# SERRE WEIGHT CONJECTURES

Abstract. Serre's modularity conjecture, as originally formulated in 1973, predicted that every odd irreducible two dimensional continuous Galois representation over a finite field $\overline{r} : G_{\mathbb{Q}} \to GL_2(\overline{\mathbb{F}}_p)$ arises from a modular form. A stronger version of this conjecture specifies the level and weight of the modular form. It is now known as the modularity theorem, after being proved by Khare and Winterberger in 2009. Much of the early work concerning Serre's conjecture was focussed on proving that the weak form implies the strong form, and the eventual proof of Serre's conjecture relied on the work that had been done to prove their equivalence. Therefore, if one wishes to generalize Serre's conjecture to the case of $GL_n$ over a number field, it is advisable to begin by trying to describe a generalization of the strong form of the conjecture. In this talk, we will briefly review the original conjecture, relate it to the Langlands programme, and introduce several generalizations done by Toby Gee, Florian Herzig and David Savitt. If time allows, we will discuss the difference between generic and non-generic representations and talk about work in progress trying to extend the picture to non-generic representations.

## 1. Background - Serre's Conjecture

### 1.1. Modular Forms.

Recall first the definition of a modular form.

**Definition 1.1.** Let $1 \leq N \in \mathbb{Z}$. We consider the following subgroups of $SL_2(\mathbb{Z})$.

$$\Gamma(N) = \ker\left(SL_2(\mathbb{Z}) \to SL_2(\mathbb{Z}/N\mathbb{Z})\right)$$

$$\Gamma_1(N) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL_2(\mathbb{Z}) \mid a \equiv d \equiv 1 \mod N, \quad c \equiv 0 \mod N \right\}$$

$$\Gamma_0(N) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL_2(\mathbb{Z}) \mid c \equiv 0 \mod N \right\}$$

with $\Gamma(N) \subseteq \Gamma_1(N) \subseteq \Gamma_0(N)$.

**Definition 1.2.** Let $\mathcal{H} = \{z \in \mathbb{C} \mid \Im(z) > 0\}$. Let $f : \mathcal{H} \to \mathbb{C}$, $k \in \mathbb{Z}$, and $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL_2(\mathbb{R})$. Denote (sometimes denoted $f[\gamma]_k$)

$$(f|_k\gamma)(z) = (cz+d)^{-k} f(\gamma z), \quad \gamma z = \frac{az+b}{cz+d}$$

**Definition 1.3.** Let $k \in \mathbb{Z}$, and let $\Gamma \subseteq SL_2(\mathbb{Z})$ be a subgroup containing $\Gamma(N)$. A *modular form of weight $k$ for $\Gamma$* is a holomorphic function $f : \mathcal{H} \to \mathbb{C}$ satisfying:

(1) $f|_k\gamma = f$ for all $\gamma \in \Gamma$.
(2) $f$ is "holomorphic at the cusps", i.e. for every $\sigma \in SL_2(\mathbb{Z})$, there exist $a_n \in \mathbb{C}$ such that, if one writes $q = e^{2\pi i z/N}$, then

$$(f|_k\sigma)(z) = a_0 + a_1 q + \ldots + a_n q^n + \ldots$$

where the series is absolutely convergent for $z \in H$, i.e. for $|q| < 1$.

If one requires that $a_0 = 0$ in every such expansion, i.e. $f$ vanishes at all the cusps, we say that $f$ is *cuspidal.*

*Remark* 1.1. Note the following:

(1) Since $\begin{pmatrix} 1 & N \\ 0 & 1 \end{pmatrix} \in \Gamma(N) \subseteq \Gamma$, condition (1) implies that $f(z+N) = f(z)$, which in turn gives a meromorphic Fourier expansion as above. Holomorphicity shows then that it suffices to check condition (2) at $\sigma$ such that $\overline{\sigma \cdot \infty} \in \partial(\Gamma \backslash \mathcal{H}) \subseteq \mathbb{Q} \cup \{\infty\}$ (the cusps).
(2) If $\Gamma = \Gamma(1) = SL_2(\mathbb{Z})$, $\Gamma$ is generated by $T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ and $S = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, so condition 1 is equivalent to $f(z) = f(z+1)$ and $f(-1/z) = z^k f(z)$. Also, $\Gamma$ has on ly one cusp at $\infty$, so that it suffices to check condition (2) for $\sigma = id$.
(3) If $\Gamma = \Gamma(1)$ and $f \neq 0$, $k$ is necessarily even, $\geq 0$ (the action of $S$ is of order 2).
(4) If $f$ is a modular form of weight $k$ for $\Gamma(N)$, a necessary and sufficient condition for it to be modular for $\Gamma_1(N)$ is that $f(z+1) = f(z)$ for all $z$, hence $f$ has a Fourier expansion with $q = e^{2\pi i z}$.
(5) If $f$ is a modular form of weight $k$ for $\Gamma_1(N)$, and $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_0(N)$, then $f|_k\gamma$ depends only on the image of $d$ in $(\mathbb{Z}/N\mathbb{Z})^\times$. We set $\langle d \rangle f = f|_k\gamma$. One has $\langle -1 \rangle f = (-1)^k f$.

**Definition 1.4.** Let $\varepsilon$ a Dirichlet character modulo $N$, i.e. a homomorphism

$$\varepsilon : (\mathbb{Z}/N\mathbb{Z})^\times \to \mathbb{C}^\times$$

We say that $\varepsilon$ is *even* (resp. *odd)* if $\varepsilon(-1) = 1$ (resp. if $\varepsilon(-1) = -1$). Let $k$ be an integer of the same parity as $\varepsilon$, i.e. $\varepsilon(-1) = (-1)^k$. We say that $f$ is a *modular form of level $N$, weight $k$ and Nebentypus $\varepsilon$* if $f$ is a modular form of weight $k$ for $\Gamma_1(N)$ such that

$$\langle d \rangle f = \varepsilon(d) \cdot f$$

for all $d \in (\mathbb{Z}/N\mathbb{Z})^\times$, i.e.

$$f\left(\frac{az+b}{cz+d}\right) = \varepsilon(d) \cdot (cz+d)^k f(z), \quad \forall \gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_0(N)$$

Note that if $k$ and $\varepsilon$ are not of the same parity, this formula implies $f = 0$, by considering the action of $-1$.

**Example 1.1.** If $k \geq 4$ is even, an example of such a function is the *Eisenstein series* of weight $k$, which is

$$G_k = \frac{1}{2}\zeta(1-k) + \sum_{n=1}^{\infty} \sigma_{k-1}(n) q^n$$

where $\zeta$ is the Riemann zeta function, and $\sigma_{k-1}(n) = \sum_{d|n} d^{k-1}$. Since $\zeta(1-k) = -\frac{B_k}{k}$, where $B_k$ is the $k$-th Bernoulli number $\left(\frac{t}{e^t-1} = \sum_{k=0}^{\infty} \frac{B_k}{k!} t^k\right)$, $G_k$ has rational coefficients (in fact, integral save the constant term). It is sometimes convenient to normalize it such that the constant term is 1, resulting in

$$E_k = -\frac{2k}{B_k}G_k = 1 - \frac{2k}{B_k}\sum_{n=1}^{\infty}\sigma_{k-1}(n)q^n$$

In particular, as $B_4 = -\frac{1}{30}$ and $B_6 = \frac{1}{42}$, $E_4$ and $E_6$ have integral coefficients.

**Theorem 1.1.** *The algebra of modular forms is generated over $\mathbb{C}$ by $E_4$ and $E_6$.*

**Example 1.2.** Consider the function

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

Then calculation shows that $\Delta$ is a modular form of weight 12 and indeed

$$\Delta = \frac{E_4^3 - E_6^2}{1728}$$

1.2. **Geometric perspective.** From a geometric point of view, a modular form of weight $k$ over $\Gamma_1(N)$ could be seen as a law, which associates to each elliptic curve $E$, equipped with an embedding $\alpha : \mu_N \to E$ of the $N$-th roots of unity, a section of $\omega_E^{\otimes k}$, where $\omega_E = Lie(E)^\vee$.
More precisely:

**Definition 1.5.** An *elliptic curve* over a scheme $S$ is a proper smooth morphism $E \to S$, equipped with a section $e : S \to E$, where the geometric fibers are elliptic curves. When $S = \mathrm{Spec}A$ for some commutative ring $A$, we say that $E$ is an elliptic curve over $A$. One sets $\omega_E = e^*\Omega^1_{E/S}$. When $S = \mathrm{Spec}A$, $\omega_E$ is an invertible $A$-module.

**Definition 1.6.** Let $R$ be a commutative ring, in which $N$ is invertible. A *modular form of weight $k$ for $\Gamma_1(N)$, meromorphic at $\infty$, defined over $R$* is a law which, for each elliptic curve $E$ over an $R$-algebra $A$, equipped with an embedding $\alpha : \mu_N \to E$, associates an element $f(E,\alpha) \in \omega_E^{\otimes k}$. This law should be compatible with isomorphisms and extension by scalars.

**Definition 1.7.** One says that $f$ is *holomorphic at $\infty$* if it can be extended to a law $\tilde{f}$ over pairs $(E,\alpha)$ where $E$ is a generalized elliptic curve (i.e. a proper flat morphism $E \to S$ with a group structure on $E^{reg}$), and $\alpha$ is such that its image meets every irreducible component of every geometric fiber. If it exists, the law $\tilde{f}$ is unique.

**Definition 1.8.** Let $R$ be a field. We say that $f$ is *cuspidal* if is holomorphic at infinity and $\tilde{f}(E,\alpha) = 0$ whenever $E$ is a degenerated elliptic curve (i.e. non-smooth) over an algebraically closed extension of $R$.

**Definition 1.9.** Let $\varepsilon : (\mathbb{Z}/N\mathbb{Z})^\times \to R^\times$ be a homomorphism. We say that $f$ is a *modular form of level $N$, weight $k$ and Nebentypus $\varepsilon$* if $f(E, d\alpha) = \varepsilon(d)\cdot f(E,\alpha)$ for all $d \in (\mathbb{Z}/N\mathbb{Z})^\times$.

**Example 1.3.** Let $R = \mathbb{C}$. Specifying an embedding $\alpha : \mu_N \to E$ amounts to choosing the image $\alpha(\zeta_N)$, which should be a point of order $N$. With a modular form $f$ defined as above, we can assoicate a function on $\mathcal{H}$ via

$$f(z) = f\left(E_z, 1/N\right)/(2\pi i \cdot du)^{\otimes k}$$

where $E_z = \mathbb{C}/(\mathbb{Z} \oplus z\mathbb{Z})$. If we exponentiate, we obtain, for $q = e^{2\pi i z}$

$$f(z) = \sum_{n=0}^{\infty} a_n q^n = f\left(\mathbb{C}^\times/q^{\mathbb{Z}}, Id\right)/(dt/t)^{\otimes k}, \quad 0 < |q| < 1$$

where $Id : \mu_N \hookrightarrow \mathbb{C}^\times$ is the natural embedding.

1.3. **Reduction of modular forms.**

**Definition 1.10.** Let $p$ be a prime number, and let $v_p$ be the $p$-adic valuation on $\mathbb{Q}$. A formal series

$$f = \sum_{n \geq 0} a_n q^n, \quad a_n \in \mathbb{Q}$$

is said to be *p-integral* if $v_p(a_n) \geq 0$ for all $n$. Its reduction modulo $p$ is the formal power series

$$\tilde{f} = \sum_{n \geq 0} \tilde{a}_n q^n \in \mathbb{F}_p[\![q]\!]$$

where $\tilde{a}_n$ is the image of $a_n$ in $\mathbb{F}_p$. If $f$ is a modular form of weight $k$ with a $p$-integral power series expansion, $\tilde{f}$ is called a *modular form mod p.*

We denote the space of modular forms mod $p$ which are images of a modular form of weight $k$ by $\tilde{M}_k$, and the space of all modular forms mod $p$ (which is a subalgebra of $\mathbb{F}_p[\![q]\!]$) by $\tilde{M}$.

**Theorem 1.2.** *(Swinnerton-Dyer, 1971) If $p \in \{2,3\}$, then $\tilde{M} = \mathbb{F}_p[\tilde{\Delta}]$. If $p \geq 5$, then $\tilde{M} = F_p[\tilde{E}_4, \tilde{E}_6]/\left(\tilde{A}(\tilde{E}_4, \tilde{E}_6) - 1\right)$ where $A \in \mathbb{Q}[X,Y]$ is the $(4,6)$-homogeneous polynomial of degree $p-1$ such that $A(E_4, E_6) = E_{p-1}$.*

**Example 1.4.** Let $p = 5$. Then $E_{p-1} = E_4$, hence $A(X,Y) = X$, and the ideal of relations is generated by $\tilde{E}_4 = 1$. Then $\tilde{M} \cong \mathbb{F}_5[\tilde{E}_6]$. Similarly, if $p = 7$, $\tilde{M} \cong \mathbb{F}_7[\tilde{E}_4]$. If $p = 13$, then

$$E_{12} = \frac{441E_4^3 + 250E_6^2}{691} \equiv 6E_4^3 - 5E_6^2 \mod 13$$

which gives us the fundamental relation $6\tilde{E}_4^3 - 5\tilde{E}_6^2 = 1$.

*Remark* 1.2. We equip $\mathbb{F}_p[X,Y]$ with a grading into $\mathbb{Z}/(p-1)\mathbb{Z}$ obtained by a quotient of the grading where $X$ is of weight 4 and $Y$ is of weight 6. Then $\tilde{A} - 1$ is homogeneous of weight 0, and the ideal it generates is therefore graded, hence the quotient $\tilde{M}$ is also graded, with a grading into $\mathbb{Z}/(p-1)\mathbb{Z}$. Therefore $\tilde{M} = \bigoplus_{\alpha \in \mathbb{Z}/(p-1)\mathbb{Z}} \tilde{M}^\alpha$ , with $\tilde{M}^\alpha = \bigcup_{k \equiv \alpha} \tilde{M}_k$.
Therefore, a modular form mod $p$ has a weight in $\mathbb{Z}/(p-1)\mathbb{Z}$.

**Definition 1.11.** Let $\alpha \in \mathbb{Z}/(p-1)\mathbb{Z}$, let $\varepsilon : (\mathbb{Z}/N\mathbb{Z})^\times \to \mathbb{F}_p^\times$, and let $f = \sum a_n q^n \in \tilde{M}^\alpha$ (if $p < 5$, just let $f \in \tilde{M}$ and $\alpha = k$). We let

$$f \mid U = \sum a_{pn} q^n$$

$$f \mid T_l = \sum a_{ln} q^n + \varepsilon(l) \cdot l^{\alpha - 1} \sum a_n q^{ln} \quad (l \neq p)$$

Then if $f \in \tilde{M}_k$, we have $f \mid U, f \mid T_l \in \tilde{M}_k$. We call $U$ and the $T_l$ *Hecke operators.*

The Hecke operators commute with each other and respect the filtration (note the map $f \mapsto f \cdot E_{p-1}$)

$$\tilde{M}_k \subseteq \tilde{M}_{k+p-1} \subseteq \dots$$

of each $\tilde{M}^{\alpha}$. In contrary to the classical case, these operators **are not semisimple**. For $p = 2$, they are in fact nilpotent.

*Remark* 1.3. One may allow $\varepsilon$ to attain values in a finite extension $F^{\times}$ of $\mathbb{F}_p$, extending by scalars to $\tilde{M} \otimes F$.

1.4. **Systems of Eigenvalues and Represntations of** $Gal(\overline{\mathbb{Q}}/\mathbb{Q})$**.** Let $F$ be a finite extension of $\mathbb{F}_p$. The operators $U$ and $T_l$ above can be linearly extended to the algebra $\tilde{M} \otimes F$ of modular forms with coefficients in $k$.

**Definition 1.12.** Let $0 \neq f \in \tilde{M}_k \otimes F$ be an eigenvector for the $T_l$ for all $l \neq p$, i.e. such that

$$f \mid T_l = a_l f, \quad a_l \in F$$

Then $f$ is called an *eigenform*.

Recall also the following.

**Definition 1.13.** Let $\rho : Gal(\overline{\mathbb{Q}}/\mathbb{Q}) \to GL_2(F)$ be a continuous representation. We say that $\rho$ is unramified at a prime $l$ if $\rho(I_l) = 1$ where $I_l$ is the inertia group at $p$.

We then have the following theorem.

**Theorem 1.3.** *(Deligne, 6.7) If $f$ is an eigenform of weight $k \geq 2$, Nebentypus $\varepsilon$ and level $N$, prime to $p$, with eigenvalues $a_l \in F$, then there exists a semi-simple continuous representation*

$$\rho : Gal(\overline{\mathbb{Q}}/\mathbb{Q}) \to GL_2(F)$$

*which is unramified away from $pN$, and such that*

$$Tr(\rho(Frob_l)) = a_l, \quad \det(\rho(Frob_l)) = \varepsilon(l) \cdot l^{k-1}$$

*for any prime $l \nmid pN$.*
*Moreover, this representation is unique up to a unique isomorphism, and it is irreducible iff $f$ is cuspidal.*

*Remark* 1.4. We note:

(1) If one writes $\chi_p : G_{\mathbb{Q}} \to \mathbb{F}_p^{\times}$ for the fundamental character (given by the action of $G_{\mathbb{Q}}$ on the $(p-1)$-th roots of unity), then $\det(\rho) = \chi_p^{k-1}$. This is an **odd** character, i.e. it sends the complex conjugation $c$ to $-1 \in \mathbb{F}_p^{\times}$. If $p = 2$, this is the trivial character.
(2) If $f$ is constant, then (as $\alpha = 0$ and $f \mid T_l = (1 + p^{-1}) \cdot l$), we have $a_l \equiv 1 + l^{-1} \mod p$, and the corresponding representation is $1 \oplus \chi_p^{-1}$. All other systems of eigenvalues can be obtained from a normalized eigenform, i.e. with $a_1 = 1$.
(3) For any integer $m \geq 0$, there exists a nonzero eigenvector $f_m$ of the $T_l$ with eigenvalues $l^m a_l$ (if $f = \sum a_n q^n$ is normalized, one can take $f_m = \theta^m f = \sum n^m a_n q^n$). The representation $\rho_m$ associated to $(l^m a_l)$ is $\rho \otimes \chi_p^m$. We say that it is obtained from $\rho$ by *twisting* by the character $\chi_p^m$.

1.5. **Serre's conjecture for $GL_2$ over $\mathbb{Q}$.** In 1973, Serre conjectured an inverse to this Theorem of Deligne.

**Conjecture 1.1.** *(Serre, 1973) Let $\rho : G_{\mathbb{Q}} \to GL_2(F)$ be a semi-simple continuous representation, unramified away from $p$. Since $\det(\rho)$ is a character into $\mathbb{F}_p^{\times}$, it is necessarily of the form $\chi_p^{\alpha-1}$ for some $\alpha \in \mathbb{Z}/(p-1)\mathbb{Z}$. Assume $\alpha$ is even, or equivalently that $\det(\rho)$ is odd. Then there exists a system of eigenvalues $(a_l)$ of the Hecke operators $T_l$ on some space $\tilde{M}_k$ with $k \in \alpha$ for which $\rho$ is the associated representation as above.*

(Why? First note that if $K$ is a local field, and $L$ is a Galois extension, we may define inertia $I$ and wild inertia $P$. We have

$$0 \to I/P \to G/P \to G/I \cong Gal(k_L/k) \to 0$$

hence $G/P$ acts on $I/P$ by conjugation. Since $I/P$ is abelian (in fact cyclic), this factors through an action of $Gal(k_L/k)$. Recall we have an isomorphism

$$I/P \cong \mu_N(k_L), \quad \sigma \mapsto [\sigma(\pi)/\pi]$$

which induces the natural action. Taking limits, one obtains

$$I_K/P_K \cong \prod_{l \neq p} \mathbb{Z}_l(1)$$

This should show that the restriction of $\rho$ to $I_p$ the inertia at $p$ is a power of the fundamental character. (Why?). What about the rest?)

Moreover, we have the following theorem.

**Theorem 1.4.** *(Atkin) Any system of eigenvalues for the $T_l$ can be obtained by a twist from a system coming from a modular form of weight $\leq p+1$.*

**Example 1.5.** Consider the case $p = 2$. There is a single system of eigenvalues, that is $a_l = 0$ for all $l$. It corresponds to the trivial representation (since $2 = 0$ in $\mathbb{F}_2$). This is because (it can be shown that) the $T_l$ are nilpotent, and furthermore, for all $i \geq 0$, $\Delta^i \mid T_l$ is a linear combination of $\Delta^j$ with $j < i$. This last result can be made precise:

$$\Delta^i \mid T_l \equiv (l+1)\Delta^i + a_1\Delta^{i-1} + \ldots + a_i \mod 8, \quad a_1, \ldots, a_i \in \mathbb{Z}$$

**Example 1.6.** Let $p \in \{3, 5, 7\}$. The only systems of eigenvalues are

$$a_l \equiv l^m + l^n \mod p \quad m, n \in \mathbb{Z}/(p-1)\mathbb{Z}, \quad 2 \nmid m+n$$

These correspond to the reducible representations $\chi_p^m \oplus \chi_p^n$ of odd determinant. Their number is $(p-1)^2/4$.

**Example 1.7.** $p \in \{11, 13, 17, 19\}$. Other than the systems $a_l \equiv l^m + l^n \mod p$ as above, one finds systems corresponding to irreducible representations $\rho$ with values in $GL_2(\mathbb{F}_p)$. Up to twist, these are:

(1) For $p = 11, 13$ the system $a_l = \tilde{\tau}(l)$ associated to the cuspidal form $\Delta$ of weight 12. The corresponding representation $\rho : G_{\mathbb{Q}} \to GL_2(\mathbb{F}_p)$ is surjective.
(2) For $p = 17$ there are 3 systems $a_l$, associated to cuspidal forms of weights $12, 16, 18$. The corresponding representations $\rho : G_{\mathbb{Q}} \to GL_2(\mathbb{F}_p)$ are surjective.

(3) For $p = 19$, there are 4 systems $a_l$, associated to cuspidal forms of weights $12, 16, 18, 20$. The corresponding representations $\rho : G_{\mathbb{Q}} \to GL_2(\mathbb{F}_p)$ are surjective, except the one corresponding to the weight 16 form, whose image lies in the subgroup of matrices with determinant which is a cube in $\mathbb{F}_{19}^{\times}$.

**Example 1.8.** $p = 23$. Other than the systems $a_l \equiv l^m + l^n \mod p$ as above, One finds:

(1) 5 systems with values in $\mathbb{F}_p$, associated to cuspidal forms of weights $12, 16, 18, 20, 22$; The corresponding representations are surjective, save the first whose image is isomorphic to the symmetric group $S_3$.
(2) 2 systems with values in $\mathbb{F}_{p^2}$, associated to two cuspidal forms of weight 24 conjugate over $\mathbb{Q}\left(\sqrt{144169}\right)$ (note that 144169 is not a square modulo 23). The corresponding representations have as image the subgroup of $GL_2(\mathbb{F}_{p^2})$ of elements whose determinant is in $\mathbb{F}_p^{\times}$.

The examples above for $p = 2, 3$ imply that to verify the conjecture it suffices to show that there are no such representations of $G_{\mathbb{Q}}$mod 2, and no such irreducible representations mod 3. This fact was established using bounds on the discriminant by Tate and Serre.

1.6. **Serre's weight conjecture - the recipe.** Recall that given a Galois representation, we wish to describe the level $N$ and the weight $k$ of the corresponding modular form. In order to understand Serre's recipe, we should first recall some definitions from Galois theory.

The recipe for $N$ will be simply the conductor of $\rho$:

**Definition 1.14.** Let $\rho : G_{\mathbb{Q}} \to GL(V)$ be a continuous homomorphism, where $V$ is a finite dimensional vector space over $\overline{\mathbb{F}}_p$. Then $\ker \rho = Gal(\overline{\mathbb{Q}}/K)$ for some finite Galois extension $K/\mathbb{Q}$. Let $l \neq p$ be a prime. Choose an extension to $\overline{\mathbb{Q}}$ of the $l$-adic valuation on $\mathbb{Q}$, and let

$$G_0 \supset G_1 \supset \cdots \supset G_i \supset \cdots$$

be the ramification groups of $G = Gal(K/\mathbb{Q})$ corresponding to that valuation. (recall $G_i = \{\sigma \in G \mid v_l(\sigma(x) - x) > i \quad \forall x\}$). Let $V_i = V^{G_i}$, and

$$n(l, \rho) = \sum_{i=0}^{\infty} \frac{1}{(G_0 : G_i)} \dim(V/V_i) =: \dim(V/V_0) + b(V)$$

We say that $b(V)$ is the *wild invariant* of the $G_0$-module $V$. Then

$$N = \prod_{l \neq p} l^{n(l,\rho)}$$

is the *Artin conductor of* $\rho$. (sometimes, simply the *conductor of* $\rho$)

*Remark* 1.5. Note that:

(1) $n(l, \rho) \geq 0$ is an integer.
(2) $n(l, \rho) = 0$ iff $G_0 = \{1\}$, i.e. iff $\rho$ is unramified at $l$.
(3) $n(l, \rho) = \dim V/V_0$ iff $G_1 = \{1\}$, i.e. iff $\rho$ is tamely ramified at $l$.

Therefore, $N$ is an integer (there are only finitely many ramified primes, hence nonzero powers), prime to $p$.

*Remark* 1.6. Note that $\det\rho : G_{\mathbb{Q}} \to \overline{\mathbb{F}}_p^{\times}$ is a character with conductor dividing $pN$, therefore it can be viewed as a homomorphism $(\mathbb{Z}/pN\mathbb{Z})^{\times} \to \overline{\mathbb{F}}_p^{\times}$. We write $\rho = \varphi\varepsilon$, where

$$\varphi : (\mathbb{Z}/p\mathbb{Z})^{\times} \to \overline{\mathbb{F}}_p^{\times}, \quad \varepsilon : (\mathbb{Z}/N\mathbb{Z})^{\times} \to \overline{\mathbb{F}}_p^{\times}$$

As $(\mathbb{Z}/p\mathbb{Z})^{\times}$ is cyclic of order $p-1$, the homomorphism $\varphi$ is of the form $x \mapsto x^h$ for some $h \in \mathbb{Z}/(p-1)\mathbb{Z}$, so that $\varphi = \chi_p^h$. Also, if $l \nmid pN$, then

$$\det(\rho(\mathrm{Frob}_l)) = l^h \varepsilon(l)$$

For the weight, we will need some more.

**Definition 1.15.** Let $G_p = Gal(\overline{\mathbb{Q}}_p/\mathbb{Q}_p)$, and let $\rho_p : G_p \to GL_2(\overline{\mathbb{F}}_p)$ be a continuous two dimensional representation of $G_p$. Let $I$ be the inertia group of $G_p$, and $I_p$ the largest pro-$p$-subgroup of $I$ (the group of wild inertia). The quotient $I_t := I/I_p$ is the *tame inertia group* of $G_p$. It can be identified with $\varprojlim \mathbb{F}_{p^n}^{\times}$. A character of $I_t$ is of *level* $n$ if ot factorizes through $\mathbb{F}_{p^n}^{\times}$, and does not factor through any $\mathbb{F}_{p^m}^{\times}$ where $m \mid n$ $(m \neq n)$.

Let $V^{ss}$ be the semi-simplification of $V$ with respect to the action of $G_p$. Then $I_p$ acts trivially on $V^{ss}$ (it is pro-$p$), so that it has an action of $I_t$. This action is diagonalizable and given by two characters

$$\varphi, \varphi' : I_t \to \overline{\mathbb{F}}_p^{\times}$$

It turns out that

**Proposition 1.1.** *(Serre) The characters $\varphi, \varphi'$ describing the action of $I_t$ on $V^{ss}$ are both of level 1 or both of level 2. If they are of level 2, they are conjugate: $\varphi' = \varphi^p$, $\varphi = \varphi'^p$.*

We can now consider separately each of the two cases.

**Definition 1.16.** Let $\rho : G_{\mathbb{Q}} \to GL_2(\overline{\mathbb{F}}_p)$ be a continuous Galois representation such that the associated characters $\varphi, \varphi'$ of $\rho_p$ are of level 2. In this case, $V$ is irreducible (or else it would contain a stable one dimensional subspace, and the action of $I_t$on that subspace would be by a character that can be extended to $G_p$ hence of level 1). Let $\psi, \psi' = \psi^p$ be the two fundamental characters of level 2 of $I_t$, i.e. the two characters $I_t \to \mathbb{F}_{p^2}^{\times} \to \overline{\mathbb{F}}_p^{\times}$ corresponding to the two embeddings of $\mathbb{F}_{p^2}$ into $\overline{\mathbb{F}}_p$. Then one can write $\varphi = \psi^{a+pb} = \psi^a \cdot \psi'^b$ for some $0 \leq a, b \leq p-1$. Then $b \neq a$, or else $\varphi = (\psi\psi')^a = \chi^a$, where $\chi$ is the restriction to $I$ of the cyclotomic character, contradicting the assumption on the level. Moreover, as $\varphi'$ is conjugate to $\varphi$, we have $\varphi' = \psi^b\psi'^a$. Thus, w.l.o.g. we may assume $0 \leq a < b \leq p-1$. We define the *weight associated to* $\rho$ to be

$$k := 1 + pa + b$$

*Remark* 1.7. a) The minimal possible value of $k$ is $k = 2$, which is obtained with $a = 0, b = 1$, i.e. in this case $\varphi$ and $\varphi'$ are equal to the two fundamental characters of level 2.

b) In the particular case of $a = 0$, one has $(\varphi, \varphi') = (\psi^b, \psi'^b)$ with $1 \leq b \leq p-1$, and by definition $k = b+1$, hence $2 \leq k \leq p$. The general can be obtained from this one by twisting. In fact, we can write

$$\rho_p = \chi^a \otimes \rho_p^{'}$$

where $\chi$ is the cyclotomic character (viewed as a character of $G_p$). The couple associated to $\rho_p^{'}$ is therefore $(0, b-a)$, and $k' = b-a+1$. One can write $k = k' + a(p+1)$.

**Definition 1.17.** Let $\rho : G_{\mathbb{Q}} \to GL_2(\overline{\mathbb{F}}_p)$ be a continuous Galois representation such that the associated characters $\varphi, \varphi'$ of $\rho_p$ are of level 1, and $I_p$ acts trivially. We may assume that the action of $I$ on $V$ is semi-simple, and given by characters $\varphi, \varphi'$ which are powers of the cyclotomic character $\chi^a, \chi^b$:

$$\rho_p \mid_I = \begin{pmatrix} \chi^a & 0 \\ 0 & \chi^b \end{pmatrix}$$

The integers $a, b$ are determined modulo $p-1$. We normalize them so that $0 \leq a, b \leq p-2$, and up to permuting the order, we may assume $0 \leq a \leq b \leq p-2$. We then define the *weight associated to* $\rho$ to be

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

*Remark* 1.8. a) Again, the minimal possible value of $k$ is $k = 2$, obtained when $a = 0, b = 1$, corresponding to the case $\varphi = 1, \varphi' = \chi$.
b) The case $(a,b) = (0,0)$, is the case where $I$ acts trivially on $V$, that is to say that $\rho_p$ is unramified. The general formula then gives $k = 1$. Since modular forms of weight 1 are somewhat exceptional, Serre is avoiding them explicitly.

Assume that $I_p$ acts non-trivially, i.e. that the action of $I$ is not tame. The elements of $V$ fixed by $I_p$ form a line $D$, which is stable by $G_p$. The action of $G_p$ on $V/D$ (resp. on $D$) is given by a character $\theta_1$ (resp. $\theta_2$) of $G_p$:

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

One can write $\theta_1, \theta_2$ uniquely in the form

$$\theta_1 = \chi^\alpha \varepsilon_1, \theta_2 = \chi^\beta \varepsilon_2 \quad \alpha, \beta \in \mathbb{Z}/(p-1)\mathbb{Z}$$

where $\varepsilon_1, \varepsilon_2$ are unramified characters of $G_p$ with values in $\overline{\mathbb{F}}_p^\times$. Restricting to $I$, we have

$$\rho_p \mid_I = \begin{pmatrix} \chi^\beta & * \\ 0 & \chi^\alpha \end{pmatrix}$$

We normalize the exponents so that $0 \leq \alpha \leq p-2$ and $1 \leq \beta \leq p-1$ (note that here they are not symmetric). Let $a = \inf(\alpha, \beta)$, $b = \sup(\alpha, \beta)$. Then
$(i)$If $\beta \neq \alpha + 1$, we let $k = 1 + pa + b$, as before.
$(ii)$ If $\beta = \alpha + 1$, the definition of the associated weight depends on the type of the wild ramification, which we proceed to define.
Let $K_0 = \mathbb{Q}_p^{nr}$ be the maximal unramified extension of $\mathbb{Q}_p$. It is the fixed field of $I$.
The group $\rho_p(I)$ is isomorphic to a Galois group of a certain totally ramified extension $K$ of $K_0$, and the wild inertia is mapped to $\rho_p(I_p)$, the Galois group of $K/K_t$, where $K_t$ is the maximal tamely ramified extension of $K_0$ contained in $K$.
Since $\beta = \alpha + 1$, one deduces that $Gal(K_t/K_0) = (\mathbb{Z}/p\mathbb{Z})^\times$, therefore $K_t = K_0(z)$ where $z$ is a primitive $p$-th root of unity. On the other hand, $Gal(K/K_t) = \rho_p(I_p)$ is an elementary abelian group of type $(p, p, \ldots, p)$, represented matricially by

$\begin{pmatrix} 1 & * \\ 0 & 1 \end{pmatrix}$. Moreover, from $\beta = \alpha + 1$ we see that the action via conjugation of $Gal(K_t/K_0) = (\mathbb{Z}/p\mathbb{Z})^\times$ on $Gal(K/K_t)$ is the natural one. By Kummer theory, we deduce that

$$K = K_t\left(x_1^{1/p}, x_2^{1/p}, \dots, x_m^{1/p}\right)$$

where $p^m = [K : K_t]$, and the $x_i$ are elements of $K_0^\times / \left(K_0^\times\right)^p$. If $v_p$ is the valuation on $K_0$, normalized so that $v_p(p) = 1$, we say that the extension $K$ is *peu ramifieé* if

$$v_p(x_i) \equiv 0 \mod p \quad \forall i = 1, \dots, m$$

i.e. if the $x_i$ could be chosen to be units of $K_0$. If not, we say that $K$ and $\rho_p$ are *trés ramifieé*s.

*Remark* 1.9. a) The trés ramifié case is only possible if $\varepsilon_1 = \varepsilon_2$, and then either $m = 1$ or $m = 2$. It can be seen by looking at the conjugation action of $G_p$ on $\rho_p(I_p)$.
b) Let $\pi$ be a uniformizer of $K_t$, for example $\pi = 1 - z$, or $\pi = p^{1/(p-1)}$. If $K/K_t$ is peu ramifiée, the $p^m - 1$ characters of order $p$ associated to that extension are all of conductor $(\pi^2)$. In the trés ramifié case, $p^m - p^{m-1}$ are of conductor $(\pi^{p+1}) = (p\pi^2)$ and the $p^{m-1} - 1$ others are of conductor $(\pi^2)$.

**Definition 1.18.** If $\beta = \alpha + 1$, and $\rho_p$ is peu ramifié, the associated weight is

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

If $\beta = \alpha + 1$, and $\rho_p$ is trés ramifié, the associated weight is

$$k = \begin{cases} 1 + pa + b + p - 1 = (\alpha + 1)(p + 1) & p \neq 2 \\ 4 & p = 2 \end{cases}$$

We now have finally a complete definition of the associated weight $k$.
This definition makes sense, due to the following:

**Proposition 1.2.** *If $k$ is the weight associated to $\rho_p$, we have*

$$\det \rho_p \mid I = \chi^{k-1}$$

Therefore, we have $\det \rho_p = \varepsilon_p \cdot \chi^{k-1}$, where $\varepsilon_p$ is an unramified character of $G_p$ with values in $\overline{\mathbb{F}}_p^\times$. When $\rho_p$ is the $p$-part of a global representation $\rho$, then $\varepsilon_p$ is the $p$-component of the character $\varepsilon$ defined above:

$$\varepsilon_p(\mathrm{Frob}_p) = \varepsilon(p)$$

For $p \neq 2$, $k$ attains the values in $[2, p^2 - 1]$ which can be written as $k = 1 + a_0 + pa_1$, with $0 \leq a_0, a_1 \leq p - 1$, with $a_1 \leq a_0 + 1$. If $p = 2$, $k = 2$ if $I_p$ acts trivially or peu ramifié, and $k = 4$ if it is trés ramifié.

**Example 1.9.** Let $p = 2$. Let $u : G_2 \to \mathbb{Z}/2\mathbb{Z}$ be a surjective homomorphism, and let $\rho_2 : G_2 \to GL_2(\mathbb{F}_2)$ be the representation defined by

$$s \mapsto \begin{pmatrix} 1 & u(s) \\ 0 & 1 \end{pmatrix}$$

Let $K/\mathbb{Q}_2$ be the quadratic extension corresponding to the kernel of $u$. The one has:

(1) $k = 2$ if $K/\mathbb{Q}_2$ is unramified, i.e. $K = \mathbb{Q}_2(\sqrt{5})$.
(2) $k = 2$ if $\mathrm{disc}(K/\mathbb{Q}_2) = (4)$, i.e. $K = \mathbb{Q}_2(\sqrt{-1})$ or $\mathbb{Q}_2(\sqrt{-5})$.
(3) $k = 4$ if $\mathrm{disc}(K/\mathbb{Q}_2) = (8)$, i.e. $K = \mathbb{Q}_2(\sqrt{2}), \mathbb{Q}_2(\sqrt{-2}), \mathbb{Q}_2(\sqrt{10}), \mathbb{Q}_2(\sqrt{-10})$.

**Conjecture 1.2.** *(Serre's weight conjecture)*
*Let $\overline{\rho} : G_{\mathbb{Q}} \to GL_2(\overline{\mathbb{F}}_p)$ be an odd irreducible continuous Galois representation. Then there exists a cuspidal eigenform $f \mod p$, and such that the associated representation$\rho_f$ is isomorphic to $\rho$. Moreover, $f$ can be chosen to be of level $N$, weight $k$ and Nebentypus $\varepsilon$, where $N$ is the (prime to $p$) conductor of $\rho$, $k$ is the associated weight of $\rho_p$, and $\varepsilon$ is such that $\det \rho = \varepsilon \cdot \chi^{k-1}$.*

## 2. $p$-adic Hodge Theory

The construction of the associated weight might seem quite arbitrary, as it comes from known results (at the time) about the reduction mod $p$ of the associated representation $\rho_f$. Using modern terminology, it is simplified.

Fontaine came up with the idea of classifying the $p$-adic Galois representations (restriction to decomposition groups of global Galois representations) using only linear algebraic data, and the rings of periods which supply us with a Hodge filtration, monodromy and Frobenius. I will not go into detail about the construction of these period rings, but using them one can classify $p$-adic Galois representations, so that the analogue of $l$-adic representations with good reduction are crystalline representations.

More precisely, there exists a certain ring of periods, a $\mathbb{Q}_p$-algebra, $\mathbb{B}_{cris}$, equipped with a filtration and a Frobenius, such that

**Definition 2.1.** A continuous representation $\rho : G_{\mathbb{Q}_p} \to GL(V)$ on a finite dimensional vector space $V$ over $\mathbb{Q}_p$ is *crystalline* if

$$\dim_{\mathbb{Q}_p} \left(\mathbb{B}_{cris} \otimes_{\mathbb{Q}_p} V\right)^{G_{\mathbb{Q}_p}} = \dim_{\mathbb{Q}_p} V$$

we set $D_{cris}(V) = (\mathbb{B}_{cris} \otimes_{\mathbb{Q}_p} V)^{G_{\mathbb{Q}_p}}$.

We then have the following result.

**Theorem 2.1.** *(Scholl, Faltings) Let$k \geq 2$, and let $f$ be a normalised cuspidal eigenform of level $N$ prime to $p$. Then $\rho_f |_{G_{\mathbb{Q}_p}}$ is crystalline. Moreover, if $v(a_p) > 0$ and $a_p^2 \neq 4p^{k-1}$, then $D_{cris}\left(\rho_f |_{G_{\mathbb{Q}_p}}\right) \cong D_{k,a_p}$ is equipped with a Frobenius having characteristic polynomial $X^2 - a_p x + p^{k-1}$ and a weakly admissible filtration with jumps at $0$ and $k-1$. (Maybe say a few words and draw the Newton and Hodge polygons).*

Therefore, any results on the reduction mod $p$ of crystalline representations of $G_{\mathbb{Q}_p}$ with Hodge-Tate weights (jumps in the above filtration) $k-1$ and $0$, give information about the possible associated weights.

In particular, for the case at hand, a result coming from the local-global compatiblity of the $p$-adic Langlands correspondence for $GL_2(\mathbb{Q}_p)$ is coming to our aid:

**Theorem 2.2.** *(Berger) If $V$ is an irreducible crystalline two-dimensional representation of $G_{\mathbb{Q}_p}$, whose reduction mod $p$ is reducible, then*

$$\overline{V} \cong \mu_1 \chi^r \oplus \mu_2$$

*where $\mu_1, \mu_2$ are unramified characters, and $\chi$ is the cyclotomic character. Moreover, $V$ has Hodge-Tate weights $\{0, r+1\}$.*

**Example 2.1.** Suppose that $\overline{\rho} : G_{\mathbb{Q}} \to GL_2(\overline{\mathbb{F}}_p)$ satisfies

$$\overline{\rho}\,|_{I_{\mathbb{Q}_p}} \cong \begin{pmatrix} \chi^{k-1} & * \\ 0 & 1 \end{pmatrix}$$

where $I_{\mathbb{Q}_p}$ is the inertia group at $p$, and $2 < k < p+1$. Then Serre's recipe predicts the minimal associated weight to be $k$. But by the above theorem, we know that any crystalline representation with Hodge-Tate weights $\{0, k-1\}$ which is reducible mod $p$ , must be of this form.

In general, the general theory of "change of weight" of Galois representations shows that it is reasonable to expect that the only obstructions to producing automorphic lifts of particular weights will be the local ones presecribed by $p$-adic Hodge theory.

*Remark* 2.1. Assume that in this example, $k < p-1$ and $*$ vanishes. Then

$$\overline{\rho} \otimes \chi^{1-k}\,|_{I_{\mathbb{Q}_p}} \cong \begin{pmatrix} \chi^{p-k} & 0 \\ 0 & 1 \end{pmatrix}$$

Note that for that representation, Serre's recipe gives conjectural weight $p+1-k$. Thus, although the conjecture predicts that any $\overline{\rho}$ has a twist which is modular with weight at most $p+1$, in the split case there are actually two such twists. This phenomenon is known as "companion forms".

## 3. Representation Theory

**Problem 3.1.** How does one generalize Serre's conjecture to $n$-dimensional Galois representations?

The idea is to pass to a representation-theoretic formulation, which appears first in the work of Ash. First, we would like to point out that the Hecke algebra (the algebra generated by the $T_l$, $U$ and $\langle d \rangle$) has a natural action on some cohomology groups.

**Definition 3.1.** Let $V$ be an $\overline{\mathbb{F}}_p$-representation of $GL_2(\mathbb{F}_p)$, and let $N$ be prime to $p$. Then $V$ is also a $\Gamma_1(N)$-module naturally. Let $\beta \in H^1(\Gamma_1(N), V)$.

For $g \in GL_2(\mathbb{Q})$, we write $\Gamma_g = g^{-1}\Gamma_1(N)g \cap \Gamma_1(N)$. Then we have morphisms $i(g), j(g) : (\Gamma_g, V) \to (\Gamma_1(N), V)$ given by $i(\gamma, v) = (\gamma, v)$ and $j(\gamma, v) = (g\gamma g^{-1}, g^{-1}v)$. For $l \in \mathbb{Z}$, we let $\sigma_l = diag(l^{-1}, 1)$, and $T_l = i(\sigma_l)_* j(\sigma_l)^*$ is an operator on $H^1(\Gamma_1(N), V)$, and for $r \in (\mathbb{Z}/N\mathbb{Z})^\times$, fix $\gamma_r \in \Gamma_0(N) \cap \Gamma_1(p)$ with $d \equiv r \mod N$. Then

$$\langle r \rangle = i(\gamma_r)_* j(\gamma_r)^* : H^1(\Gamma_1(N), V) \to H^1(\Gamma_1(N), V)$$

Now, given an eigenclass in $H^1(\Gamma_1(N), V)$, giving rise to a system of eigenvalues, we may wish to associate with it a continuous representation $\overline{\rho} : G_{\mathbb{Q}} \to GL_2(\overline{\mathbb{F}}_p)$. We also recall

**Theorem 3.1.** *(Eichler-Shimura - Fontatine, Messing, Faltings, Edixhoven) Let $2 \leq k < p$ be an integer. Let $p$ be a prime, $N$ an integer prime to $N$. Then there is an exact sequence of $\mathscr{H}(\Gamma_1(N), GL_2(\mathbb{Q}))$-modules*

$$0 \to S_k(\Gamma_1(N), \mathbb{F}_p) \to H^1\left(\Gamma_1(N), Sym^{k-2}\overline{\mathbb{F}}_p^2\right) \to S_k(\Gamma_1(N), \mathbb{F}_p)^\vee \to 0$$

Therefore, if $\overline{\rho}$ is odd and irreducible, we know that it is modular of weight $k$ and prime to $p$ level $N$ iff it is associated to an eigenclass in $H^1\left(\Gamma_1(N), \mathrm{Sym}^{k-2}\overline{\mathbb{F}}_p^2\right)$.
By devissage this is iff it is associated to an eigenclass in $H^1(\Gamma_1(N), V)$ for some JH factor $V$ of $\mathrm{Sym}^{k-2}\overline{\mathbb{F}}_p^2$. In fact, we may write $W(\overline{\rho})$ as the set of irreducible representations over $\overline{\mathbb{F}}_p$ of $GL_2(\mathbb{F}_p)$ such that $\overline{\rho}$ is associated to an eigenclass in $H^1(\Gamma_1(N), V)$ for some prime-to-$p$ level $N$. This is a finite set, and it determines all weights in which $\overline{\rho}$ occurs in prime-to-$p$ level. (not just the minimal such weight).

**Definition 3.2.** Let $\overline{\rho} : G_{\mathbb{Q}} \to GL_n(\overline{\mathbb{F}}_p)$ be continuous odd irreducible Galois representation. We call $W(\overline{\rho})$ the *Serre weights* of $\overline{\rho}$. A *Serre weight* is an irreducible representation over $\overline{\mathbb{F}}_p$ of $GL_n(\mathbb{F}_p)$.

**Example 3.1.** Return to our example from before, with $2 < k < p-1$. If $* \neq 0$, i.e. this is a non-split extension, we have $W(\overline{\rho}) = \{\mathrm{Sym}^{k-2}\overline{\mathbb{F}}_p^2\}$. However, when $* = 0$, we have

$$W(\overline{\rho}) = \left\{\mathrm{Sym}^{k-2}\overline{\mathbb{F}}_p^2, \det{}^{k-1} \otimes \mathrm{Sym}^{p-1-k}\overline{\mathbb{F}}_p^2\right\}$$

One could see the second weight from observing that $p-1-k$ should lie in $W(\overline{\rho} \otimes \chi^{1-k})$ and undoing the twist. This in fact follows from the mod $p$ Langlands correspondence for $GL_2(\mathbb{Q}_p)$.
Quite more generally, we let

**Definition 3.3.** Let $q$ be a power of a prime $p$. A *Serre weight* is an isomorphism class of irreducible representations over $\overline{\mathbb{F}}_p$ of $GL_n(\mathbb{F}_q)$. Let $W(\mathbb{F}_q, n)$ be the set of all Serre weights.

*Remark* 3.1. The set $W(\mathbb{F}_q, n)$ admits a simple description, by some results from modular representation theory. It is indexed by the set $\left(X_1^{(n)}\right)^S / \sim$, where $S$ is the set of embedding $\mathbb{F}_q \hookrightarrow \overline{\mathbb{F}}_p$ and

$$X_1^{(n)} = \left\{(a_i)_{i=1}^n \in \mathbb{Z}_+^n \mid a_i - a_{i+1} \leq p-1\right\}$$

## 4. Generalization of Serre's conjecture

4.1. **Global Setting.** Although we are mainly interested today in the generalization of the weight part, which is expected to be purely local, I'll briefly sketch the global setting.
Let $F$ be a number field, and let

$$\overline{\rho} : G_F \to GL_n(\overline{\mathbb{F}}_p)$$

be an irreducible representation. Recall that we wanted to understand the weights appearing in such representations coming from modular forms. Also, we have seen that modular forms are equivalent to systems of Hecke eigenvalues, which, in turn, are equivalent, by Eichler-Shimura, to cohomology classes of the level group (same as cohomology of corresponding modular curve) with coefficients in an irreducible representation.

Therefore, in order to generalize this notion, we should look at cohmology of higher-dimensional local symmetric spaces with some fixed level, and coefficients in an irreducbile representation.

Instead of the modular curve, we may consider

$$Y(U) := GL_n(F)\backslash GL_n(\mathbb{A}_F)/UA_\infty^\circ U_\infty^\circ$$

where $\mathbb{A}_F$ are the adeles of $F$, $U = U^pU_p$ is a compact open subgroup of $GL_n(\mathbb{A}_F^\infty)$ (finite adeles), where $U^p$ is sufficiently small, and $U_p = GL_n(\mathcal{O}_F \otimes \mathbb{Z}_p)$ (this corresponds to $p \nmid N$ in the modular curve case). Also $A_\infty^\circ = \mathbb{R}_{>0}^\times$, embedded diagonally in $\prod_{v|\infty} GL_n(F_v)$, and $U_\infty^\circ = \prod_{v|\infty} U_v^\circ$ where $U_v^\circ = SO_n(\mathbb{R})$ if $v$ is real and $U_v^\circ = U_n(\mathbb{R})$ if $v$ is complex.

Let $W$ be an irreducible smooth $\overline{\mathbb{F}}_p$-representation of $U_p$. The action of $U_p$ factors through $\prod_{v|p} GL_n(k_v)$ (the kernel of the reduction is a normal pro-$p$ subgroup). Write $W = \bigotimes_{v|p} W_v$ with $W_v$ an irreducible $\overline{\mathbb{F}}_p$-representation of $GL_n(k_v)$. Let

$$\mathcal{W} := \left((GL_n(F)\backslash GL_n(\mathbb{A}_F)/U^pA_\infty^\circ U_\infty^\circ) \times W\right)/U_p$$

be a local system on $Y(U)$.

Let $\Sigma_0$ be a finite set of places of $F$ (depending on $U$) which contains all places $v \mid p$, and such that if $v \notin \Sigma_0$ is finite, then $U_v = GL_n(\mathcal{O}_{F_v})$ ($Y(U)$ is unramified outside $\Sigma_0$).

Then, for each $v \notin \Sigma_0$, the spherical Hecke algebra

$$\mathcal{H}_v := C\left(GL_n(\mathcal{O}_{F_v})\backslash GL_n(F_v)/GL_n(\mathcal{O}_{F_v}), \overline{\mathbb{Z}}_p\right)$$

acts naturally on the cohomolgy groups $H^i(Y(U), \mathcal{W})$.

Let $\Sigma$ be a finite set of places of $F$ containing $\Sigma_0$ and the places where $\overline{\rho}$ is ramified. (non-trivial on the inertia group).

We may now define a maximal ideal $\mathfrak{m} = \mathfrak{m}(\overline{\rho}, U, \Sigma)$ of the Hecke algebra $\mathbb{T}_\Sigma := \otimes'_{v\notin\Sigma}\mathcal{H}_v$ with residue field $\overline{\mathbb{F}}_p$ by demanding that for all $v \notin \Sigma$, the semisimple part of $\overline{\rho}(\mathrm{Frob}_v^{-1})$ be conjugate to the class defined by the $\mathcal{H}_v$-eigenvalues determined by $\mathfrak{m}$. (equivalently, for $GL_n$, specify the characteristic polynomial of $\overline{\rho}(\mathrm{Frob}_v)$ or its eigenvalues).

**Definition 4.1.** We say that $\overline{\rho}$ is *automorphic* if there are some $W, U, \Sigma$ as above such that $H^i(Y(U), \mathcal{W})_\mathfrak{m} \neq 0$ for some $i \geq 0$.

**Example 4.1.** For $F = \mathbb{Q}$, $n = 2$, with $U = \Gamma_1(N)$, $p \nmid N$, $Y(U) = X_1(N)$ is the modular curve of level $N$. Take $W = \mathrm{Sym}^{k-2}\overline{\mathbb{F}}_p^2$ for some $k \geq 2$, and an eigenform determines a system of Hecke eigenvalues, giving rise to a morphism $a : \mathbb{T}_\Sigma \to \overline{\mathbb{F}}_p$, and a maximal ideal$\mathfrak{m} = \ker a$ with residue field $\overline{\mathbb{F}}_p$. Moreover, such an eigenform corresponds to an eigenclass in $H^1(Y(U), \mathcal{W})$ by Eichler-Shimura, therefore a nonzero element in $H^1(Y(U), \mathcal{W})_\mathfrak{m}$. Conversely, for $i = 0, 2$ these vanish for all $\mathfrak{m}$, and for $i = 1$, such an element is an eigenclass, hence corresponds to an eigenform.

**Definition 4.2.** Let $\overline{\rho} : G_F \to GL_n(\overline{\mathbb{F}}_p)$ be automorphic. Let $W(\overline{\rho})$ denote the set of isomorphism classes of irreducible representations $W$ of $\prod_{v|p} GL_n(k_v)$ for which $H^i(Y(U), \mathcal{W})_\mathfrak{m} \neq 0$. We say that $W(\overline{\rho})$ is the set of *Serre weights* of $\overline{\rho}$.

We may now formulate a generalization of the "weak" Serre conjecture.

**Conjecture 4.1.** *Suppose that $\overline{\rho}$ is automorphic. Then we may write $W(\overline{\rho}) = \bigotimes_{v|p} W_v(\overline{\rho})$, where $W_v(\overline{\rho})$ is a set of isomorphism classes of irreducible representations of $GL_n(k_v)$ which depends only on $\overline{\rho}\mid_{G_{F_v}}$.*

In fact, as in the $GL_2(\mathbb{Q}_p)$ case, one expects that this set will only depend on $\overline{\rho}\mid_{I_{F_v}}$.

## 5. The Breuil-Mezard Conjecture

We now move to a local setting. Let $K$ be a finite extension of $\mathbb{Q}_p$ (corresponding to $F_v$ for some $v \mid p$), with ring of integers $\mathcal{O}_K$ and residue field $k$. Let $I_K$ be the inertia subgroup of the absolute Galois group $G_K = Gal(\overline{K}/K)$. Let $\mathrm{Frob}_K$ be a geometric Frobenius element of $G_K$. Let $E$ be a finite extension of $\mathbb{Q}_p$, with ring of integers $\mathcal{O} = \mathcal{O}_E$, and residue field $\mathbb{F}$. It is now natural the ask

**Problem 5.1.** Given $\overline{\rho} : G_K \to GL_n(k)$, what is $W_v(\overline{\rho})$, the set of corresponding Serre weights?

In order to describe the possible sets, we would need to recall some more definitions.

**Definition 5.1.** Let $\mathbb{Z}_+^n$ denote the set of tuples $\lambda = (\lambda_1, \ldots, \lambda_n)$ with $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$. A *Hodge type* is an element of $\left(\mathbb{Z}_+^n\right)^{S_K}$, where $S_K = \{K \hookrightarrow E\}$.

**Definition 5.2.** An *inertial type* is a representation $\tau : I_K \to GL_n(E)$ with open kernel, which can be extended to a representation of the Weil group $W_K$.

We have the following result (recall $-i \in HT_\kappa(\rho)$ with multiplicity $\dim_E \left(\rho \otimes_{\kappa,K} \hat{\overline{K}}(i)\right)^{G_K}$)

**Theorem 5.1.** *(Kisin) Let $\lambda$ be a Hodge type and $\tau$ be an inertial type. There is a unique reduced and p-torsion free quotient $R^{\lambda,\tau}_{\overline{\rho},\mathcal{O}}$ of the the universal lifting $\mathcal{O}$-algebra $R_{\overline{\rho},\mathcal{O}}$ whose points parametrise lifts of $\overline{\rho}$ that are potentially crystalline with inertial type $\tau$ and Hodge-Tate weights*

$$HT_\kappa(\rho) = \{\lambda_{\kappa,1} + n - 1, \ldots, \lambda_{\kappa,n-1} + 1, \lambda_{\kappa,n}\}$$

*The ring $R^{\lambda,\tau}_{\overline{\rho},\mathcal{O}}[1/p]$ is regular.*

Given an inertial type, we may associate with it a finite-dimensional smooth irreducible $\overline{\mathbb{Q}}_p$-representation $\sigma(\tau)$ of $GL_n(\mathcal{O}_K)$ by "inertial local Langlands correspondence", namely

**Theorem 5.2.** *(Caraiani, Emerton, Gee)*
*Let $\tau$ be an inertial type. Then there is a finite dimensional smooth irreducible $\overline{\mathbb{Q}}_p$-representation $\sigma(\tau)$ of $GL_n(\mathcal{O}_K)$ such that if $\pi$ is any irreducible smooth $\overline{\mathbb{Q}}_p$-representation of $G$, then $\pi\mid_{GL_n(\mathcal{O}_K)}$ contains an isomorphic copy of $\sigma(\tau)$ as a subrepresentation iff $rec_p(\pi)\mid_{I_K} \sim \tau$ and $N = 0$ on $rec_p(\pi)$. Furthermore, in this case the restriction of $\pi$ to $GL_n(\mathcal{O}_K)$ contains a unique copy of $\sigma(\tau)$.*

*Remark* 5.1. In particular, if $\tau$ is the trivial inertial type, then $\sigma(\tau) \cong \overline{\mathbb{Q}}_p$ is the trivial one-dimensional representation of $GL_n(\mathcal{O}_K)$.

Moreover, for $GL_2(\mathbb{Q}_p)$, we know that this deformation ring has an intimiate connection with the Serre weights. This is a result following from $p$-adic local Langlands.

**Definition 5.3.** Let $(A, \mathfrak{m})$ be a local ring of dimension $d$. Let $P(n)$ be the Hilbert polynomial of $A$, i.e. $P(n) = \mathrm{length}(A/\mathfrak{m}^{n+1}A)$. Let $P(X) = \sum_i a_i X^i$. Then $e(A) = d! \cdot a_d$ is the Hilbert-Samuel multiplicity of $A$.

**Theorem 5.3.** *(Kisin) For $\overline{\rho}$ sufficiently generic (*), the Hilbert-Samuel multiplicity of the local ring $R^{k,\tau}_{\overline{\rho},\mathcal{O}}/(\varpi_E)$ is equal to the number of Serre weights of $\overline{\rho}$ (computes with multiplicity) that appear in $\overline{\sigma}(k,t)$, where $\sigma(k,t) = \sigma(\tau) \otimes_E Sym^{k-2}E^2$.*

In fact, this is more than mere numerology

**Theorem 5.4.** *(Breuil, Mezard) If $E$ is large enough, there exists a bijection, which respects multiplicities, between the irreducible components of $R^{k,\tau}_{\overline{\rho},\mathcal{O}}/(\varpi_E)$ and the collection of distinct Serre weights of $\overline{\rho}$ in $\overline{\sigma}(k,t)$*

Therefore, if we would like to know the expected weights of a certain modular Galois representation, we should understand the irreducible components of the reduction of the deformation ring.

In order to generalize the representations $\sigma(k,t)$, we have to consider the reductions of irreducible algebraic representations.

In general, given an element $\lambda \in \mathbb{Z}^n_+$, it can be viewed as a dominant weight for $GL_n$, hence we can construct the corresponding algebraic $\mathcal{O}_K$-representation of $GL_n$

$$H^0_{\mathcal{O}_K}(\lambda) := Ind_{B_n}^{GL_n}(w_0\lambda)_{/\mathcal{O}_K}$$

where $B_n$ is the Borel subgroup of upper triangular matrices of $GL_n$ and $w_0$ is the longest element of the Weyl group. We write $M_\lambda$ for its $\mathcal{O}_K$-points. Then for a Hodge type $\lambda$, we may write

$$L_\lambda = \bigotimes_{\kappa \in S_K} \left( M_{\lambda_\kappa} \otimes_{\mathcal{O}_{K,\kappa}} \mathcal{O} \right)$$

Now, $\sigma(\tau)$ is a finite-dimensional of the compact group, hence contains a $GL_n(\mathcal{O}_K)$-stable $\mathcal{O}$-lattice $L_\tau$. Set $L_{\lambda.\tau} := L_\lambda \otimes_{\mathcal{O}} L_\tau$. Then the collection of distinct Serre weights in its reduction can be given by multiplicities by

$$(L_{\lambda,\tau} \otimes_{\mathcal{O}} \mathbb{F})^{ss} \cong \bigoplus_{a \in W(k,n)} F_a^{n_{\lambda,\tau}(a)}$$

where $n_{\lambda,\tau}(a)$ are nonnegative integers.

We may then generalize the result of Breuil-Mezard and obtain

**Conjecture 5.1.** *(Breuil-Mezard, generalized by Gee, Herzig, Savitt) There exist non-negative integers $\mu_a(\overline{\rho})$ depending only on $a$ and $\overline{\rho}$ such that for all Hodge types $\lambda$ and all inertial types $\tau$ we have*

$$e\left(R^{\lambda,\tau}_{\overline{\rho}}/\varpi\right) = \sum_a n_{\lambda,\tau}(a)\mu_a(\overline{\rho})$$

*Remark* 5.2. One may wonder why do we not conjecture the integers to be all in $\{0,1\}$. In fact, if we restrict the range of $a$ to be Fontaine-Laffaille regular, and $K/\mathbb{Q}_p$ is unramified, this is true. But we have

**Example 5.1.** In the case of $K = \mathbb{Q}_p$, $n = 2$, if$k = p$ and $\overline{\rho}$ is split, then $\mu_k(\overline{\rho}) = 2$. This multiplicity is coming from the fact that $\chi^{p-1}$ and 1 have both the same reduction modulo $p$.

Therefore, if we wish to extend the conjecture to all representations, such a statement is better.

This conjecture, in fact, is now known to hold (Gee, Kisin) for a certain type of representations (potentially Barsotti-Tate).

We may now define

**Definition 5.4.** We define $W_{BM}(\overline{\rho})$, the set of *Breuil-Mezard predicted weights* for $\overline{\rho}$ to be the set of Serre weights $a$ such that $\mu_a(\overline{\rho}) > 0$.

This leads us to our first possible generalization of the weight part of Serre's conjecture

**Conjecture 5.2.** *One has* $W_v(\overline{\rho}) = W_{BM}(\overline{\rho})$.

In fact, we can generalize this statement. Since we are allowing for linear combinations of the multiplicities, all we need for them is to be uniquely defined by these equations, so a natural definition is: (we don't need all types)

**Definition 5.5.** We say that a set $\mathcal{S} = \{(\lambda, \tau)\}$ where $\lambda$ is a Hodge type and $\tau$ is an inertial type is a *Breuil-Mezard system* if the map $\mathbb{Z}^{W(k,n)} \to \mathbb{Z}^{\mathcal{S}}$ given by

$$(x_a)_{a \in W(k,n)} \mapsto \left( \sum_a n_{\lambda,\tau}(a) x_a \right)_{(\lambda,\tau) \in \mathcal{S}}$$

is injective.

In particular, the above equations can have at most one solution.

**Example 5.2.** Let $n = 2$ and let $BT = \{(0, \tau)\}$ so that $BT$ is the set of potentially Barsotti-Tate types. Then Gee and Kisin show that $BT$ is a Breuil-Mezard system, even if one restrict to $\tau$ such that $\det \tau$ is tame.

**Definition 5.6.** We say that a Hodge type $\lambda$ is a *lift* of an element $a \in \left(\mathbb{Z}_+^n\right)^{S_k}$ if for all $\sigma \in S_k$ there exists a $\kappa_\sigma \in S_K$ lifting $\sigma$ such that $\lambda_{\kappa_\sigma} = a_\sigma$, and $\lambda_{\kappa'} = 0$ for all other $\kappa' \neq \kappa_\sigma$ in $S_K$ lifting $\sigma$. In that case we say that the lift $\lambda$ is taken with respect to the choice of embeddings $(\kappa_\sigma)$. When $a \in \left(X_1^{(n)}\right)^{S_k}$, we will also say that $\lambda$ is a *lift* of the Serre weight represented by $a$.

**Example 5.3.** For each Serre weight $b$, fix a lift $\lambda_b$, and let $\tilde{\text{cr}}$ be the set of pairs $(\lambda_b, triv)$, where $triv$ denotes the trivial type. Then $\tilde{cr}$ is a Breuil-Mezard system.

Thus, we may replace our hopeful conjecture by one, which is more modest, call this set of weights $W_{\mathcal{S}}(\overline{\rho})$.

In this case, a weak version of the Breuil-Mezard conjecture for representations of type $\tilde{\text{cr}}$ is known to hold - there are uniquely determined integers $\mu_a(\overline{\rho})$ satisfying the equations. (For the existence of rational solutions, a counting argument suffices). For $n = 2$, it follows trivially that these are nonnegative integers, and that $a \in W_{\tilde{\text{cr}}}(\overline{\rho})$ iff $\overline{\rho}$ has a crystalline lift of Hodge type $\lambda_a$.

## 6. Crystalline Lifts and Serre Weights

The Breuil-Mezard version of the weight part of Serre's conjecture has the obvious drawback that even the definition of the conjectural set of weights $W_{BM}(\overline{\rho})$ depends on the Breuil-Mezard conjecture. (Of course, in theory it is possible to determine the conjectural values of $\mu_a(\overline{\rho})$ without proving the conjecture by computing the HS multiplicity $e(R/\omega)$ for enough choices of $\lambda, \tau$ but in practive that seems very difficult. In the case where $\overline{\rho}\mid_{I_K}$ is semisimple, it is possible to define the set of weights in a different fashion.

**Definition 6.1.** Suppose that $\lambda \in (\mathbb{Z}_+^n)^{S_K}$. A *crystalline lift* of $\overline{\rho}$ of Hodge type $\lambda$ is a representation $\rho : G_K \to GL_n\left(\overline{\mathbb{Z}}_p\right)$ such that

- $\rho \otimes_{\overline{\mathbb{Z}}_p} \overline{\mathbb{F}}_p \cong \overline{\rho}$
- $\overline{\rho} \otimes_{\overline{\mathbb{Z}}_p} \overline{\mathbb{Q}}_p$ is crystalline and regular of weight $\lambda$.

**Lemma 6.1.** *Assume that the generalized Breuil-Mezard conjecture holds. Then $\overline{\rho}$ has a crystalline lift of Hodge type $\lambda$ iff $W_{BM}(\overline{\rho}) \cap JH_{GL_n(k)}\left(L_\lambda \otimes_{\overline{\mathbb{Z}}_p} \overline{\mathbb{F}}_p\right) \neq \emptyset$.*

**Corollary 6.1.** *Assume that the generalized Breuil-Mezard conjecture holds, and let $\lambda$ be a lift of the Serre weight $a$. If $a \in W_{BM}(\overline{\rho})$, then $\overline{\rho}$ has a crystalline lift of Hodge type $\lambda$.*

This motivates the following definition

**Definition 6.2.** Let $W^{\exists}_{cris}(\overline{\rho})$ be the set of *crystalline weights for* $\overline{\rho}$, the set of Serre weights $a$ such that $\overline{\rho}$ has a crystalline lift of Hodge type $\lambda$ for some lift $\lambda$ of $a$. Let $W^{\forall}_{cris}(\overline{\rho})$ be the set of Serre weights $a$ such that $\overline{\rho}$ has a crystalline lift of Hodge type $\lambda$ for every lift $\lambda$ of $a$.

Assuming the generalized Breuil-Mezard conjecture, we have $W_{BM}(\overline{\rho}) \subseteq W^{\forall}_{cris}(\overline{\rho}) = W^{\exists}_{cris}(\overline{\rho})$. If $\overline{\rho}\mid_{I_K}$ is semisimple, we may conjecture:

**Conjecture 6.1.** *One has $W^{\exists}_{cris}(\overline{\rho}) = W^{\forall}_{cris}(\overline{\rho})$. Moreover:*
*(i) If $\overline{\rho}\mid_{I_K}$ is semisimple, then $W_{BM}(\overline{\rho}) = W^{\exists}_{cris}(\overline{\rho}) = W^{\forall}_{cris}(\overline{\rho})$.*
*(ii) If $\overline{\rho}\mid_{I_{F_v}}$ is semisimple for all $v \mid p$, then $W_v(\overline{\rho}) = W^{\exists}_{cris}(\overline{\rho}\mid_{G_{F_v}})$.*

In general, we dare not hope that this happens when dropping the semisimplicity assumption. If time allows we will see an example later on (it will require some work to identify these sets).
There is much evidence for this conjecture in the case of $GL_3(\mathbb{Q}_p)$ which is quite striking.

*Remark* 6.1. When $n = 1$, this is a consequence of CFT and analysis of the reduction of crystalline characters. e.g. the following Lemma.

**Lemma 6.2.** *Let $\Lambda = \{\lambda_\kappa\}_{\kappa \in S_K}$ be a collection of integers.*

*(i) There is a crystalline character $\psi^K_\Lambda : G_K \to \overline{\mathbb{Z}}_p^\times$ such that for each $\kappa \in S_K$ we have $HT_\kappa(\psi^K_\Lambda) = \lambda_\kappa$. It is uniquely determined up to unramified twists.*

*(ii) We have $\overline{\psi}^K_\Lambda \mid_{I_K} = \prod_{\sigma \in S_k} \chi_\sigma^{b_\sigma}$, where $b_\sigma = \sum_{\kappa \in S_K : \overline{\kappa} = \sigma} \lambda_\kappa$.*

*Remark* 6.2. When $n = 2$, $p > 2$, part (i) is known, some analogue of part (iii) is known (for quaternion algebras). If furthermore $K = \mathbb{Q}_p$, then part (ii) is known whenever the Breuil-Mezard conjecture is known. All these results hold without the assumption on semisimplicity.
We also note that the weights in $W_{BM}(\overline{\rho})$ and $W^{\exists}_{cris}(\overline{\rho})$ which are in the closure of the lowest alcove (i.e. $a_{\sigma,1} - a_{\sigma,n} + (n-1) \leq p$ for all $\sigma$) must always coincide. Indeed, if $\lambda$ is a lift of such a weight, then $L_\lambda \otimes \overline{\mathbb{F}}_p$ is irreducible. In particular, when $n \leq 2$ this is true for all Serre weights, hence the progress for $n \leq 2$ provides very weak evidence.
Combining everything we get the following conjecture:

**Conjecture 6.2.** *Suppose that $\overline{\rho}\mid_{I_K}$ is semisimple. If $W^{\exists}_{cris}(\overline{\rho}) \cap JH_{GL_n(k)}(L_\lambda \otimes_{\overline{\mathbb{Z}}_p} \overline{\mathbb{F}}_p) \neq \emptyset$ for some lift $\lambda$ of the Serre weight $a$, then $a \in W^{\exists}_{cris}(\overline{\rho})$.*

## 7. The Geometric Picture

In a series of papers, Gee and Emerton construct a finite type equidimensional Artin stack $\overline{\mathcal{X}}$ over $\mathbb{F}_p$ whose $\overline{\mathbb{F}}_p$-points correspond to isomorphism classes of representations $\overline{\rho} : G_K \to GL_n(\overline{\mathbb{F}}_p)$ that admit a de-Rham lift to $GL_n(\overline{\mathbb{Z}}_p)$. Furthermore, for each Hodge type $\lambda$ and inertial type $\tau$, there is a finite type formal Artin stack $\mathcal{X}_{\lambda,\tau}$ over $\mathrm{Spf}\mathbb{Z}_p$ , whose $\overline{\mathbb{Z}}_p$-points are in natural bijection with the isomorphism classes of the de-Rham representations of type $(\lambda, \tau)$. There is a specialization morphism $\pi$ which on points is just the reduction. The underlying reduced substack of $\pi(\mathcal{X}_{\lambda,\tau})$ is a union of irreducible components of $\overline{\mathcal{X}}$. Each irreducible component of $\overline{\mathcal{X}}$ has a dense open subset of closed points that lie only on that component, and which correspond to certain maximally non-split upper-triangular representations with characters $\chi_1, \dots, \chi_n$ on the diagonal such that the characters are fixed. These are the *generic* $\overline{\mathbb{F}}_p$*-points* of this component.

**Example 7.1.** Let $n = 2$. Fix characters $\psi_i : I_K \to \overline{\mathbb{F}}_p^\times$ for $i = 1, 2$ that extend to $G_K$. Then whenever $\psi_1\psi_2^{-1} \neq \overline{\varepsilon}$, there is a unique component whose generic $\overline{\mathbb{F}}_p$-points correspond to extensions of $\chi_2$ by $\chi_1$ with $\chi_i \mid_{I_K} \cong \psi_i$, and these representations have a unique Serre weight. This Serre weight can be read off directly from the tame inertial weights.

When $\psi_1\psi_2^{-1} = \overline{\varepsilon}$, suppose that $K = \mathbb{Q}_p$. Then there is one component of $\overline{\mathcal{X}}$ whose generic $\overline{\mathbb{F}}_p$-points are tres ramifiee extensions of $\chi$ by $\chi\overline{\varepsilon}$, where $\chi$ is any unramified character, and another component whose generic points are extensions of $\chi_2$ by $\chi_1\overline{\varepsilon}$ where $\chi_1 \neq \chi_2$ are any unramified characters. The peu ramifiee extensions of $\chi$ by $\chi\overline{\varepsilon}$ lie on both components (and so are not generic of neither). We label the first component by the Serre weight $\mathrm{Sym}^{p-1}\overline{\mathbb{F}}_p^2$ (note that this corresponds to the weight $k + 1$, which is Serre's prediction), and the second component by both 1 and $\mathrm{Sym}^{p-1}\overline{\mathbb{F}}_p^2$ (the first coming from peu ramifiee extensions, where the expected weight is 2, and the second its companion) - the two Serre weights of a generic point on this component. In particular, every component of $\overline{\mathcal{X}}$ labeled by 1 is also labeled by $\mathrm{Sym}^{p-1}\overline{\mathbb{F}}_p^2$. All other components of $\overline{\mathcal{X}}$ are labeled by a single Serre weight, and in fact, each other Serre weight is the label for a unique irreducible component.

In general, we expect that to each component there will be set of weights, and the Serre weights of $\overline{\rho}$ will be the union of the sets of weights associated to the components it lies on. In particular, the labels of a component must therefore be the Serre weights of its generic points. This structure should be a consequence of the Breuil-Mezard conjecture (and in fact it is for $n = 2$).

Accordingly, to understand the weight part of Serre's conjecture, should reduce to understanding the components of $\overline{\mathcal{X}}$ on which a given representation lies, and understanding what the Serre weights are for maximally non-split upper-triangular representations (that are generic enough to lie on a single component).

We expect that most components are labeled by a single weight, and that in the cases where there are multiple weights labeling a component, they are frequetly related in a simple way. For example, if $K = \mathbb{Q}_p$, and a component has $F(a_1, \dots, a_n)$ as a label, then the generic representations on the component are of the form upper triangular with $\chi_i$ on the diagonal, where $\chi_i \mid_{I_{\mathbb{Q}_p}} = \omega^{a_i+n-i}$.

Furthermore, if non of the $a_i - a_{i+1}$ are 0 or $p-1$, we expect there to be a unique component labeled by this weight, and this component should be labeled only by $F(a_1, \ldots, a_n)$.

*Remark* 7.1. For $n = 3$ and $K = \mathbb{Q}_p$, work of Bao, Levin, Le and Morra shows that if $a$ is in the upper alcove and is suitably generic, then the two JH factors $F_a, F_b$ of $L_\lambda \otimes \overline{\mathbb{F}}_p$ correspond to two components of $\overline{\mathcal{X}}$, labeled by a single weight $F_a$ (resp. $F_b$) which meet in a codimension one substack. Thus the generic $\overline{\rho}$ on the component labeled by $F_b$ (with $b$ in the lower alcove) do not satisfy the crystalline lift conjecture, as it is not labeled by $F_a$(should be a label of each of the ones having some JH factor of $L_\lambda \otimes \overline{\mathbb{F}}_p$ among their labels).

## 8. Current Work

By considering the Langlands correspondence, given our inertial type, one could argue that there should be a connection between $W(\overline{\rho})$ and the set of the Jordan-Holder factor of the corresponding representation $\overline{V(\rho\,|_{I_p})}$, where $\rho$ is any lift, where $V$ is the association of Deligne-Lusztig. In particular, if $\rho\,|_{I_p}$ is trivial, $V(\rho\,|_{I_p}) = Ind_B^G 1$. Then we would be interested in decomposing it.