chi(\alpha)$.

If L/F is unramified and $c > 1$ is odd, then η is the unique irreducible representation of J^0 of dimension q satisfying $\operatorname{Tr}\eta(\alpha+\beta\sigma) = -\chi(\alpha)$ when $\beta \in \mathfrak{p}^{(c+1)/2}$ and $\alpha \pmod{\mathfrak{p}_E} \notin k^*$.

If L/F is ramified, then $\eta\colon J^0 \to \mathbf{C}^*$ is the character $\alpha + \beta\sigma \mapsto \chi(\alpha)$.

Theorem 2.5.3. *In all cases the type of ρ_0 is $\operatorname{Ind}_{J^0}^K \eta$.*

This is essentially Lemma 1.2.1 of [BCDT01]. Note that τ vanishes on $K(n)$ but not on $K(n-1)$. We now check the final three columns of Figure 2.1. Before going on we record a lemma about the coset space $\mathfrak{A}^*/J^0$ analogous to the familiar identification of the coset space $\mathrm{GL}_2(\mathbf{R})/\mathrm{O}_2(\mathbf{R})$ with (two copies of) the complex upper half plane via $g \mapsto g\sqrt{-1}$. Besides the action of $\mathfrak{A}^*$ on L coming from the inclusion $\mathfrak{A} \subset \operatorname{End}_F L$, there is another action by fractional linear transformations, which we will notate as $(g,\beta) \mapsto g\cdot\beta$. Then there exists a $b \in F^*$ for which $\mathcal{O}_L = \{g \in \mathfrak{A} | g\cdot b = b\}$. We can take $v(b) = 0$ if $e = 1$ and $v(b) = -1$ if $e = 2$. Let $\mathcal{H}_m$ be defined by

$$\begin{aligned}
\mathcal{H}_m &= \{\beta \in (\mathcal{O}_L/\mathfrak{p}_L^m)^* \mid \overline{\alpha} \notin k^*\}\,,\ \text{if } e = 1 \\
\mathcal{H}_m &= b\mathcal{O}_L^*/U_L^m,\ \text{if } e = 2
\end{aligned}$$

These sets are preserved by the "dot" action of $\mathfrak{A}^*$, and we have a bijection of $\mathfrak{A}^*$-sets $\mathfrak{A}^*/J^0 \tilde{\to} \mathcal{H}_m$ given by $g \mapsto g\cdot b$.

If $e = 1$ then $K/J^0 = \mathcal{H}_m$; whereas if $e = 2$, K/J^0 is a disjoint union of copies of $\mathcal{H}_m$ indexed by $K/\mathfrak{A}^*$, with the obvious action of K. It follows that $\dim\tau = \#\mathcal{H}_m = q^{n-1}(q-1)$ if $e = 1$ and $\dim\tau = [K:\mathfrak{A}^*]\#\mathcal{H}_m = q^{n-2}(q^2-1)$ if $e = 2$. We are now ready to compute traces.

If $g = u(\varpi^k)$ is parabolic, with $k < n$, the result follows from the discussion in [Cas73b]. Indeed, the restriction of $\tau(\rho_0)$ to $\left(\begin{smallmatrix} 1 & \mathcal{O}_F \\ & 1 \end{smallmatrix}\right)$ decomposes into a sum of the characters of $\mathcal{O}_F$ of conductor n if $e = 1$, and into a sum of the characters of conductor n and $n-1$ if $e = 2$.

If $g = \delta(x)$ is hyperbolic with $x \in \mathcal{O}_F^*$ not congruent to $1 \pmod{\mathfrak{p}}$, then g has no fixed points on K/J^0, so that $\operatorname{Tr} \tau(g) = 0$.

Finally let $g = \iota(\alpha)$ be elliptic, with $\alpha \in \mathcal{O}_{F'}^*$ having reduction mod $\mathfrak{p}_{F'}$ not in k^* if $e = 1$. First assume $e = 1$ and identify F' with L. Then α can play the role of β in the definition of $\mathcal{H}_m$, and then the action of g on $\mathcal{H}_m$ has fixed points exactly α and α^σ. Lemma 2.5.2 gives $\operatorname{Tr} \tau(g) = \operatorname{Tr} \eta(\alpha) + \operatorname{Tr} \eta(\alpha^\sigma) = (-1)^n(\varepsilon_0(\alpha) + \varepsilon_0(\alpha^\sigma))$. If on the other hand $e = 2$, g will not fix any cosets in $K/\mathfrak{A}^*$, let alone cosets in K/J^0, so that $\operatorname{Tr} \tau(g) = 0$. For completeness, we remark that if instead g is the image of an element $\alpha \in \mathcal{O}_L^*$ in the embedding $\mathcal{O}_L \hookrightarrow \mathfrak{A}$, then $\operatorname{Tr} \tau(g) = \varepsilon_0(\alpha) + \varepsilon_0(\alpha^\iota)$ for the nontrivial automorphism $\iota \in \operatorname{Gal}(L/F)$.

Chapter 3

Global arguments

3.1 Automorphic forms on the adele group

The theory of modular forms as functions on the upper half plane is in hindsight only a corner of the grander theory of automorphic representations of adelic groups. In this section we review the bridge between the classical and modern theories in a manner that suits our purposes for the following sections.

Let $N \geq 1$ and $k \geq 2$ be two integers. Recall that the space of cusp forms $S_k(\Gamma_1(N))$ comes equipped with an action of a Hecke algebra $\mathbf{T}$, another action by the diamond operators $\langle a \rangle$ for $a \in (\mathbf{Z}/N\mathbf{Z})^*$, and a positive definite Petersson inner product $(\ , \ )$ for which $(T_n f, g) = (f, \langle n \rangle T_n g)$ for all $f, g \in S_k(\Gamma_1(N))$ and for $(n, N) = 1$. Inside of $S_k(\Gamma_1(N))$ is a subspace $S_k(\Gamma_1(N))^{\mathrm{old}}$ which we define to be the sum of the images of the maps

$$\begin{array}{ccc} S_k(\Gamma_1(d)) & \to & S_k(\Gamma_1(N)) \\ f(z) & \mapsto & f(nz) \end{array}$$

for all proper divisors d of N and all divisors n of N/d.

Definition 3.1.1. The *new subspace* $S_k(\Gamma_1(N))^{\mathrm{new}}$ is the orthogonal complement of $S_k(\Gamma_1(N))^{\mathrm{old}}$ in $S_k(\Gamma_1(N))$ with respect to the Petersson inner product. A *newform* of weight k, level N and character ε is a cusp form $f \in S_k(\Gamma_1(N))^{\mathrm{new}}$ which is an eigenvector with respect to the Hecke operators, for which $\langle a \rangle f = \varepsilon(a) f$, $a \in (\mathbf{Z}/N\mathbf{Z})^*$, and for which the q-expansion has q-coefficient 1.

The following discussion is gathered from [Gel75], §5.

Let $\mathbf{A}$ be the adele ring of $\mathbf{Q}$. To a newform f of level N, weight $k \geq 2$ and character ε we associate a certain representation of $G_{\mathbf{A}} = \mathrm{GL}_2(\mathbf{A})$ as follows. Identify ε with a character of $\mathbf{Q}\backslash\mathbf{A}^* = \hat{\mathbf{Z}}^* \times \mathbf{R}_{\geq 0}$ by having it act trivially on $\mathbf{R}_{\geq 0}$. Define a subgroup $K_0(N) \subset \mathrm{GL}_2(\hat{\mathbf{Z}})$ as those matrices which are upper-triangular mod N, and extend ε to $K_0(N)$ by $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto \varepsilon(a)$. By strong approximation for SL_2,

$$\mathrm{GL}_2(\mathbf{A}) = \mathrm{GL}_2(\mathbf{Q})\, \mathrm{GL}_2^+(\mathbf{R}) K_0(N).$$

Define a function ϕ_f on $G_{\mathbf{A}}$ by

$$\phi_f(g) = \varepsilon(k)(ci+d)^{-k} f\left(\frac{ai+b}{ci+d}\right),$$

where $g = g_{\mathbf{Q}} g_\infty k$ for $g_{\mathbf{Q}} \in G_{\mathbf{Q}}$, $g_\infty = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{GL}_2^+(\mathbf{R})$ and $k \in K_0(N)$. Then ϕ_f lies in the space $R_0(\varepsilon)$ consisting of functions $\phi \in L^2(G_{\mathbf{Q}}\backslash G_{\mathbf{A}})$ satisfying $\phi(gz) = \varepsilon(z)\phi(g)$ for $z \in \mathbf{A}^*$ and the cuspidality condition

$$\int_{\mathbf{Q}\backslash\mathbf{A}} \phi(u(x)g)dx = 0, \text{ almost all } g.$$

Then $R_0(\varepsilon)$ admits an action of $G_{\mathbf{A}}$ by right translation.

Definition 3.1.2. The *automorphic representation* π_f attached to f is the span of the $G_{\mathbf{A}}$-translates of f, considered as a $G_{\mathbf{A}}$-submodule of $R_0(\varepsilon)$.

Theorem 3.1.3. *The π_f are irreducible and admit a complete tensor product decomposition $\pi = \bigotimes_v \pi_{f,v}$ with*

$$\begin{aligned} \pi_{f,p} &\in \mathcal{A}_2(\mathbf{Q}_p) \textit{ for every prime } p, \\ \pi_{f,\infty} &= \textit{the discrete series representation } \sigma_k := \sigma\left(|\,|^{\frac{k-1}{2}}, |\,|^{\frac{1-k}{2}} \mathrm{sgn}^k\right). \end{aligned}$$

Furthermore, the space $R_0(\varepsilon)$ is the discrete direct sum of irreducible representations π of $G_{\mathbf{A}}$, each occurring with multiplicity one; those components π whose archimedean component is σ_k are exactly the π_f as f runs through all newforms of weight k and character ε.

Sketch of proof. The claim that $R_0(\varepsilon)$ splits up as a discrete direct sum of representations of the form $\pi = \bigotimes_v \pi_v$ has a functional-analytic proof. Suppose $\pi = \otimes_v \pi_v$ is one of the constituents of $R_0(\varepsilon)$. For each prime p, let p^{c_p} be the conductor of π_p and choose a new vector w_p in the representation space of π_p. Also choose a lowest weight vector $w_\infty \in \sigma_k$. Then $f = \otimes_v w_v \in R_0(\varepsilon)$ is easily seen to lie in $S_k(\Gamma_1(N))$, where $N = \prod_p p^{c_p}$; with some more work one can show that f is a newform and that $\pi = \pi_f$ (see [Gel75], Lemma 5.16). $\square$

Let $S_k(\Gamma(N))$ be the space of modular forms of weight k for the principal congruence subgroup $\Gamma(N)$. This admits an action of $\Gamma(1)/\Gamma(N) = \mathrm{SL}_2(\mathbf{Z}/N\mathbf{Z})$.

Theorem 3.1.4. *Let λ be an irreducible complex representation of $\mathrm{GL}_2(\hat{\mathbf{Z}})$ which factors through $\mathrm{GL}_2(\mathbf{Z}/N\mathbf{Z})$. The multiplicity of $\lambda|_{\mathrm{SL}_2(\mathbf{Z}/N\mathbf{Z})}$ in $S_k(\Gamma(N))$ counts the number of newforms f of weight k satisfying*

$$\bigotimes_{p<\infty} \pi_{f,p}|_{\mathrm{GL}_2(\hat{\mathbf{Z}})} \supset \lambda.$$

Proof. Let N_λ be the number of newforms f satisfying the property required by the theorem. Let ε be the central character of λ. By Theorem 3.1.3 the space $R_0(\varepsilon)$ decomposes into a discrete direct sum of irreducible representations π of $\mathrm{GL}_2(\mathbf{A})$; then N_λ is the number of π for which $\pi_\infty = \sigma_k$ and for which $\lambda \subset \bigotimes_p \pi_{f,p}$. This means that

$$N_\lambda = \dim \mathrm{Hom}_{\mathrm{GL}_2(\hat{\mathbf{Z}}) \times \mathrm{GL}_2(\mathbf{R})} \left(\lambda \otimes \sigma_k, R_0(\varepsilon)\right). \tag{3.1}$$

But since λ factors through $\mathrm{GL}_2(\mathbf{Z}/N\mathbf{Z})$, this multiplicity is the same as the multiplicity of $\lambda \otimes \sigma_k$ in the subspace $R_0(\varepsilon)^{K(N)}$ of $R_0(\varepsilon)$ fixed by $K(N) = \ker(\mathrm{GL}_2(\hat{\mathbf{Z}}) \to \mathrm{GL}_2(\mathbf{Z}/N\mathbf{Z}))$.

We need to identify the adelic object $R_0(\varepsilon)^{K(N)}$ with an object that appears in the classical study of the representation theory of $\mathrm{SL}_2(\mathbf{R})$. Let $L = L_0(\Gamma(N)\backslash \mathrm{SL}_2(\mathbf{R}))$ be the space of square-integrable functions f on $\Gamma(N)\backslash \mathrm{SL}_2(\mathbf{R})$ satisfying the cuspidality condition

$$\int_{\mathbf{Z}\backslash\mathbf{R}} f(u(x)g)dx = 0, \text{ almost all } g.$$

Then L admits an action of $S_N \times \mathrm{SL}_2(\mathbf{R})$.

The following abbreviations will be useful: $S_N = \mathrm{SL}_2(\mathbf{Z}/N\mathbf{Z})$, $G_N = \mathrm{GL}_2(\mathbf{Z}/N\mathbf{Z})$. If M is a G_N-module, let M^ε denote the largest submodule on which the center of G_N acts as ε.

Lemma 3.1.5. *We have an isomorphism of $G_N \times \mathrm{SL}_2(\mathbf{R})$-modules*

$$\left[\operatorname{Ind}_{S_N}^{G_N} L\right]^\varepsilon \xrightarrow{\sim} R_0(\varepsilon)^{K(N)}|_{\mathrm{SL}_2(\mathbf{R})}.$$

Proof. Let $L = L_0^2(\Gamma(N)\backslash \mathrm{SL}_2(\mathbf{R}))$ with its action of $S_N \times \mathrm{SL}_2(\mathbf{R})$. The space on the left is the space of functions $F\colon G_N \to L$, $g \mapsto F_g$ satisfying $F_{zgh} = \varepsilon(z)gF_h$ for $z \in (\mathbf{Z}/N\mathbf{Z})^* \operatorname{Id}_2$, $g \in S_N$. Our map is $F \mapsto \phi_F \in R_0(\varepsilon)^{K(N)}$, where ϕ_F is defined as follows: for $g \in G_{\mathbf{A}}$, write $g = rg_{\mathbf{Q}}g_{\mathbf{R}}k$, for $r \in \mathbf{R}_{\geq 0}$, $g_{\mathbf{Q}} \in \mathrm{GL}_2(\mathbf{Q})$, $g_{\mathbf{R}} \in \mathrm{SL}_2(\mathbf{R})$ and $k \in K(N)$. Then $\phi_F(g) = F_k(g_{\mathbf{R}})$. One checks that ϕ_F is $K(N)$-invariant and independent of the decomposition $g = rg_{\mathbf{Q}}g_{\mathbf{R}}k$ chosen.

In the other direction, if $\phi \in R_0(\varepsilon)^{K(N)}$, define $F\colon G_N \to L$ as follows. If $g \in G_N$, let $\hat{g} \in \mathrm{GL}_2(\hat{\mathbf{Z}})$ be a lift of g and let $F_g(h) = \phi(\hat{g}h)$ for $h \in \mathrm{SL}_2(\mathbf{R})$. $\square$

We resume the proof of Theorem 3.2.4. The discrete series representation σ_k is induced from a discrete series representation σ_k^+ of $\mathrm{SL}_2(\mathbf{R})$:

$$\sigma_k = \operatorname{Ind}_{\mathrm{SL}_2(\mathbf{R})}^{\mathrm{GL}_2(\mathbf{R})} \sigma_k^+. \tag{3.2}$$

The space L decomposes as a direct sum of irreducible representations of $\mathrm{SL}_2(\mathbf{R})$, and we have

$$\operatorname{Hom}_{\mathrm{SL}_2(\mathbf{R})}\left(\sigma_k^+, L\right) \cong S_k(\Gamma(N)). \tag{3.3}$$

(See [Gel75], Theorem 2.10: the image of a lowest-weight vector from σ_k^+ in L is forced to be the function on $\mathrm{SL}_2(\mathbf{R})$ arising from a classical modular form for $\Gamma(N)$.) We have

$$\begin{aligned}
N_\lambda &= \dim \mathrm{Hom}_{G_N \times \mathrm{GL}_2(\mathbf{R})} \left(\lambda \otimes \sigma_k, R_0(\varepsilon)^{K(N)}\right) && \text{Eq. 3.1} \\
&= \dim \mathrm{Hom}_{G_N \times \mathrm{GL}_2(\mathbf{R})} \left(\lambda \otimes \mathrm{Ind}_{\mathrm{SL}_2(\mathbf{R})}^{\mathrm{GL}_2(\mathbf{R})} \sigma_k^+, R_0(\varepsilon)^{K(N)}\right) && \text{Eq. 3.2} \\
&= \dim \mathrm{Hom}_{G_N \times \mathrm{SL}_2(\mathbf{R})} \left(\lambda \otimes \sigma_k^+, R_0(\varepsilon)^{K(N)}|_{\mathrm{SL}_2(\mathbf{R})}\right) && \text{Frob. reciprocity} \\
&= \dim \mathrm{Hom}_{G_N \times \mathrm{SL}_2(\mathbf{R})} \left(\lambda \otimes \sigma_k^+, \left[\mathrm{Ind}_{S_N}^{G_N} L\right]^\varepsilon\right) && \text{Lemma 3.1.5} \\
&= \dim \mathrm{Hom}_{G_N \times \mathrm{SL}_2(\mathbf{R})} \left(\lambda \otimes \sigma_k^+, \mathrm{Ind}_{S_N}^{G_N} L\right),
\end{aligned}$$

because the central character of λ is already ε. By Frobenius reciprocity,

$$N_\lambda = \dim \mathrm{Hom}_{S_N \times \mathrm{SL}_2(\mathbf{R})} \left(\lambda \otimes \sigma_k^+, L\right),$$

and because the functor $- \otimes M$ is adjoint to $\mathrm{Hom}(-, M)$ in the category of $\mathrm{SL}_2(\mathbf{R})$-modules, this becomes

$$\begin{aligned}
N_\lambda &= \dim \mathrm{Hom}_{S_N} \left(\lambda, \mathrm{Hom}_{SL_2(\mathbf{R})}(\sigma^+, L)\right) \\
&= \langle \lambda, S_k(\Gamma(N)) \rangle_{S_N}
\end{aligned}$$

by Equation 3.3. □

In the next section we will see how to apply Theorem 3.1.4 to count newforms whose Galois representations have prescribed local properties.

3.2 The theorem of Deligne-Carayol

Let $N \geq 1$ and $k \geq 2$ be two integers, and let ℓ be prime. Take f to be a cusp form with coefficients in $\overline{\mathbf{Q}}_\ell$ of weight k and character ε for $\Gamma_1(N)$ which is a newform for the action of the Hecke operators. Let $a_n \in \overline{\mathbf{Q}}_\ell$ be the eigenvalue of the operator T_n on f. One associates to f a Galois representation

$$\rho_f \colon \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathrm{GL}_2(\overline{\mathbf{Q}}_\ell)$$

unramified outside of ℓN satisfying

$$\begin{aligned}\mathrm{Tr}(\rho_f(\mathrm{Frob}_p)) &= a_p \\ \det(\rho_f(\mathrm{Frob}_p)) &= p^{k-1}\varepsilon(p)\end{aligned}$$

whenever p is a prime not dividing $N\ell$. These conditions determine the restriction $\rho_{f,p}$ of ρ_f to $G_{\mathbf{Q}_p}$ when $p \nmid N\ell$. When $p|N\ell$, however, the shape of $\rho_{f,p}$ cannot be determined directly from the Hecke eigenvalues of f.

To at least treat the case when $p|N$ but $p \neq \ell$, we pass to the irreducible admissible $\overline{\mathbf{Q}}_\ell$-valued representation $\otimes_p \pi_p$ of $\mathrm{GL}_2(\mathbf{A}^{\mathrm{f}})$, with each $\pi_p \in \mathcal{A}_2(\mathbf{Q}_p)_{\overline{\mathbf{Q}}_\ell}$. When $p \neq \ell$, $\rho_{f,p}^{\mathrm{WD}} \in \mathcal{G}_2(\mathbf{Q}_p)$ be the WD representation associated by Theorem 2.1.4 to the ℓ-adic representation $\rho_{f,p}|_{W_{\mathbf{Q}_p}}$.

For $s > 0$, let ω^s be the character $g \mapsto |\det g|^s$ of $\mathrm{GL}_2(\mathbf{Q}_p)$.

Theorem 3.2.1 ([Car83], Théorème A)**.** *For $p \neq \ell$, we have*

$$\pi_{f,p} = \omega^{\frac{1}{2}} \operatorname{rec} \rho_{f,p}^{WD}.$$

Remark 3.2.2. In particular

$$(\pi_{f,p})|_{\mathrm{GL}_2(\mathbf{Z}_p)} \cong \operatorname{rec}(\rho_{f,p}^{\mathrm{WD}})|_{\mathrm{GL}_2(\mathbf{Z}_p)}.$$

Remark 3.2.3. When $p = \ell$, the theorem is not at all true, even for $p \nmid N$: the p-adic Galois representation contains certain "period" data in addition to the structural data contained in $\pi_{f,p}$.

Theorem 3.2.1 has a certain "independence of ℓ" consequence. Suppose now that f begins life as a form with complex coefficients, and let p be a prime. We may simply define the WD representation $\rho_{f,p}^{\mathrm{WD}}$ by

$$\rho_{f,p}^{\mathrm{WD}} := \operatorname{rec}^{-1}\left(\omega^{-\frac{1}{2}}\pi_{f,p}\right).$$

Let $\ell \neq p$ be prime and let $\iota\colon \mathbf{C} \xrightarrow{\sim} \overline{\mathbf{Q}}_\ell$ be an isomorphism. Then composing $\rho_{f,p}^{\mathrm{WD}}$ with ι

gives an E-valued WD representation $\iota\rho_{f,p}^{\mathrm{WD}}$, which equals

$$\begin{aligned}
\iota\rho_{f,p}^{\mathrm{WD}} &= \iota\,\mathrm{rec}^{-1}\left(\omega^{-\frac{1}{2}}\pi_{f,p}\right) && \text{by definition} \\
&= \mathrm{rec}^{-1}\left(\omega^{-\frac{1}{2}}\pi_{\iota f,p}\right) && \text{because } \iota \text{ and rec commute} \\
&= \rho_{\iota(f),p} && \text{by Theorem 3.2.1,}
\end{aligned}$$

this last being the WD representation associated as in the beginning of the section to the ℓ-adic newform f. Consequently the ℓ-adic representations $\rho_{\iota(f),p}$ carry the same information regardless of the isomorphism ι and even the prime ℓ, so long as $\ell \neq p$.

For each prime p, let $\rho_p \in \mathcal{G}_2^I(\mathbf{Q}_p)$ be an inertial WD representation of $\mathbf{Q}_p$, with $\rho_p = 1 \oplus 1$ for all but finitely many p. We wish to determine how many newforms f have $\rho_{f,p}^{\mathrm{WD}}|_{I_{\mathbf{Q}_p}} \cong \rho_p$ for all p. For each p, let $\lambda_p = \tau\left(\rho_p^{\mathrm{WD}}\right)$ be the inertial type. This is a family of finite-dimensional representations of the group $\mathrm{GL}_2(\mathbf{Z}_p)$, trivial for almost all p, so it makes sense to define the *global inertial type* $\lambda = \bigotimes_p \lambda_p$ as a representation of the group $\mathrm{GL}_2(\hat{\mathbf{Z}})$. Let N be the products of the levels of the λ_p, so that λ factors through $\mathrm{GL}_2(\mathbf{Z}/N\mathbf{Z})$ and through no smaller such group.

The next theorem shows that counting the number of f whose Galois representations restrict to inertia as ρ_p amounts to finding the multiplicity of λ in the space of cusp forms $S_k(\Gamma(N))$.

Theorem 3.2.4. *Assume that each $\rho_p \in \mathcal{G}_2^I(\mathbf{Q}_p)^{ss}$. Define λ and N as above. Then*

$$\#\left\{\text{wt. } k \text{ newforms } f\colon \rho_{f,p}^{WD}|_{I_{\mathbf{Q}_p}} \cong \rho_p \text{ for all } p\right\} = \langle\lambda, S_k(\Gamma(N))\rangle_{\mathrm{SL}_2(\mathbf{Z}/N\mathbf{Z})}. \tag{3.4}$$

Proof. For each prime p, Theorem 2.4.1 asserts that $\rho_{f,p}^{\mathrm{WD}}|_{I_{\mathbf{Q}_p}} \cong \rho_p$ if and only if λ_p is contained in $\mathrm{rec}(\rho_{f,p}^{\mathrm{WD}})|_{\mathrm{GL}_2(\mathbf{Z}_p)}$. (Since ρ_p is semi-simple, we avoid the "unless" clause of that theorem.) But by Remark 3.2.2, $\mathrm{rec}(\rho_{f,p}^{\mathrm{WD}})|_{\mathrm{GL}_2(\mathbf{Z}_p)} \cong \pi_{f,p}|_{\mathrm{GL}_2(\mathbf{Z}_p)}$. Therefore $\mathrm{rec}\, {f_p}^{\mathrm{WD}}|_{I_{\mathbf{Q}_p}} \cong \rho_p$ for all p if and only if $\bigotimes_{p<\infty} \pi_{f,p}|_{\mathrm{GL}_2(\hat{\mathbf{Z}})}$ contains $\lambda = \bigotimes_{f,p} \lambda_p$. The number of newforms f of weight k for which this is true is $\langle\lambda, S_k(\Gamma(N))\rangle_{\mathrm{SL}_2(\mathbf{Z}/N\mathbf{Z})}$ by Theorem 3.1.4. ∎

Theorem 3.2.4 reduces the proof of our main Theorem 1.0.1 to the problem of determining $S_k(\Gamma(N))$ as a $\mathrm{SL}_2(\mathbf{Z}/N\mathbf{Z})$-module. We do this in the next section, at least for k even.

3.3 Generalities on Galois covers of curves

Let $N > 1$. The space $S_k(\Gamma(N))$ of cusp forms for the principal congruence subgroup $\Gamma(N)$ admits an action of the quotient $\mathrm{SL}_2(\mathbf{Z}/N\mathbf{Z})/\{\pm 1\}$ when k is even. Using the equivariant Riemann-Roch formula, we will compute this action explicitly. We will see that an irreducible representation π of $\mathrm{SL}_2(\mathbf{Z}/N\mathbf{Z})/\{\pm 1\}$ is contained in $S_k(\Gamma(N))$ with multiplicity equal to $k/12 \times \dim \pi$ to within a constant depending on π and on k (mod 12).

Let X and Y be connected projective curves over an algebraically closed field K, and let $p\colon X \to Y$ be a Galois cover with group G. The action of G on the cohomology of X with coefficients in a line bundle seems to be well studied in the literature; in this section, we borrow the notation and results gathered in [Bor05] as they are in a form well-suited to our applications.

Definition 3.3.1. A *G-equivariant line bundle* is a pair $(\mathcal{L}, \{\phi_g\})$ consisting of a line bundle $\mathcal{L}$ on X together with isomorphisms

$$\phi_g\colon g^*\mathcal{L} \to \mathcal{L}$$

for each $g \in G$, satisfying the cocycle condition

$$\phi_{gh} = h^*(\phi_g) \circ \phi_h.$$

We can recast the notion in terms of divisor classes as follows. The group $\operatorname{Div} X$ admits an action of G. If $\mathcal{L}$ is a G-equivariant line bundle, let s be a nonzero meromorphic section of $\mathcal{L}$. Then for each $g \in G$, the quotient $\phi_g(s(gx))/s(x)$ is a meromorphic function on X. By the cocycle condition, $g \mapsto \phi_g(s(gx))/s(x)$ is a cocycle

ξ in $H^1(G, K(X)^*)$, which vanishes by Hilbert's Theorem 90. Therefore there is a meromorphic function $h \in K(X)^*$ which splits ξ, meaning that

$$\frac{\phi_g(s(gx))}{s(x)} = \frac{h(g(x))}{h(x)}.$$

Thus $x \mapsto s(x)/h(x)$ is another meromorphic section whose divisor D is visibly G-invariant. The choice of another section s or another splitting of ξ replaces D by the divisor of an element in $K(X)^G = K(Y)$. The line bundle $\mathcal{L}$ can be recovered from D as the line bundle whose fibers are $\mathcal{L}(U) = \{f \in k(X)^* |\operatorname{div} f \geq D \text{ on } U\}$, with transition maps ϕ_g given by $\phi_g(f) = f \circ g$. Therefore the group of G-equivariant line bundles up to linear equivalence can be identified with the quotient $(\operatorname{Div} X)^G / \operatorname{Div} K(Y)^*$.

For each $k \geq 0$, the line bundle of differentials Ω_X^k is a G-equivariant line bundle. In the case of $k = 0$, $\Omega_X^0 = \mathcal{O}_X$ has associated divisor class 0.

In this context, the sheaf cohomology groups $H^i(X, \mathcal{L})$ admit an action of G, $i = 0, 1, 2$. Let $R(G)$ be the representation ring of G; that is, the K-vector space with basis the irreducible characters of G. Then we can define the *equivariant Euler characteristic*

$$\chi_{\mathrm{eq}}(\mathcal{L}) = \left[H^0(X, \mathcal{L})\right] - \left[H^1(X, \mathcal{L})\right].$$

The *equivariant degree* of $\mathcal{L}$ is an element of $R(G)$ defined as follows: Let D be the G-invariant divisor attached to $\mathcal{L}$. First assume D is a multiple of an orbit $\sum_{\sigma \in G/G_P} \sigma P$ with multiplicity μ_P, where G_P is the decomposition group at $P \in X$. The group G_P acts on the tangent space $T_P X$ through a character ψ_P with values in K^*. Then define

$$\deg_{\mathrm{eq}} D = \begin{cases} \sum_{c=1}^{\mu_P} \operatorname{Ind}_{G_P}^G \psi_P^{-c}, & \mu_P > 0, \\ 0, & \mu_P = 0, \\ -\sum_{c=0}^{-(\mu_P+1)} \operatorname{Ind}_{G_P}^G \psi_P^{c}, & \mu_P < 0. \end{cases} \tag{3.5}$$

Extend $\deg_{\mathrm{eq}}$ to general D in such a way that $\deg_{\mathrm{eq}}(D + D') = \deg_{\mathrm{eq}} D + \deg_{\mathrm{eq}} D'$ when D and D' have disjoint support.

The association $\deg_{\mathrm{eq}} : (\operatorname{Div} X)^G \to R(G)$ is *not* a linear map. However, the composition of $\deg_{\mathrm{eq}}$ with the dimension map $R(G) \to \mathbf{Z}$ is the usual degree map on divisors. Further, the equivariant degree vanishes on the group Div $f^*(K(Y)^*)$ of principal divisors, so it really is well-defined on G-equivariant line bundles.

We are now ready to state the equivariant Riemann-Roch formula:

Theorem 3.3.2. *In the ring $R(G)$:*

$$\chi_{eq}(\mathcal{L}) = \chi_{eq}(\mathcal{O}_X) + \deg_{eq}(\mathcal{L}).$$

3.4 Galois action on cusp forms

Here we carry out a calculation essentially equivalent to the calculation of the dimension of the space of cusp forms given in [Shi71] but using the results gathered in the previous section to keep track of the action of SL_2.

The map of modular curves $f\colon X(N) \to X(1) \cong \mathbf{P}^1$ plays the role of $X \to Y$ in the previous section, with $G = \mathrm{SL}_2(\mathbf{Z}/N\mathbf{Z})/\{\pm 1\}$. We assume $N > 1$, so that $\Gamma(N)$ has no elliptic elements. We work over the field of complex numbers.

Recall that the map f is ramified over the points $j = 1728, 0, \infty$ with inertia groups (conjugate to) $G_{1728}, G_0, G_\infty \subset G$ having generators $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ -1 & -1 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ of order $2, 3, N$, respectively. For each $i \in \{1728, 0, \infty\}$, the character ψ_{G_i} is the one taking the generator of G_i to $\exp(2\pi i/\#G_i)$. Let $\Psi_{i,j} = \mathrm{Ind}_{G_i}^G \psi_{G_i}^{-j}$ for j taken modulo $\#G$. Note that

$$\begin{aligned} \sum_{j=0}^{\#G_i - 1} \Psi_{i,j} &= \mathrm{Ind}_{G_i}^G \sum_{j=0}^{\#G_i - 1} \psi^{-j} && (3.6) \\ &= \mathrm{Ind}_{G_i}^G \mathrm{Ind}_1^{G_i} 1 && (3.7) \\ &= \mathbf{C}G, && (3.8) \end{aligned}$$

the regular representation of G.

Fix a point $S \in X(1) \setminus \{1728, 0, \infty\}$. Since $X(1)$ is a rational curve, there exists a nonzero meromorphic differential $\omega_1 \in H^0(X(1), \Omega_{X(1)})$ with divisor $-2S$. Then the pullback $\omega = f^*(\omega_1)$ has divisor

$$\operatorname{div}\omega = -2\sum_{g\in G} gR + \sum_{g\in G/G_{1728}} gP_{1728} + 2\sum_{g\in G/G_0} gP_0 + (N-1)\sum_{g\in G/G_\infty} gP_\infty, \quad (3.9)$$

where R is a point in the preimage of S and P_i is fixed by G_i for $i \in \{1728, 0, \infty\}$.

Our goal is to calculate the action of G on $S_k(\Gamma(N))$.

When $k = 2$, $S_2(\Gamma(N)) = H^0(X(N), \Omega^1_{X(N)})$. By Serre duality, this is isomorphic to the linear dual of $H^1(X(N), \mathcal{O}_{X(N)})$. In fact, since Serre duality is natural, the isomorphism between the two spaces is G-equivariant. The Petersson inner product on $S_2(\Gamma(N))$ is also G-equivariant, meaning that $S_2(\Gamma(N))$ admits a G-equivariant isomorphism to its linear dual. Therefore $[S_2(\Gamma(N))] = [H^1(X(N), \mathcal{O}_{X(N)})]$ in $R(G)$. Similarly,

$$[H^1(X(N), \Omega^1_{X(N)}] = [H^0(X(N), \mathcal{O}_{X(N)}] = [1].$$

Therefore the trivial bundle on $X(N)$ has equivariant Euler characteristic

$$\begin{aligned}
\chi_{\text{eq}}(\mathcal{O}_{X(N)}) &= [H^0(X(N), \mathcal{O}_{X(N)})] - [H^1(X(N), \mathcal{O}_{X(N)}] \\
&= [1] - [S_2(\Gamma(N))]
\end{aligned}$$

and the canonical divisor $\Omega^1_{X(N)}$ has equivariant Euler characteristic

$$\begin{aligned}
\chi_{\text{eq}}(\Omega^1_{X(N)}) &= [H^0(X(N), \Omega^1_{X(N)}] - [H^1(X(N), \Omega^1_{X(N)}] \\
&= [S_2(\Gamma(N))] - [1] \\
&= -\chi_{\text{eq}}(\Omega^1_{X(N)})
\end{aligned}$$

On the other hand, the Riemann-Roch formula gives $\chi(\Omega^1_{X(N)}) = \chi(\mathcal{O}_{X(N)}) + \deg_{\text{eq}}(\Omega^1_{X(N)})$, so we find that

$$\deg_{\text{eq}}\left(\Omega^1_{X(N)}\right) = 2\chi(\Omega^1_{X(N)}) = 2\left([S_2(\Gamma(N))] - [1]\right)$$

The equivariant degree of $\Omega^1_{X(N)}$ is also the degree of the divisor in Eq. (3.9), which is

$$\begin{aligned}
\deg_{\text{eq}}\left(\Omega^1_{X(N)}\right) &= \deg_{\text{eq}}\operatorname{div}\omega \\
&= -2\mathbf{C}G + \Psi_{1728,1} + \Psi_{0,1} + \Psi_{0,2} + \sum_{i=1}^{N-1}\Psi_{\infty,i} \\
&= \mathbf{C}G - \Psi_{1728,0} - \Psi_{0,0} - \Psi_{\infty,0}.
\end{aligned}$$

Solving for $[S_2(\Gamma(N))]$, we find

Proposition 3.4.1.

$$[S_2(\Gamma(N))] = \frac{1}{2}\left(\mathbf{C}G - \Psi_{1728,0} - \Psi_{0,0} - \Psi_{\infty,0}\right) + [1]$$

Now suppose $k > 2$.

Definition 3.4.2. An *automorphic form* of weight w for $\Gamma(N)$ is a meromorphic function F on the upper half-plane with $F(\gamma z) = (cz+d)^w F(z)$ for each $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma(N)$. The space of automorphic functions is denoted $A_w(N)$.

Then $A_0(N) = \mathbf{C}(X(N))$ and generally $A_w(N)$ is a one-dimensional vector space over $\mathbf{C}(X(N))$, because the quotient of any two nonzero automorphic forms of weight w has weight 0. Recall that k is even. We choose a basis vector F for $A_k(N)$ by setting $\omega^{k/2} = F(z)dz^{k/2}$. Then the divisor of F is

$$\operatorname{div} F = (k/2)\operatorname{div}\omega + (k/2)\sum_{g\in G/G_\infty} gP_\infty. \tag{3.10}$$

(Indeed, the coordinate in the neighborhood of P_∞ is $q = \exp(2\pi iz/N)$, whereby $dz = Ndq/2\pi iq$ has a simple pole at P_∞ and therefore at all of its G-translates.)

$$\begin{aligned} S_k(\Gamma(N)) &= \left\{F' \in A_k(N) | \operatorname{div} F' \geq \sum_{g\in G/G_\infty} gP_\infty\right\} \\ &\tilde{\to} \left\{f \in \mathbf{C}(X(N)) | \operatorname{div} f \geq -\operatorname{div} F + \sum_{g\in G/G_\infty} gP_\infty\right\} \end{aligned}$$

via the map $F' \mapsto f = F'/F$. Putting together Equations 3.9 and 3.10 we find that $S_k(\Gamma(N)) \cong H^0(X(N), \mathcal{L}_k)$, where $\mathcal{L}_k$ is the line bundle with G-invariant divisor

$$-k\sum_{g\in G} gR + \frac{k}{2}\sum_{g\in G/G_{1728}} gP_{1728} + k\sum_{g\in G/G_0} gP_0 + \left(\frac{k}{2}N - 1\right)\sum_{g\in G/G_\infty} gP_\infty. \tag{3.11}$$

For $k > 2$, the line bundle $\mathcal{L}_k$ has vanishing H^1, so that the Riemann-Roch formula gives

$$[S_k(\Gamma(N))] = \deg_{\mathrm{eq}}\mathcal{L}_k + \chi_{\mathrm{eq}}(\mathcal{O}_{X(N)}),\ k > 2. \tag{3.12}$$

Applying the definition of degree in Equation 3.5 to Equation 3.11 and substituting into Equation 3.12 gives

Theorem 3.4.3. *Let $k > 2$ be even. In the ring $R(G)$, we have*

$$[S_k(\Gamma(N))] = \left\lfloor \frac{k}{12} \right\rfloor \mathbf{C}G + \varepsilon_k,$$

where

$$\varepsilon_k = \begin{cases} -\Psi_{1728,1} + \Psi_{0,0}, & k \equiv 0 \pmod{12} \\ 0, & k \equiv 2 \pmod{12} \\ \Psi_{1728,0} - \Psi_{0,2}, & k \equiv 4 \pmod{12} \\ \Psi_{0,0}, & k \equiv 6 \pmod{12} \\ \Psi_{1728,0}, & k \equiv 8 \pmod{12} \\ \Psi_{0,0} + \Psi_{0,1}, & k \equiv 10 \pmod{12} \end{cases}$$

This is a finer form of Theorem 1.0.3 from the introduction.

It follows immediately that ε_4 is effective, and we find

Corollary 3.4.4. *For $k \geq 14$ even, $S_k(\Gamma(N))$ contains every irreducible representation of G at least once.*

A compact way of summarizing Proposition 3.4.1 and Theorem 3.4.3 is as follows:

Corollary 3.4.5. *Let $k \geq 2$ be even. Then $[S_k(\Gamma(N))] \in R(G)$ is a sum of terms of the form $\mathbf{C}G$, $\Psi_{1728,j}$, $\Psi_{0,j}$ and $\Psi_{\infty,j}$, say with respective multiplicities c, c_{1728}, c_0, and c_∞ (collecting together terms over all j). The multiplicities satisfy*

$$c + \frac{1}{2}c_{1728} + \frac{1}{3}c_0 = \begin{cases} \frac{1}{12}, & k = 2 \\ \frac{k-2}{12}, & k > 2. \end{cases}$$

3.5 Multiplicity of inertial types in the space of cusp forms

For each prime p, let $\rho_p \in \mathcal{G}_2^I(\mathbf{Q}_p)$ be a WD inertial representation of $\mathbf{Q}_p$. Assume that $\rho_p = 1 \oplus 1$ for almost every p, including 2 and 3. Let $\lambda_p = \tau(\rho_p)$ be the inertial type and let $\lambda = \bigotimes_p \lambda_p$ be the global inertial type. Let $N = \prod_p p^{n_p}$ be the level of λ, so that λ factors through $\mathrm{GL}_2(\mathbf{Z}/N\mathbf{Z})$ and no smaller such group. Assume that $\lambda(-\mathrm{Id}_2)$ is the identity, or equivalently that $\prod_p \det \rho_p(-1) = 1$.

Let $k \geq 2$ be even. The goal of this section is to determine the multiplicity of λ in $S_k(\Gamma(N))$ as a module for the group $G = \mathrm{GL}_2(\mathbf{Z}/N\mathbf{Z})/\{\pm 1\}$, with a view towards applying Theorem 3.2.4.

Lemma 3.5.1. *For $i \in \{1728, 0, \infty\}$ and $j \in \mathbf{Z}/\#G_i\mathbf{Z}$, let $\mu_{i,j}(\lambda)$ be the multiplicity of λ in $\Psi_{i,j}$ as a G-module. Then*

$$\begin{aligned} \left|\mu_{1728,j}(\lambda) - \frac{\dim \lambda}{2}\right| &\leq \frac{1}{2} \times 2^{\nu(N)} \\ \left|\mu_{0,j}(\lambda) - \frac{\dim \lambda}{3}\right| &\leq \frac{2}{3} \times 2^{\nu(N)} \\ \mu_{\infty,j} &\leq 2^{\nu(N)} \end{aligned}$$

Proof. First note that the elements of each G_i are each elliptic or hyperbolic or parabolic when considered as elements of $\mathrm{GL}_2(\mathbf{Z}_p)$, for all $p \neq 2, 3$. If $i \in \{1728, 0\}$ and $g \in G_i$ is nontrivial then the inequality in Equation 2.3 gives $|\lambda(g)| \leq 2$. Therefore in those cases

$$|\mathrm{Tr}\,\lambda(g)| = \prod_p |\mathrm{Tr}\,\lambda_p(g)| \leq 2^{\nu(N)}. \tag{3.13}$$

The multiplicity of λ in $\Psi_{i,j}$ is

$$\begin{aligned} \mu_{i,j}(\lambda) &= \left\langle \lambda, \mathrm{Ind}_{G_i}^G \psi^{-j} \right\rangle_G & (3.14) \\ &= \left\langle \lambda|_{G_i}, \psi^{-j} \right\rangle_{G_i} & (3.15) \\ &= \frac{1}{\#G_i} \sum_{r=0}^{\#G_i - 1} \mathrm{Tr}\,\lambda(g^r) e^{2\pi\sqrt{-1}j/\#G_i}, & (3.16) \end{aligned}$$

where now g is the generator of G_i. In the cases $i = 0, 1728$, we bring the $r = 0$ term to the left side:

$$\begin{aligned} \left| \mu_{i,j}(\lambda) - \frac{1}{\#G_i} \dim \lambda \right| &\leq \frac{1}{\#G_i} \sum_{j=1}^{\#G_i - 1} |\mathrm{Tr}\, \lambda_p(g^r)| \\ &\leq \left(1 - \frac{1}{\#G_i}\right) 2^{\nu(N)} \end{aligned}$$

When $i = \infty$, the sum in (3.16) factors as

$$\mu_{\infty,j} = \prod_{p|N} \frac{1}{p^{n_p}} \sum_{r=0}^{p^{n_p}-1} \mathrm{Tr}\, \lambda_p(u(r)) e^{2\pi\sqrt{-1}rj/p^{n_p}},$$

and referring to the "parabolic" row of Figure 2.1 it is easily checked that each factor in this product has maximum absolute value at most 2.

□

We are now ready to prove the main theorem of this chapter.

Theorem 3.5.2. *For all but finitely many pairs $(\{\rho_p\}, k)$ with $\rho_p \in \mathcal{G}_2^I(\mathbf{Q}_p)^{ss}$ minimal and trivial for almost every p, there exists a newform f for which $\rho_f^{WD} = \rho$.*

Proof. Let $\lambda = \bigotimes_p \tau(\rho_p)$ have level N. By Theorem 3.2.4, the number of such newforms is $\langle \lambda, S_k(\Gamma(N)) \rangle$. By Corollary 3.4.4, we may assume that $k < 14$. It will be enough now to show that $\langle \lambda, S_k(\Gamma(N)) \rangle > 0$ for N large enough, since there are only finitely many λ of level N.

By Corollary 3.4.5, $[S_k(\Gamma(N))]$ is a sum of terms of the form $\mathbf{C}G$, $\Psi_{1729,j}$, $\Psi_{0,j}$ and $\Psi_{\infty,j}$ with multiplicities c, c_{1728}, c_0 and c_∞. By Theorem 3.4.3 we may assume $c_\infty = 0$ if $k > 2$. Then for $i \in \{1728, 0\}$ we have by Lemma 3.5.1

$$\mu_{i,j}(\lambda) = \frac{1}{\#G_i} \dim \lambda + O\left(2^{\nu(N)}\right)$$

and $\mu_{\infty,j}(\lambda) = O\left(2^{\nu(N)}\right)$, so that

$$\begin{aligned} \langle \lambda, S_k(\Gamma(N)) \rangle &= \left(c + \frac{1}{2} c_{1728} + \frac{1}{3} c_0\right) \dim \lambda + O\left(2^{\nu(N)}\right) \\ &= t \dim \lambda + O\left(2^{\nu(N)}\right), \end{aligned}$$

where $t = 1/12$ if $k = 2$ and $(k-2)/12$ if $k > 2$. Referring to Theorem 2.5.1 we have the lower bound

$$\dim \lambda \geq \phi(N) \geq 2^{v(N)} \prod_{p|N} \frac{p^{e_p} - p^{e_p - 1}}{2},$$

where $N = \prod_{p|N} p^{e_p}$. Thus $\dim \lambda / 2^{\nu(N)} \to \infty$ as $N \to \infty$ and $\langle \lambda, S_k(\Gamma(N)) \rangle > 0$ for $N \gg 0$. $\square$

3.6 Abelian varieties with everywhere good reduction

In this section we use the methods of the previous section to find a sufficient condition on a number field F for there to exist a modular abelian variety $A/\mathbf{Q}$ for which A_F admits a model with good reduction at all places.

We start with some generalities on modular abelian varieties. Let f be a newform of weight 2 for $\Gamma_1(N)$ whose Hecke eigenvalues $a_n \in \mathbf{C}$ generate a number field L. Then there exists an abelian variety quotient A_f of $J_1(N) = \operatorname{Jac} X_1(N)$ of dimension $[L : \mathbf{Q}]$ with an action $L \hookrightarrow \operatorname{End} A_f \otimes \mathbf{Q}$ for which the Hecke operator T_n on $J_1(N)$ induces a_n on A_f. Let $\ell \nmid N$ be prime and consider the contravariant Tate module

$$V_\ell(A_f) = \operatorname{Hom}_{\mathbf{Z}_\ell}\left(\varinjlim_n A[\ell^n], \mathbf{Q}_\ell \right).$$

Then $V_\ell(A_f)$ is a free $E \otimes \mathbf{Q}_\ell$-module of rank 2. Choosing a place λ of E above ℓ we obtain a Galois representation

$$\rho^\lambda_{A_f} \colon \operatorname{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \operatorname{GL}_2(E_\lambda)$$

agreeing with the representation ρ_f considered in more generality in Section 3.2. Let F be a number field, and let $\mathfrak{p}|p$ be a prime of F with $p \neq \ell$; let $G_\mathfrak{p} \subset \operatorname{Gal}(\overline{\mathbf{Q}}/F) \subset \operatorname{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ be a decomposition group at $\mathfrak{p}$, and let $I_\mathfrak{p} \subset G_\mathfrak{p}$ be the inertia group. By the criterion of Néron-Ogg-Shafarevich:

$$\begin{aligned} A_f \text{ has good reduction at } \mathfrak{p} \quad &\Longleftrightarrow \quad \rho^\lambda_{A_f}|_{G_\mathfrak{p}} \text{ is unramified} \\ &\Longleftrightarrow \quad \rho^{\mathrm{WD}}_{f,p}|_{W_{F_\mathfrak{p}}} \text{ is semi-simple and unramified.} \end{aligned}$$

As usual we have transferred the discussion to complex WD representations to free ourselves of the auxiliary prime λ.

Theorem 3.6.1. *Assume $F/\mathbf{Q}$ is Galois, and that there is a ramified prime $\mathfrak{p}|p$ of F for which one of the following holds:*

1. *$p \geq 29$, $p \equiv 1 \pmod 4$, and $F_\mathfrak{p}/\mathbf{Q}_p$ is ramified quadratic,*
2. *$p \geq 23$ and $F_\mathfrak{p}/\mathbf{Q}_p$ is cyclic with ramification degree at least 3, or*
3. *$p \geq 17$, $p \equiv 1 \pmod 4$ and $\mathrm{Gal}(F_\mathfrak{p}/\mathbf{Q}_p)$ is a dihedral group of order at least 6.*

Then there exists a modular abelian variety $A/\mathbf{Q}$ for which A_F has everywhere good reduction.

Remark 3.6.2. The theorem applies to any subfield $F \neq \mathbf{Q}$ of $\mathbf{Q}(\zeta_p)$, $p \geq 29$, so long as $F \neq \mathbf{Q}(\sqrt{-p})$.

Proof. . Under the hypotheses of the theorem, we will construct a WD representation $\rho \in \mathcal{G}_2(\mathbf{Q}_p)$ for which

1. $\rho|_{G_\mathfrak{p}}$ is unramified,
2. $\det \rho(-1) = 1$, and
3. The inertial type $\tau\left(\rho|_{I_{\mathbf{Q}_p}}\right)$ appears in the appropriate space of weight 2 cusp forms.

Then by Theorem 3.2.4, there will exist a newform f of weight 2 for which $\rho_{f,q}^{\mathrm{WD}}$ is unramified for all primes q except for $q = p$, and $\rho_{f,p}^{\mathrm{WD}}$ will be unramified upon restriction to $W_{F_\mathfrak{p}}$. Then by the criterion of Néron-Ogg-Shafarevich, A_f will have everywhere good reduction over F.

In case (1), let ε be a ramified quadratic character of $W_{\mathbf{Q}_p}$. Then $\varepsilon(-1) = (-1)^{(p-1)/2} = 1$; let $\rho = \varepsilon \oplus 1 \in \mathcal{G}_2(\mathbf{Q}_p)$; then ρ satisfies $\det \rho(-1) = 1$. In case

(2), the inertia group $I(F_{\mathfrak{p}}/\mathbf{Q}_p)$ is identified via local class field theory with a quotient of $\mathbf{Z}_p^*$; let χ be a character of $\mathbf{Z}_p^*$ factoring through an injective character of $I(F_{\mathfrak{p}}/\mathbf{Q}_p)$. By hypothesis, ε has order at least 3, and in particular $\varepsilon \neq \varepsilon^{-1}$. Let $\rho = \varepsilon \oplus \varepsilon^{-1}$; then ρ is minimal and $\det \rho(-1) = 1$. Finally in case (3), let ρ be a faithful irreducible representation of the dihedral group $\mathrm{Gal}(F_{\mathfrak{p}}/\mathbf{Q}_p)$, considered as a representation of $W_{\mathbf{Q}_p}$. Then ρ is minimal with $\det \rho(-1) = 1$ because -1 is a square in $\mathbf{Q}_p^*$. In each case, it is apparent that $\rho|_{W_{\mathfrak{p}}}$ is unramified.

Let $\lambda = \tau\left(\rho|_{I_{\mathbf{Q}_p}}\right)$ be the inertial type of ρ, considered as a representation of $\mathrm{GL}_2(\hat{\mathbf{Z}})$. We shall show that λ appears in the appropriate space of cusp forms. Proposition 3.4.1 implies that the number μ of newforms f of weight 2 and conductor a power of p for which $\rho_{f,p}^{\mathrm{WD}}|_{I_{\mathbf{Q}_p}} \cong \rho|_{I_{\mathbf{Q}_p}}$ is

$$\mu = \frac{1}{2}\left(\dim \lambda - \mu_{1728,0}(\lambda) - \mu_{0,0}(\lambda) - \mu_{\infty,0}(\lambda)\right),$$

where for $i \in \{1728, 0, \infty\}$, $\mu_{i,0}(\lambda)$ is the multiplicity of λ in $\Psi_{i,0}$. The rest is a calculation using Figure 2.1. In case (1), $\dim \lambda = p+1$, and by Lemma 3.5.1 we have the inequalities $\mu_{0,0} - (p+1)/3 \leq 4/3$ and $\mu_{1728,0} - (p+1)/2 \leq 1$, and $\mu_{\infty,0} = 2$, so for μ to be positive it is sufficient that $p \geq 29$. Case (2) is similar but we include the possibility $p = 23$, for then both G_0 and G_{1728} are elliptic and therefore $\mu_{0,0} = (p+1)/2$ and $\mu_{1728,0} = (p+1)/3$. Case (3) is again very similar, only this time $\dim \lambda \geq p-1$ and $\mu_{\infty,0} = 0$. $\square$

Chapter 4

Stable reduction of modular curves

4.1 The field over which $J_1(p^n)$ is semi-stable

Let $F/\mathbf{Q}_p$ be an algebraic extension which is finitely ramified. Let A/F be an abelian variety of dimension d.

Definition 4.1.1. The variety A is *semi-stable* if the connected component of the special fiber of its Néron model is an extension of an abelian variety by a torus.

Let $\ell \neq p$ be prime. Recall the definition of the contravariant Tate module:

$$V_\ell(A) = \mathrm{Hom}_{\mathbf{Z}_\ell}\left(\varprojlim_n A[\ell^n], \mathbf{Q}_\ell\right),$$

and let

$$\rho_{A,\ell} \colon G_F \to \mathrm{Aut}(V_\ell(A)) \cong \mathrm{GL}_{2d}(\mathbf{Q}_\ell)$$

be the associated Galois representation.

Theorem 4.1.2 ([Gro72], exp. IX)**.** *A is semi-stable if and only if $\rho_{A,\ell}|_{I_F}$ is unipotent.*

Let $p \geq 5$ be prime, let $n \geq 1$, and let $J = \operatorname{Jac} X_1(p^n)$, considered as an abelian variety defined over the field $K = \mathbf{Q}_p(\zeta_{p^n})$. In [Kri96] an explicit extension field of $\mathbf{Q}_p^{\mathrm{nr}}$ is given over which the Jacobian of $X_0(N)$ is assured to become semi-stable. The result of this section is a converse to this sort of theorem, whereby we construct an explicit extension field M of K^{nr} such that any extension field of K^{nr} over which J becomes semi-stable must contain M.

The field M is constructed as follows. Following the notation of [Kri96], let $\Omega_i/\mathbf{Q}_p,\ i = 1, 2, 3$ be the three quadratic extensions of $\mathbf{Q}_p$, with $\Omega_1/\mathbf{Q}_p$ unramified. One realization of this scenario is $\Omega_1 = \mathbf{Q}_p(\sqrt{D})$, $\Omega_2 = \mathbf{Q}_p(\sqrt{p})$, $\Omega_3 = \mathbf{Q}_p(\sqrt{Dp})$, where $D \in \mathbf{Z}_p^*$ is a quadratic nonresidue. For each i let $M_i/\Omega_i^{\mathrm{nr}}$ be the class field with norm subgroup U_i defined by

$$U_i = \begin{cases} \pm(1 + \mathfrak{p}_{\Omega_i}^{\lfloor n/2 \rfloor}) & i = 1 \\ 1 + \mathfrak{p}_{\Omega_i}^{n-1}, & i = 2, 3. \end{cases}$$

Finally let $M = M_1 M_2 M_3 K^{\mathrm{nr}}$.

Theorem 4.1.3. *J is semi-stable over M. Conversely, for all but finitely many values*[1] *of p^n, M is the minimal extension of K^{nr} over which J becomes semi-stable.*

Proof. The variety J is isogenous to $\prod_f A_f$, where f runs over Galois orbits of newforms of conductor dividing p^n.

For each i, let Θ_i be the set of characters θ of Ω_i^* satisfying

1. θ is minimal of conductor $\lfloor n/2 \rfloor$ if $i = 1$ and $n - 1$ if $i = 2, 3$, and

2. $\theta(-1) = 1$ if $i = 1$, and $\theta(-1) = (-1)^{(p-1)/2}$ if $i = 2, 3$.

For $\theta \in \Theta_i$, let $\rho_\theta = \operatorname{Ind}_{\Omega_i/\mathbf{Q}_p} \theta_i|_{I_{\mathbf{Q}_p}} \in \mathcal{G}_2^I(\mathbf{Q}_p)$. Then $\det \rho_\theta(-1) = 1$ and

$$\bigcap_{\theta \in \Theta_i} \ker \rho_\theta \text{ has fixed field precisely } M_i. \tag{4.1}$$

[1] *Quantifying over all triples (p, n) with $n \geq 1$, $p \geq 5$ prime. Certainly the theorem can be extended to* $\operatorname{Jac} X_1(Np^n)$*, where it would apply to* all *prime powers p^n so long as N is large enough.*

Further, if f is a newform of conductor dividing p^n, then $\rho := \rho_{f,p}^{\mathrm{WD}} \in \mathcal{G}_2^I(\mathbf{Q}_p)$ has conductor dividing p^n and satisfies $\det \rho(-1) = 1$. Then ρ must be one of the following:

1. decomposable as $\varepsilon_1 \oplus \varepsilon_2$, where the ε_i have conductor dividing p^n,
2. $\varepsilon \otimes \mathrm{Sp}(2)$, where ε has conductor dividing p^n, or
3. $\mathrm{Ind}_{\Omega_i/\mathbf{Q}_p} \theta$, where $i \in \{1, 2, 3\}$ and $\theta \in \Theta_i$.

In every case, $\rho|_{W_M}$ is unipotent. It follows that each A_f, and therefore J, is semi-stable over M.

For the converse statement, suppose J is semi-stable over $L \supset K^{\mathrm{nr}}$. Let $i \in \{1, 2, 3\}$ and let $\theta \in \Theta_i$. Theorem 3.5.2 implies that as long as p^n is large enough, there will be an eigenform $f \in S_2(\Gamma_1(p^n))$ for which $\rho_{f,p}^{\mathrm{WD}}|_{I_{\mathbf{Q}_p}} = \rho_\theta$. The assumption on L then implies $W_L \subset \ker \rho_\theta$. By (4.1) we have $L \supset M_i$ for each i, whence the theorem. □

4.2 Generalities on stable reduction

Let $F \subset \overline{\mathbf{Q}}$be as in the previous section, with residue field k. Let X/F be a smooth projective curve.

Definition 4.2.1. A *semi-stable model* for X is a pair $(\mathfrak{X}, \phi)$ consisting of a flat proper curve $\mathfrak{X}/\mathcal{O}_F$ together with an F-isomorphism $\phi \colon \mathfrak{X} \otimes_{\mathcal{O}_F} F \to X$ such that the singularities of $\overline{\mathfrak{X}} := \mathfrak{X} \otimes_{\mathcal{O}_F} k$ are ordinary double crossings of degree 1.

The definition of morphism between semi-stable models of X being evident, we define

Definition 4.2.2. A *stable* model for X is a final object in the category of semi-stable models of X.

Of course, the stable model is unique if it exists. Suppose there is a Galois extension field L/F for which X_L admits a stable model $(\mathfrak{X}, \phi)$. Let $G = \mathrm{Gal}(L/F)$. Then for

$\sigma \in G$, let $\mathfrak{X}^\sigma = \mathfrak{X} \otimes_{\mathcal{O}_L,\sigma} \mathcal{O}_L$ and let $\phi^\sigma \colon \mathfrak{X}^\sigma \otimes_{\mathcal{O}_L} L \to X_L$ be the composition

$$\mathfrak{X}^\sigma \otimes_{\mathcal{O}_L} L \longrightarrow (\mathfrak{X} \otimes_{\mathcal{O}_L} L) \otimes_{L,\sigma} L \xrightarrow{\phi\otimes 1} X_L \otimes_{L,\sigma} L \longrightarrow X_L,$$

the last arrow appearing because X is defined over F. Then $(\mathfrak{X}^\sigma, \phi^\sigma)$ is also a stable model for X_L, so that there is a canonical isomorphism $\mathfrak{X}^\sigma \to \mathfrak{X}$. Composing this with $1 \otimes \sigma \colon \mathfrak{X} \to \mathfrak{X}^\sigma$ gives an isomorphism $\sigma \colon \mathfrak{X} \to \mathfrak{X}$ lying over $\operatorname{Spec} \sigma \colon \operatorname{Spec} \mathcal{O}_L \to \operatorname{Spec} \mathcal{O}_L$. We have therefore described an $\mathcal{O}_L$-semilinear action of G on the curve $\mathfrak{X}$. Notice that the action of the inertia subgroup of G acts $\overline{\mathbb{F}}_p$-linearly on the special fiber $\mathfrak{X}_s = \mathfrak{X} \otimes_{\mathcal{O}_L} \overline{\mathbb{F}}_p$.

Let $\tilde{\mathfrak{X}}$ be the normalization of $\mathfrak{X}_s$; that is, the disjoint union of the irreducible components $\mathfrak{X}_i$ of $\mathfrak{X}_s$. Let Γ be the *dual graph* of $\mathfrak{X}_s$: this is the graph with one vertex v_i for each $\mathfrak{X}_i$ and an edge joining v_i to v_j for each point of $\mathfrak{X}_i \cap \mathfrak{X}_j$. Both $\tilde{\mathfrak{X}}$ and Γ admit actions of G. For an abelian group A, we have the cohomology group $H^1(\Gamma, A)$, defined as follows: Choose an orientation of the edges of Γ. Let $C^0(A)$ (respectively, $C^1(A)$) be the group of A-valued functions on the set of vertices (respectively, edges) of A. There is a map $\delta \colon C^0(A) \to C^1(A)$ sending a function f to the function $e \mapsto f(v_1) - f(v_0)$, where v_0 and v_1 are the origin and target of e respectively. Then $H^1(\Gamma, A) = \operatorname{coker} \delta$. The image of δ doesn't depend on the choice of orientation of Γ, so neither does $H^1(\Gamma, A)$.

Let $\ell \neq p$ and let $\eta = \operatorname{Spec} \overline{F}$. The data of $\tilde{\mathfrak{X}}$ and Γ essentially determine the ℓ-adic étale cohomology of the original curve X as a G_F-module:

Theorem 4.2.3. *The space $H^1(\mathfrak{X}_s, \mathbf{Q}_\ell)$ is the space of I_L-invariants in $H^1(X_\eta, \mathbf{Q}_\ell)$.*

Further, $H^1(X_\eta, \mathbf{Q}_\ell)$ *fits naturally into a diagram of* $\mathrm{Gal}(\overline{F}/F)$*-modules*

$$\begin{tikzcd}
 & 0 \arrow[d] & & & \\
 & H^1(\Gamma, \mathbf{Q}_\ell) \arrow[d] & & & \\
0 \arrow[r] & H^1(\mathfrak{X}_s, \mathbf{Q}_\ell) \arrow[r] \arrow[d] & H^1(X_\eta, \mathbf{Q}_\ell) \arrow[r] & H^1(\Gamma, \mathbf{Q}_\ell)(-1) \arrow[r] & 0 \\
 & H^1(\tilde{\mathfrak{X}}, \mathbf{Q}_\ell) \arrow[d] & & & \\
 & 0 & & &
\end{tikzcd} \tag{4.2}$$

with both sequences exact.

Remark 4.2.4. In particular, we have $\mathrm{genus}(X) = \sum_i \mathrm{genus}(\mathfrak{X}_i) + \dim H^1(\Gamma, \mathbf{Q}_\ell)$, where i indexes the irreducible components $\mathfrak{X}_i$ of $\tilde{\mathfrak{X}}$.

Proof. We start with the assertion that $H^1(\mathfrak{X}_s, \mathbf{Q}_\ell) = H^1(X_\eta, \mathbf{Q}_\ell)^{I_L}$. This argument is adapted from the proof of Prop. 3.13 in [Edi], §3.12. Let $S = \mathrm{Spec}\, \mathcal{O}_L$ and let Let J/S be the Néron model of the Jacobian of X_L. Then the connected component J^0 of J is equal to $\mathrm{Pic}^0_{\mathfrak{X}/S}$ by a property of Néron models, see [BLR90], Chap. 9. Let $n \geq 0$. By the Kummer exact sequence,

$$H^1(\mathfrak{X}_s, \mu_{\ell^n}) = \mathrm{Pic}^0_{\mathfrak{X}_s/s}[\ell^n] = J^0(s)[\ell^n] = J^0(L^{\mathrm{nr}})[\ell^n]$$

and

$$H^1(X_\eta, \mu_{\ell^n})^{I_L} = \mathrm{Pic}^0_{X_\eta/\eta}[\ell^n]^{I_L} = J(\overline{F})[\ell^n]^{I_L} = J(L^{\mathrm{nr}})[\ell^n].$$

Therefore the cokernel of the injection $H^1(X_s, \mu_{\ell^n}) \hookrightarrow H^1(X_\eta, \mu_{\ell^n})^{I_F}$ has order bounded by the order of the group of connected components of J. This order is independent of n. Taking inverse limits along n we conclude that the cokernel of $H^1(X_s, \mathbf{Q}_\ell(-1)) \hookrightarrow H^1(X_\eta, \mathbf{Q}_\ell(-1))^{I_F}$ is torsion, therefore 0.

The filtration of $H^1(X_\eta, \mathbf{Q}_\ell)$ shown in the diagram is an application of the machinery of the weight spectral sequence, which is treated in [RZ80], Satz 2.10 in the case of semi-stable reduction of varieties of arbitrary dimension. In our specific case, the technology unravels as follows: Let $Y^{(0)} = \tilde{\mathfrak{X}}$ and let $Y^{(1)}$ be the disjoint union of the singular points of $\mathfrak{X}_s$, considered as a variety over $s = \operatorname{Spec} \overline{\mathbb{F}}_p$. The weight spectral sequence begins on page one as

$$E_1^{pq} = \bigoplus_i H^{q-2i}(Y^{(2i+p)}, \mathbf{Q}_\ell)(-i) \implies H^{p+q}(X_\eta, \mathbf{Q}_\ell),$$

where the sum runs over integers $i \geq \max\{0, -p\}$ satisfying $2i+p \in \{0,1\}$. The picture on page one is

$$\begin{array}{ll}
E_1^{-1,2} = H^0(Y^{(1)}, \mathbf{Q}_\ell)(-1) \xrightarrow{\alpha} H^2(Y^{(0)}, \mathbf{Q}_\ell) & \\
\\
\qquad\qquad\qquad\qquad\qquad\qquad\qquad H^1(Y^{(0)}, \mathbf{Q}_\ell) & \\
\\
\qquad\qquad\qquad\qquad\qquad\qquad\qquad H^0(Y^{(0)}, \mathbf{Q}_\ell) \xrightarrow{\beta} H^0(Y^{(1)}, \mathbf{Q}_\ell) = E_1^{1,0}. &
\end{array}$$

Then $\ker \alpha = H^1(\Gamma, \mathbf{Q}_\ell)(-1)$ and $\operatorname{coker} \beta = H^1(\Gamma, \mathbf{Q}_\ell)$. Therefore the diagonal line $p + q = 1$ in page two has entries $H^1(\Gamma, \mathbf{Q}_\ell)(-1)$, $H^1(\tilde{\mathfrak{X}}, \mathbf{Q}_\ell)$ and $H^1(\Gamma, \mathbf{Q}_\ell)$. The spectral sequence degenerates on page two because there are no arrows between nonzero entries; therefore $H^1(X_\eta, \mathbf{Q}_\ell)$ admits a filtration with these three subquotients, as in the diagram of the theorem. □

4.3 Deligne-Carayol revisited

To motivate our discussion of the stable reduction of modular curves, we need to return to the setting of Theorem 3.2.1, where the cohomology of modular curves is discussed in detail. This section reviews the results of [Car83]. Let p be prime and let $n \geq 1$.

Let $N \geq 4$ be prime to p, and let X_n be the compactification of the moduli scheme over $S := \operatorname{Spec} \mathbf{Z}_p^{\mathrm{nr}}$ of elliptic curves with a $\Gamma(p^n)$-structure and a $\Gamma_1(N)$-structure, as defined in [KM85]. Then over $T = S[\zeta_{p^n}]$, X_n breaks up into a disjoint union

$\coprod_{a\in(\mathbf{Z}/p^n\mathbf{Z})^*} X_n^a$ of regular curves over T, with each X_n^a having generic fiber isomorphic to the classical modular curve $X(\Gamma(p^n)\cap\Gamma_1(N))$.

Let $s = \operatorname{Spec}\overline{\mathbb{F}}_p$. The special fiber $X_{n,s}^a$ of X_n^a admits the following description. For $b\in\mathbf{P}^1(\mathbf{Z}/p^n\mathbf{Z})$, let $C_n^{a,b}$ be the moduli scheme over $\overline{\mathbb{F}}_p$ parametrizing pairs $(E/S'/S,\phi,\psi)$ consisting of elliptic curves E over a base S', a $\Gamma_1(N)$-level structure ψ, and a $\Gamma(p^n)$-level structure $\phi\colon (\mathbf{Z}/p^n\mathbf{Z})^{\oplus 2}_{S'}\to E$ satisfying $\phi((1,0),(0,1))=\zeta_{p^n}^a$ and $\phi(b)=0$. Then naturally $C_n^{a,b}\subset X_{n,s}^a$ and

$$X_{n,s}^a = \bigcup_{b\in\mathbf{P}^1(\mathbf{Z}/p^n\mathbf{Z})} C_n^{a,b}$$

is the union of the $C_n^{a,b}$ meeting simultaneously over each supersingular point of the j-line. Each $C_n^{a,b}$ is abstractly isomorphic to the Igusa curve $\mathrm{Ig}(p^n)$ ([KM85], Chap. 10).

Let $\eta = \operatorname{Spec}\overline{\mathbf{Q}}_p$ and let

$$H^1 = \lim_{n\to\infty} H^1(X_{n,\eta},\overline{\mathbf{Q}}_\ell)$$

be the limit along n of the cohomology of the geometric generic fiber of X_n. Let also

$$\begin{aligned} H_s^1 &= \lim_{n\to\infty} H^1(X_{n,s},\overline{\mathbf{Q}}_\ell)\\ \tilde{H}_s^1 &= \lim_{n\to\infty} H^1(\tilde{X}_{n,s},\overline{\mathbf{Q}}_\ell)\end{aligned}$$

The various cohomology spaces assemble into an exact diagram

$$\begin{tikzcd} & 0 \arrow[d] & & \\ & K \arrow[d] & & \\ 0 \arrow[r] & H_s^1 \arrow[r] \arrow[d] & H^1 \arrow[r] & H_{\mathrm{e}}^1 \\ & \tilde{H}_s^1 \arrow[d] & & \\ & 0 & & \end{tikzcd} \tag{4.3}$$

in which every space appearing[2] admits an action of $\mathrm{GL}_2(\mathbf{Q}_p)\times W_{\mathbf{Q}_p}$. Carayol's theorem realizes the LLC inside the cohomology of X_n in the following sense.

Theorem 4.3.1 ([Car83], §11.1)**.** *We have a decomposition*

$$H^1 = \bigoplus_f \pi_{f,p} \otimes \rho_{f,p}^\vee,$$

where f runs over all newforms of weight 2 whose prime-to-p-conductor divides N. With reference to the diagram in 4.6, the "locations" of the terms in the sum are given by

1. *$\rho_{f,p}$ is decomposable if and only if $\pi_{f,p} \otimes \rho_{f,p}^\vee \subset H_s^1$, in which case $\pi_{f,p} \otimes \rho_{f,p}^\vee$ is isomorphic to its image in $\tilde{H}_s^1$,*
2. *$\rho_{f,p}$ is irreducible if and only if $\pi_{f,p} \otimes \rho_{f,p}^\vee$ is isomorphic to its image in H_e^1, and*
3. *If $\rho_{f,p} = \chi \otimes \mathrm{Sp}(2)$, the image of $\pi_{f,p} \otimes \rho_{f,p}^\vee$ in H_e^1 is isomorphic to $\pi_{f,p} \otimes \chi^{-1}(-1)$; the kernel lies in K and is isomorphic to $\pi_{f,p} \otimes \chi^{-1}$.*

4.4 Stable models for modular curves

In the proof of the Theorem 3.2.1 of Deligne-Carayol, the cases where $\rho_{f,p}$ is decomposable are treated by a study of the smooth curves $C_n^{a,b}$ in characteristic p. However, the treatment of the cases where $\rho_{f,p}$ is non-semisimple or irreducible involves an analysis of the vanishing cycles H_{e}^1, which carry an action of a certain quaternion group, and in particular it relies upon the Jacquet-Langlands correspondence. If on the other hand one could calculate a stable model $\mathfrak{X}_n$ for X_n over an appropriate ramified extension of S, then in view of Theorem 4.2.3, one would be able to "see" the theorem of Deligne-Carayol, and therefore the LLC for GL_2 over $\mathbf{Q}_p$, in the structure of the smooth curve $\tilde{\mathfrak{X}}_n$ and the dual graph Γ!

[2] The space H_{e}^1 admits an interpretation as a space of "vanishing cycles," but we will not need to define these here.

In this final section of our thesis, we present a known result on the stable reduction of the curve X_1. We then "reverse-engineer" a conjectural stable reduction of X_2 in a way that is consistent with Theorem 4.3.1.

Let X_{DL} be the smooth projective curve containing a dense open subscheme isomorphic to the affine curve $(x^p y - x^p y)^{p-1} = 1$. Then X_{DL} admits an action of $\mathrm{GL}_2(\mathbb{F}_p)$ given by

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} (x, y) = (ax + by, cx + dy).$$

For $a \in \mathbb{F}_p^*$, suppose X_{DL}^a is the component of this curve whose equation is $xy^p - x^p y = a$. Then X_{DL}^a has points at infinity $\infty^{a,b}$ indexed by $b \in \mathbf{P}^1(\mathbb{F}_p)$.

Theorem 4.4.1. *The curve X_1 admits a stable model $\mathfrak{X}_1$ over the tamely ramified extension of $\mathbf{Q}_p^{nr}$ of degree $p^2 - 1$. The reduction $\overline{\mathfrak{X}}_1$ over $\overline{\mathbb{F}}_p$ consists of*

1. *A disjoint union of Igusa curves $C_1^{a,b}$ for $a \in \mathbb{F}_p^*$ and $b \in \mathbf{P}^1(\mathbb{F}_p)$. An element $g \in \mathrm{GL}_2(\mathbb{F}_p)$ carries $C_1^{a,b}$ onto $C_1^{(\det g)a, gb}$.*

2. *For each supersingular point in $X_0 = X_1(N)$, a copy of the curve X_{DL}^a, attached at $\infty^{a,b}$ to the corresponding supersingular point of $C_1^{a,b}$ for all b.*

Proof. The fact that the supersingular components are the curves X_{DL} with this particular action of $\mathrm{GL}_2(\mathbb{F}_p)$ is in [BW04]. The complete case of the modular curve $X_1(Np^2)$, which is a cover of each X_1^a, is treated in [Joy06], Theorem 1. □

Remark 4.4.2. The curve X_{DL} is the Deligne-Lusztig curve for the group GL_2. Its appearance in the stable reduction of X_1 is a special case of the work carried out in [Yos04], whereby the "depth 0" case of the LLC for GL_n is realized in the vanishing cycles of a deformation space of formal groups.

We now turn to X_2. Assume it admits a stable model $\mathfrak{X}_2$. As was the case for X_1, we expect the quotient map $\mathfrak{X}_{2,s} \to X_1(N)_s$ on the special fiber to be flat outside of the supersingular locus of $X_1(N)_s$. The preimage of each supersingular point in

$X_1(N)_s$ ought to have components whose cohomology carries representations of the group $\mathrm{GL}_2(\mathbf{Z}/p^2\mathbf{Z}) \times I_{\mathbf{Q}_p}$ in accordance with Theorem 4.3.1.

First we classify the irreducible inertial WD representations $\rho \in \mathcal{G}_2^I(\mathbf{Q}_p)^0_{\overline{\mathbf{Q}}_\ell}$ whose types have level ≤ 2. Assume $p \geq 3$. Then any lift of ρ to $W_{\mathbf{Q}_p}$ is of the form $\mathrm{Ind}_{\Omega/\mathbf{Q}_p} \theta$ for one of the quadratic extensions $\Omega/\mathbf{Q}_p$ and a character θ of Ω^*. We say that such a θ has *essential conductor* $\mathfrak{p}_\Omega^n$ if $\theta = \theta_0 \times \chi \circ N_{\Omega/\mathbf{Q}_p}$ for a character χ of $\mathbf{Q}_p^*$ and θ_0 minimal of conductor $\mathfrak{p}_\Omega^n$. We divide the pairs (Ω, θ) for which $\ell(\rho) \leq 2$ into "species" as follows:

1. $\mathfrak{S}_2$: $\Omega = \Omega_1$, θ has essential conductor $\mathfrak{p}_{\Omega_1}$,
2. $\mathfrak{S}_3^i$, $i = 2, 3$: $\Omega = \Omega_i$ and θ has essential conductor $\mathfrak{p}_{\Omega_i}^2$,
3. $\mathfrak{S}_4$: $\Omega = \Omega_1$ and θ has essential conductor $\mathfrak{p}_{\Omega_1}^2$.

The subscript on the $\mathfrak{S}$ refers to the minimal conductor of a twist of ρ. In view of part 3 of Theorem 2.4.1, $\mathrm{rec}(\rho)$ is well-defined as a representation of $\mathrm{GL}_2(\mathbf{Z}_p)$. Let $K(p^2) = \left\{ g \in \mathrm{GL}_2(\mathbf{Z}_p) \colon g \equiv \mathrm{Id}_2 \pmod{p^2} \right\}$.

Proposition 4.4.3. *Let $G_2 = \mathrm{GL}_2(\mathbf{Z}/p^2\mathbf{Z})$. For each of the species $\mathfrak{S}$ above, there exists a smooth projective curve $X_{\mathfrak{S}}$ over $\overline{\mathbb{F}}_p$ admitting an action of $G_2 \times I_{\overline{\mathbf{Q}}_p}$ for which*

$$H^1(X_{\mathfrak{S}}, \overline{\mathbf{Q}}_\ell) \cong \left[\bigoplus_{\rho \in \mathfrak{S}} \mathrm{rec}(\rho)^{K(p^2)} \otimes \rho^\vee \right]^{\oplus n_{\mathfrak{S}}}$$

for a certain multiplicity $n_{\mathfrak{S}}$.

Remark 4.4.4. The form of the curves $X_{\mathfrak{S}}$ in the proof below is suggested by recent unpublished work of Wewers, and by the forms of the curves appearing in [CM06]. They are probably not unique for the property above.

Proof. For each $i = 1, 2, 3$, let $\psi_i \colon I_{\Omega_i} \to \mathcal{O}_{\Omega_i}^*$ be the restriction to inertia of the inverse to the reciprocity map $\mathrm{Art} : \Omega_i^* \to W_{\Omega_i}^{\mathrm{ab}}$.

When $\rho \in \mathfrak{S} = \mathfrak{S}_2$, we have

$$\operatorname{rec}(\rho)|_K = \bigoplus_g \operatorname{Ind}_{K \cap gKg^{-1}}^K \tau(\rho), \tag{4.4}$$

where g runs over the coset representatives $\begin{pmatrix} p^a & \\ & 1 \end{pmatrix}$ of $\Omega_1^* K \backslash G / K$, $a \geq 0$, as in [Hen02], A.3.6. Therefore

$$\operatorname{rec}(\tilde{\rho})^{K(2)} = \tau(\rho) \oplus \operatorname{Ind}_{K_0(1)}^K \tau(\rho)^{g_0}, \tag{4.5}$$

where $g_0 = \begin{pmatrix} p & \\ & 1 \end{pmatrix}$.

Let $X^0_{\mathfrak{S}_2} = X_{DL}$ be the smooth projective curve over $\overline{\mathbb{F}}_p$ with affine part

$$(x^p y - y^p x)^{p-1} = 1.$$

This admits an action of the product group $\mathrm{GL}_2(\mathbb{F}_p) \times \mathbb{F}_{p^2}^*$, by the formula

$$(g, \beta)(x, y) = (\beta(ax + by), \beta(cx + dy)),$$

for every $g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{GL}_2(\mathbb{F})_p$ and $\beta \in \mathbb{F}_{p^2}^*$. A calculation using the Lefshetz fixed-point formula gives

$$[H^1(X_{DL}, \overline{\mathbf{Q}}_\ell)] = \bigoplus_\theta [\tau(\theta) \otimes \theta^{-1}]$$

as an equality in the Grothendieck ring $R(\mathrm{GL}_2(\mathbb{F}_p) \times \mathbb{F}_{p^2}^*)$. Here θ runs over the characters of $\mathbb{F}_{p^2}^*$ satisfying $\theta \neq \theta^p$ and $\tau(\theta)$ is the unique irreducible representation of $\mathrm{GL}_2(\mathbb{F}_p)$ whose trace on an element g is $-(\theta(\alpha) + \theta(\alpha^p))$ whenever g has eigenvalues $\alpha, \alpha^p \in \mathbb{F}_{p^2}^* \backslash \mathbb{F}_p^*$. In fact, if θ is identified via ψ_1 with a character of I_{Ω_1}, then $\tau(\theta)$ is the inertial type for $\operatorname{Ind}_{\Omega_1/\mathbf{Q}_p} \tilde{\theta}$ for any lift of θ to W_{Ω_1}.

Let $H_2 \subset \mathrm{GL}_2(\mathbb{F}_p) \times I_{\mathbf{Q}_p}$ be the subgroup of pairs (g, α) satisfying $\det g^{-1} \equiv N_{\Omega_1/\mathbf{Q}_p} \psi_1(\alpha) \pmod{p^2}$. We define an action of H_2 on $X_{\mathfrak{S}_2}$ by pullback from the map $H_2 \to \mathrm{GL}_2(\mathbb{F}_p) \times \mathbb{F}_{p^2}^*$, $(g, \alpha) \mapsto (g, \psi_1(\alpha))$.

Let

$$X^1_{\mathfrak{S}_2} = (G_2 \times I_{\mathbf{Q}_p}) \times_H X^0_{\mathfrak{S}_2}.$$

Then

$$H^1(X^1_{\mathfrak{S}_2}, \mathbf{Q}_\ell) = \mathrm{Ind}_{H_2}^{G_2 \times I_{\mathbf{Q}_p}} H^1(X^0_{\mathfrak{S}_2}, \mathbf{Q}_\ell)$$

is the sum of $\tau(\rho) \otimes \rho^\vee$ for $\rho \in \mathfrak{S}_2$. Finally, let

$$X_{\mathfrak{S}_2} = X^1_{\mathfrak{S}_2} \coprod \left(G_2 \times_{K_0(1)} (X^1_{\mathfrak{S}_2})^{g_0} \right),$$

where $(X^1_{\mathfrak{S}_2})^{g_0}$ is the same underlying curve as $X^1_{\mathfrak{S}_2}$ but with an action of $K_0(1) = K \cap gKg^{-1}$ conjugate through g to the action of K on $X^1_{\mathfrak{S}_2}$. Then by Equation 4.5,

$$H^1(X_{\mathfrak{S}_2}, \overline{\mathbf{Q}}_\ell) \cong \bigoplus_{\rho \in \mathfrak{S}_2} \mathrm{rec}(\rho)^{K(2)} \otimes \rho^\vee.$$

When $\mathfrak{S} = \mathfrak{S}_3^i$, $\mathrm{rec}(\rho)^{K(2)} = \tau(\rho)$: This follows from Theorem 3 of [Cas73b]. Let $\mathfrak{A} \subset M_2(\mathbf{Z}_p)$ be the hereditary chain order with $e = 2$ as in Section 2.5.3. Choose a uniformizer ϖ of Ω_i. We identify $\mathfrak{A}$ with $\mathcal{O}_{\Omega_i} \oplus \mathcal{O}_{\Omega_i}\sigma$ and define the subgroup $U^1 = (1 + \mathfrak{p}_{\Omega_i}) + \mathfrak{p}_{\Omega_i}\sigma$. Let $X^0_{\mathfrak{S}_3^i}$ be the smooth projective curve over $\mathbb{F}_p$ with affine part

$$y^2 = x^p - x,$$

and let U^1 act on $X^0_{\mathfrak{S}_3^i}$ as follows: if $u \in U^1$, write $u = r + s\sigma$ with $r = 1 + \varpi t$ and let $u(x, y) = (x + \mathrm{Tr}_{\Omega_i/\mathbf{Q}_p} t, y)$. Then one can check that as a U^1-module, $H^1(X^0_{\mathfrak{S}_3^i}, \mathbf{Q}_\ell) = \sum_\varepsilon [\varepsilon]$, where ε runs over the nontrivial characters of $1 + \mathfrak{p}_{\Omega_i}$ vanishing on $1 + \mathfrak{p}^2_{\Omega_i}$.

Define

$$H_3 = \left\{ (g, \alpha) \in G_2 \times I_{\Omega_i} | g\psi_i(\alpha)^{-1} \in U^1 \cap S_2 \right\},$$

and let H_3 act on $X_{\mathfrak{S}_3^i}$ by pullback from $U^1 \cap S_2$ through the map $(g, \alpha) \mapsto g\psi_i(\alpha)^{-1} \in U_1 \cap S_2$ and let $X_{\mathfrak{S}_3^i} = (G_2 \times I_{\mathbf{Q}_p}) \times_{H_3} X^0_{\mathfrak{S}_3^i}$. Let $J^0 = \mathcal{O}^*_{\Omega_i} U^1$. Then

$$\begin{aligned} H^1(X_{\mathfrak{S}_3^i}, \mathbf{Q}_\ell) &= \bigoplus_\varepsilon \mathrm{Ind}_{H_3}^{G_2 \times I_{\mathbf{Q}_p}} \varepsilon \\ &= \bigoplus_\varepsilon \mathrm{Ind}_{J^0 \times I_{\Omega_i}}^{G_2 \times I_{\mathbf{Q}_p}} \mathrm{Ind}_H^{J^0 \times I_{\Omega_i}} \varepsilon \\ &= \bigoplus_\varepsilon \mathrm{Ind}_{J^0 \times I_{\Omega_i}}^{G_2 \times I_{\mathbf{Q}_p}} \bigoplus_\theta [\theta \otimes \theta^{-1}], \end{aligned}$$

where θ runs over the characters of $\mathcal{O}_E^*$ which restrict to $1+\mathfrak{p}$ as ε up to a twist by some $\chi \circ N_{\Omega_i/\mathbf{Q}_p}$, and $[\theta\otimes\theta^{-1}]$ means the character of $J^0 \times I_{\Omega_i}$ defined by $(g,\alpha)\mapsto \theta(g)\theta(\psi_i(\alpha))^{-1}$. As ε runs over the nontrivial characters of $(1+\mathfrak{p}_{\Omega_i})/(1+\mathfrak{p}_{\Omega_i}^2)$, θ runs over the characters of $\mathcal{O}_E^*$ with minimal conductor $\mathfrak{p}_{\Omega_i}^2$. Noting that $\tau(\mathrm{Ind}_{\Omega_i/\mathbf{Q}_p}\tilde{\theta}) = \mathrm{Ind}_{J^0}^{G_2}\theta$ for any lift $\tilde{\theta}$ of θ to W_{Ω_i} (see Theorem 2.5.3), we find

$$\begin{aligned} H^1(X_{\mathfrak{S}_3^i}, \overline{\mathbf{Q}}_\ell) &= \bigoplus_\theta \mathrm{Ind}_{J^0\times I_{\Omega_i}}^{G_2\times I_{\mathbf{Q}_p}}[\theta\otimes\theta^{-1}] \\ &= \bigoplus_\theta \tau\left(\mathrm{Ind}_{\Omega_i/\mathbf{Q}_p}\theta\right)\otimes \mathrm{Ind}_{\Omega_i/\mathbf{Q}_p}\theta^{-1} \\ &= \bigoplus_{\rho\in\mathfrak{S}_3^i} \tau(\rho)\otimes\rho^\vee \end{aligned}$$

When $\mathfrak{S}=\mathfrak{S}_4$, $\mathrm{rec}(\tilde{\rho})^{K(2)}=\tau(\rho)$. Let $\mathfrak{A}=M_2(\mathbf{Z}_p)=\mathcal{O}_{\Omega_1}\oplus\mathcal{O}_{\Omega_1}\sigma$ be the hereditary chain order with $e=1$. Say $\mathcal{O}_{\Omega_1}=\mathbf{Z}_p[\sqrt{d}]$. This time

$$U^1=(1+\mathfrak{p}_{\Omega_1})+\mathfrak{p}_{\Omega_1}\sigma\subset\mathfrak{A}^*$$

equals $K(1)$, the subgroup of matrices congruent to Id_2 mod p. Define $X_{\mathfrak{S}_4}^0$ as the smooth projective curve with affine part $x^py-y^px=1$, with an action of $U^1\cap S_2$ as follows: if $r+s\sigma\in U^1\cap S_2$, then $r=1+p\sqrt{d}t$ with $t\in\mathbf{Z}_p$; let this element act on $X_{\mathfrak{S}_4}^0$ as $(x,y)\mapsto(x+\bar{t}y,y)$. Then as a $U^1\cap S_2$-module, $H^1(X_{\mathfrak{S}_4}^0,\mathbf{Q}_\ell)=\bigoplus_\varepsilon p[\varepsilon]$, where ε runs over the $p-1$ nontrivial characters of $(1+\mathfrak{p}_{\Omega_1})^{\mathrm{Tr}=0}$ vanishing on $1+\mathfrak{p}_{\Omega_1}^2$.

Let

$$H_4=\left\{(g,\alpha)\in G_2\times I_{\mathbf{Q}_p}\colon g\psi_1(\alpha)^{-1}\in U_1\cap S_2\right\}$$

and extend the action of $U^1\cap S_2$ on $X_{\mathfrak{S}_4}^0$ to H_4 via the map $(g,b)\mapsto g\psi_1(b)^{-1}$. Let $X_{\mathfrak{S}_4}=(G_2\times I_{\mathbf{Q}_p})\times_{H_4}X_{\mathfrak{S}_4}^0$. The calculation precedes almost word-for-word as in the case of $\mathfrak{S}_3^i$, resulting in

$$H^1(X_{\mathfrak{S}_4},\overline{\mathbf{Q}}_\ell)=\bigoplus_{\rho\in\mathfrak{S}_4}\left[\tau(\rho)\oplus\rho^\vee\right]^{\oplus p}.$$

$\square$

Remark 4.4.5. We note that the dimensions of $\tau(\rho)$ for $\rho \in \mathfrak{S}_2, \mathfrak{S}_3^i, \mathfrak{S}_4$ are $p-1$, p^2-1 and p^2-p, respectively, and that $n_{\mathfrak{S}_2} = n_{\mathfrak{S}_3^i} = 1$, while $n_{\mathfrak{S}_4} = p$.

Let X_{LT} for ("Lubin-Tate") be any semi-stable curve over $\overline{\mathbb{F}}_p$ with an action of $G_2 \times I_{\mathbf{Q}_p}$ whose components of genus greater than 0 are

1. $X_{\mathfrak{S}_2}$,
2. $(p+1)$ copies of $X_{\mathfrak{S}_3^i}$ for each $i \in \{2,3\}$,
3. $X_{\mathfrak{S}_3}$,

satisfying the conditions

1. X_{LT} is the disjoint union of connected components X_{LT}^a, where a runs through $(\mathbf{Z}/p^2\mathbf{Z})^*$,
2. Each $(g,\alpha) \in G_2 \times I_{\mathbf{Q}_p}$ carries X_{LT}^a isomorphically onto $X_{\mathrm{LT}}^{\det(g)\chi_{p^2}(\alpha)a}$, where $\chi_{p^2}\colon I_{\mathbf{Q}_p} \to (\mathbf{Z}/p^2\mathbf{Z})^*$ is the cyclotomic character, and
3. The dual graph of X_{LT} is contractible.

It is not obvious that such a curve should exist. In Figure 4.4 we present a possible configuration of one of the connected components of X_{LT} in the case of $p=3$. Here, the blue line represents a connected component of $X_{\mathfrak{S}_2}^1$, the cyan lines represent $X_{\mathfrak{S}_2} \backslash X_{\mathfrak{S}_2}^1$, the green lines represent components of $X_{\mathfrak{S}_3}^i$ for $i=2,3$, the red lines represent $X_{\mathfrak{S}_4}$, and finally the black lines are rational curves. We think, but we do not check it here, that $G_2 \times I_{\mathbf{Q}_p}$ acts on this configuration in a consistent manner.

Conjecture 4.4.6. *There is a stable model $\mathfrak{X}_2$ of the modular curve X_2 over the field*

$$M = \bigcap_{\rho \in \mathfrak{S}} \ker \rho \supset \mathbf{Q}_p^{nr}(\zeta_{p^2}).$$

The special fiber $\mathfrak{X}_s$, considered as a curve over $\overline{\mathbb{F}}_p$ together with an action of $G_2 \times I_{\mathbf{Q}_p}$, consists of:

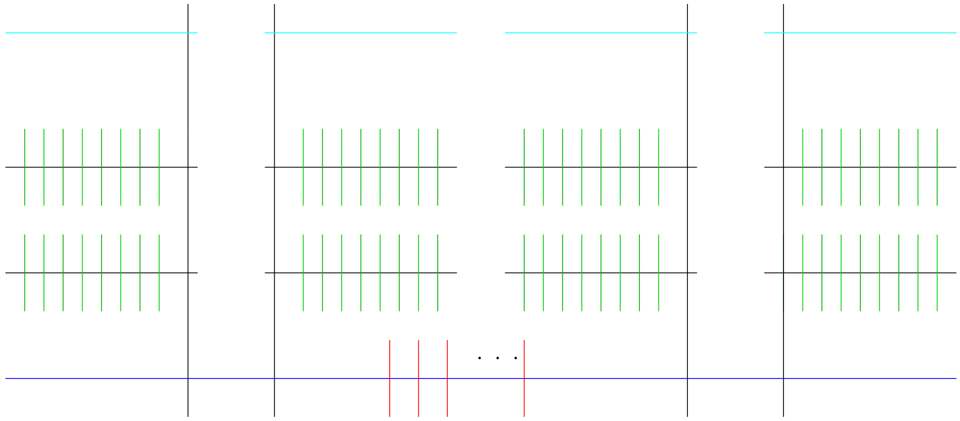

Figure 4.1. A connected component of the curve X_{LT}, shown in the case of $p=3$.

1. *The Igusa curves* $C_2^{a,b}$ *for* $a \in (\mathbf{Z}/p^2\mathbf{Z})^*$, $b \in \mathbf{P}^1(\mathbf{Z}/p^2\mathbf{Z})$, *and*

2. *For each supersingular point* A *of* $X_1(N)$ *and each* $a \in (\mathbf{Z}/p^2\mathbf{Z})^*$, *a copy of the curve* X_{LT}^a *is glued at one point to* $C_2^{a,b}$ *for each* $b \in \mathbf{P}^1(\mathbf{Z}/p^2\mathbf{Z})$.

The conjecture is consistent with Theorem 4.3.1:

Theorem 4.4.7. *Let* $Y/\mathbf{Q}_p^{nr}$ *be a smooth Galois cover of* $X_1(N)$ *with group* G_2. *Assume that* Y *admits a stable model* $\mathfrak{Y}$ *over some extension field of* $\mathbf{Q}_p^{nr}$ *for which the special fiber* $\mathfrak{Y}_s$ *of the form appearing in Conjecture 4.4.6. Then as* $G_2 \times I_{\mathbf{Q}_p}$*-modules,*

$$H^1(Y_\eta, \overline{\mathbf{Q}}_\ell) \cong H^1(X_{2,\eta}, \overline{\mathbf{Q}}_\ell).$$

Proof. Theorem 4.3.1 predicts the structure of $H^1(X_{2,\eta}, \overline{\mathbf{Q}}_\ell)$ as a module for $G_2 \times W_{\mathbf{Q}_p}$:

$$H^1(X_{2,\eta}, \overline{\mathbf{Q}}_\ell) \cong \bigoplus_f \pi_{f,p}^{K(p^2)} \otimes \rho_{f,p}^\vee,$$

where f runs over newforms whose prime-to-p conductor divides N and for which $\pi_{f,p}$ admits a nonzero vector fixed by $K(p^2)$. We will show that for each $\rho^\vee \in \mathcal{G}_2^I(\mathbf{Q}_p)$, the $\rho^\vee$-isotypic part of $H^1(Y_\eta, \overline{\mathbf{Q}}_\ell)$ is $\mathrm{rec}(\rho)^{K(2)} \otimes \rho^\vee$ with the correct multiplicity.

By Theorem 4.2.3, we have a diagram

$$
\begin{tikzcd}
 & 0 \arrow[d] & & & \\
 & H^1(\Gamma,\overline{\mathbf{Q}}_\ell) \arrow[d] & & & \\
0 \arrow[r] & H^1(\mathfrak{Y}_s,\overline{\mathbf{Q}}_\ell) \arrow[r] \arrow[d] & H^1(Y_\eta,\overline{\mathbf{Q}}_\ell) \arrow[r] & H^1(\Gamma,\overline{\mathbf{Q}}_\ell)(-1) \arrow[r] & 0, \\
 & H^1(\tilde{\mathfrak{Y}}_s,\overline{\mathbf{Q}}_\ell) \arrow[d] & & & \\
 & 0 & & &
\end{tikzcd}
\tag{4.6}
$$

where $\tilde{\mathfrak{Y}}$ is the normalization of $\mathfrak{Y}_s$ and Γ is its dual graph, and

$$H^1(\mathfrak{Y}_s,\overline{\mathbf{Q}}_\ell) = H^1(Y_\eta,\overline{\mathbf{Q}}_\ell)^{I_M}. \tag{4.7}$$

Let S be the set of supersingular points of $X_1(N)_s$.

First we consider those $\rho^\vee \subset H^1(Y_\eta,\overline{\mathbf{Q}}_\ell)$ which are contained in the cohomology of the special fiber $H^1(\mathfrak{Y}_s,\overline{\mathbf{Q}}_\ell)$ and which are isomorphic to their image in $H^1(\tilde{\mathfrak{Y}},\overline{\mathbf{Q}})$. The normalization is

$$\tilde{\mathfrak{Y}} = \left(\coprod_{a,b} C_2^{a,b}\right) \coprod \left(\#S \text{ copies of } X_{LT}\right),$$

where a runs over $(\mathbf{Z}/p^2\mathbf{Z})^*$, b runs over $\mathbf{P}^1(\mathbf{Z}/p^2\mathbf{Z})$ and ρ runs over all the inertial WD representations belonging to the species $\mathfrak{S}_2$, $\mathfrak{S}_3^i$ and $\mathfrak{S}_4$. Already by part 1 of Theorem 4.3.1 we have

$$\bigoplus_{a,b} H^1(C_2^{a,b},\overline{\mathbf{Q}}_\ell) \cong \bigoplus_f \pi_{f,p}^{K(p^2)} \otimes \rho_{f,p}^\vee,$$

where f ranges over those forms with $\rho_{f,p}$ decomposable.

Now let $\rho^\vee$ be an *irreducible* inertial WD representation appearing in $H^1(\tilde{\mathfrak{Y}},\overline{\mathbf{Q}}_\ell)$. We claim that

$$\#S \times \left\langle \tau(\rho), H^1(X_{LT},\overline{\mathbf{Q}}_\ell)\right\rangle_{G_2} = \left\langle \tau(\rho), H^1(X_{2,\eta},\overline{\mathbf{Q}}_\ell)\right\rangle_{G_2}. \tag{4.8}$$

In view of Theorem 4.3.1 and Proposition 4.4.3, this would mean that $\tau(\rho)\otimes\rho^\vee$ appears in the cohomology of Y with the same multiplicity as it does for X_2. By Remark 4.4.5, the LHS of Equation 4.8 is $2\#S\dim\tau(\rho)/(p-1)$, regardless of the species to which ρ belongs. We are reduced to showing that

$$\left\langle \tau(\rho), H^1(X_{2,\eta}, \overline{\mathbf{Q}}_\ell)\right\rangle = 2\#S\dim\tau(\rho)/(p-1) \tag{4.9}$$

For the RHS, we consider the generic fiber of the cover $X_2 \to X_1(N)$ with Galois group G_2. In characteristic zero, this cover is unramified away from the cusps. Let c be the number of cusps of $X_1(N)$. The preimage of each of these cusps in X_2 can be identified as a G_2-set with G_2/U, where up to conjugacy $U = \begin{pmatrix} 1 & * \\ & 1 \end{pmatrix} \subset G_2$. There exists an integer m for which we have as an equality in $R(G_2)$:

$$[H^*(X_{2,\eta}, \mathbf{Q}_\ell)] = m\mathbf{Q}_\ell G_2 + c\operatorname{Ind}_U^{G_2} 1, \tag{4.10}$$

obtained by comparing the traces of both sides on an element $g \in G_2\setminus\{\mathrm{Id}_2\}$ and applying the Lefshetz fixed-point formula. Let g be the genus of $X_1(N)$. By Riemann-Hurwitz,

$$-\dim H^*(X_{2,\eta}, \mathbf{Q}_\ell) = (2g-2)\#G_2 + c(\#G_2 - \#G_2/U); \tag{4.11}$$

using this to solve for m in Equation 4.10 gives $m = 2-2g-c$ and

$$[H^1(X_{2,\eta}, \mathbf{Q}_\ell)] = (2g-2+c)\mathbf{Q}_\ell G_2 - c\operatorname{Ind}_U^{G_2} 1 + [H^0(X_{2,\eta}, \mathbf{Q}_\ell)] + [H^1(X_{2,\eta}, \mathbf{Q}_\ell)]. \tag{4.12}$$

Now $\left\langle \tau(\rho), \operatorname{Ind}_U^{G_2} 1\right\rangle_{G_2} = \langle \tau(\rho)|_U, 1\rangle_U = 0$ by [Cas73b]: The restriction of $\tau(\rho)$ to U only contains nontrivial characters of U. The action of G_2 on $H^i(X_{2,\eta}, \mathbf{Q}_\ell)$ for $i = 0, 2$ factors through the determinant map, so these spaces cannot contain $\tau(\rho)$. We find that the multiplicity of $\tau(\rho)$ in $H^1(X_{2,\eta}, \mathbf{Q}_\ell)$ is

$$\left\langle \tau(\rho), H^1(X_{2,\eta}, \mathbf{Q}_\ell)\right\rangle = (2g-2+c)\dim\tau(\rho). \tag{4.13}$$

To deduce Equation 4.9 from Equation 4.13, we are down to showing that

$$2\#S = (2g-2+c)(p-1). \tag{4.14}$$

Consider the curve $X = X(\Gamma_0(p) \cap \Gamma_1(N))$; by a well-known result of Deligne-Rapoport, X admits a model over $\mathbf{Z}_p^{\mathrm{nr}}$ whose reduction consists of two copies of $X_1(N)$ crossing transversely at each of the $\#S$ supersingular points. On the one hand, by Remark 4.2.4 the genus of X is

$$g_X = 2g + \#S - 1, \tag{4.15}$$

because in this situation the dual graph of the special fiber of the Deligne-Rapoport model has Betti number $\#S - 1$. On the other hand, the cover $X \to X_1(N)$ has degree $(p+1)$ and is ramified at exactly one point in the preimage of each cusp of $X_1(N)$, with index p. Therefore Riemann-Hurwitz gives

$$2g_X - 2 = (2g-2)(p+1) + c(p-1). \tag{4.16}$$

Putting together Equations 4.15 and 4.16 establishes Equation 4.14.

We turn now to the graph Γ. It has connected components Γ^a for $a \in (\mathbf{Z}/p^2\mathbf{Z})^*$, and each Γ^a is homotopic to the graph with vertices $v^{a,b}$ for each $b \in \mathbf{P}^1(\mathbf{Z}/p^2\mathbf{Z})$ (these represent the Igusa curves $C_2^{a,b}$) and w_x^a for each $x \in S$ (these represent X_{LT}^a, whose dual graph was assumed contractible), with an edge connecting each v_b^a to each w_x^a. An element $\alpha \in I_{\mathbf{Q}_p}$ takes Γ^a to $\Gamma^{\chi_{p^2}(\alpha)(a)}$, where $\chi\colon W_{\mathbf{Q}_p} \to (\mathbf{Z}/p^2\mathbf{Z})^*$ is the cyclotomic character. An element $g \in G_2$ sends the vertex $v^{a,b}$ to $v^{a \det g,\, gb}$. Therefore as a $G_2 \times I_{\mathbf{Q}_p}$-module we have

$$H^1(\Gamma, \overline{\mathbf{Q}}_\ell) \cong \bigoplus_{\chi} \left[\chi \cdot \mathrm{St}_2 \otimes \chi^{-1}\right]^{\oplus(\#S-1)}, \tag{4.17}$$

where χ runs over characters of $(\mathbf{Z}/p^2\mathbf{Z})^*$ and St_2 is the complement of the G_2-fixed line in the permutation representation of G_2 on $\mathbf{P}^1(\mathbf{Z}/p^2\mathbf{Z})$.

Now consider the case where an inertial representation $\rho^\vee \subset H^1(Y_\eta, \overline{\mathbf{Q}}_\ell)$ is indecomposable and not contained in $H^1(\mathfrak{Y}_s, \overline{\mathbf{Q}})$. Such a $\rho^\vee$ must remain ramified when restricted to M, by Equation 4.7. But then by the above description of the $I_{\mathbf{Q}_p}$-module structure of $H^1(\Gamma, \overline{\mathbf{Q}}_\ell)$, the image of $\rho^\vee$ in $H^1(\Gamma, \overline{\mathbf{Q}}_\ell)$ must be a single character of $I_{\mathbf{Q}_p}$ factoring through a character χ^{-1} of $\mathbf{Q}_p^{\mathrm{nr}}(\zeta_{p^2}) \subset M$. It follows that $\rho^\vee \cap H^1(\mathfrak{Y}_s, \overline{\mathbf{Q}}_\ell) \neq 0$

and that $\rho^\vee$ is reducible; *i.e.* ρ is the ℓ-adic representation corresponding to $\chi\,\mathrm{Sp}(2)$. We also find that $\rho^\vee \cap H^1(\mathfrak{Y}_s, \overline{\mathbf{Q}}_\ell)$ is a copy of χ^{-1}; the image of this in $H^1(\tilde{\mathfrak{Y}}, \overline{\mathbf{Q}})$ is 0 because the latter space is already "filled up" by the irreducible WD representations appearing in $H^1(Y_\eta, \overline{\mathbf{Q}}_\ell)$. Finally, $\mathrm{rec}(\mathrm{Sp}(2))$ is the Steinberg representation St of $\mathrm{GL}_2(\mathbf{Q}_p)$ and it is easy to check that $\mathrm{St}^{K(2)} = \mathrm{St}_2$. We conclude that $\mathrm{rec}(\rho)^{K(2)} \otimes \rho^\vee$ appears in $H^1(Y_\eta, \overline{\mathbf{Q}}_\ell)$ with multiplicity $\#S - 1$.

The proof of the theorem is complete as soon as we establish that the number of weight 2 newforms f with $\pi_{f,p}$ Steinberg and with prime-to-p conductor dividing N is $\#S - 1$. This number of newforms is simply the dimension of $S_2(\Gamma_0(p) \cap \Gamma_1(N))^{p-\mathrm{new}}$, and we have

$$\begin{aligned}\dim S_2(\Gamma_0(p) \cap \Gamma_1(N))^{p-\mathrm{new}} &= \dim S_2(\Gamma_0(p) \cap \Gamma_1(N)) - 2 \dim S_2(\Gamma_1(N)) \\ &= g_X - 2g \\ &= \#S - 1\end{aligned}$$

by Equation 4.15. $\square$

Remark 4.4.8. There is no obstacle to generalizing Conjecture 4.4.6 to the modular curves X_n for higher n. In fact, we suspect that one can write down a plausible candidate for the stable reduction of the projective limit $X_\infty = \varprojlim X_n$, and that it will be *simpler* to describe than it would be for any particular X_n!

The long-term hope is that one can prove a form of Conjecture 4.4.6 in all levels and then use an analogue of Theorem 4.4.7 to *deduce* the theorems of Deligne-Carayol (3.2.1 and 4.3.1) in a rather new manner, at least in the case of supercuspidal $\pi_{f,p}$. A promising method for carrying out this program is that of R. Coleman and K. McMurdy, where the structure of a rigid model of a curve is obtained from a special type of covering of the rigid-analytic space associated to its generic fiber. This method has worked successfully in the case of $X_0(Np^3)$, see [CM06]. There, the rigid-analytic space associated to X_n contains certain points (called "fake CM points" in [CM06]) corresponding to formal groups with endomorphisms by a ring of integers in a quadratic extension $\Omega/\mathbf{Q}_p$. The action of inertia on such points is already known by classical Lubin-Tate theory.

Generally, one might hope to find within the rigid-analytic space associated to X_n certain affinoid neighborhoods of the fake CM points whose reductions are the curves $X_{\mathfrak{S}}$ considered appearing in Prop. 4.4.3. In the case of $n = 2$, one hopes the peculiar subgroups $H \subset G_2 \times I_{\mathbf{Q}_p}$ appearing as the stabilizers of these curves in Proposition 4.4.3 might have a natural interpretation in terms of classical Lubin-Tate theory for the fields Ω_i. There are two potential problems with this approach. One is that we are only considering the action of inertia on the special fiber and therefore cannot hope to recover the full theory of Deligne-Carayol, rather only "up to inertia." The other is the troublesome case of $p = 2$, wherein primitive WD representations are admitted. There, one might have to use some sort of base-change technique of the sort employed for primitive representations in [Car83].

Yet more speculative is the possibility of applying these techniques to the case of GL_n. Here, the correct object of study seems to be the formal scheme parametrizing deformations with level structure of a fixed one-dimensional formal group of height $n-1$, as in [HT01]. Now the hope is that a model for this sort of object can be found which is strictly semi-stable; then, an appropriate form of Theorem 4.2.3 could be applied to calculate cohomology. In the best possible scenario, the LLC for GL_n ought to be lurking in this cohomology space and a purely local proof for it might be found which avoids the use of techniques in the theory of global automorphic representations.

Bibliography

[BCDT01] C. Breuil, B. Conrad, F. Diamond, and R. Taylor. On the modularity of elliptic curves over $\mathbf{Q}$: Wild 3-adic exercises. *Journal of the Amer. Math. Soc.*, 14:843–939, 2001.

[BLR90] S. Bosch, W. Lütkebohmert, and M. Raynaud. *Néron models*. 3. Folge, Band 21. Ergebnisse des Mathematik und ihrer Grenzgebiete, 1990.

[Bor05] Niels Borne. Une formule de Riemann-Roch équivariante pour les courbes. *Canad. J. Math.*, 55(4):693–710, 2005.

[BW04] I. Bouw and S. Wewers. Stable reduction of modular curves. In *Modular Curves and abelian varieties*. Birkhauser, 2004.

[Car83] H. Carayol. Sur les représentations ℓ-adiques attachees aux formes modulaires de Hilbert. *C. R. Acad. Sci. Paris.*, 296(15):629–632, 1983.

[Cas73a] W. Casselman. On some results of Atkin and Lehner. *Math. Ann.*, 201:301–314, 1973.

[Cas73b] W. Casselman. The restriction of a representation of $\mathrm{GL}_2(k)$ to $\mathrm{GL}_2(\mathfrak{O})$. *Mathematischen Annalen*, 206(4), 1973.

[CDT99] B. Conrad, F. Diamond, and R. Taylor. Modularity of certain potentially Barsotti-Tate Galois representations. *Journal of the A.M.S.*, 12(2):521–567, April 1999.

[CM06] R. Coleman and K. McMurdy. Fake CM and the stable model of $X_0(Np^3)$. *Documenta Math.*, Extra Volume: John H. Coates' Sixtieth Birthday:261–300, 2006.

[DR73] P. Deligne and M. Rapoport. Les schémas de modules de courbes elliptiques. *Lecture Notes in Mathematics*, 349:143–316, 1973.

[Edi] B. Edixhoven. Modular forms, Galois representations and local Langlands. http://www.crm.es/Publications/Quaderns/Quadern20-2.pdf.

[Edi90] B. Edixhoven. Minimal resolution and stable reduction of $X_0(N)$. *Annales de l'institut Fourier*, 40(1):31–67, 1990.

[Gel75] S. Gelbart. *Automorphic forms on adele groups*. Princeton University Press, 1975.

[Gro72] A. Grothendieck. *Séminaire de Géometrie Algébrique*. Springer-Verlag, 1972.

[HB06] G. Henniart and Bushnell. *The Local Langlands Conjecture for GL(2)*. Springer-Verlag, 2006.

[Hen02] G. Henniart. Sur l'unicité des types pour GL(2). *Duke Math. J.*, 155(2):205–310, 2002.

[HT01] M. Harris and R. Taylor. The geometry and cohomology of some simple Shimura varieties. *Annals of Mathematics Studies*, 151, 2001.

[Joy06] A. Joyce. *The Manin Constants of Modular Abelian Varieties*. PhD thesis, Imperial College, 2006.

[KM85] N. Katz and B. Mazur. *Arithmetic moduli of elliptic curves*. Princeton University Press, 1985.

[KP96] C.S. Khare and D. Prasad. Extending local representations to global representations. *Kyoto J. of Math.*, 36:471–480, 1996.

[Kri96] M. Krir. Degré d'une extension de $\mathbf{Q}_p^{\mathrm{nr}}$ sur laquelle $J_0(N)$ est semi-stable. *Annales de l'institut Fourier*, 2(46):279–291, 1996.

[Kut80] P.C. Kutzko. The Langlands conjecture for GL_2 of a local field. *Annals of Math.*, 112:381–412, 1980.

[Kut84] P.C. Kutzko. The exceptional representations of GL_2. *Compositio Math.*, 51:3–14, 1984.

[Lan70] R. Langlands. Problems in the theory of automorphic forms. *Lectures in modern analysis and applications III, Lectures Notes in Math.*, 170:18–86, 1970.

[MW84] B. Mazur and A. Wiles. Class fields of abelian extensions of $\mathbf{Q}$. *Invent. Math.*, 76:179–330, 1984.

[Pas05] V. Paskunas. Unicity of types for supercuspidal representations. *Proc. London Math. Soc.*, 91(3):623–654, 2005.

[RZ80] M. Rapoport and T. Zink. Über die lokale Zetafunktion von Shimura varietäten, Monodromiefiltrations und verschwindende Zyklen in ungleicher Characteristik. *Inv. Math.*, 68:21–101, 1980.

[Ser67] J.P. Serre. Local class field theory. In *Algebraic Number Theory*. Academic Press, 1967.

[Shi71] G. Shimura. *Introduction to the Arithmetic Theory of Automorphic Forms*. Princeton University Press, 1971.

[Tat79] J. Tate. Number theoretic background. *Proc. Symp. Pure Math.*, 33, part 2:3–26, 1979.

[Yos04] Teruyoshi Yoshida. On non-abelian Lubin-Tate theory via vanishing cycles, 2004.