Serre's conjecture

Alex J. Best

June 2015

Contents

1 Introduction

In 1987 Jean-Pierre Serre published a paper [Ser87], "Sur les représentations modulaires de degré 2 de $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$", in the Duke Mathematical Journal. In this paper Serre outlined a conjecture detailing a precise relationship between certain mod p Galois representations and specific mod p modular forms. This conjecture and its variants have become known as Serre's conjecture, or sometimes *Serre's modularity conjecture* in order to distinguish it from the many other conjectures Serre has made. The conjecture has since been proven correct by the work of numerous people, culminating with that of Khare–Wintenberger and Kisin, published in 2009 [KW09a, KW09b, Kis09].

Here we provide a motivated account of the original form of the conjecture before going on to compute some explicit examples and examining some interesting consequences.

Beyond Serre's paper there are many very good accounts of his statement, including articles by Darmon [Dar95] and by Ribet and Stein [RS99] who also have a book chapter on the topic [RS11]. In 1992 Bas Edixhoven wrote a paper [Edi92] in which he gave a slight alteration of Serre's conjecture. There are additional articles covering this version of the conjecture, such as Cais [Cai09] and Edixhoven [Edi97]. Finally Alex Ghitza has prepared a translation of part of Serre's paper [Ghi] which has been helpful. These articles were of great help when preparing the current essay and many of the ideas used here are contained in at least one of them.

2 Background

We begin by fixing several definitions and key results that will be relevant when discussing Serre's conjecture.

2.1 Modular forms

In the interests of space we assume material relating to classical modular forms. Here we only look at the passage to *mod p modular forms*, as there is some amount of choice in how these forms are defined and they form a key part of Serre's conjecture. We use Serre and Swinnerton-Dyer's original approach to define mod p modular forms, this is detailed in [SD73, Ser73a, Ser73b]. This approach is a simple one and it allows us to talk about

congruences involving q-expansions and their coefficients without getting too bogged down with notation.

Definition 2.1.1. Given a subring R of $\mathbf{C}$ we let $S_k(N, \varepsilon; R)$ be the space of cusp forms of weight k, level N and character $\varepsilon\colon (\mathbf{Z}/N\mathbf{Z})^* \to R$, whose q-expansion coefficients lie in R.

Now fix a prime $p \nmid N$. Given a character

$$\varepsilon\colon (\mathbf{Z}/N\mathbf{Z})^* \to \overline{\mathbf{F}}_p^*$$

we may lift to a character

$$\hat{\varepsilon}\colon (\mathbf{Z}/N\mathbf{Z})^* \to \overline{\mathbf{Z}}^*,$$

with values in the prime to p roots of unity.

Additionally fix a place v of $\overline{\mathbf{Q}}$ above p, this place gives us a reduction map $\overline{\mathbf{Z}} \to \overline{\mathbf{F}}_p$. We let the space of *cuspidal mod p modular forms* of weight k, level N and character $\varepsilon\colon (\mathbf{Z}/N\mathbf{Z})^* \to \overline{\mathbf{F}}_p^*$ be the subspace of $\overline{\mathbf{F}}_p[[q]]$ obtained by reducing the q-expansions of forms in $S_k(N, \hat{\varepsilon}; \overline{\mathbf{Z}})$ using our place v. We denote this space by

$$S_k(N, \varepsilon; \overline{\mathbf{F}}_p).$$

Taking the union over all characters ε as above gives us the space of all mod p cusp forms of weight k and level N,

$$S_k(N; \overline{\mathbf{F}}_p).$$

We can in the same way define the full (non-cuspidal) space of mod p modular forms, along with mod p modular forms for more general congruence subgroups. But we don't need to consider such forms in this essay so we restrict to cusp forms for Γ_1 to keep things concise.

Many notions defined for normal modular forms descend to mod p modular forms in the natural way.

Definition 2.1.2. The standard Hecke operators act on q-expansions in a way that preserves each space $S_k(N, \varepsilon; \overline{\mathbf{Z}})$ (this action is recalled in the proof of **??** 2.1.9). So we may define the action of the Hecke operators on a mod p modular form f by letting them act on a lift of f and then reducing the q-expansion mod p again. Equivalently we may just define the action on q-expansions to be given by the same formulae as for normal modular forms.

Definition 2.1.3. As for standard modular forms, we say a mod p cusp form $f = \sum_{n \geq 1} a_n q^n$ is *normalised* if $a_1 = 1$.

An interesting difference between the world of standard modular forms and that of mod p modular forms comes from the fact that distinct modular forms may reduce mod p to have exactly the same q-expansion.

Example 2.1.4. Using Sage [S$^+$15] we find the following example, let

$$f = q - q^2 - 2q^3 - 7q^4 + 16q^5 + 2q^6 - 7q^7 + O(q^8) \in S_4(7, \mathrm{Id}; \mathbf{Z}),$$
$$g = q - 6q^2 - 42q^3 - 92q^4 - 84q^5 + 252q^6 + 343q^7 + O(q^8) \in S_8(7, \mathrm{Id}; \mathbf{Z}),$$

then if we reduce mod 5 we see that

$$\overline{f} = q + 4q^2 + 3q^3 + 3q^4 + q^5 + 2q^6 + 3q^7 + O(q^9) \in S_4(7, \mathrm{Id}; \overline{\mathbf{F}}_5),$$
$$\overline{g} = q + 4q^2 + 3q^3 + 3q^4 + q^5 + 2q^6 + 3q^7 + O(q^9) \in S_8(7, \mathrm{Id}; \overline{\mathbf{F}}_5),$$

which are indeed equal up to this precision.

This means that we need to be careful if we are just given a q-expansion corresponding to some mod p modular form as there is no longer necessarily a single weight attached that form. There is a useful necessary condition controlling when this sort of behaviour can happen however.

Proposition 2.1.5. *If f and g are two non-zero mod p modular forms of weights k and k' respectively, whose q-expansions are equal, then*

$$k \equiv k' \pmod{p-1}.$$

Proof. See [Ser73a]. □

In fact for $p \geq 5$ it is always the case that $S_k(N; \overline{\mathbf{F}}_p) \subset S_{k+p-1}(N; \overline{\mathbf{F}}_p)$ [Ser73b]. Due to this behaviour the concept of weight is not particularly well defined for mod p modular forms, so we introduce the notion of a *filtration* to help us deal with the ambiguity.

Definition 2.1.6. The *filtration* of a mod p cusp form f of level N is the minimal $k \geq 0$ for which $f \in S_k(N; \overline{\mathbf{F}}_p)$. We denote this by $w(f)$.

Now we look at an important operator on the space of mod p modular forms, which we shall study more in Section 4.3.

Definition 2.1.7. The Θ operator is defined on (cuspidal) mod p modular forms via its action on q-expansions, which is

$$\Theta\left(\sum_{n\geq 0} a_n q^n\right) = q\frac{\mathrm{d}}{\mathrm{d}q}\left(\sum_{n\geq 0} a_n q^n\right) = \sum_{n\geq 0} n a_n q^n.$$

It is not clear from this definition that the resulting q-expansion should always correspond to another modular form, but this is indeed the case. In fact many of the properties of the cusp form are preserved by Θ.

Proposition 2.1.8. *If f is a mod p cusp form of filtration $w(f) = k$, then $\Theta(f)$ is also a mod p cusp form of the same level and character and has filtration*

$$w(\Theta(f)) = \begin{cases} k+p+1 & \text{if } p \nmid k, \\ k+p+1-n(p-1),\ n \geq 1 & \text{if } p \mid k. \end{cases}$$

Proof. See [Ser73a] and also [Joc82] for more detail about how the filtration lowers in the $p \mid k$ case. □

We see straight away from the definition of the action that Θ preserves the set of normalised mod p cusp forms.

Proposition 2.1.9. *Θ semicommutes with the Hecke operators T_ℓ (specifically we have $T_\ell\Theta = \ell\Theta T_\ell$), and hence Θ preserves eigenforms.*

Proof. The Hecke operators T_ℓ on $S_k(N, \varepsilon; \overline{\mathbf{F}}_p)$ act on q-expansions by

$$T_\ell\left(\sum_{n\geq 1} a_n q^n\right) = \begin{cases} \sum_{n\geq 1} a_{\ell n} q^n + \ell^{k-1}\varepsilon(\ell)\sum_{n\geq 1} a_n q^{\ell n} & \text{if } \ell \nmid N, \\ \sum_{n\geq 1} a_{\ell n} q^n & \text{if } \ell \mid N. \end{cases}$$

We let $f = \sum_{n\geq 1} a_n q^n \in S_k(N, \varepsilon; \overline{\mathbf{F}}_p)$ and calculate

$$\Theta T_\ell f = \begin{cases} \sum_{n\geq 1} n a_{\ell n} q^n + \ell^{k-1}\varepsilon(\ell)\sum_{n\geq 1} \ell n a_n q^{\ell n} & \text{if } \ell \nmid N, \\ \sum_{n\geq 1} n a_{\ell n} q^n & \text{if } \ell \mid N, \end{cases}$$

and

$$T_\ell\Theta f = \begin{cases} \sum_{n\geq 1} \ell n a_{\ell n} q^n + \ell^{k+p+1-1}\varepsilon(\ell)\sum_{n\geq 1} n a_n q^{\ell n} & \text{if } \ell \nmid N, \\ \sum_{n\geq 1} \ell n a_{\ell n} q^n & \text{if } \ell \mid N. \end{cases}$$

As we are working in characteristic p here $\ell^{k+p} = \ell^{k+1}$, so

$$T_\ell\Theta = \ell\Theta T_\ell,$$

thus if f is an eigenform for the T_ℓ then Θf is an eigenform too. However the eigenvalue for each T_ℓ is ℓ times the original. □

2.2 Galois representations

We now move on to the next major component of Serre's conjecture, Galois representations. As above we mostly concern ourselves with fixing definitions and recalling important results that shall be needed later. There are many good references for this type of material, for example [DDT95, Wie12].

Definition 2.2.1. An *n-dimensional mod p Galois representation* is a homomorphism

$$\rho\colon \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathrm{GL}_n(\overline{\mathbf{F}}_p).$$

Similarly, an *n-dimensional p-adic Galois representation* is a homomorphism

$$\rho\colon \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathrm{GL}_n(\overline{\mathbf{Q}}_p).$$

Unless stated otherwise the term Galois representation will refer to a mod p Galois representation.

Recall that $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ is defined as the inverse limit of $\mathrm{Gal}(K/\mathbf{Q})$ as K ranges over all number fields. So the group $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ naturally has the profinite topology, where the open subgroups are the subgroups of finite index. We demand that all of our mod p representations be continuous with respect to this topology and the discrete topology on $\mathrm{GL}_n(\overline{\mathbf{F}}_p)$.

Remark 2.2.2. The continuity condition for mod p Galois representations reduces to having an open kernel, so continuous mod p Galois representations always have finite image.

We deal mostly with 1 and 2 dimensional mod p Galois representations. Those of dimension 1 (i.e. maps $\phi\colon \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \overline{\mathbf{F}}_p^*$) are called *characters*. Given a 2-dimensional mod p representation $\rho\colon G \to \mathrm{GL}_2(\overline{\mathbf{F}}_p)$ we often use the notation

$$\rho \sim \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix},$$

where α, β, γ and δ are functions $G \to \overline{\mathbf{F}}_p$, to indicate that there is some $A \in \mathrm{GL}_2(\overline{\mathbf{F}}_p)$ such that for every $\sigma \in G$

$$\rho(\sigma) = A \begin{pmatrix} \alpha(\sigma) & \beta(\sigma) \\ \gamma(\sigma) & \delta(\sigma) \end{pmatrix} A^{-1}.$$

Definition 2.2.3. Let ρ be a mod p Galois representation and ϕ be a mod p Galois character. We can form a new mod p Galois representation of the same dimension as ρ by taking the product of the images for each element of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$. This is called the *twist* of ρ by ϕ, and is denoted $\phi \otimes \rho$.

Definition 2.2.4. Let $\phi\colon \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to K^*$ be a character for some field K and fix an embedding $\overline{\mathbf{Q}} \hookrightarrow \mathbf{C}$. We may then view complex conjugation as an element $c \in \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$, looking at its image $\phi(c)$ we see it is an element of order 2 in K^*, so $\phi(c)$ must be ± 1. If $\phi(c) = -1$ we say ϕ is *odd*, otherwise we say ϕ is *even* (though we shall mostly be concerned with distinguishing odd representations here).

Now given any Galois representation

$$\rho\colon \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathrm{GL}_n(K),$$

we define the parity of ρ to be that of the character $\det \rho$.

Due to **??** 2.2.2 all of our Galois representations will factor through finite extensions, and so our understanding of finite Galois groups will be essential in what follows. The next few definitions and results will allow us to talk about the properties of finite Galois extensions of $\mathbf{Q}$ and $\mathbf{Q}_p$.

Definition 2.2.5. Given a finite Galois extension $K/\mathbf{Q}$ we may define a series of subgroups of $G = \mathrm{Gal}(K/\mathbf{Q})$ that measure the ramification of $K/\mathbf{Q}$ at a prime ℓ. Let $\mathfrak{L}$ be a prime of K above ℓ and let

$$D_\ell = \{\sigma \in G : \sigma(\mathfrak{L}) \subset \mathfrak{L}\}$$

be the *decomposition group* at ℓ of G. Now take i to be a non-negative integer, we define the ith *higher ramification group* to be

$$G_{\ell,i} = \{\sigma \in D_\ell : \sigma(x) - x \in \mathfrak{L}^{i+1}\ \forall x \in \mathcal{O}_K\}.$$

These groups tell us how much of the whole of G we are seeing if we just look modulo powers of $\mathfrak{L}$. They form a descending chain as i increases

$$D_\ell \supseteq G_{\ell,0} \supseteq G_{\ell,1} \supseteq G_{\ell,2} \supseteq \cdots.$$

Now we recall a result of Kummer theory, which tells us how to obtain abelian Galois extensions of certain fields, see [Bir67] for the full story.

Proposition 2.2.6. *Suppose K is a field containing all nth roots of unity, then there is a bijection*

$$\{\text{subgroups } H \subset K^*/(K^*)^n\} \leftrightarrow \{\text{abelian extensions } L|K : \mathrm{Gal}(L/K)^n = 1\},$$

defined by sending $H \mapsto K(\sqrt[n]{H})$ and $K^* \cap (L^*)^n \hookleftarrow L$. The notation $K(\sqrt[n]{H})$ simply means we adjoin nth roots of all elements of H to K.

Although our main objects of study are representations $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ it will be very useful for us to take a prime ℓ and also consider representations of

$$\mathrm{G}_\ell = \mathrm{Gal}(\overline{\mathbf{Q}}_\ell/\mathbf{Q}_\ell).$$

Such representations can be obtained from those of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ using an inclusion

$$\overline{\mathbf{Q}} \hookrightarrow \overline{\mathbf{Q}}_\ell$$

to define a restriction map

$$\mathrm{G}_\ell \to \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}).$$

In fact due to Krasner's lemma [Coh08, p. 238] the map $\mathrm{G}_\ell \to \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ is injective and so we may view G_ℓ as a subgroup of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$. The way this subgroup sits inside $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ depends on the choice of embedding $\overline{\mathbf{Q}} \hookrightarrow \overline{\mathbf{Q}}_\ell$ and varies by conjugation as this embedding changes.

The group G_ℓ has several subquotients which will be helpful for us to study restrictions of representations to.

Definition 2.2.7. The ring of integers of $\overline{\mathbf{Q}}_\ell$ is stable under the action of G_ℓ, as is its unique maximal ideal. So we get an action of G_ℓ on the residue field, this field may be identified with $\overline{\mathbf{F}}_p$. We therefore obtain a map

$$\mathrm{G}_\ell \to \mathrm{Gal}(\overline{\mathbf{F}}_\ell/\mathbf{F}_\ell).$$

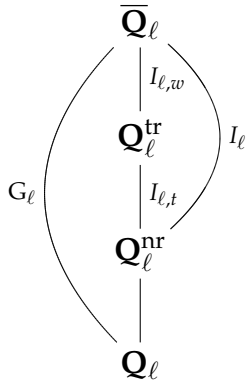
The *inertia subgroup* I_ℓ is defined to be the kernel of this map.

The group $\mathrm{Gal}(\overline{\mathbf{F}}_\ell/\mathbf{F}_\ell)$ is topologically cyclic, generated by the Frobenius morphism $x \mapsto x^\ell$. In fact the above map $\mathrm{G}_\ell \to \mathrm{Gal}(\overline{\mathbf{F}}_\ell/\mathbf{F}_\ell)$ is surjective and so we may let $\mathrm{Frob}_\ell \in \mathrm{G}_\ell$ be a preimage of this morphism under the restriction map, we call such elements *Frobenius elements*.

Next the *wild* inertia group $I_{\ell,w}$ is the maximal pro-ℓ-subgroup of I_ℓ and the *tame* inertia group is the quotient

$$I_{\ell,t} = I_\ell / I_{\ell,w}.$$

The groups we have been looking at give rise to Galois extensions of $\mathbf{Q}_\ell$. The fixed field of I_ℓ is the maximal non-ramified extension of $\mathbf{Q}_\ell$, denoted $\mathbf{Q}_\ell^{\mathrm{nr}}$. Similarly, the fixed field of $I_{\ell,w}$ is the maximal tamely-ramified extension of $\mathbf{Q}_\ell$, denoted $\mathbf{Q}_\ell^{\mathrm{tr}}$. So the full set-up looks as follows



The field $\mathbf{Q}_\ell^{\mathrm{tr}}$ is generated by the fields $\mathbf{Q}_\ell^{\mathrm{nr}}(\sqrt[n]{\ell})$ [Frö67] and so, using **??** 2.2.6 the tame inertia $I_{\ell,t}$ may be identified with

$$\varprojlim \mathrm{Gal}(\mathbf{Q}_\ell^{\mathrm{nr}}(\sqrt[n]{\ell})/\mathbf{Q}_\ell^{\mathrm{nr}}) = \varprojlim \mathbf{F}_{\ell^n}^*,$$

where the maps in the last inverse limit are the norm maps.

Definition 2.2.8. We say a Galois representation ρ is *unramified* at ℓ if $\rho|_{I_\ell}$ is trivial. Otherwise, we say ρ is *ramified* at ℓ. Similarly we say ρ is *tamely ramified* at ℓ if $\rho|_{I_{\ell,w}}$ is trivial.

The usefulness of the Frobenius elements for us stems in part from the following theorem.

Theorem 2.2.9. *Let S be a finite set of primes, then:*

1. *A semi-simple mod p representation*

$$\rho\colon \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathrm{GL}_2(\overline{\mathbf{F}}_p)$$

is determined by the values of $\mathrm{tr}\,\rho(\mathrm{Frob}_\ell)$ *and* $\det\rho(\mathrm{Frob}_\ell)$ *for all $\ell \notin S$ at which ρ is unramified.*

2. *A mod p character*

$$\phi\colon \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \overline{\mathbf{F}}_p^*$$

is determined by the values of $\phi(\mathrm{Frob}_\ell)$ for all $\ell \notin S$ at which ϕ is unramified.

Proof. See [DDT95, prop. 25]. □

Definition 2.2.10. Each character

$$\phi\colon \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \overline{\mathbf{F}}_p^*$$

has finite image and so factors through some $\mathbf{F}_{p^n}^*$, the smallest n for which this can happen is called the *level* of the character.

For each $n \geq 1$ we now distinguish n special mod p characters of $I_{p,t}$ of level n, these will allow us to describe all such characters of a particular level.

Definition 2.2.11. The identification

$$I_{p,t} = \varprojlim \mathbf{F}_{p^n}^*$$

gives us a natural map

$$\psi_n\colon I_{p,t} \to \mathbf{F}_{p^n}^*$$

for each n. The *fundamental characters* of level n are defined by extending ψ_n to an $\overline{\mathbf{F}}_p$-character of $I_{p,t}$ via the n embeddings $\mathbf{F}_{p^n} \hookrightarrow \overline{\mathbf{F}}_p$.

While any individual fundamental character is not canonically defined, the set of all of them of a particular level is.

Remark 2.2.12. The embeddings are all obtained from any chosen one by applying Frobenius and as such the product of all fundamental characters of level n is the same as the composition of the norm map $\mathbf{F}_{p^n}^* \to \mathbf{F}_p^*$ with any one. So this product will always be the unique fundamental character of level 1.

Proposition 2.2.13. *The fundamental characters of level n generate the set of all characters of level n.*

Proof. See [Ser72, prop. 5]. □

As any character $\phi\colon \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \overline{\mathbf{F}}_p^*$ factors through an abelian quotient, the Kronecker–Weber theorem tells us that any such character factors as

$$\phi\colon \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathrm{Gal}(\mathbf{Q}(\zeta_N)/\mathbf{Q}) \cong (\mathbf{Z}/N\mathbf{Z})^* \xrightarrow{\phi'} \overline{\mathbf{F}}_p^*,$$

where ζ_N is a primitive Nth root of unity. We can also use this factorisation to extend any Dirichlet character to a character of the absolute Galois group. Thus characters of the absolute Galois group are in bijection with Dirichlet characters

$$(\mathbf{Z}/N\mathbf{Z})^* \xrightarrow{\phi} \overline{\mathbf{F}}_p^*.$$

Definition 2.2.14. The identity map

$$(\mathbf{Z}/p\mathbf{Z})^* \to \mathbf{F}_p^* \hookrightarrow \overline{\mathbf{F}}_p^*$$

is a Dirichlet character and thus gives us a character of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$. This character is called the *mod p cyclotomic character*, and denoted χ_p.

Remark 2.2.15. Taking any $\sigma \in \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ the definition above is saying that σ sends

$$\zeta_p \mapsto \zeta_p^{\chi_p(\sigma)},$$

where ζ_p is a primitive pth root of unity. If we take some $\ell \neq p$ prime and denote reduction mod ℓ of $\mathbf{Z}[\zeta_p]$ by $\overline{\cdot}$ we have that

$$\overline{\mathrm{Frob}_\ell(\zeta_p)} = \overline{\zeta_p}^{\ell},$$

and so as χ_p takes values mod p we find that

$$\chi_p(\mathrm{Frob}_\ell) = \ell.$$

Now if we fix an embedding $\overline{\mathbf{Q}} \hookrightarrow \mathbf{C}$ and consider complex conjugation as some element $c \in \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ we see that it takes $\zeta_p \mapsto \zeta_p^{-1}$ and hence

$$\chi_p(c) = -1,$$

so χ_p is an example of an odd character.

Finally, χ_p was defined to factor through $\mathrm{Gal}(\mathbf{Q}(\zeta_p)/\mathbf{Q})$, this extension is tamely ramified and hence so is χ_p. Therefore χ_p passes to a character on the group $I_p/I_{p,w} = I_{p,t}$. Straight from the definition we see that χ_p, when viewed in this way, is an example of a level 1 character of $I_{p,t}$, but in fact more is true.

Proposition 2.2.16. *The fundamental character of level 1 is the cyclotomic character χ_p (or rather its induced character on $I_{p,t}$).*

Proof. See [Ser72, prop. 8]. □

Definition 2.2.17. We now classify Galois representations

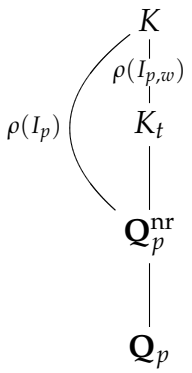
$$\rho\colon \mathrm{G}_p \to \mathrm{GL}_2(\overline{\mathbf{F}}_p)$$

of the form

$$\rho \sim \begin{pmatrix} \chi_p\varepsilon_1 & * \\ 0 & \varepsilon_2 \end{pmatrix},$$

for some unramified characters ε_1 and ε_2.

Let K be the fixed field of the kernel of ρ. There is a unique maximal tamely ramified extension K_t of $\mathbf{Q}_p^{\mathrm{nr}}$ that is contained inside of K. These fields fit together as follows



As

$$\mathrm{Gal}(K_t/\mathbf{Q}_p^{\mathrm{nr}}) = (\mathbf{Z}/p\mathbf{Z})^*$$

we may write

$$K_t = \mathbf{Q}_p^{\mathrm{nr}}(z),$$

where z is a primitive pth root of unity. If we look at $\mathrm{Gal}(K/K_t)$ we see that

$$\mathrm{Gal}(K/K_t) = \rho_p(I_{p,w})$$

consists only of elements of the form

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

So $\mathrm{Gal}(K/K_t)$ is a finite elementary abelian p-group and hence isomorphic to $(\mathbf{Z}/p\mathbf{Z})^m$ for some m. Applying **??** 2.2.6 we see that K is in fact given by

$$K = K_t(x_1^{1/p}, \ldots, x_m^{1/p}),$$

where in fact we may take the $x_i \in \mathbf{Q}_p^{\mathrm{nr}}$. The valuations of these x_i will determine our classification. If

$$v_p(x_i) \equiv 0 \pmod{p}$$

for all i then we say that ρ is *peu ramifé*, otherwise if any of the $v_p(x_i)$ is coprime to p then we say it is *très ramifé*.

Finally we introduce a notion that will allow us to obtain semisimple representations from arbitrary Galois representations.

Definition 2.2.18. The *semisimplification* of a 2-dimensional representation ρ is another representation, denoted ρ^{ss}, that is obtained as follows. If ρ is irreducible (and hence semisimple) we leave it as it is and set $\rho^{\mathrm{ss}} = \rho$. Otherwise if ρ is reducible we know that there is some subspace on which ρ acts via a character and so we may write

$$\rho \sim \begin{pmatrix} \phi_1 & * \\ 0 & \phi_2 \end{pmatrix}.$$

The semisimplification ρ^{ss} is then the representation given by

$$\begin{pmatrix} \phi_1 & 0 \\ 0 & \phi_2 \end{pmatrix},$$

conjugated in the same way ρ was. This is indeed semisimple.

In general the process of semisimplification is analogous, it is obtained by taking the direct sum of the Jordan–Hölder constituents of a representation, though for us the above description suffices.

3 Obtaining Galois representations from modular forms

The two concepts just introduced, modular forms and Galois representation, appear at first glance not to be particularly related to each other. However in reality they are inextricably linked, and exploring some of the links between them will be the goal of the rest of this essay.

3.1 Congruences for Ramanujan's τ function

We will start with a historically important example that provides the first glimpse of the sort of connection we will be looking at in this essay.

Example 3.1.1. Let

$$\Delta = \sum_{n \geq 1} \tau(n) q^n$$

be the unique normalised cusp form of weight 12 for $\Gamma_1(1) = \mathrm{SL}_2(\mathbf{Z})$. The coefficients of this q-expansion were studied in detail by Ramanujan,

who made many influential conjectures concerning them. The function $n \mapsto \tau(n)$ is now known as the Ramanujan τ function. The properties of this function provide the first glimpses of behaviours occurring for more general systems of Hecke eigenvalues.

Various people, including Ramanujan in the mod 691 case, found congruences involving the coefficients $\tau(\ell)$ modulo powers of primes for prime ℓ. Below are a few examples, though others do exist for higher powers of these primes.

$$\tau(\ell) \equiv 1 + \ell^{11} \pmod{2^8}, \text{ if } \ell \neq 2, \tag{1}$$

$$\tau(\ell) \equiv \ell^2 + \ell^9 \pmod{3^3}, \text{ if } \ell \neq 3, \tag{2}$$

$$\tau(\ell) \equiv \ell + \ell^{10} \pmod{5^2}, \tag{3}$$

$$\tau(\ell) \equiv \ell + \ell^4 \pmod{7}, \tag{4}$$

$$\tau(\ell) \equiv \left.\begin{cases} 0 \pmod{23} & \text{if } \left(\frac{\ell}{23}\right) = -1, \\ 2 \pmod{23} & \text{if } \ell \text{ is of the form } u^2 + 23v^2, \\ -1 \pmod{23} & \text{otherwise,} \end{cases}\right\} \text{ if } \ell \neq 23, \tag{5}$$

$$\tau(\ell) \equiv 1 + \ell^{11} \pmod{691}. \tag{6}$$

The original proofs of these congruences were in many cases quite involved and did not all work in the same manner. Indeed Peter Swinnerton-Dyer writes in [SD73] that these proofs "do little to explain why such congruences occur". So in order to try to explain all of the congruences in a unified manner, Serre predicted [Ser67] for each prime p the existence of a p-adic Galois representation

$$\rho_p \colon \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathrm{GL}_2(\mathbf{Q}_p)$$

such that

1. $\mathrm{tr}(\rho_p(\mathrm{Frob}_\ell)) = \tau(\ell)$ for all $\ell \neq p$,
2. $\det(\rho_p(\mathrm{Frob}_\ell)) = \ell^{11}$ for all $\ell \neq p$.

The congruences would then follow from these Galois representations being of specific forms. For example Eqs. (1) to (4) and (6) can all be obtained from these Galois representations if the ρ_p satisfy

$$\rho_p \equiv \begin{pmatrix} \chi_p^a & * \\ 0 & \chi_p^{11-a} \end{pmatrix} \pmod{p^b},$$

where a is 0, 2, 1, 1 or 0 respectively and b is as in the original congruences in each case. Here in each case we can see that $\det \rho_p \equiv \chi_p^{11}$, which is consistent with Item 2 above, and knowing Item 1 would give us the desired congruences.

Serre's prediction for the representation ρ_{23} has a more interesting form, but nevertheless the images of Frobenius elements can be described explicitly. Following Serre we take K to be the splitting field of $x^3 - x - 1$, this is ramified only at 23 and has Galois group S_3. We then let r be the unique irreducible degree 2 representation of S_3 taken with coefficients in $\mathbf{Q}_{23}$, this satisfies

$$\operatorname{tr}(r(\sigma)) = \begin{cases} 0 & \text{if } |\sigma| = 2, \\ 2 & \text{if } |\sigma| = 1, \\ -1 & \text{if } |\sigma| = 3, \end{cases}$$

for each $\sigma \in S_3$. As $\mathrm{Gal}(K/\mathbf{Q})$ is a quotient of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ the representation r extends to a representation of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$. If some ρ_{23} exists satisfying the conditions above it being isomorphic to r then gives rise to Eq. (5) in the same way as before.

Shortly after Serre hypothesised their existence Pierre Deligne constructed the representations ρ_p for all primes p [Del69]. In doing so Deligne also reduced another conjecture of Ramanujan's, that $|\tau(p)| \leq 2p^{11/2}$ for all p, to the Weil conjectures. This provided an indication that relating questions regarding τ to Galois representations is a natural and effective way of working on these questions. Being able to compute these associated representations makes it possible to read off many more congruences for $\tau(n)$ (see, for example, [Mas13]).

3.2 Attaching Galois representations to general eigenforms

Given the above example it is natural to wonder whether such a relationship holds more generally. Indeed Serre also asked if one could associate to each normalised cuspidal eigenform a Galois representation whose traces of Frobenius elements match the q-expansion coefficients. Serre's conjectures on this led to the following more general theorem, which we state for mod p modular forms now.

Theorem 3.2.1 (Deligne)**.** *Let p be a prime, $k \geq 2$, $N \geq 1$ an integer coprime to p and $\varepsilon\colon (\mathbf{Z}/N\mathbf{Z})^* \to \overline{\mathbf{F}}_p^*$. Given a normalised eigenform $f \in S_k(N, \varepsilon; \overline{\mathbf{F}}_p)$ which has q-expansion coefficients a_i, there exists a two-dimensional mod p Galois representation ρ_f such that*

(i) ρ_f *is semi-simple,*

(ii) ρ_f *is unramified outside* Np*,*

(iii) $\mathrm{tr}(\rho_f(\mathrm{Frob}_\ell)) = a_\ell$ *for all* $\ell \nmid Np$*,*

(iv) $\det(\rho_f(\mathrm{Frob}_\ell)) = \varepsilon(\ell)\ell^{k-1}$ *for all* $\ell \nmid Np$*.*

We often refer to the representation ρ_f *as arising from, or being attached to,* f*.*

The construction of these representations in this generality is due to Deligne [Del69], building on work of Shimura and others. There is also a similar statement for weight 1 due to both Deligne and Serre [DS74]. However, as we will discuss Serre's original statement of his conjecture here, we will ignore weight 1 forms. (There is a translation of Deligne's paper available from the IAS [Del04], it has nicer typesetting too.)

In fact the representations obtained in these constructions are p-adic Galois representations $\rho_f\colon \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathrm{GL}_2(\overline{\mathbf{Q}}_p)$, as they were in **??** 3.1.1. The representations of the theorem are then obtained from the p-adic ones via a process of mod p reduction and semisimplification. The mod p representations are the ones that we will be most interested in from here on though.

Remark 3.2.2. Let ρ_f be a Galois representation attached to a normalised eigenform $f \in S_k(N, \varepsilon; \overline{\mathbf{F}}_p)$ as in the theorem. As ρ_f is semisimple and

$$\det(\rho_f(\mathrm{Frob}_\ell)) = \chi_p^{k-1}(\mathrm{Frob}_\ell)\varepsilon(\mathrm{Frob}_\ell)$$

for all $\ell \nmid Np$ (here viewing ε as character of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ now) we may apply **??** 2.2.9 to see that condition (iv) of the theorem is equivalent to the statement that

$$\det \rho_f = \varepsilon\chi_p^{k-1}.$$

By looking at the action of $\langle -1\rangle$ on f we find $\varepsilon(-1)f = \langle -1\rangle f = (-1)^k f$, and so

$$\varepsilon(c)\chi_p^{k-1}(c) = (-1)^k(-1)^{k-1} = -1,$$

hence $\det\rho_f$ must be odd (i.e ρ_f is odd).

We will look at some more properties of this construction in Section 4.3, but first we move on to the conjecture itself.

4 Serre's conjecture

4.1 The qualitative form

Given the above result one might wonder about a converse statement, given a mod p Galois representation satisfying some necessary conditions, does it arise from an eigenform? Serre's conjecture was that the answer to this question is yes, all Galois representations that could possibly arise from an eigenform, based on **??** 3.2.1 and the remark following it, do.

The conjecture naturally comes into two parts, a weaker existence statement, and another refined form that makes exact predictions about an extremal eigenform from which the Galois representation arises. We look at the existence statement, or *qualitative form* first.

Conjecture 4.1.1 (Serre's conjecture, qualitative form)**.** *Let*

$$\rho\colon \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathrm{GL}_2(\overline{\mathbf{F}}_p)$$

be a continuous, odd, irreducible Galois representation. Then there exists a normalised cuspidal mod p eigenform f, such that ρ is isomorphic to ρ_f, the Galois representation associated to f.

This is already a very useful thing to know, any statement one could prove about Galois representations attached to mod p modular forms, by using the theory of these forms for example, would hold for all odd 2-dimensional mod p Galois representations. One interesting consequence of this type stems from the fact that Deligne's construction of Galois representations from modular forms is actually of p-adic representations. If we were to assume **??** 4.1.1 and then apply this construction, we would be able to lift all irreducible odd 2-dimensional mod p Galois representations to p-adic Galois representations.

A similar statement to the one above also holds for reducible representations, which correspond to Eisenstein series instead. We will not consider this more general setup here as it is not what the refined form of Serre's conjecture deals with.

This conjecture (at least for Galois representations unramified outside p) appeared much earlier than the Duke paper and is mentioned by Serre in a 1975 paper [Ser75, sec. 3]. It was computations performed by J.-F. Mestre that convinced Serre that strengthening this conjecture was plausible, and this led to the form we are about to see.

4.2 The refined form

After seeing the qualitative form of Serre's conjecture it is natural to also ask about the properties of the form f whose existence is claimed. Can anything be said about the weight and level of f, based only on the properties of ρ? Serre also conjectured that the answer to this question is yes. He defined a weight, level and character for each ρ, such that there should be a form f of that weight, level and character that ρ is attached to. In a slightly backwards manner we will first state this refined form of the conjecture more precisely, before moving on to motivate and define the integers $N(\rho)$, $k(\rho)$ and character

$$\varepsilon(\rho)\colon (\mathbf{Z}/N(\rho)\mathbf{Z})^* \to \overline{\mathbf{F}}_p^*$$

used in the statement.

Conjecture 4.2.1 (Serre's conjecture, refined form)**.** *Let*

$$\rho\colon \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathrm{GL}_2(\overline{\mathbf{F}}_p)$$

be a continuous, odd, irreducible Galois representation. Then there exists a normalised eigenform

$$f \in S_{k(\rho)}(N(\rho), \varepsilon(\rho); \overline{\mathbf{F}}_p)$$

whose associated Galois representation ρ_f is isomorphic to ρ.

Moreover $N(\rho)$ and $k(\rho)$ are the minimal weight and level for which there exists such a form f.

From now on we refer to a Galois representation ρ satisfying the hypotheses of this conjecture as being of *Serre-type*.

This conjecture is very bold, even given the existence statement of **??** 4.1.1 it is not clear that there should be a simultaneously minimal weight and level, let alone that they should be given by the relatively straightforward (though intricate) description that we will soon see.

If correct however, the refined form is eminently more useful than the qualitative form for specific applications. The precise definition of the level, weight and character often allow one to actually find an associated eigenform and work explicitly with it to study a particular Galois representation. Doing this would be far more difficult in general if no knowledge of the modular forms side could be obtained from the representation.

We will soon give Serre's explicit recipe for the weight, level and character. First however we will look at some results that might lead one to make these definitions in the first place.

4.3 Results on Galois representations associated to modular forms

In order to try and understand which eigenforms can give rise to a particular Galois representation, it is useful to take an arbitrary eigenform and study the properties of the representation attached to it, in an attempt to see what information about the eigenform may be recovered. Several people have obtained interesting results of this type which will be helpful for our definition of the weight and level.

We fix a prime p and a normalised eigenform $f \in S_k(N, \varepsilon; \overline{\mathbf{F}}_p)$ with q-expansion

$$f = \sum_{n \geq 1} a_n q^n.$$

Let ρ_f be the mod p Galois representation attached to f by **??** 3.2.1. Concerning the conductor of ρ_f there is the following result due to Carayol and Livné [Car86, Liv89].

Theorem 4.3.1. *Let $N(\rho_f)$ be the level associated to ρ_f (which we will define explicitly in Section 4.4), then*

$$N(\rho_f) | N.$$

Given this it is natural to hope that any Galois representation ρ of Serre-type arises from a form of level exactly $N(\rho)$ as this is the minimal possible, of course we still have yet to define this quantity!

We can also make useful observations concerning the restriction of ρ_f to G_p, and its subgroups, these have implications for our definition of the weight. There are two main cases here depending on whether $a_p \neq 0$ (the *ordinary* case) or otherwise (the *supersingular* case). We start with the supersingular case.

Theorem 4.3.2 (Fontaine)**.** *Suppose that $2 \leq k \leq p+1$ and that $a_p = 0$ then $\rho_f|_{G_p}$ is irreducible, moreover, letting ψ_1 and ψ_2 be the two fundamental characters of level 2, we have*

$$\rho_f|_{I_p} \sim \begin{pmatrix} \psi_1^{k-1} & 0 \\ 0 & \psi_2^{k-1} \end{pmatrix}.$$

This was originally proved by Fontaine in letters to Serre in 1979. There is a published proof in [Edi92, sec. 6]. So in this case we can certainly recover some information about the weight by looking at the restriction to inertia.

In the ordinary case we have slightly different behaviour.

Theorem 4.3.3 (Deligne). *Suppose* $2 \leq k \leq p+1$ *and* $a_p \neq 0$ *then* $\rho_f|_{G_p}$ *is reducible. Moreover, letting* $\lambda(a)\colon G_p \to \overline{\mathbf{F}}_p^*$ *be the unramified character of* G_p *that takes each* $\mathrm{Frob}_p \in G_p/I_p$ *to some* $a \in \overline{\mathbf{F}}_p^*$*, we have*

$$\rho_f|_{G_p} \sim \begin{pmatrix} \chi_p^{k-1}\lambda(\varepsilon(p)/a_p) & * \\ 0 & \lambda(a_p) \end{pmatrix}.$$

In particular when we look at the restriction to inertia we get

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

A proof of this result when $k \leq p$ is given in [Gro90] and the general case was originally proved in an unpublished letter from Deligne to Serre.

Once again we can recover some information about the weight by looking at the restriction to I_p. As $\chi_p^2 = \chi_p^{p+1}$ eigenforms of weight 2 and $p+1$ give us similar looking representations when restricted to inertia. This makes it more difficult to tell exactly what k was just by looking at restriction to inertia in this case. This will cause us some issues later, fortunately the following result will allow us to distinguish these two cases.

Theorem 4.3.4 (Mazur). *Let* f *have* filtration $p+1$ *and be such that* ρ_f *is irreducible, then* $\rho_f|_{G_p}$ *très ramifé.*

For $p > 2$ and trivial character this is due to Mazur [Rib90, sec. 6]. In [Edi92, sec. 2] Edixhoven gives a modification to the general case.

In all the theorems we have just introduced the weight k is at most $p+1$, of course we can also make Galois representation from higher weight eigenforms. So we need some way of finding out about the representations attached to higher weight forms, using what we know about low weight ones. To do this we use the Θ operator. Which preserves the set of mod p normalised cuspidal eigenforms of a particular level, but in most cases increases the filtration of the eigenform.

We then need to know how the action of Θ affects the associated Galois representation, it turns out that Θ changes these representations in a very simple way.

Proposition 4.3.5. *Let*

$$\Theta\colon S_k(N, \epsilon; \overline{\mathbf{F}}_p) \to S_{k+p+1}(N, \epsilon; \overline{\mathbf{F}}_p)$$

be the operator defined in **??** *2.1.7. Then if* $f \in S_k(N, \epsilon; \overline{\mathbf{F}}_p)$ *is a normalised eigenform the Galois representation associated to* $\Theta(f)$ *is*

$$\rho_{\Theta(f)} \cong \chi_p \otimes \rho_f.$$

Proof. In **??** 2.1.9 we saw that Θ took eigenforms to eigenforms, but with the eigenvalue for each T_ℓ being ℓ times the original. So

$$\mathrm{tr}(\rho_{\Theta(f)}(\mathrm{Frob}_\ell)) = \ell a_\ell = \mathrm{tr}((\chi_p \otimes \rho_f)(\mathrm{Frob}_\ell))$$

and

$$\begin{aligned}\det(\rho_{\Theta(f)}(\mathrm{Frob}_\ell)) &= \ell^{k+p+1}\varepsilon(\ell)\\ &= \ell^{k+1}\varepsilon(\ell)\\ &= \ell^2\ell^{k-1}\varepsilon(\ell)\\ &= \det((\chi_p \otimes \rho_f)(\mathrm{Frob}_\ell)).\end{aligned}$$

By **??** 2.2.9 the representations $\rho_{\Theta(f)}$ and $\chi_p \otimes \rho_f$ are isomorphic. □

So applying Θ twists the associated representation by χ_p and if $p \nmid w(f)$ then Θ increases the filtration by $p+1$. Hence we can use the Θ operator to transfer our understanding of Galois representations attached to modular forms of weight at most $p+1$ to modular forms of higher weight.

It is worth noting that the proofs of some of the above theorems actually came after Serre's paper. However it seems likely that observations of these results in specific examples informed the recipe below.

4.4 The level

Assume that we have a Galois representation $\rho\colon \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathrm{GL}_2(\overline{\mathbf{F}}_p)$ of Serre-type. We now define the integer $N(\rho) \geq 1$ which plays the role of the level in the refined conjecture.

We can view our representation ρ as a homomorphism

$$\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathrm{Aut}(V),$$

where V is a two-dimensional $\overline{\mathbf{F}}_p$ vector space. Let $G \cong \mathrm{im}\,\rho$ be the finite Galois group obtained by quotienting out by the kernel of ρ. Then for each

prime ℓ we let $G_{\ell,i} \subset G$ be the ith ramification group at ℓ of G, as defined in **??** 2.2.5. Now, for each ℓ and i, consider the fixed subspace of V

$$V^{\ell,i} = \{\mathbf{v} \in V : \rho(\sigma)\mathbf{v} = \mathbf{v}\ \forall \sigma \in G_{\ell,i}\}.$$

For each ℓ we define

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

this quantity is (non-trivially) an integer [GS95, p. 99]. We then set the level to be

$$N(\rho) = \prod_{\substack{\ell \neq p \\ \ell \text{ prime}}} \ell^{\nu_\ell(\rho)}.$$

This is indeed a positive integer, and by construction it is coprime to p, this is necessary for it to be the level of a mod p cusp form. This number serves to quantify the higher ramification of a Galois representation at all primes other than p, the larger the ramification groups for a particular prime, the larger the power of that prime in $N(\rho)$. The definition here is almost that of the *Artin conductor* of a representation, but here the p-part is ignored.

Remark 4.4.1. Unwinding this definition when ρ is unramified at some ℓ, we see that each $V^{\ell,i}$ is in fact the whole of V, as all the ramification groups involved are trivial. Hence in this case $\nu_\ell(\rho) = 0$ and so $N(\rho)$ is only divisible by the primes $\ell \neq p$ at which ρ is ramified.

?? 4.3.1 stated that when ρ comes from an eigenform f the integer $N(\rho)$ defined here divides the level of f. With that in mind conjecturing that any Serre-type representation comes from an eigenform of level exactly $N(\rho)$ is fairly logical, though perhaps optimistic without more evidence.

4.5 The character and the weight mod $p - 1$

Beginning with a Galois representation of Serre-type, as before, we now define the character

$$\varepsilon(\rho)\colon (\mathbf{Z}/N(\rho)\mathbf{Z})^* \to \overline{\mathbf{F}}_p^*.$$

We also state the class of $k(\rho)$ mod $p - 1$, though the full definition of $k(\rho)$ will be given in the following section.

Given a Serre-type Galois representation ρ we can compose with the determinant map to obtain a continuous character

$$\det\rho\colon \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \overline{\mathbf{F}}_p^*.$$

The Artin conductor of a 1-dimensional Galois representation (i.e. the level $N(\rho)$ introduced in Section 4.4, but including the p-part now) is actually equal to the conductor of the associated Dirichlet character [GS95, p. 228]. So we now partially compute the Artin conductor of $\det\rho$ in order to find a number field that $\det\rho$ factors through.

Let V_1 be the 2-dimensional vector space realising ρ and V_2 be the 1-dimensional vector space realising $\det\rho$. If $\rho|_{G_{\ell,i}}$ is trivial then $\det\rho|_{G_{\ell,i}}$ is also trivial. So having $\dim(V_2/V_2^{\ell,i}) > 0$ implies that $\dim(V_1/V_1^{\ell,i}) > 0$. As $0 \leq \dim(V_2/V_2^{\ell,i}) \leq 1$ we get that

$$\dim(V_2/V_2^{\ell,i}) \leq \dim(V_1/V_1^{\ell,i})$$

for all ℓ and i and so

$$\nu_\ell(\det\rho) \leq \nu_\ell(\rho).$$

This gives us that

$$N(\det\rho) \mid N(\rho).$$

As the restriction of $\det\rho$ to $I_{p,w}$ is trivial (see the proof of **??** 4.6.1, using that characters are simple) we find that $\nu_p(\det\rho) \leq 1$. So the full Artin conductor of $\det\rho$ divides $pN(\rho)$.

We can therefore identify $\det\rho$ with a homomorphism

$$(\mathbf{Z}/pN(\rho)\mathbf{Z})^* \to \overline{\mathbf{F}}_p^*,$$

or equivalently with a pair of homomorphisms

$$\phi\colon (\mathbf{Z}/p\mathbf{Z})^* \to \overline{\mathbf{F}}_p^*,$$
$$\varepsilon\colon (\mathbf{Z}/N(\rho)\mathbf{Z})^* \to \overline{\mathbf{F}}_p^*.$$

The group $(\mathbf{Z}/p\mathbf{Z})^*$ is cyclic of order $p-1$ and so the image of ϕ lies inside $\mathbf{F}_p^*$. So ϕ is an endomorphism of the cyclic group $(\mathbf{Z}/p\mathbf{Z})^*$ and hence is of the form

$$x \mapsto x^h,$$

for some $h \in \mathbf{Z}/(p-1)\mathbf{Z}$. We have expressed the fact that $\phi = \chi_p^h$, where χ_p is the mod p cyclotomic character.

This gives us the factorisation

$$\det \rho = \varepsilon \chi_p^h$$

and so, comparing with **??** 3.2.2, we set $\varepsilon(\rho)$ to be the ε obtained here. We also see that h had better be the same as $k(\rho) - 1$ modulo $p - 1$.

4.6 The weight

We now come to the final ingredient in Serre's recipe, the weight $k(\rho)$. The general strategy of our approach here is to express a representation of Serre-type as a twist of another representation, one that looks like it comes from a cusp form of weight at most $p + 1$. We then read off the minimal weight of a cusp form that could give this twisted representation. Then we can apply the results above regarding the Θ operator (**??** 2.1.8 and **??** 4.3.5) to define the weight of the original representation.

Given our Galois representation

$$\rho \colon \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathrm{GL}_2(\overline{\mathbf{F}}_p)$$

we restrict to the subgroup G_p to form

$$\rho_p \colon \mathrm{G}_p \to \mathrm{GL}_2(\overline{\mathbf{F}}_p).$$

The definition of $k(\rho)$ will only depend on this ρ_p (in fact it will only depend on $\rho_p|_{I_p}$). We will from here on refer to $k(\rho)$ as $k(\rho_p)$ to emphasise this fact. As such the weight will only reflect the behaviour at p of the representation, whereas the level reflected the behaviour away from p.

Proposition 4.6.1. *(Serre [Ser72, prop. 4]) The semisimplification ρ_p^{ss} of ρ_p is trivial when restricted to $I_{p,w}$.*

Proof. It suffices to prove this for simple representations ρ_p, as a sum of trivial representations is trivial.

The wild inertia $I_{p,w}$ is a pro-p-group, and so the image is also a pro-p-group. This group is finite, so it is simply a p-group, and defined over some finite field $\mathbf{F}_q$. Consider an $\mathbf{F}_q$-vector space V that realises $\rho_p^{\mathrm{ss}}|_{I_{p,w}}$. Additively V is a p-group too, of order q^n, where n is the dimension of V. Looking at the action of $\rho_p^{\mathrm{ss}}|_{I_{p,w}}$ on V we see that there is a singleton orbit $\{0\}$. As all orbits are of p-power order there must be an additional $p - 1$ singleton orbits at least, else the orbits could not partition V. Therefore the fixed subspace W of V is non-trivial. However as $I_{p,w}$ is normal in G_p the subspace W is stable under G_p, hence W must equal V by simplicity. □

We may therefore view ρ_p^{ss} as a representation of $I_{p,t}$, we shall write ρ_t^{ss} for this new representation. The tame inertia group is abelian, and so this representation is diagonalisable. The representation ρ_t^{ss} is therefore given by a pair of characters

$$\phi_1,\, \phi_2 \colon I_{p,t} \to \overline{\mathbf{F}}_p^*.$$

Proposition 4.6.2. *Both of the characters ϕ_1 and ϕ_2 are of the same level, and that level is either 1 or 2.*

Moreover if they are both of level 2 then they are pth powers of each other.

Proof. Letting a Frobenius element at p act by conjugation on $\sigma \in I_p$ we have that [RS99, p. 167]

$$\mathrm{Frob}_p\, \sigma\, \mathrm{Frob}_p^{-1} \equiv \sigma^p \pmod{I_{p,w}},$$

and so

$$\rho_t^{\mathrm{ss}}(\mathrm{Frob}_p\, \sigma\, \mathrm{Frob}_p^{-1}) = \rho_t^{\mathrm{ss}}(\sigma^p) = \rho_t^{\mathrm{ss}}(\sigma)^p,$$

so we have an equivalence of representations

$$\rho_t^{\mathrm{ss}} \cong (\rho_t^{\mathrm{ss}})^p.$$

Hence the set $\{\phi_1, \phi_2\}$ must be fixed by pth powering.

We then have two possibilities, either taking the pth power fixes each of ϕ_1 and ϕ_2 or it swaps them. If they are both fixed then their images lie in the prime field, so they are of level 1. Otherwise, if they swap under pth powering, each of them is fixed under powering by p^2, and hence they are of level 2. □

We now treat three different cases separately, based on the levels of the characters just obtained and whether or not $\rho|_{I_{p,w}}$ is trivial.

4.6.1 The level 2 case

If the two characters ϕ_1 and ϕ_2 are of level 2 then ρ_p is irreducible.

To see this, assume otherwise and consider a stable 1-dimensional subspace of the vector space realising ρ_p. The representation ρ_p acts by a character on this subspace, which is tamely ramified by the argument in **??** 4.6.1. So this gives a character ϕ of $I_{p,t}$ which extends to all of G_p, so we have

$$\phi(\sigma) = \phi(\mathrm{Frob}_p\, \sigma\, \mathrm{Frob}_p^{-1}) = \phi^p(\sigma).$$

Hence ϕ must factor through $\mathbf{F}_p^*$, i.e. ϕ is of level 1. This is a contradiction as ϕ must be one of ϕ_1 or ϕ_2.

So ρ_p is irreducible and hence $\rho_p = \rho_p^{\mathrm{ss}}$, therefore the characters ϕ_1 and ϕ_2 above define the representation $\rho_p|_{I_p}$. We can write them in terms the fundamental characters of level 2, ψ_1 and ψ_2 (as defined in **??** 2.2.11) and use this description to define $k(\rho_p)$. Specifically we can write ϕ_1 as

$$\phi_1 = \psi_1^a \psi_2^b$$

with $0 \leq a, b \leq p-1$. If $a = b$ then $\phi_1 = (\psi_1\psi_2)^a = \chi_p^a$ (recall **??** 2.2.12), which contradicts ϕ_1 being of level 2 as the mod p cyclotomic character is of level 1. Now we observe that

$$\phi_2 = \phi_1^p = (\psi_1^a \psi_2^b)^p = \psi_2^a \psi_1^b,$$

so by switching the places of ϕ_1 and ϕ_2 if necessary we may assume that in fact $0 \leq a < b \leq p-1$.

Now when restricting to inertia we have

$$\rho_p|_{I_p} \sim \begin{pmatrix} \psi_1^b \psi_2^a & 0 \\ 0 & \psi_1^a \psi_2^b \end{pmatrix}.$$

This looks a bit like the supersingular case of **??** 4.3.2. So we massage our representation into the form seen in the theorem by factoring out a character to get

$$\rho_p|_{I_p} \sim \begin{pmatrix} \psi_1^b \psi_2^a & 0 \\ 0 & \psi_1^a \psi_2^b \end{pmatrix} = \psi_2^a \psi_1^a \begin{pmatrix} \psi_1^{b-a} & 0 \\ 0 & \psi_2^{b-a} \end{pmatrix} = \chi_p^a \begin{pmatrix} \psi_1^{b-a} & 0 \\ 0 & \psi_2^{b-a} \end{pmatrix}$$

If we were just considering the rightmost matrix we would like to set $k(\rho_p) - 1 = b - a$ to be consistent with the theorem, however we have twisted by χ_p^a. Recalling **??** 2.1.8 and **??** 4.3.5 we make the definition

$$k(\rho_p) - 1 = b - a + a(p+1),$$

or equivalently

$$k(\rho_p) = 1 + pa + b. \tag{7}$$

As we have $0 \leq a < b \leq p-1$ we see that

$$2 \leq k(\rho_p) \leq 1 + p(p-2) + p - 1 = p^2 - p.$$

4.6.2 The level 1 tame case

Assuming ϕ_1 and ϕ_2 are of level 1 and that $\rho_p|_{I_{p,w}}$ is trivial we can write

$$\rho_p|_{I_p} \sim \begin{pmatrix} \phi_1 & 0 \\ 0 & \phi_2 \end{pmatrix} = \begin{pmatrix} \chi_p^b & 0 \\ 0 & \chi_p^a \end{pmatrix}.$$

For some integers a and b defined modulo $p-1$, we can assume that $0 \leq a \leq b \leq p-2$ by switching ϕ_1 and ϕ_2 if necessary. This looks similar to the ordinary case we covered in **??** 4.3.3. So we factor out a character again to get something that looks exactly like that theorem,

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

If we just had the right hand matrix we would want to set $k(\rho_p) - 1 = b - a$, but once again we have a twist. Taking this into account we try to set

$$k(\rho_p) - 1 = b - a + a(p+1),$$

or equivalently

$$k(\rho_p) = 1 + pa + b$$

as above, but there is a small issue this time. It is possible here that $a = b = 0$, in which case this definition would give us $k(\rho_p) = 1$. However we do not want to consider weight 1 modular forms at all (indeed we only attached Galois representations to eigenforms of weight at least 2), so our formula needs modifying in this case. Looking at Section 4.5 and **??** 2.1.5 we see that it is only permissible to change the weight by multiples of $p-1$. So to remedy the situation we add $p-1$ when we are in the problem case. The definition in this case is then

$$k(\rho_p) = \begin{cases} 1 + pa + b & \text{if } (a,\, b) \neq (0,\, 0), \\ p & \text{if } (a,\, b) = (0,\, 0). \end{cases} \tag{8}$$

With this definition we have

$$2 \leq k(\rho_p) \leq 1 + p(p-2) + p - 2 = p^2 - p - 1,$$

unless $p = 2$, where the above inequality makes no sense, in which case $k(\rho_p) = 2$ is the only possibility.

4.6.3 The level 1 non-tame case

The final case is where ϕ_1 and ϕ_2 are of level 1 but the action of $I_{p,w}$ on V is non-trivial.

If we consider the subspace of V fixed by $I_{p,w}$ the same argument we used in **??** 4.6.1 shows that this subspace is always non-trivial. However $I_{p,w}$ is assumed to act non-trivially and so $V^{I_{p,w}}$ must in fact be a 1-dimensional subspace. This subspace is stable under the action of G_p, as is the space $V/V^{I_{p,w}}$, so we may write

$$\rho_p \sim \begin{pmatrix} \theta_2 & * \\ 0 & \theta_1 \end{pmatrix},$$

where θ_1 and θ_2 are characters of G_p.

Using this decomposition we see that upon restricting to I_p we have

$$\rho_p|_{I_p} \sim \begin{pmatrix} \chi_p^{\beta} & * \\ 0 & \chi_p^{\alpha} \end{pmatrix},$$

for some α, $\beta \in \mathbf{Z}/(p-1)\mathbf{Z}$ as the restrictions of θ_1 and θ_2 to I_p must be of level 1.

We fix representatives α and β such that

$$0 \leq \alpha \leq p-2,$$
$$1 \leq \beta \leq p-1.$$

We can then proceed in the same way as we did before, factoring out a twist by χ_p^a to get

$$\rho_p|_{I_p} \sim \chi_p^{\alpha} \begin{pmatrix} \chi_p^{\beta-\alpha} & * \\ 0 & 1 \end{pmatrix}.$$

As before this looks like **??** 4.3.3, and at first glance it seems as if the representation given by the right hand matrix comes from an eigenform of weight $\beta - \alpha + 1$. However if $\beta - \alpha = 1$ this representation could have come from a form of weight $\beta - \alpha + p + 1$ instead, because $\chi_p = \chi_p^p$ and we cannot determine what the correct power of χ_p is.

This is a real problem as it is incorrect to simply use the smallest weight here (or the largest for that matter). For example, let ρ be the mod 11 representation arising from the mod 11 reduction of the eigenform Δ of weight 12. This representation is unramified outside of 11 and so $N(\rho) = 1$.

Then, as the 11th Fourier coefficient of Δ is $534612 \equiv 1 \pmod{11}$, **??** 4.3.3 tells us that

$$\rho|_{I_{11}} \sim \begin{pmatrix} \chi_{11}^{11} & * \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} \chi_{11} & * \\ 0 & 1 \end{pmatrix}.$$

So if we use $\beta - \alpha + 1$ for our definition here this would predict the existence of a mod 11 eigenform of weight 2 and level 1 from which ρ arises, but there are no such forms.

If $\beta \neq \alpha + 1$ we do not have this problem, so Serre lets

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

and defines, as we did in Section 4.6.1,

$$k(\rho_p) = 1 + pa + b. \tag{9}$$

When $\beta = \alpha + 1$ we have to deal with the ambiguity by finding some way of distinguishing representations that come from eigenforms of weight 2 and $p + 1$. In order to decide case we are in we can make use of **??** 4.3.4. This theorem stated that if a Galois representation arises from a filtration $p + 1$ form, then the representation at p is très ramifé.

So if ρ_p is peu ramifé this cannot be the case, and the twist of ρ_p looks like it came from an eigenform of weight 2 rather than $p + 1$. We then define $k(\rho_p)$ as we did earlier via

$$k(\rho_p) = 1 + pa + b = 2 + \alpha(p + 1). \tag{10}$$

If ρ_p is très ramifé then it looks as if the twist comes from a form of weight $p + 1$. So we make an analogous definition to what we have done before, simply accounting for the twist starting from a weight $p + 1$ form. We have one final adjustment to make, if $p = 2$ this definition would give $k(\rho_p) = 3$, in Serre's definition this is set to be 4 instead. In the end we obtain the following definition

$$k(\rho_p) = \begin{cases} 1 + pa + b + p - 1 = p + 1 + \alpha(p + 1) & \text{if } p \neq 2, \\ 4 & \text{if } p = 2. \end{cases} \tag{11}$$

Looking at the bounds for $k(\rho_p)$ now we see that if $\beta \neq \alpha + 1$ or if ρ_p is peu ramifé then for all p

$$2 \leq k(\rho_p) \leq 1 + p(p - 2) + p - 1 = p^2 - p.$$

Otherwise for the très ramifé case we get

$$2 \leq k(\rho_p) \leq p^2 - p + p - 1 = p^2 - 1,$$

unless $p = 2$ where $k(\rho_p) = 4$ is the only possibility.

Considering all cases together we see that $k(\rho_p)$ has range of

$$2 \leq k(\rho_p) \leq p^2 - 1$$

for odd p, and $k(\rho_p) \in \{2, 4\}$ for $p = 2$.

To see why this approach might be expected to produce the minimal possible weight when twisting is involved depends on analysis of the sequences of filtrations

$$w(\Theta^i f) \text{ for } 0 \leq i \leq p + 1,$$

for mod p eigenforms f. These sequences are known as Θ-cycles and they are studied in detail in [Joc82].

4.7 A counterexample

In fact the conjecture exactly as stated above is in fact *incorrect*, this was noted by Serre in a letter to Ken Ribet in 1987. The following counterexample is due to Serre and is given in [Rib95, sec. 2] and also in [RS11, sec. 21.6.1] which we are following here (see also the notes for Serre's paper introducing the conjecture in his collected works).

Example 4.7.1. Let α be a root of $x^2 + 3x + 3$, so that $\mathbf{Q}(\alpha) = \mathbf{Q}(\sqrt{-3})$. The space $S_2(13; \overline{\mathbf{Z}})$ is spanned by the normalised eigenform

$$f = q + (-\alpha - 3)q^2 + (2\alpha + 2)q^3 + (\alpha + 2)q^4 + (-2\alpha - 3)q^5 + O(q^6)$$

and its $\mathrm{Gal}(\mathbf{Q}(\alpha)/\mathbf{Q})$ conjugate form

$$g = q + \alpha q^2 + (-2\alpha - 4)q^3 + (-\alpha - 1)q^4 + (2\alpha + 3)q^5 + O(q^6),$$

which is the other normalised eigenform in $S_2(13; \overline{\mathbf{Z}})$.

The mod 3 Galois representation attached to f has determinant $\chi_3\phi$, where ϕ is the non-trivial Galois character coming from the extension $\mathbf{Q}(\sqrt{13})/\mathbf{Q}$. We can view ϕ as a Dirichlet character $(\mathbf{Z}/13\mathbf{Z})^* \to \overline{\mathbf{F}}_3$ using the arguments of Section 4.4. Serre's conjecture tells us that this character is our $\varepsilon(\rho)$ and so ρ should arise from some eigenform f in $S_2(13, \phi; \overline{\mathbf{F}}_3)$.

We claim that no such non-zero f can exist. First, let H be the group of squares in $(\mathbf{Z}/13\mathbf{Z})^*$, i.e. the kernel of ϕ. Doing this we may view f as a mod 3 cusp form of weight 2 for the group

$$\Gamma_H(13) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbf{Z}) : c \equiv 0 \pmod{13},\, d \in H \right\}.$$

Although we did not define mod p modular forms for more general congruence subgroups the definition is the same. We can see that the order of the character of such a form must divide 2. However the space of weight 2 cusp forms on $\Gamma_H(13)$ is a subspace of $S_2(13; \overline{\mathbf{F}}_3)$. This larger space is spanned by $\overline{f}$ and $\overline{g}$, both of which have a character of order 6. So the space of mod 3 weight 2 cusp forms on $\Gamma_H(13)$ is zero, and f cannot exist.

This problem is fairly isolated and only arises when we work with mod 2 Galois representations, or mod 3 Galois representations that have abelian restriction to $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}(\sqrt{-3}))$. In order to fix this issue whilst sticking with our definition of modular forms it is necessary to change the definition of the character in these cases. This problem is not a serious issue with the validity of the conjecture in general however, if one simply ignores the character completely or is a bit careful in the situations mentioned this can be completely resolved.

4.8 The proof

As mentioned at the start, this conjecture is in fact now a theorem, due to Khare and Wintenberger using results of Kisin and others.

For $p \neq 2$ it was known that the qualitative and refined forms were equivalent before either was known in general. This reduction of the refined form to the qualitative form was incremental and is due to a large number of people, for example Ribet [Rib94] lists N. Boston, H. Carayol, F. Diamond, B. Edixhoven, G. Faltings, B. H. Gross, B. Jordan, K. Ribet, H. W. Lenstra, Jr., R. Livné, B. Mazur and J-P. Serre. The reduction for $p \neq 2$ was completed by Diamond around 1993 [Dia95]. The missing case of $p = 2$ was completed by Khare and Wintenberger as part of their proof of the conjecture.

Many special cases of Serre's conjecture were also known long before the general case. Indeed in 1973, 14 years before the Duke paper, John Tate proved the conjecture when $p = 2$ for Serre-type representations unramified outside 2 [Tat94]. Serre himself used similar techniques to

prove the same thing for $p = 3$ (published as a note on page 710 of volume III of his collected works). As we will see in Section 6.2, for these cases what needs to be shown in these proofs is that no such Galois representations exist, to match the lack of cusp forms of level 1 for weights less than 12. It is interesting to note that these proofs were not completely subsumed or rendered obsolete by the proof of Khare and Wintenberger. Rather they help form the base case for an induction type argument which gives the full conjecture.

5 Examples

One of the great things about Serre's conjecture, even if it were not yet known to be correct, is the fact that it can be used in concrete cases easily. Specifically, given a Galois representation of Serre-type we can calculate the weight and level along with the character as detailed above, then in many instances we can compute the associated space of eigenforms and look for a form from which our Galois representation arises.

5.1 A Galois representation arising from Δ

Let's return first to **??** 3.1.1, which concerned the cusp form Δ, and check that everything we have just done is consistent with what we saw there.

Example 5.1.1. We consider the 23-adic Galois representation ρ_{23}, as out of the representations we considered there it is the only irreducible one and so the only case in which Serre's conjecture may have something to say. Recall that this was defined by taking K to be the splitting field of $x^3 - x - 1$, this is ramified only at 23 and has Galois group S_3. We then took r be the unique irreducible degree 2 representation of S_3 taken with coefficients in $\mathbf{Q}_{23}$, this representation satisfies

$$\mathrm{tr}(r(\sigma)) = \begin{cases} 0 & \text{if } |\sigma| = 2, \\ 2 & \text{if } |\sigma| = 1, \\ -1 & \text{if } |\sigma| = 3, \end{cases}$$

for each $\sigma \in S_3$. The Galois representation ρ_{23} was then the composition

$$\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathrm{Gal}(K/\mathbf{Q}) \xrightarrow{\sim} S_3 \to \mathrm{GL}_2(\mathbf{Q}_{23}).$$

In order to use Serre's conjecture we need a mod p representation so we reduce mod 23 to obtain $\rho = \overline{\rho}_{23}$.

Straight away we see that $N(\rho) = 1$ due to the fact K is unramified outside 23 (see **??** 4.4.1). The determinant is the sign character ϕ of S_3, taken with values mod 23 and viewed as a character of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$. As ϕ is of order exactly 2 it must be the 11th power of χ_{23}.

The inertia group is of order 2, and the wild inertia is trivial. Additionally, as our whole representation lands in the prime field we are in the level 1 tame case. The two diagonal characters must be either 1 or $\phi|_{I_{23}}$, we know their product is $\phi|_{I_{23}}$ so we have

$$\rho_{23}|_{I_{23}} \sim \begin{pmatrix} \chi_{23}^{11} & 0 \\ 0 & 1 \end{pmatrix}$$

hence our weight is $k(\rho_{23}) = 11 + 1 = 12$.

So ρ should have arisen from a normalised eigenform in

$$S_{12}(1, \mathrm{Id}; \overline{\mathbf{F}}_p) = \overline{\mathbf{F}}_p \cdot \overline{\Delta},$$

as we would expect.

5.2 A Galois representation arising from a D_4 extension

Now we move to a new example, once again arising from the Galois group of a number field.

Example 5.2.1. Take the K to be the splitting field of

$$f = x^4 - 3,$$

so $K = \mathbf{Q}(\sqrt[4]{3}, i)$. This extension has Galois group D_4, generated by some σ and τ, where

$$\sigma(\sqrt[4]{3}) = \sqrt[4]{3},\ \sigma(i) = -i,\ \tau(\sqrt[4]{3}) = i\sqrt[4]{3},\ \tau(i) = i.$$

We can turn this into a mod 5 Galois representation using the fact that D_4 embeds into $\mathrm{GL}_2(\mathbf{F}_5)$ via its natural 2-dimensional faithful representation

$$\sigma \mapsto \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},\ \tau \mapsto \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

So we obtain an irreducible a mod 5 Galois representation

$$\rho\colon\ \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathrm{Gal}(K/\mathbf{Q}) \xrightarrow{\sim} D_4 \to \mathrm{GL}_2(\mathbf{F}_5).$$

What does Serre have to say about ρ? After fixing an embedding $K \hookrightarrow \mathbf{C}$, complex conjugation corresponds to a reflection which maps to a matrix of determinant -1, so we see that this representation is indeed odd, and Serre's conjecture will apply.

The determinant of ρ takes τ to 1 and σ to -1, so it factors through $\mathrm{Gal}(\mathbf{Q}(i)/\mathbf{Q})$ and corresponds to the non-trivial Dirichlet character of conductor 4.

The conductor $N(\rho)$ will only be divisible by 2 and 3, as these are the primes dividing disc K and so ρ is unramified outside of these two primes. We first note that any order 4 subgroup of D_4 contains τ^2. The image of such a subgroup will therefore contain $-I_2$ and hence will have trivial fixed subspace.

For $\ell = 2$ we find the following ramification groups

$$\begin{aligned} G_{2,0} &= G_{2,1} = V_4, \\ G_{2,2} &= G_{2,3} = \langle \tau^2 \rangle, \\ G_{2,4} &= 1. \end{aligned}$$

and so

$$\begin{aligned} \dim V/V^{2,0} &= \dim V/V^{2,1} = 2, \\ \dim V/V^{2,2} &= \dim V/V^{2,3} = 2, \\ \dim V/V^{2,4} &= 0. \end{aligned}$$

Giving $\nu_2(\rho) = 2 \cdot 2 + 2 \cdot \frac{1}{2} \cdot 2 = 6$.

For $\ell = 3$ we find the following ramification groups

$$\begin{aligned} G_{3,0} &= C_4, \\ G_{3,1} &= 1. \end{aligned}$$

and so

$$\begin{aligned} \dim V/V^{3,0} &= 2, \\ \dim V/V^{3,1} &= 0. \end{aligned}$$

Giving $\nu_3(\rho) = 2$. Together we get that $N(\rho) = 2^6 3^2 = 576$.

As K is unramified at 5 the restriction to inertia is trivial. So we are in the level 1 tame case, with trivial characters, and hence the weight $k(\rho_p)$ is 5. Finally we let ε be the extension of the non-trivial Dirichlet character of conductor 4 to a Dirichlet character for $N(\rho)$.

So ρ should come from some eigenform $f \in S_5(576, \varepsilon; \overline{\mathbf{F}}_5)$. We can explicitly compute this space using, for example, Sage [S$^+$15]. Doing this gives us a number of eigenforms, including

$$q + 2q^5 + 2q^{13} + 4q^{25} + 3q^{37} + q^{49} + 3q^{61} + 4q^{65} + 3q^{73} + 3q^{97} + O(q^{109}),$$

which we verify has the expected q-expansion coefficients a_p for prime p up to q^{1000} (i.e. 2 if $\mathrm{Frob}_p = 1$, -2 if $\mathrm{Frob}_p = \tau^2$ and 0 otherwise).

6 Consequences

Serre's conjecture is a strong statement that implies many other results within number theory. We now mention briefly a few of these. The first two are little more than examples but interesting ones nonetheless, whereas the third is a very deep statement within arithmetic geometry.

6.1 Finiteness of classes of Galois representations

First let us examine a very direct consequence. Fix a prime p and an integer N and consider Serre-type Galois representations

$$\rho\colon \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathrm{GL}_2(\overline{\mathbf{F}}_p)$$

whose associated levels $N(\rho)$ divide N. Serre's conjecture states that each corresponds to some normalised mod p eigenform of level $N(\rho)|N$ and weight $k(\rho)$ in the range $[2, p^2 - 1]$ (or $\{2, 4\}$ for $p = 2$). However there are only finitely many spaces of forms satisfying these requirements and only finitely many normalised eigenforms in each. Therefore for each prime p and integer N there are only finitely many isomorphism classes of mod p Serre-type Galois representations of conductor dividing N. Apparently there are no alternative methods of proving this result currently known [Wie13].

6.2 Unramified mod p Galois representations for small p

We can specialise the previous type of direct argument further to get more control over the number of representations with particular properties. In fact we can get enough control to prove the following non-existence result.

Let ρ be a Serre-type mod p Galois representation for some $p \leq 7$ that is unramified outside of p. In this case, due to the absence of ramification, $N(\rho)$ is simply 1 (recall **??** 4.4.1). The idea of our definition of the weight was that each Galois representation ρ should be the twist by a power of the cyclotomic character of another form ρ', such that $2 \leq k(\rho') \leq p+1$.

So Serre's conjecture predicts there is some mod p cusp form of level 1 and weight ≤ 8 from which some twist of ρ arises. But there are no cusp forms of level 1 of weight < 12 and so such a twisted representation cannot exist, hence the original ρ cannot exist either. As mentioned in Section 4.8 this sort of statement was shown via other methods prior to the proof of the full conjecture and provided important theoretical evidence in support of Serre's conjecture.

6.3 Modularity of abelian varieties

In this section we look briefly at another strong result that follows from Serre's conjecture, despite it not obviously concerning the objects related in the conjecture. This was a genuinely new result that was not known before the proof of Serre's conjecture and so serves as a good example of the usefulness of the conjecture outside of its immediate domain. Going into detail would take us too far afield so this section is necessarily sketch-like and without background material. For more details see [Rib04] or [RS11, chap. 15].

Definition 6.3.1. An abelian variety A over $\mathbf{Q}$ is *modular* if there exists an N such that there is a surjective map defined over $\mathbf{Q}$

$$J_1(N) \twoheadrightarrow A.$$

Definition 6.3.2. An abelian variety A defined over $\mathbf{Q}$ is said to be of GL_2-*type* if it is simple and its endomorphism algebra

$$\mathbf{Q} \otimes \mathrm{End}_{\mathbf{Q}}(A)$$

contains a number field E whose degree is equal to the dimension of A.

Example 6.3.3. Elliptic curves are of GL_2-type because all endomorphism rings over characteristic 0 fields of elliptic curves contain $\mathbf{Z}$ and hence

$$\mathbf{Q} \subset \mathbf{Q} \otimes \mathrm{End}_{\mathbf{Q}}(E).$$

Ken Ribet has shown [Rib04] that Serre's conjecture implies the following nice classification of modular abelian varieties.

Theorem 6.3.4. *Every abelian variety of* GL_2*-type is modular.*

This theorem is a generalisation of the Taniyama–Shimura–Weil conjecture, or modularity theorem, first proved by Breuil, Conrad, Diamond and Taylor in 2001 [BCDT01]. However this modularity statement is significantly stronger and uses the full power of Serre's conjecture. Indeed this theorem was not known prior to the proof of the conjecture.

7 References

[BCDT01] Christophe Breuil, Brian Conrad, Fred Diamond, and Richard Taylor. On the modularity of elliptic curves over **Q**: wild 3-adic exercises. *Journal of the American Mathematical Society*, pages 843–939, 2001.

[Bir67] Brian J. Birch. Cyclotomic fields and Kummer extensions. In *Algebraic Number Theory (Proc. Instructional Conf., Brighton, 1965), Thompson, Washington, DC*, pages 85–93, 1967.

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

[Car86] Henri Carayol. Sur les représentations l-adiques associées aux formes modulaires de Hilbert. *Annales Scientifiques de l'Ecole Normale Supérieure*, 19(3):409–468, 1986.

[Coh08] Henri Cohen. *Number Theory: Volume I: Tools and Diophantine Equations*. Graduate Texts in Mathematics. Springer New York, 2008.

[Dar95] Henri Darmon. Serre's conjectures. In Vijaya Kumar Murty, editor, *Seminar on Fermat's Last Theorem: 1993-1994, the Fields Institute for Research in the Mathematical Sciences, Toronto, Ontario, Canada*, volume 17 of *CMS conference proceedings*, pages 135–153. American Mathematical Society, 1995.

[DDT95] Henri Darmon, Fred Diamond, and Richard Taylor. Fermat's last theorem. In J. Coates and S. Yau, editors, *Elliptic curves, modular forms, and Fermat's last theorem*, pages 1–154, 1995.

[Del69] Pierre Deligne. Formes modulaires et représentations ℓ-adiques. *Séminaire Bourbaki*, 11:139–172, 1968-1969.

[Del04] Pierre Deligne. Formes modulaires et représentations ℓ-adiques (translation). `http://publications.ias.edu/sites/default/files/5_FormesModulaires.pdf`, June 2004.

[Dia95] Fred Diamond. The refined conjecture of Serre. In J. Coates and S. Yau, editors, *Elliptic curves, modular forms, and Fermat's last theorem*, pages 172–186, 1995.

[DS74] Pierre Deligne and Jean-Pierre Serre. Formes modulaires de poids 1. *Annales scientifiques de l'cole Normale Suprieure*, 7(4):507–530, 1974.

[Edi92] Bas Edixhoven. The weight in Serre's conjectures on modular forms. *Inventiones mathematicae*, 109(1):563–594, 1992.

[Edi97] Bas Edixhoven. Serre's conjecture. In Gary Cornell, Joseph H. Silverman, and Glenn Stevens, editors, *Modular Forms and Fermat's Last Theorem*, pages 209–242. Springer New York, 1997.

[Frö67] A Fröhlich. Local fields. In *Algebraic Number Theory (Proc. Instructional Conf., Brighton, 1965), Thompson, Washington, DC*, pages 1–41, 1967.

[Ghi] Alex Ghitza. On the two-dimensional modular representations of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ (translation). `http://www.ms.unimelb.edu.au/~aghitza/nuts/serre.pdf`.

[Gro90] Benedict H. Gross. A tameness criterion for Galois representations associated to modular forms (mod p). *Duke Mathematical Journal*, 61(2):445–517, 1990.

[GS95] Marvin J. Greenberg and Jean-Pierre Serre. *Local Fields*. Graduate Texts in Mathematics. Springer New York, 1995.

[Joc82] Naomi Jochnowitz. A study of the local components of the Hecke algebra mod l. *Transactions of the American Mathematical Society*, 270(1):253–267, 1982.

[Kis09] Mark Kisin. Modularity of 2-adic Barsotti-Tate representations. *Inventiones mathematicae*, 178(3):587–634, 2009.

[KW09a] Chandrashekhar Khare and Jean-Pierre Wintenberger. Serre's modularity conjecture (I). *Inventiones mathematicae*, 178(3):485–504, 2009.

[KW09b] Chandrashekhar Khare and Jean-Pierre Wintenberger. Serre's modularity conjecture (II). *Inventiones mathematicae*, 178(3):505–586, 2009.

[Liv89] Ron Livné. On the conductors of mod l Galois representations coming from modular forms. *Journal of Number Theory*, 31(2):133–141, 1989.

[Mas13] Nicolas Mascot. Computing modular Galois representations. *Rendiconti del Circolo Matematico di Palermo*, 62(3):451–476, 2013.

[Rib90] Kenneth A. Ribet. On modular representations of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ arising from modular forms. *Inventiones mathematicae*, 100(1):431–476, 1990.

[Rib94] Kenneth A. Ribet. Report on mod l representations of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$. In *Proceedings of Symposia in Pure Mathematics*, volume 55, pages 639–676, 1994.

[Rib95] Kenneth A. Ribet. Irreducible Galois representations arising from component groups of Jacobians. In J. Coates and S. Yau, editors, *Elliptic curves, modular forms, and Fermat's last theorem*, pages 131–147, 1995.

[Rib04] Kenneth A. Ribet. Abelian Varieties over $\mathbf{Q}$ and Modular Forms. In John E. Cremona, Joan-Carles Lario, Jordi Quer, and Kenneth A. Ribet, editors, *Modular Curves and Abelian Varieties*, volume 224 of *Progress in Mathematics*, pages 241–261. Birkhuser Basel, 2004.

[RS99] Kenneth A. Ribet and William A. Stein. Lectures on Serre's conjectures. In B.D. Conrad and K. Rubin, editors, *Arithmetic Algebraic Geometry*, IAS Park City mathematics series: Institute for Advanced Study, pages 143–232. American Mathematical Soc., 1999.

[RS11] Kenneth A. Ribet and William A. Stein. Lectures on modular forms and Hecke operators. `http://wstein.org/books/ribet-stein/main.pdf`, November 2011.

[S+15] W. A. Stein et al. *Sage Mathematics Software (Version 6.5)*. The Sage Development Team, 2015. http://www.sagemath.org.

[SD73] Peter Swinnerton-Dyer. On ℓ-adic representations and congruences for coefficients of modular forms. In *Modular functions of one variable III*, pages 3–55. Springer, 1973.

[Ser67] Jean-Pierre Serre. Une interprétation des congruences relatives à la fonction τ de Ramanujan. *Séminaire Delange-Pisot-Poitou. Théorie des nombres*, 9(1):1–17, 1967.

[Ser72] Jean-Pierre Serre. Propriétés galoisiennes des points d'ordre fini des courbes elliptiques. *Inventiones mathematicae*, 15(4):259–331, 1972.

[Ser73a] Jean-Pierre Serre. Congruences et formes modulaires (d'après H.P.F. Swinnerton-Dyer). In *Séminaire Bourbaki vol. 1971/72 Exposés 400–417*, pages 319–338. Springer, 1973.

[Ser73b] Jean-Pierre Serre. Formes modulaires et fonctions zêta p-adiques. In *Modular functions of one variable III*, pages 191–268. Springer, 1973.

[Ser75] Jean-Pierre Serre. Valeurs propres des opérateurs de Hecke modulo l. *Astérisque*, 24:109–117, 1975.

[Ser87] Jean-Pierre Serre. Sur les représentations modulaires de degré 2 de $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$. *Duke Mathematical Journal*, 54(1):179–230, 1987.

[Tat94] John Tate. The non-existence of certain Galois extensions of $\mathbf{Q}$ unramified outside 2. *Contemporary Mathematics*, 174:153–153, 1994.

[Wie12] Gabor Wiese. Galois Representations. http://math.uni.lu/~wiese/notes/GalRep.pdf, February 2012.

[Wie13] Gabor Wiese. Modular Galois Representations and Applications. http://math.uni.lu/~wiese/notes/2013-ModGalRepApp.pdf, April 2013.