Canadian Mathematical Society
Conference Proceedings
Volume 17, 1995

Ribet's Theorem: Shimura-Taniyama-Weil implies Fermat

DIPENDRA PRASAD

1. Introduction

The aim of these notes is to give some idea on the steps involved, and some of the proofs, in the work of Ribet which implies Fermat's Last "Theorem" assuming the Shimura-Taniyama-Weil conjecture for semi-stable elliptic curves. Except perhaps for remarks 3.2, 3.4, and proposition 3.6 nothing contained here is new, and in fact these notes are incomplete at various places for which we refer the reader to the original sources, cf. [**R3**], [**R6**]. For other expositions to this work of Ribet, we refer to the Bourbaki talk of Oesterlé [**O**], and the article [**R5**] by Ribet himself.

Let $S_k(\Gamma_0(N), \varepsilon)$ be the space of weight k cusp forms on the Poincaré upper half plane for the group $\Gamma_0(N)$ of Nebentypus character $\varepsilon : (\mathbf{Z}/N)^* \to \mathbf{C}^*$, where $\Gamma_0(N)$ is

$$\Gamma_0(N) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL(2, \mathbf{Z}) \mid \ c \equiv 0 \bmod N \right\}.$$

When the character ε is trivial, it is omitted from the notation in $S_k(\Gamma_0(N), \varepsilon)$.

A function f in $S_k(\Gamma_0(N), \varepsilon)$ has Fourier expansion

$$f(z) = \sum_{n=1}^{\infty} a_n e^{2\pi i n z}.$$

The space $S_k(\Gamma_0(N))$ has a $\mathbf{Z}$-structure coming from the Fourier expansion: If $S_k(\Gamma_0(N); \mathbf{Z})$ denotes the space of cusp forms whose Fourier expansion $f(z) = \sum_{n \geq 1} a_n e^{2\pi i n z}$ has all the $a_n \in \mathbf{Z}$, then $S_k(\Gamma_0(N)) = S_k(\Gamma_0(N); \mathbf{Z}) \otimes \mathbf{C}$. This

1991 *Mathematics Subject Classification.* Primary 11F80, 11G18; Secondary 11F11, 11F33.

The author gratefully acknowledges the generous help received from Chandrashekhar Khare in understanding this work of Ribet, and other related works. The author thanks the referee for pointing out several errors in the manuscript and for making other remarks to improve the exposition. The author also thanks K. Ribet for looking at the paper and offering several suggestions.

$\mathbf{Z}$-structure on the space of cusp forms allows one to define cusp forms with coefficients in any commutative ring R, denoted $S_k(\Gamma_0(N); R)$ by $S_k(\Gamma_0(N); R) = S_k(\Gamma_0(N); \mathbf{Z}) \otimes R$. In particular, one has the concept of modular forms with trivial Nebentypus character on $\Gamma_0(N)$ with coefficients in a finite field. For general Nebentypus character ε, if $\mathbf{Z}[\varepsilon]$ denotes the subring of $\mathbf{C}$ generated by the values of ε, then $S_k(\Gamma_0(N), \varepsilon)$ has a $\mathbf{Z}[\varepsilon]$ structure again coming from the Fourier expansion.

For each integer $n \geq 1$, there are Hecke operators T_n acting on the space $S_k(\Gamma_0(N), \varepsilon)$. For primes $p \not| N$, these are defined by

$$T_p f = \sum_{n \geq 1} a_{np} e^{2\pi i n z} + \varepsilon(p) p^{k-1} \sum_{n \geq 1} a_n e^{2\pi i n p z} \tag{1.1}$$

and for $p \mid N, T_p$ is defined by :

$$T_p f = \sum_{n \geq 1} a_{np} e^{2\pi i n z}. \tag{1.2}$$

The Hecke operators T_n for general n can be defined in terms of T_p for p prime, using the following formal identity:

$$\sum_{n=1}^{\infty} \frac{T_n}{n^s} = \prod_{p \not| N} \frac{1}{[1 - \frac{T_p}{p^s} + \frac{\varepsilon(p) p^{k-1}}{p^{2s}}]} \prod_{p | N} \frac{1}{[1 - \frac{T_p}{p^s}]}. \tag{1.3}$$

The Hecke operators T_n preserve the space of modular forms with coefficients in a ring R whenever R contains the value group of the character $\varepsilon : (\mathbf{Z}/N)^* \to \mathbf{C}^*$. The formal q-expansion $\sum_{n=1}^{\infty} T_n q^n$ is an eigenfunction of all the Hecke operators T_n with eigenvalue T_n.

Let $f = \sum_{n \geq 1} a_n e^{2\pi i n z}$ be an eigenfunction of all the Hecke operators T_n. Then $a_1 \neq 0$, and therefore we can assume $a_1 = 1$ after scaling. In such a situation, the Fourier coefficients a_n belong to the ring of integers $\mathcal{O}_K$ of a finite extension field K of $\mathbf{Q}$. It is a theorem of Deligne that corresponding to such an eigenform f, and a finite place λ of K, there is a representation

$$\rho_{f,\lambda} : \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathrm{GL}(2, K_\lambda) \tag{1.4}$$

where K_λ is the completion of K at the place λ. The representation $\rho_{f,\lambda}$ is irreducible, and is unique up to isomorphism. It is unramified outside the places dividing N and λ, and has the property that

$$\begin{aligned} \mathrm{tr}(\rho_{f,\lambda}(\mathrm{Frob}_p)) &= a_p \\ \det(\rho_{f,\lambda}(\mathrm{Frob}_p)) &= \varepsilon(p) p^{k-1}, \end{aligned} \tag{1.5}$$

where $\mathrm{Frob}_p \in \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ is a Frobenius element at p.

Conjugating by a matrix in $\mathrm{GL}(2, K_\lambda)$, one can assume that the image of $\rho_{f,\lambda}$ lands inside $\mathrm{GL}(2, \mathcal{O}_{K,\lambda})$. Reducing this representation with values in $\mathrm{GL}(2, \mathcal{O}_{K,\lambda})$ modulo λ, we get a representation of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ with finite image:

$$\overline{\rho}_{f,\lambda} : \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to GL(2, \mathcal{O}_K/\lambda). \tag{1.6}$$

The representation $\overline{\rho}_{f,\lambda}$ is well defined up to semi-simplification and depends only on the modular form f modulo λ.

Representations of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ with values in $\mathrm{GL}(2,\mathbf{F})$ where $\mathbf{F}$ is a finite field, which arise as $\overline{\rho}_{f,\lambda}$ for some eigen cusp form f of weight k and level N are called modular of weight k and level N (as there is no uniqueness of f, nor of (k,N), when we speak of a modular representation of weight k and level N, it just means that it arises from some eigen cusp form of weight k and level N). It is easy to see, cf. Lemma 6.11 of [**DS**], that the Hecke eigenvalues of an eigen cusp form with values in a finite field can be lifted to Hecke eigenvalues of an eigen cusp form with values in a number field, and therefore to an eigen cusp form with values in a finite field $\mathbf{F}$, one can associate a representation of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ with values in $GL(2,\mathbf{F})$ which is unique up to semi-simplification.

The representations $\overline{\rho}_{f,\lambda}$ have the property that $\det\rho(c) = -1$ where c is a complex conjugation in $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$. Conversely, it has been conjectured by Serre [**S**] that any irreducible representation ρ of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ with values inside $\mathrm{GL}(2,\mathbf{F})$ where $\mathbf{F}$ is a finite field such that $\det\rho(c) = -1$ arises from a cusp form in $S_k(\Gamma_0(N),\varepsilon)$ for some (k,N,ε) which is an eigenform of all the Hecke operators. Moreover, Serre gives an explicit procedure to determine a value of (k,N,ε) from the representation $\rho : \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to GL(2,\mathbf{F})$. The value (k,N,ε) predicted by Serre has the property that if ρ is also associated to a modular form on $S_{k'}(\Gamma_0(N'),\varepsilon')$, with $(\ell,N') = 1$ where ℓ is the residue characteristic of $\mathbf{F}$ and $k' \geq 2$, then $k' \geq k$ and $N \mid N'$. As a consequence of this conjecture of Serre, if ρ arises from a cusp form, then it arises from one of predicted level and weight. Ribet's theorem – which is the aim of these notes – answers a question of this form, a special case of which was earlier handled by B. Mazur. We now come to the statement of these theorems of Mazur and Ribet.

Let $\ell \geq 3$ be a prime number, and $\rho : \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to GL(2,\mathbf{F})$ be an irreducible representation of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ with values in $\mathrm{GL}(2,\mathbf{F})$ for $\mathbf{F}$ a finite field of characteristic ℓ which is modular of weight 2 and level N and trivial Nebentypus character. Since $\det\rho(c) = -1$ for c a complex conjugation in $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$, the eigenvalues of $\rho(c)$ are ± 1. It follows that such a representation is absolutely irreducible, and can be defined over the subfield of $\mathbf{F}$ containing $\mathrm{tr}\rho(x)$ for $x \in \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$. We assume that this subfield is $\mathbf{F}$ itself.

A representation $\rho : \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathrm{GL}(2,\mathbf{F})$ is said to be finite at a prime p of $\mathbf{Q}$ if there exists a finite, flat, $\mathbf{F}$-vector space scheme H over $\mathbf{Z}_p$ such that the representation of $\mathrm{Gal}(\overline{\mathbf{Q}}_p/\mathbf{Q}_p)$ given by the $\mathbf{F}$-vector space $H(\overline{\mathbf{Q}}_p)$ is isomorphic to the restriction of ρ to $\mathrm{Gal}(\overline{\mathbf{Q}}_p/\mathbf{Q}_p)$. If $\mathbf{F}$ is of characteristic $\ell \neq p$, then ρ is finite at p if and only if its restriction to $\mathrm{Gal}(\overline{\mathbf{Q}}_p/\mathbf{Q}_p)$ is unramified.

THEOREM 1.7. *Let $\rho : Gal(\overline{\mathbf{Q}}/\mathbf{Q}) \to GL(2,\mathbf{F})$ where $\mathbf{F}$ is a finite field of characteristic ℓ be an irreducible representation which is modular of weight 2, level N and trivial Nebentypus character. Let p be a prime dividing N such that p^2 does not divide N, and assume that the representation ρ is finite at p. Then*

if either

(1) $p \not\equiv 1 \bmod \ell$, *or*

(2) ℓ *is prime to* N,

then ρ *is modular of level* N/p.

REMARK 1.8. The above theorem basically says that if $f \in S_2(\Gamma_0(N))$ is an eigenform of Hecke operators, then under conditions of flatness at $p \mid N$ on the Galois representation attached to f modulo λ, there is $g \in S_2(\Gamma_0(N/p))$ which is also an eigenform of the Hecke operators such that if $f = \sum_{n \geq 1} f_n e^{2\pi i n z}$, and $g = \sum_{n \geq 1} g_n e^{2\pi i n z}$ with f_n, g_n belonging to the ring of integers $\mathcal{O}_k$ of a number field k with prime ideal λ, with $f_q \equiv g_q \bmod \lambda$ for all but finitely many primes q.

REMARK 1.9. Ribet has strengthened theorem 1.7 above in [**R6**] by only assuming l coprime to p (instead of l coprime to N in the version here), and a recent preprint of Diamond [**Di2**] has generalised it to non-trivial Nebentypus characters.

REMARK 1.10. Carayol has strengthened theorem 1.7 above in [**Ca**] to prove that if ρ comes from a modular representation of weight 2, level N which is coprime to ℓ, and trivial Nebentypus, then it also comes from a modular representation of weight 2, level M which is the Artin conductor of ρ, cf. [**S**], and trivial Nebentypus.

REMARK 1.11. It is a theorem of Langlands and Carayol that the restriction of the λ-adic representation $\rho : \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to GL(2, \mathcal{O}_{K,\lambda})$ associated to a new form f to the decomposition group $\mathrm{Gal}(\overline{\mathbf{Q}}_p/\mathbf{Q}_p)$ at p (for $(p, \ell) = 1$) is as required by the local Langlands correspondence; in particular, the conductor of the new form f and the Galois representation ρ are the same (away from ℓ). From this, it is easy to conclude that for $p \neq 2, \ell$, the conductor of f is not divisible by p^3 if the reduction of ρ modulo λ is unramified at p.

REMARK 1.12. As a trivial corollary of theorem 1.7, observe that if $\rho : \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to GL(2, \mathbf{F})$ is irreducible, then if ρ is modular of weight 2, trivial Nebentypus character and level M, then it can't be modular of weight 2, trivial Nebentypus character and level N with $(M, N) = 1$.

We now indicate how Theorem 1.7 can be used to deduce Fermat's last theorem from the Shimura-Taniyama-Weil conjecture on the modularity of all elliptic curves over $\mathbf{Q}$ (actually one needs the modularity of only semi-stable elliptic curves).

Assume that there are integers a, b, c with $abc \neq 0$, and $gcd(a, b, c) = 1$ such that $a^\ell + b^\ell + c^\ell = 0$ and $a \equiv -1 \bmod 4$, $b \equiv 0 \bmod 2$, where ℓ is a prime ≥ 5. Following Frey, one defines the elliptic curve E by the equation,

$$y^2 = x(x - a^\ell)(x + b^\ell).$$

As the roots of the cubic polynomial $x(x - a^\ell)(x + b^\ell)$ are not all the same modulo any prime number, the curve $y^2 = x(x - a^\ell)(x + b^\ell)$ has semi-stable

reduction at all primes $p \neq 2$. By constructing an appropriate model over $\mathbf{F}_2$ as done by Serre [**S**], one can prove semi-stability at 2 also. Assuming the modularity of E, there is a modular form f for $\Gamma_0(N)$ where N is a square free integer such that its Mellin transform is the L-function of E over $\mathbf{Q}$. In particular, the mod-ℓ representation ρ of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ attached to f is realised by the vector space $E[\ell]$ of ℓ-torsion points of E. Since $\ell \geq 5$, and the 2-torsion of E is defined over $\mathbf{Q}$, it follows from the results of Mazur that the representation of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ on $E[\ell]$ is irreducible. We now check using the Tate parametrization that the representation ρ is finite at every prime. The Tate curve $E = G_m/(q^{\mathbf{Z}})$ over a local field of residue characteristic p (which is either $\mathbf{Q}_p$ or an unramified quadratic extension of $\mathbf{Q}_p$) where the semi-stable elliptic curve has bad reduction, shows that the representation of $\mathrm{Gal}(\overline{\mathbf{Q}}_p/\mathbf{Q}_p)$ on $E[\ell]$ is unramified if and only if $p \neq \ell$, and $\ell \mid \mathrm{val}(q)$, and $E[\ell]$ is finite if and only if $\ell \mid \mathrm{val}\ (q)$. As $\mathrm{val}(q) = -\mathrm{val}(j_E)$, it can be checked by explicit calculation that $\ell \mid \mathrm{val}(j_E)$ for $E : y^2 = x(x - a^\ell)(x + b^\ell)$. By the theorem of Mazur and Ribet, the representation of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ on $E[\ell]$ is modular of weight 2, level 2 and trivial Nebentypus character. As there are no non-zero cusp forms of weight 2 on $\Gamma_0(2)$, this contradiction proves Fermat's last theorem.

2. Modular Curves and their Jacobians

Let $X_0(N)$ denote the classical modular curve over $\mathbf{Q}$ which is the compactification of $Y_0(N)$ which classifies pairs (E, C) where E is an elliptic curve together with a cyclic subgroup C of order N. The curve $X_0(N)$ has a model over $\mathbf{Z}$ which has good reduction at all primes p with $(p, N) = 1$, and for primes dividing N, its reduction modulo p can be explicitly described, and is due to Deligne and Rapoport [**DR**] in the case when p^2 does not divide N. We recall these results assuming that p^2 does not divide N in what follows.

The reduction modulo p of the curve $X_0(N)$ is the union of two copies of $X_0(N/p)$ over $\mathbf{F}_p$ intersecting transversally at their supersingular points; a point x on the first copy of $X_0(N/p)$ being glued to its Frobenius transform $x^{(p)}$ on the second copy of $X_0(N/p)$.

Using this description of the reduction modulo p of the curve $X_0(N)$, it follows from a general theorem due to Raynaud that the connected component of the reduction modulo p of the Néron model of the Jacobian of $X_0(N)$ is an extension of the product of two copies of the Jacobian of $X_0(M)$ $(M = N/p)$ by a torus whose character group is the group of degree zero divisors on $X_0(M)$ over $\overline{\mathbf{F}}_p$ supported at the supersingular points:

$$0 \to T \to J_0(N)^0 \to J_0(M) \times J_0(M) \to 0 \qquad (2.1).$$

The action of $\mathrm{Gal}(\overline{\mathbf{F}}_p/\mathbf{F}_p)$ on the character group of the torus is via the natural action of $\mathrm{Gal}(\overline{\mathbf{F}}_p/\mathbf{F}_p)$ on the supersingular points of $X_0(M)$ over $\overline{\mathbf{F}}_p$. Since any supersingular point is defined over the quadratic field extension of $\mathbf{F}_p$, the torus splits over $\mathbf{F}_{p^2}$.

Also, from another general theorem of Raynaud, the group of connected components of the reduction modulo p of the Néron model of $J_0(N)$ can be described as follows.

Let Λ be the free abelian group on the set of supersingular points on $X_0(N/p)$ over $\overline{\mathbf{F}}_p$, and let $X \subseteq \Lambda$ be the subgroup of Λ consisting of elements of degree 0. Thus X is the character group of the torus T described above. Define a bilinear form $B : \Lambda \times \Lambda \to \mathbf{Z}$ by $B(e(i), e(j)) = 0$ if $i \neq j$, and $B(e(i), e(i)) = \#\frac{1}{2}\mathrm{Aut}\ e(i)$ where $e(i)$ is the supersingular point on $X_0(N/p)$ which is represented by a supersingular elliptic curve $E(i)$ over $\overline{\mathbf{F}}_p$ together with a cyclic subgroup $\sqsubset$ of order N/p, and $\#\mathrm{Aut}\ e(i)$ is the cardinality of the automorphism group of the elliptic curve $E(i)$ which preserves $\sqsubset$. Restricting the bilinear form to X, we get a mapping $X \to \mathrm{Hom}\ (X, \mathbf{Z})$, the cokernel of which is finite and isomorphic to the component group ϕ of the reduction modulo p of the Néron model of $J_0(N)$.

The Hecke operators operate on the modular curve $X_0(N)$ by correspondences. Therefore using the isomorphism of the Picard variety of $X_0(N)$ with its Albanese variety, one has two actions of the Hecke operators on the Jacobian of $X_0(N)$. If $\begin{matrix} & X_0(nN) & \\ \alpha \swarrow & & \searrow \beta \\ X_0(N) & & X_0(N) \end{matrix}$ represents the correspondence defining the Hecke operator T_n, then the two actions are $\alpha_* \circ \beta^*$ and $\beta_* \circ \alpha^*$. The first of these two actions is called the action via Picard functoriality (as it is contravariant) and the second is called the action via Albanese functoriality; unless otherwise mentioned, we take the action via Picard functoriality. These two actions are the same for T_n when $(n, N) = 1$ but not in general.

By the universal property of the Néron model, the Hecke operators also act on the Néron model of $J_0(N)$ over $\mathbf{Z}$. This gives the action of the Hecke operators on $J_0(N)$ over $\overline{\mathbf{F}}_p$ which preserves the connected component $J_0(N)^0$ passing through the origin, and therefore gives an action of the Hecke operators on the group of connected components of $J_0(N)$ over $\overline{\mathbf{F}}_p$. The action of the Hecke operators on $J_0(N)^0$, preserves the toric part, and the corresponding action of the Hecke operators on the character group X of the torus is the same as the one on supersingular points on $X_0(N/p)$ over $\overline{\mathbf{F}}_p$. The induced action on $\phi = X^*/X$ gives the action of Hecke operators on the group of connected components. From this description it is easy to see the following lemma.

LEMMA 2.2. *The Hecke operator T_p acts on ϕ by multiplication by $(p+1)$ for a prime p with $(p, N) = 1$.*

Finally, we can use the torus T over $\mathbf{F}_p$ to construct a certain subgroup of $J_0(N)(\overline{\mathbf{Q}}_p)$ which is stable under both the action of Galois group and the Hecke operators. This is done by using the surjection $\mathrm{Gal}(\overline{\mathbf{Q}}_p/\mathbf{Q}_p) \to \mathrm{Gal}(\overline{\mathbf{F}}_p/\mathbf{F}_p)$ to lift the character group X to a module for $\mathrm{Gal}(\overline{\mathbf{Q}}_p/\mathbf{Q}_p)$; this lifts the torus T over $\mathbf{F}_p$ to one over $\mathbf{Z}_p$ which we abuse notation to again denote by T. Then, it is a theorem of Grothendieck that $T(\overline{\mathbf{Z}}_p)$ is a subgroup of $J_0(N)(\overline{\mathbf{Q}}_p)$ respecting both the Hecke and Galois actions.

3. Old and New forms

Let p be a prime such that p divides N but p^2 does not. Any modular form f on $\Gamma_0(N/p)$ gives rise to two modular forms on $\Gamma_0(N)$: one f itself, and the other $z \to f(pz)$. Thus, we have two embeddings of $S_k(\Gamma_0(N/p))$ into $S_k(\Gamma_0(N))$, giving rise to a map $S_k(\Gamma_0(N/p)) \oplus S_k(\Gamma_0(N/p)) \to S_k(\Gamma_0(N))$, the image of which is called the space of p-old forms on $\Gamma_0(N)$, and denoted by $S_k(\Gamma_0(N))^{p\text{-old}}$. The space $S_k(\Gamma_0(N))^{p\text{-old}}$ is stable under all the Hecke operators though the map $S_k(\Gamma_0(N/p)) \oplus S_k(\Gamma_0(N/p)) \to S_k(\Gamma_0(N))$ commutes only with those Hecke operators T_n for which $(p,n) = 1$; the action of the Hecke operator T_p on the subspace $S_k(\Gamma_0(N/p)) \oplus S_k(\Gamma_0(N/p)) \subseteq S_k(\Gamma_0(N))$ is given by the 2×2 matrix $\begin{pmatrix} T_p & 1 \\ -p^{k-1} & 0 \end{pmatrix}$ where T_p in the matrix is the p-th Hecke operator on $S_k(\Gamma_0(N/p))$. The orthogonal complement under the Petersson product gives a Hecke equivariant complement $S_k(\Gamma_0(N))^{p\text{-new}}$ to $S_k(\Gamma_0(N))^{p\text{-old}}$ in $S_k(\Gamma_0(N))$.

Let $\mathbf{T}_N$ denote the commutative subalgebra of $\mathrm{End}(S_2(\Gamma_0(N)))$ generated by the Hecke operators $T_n, n \geq 1$. It is a torsion free commutative $\mathbf{Z}$ algebra of rank $= \dim(S_2(\Gamma_0(N))$ which is an order in a product of number fields. Let the image of the Hecke algebra $\mathbf{T}_N$ on the p-old forms (resp. p-new forms) be denoted by $\mathbf{T}_N^{p\text{-old}}$ (resp. $\mathbf{T}_N^{p\text{-new}}$). We get surjective maps

$$\mathbf{T}_N \to \mathbf{T}_N^{p\text{-old}},$$

and

$$\mathbf{T}_N \to \mathbf{T}_N^{p\text{-new}},$$

and an injective map

$$\mathbf{T}_N \to \mathbf{T}_N^{p\text{-old}} \times \mathbf{T}_N^{p\text{-new}}.$$

A maximal ideal m of $\mathbf{T}_N$ whose image under $\mathbf{T}_N \to \mathbf{T}_N^{p\text{-old}}$ is not the unit ideal - equivalently, a maximal ideal m in $\mathbf{T}_N$ which comes as the pull back of a maximal ideal in $\mathbf{T}_N^{p\text{-old}}$- is called a p-old ideal. Similarly, one defines p-new ideals. As the map $\mathbf{T}_N \to \mathbf{T}_N^{p\text{-old}} \times \mathbf{T}_N^{p\text{-new}}$ is not necessarily an isomorphism, it is possible that a maximal ideal m of $\mathbf{T}_N$ is both p-old and p-new. This is a reflection of the possibility that the Galois representation associated to an old form in $\Gamma_0(N)$ and a new form in $\Gamma_0(N)$ may be the same modulo a prime ℓ. Such primes ℓ (or, the maximal ideals m of $\mathbf{T}_N$ of residue characteristic ℓ) are called primes of fusion between p-old and p-new forms.

REMARK 3.1. The general context for the concept of primes of fusion is the following. Suppose that V is a finite dimensional vector space over $\mathbf{Q}$ together with a choice of a lattice $V(\mathbf{Z}) \subseteq V$. Suppose $V = X \oplus Y$ where X and Y are vector spaces over $\mathbf{Q}$. Then one can ask if $V(\mathbf{Z}) = X \cap V(\mathbf{Z}) \bigoplus Y \cap V(\mathbf{Z})$? The primes which divide the order of the finite abelian group $[V(\mathbf{Z}) : X \cap V(\mathbf{Z}) \bigoplus Y \cap V(\mathbf{Z})]$ are called primes of fusion. If $V_X(\mathbf{Z})$ is the image of $V(\mathbf{Z})$ in X, and if $V_Y(\mathbf{Z})$ is similarly defined, then $V(\mathbf{Z}) \subseteq V_X(\mathbf{Z}) \oplus V_Y(\mathbf{Z})$ and it can be seen that the primes which divide the orders of the finite abelian groups

$[V(\mathbf{Z}) : X \cap V(\mathbf{Z}) \bigoplus Y \cap V(\mathbf{Z})]$ and $[V_X(\mathbf{Z}) \oplus V_Y(\mathbf{Z}) : V(\mathbf{Z})]$ are the same. In the context of modular forms, a basic example of the situation $V = X \oplus Y$ is given by V, the space of cusp forms of weight k and level N, X the space of old forms, and Y the space of new forms. There is a natural $\mathbf{Z}$ structure on V given by the space of cusp forms for which the co-efficients of q-expansion lies in $\mathbf{Z}$. Interpreting the space of cusp forms as the parabolic cohomology, $V(\mathbf{C}) \oplus \overline{V}(\mathbf{C})$ has an entirely different $\mathbf{Q}$ and $\mathbf{Z}$ structures. One can define spaces X and Y in this set-up too, and it is a theorem of Ribet [**R1**] that the primes of fusion for these two examples are essentially the same.

REMARK 3.2. It is a particular case of a result of Mazur [**M,Prop 10.6**] that for $N = pM$ with $(p, M) = 1$, the exact sequence of abelian varieties

$$0 \to J_0(N)^{\text{new}} \to J_0(N) \to J_0(N)' \to 0 \tag{3.3}$$

where $J_0(N)^{\text{new}}$ is the connected component of the kernel of the natural map $J_0(N) \to J_0(M) \times J_0(M)$, and $J_0(N)'$ is defined by the exact sequence above, does not split. We use the results of Deligne-Rapoport to prove this result. We first note that if the sequence above splits over $\mathbf{C}$, then as the splitting is unique (because $\text{Hom}[J_0(N)', J_0(N)^{\text{new}}] = 0$), the sequence will split over $\mathbf{Q}$ too. This will imply $J_0(N) = J_0(N)' \times J_0(N)^{\text{new}}$ over $\mathbf{Q}$, and therefore we will have the product decomposition of the corresponding Néron models over $\mathbf{Z}$. Now, $J_0(N)'$ is isogenous to $J_0(M) \times J_0(M)$ with $J_0(M)$ having good reduction at p, and therefore $J_0(N)'$ also has good reduction at p. From the exact sequence (2.1), we find that $J_0(N)^{\text{new}}$ has multiplicative reduction, and moreover $J_0(N) = J_0(N)' \times J_0(N)^{\text{new}}$ forces the exact sequence (2.1) to split. So we need to prove that the exact sequence (2.1) does not split. This is a general lemma about Pic^0 of singular curves which we briefly describe; the author has however not been able to find a reference to this in the literature. First of all, we note that $\text{Ext}^1[A, \mathbf{G}_m] \cong \text{Pic}^0 A$ for any abelian variety A. This can be generalised to any torus T with character group $X^*(T)$ as,

$$\text{Ext}^1[A, T] \cong \text{Hom}[X^*(T), \text{Pic}^0(A)].$$

The exact sequence (2.1) gives rise to an element e in

$$\text{Ext}^1[J_0(M) \times J_0(M), T] \\ \cong \text{Hom}[X^*(T), \text{Pic}^0(J_0(M))] \oplus \text{Hom}[X^*(T), \text{Pic}^0(J_0(M))].$$

Now $X^*(T)$ is the group of degree 0 divisors on $X_0(M)$ in characteristic p which is supported at the supersingular points, and as such there is a natural mapping of $X^*(T)$ into $\text{Pic}^0(J_0(M))$ which is non-trivial as soon as $J_0(M)$ and T have positive dimensions. The element e is

$$e = e_1 + e_2 \in \text{Hom}[X^*(T), \text{Pic}^0(J_0(M))] \oplus \text{Hom}[X^*(T), \text{Pic}^0(J_0(M))]$$

where e_1 is the natural map of $X^*(T)$ into $\text{Pic}^0(J_0(M))$, and e_2 is this map followed by the Frobenius.

REMARK 3.4. One can calculate the order of the intersection of the p-old and p-new abelian subvarieties in $J_0(pM)$ which is responsible for the non-splitting of the exact sequence (3.3). We recall that using the two degeneracy maps $X_0(pM) \rightrightarrows X_0(M)$, we get maps $J_0(M) \times J_0(M) \to J_0(pM)$, and $J_0(pM) \to J_0(M) \times J_0(M)$ using the Picard and the Albanese functoriality of the Jacobians. The image of $J_0(M) \times J_0(M)$ in $J_0(pM)$ under the map $J_0(M) \times J_0(M) \to J_0(pM)$ is called the p-old subvariety of $J_0(pM)$, and the connected component of the kernel of the map $J_0(pM) \to J_0(M) \times J_0(M)$ is called the p-new subvariety of $J_0(pM)$. We denote the p-old and p-new subvarieties of $J_0(pM)$ by $J_0(pM)^{\text{old}}$ and $J_0(pM)^{\text{new}}$ respectively. The composition of the maps $J_0(M) \times J_0(M) \to J_0(pM) \to J_0(M) \times J_0(M)$ is represented by the 2×2 matrix of endomorphisms of $J_0(M) \times J_0(M)$ given by $\begin{pmatrix} p+1 & T_p \\ T_p & p+1 \end{pmatrix}$ where T_p is the pth Hecke operator on $J_0(M)$. As the degree of an endomorphism of an abelian variety is the square of the determinant of the endomorphism on the tangent space, the degree of the endomorphism of $J_0(M) \times J_0(M)$ given by $\begin{pmatrix} p+1 & T_p \\ T_p & p+1 \end{pmatrix}$ is

$$\prod_f \left[a_p^2(f) - (p+1)^2\right]^2, \tag{3.5.1}$$

where $f = \sum a_n(f) q^n$ varies over all the new forms of level a divisor of M.

The group of connected components of the kernel of the map $J_0(pM) \to J_0(M) \times J_0(M)$ is in duality with the kernel $\Sigma(M)$ (which is finite, and is isomorphic to the Shimura subgroup of $J_0(M)$; see [**LO**] for the structure of this group) of the map $J_0(M) \times J_0(M) \to J_0(pM)$; in particular, the two groups have the same order. We can write the composition of the maps, $J_0(M) \times J_0(M) \to J_0(pM) \to J_0(M) \times J_0(M)$ as

$$J_0(M) \times J_0(M) \to J_0(pM)^{\text{old}} \to J_0(pM)/J_0(pM)^{\text{new}} \to J_0(M) \times J_0(M).$$

Using this, we find that the order of the intersection of the p-old and p-new abelian subvarieties in $J_0(pM)$ is equal to $\prod_f \left[a_p^2(f) - (p+1)^2\right]^2 / \#\Sigma(M)^2$. (We recall that if A and B are two abelian subvarieties of an abelian variety J with $A+B = J$, $A \cap B$ finite, then the degree of the map $A \to J/B$ is clearly the order of $A \cap B$.) On the other hand, it can be easily seen that the order of $J_0(M)(\mathbf{F}_{p^2})$ is given by

$$\prod_f \left[a_p^2(f) - (p+1)^2\right], \tag{3.5.2}$$

where $f = \sum a_n(f) q^n$ varies over all the new forms of level a divisor of M. Therefore the order of the intersection of the p-old and p-new abelian subvarieties in $J_0(pM)$ is the same as $[\#J_0(M)(\mathbf{F}_{p^2})/\#\Sigma(M)]^2$.

One can describe the subgroup of $J_0(M)$ of degree zero divisors supported at the supersingular points of $X_0(M)$ in characteristic p for $(p, M) = 1$ which appears in remark 3.2. Before we do this, we recall that the kernel of the map $J_0(M) \to J_1(M)$ where $J_1(M) = \text{Pic}^0(X_1(M))$ is a finite group scheme $\Sigma(M)$,

called the Shimura group, such that $\Sigma(M)(\overline{k})$ is the Galois group of the maximal unramified cover of $X_0(M)$ contained in $X_1(M)$ over any algebraically closed field $\overline{k}$ of characteristic coprime to M.

In the statement of the next proposition, and its proof, all the geometric objects are defined over the finite field with p elements.

PROPOSITION 3.6. *The subgroup $S(M)$ of $J_0(M)$ of degree zero divisors supported at the supersingular points of $X_0(M)$ in characteristic p for $(p, M) = 1$ is contained in $J_0(M)(\mathbf{F}_{p^2})$, and in fact, $S(M)$ is precisely the image of $J_1(M)(\mathbf{F}_{p^2})$ into $J_0(M)(\mathbf{F}_{p^2})$ under the natural map $J_1(M) \to J_0(M)$. Moreover, there is an isomorphism of group schemes over $\mathbf{F}_p$,*

$$J_0(M)(\mathbf{F}_{p^2})/S(M) \cong \Sigma(M)^\vee,$$

where by abuse of notation, $J_0(M)(\mathbf{F}_{p^2})$ denotes the group scheme over $\mathbf{F}_p$ whose $\overline{\mathbf{F}}_p$-valued points are $J_0(M)(\mathbf{F}_{p^2})$ with the obvious action of the Frobenius, and $\Sigma(M)^\vee$ is the Cartier dual of the group $\Sigma(M)$.

PROOF. Let K be the kernel of the natural map $J_1(M) \to J_0(M)$. From the exact sequence,

$$0 \to K \to J_1(M) \xrightarrow{j} J_0(M) \to 0$$

of group schemes over $\mathbf{F}_{p^2}$, we get using Lang's theorem on the triviality of the first Galois cohomology of a connected group over a finite field,

$$J_1(M)(\mathbf{F}_{p^2}) \xrightarrow{j} J_0(M)(\mathbf{F}_{p^2}) \to H^1(\mathrm{Gal}(\overline{\mathbf{F}}_p/\mathbf{F}_{p^2}), K) \to 0.$$

The connected component of K is an abelian variety, and the group of connected components of K is Cartier dual $\Sigma(M)^\vee$ to the Shimura group $\Sigma(M)$ and is therefore a constant group scheme over $\mathbf{F}_p$ [**M,Prop. 11.6**]. Therefore by one more application of Lang's theorem, $H^1(\mathrm{Gal}(\overline{\mathbf{F}}_p/\mathbf{F}_{p^2}), K)$ is isomorphic to $\Sigma(M)^\vee(\mathbf{F}_{p^2})$. As the exact sequence above involving Galois cohomology is equivariant under the action of $\mathrm{Gal}(\mathbf{F}_{p^2}/\mathbf{F}_p)$, and the action of $\mathrm{Gal}(\mathbf{F}_{p^2}/\mathbf{F}_p)$ on $H^1(\mathrm{Gal}(\overline{\mathbf{F}}_p/\mathbf{F}_{p^2}), K)$ is trivial, we have the isomorphism of group schemes,

$$J_0(M)(\mathbf{F}_{p^2})/j[J_1(M)(\mathbf{F}_{p^2})] \cong \Sigma(M)^\vee. \tag{3.7}$$

As all the supersingular points of $X_1(M)$ are defined over $\mathbf{F}_{p^2}$, the image of $J_1(M)(\mathbf{F}_{p^2})$ in $J_0(M)(\mathbf{F}_{p^2})$ contains the subgroup $S(M)$ of $J_0(M)$. It therefore suffices to prove that the order of $J_0(M)(\mathbf{F}_{p^2})/S(M)$ is less than or equal to the order of the subgroup $\Sigma(M)$. For this we use a theorem of Ihara [**Ih,Theorem on page 161**] according to which the modular curve $X(M)$ over $\mathbf{F}_{p^2}$ corresponding to the full congruence subgroup $\Gamma(M)$ has no unramified coverings over $\mathbf{F}_{p^2}$ which are completely split at the supersingular points of $X(M)$. This implies by base change that all the unramified coverings of $X_0(M)$ over $\mathbf{F}_{p^2}$ which are completely split at the supersingular points are contained in $X(M)$ over $\mathbf{F}_{p^2}$. The maximal abelian cover of $X_0(M)$ in $X(M)$ is $X_1(M)$, and the Galois group of the maximal unramified covering of $X_0(M)$ which is contained in $X_1(M)$ is $\Sigma(M)$.

Since subgroups H of $J_0(M)(\mathbf{F}_{p^2})$ give rise to unramified abelian covers of $X_0(M)$ over $\mathbf{F}_{p^2}$ with Galois group $J_0(M)(\mathbf{F}_{p^2})/H$, taking H to be the subgroup $S(M)$, we get a covering of $X_0(M)$ with Galois group $J_0(M)(\mathbf{F}_{p^2})/S(M)$ in which all the supersingular points of $X_0(M)$ are split. Therefore $J_0(M)(\mathbf{F}_{p^2})/S(M)$ is a quotient of $\Sigma(M)$. Combined with (3.7), this gives the desired isomorphism $J_0(M)(\mathbf{F}_{p^2})/S(M) \cong \Sigma(M)^\vee$.

REMARK 3.8. Ken Ribet has pointed out that the above proposition, although never written down in print, has been known to him in principle; in fact, it was supposed to be a major step for his ICM, 83 paper [**R2**] in a preliminary version but in the final version it was short-circuited by directly using the result of Ihara used in the proof above.

REMARK 3.9. The Hecke algebra $\mathbf{T}_N$ operates on $J_0(N)$ over $\mathbf{Q}$, and therefore by the universal property of the Néron model, it also acts on the reduction modulo p of $J_0(N)$, preserving the connected component $J_0(N)^0$ passing through the origin. In the exact sequence (2.1), $\mathbf{T}_N$ preserves T, and therefore acts also on $J_0(N/p) \times J_0(N/p)$. It can be seen, cf. Theorems 3.10 and 3.11 of [**R3**], that these actions of $\mathbf{T}_N$ factor through $\mathbf{T}_N^{p\text{-new}}$, and $\mathbf{T}_N^{p\text{-old}}$, and these act faithfully.

4. Reformulation of a theorem of Deligne

In this section we reformulate the theorem of Deligne which attaches representations of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ into $GL(2, \mathbf{F})$ to Hecke eigenforms with coefficients in the finite field $\mathbf{F}$ in terms of the Hecke algebra $\mathbf{T}_N$.

THEOREM 4.1. *Let $\mathfrak{m}$ be a maximal ideal of $\mathbf{T}_N$. Then there exists a representation $\rho_{\mathfrak{m}} : Gal(\overline{\mathbf{Q}}/\mathbf{Q}) \to GL(2, \mathbf{T}_N/\mathfrak{m})$ which is unramified outside the primes dividing N and the residue characteristic of $\mathbf{T}_N/\mathfrak{m}$ such that for such primes*

(1) $tr\ \rho_{\mathfrak{m}}(Frob_p) = T_p$,
(2) $\det \rho_{\mathfrak{m}}(Frob_p) = p$.

PROOF. We begin by noting the well-known fact that the bilinear form $\mathbf{T}_N \otimes S_k(\Gamma_0(N), \mathbf{Z}) \to \mathbf{Z}$ which is obtained by sending $T \otimes f \in \mathbf{T}_N \otimes S_k(\Gamma_0(N), \mathbf{Z})$ to the coefficient of q in Tf, is a non-degenerate bilinear form. This implies that for any $\phi \in \mathrm{Hom}(\mathbf{T}_N, R)$ where R is any commutative ring, $\sum \phi(T_n) q^n$ is a modular form with coefficients in R. This modular form is an eigenform of all the Hecke operators T_n with eigenvalues $\phi(T_n)$ whenever $\phi \in \mathrm{Hom}(\mathbf{T}_N, R)$ is a ring homomorphism. Taking $R = \mathbf{T}_N/\mathfrak{m}$, and ϕ to be the natural element of $\mathrm{Hom}(\mathbf{T}_N, \mathbf{T}_N/\mathfrak{m})$, we find that $\sum_{n \geq 1} T_n q^n$ is an eigenform of the Hecke operators T_n with eigenvalues $T_n \bmod \mathfrak{m}$. Therefore, the theorem of Deligne which attaches a representation of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ into $\mathrm{GL}(2,\mathbf{F})$ to a cusp form with coefficients in the finite field $\mathbf{F}$ which is an eigenform of all the Hecke operators completes the proof.

LEMMA 4.2. *Let q be a prime and N an integer such that $(q, N) = 1$. Let $\mathfrak{m}$ be a maximal ideal in $\mathbf{T}_{qN}$ which is q-old; i.e., assume that $\mathfrak{m}$ arises from the pull back of a maximal ideal $\mathfrak{m}'$ of $\mathbf{T}_{qN}^{q\text{-old}}$ under the natural surjection $\mathbf{T}_{qN} \to \mathbf{T}_{qN}^{q\text{-old}}$. Then the Galois representation $\rho_{\mathfrak{m}}$ is modular of level N, i.e., $\rho_{\mathfrak{m}}$ can be realised (up to extension of scalars) as $\rho_{\mathfrak{m}''}$ with values in $GL(2, \mathbf{T}_N/\mathfrak{m}'')$ for some maximal ideal $\mathfrak{m}''$ in $\mathbf{T}_N$.*

PROOF. We begin by noting that as $\mathbf{T}_{qN}/\mathfrak{m} \cong \mathbf{T}_{qN}^{q\text{-old}}/\mathfrak{m}'$, the representation $\rho_{\mathfrak{m}}$ can be assumed to take values in $\mathrm{GL}(2, \mathbf{T}_{qN}^{q\text{-old}}/\mathfrak{m}')$. Let $\mathbf{T}_{qN}^{(q)}$ be the subring of $\mathbf{T}_{qN}^{q\text{-old}}$ generated by the Hecke operators T_n for n coprime to q. Clearly, $\mathbf{T}_{qN}^{q\text{-old}}$ is an integral extension of $\mathbf{T}_{qN}^{(q)}$. Therefore, $\mathfrak{m}' \cap \mathbf{T}_{qN}^{(q)} = \mathfrak{m}^{(q)}$ is a maximal ideal of $\mathbf{T}_{qN}^{(q)}$. As the Hecke operators T_n for n coprime to q act diagonally on the subspace $S_2(\Gamma_0(N)) \oplus S_2(\Gamma_0(N))$ of $S_2(\Gamma_0(Nq))$ according to the action of T_n on $S_2(\Gamma_0(N))$, there is an injection $\mathbf{T}_{qN}^{(q)} \hookrightarrow \mathbf{T}_N$ which is again an integral extension. By the going-up theorem, there is a maximal ideal $\mathfrak{m}''$ of $\mathbf{T}_N$ such that $\mathfrak{m}'' \cap \mathbf{T}_{qN}^{(q)} = \mathfrak{m}^{(q)}$. As for all primes p different from q, $\mathrm{tr}\ \rho_{\mathfrak{m}}(\mathrm{Frob}_p) = \mathrm{tr}\ \rho_{\mathfrak{m}''}(\mathrm{Frob}_p) = T_p$ in $\mathbf{T}_{qN}^{(q)}/\mathfrak{m}^{(q)}$ which is a subfield of both $\mathbf{T}_{qN}^{q\text{-old}}/\mathfrak{m}'$ and $\mathbf{T}_N/\mathfrak{m}''$, the representations $\rho_{\mathfrak{m}}$ and $\rho_{\mathfrak{m}''}$ are isomorphic after extension of scalars.

5. Realisation of $\rho_{\mathfrak{m}}$ on the Jacobian of modular curve

Let $\mathfrak{m}$ be a maximal ideal in the Hecke algebra $\mathbf{T}_N$. The Hecke algebra $\mathbf{T}_N$ operates faithfully by endomorphisms on $J_0(N)$, the Jacobian of the modular curve $X_0(N)$.

Let $J_0(N)[\mathfrak{m}] = \left\{x \in J_0(N)(\overline{\mathbf{Q}}) \mid Tx = 0 \text{ for all } T \in \mathfrak{m}\right\}$ be the $\mathfrak{m}$-torsion subgroup of $J_0(N)$. Then $J_0(N)[\mathfrak{m}] \neq 0$. To see this, write $J_0(N)(\mathbf{C}) = \mathbf{C}^n/L$ for a lattice L in $\mathbf{C}^n$. The Hecke algebra $\mathbf{T}_N$ operates on L, and also on $L^* = \mathrm{Hom}(L, \mathbf{Z})$. Since the action of $\mathbf{T}_N$ is faithful on L^*, $\mathfrak{m}L^* \neq L^*$ by Nakayama's lemma. Therefore by the elementary divisor theorem, the $\mathbf{Z}$-subspace of $L \otimes \mathbf{Q}$ which under the natural pairing $(L \otimes \mathbf{Q}) \otimes L^* \to \mathbf{Q}$ takes integral values on $\mathfrak{m}L^*$ is strictly larger than L. Therefore there exists x in $L \otimes \mathbf{Q}$ which does not lie in L but for which $(\mathfrak{m}f)(x) \subseteq \mathbf{Z}$ for all $f \in L^*$. This implies that $f(\mathfrak{m}x) \subseteq \mathbf{Z}$ for all $f \in L^*$, i.e., $\mathfrak{m}x \subseteq L$. Such an x gives rise to a non-zero element of $J_0(N)[\mathfrak{m}]$.

So $J_0(N)[\mathfrak{m}]$ is a non-zero vector space over the finite field $\mathbf{T}_N/\mathfrak{m}$, and since the endomorphisms coming from $\mathbf{T}_N$ are defined over $\mathbf{Q}$, $J_0(N)[\mathfrak{m}]$ is stable under $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ and gives rise to a representation of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ in a vector space over $\mathbf{T}_N/\mathfrak{m}$. Denote by $\rho_{\mathfrak{m}}$, the representation of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ into $\mathrm{GL}(2, \mathbf{T}_N/\mathfrak{m})$ arising in Theorem 4.1.

The next lemma is proved by an argument due to Mazur [**M,Prop. 14.2**].

LEMMA 5.1. *Assume that $\rho_{\mathfrak{m}}$ is irreducible. Then the semi-simplification of the representation of $Gal(\overline{\mathbf{Q}}/\mathbf{Q})$ on the $\mathbf{T}_N/\mathfrak{m}$ vector space $J_0(N)[\mathfrak{m}]$ is isomorphic to $\rho_{\mathfrak{m}}^d$ for some $d \geq 1$.*

PROOF. Let p be a prime different from the residue characteristic of $\mathbf{T}_N/\mathfrak{m}$ which does not divide N. If F_p denotes the Frobenius endomorphism of the Jacobian of $J_0(N)$ over $\mathbf{F}_p$, then by the Eichler-Shimura relations

$$F_p^2 - T_p\, F_p \;+\; p = 0.$$

We consider this equation over the $\mathbf{T}_N/\mathfrak{m}$ vector space $J_0(N)(\overline{\mathbf{F}}_p)[\mathfrak{m}]$ where T_p is now an element of the finite field $\mathbf{T}_N/\mathfrak{m}$. It follows that the eigenvalues of the operator F_p acting on $J_0(N)(\overline{\mathbf{F}}_p)[\mathfrak{m}]$ satisfy the equation $X^2 - T_pX + p = 0$. We want to use this to conclude that the characteristic polynomial of F_p is a power of $[X^2 - T_pX + p]$. However if α is an eigenvalue of F_p acting on $J_0(N)(\overline{\mathbf{F}}_p)[\mathfrak{m}]$, it is not clear that p/α is also an eigenvalue. To deduce this, we will have to use the irreducibility of the representation $\rho_{\mathfrak{m}}$. Let W be the group scheme $J_0(N)[\mathfrak{m}]$ over $\mathbf{Q}$, and let $W^\vee = \mathrm{Hom}(W, \mathcal{G}_m)$ be its Cartier dual. Now whenever α is an eigenvalue of a Frobenius element for W, p/α is one for $W^\vee$. This implies that the characteristic polynomial satisfied by the action of the Frobenius at p on $W \oplus W^\vee$ is $[X^2 - T_pX + p]^{\dim\ W}$. Since the characteristic polynomial of the Frobenius for the representation $\rho_{\mathfrak{m}}$ is $X^2 - T_pX + p = 0$, we find that the semi-simplification of $W \oplus W^\vee$ is $\rho_{\mathfrak{m}}^{\dim\ W}$. Since $\rho_{\mathfrak{m}}$ is assumed to be irreducible it follows that the semi-simplification of W itself is $\rho_{\mathfrak{m}}^d$ where $d = \frac{1}{2}\ \dim\ W$.

REMARK 5.2. It has been proved by Boston, Lenstra and Ribet in [**BLR**] that the representation of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ on the $\mathbf{T}_N/\mathfrak{m}$ vector space $J_0(N)[\mathfrak{m}]$ is semi-simple.

REMARK 5.3. If A is an abelian subvariety of $J_0(N)$ which is defined over $\mathbf{Q}$, and is invariant under the Hecke algebra $\mathbf{T}_N$, then $\mathbf{T}_N$ operates on $J_0(N)/A$ also. The proof of lemma 5.1 in fact works for $(J_0(N)/A)$ too, and gives that $(J_0(N)/A)[\mathfrak{m}]$ has as its semi-simplification $\rho_{\mathfrak{m}}^d$. If A, for example, is the abelian subvariety of $J_0(N)$ corresponding to the space of p-new forms where $p|N$, then A is preserved under $\mathbf{T}_N$, and $J_0(N)/A$ corresponds to the space of p-old forms. If the maximal ideal $\mathfrak{m}$ of $\mathbf{T}_N$ when considered as endomorphisms of $J_0(N)/A$ does not contain the identity endomorphism of $J_0(N)/A$ then the representation $\rho_{\mathfrak{m}}$ arises as $\rho_{\mathfrak{m}_0}$ for a maximal ideal $\mathfrak{m}_0$ in $T_{N/p}$ by lemma 4.2.

In certain situations the representation $J_0(N)[\mathfrak{m}]$ of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ is irreducible as in the following theorem due to Mazur [**M**].

THEOREM 5.4. *Assume that $\rho_{\mathfrak{m}}$ is irreducible, and that the residue characteristic of $\mathfrak{m}$ is coprime to $2N$. Then the representation of Gal $(\overline{\mathbf{Q}}/\mathbf{Q})$ on the $\mathbf{T}_N/\mathfrak{m}$ vector space $J_0(N)[\mathfrak{m}]$ is irreducible, and therefore isomorphic to $\rho_{\mathfrak{m}}$.*

REMARK 5.5. Clearly the representation $\rho_{\mathfrak{m}}$ of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ is reducible if $\mathfrak{m}$ is an Eisenstein ideal, i.e., if $\mathfrak{m}$ contains $T_l - (1 + l)$ for all but finitely many l. The converse is also true, cf. [**M,Prop 14.2**].

REMARK 5.6. The theorem 5.4 above goes under the name of "multiplicity one theorem" in the subject. More situations of such multiplicity one theorem have been studied by Mazur and Ribet in [**MR**], where they also construct

some examples where multiplicity one fails. The Hecke algebras also act on the Jacobians of various Shimura curves, and again the $\mathfrak{m}$-torsion gives rise to representations of Galois group which is isomorphic to $\rho_{\mathfrak{m}}^d$. Ribet [**R4**] has studied the multiplicity d for the case of Shimura curve arising out of a maximal order in a division algebra ramified at two primes p and q, but the case of general Shimura curves, or of non-maximal orders even in the case when $\mathfrak{m}$ is coprime to the (reduced) discriminant of the order (as in Mazur's theorem) seems open.

6. Theorem of Cerednik and Drinfeld

Let B be the quaternion division algebra over $\mathbf{Q}$ ramified at two finite primes p and q, and unramified everywhere else. Let $\mathbf{A}$ denote the adele ring of $\mathbf{Q}$, and $\mathbf{A}_f$ the finite adeles. Let R be an Eichler order of reduced discriminant Mpq where M is coprime to pq so that $\widehat{R}^* = (R \otimes \widehat{\mathbf{Z}})^* = \Pi R_\ell^*$ is a compact open subgroup of $B^*(\mathbf{A}_f)$, such that R_ℓ^* is a maximal compact subgroup of $(B \otimes \mathbf{Q}_\ell)^*$ for $\ell \not| M$, and is conjugate to the usual $\Gamma_0(M)$ subgroup at places dividing M.

The set

$$X(\mathbf{C}) = B^* \backslash [GL(2, \mathbf{R})/\mathbf{C}^* \times B^*(\mathbf{A}_f)/\widehat{R}^*] \tag{6.1}$$

is the set of complex points of a projective algebraic curve X defined over $\mathbf{Q}$. The work of Cerednik and Drinfeld, cf. [**Ce**] and [**Dr**], gives a model of X over $\mathbf{Z}_p$ such that the irreducible components of the reduction modulo p of X are all $\mathbf{P}^1$'s. This reduction modulo p of the curve X is really what is relevant to Ribet's work, and we will concentrate only on this aspect of the work of Cerednik-Drinfeld which is most conveniently expressed in terms of what is called the dual graph of the configuration whose vertices are in one-to-one correspondence with the $\mathbf{P}^1$'s occuring in the reduction modulo p of the curve X, and the edges between two vertices correspond to points of intersection of the corresponding $\mathbf{P}^1$'s. The result is that in the case of reduction modulo p of the curve X, the dual graph is

$$\mathcal{G} = GL(2, \mathbf{Q}_p)^+ \backslash [\Delta \times \{\widehat{R'}_p^* \backslash H^*(\mathbf{A}_f)/H^*\}] \tag{6.2}$$

where $GL(2, \mathbf{Q}_p)^+ = \{g \in GL(2, \mathbf{Q}_p) \mid \text{val}(\det g) \in 2\ \mathbf{Z}\}$, H is the quaternion division algebra ramified at q and ∞, and unramified elsewhere, R' is an Eichler order of level M in H, and $\widehat{R'}_p^* = (R' \otimes \widehat{\mathbf{Z}})^*$ except that one omits the factor at the place p; Δ is the tree associated to $SL(2, \mathbf{Q}_p)$.

We observe that the Eichler order R' can be realised as the endomorphism ring of a pair $(E, \sqsubset)$ where E is a supersingular elliptic curve over $\overline{\mathbf{F}}_q$, and $\sqsubset$ is a cyclic subgroup of $E(\overline{\mathbf{F}}_q)$ of order M.

Since $GL(2, \mathbf{Q}_p)^+$ operates on the vertices of Δ with two orbits and operates transitively on the edges of Δ, it follows that the set of vertices of $\mathcal{G}$ is a disjoint union of two copies of

$$v(\mathcal{G}) = \widehat{R'}^* \backslash H^*(\mathbf{A}_f)/H^*, \tag{6.3}$$

and the set of edges of $\mathcal{G}$ is

$$e(\mathcal{G}) = \widehat{R''}^* \backslash H^*(\mathbf{A}_f)\ /\ H^*, \tag{6.4}$$

where $R'' \subseteq R'$ is an Eichler order of level Mp.

There are degeneracy maps (corresponding to the two vertices of an oriented edge) :

$$e(\mathcal{G}) = \widehat{R''}^* \backslash H^*(\mathbf{A}_f)/H^* \underset{\beta}{\overset{\alpha}{\rightrightarrows}} v(\mathcal{G}) = \widehat{R'}^* \backslash H^*(\mathbf{A}_f) \ / \ H^*, \tag{6.5}$$

where α is the natural projection, and β is obtained by conjugating this by an element of H^* which normalises $\widehat{R''}^*$ but does not lie in it. By the definition of $H_1(\mathcal{G}, \mathbf{Z})$, we have the exact sequence

$$\begin{array}{rcl} 0 \to H_1(\mathcal{G}, \mathbf{Z}) \to & \mathbf{Z}^{e(\mathcal{G})} & \to \mathbf{Z}^{v(\mathcal{G})} \oplus \mathbf{Z}^{v(\mathcal{G})} \\ & e & \to \alpha(e) - \beta(e) \end{array} \tag{6.6}$$

Clearly $H_1(\mathcal{G}, \mathbf{Z})$ lies in $(\mathbf{Z}^{e(\mathcal{G})})_0$, the group of degree 0 divisors on $e(\mathcal{G})$. Defining $(\mathbf{Z}^{v(\mathcal{G})})_0$ similarly, we have the exact sequence

$$0 \to H_1(\mathcal{G}, \mathbf{Z}) \to (\mathbf{Z}^{e(\mathcal{G})})_0 \to (\mathbf{Z}^{v(\mathcal{G})})_0 \oplus (\mathbf{Z}^{v(\mathcal{G})})_0. \tag{6.7}$$

¿From strong approximation, $GL(2, \mathbf{Q}_p)^+ \times \widehat{R'}_p^* \backslash H^*(\mathbf{A}_f) \ / H^*$ consists of a single point, and therefore as the tree Δ is connected, so is the graph $\mathcal{G}$. We therefore have the exact sequence

$$0 \to H_1(\mathcal{G}, \mathbf{Z}) \to (\mathbf{Z}^{e(\mathcal{G})})_0 \to (\mathbf{Z}^{v(\mathcal{G})})_0 \oplus (\mathbf{Z}^{v(\mathcal{G})})_0 \to 0. \tag{6.8}$$

By a theorem of Raynaud, $H_1(\mathcal{G}, \mathbf{Z})$ is the character group $\widehat{T}_{X,p}$ of the torus associated to the reduction modulo p of the Shimura curve X. The other terms in the exact sequence (6.8), viz. $(\mathbf{Z}^{e(\mathcal{G})})_0$ and $(\mathbf{Z}^{v(\mathcal{G})})_0$ can be interpreted as the degree zero divisors supported at the supersingular points in characteristic q of the curves $X_0(Mp)$ and $X_0(M)$ respectively, which as recalled earlier are isomorphic to the character groups $\widehat{T}_{Mpq,q}$ and $\widehat{T}_{Mq,q}$ of the tori associated to the reduction modulo q of the modular curves $X_0(Mpq)$ and $X_0(Mq)$, respectively. We will not prove this well known fact which identifies the supersingular points of $X_0(M)$ in characteristic q (equivalently, the isomorphism classes of pairs $(E, \sqsubset)$ where E is a supersingular elliptic curve over $\overline{\mathbf{F}}_q$, and $\sqsubset$ is a cyclic subgroup of order M) to the double coset

$$\widehat{R'}^* \backslash H^*(\mathbf{A}_f) \ / \ H^*. \tag{6.9}$$

Using these interpretations the short exact sequence (6.8) may be re-written as

$$0 \to \widehat{T}_{X,p} \to \widehat{T}_{Mpq,q} \to (\widehat{T}_{Mq,q})^2 \to 0. \tag{6.10}$$

Just as for the modular curves $X_0(N)$, one can define Hecke correspondence T'_ℓ on the Shimura curve X by the action of the standard double cosets on

$$X(\mathbf{C}) = B^* \backslash [GL(2, \mathbf{R}) \ / \ \mathbf{C}^* B^*(\mathbf{A}^f) \ / \widehat{R}^*] \tag{6.11}$$

for $l \neq p, q$, and for $l = p$ or q, T'_l is an involution which is obtained by multiplication by a uniformising parameter of the local division algebra at l.

The curve X can be realised as the moduli space of certain abelian surfaces with level structure on which the Eichler order R acts, and one can interpret the

action of Hecke operators, and the reduction modulo p of the curve X from this point of view from which one deduces that the inclusion $i : \widehat{T}_{X,p} \to \widehat{T}_{Mpq,q}$ of the exact sequence (6.10) is Hecke equivariant, i.e., has the property $i(T'_n x) = T_n(i\,x)$ for all x in $\widehat{T}_{X,p}$. We do not go into it here, and refer to Ribet [**R3, sect.4**].

The injection $\widehat{T}_{X,p} \hookrightarrow \widehat{T}_{Mpq,q}$ in the exact sequence (6.10) respects the inner products on these groups. We again refer to Ribet [**R3, sect.4**] for the proof of this.

7. On the component group of the Jacobian of the Shimura curve

We use the exact sequence (6.10) of character groups (interchanging p and q, and using the abbreviations $Y_q = \widehat{T}_{X,q}, L_p = \widehat{T}_{Mpq,p}, X_p = (\widehat{T}_{Mp,p})^2$)

$$0 \to Y_q \to L_p \to X_p \to 0,$$

to get some information about the component group as a module for the Hecke algebra $\mathbf{T}_{Mpq}$ of the reduction modulo q of the Néron model of the Jacobian of the Shimura curve X.

We begin by noting that the map $\delta : L_p \to X_p$ is obtained from the two degeneracy maps $X_0(Mpq) \rightrightarrows X_0(Mq)$ from the Picard functoriality. Similarly there is a map $\sigma : X_p \to L_p$ obtained from the two degeneracy maps $X_0(Mpq) \rightrightarrows X_0(Mq)$ from the Albanese functoriality. It can be checked that:

(i) σ and δ are adjoint to each other under the pairings on the character groups X_p and L_p defined in section 2.

(ii) $\delta \circ \sigma$ is an endomorphism of X_p which is represented by the 2×2 matrix $\begin{pmatrix} p+1 & T_p \\ T_p & p+1 \end{pmatrix}$.

Let $\Phi_p = X_p^*/X_p$, and $\Theta_p = L_p^*/L_p$ denote the group of connected components associated to $J_0(Mq) \times J_0(Mq)$ and $J_0(Mpq)$ respectively in characteristic p. It follows from (i) above that the following is a commutative diagram

$$\begin{array}{ccccccccc} 0 & \to & L_p & \to & L_p^* & \to & \Theta_p & \to & 0 \\ & & \uparrow & & \uparrow & & \uparrow & & \\ 0 & \to & X_p & \to & X_p^* & \to & \Phi_p & \to & 0. \end{array}$$

As $Y_q \subseteq L_p$ with X_p as the quotient, we have

$$\begin{array}{ccccccccc} 0 & \to & X_p & \to & L_p^*/Y_q & \to & \Theta_p & \to & 0 \\ & & \uparrow \mu & & \uparrow & & \uparrow \theta & & \\ 0 & \to & X_p & \to & X_p^* & \to & \Phi_p & \to & 0, \end{array}$$

where the map $\mu : X_p \to X_p$ in the above diagram is the map

$$\mu = \begin{pmatrix} p+1 & T_p \\ T_p & p+1 \end{pmatrix}$$

of (ii). This commutative diagram yields via the snake lemma, the following exact sequence

$$0 \to \ker(\theta) \to X_p/\mu X_p \to Y_q^*/Y_q \to \operatorname{coker}(\theta) \to 0. \tag{7.1}$$

As the pairing $L_p \times L_p \to \mathbf{Z}$ restricts to give the natural pairing on $Y_q \times Y_q \to \mathbf{Z}$, Y_q^*/Y_q is isomorphic to the component group Ψ_q of the Jacobian of the Shimura curve X over $\overline{\mathbf{F}}_q$. Also, if λ is the automorphism $\begin{pmatrix} -1 & T_p \\ 0 & -1 \end{pmatrix}$, then $\lambda \circ \mu = \gamma = U_p^2 - 1$, where U_p is the endomorphism of X_p represented by the matrix $\begin{pmatrix} T_p & p \\ -1 & 0 \end{pmatrix}$. Therefore we can write the above exact sequence as

$$0 \to \ker(\theta) \to X_p/\gamma X_p \to \Psi_q \to \operatorname{coker}(\theta) \to 0. \tag{7.2}$$

LEMMA 7.3. *Let $\mathfrak{m}$ be a maximal ideal in $\mathbf{T}_{Mpq}$, and $\rho_{\mathfrak{m}}$ the associated two-dimensional Galois representation. Assume that $\rho_{\mathfrak{m}}$ is irreducible. Then the following are equivalent.*

(1) $\Psi_q/\mathfrak{m}\Psi_q \neq 0$.
(2) $X_p/\mathfrak{m}X_p \neq 0$, *and* $U_p^2 - 1 \in \mathfrak{m}$.

Moreover, either of the two conditions implies that $\rho_{\mathfrak{m}}$ is modular of level Mp.

PROOF. The equivalence of (i) and (ii) is immediate from (7.2) after we have noted that by lemma 2.2, $\ker(\theta)$ and $\operatorname{coker}(\theta)$ are Eisenstein, i.e., annihilated by $T_r - (r+1)$ for all but finitely many primes r but as $\rho_{\mathfrak{m}}$ is irreducible, it follows from remark 5.5 that $\ker(\theta)[\mathfrak{m}] = 0$, and $\operatorname{coker}(\theta)[\mathfrak{m}] = 0$ (this is equivalent to say that $\mathfrak{m}$ thought of as endomorphisms of $\ker(\theta)$, and $\operatorname{coker}(\theta)$ contains the identity endomorphism of these). By lemma 4.2, $X_p/\mathfrak{m}X_p \neq 0$ implies that $\rho_{\mathfrak{m}}$ is modular of level Mp.

8. Raising the level

Let $f = \sum_{n \geq 1} a_n e^{2\pi i n z}$ be a Hecke eigenform for the group $\Gamma_0(N)$ and q and ℓ primes such that $(q, n) = (\ell, qN) = 1$. Our goal in this section is to find a Hecke eigenform $g = \sum b_n e^{2\pi i n z}$ of level qN which is new at q such that the mod ℓ Galois representations associated to f and g are the same; equivalently, if $\mathcal{O}_K$ denotes the ring of integers in a number field K containing the eigenvalues a_n and b_n for all n, then for a prime $\lambda \mid \ell$ in $\mathcal{O}_K$, $a_n \equiv b_n \bmod \lambda$ for all n with $(n, qN) = 1$. We refer to [**Di1**] and [**R6**] for the following theorem.

THEOREM 8.1. *Assume that l is a prime with $(l, N\phi(N)q) = 1$ where $\phi(N)$ is the Euler ϕ function, then with the notation as above, there exists a form g if and only if*

$$a_q^2 \equiv (q+1)^2 \mod \lambda.$$

PROOF. We first prove that if the form g exists, then

$$a_q^2 \equiv (q+1)^2 \mod \lambda \tag{8.2}.$$

Since the form g is new at q, it follows from a theorem of Langlands [**L**] that the ℓ-adic representation associated to g when restricted to the decomposition

group of Gal $(\overline{\mathbf{Q}}/\mathbf{Q})$ at a place over q is of the form

$$\begin{pmatrix} \mu\chi_\ell & * \\ 0 & \mu \end{pmatrix}$$

where μ is an unramified character of $\mathrm{Gal}(\overline{\mathbf{Q}_q}/\mathbf{Q}_q)$ of order ≤ 2, and χ_ℓ is the ℓ-th cyclotomic character. Taking the trace of the element Frob_q in the reduction modulo ℓ of this representation, we find from the equality of mod ℓ representations associated to f and g that

$$a_q \equiv \mu(\mathrm{Frob}_q)(q+1)(\mathrm{mod}\ \ell),$$

i.e.,

$$a_q^2 \equiv (q+1)^2 \mod \ell.$$

We now prove that the condition $a_q^2 \equiv (q+1)^2 \mod \ell$ is also sufficient to guarantee the existence of a g.

Let α and β be the roots of the equation $X^2 - a_q X + q = 0$. From the condition $a_q^2 \equiv (q+1)^2 \mod \ell$, we find that α and β mod ℓ can be written as $\alpha = a,\ \beta = a\,q$ where $a = \pm 1$.

Consider the form $g = f(z) - \beta f(qz)$. This form can be easily seen to be an eigenvector of the Hecke algebra $\mathbf{T}_{qN}$. The action of $\mathbf{T}_{qN}$ on g gives a homomorphism $\phi : \mathbf{T}_{qN} \to \mathbf{C}$ such that $\phi(T_n) = a_n$ for all $(n,q) = 1$, and $\phi(T_q) = \alpha$.

Viewing this homomorphism into the ring of integers of a number field K and taking the reduction modulo a prime λ of K which lies over ℓ, we get a ring homomorphism $\overline{\phi} : \mathbf{T}_{qN} \to \overline{\mathbf{F}}_l$ such that $\overline{\phi}(T_q^2 - 1) = (a^2 - 1) = 0$. The kernel of $\overline{\phi}$ is therefore a maximal ideal m in $\mathbf{T}_{qN}$ which contains $\mu = (T_q^2 - 1)$. It suffices to prove that m is q-new, i.e., comes from the pull back of a maximal ideal in $\mathbf{T}_{qN}^{q\text{-new}}$ under the surjection $\mathbf{T}_{qN} \to \mathbf{T}_{qN}^{q\text{-new}}$. The proof given below for this is due to Ribet, cf. [**R2**].

Consider the exact sequence

$$0 \to X \to L \to Y \to 0,$$

where $X = H^1(X_0(N), \mathbf{Z}_\ell) \oplus H^1(X_0(N), \mathbf{Z}_\ell)$, and $L = H^1(X_0(qN), \mathbf{Z}_\ell)$, and the two maps from $H^1(X_0(N), \mathbf{Z}_\ell)$ to $H^1(X_0(qN), \mathbf{Z}_\ell)$ correspond to the two degeneracy maps $X_0(qN) \rightrightarrows X_0(N)$. Using results of Ihara, one can prove that Y is a torsion free $\mathbf{Z}_\ell$-module. We assume this result for the time being, and come back to its proof at the end.

The Hecke algebra $\mathbf{T}_{qN}$ operates on this exact sequence, and its image in $\mathrm{End}(X)$ and $\mathrm{End}(Y)$ gives respectively $\mathbf{T}_{qN}^{q\text{-old}}$ and $\mathbf{T}_{qN}^{q\text{-new}}$. To prove that the maximal ideal $\mathfrak{m}$ of $\mathbf{T}_{qN}$ with $\mu = T_q^2 - 1 \in \mathfrak{m}$ comes from $\mathbf{T}_{qN}^{q\text{-new}}$, it suffices to prove that $Y/\mathfrak{m}Y \neq 0$. Since $\mathbf{T}_{qN}$ acts faithfully on L, $L/\mathfrak{m}L \neq 0$. This implies that if $X/\mathfrak{m}X = 0$, $Y/\mathfrak{m}Y \neq 0$. So assume that $X/\mathfrak{m}X \neq 0$. Since $\mu \in \mathfrak{m}$, this means $(X/\mu X)/\mathfrak{m}(X/\mu X) \neq 0$. From this we conclude $Y/\mathfrak{m}Y$ is non-zero by constructing a $\mathbf{T}_{qN}$-equivariant surjective map $Y \to X/\mu X$. To construct

this map, we begin by observing that by Poincare duality, X and L are self-dual $\mathbf{Z}_\ell$-modules. Therefore applying $\mathrm{Hom}(-, \mathbf{Z}_\ell)$ to the exact sequence above, we have

$$0 \to \mathrm{Hom}(Y, \mathbf{Z}_\ell) \to L \to X \to 0.$$

After applying Poincare duality to the map $X \xrightarrow{\alpha} L$ which was by definition $\mathbf{T}_{qN}$-equivariant, the dual map $L \xrightarrow{\beta} X$ does not remain $\mathbf{T}_{qN}$-equivariant (with the previous $\mathbf{T}_{qN}$ structure on L and X).

A computation shows that $\omega \circ \beta : L \to X$ is $\mathbf{T}_{qN}$-equivariant where ω is the automorphism $\begin{pmatrix} -1 & T_q \\ 0 & -1 \end{pmatrix}$ of X, where T_q is the standard Hecke operator on X (and not the one in $\mathbf{T}_{qN}$). By a simple calculation, $\omega \circ \beta \circ \alpha = \mu : X \to X$. It follows that $\omega \circ \beta$ induces a surjection $L/\alpha X = Y \to X/\mu X$ proving the theorem, except that we still need to prove that Y is torsion free which we take up now.

To prove that the cokernel of the map $\alpha : X \to L$ is torsion free, it suffices to prove that the map

$$H^1(X_0(N), \mathbf{Z}/\ell) \oplus H^1(X_0(N), \mathbf{Z}/\ell) \to H^1(X_0(qN), \mathbf{Z}/\ell)$$

is injective. As $H^1(X_0(N), \mathbf{Z}/\ell)$ injects into the cohomology of the open curve $H^1(Y_0(N), \mathbf{Z}/\ell)$ (and similarly for N replaced by qN), it suffices to prove the injectivity of the following map:

$$H^1(Y_0(N), \mathbf{Z}/\ell) \oplus H^1(Y_0(N), \mathbf{Z}/\ell) \to H^1(Y_0(qN), \mathbf{Z}/\ell).$$

As $H^1(Y_0(N), \mathbf{Z}/\ell)$ is the group cohomology $H^1(\Gamma_0(N), \mathbf{Z}/\ell)$ (when $\ell > 3$; the primes ≤ 3 need to be omitted because of the possibility of fixed points of $\Gamma_0(N)$ on the upper half plane), we need to prove the injectivity of

$$H^1(\Gamma_0(N), \mathbf{Z}/\ell) \oplus H^1(\Gamma_0(N), \mathbf{Z}/\ell) \to H^1(\Gamma_0(Nq), \mathbf{Z}/\ell).$$

It is a theorem of Ihara that

$$\Gamma_0(N) \ *_{\Gamma_0(Nq)} \Gamma_0(N) = \Gamma_0(N, \mathbf{Z}[\frac{1}{q}]),$$

where for any ring R, $\Gamma_0(N, R) = \left\{ g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in GL(2, R) \text{ such that } N \mid c \right\}$.

Using this, the Lyndon exact sequence gives

$$H^1(\Gamma_0(N), \mathbf{Z}[\tfrac{1}{q}]), \mathbf{Z}/\ell) \to H^1(\Gamma_0(N), \mathbf{Z}/\ell) \oplus H^1(\Gamma_0(N), \mathbf{Z}/\ell) \to H^1(\Gamma_0(Nq), \mathbf{Z}/\ell).$$

It therefore suffices to show that $H^1(\Gamma_0(N, \mathbf{Z}[\frac{1}{q}]), \mathbf{Z}/\ell) = 0$. An element of the group $H^1(\Gamma_0(N, \mathbf{Z}[\frac{1}{q}]), \mathbf{Z}/\ell)$ is given by a homomorphism $\Gamma_0(N, \mathbf{Z}[\frac{1}{q}]) \to \mathbf{Z}/\ell$. Using the congruence subgroup property for $SL(2, \mathbf{Z}[\frac{1}{q}])$ due to Mennicke and Serre, the kernel of the homomorphism contains a principal congruence subgroup, and since $SL(2, R)$ is a perfect group whenever R is an Artinian local ring of residue characteristic > 2, $\mathrm{Hom}(\Gamma_0(N, \mathbf{Z}[\frac{1}{q}]), \mathbf{Z}/\ell) = 0$ whenever $(l, N\phi(N)) = 1$.

9. Mazur's theorem

The aim of this section is to prove the following theorem due to Mazur.

THEOREM 9.1. *Let $\mathfrak{m}$ be a maximal ideal in $\mathbf{T}_N$ of residue characteristic $l \neq 2$ such that $\rho_{\mathfrak{m}}$ is irreducible, and finite at a place p which divides N but such that p^2 does not. Then if $p \not\equiv 1 \bmod (l)$, $\rho_{\mathfrak{m}}$ is modular of level N/p.*

PROOF. We will prove this theorem only in the case when $l \neq p$; the case when $l = p$ is similar though technically more difficult. For any $\mathbf{T}_N$-module M, let $M[\mathfrak{m}] = \{x \in M | a.x = 0 \text{ for all } a \in \mathfrak{m}\}$ denote the $\mathfrak{m}$-torsion subgroup of M. Let $V \subseteq J(\overline{\mathbf{Q}})[\mathfrak{m}]$ ($J = J_0(N)$) be a $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ stable vector space over $\mathbf{T}_N/\mathfrak{m}$ realising the representation $\rho_{\mathfrak{m}}$ of Theorem 4.1. As $p \neq l$, the representation $\rho_{\mathfrak{m}}$ is unramified at p, and therefore from the properties of the Néron model, the reduction modulo p gives an injective map of V into $J(\overline{\mathbf{F}}_p)[\mathfrak{m}]$. If ϕ denotes the group of connected components of J (over $\overline{\mathbf{F}}_p$), with J^0 the connected component passing through the origin, we have the exact sequence

$$0 \to J^0(\overline{\mathbf{F}}_p) \to J(\overline{\mathbf{F}}_p) \to \phi \to 0.$$

If the image of V under the natural map $J(\overline{\mathbf{F}}_p) \to \phi$ is non-zero, $\phi[\mathfrak{m}] \neq 0$. Since $\phi[\mathfrak{m}]$ is annihilated by the maximal ideal $\mathfrak{m}$, any other annihilator of $\phi[\mathfrak{m}]$ like $T_\ell - (\ell+1)$ should also belong to $\mathfrak{m}$, i.e., $T_\ell - (\ell+1) \in \mathfrak{m}$ for all but finitely many ℓ, and therefore the remark (5.5) which implies that under such a condition $\rho_{\mathfrak{m}}$ is not irreducible, leads to a contradiction, and forces V to lie in $J^0(\overline{\mathbf{F}}_p)$.

Now, J^0 itself is an extension of an abelian variety A by a torus T:

$$0 \to T(\overline{\mathbf{F}}_p) \to J^0(\overline{\mathbf{F}}_p) \to A(\overline{\mathbf{F}}_p) \to 0.$$

If the image of $V \subseteq J^0(\overline{\mathbf{F}}_p)[\mathfrak{m}]$ into $A(\overline{\mathbf{F}}_p)$ is non-zero, then in particular $A[\overline{\mathbf{F}}_p][\mathfrak{m}] \neq 0$, i.e., the image of $\mathfrak{m}$ into $\mathbf{T}_N^{p\text{-old}}$ is not the unit ideal, proving that $\rho_{\mathfrak{m}}$ is of level N/p. Assume, therefore, that $V \subseteq T(\overline{\mathbf{F}}_p)$. By the theorem of Grothendieck recalled at the end of section 2, $T(\overline{\mathbf{Z}}_p) \subseteq J(\overline{\mathbf{Q}}_p)$, and since l is coprime to p, $T(\overline{\mathbf{Z}}_p)[\mathfrak{m}] \cong T(\overline{\mathbf{F}}_p)[\mathfrak{m}]$ under the natural projection. If $\widehat{T}$ denotes the character group of T, then clearly,

$$T(\overline{\mathbf{Z}}_p)[\mathfrak{m}] = \mathrm{Hom}(\widehat{T}/\mathfrak{m}\widehat{T}, \mu_l).$$

The action of $\mathrm{Frob}_p \in \mathrm{Gal}(\overline{\mathbf{Q}}_p/\mathbf{Q}_p)$ on $\widehat{T}$ is via the action of T_p whose square is 1. Since $\widehat{T}/\mathfrak{m}\widehat{T}$ is a vector space over the field $\mathbf{T}_N/\mathfrak{m}$, the action of T_p is either $+1$ or -1 on $\widehat{T}/\mathfrak{m}\widehat{T}$. The determinant of the $\mathrm{Gal}(\overline{\mathbf{Q}}_p/\mathbf{Q}_p)$ representation $\rho_{\mathfrak{m}}$ is χ_l, and by the considerations above if $V \subseteq \mathrm{Hom}(\widehat{T}/\mathfrak{m}\widehat{T}, \mu_l)$, its determinant will be χ_l^2. Therefore $\chi_l^2 = \chi_l$. This is true only if $p \equiv 1 \bmod (l)$. Therefore if $p \not\equiv 1 \bmod (l)$, then V does not lie in $T(\overline{\mathbf{F}}_p)[\mathfrak{m}]$, and therefore as seen earlier $\rho_{\mathfrak{m}}$ is modular of level N/p.

REMARK 9.2. Observe that the proof above shows that if $\rho_{\mathfrak{m}}$ is irreducible with $l \neq 2$, finite at p, and not modular of level N/p, then $\widehat{T} \neq \mathfrak{m}\widehat{T}$.

Using the multiplicity one theorem, we can prove the following too by a similar method:

THEOREM 9.3. *Suppose that $\rho_{\mathfrak{m}}$ is irreducible and $\widehat{T}/\mathfrak{m}\widehat{T}$ is of dimension ≥ 2 over $\mathbf{T}_N/\mathfrak{m}\mathbf{T}_N$ and $(l, 2N) = 1$. Then $p \equiv 1 \bmod l$.*

PROOF. By theorem 5.4, the $\mathbf{T}_N/\mathfrak{m}\mathbf{T}_N$-module

$$J_0(N)(\overline{\mathbf{Q}}_p)[\mathfrak{m}] = J_0(N)(\overline{\mathbf{Q}})[\mathfrak{m}]$$

is an irreducible 2-dimensional representation of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ which realises the representation $\rho_{\mathfrak{m}}$. It contains $T(\overline{\mathbf{Z}}_p)[\mathfrak{m}] = \mathrm{Hom}(\widehat{T}/m\widehat{T}, \mu_l)$ which is by hypothesis of dimension ≥ 2 over $\mathbf{T}_N/\mathfrak{m}\mathbf{T}_N$ (note that for a finite field k of characteristic l, $\mathrm{Hom}(k, \mu_l)$ is naturally a one dimensional k-vector space). Therefore $V = T(\overline{\mathbf{Z}}_p)[\mathfrak{m}]$, and considerations in the above theorem (of calculating the determinant of V in two ways) prove the theorem.

10. Ribet's Theorem

We are finally ready to prove Ribet's Theorem which involves a very delicate consideration of the reduction modulo p of the modular and Shimura curves, and the interplay between these two as evidenced in the exact sequence (6.10).

THEOREM 10.1. *Let $\rho_{\mathfrak{m}_0}$ be an irreducible modular representation*

$$\rho_{\mathfrak{m}_0} \ : \ Gal(\overline{\mathbf{Q}}/\mathbf{Q}) \longrightarrow GL(2, \mathbf{T}_N/\mathfrak{m}_0)$$

where $N = Mp$ with $(p, M) = 1$, and the residue characteristic of the maximal ideal $\mathfrak{m}_0$ in $\mathbf{T}_N$ is l such that $(l, N) = 1$. Assume that $\rho_{\mathfrak{m}_0}$ is unramified at p. Then $\rho_{\mathfrak{m}_0}$ is modular of level M.

PROOF. The image of $\rho_{\mathfrak{m}_0}$ contains the matrix $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$, and therefore by Cebotaraev density theorem, there exists a prime q coprime to $lN\phi(N)$ such that $T_q \equiv 0 \bmod \mathfrak{m}_0$, and $q \equiv -1 \bmod \mathfrak{m}_0$. By the proof of theorem 8.1 about raising of levels, the representation $\rho_{\mathfrak{m}_0}$ is isomorphic to a representation $\rho_{\mathfrak{m}}$ arising out of a maximal ideal $\mathfrak{m}$ in $\mathbf{T}_{Nq}$ containing $T_q^2 - 1$, and such that $\mathfrak{m}$ belongs to the support of the $\mathbf{T}_{Nq}$ module $(\widehat{T}^2_{N,p})$. Let X be the Shimura curve constructed as in section 6 from an Eichler order of reduced discriminant Mpq in a division algebra ramified at p, q. From lemma 7.3, it follows that $\Psi_{X,q} \neq \mathfrak{m}\Psi_{X,q}$. If we let $\mathbf{T}(X)$ denote the quotient of $\mathbf{T}_{Nq}$ by which $\mathbf{T}_{Nq}$ operates on $J(X)$, then from $\Psi_{X,q} \neq \mathfrak{m}\Psi_{X,q}$, we find that $\mathbf{T}(X) \neq \mathfrak{m}\mathbf{T}(X)$, and therefore $J(X)(\overline{\mathbf{Q}})[\mathfrak{m}] \neq 0$. By the Eichler-Shimura theorem for $J(X)$, just as in the case of the modular curve $X_0(N)$ in Lemma 5.1, it follows that the semi-simplification of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$-module $J(X)(\overline{\mathbf{Q}})[\mathfrak{m}]$ over $\mathbf{T}_{Nq}/\mathfrak{m}\mathbf{T}_{Nq} \cong \mathbf{T}(X)/\mathfrak{m}\mathbf{T}(X)$ is isomorphic to $\rho_{\mathfrak{m}}^d$ for some $d \geq 1$. Therefore $J(X)(\overline{\mathbf{Q}})[\mathfrak{m}]$ contains a two-dimensional $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$-invariant $\mathbf{T}_{Nq}/\mathfrak{m}\mathbf{T}_{Nq}$-module V which realises the representation $\rho_{\mathfrak{m}}$. By hypothesis, the representation $\rho_{\mathfrak{m}}$ is unramified at p. Therefore V maps injectively into the

reduction modulo p, $J(X)(\overline{\mathbf{F}}_p)$, of the Jacobian of X by the properties of the Néron model.

If the representation $\rho_{\mathfrak{m}}$ is modular of level Mq, then by Mazur's theorem $\rho_{\mathfrak{m}}$ is in fact of level M. Assume therefore that $\rho_{\mathfrak{m}}$ is not modular of level Mq. From lemma 7.3, it follows that $\Psi_{X,p}[\mathfrak{m}] = 0$. Therefore, the points of $J(X)(\overline{\mathbf{F}}_p)[\mathfrak{m}]$ are located on the connected component passing through the origin which is a torus with character group $\widehat{T}_{X,p}$. Therefore, $J(X)(\overline{\mathbf{F}}_p)[\mathfrak{m}] = \mathrm{Hom}(\widehat{T}_{X,p}/\mathfrak{m}\widehat{T}_{X,p}, \overline{\mathbf{F}}_p^*)$. It follows that $\widehat{T}_{X,p}/\mathfrak{m}\widehat{T}_{X,p}$ is of dimension ≥ 2 over $\mathbf{T}_{Nq}/\mathfrak{m}\mathbf{T}_{Nq}$. From the exact sequence :

$$0 \to \widehat{T}_{X,p} \to \widehat{T}_{Mpq,q} \to (\widehat{T}_{Mq,q})^2 \to 0,$$

and the fact that $\rho_{\mathfrak{m}}$ is not modular of level Mq (so $\mathfrak{m}(\widehat{T}_{Mq,q})^2 = (\widehat{T}_{Mq,q})^2$), it follows that $\widehat{T}_{Nq,q}/\mathfrak{m}\widehat{T}_{Nq,q}$ is of dimension ≥ 2 over $\mathbf{T}_{Nq}/\mathfrak{m}\mathbf{T}_{Nq}$. Since $q \equiv -1 \bmod l$, and $l \nmid N$, Theorem 9.3 leads to a contradiction, and this completes the proof of the theorem.

REMARK 10.2. Ribet has given another proof of the above theorem in [**R6**] which does not rely on the multiplicity one theorem but instead uses the semi-simplicity of Galois representation on $J_0(N)[\mathfrak{m}]$ (and also on the Jacobian of the Shimura curve X) proved in [**BLR**].

REFERENCES

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

[Ca] H. Carayol, *Sur les représentations galoisiennes modulo l attachees aux formes modulaires*, Duke Math. J. **59** (1989), 785-801.

[Ce] I.V. Cerednik, *Uniformization of algebraic curves by discrete arithmetic subgroups of PGL_2 with compact quotients*, Mat. Sb. **100** (1976), 59-88; English transl. in Math USSR Sb. **29** (1976).

[DR] P. Deligne and M. Rapoport, *Les schémas de modules de courbes elliptiques*, Lecture Notes in Math., Springer-Verlag, vol. 349, 1973, pp. 143-316.

[DS] P. Deligne, and J.-P. Serre, *Formes Modulaires de poids 1*, Ann. Sci. École Norm. Sup. **7** (1974), 507-530.

[Di1] F.Diamond, *Congruence primes for cusp forms of weight $k \geq 2$*, Astérisque **196-197** (1991), 205-213.

[Di2] ———, *The refined conjecture of Serre*, Preprint.

[Dr] V.G. Drinfeld, *Coverings of p-adic symmetric regions*, Functional Anal. i Prilozen. **10** (1976), 29-40; English transl. in Functional Anal. Appl. **10** (1976).

[Gr] A. Grothendieck, SGA7 I, Exposé IX , Lecture Notes in Math. vol. 288, Springer-Verlag, 1972, pp. 313-523..

[Ih] Y. Ihara, *On modular curves over finite fields*, Discrete subgroups of Lie groups and applications to moduli, Bombay Colloquium (1973) pp. 161-202.

[L] R.P. Langlands, *Modular forms and l-adic representations*, Lecture Notes in Math., vol. 349, Springer-Verlag, 1973, pp. 361-500.

[LO] S.Ling and J.Oesterlé, *The Shimura subgroup of $J_0(N)$*, Astérisque **196-197** (1991), 171-203.

[M] B. Mazur, *Modular curves and the Eisenstein ideal*, Publ. Math de IHES **47** (1977), 33-186.

[MR] B. Mazur and K.A.Ribet, *Two-dimensional representations in the arithmetic of modular curves*, Astérisque **196-197** (1991), 215-255.

[O] J. Oesterlé, *Nouvelles Approaches du "Théorème" de Fermat*, Seminaire Bourbaki, Feb. 88, Astérisque 161-162 (1988) pp. 165-186.

[R1] K.A. Ribet, *Mod p Hecke operators and congruences between modular forms*, Inv. Math. **71** (1983), 193-205.

[R2] ———, *Congruence relations between modular forms*, Proc. ICM (1983) 503-514.

[R3] ———, *On modular representations of Gal($\overline{\mathbf{Q}}/\mathbf{Q}$) arising from modular forms*, Inv. Math. **100** (1990), 431-476.

[R4] ———, *Multiplicities of Galois Representations in Jacobians of Shimura Curves*, Israel Mathematical Conference Proceedings **3** (1990), 221-236.

[R5] ———, *From the Taniyama-Shimura Conjecture to Fermat's Last Theorem*, Annales de la Faculté des Sciences de l'université de Toulouse (1990).

[R6] ———, *Report on Mod l Representations of Gal($\overline{\mathbf{Q}}/\mathbf{Q}$)*, Motives, Proceedings of Symposia in Pure Mathematics, vol 55, (1994), pp. 639-676.

[S] J.-P. Serre, *Sur les représentations modulaires de degre 2 de Gal($\overline{\mathbf{Q}}/\mathbf{Q}$)*, Duke Math. J. **54** (1987), 179-230.

TATA INSTITUTE OF FUNDAMENTAL RESEARCH, BOMBAY, 400005, INDIA

Current address: Mehta Research Institute, Allahabad, 211002, INDIA

E-mail address: dprasad@mri.ernet.in