Serre's Conjecture

Contents

1 Introduction

Let p be a prime and f be a modular form with coefficients in $\overline{\mathbf{F}}_p$. Suppose that f is an eigenform for the Hecke operators. A construction of Deligne (see §1.6) allows us to associate to f a Galois representation

$$\rho_f : G_{\mathbf{Q}} \longrightarrow \mathrm{GL}_2(\overline{\mathbf{F}}_p),$$

where $G_{\mathbf{Q}} = \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ is the absolute Galois group of the rationals.

Two-dimensional mod p Galois representations arise from many different contexts. For example (see §4.1), if $E/\mathbf{Q}$ is an elliptic curve, then the action of $G_{\mathbf{Q}}$ on the p-torsion points $E[p]$ of $E(\overline{\mathbf{Q}})$ gives rise to a representation

$$\overline{\rho}_{E,p} : G_{\mathbf{Q}} \longrightarrow \mathrm{Aut}(E[p]) \cong \mathrm{GL}_2(\mathbf{F}_p).$$

Given an arbitrary two-dimensional mod p Galois representation ρ, does it arise from a modular form? Serre's conjecture provides an answer to this question: if ρ is irreducible and odd (i.e. $\rho(c) = -1$ for any complex conjugation $c \in G_{\mathbf{Q}}$), then ρ is indeed equivalent to ρ_f for some cusp form f.

The strength of the conjecture comes from the fact that as well as predicting the existence of such a cusp form, it comes with a precise recipe for calculating its weight, level and character. Since the spaces of modular forms of given weight, level and character are well understood, the precision of Serre's conjecture allows us to use it to prove some deep results. For example, in §5.1 we show how to use a supposed counterexample to Fermat's last theorem to produce a Galois representation, which by Serre's conjecture should arise from a modular form of weight 2 and level 2. The fact that no such modular forms exist is a contradiction. Hence, Serre's conjecture implies Fermat's last theorem.

The purpose of this essay is to provide a detailed description of Serre's conjecture in its context as a converse to the construction of Deligne. In particular, we will show how the weight, level and character that Serre assigns to a representation can be motivated by looking at the properties of representations which arise from modular forms.

The remainder of this chapter will focus on the technical background to Serre's conjecture. We will begin by defining the topology of $G_{\mathbf{Q}}$ and the structure of its ramification groups; we will then discuss the basic properties of Galois representations. The characterisation of the one-dimensional case in §1.3 will be particularly useful for studying the determinants of two-dimensional representations. We will then provide a brief summary of the results we will require about modular forms and Hecke operators. Finally, at the end of this chapter, we will state the theorem of Deligne mentioned above, and state the properties of representations obtained in this way; these properties will be used to motivate the statement of the weak version of Serre's conjecture.

In chapter two, we will state Serre's conjecture, and justify Serre's definition of the level and character by using the ramification properties and the characteristic equation of representations arising from modular forms. The calculation of the weight is much more involved and is deferred to chapter three. In §3.1 and §3.2, we will explain why the weight depends only on the restriction of the representation ρ to an inertia group at p. In particular, the weight reflects the kind of ramification that ρ has at p, whereas the level reflects its ramification away from p. Sections §3.3-5 will explain results about the structure of the group $G_p = \mathrm{Gal}(\overline{\mathbf{Q}}_p/\mathbf{Q}_p)$ and the restriction $\rho|_{G_p}$, which are then used to give an explicit construction of the weight in §3.6-8.

In chapter four, we will apply the results of the previous chapters to Galois representations arising from elliptic curves, and we will provide some computational examples.

Finally, in chapter five, using the results on elliptic curves from chapter four, we will show how Serre's conjecture can be used to prove Fermat's last theorem and the modularity theorem. We will finish with an example showing the limitations of the conjecture.

Our main source is the paper by Serre [Ser87] detailing his conjecture. The definitions and proofs in the first chapter follow [DS05], with the exception of §1.5, which is taken from [Ser87]. Chapters two, three and the first section of chapter four are derived from [Ser87], as well as from the survey papers [Cai09], [RS99] and Edixhoven's article in [CSS97]. The examples in chapter four were obtained using Cremona's tables featured on his website. The proofs in chapter five are based on those in [Ser87] and [Dar95].

1.1 Absolute Galois groups

Definition and topology

Let K be a perfect field, and let $\overline{K}$ denote the algebraic closure of K.

Definition 1.1. The K-automorphisms of $\overline{K}$ form the *absolute Galois group* of K

$$G_K = \mathrm{Aut}_K(\overline{K}) = \mathrm{Gal}(\overline{K}/K).$$

Observe that as the union of the Galois groups of all finite Galois extensions of K, G_K is a profinite group. Indeed, since $\overline{K}$ is the union of all finite Galois extensions of K, the restriction map

$$\begin{aligned} G_K &\longrightarrow \mathrm{Gal}(L/K) \\ \sigma &\longrightarrow \sigma|_L \end{aligned}$$

is surjective for each finite Galois extension L, and the restriction maps compose over subextensions. Conversely, every compatible system of automorphisms

$$\big\{\sigma|_L : L/K \text{ is finite Galois}\big\}$$

gives an automorphism of $\overline{K}$ - i.e. an element of G_K. Therefore, we can define

$$G_K = \varprojlim_{L} \mathrm{Gal}(L/K).$$

Since each Galois group $\mathrm{Gal}(L/K)$ is finite, G_K is profinite. We can therefore define a natural topology on G_K, the *profinite* or *Krull* topology, to be the coarsest topology for which the restriction maps are continuous (each $\mathrm{Gal}(L/K)$ is given the discrete topology). Explicitly, we take as a basis of open sets all cosets of finite index normal subgroups of G_K. By the fundamental theorem of Galois theory, we can show that if $H \lhd G_{\mathbf{Q}}$ has finite index, then

$$H = \ker(G_{\mathbf{Q}} \longrightarrow \mathrm{Gal}(L/K)),$$

where L is a finite Galois extension of K. Hence, this basis of open sets is exactly the set of cosets of every Galois group $\mathrm{Gal}(\overline{K}/L)$ where L is a finite Galois extension of K.

Since G_K is the inverse limit of finite groups it is compact. This fact will prove useful later in allowing us to show that any representation

$$\rho : G_K \longrightarrow \mathrm{GL}_d(\overline{\mathbf{F}}_p)$$

has finite image.

The structure of $G_{\mathbf{Q}}$

In the case that $K = \mathbf{Q}$, we would like to make use of the number theoretic structure of $\mathbf{Q}$. In particular, we would like to understand the behaviour of maximal ideals $\mathfrak{p} \subseteq \overline{\mathbf{Z}}$ and their interaction with $G_{\mathbf{Q}}$, as well as the structure $\mathbf{Q}$ obtains from its p-adic valuations. Many of the results that are true for finite Galois extensions $L/\mathbf{Q}$ will also hold for the algebraic extension $\overline{\mathbf{Q}}/\mathbf{Q}$.

Let $p \in \mathbf{Z}$ be prime, and let $\mathfrak{p} \subseteq \overline{\mathbf{Z}}$ be a maximal ideal lying over p ($\overline{\mathbf{Z}}$ is the integral closure of $\mathbf{Z}$ in $\overline{\mathbf{Q}}$). We can identify the residue field of $\overline{\mathbf{Q}}$ with $\overline{\mathbf{F}}_p$ by using $\mathfrak{p}$ as the kernel of the map

$$\overline{\mathbf{Z}} \longrightarrow \overline{\mathbf{Z}}/\mathfrak{p} \cong \overline{\mathbf{F}}_p. \tag{1.1}$$

As in the case of a finite extension, define the *decomposition group* of $\mathfrak{p}$ to be

$$D_{\mathfrak{p}} = \left\{\sigma \in G_{\mathbf{Q}} : \sigma(\mathfrak{p}) = \mathfrak{p}\right\},$$

the group of elements of $G_{\mathbf{Q}}$ which fix $\mathfrak{p}$. This group acts on $\overline{\mathbf{Z}}/\mathfrak{p}$ via $\sigma(x+\mathfrak{p}) = \sigma(x)+\mathfrak{p}$, and hence, using (1.1), it can be used to define a reduction map

$$D_{\mathfrak{p}} \longrightarrow \mathrm{Gal}(\overline{\mathbf{F}}_p/\mathbf{F}_p) = G_{\mathbf{F}_p}.$$

The group $G_{\mathbf{F}_p}$ is topologically generated by the Frobenius automorphism

$$\sigma_p : x \longmapsto x^p.$$

We would like to define an *absolute Frobenius element* $\mathrm{Frob}_{\mathfrak{p}}$ to be the pre-image of this automorphism. This will be well defined up to the kernel of the reduction map

$$I_{\mathfrak{p}} = \ker(D_{\mathfrak{p}} \longrightarrow G_{\mathbf{F}_p}),$$

called the *inertia group* of $\mathfrak{p}$. We also define the *higher ramification groups* $G_{\mathfrak{p},i}$ by

$$G_{\mathfrak{p},i} = \ker\left(D_{\mathfrak{p}} \longrightarrow \mathrm{Aut}(\overline{\mathbf{Z}}/\mathfrak{p}^{i+1})\right)$$

for $i \geq -1$. We have $G_{\mathfrak{p},-1} = D_{\mathfrak{p}}$ and $G_{\mathfrak{p},0} = I_{\mathfrak{p}}$.

Whilst the objects $\mathrm{Frob}_{\mathfrak{p}}$ and the groups

$$G_{\mathfrak{p},i} \lhd I_{\mathfrak{p}} \lhd D_{\mathfrak{p}}$$

depend on the choice of ideal $\mathfrak{p}$ lying over p, as in the case of finite extensions, $G_{\mathbf{Q}}$ acts transitively on the set of maximal ideals lying over p, and

$$\begin{aligned} \mathrm{Frob}_{\sigma(\mathfrak{p})} &= \sigma^{-1}\mathrm{Frob}_{\mathfrak{p}}\sigma \qquad \sigma \in G_{\mathbf{Q}}, \\ D_{\sigma(\mathfrak{p})} &= \sigma^{-1}D_{\mathfrak{p}}\sigma. \end{aligned}$$

Hence, these objects are well defined up to conjugation. As such, we will often refer to Frob_p or I_p to denote any member of the conjugacy class.

The following theorem will be extremely useful when studying Galois representations:

Theorem 1.2 (Chebotarev's Density Theorem)**.** *For each maximal ideal $\mathfrak{p} \subseteq \overline{\mathbf{Z}}$ lying over all but a finite set of primes, let* $\mathrm{Frob}_{\mathfrak{p}}$ *be an absolute Frobenius element. The set of such elements forms a dense subset of $G_{\mathbf{Q}}$.*

Since we will be considering Galois representations that are continuous, this theorem will enable us to describe a Galois representation by only considering absolute Frobenius elements.

Let $\mathbf{Q}_p$ be the completion of $\mathbf{Q}$ with respect to the p-adic valuation v_p, and let

$$G_p = \mathrm{Gal}(\overline{\mathbf{Q}}_p/\mathbf{Q}_p).$$

Identifying G_p with $D_{\mathfrak{p}}$, we can form local analogues

$$G_{p,i} \lhd I_p \lhd G_p$$

of the inertia group and the higher ramification groups defined above.

1.2 Galois representations

We are now ready to define a Galois representation.

Definition 1.3. Let K be a field. A *d-dimensional Galois representation* is a homomorphism

$$\rho : G_{\mathbf{Q}} \longrightarrow \mathrm{GL}_d(K).$$

If $\rho' : G_{\mathbf{Q}} \longrightarrow \mathrm{GL}_d(K)$ is another such representation and $m \in \mathrm{GL}_d(K)$ is a matrix such that

$$\rho'(\sigma) = m^{-1}\rho(\sigma)m$$

for all $\sigma \in G_{\mathbf{Q}}$, then we say that ρ and ρ' are *equivalent*, which we denote $\rho \sim \rho'$.

When K is a topological field, we require that ρ be continuous; this enables us to utilise the profinite structure of $G_{\mathbf{Q}}$. In particular:

Lemma 1.4. *Let $K = \overline{\mathbf{F}}_p$ or $\mathbf{C}$, and let*

$$\rho : G_{\mathbf{Q}} \longrightarrow \mathrm{GL}_d(K)$$

be a (continuous) Galois representation. Then ρ has finite image.

Proof. If $K = \overline{\mathbf{F}}_p$, then K, and hence $\mathrm{GL}_d(K)$, has the discrete topology. The group $G_{\mathbf{Q}}$ is compact, so its image $\rho(G_{\mathbf{Q}})$ is both compact and discrete; therefore it is finite.

If $K = \mathbf{C}$, we use the fact that there is an open neighbourhood V of $\mathrm{GL}_d(\mathbf{C})$ which

- contains the identity I,
- contains no non-trivial subgroups.

The pre-image $U = \rho^{-1}(V)$ is an open neighbourhood of $\mathbf{1} \in G_{\mathbf{Q}}$, and therefore contains an open subgroup

$$U(\mathbf{F}) = \ker(G_{\mathbf{Q}} \longrightarrow \mathrm{Gal}(\mathbf{F}/\mathbf{Q}))$$

for some Galois extension $\mathbf{F}/\mathbf{Q}$. The image $\rho(U(\mathbf{F}))$ is a subgroup group of $\mathrm{GL}_d(\mathbf{C})$ contained in V, and hence is trivial. This shows that ρ factors through $\mathrm{Gal}(\mathbf{F}/\mathbf{Q})$. □

Corollary 1.5. *Let $K = \overline{\mathbf{F}}_p$ or $\mathbf{C}$, and let*

$$\rho : G_{\mathbf{Q}} \longrightarrow \mathrm{GL}_d(K)$$

be a Galois representation. Then $\ker \rho$ *is a finite index normal subgroup of $G_{\mathbf{Q}}$, and is therefore open. Hence, there exists a finite Galois extension $\mathbf{F}/\mathbf{Q}$, such that ρ factors through* $\mathrm{Gal}(\mathbf{F}/\mathbf{Q})$.

As a result, with K as above, a Galois representation $\rho : G_{\mathbf{Q}} \longrightarrow \mathrm{GL}_d(K)$ is really just a representation $\mathrm{Gal}(\mathbf{F}/\mathbf{Q}) \longrightarrow \mathrm{GL}_d(K)$ for some finite extension $\mathbf{F}/\mathbf{Q}$. We will often switch between these two viewpoints.

Given a Galois representation ρ, we would like to know the value of $\rho(\sigma)$ for each $\sigma \in G_{\mathbf{Q}}$. By Chebotarev's density theorem, the absolute Frobenius elements $\mathrm{Frob}_{\mathfrak{p}}$ form a dense subset of $G_{\mathbf{Q}}$; we can therefore obtain a lot of information about ρ by evaluating it at these elements. However, an absolute Frobenius element $\mathrm{Frob}_{\mathfrak{p}}$ is only defined up to the inertia group $I_{\mathfrak{p}}$ - i.e.

$$\rho(\mathrm{Frob}_{\mathfrak{p}}) \text{ is well defined if and only if } I_{\mathfrak{p}} \subseteq \ker \rho.$$

This suggests the following definition:

Definition 1.6. Let ρ be a Galois representation and p be a rational prime. If $I_{\mathfrak{p}} \subseteq \ker \rho$ for any maximal ideal $\mathfrak{p} \subseteq \overline{\mathbf{Z}}$ lying over p, then we say that ρ is *unramified* at p.

If $I_{\mathfrak{p}} \subseteq \ker \rho$ for some maximal ideal $\mathfrak{p} \subseteq \overline{\mathbf{Z}}$ lying over p, then this will be true for every maximal ideal lying over p. This is because all inertia groups $I_{\mathfrak{p}}$, where $\mathfrak{p}$ lies over a given prime $p \in \mathbf{Z}$, are conjugate, and $\ker \rho$ is a normal subgroup of $G_{\mathbf{Q}}$.

Remark 1.7. If ρ is unramified at p, then whilst $\rho(\mathrm{Frob}_{\mathfrak{p}})$ will depend on the choice of $\mathfrak{p}$, its characteristic polynomial will depend only on its conjugacy class, and therefore only on p. Hence, it makes sense to talk about the characteristic polynomial of $\rho(\mathrm{Frob}_p)$.

Switching viewpoints, in the case that $K = \overline{\mathbf{F}}_p$ or $\mathbf{C}$, ρ factors through some Galois group $\mathrm{Gal}(\mathbf{F}/\mathbf{Q})$. The representation ρ is unramified at a prime p if and only if $I_{\mathfrak{p}} \subseteq \ker \rho$ for some $\mathfrak{p}$ lying over p, which occurs if and only if

$$I_{\mathfrak{p}}\big|_{\mathbf{F}} = \{1\},$$

i.e. if the finite extension $\mathbf{F}/\mathbf{Q}$ is unramified at p. Hence, saying that ρ is unramified at p is equivalent to saying that the extension $\mathbf{F}/\mathbf{Q}$ is unramified at p.

1.3 Example: one-dimensional Galois representations

Let $K = \overline{\mathbf{F}}_p$ or $\mathbf{C}$. In this section, we will classify the one-dimensional Galois representations

$$\rho : G_{\mathbf{Q}} \longrightarrow \mathrm{GL}_1(K) = K^{\times}$$

by showing that they correspond exactly to primitive Dirichlet characters

$$\chi : (\mathbf{Z}/N\mathbf{Z})^{\times} \longrightarrow K^{\times}.$$

By Corollary 1.5, ρ factors through the Galois group $\mathrm{Gal}(\mathbf{F}/\mathbf{Q})$ of a finite Galois extension $\mathbf{F}/\mathbf{Q}$; since $\mathrm{Gal}(\mathbf{F}/\mathbf{Q}) \leq K^{\times}$, this group is abelian. By the Kronecker-Webber theorem, $\mathbf{F} \subseteq \mathbf{Q}(\mu_N)$ for some integer N; enlarging $\mathbf{F}$ we can take $\mathbf{F} = \mathbf{Q}(\mu_N)$.

Since $\mathrm{Gal}(\mathbf{Q}(\mu_N)/\mathbf{Q}) \cong (\mathbf{Z}/N\mathbf{Z})^\times$, the following diagram commutes:

$$\begin{tikzcd}
G_{\mathbf{Q}} \arrow[rr, "\rho"] \arrow[dr, "\pi_N"'] & & K^\times \\
& \mathrm{Gal}(\mathbf{Q}(\mu_N)/\mathbf{Q}) \arrow[r, "\sim"] & (\mathbf{Z}/N\mathbf{Z})^\times \arrow[u, "\chi"']
\end{tikzcd}$$

and we obtain a character $\chi : (\mathbf{Z}/N\mathbf{Z})^\times \longrightarrow K^\times$. Taking N to be minimal, we obtain a primitive character χ.

Conversely, if $\chi : (\mathbf{Z}/N\mathbf{Z})^\times \longrightarrow K^\times$ is a primitive character, let $\rho_{\chi,N}$ be the unique homomorphism making the following diagram commute:

$$\begin{tikzcd}
G_{\mathbf{Q}} \arrow[rr, "\rho_\chi"] \arrow[dr, "\pi_N"'] & & K^\times \\
& \mathrm{Gal}(\mathbf{Q}(\mu_N)/\mathbf{Q}) \arrow[r, "\sim"] \arrow[ur, "\rho_{\chi,N}"] & (\mathbf{Z}/N\mathbf{Z})^\times \arrow[u, "\chi"']
\end{tikzcd}$$

We can show that $\rho_{\chi,N}$ is continuous. Hence, χ determines a unique homomorphism

$$\rho_\chi = \rho_{\chi,N} \circ \pi_N : G_{\mathbf{Q}} \longrightarrow K^\times.$$

We deduce that the one-dimensional Galois representations with $K = \overline{\mathbf{F}}_p$ and $\mathbf{C}$ correspond exactly to primitive characters.

Now let p be a prime not dividing N. Since the extension $\mathbf{Q}(\mu_N)/\mathbf{Q}$ is unramified for such p, so is ρ_χ. The facts that the isomorphism $\mathrm{Gal}(\mathbf{Q}(\mu_N)/\mathbf{Q}) \longrightarrow (\mathbf{Z}/N\mathbf{Z})^\times$ maps $\mathrm{Frob}_p \longmapsto p$, and that for any maximal ideal $\mathfrak{p} \subseteq \overline{\mathbf{Z}}$ lying over p, we have

$$\mathrm{Frob}_{\mathfrak{p}}\big|_{\mathbf{Q}(\mu_N)} = \mathrm{Frob}_{\mathfrak{p}\cap\mathbf{Q}(\mu_N)},$$

imply that

$$\rho_\chi(\mathrm{Frob}_{\mathfrak{p}}) = \chi(p). \tag{1.2}$$

1.4 Modular forms and Hecke operators

Modular forms

The *modular group* $\mathrm{SL}_2(\mathbf{Z})$ is the group of 2×2 invertible matrices with integer entries. For each positive integer N, define

$$\Gamma_0(N) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbf{Z}) : \begin{pmatrix} a & b \\ c & d \end{pmatrix} \equiv \begin{pmatrix} * & * \\ 0 & * \end{pmatrix} \pmod N \right\}$$

and

$$\Gamma_1(N) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbf{Z}) : \begin{pmatrix} a & b \\ c & d \end{pmatrix} \equiv \begin{pmatrix} 1 & * \\ 0 & 1 \end{pmatrix} \pmod N \right\}.$$

Let $M_k(N) = M_k(\Gamma_1(N))$ be the space of modular forms of weight k and level $\Gamma_1(N)$. and let $\epsilon : (\mathbf{Z}/N\mathbf{Z})^\times \longrightarrow \mathbf{C}^\times$ be a Dirichlet character. The space $M_k(N)$ admits a decomposition into subspaces $M_k(N,\epsilon)$ which are stable under the Hecke operators. A modular form $f \in M_k(N,\epsilon)$ is defined as follows:

Definition 1.8. A *modular form* of weight k, level N and character ϵ is a modular form $f \in M_k(\Gamma_1(N))$ such that for all $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_0(N)$,

$$f(\gamma\tau) = f\left(\frac{a\tau + b}{c\tau + d}\right) = \epsilon(d)(c\tau + d)^k f(\tau)$$

We denote the space of such forms $M_k(N, \epsilon)$. This is a finite-dimensional complex vector space and

$$M_k(N) = \bigoplus_{\epsilon:(\mathbf{Z}/N\mathbf{Z})^\times \to \mathbf{C}^\times} M_k(N, \epsilon). \tag{1.3}$$

The subspace $S_k(N) = S_k(\Gamma_1(N))$ of cusp forms decomposes in the same way:

$$S_k(N) = \bigoplus_{\epsilon:(\mathbf{Z}/N\mathbf{Z})^\times \to \mathbf{C}^\times} S_k(N, \epsilon).$$

Observe that since $\gamma = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \in \Gamma_0(N)$, if $f \in M_k(N, \epsilon)$ and $\tau \in \mathbf{C}$, then

$$f(\tau) = f(\gamma\tau) = \epsilon(-1)(-1)^k f(\tau).$$

It follows that $M_k(N, \epsilon)$ is zero-dimensional unless

$$\epsilon(-1) = (-1)^k. \tag{1.4}$$

Hecke operators

The space $M_k(N)$ is acted on by the *Hecke operators*; these operators preserve the character decomposition (1.3) of $M_k(N)$:

- The *diamond operator*: for each $d \in (\mathbf{Z}/N\mathbf{Z})^\times$, define

  $$\langle d\rangle f(\tau) = (c\tau + \delta)^{-k} f\left(\frac{a\tau + b}{c\tau + \delta}\right),$$

  where $\begin{pmatrix} a & b \\ c & \delta \end{pmatrix} \in \Gamma_0(N)$ and $\delta \equiv d \pmod{N}$.

  Noting that

  $$M_k(N, \epsilon) = \left\{ f \in M_K(N) : \langle d\rangle f = \epsilon(d) f \ \forall d \in (\mathbf{Z}/N\mathbf{Z})^\times \right\},$$

  we see that the diamond operator respects the character decomposition.

- The n^{th} *Hecke operator* is an endomorphism T_n of $M_k(N)$. If ℓ is a prime number then T_ℓ is defined on the q-expansion of a modular form $f \in M_k(N)$ by:

  $$T_\ell : \sum_{n=0}^{\infty} a_n(f) q^n \longmapsto \begin{cases} \displaystyle\sum_{n=0}^{\infty} a_{n\ell}(f) q^n + \ell^{k-1} \sum_{n=0}^{\infty} a_n(\langle \ell\rangle f) q^{n\ell} & \text{if } \ell \nmid N \\ \displaystyle\sum_{n=0}^{\infty} a_{n\ell}(f) q^n & \text{if } \ell \mid N \end{cases}$$

The operator T_n is defined on prime powers by

$$T_{\ell^r} = T_\ell T_{\ell^{r-1}} - \ell^{k-1}\langle \ell \rangle T_{\ell^{r-2}}$$

for integers $r \geq 2$, and is then defined multiplicatively so that T_n and T_m commute if n and m are coprime.

If $f \in M_k(N, \epsilon)$, then for ℓ prime,

$$T_\ell : \sum_{n=0}^{\infty} a_n(f) q^n \longmapsto \begin{cases} \displaystyle\sum_{n=0}^{\infty} a_{n\ell}(f) q^n + \epsilon(\ell)\ell^{k-1} \sum_{n=0}^{\infty} a_n(f) q^{n\ell} & \text{if } \ell \nmid N \\ \displaystyle\sum_{n=0}^{\infty} a_{n\ell}(f) q^n & \text{if } \ell \mid N \end{cases}$$

and we see that $M_k(N, \epsilon)$ is indeed stable under the action of T_n.

Eigenforms

Definition 1.9. A non-zero modular form $f = \sum_{n=0}^{\infty} a_n(f) q^n \in M_k(N)$ is called an *eigenform* if it is a simultaneous eigenvector for all the Hecke operators. If $a_1 = 1$, we say that f is *normalised.*

Let $f = \sum_{n=0}^{\infty} c_n q^n$ be a normalised eigenform. Using the above formulae, we see that for each $n \in \mathbf{N}$,

$$a_1(T_n(f)) = c_n = c_n \cdot c_1,$$

i.e. the eigenvalue of T_n is exactly the Foruier coefficient c_n. Hence, a normalised eigenform is completely determined by its eigenvalues. In fact, we need only know the eigenvalues of T_p for p prime, since the formal Dirichlet series

$$L(s, f) = \sum_{n=1}^{\infty} c_n n^{-s}$$

is given by an Euler product

$$L(s, f) = \prod_p \left(1 - c_p p^{-s} + \epsilon(p) p^{k-1} p^{-2s}\right)^{-1}.$$

Since f is an eigenform, the action of the diamond bracket operators defines a character

$$\begin{aligned} \epsilon : (\mathbf{Z}/N\mathbf{Z})^\times &\longrightarrow \mathbf{C}^\times \\ d &\longmapsto \frac{\langle d \rangle f}{f}. \end{aligned}$$

We deduce that if f is an eigenform, then $f \in M_k(N, \epsilon)$.

Example 1.10. The *Ramanujan delta function*

$$\Delta(z) = q \prod_{n \geq 1} (1 - q^n)^{24} = \sum_{n \geq 1} \tau(n) q^n$$

is a normalised cusp form of weight 12 and level 1. Since the space $S_{12}(1)$ is one-dimensional and is preserved by the Hecke operators, Δ must be a Hecke eigenform with eigenvalues $\tau(n)$.

1.5 Modular forms in characteristic p

In this section, we will define the notion of a modular form with coefficients in $\overline{\mathbf{F}}_p$ using Serre's definition in [Ser87].

Let p be a prime number, and let

- $N \geq 1$ be a positive integer such that $p \nmid N$,
- $k \geq 2$ be a positive integer and
- $\epsilon : (\mathbf{Z}/N\mathbf{Z})^\times \longrightarrow \overline{\mathbf{F}}_p^\times$ be a character such that $\epsilon(-1) = (-1)^k$ (so that the space $M_k(N, \epsilon)$ is not automatically empty).

We define a modular form of type (N, k, ϵ) with coefficients in $\overline{\mathbf{F}}_p$ to be the reduction modulo p of a modular form $f \in M_k(N, \epsilon)$ with coefficients in the algebraic integers.

More precisely, choose a place of $\overline{\mathbf{Q}}$ above p, and let $\overline{\mathbf{Z}}_p$ denote the ring of integers of $\overline{\mathbf{Q}}_p$. The choice of place determines an embedding $\overline{\mathbf{Z}} \hookrightarrow \overline{\mathbf{Z}}_p$. The ring $\overline{\mathbf{Z}}_p$ has a natural reduction map to its residue field $\overline{\mathbf{F}}_p$, so we obtain a homomorphism

$$\begin{aligned} \overline{\mathbf{Z}} &\longrightarrow \overline{\mathbf{F}}_p \\ z &\longmapsto \tilde{z}. \end{aligned} \tag{1.5}$$

The character ϵ has finite image lying in a finite field $\mathbf{F}_q$. Let

$$\epsilon_0 : (\mathbf{Z}/N\mathbf{Z})^\times \longrightarrow \overline{\mathbf{Z}}_p^\times$$

denote the Teichmuller lift of ϵ, the unique character ϵ_0 such that for all $x \in (\mathbf{Z}/N\mathbf{Z})^\times$

$$\epsilon(x) = \epsilon_0(x) \pmod{\mathbf{F}_q}$$

and

$$\epsilon_0(x)^q = \epsilon_0(x).$$

Since ϵ_0 takes values in the N^{th} roots of unity, the image of ϵ_0 lies in $\overline{\mathbf{Z}}^\times \subseteq \mathbf{C}^\times$.

Definition 1.11. A *modular form* of type (N, k, ϵ) with coefficients in $\overline{\mathbf{F}}_p$ is a formal power series

$$f = \sum_{n=0}^{\infty} a_n q^n \quad a_n \in \overline{\mathbf{F}}_p$$

for which there exists a modular form

$$F = \sum_{n=0}^{\infty} A_n q^n, \quad A_n \in \overline{\mathbf{Z}}$$

in $M_k(N, \epsilon_0)$ such that $\widetilde{A}_n = a_n$ for all n. A cusp form of type (N, k, ϵ) with coefficients in $\overline{\mathbf{F}}_p$ is defined analogously.

We denote the space of modular forms of type (N, k, ϵ) with coefficients in $\overline{\mathbf{F}}_p$ by $\widetilde{M_k}(N, \epsilon)$ and the subspace of cusp forms by $\widetilde{S}_k(N, \epsilon)$. These spaces have the following properties (see [Ser87] §3.1):

1. The dimension of $\widetilde{S}_k(N, \epsilon)$ as a vector space over $\overline{\mathbf{F}}_p$ is equal to the dimension of the corresponding space $S_k(N, \epsilon_0)$ as a vector space over $\mathbf{C}$.

2. The spaces $\widetilde{M}_k(N,\epsilon)$ and $\widetilde{S}_k(N,\epsilon)$ are stable under the action of the Hecke operators T_ℓ for prime ℓ, defined by

$$T_\ell : \sum_{n=0}^{\infty} a_n q^n \longmapsto \begin{cases} \displaystyle\sum_{n=0}^{\infty} a_{n\ell} q^n + \epsilon(\ell)\ell^{k-1} \sum_{n=0}^{\infty} a_n q^{n\ell} & \text{if } \ell \nmid pN \\ \displaystyle\sum_{n=0}^{\infty} a_{n\ell} a^n & \text{if } \ell \mid pN \end{cases} \tag{1.6}$$

Moreover, for n, m coprime, the Hecke operators T_n, T_m commute.

For $\ell \neq p$, this follows from the similar properties in characteristic zero. For $\ell = p$, observe that T_p is the reduction mod p of the characteristic zero Hecke operator

$$T_p : \sum_{n=0}^{\infty} a_n q^n \longmapsto \sum_{n=0}^{\infty} a_{pn} q^n + \epsilon_0(p) p^{k-1} \sum_{n=0}^{\infty} a_n q^{pn}$$

thanks to the hypothesis $k \geq 2$.

3. If

$$f = \sum_{n=1}^{\infty} a_n q^n \in \widetilde{S}_k(N,\epsilon)$$

is a non-zero normalised Hecke eigenform, then by definition, f is the reduction mod p of some normalised cusp form

$$F = \sum_{n=1}^{\infty} A_n q^n$$

of type (N, k, ϵ_0) with coefficients in $\overline{\mathbf{Z}}$. This F will also be a Hecke eigenform, with coefficients satisfying

$$\widetilde{A}_\ell = a_\ell$$

for any prime number ℓ.

1.6 Galois representations arising from mod p modular forms

Let $f = \sum_{n=0}^{\infty} a_n q^n$ be an eigenform with coefficients in $\overline{\mathbf{F}}_p$. By a construction of Deligne ([DS74], Theorem 6.7), we can associate to f a mod p Galois representation ρ_f. Specifically:

Theorem 1.12 (Deligne)**.** *Let N be an integer, p be a prime not dividing N, and let f be a non-zero normalised eigenform of type (N, k, ϵ) with coefficients in $\overline{\mathbf{F}}_p$. Then there exists a continuous, semisimple Galois representation*

$$\rho_f : G_{\mathbf{Q}} \to \mathrm{GL}_2(\overline{\mathbf{F}}_p)$$

such that for any prime $\ell \nmid pN$

- *ρ_f is unramified at ℓ*
- *$\rho_f(\mathrm{Frob}_\ell)$ has characteristic polynomial*

$$X^2 - a_\ell X + \epsilon(\ell)\ell^{k-1}. \tag{1.7}$$

Observe that since ρ_f is unramified at ℓ, Remark 1.7 shows that the characteristic polynomial of $\rho_f(\mathrm{Frob}_\ell)$ is well defined. We can state formula (1.7) equivalently as

$$\det \rho_f(\mathrm{Frob}_\ell) = \epsilon(\ell)\ell^{k-1} \quad \text{and} \quad \mathrm{tr}\ \rho_f(\mathrm{Frob}_\ell) = a_\ell. \tag{1.8}$$

Remarks 1.13.

1. Let $\chi_p : G_{\mathbf{Q}} \longrightarrow (\mathbf{Z}/p\mathbf{Z})^\times$ be the p^{th} *cyclotomic character*, obtained by the action of $G_{\mathbf{Q}}$ on the p^{th} roots of unity - i.e. the character such that for every $\sigma \in G_{\mathbf{Q}}$, we have
$$\sigma(\zeta_p) = \zeta_p^{\chi(\sigma)}$$
where ζ_p is a primitive p^{th} root of unity. Recall equation (1.2), that if ρ_χ is the representation induced by χ, then
$$\rho_\chi(\mathrm{Frob}_p) = \chi(p).$$
By Chebotarev's density theorem and the continuity of ρ_f, it follows that (1.8) can be rewritten as
$$\det \rho_f = \epsilon\chi_p^{k-1} \quad \text{and} \quad \mathrm{tr}\ \rho_f(\mathrm{Frob}_\ell) = a_\ell \tag{1.9}$$
Here, we are viewing ϵ as a character $G_{\mathbf{Q}} \longrightarrow \overline{\mathbf{F}}_p^\times$ obtained by lifting
$\epsilon : (\mathbf{Z}/N\mathbf{Z})^\times \longrightarrow \overline{\mathbf{F}}_p^\times$ to $G_{\mathbf{Q}}$.

2. Let $c \in G_{\mathbf{Q}}$ be any complex conjugation. Then by (1.4)
$$\det \rho_f(c) = \epsilon(-1)(-1)^{k-1} = (-1)^k(-1)^{k-1} = -1.$$
A representation with this property is called an *odd* representation.

3. The representation ρ_f need not be irreducible - indeed, if $f \in \widetilde{M}_k(N,\epsilon)$ is a non-cuspidal eigenform then ρ_f must be reducible. For example, if $N = 1$ and $k > 2$, then f is the reduction mod p of an Eisenstein series
$$E_k(\tau) = \frac{B_k}{2k} - \sum_{n=1}^{\infty} \sigma_{k-1}(n)q^n,$$
where
$$\sigma_{k-1}(n) = \sum_{\substack{m|n \\ m>0}} m^{k-1}$$
and B_k is the k^{th} Bernoulli number. E_k is a Hecke eigenform, so we only need to calculate σ_{k-1} at primes. Evaluating at a prime ℓ gives
$$a_\ell = \sigma_{k-1}(\ell) = 1 + \ell^{k-1}$$
Let ρ_f be the mod p representation corresponding to f. If $\ell \neq p$ is prime, then ρ_f is unramified at ℓ, and $\rho_f(\mathrm{Frob}_\ell)$ has characteristic polynomial
$$X^2 - (1 + \ell^{k-1})X + \ell^{k-1}.$$
We deduce that
$$\rho_f = \begin{pmatrix} 1 & 0 \\ 0 & \chi_p^{k-1} \end{pmatrix} = \mathbf{1} \oplus \chi_p^{k-1}$$
is reducible. The proof for the general case is similar (see [DS05] Theorem 9.6.6).

2 Serre's conjecture

We stated in §1.6 that if $f \in \widetilde{M}_k(N,\epsilon)$ is a eigenform with coefficients in $\overline{\mathbf{F}}_p$, then we can associate to f a Galois representation

$$\rho_f : G_{\mathbf{Q}} \longrightarrow \mathrm{GL}_2(\overline{\mathbf{F}}_p).$$

Suppose now that we are given a Galois representation

$$\rho : G_{\mathbf{Q}} \longrightarrow \mathrm{GL}_2(\overline{\mathbf{F}}_p).$$

We would like to investigate the converse of Deligne's theorem: does ρ arise from a modular form - i.e. is $\rho \sim \rho_f$ for some eigenform f?

The remarks at the end of the previous section show that in order for this to be possible ρ must certainly be odd (i.e. $\det \rho(c) = -1$ where c is any complex conjugation). Moreover, restricting our attention to cuspidal eigenforms, we can assume that ρ is irreducible. The weak version of Serre's conjecture states that in this situation, ρ is indeed modular:

Theorem 2.1 (Serre's Conjecture, Weak Version)**.** *Let*

$$\rho : G_{\mathbf{Q}} \longrightarrow \mathrm{GL}(V) \cong \mathrm{GL}_2(\overline{\mathbf{F}}_p)$$

be a Galois representation, V being a two-dimensional vector space over $\overline{\mathbf{F}}_p$. Suppose that ρ is irreducible and odd. Then there exists a cuspidal eigenform f with coefficients in $\overline{\mathbf{F}}_p$ whose associated representation ρ_f is equivalent to ρ.

Beyond conjecturing the existence of such a modular form, Serre gives a precise recipe for its level $N(\rho)$, weight $k(\rho)$ and character $\epsilon(\rho)$. This chapter will explain Serre's recipe for the level and character of the representation; the next chapter will explain the construction of the weight.

2.1 The level $N(\rho)$

In this section, we will define the level $N(\rho)$ of the representation ρ to be the prime-to-p part of the *global Artin conductor* of ρ. The results of Carayol ([Car89]) and Livné ([Liv89]) show that this value is optimal, in the sense that if ρ arises from a cuspidal eigenform $f \in \widetilde{S}_k(N,\epsilon)$, then $N(\rho) \mid N$.

Deligne's theorem states that if $\rho_f : G_{\mathbf{Q}} \longrightarrow \mathrm{GL}_2(\overline{\mathbf{F}}_p)$ is the Galois representation corresponding to an eigenform $f \in \widetilde{S}_k(N,\epsilon)$ where $(N,p)=1$, then f is unramified at every prime $\ell \nmid pN$. The level $N(\rho)$ must therefore depend on the ramification of ρ at primes $\ell \neq p$, and should be of the form

$$N(\rho) = \prod_{\ell \neq p \text{ prime}} \ell^{n(\ell,\rho)}, \tag{2.1}$$

where $n(\ell,\rho) \in \mathbf{Z}$ and

$$\begin{aligned} n(\ell,\rho) &= 0 \qquad \text{if } \rho \text{ is unramified at } \ell, \\ n(\ell,\rho) &> 0 \qquad \text{if } \rho \text{ is ramified at } \ell. \end{aligned}$$

Thus, it makes sense to consider the local representation

$$\rho_\ell : G_\ell = \mathrm{Gal}(\overline{\mathbf{Q}}_\ell/\mathbf{Q}_\ell) \longrightarrow \mathrm{GL}_2(V)$$

formed by restricting ρ to G_ℓ. The group G_ℓ comes with a filtration

$$G_\ell = G_{\ell,-1} \supset G_{\ell,0} \supset G_{\ell,1} \supset \cdots$$

of higher ramification groups. It is natural to consider the action of each $G_{\ell,i}$ on V. In particular, let

$$V^{G_{\ell,i}} = \left\{ v \in V : \rho(\sigma)v = v \quad \forall \sigma \in G_{\ell,i} \right\}$$

be the subspace of V of elements invariant under the action of $G_{\ell,i}$. For example, $G_{\ell,0} = I_\ell$ is the inertia group at ℓ and by definition,

$$\rho \text{ is unramified at } \ell \iff V = V^{I_\ell} \tag{2.2}$$

Similarly,

$$G_{\ell,i} \subseteq \ker \rho \iff V = V^{G_{\ell,i}}$$

Hence, the codimension of $V^{G_{\ell,i}}$ in V is a good measure of the level of higher ramification of ρ.

We define $n(\ell,\rho)$ to be the *local Artin conductor*

$$n(\ell,\rho) = \sum_{i=0}^{\infty} \frac{1}{[G_{\ell,0}, G_{\ell,i}]} \dim(V/V^{G_{\ell,i}})$$

or equivalently as

$$n(\ell,\rho) = \dim(V/V^{I_\ell}) + b(V)$$

where $b(V)$ is the *Swan conductor* or the *wild invariant* as defined in [Ser77]. Serre shows in [Ser79] chapter VI that $n(\ell,\rho)$ is a non-negative integer. Observe that:

1. $n(\ell,\rho) = 0$ if and only if $V = V^{G_{\ell,0}}$. By (2.2), this occurs if and only if ρ is unramified at ℓ.

2. Similarly, $n(\ell,\rho) = 1$ if and only if ρ is tamely ramified at ℓ.

2.2 The character $\epsilon(\rho)$

Deligne's theorem states that if $f \in \widetilde{S}_k(N,\epsilon)$ is an eigenform with corresponding Galois representation ρ_f, then

$$\det \rho_f = \epsilon \chi_p^{k-1}, \tag{2.3}$$

where χ_p is the p^{th} cyclotomic character. Since ϵ and χ_p are characters of level N and p respectively, we can view $\det \rho_f$ as a character of level pN. By the Chinese remainder theorem, the group $(\mathbf{Z}/pN\mathbf{Z})^\times$ is canonically isomorphic to $(\mathbf{Z}/p\mathbf{Z})^\times \times (\mathbf{Z}/N\mathbf{Z})^\times$. Under this decomposition, the restriction of $\det \rho_f$ to $(\mathbf{Z}/p\mathbf{Z})^\times$ is χ_p^{k-1}, and the restriction to $(\mathbf{Z}/N\mathbf{Z})^\times$ is ϵ. This is shown in the following diagram:

$$\begin{tikzcd}
 & & (\mathbf{Z}/N\mathbf{Z})^\times \arrow[rd, "\epsilon"] & \\
G_{\mathbf{Q}} \arrow[r, "\pi_{pN}"] & (\mathbf{Z}/pN\mathbf{Z})^\times \arrow[ru] \arrow[rr] \arrow[rd] & & \overline{\mathbf{F}}_p^\times \\
 & & (\mathbf{Z}/p\mathbf{Z})^\times \arrow[ru, "\chi_p^{k-1}"'] &
\end{tikzcd}$$

Now let us return to our representation ρ. We will extract the character $\epsilon(\rho)$ from $\det\rho$ in an analogous way. By the results of §1.3, the one-dimensional representation

$$\det\rho : G_{\mathbf{Q}} \longrightarrow \overline{\mathbf{F}}_p$$

corresponds to a character

$$\chi : \mathbf{Z}/M\mathbf{Z} \longrightarrow \overline{\mathbf{F}}_p$$

for some integer $M \geq 1$, such that $\det\rho$ is unramified for all primes $\ell \nmid M$. Assuming that ρ arises from a modular form of level $N(\rho)$, where $N(\rho)$ is as in the previous section, we see that ρ is unramified away from $pN(\rho)$. It follows that $\det\rho$ can be considered as a character of $(\mathbf{Z}/p^n\mathbf{Z})^\times \times (\mathbf{Z}/N(\rho)^m\mathbf{Z})^\times$ for some $n, m \geq 1$. We show that we can take $n, m = 1$:

- Since
$$(\mathbf{Z}/p^n\mathbf{Z})^\times \cong \mathbf{Z}/p^{n-1}\mathbf{Z} \times \mathbf{Z}/(p-1)\mathbf{Z}$$
and χ takes values in $\overline{\mathbf{F}}_p^\times$, which contains no p-power roots of unity, we can take $n = 1$.

- Comparing formula (2.1) for ρ and $\det\rho$, we see that $N(\det\rho) \mid N(\rho)$; the relationship of the conductor $N(\det\rho)$ with the class field theoretic conductor (see [AT68] Theorem XI.14) then shows that we can also take $m = 1$.

We therefore obtain a character

$$\chi : (\mathbf{Z}/pN(\rho)\mathbf{Z})^\times \cong (\mathbf{Z}/p\mathbf{Z})^\times \times (\mathbf{Z}/N(\rho)\mathbf{Z})^\times \longrightarrow \overline{\mathbf{F}}_p^\times,$$

and as before we define, by the restriction of χ to $(\mathbf{Z}/N(\rho)\mathbf{Z})^\times$ and $(\mathbf{Z}/p\mathbf{Z})^\times$ respectively, the character

$$\epsilon(\rho) : (\mathbf{Z}/N(\rho)\mathbf{Z})^\times \longrightarrow \overline{\mathbf{F}}_p^\times \tag{2.4}$$

and a corresponding homomorphism

$$\varphi : (\mathbf{Z}/p\mathbf{Z})^\times \longrightarrow \overline{\mathbf{F}}_p^\times.$$

Since $(\mathbf{Z}/p\mathbf{Z})^\times$ is a cyclic group of order $p-1$, viewing $(\mathbf{Z}/p\mathbf{Z})^\times$ as a subgroup of $\overline{\mathbf{F}}_p^\times$, the homomorphism φ is of the form

$$x \longmapsto x^h \qquad h \in \mathbf{Z}/(p-1)\mathbf{Z}.$$

We will see in §3.4 that the homomorphisms $G_{\mathbf{Q}} \longrightarrow \overline{\mathbf{F}}_p$ which factor through $\mathbf{F}_p$ form a cyclic group generated by the cyclotomic character χ_p. Hence, we can write

$$\varphi = \chi_p^h,$$

and we deduce that

$$\det\rho = \epsilon\chi_p^h.$$

Comparing this to (2.3), we see that h should be related to $k(\rho)$ by the congruence

$$h \equiv k(\rho) - 1 \pmod{p-1}. \tag{2.5}$$

3 The weight

3.1 Motivation

In this chapter, we will define the weight $k(\rho)$ attached to the representation ρ, following Serre's definitions. At the end of the previous chapter, we showed that the determinant $\det \rho$ is given by a product $\epsilon\chi_p^h$, where ϵ is a character of level $N(\rho)$. Since $p \nmid N(\rho)$, ϵ is unramified at p, and restricting to an inertia group I_p lying over p we have

$$\det \rho|_{I_p} = \chi_p^h$$

Comparing this to equation (2.3), we see that if ρ arises from a modular form of weight $k(\rho)$, then h must satisfy

$$h \equiv k(\rho) - 1 \pmod{p-1}.$$

In particular, if we knew that $2 \leq k(\rho) \leq p+1$, then up to some ambiguity with the cases $k(\rho) = 2$ and $k(\rho) = p+1$ (see §3.8), we would be able determine the weight directly from $\det \rho|_{I_p}$.

Suppose that $f = \sum_{n=1}^{\infty} a_n q^n$ is a normalised cuspidal eigenform of type (N, k, ϵ) with coefficients in $\overline{\mathbf{F}}_p$, and that $2 \leq k \leq p+1$. Let

- ρ_f be the mod p Galois representation associated to f,
- $\rho_{f,p} = \rho|_{G_p}$ be its restriction to G_p,
- $\lambda(a)$ denote the unramified character $G_p \longrightarrow \overline{\mathbf{F}}_p^\times$ such that $\lambda(\mathrm{Frob}_p) = a$ for any $a \in \overline{\mathbf{F}}_p^\times$.

The following two theorems (see [Edi92] §2.4) tell us the form that $\rho_f|_{I_p}$ should take:

Theorem 3.1 (Deligne). *Suppose that $a_p \neq 0$. Then $\rho_{f,p}$ is reducible and*

$$\rho_{f,p} = \begin{pmatrix} \chi_p^{k-1}\lambda(\epsilon(p)/a_p) & * \\ 0 & \lambda(a_p) \end{pmatrix}.$$

In this case, we have

$$\rho_f|_{I_p} = \begin{pmatrix} \chi_p^{k-1} & * \\ 0 & 1 \end{pmatrix}.$$

Theorem 3.2 (Fontaine). *Suppose that $a_p = 0$. Then $\rho_{f,p}$ is irreducible and*

$$\rho_f|_{I_p} = \begin{pmatrix} \psi'^{k-1} & 0 \\ 0 & \psi^{k-1} \end{pmatrix}.$$

where ψ, ψ' are the two fundamental characters of level 2 (see §3.4).

If ρ takes one of the above forms, then we set $k(\rho)$ to be k. In §3.2, we will show that for every representation ρ, there is a twist of ρ by a power of the cyclotomic character that takes one of these forms. In particular, the weight depends only on the restriction of ρ to an inertia subgroup I_p lying over p. We see that whilst the level depends on the ramification of ρ away from p, the weight depends on the ramification of ρ at p.

3.2 Twists of representations

In [Kat76], Katz shows that there is a derivation

$$\theta : \widetilde{M}_k(N,\epsilon) \longrightarrow \widetilde{M}_{k+p+1}(N,\epsilon),$$

whose action on q-expansions is given by

$$q\frac{d}{dq} : \sum_{n=0}^{\infty} a_n q^n \longmapsto \sum_{n=0}^{\infty} n a_n q^n.$$

This map preserves the subspace of cusp forms. Moreover, by computing the action of the Hecke operators using (1.6), we observe that for all primes ℓ

$$T_\ell(\theta f) = \ell\theta(T_\ell(f)). \tag{3.1}$$

As a result, if f is a normalised eigenform of type (N,k,ϵ) with eigenvalues a_ℓ, then θf is an eigenform of type $(N, k+p+1, \epsilon)$ with eigenvalues ℓa_ℓ.

Let $f \in \widetilde{S}_k(N,\epsilon)$ be a normalised eigenform, and let ρ_f and $\rho_{\theta f}$ be the mod p Galois representations corresponding to f and θf. Let $\ell \nmid pN$ be prime, and let Frob_ℓ be an absolute Frobenius element corresponding to ℓ. Equation (1.8) shows that $\rho_{\theta f}(\mathrm{Frob}_\ell)$ has characteristic polynomial

$$X^2 - \ell a_\ell X + \epsilon(\ell)\ell^{k+p}.$$

Similarly, $\chi_p \otimes \rho_f(\mathrm{Frob}_\ell)$ has characteristic polynomial

$$X^2 - \chi_p(\mathrm{Frob}_\ell)a_\ell X + \epsilon(\ell)\chi_p^2(\mathrm{Frob}_\ell)\ell^{k-1},$$

and the fact that the cyclotomic character χ_p satisfies

$$\chi_p(\mathrm{Frob}_\ell) = \ell \pmod p \quad \ell \neq p$$

shows that these polynomials are equal mod p. By Chebotarev's density theorem, this means that $\rho_{\theta f}$ and $\chi_p \otimes \rho_f$ have the same characteristic polynomial. It follows by the Brauer-Nesbitt theorem, which states that semisimple representations are determined by their characteristic polynomials, that

$$\rho_{\theta f} \sim \chi_p \otimes \rho_f.$$

We call the representation $\chi_p \otimes \rho_f$ a *twist* of ρ_f by χ_p. The usefulness of these twists comes from the following theorem (see [Edi92] Theorem 3.4):

Theorem 3.3. *Let f be an eigenform of type (N,k,ϵ) with coefficients in $\overline{\mathbf{F}}_p$. Then there exist integers i and k' with*

$$0 \le i \le p-1 \text{ and } 2 \le k' \le p+1$$

and an eigenform g of type (N,k',ϵ), such that f and $\theta^i g$ have the same eigenvalues for all the operators T_ℓ $(\ell \neq p)$.

In particular, we have

$$\rho_f \sim \rho_{\theta^i g} \sim \chi_p^i \otimes \rho_g.$$

3.3 The structure of $\mathrm{Gal}(\overline{\mathbf{Q}}_p/\mathbf{Q}_p)$

In the previous section, we showed that the weight $k(\rho)$ depends on the restriction of ρ to the inertia group at p. In order to understand this, we will need to understand the structure of $\mathrm{Gal}(\overline{\mathbf{Q}}_p/\mathbf{Q}_p)$. Recall from §1.1 that we have a sequence of subgroups

$$G_{p,i} \subseteq I_p \subseteq G_p.$$

Let

- $I_w \cong G_{p,1} \lhd I_p$ be the *wild inertia group*, the largest pro-p subgroup of I_p.
- $I_t = I_p/I_w$ be the *tame inertia group.*
- $\mathbf{Q}_p^{\mathrm{unr}}$ be the maximal unramified extension of $\mathbf{Q}_p$.
- $\mathbf{Q}_p^{\mathrm{tr}}$ be the maximal tamely ramified extension of $\mathbf{Q}_p$.

We obtain the following tower of fields with their corresponding Galois groups:

$$\begin{tikzcd}
\overline{\mathbf{Q}}_p \arrow[d, dash, "I_w"] \arrow[dd, dash, bend left, "I_p"] \arrow[ddd, dash, bend right, "G_p"'] \\
\mathbf{Q}_p^{\mathrm{tr}} \arrow[d, dash, "I_t"] \\
\mathbf{Q}_p^{\mathrm{unr}} \arrow[d, dash] \\
\mathbf{Q}_p
\end{tikzcd}$$

Recall that since I_p is the kernel of the reduction homomorphism $G_p \longrightarrow \mathrm{Gal}(\overline{\mathbf{F}}_p/\mathbf{F}_p)$, we have

$$\mathrm{Gal}(\overline{\mathbf{F}}_p/\mathbf{F}_p) \cong G_p/I_p.$$

Hence, $\mathrm{Gal}(\overline{\mathbf{F}}_p/\mathbf{F}_p)$ acts naturally on I_t, so we can view I_t as a $\mathrm{Gal}(\overline{\mathbf{F}}_p/\mathbf{F}_p)$-module.

Lemma 3.4. *There is an identification of* $\mathrm{Gal}(\overline{\mathbf{F}}_p/\mathbf{F}_p)$*-modules*

$$I_t = \varprojlim_{n\geq 1} \mathbf{F}_{p^n}^\times,$$

where the transition maps are the norm maps

$$\begin{aligned} \mathbf{F}_{q^m}^\times &\longrightarrow \mathbf{F}_q^\times \\ \zeta &\longmapsto \zeta^{\frac{q^m-1}{q-1}} = \mathbf{N}_{\mathbf{F}_{q^m}/\mathbf{F}_q}(\zeta). \end{aligned}$$

Proof. The field $\mathbf{Q}_p^{\mathrm{tr}}$ consists of all totally and tamely ramified extensions of $\mathbf{Q}_p^{\mathrm{unr}}$; since all such extensions are of the form $\mathbf{Q}_p^{\mathrm{unr}}(\sqrt[n]{p})$ for integers n such that $p \nmid n$, we deduce that

$$\mathbf{Q}_p^{\mathrm{tr}} = \varinjlim_{p\nmid n} \mathbf{Q}_p^{\mathrm{unr}}(\sqrt[n]{p}). \tag{3.2}$$

Similarly, since the unramified extensions of $\mathbf{Q}_p$ are the extensions $\mathbf{Q}_p(\zeta_n)$ for integers n for which $p \nmid n$, we have

$$\mathbf{Q}_p^{\mathrm{unr}} = \varinjlim_{p \nmid n} \mathbf{Q}_p(\zeta_n). \tag{3.3}$$

For each n not divisible by p, Kummer theory gives us a canonical isomorphism

$$\begin{aligned} \mathrm{Gal}(\mathbf{Q}_p^{\mathrm{unr}}(\sqrt[n]{p})/\mathbf{Q}_p^{\mathrm{unr}}) &\longrightarrow \mu_n \\ \sigma &\longmapsto \frac{\sigma(\sqrt[n]{p})}{\sqrt[n]{p}}, \end{aligned}$$

where $\mu_n = \mu_n(\overline{\mathbf{Q}}_p)$ is the group of n^{th} roots of unity, which are contained in $\mathbf{Q}_p^{\mathrm{unr}}$ by (3.3). Each such isomorphism lifts to a map $I_p \longrightarrow \mu_n$, which factors through I_t due to identification (3.2). Write $\theta_n : I_t \longrightarrow \mu_n$ for this map.

We obtain an identification

$$I_t = \mathrm{Gal}(\mathbf{Q}_p^{\mathrm{tr}}/\mathbf{Q}_p^{\mathrm{unr}}) = \varprojlim_{p \nmid n} \mathrm{Gal}(\mathbf{Q}_p^{\mathrm{unr}}(\sqrt[n]{p})/\mathbf{Q}_p^{\mathrm{unr}}) = \varprojlim_{p \nmid n} \mu_n.$$

We can show that this is an identification of $\mathrm{Gal}(\overline{\mathbf{F}}_p/\mathbf{F}_p)$-modules, and that we obtain an identification

$$I_t = \varprojlim_{n \geq 1} \mathbf{F}_{p^n}$$

with the transition maps as claimed. □

3.4 Fundamental characters

Definition 3.5. Let $\phi : I_t \longrightarrow \overline{\mathbf{F}}_p$ be any continuous character. We say that ϕ is of *level n* if n is the smallest integer such that ϕ factors through $\mathbf{F}_{p^n}^\times$:

$$\begin{tikzcd} I_t \cong \varprojlim_n \mathbf{F}_{p^n}^\times \arrow[rr, "\phi"] \arrow[dr] & & \overline{\mathbf{F}}_p^\times \\ & \mathbf{F}_{p^n}^\times \arrow[ur] & \end{tikzcd}$$

The continuity of ϕ ensures that such an n exists.

In the previous section, we defined a map

$$\theta_n : I_t \longrightarrow \mu_n(\overline{\mathbf{Q}}_p).$$

The group $\mu_n(\overline{\mathbf{Q}}_p)$ lies in the ring of integers $\overline{\mathbf{Z}}_p$ of $\overline{\mathbf{Q}}_p$. Composing this map with reduction modulo the maximal ideal of $\overline{\mathbf{Z}}_p$ gives a mod p character of I_t as illustrated:

$$\begin{tikzcd} I_t \arrow[rrr] \arrow[dr, "\theta_n"'] & & & \overline{\mathbf{F}}_p^\times \\ & \mu_n(\overline{\mathbf{Q}}_p) \arrow[r, "\sim"] & \mu_n(\overline{\mathbf{F}}_p) \arrow[ur, hook] & \end{tikzcd}$$

The embedding $\mu_n(\overline{\mathbf{F}}_p) \hookrightarrow \overline{\mathbf{F}}_p^\times$ is in general non-canonical and non-unique. Let $q = p^n$ with $n \geq 1$. Since

$$\mu_{q-1}(\mathbf{F}_q) = \mathbf{F}_q^\times,$$

the map $\theta_{q-1} : I_t \longrightarrow \mu_{q-1}$ allows us to define a character

$$\phi : I_t \longrightarrow \mathbf{F}_q^\times.$$

Each of the n embeddings $\mathbf{F}_{p^n} \hookrightarrow \overline{\mathbf{F}}_p$ induces different embedding $\mathbf{F}_{p^n}^\times \hookrightarrow \overline{\mathbf{F}}_p^\times$.

Definition 3.6. The *fundamental characters of level* n are the characters induced by composing ϕ with one of the n embeddings $\mathbf{F}_{p^n} \hookrightarrow \overline{\mathbf{F}}_p$

$$\begin{tikzcd}
I_t \arrow[rr] \arrow[dr, "\phi"'] & & \overline{\mathbf{F}}_p^\times \subseteq \overline{\mathbf{F}}_p \\
& \mathbf{F}_{p^n}^\times \subseteq \mathbf{F}_{p^n} \arrow[ur, shift left] \arrow[ur] \arrow[ur, shift right, "n \text{ embeddings}"'] &
\end{tikzcd}$$

These characters form a generating set for the group of characters of I_t of level n.

Example 3.7. The mod p cyclotomic character χ_p is tamely ramified (since $\mathbf{Q}_p(\zeta_p)/\mathbf{Q}_p$ is), and hence gives an character of I_t. Let ψ be the fundamental character of level 1. By definition, we have a commutative diagram

$$\begin{tikzcd}
I_t \arrow[r] \arrow[dr, "\theta_{p-1}"'] \arrow[rrr, bend left, "\psi"] & \mathrm{Gal}(\mathbf{Q}_p^{\mathrm{unr}}(\sqrt[p-1]{p})/\mathbf{Q}_p^{\mathrm{unr}}) \arrow[rr] \arrow[d] & & \mathbf{F}_p^\times \\
& \mu_{p-1}(\overline{\mathbf{Q}}_p) \arrow[r, "\sim"] & \mu_{p-1}(\overline{\mathbf{F}}_p^\times) \arrow[ur] &
\end{tikzcd}$$

Similarly, the character χ_p satisfies the following commutative diagram

$$\begin{tikzcd}
I_t \arrow[r] \arrow[dr] \arrow[rrr, bend left, "\chi_p"] & \mathrm{Gal}(\mathbf{Q}_p^{\mathrm{unr}}(\zeta_p)/\mathbf{Q}_p^{\mathrm{unr}}) \arrow[rr] \arrow[d] & & \mathbf{F}_p^\times \\
& \mu_{p-1}(\overline{\mathbf{Q}}_p) \arrow[r, "\sim"] & \mu_{p-1}(\overline{\mathbf{F}}_p^\times) \arrow[ur] &
\end{tikzcd}$$

The fact that $\mathbf{Q}_p^{\mathrm{unr}}(\zeta_p) = \mathbf{Q}_p^{\mathrm{unr}}(\sqrt[p-1]{p})$ shows that $\psi = \chi_p$.

Example 3.8. There are two fundamental characters of level 2, which we denote ψ and ψ'. We have

$$\psi^p = \psi' \quad \psi'^p = \psi.$$

3.5 The local representation at p

Let ρ_p be the restriction of ρ to G_p, and let V^{ss} denote the semisimplification of V with respect to the action of G_p.

Lemma 3.9. *The wild ramification group I_w acts trivially on V^{ss}.*

Proof. We follow the proof of Proposition 4 in [Ser72]. The key fact is that I_w is a pro-p group, and hence that the orbits of its action on any finite set X must have p-power order. In particular, if $\#X$ is itself a power of p, then by a combinatorial argument, this action cannot be trivial.

We need show that I_w acts trivially on each of the direct summands of V^{ss}; hence, without loss of generality, we can assume that $V = V^{ss}$ is simple. By Theorem 1.4, the image of ρ_p is finite; therefore, it can be realised over a finite extension K of $\mathbf{F}_p$.

Let $V' = K^{I_w}$ be the space upon which I_w acts trivially. I_w acts on K and its orbits have p-power order. The orbit $\{0\}$ has size 1, and $\#K$ is a power of p; since the orbits partition K, there must be at least $p-1$ other singleton orbits - i.e. at least $p-1$ non-trivial points fixed by I_w. Hence, V' is non-trivial.

Since $I_w \lhd G_p$, the space V' is stable under the action of G_p. But V' is a non-trivial subspace, so it follows that $V' = K$. This proves the lemma. □

Corollary 3.10. *The tame ramification group $I_t \cong I_p/I_w$ acts on V^{ss}.*

The group

$$I_t \cong \varprojlim_{m \geq 1} \mathbf{F}_{p^m}^{\times}$$

is abelian; hence all its irreducible representations are one-dimensional. Therefore, the two-dimensional representation $\rho^{ss}|_{I_t}$ is reducible, and can be written as a direct sum of two characters

$$\varphi, \varphi' : I_t \longrightarrow \overline{\mathbf{F}}_p^{\times}.$$

Proposition 3.11. *The characters φ and φ' have level 1 or 2. If they have level 2, then they are conjugate, and we have $\varphi' = \varphi^p$ and $\varphi = \varphi'^p$.*

Proof. Let σ be any element of G_p whose image in $G_p/I_p \cong \mathrm{Gal}(\overline{\mathbf{F}}_p/\mathbf{F}_p) = G_{\mathbf{F}_p}$ is the Frobenius automorphism

$$\mathrm{Frob}_p : x \longmapsto x^p.$$

The group $G_{\mathbf{F}_p}$ acts on I_t by conjugation. Since

$$I_t = \varprojlim_{n \geq 1} \mathbf{F}_{p^n}^{\times}$$

as $G_{\mathbf{F}_p}$-modules, conjugation by σ acts on $I_t = I_p/I_w$ via $u \longmapsto u^p$. Hence, for each $u \in I_p$, we have

$$\sigma u \sigma^{-1} = u^p \pmod{I_w}.$$

The two representations $u \longmapsto \rho^{ss}(u)^p$ and $u \longmapsto \rho^{ss}(u)$ are therefore equivalent via conjugation by $\rho^{ss}(\sigma)$:

$$\begin{aligned}\rho^{ss}(\sigma)\rho^{ss}(u)\rho^{ss}(\sigma^{-1}) &= \rho^{ss}(\sigma u \sigma^{-1}) \\ &= \rho^{ss}(u^p) = \rho^{ss}(u)^p\end{aligned}$$

and hence, both representations can be expressed as a direct sum of φ and φ'. We deduce that

$$\{\varphi, \varphi'\} = \{\varphi^p, \varphi'^p\}.$$

There are then two cases:

1. We have $\varphi^p = \varphi$ and $\varphi'^p = \varphi'$. In this case, the two characters φ and φ' have level 1.

2. We have $\varphi^p = \varphi'$ and $\varphi'^p = \varphi$, and the two characters φ and φ' have level 2.

□

3.6 Definition of $k(\rho)$ when φ and φ' have level 2

Suppose that we are in case 2 of Proposition 3.11: that φ and φ' have level 2. We can write φ and φ' uniquely as

$$\begin{aligned} \varphi &= \psi^a \psi'^b \\ \varphi' = \varphi^p &= (\psi^p)^a (\psi'^p)^b = \psi^b \psi'^a \end{aligned}$$

where ψ, ψ' are the fundamental characters of level 2 and $0 \leq a, b \leq p-1$. Since φ has level 2, and $\psi\psi' = \chi_p$, we have $a \neq b$. Interchanging φ and φ' as necessary, we can therefore assume that $0 \leq a < b \leq p-1$.

Lemma 3.12. *The representation $\rho_p : G_p \longrightarrow \mathrm{GL}(V)$ is irreducible.*

Proof. Suppose not. Then V contains a stable one-dimensional subspace and the action of I_t on this subspace is given by one of the characters φ, φ'. This character can be extended to a tame character Φ of G_p. Note that since

$$\sigma u \sigma^{-1} \equiv u^p \pmod{I_w}$$

for any $u \in I_t$ and $\sigma \in G_p$ a lift of $\mathrm{Frob}_p \in G_{\mathbf{F}_p}$, the fact that Φ is tame (so that $\Phi|_{I_w}$ is trivial) means that $\Phi|_{I_p}$ takes values in $\mathbf{F}_p^\times$. Hence, it is of level 1, contradicting our assumption that φ and φ' have level 2. □

Hence, by Lemma 3.9, the wild inertia group I_w acts trivially on V. Therefore, the restriction of ρ_p to the inertia group takes the form

$$\rho_p\big|_{I_p} = \rho_p\big|_{I_t} = \begin{pmatrix} \psi^a\psi'^b & 0 \\ 0 & \psi'^a\psi^b \end{pmatrix} = \chi_p^a \begin{pmatrix} \psi'^{b-a} & 0 \\ 0 & \psi^{b-a} \end{pmatrix}.$$

This is exactly of the form of a twist of a representation arising from an eigenform with $a_p = 0$ as in Fontaine's theorem (Theorem 3.2). We write

$$\rho_p = \chi_p^a \otimes \rho_p',$$

where ρ_p' is a representation whose restriction to I_p is

$$\begin{pmatrix} \psi'^{b-a} & 0 \\ 0 & \psi^{b-a} \end{pmatrix}.$$

The weight of the untwisted representation ρ_p' should be $k = b - a + 1$, and hence, by the results of §3.2, we define

$$k(\rho) = b - a + 1 + a(p+1) = 1 + pa + b.$$

3.7 Definition of $k(\rho)$ when φ and φ' have level 1, and I_w acts trivially

Suppose that we are in case 1 of Proposition 3.11, but that I_w acts trivially on V. We have

$$(\varphi, \varphi') = (\chi_p^a, \chi_p^b)$$

for some $0 \leq a, b \leq p-2$. Interchanging φ and φ' as necessary, we can assume that $a \leq b$. The restriction of ρ_p to the inertia takes the form

$$\rho_p\big|_{I_p} = \rho_p\big|_{I_t} = \begin{pmatrix} \chi_p^b & 0 \\ 0 & \chi_p^a \end{pmatrix} = \chi_p^a \begin{pmatrix} \chi_p^{b-a} & 0 \\ 0 & 1 \end{pmatrix}.$$

This is of the form

$$\chi_p^a \otimes \rho_p',$$

where ρ_p' arises from an eigenform with $a_p \neq 0$ as in Deligne's theorem (Theorem 3.1). The weight of ρ_p should be $k = b - a + 1$. Hence we define

$$k(\rho) = \begin{cases} 1 + pa + b & \text{if } (a,b) \neq (0,0) \\ p & \text{if } (a,b) = (0,0) \end{cases}$$

Remark 3.13. The case $(a,b) = (0,0)$ corresponds to the case where I_p acts trivially on V - i.e. where ρ_p is unramified. The formula $k(\rho) = 1 + pa + b$ would give $k(\rho) = 1$ in this case. However, in Serre's definition of mod p modular forms, the forms of weight 1 have different properties to those of other weights; we avoid these by translating the weight by $p-1$ to give $k(\rho) = p$.

3.8 Definition of $k(\rho)$ when φ and φ' have level 1, and I_w does not act trivially

Suppose that I_w does not act trivially, so that the action of I_p is not tame. The proof of Lemma 3.9 shows that V^{I_w} forms a stable one-dimensional subspace of V. Hence G_p acts on the spaces V/V^{I_w}, V^{I_w} via two characters, θ_1, θ_2:

$$\rho_p = \begin{pmatrix} \theta_2 & * \\ 0 & \theta_1 \end{pmatrix}.$$

We can write θ_1 and θ_2 uniquely as

$$\theta_1 = \chi_p^\alpha \epsilon_1, \quad \theta_2 = \chi_p^\beta \epsilon_2,$$

where $0 \leq \alpha \leq p-2$ and $1 \leq \beta \leq p-1$. The restriction of ρ_p to I_p is therefore

$$\rho_p\big|_{I_p} = \begin{pmatrix} \chi_p^\beta & * \\ 0 & \chi_p^\alpha \end{pmatrix}.$$

In this case, the roles of α and β are not symmetric, so we cannot just swap them to ensure that $\alpha < \beta$. Hence, we let

$$a = \min(\alpha, \beta), \quad b = \max(\alpha, \beta).$$

There are now two cases:

(i) If $\beta \neq \alpha + 1$, then we proceed as in §3.7 and set

$$k(\rho) = 1 + pa + b.$$

(ii) If $\beta = \alpha + 1$, then $\chi_p^{\beta-\alpha} = \chi_p$. Hence the restriction of ρ_p to the inertia is of the form

$$\rho_p\big|_{I_p} = \chi_p^{\alpha} \begin{pmatrix} \chi_p & * \\ 0 & 1 \end{pmatrix}, \tag{3.4}$$

and it is unclear whether the untwisted representation should have weight 2 or $p+1$. Which it should be will depend on the type of wild ramification - whether ρ is *peu ramifié* or *très ramifié*.

Recall from section §3.3 that we have the following tower of fields with their corresponding Galois groups:

$$\begin{tikzcd}
\overline{\mathbf{Q}}_p \arrow[d, no head, "I_w"] \arrow[dd, no head, bend left, "I_p"] \arrow[ddd, no head, bend right=60, "G_p"'] \\
\mathbf{Q}_p^{\mathrm{tr}} \arrow[d, no head, "I_t"] \\
\mathbf{Q}_p^{\mathrm{unr}} \arrow[d, no head] \\
\mathbf{Q}_p
\end{tikzcd}$$

The representation $\rho_p\big|_{I_p}$ factors through some Galois group $\mathrm{Gal}(K/\mathbf{Q}_p^{\mathrm{unr}})$ of a totally ramified finite extension $K/\mathbf{Q}_p^{\mathrm{unr}}$. The wild inertia group $\rho_p(I_p)$ is the Galois group of K/K_t, where K_t is the largest tamely ramified extension of $\mathbf{Q}_p^{\mathrm{unr}}$ contained in K.

$$\begin{tikzcd}
K \arrow[d, no head, "\rho_p(I_w)"] \arrow[dd, no head, bend right=60, "\rho_p(I_p)"'] \\
K_t \arrow[d, no head] \\
\mathbf{Q}_p^{\mathrm{unr}}
\end{tikzcd}$$

Since $\beta = \alpha + 1$, the action of $\mathrm{Gal}(K_t/\mathbf{Q}_p^{\mathrm{unr}}) = \rho_p(I_p)/\rho_p(I_w)$ on V^{ss} gives a faithful representation of the form

$$\begin{pmatrix} \chi_p^{\alpha+1} & 0 \\ 0 & \chi_p^{\alpha} \end{pmatrix},$$

and we deduce that $\mathrm{Gal}(K_t/\mathbf{Q}_p^{\mathrm{unr}})$ has order $p-1$. Since $K_t \supseteq \mathbf{Q}_p^{\mathrm{unr}}(\zeta_p)$, we conclude that $K_t = \mathbf{Q}_p^{\mathrm{unr}}(\zeta_p)$, and that $\mathrm{Gal}(K_t/\mathbf{Q}_p^{\mathrm{unr}}) \cong (\mathbf{Z}/p\mathbf{Z})^{\times}$.

On the other hand, the group $\mathrm{Gal}(K/K_t) = \rho_p(I_w)$ is a finite abelian p-group. Since χ_p acts trivially on I_w, we see that $\rho_p(I_w)$ is of the form

$$\begin{pmatrix} 1 & * \\ 0 & 1 \end{pmatrix}$$

and is therefore killed by p. Hence, $\rho_p(I_p)$ is an abelian group of type $(p, p, \ldots, p)$.

The group $\mathrm{Gal}(K_t/\mathbf{Q}_p^{\mathrm{unr}}) \cong (\mathbf{Z}/p\mathbf{Z})^\times$ acts on $\mathrm{Gal}(K/K_t)$ by conjugation via the character $\chi_p^{\beta-\alpha} = \chi_p$. Using Kummer theory, we deduce that

$$K = K_t(x_1^{1/p}, \ldots x_m^{1/p}), \quad \text{where } p^m = [K : K_t], \tag{3.5}$$

and the x_i are elements of $(\mathbf{Q}_p^{\mathrm{unr}})^\times/(\mathbf{Q}_p^{\mathrm{unr}})^{\times p}$.

Definition 3.14. Let v_p denote the p-adic valuation of $\mathbf{Q}_p^{\mathrm{unr}}$, normalised so that $v_p(p) = 1$.

a) If the x_i can be chosen to satisfy

$$v_p(x_i) \equiv 0 \pmod p \text{ for } i = 1, \ldots m,$$

then we say that the extension K (and hence the representation ρ_p) is *peu ramifié*.

b) Otherwise, we say that K and ρ_p are both *très ramifié*.

We can now define the integer $k(\rho)$:

(ii_1) If ρ_p is *peu ramifié*, then we let the untwisted part of (3.4) have $k = 2$ so that

$$k(\rho) = 2 + \alpha(p + 1) = 1 + pa + b$$

(ii_2) If ρ_p is *très ramifié* and $p \geq 3$, then we let the untwisted part of (3.4) have $k = p + 1$ so that

$$k(\rho) = p + 1 + p(\alpha + 1) = 1 + pa + b + p - 1.$$

In the case that $p = 2$, we let $k(\rho) = 4$.

We can now state the strong version of Serre's conjecture:

Theorem 3.15 (Serre's Conjecture, Strong Version). *Let*

$$\rho : G_{\mathbf{Q}} \longrightarrow \mathrm{GL}(V) \cong \mathrm{GL}_2(\overline{\mathbf{F}}_p)$$

be a Galois representation, V being a two-dimensional vector space over $\overline{\mathbf{F}}_p$. Suppose that ρ is irreducible and odd. Then there exists a cuspidal eigenform f with coefficients in $\overline{\mathbf{F}}_p$ of type $(N(\rho), k(\rho), \epsilon(\rho))$ whose associated representation ρ_f is equivalent to ρ.

4 Examples

4.1 Galois representations arising from semistable elliptic curves

Let E be a semistable elliptic curve over $\mathbf{Q}$ - i.e. a curve which has good or multiplicative reduction at every prime p. Let j_E be its modular invariant, and $E[p] = E[p](\overline{\mathbf{Q}})$ be the group of its p-torsion points. As an abstract group, $E[p]$ is isomorphic to $\mathbf{Z}/p\mathbf{Z} \times \mathbf{Z}/p\mathbf{Z}$, and can therefore be viewed as a two-dimensional vector space over $\mathbf{F}_p$. The action of $G_{\mathbf{Q}}$ on $E[p]$ defines a representation

$$\overline{\rho}_{E,p} : G_{\mathbf{Q}} \longrightarrow \mathrm{Aut}(E[p]) \cong \mathrm{GL}_2(\mathbf{F}_p).$$

In this section, we will show how to explicitly calculate the predicted level, weight and character of $\overline{\rho}_{E,p}$:

Theorem 4.1. *Let $E/\mathbf{Q}$ be a semistable elliptic curve with minimal discriminant Δ_E. Let v_ℓ be the normalised ℓ-adic valuation of $\mathbf{Q}$. Then*

1. $\epsilon(\overline{\rho}_{E,p}) = \mathbf{1}$,
2. $N(\overline{\rho}_{E,p}) = \prod_{\ell \neq p} \ell^{n(\ell,\overline{\rho}_{E,p})}$ where $n(\ell,\overline{\rho}_{E,p}) = \begin{cases} 0 & \text{if } v_\ell(\Delta_E) \equiv 0 \pmod p \\ 1 & \text{if } v_\ell(\Delta_E) \not\equiv 0 \pmod p \end{cases}$
3. $k(\overline{\rho}_{E,p}) = \begin{cases} 2 & \text{if } v_p(\Delta_E) \equiv 0 \pmod p \\ p+1 & \textit{otherwise.} \end{cases}$

Proof.

1. We can define the *Weil pairing* on E

$$e_p : E[p] \times E[p] \longrightarrow \mu_p.$$

This pairing is bilinear, alternating (i.e. $e_p(T,T) = 1$), non-degenerate and Galois invariant (see [CSS97] §II.8). Hence, e_p induces an isomorphism of Galois modules

$$E[p] \wedge E[p] \longrightarrow \mu_p.$$

Therefore, for each $\sigma \in G_{\mathbf{Q}}$, $S, T \in E[p]$, we have with this identification

$$\begin{aligned} \chi_p(\sigma)(S \wedge T) &= \sigma(S \wedge T) && \text{by the definition of } \chi_p \\ &= \sigma(S) \wedge \sigma(T) && \text{by Galois invariance} \\ &= \overline{\rho}_{E,p}(\sigma)S \wedge \overline{\rho}_{E,p}(\sigma)T. \end{aligned}$$

We deduce that

$$\det \overline{\rho}_{E,p} = \chi_p.$$

and hence that $\epsilon(\overline{\rho}_{E,p}) = \mathbf{1}$.

2. To calculate $N(\overline{\rho}_{E,p})$, we will split into two cases, depending on whether or not E has a good reduction mod ℓ.

Suppose first that E has a good reduction mod ℓ - i.e. that $v_\ell(\Delta_E) = 0$. We will show that $\overline{\rho}_{E,p}$ is unramified at ℓ; from here, we can deduce that $\ell \nmid N(\overline{\rho}_{E,p})$.

Let λ be a choice of prime lying above ℓ; this choice of prime allows us to define a reduction map $D_\lambda \longrightarrow \mathrm{Gal}(\overline{\mathbf{F}}_\ell/\mathbf{F}_\ell)$ where D_λ is the decomposition group of λ. Let $\widetilde{E}$ be the reduction of E mod ℓ. Then there is a commutative diagram (see [DS05] p383)

$$\begin{tikzcd}
D_\lambda \arrow[r] \arrow[d] & \mathrm{Aut}(E[p]) \arrow[d] \\
\mathrm{Gal}(\overline{\mathbf{F}}_\ell/\mathbf{F}_\ell) \arrow[r] & \mathrm{Aut}(\widetilde{E}[p])
\end{tikzcd}$$

where the top map is the restriction of $\overline{\rho}_{E,p}$ to D_λ and the bottom map is given by the action of $\mathrm{Gal}(\overline{\mathbf{F}}_\ell/\mathbf{F}_\ell)$ on $\widetilde{E}$. The map $D_\lambda \longrightarrow \mathrm{Gal}(\overline{\mathbf{F}}_\ell/\mathbf{F}_\ell)$ has kernel I_λ; hence, the inertia group I_λ lies in the kernel of the composite map

$$D_\lambda \longrightarrow \mathrm{Aut}(E[p]) \longrightarrow \mathrm{Aut}(\widetilde{E}[p]).$$

Since E has good reduction at ℓ, the second map is an isomorphism. Hence, $I_\lambda \subseteq \ker \overline{\rho}_{E,p}$, and $\overline{\rho}_{E,p}$ is unramified at ℓ as claimed.

Now suppose that E has multiplicative reduction mod ℓ. Then there exists an unramified extension K of $\mathbf{Q}_\ell$ over which E has *split multiplicative reduction* at ℓ - i.e. the two tangent lines to the node on $\widetilde{E}(\mathbf{F}_p)$ have slopes defined over $\mathbf{F}_p$.

By consideration of the Tate curve (see [DDT94] Proposition 1.5, [Cai09] Theorem 6.2.1), there is a ℓ-adic analytic isomorphism of $\mathrm{Gal}(\overline{\mathbf{Q}}_\ell/K)$-modules

$$\Phi : \overline{\mathbf{Q}}_\ell^\times/q^{\mathbf{Z}} \longrightarrow E(\overline{\mathbf{Q}}_\ell), \tag{4.1}$$

where q is defined in terms of the j-invariant j_E of E:

$$j_E = \frac{1}{q} + 744 + 196884q + \cdots .$$

Under this isomorphism, $E[p]$ corresponds to the subgroup

$$\langle \zeta_p, q^{1/p} \rangle = \left\{ \zeta_p^a (q^{1/p})^b : 0 \leq a, b < p \right\}$$

of elements of $\overline{\mathbf{Q}}_\ell^\times/q^{\mathbf{Z}}$ of order p.

We need to show that $\mathbf{Q}_\ell(E[p])$ is tamely ramified, and unramified if and only if $p \mid v_\ell(\Delta_E)$. This will hold if and only if it holds for the extension $K(\zeta_p, q^{1/p})/K$ (since $K/\mathbf{Q}_\ell$ is unramified).

The extension has degree p^2; hence if $\ell \neq p$, then since the wild inertia group I_w is a ℓ-group contained in $\mathrm{Gal}(K(\zeta_p, q^{1/p})/K)$, it must be trivial, so the extension is indeed tamely ramified. It will be unramified if and only if $v_\ell(q^{1/p}) \in \mathbf{Z}$ (here v_ℓ is the extension of v_ℓ to $K(q^{1/p})$). This will occur if and only if $p \mid v_\ell(q)$. Since

$$v_\ell(q) = -v_\ell(j_E) = v_\ell(\Delta_E),$$

the result follows.

3. Since $\det \overline{\rho}_{E,p} = \chi_p$, the weight $k(\overline{\rho}_{E,p})$ must satisfy

$$k(\overline{\rho}_{E,p}) - 1 \equiv 1 \pmod{p-1}.$$

If E has good reduction mod p, then $\overline{\rho}_{E,p}$ will be *finite at* p, which is an equivalent definition to $\overline{\rho}_{E,p}$ being *peu ramifié* (see [Ser87] Propositon 4).

If E has multiplicative reduction mod p, then working over an unramified quadratic extension K of $\mathbf{Q}_p$ and using the Tate model (see [Ser87] §2.9), we have an exact sequence of Galois modules over K

$$0 \longrightarrow \mu_p \longrightarrow E[p] \longrightarrow \mathbf{Z}/p\mathbf{Z} \longrightarrow 0.$$

We deduce that

$$\overline{\rho}_{E,p}\Big|_{I_p} \cong \begin{pmatrix} \chi_p & * \\ 0 & 1 \end{pmatrix},$$

and therefore that

$$k(\overline{\rho}_{E,p}) = \begin{cases} 2 & \text{if } \overline{\rho}_{E,p} \text{ is } \textit{peu ramifié}, \\ p+1 & \text{if } \overline{\rho}_{E,p} \text{ is } \textit{très ramifié}. \end{cases}$$

As before, under the isomorphism (4.1), $E[p]$ corresponds to $\langle \zeta_p, q^{1/p} \rangle$. By equation (3.5), we need to consider the field extension

$$K_t(\zeta_p, q^{1/p}) = K_t(q^{1/p}).$$

This will be *peu ramifié* if and only if $v_p(q) \equiv 0 \pmod{p}$ - i.e. if $p \nmid v_p(\Delta_E)$, as required.

□

4.2 Computational examples

1. Consider the elliptic curve

$$A : y^2 + y = x^3 + x^2 - 23x - 50,$$

which has conductor $N_A = 37$ and minimal discriminant $\Delta_A = 37^3$.

Let $\overline{\rho}_{A,3}$ be the mod 3 representation corresponding to A. If $\overline{\rho}_{A,3}$ were irreducible, then by Theorem 4.1, it would arise from a modular form of weight 2 and level 1. Since the space of such forms is empty, we deduce that $\overline{\rho}_{A,3}$ is reducible.

2. More generally, suppose $p < 11$ and that

$$\rho : G_{\mathbf{Q}} \longrightarrow \mathrm{GL}_2(\overline{\mathbf{F}}_p)$$

is an odd Galois representation that is unramified away from p, so that $N(\rho) = 1$. By Theorem 3.3, there is a twist of ρ such that $2 \leq k(\rho) \leq p+1$. Since the spaces $S_k(1)$ are all zero for $k \leq 11$, we deduce that ρ cannot be irreducible. Therefore, there are no two-dimensional irreducible unramified mod p Galois representations when $p < 11$.

3. The above result is false for $p = 11$. Indeed, consider the elliptic curve

$$B : y^2 + y = x^3 - x^2,$$

which has conductor $N_B = 11$ and minimal discriminant $\Delta_B = -11$.

Let $\overline{\rho}_{B,11}$ be the mod 11 representation corresponding to B. Since B is semistable (N_B is square-free), it follows from Mazur's theorem (Theorem 5.3) that $\overline{\rho}_{B,11}$ is irreducible (the proof is similar to that of Corollary 5.4). We have $N(\overline{\rho}_{B,11}) = 1$ and $k(\overline{\rho}_{B,11}) = 12$. The space $S_{12}(1)$ has dimension 1, and contains a unique eigenform

$$\Delta(z) = \sum_{n\geq 1} \tau(n) q^n = q \prod_{n\geq 1} (1-q^n)^{24},$$

where τ is the *Ramanujan tau function.* We can compare the values of τ_ℓ with the traces of absolute Frobenius elements on $B[11]$:

ℓ	2	3	5	7	11	13	17	19
$\tau(\ell)$	-24	252	4830	-16744	534612	-577738	-6905934	10661420
$a_\ell(B)$	-2	-1	1	-2	1	4	-2	0

We check that

$$\text{tr}\ (\overline{\rho}_{A,11}(\text{Frob}_\ell)) = a_\ell(B) \equiv \tau(n) \pmod{11}$$

for all the values calculated (including for $\ell = 11$, although this is not expected).

5 Applications

5.1 Fermat's last theorem

Theorem 5.1 (Fermat's Last Theorem)**.** *The equation*

$$x^n + y^n = z^n, \quad xyz \neq 0$$

has no integer solutions when $n \geq 3$.

In this section, we will show how to deduce Fermat's last theorem from Serre's conjecture in the case that n is a prime number greater than or equal to 5.

Suppose that

$$a^p + b^p = c^p, \quad abc \neq 0,\ p \geq 5$$

is a solution to Fermat's equation. The idea of the proof is to use this proposed solution to write down a semistable elliptic curve, known as the *Frey curve*, and to apply Serre's conjecture to its corresponding mod p Galois representation. However, due to the properties of the chosen elliptic curve, this Galois representation cannot be modular, as there are no non-zero modular forms of the prescribed weight and level. This contradicts Serre's conjecture.

Definition 5.2. The *Frey curve* associated with this proposed solution to Fermat's equation is the elliptic curve

$$E : y^2 = x(x - a^p)(x + b^p).$$

In order to be able to apply the results of §4.1, we would like E to be semistable. Without loss of generality, we can assume that a, b, c are pairwise coprime. This ensures that E has good or multiplicative reduction at any prime $\ell \neq 2$ (since at most one of a, b can be divisible by ℓ). By swapping a and b, and by changing signs as needed, we can also assume that

$$a \equiv -1 \pmod 4, \quad b \equiv 0 \pmod 4$$

thereby ensuring that E has at worst multiplicative reduction at $\ell = 2$.

We can also check that the corresponding Galois representation

$$\overline{\rho}_{E,p} : G_{\mathbf{Q}} \longrightarrow \mathrm{GL}_2(\mathbf{F}_p)$$

is irreducible. This follows from the following theorem of Mazur ([Maz78] Theorem 2):

Theorem 5.3 (Mazur)**.** *If $E/\mathbf{Q}$ is an elliptic curve, then its torsion subgroup is isomorphic to one of the following:*

- $\mathbf{Z}/N\mathbf{Z},\ 1 \leq N \leq 10,\ N = 12,$
- $\mathbf{Z}/2N\mathbf{Z} \times \mathbf{Z}/2\mathbf{Z},\ 1 \leq n \leq 4.$

Corollary 5.4. *The representation $\overline{\rho}_{E,p}$ is irreducible.*

Proof. Any semistable elliptic curve with a rational subgroup of order ℓ is necessarily isogenous to an elliptic curve with a rational point of order ℓ (see [DDT94] Theorem 2.9).

Since

$$E[2] = \big\{\mathcal{O}, (0,0), (a^p, 0), (-b^p, 0)\big\}$$

consists of only rational points, we use deduce that E is isogenous to an elliptic curve whose torsion subgroup has order divisible by 4.

Suppose that $\overline{\rho}_{E,p}$ is reducible - i.e. that the torsion subgroup of E contains a subgroup of order p. By the same argument as above, the torsion subgroup of E must have order divisible by p, and hence by $4p \geq 20$ (here we use the fact that $p \geq 5$). This contradicts Mazur's theorem. □

We've now shown that the Galois representation $\overline{\rho}_{E,p}$ is irreducible. Since every elliptic curve has the automorphism -1, we can check that $\overline{\rho}_{E,p}$ is odd, and hence that it satisfies the conditions of Serre's conjecture. The key fact about this representation is that it has very little ramification. Indeed, E has minimal discriminant

$$\Delta_E = 2^{-8}(abc)^{2p},$$

so the results of §4.1 show that $\overline{\rho}_{E,p}$ is unramified outside 2 and p, and that Serre's conjecture predicts the weight and level of $\overline{\rho}_{E,p}$ to be

$$N(\overline{\rho}_{E,p}) = 1 \text{ or } 2 \quad k(\overline{\rho}_{E,p}) = 2$$

where the value of N depends on the ramification at 2. But there are no non-trivial eigenforms of weight 2 and level 2, contradicting Serre's conjecture. Thus, Serre's conjecture implies Fermat's last theorem.

5.2 The modularity theorem

Let $f \in S_2(N, \mathbf{1})$ be a normalised eigenform of weight 2 and level N with trivial character. Since f is a normalised eigenform, its Fourier coefficients are algebraic integers.

Suppose that the Fourier coefficients of f are rational integers. Then there exists a holomorphic map from the upper half plane

$$\mathcal{H} \cup \{\text{cusps}\} \longrightarrow \mathbf{C}/\Lambda_f$$

where Λ_f is a rank 2 lattice in $\mathbf{C}$, and hence $E = \mathbf{C}/\Lambda_f$ is an elliptic curve. In fact, this curve can be defined over $\mathbf{Q}$, and has a good reduction for all primes $p \nmid N$. We obtain a map

$$\left\{\begin{array}{c}\text{Normalised eigenforms}\\ \text{of level } N \text{ with}\\ \text{coefficients in } \mathbf{Z}\end{array}\right\} \rightsquigarrow \left\{\begin{array}{c}\text{Elliptic curves over}\\ \mathbf{Q} \text{ of conductor } N\end{array}\right\}$$

The modularity theorem states that every elliptic curve over $\mathbf{Q}$ arises in this way:

Theorem 5.5 (Modularity Theorem)**.** *Let E be an elliptic curve over $\mathbf{Q}$ with conductor N. Then there exists a normalised eigenform $f = \sum_{n=1}^{\infty} a_n(f)q^n \in S_2(N, \mathbf{1})$, such that E is isogenous to $\mathbf{C}/\Lambda_f$.*

Proof. For simplicity, we will prove this in the case that E is semistable. The proof of the general case is almost identical, but requires more care when defining the weight and level of the representation $\overline{\rho}_{E,p}$, and in determining when it is irreducible.

Let p be a prime number not dividing N, and let

$$\overline{\rho}_{E,p} : G_{\mathbf{Q}} \longrightarrow \mathrm{GL}_2(\mathbf{F}_p)$$

be the Galois representation arising from the action of $G_{\mathbf{Q}}$ on $E[p]$. In order to apply Serre's conjecture to $\overline{\rho}_{E,p}$, we need it to be irreducible. Since E is semistable, the same proof as that in Corollary 5.4 shows that $\overline{\rho}_{E,p}$ will be irreducible if $p > 7$.

Let Δ_E be the minimal discriminant of E. The results of §4.1 show that for all but finitely many primes, the Serre invariants attached to $\overline{\rho}_{E,p}$ will be $(N, 2, \mathbf{1})$.

Serre's conjecture says that there exists a cuspidal eigenform

$$f_p = \sum_{n=1}^{\infty} a_{n,p}\ q^n$$

of type $(N, 2, \mathbf{1})$ with coefficients in $\overline{\mathbf{F}}_p$ such that $\rho_f \sim \overline{\rho}_{E,p}$. This will in turn arise as a reduction mod p of a cuspidal eigenform

$$F_p = \sum_{n=1}^{\infty} A_n q^n \in S_2(N, \mathbf{1}).$$

The crucial fact is that the weight and level of these modular forms are independent of p. Since, there are only a finite number of normalised cuspidal eigenforms of weight 2 and level N, there exists an infinite set of primes $\mathcal{P}$ and a choice of F such that

$$\widetilde{F} = f_p \quad \text{for all } p \in \mathcal{P}.$$

We show that F has coefficients in $\mathbf{Z}$. Let ℓ be a prime not dividing N, so that the curve E has good reduction at ℓ. Let

$$a_\ell = \ell + 1 - \#E(\mathbf{F}_\ell)$$

It follows from the corresponding result for p-adic representations (see [DS05] Theorem 9.4.1) that

$$a_\ell \equiv \mathrm{tr}\ \overline{\rho}_{E,p}(\mathrm{Frob}_\ell) \pmod{p} \quad \text{for all } p \neq \ell,$$

and hence by Deligne's theorem (Theorem 1.12) that

$$a_\ell \equiv a_{\ell,p} \pmod{p} \quad \text{for all } p \neq \ell.$$

Hence, for every prime $p \in \mathcal{P}$ with $p \neq \ell$, the algebraic integer $A_\ell - a_\ell$ has image in $\overline{\mathbf{F}}_p$ equal to 0. Since $\mathcal{P}$ is infinite, this means that

$$A_\ell = a_\ell \quad \text{for all } \ell \nmid N. \tag{5.1}$$

In particular, the Fourier coefficients of F lie in $\mathbf{Z}$, and hence, we can associate to F an elliptic curve $E_F/\mathbf{Q}$.

One can show that E and E_F are actually isogenous. Indeed, equation (5.1) shows that if $\rho_{E,p}$ and $\rho_{E_F,p}$ are the p-adic Galois representations attached to E and E_F, then both have the same characteristic equation and are therefore equivalent. This proves strong version R of the modularity theorem in [DS05] (see Theorem 9.6.3 there). The proof that E and E_F are isogenous follows from this. □

5.3 The equation $Ax^n + By^n = Cz^n$

In solving Diophantine equations, methods based on Serre's conjecture usually involve proving that a particular equation has no solutions, by constructing a modular form that does not exist. As a result, it is much harder to apply techniques based on Serre's conjecture to equations such as

$$Ax^n + Bx^n = Cz^n \quad A, B, C \in \mathbf{Z},\ n \geq 3,$$

where there should be only finitely many solutions, but for example, $(-1)^5 + 2^5 = 31 \cdot 1^5$, so non-trivial solutions do exist.

In this section, by combining Serre's conjecture with other conjectural results about Galois representations, we will show that the above equation has only finitely many solutions (x, y, z, n) where $n > 3$. For a fixed choice $n > 3$, it follows from the Faltings' theorem on curves of genus $g > 1$ that there are only finitely many solutions. By combining Serre's conjecture with the following conjecture due to Frey ([Dar95] Conjecture 4.3), we can prove the stronger result where n is not fixed.

Conjecture 5.6 (Frey)**.** *Let $A/\mathbf{Q}$ be an elliptic curve. There are only finitely many pairs (E, p) consisting of an elliptic curve $E/\mathbf{Q}$ which is not isogenous to A and a prime number $p > 5$, such that the corresponding representations*

$$\overline{\rho}_{E,p} : G_{\mathbf{Q}} \longrightarrow \mathrm{Aut}(E[p])$$

and

$$\overline{\rho}_{A,p} : G_{\mathbf{Q}} \longrightarrow \mathrm{Aut}(A[p])$$

are equivalent.

Theorem 5.7. *Assume Serre's conjecture and Frey's conjecture. Then the equation*

$$Ax^n + By^n = Cz^n, \quad n > 3, \quad \gcd(x, y, z) = 1$$

has only finitely many integer solutions (x, y, z, n).

Sketch of proof. Suppose for contradiction that there are infinitely many solutions. Since for each fixed $n > 3$, the equation has only finitely many solutions, we can construct an infinite set of solutions of the form (a_i, b_i, c_i, p_i) where the $p_i \geq 5$ are distinct primes. Moreover, we can assume that for each i, $p_i \nmid 2ABC$.

The idea of the proof is to attach to each solution (a_i, b_i, c_i, p_i) a corresponding Frey curve

$$E_i : y^2 = x(x - Aa_i^{p_i})(x + Bb_i^{p_i}).$$

Unlike in the proof of Fermat's last theorem, the lack of symmetry means that we cannot assume that E_i is semistable, but a generalisation of the results in §4.1 (see for example [Cai09] Remark 6.2.2) allows us to show that the level $N(\overline{\rho}_{E_i,p_i})$ of the corresponding mod p_i Galois representation divides $32(ABC)^2$, and that this condition is independent of i. Moreover, for p sufficiently large (indeed $p > 163$ will suffice), $\overline{\rho}_{E_i,p_i}$ will be irreducible.

The proof now proceeds similarly to the proof of the modularity theorem. Using Serre's conjecture, we obtain from each $\overline{\rho}_{E_i,p_i}$ a corresponding mod p_i eigenform f_i, which itself arises from a normalised eigenform F_i with coefficients in $\overline{\mathbf{Z}}_p$. The level of F_i must divide $32(ABC)^2$, and since the set of such eigenforms is finite, we deduce that there is an eigenform F whose reduction mod p_i is equal to f_i for infinitely many i.

As before, we show that F has integer coefficients. Indeed, for $\ell \nmid 2ABC$, the curve E_i has either a good or multiplicative reduction mod ℓ. If it has a good reduction, then the Hasse bound gives

$$|a_\ell(E_i)| = |\ell + 1 - \#E_i(\mathbf{F}_\ell)| \leq 2\sqrt{\ell},$$

and if it has a multiplicative reduction, then

$$\widetilde{a_\ell(E_i)} = \pm(\ell + 1)$$

in $\overline{\mathbf{F}}_p$, where $\tilde{z}$ denotes the reduction map defined in equation (1.5). Hence, as i varies, $\widetilde{a_\ell(E_i)}$ can only take finitely many values, and therefore, it will take some value a infinitely many times. Since $\widetilde{a_\ell(E_i)} = a_\ell(f_i)$, it follows that

$$\widetilde{a_\ell(f)} = a$$

for infinitely many i, and therefore that $a_\ell(f) = a$ is an integer.

Hence, F has a corresponding elliptic curve $E_F/\mathbf{Q}$, and

$$E_F[p_i] \cong E_i[p_i]$$

as $G_{\mathbf{Q}}$ modules. This contradicts Frey's conjecture. □

References

[AT68] Emil Artin and John Tate. *Class Field Theory.* American Mathematical Society, 1968.

[Cai09] Bryden Cais. Serre's conjectures. *http://math.arizona.edu/∼cais/Papers/Expos/Serre05.pdf*, 2009.

[Car89] Henri Carayol. Sur les représentations galoisiennes modulo l attachées aux formes modulaires. *Duke Math. J.*, pages 785–801, 1989.

[CSS97] Gary Cornell, Joseph Silverman, and Glenn Stevens. *Modular forms and Fermats last theorem.* Springer Science & Business Media, 1997.

[Dar95] Henri Darmon. Serre's conjectures. *Seminar on Fermat's Last Theorem (Toronto, ON, 1993-1994)*, pages 135–153, 1995.

[DDT94] Henri Darmon, Fred Diamond, and Richard Taylor. Fermat's last theorem. *Current developments in mathematics, International Press*, pages 1–154, 1994.

[DS74] Pierre Deligne and Jean-Pierre Serre. Formes modulaires de poids 1. *Ann. Scient. École Norm. Sup. (4)*, pages 507–530, 1974.

[DS05] Fred Diamond and Jerry Shurman. *A First Course in Modular Forms.* Springer-Verlag, New York, 2005.

[Edi92] Bas Edixhoven. The weight in serre's conjectures on modular forms. *Invent. Math.*, pages 563–594, 1992.

[Kat76] Nick Katz. A result on modular forms in characteristic p. *Lecture notes in Math.*, pages 53–61, 1976.

[Liv89] Ron Livné. On the conductors of mod/galois representations coming from modular forms. *Journal of Number Theory*, pages 133–141, 1989.

[Maz78] Barry Mazur. Rational isogenies of prime degree. *Invent. Math.*, pages 129–162, 1978.

[RS99] Kenneth Ribet and William Stein. Lectures on serre's conjectures. *Arithmetic algebraic geometry (Park City, UT, 1999)*, 9:143–232, 1999.

[Ser72] Jean-Pierre Serre. Propriétés galoisiennes des points d'ordre fini des courbes elliptiques. *Invent. Math*, 15.4:259–331, 1972.

[Ser77] Jean-Pierre Serre. *Linear Representations of Finite Groups.* Springer-Verlag, New York, 1977.

[Ser79] Jean-Pierre Serre. *Local Fields.* Springer-Verlag, New York, 1979.

[Ser87] Jean-Pierre Serre. Sur les représentations modulaires de degré 2 de $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$. *Duke Math. J.*, pages 179–230, 1987.