Serre's conjecture

Alex J. Best

April 6, 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. 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 standard form of the conjecture before going on to compute some explicit examples and examining some interesting consequences.

Beyond the original paper there are many very good accounts of Serre's statement, including Cais [Cai09], Edixhoven [Edi97] (both of which use Katz's definition of mod p modular forms), Ribet–Stein [RS99] and Darmon [Dar95] (which stay closer to the original article). Alex Ghitza has prepared a translation of part of Serre's paper [Ghi] which has been helpful.

2 Background

Here we fix several definitions and key results that will be relevant to our study of Serre's conjecture.

2.1 Modular forms

Definition 2.1. Let

$$S_k(N, \epsilon, R)$$

be the space of cusp forms of weight k for $\Gamma_1(N)$ with character ϵ and q-expansion coefficients in the ring R.

2.2 Galois representations

Definition 2.2. A *n-dimensional mod p Galois representation* is a homomorphism

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

Here 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*.

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. Our representations are always continuous (where we give $\mathrm{GL}_n(\overline{\mathbf{F}}_p)$ the discrete topology), though we will often still state as such to remind ourselves. The continuity condition for mod p representations then reduces to having an open kernel and thus our representations have finite image.

Although our main object of study is $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ it will be very useful for us to also consider representations of

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

Indeed such representations can be obtained from those of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ via restriction.

The group G_p has several important subquotients which will be helpful to study restrictions of representations to.

Definition 2.3. The first is the *inertia subgroup* at I_p which is defined as the kernel of the map

$$\mathrm{G}_p \twoheadrightarrow \mathrm{Gal}(\overline{\mathbf{F}}_p/\mathbf{F}_p)$$

obtained by quotienting out $\overline{\mathbf{Q}}_p$ by its maximal ideal.

Next the *wild* inertia group $I_{p,w}$ is the maximal pro-p-subgroup of I_p and the *tame* inertia group is the quotient

$$I_{p,t} = I_p/I_{p,w}.$$

Finally we may define a series of subgroups of inertia, the *higher inertia groups*

$$I_p^u = \{\} \subseteq I_p.$$

These form an ascending chain

$$I_p = I_p^0 \subseteq I_p^1 \subseteq \cdots.$$

Definition 2.4. Given a character $\phi\colon \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to K^*$ for some field K and after fixing an embedding $\overline{\mathbf{Q}} \hookrightarrow \mathbf{C}$. The image of complex conjugation, viewed as an element $c \in \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$, under ϕ must be an element of order 2 in K^*, so must be ± 1. If $\phi(c) = -1$ we say ϕ is *odd*, otherwise we say ϕ is *even* (though we shall be concerned mostly 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 can define the parity of ϕ to be that of the character $\det\phi$.

Definition 2.5. 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 characters of $I_{t,p}$ of level n which will allow us to describe all characters of a particular level.

Definition 2.6. The *fundamental characters* of level n are defined by extending the natural character

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

to an $\overline{\mathbf{F}}_p$-character via the n embeddings of $\mathbf{F}_{p^n}^* \hookrightarrow \overline{\mathbf{F}}_p^*$.

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 shall be the goal of the rest of this essay.

We shall start with a historically important example that provides the first glimpse of the behaviour we will be looking at.

Example 3.1. Let

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

be the Ramanujan Δ function, the unique normalised cusp form of weight 12 for $\Gamma_1(1) = \mathrm{SL}_2(\mathbf{Z})$.

Given examples such as the above it is natural to wonder whether such a relationship holds more generally. Indeed Serre asked how one could associate to each ???? eigenform a Galois representation whose traces of Frobenius elements match the Hecke eigenvalues mod p. More precisely, Serre conjectured the following:

Theorem 3.1. *Given a normalised cuspidal mod p eigenform f of weight k and character ϵ there exists a two-dimensional mod p Galois representation ρ_f such that for all primes ℓ not dividing pN*

(i) $\operatorname{tr}(\rho_f(\mathrm{Frob}_\ell)) = a_\ell$,

(ii) $\det(\rho_f(\mathrm{Frob}_\ell)) = \ell^{k-1}\epsilon(\ell)$.

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

The proof of this theorem is due to Shimura when $k = 2$ [**?**], Deligne when $k > 2$ [Del69] and Deligne–Serre when $k = 1$ [DS74]. In fact the constructions obtained by these authors are of p-adic Galois representations $\rho_f\colon \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathrm{GL}_2(\mathbf{Q}_p)$ and the representation of the theorem is then obtained from the p-adic one by a process of reduction and semisimplification.

If we consider $c \in \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ corresponding to complex conjugation we see that

$$\det \rho_f(c) = ?? = -1.$$

So such a representation ρ_f is necessarily *odd*.

4 Serre's Conjecture

Given the above result it is natural to ask about the converse statement, given a Galois representation satisfying some necessary conditions, does it arise from a newform? Serre's conjecture is that the answer to this question is yes, all Galois representations that could possibly arise from a newform as in Theorem 3.1 do. The conjecture naturally comes into two parts, one weaker existence statement, and another refined form that makes exact predictions about the quantities involved.

Conjecture 4.1 (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 mod p cusp form f such that ρ is isomorphic to ρ_f, the Galois representation associated to f defined in Theorem 3.1.*

This conjecture (at least for $N = 1$) appeared much earlier than the Duke paper and is mentioned by Serre in 1975 [Ser75].

Given the above statement it is natural to 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 as above of that weight, level and character. In a slightly backwards manner we will first state this refined form of the conjecture, before moving on to motivate and define the quantities $N(\rho)$, $k(\rho)$ and character $\epsilon(\rho)$ used in the statement.

Conjecture 4.2 (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 mod* p *cusp form* f *of weight* $k(\rho)$*, level* $N(\rho)$ *and character*

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

whose associated Galois representation ρ_f *is isomorphic to* ρ*.*

From now on we refer to a Galois representation ρ satisfying the hypotheses of this conjecture as being of *Serre type*. We also call a normalised cuspidal eigenform of weight k, level N and character ϵ a form of type (k, N, ϵ).

4.1 Results on Galois representations associated to modular forms

In order to try and understand which types of forms can give rise to a particular representation it is useful to take an arbitrary form and study the properties of the representation attached to it. This has been done by several people and the information which will be important to us is contained in the following theorems.

Fix a prime p and a normalised eigenform $f \in S_k(\Gamma_1(N), \epsilon, \mathbf{F}_p)$ 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 Theorem 3.1. Concerning the conductor of ρ_f we have the following result of Carayol [**?**].

Theorem 4.1. *Let* $N(\rho_f)$ *be the level associated to* ρ_f *(which we will define explicitly shortly), then*

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

Given this it is natural to hope that any Galois representation of Serre type arises from a form of weight exactly $N(\rho)$.

We can also consider the restriction of ρ_f to I_p, for this there are two different theorems depending on whether $a_p \neq 0$ (the *ordinary case*) or otherwise.

Theorem 4.2 (Deligne). *Suppose $k \geq 2$ and $a_p \neq 0$ and let $\lambda(a)\colon \mathrm{G}_p \to \overline{\mathbf{F}}_p^*$ be the unramified character of G_p that takes all $\mathrm{Frob}_p \in \mathrm{G}_p/I$ to a, then we have*

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

up to conjugation in $\mathrm{GL}_2(\overline{\mathbf{F}}_p)$.

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.

Theorem 4.3 (Fontaine). *Suppose $k \geq 2$ and $a_p = 0$ and let ψ_1 and ψ_2 be the two fundamental characters of level 2 then we have*

$$\rho_{f,p}|I = \begin{pmatrix} \psi_1^{k-1} & 0 \\ 0 & \psi_2^{k-1} \end{pmatrix}$$

up to conjugation in $\mathrm{GL}_2(\overline{\mathbf{F}}_p)$.

It is worth noting that proofs of some of these theorems came after Serre's definition of the weight, level and character. However it seems likely that observations of the above results in specific examples informed the recipe below.

4.2 The optimal 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 optimal level in the conjecture.

We can equivalently 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. Now letting $I_{\ell,i} \subset \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ be the ith inertia subgroup at ℓ for each prime ℓ, as defined in Definition 2.3, we can consider the fixed subspace of V

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

We can then define integers $v_\ell(\rho)$ by

$$v_\ell(\rho) = \sum_{i=0}^{\infty} \frac{1}{[I_{\ell,0} : I_{\ell,i}]} \dim \left(V / V^{I_{\ell,i}} \right),$$

and set

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

Note that this is indeed a positive integer, and by construction it is coprime to ℓ.

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

The definition above is that of the *Artin conductor* of a representation, but with the p part ignored.

Given a continuous mod p Galois representation ρ we can compose with the determinant map to obtain a homomorphism

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

As ρ factors through some finite Galois group the image of ρ in $\mathrm{GL}_n(\overline{\mathbf{F}}_p)$ is finite, and hence so is the image of $\det \rho$. So the image of $\det \rho$ is cyclic.

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

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

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

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

4.4 The optimal weight

We now come to the final ingredient in Serre's recipe, that of the weight $k(\rho)$. Given our Galois representation

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

we can form a representation of G_p by composing with the restriction map $G_p \to \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$, to obtain

$$\rho_p \colon G_p \to \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathrm{Aut}(V).$$

The definition of $k(\rho)$ will in fact only depend on this ρ_p and therefore the weight will only reflect the behaviour at p of the representation. We will from here on refer to $k(\rho)$ as $k(\rho_p)$ to emphasise this.

After choosing a particular basis the action of G_p on V is given by The semisimplification of V is then obtained by replacing this action with vector space V with G_p action blah, denoted V^{ss}. In general the semisimplification is obtained by taking the direct sum of the Jordan–Hölder constituents of a representation, though for us the above description suffices.

Now consider, instead of V, this action of G_p on the semisimplification of V. The action of $I_{p,w}$ on V^{ss} is trivial and therefore the quotient $I_{p,t}$ has a well defined action on V^{ss}. This action is diagonalisable and so it is given by a pair of characters

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

Proposition 4.1. *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. We have two possibilities, either taking the pth power fixes both ϕ_1 and ϕ_2 or it swaps them. If they are both fixed then they must be of level 1. Otherwise, if they swap under pth powering, then they 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 $I_{p,w}$ acts trivially on V.

4.4.1 The level 2 case

When the characters are of level 2 we can write them in terms the fundamental characters ψ_1 and ψ_2 of level 2 (as defined in Definition 2.6) 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 $\phi_1 = (\psi_1\psi_2)^a$, which contradicts ϕ_1 being of level 2 as $\psi_1\psi_2$ is the level 1 cyclotomic character. Now we observe that

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

and so by switching the places of ϕ_1 and ϕ_2 we may assume that in fact $0 \leq a < b \leq p-1$. In this case we then set

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

4.4.2 The level 1 tame case

Assuming ϕ_1 and ϕ_2 are of level 1 and the action of $I_{p,w}$ on V is semisimple we can write

$$\rho_p|_{I_{p,w}} = \begin{pmatrix} \phi_1 & 0 \\ 0 & \phi_2 \end{pmatrix} = \begin{pmatrix} \chi^a & 0 \\ 0 & \chi^b \end{pmatrix}.$$

So we obtain integers a and b defined modulo $p-1$, we can then assume that $0 \leq a \leq b \leq p-2$ by switching ϕ_1 and ϕ_2 if necessary. We then set

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

4.4.3 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 not semisimple.

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 specific cases easily. Specifically, given a Galois representation of Serre type we can calculate the optimal weight and level as detailed above, and then in many instances we can compute the associated space of modular forms and look for a form from which our Galois representation arises.

Example 5.1. Let's return first to Example 3.1 and check that everything we have just done is consistent with what we saw there.

6 Consequences

Serre's conjecture is a strong statement that implies many other difficult results within number theory. We now mention briefly a few of these. While

many of these results were obtained via other means long before Serre's conjecture was shown in general they are worth mentioning to demonstrate the power of the conjecture.

The next result we can deduce from Serre's conjecture is not as obviously related to the objects related in the conjecture at all.

7 The proof

It needs mentioning that the conjecture exactly as stated above is in fact *incorrect*.

8 References

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

[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.

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

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

[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.

[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.

[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.

[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.

[Ser75] Jean-Pierre Serre. Valeurs propres des opérateurs de Hecke modulo ℓ. *Astérisque*, 24(25):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.