The weight in Serre's conjectures on modular forms.

Bas Edixhoven*

December 24, 1998

Contents

1 Introduction.

Let p be a prime number and let f be a modular form of level N with $p\nmid N$, weight k and character ε, with coefficients in $\overline{\mathbf{F}}_p$. Suppose that f is an eigenform for all Hecke operators T_l^*, l prime, say with eigenvalues $a_l\in\overline{\mathbf{F}}_p$. Then there exists a unique 2-dimensional semi-simple continuous representation $\rho_f\colon G_{\mathbf{Q}}=\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})\to\mathrm{GL}_2(\overline{\mathbf{F}}_p)$, which is unramified outside pN and has the property that $\mathrm{trace}(\rho_f(\mathrm{Frob}_l))=a_l$ and $\det(\rho_f(\mathrm{Frob}_l))=\varepsilon(l)l^{k-1}$ for all $l\nmid pN$. For a historical account of this result see [24], §6, see also [3]. It follows from the identities $\det(\rho_f(\mathrm{Frob}_l))=\varepsilon(l)l^{k-1}$ that ρ_f is odd, i.e., $\det(\rho_f(c))=-1$ for $c\in G_{\mathbf{Q}}$ a complex conjugation.

Serre conjectured in 1973 that every continuous semi-simple odd representation $\rho\colon G_{\mathbf{Q}}\to\mathrm{GL}_2(\overline{\mathbf{F}}_p)$ arises in this way. In his article [26] he stated a more precise version of this conjecture

*Supported by the Netherlands Organisation for Scientific Research (N.W.O.)

for irreducible ρ: ρ should arise from a form of level $N(\rho)$, weight k_ρ and character $\varepsilon(\rho)$, where $N(\rho)$, k_ρ and $\varepsilon(\rho)$ are described in terms of ρ; $N(\rho)$ and k_ρ are meant to be as small as possible. The aim of this paper is to show that *if* ρ comes from a modular form at all, say of some type (N, k, ε), then ρ also comes from a modular form of type (N, k_ρ, ε) and k_ρ is (almost) minimal. For a precise statement see Thm. 4.5. For results on $N(\rho)$ we refer to [22] and [11].

Let us now briefly describe the proof of Thm. 4.5. Suppose that $\rho \cong \rho_g$, with g of type (N, k, ε). According to a well known result there exists an eigenform f_1 of some type (N, k_1, ε) with $k_1 \leq p+1$ such that $\rho_g \cong \rho_{f_1} \otimes \chi^a$ for some $a \in \mathbf{Z}/(p-1)\mathbf{Z}$; here χ is the p-cyclotomic character (see Thm. 3.4). Let G_p be a decomposition group of $G_{\mathbf{Q}}$ at p. Since $k_1 \leq p+1$, a lot is known about the local representation $\rho_{f_1}|_{G_p}$, see Thm. 2.5–2.9. One finds k_1 and a in terms of $\rho|_{G_p}$. The desired form f of weight k_ρ is then found by "untwisting" the form f_1, i.e., one applies the operator θ (see §3) a times and then divides as many times as possible by the Hasse-invariant.

The main complication in the proof is that k_1 is not necessarily unique. This can arise in two ways: there might exist a "companion form" f_1' of weight $p+1-k_1$, or f_1 might be a multiple of the Hasse-invariant (this can only happen if $k_1 = p$ or $k_1 = p+1$). For $k_1 = p$ these two ways are essentially the same. What one needs is a criterion in terms of ρ whether or not this non-uniqueness occurs. In both cases Serre conjectured such a criterion. The one concerning the "companion form" was proved by Gross in [10] (with some exception, however, see 2.9). For $k = p+1$ the problem, which has to do with "finiteness at p", has been settled by Mazur, see Thm. 2.8. It should be clear to the reader that Serre must have known the proof of Thm. 4.5, modulo Thm.'s 2.8 and 2.9, in order to give the recipe for computing k_ρ in [26].

As explained above, our results depend on Gross' article [10], in which several isomorphisms between different p-adic cohomology theories have not been checked to be compatible with Hecke operators (cf. [10], Introduction). R. Coleman has checked that the map η in (16.6) of [10] commutes with the action of Hecke operators; this result can be extracted from [4]. The remaining compatibilities (pages 504 and 505 of [10]) have not been checked.

Let us now make some remarks on forms of weight 1. In [26], Serre considers modular forms mod p as reductions of forms in characteristic 0. Since forms mod p of weight 1 cannot necessarily be lifted to characteristic 0 they had to be avoided. Later, for another reason ([27]), Serre replaced his definition of mod p modular forms by Katz's (we will use Katz's definition, see §2). Now one can allow forms of weight 1 without problem. As a consequence we have to change the definition of k_ρ if we want the conjecture to specify the minimal possible weight $k(\rho)$ of a form f with $\rho \cong \rho_f$ (see Def. 4.3 and Rem. 4.4). Serre's conjectures then suggest that $k(\rho)$ should be 1 precisely when ρ is unramified at p (in [26] k_ρ is defined to be p if ρ is unramified at p). If f is an eigenform of type (N, p, ε), with $T_p^* f = a_p f$, such that ρ is unramified at p and $a_p^2 \neq \varepsilon(p)$, then Gross's Thm. 2.9 shows that ρ_f comes from weight 1. It is not known if this remains true without the assumption $a_p^2 \neq \varepsilon(p)$.

Let us end with some technicalities. The reader will note that most formulas concerning the modular curves $X_1(N)$ differ from those in [10]. The reason for this is that we work with points of exact order N in the terminology of [16] whereas Gross works with embeddings of μ_N.

We apply the "Drinfeldian" moduli problems of [16] to the generalized elliptic curves of [8]; in the cases we consider this gives algebraic stacks. We denote the Hecke operators U_l^* for $l|N$ by T_l^*.

Finally, it is a pleasure to thank R.F. Coleman, B.H. Gross, J.-P. Serre, D.L. Ulmer and especially K.A. Ribet for conversations on this subject and their comments on this article.

After this article was written R.F. Coleman and J.F. Voloch have obtained new results concerning companion forms [5]. Their main result, that does not depend on unproven compatibilities, states that a cuspidal eigenform f over $\overline{\mathbf{F}}_p$ of type (N,k,ε) with $T_p^*f \neq 0$ and $2 < k \leq p$ has a companion form if and only if ρ_f is tamely ramified at p (compare Thm. 2.9). The consequences for this article are:

1. For $p > 2$ our results do not depend anymore on the unproven compatibilities mentioned above.

2. For $p > 2$ and f a weight 1 cuspidal eigenform over $\overline{\mathbf{F}}_p$ the representation ρ_f is unramified at p (compare Prop. 2.7).

3. All statements in Thm. 4.5 are true for $p > 2$ without the condition "ρ is not exceptional".

2 Modular forms mod p and Galois representations.

2.1 Modular forms. Let p be a prime number and let $N \geq 1$ be prime to p, let k be an integer, and let $\varepsilon : (\mathbf{Z}/N\mathbf{Z})^* \to \overline{\mathbf{F}}_p^{\,*}$ be a character. By a modular form of type (N,k) over $\overline{\mathbf{F}}_p$ we will mean a section of the invertible sheaf $\underline{\omega}^{\otimes k}$ over the algebraic stack $[\Gamma_1(N)]_{\overline{\mathbf{F}}_p}$ (denoted $\mathcal{M}_{\Gamma_{00}(N)} \otimes_{\mathbf{Z}} \overline{\mathbf{F}}_p$ in [8]). The objects of this algebraic stack are pairs $(E/S/\overline{\mathbf{F}}_p, \alpha)$, where E is a generalized elliptic curve over S and $\alpha : (\mathbf{Z}/N\mathbf{Z})_S \hookrightarrow E$ is an embedding of group schemes whose image meets all irreducible components of all geometric fibres. The invertible sheaf $\underline{\omega}$ associates to each pair $(E/S/\overline{\mathbf{F}}_p, \alpha)$ the invertible $\mathcal{O}_S$-module $\underline{\omega}_{E/S} = 0^*\Omega^1_{E/S}$. In more down-to-earth terms this means that a modular form f of type (N,k) over $\overline{\mathbf{F}}_p$ is a rule, compatible with cartesian squares, that associates to each $(E/S/\overline{\mathbf{F}}_p, \alpha)$ a section $f(E/S,\alpha)$ of $\underline{\omega}^{\otimes k}_{E/S}$. If $n \geq 3$ is prime to p then the stack $[\Gamma_1(N), \Gamma(n)]_{\overline{\mathbf{F}}_p}$ is represented by a non-connected smooth proper curve over $\overline{\mathbf{F}}_p$, say X, on which $\mathrm{GL}_2(\mathbf{Z}/n\mathbf{Z})$ acts. Modular forms of type (N,k) are then $\mathrm{GL}_2(\mathbf{Z}/n\mathbf{Z})$-invariant sections of $\underline{\omega}^{\otimes k}$ over X. The proof of this goes as follows. It is clear that a form f of type (N,k) gives an invariant section on X. Suppose now that we have an invariant section on X. Let $(E/S/\overline{\mathbf{F}}_p, \alpha)$ be an object of $[\Gamma_1(N)]_{\overline{\mathbf{F}}_p}$ such that the induced morphism $S \to [\Gamma_1(N)]_{\overline{\mathbf{F}}_p}$ is etale. Let $S' \to S$ be the universal S-scheme over which E/S acquires a $\Gamma(n)$-structure. Then $S' \to S$ is finite and flat and it is the quotient morphism for the action of $\mathrm{GL}_2(\mathbf{Z}/n\mathbf{Z})$. Let E' be the pullback of E to S'. Then $\underline{\omega}_{E'/S'}$ is the pullback to S' of $\underline{\omega}_{E/S}$. Since $S' \to S$ is the quotient for the action of $\mathrm{GL}_2(\mathbf{Z}/n\mathbf{Z})$ and $\underline{\omega}_{E/S}$ is a locally free $\mathcal{O}_S$-module, invariant sections descend.

One obtains the q-expansions of f at the various cusps of $[\Gamma_1(N)]_{\overline{\mathbf{F}}_p}$ by evaluating f on pairs $(\mathrm{Tate}(q^d), \alpha)$, where $\mathrm{Tate}(q^d)$ is the Tate curve $\mathbf{G}_m/q^{d\mathbf{Z}}$ over $\overline{\mathbf{F}}_p[[q]](q^{-1})$ and $d|N$. Explicitly: the q-expansion $f_{d,\alpha}(q)$ of f at the cusp $(\mathrm{Tate}(q^d),\alpha)$ is the power series $f(\mathrm{Tate}(q^d),\alpha)/(dt/t)^{\otimes k}$

in $\overline{\mathbf{F}}_p[[q]]$.

If $N > 4$ then the stack $[\Gamma_1(N)]_{\overline{\mathbf{F}}_p}$ is represented by the curve $X_1(N)_{\overline{\mathbf{F}}_p}$ and modular forms of type (N,k) are just global sections of the invertible sheaf $\underline{\omega}^{\otimes k}$ on $X_1(N)_{\overline{\mathbf{F}}_p}$.

If N is 3 or 4 then there is up to isomorphism exactly one object in $[\Gamma_1(N)]_{\overline{\mathbf{F}}_p}(\overline{\mathbf{F}}_p)$ with non-trivial automorphisms. For $N = 4$ this object is the generalized elliptic curve with two components equipped with a point of order 4; its automorphism group has order 2. For $N = 3$ this object is the elliptic curve with j-invariant 0 equipped with a point of order 3 that is fixed by the automorphisms of order 3 of the curve. In these cases modular forms of type (N,k) can be identified with sections of $\underline{\omega}^{\otimes k}$ over the complement of these points in the curve $X_1(N)_{\overline{\mathbf{F}}_p}$ that extend after suitable ramified coverings (compare [17], p.73–74).

If N equals 1 or 2 then every pair $(E/S,\alpha)$ has the automorphism -1, which implies that $(-1)^k f = f$. It follows that in these cases $f = 0$ unless k is even or $p = 2$. In this case too modular forms can be identified with sections of $\underline{\omega}^{\otimes k}$ over the complement of the points with extra automorphisms in the curve $X_1(N)_{\overline{\mathbf{F}}_p}$ that extend after suitable ramified coverings.

From the fact that there exists a modular form Δ of type $(1,12)$, with a simple zero at infinity, it follows that there are no non-zero forms with $k < 0$. We let A denote the Hasse invariant ([16], §12.4); it is the modular form of type $(1,p-1)$ which has q-expansion 1 at all cusps.

Let $M(N,k)$ denote the $\overline{\mathbf{F}}_p$-vector space of modular forms over $\overline{\mathbf{F}}_p$ of type (N,k) and let $M^0(N,k)$ be its subspace of cusp forms, i.e. the space of forms that have value zero on degenerate curves.

The group $(\mathbf{Z}/N\mathbf{Z})^*$ acts on $[\Gamma_1(N)]_{\overline{\mathbf{F}}_p}$ by:

$$\langle a\rangle : \ (E/S,\alpha) \mapsto (E/S, a\alpha),$$

for $a \in (\mathbf{Z}/N\mathbf{Z})^*$. This gives an action by $(\mathbf{Z}/N\mathbf{Z})^*$ on modular forms:

$$(\langle a\rangle^* f)(E/S,\alpha) = f(E/S, a\alpha).$$

A modular form f of type (N,k) is said to be of type (N,k,ε) if $\langle a\rangle^* f = \varepsilon(a) f$ for all $a \in (\mathbf{Z}/N\mathbf{Z})^*$. Of course, f can be non-zero only if $\varepsilon(-1) = (-1)^k$.

Let R be any commutative ring in which N is invertible. By replacing "$\overline{\mathbf{F}}_p$" above by "R" one gets the definition of a modular form over R. For more details concerning modular forms in this setting see [14], Chapter 1 or [8], VII, §3.

2.2 Hecke operators. The $\overline{\mathbf{F}}_p$-vector spaces of modular forms of type (N,k,ε) are equipped with certain endomorphisms, called Hecke operators. For each prime number l we have an operator T_l^*. For the construction of these operators we refer to [10]. Let us just mention that for $l \not| pN$ and $(E/\overline{\mathbf{F}}_p,\alpha)$, with E a genuine (i.e., not generalized) elliptic curve, we have:

$$(T_l^* f)(E/\overline{\mathbf{F}}_p,\alpha) = \frac{1}{l}\sum_{\phi} \phi^*(f(E'/\overline{\mathbf{F}}_p, \phi\circ\alpha)), \tag{2.2.1}$$

where the sum is over the $l+1$ isogenies $\phi : E \to E'$ of degree l and source E. For $l|N$ the operator T_l^* is given by the same formula, but with the sum restricted to those ϕ such that $\phi\circ\alpha$

is injective. The construction of T_p^*, especially for $k=1$, is somewhat harder. The operators T_l^* all commute with each other; the common eigenspaces are one dimensional since the eigenvalues determine the q-expansions up to a constant factor.

2.3 Galois representations. Let f be of type (N,k,ε) and suppose that f is a common eigenvector for all the T_l^* ($l \not| pN$) with eigenvalues $a_l \in \overline{\mathbf{F}}_p$. Let $G_{\mathbf{Q}} = \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ denote the absolute Galois group of $\mathbf{Q}$. According to a theorem of Deligne ([9], Theorem 6.7), there exists a unique (up to isomorphism) continuous semi-simple representation $\rho_f : G_{\mathbf{Q}} \to \mathrm{GL}(V)$, with V a 2-dimensional $\overline{\mathbf{F}}_p$-vector space, which is unramified outside pN and has the property that $\mathrm{trace}(\rho_f(\mathrm{Frob}_l)) = a_l$ and $\det(\rho_f(\mathrm{Frob}_l)) = \varepsilon(l)l^{k-1}$ for all $l \not| pN$. In this last equality $\rho_f(\mathrm{Frob}_l)$ denotes a Frobenius element at l (unique up to conjugation). If f is not a cusp form then ρ_f is reducible.

2.4 The local representation at p. Let $\overline{\mathbf{Z}}$ be the integral closure of $\mathbf{Z}$ in $\overline{\mathbf{Q}}$. Let us choose a surjection $\overline{\mathbf{Z}} \to \overline{\mathbf{F}}_p$ (these are permuted transitively by $G_{\mathbf{Q}}$). Then we have a sequence of subgroups:

$$I_p \lhd I \lhd G_p \subset G_{\mathbf{Q}},$$

where G_p is the decomposition subgroup, I the ramification subgoup, and I_p the wild ramification subgroup. We can identify G_p with $\mathrm{Gal}(\overline{\mathbf{Q}}_p/\mathbf{Q}_p)$, G_p/I with $\mathrm{Gal}(\overline{\mathbf{F}}_p/\mathbf{F}_p)$, and $I_t := I/I_p$ (the tame ramification group) with $\mathrm{Gal}(\mathbf{Q}_p^{\mathrm{tr}}/\mathbf{Q}_p^{\mathrm{unr}})$ and with $\varprojlim \mathbf{F}_{p^n}^*$. A character $\phi : I_t \to \overline{\mathbf{F}}_p^{\,*}$ is said to be of level n if n is the smallest integer such that ϕ factors through $\mathbf{F}_{p^n}^*$. The n characters $I_t \to \mathbf{F}_{p^n}^* \to \overline{\mathbf{F}}_p^{\,*}$ that are induced by embeddings of fields $\mathbf{F}_{p^n} \to \overline{\mathbf{F}}_p$ are called the fundamental characters of level n. We let $\chi : G_{\mathbf{Q}} \to \overline{\mathbf{F}}_p^{\,*}$ denote the cyclotomic character giving the action on the p^{th} roots of unity: $\sigma(\zeta) = \zeta^{\chi(\sigma)}$.

Let $\rho_p : G_p \to \mathrm{GL}(V)$ be any continuous 2-dimensional representation on a $\overline{\mathbf{F}}_p$-vector space. 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} and the action of I_t on V^{ss} is given by two characters ϕ, $\phi' : I_t \to \overline{\mathbf{F}}_p^{\,*}$. Since $\mathrm{Gal}(\overline{\mathbf{F}}_p/\mathbf{F}_p)$ acts by conjugation on I_t it follows that $\{\phi^p, \phi'^p\} = \{\phi, \phi'\}$. This means that there are two cases:

2.4.1 ϕ and ϕ' are both of level 1, ρ_p is reducible.

2.4.2 ϕ and ϕ' are both of level 2, $\phi^p = \phi'$, $\phi'^p = \phi$, and ρ_p is irreducible.

Now suppose that f, ρ_f and V are as in (2.3). Then we have the following results on $\rho_{f,p} := \rho_f|_{G_p}$. For any $a \in \overline{\mathbf{F}}_p^{\,*}$ let $\lambda(a)$ denote the unramified character $G_p \to \overline{\mathbf{F}}_p^{\,*}$ such that $\lambda(\mathrm{Frob}_p) = a$.

2.5 Theorem. *(Deligne) Let f be a cusp form of type (N,k,ε) with $2 \le k \le p+1$ which is an eigenform for all T_l^* with eigenvalues a_l. Suppose that $a_p \neq 0$. Then $\rho_{f,p}$ is reducible and*

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

2.6 Theorem. *(Fontaine) Let f be a cusp form of type (N,k,ε) with $2 \leq k \leq p+1$ which is an eigenform for all T_l^* with eigenvalues a_l. Suppose that $a_p = 0$. Then $\rho_{f,p}$ is irreducible and*

$$\rho_f|_I = \begin{pmatrix} \psi^{k-1} & 0 \\ 0 & \psi'^{k-1} \end{pmatrix},$$

where ψ and ψ' are the two fundamental characters of level 2.

For the proof of Theorem 2.5 we refer to [10]. Theorem 2.6 was proved by Fontaine in 1979 in two letters to Serre, but that proof has never been published; we will give a proof in §6.

2.7 Proposition. *Let f be a cuspidal eigenform of type $(N,1,\varepsilon)$ with $T_p^* f = a_p f$. Suppose that the polynomial $X^2 - a_p X + \varepsilon(p)$ has two distinct roots in $\overline{\mathbf{F}}_p$, or that f is the reduction of a characteristic zero eigenform. Then ρ_f is unramified at p.*

Proof. Let $f_{d,\alpha}(q) = \sum b_n q^n$ be the q-expansion of f at some cusp $(\mathrm{Tate}(q^d), \alpha)$. Let $V_p^* f$ be the form of type (N,p,ε) with $(V_p^* f)_{d,\alpha}(q) = \sum b_n q^{pn}$. Then Af and $V_p^* f$ are eigenforms for all T_l, $l \neq p$, with the same eigenvalues as f. The action of T_p^* on the two dimensional subspace of $M^0(N,p,\varepsilon)$ spanned by Af and $V_p^* f$ is given by: $T_p^*(Af) = a_p Af - \varepsilon(p) V_p^* f$, $T_p^*(V_p^* f) = Af$, where $T_p^* f = a_p f$ (see formula 4.7 in [10]). It follows that the characteristic polynomial of T_p^*, acting on this two-dimensional vector space, equals $X^2 - a_p X + \varepsilon(p)$; let a and a' be the two roots, and let g and g' be the corresponding two normalized eigenforms. The representations ρ_f, ρ_g and $\rho_{g'}$ are isomorphic, since f, g and g' have the same eigenvalues for all T_l^*, $l \neq p$. According to Theorem 2.5 we have:

$$\rho_{g,p} = \begin{pmatrix} \lambda(a') & * \\ 0 & \lambda(a) \end{pmatrix}, \qquad \rho_{g',p} = \begin{pmatrix} \lambda(a) & * \\ 0 & \lambda(a') \end{pmatrix}.$$

From this we see that ρ_f is unramified at p if $a \neq a'$. One also knows that ρ_f is unramified at p if f is the reduction mod p of a weight 1 form of level N in characteristic zero (see Theorem 4.1 in [9]). It is not known whether ρ_f is unramified or not if $a = a'$ and f is not the reduction of a characteristic zero eigenform. □

2.8 Theorem. *(Mazur) Let f be a cuspidal eigenform of type $(N,p+1,\varepsilon)$. Suppose that ρ_f is irreducible and that there exists no form g of type $(N,2)$ such that $f = Ag$. Then $a_p^2 = \varepsilon(p)$, and the extension in Theorem 2.5 is not "finite at p" ([26], §2.8). I.e., the corresponding group scheme over $\mathbf{Q}_p$ corresponding to a realization of ρ_f over a finite field does not admit a finite extension over $\mathbf{Z}_p$.*

See [22], §6 for a proof in the case where ε is trivial and $p > 2$. That proof generalizes without problems to arbitrary ε. For $p = 2$ one argues as follows. Lemma 6.2 of [22] and its proof only require a slight adaptation: if V is unramified then the result is immediate, so we may assume that $\mathcal{V}^{\mathrm{l}}$ and $\mathcal{V}^{\mathrm{et}}$ are 1-dimensional k-vector space schemes; $\mathcal{V}^{\mathrm{l}}_{\mathbf{Q}_2}$ has finite Zariski closure in $\mathcal{J}$ by the Néronian property (since p=2); this Zariski closure is multiplicative since it is contained in the toric part. As in the proof of Lemma 6.3 of [22] one finds that $\mathcal{V}$ lands in the toric part of $\mathcal{J}$, but this is a contradiction with the 1-dimensionality of U^{t} in the proof of Prop. 18 in [18].

In §8 we will give a proof that the notions "finite at p" and "peu ramifié" ([26], 2.4.7) are equivalent for representations ρ_f where f is a cuspidal eigenform of type $(N, p+1, \varepsilon)$. Let us also remark that it follows from the construction of ρ_f that ρ_f is finite at p if f has weight 2.

The following theorem, which gives a criterion for $\rho_{f,p}$ to be tamely ramified, was conjectured by Serre and recently proved by Gross ([10], Theorem 13.10).

2.9 Theorem. *(Gross) Let f be as in Theorem 2.5, and suppose that $a_p^2 \neq \varepsilon(p)$ if $k = p$. Then $\rho_{f,p}$ is tamely ramified (i.e., its restriction to the inertia group is diagonalizable) if and only if there exists a cuspidal eigenform f' of type (N, k', ε), with $k' = p+1-k$, such that $la'_l = l^{k'} a_l$ for all l.*

The form f' is called a "companion form" of f. The hard part of the proof of the theorem is to show the existence of a companion f' if $\rho_{f,p}$ is tamely ramified. It seems not to be known whether a similar statement is still true if $k = p$ and $a_p^2 = \varepsilon(p)$.

3 Tate's θ-cycles.

3.1 The operator θ on modular forms mod p. Let $M(N,k)$ denote the $\overline{\mathbf{F}}_p$-vector space of modular forms over $\overline{\mathbf{F}}_p$ of type (N,k) and let $M(N) = \oplus_{k \geq 0} M(N,k)$. Then $M(N)$ is a graded $\overline{\mathbf{F}}_p$-algebra. Each $M(N,k)$ has a descending filtration consisting of the subspaces $A^i M(N, k - i(p-1))$, $i \geq 0$, where A is the Hasse-invariant. The *filtration* $w(f)$ of $f \in M(N,k)$ is defined by:

$$w(f) = \min\{k - i(p-1) \mid f \in A^i M(N, k - i(p-1))\}. \tag{3.1.1}$$

Equivalently, $w(f)$ is the smallest integer k' for which there exists a form in $M(N,k')$ which, at some cusp, has the same q-expansion that f does. This last equivalence follows from the fact that for any cusp (d,α) the kernel of the q-expansion map $M(N) \to \overline{\mathbf{F}}_p[[q]]$ is the ideal generated by $A-1$ (see [15], §1). Note that if $f \in M(N,k)$ and $f' \in M(N,k')$ have the same q-expansion at some cusp then $k' \equiv k \pmod{p-1}$. Taking into account our desription of modular forms in §2.1 Katz proves in [15] that there exists a derivation

$$\theta : \ M(N) \to M(N)$$

that increases degrees by $p+1$ and whose effect on q-expansions (at all cusps) is qd/dq: if $f_{d,\alpha}(q) = \sum a_n q^n$, then $(\theta f)_{d,\alpha}(q) = \sum n a_n q^n$. If $f \in M(N,k)$ has filtration k, and $p \nmid k$, then θf has filtration $k+p+1$. If $f \in M(N,pk)$ and $\theta f = 0$, then $f = g^p$ for a unique $g \in M(N,k)$. The commutation relations of θ with the Hecke operators are:

$$T_l^*(\theta f) = l\,\theta(T_l^* f).$$

In particular, $T_p^*(\theta f) = 0$. Note that θ preserves the subspace of cusp forms.

If f is an eigenform of type (N,k,ε) with eigenvalues a_l, then θf is an eigenform of type $(N, k{+}p{+}1, \varepsilon)$ with eigenvalues la_l. This implies that

$$\rho_{\theta f} = \rho_f \otimes \chi.$$

3.2 θ-cycles. Let f be a homogeneous element in $M(N)$ such that $\theta(f) \neq 0$. The θ-cycle of f is by definition the sequence of $p{-}1$ integers

$$(w(\theta f), \ldots, w(\theta^{p-1} f)).$$

Note that always $w(\theta^p f) = w(\theta f)$, but $w(\theta^{p-1} f) = w(f)$ only if the q-expansion of f at some cusp has $a_{pn} = 0$ for all n. It is straightforward but surprising that up to a cyclic permutation a θ-cycle is always of one of the forms

$$\begin{array}{ll} (k, k+p+1, .., k+(p-2)(p+1)) & k \equiv 2\ (p) \\ (k, k+p+1, .., k+(p-k_0)(p+1), k_1, k_1+p+1, .., k_1+(k_0-3)(p+1)) & k \equiv k_0\ (p), \end{array}$$

where $3 \le k_0 \le (p+3)/2$ and $k_1 = k+p+3-2k_0$, at least for $p>3$. A proof of this fact for level 1 can be found in [13], §7. For $p=2$ the only possibility is (k) with $k \equiv 0\ (2)$. For $p=3$ the possibilities are (again up to cyclic permutation): $(k, k+4)$ with $k \equiv 2\ (3)$ and (k,k) with $k \equiv 0\ (3)$.

3.3 Proposition. *Let f be a cuspidal eigenform of type (N,k,ε) with $1 \le k \le p{+}1$, with eigenvalues a_l and with $w(f) = k$. Then the θ-cycle of f is given as follows:*

$$a_p = 0 \left\{ \begin{array}{ll} k=1 & (p+2, .., p+2+(p-2)(p+1)) \\ k=2 & (2+p+1, .., 2+(p-2)(p+1), 2) \\ 3 \le k \le p-1 & (k+p+1, .., k+(p-k)(p+1), k_1, .., k_1+(k-3)(p+1), k) \\ k=p & (3, .., 3+(p-3)(p+1), p) \\ k=p+1 & \textit{does not occur,} \end{array} \right.$$

$$a_p \neq 0 \left\{ \begin{array}{ll} k=1 & (p+2, .., p+2+(p-2)(p+1)) \\ 2 \le k \le p-1 & (k+p+1, .., k+(p-k)(p+1), k'+p+1, .., k'+(k-1)(p+1)) \\ k=p & (p+2, .., p+2+(p-2)(p+1)) \\ k=p+1 & (2p+2, .., 2p+2+(p-2)(p+1)), \end{array} \right.$$

where $k_1 = p{+}3{-}k$ and $k' = p{+}1{-}k$. The correct interpretation of these formulas for $p=2$ is:

$$a_2 = 0 \left\{ \begin{array}{ll} k=1 & (4) \\ k=2 & (2) \\ k=3 & \textit{does not occur,} \end{array} \right. \qquad a_2 \neq 0 \left\{ \begin{array}{ll} k=1 & (4) \\ k=2 & (4) \\ k=3 & (6) \end{array} \right.$$

and for $p=3$:

$$a_3 = 0 \left\{ \begin{array}{ll} k=1 & (5,9) \\ k=2 & (6,2) \\ k=3 & (3,3) \\ k=4 & \textit{does not occur,} \end{array} \right. \qquad a_3 \neq 0 \left\{ \begin{array}{ll} k=1 & (5,9) \\ k=2 & (6,6) \\ k=3 & (5,9) \\ k=4 & (8,12) \end{array} \right.$$

Proof. This follows from the classification above, except for the case $a_p \neq 0$ and $k = p$. In that case one has to show that $w(\theta f) \neq 3$. But if $w(\theta f) = 3$ then $w(\theta^{p-1} f) = p = w(f)$, $f - \theta^{p-1} f = V_p g$ for some non-zero g of weight 1, f is a linear combination of Ag and $V_p g$, $w(\theta f) = p+2$, contradiction. See also [13], §7. □

The reader will probably have noticed by now that we have been gathering results concerning forms of weight at most $p+1$. Fortunately, there is the following result.

3.4 Theorem. *Let f be an eigenform of some type (N, k, ε), then there exist integers i and k' with $0 \leq i \leq p-1$, $k' \leq p+1$ and an eigenform g of type (N, k', ε) such that f and $\theta^i g$ have the same eigenvalues for all T_l^* $(l \neq p)$.*

The fact that eigenforms of weight at most $p+1$ give, up to twist, all systems of eigenvalues, has been known to be true for a long time. A possible proof, which has never been published, uses the construction of ρ_f in terms of the étale cohomology of the locally constant sheaf $\mathrm{Sym}^{k-2}\mathcal{F}$ on the curve $X = X_1(N)_{\mathbf{Q}}-$cusps, where $\mathcal{F}$ is the p-torsion of the universal elliptic curve. Locally constant sheaves on X correspond to representations of the fundamental group of X (for some chosen base point). The mod p representation theory of $\mathrm{GL}_2(\mathbf{F}_p)$ then shows that $\mathrm{Sym}^{k-2}\mathcal{F}$ has a Jordan-Hölder sequence with successive quotients of the form $(\mathrm{Sym}^i\mathcal{F}) \otimes (\det \mathcal{F})^{\otimes j}$ with $0 \leq i \leq p-1$ and $0 \leq j \leq p-2$. For example, let $V = \mathbf{F}_p \oplus \mathbf{F}_p$ be the standard representation of $\mathrm{GL}_2(\mathbf{F}_p)$. Then $\mathrm{Sym}^p(V)$ has a subrepresentation $V \hookrightarrow \mathrm{Sym}^p(V)$ by sending v to v^p. The quotient is isomorphic to $\mathrm{Sym}^{p-2}(V) \otimes \det$. This induces a short exact sequence

$$0 \longrightarrow \mathcal{F} \longrightarrow \mathrm{Sym}^p(\mathcal{F}) \longrightarrow \mathrm{Sym}^{p-2}(\mathcal{F}) \otimes \det(\mathcal{F}) \longrightarrow 0.$$

The first published proof of Thm. 3.4, for $p \geq 5$, is due to Ash and Stevens in [1], Thm. 3.4, 3.5. We will give another, modular form theoretic proof for arbitrary p in §7.

4 The weight in Serre's conjectures.

4.1 Let $\rho : G_{\mathbf{Q}} \to \mathrm{GL}_2(\overline{\mathbf{F}}_p)$ be a continuous representation of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$. Suppose that ρ is irreducible and that $\det(\rho)$ is *odd*: $\det(\rho(c)) = -1$, where $c \in G_{\mathbf{Q}}$ is the complex conjugation induced by an embedding of $\overline{\mathbf{Q}}$ into $\mathbf{C}$. In [26], §1–2 Serre attaches a triple $(N(\rho), k_\rho, \varepsilon(\rho))$ to ρ, where $N(\rho)$ is a positive integer which is prime to p; k_ρ is a positive integer and $\varepsilon(\rho) : (\mathbf{Z}/N(\rho)\mathbf{Z})^* \to \overline{\mathbf{F}}_p^*$ a character. By definition, $N(\rho)$ is the prime to p part of the Artin conductor of ρ (defined as in characteristic zero). To define $k_\rho \bmod p-1$ and $\varepsilon(\rho)$ one considers the character $\det(\rho) : G_{\mathbf{Q}} \to \overline{\mathbf{F}}_p^*$. Since the conductor of $\det(\rho)$ divides $pN(\rho)$ one can identify $\det(\rho)$ with a homomorphism $(\mathbf{Z}/pN(\rho)\mathbf{Z})^* \to \overline{\mathbf{F}}_p^*$, or equivalently, with a couple of homomorphisms $\varepsilon(\rho) : (\mathbf{Z}/N(\rho)\mathbf{Z})^* \to \overline{\mathbf{F}}_p^*$ and $\chi^{k_\rho - 1} : (\mathbf{Z}/p\mathbf{Z})^* \to \overline{\mathbf{F}}_p^*$, where χ denotes the cyclotomic character as in §2.4. For the complete definition of k_ρ, which depends only on $\rho|_I$, we refer to [26], §2.

4.2 Conjecture. *(Serre [26], 3.2.4) Let ρ be as above. Then there exists a cusp form f of type $(N(\rho), k_\rho, \varepsilon(\rho))$, which is an eigenform for all T_l^*, such that ρ is isomorphic to ρ_f.*

The mod p modular form f in this conjecture was originally meant to be the reduction mod p of some modular form in characteristic zero. Because of this definition of modular form mod p, Serre avoided the possibility $k_\rho = 1$. Later, in [27], Serre replaced this definition of modular form mod p by the definition that we are using in this article (i.e., Katz's definition). This change in definition makes it natural to allow forms of weight 1. As a consequence we have to change the definition of k_ρ if we want the conjecture to specify the minimal possible weight of a form f with $\rho \cong \rho_f$. We propose the following definition.

4.3 Definition. Let $\rho_p\colon G_p \to \mathrm{GL}(V)$, V^{ss}, ϕ and ϕ' be as in §2.4. We associate an integer $k(\rho_p)$ to ρ_p as follows.

1. Suppose that ϕ and ϕ' are of level 2. We have:

$$\rho_p|_I \cong \begin{pmatrix} \phi & 0 \\ 0 & \phi' \end{pmatrix}$$

   After interchanging ϕ and ϕ' if necessary, we have (uniquely) $\phi = \psi^{a+pb} = \psi^a\psi'^b$ and $\phi' = \psi'^a\psi^b$ with $0 \leq a < b \leq p{-}1$. We set $k(\rho_p) = 1 + pa + b$.

2. Suppose that ϕ and ϕ' are of level 1.

   (a) If $\rho_p|_{I_p}$ is trivial then we have:

$$\rho_p|_I \cong \begin{pmatrix} \chi^a & 0 \\ 0 & \chi^b \end{pmatrix}$$

   with $0 \leq a \leq b \leq p{-}2$. We set $k(\rho_p) = 1 + pa + b$.

   (b) Suppose that $\rho_p|_{I_p}$ is not trivial. We have:

$$\rho_p|_I \cong \begin{pmatrix} \chi^\beta & * \\ 0 & \chi^\alpha \end{pmatrix}$$

   for unique α and β with $0 \leq \alpha \leq p{-}2$ and $1 \leq \beta \leq p{-}1$. We set $a = \min(\alpha, \beta)$, $b = \max(\alpha, \beta)$. If $\chi^{\beta-\alpha} = \chi$ and $\rho_p \otimes \chi^{-\alpha}$ is not finite at p (see §8.1) then we set $k(\rho_p) = 1 + pa + b + p - 1$; otherwise we set $k(\rho_p) = 1 + pa + b$.

4.4 Remark. 1. By comparing the definitions of k_ρ and $k(\rho)$ one sees that always $k(\rho) \leq k_\rho$ and that they differ only in two cases. In both cases ϕ and ϕ' are of level 1. In the first case $\rho_p|_{I_p}$ is trivial and $a = 0 = b$: then $k(\rho) = 1$ and $k_\rho = p$. In the second case $p = 2$, $\rho_p|_{I_p}$ is non-trivial, $\alpha = 0$, $\beta = 1$ and ρ_p is not finite at p: then $k(\rho) = 3$ and $k_\rho = 4$.

2. We call ρ *exceptional* if ρ_p is isomorphic to an extension (either split or non-split) of two copies of an unramified character ε of G_p:

$$\rho_p \cong \begin{pmatrix} \varepsilon & * \\ 0 & \varepsilon \end{pmatrix}$$

4.5 Theorem. *Let $\rho : G_{\mathbf{Q}} \to \mathrm{GL}_2(\overline{\mathbf{F}}_p)$ be continuous, irreducible and odd. Let $k(\rho)$ and k_ρ be defined as in 4.3 and 4.4. Suppose that there exists a cusp form g of some type (N, k, ε) with $p \nmid N$, which is an eigenform for all T_l^*, such that ρ is isomorphic to ρ_g. Then there exists a cuspidal eigenform f of type (N, k_ρ, ε) which has the same eigenvalues for T_l^* $(l \neq p)$ as g has, such that ρ is isomorphic to ρ_f. If ρ is not exceptional (see 4.4) then there exists an eigenform f of type $(N, k(\rho), \varepsilon)$ with the same eigenvalues for T_l^* $(l \neq p)$ as g has, such that ρ is isomorphic to ρ_f. If ρ is not exceptional then there is no eigenform of level prime to p and of weight less than $k(\rho)$ whose associated Galois representation is isomorphic to ρ.*

Proof. First we will show the existence of f. The proof will distinguish between the various possible "types" for ρ_p. As in §2.4 let ϕ, ϕ' denote the two characters: $I_t \to \overline{\mathbf{F}}_p^{\,*}$ associated to ρ_p.

Suppose that ϕ and ϕ' are of level 2. Then $\phi \neq \phi^p = \phi'$; ρ_p is irreducible and

$$\rho|_I = \begin{pmatrix} \phi & 0 \\ 0 & \phi' \end{pmatrix}.$$

Let ψ, ψ' be the two fundamental characters of level 2. After permuting ϕ and ϕ' if necessary we can write uniquely $\phi = \psi^a \psi'^b$ with $0 \leq a < b \leq p{-}1$. According to Theorem 3.4 there exists at least one $\alpha \in \mathbf{Z}/(p-1)\mathbf{Z}$ such that $\rho \otimes \chi^{-\alpha}$ is isomorphic to the Galois representation corresponding to an eigenform f_1 of type (N, k_1, ε) with $k_1 \leq p{+}1$. We may suppose that $k_1 = w(f_1)$ and that f_1 has eigenvalues $l^{-\alpha} a_l$ $(l \neq p)$, where a_l are the eigenvalues of g. By Thm.'s 2.5, 2.6, 2.8 and Prop. 2.7 we have $k_1 \geq 2$; $T_p^* f_1 = 0$; $k_1 \leq p$. Since $\chi = \psi\psi'$ it follows from Theorem 2.6 that the only possible values for (α, k_1) are:

$$\begin{cases} (a, 1+b-a),\ (b-1, p+2+a-b) & \text{if } b-a \neq 1 \\ (a, 2) & \text{if } b-a = 1. \end{cases}$$

Because of Prop. 3.3 these possible values all occur. Let f_1 be an eigenform of type $(N, 1{+}b{-}a, \varepsilon)$ with eigenvalues $l^{-a} a_l$ for all T_l^* $(l \neq p)$ (this f_1 is unique up to multiplication by a scalar). Then we can take $f := \theta^a f_1$. We have

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

We can now suppose that ϕ and ϕ' are of level 1. Then we have

$$\rho_p = \begin{pmatrix} \chi^\beta \varepsilon_1 & * \\ 0 & \chi^\alpha \varepsilon_2 \end{pmatrix},$$

where ε_1 and ε_2 are unramified characters $G_p \to \overline{\mathbf{F}}_p^{\,*}$. Let us first suppose that ρ_p is *not* tamely ramified, i.e., "$* \neq 0$". In that case there exists exactly one twist of ρ by a power of χ that comes from an eigenform f_1 of type (N, k_1, ε) with $2 \leq k_1 \leq p{+}1$, namely $\rho \otimes \chi^{-\alpha}$ (see Thm. 2.5 and Prop. 2.7). We may suppose that f_1 has eigenvalues $l^{-\alpha} a_l$ $(l \neq p)$ and that $k_1 = w(f_1)$ (even if $k_1 = p$, see Prop. 2.7). We normalize α and β as Serre does:

$$0 \leq \alpha \leq p-2 \ \text{ and } \ 1 \leq \beta \leq p-1.$$

By Thm. 2.5 we have the congruence $k_1{-}1 \equiv \beta{-}\alpha \pmod{p{-}1}$; together with Thm. 2.8 this determines k_1. Let us deal with the case $p = 2$ first. Then $\alpha = 0$, $\beta = 1$, $k_1 = 2$ if ρ_p is "peu ramifié" and $k_1 = 3$ otherwise. In this case f_1 is the form f of weight $k(\rho)$ we are looking for; clearly the desired form of weight k_ρ exists too. Now suppose $p > 2$. If $\alpha < \beta$ then $k_1 = 1{+}\beta{-}\alpha{+}p{-}1$ if $\beta = \alpha{+}1$ and ρ is "très ramifié" and $k_1 = 1{+}\beta{-}\alpha$ otherwise. It follows from Prop. 3.3 that

$$w(\theta^\alpha f_1) = k_1 + (p+1)\alpha = k(\rho) = k_\rho.$$

If $\alpha \geq \beta$ then $k_1 = \beta{-}\alpha{+}p$ and Prop. 3.3 shows:

$$w(\theta^\alpha f_1) = w(\theta^{\beta-1}\theta^{\alpha-\beta+1} f_1) = 1 + p\beta + \alpha = k(\rho) = k_\rho.$$

Finally we have the case where "$* = 0$". Then we can normalize α and β such that $0 \leq \alpha \leq \beta \leq p{-}2$. According to Thm.'s 3.4, 2.5, 2.6, 2.8, 2.9 andthe proof of Prop. 2.7 $\rho \otimes \chi^{-\alpha}$ comes from a form f_1 of type (N, k_1, ε) with $2 \leq k_1 < p{+}1$. As above we have $k_1 \equiv 1{+}\beta{-}\alpha \pmod{p{-}1}$. If $\alpha < \beta$ then we find the desired form f of type $(N, k(\rho), \varepsilon)$ with $k(\rho) = 1{+}p\alpha{+}\beta = k_\rho$ as above. If $\alpha = \beta \neq 0$ then $k_1 = p$ and $\theta^\alpha f_1$ has weight $k(\rho) = k_\rho$. If $\alpha = \beta = 0$ then, since we are assuming that ρ is not exceptional, the desired form f of weight $k(\rho) = 1$ exists by Thm. 2.9.

Now we prove the minimality statement in the theorem. Suppose that ρ is not exceptional and that g is an eigenform of some type (M, k, ε') with $p \nmid M$ and $\rho \cong \rho_g$. Then the preceding part of the proof gives us a form f of type $(M, k(\rho), \varepsilon')$ with $\rho \cong \rho_f$ such that f has the same eigenvalues a_l for all $l \neq p$ as g has. If both $T_p^* f = 0$ and $T_p^* g = 0$ then the q-expansions of f and g at a fixed cusp differ by a constant factor and we have $k(\rho) = w(f) = w(g) \leq k$. If $T_p^* g \neq 0$ then $k \leq p{+}1$ by Proposition 4.12 in [10] and a "case by case" check shows that $k(\rho) \leq k$. □

5 Group schemes of type $(p, \ldots, p)$ when $e = p{-}1$.

5.1 Raynaud's classification. Let p be a prime number, let $r \geq 1$ and let $\mathbf{F}$ be a field with $q = p^r$ elements. As in [21], a $\mathbf{F}$-vector space scheme over a scheme S is a commutative group scheme G/S given with a homomorphism of $\mathbf{F}$ to the ring of S-endomorphisms of G. Let D be a mixed characteristic strictly henselian discrete valuation ring with residue field k of characteristic p. Let $v : D \to \mathbf{Z}$ be the valuation and let $e = v(p)$ be the absolute ramification index. If S is a D-scheme and G is affine over S we can decompose $\mathcal{O}_G = \mathcal{O}_S \oplus (\oplus_\chi \mathcal{I}_\chi)$ as a direct sum of $\mathcal{O}_S$-modules, where $\chi \colon \mathbf{F}^* \to D^*$ is a character and $\mathcal{I}_\chi$ is the subsheaf of the augmentation ideal $\mathcal{I}$ consisting of sections f such that for all $\lambda \in \mathbf{F}^*$ one has $\lambda^\# f = \chi(\lambda) f$. Here $\lambda^\#$ denotes the pull-back map on functions induced by the multiplication by λ on G (note that $q{-}1$ is invertible in D and that D contains the $(q{-}1)^{\text{th}}$ roots of unity). The $\mathbf{F}$-vector space scheme G/S is then said to satisfy condition $(**)$ ([21], §1.5) if all $q{-}1$ $\mathcal{O}_S$-modules $\mathcal{I}_\chi$ are invertible. Note that this condition is automatically satisfied if G/S is finite locally free of rank q and S is integral with fraction field of characteristic 0 ([21], 1.2.2). Note that if

S is the spectrum of an algebraically closed field of characteristic 0 and G/S is finite, then G/S is a constant $\mathbf{F}$-vector space scheme and $(**)$ means that it has dimension 1. A complete classification up to isomorphism of the $G/S/D$ satisfying $(**)$ is given in [21], Theorem 1.4.1 and [21], §1.5 gives more information for more specific S. The results contained here in sections 5.1, 5.2 and 5.3 can be found in greater detail in [21].

A character $\chi\colon \mathbf{F}^* \to \mu_{q-1}(D) \subset D^*$ is called fundamental if, when we define $\chi(0) = 0$, the composition $\chi\colon \mathbf{F} \to D \to k$ is a homomorphism of rings. The r fundamental characters play a key role in the classification mentioned above. If χ is a fundamental character then the r fundamental characters are the χ^{p^i}, with $i \in \mathbf{Z}/r\mathbf{Z}$. Let us fix a fundamental character χ_0, and let $\chi_i := \chi_0^{p^i}$. The fundamental characters $\chi\colon \mathbf{F}^* \to \mu_{q-1}(D)$ are essentially the inverses of the fundamental characters $I_t \to \overline{\mathbf{F}}_p^{\,*}$ of §2.4.

5.2 The classification over K. Let K be the field of fractions of D. The classification for $S = \mathrm{Spec}(K)$ is as follows. Each $\mathbf{F}$-vector space scheme G over K can be described by r equations in r variables X_i, $i \in \mathbf{Z}/r\mathbf{Z}$

$$X_i^p = \delta_i X_{i+1}, \quad i \in \mathbf{Z}/r\mathbf{Z}, \quad 0 \neq \delta_i \in K. \tag{5.2.1}$$

Here, for all i, X_i is a generator of $\mathcal{I}_{\chi_i}$. Let G'/K be given by equations $X_i'^p = \delta_i' X_{i+1}'$. Any morphism of $\mathbf{F}$-vector space schemes $f\colon G \to G'$ is necessarily of the form $f^\#\colon X_i' \mapsto u_i X_i$, with $u_i \in K$. Such a f is an isomorphism if and only if $u_i \in K^*$ for all i. In that case the δ_i' and the u_i have to satisfy

$$\delta_i' = u_i^p \delta_i u_{i+1}^{-1}, \quad i \in \mathbf{Z}/r\mathbf{Z}. \tag{5.2.2}$$

It follows that the map

$$(\delta_0, \ldots, \delta_{r-1}) \mapsto (v(\delta_0), \ldots, v(\delta_{r-1})) \tag{5.2.3}$$

induces a bijection between the set of isomorphism classes of G/K satisfying $(**)$ and the set $\mathbf{Z}^r/\Lambda$, where Λ is the lattice

$$\Lambda = \mathbf{Z}(p, 0, \ldots, 0, -1) \oplus \mathbf{Z}(-1, p, 0, 0, \ldots, 0) \oplus \cdots \oplus \mathbf{Z}(0, \ldots, 0, -1, p). \tag{5.2.4}$$

The set

$$\{(a_0, \ldots, a_{r-1}) \in \mathbf{Z}^r \mid 0 \leq a_i \leq p-1,\ (a_0, \ldots, a_{r-1}) \neq (p-1, \ldots, p-1)\} \tag{5.2.5}$$

is a set of representatives and hence corresponds bijectively to the set of isomorphism classes of G/K satisfying $(**)$.

Of course, a finite group scheme G/K is determined by the action of $\mathrm{Gal}(\overline{K}/K)$ on $G(\overline{K})$. Let G/K be the $\mathbf{F}$-vector space scheme given by equations 5.2.1 with $a_i := v(\delta_i)$. Then the action of $\mathrm{Gal}(\overline{K}/K)$ on $G(\overline{K})$ is given by a character $\mathrm{Gal}(\overline{K}/K) \to \mathbf{F}^*$. As in 2.4 we identify the tame ramification group I_t of $\mathrm{Gal}(\overline{K}/K)$ with $\varprojlim \mu_n(K)$. The character above is then identified with a character $\phi\colon \mu_{q-1}(K) \to \mathbf{F}^*$. For all $i \in \mathbf{Z}/r\mathbf{Z}$ let $\psi_i\colon \mu_{q-1}(K) \to \mathbf{F}^*$ denote the inverse mapping of $\chi_i\colon \mathbf{F}^* \to \mu_{q-1}(K)$; the ψ_i are the fundamental characters of level r in the sense of 2.4. Note that $\psi_i^p = \psi_{i-1}$.

Elimination of the X_i, $i \neq 0$, from the equations 5.2.1 defining G/K gives the equation

$$X_0^q = \delta X_0, \quad \text{where } \delta = \delta_0^{p^{r-1}} \delta_1^{p^{r-2}} \cdots \delta_{r-1}. \tag{5.2.6}$$

We can now identify $G(\overline{K})$ with the roots of 5.2.6 in $\overline{K}$. Then for all $\lambda \in \mathbf{F}^*$, multiplication by λ on $G(\overline{K})$ corresponds to multiplication by $\chi_0(\lambda)$ on these roots (since $\lambda^\# X_0 = \chi_0(\lambda) X_0$). An element $\zeta \in \mu_{q-1}(K)$, viewed as an element of the Galois group of the degree $q-1$ extension of K, acts on the roots by multiplication by $\zeta^{v(\delta)}$. From $v(\delta) = p^{r-1}a_0 + \cdots + pa_{r-2} + a_{r-1}$ it follows that

$$\phi = \psi_1^{a_0} \psi_2^{a_1} \cdots \psi_r^{a_{r-1}}. \tag{5.2.7}$$

For more details see [21], §3.4.

5.3 The classification over D. Each $\mathbf{F}$-vector space scheme G over D satisfying $(**)$ is given by equations

$$X_i^p = \delta_i X_{i+1}, \quad i \in \mathbf{Z}/r\mathbf{Z}, \quad \delta_i \in D, \quad 0 \le v(\delta_i) \le e. \tag{5.3.1}$$

The reason the conditions "$v(\delta_i) \le e$" show up is that the p/δ_i are needed to define the group structure. If G'/D is given by equations $X_i'^p = \delta_i' X_{i+1}'$ then G and G' are isomorphic if and only if there exist units $u_i \in D^*$ satisfying the equations 5.2.2. As above, the map 5.2.3 induces a bijection between the set of isomorphism classes of $\mathbf{F}$-vector space schemes over D satisfying $(**)$ and the set

$$\{(a_0, \ldots, a_{r-1}) \in \mathbf{Z}^r \mid 0 \le a_i \le e\}. \tag{5.3.2}$$

Note that in the case $e = p-1$ this shows that every $\mathbf{F}$-vector space scheme G over K satisfying $(**)$ can be extended over D; that any two such extensions are uniquely isomorphic except when G corresponds to $(0, \ldots, 0)$ in the set 5.2.5 (i.e., G is constant); if G is constant then there are exactly two extensions corresponding to $(0, \ldots, 0)$ and $(p-1, \ldots, p-1)$ which are étale and multiplicative, respectively.

5.4 From $e = 1$ to $e' = p-1$. Let $D := W(\overline{\mathbf{F}}_p)$ be the ring of Witt vectors of $\overline{\mathbf{F}}_p$, i.e., D is a complete discrete valuation ring with residue field $\overline{\mathbf{F}}_p$ and uniformizer p. Let $D' := D[\zeta_p]$ where ζ_p is a primitive p^{th} root of unity. Note that D and D' are strictly henselian discrete valuation rings with absolute ramification indices $e = 1$ and $e' = p-1$, respectively; let K and K' denote their fields of fractions and let v and v' denote their valuations.

Let G_K be a $\mathbf{F}$-vector space scheme given by equations 5.2.1 in the isomorphism class corresponding to an element $a = (a_0, \ldots, a_{r-1})$ in the set 5.2.5. Let $G_{K'}$ be the pull-back to $\mathrm{Spec}(K')$ of G_K. Of course, $G_{K'}$ is described by the same equations as G_K, but note that $v'(\delta_i) = (p-1)v(\delta_i)$ for all i. As remarked above in §5.3, $G_{K'}$ can be extended (uniquely in all cases except one) to a $\mathbf{F}$-vector space scheme $G_{D'}$ over D' satisfying $(**)$. Let $G_{D'}$ be such an extension given by equations $X_i'^p = \delta_i' X_{i+1}'$; let $a_i' = v'(\delta_i')$ and let $a' = (a_0', \ldots, a_{r-1}')$. Note that $0 \le a_i' \le p-1$ for all i and that $(p-1)a - a' \in \Lambda$ with Λ as in 5.2.4. Let

$$f \colon G_{K'} \xrightarrow{\sim} G_{D',K'} \quad f^\# \colon X_i' \mapsto u_i X_i \quad (u_i \in K'^*) \tag{5.4.1}$$

be an isomorphism (unique up to $\mathbf{F}^*$). Let $b_i = -v'(u_i)$ (note the sign). The b_i are independent of the choice of f and from 5.2.2 it follows that the b_i are the solution to the set of equations

$$a'_i = -pb_i + (p-1)a_i + b_{i+1} \quad (i \in \mathbf{Z}/r\mathbf{Z}). \tag{5.4.2}$$

In order to describe the action (yet to be defined) of $\mathrm{Gal}(\overline{K}/K)$ on the special fibre $G_{\overline{\mathbf{F}}_p}$ of $G_{D'}$ in terms of the a_i we need to have a more explicit description of the a'_i and the b_i. In §5.2, after identifying I_t with $\varprojlim \mu_n(K)$, it was shown that the action of $\mathrm{Gal}(\overline{K}/K)$ on $G_K(\overline{K})$ is given by the character ϕ in 5.2.7. Let I'_t denote the tame ramification group of $\mathrm{Gal}(\overline{K}/K')$. Just as for I_t, we identify I'_t with $\varprojlim \mu_n(K') = \varprojlim \mu_n(\overline{\mathbf{F}}_p) = \varprojlim \mu_n(K)$. Then the canonical injection $I'_t \subset I_t$ corresponds to multiplication by $p{-}1$ on $\varprojlim \mu_n(K)$. It follows that the $\mathbf{F}$-vector space scheme $G_{K'}$ corresponds to the character $\phi' = \phi^{p-1}$. From equation 5.2.7 for both ϕ and ϕ' it the follows that

$$p^{r-1}a'_0 + \cdots + pa'_{r-2} + a'_{r-1} \equiv (p-1)(p^{r-1}a_0 + \cdots + pa_{r-2} + a_{r-1}) \pmod{p^r - 1}. \tag{5.4.3}$$

For any set of x_i, $i \in \mathbf{Z}/r\mathbf{Z}$, with $0 \leq x_i \leq p-1$ let $[x_{r-1}\cdots x_1x_0]_p$ denote the number $x_{r-1}p^{r-1} + \cdots + x_1p + x_0$, written in base p. Then the equation 5.4.3 means that one obtains $[a'_0 \cdots a'_{r-2}a'_{r-1}]_p$ by summing $[a_1 \cdots a_{r-1}a_0]_p$ and $[(p{-}1{-}a_0)\cdots(p{-}1{-}a_{r-1})]_p$ with a "*cyclic* carry", i.e., one adds them as if they were numbers written out in base p but the carry that shows up on the left must be added on the right to the result.

Let us now discuss how to find the unique solution of equation 5.4.2 for the b_i. One sees immediately that $b_i \equiv a'_{i-1} + a_{i-1} \pmod p$ for all i. If j is such that $|b_j|$ is maximal among the $|b_i|$ we see from 5.4.2 that $|pb_j| \leq (p-1)^2 + |b_j|$, i.e., $|b_i| \leq p-1$ for all i. Combining equation 5.4.2 and the fact that $b_i \geq -(p-1)$ for all i shows that $b_i \geq -1$ for all i. Now remark that $b_i = -1$ implies that $a_i = 0$, $a'_i = p-1$ and $b_{i+1} = -1$. So we can conclude that

$$\begin{aligned} &b_i \equiv a'_{i-1} + a_{i-1} \pmod p \quad \text{and} \quad 0 \leq b_i \leq p-1 \quad \text{for all } i, \text{ or} \\ &a_i = 0, \quad a'_i = p-1, \quad \text{and} \quad b_i = -1 \quad \text{for all } i. \end{aligned} \tag{5.4.4}$$

For later use we give the following explicit formulas in the case $r = 2$:

	a'_0	a'_1	b_0	b_1
$a_0 > a_1$	$p + a_1 - a_0$	$a_0 - a_1 - 1$	$a_0 - 1$	a_1
$a_0 < a_1$	$a_1 - a_0 - 1$	$p + a_0 - a_1$	a_0	$a_1 - 1$
$a_0 = a_1$	0	0	a_1	a_0
	$p - 1$	$p - 1$	$a_1 - 1$	$a_0 - 1$

(5.4.5)

5.5 Action of inertia. The notation is as in §5.4. The cyclotomic character χ (see 2.4) gives an isomorphism $\mathrm{Gal}(K'/K) \to \mathbf{F}_p{}^*$. For $\sigma \in \mathrm{Gal}(K'/K)$ we let $\mathrm{Spec}(\sigma)$ denote the automorphisms of $\mathrm{Spec}(K')$ and $\mathrm{Spec}(D')$ that it induces. We let $g(\sigma)$ be the automorphism $\mathrm{id}_G \times \mathrm{Spec}(\sigma)$

of $G_{K'}$. Then there is a unique automorphism $g(\sigma)$ of $G_{D',K'}$, such that the isomorphism f in 5.4.1 commutes with $\mathrm{Spec}(\sigma)$. Then we have:

$$f^{\#}(g(\sigma)^{\#}X'_i) = g(\sigma)^{\#}f^{\#}X'_i = g(\sigma)^{\#}(u_iX_i) = \sigma(u_i)X_i = f^{\#}(\sigma(u_i)u_i^{-1}X'_i),$$

from which we conclude that

$$g(\sigma)^{\#}X'_i = \sigma(u_i)u_i^{-1}X'_i = \chi(\sigma)^{-b_i}X'_i \quad (i \in \mathbf{Z}/r\mathbf{Z}). \tag{5.5.1}$$

From this formula it follows that the automorphisms $g(\sigma)$ extend (uniquely, of course) to automorphisms $g(\sigma)$ of $G_{D'}$. Restricting these to the special fibre then gives us an action of $\mathrm{Gal}(K'/K)$ on $G_{\overline{\mathbf{F}}_p}$ by $\overline{\mathbf{F}}_p$-automorphisms $g(\sigma)$. We are interested in the action by $\mathrm{Gal}(K'/K)$ on $\mathrm{Cot}(G_{\overline{\mathbf{F}}_p})$, the cotangent space at 0 of $G_{\overline{\mathbf{F}}_p}$. The $\overline{\mathbf{F}}_p$-vector space $\mathrm{Cot}(G_{\overline{\mathbf{F}}_p})$ has the following presentation:

$$\mathrm{Cot}(G_{\overline{\mathbf{F}}_p}) = \Big(\bigoplus_{i\in\mathbf{Z}/r\mathbf{Z}} \overline{\mathbf{F}}_p\cdot dX'_i\Big)\Big/\Big(\bigoplus_{i\in\mathbf{Z}/r\mathbf{Z}} \overline{\mathbf{F}}_p\cdot \bar{\delta}'_i dX'_{i+1}\Big) = \bigoplus_{\substack{i\in\mathbf{Z}/r\mathbf{Z}\\ a'_{i-1}>0}} \overline{\mathbf{F}}_p\cdot dX'_i, \tag{5.5.2}$$

where $\bar{\delta}'_i$ denotes the image of δ'_i in $\overline{\mathbf{F}}_p$. From 5.5.1 we see that $\mathrm{Gal}(K'/K)$ acts on the generator dX'_i via the character χ^{-b_i}.

6 The local representation in the supersingular case.

6.1 The aim of this section is to give a proof of Theorem 2.6. So we let f be a cusp form of type (N,k,ε) over $\overline{\mathbf{F}}_p$ which is an eigenform for all T_l^* with eigenvalues a_l; we suppose that $2 \le k \le p{+}1$ and that $a_p = 0$. By Theorem 2.8 we may suppose that $2 \le k \le p$. In order to prove Theorem 2.6 we may suppose, after replacing N by a multiple of N, that the curve $X_1(N)$ over $\mathbf{Z}[1/N]$ represents the stack $[\Gamma_1(N)]_{\mathbf{Z}[1/N]}$. For $k>2$ we will use a construction of ρ_f using models of the modular curve $X_1(pN)$ over various rings. All details of this construction can be found in [10]. Unfortunately Gross [10] works with embeddings of μ_{pN} into elliptic curves whereas we prefer to work with *points* of exact order pN in the terminology of [16], since then we can apply all the results of [16] without change. This explains that most formulas in [10] differ from ours. The construction, very briefly, goes as follows. First one constructs a differential form $\omega(f)$ on the so-called Igusa curve over $X_1(N)_{\overline{\mathbf{F}}_p}$ which has the same eigenvalues for all T_l^* as f has. The Igusa curve occurs in the characteristic p fibre of the stable model of $X_1(pN)$ over $\mathbf{Z}_p[\zeta_p]$, where ζ_p is a p^{th} root of unity. This allows one to lift $\omega(f)$ to a characteristic 0 eigenform on $X_1(pN)_{\overline{\mathbf{Q}}}$. Let $J_{\mathbf{Q}}$ denote the jacobian of $X_1(pN)_{\mathbf{Q}}$. Then ρ_f is obtained as the "reduction mod p" of one of the factors of the $\mathbf{Q}_p[G_{\mathbf{Q}}]$-module $(\varprojlim J_{\mathbf{Q}}[p^n](\overline{\mathbf{Q}}))\otimes_{\mathbf{Z}_p}\mathbf{Q}_p$.

6.2 The reduction mod p of $X_1(pN)$. Let $[\Gamma_1(p),\Gamma_1(N)]$ denote the algebraic stack over $\mathbf{Z}$ classifying triples $(E/S,\phi,\alpha)$ where E/S is a generalized elliptic curve, where $\phi\colon \mathbf{F}_p \to E(S)$ and $\alpha\colon \mathbf{Z}/N\mathbf{Z} \to E(S)$ are a $\mathbf{F}_p$-structure and a $\mathbf{Z}/N\mathbf{Z}$-structure (see [16], §1.5) such that the group generated by the images of ϕ and α meets all irreducible components of all geometric fibres.

Let $[\Gamma(p),\Gamma_1(N)]$ denote the algebraic stack over $\mathbf{Z}$ classifying triples $(E/S,\phi,\alpha)$, where E/S and α are as above and where $\phi\colon \mathbf{F}_p\times\mathbf{F}_p\to E[p]$ is a Drinfeld basis ([16], §1.4), such that the group generated by the images of ϕ and α meets all irreducible components of all geometric fibres. The group $\mathrm{GL}_2(\mathbf{F}_p)$ acts on $[\Gamma(p),\Gamma_1(N)]$ from the right by:

$$g\in \mathrm{GL}_2(\mathbf{F}_p):\quad (E/S,\phi,\alpha)\mapsto (E/S,\phi\circ g,\alpha).$$

The map $(E/S,\phi,\alpha)\mapsto (E'/S,\phi|_{\mathbf{F}_p\times\{0\}},\alpha)$, where E'/S is obtained from E/S by contractions (see [8], IV, §1), from $[\Gamma(p),\Gamma_1(N)]$ to $[\Gamma_1(p),\Gamma_1(N)]$ is the quotient, in the sense of [16], §7.1, by the action of the group $\Gamma:=\{g\in\mathrm{GL}_2(\mathbf{F}_p)\mid g(1,0)=(1,0)\}$ ([16], Theorem 7.4.2).

For a scheme or an algebraic stack X over $\mathbf{Z}$ let $X\widetilde{\otimes}_{\mathbf{Z}}\mathbf{Z}[\zeta_p]$ denote the normalization of $X\otimes_{\mathbf{Z}}\mathbf{Z}[\zeta_p]$. Then we get an isomorphism (see [16], §9 and §5.1.1):

$$\left(\coprod_{\sigma\in\mathrm{Gal}(\mathbf{Q}(\zeta_p)/\mathbf{Q})}[\Gamma(p)^{\sigma(\zeta_p)-\mathrm{can}},\Gamma_1(N)]\right)\Big/\Gamma\ \xrightarrow{\ \sim\ }\ [\Gamma_1(p),\Gamma_1(N)]\widetilde{\otimes}_{\mathbf{Z}}\mathbf{Z}[\zeta_p].$$

Since Γ acts transitively on the terms in the disjoint union we get an isomorphism:

$$[\Gamma(p)^{\zeta_p-\mathrm{can}},\Gamma_1(N)]/\left(\begin{smallmatrix}1&*\\0&1\end{smallmatrix}\right)\to[\Gamma_1(p),\Gamma_1(N)]\widetilde{\otimes}_{\mathbf{Z}}\mathbf{Z}[\zeta_p].\tag{6.2.1}$$

Let $X/\mathbf{Z}_p[\zeta_p]$ be the moduli scheme over $\mathbf{Z}_p[\zeta_p]$ representing $[\Gamma_1(p),\Gamma_1(N)]\widetilde{\otimes}_{\mathbf{Z}}\mathbf{Z}[\zeta_p]$. Using the results from [16], Chapter 13 it is not hard to obtain the following description of X.

6.2.2 X is a regular stable model of $X_1(pN)$ in the sense of [7],

6.2.3 $X_{\mathbf{F}_p}=I\cup E$, where $I\cong\overline{\mathcal{M}}([\mathrm{Ig}(p),\Gamma_1(N)])$ and $E\cong\overline{\mathcal{M}}([\mathrm{ExIg}(p,1),\Gamma_1(N)])$ (these are isomorphisms over $X_1(N)_{\mathbf{F}_p}$; they are not unique),

6.2.4 $I\cap E$ is the set of supersingular points.

Here $[\mathrm{Ig}(p)]$ is the $\mathbf{F}_p$-stack classifying pairs $(E/S,P)$ where $P\in E^{(p)}(S)$ generates $\ker V\colon E^{(p)}\to E$, and $[\mathrm{ExIg}(p,1)]$ classifies pairs $(E/S,P)$ where $P\in E(S)$ is such that $(P,0)$ is a Drinfeld basis of $E[p]$. The curves I and E are smooth, proper and geometrically irreducible over $\mathbf{F}_p$; the assignment

$$(E/S,P,Q)\mapsto(E^{(p)}/S,P,Q^{(p)})\tag{6.2.5}$$

induces an "exotic" ([16], §11) isomorphism $[\mathrm{Ig}(p),\Gamma_1(N)]\to[\mathrm{ExIg}(p,1),\Gamma_1(N)]$.

6.3 The reduction mod p of $J_{\mathbf{Q}}$. Let J be the Néron model over $\mathbf{Z}_p[\zeta_p]$ of the jacobian of $X_1(pN)_{\mathbf{Q}_p(\zeta_p)}$. Since $X/\mathbf{Z}_p[\zeta_p]$ is a stable curve, $\mathrm{Pic}^o_{X/\mathbf{Z}_p[\zeta_p]}$ is semi-abelian and $J^o=\mathrm{Pic}^o_{X/\mathbf{Z}_p[\zeta_p]}$. Let $\overline{J}=J_{\mathbf{F}_p}$. The pullbacks of invertible sheaves from $X_{\mathbf{F}_p}$ to I and E induce an exact sequence:

$$0\longrightarrow T\longrightarrow\overline{J}^o\longrightarrow J_I\oplus J_E\longrightarrow 0,\tag{6.3.1}$$

where J_I and J_E are the jacobians of I and E, and where T is a torus whose character group $\mathrm{X}(T)=\mathrm{Hom}_{\overline{\mathbf{F}}_p}(T_{\overline{\mathbf{F}}_p},\mathrm{G}_{m\overline{\mathbf{F}}_p})$ sits in an exact sequence:

$$0\longrightarrow\mathrm{X}(T)\longrightarrow\mathbf{Z}^{\mathrm{s.s.}X_1(N)(\overline{\mathbf{F}}_p)}\xrightarrow{\ +\ }\mathbf{Z}\longrightarrow 0.\tag{6.3.2}$$

Here s.s.$X_1(N)(\overline{\mathbf{F}}_p)$ denotes the set of supersingular points.

6.4 The ring of Hecke operators. For each prime number l we have a correspondence T_l on $[\Gamma_1(p), \Gamma_1(N)]$:

$$T_l\colon (E/S, \phi, \alpha) \mapsto \sum_{\beta\colon E\to E'} (E'/S, \beta\circ\phi, \beta\circ\alpha), \tag{6.4.1}$$

where the sum is over the $\beta \in [\Gamma_0(l)](E/S)$ such that $(E'/S, \beta\circ\phi, \beta\circ\alpha) \in [\Gamma_1(p), \Gamma_1(N)]$. Here we assume that S is irreducible, the j-invariant $j\colon S \to \mathrm{P}^1_{\mathbf{Z}}$ is flat at infinity and at the points with residue characteristic l, and that $[\Gamma_0(l)]_{E/S}$ is set theoretically the union of its S-valued points; we have identified the stacks $[\Gamma_0(l)]$ and $[l\text{-Isog}]$.

We also have the "diamond operators" $\langle a\rangle_N$, $a \in (\mathbf{Z}/N\mathbf{Z})^*$, and $\langle b\rangle_p$, $b \in \mathbf{F}_p{}^*$. They act on $[\Gamma_1(pN)]$ by automorphisms:

$$\langle a\rangle_N\colon (E/S, \phi, \alpha) \mapsto (E/S, \phi, a\alpha), \qquad \langle b\rangle_p\colon (E/S, \phi, \alpha) \mapsto (E/S, b\phi, \alpha). \tag{6.4.2}$$

We let the T_l, $\langle a\rangle_N$ and $\langle b\rangle_p$ act on $J_{\mathbf{Q}}$ via Albanese functoriality (i.e., push-forward of divisors). Let H be the subring of $\mathrm{End}(J_{\mathbf{Q}})$ generated by them; H is commutative.

We define correspondences T_l and automorphisms $\langle a\rangle_N$ and $\langle b\rangle_p$ on the stacks $[\mathrm{Ig}(p), \Gamma_1(N)]$ and $[\mathrm{ExIg}(p,1), \Gamma_1(N)]$ by formulas analogous to 6.4.1 and 6.4.2. They induce operators T_l, $\langle a\rangle_N$ and $\langle b\rangle_p$ on J_I and J_E (independent of the choice of isomorphisms in 6.2.3).

Finally we have the automorphism $w = w_{\zeta_{pN}}$ of $[\Gamma_1(p), \Gamma_1(N)]_{\mathbf{Z}[1/pN, \zeta_{pN}]}$ defined by

$$w\colon (E/S, \phi, \alpha) \mapsto (E'/S, \phi', \alpha'), \tag{6.4.3}$$

where $\beta\colon E \to E'$ is the pN-isogeny with $\ker(\beta) = \mathrm{im}(\phi) + \mathrm{im}(\alpha)$, where $\mathrm{im}(\phi') + \mathrm{im}(\alpha') = \ker(\beta^t)$ and $\langle \phi(1) + \alpha(1), \phi'(1) + \alpha'(1)\rangle_\beta = \zeta_{pN}$. Here $\langle *, *\rangle_\beta$ denotes the pairing from [16], §2.8 and $\beta^t\colon E \to E'$ denotes the dual of β. The use of w is that it gives the Rosati involution on H:

$$\forall h \in H \subset \mathrm{End}_{\mathbf{Q}(\zeta_{pN})}(J): \quad h^t = whw^{-1}. \tag{6.4.4}$$

In fact, one can check that for all l, wT_lw^{-1} *is* the correspondence dual to T_l.

6.5 Action of inertia. For every $\sigma \in \mathrm{Gal}(\mathbf{Q}(\zeta_p)/\mathbf{Q})$ the automorphism

$$\mathrm{id} \times \mathrm{Spec}(\sigma)\colon \quad [\Gamma_1(p), \Gamma_1(N)] \otimes_{\mathbf{Z}} \mathbf{Q}(\zeta_p)$$

extends uniquely to an automorphism $g(\sigma)$ of $[\Gamma_1(p), \Gamma_1(N)] \tilde{\otimes}_{\mathbf{Z}} \mathbf{Z}[\zeta_p]$. This gives an automorphism $g(\sigma)$ of X. Restricting the $g(\sigma)$ to $X_{\mathbf{F}_p}$ gives a right-action of $\mathrm{Gal}(\mathbf{Q}(\zeta_p)/\mathbf{Q})$ on $X_{\mathbf{F}_p}$ by $\mathbf{F}_p$-automorphisms. The automorphism $g(\sigma)$ of X induces via Albanese functoriality an automorphism $g(\sigma)$ of J, whose restriction to $\overline{J}$ we still denote by $g(\sigma)$.

6.6 Some formulas. By the Néron property of J all endomorphisms of $J_{\mathbf{Q}_p(\zeta_p)}$ extend to endomorphisms of J; their restrictions to $\overline{J}^o$ preserve the exact sequence 6.3.1. We write an endomorphism of $J_I \oplus J_E$ as a matrix $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$, with $a \in \mathrm{End}(J_I)$, $b \in \mathrm{Hom}(J_E, J_I)$, $c \in \mathrm{Hom}(J_I, J_E)$ and $d \in \mathrm{End}(J_E)$. Then we have the following formulas:

$$
\begin{gathered}
T_l = \begin{pmatrix} \langle l \rangle_p^{-1} T_l & 0 \\ 0 & T_l \end{pmatrix} \quad \text{if } l \neq p, \qquad T_p = \begin{pmatrix} \langle p \rangle_N V & \sum\limits_{a \in \mathbf{F}_p{}^*} \langle a \rangle_p \langle p \rangle_N \iota^{-1} \\ 0 & F \end{pmatrix}, \\
\langle a \rangle_N = \begin{pmatrix} \langle a \rangle_N & 0 \\ 0 & \langle a \rangle_N \end{pmatrix}, \quad \langle a \rangle_p = \begin{pmatrix} \langle a \rangle_p^{-1} & 0 \\ 0 & \langle a \rangle_p \end{pmatrix}, \quad g(\sigma) = \begin{pmatrix} \langle \chi(\sigma) \rangle_p & 0 \\ 0 & \mathrm{id} \end{pmatrix}.
\end{gathered} \tag{6.6.1}
$$

where V and F denote “Verschiebung” and “Frobenius”, and where $\iota\colon J_I \to J_E$ is induced by the isomorphism 6.2.5. For a proof of these identities (in a slightly different setting) we refer to [10]. A detailed computation of T_p, in a more general case, can be found in [30], §7–8.

The action of the operators above on the torus T can be described in terms of the effect of the correspondences they are induced from on the supersingular points in I (or in E). It follows from 6.3.2 that the $\langle a \rangle_p$ and the $g(\sigma)$ act trivially on T.

Over $\mathbf{F}_p(\zeta_N)$ we define automorphisms w_{ζ_N} of $[\mathrm{Ig}(p), \Gamma_1(N)]$ and $[\mathrm{ExIg}(p,1), \Gamma_1(N)]$ by

$$
(E/S, Q, R) \mapsto (E'/S, \beta^{(p)} Q, R') \quad \text{and} \quad (E/S, P, R) \mapsto (E'/S, \beta P, R'), \tag{6.6.2}
$$

where $\beta\colon E \to E'$ is the N-isogeny with kernel generated by R, where R' generates $\ker(\beta^t)$ and where $\langle R, R' \rangle_\beta = \zeta_N$. Then, for a particular choice of isomorphisms in 6.2.3, we have

$$
w = w_{\zeta_p N} = \begin{pmatrix} 0 & w_{\zeta_N} \langle -1 \rangle_p \langle p \rangle_N \iota^{-1} \\ \langle N \rangle_p^{-1} w_{\zeta_N} \iota & 0 \end{pmatrix}. \tag{6.6.3}
$$

6.7 The construction of $\omega(f)$**.** Let f be as in 6.1, but suppose that $k > 2$. Let $A \in M(1, p-1)$ be the Hasse invariant. In [16], §12.8 a construction is given of a $(p-1)^{\text{th}}$ root of A on $[\mathrm{Ig}(p)]$, i.e., a global section $A^{1/p-1}$ of $\underline{\omega}_{\mathbf{E}/[\mathrm{Ig}(p)]}$ such that $A = (A^{1/p-1})^{p-1}$ in $\underline{\omega}^{\otimes p-1}_{\mathbf{E}/[\mathrm{Ig}(p)]}$. The form $A^{1/p-1}$ has simple zeroes at the supersingular points and for all $a \in \mathbf{F}_p{}^*$ we have $\langle a \rangle_p^* A^{1/p-1} = a A^{1/p-1}$.

For $2 \leq k \leq p$ we have an isomorphism of $\overline{\mathbf{F}}_p$-vector spaces:

$$
\begin{gathered}
\bigoplus_{k=2}^p M(N,k)^0 \xrightarrow{\sim} \mathrm{H}^0([\mathrm{Ig}(p), \Gamma_1(N)]_{\overline{\mathbf{F}}_p}, \Omega^1), \\
\phi_k \mapsto \omega(\phi_k) = \mathrm{KS}(\mathrm{pr}^*(\phi_k)/(A^{1/p-1})^{k-2}),
\end{gathered} \tag{6.7.1}
$$

where $\mathrm{pr}\colon [\mathrm{Ig}(p), \Gamma_1(N)] \to [\Gamma_1(N)]$ is the canonical projection and where $\mathrm{KS}\colon \underline{\omega}^{\otimes 2} \to \Omega^1(\text{cusps})$ is the Kodaira-Spencer morphism on $[\mathrm{Ig}(p), \Gamma_1(N)]$ ([16], §12.8.8).

Now let $\omega_{X/\mathbf{Z}_p[\zeta_p]}$ be the dualizing sheaf on X (see [8], I, §2.3) and let $(\omega(f), 0)$ be its section over $X_{\overline{\mathbf{F}}_p}$ whose restrictions to $I_{\overline{\mathbf{F}}_p}$ and $E_{\overline{\mathbf{F}}_p}$ are $\omega(f)$ and 0, respectively. Since $J^o = \mathrm{Pic}^o_{X/\mathbf{Z}_p[\zeta_p]}$ we have canonical isomorphisms:

$$
\mathrm{Cot}(J^o) = \mathrm{H}^0(X, \omega_{X/\mathbf{Z}_p[\zeta_p]}), \qquad \mathrm{Cot}(\overline{J}^o_{\overline{\mathbf{F}}_p}) = \mathrm{H}^0(X_{\overline{\mathbf{F}}_p}, \omega_{X_{\overline{\mathbf{F}}_p}/\overline{\mathbf{F}}_p}), \tag{6.7.2}
$$

hence we can view $(\omega(f), 0)$ as an element of $\mathrm{Cot}(\overline{J}^o_{\overline{\mathbf{F}}_p})$. One can check that as such, $(\omega(f), 0)$ is an eigenvector for the subring H of $\mathrm{End}(J^o)$ defined in 6.4:

$$\begin{aligned} &\forall l: \ \ T_l^*((\omega(f),0)) = a_l(\omega(f),0)), \\ &\forall a \in \mathbf{F}_p{}^*: \ \ \langle a\rangle_p^*(\omega(f),0)) = a^{k-2}(\omega(f),0)), \\ &\forall a \in (\mathbf{Z}/N\mathbf{Z})^*: \ \ \langle a\rangle_N^*(\omega(f),0)) = \varepsilon(a)(\omega(f),0)). \end{aligned} \tag{6.7.3}$$

To check this one needs to work out the relation between the isomorphism in 6.7.1 and the T_l, $\langle a\rangle_N$ and $\langle a\rangle_p$, and one uses the formulas in 6.6.

6.8 The proof of Theorem 2.6. Let f be as in 6.7. The action of H on $(\omega(f), 0)$ gives us a character $\lambda\colon H \to \overline{\mathbf{F}}_p$ such that for all h in H we have $h^*(\omega(f), 0) = \lambda(h)(\omega(f), 0)$. Let $\mathbf{F} = \mathrm{im}(\lambda)$ and let $m = \ker(\lambda)$; hence m is a maximal ideal of H. Let H_p be the p-adic completion of H and let H_m be the m-adic completion. Since H, being a subring of $\mathrm{End}(J_{\mathbf{Q}})$, is a free $\mathbf{Z}$-module of finite rank, we have

$$H_p = H \otimes_{\mathbf{Z}} \mathbf{Z}_p = H_m \times H'_m, \tag{6.8.1}$$

where H'_m is the product of the completions of H at the maximal ideals over p other than m. Let ϵ and ϵ' be the corresponding idempotents in H_p: $1 = \epsilon + \epsilon'$, $\epsilon^2 = \epsilon$, $\epsilon'^2 = \epsilon'$, $\epsilon\epsilon' = 0$, $\epsilon H_p = H_m$, $\epsilon' H_p = H'_m$. Note that for any H_p-module M (for example $J^o[p^n](T)$ for some $\mathbf{Z}_p[\zeta_p]$-scheme T) we have $M = \epsilon M \oplus \epsilon' M = M[\epsilon'] \oplus M[\epsilon]$.

Let $\eta = \frac{1}{p-1} \sum_{a \in \mathbf{F}_p{}^*} \langle a\rangle_p \in H_p$; let $\eta' = 1 - \eta$. Then η and η' are orthogonal idempotents; for any H_p-module M we have $M = \eta M \oplus \eta' M$ and ηM is the submodule of $\langle a\rangle_p$-invariants of M. Since $\lambda(\langle a\rangle_p) = a^{k-2}$ for all $a \in \mathbf{F}_p{}^*$, and $3 \le k \le p$, we have $\epsilon\eta = \epsilon$, $\epsilon\eta' = 0$.

Let $J^o[p^\infty]$ be the "p-divisible group" $(J^o[p^n])_{n\ge 0}$ over $\mathbf{Z}_p[\zeta_p]$ (for the definition of a p-divisible group see for example [29], [6], III or [28], VI). Note that the $J^o[p^n]$ are quasi-finite and flat but not necessarily finite over $\mathbf{Z}[\zeta_p]$. Since H_p acts on $J^o[p^\infty]$, it splits as a direct sum:

$$J^o[p^\infty] = G \oplus G', \tag{6.8.2}$$

with $G_n = J^o[p^n][\epsilon']$ and $G'_n = J^o[p^n][\epsilon]$. We also have the direct sum decomposition $J^o[p^\infty] = J^o[p^\infty][\eta'] \oplus J^o[p^\infty][\eta]$ and $G = J^o[p^\infty][\eta', \epsilon']$. Let $J^o[p^\infty]^{\mathrm{f}} = (J^o[p^n]^{\mathrm{f}})_n$ be the finite part (i.e., "la partie fixe" in [12], 7, IX, §2.2.3) of $J^o[p^\infty]$. Since the $\langle a\rangle_p$ act trivially on the torus T in $\overline{J}^o$ (see 6.6), they act trivially on $J^o[p^\infty]/J^o[p^\infty]^{\mathrm{f}}$ by [12], 7, IX, 5.5.8. It follows that all the G_n are finite over $\mathbf{Z}_p[\zeta_p]$.

Let $\overline{G} = G \times_{\mathbf{Z}_p[\zeta_p]} \mathbf{F}_p = \overline{J}^o[p^\infty][\epsilon']$. The exact sequence 6.3.1 induces an exact sequence of p-divisible groups over $\mathbf{F}_p$:

$$0 \longrightarrow T[p^\infty] \longrightarrow \overline{J}^o[p^\infty] \longrightarrow J_I[p^\infty] \oplus J_E[p^\infty] \longrightarrow 0. \tag{6.8.3}$$

Again, since the $\langle a\rangle_p$ act trivially on T, we get an isomorphism

$$\overline{J}^o[p^\infty][\eta'] \xrightarrow{\sim} (J_I[p^\infty] \oplus J_E[p^\infty])[\eta'] = J_I[p^\infty][\eta'] \oplus J_E[p^\infty][\eta'], \tag{6.8.4}$$

where the last equality follows from the fact that the $\langle a\rangle_p$ preserve J_I and J_E (see 6.6.1). It also follows that H_p preserves the direct sum decomposition in 6.8.4. Therefore we finally get:

$$\overline{G} \;=\; J_I[p^\infty][\eta',\epsilon'] \oplus J_E[p^\infty][\eta',\epsilon'] \;=\; J_I[p^\infty][\epsilon'] \oplus J_E[p^\infty][\epsilon'] \;=\; \overline{G}_I \oplus \overline{G}_E, \tag{6.8.5}$$

where the last equality is a *definition*. Since the $\overline{G}_{I,n}$ and the $\overline{G}_{E,n}$ are finite group schemes over the perfect field $\mathbf{F}_p$, we get direct sum decompositions into étale, multiplicative and local-local parts:

$$\overline{G}_I = \overline{G}_I^{\mathrm{e}} \oplus \overline{G}_I^{\mathrm{m}} \oplus \overline{G}_I^{\mathrm{l}}, \qquad \overline{G}_E = \overline{G}_E^{\mathrm{e}} \oplus \overline{G}_E^{\mathrm{m}} \oplus \overline{G}_E^{\mathrm{l}}. \tag{6.8.6}$$

From the fact that T_p acts on J_I as $\langle p\rangle_N V$, on J_E as F (see 6.6.1) and that $\lambda(T_p) = a_p = 0$, we conclude that $\overline{G}_I^{\mathrm{m}} = 0 = \overline{G}_E^{\mathrm{e}}$ since V (resp. F) acts both as a unit and a non-unit on $\overline{G}_I^{\mathrm{m}}$ (resp. $\overline{G}_E^{\mathrm{e}}$).

Let us now argue that $\overline{G}_I^{\mathrm{l}} \neq 0$. Since $\overline{G}_I^{\mathrm{m}} = 0$, it suffices to show that $\mathrm{Cot}(\overline{G}_I)$, the cotangent space of $\overline{G}_I$ at 0, is not zero (here $\mathrm{Cot}(\overline{G}_I) = \varprojlim \mathrm{Cot}(\overline{G}_{I,n})$ which equals $\mathrm{Cot}(\overline{G}_{I,1})$ since the $\overline{G}_I$ is a p-divisible group over $\mathbf{F}_p$). But we have:

$$0 \neq \omega(f) \in \mathrm{Cot}(J_{I,\overline{\mathbf{F}}_p})[\epsilon'^*] = \mathrm{Cot}(J_I[p^\infty]_{\overline{\mathbf{F}}_p})[\epsilon'^*] = \mathrm{Cot}(J_I[p^\infty][\epsilon']_{\overline{\mathbf{F}}_p}) = \mathrm{Cot}(\overline{G}_I) \otimes_{\mathbf{F}_p} \overline{\mathbf{F}}_p, \tag{6.8.7}$$

where we used that "Cot" is an additive functor and that ϵ' is an idempotent.

The next thing we want to show is that $\overline{G}_I$ and $\overline{G}_E$ become Cartier dual over $\mathbf{F}_p(\zeta_N)$. Let $(\cdot,\cdot)_n$ denote the Weil pairing on $J[p^n]_{\mathbf{Q}}$, and let $\langle\cdot,\cdot\rangle_n$ be the pairing on $J[p^n]_{\mathbf{Q}(\zeta_{pN})}$ given by $\langle x,y\rangle_n = (x, w(y))_n$, where w is the automorphism defined in 6.4. It follows from 6.4.4 that the $h \in H_p$ are self-adjoint with respect to the $\langle\cdot,\cdot\rangle_n$: $\langle h(x), y\rangle_n = \langle x, h(y)\rangle_n$ for all x, y and h. In particular this is true for ϵ and ϵ', which shows that the decomposition 6.8.2, pulled back to $\mathbf{Q}(\zeta_{pN})$, is orthogonal with respect to the $\langle\cdot,\cdot\rangle_n$.

According to [12], 7, IX, Thm. 5.4, the $(\cdot,\cdot)_n$ extend to $J^o[p^\infty]$ and their restrictions to $\overline{J}^o[p^\infty]$ are induced from a principal polarization on $J_I \oplus J_E$ which one can check to be equal to the product of the canonical principal polarizations on J_I and J_E. Since w interchanges I and E it follows that the $\langle\cdot,\cdot\rangle_n$ induce a perfect duality between $\overline{G}_{I,\mathbf{F}_p(\zeta_N)}$ and $\overline{G}_{E,\mathbf{F}_p(\zeta_N)}$. We conclude that $\overline{G}_E = \overline{G}_E^{\mathrm{l}}$ and that $\mathrm{Cot}(\overline{G}_E) \neq 0$. From 6.6.1 and 6.7.3 it follows that $g(\sigma)$ acts on $\mathrm{Cot}(\overline{G}_I)$ and $\mathrm{Cot}(\overline{G}_E)$ by $\chi(\sigma)^{2-k}$ and 1, respectively.

Let $\rho_f\colon G_{\mathbf{Q}} \to \mathrm{GL}_2(\mathbf{F})$ be the continuous semi-simple representation corresponding to f as in 2.3. Let W be the 2-dimensional $\mathbf{F}$-vector space scheme over $\mathbf{Q}$ corresponding to ρ_f. Let K, K', D and D' be as in 5.4. By construction, each $G[p^n](\overline{\mathbf{Q}})$ has all Jordan-Hölder quotients isomorphic to irreducible subrepresentations of ρ_f. Now $\rho_{f,p}$ has to be either in case 2.4.1 or 2.4.2. Suppose that $\rho_{f,p}$ is in case 2.4.1. Then there is an exact sequence

$$0 \longrightarrow W_K' \longrightarrow W_K \longrightarrow W_K'' \longrightarrow 0,$$

where W_K' and W_K'' are 1-dimensional $\mathbf{F}$-vector space schemes over K, corresponding to characters $\mathrm{Gal}(\overline{K}/K) \to \mathbf{F}^*$ of level 1. But then $W_{K'}'$ and $W_{K'}''$ are *constant*, hence all $G[p^n]_{K'}$ have all Jordan-Hölder quotients isomorphic to $\mathbf{F}_{K'}$. It follows from [21], §3.3.5 that then $\overline{G}[p^n] = \overline{G}[p^n]^{\mathrm{e}} \oplus \overline{G}[p^n]^{\mathrm{m}}$, but this is nonsense since we know that $\overline{G}[p^n] = \overline{G}[p^n]^{\mathrm{l}}$. We conclude that $\rho_{f,p}$ is in case 2.4.2, say with characters ϕ and ϕ' of level 2. In that case $\rho_{f,p}$ is absolutely irreducible hence certainly ρ_f is absolutely irreducible. A recent result by Boston, Lenstra and Ribet ([2], Theorems 1 and 2) states that $G[m](\overline{\mathbf{Q}})$ is isomorphic to a direct sum of d copies of ρ_f for some $d>0$. By [21], §3.3.5 both $W_{K'}$ and $G[m]_{K'}$ have *unique* extensions to $W_{D'}$ and $G[m]$ over D'. It follows that $G[m]\cong\oplus_{i=1}^{d}W_{D'}$ and that this isomorphism is compatible with the action of $\mathrm{Gal}(K'/K)$.

Let $\mathbf{F}'$ be a finite extension of $\mathbf{F}$ containing $\mathbf{F}_{p^2}$. Then $W_K\otimes_{\mathbf{F}}\mathbf{F}'$ splits as $\mathbf{F}'$-vector space scheme over K into a direct sum $V\oplus V'$ where V and V' correspond to the characters ϕ and $\phi'\colon\mathrm{Gal}(\overline{K}/K)\to\mathbf{F}_{p^2}{}^*\subset\mathbf{F}'^*$, respectively. Let $V(\phi)$ and $V(\phi')$ be the $\mathbf{F}_{p^2}$-vector space schemes over K corresponding to ϕ and $\phi'\colon\mathrm{Gal}(\overline{K}/K)\to\mathbf{F}_{p^2}{}^*$. Let $r'=[\mathbf{F}':\mathbf{F}_{p^2}]$. Then we have $V\cong\oplus_1^{r'}V(\phi)$ and $V'\cong\oplus_1^{r'}V(\phi')$ as $\mathbf{F}_{p^2}$-vector space schemes over K. Again, since local-local group schemes extend uniquely over D', we get an isomorphism, compatible with the action of $\mathrm{Gal}(K'/K)$, of $\mathbf{F}_{p^2}$-vector space schemes over D':

$$G[m]\otimes_{\mathbf{F}}\mathbf{F}' \;\cong\; \big(\oplus_1^d\oplus_1^{r'}V(\phi)_{D'}\big)\bigoplus\big(\oplus_1^d\oplus_1^{r'}V(\phi')_{D'}\big). \tag{6.8.8}$$

Since $\overline{G}[m]=\overline{G}_I[m]\oplus\overline{G}_E[m]$, the action of $\mathrm{Gal}(K'/K)$ on $\mathrm{Cot}(\overline{G}[m])$ consists of the characters χ^{2-k} and 1, both with positive multiplicities. Since $V(\phi)$ and $V(\phi')$ are isomorphic as group schemes over K, the $\mathrm{Gal}(K'/K)$-representations $\mathrm{Cot}(V(\phi)_{\overline{\mathbf{F}}_p})$ and $\mathrm{Cot}(V(\phi')_{\overline{\mathbf{F}}_p})$ are isomorphic. Now write $\phi=\psi^{a_0}\psi'^{a_1}$ with ψ and ψ' the fundamental characters of level 2 and $a_0>a_1$. It follows that $\mathrm{Gal}(K'/K)$ must act on $\mathrm{Cot}(V(\phi)_{\overline{\mathbf{F}}_p})$ and $\mathrm{Cot}(V(\phi')_{\overline{\mathbf{F}}_p})$ each via the characters χ^{2-k} and 1, both with positive multiplicities. By 5.5.2 and 5.5.1 this means that we must have $a_0=k-1$ and $a_1=0$, i.e., $\phi=\psi^{k-1}$, $\phi'=\psi'^{k-1}$. This completes the proof of Theorem 2.6 for $k>2$.

The proof for $k=2$ is similar but much easier, so we only give a sketch. Instead of $X_1(pN)$ we use the curve $X:=X_1(N)$ and its jacobian J over $\mathbf{Z}_p$; X and J have good reduction over $\mathbf{Z}_p$. The Kodaira-Spencer isomorphism then gives us a differential form $\omega(f)$ on $X_{\overline{\mathbf{F}}_p}$. Let H be the subring of $\mathrm{End}(J)$ generated by all T_l and $\langle a\rangle_N$. We have $T_l^*\omega(f)=a_l\omega(f)$ and $\langle a\rangle_N^*\omega(f)=\varepsilon(a)\omega(f)$ for all l and a. We define λ, $\mathbf{F}$, m, H_m, ϵ, ϵ', G and $\overline{G}$ as above. Since T_p acts on $\overline{J}$ as $F+\langle p\rangle_N V$ and $a_p=0$, we have $\overline{G}=\overline{G}^1$ and we get two characters $\phi=\psi^{a_0}\psi'^{a_1}$ and $\phi'=\phi^p$ as above. The fact that G has good reduction over $\mathbf{Z}_p$ then implies that $a_0=1$ and $a_1=0$. □

The proof above actually also proves the following result, that we will use in §9.

6.9 Proposition. *With the notation as above in the case $k>2$, let $\overline{W}=W_{\mathbf{Z}_p[\zeta_p]}\otimes_{\mathbf{Z}_p[\zeta_p]}\mathbf{F}_p$. Then $\mathrm{Gal}(\mathbf{Q}_p(\zeta_p)/\mathbf{Q}_p)$ acts on $\overline{W}$ via the characters χ^{2-k} and $\chi^0=1$. Let $\overline{W}=\overline{W}(2-k)\oplus\overline{W}(0)$ be the decomposition into eigenspaces. Then $\overline{W}(2-k)$ and $\overline{W}(0)$ are "1-dimensional" $\mathbf{F}$-vector space schemes (i.e., they satisfy condition $(**)$ in §5.1) over $\mathbf{F}_p$ that are annihilated by F and by V. Let d be the multiplicity of ρ_f in $J_{\mathbf{Q}}[m](\overline{\mathbf{Q}})$ as above. Then*

$$J_I[m]\;\cong\;\oplus_1^d\,\overline{W}(2-k)\qquad\textit{and}\qquad J_E[m]\;\cong\;\oplus_1^d\,\overline{W}(0),$$

where the action of H on J_I and J_E is described by the diagonal entries of the matrices in 6.6.1.

With the notation as above in the case $k = 2$, let $\overline{W} = W_{\mathbf{Z}_p} \otimes_{\mathbf{Z}_p} \mathbf{F}_p$. Then $\overline{W}[F] = \overline{W}[V]$ is a "1-dimensional" $\mathbf{F}$-vector space scheme over $\mathbf{F}_p$ and

$$\overline{J}[m] \;\cong\; \oplus_1^d \overline{W}.$$

7 Reduction to weight at most $p+1$.

7.1. In this section we will give a proof of Thm. 3.4, using certain spaces $S(N, k)$ that were introduced in [27]. Let p be any prime number and let $N \geq 5$ be prime to p; let $X := X_1(N)_{\overline{\mathbf{F}}_p}$. For any integer k, multiplication by the Hasse-invariant gives a short exact sequence of sheaves on X:

$$0 \longrightarrow \underline{\omega}^{\otimes k-p+1} \xrightarrow{\cdot A} \underline{\omega}^{\otimes k} \longrightarrow \oplus_{x \text{ s.s.}} \omega_{E_x}^{\otimes k} \longrightarrow 0. \tag{7.1.1}$$

Let $S(N, k) = \mathrm{H}^0(X, \oplus_{x \text{ s.s.}} \omega_{E_x}^{\otimes k})$; this is the space of modular forms of type (N, k) on supersingular elliptic curves. We let the T_l, $l \neq p$, act on the $S(N, k)$ by the usual formulas (see 2.2). The long exact sequences arising from 7.1.1, together with the following proposition, show that $\oplus_k M(N, k)$ and $\oplus_k S(N, k)$ have the same systems of eigenvalues for the T_l^*, $l \neq p$. To see this, consider the $S(N, k)$ and $M(N, k)$ as Artinian modules over the polynomial ring $\overline{\mathbf{F}}_p[T_l \mid l \neq p]$; systems of eigenvalues correspond to maximal ideals, etc.

7.2 Proposition. *There exists an element B in $S(N, p{+}1)$ inducing isomorphisms*

$$\cdot B\colon\;\; S(N, k) \xrightarrow{\sim} S(N, k + p + 1),$$

for all $k \in \mathbf{Z}$, such that $T_l^*(Bf) = lBT_l^*(f)$ for all $f \in S(N, k)$ and $l \neq p$. For all $k \in \mathbf{Z}$ there exists an isomorphism, compatible with the T_l^*, $l \neq p$, between $S(N, k{+}p^2{-}1)$ and $S(N, k)$.

Proof. We will give three constructions of such a B.

1. According to Robert, [23], Thm. B, one may take B to be the reduction mod p of the Eisenstein series E_{p+1}, if $p \geq 5$.

2. Let $x \in \text{s.s.}X(\overline{\mathbf{F}}_p)$. Since A has a simple zero at x, it gives a non-zero element in $\omega_{E_x}^{\otimes p-1} \otimes m_x/m_x^2$ (here m_x is the maximal ideal in $\mathcal{O}_{X,x}$). The Kodaira-Spencer isomorphism then gives a non-zero element in $\omega_{E_x}^{\otimes p+1}$.

3. (See [15].) Let E be a supersingular elliptic curve over $\overline{\mathbf{F}}_p$. There is an exact sequence $0 \to \mathrm{H}^0(E, \Omega) \to \mathrm{H}^1_{\mathrm{DR}}(E) \to \mathrm{H}^1(E, \mathcal{O}) \to 0$. Since $F\colon E \to E^{(p)}$ induces $0\colon \mathrm{H}^0(E^{(p)}, \Omega) \to \mathrm{H}^0(E, \Omega)$ and $0\colon \mathrm{H}^1(E^{(p)}, \mathcal{O}) \to \mathrm{H}^1(E, \mathcal{O})$, it induces a map $\mathrm{H}^1(E^{(p)}, \mathcal{O}) \to \mathrm{H}^0(E, \Omega)$. This map is an isomorphism; via Serre duality we can interpret it as an element of $\omega_{E_x}^{\otimes p+1}$.

It follows easily from lemma 7.4 that the elements B as described in (2) and (3) have the desired property. Multiplication by B^{p-1} is the desired isomorphism $S(N, k) \to S(N, k{+}p^2{-}1)$. Another way to get such an isomorphism is as follows (see [27]). Every supersingular elliptic curve over $\overline{\mathbf{F}}_p$ has a *unique* $\mathbf{F}_{p^2}$-structure for which the Frobenius endomorphism is $-p$. Every morphism between supersingular elliptic curves over $\overline{\mathbf{F}}_p$ descends to $\mathbf{F}_{p^2}$. It follows that each $\omega_{E_x}^{\otimes p^2-1}$ has a *canonical* $\overline{\mathbf{F}}_p$-basis; these bases are respected by all isogenies of degree prime to p. □

Proposition 7.2 shows that all systems of eigenvalues occurring in $\oplus_k M(N,k)$ occur, up to twist, in $\oplus_{1\leq k\leq p+1}S(N,k)$; we want to prove the same for $\oplus_{1\leq k\leq p+1}M(N,k)$. The relevant part of the long exact sequence arising from 7.1.1 is:

$$M(N,k) \longrightarrow S(N,k) \longrightarrow \mathrm{H}^1(X,\underline{\omega}^{\otimes k-p+1}). \tag{7.2.1}$$

Let $k'=p+1-k$, then Serre duality and the Kodaira-Spencer map give an isomorphism:

$$\mathrm{H}^1(X,\underline{\omega}^{\otimes k-p+1}) \xrightarrow{\mathrm{SD}} \mathrm{H}^0(X,\Omega\otimes\underline{\omega}^{\otimes p-1-k})^\vee \xrightarrow{\mathrm{KS}^\vee} \mathrm{H}^0(X,\underline{\omega}^{\otimes k'}(-\mathrm{cusps}))^\vee = M^0(N,k')^\vee.$$

Let $\zeta\in\overline{\mathbf{F}}_p^{\,*}$ have order N. Let w be the automorphism of $[\Gamma_1(N)]_{\overline{\mathbf{F}}_p}$ sending $(E/S,\alpha)$ to $(wE/S,w\alpha)$, where $\beta\colon E\to wE$ is the N-isogeny with $\ker(\beta)=\mathrm{im}(\alpha)$, $\mathrm{im}(w\alpha)=\ker(\beta^t)$ and $\langle\alpha(1),(w\alpha)(1)\rangle_\beta=\zeta$ (compare 6.4.3). We let w act on modular forms by: $(w^*f)(E,\alpha)=\beta^*f(wE,w\alpha)$.

7.3 Proposition. *Let Φ be the composition of the map $S(N,k)\to\mathrm{H}^1(X,\underline{\omega}^{\otimes k-p+1})$ as in 7.2.1 and $w^{*\vee}\circ(\mathrm{KS})^\vee\circ\mathrm{SD}$. For all k and all $l\neq p$ the following diagram is commutative:*

$$\begin{array}{ccccc}
M(N,k) & \longrightarrow & S(N,k) & \xrightarrow{\Phi} & M^0(N,k')^\vee \\
\downarrow{\scriptstyle T_l^*} & & \downarrow{\scriptstyle T_l^*} & & \downarrow{\scriptstyle l^{k-1}T_l^{*\vee}} \\
M(N,k) & \longrightarrow & S(N,k) & \xrightarrow{\Phi} & M^0(N,k')^\vee.
\end{array}$$

Proof. Let $l\neq p$. Let Y be the smooth proper connected curve over $\overline{\mathbf{F}}_p$ parametrizing pairs $(\phi\colon E\to E',\alpha)$, with ϕ an l-isogeny of elliptic curves over an $\overline{\mathbf{F}}_p$-scheme S and $\alpha\in[\Gamma_1(N)](E/S)$, such that $\phi\circ\alpha$ is injective. There are two maps, S and T (for "source" and "target") from Y to X corresponding to $(\phi\colon E\to E',\alpha)\mapsto(E,\alpha)$ and $(\phi\colon E\to E',\alpha)\mapsto(E',\phi\circ\alpha)$, respectively. Let E/X be the universal curve over X and let $\phi\colon S^*E\to T^*E$ be the universal isogeny over Y. Then ϕ induces an isomorphism $\phi^*\colon T^*\underline{\omega}^{\otimes k}\tilde{\to}S^*\underline{\omega}^{\otimes k}$. For $i\geq 0$ we define a Hecke operator T_l^* on $\mathrm{H}^i(X,\underline{\omega}^{\otimes k})$ by:

$$lT_l^*\colon\quad \mathrm{H}^i(X,\underline{\omega}^{\otimes k}) \xrightarrow{T^*} \mathrm{H}^i(Y,T^*\underline{\omega}^{\otimes k}) \xrightarrow{\phi^*} \mathrm{H}^i(Y,S^*\underline{\omega}^{\otimes k}) \xrightarrow{S_*} \mathrm{H}^i(X,\underline{\omega}^{\otimes k}),$$

where S_* is induced by the trace map $S_*S^*\underline{\omega}^{\otimes k}=\underline{\omega}^{\otimes k}\otimes S_*\mathcal{O}_Y\to\underline{\omega}^{\otimes k}$. For $i=0$ these are the T_l^* defined as before. The exact sequence 7.2.1 is compatible with the T_l^*. Now consider the following diagram:

$$\begin{array}{ccccc}
\mathrm{H}^1(X,\underline{\omega}^{\otimes k-p+1}) & \xrightarrow{\mathrm{SD}} & \mathrm{H}^0(X,\Omega_X\otimes\underline{\omega}^{\otimes p-1-k})^\vee & \xrightarrow{\mathrm{KS}^\vee} & \mathrm{H}^0(X,\underline{\omega}^{\otimes k'}(-\mathrm{cusps}))^\vee \\
\downarrow{\scriptstyle T^*} & & \downarrow{\scriptstyle T_*^\vee} & & \downarrow{\scriptstyle T_*^\vee} \\
\mathrm{H}^1(Y,T^*\underline{\omega}^{\otimes k-p+1}) & \xrightarrow{\mathrm{SD}} & \mathrm{H}^0(Y,\Omega_Y\otimes T^*\underline{\omega}^{\otimes p-1-k})^\vee & \xrightarrow{\mathrm{KS}^\vee_{T^*E}} & \mathrm{H}^0(Y,T^*\underline{\omega}^{\otimes k'}(-\mathrm{cusps}))^\vee \\
\downarrow{\scriptstyle \phi^*} & & \downarrow{\scriptstyle l^k(\mathrm{id}\otimes\phi^{t*})^\vee} & & \downarrow{\scriptstyle l^{k-1}\phi^{t*\vee}} \\
\mathrm{H}^1(Y,S^*\underline{\omega}^{\otimes k-p+1}) & \xrightarrow{\mathrm{SD}} & \mathrm{H}^0(Y,\Omega_Y\otimes S^*\underline{\omega}^{\otimes p-1-k})^\vee & \xrightarrow{\mathrm{KS}^\vee_{S^*E}} & \mathrm{H}^0(Y,S^*\underline{\omega}^{\otimes k'}(-\mathrm{cusps}))^\vee \\
\downarrow{\scriptstyle S_*} & & \downarrow{\scriptstyle S^{*\vee}} & & \downarrow{\scriptstyle S^{*\vee}} \\
\mathrm{H}^1(X,\underline{\omega}^{\otimes k-p+1}) & \xrightarrow{\mathrm{SD}} & \mathrm{H}^0(X,\Omega_X\otimes\underline{\omega}^{\otimes p-1-k})^\vee & \xrightarrow{\mathrm{KS}^\vee} & \mathrm{H}^0(X,\underline{\omega}^{\otimes k'}(-\mathrm{cusps}))^\vee.
\end{array} \tag{7.3.1}$$

We claim that this diagram is commutative. For the upper left square, this follows from $\mathrm{tr}_Y = \mathrm{tr}_X \circ T_*$, where $\mathrm{tr}_X\colon \mathrm{H}^1(X, \Omega_X) \tilde{\to} \overline{\mathbf{F}}_p$ and $\mathrm{tr}_Y\colon \mathrm{H}^1(Y, \Omega_Y) \tilde{\to} \overline{\mathbf{F}}_p$ are the trace maps used in the construction of Serre duality. For the lower left square one uses $\mathrm{tr}_Y = \mathrm{tr}_X \circ S_*$. Consider the middle left square. Let $\langle \cdot, \cdot \rangle_T$ be the natural pairing

$$T^* \underline{\omega}^{\otimes k-p+1} \times \Omega_Y \otimes T^* \underline{\omega}^{\otimes p-1-k} \longrightarrow \Omega_Y$$

and let $\langle \cdot, \cdot \rangle_S$ be the analog of this for S. Then one has:

$$\langle \phi^* \cdot, \cdot \rangle_S = \langle \phi^{t*} \phi^* \cdot, (\mathrm{id} \otimes \phi^{t*}) \cdot \rangle_T = l^{k-p+1} \langle \cdot, (\mathrm{id} \otimes \phi^{t*}) \cdot \rangle_T.$$

The upper and lower right squares are commutative just because the Kodaira-Spencer morphism is compatible with base change. The commutativity of the middle right square follows from the following lemma.

7.4 Lemma. *Let $S \to T$ be any morphism of schemes; let $\phi\colon E \to E'$ be a morphism of elliptic curves over S. Then the following diagram is commutative:*

$$\begin{tikzcd} \omega_{E/S}^{\otimes 2} \arrow[dr, "\mathrm{KS}_{E/S}"'] & & \omega_{E'/S}^{\otimes 2} \arrow[ll, "\phi^*"'] \arrow[dl, "(\deg \phi)\mathrm{KS}_{E'/S}"] \\ & \Omega_{S/T} & \end{tikzcd}$$

Proof. One uses the formula for the Kodaira-Spencer map in terms of the Gauss-Manin connection and the cup product on $\mathrm{H}^*_{\mathrm{DR}}$, compare [10], Prop. 3.15. $\square$

Let T_l^{t*} be the operator $l^{-1} T_* \phi^{t*} S^*$ on $M^0(N, k')$. The commutativity of 7.3.1 implies that $\mathrm{KS}^\vee \circ \mathrm{SD} \circ T_l^* = l^{k-1} T_l^{t*\vee} \circ \mathrm{KS}^\vee \circ \mathrm{SD}$. The formula for T_l^{t*} in the style of 2.2.1 is:

$$(T_l^{t*} f)(E/\overline{\mathbf{F}}_p, \alpha) = \frac{1}{l} \sum_{\phi} \phi^{t*}(f(E'/\overline{\mathbf{F}}_p, \alpha')), \tag{7.4.1}$$

where the sum is over the l-isogenies $\phi\colon E' \to E$ of elliptic curves with $\Gamma_1(N)$-structures. We leave it to the reader to check that $w^* T_l^{t*} = T_l^* w^*$. This finishes the proof of Prop. 7.3. $\square$

7.5 Proof of Theorem 3.4. Let us first note that the functor $\mathrm{Hom}_{\overline{\mathbf{F}}_p}(\cdot, \overline{\mathbf{F}}_p)$, on the category of modules of finite $\overline{\mathbf{F}}_p$-dimension over the polynomial ring $\overline{\mathbf{F}}_p[T_l \mid l \neq p]$, preserves supports. Using this, Theorem 3.4 is a consequence of the existence of the long exact cohomology sequence of 7.1.1 combined with Propositions 7.2 and 7.3, if $N \geq 5$. We are now left with the cases $N = 1$, 2, 3 or 4. If $N = 4$ or $N = 2$ (hence $p \neq 2$) we have $M(N, k) = \mathrm{H}^0(X, \underline{\omega}^{\otimes k})^G$, where $X = X(4)^{\zeta-\mathrm{can}}$ (ζ some 4^{th} root of 1 in $\overline{\mathbf{F}}_p$) and G the subgroup $\{(\begin{smallmatrix} 1 & * \\ 0 & 1 \end{smallmatrix})\}$ or $\{(\begin{smallmatrix} 1+2* & * \\ 0 & 1+2* \end{smallmatrix})\}$ of $\mathrm{SL}_2(\mathbf{Z}/4\mathbf{Z})$. Note that $p \not| \#G$. If $N = 3$ we are in the same situation by taking $X = X(3)^{\zeta-\mathrm{can}}$ and $G = \{(\begin{smallmatrix} 1 & * \\ 0 & 1 \end{smallmatrix})\}$. Finally, if $N = 1$ and $p > 3$ we take $X = X(3)^{\zeta-\mathrm{can}}$ and $G = \mathrm{SL}_2(\mathbf{Z}/3\mathbf{Z})$ or

$X = X(4)^{\zeta-\mathrm{can}}$ and $G = \mathrm{SL}_2(\mathbf{Z}/4\mathbf{Z})$. In all these cases the functor "G-invariants" from $\overline{\mathbf{F}}_p[G]$-modules to $\overline{\mathbf{F}}_p$-vector spaces is exact and for any $\overline{\mathbf{F}}_p[G]$-module V, the canonical map $V^G \to V_G$ from invariants to co-invariants is an isomorphism. We have an exact sequence as in Prop. 7.3:

$$\mathrm{H}^0(X, \underline{\omega}^{\otimes k}) \longrightarrow \oplus_{x \mathrm{s.s.}} \omega_{E_x}^{\otimes k} \xrightarrow{\Phi} \mathrm{H}^0(X, \underline{\omega}^{\otimes k'}(-\mathrm{cusps}))^{\vee} \tag{7.5.1}$$

where Φ is the composition of $\oplus_{x \mathrm{s.s.}} \omega_{E_x}^{\otimes k} \to \mathrm{H}^1(X, \underline{\omega}^{\otimes k-p+1})$ and $\mathrm{KS}^{\vee} \circ \mathrm{SD}$. We let the T_l, $l \neq p$ (and $l \neq 2$ or 3 if $N = 1$), act on the last term in 7.5.1 by $l^{k-1} T_l^{*\vee}$. We let $g \in G$ act on the three terms in 7.5.1 by g^*, g^* and $g^{-1*\vee}$ respectively. The sequence 7.5.1 is then compatible with all these actions. Theorem 3.4 now follows by taking G-invariants, except for $N = 1$ and $p = 2$ or 3. For these remaining two cases we refer to [25], Théorème 3. Let us remark that it is not hard to give a proof in these last two cases that is similar to the proof above (just compute the systems of eigenvalues in the sum of all $S(N, k)$ and show directly that they occur in $\oplus_{k=2}^{p+1} M(N, k)$). □

8 Finiteness at p.

8.1. Let p be a prime, let $\mathbf{F}$ be a finite extension of $\mathbf{F}_p$ and let $r := [\mathbf{F} : \mathbf{F}_p]$. Let $\rho_p \colon G_p \to \mathrm{GL}_2(\mathbf{F})$ be a continuous representation of a decomposition group G_p in $G_{\mathbf{Q}}$ (see §2.4). We suppose that ρ_p is of the form:

$$\rho_p = \begin{pmatrix} \chi\varepsilon_2 & * \\ 0 & \varepsilon_1 \end{pmatrix}, \tag{8.1.1}$$

with ε_1 and ε_2 unramified $\mathbf{F}^*$-valued characters of G_p. Let $V_{\mathbf{Q}_p}$ be the $\mathbf{F}$-vector space scheme over $\mathbf{Q}_p$ defined by ρ_p (see §5.1). Let D be the integral closure of $\mathbf{Z}_p$ in the maximal unramified extension K of $\mathbf{Q}_p$ in $\overline{\mathbf{Q}}_p$.

8.2 Proposition. *The following conditions are equivalent:*

1. *ρ_p is "peu ramifié" (see [26], 2.4.7),*
2. *$V_{\mathbf{Q}_p}$ can be extended to a finite flat $\mathbf{F}$-vector space scheme V_D over D,*
3. *$V_{\mathbf{Q}_p}$ can be extended to a finite flat $\mathbf{F}$-vector space scheme $V_{\mathbf{Z}_p}$ over $\mathbf{Z}_p$,*
4. *$V_{\mathbf{Q}_p}$ can be extended to a finite flat group scheme V_D over D,*
5. *$V_{\mathbf{Q}_p}$ can be extended to a finite flat group scheme $V_{\mathbf{Z}_p}$ over $\mathbf{Z}_p$.*

If these conditions are satisfied we say that ρ_p is finite.

Proof. The implications (3) ⇒ (2) ⇒ (4) and (3) ⇒ (5) ⇒ (4) are obvious. Now suppose that (4) holds. Let V_D be the *maximal* (see [21], 2.2.3) extension of V_K to a finite flat group scheme over D. Then the actions of $\mathbf{F}$ and $\mathrm{Gal}(K/\mathbf{Q}_p)$ on V_K extend uniquely to V_D. This shows that V_D extends V_K as $\mathbf{F}$-vector space scheme and that the descent data from K to $\mathbf{Q}_p$ for V_K extend to descent data for V_D. We have now shown that (2), (3), (4) and (5) are all equivalent.

8.3 Lemma. *Let W be an $\mathbf{F}$-vector space scheme over a scheme S. For any extension $0 \to W \to V \to \mathbf{F}_S \to 0$ of $\mathbf{F}$-vector space schemes (i.e., the sequence is an exact sequence of* f.p.p.f.*-sheaves on S) and an element a in $\mathbf{F}$ let V_a be the inverse image in V of the section a of $\mathbf{F}_S$. Then the assignment $V \mapsto V_1$ induces an equivalence between the category of extensions of $\mathbf{F}_S$ by W and the category of torsors, on S, for the underlying additive group scheme of W.*

Proof. For V such an extension and $a \in \mathbf{F}^*$, multiplication by a on V induces an isomorphism $V_1[a] \to V_a$, where $V_1[a]$ is the torsor obtained from V_1 by twisting the action of W by multiplication by a on W. This shows how to reconstruct the extension from V_1. □

Because of 8.1.1, V_K sits in an exact sequence:

$$0 \longrightarrow \mathbf{F}_K^D \longrightarrow V_K \longrightarrow \mathbf{F}_K \longrightarrow 0, \tag{8.3.1}$$

where $\mathbf{F}_K^D$ is the Cartier dual of $\mathbf{F}_K$. Let us choose an $\mathbf{F}_p$-basis of $\mathbf{F}$, i.e., an isomorphism $\oplus_{i=1}^r \mathbf{F}_p \to \mathbf{F}$ of additive groups. By Lemma 8.3, extensions as 8.3.1 are classified up to isomorphism by $\mathrm{H}^1_{\text{f.p.p.f.}}(K, \mathbf{F}^D) \cong \oplus_{i=1}^r \mathrm{H}^1_{\text{f.p.p.f.}}(K, \mu_p)$. The long exact cohomology sequence arising from $0 \to \mu_p \to \mathrm{G}_m \to \mathrm{G}_m \to 0$, plus the fact that $\mathrm{Pic}(K) = \{0\}$, show that $\mathrm{H}^1_{\text{f.p.p.f.}}(K, \mu_p) = K^*/K^{*p}$. Under this identification, $[x] \in K^*/K^{*p}$ corresponds to the μ_p-torsor of p^{th} roots of x.

In the same way, extensions of $\mathbf{F}_D$ by $\mathbf{F}_D^D$ are classified by $\oplus_{i=1}^r D^*/D^{*p}$. If V_K corresponds to the element $([x_1], \ldots, [x_r])$, then V_K can be extended to an extension of $\mathbf{F}_D$ by $\mathbf{F}_D^D$ if and only if ρ_p is "peu ramifié", i.e., $v(x_i) \equiv 0$ (p) for all i, where v is the valuation on K. For $p > 2$ any extension of V_K to V_D as in (2) is automatically an extension of $\mathbf{F}_D$ by $\mathbf{F}_D^D$ (see [21], 3.3.3). If $p = 2$ and an extension V_D as in (2) exists, the functoriality of Ext and a non-zero morphism $\mathbf{F}_D \to \mathbf{F}_D^D$ give an extension V_D of $\mathbf{F}_D$ by $\mathbf{F}_D^D$. This completes the proof of Proposition 8.2. □

8.4. Let ρ_p and $V_{\mathbf{Q}_p}$ be as in §8.1. The fact that the extension 8.3.1 comes from an extension

$$0 \longrightarrow V''_{\mathbf{Q}_p} \longrightarrow V_{\mathbf{Q}_p} \longrightarrow V'_{\mathbf{Q}_p} \longrightarrow 0,$$

with $V'_{\mathbf{Q}_p}$ and $V''_{\mathbf{Q}_p}$ twists of $\mathbf{F}^D_{\mathbf{Q}_p}$ and $\mathbf{F}_{\mathbf{Q}_p}$, gives restrictions on the torsor $V_{K,1}$ corresponding to 8.3.1. To be precise, one finds that

$$\sigma^* V_{K,1} \;\cong\; V_{K,1}[\lambda_1 \lambda_2^{-1}], \tag{8.4.1}$$

where σ is the automorphism of $\mathrm{Spec}(D)$ with $\sigma^{\#}(x) \equiv x^p \bmod p$ and $\lambda_i = \varepsilon_i(\sigma^{\#})$, $i \in \{1, 2\}$. Let $\lambda = \lambda_1 \lambda_2^{-1} \in \mathbf{F}^*$ and let $X^s + a_{n-1} X^{s-1} + \cdots + a_0$ be its minimal polynomial over $\mathbf{F}_p$. Let e_j, $1 \le j \le r/s$, be an $\mathbf{F}_p(\lambda)$-basis of $\mathbf{F}$. Then $\mathbf{F}$ has an $\mathbf{F}_p$-basis $\lambda^i e_j$, with $0 \le i < s$ and $1 \le j \le r/s$. Now $\mathbf{F} = \oplus_{j=1}^{r/s} \mathbf{F}_p(\lambda) e_j$ is a decomposition of $\mathbf{F}$ into λ-stable $\mathbf{F}_p$-subspaces of $\mathbf{F}$. Let $V_{K,1}$ correspond to $([x_{i,j}])_{i,j} \in \oplus_{i,j} K^*/K^{*p}$. Then $V_{K,1}[\lambda]$ and $\sigma^* V_{K,1}$ correspond to $([y_{i,j}])_{i,j}$ and $([\sigma^{\#} x_{i,j}])_{i,j}$ respectively, where $y_{i,j} = x_{i+1,j}$ if $i < s-1$ and $y_{s-1,j} = x_{0,j}^{-a_0} x_{1,j}^{-a_1} \cdots x_{s-1,j}^{-a_{s-1}}$. Hence we have the equations:

$$\forall j: \quad [\sigma^{\#} x_{i,j}] = [x_{i+1,j}] \;\text{ if } i < s-1 \quad \text{and } [\sigma^* x_{s-1,j}] = [x_{0,j}^{-a_0} x_{1,j}^{-a_1} \cdots x_{s-1,j}^{-a_{s-1}}]. \tag{8.4.2}$$

Now let $\alpha_{i,j} = v([x_{i,j}]) \in \mathbf{F}_p$. Then equations 8.4.2 give:

$$\forall j: \quad \alpha_{0,j} = \alpha_{1,j} = \cdots = \alpha_{s-1,j} \quad \text{and} \quad \alpha_{0,j}(1 + a_{s-1} + \cdots + a_0) = 0. \tag{8.4.3}$$

From this we conclude that either $\alpha_{i,j} = 0$ for all (i,j) (i.e., ρ_p is "peu ramifié") or $1 + a_{s-1} + \cdots + a_0 = 0$ (i.e., $\lambda = 1$ and ε is trivial).

8.5 Proposition. *(Compare [26], §2.4, Remarques.) If ρ_p is "très ramifié" then $\varepsilon_1 = \varepsilon_2$, all x_i in [26], (2.4.6) can be chosen in $\mathbf{Q}_p^*$ and the integer m in [26], (2.4.6), equals 1 or 2 if $p > 2$ and 1 or 2 or 3 if $p = 2$.* □

Suppose now that ρ_p is "peu ramifié", i.e., $\alpha_{i,j} = 0$ for all i and j. Let $U := \{x \in D^* \mid x \equiv 1\ (p)\}$ and let $\tau\colon \overline{\mathbf{F}}_p \to D$ be the Teichmüller lift. Then we have an isomorphism $\mathbf{Z} \times U \times \overline{\mathbf{F}}_p^{\,*} \to K^*$, sending (n, u, a) to $p^n u \tau(a)$. Moreover, if $p > 2$, we can view the $[x_{i,j}]$ as elements of $\overline{\mathbf{F}}_p$ via the isomorphism

$$D^*/D^{*p} \;=\; U/U^p \stackrel{\sim}{\longrightarrow} \overline{\mathbf{F}}_p, \quad [1 + pa] \mapsto a \bmod p. \tag{8.5.1}$$

The equations 8.4.2 then give the following proposition.

8.6 Proposition. *Suppose that $p > 2$ and that ρ_p is "peu ramifié". Let $\lambda = (\varepsilon_1\varepsilon_2^{-1})(\mathrm{Frob}_p)$, let $s = [\mathbf{F}_p(\lambda) : \mathbf{F}_p]$ and let n be the order of λ in $\mathbf{F}^*$. Let $f \in \mathbf{F}_p[X]$ be the minimal polynomial of λ. Then the $[x_{i,j}] \in D^*/D^{*p}$, defined in §8.4, satisfy $[x_{i,j}] = (\sigma^{\#})^i [x_{0,j}]$ for all j, and the image y of any of them under the isomorphism 8.5.1 satisfies $(f(\mathrm{Frob}_p))(y) = 0$. This last equation has all p^s solutions in $\mathbf{F}_{p^n}$ since f divides $X^n - 1$ in $\mathbf{F}_p[X]$. Conversely, any set of $[x_{i,j}]$ satisfying the equations just mentioned comes from a ρ_p that is "peu ramifié".* □

9 Multiplicity one results.

9.1. In this section we will extend some results of Mazur, Ribet and Gross that state that in certain cases the Galois representation ρ_f occurs with multiplicity one in a certain subspace of the p-torsion of the jacobian of the modular curve that is used to construct ρ_f. These results are not used in this article.

9.2 Theorem. *Let f be a cuspidal eigenform of type (N, k, ε), defined over $\overline{\mathbf{F}}_p$, with $p \nmid N$ and $2 \leq k \leq p+1$. Let $J_{\mathbf{Q}}$ be the jacobian of the curve $X_1(pN)_{\mathbf{Q}}$ if $k > 2$ and let $J_{\mathbf{Q}}$ be the jacobian of $X_1(N)_{\mathbf{Q}}$ if $k = 2$. Let H be the subring of $\mathrm{End}(J_{\mathbf{Q}})$ generated by all T_l and $\langle a \rangle_N$, and $\langle b \rangle_p$ if $k > 2$. Let m be the maximal ideal of H corresponding to f, and let $\mathbf{F} = H/m \subset \overline{\mathbf{F}}_p$. Suppose that the representation $\rho_f\colon G_{\mathbf{Q}} \to \mathrm{GL}_2(\overline{\mathbf{F}}_p)$ is irreducible. Then $J_{\mathbf{Q}}(\overline{\mathbf{Q}})[m]$ is an $\mathbf{F}$-vector space of dimension two in each of the following cases:*

1. *$2 \leq k < p$,*
2. *$k = p$ and $a_p^2 \neq \varepsilon(p)$, where $T_p^* f = a_p f$,*
3. *$k = p$ and ρ_f is ramified at p,*

4. $k = p{+}1$ *and there is no form* g *of type* $(N, 2, \varepsilon)$ *with* $\rho_g \cong \rho_f$.

Proof. Case (4) is dealt with in [18], Main Theorem, for $\varepsilon = 1$, but their proof works in general. Cases (1) and (2) are dealt with in [10], Thm. 12.10(1), assuming that $a_p \neq 0$ and that f is a newform. The assumption that f is a newform is not used in the proof of [10], Thm. 12.10(1).

It remains to prove (1) and (2) in the case $a_p = 0$, and (3). To do this, let us first take care of all cases with $a_p = 0$. Suppose that $a_p = 0$ and that $2 < k \leq p$. Let G, $\overline{G}$, $\overline{G}_I$ etc. be as in §6.8 and Prop. 6.9. The arguments that follow are motivated by pages 116–118 in [17]. Let $M(-)$ denote the contravariant Dieudonné module functor of [19]. Note that since we work with group schemes over $\mathbf{F}_p$, the Dieudonné ring $\mathbf{Z}_p[F,V]$ is commutative. Let $M = M(J_E[m])$; the reason to work with J_E and not with J_I will be explained below. By Prop. 6.9 we have $M \cong \oplus_1^d M(\overline{W}(0))$, where $M(\overline{W}(0))$ is a one-dimensional $\mathbf{F}$-vector space on which F and V act by zero. Since $M(-)$ is contravariant and exact we have $M = M(J_E[p])/mM(J_E[p])$. According to Cor. 5.11 in [19], there is a functorial isomorphism

$$M(J_E[p]) \xrightarrow{\sim} \mathbf{H}^1_{\mathrm{DR}}(J_E).$$

Under this isomorphism, the Hodge filtration

$$0 \longrightarrow \mathrm{H}^0(J_E, \Omega^1) \longrightarrow \mathbf{H}^1_{\mathrm{DR}}(J_E) \longrightarrow \mathrm{H}^1(J_E, \mathcal{O}) \longrightarrow 0$$

corresponds to the filtration

$$0 \longrightarrow M(J_E[F]) \longrightarrow M(J_E[p]) \longrightarrow M(J_E[V]) \longrightarrow 0.$$

Applying $-\otimes_H \mathbf{F}$ to the isomorphism $\mathrm{H}^1(J_E, \mathcal{O}) \xrightarrow{\sim} M(J_E[V]) = M(J_E[p])/VM(J_E[p])$ gives an isomorphism

$$\mathrm{H}^1(J_E, \mathcal{O}) \otimes_H \mathbf{F} \xrightarrow{\sim} M/VM \;\cong\; \oplus_1^d (M(\overline{W}(0))/VM(\overline{W}(0))) \;=\; \oplus_1^d M(\overline{W}(0)).$$

This isomorphism shows that $d{=}1$ if and only if $\mathrm{H}^1(J_E, \mathcal{O}) \otimes_H \mathbf{F}$ is a one-dimensional $\mathbf{F}$-vector space. Serre's duality on E gives an isomorphism of H-modules between the $\mathbf{F}_p$-dual of $\mathrm{H}^1(J_E, \mathcal{O}) \otimes_H \mathbf{F}$ and $\mathrm{H}^0(E, \Omega^1_E)[m]$, where $h \in H$ acts on $\mathrm{H}^0(E, \Omega^1_E) = \mathrm{H}^0(J_E, \Omega^1_{J_E})$ by h^{t*}. In particular, T_p acts on $\mathrm{H}^0(E, \Omega^1_E)$ by $F^{t*} = V^*$. This explains why we work with J_E; working with J_I results in T_p acting on $\mathrm{H}^0(I, \Omega^1_I)$ by $V^{t*} = F^* = 0$. Lemma 9.3 finishes the proof for $a_p{=}0$ and $2 < k \leq p$.

Suppose now that $a_p{=}0$ and $k{=}2$. Let $X := X_1(N)_{\mathbf{F}_p}$ and let J be the jacobian of X. Let $M = M(\overline{G}[m]) = M(J[m])$. By Prop. 6.9 we have $M \cong \oplus_1^d M(\overline{W})$, where $M(\overline{W})$ is a two-dimensional $\mathbf{F}$-vector space on which F and V have rank one. As above we get an isomorphism

$$\mathrm{H}^1(J, \mathcal{O}) \otimes_H \mathbf{F} \xrightarrow{\sim} M/VM \;\cong\; \oplus_1^d M(\overline{W})/VM(\overline{W}).$$

We conclude that $d{=}1$ if and only if $\mathrm{H}^0(X, \Omega)[m]$ is a one-dimensional $\mathbf{F}$-vector space, where $h \in H$ acts on $\mathrm{H}^0(X, \Omega)[m]$ by h^{t*}. Let $\zeta \in \overline{\mathbf{F}}_p$ be a primitive N^{th} root of unity and let w be the automorphism of $X_{\overline{\mathbf{F}}_p}$ as defined in §6.4. Then we have $h^t = whw^{-1}$ for all h in H. Hence

$w^*\colon \mathrm{H}^0(X,\Omega) \to \mathrm{H}^0(X,\Omega)$ is an isomorphism of H-modules, with $h\in H$ acting on source and target by h^{t*} and h^*, respectively. The result now follows from the "standard" formulas for the action of the T_l^* on q-expansions.

To prove that $d{=}1$ in case (3), we may assume that $a_p{\neq}0$. Let W be an $\mathbf{F}$-vector space scheme over $\mathbf{Q}$ such that $W(\overline{\mathbf{Q}})$ gives ρ_f. Let $J_{\mathbf{Q}}$ be the jacobian over $\mathbf{Q}$ used in the construction of ρ_f. Let H and m be as before. By [2], Thm's. 1 and 2, we have $J_{\mathbf{Q}}[m]\cong\oplus_1^d W$. From the construction of ρ_f as in [10], §11–12 (compare our §6) it follows that for any embedding $W\hookrightarrow J_{\mathbf{Q}}[m]$, the Zariski closure $W_{\mathbf{Z}_p[\zeta_p]}$ of W in $J_{\mathbf{Z}_p[\zeta_p]}$ is finite and flat over $\mathbf{Z}_p[\zeta_p]$. By Thm. 2.5 such a $W_{\mathbf{Z}_p[\zeta_p]}$ sits in a short exact sequence

$$0\longrightarrow W'\longrightarrow W_{\mathbf{Z}_p[\zeta_p]}\longrightarrow W''\longrightarrow 0, \tag{9.2.1}$$

where W' and W'' are "one-dimensional" $\mathbf{F}$-vector space schemes over $\mathbf{Z}_p[\zeta_p]$. By Raynaud's result [21], §3.3.5, W' and W'' can only be etale or multiplicative. The assumption that ρ_f is ramified at p means that the exact sequence 9.2.1, when restricted to $\mathbf{Q}_p(\zeta_p)$, does not split. Now one remarks that this means that W'' is etale and W' is multiplicative. Indeed, an extension of etale by etale is etale, hence unramified; an extension of multiplicative by multiplicative splits (consider the Cartier duals); an extension of multiplicative by etale splits (use the connected component of identity), compare [26], page 190. It follows that $J_{\mathbf{Z}_p[\zeta_p]}[m]^{\mathrm{e}}$ is an $\mathbf{F}$-vector space scheme of dimension d. But according to [10], Thm.'s 12.8 and 12.9 the dimension of the etale part is one, hence $d=1$. □

9.3 Lemma. *Let H act on $\mathrm{H}^0(E,\Omega_E)=\mathrm{H}^0(J_E,\Omega_{J_E})$ as above: $h\in H$ acts by h^{t*}. Let $m\subset H$ be a maximal ideal with $\mathbf{F}=H/m$ of characteristic p. Then $\mathrm{H}^0(E,\Omega_E)[m]$ is a $\mathbf{F}$-vector space of dimension at most one.*

Proof. (Compare [17], Chapter II, Prop. 9.3.) Let $\lambda\colon\mathbf{F}\to\overline{\mathbf{F}}_p$. For l prime let $a_l=\lambda(T_l)$. Let $\varepsilon\colon(\mathbf{Z}/N\mathbf{Z})^*\to\overline{\mathbf{F}}_p^{\,*}$ be given by $\varepsilon(a)=\lambda(\langle a\rangle_N)$. Let $d\in\mathbf{Z}/(p{-}1)\mathbf{Z}$ be determined by $\lambda(\langle b\rangle_p)=b^d$ for all $b\in\mathbf{F}_p^{\,*}$. What we must show is that the intersection, in $\mathrm{H}^0(E_{\overline{\mathbf{F}}_p},\Omega_{E_{\overline{\mathbf{F}}_p}})$, of the kernels of $T_l^{t*}-a_l$, $\langle a\rangle_N^{t*}-\varepsilon(a)$ and $\langle b\rangle_p^{t*}-b^d$ for all l prime, $a\in(\mathbf{Z}/N\mathbf{Z})^*$ and $b\in\mathbf{F}_p^{\,*}$ has dimension at most one over $\overline{\mathbf{F}}_p$. To do this, we consider q-expansions at the cusp $(\mathrm{Tate}(q^{pN}),q^N,q^p)$ of E. Let ω be a differential form on $E_{\overline{\mathbf{F}}_p}$ and let $\omega(q)=(\sum b_nq^n)dq/q$ be its q-expansion. Suppose that $\langle a\rangle_N^{t*}\omega=\varepsilon(a)\omega$ and $\langle b\rangle_p^{t*}\omega=b^d\omega$ for all a and b. Then one finds the following formulas (compare [14], §1.11):

$$(T_l^{t*}\omega)(q)=\begin{cases}(\sum_{n\geq1}(b_{nl}+\varepsilon(l)^{-1}l^{1-d}b_{n/l})q^n)dq/q & \text{if } l\nmid pN\\ (\sum_{n\geq1}b_{nl}q^n)dq/q & \text{if } l|pN\end{cases} \tag{9.3.1}$$

These formulas clearly show what we need.

Another way to prove this lemma is to use an automorphism w of $E_{\overline{\mathbf{F}}_p}$ with the property that $T_l^t=wT_lw^{-1}\langle l^{-1}\rangle_p$ for all $l\neq p$ (compare §6.4). Then one can use the "standard" formulas for the action of the T_l on q-expansions. □

References

[1] A. Ash and G. Stevens. *Modular forms in characteristic l and special values of their L-functions.* Duke Mathematical Journal 53, No. 3, (1986).

[2] N. Boston, H.W. Lenstra and K.A. Ribet. *Quotients of group rings arising from two-dimensional representations.* C.R. Acad. Sci. Paris, t. 312, Série I, p. 323–328 (1991).

[3] H. Carayol. *Sur les représentations l-adiques associées aux formes modulaires de Hilbert.* Ann. Sci. Ecole Norm. Sup. (4) 19, 409–468 (1986).

[4] R.F. Coleman. *A p-adic Shimura isomorphism and p-adic periods of modular forms.* Preprint.

[5] R.F. Coleman and J.F. Voloch. *Companion forms and Kodaira-Spencer theory.* To appear.

[6] G. Cornell, J.H. Silverman. *Arithmetic geometry.* Springer-Verlag (1986).

[7] P. Deligne and D. Mumford. *The irreducibility of the space of curves of given genus.* Publications Mathématiques de l'I.H.E.S. 36, 75–125, (1969).

[8] P. Deligne and M. Rapoport. *Les schémas de modules des courbes elliptiques.* In Modular Functions of One Variable II. Springer Lecture Notes in Mathematics 349 (1973).

[9] P. Deligne and J.-P. Serre. *Formes modulaires de poids 1.* Ann. Sci. Ecole Norm. Sup. (4) 7, 507–530 (1974).

[10] B.H. Gross. *A tameness criterion for Galois representations associated to modular forms (mod p).* Duke Mathematical Journal 61, No. 2, (1990).

[11] B. Jordan and R. Livné. *Conjecture "epsilon" for weight $k > 2$.* Bull. AMS 21, 51–56 (1989).

[12] A. Grothendieck. *Séminaire de géométrie algébrique.* Springer Lecture Notes in Mathematics 151, 152, 153, 224, 225, 269, 270, 288, 305, 340, 589.

[13] N. Jochnowitz. *The local components of the Hecke algebra mod l.* Transactions of the AMS 270, 253–267 (1982).

[14] N.M. Katz. *p-adic properties of modular schemes and modular forms.* In Modular Functions of One Variable III. Springer Lecture Notes in Mathematics 350 (1973).

[15] N.M. Katz. *A result on modular forms in characteristic p.* Springer Lecture Notes in Mathematics 601, 53–61 (1976).

[16] N.M. Katz and B. Mazur. *Arithmetic moduli of elliptic curves.* Annals of Mathematics Studies 108, Princeton University Press (1985).

[17] B. Mazur. *Modular curves and the Eisenstein ideal.* Publications Mathématiques de l'I.H.E.S. 47 (1977).

[18] B. Mazur and K.A. Ribet. *Two-dimensional representations in the arithmetic of modular curves.* To appear in Astérisque.

[19] T. Oda. *The first De Rham cohomology group and Dieudonné modules.* Ann. Sc. Ecole Norm. Sup. (4) 2, 63–135 (1969).

[20] M. Raynaud. *Spécialisation du foncteur de Picard.* Publications Mathématiques de l'I.H.E.S. 38 (1970).

[21] M. Raynaud. *Schémas en groupes de type* $(p, ..., p)$. Bull. Soc. Math. France 102, 241-280 (1974).

[22] K.A. Ribet. *On modular representations of* $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ *arising from modular forms.* Invent. Math. 100, 431–476 (1990).

[23] G. Robert. *Congruences entres séries d'Eisenstein, dans le cas supersingulier.* Inv. Math. 61, 103–158 (1980).

[24] J.-P. Serre. *Une interprétation des congruences relatives à la fonction* τ *de Ramanujan.* Séminaire Delange-Pisot-Poitou 1967/68, 14. (Oeuvres 80.)

[25] J.-P. Serre. *Valeurs propres des opérateurs de Hecke modulo* l. Journées arithmétiques Bordeaux, Astérisque 24–25, 109–117 (1975). (Oeuvres 104.)

[26] J.-P. Serre. *Sur les représentations de degré 2 de* $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$. Duke Mathematical Journal 54, No. 1, (1987).

[27] J.-P. Serre. Résumé des cours au Collège de France, 1987–1988.

[28] L. Szpiro. *Séminaire sur les pinceaux arithmétiques: la conjecture de Mordell.* Astérisque 127 (1985).

[29] J. Tate. *"p-divisible groups."* Proceedings of a conference on local fields at Driebergen, Springer-Verlag, 158–184 (1967).

[30] D.L. Ulmer. *L-functions of universal curves over Igusa curves.* American Journal of Mathematics 112, 687–712 (1990).

Mathematisch Instituut Budapestlaan
Postbus 80.010
3508 TA Utrecht
The Netherlands.