PACIFIC JOURNAL OF MATHEMATICS
Vol. 181, No. 3, 1997

IMAGES OF SEMISTABLE GALOIS REPRESENTATIONS

Kenneth A. Ribet

In memory of Olga Taussky-Todd

1. Introduction.

Let p be an odd prime number. Suppose first that E is a semistable elliptic curve over $\mathbf{Q}$. The action of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ on the group of p-division points of E defines a representation

$$\rho_{E,p}\colon \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathbf{GL}(2, \mathbf{F}_p).$$

Serre has shown that this representation is surjective whenever it is irreducible [26, Prop. 21], [29, §3.1]. Serre's arguments prove more generally the surjectivity of all continuous irreducible representations

$$\rho\colon \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathbf{GL}(2, \mathbf{F}_p)$$

which are *semistable* in the sense of [15, §1]. (Recall that ρ is semistable if the determinant of ρ is the mod p cyclotomic character χ, the Serre conductor of ρ is square free, and the Serre weight of ρ is either 2 or $p+1$. Here, the weight and conductor are the invariants defined in [28].)

In this article, we treat the situation where $\mathbf{F}_p$ is replaced by a finite field of characteristic p, or more generally by a finite product $\mathbf{F} = \prod F_i$ of finite fields of characteristic p. A continuous representation $\rho\colon \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathbf{GL}(2, \mathbf{F})$ may then be viewed as a product of components $\rho_i\colon \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathbf{GL}(2, F_i)$. We shall assume that each ρ_i is irreducible and semistable. Since the determinant of each ρ_i is then the mod p cyclotomic character, the image of ρ is contained in the group

$$A := \{\, M \in \mathbf{GL}(2, \mathbf{F}) \mid \det M \in \mathbf{F}_p^* \,\}.$$

Making the supplementary hypothesis $p \geq 5$, we show below that the image of ρ is $\mathbf{GL}(2, \mathbf{F})$-conjugate to

$$\{\, M \in \mathbf{GL}(2, \mathbf{F}') \mid \det M \in \mathbf{F}_p^* \,\},$$

where $\mathbf{F}'$ is the subalgebra of $\mathbf{F}$ generated by the traces $\operatorname{tr}(\rho(\sigma))$ for $\sigma \in \operatorname{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$. (See Theorem 3.2 and Corollary 3.3.) In particular, $\rho(\operatorname{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})) = A$ if and only if $\mathbf{F}$ is generated as an $\mathbf{F}_p$-algebra by the traces.

Applying the lifting techniques of [25], we deduce an analogous result for certain p-adic representations of $\operatorname{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$.

Our results have evident relevance to the study of the Galois representations defined by semistable abelian varieties over $\mathbf{Q}$ which are products of abelian varieties with large fields of endomorphisms. (In [23], the author referred to these as abelian varieties of "$\mathbf{GL}_2$ type.") While the computation of $\mathbf{F}'$ seems to be difficult to perform in certain cases, it can be carried out for the mod p representations coming from $J_0(N)$ if N is a prime and p is a prime number satisfying some mild conditions. In fact, our original goal was to find the image of the p-adic representation attached to $J_0(N)$, thus answering questions which were formulated by R. Coleman and B. Kaskel in connection with [10] and [2]. It is a pleasure to thank them for their continuing encouragement and interest in this work.

2. Representations over a finite field.

Let $\mathbf{F}$ be a finite field, and let p be the characteristic of $\mathbf{F}$. We will assume that p is odd; most of our results will require that p be at least 5. Suppose that

$$\rho\colon \operatorname{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathbf{GL}(2, \mathbf{F})$$

is an irreducible representation whose determinant is the mod p cyclotomic character $\chi\colon \operatorname{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathbf{F}_p^* \hookrightarrow \mathbf{F}^*$. Let $k(\rho)$ and $N(\rho)$ be the weight and conductor of ρ in the sense of Serre [28, §§1-2]. As we recalled above, Oesterlé [15, §1] has defined the notion of semistability: ρ is *semistable* if the conductor $N(\rho)$ of ρ is square free and Serre's weight $k(\rho)$ is either 2 or $p{+}1$.

The definition of the conductor shows that $N(\rho)$ is square free if and only if $\rho(\sigma)$ is unipotent whenever σ belongs to an inertia subgroup of $\operatorname{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ for a prime $\ell \neq p$. To illuminate the condition on $k(\rho)$, we let I be an inertia subgroup of $\operatorname{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ for the prime p. The semisimplification of $\rho_{|I}$ is described by a pair of characters $\varphi, \varphi'\colon I \rightrightarrows \overline{\mathbf{F}}^*$, cf. [28, §2]. Since $\det\rho$ is the cyclotomic character χ, we have in particular $\varphi\varphi' = \chi$. If $k(\rho)$ is one of 2, $p{+}1$, then $\{\varphi, \varphi'\}$ is either $\{1, \chi\}$ or else the set of fundamental characters $\psi, \psi'\colon I \to \overline{\mathbf{F}}^*$ of level 2 ([28, §2]). It follows that the order of $\varphi'\varphi^{-1}$ (a character which is defined only up to inversion) is either $p{-}1$ or $p{+}1$.

Lemma 2.1. *Assume that $p > 3$ and that ρ is a semistable irreducible* *representation as above. Let* $\epsilon\colon \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathbf{F}^*$ *be a continuous character. If* $\epsilon \otimes \rho$ *is semistable, then* ϵ *is trivial.*

Proof. By definition, the representation $\epsilon \otimes \rho$ sends each $\sigma \in \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ to the product $\epsilon(\sigma)\cdot\rho(\sigma) \in \mathbf{GL}(2, \mathbf{F})$. If ρ and $\epsilon\otimes\rho$ are both semistable, then ϵ is certainly unramified outside p, since both $\rho(\sigma)$ and $\epsilon(\sigma)\cdot\rho(\sigma)$ are unipotent whenever σ belongs to an inertia group for a prime $\ell \neq p$. Moreover, ϵ^2 is the trivial character, since the determinants of ρ and $\epsilon \otimes \rho$ coincide. Hence ϵ is either the trivial character, as desired, or the quadratic character $\chi^{\frac{p-1}{2}}$. One checks easily, however, that no pair of characters drawn from the set $\{1, \chi, \psi, \psi'\}$ have a quadratic ratio. ☐

Remark. The lemma does not extend to cover the case $p = 3$. To see this, consider the modular forms $f = \sum a_n q^n$ and $g = \sum b_n q^n$ which correspond to the two strong modular elliptic curves of conductor 89. As B. Gross explains in the last paragraphs of [8], these forms define irreducible mod 3 representations ρ_f and ρ_g of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ whose restrictions to a decomposition group for 3 are direct sums of two characters. Further, although ρ_f and ρ_g are semistable, the two representations are twists of each other by the mod 3 cyclotomic character. (In other words, f and g are "companions" of each other in Serre's language.)

To verify this latter fact, we can cite [8, Th. 13.10] along with Gross, or perform a numerical calculation with `gp` [1] to check that the mod 3 congruence $b_p \equiv \left(\frac{p}{3}\right) a_p$ holds for $3 \le p \le 200$. Using the results of J. Sturm [33], we can then conclude that it holds for all $p \ge 3$.

It perhaps is worth recalling at this juncture that an irreducible mod p semistable representation ρ is absolutely irreducible. This follows easily from the fact that $\rho\big(\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})\big)$ contains a matrix with distinct eigenvalues in $\mathbf{F}_p^*$. Such a matrix is obtained by taking $\rho(c)$, where c is a complex conjugation — because $\det \rho$ is the cyclotomic character, which is odd, the eigenvalues of $\rho(c)$ are $+1$ and -1. These are distinct because p is odd.

Proposition 2.2. *If* ρ *is semistable, then the image of* ρ *has order divisible by* p.

Proof. Let $G = \rho\big(\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})\big)$, and assume that the order of G is prime to p. Since the semistability hypothesis implies that ramification subgroups of G for primes $\ell \neq p$ are unipotent, these ramification groups are forced to be trivial. In other words, we have $N(\rho) = 1$ and the representation ρ is unramified outside p. As Serre remarks in a note on page 710 of [27, Vol. III], an analogue of the argument of Tate [34] shows that there are no irreducible representations ρ with this property when $p = 3$.

Now assume that $p \geq 5$ and write $\overline{G}$ for the image of G in $\mathbf{PGL}(2, \mathbf{F})$. Group theory shows that $\overline{G}$ is either cyclic or dihedral, or else one of the three exceptional groups $\mathbf{S}_4$, $\mathbf{A}_4$, $\mathbf{A}_5$ [26, §2.5]. In fact, $\overline{G}$ cannot be cyclic, since the cyclicity of $\overline{G}$ would imply that G is abelian, and hence that ρ is not absolutely irreducible.

To rule out the other cases, one considers an inertia subgroup $\overline{I}$ of $\overline{G}$ for the prime p. We know that $\overline{I}$ is a cyclic group of order either $p+1$ or $p-1$. Indeed, $\overline{I}$ may be viewed as the image of $I \subset \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ under the character φ/φ' which we introduced above. This character has order either $p-1$ or $p+1$; its image is cyclic because it is a finite subgroup of $\overline{\mathbf{F}}^*$.

Assume now that $\overline{G}$ is dihedral, and let Z be the center of $\overline{G}$. It is evident that $\overline{I}$ is contained in Z, since $\overline{I}$ is a cyclic subgroup of a dihedral group and the order of $\overline{I}$ is greater than 2. Accordingly, the quadratic extension of $\mathbf{Q}$ corresponding to Z is everywhere unramified—this contradiction excludes the dihedral case and shows that $\overline{G}$ must be one of the three exceptional groups.

However, as observed in the proof of [26, Prop. 21], the fact that $\overline{I}$ has an element of order $p \pm 1$ rules out the groups $\mathbf{S}_4$, $\mathbf{A}_4$, $\mathbf{A}_5$ in case $p \geq 7$. Thus we are left only with the possibility that $p = 5$, in which case $\overline{G}$ is either $\mathbf{S}_4$ or $\mathbf{A}_4$, since its order is prime to 5 by assumption. The group $\overline{I}$ is then cyclic of order 4, since $\mathbf{S}_4$ has no element of order 6. Also, we have $\overline{G} \approx \mathbf{S}_4$, since $\mathbf{A}_4$ has no element of order 4. Consider the quotient $\mathbf{S}_3$ of $\mathbf{S}_4$. This quotient allows us to produce an $\mathbf{S}_3$-extension of $\mathbf{Q}$ which is ramified only at 5 and such that the inertia groups for 5 in the extension have order 2. However, there certainly is no such extension, since the class number of $\mathbf{Q}(\sqrt{5})$ is 1. □

Corollary 2.3 (cf. [29, Prop. 1])**.** *Let ρ be as in Proposition 2.2, and suppose that $p > 2$. Then the image of ρ contains a subgroup isomorphic to $\mathbf{SL}(2, \mathbf{F}_p)$. In particular, if $\mathbf{F} = \mathbf{F}_p$, then ρ is surjective.*

Proof. The second statement is a consequence of the first, since the cyclotomic character χ: $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathbf{F}_p^*$ is surjective. To prove the first statement, we proceed as in [26, §2.4]. Namely, let $g \in G$ be an element of order p, and let v be a non-zero vector in $\mathbf{F} \oplus \mathbf{F}$ which is fixed by g. Since ρ is irreducible, G cannot fix the line generated by v; therefore, there is an $r \in G$ such that v and rv form a basis of $\mathbf{F} \oplus \mathbf{F}$. With respect to this basis, g has the form $\begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix}$, while rgr^{-1} has the form $\begin{pmatrix} 1 & 0 \\ b & 1 \end{pmatrix}$. Scaling one of the vectors v, rv, we may assume that $a = 1$. Hence, in an appropriate basis, G contains the elements $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ and $\begin{pmatrix} 1 & 0 \\ b & 1 \end{pmatrix}$. A well known theorem of L. E. Dickson (cf. [7, Th. 2.8.4]) states that the group generated by these elements has a subgroup isomorphic to $\mathbf{SL}(2, \mathbf{F}_p)$. In fact, a more precise statement is true for $p \neq 3$: $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ and $\begin{pmatrix} 1 & 0 \\ b & 1 \end{pmatrix}$ generate the group $\mathbf{SL}(2, \mathbf{F}')$, where $\mathbf{F}' = \mathbf{F}_p(b)$. □

Our aim now is to complement the Corollary by determining the exact image of ρ in a situation generalizing that where $\mathbf{F} = \mathbf{F}_p$. The situation which we have in mind is that where $\mathbf{F}$ is a minimal field of definition for ρ in the sense that it is generated over $\mathbf{F}_p$ by the numbers $\operatorname{tr}(\rho(\sigma))$ for $\sigma \in \operatorname{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$. We first show in this case that $\mathbf{F}$ is also a minimal field of definition for ρ as a projective representation, at least when $p \geq 5$.

For the following lemma, let $\overline{\mathbf{F}}$ be an algebraic closure of $\mathbf{F}$ and consider an arbitrary subfield K of $\overline{\mathbf{F}}$. We view G and $\overline{G}$ inside $\mathbf{GL}(2, \overline{\mathbf{F}})$ and $\mathbf{PGL}(2, \overline{\mathbf{F}})$, respectively.

Lemma 2.4. *Suppose that ρ semistable and that $p \geq 5$. Then $\overline{G}$ lies in $\mathbf{PGL}(2, K)$ if and only if G lies in $\mathbf{GL}(2, K)$.*

Proof. If G lies in $\mathbf{GL}(2, K)$, then it is evident that $\overline{G}$ is contained in $\mathbf{PGL}(2, K)$. Conversely, suppose $\overline{G} \subseteq \mathbf{PGL}(2, K)$. Then certainly $G \subseteq \overline{\mathbf{F}}^* \cdot \mathbf{GL}(2, K)$. Let $\alpha \colon \operatorname{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \overline{\mathbf{F}}^*/K^*$ be the composite homomorphism

$$\operatorname{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \xrightarrow{\rho} G \hookrightarrow \overline{\mathbf{F}}^* \mathbf{GL}(2, K) \to \big(\overline{\mathbf{F}}^* \mathbf{GL}(2, K)\big)/\mathbf{GL}(2, K) = \overline{\mathbf{F}}^*/K^*.$$

For $\sigma \in \operatorname{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$, we have $\alpha(\sigma) = 1$ whenever the trace of $\rho(\sigma)$ is a non-zero element of K. Indeed, write $\rho(\sigma) = t \cdot M$, with $M \in \mathbf{GL}(2, K)$ and $t \in \overline{\mathbf{F}}^*$. Then $\operatorname{tr}(\rho(\sigma)) = t \cdot \operatorname{tr}(M)$ and we have $\operatorname{tr} M \in K$. If $\operatorname{tr}(\rho(\sigma))$ belongs to K^*, then t lies in K, so that $\rho(\sigma)$ is an element of $\mathbf{GL}(2, K)$. In particular, $\alpha(\sigma) = 1$ whenever the trace of $\rho(\sigma)$ is a non-zero element of $\mathbf{F}_p$.

Let M now be the finite abelian extension of $\mathbf{Q}$ which is cut out by α. We seek to show that α is identically 1, i.e., that $M = \mathbf{Q}$. We first prove that α is unramified outside p by using the remark about traces. If σ belongs to an inertia subgroup of $\operatorname{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ for a prime $\ell \neq p$, then $\rho(\sigma)$ is unipotent, so that its trace is 2. Since p is odd, 2 is a non-zero element of $\mathbf{F}_p$, and we may conclude that $\alpha(\sigma) = 1$.

Thus M is a finite abelian extension of $\mathbf{Q}$ which is unramified outside p. Moreover, $[M\colon \mathbf{Q}]$ is prime to p, since $\overline{\mathbf{F}}^*$ has no elements of order p. Hence one has $M \subseteq \mathbf{Q}(\mu_p)$; equivalently, α factors through the mod p cyclotomic character χ.

Now let I be an inertia group for p in $\operatorname{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$. It suffices to show that there is an element σ of I for which $\alpha(\sigma) = 1$ and for which $\chi(\sigma)$ is a generator of the cyclic group $\mathbf{F}_p^*$. The semisimplification of $\rho_{|I}$ is described by a pair of characters $\varphi, \varphi' \colon I \rightrightarrows \overline{\mathbf{F}}^*$, cf. [28, §2]. As we mentioned above, one has either $\{\varphi, \varphi'\} = \{1, \chi\}$ or $\{\varphi, \varphi'\} = \{\psi, \psi^p\}$, where ψ and ψ^p are the fundamental characters of level 2. Suppose first that we are in the former case, and let $\sigma \in I$ be such that $t = \chi(\sigma)$ is a generator of $\mathbf{F}_p^*$. Then $\operatorname{tr}(\rho(\sigma)) = 1 + t$ is non-zero since $p \neq 3$. Thus $\alpha(\sigma) = 1$, as required. In the latter case, choose $\sigma \in I$ so that $t = \psi(\sigma)$ generates $\psi(I) \approx \mathbf{F}_{p^2}^*$. Then $\chi(\sigma) = t^{p+1}$ is a generator of $\mathbf{F}_p^*$; note that $\chi = \psi\psi' = \psi^{p+1}$. On the other hand, $\operatorname{tr}(\rho(\sigma)) = t + t^p$. The number t^{p-1} cannot be -1, since it has order $p+1$. Since $\operatorname{tr}(\rho(\sigma))$ is non-zero, we may conclude $\alpha(\sigma) = 1$. □

Theorem 2.5. *Assume that ρ is semistable, that $p \geq 5$ and that $\mathbf{F}$ is generated over $\mathbf{F}_p$ by the set $\{ \operatorname{tr}(\rho(\sigma)) \mid \sigma \in \operatorname{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \}$. Then*

$$\rho\big(\operatorname{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})\big) = \{ M \in \mathbf{GL}(2, \mathbf{F}) \mid \det M \in \mathbf{F}_p^* \}.$$

Proof. It is clear that $G = \rho\big(\operatorname{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})\big)$ is contained in the indicated matrix group because $\det \rho$ is $\mathbf{F}_p^*$-valued. Since the image of $\det \rho$ is precisely $\mathbf{F}_p^*$, the theorem amounts to the statement that G contains $\mathbf{SL}(2, \mathbf{F})$. An equivalent assertion is that the commutator subgroup of G contains a subgroup isomorphic to $\mathbf{SL}(2, \mathbf{F})$.

Let k be a finite extension of $\mathbf{F}_p$ which contains $\mathbf{F}$ and which has even degree over $\mathbf{F}_p$. Since $\det(G) \subseteq \mathbf{F}_p^*$, the subgroup $\overline{G}$ of $\mathbf{PGL}(2, k)$ lies in $\mathbf{PSL}(2, k)$. Dickson [4, Ch. XII] has enumerated all subgroups of $\mathbf{PSL}(2, k)$; his list is summarized in [4, §260].

To situate $\overline{G}$ within Dickson's list, we recall that the order of $\overline{G}$ is divisible by p by Proposition 2.2, and that the identity representation $G \to \mathbf{GL}(2, \overline{\mathbf{F}})$ is irreducible. (The representation ρ is irreducible over $\mathbf{F}$ by hypothesis. It is then absolutely irreducible, as was noted above.) It follows that $\overline{G}$ is one of the groups enumerated in [4, §251–§253]. Since we have assumed $p \geq 5$, the final conclusion is easy to state: After replacing $\overline{G}$ by a conjugate of $\overline{G}$ inside $\mathbf{PGL}(2, k)$, we have either $\overline{G} = \mathbf{PSL}(2, K)$ or $\overline{G} = \mathbf{PGL}(2, K)$, for some subfield K of k.

Thus, in either case one has $\overline{G} \subseteq \mathbf{PGL}(2, K)$. From the Lemma, $G \subseteq \mathbf{GL}(2, K)$. In particular, $\operatorname{tr}(\rho(\sigma)) \in K$ for all $\sigma \in \operatorname{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$. Since these numbers generate $\mathbf{F}$ over $\mathbf{F}_p$, $\mathbf{F} \subseteq K$. On the other hand, $\mathbf{SL}(2, K) \subseteq k^* \cdot G$ because $\overline{G}$ contains $\mathbf{PSL}(2, K)$. On taking commutators, we obtain $\mathbf{SL}(2, K) \subseteq [G, G]$, and therefore $\mathbf{SL}(2, F) \subseteq [G, G]$. As indicated above, this proves the theorem. □

In the spirit of [26, Th. 4], let us choose an inertia group I for p in $\operatorname{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ and define X to be the smallest closed normal subgroup of $\operatorname{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ which contains I.

Theorem 2.6. *In the situation of Theorem 2.5*

$$\rho(X) = \{\, M \in \mathbf{GL}(2, \mathbf{F}) \mid \det M \in \mathbf{F}_p^* \,\}.$$

Proof. By Theorem 2.5, the group $G = \rho(\,\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}))$ is the group of matrices in $\mathbf{GL}(2, \mathbf{F})$ with determinants in $\mathbf{F}_p^*$. Let $H = \rho(X)$; thus H is a normal subgroup of G. Let $\bar{H}$ and $\bar{G}$ be the images of H and G in $\mathbf{PGL}(2, \mathbf{F})$. The group $\bar{G}$ is either $\mathbf{PSL}(2, \mathbf{F})$ or $\mathbf{PGL}(2, \mathbf{F})$, according as the degree of $\mathbf{F}$ over $\mathbf{F}_p$ is even or odd. The discussion above shows that the order of H is at least $p - 1 \geq 4$. Therefore, the intersection $H \cap \mathbf{PSL}(2, \mathbf{F})$ has order at least 2. Since $\mathbf{PSL}(2, \mathbf{F})$ is normal in $\bar{G}$, $H \cap \mathbf{PSL}(2, \mathbf{F})$ is a non-trivial normal subgroup of $\mathbf{PSL}(2, \mathbf{F})$. Accordingly, it is all of $\mathbf{PSL}(2, \mathbf{F})$; in other words, $\bar{H}$ contains $\mathbf{PSL}(2, \mathbf{F})$. On taking commutators as above, we see that H contains $\mathbf{SL}(2, \mathbf{F})$. Because the mod p cyclotomic character maps I onto $\mathbf{F}_p^*$, it follows now that $H = G$. □

Remark 1. In the context of Theorem 2.6, one may consider the more general situation where $\mathbf{F}$ is not necessarily generated by the traces $\mathrm{tr}(\rho(\sigma))$ for $\sigma \in \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$. Let $\mathbf{F}'$ be the subfield of $\mathbf{F}$ which is generated by these traces. Because complex conjugation acts in ρ as a matrix with distinct rational eigenvalues, a well known theorem of I. Schur [24, IX a] (cf. [35, Lemme I.1]) implies that ρ can be conjugated into a representation with values in $\mathbf{GL}(2, \mathbf{F}')$. The theorem applies to this latter representation, and shows that its image is the group of matrices in $\mathbf{GL}(2, \mathbf{F}')$ whose determinants lie in the multiplicative group of the prime field $\mathbf{F}_p$.

To prove the well known statement that there is a model for ρ over $\mathbf{F}'$, one may proceed alternatively by direct computation, along the lines suggested by Wiles [36, p. 483].

Remark 2. The referee has asked whether Theorem 2.5 extends to the case $p = 3$. The answer is negative; in fact, a counterexample is furnished by the abelian surface $J_0(23)$. Recall that $J_0(23)$ has "real multiplication" by the Hecke ring $\mathbf{T}$ associated with the space of weight-two cusp forms on $\Gamma_0(23)$. The algebra $\mathbf{T} \otimes \mathbf{Q}$ is the real quadratic field $\mathbf{Q}(\sqrt{5})$, and $\mathbf{T}$ is the ring of integers of $\mathbf{T} \otimes \mathbf{Q}$. The group V of 3-division points on $J_0(23)$ is a two-dimensional vector space over the field $\mathbf{F} := \mathbf{T}/3\mathbf{T}$, which has nine elements. The action of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ on V is given by a homomorphism $\rho\colon \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathrm{Aut}_{\mathbf{F}} V \approx \mathbf{GL}(2, \mathbf{F})$. As we shall see in §5, this representation is semistable; it is irreducible because 3 is prime to $23 - 1$. Moreover, the traces $\mathrm{tr}(\rho(\sigma))$ generate $\mathbf{F}$; indeed, the trace of a Frobenius element Frob_2 for the prime 2 satisfies $x^2 + x - 1 = 0$. In view of Theorem 2.5, it is very tempting to guess that the image of ρ contains $\mathbf{SL}(2, \mathbf{F})$. Equivalently, consider the composite of ρ and the natural homomorphism $\mathbf{GL}(2, \mathbf{F}) \to \mathbf{PGL}(2, \mathbf{F})$; let $\bar{G}$ be the image of this composite, so that $\bar{G}$ is a subgroup of $\mathbf{PSL}(2, \mathbf{F})$. The guess that the image of ρ contains $\mathbf{SL}(2, \mathbf{F})$ means that $\bar{G} = \mathbf{PSL}(2, \mathbf{F})$.

What information do we have about $\bar{G}$? The group $\bar{G}$ is "irreducible" (i.e., acts transitively on $\mathbf{P}^1(\mathbf{F})$) since ρ is irreducible. Also, the order of $\bar{G}$ is divisible by 3, as one sees by considering an inertia subgroup of G for the prime 23. Finally, the order of $\bar{G}$ is divisible by 5 because of the information concerning $\operatorname{tr}(\rho(\mathrm{Frob}_2))$, which implies that $\rho(\mathrm{Frob}_2)$ has order 5. The results of Dickson used above thus permit only two possibilities for $\bar{G}$: either $\bar{G}$ is all of $\mathbf{PSL}(2, \mathbf{F})$, or else the alternating group $\mathbf{A}_5$.

Rather to the author's surprise, calculations based on [14, Table B] suggested strongly that $\bar{G}$ is in fact the smaller of these two groups. The author's suspicion that this was the case deepened when he learned that there is an $\mathbf{A}_5$-extension of $\mathbf{Q}$ which is ramified only at 3 and 23: the second line of [6, Table 1] shows that the splitting field of the polynomial $x^5 + 3x^3 + 6x^2 + 9$ is such an extension. Subsequently, Jean-François Mestre carried out computations which confirm that this latter $\mathbf{A}_5$-extension is indeed the extension of $\mathbf{Q}$ which is cut out by the projective representation deduced from ρ. In particular, one has $\bar{G} \approx \mathbf{A}_5$.

3. Products.

We next consider finite products of representations as above, keeping fixed the prime number p. Thus let $F_1, \dots, F_t$ $(t \geq 1)$ be finite fields of characteristic p, where p is a prime which is different from 2 and 3. Let $\mathbf{F}$ be the finite étale $\mathbf{F}_p$-algebra $F_1 \times \cdots \times F_t$. Suppose that $\rho\colon \operatorname{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathbf{GL}(2, \mathbf{F})$ is a continuous homomorphism, so that ρ is a product of representations ρ_i $(i = 1, \dots, t)$ as above. We will assume that each ρ_i is semistable and irreducible, and also that $\det \rho_i = \chi$ for $i = 1, \dots, t$. With the evident convention, the latter hypothesis may be summarized by the formula $\det \rho = \chi$.

For each $\sigma \in \operatorname{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$, we obtain an element $\operatorname{tr}(\rho(\sigma))$ of $\mathbf{F}$ by considering the trace of the matrix $\rho(\sigma)$. Motivated by the remark above, we let $\mathbf{F}'$ be the subalgebra of $\mathbf{F}$ generated by the $\operatorname{tr}(\rho(\sigma))$.

Lemma 3.1. *The representation ρ has a model over $\mathbf{F}'$.*

Proof. As we have seen, each component ρ_i has a model over the subfield of F_i generated by the traces $\operatorname{tr}(\rho_i(\sigma))$. After replacing ρ_i by such a model (and F_i by the trace field in question), we arrive at a situation in which $\mathbf{F}'$ maps surjectively onto each factor F_i.

As one knows, $\mathbf{F}'$ is then isomorphic to a partial product of the factors F_i. To see this explicitly, we let $\pi\colon \mathbf{F} \to F_i$ be the projections $(x_1, \dots, x_t) \mapsto x_i$

and consider the following relation on the set $\{1, \dots, t\}$: $i \sim j$ if and only if the map $\pi_i \times \pi_j : \mathbf{F}' \to F_i \times F_j$ is *not* surjective. It is easy to see that this relation is an equivalence relation and that $i \sim j$ if and only if there is an isomorphism $\sigma\colon F_i \xrightarrow{\sim} F_j$ so that $\pi_j = \sigma \circ \pi_i$ on $\mathbf{F}'$. If there is such an isomorphism, it is unique; we denote it σ_{ji}. One shows that

$$\mathbf{F}' = \{(x_1, \dots, x_t) \in \mathbf{F} \mid x_j = \sigma_{ji}(x_i) \text{ for all pairs } (i,j) \text{ such that } i \sim j\}.$$

In particular, $\mathbf{F}'$ is isomorphic to the product $\prod_{i \in I} F_i$, where I is a set of representatives for the equivalence $\sim$.

By the Brauer-Nesbitt theorem, ρ_j and ${}^{\sigma_{ji}}\rho_i$ are isomorphic whenever i and j are equivalent. (The two representations have the same trace and determinant.) Replace ρ_j by ${}^{\sigma_{ji}}\rho_i$ for all equivalent pairs (i,j) with $i \in I$. Then the representation ρ, a priori with values in $\mathbf{GL}(2,\mathbf{F})$, takes values in $\mathbf{GL}(2,\mathbf{F}')$. □

Theorem 3.2. *One has*

$$\rho(\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})) = \{M \in \mathbf{GL}(2,\mathbf{F}) \mid \det M \in \mathbf{F}_p^*\}$$

if and only if $\mathbf{F}' = \mathbf{F}$.

Proof. The necessity is clear, since the trace function $\mathbf{SL}(2,\mathbf{F}) \to \mathbf{F}$ is surjective. For the sufficiency, as in the proof of Theorem 2.5, one must show that $\rho(\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}))$ contains $\mathbf{SL}(2,\mathbf{F})$. Let $H = \rho(\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})) \cap \mathbf{SL}(2,\mathbf{F})$, so that H is a subgroup of the product $\mathbf{SL}(2,F_1) \times \cdots \times \mathbf{SL}(2,F_t)$. By Theorem 2.5, H projects onto each factor $\mathbf{SL}(2,F_i)$. Because each group $\mathbf{SL}(2,F_i)$ is its own commutator subgroup, the "two principle" [19, 3.3] implies that H is the full product of the $\mathbf{SL}(2,F_i)$ if and only if H maps onto each product $\mathbf{SL}(2,F_i) \times \mathbf{SL}(2,F_j)$ for $i \neq j$.

Assume, then, that we have $i \neq j$, and suppose that the image of the product $\rho_i \times \rho_j$ does *not* contain $\mathbf{SL}(2,F_i) \times \mathbf{SL}(2,F_j)$. By exploiting results of Dieudonné and Hua, one may construct: (i) An isomorphism $\omega\colon F_i \xrightarrow{\sim} F_j$, and (ii) a continuous homomorphism $\epsilon\colon \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to F_j^*$ such that ${}^{\omega}\rho_i$ and $\rho_j \otimes \epsilon$ are isomorphic representations of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ over F_j. This statement was proved during the course of the proof of [19, Th. 3.8] — the theorem itself states merely that ${}^{\omega}\rho_i$ and $\rho_j \otimes \epsilon$ have equal traces and determinants.

By Lemma 3.1, however, we have $\epsilon = 1$; thus ${}^{\omega}\rho_i \approx \rho_j$. Accordingly, we have $\omega(\mathrm{tr}(\rho_i(\sigma))) = \mathrm{tr}(\rho_j(\sigma))$ for all $\sigma \in \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$. This means that the image of $\mathrm{tr}(\rho(\sigma))$ in $F_i \times F_j$ lies in the subalgebra $\{(x, \omega x) \mid x \in F_i\}$ of $F_i \times F_j$. Thus the quantities $\mathrm{tr}(\rho(\sigma))$ fail to generate $F_i \times F_j$ over $\mathbf{F}_p$. This contradicts the hypothesis that they generate the full product $\mathbf{F}$ over $\mathbf{F}_p$. □

Corollary 3.3. *Assume that ρ is as in Theorem 3.2. Then the image of ρ is conjugate in $\mathbf{GL}(2,\mathbf{F})$ to the group*

$$\{M \in \mathbf{GL}(2,\mathbf{F}') \mid \det M \in \mathbf{F}_p^*\}.$$

Proof. By Lemma 3.1, there is a model for ρ over $\mathbf{F}'$: after conjugating by a matrix in $\mathbf{GL}(2,\mathbf{F})$, we make ρ take values in $\mathbf{GL}(2,\mathbf{F}')$. Applying the Theorem to this model, we arrive at the desired conclusion. □

We continue the discussion begun with Theorem 2.6, letting X be the subgroup of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ which appears in the statement of that result.

Theorem 3.4. *In the situation of Theorem 3.2, suppose that $\mathbf{F} = \mathbf{F}'$. Then $\rho(X) = \{\, M \in \mathbf{GL}(2,\mathbf{F}) \mid \det M \in \mathbf{F}_p^* \,\}$.*

Proof. The group $\rho(X)$ is a normal subgroup of $G = \rho\big(\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})\big)$. By Theorem 3.2, $G = \{\, M \in \mathbf{GL}(2,\mathbf{F}) \mid \det M \in \mathbf{F}_p^* \,\}$. The intersection $\rho(X) \cap \mathbf{SL}(2,\mathbf{F})$ is then normal in $\mathbf{SL}(2,\mathbf{F})$; it maps onto each factor $\mathbf{SL}(2,F_i)$ by Theorem 2.6. On taking commutators with elements of $\mathbf{SL}(2,\mathbf{F})$ of the form $(1,\dots,1,\alpha,1,\dots,1)$, we see that $\rho(X)\cap\mathbf{SL}(2,\mathbf{F})$ contains $\mathbf{SL}(2,F_i)$ (viewed as a subgroup of $\mathbf{SL}(2,\mathbf{F})$) for each i. It follows that $\rho(X)$ contains $\mathbf{SL}(2,\mathbf{F})$. We then obtain $\rho(X) = G$, since the mod p cyclotomic character is totally ramified at p. □

4. Lifts.

Again suppose that p is a prime ≥ 5 and let $\mathcal{O}_1,\dots,\mathcal{O}_t$ be integer rings of finite-degree unramified extensions of $\mathbf{Q}_p$. For each i, let $\tilde{\rho}_i\colon \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathbf{GL}(2,\mathcal{O}_i)$ be a continuous representation whose determinant is the p-adic cyclotomic character $\tilde{\chi}\colon \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathbf{Z}_p^*$. Let $\tilde{\rho}$ be the product of the $\tilde{\rho}_i$, so that $\tilde{\rho}$ is a p-adic representation $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathbf{GL}(2,\mathcal{O})$, where $\mathcal{O} := \prod \mathcal{O}_i$. The diagonal embedding $\mathbf{Z}_p \hookrightarrow \mathcal{O}$ induces an inclusion $\mathbf{Z}_p^* \hookrightarrow \mathcal{O}^*$. We clearly have $\tilde{\rho}\big(\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})\big) \subseteq \mathcal{A}$, where

$$\mathcal{A} = \{\, M \in \mathbf{GL}(2,\mathcal{O}) \mid \det M \in \mathbf{Z}_p^* \,\}.$$

Let $\mathbf{F} = \mathcal{O}/p\mathcal{O}$, and let $\rho\colon \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathbf{GL}(2,\mathbf{F})$ be the reduction of $\tilde{\rho}$ mod p.

In the following statement, X is again the closed normal subgroup of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ which is generated by the inertia groups for p in $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$.

Theorem 4.1. *Suppose that the mod p reduction of each $\tilde{\rho}_i$ is semistable and irreducible. Assume that* $\mathbf{F}$ *is generated as an* $\mathbf{F}_p$*-algebra by the traces* $\operatorname{tr}\rho(g)$ *with* $g \in \operatorname{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$. *Then* $\tilde{\rho}(X) = \tilde{\rho}(\operatorname{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})) = \mathcal{A}$.

Remark. Using Nakayama's Lemma, one may reformulate the trace hypothesis as the apparently stronger assertion that $\mathcal{O}$ is generated as a $\mathbf{Z}_p$-algebra by the traces $\operatorname{tr}\tilde{\rho}(g)$ with $g \in \operatorname{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$.

Proof. The group $\tilde{\rho}(\operatorname{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}))$ and its subgroup $\tilde{\rho}(X)$ are closed subgroups of $\mathcal{A}$ whose determinants are equal to all of $\mathbf{Z}_p^*$. It suffices to show that $\tilde{\rho}(X)$ contains $\mathbf{SL}(2,\mathcal{O})$. We recall the following fact (see [25, p. IV-23] and [19, Th. 2.1]):

Proposition 4.2. *Let $\mathcal{G}$ be a closed subgroup of* $\mathbf{GL}(2,\mathcal{O})$ *and let $\bar{\mathcal{G}}$ be the image of $\mathcal{G}$ in* $\mathbf{GL}(2,\mathcal{O}/p\mathcal{O})$. *If $\bar{\mathcal{G}}$ contains* $\mathbf{SL}(2,\mathcal{O}/p\mathcal{O})$, *then $\mathcal{G}$ contains* $\mathbf{SL}(2,\mathcal{O})$.

Proof. On taking $\mathcal{G} = \tilde{\rho}(X)$, we now obtain the required inclusion $\tilde{\rho}(X) \supseteq \mathbf{SL}(2,\mathcal{O})$ from Theorem 3.4. □

5. Semistable abelian varieties of $\mathbf{GL}_2$-type over $\mathbf{Q}$.

Let A be an abelian variety over $\mathbf{Q}$ for which $\mathbf{Q} \otimes \operatorname{End}_{\mathbf{Q}} A$ is a number field of degree equal to the dimension of A. Suppose that the ring $R = \operatorname{End}_{\mathbf{Q}} A$ is the full ring of integers in the field $\mathbf{Q} \otimes \operatorname{End}_{\mathbf{Q}} A$. (After replacement of A by an isogenous abelian variety, this hypothesis is always verified.) Let $\mathfrak{m}$ be a maximal ideal of R and let $A[\mathfrak{m}]$ be the kernel of $\mathfrak{m}$ on A, i.e., the group of points in $A(\overline{\mathbf{Q}})$ which are annihilated by all elements of $\mathfrak{m}$. It is easy to check that $A[\mathfrak{m}]$ is free of rank two over $\mathbf{F} = R/\mathfrak{m}$, cf. [32, Prop. 10, p. 56]. In fact, the representation

$$\rho\colon\ \operatorname{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathbf{GL}(2,\mathbf{F})$$

defined by the action of $\operatorname{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ on $A[\mathfrak{m}]$ is just the mod $\mathfrak{m}$ reduction of the $\mathfrak{m}$-adic representation of $\operatorname{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ which is attached to A. As usual, we denote by p the characteristic of $\mathbf{F}$.

Proposition 5.1. *Assume that p is odd, that ρ is irreducible, and that A is semistable over* $\mathbf{Q}$. *Then ρ is a semistable representation.*

Proof. We first show that the determinant of ρ is the mod p cyclotomic character χ. According to the statement of [23, Lemma 3.1], we have $\det\rho = \epsilon\chi$, where ϵ is a Dirichlet character $\operatorname{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to R^*$ which depends only on A (i.e., is independent of $\mathfrak{m}$) and which is ramified only at primes of bad reduction for A. In fact, the proof of this Lemma shows that ϵ is unramified at a prime q whenever the following condition is satisfied: there is a prime λ of R such that the determinant of the λ-adic representation of A is unramified at q. Now fix q and take any prime λ not dividing q. By a well known result of Grothendieck [9, Prop. 3.5], each element of an inertia subgroup for q in $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ acts unipotently in the λ-adic representation for A. In particular, the determinant of this representation is unramified at q. Thus ϵ is unramified at every prime number q, so that ϵ is the trivial character.

The same proposition of Grothendieck, applied to the $\mathfrak{m}$-adic representation for A, shows that the conductor of ρ is square free. We may paraphrase this statement by saying that ρ is semistable outside p. Furthermore, results of Raynaud [16] imply that $k(\rho) = 2$, so that ρ is semistable at p, whenever ρ is finite at p. (See [28, Prop. 4, p. 189].) It remains only to show that $k(\rho) = p+1$ if ρ is not finite at p.

For this, we consider $A[\mathfrak{m}]$ as a subgroup of $A[p]$, and view both as modules for a decomposition group D_p in $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ for the prime p. As explained in the proof of [21, Lemma 6.2], one may deduce from Grothendieck's study in SGA7I that $A[p]$ is an extension of an unramified D_p-module by a subgroup $A[p]^{\mathrm{f}}$ which is finite, i.e., which extends to a finite flat group scheme over $\mathbf{Z}_p$. Were $A[\mathfrak{m}]$ contained in $A[p]^{\mathrm{f}}$, $A[\mathfrak{m}]$ would be finite, contrary to assumption. Hence $A[\mathfrak{m}]$ has an unramified quotient. This implies that the restriction of ρ to D_p has the form $\begin{pmatrix} \theta_1\chi & * \\ 0 & \theta_2 \end{pmatrix}$, where the θ_i are unramified characters. The recipe for $k(\rho)$ given in [28, §2] then sets $k(\rho) = p+1$. (Compare Remarque (1) on page 188 of [28].) □

6. Application to $J_0(N)$ for prime N.

Let N be a prime number, and consider the abelian variety $J = J_0(N)$ over $\mathbf{Q}$. In this section and the next, we will study the action of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ on torsion points of J. The first work in this direction was the investigation of Shimura [30], which concerns the mod p representation of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ defined by $J_0(11)$, when p lies between 7 and 97. Shimura's discussion was completed by Serre [26, §5.5], who determined for all p the image of the mod p representation attached to $J_0(11)$. Subsequently, Lang-Trotter [12], Part I, §8; calculated the image of the (adelic) representation of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ defined by *all* torsion points of $J_0(11)$.

In what follows, we generalize some of the results of Serre and Lang-Trotter to the case where 11 is replaced by an arbitrary prime. We exploit the insights of Mazur's article [13], which makes an extensive study of the action of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ on torsion points of $J_0(N)$, and the recent thesis of B. Kaskel [10], which determines the image of the adelic representation defined by $J_0(37)$. There is a certain amount of overlap with the joint article [2].

Let $\mathbf{T}$ be the ring of Hecke operators $\mathbf{Z}[\dots, T_n, \dots]$, considered as a ring of endomorphisms of $J_0(N)$. To orient the reader, we recall that $\mathbf{T}$ is the full ring of endomorphisms of $J_0(N)$ [13, p. 95]. To avoid the situation where N is very small, we will assume that $J_0(N) \neq 0$. This means that $N = 11$ or that $N \geq 17$.

Let p be a prime number ≥ 5.

Proposition 6.1. *The quotient $\mathbf{T}/p\mathbf{T}$ is generated by the operators T_n with n prime to pN.*

Proof. Let S be the space of cusp forms of weight 2 on $\Gamma_0(N)$ over $\mathbf{F}_p$. We view S as the mod p reduction of the space of weight-two forms on $\Gamma_0(N)$ with integral q-expansions. (See [13, Ch. II, §4] for a comparison of several possible definitions of S.) We consider the bilinear pairing $\alpha\colon \mathbf{T}/p\mathbf{T} \times S \to \mathbf{F}_p$ which maps (T, f) to the initial coefficient of q in the Fourier expansion of $f|T$. It is well known that α is a perfect pairing; this is explained, for example, in [20, §1]. To prove the Proposition, then, it suffices to prove that there is no non-zero element $f = \sum a_n q^n$ of S which satisfies $a_n = 0$ for all n prime to pN.

Suppose that $f = \sum a_n q^n$ satisfies the condition. We will show first that $a_n = 0$ for all n prime to N. If $N = p$ there is nothing extra to prove, so we will assume for the moment that N and p are distinct. Let g be the form $\sum_{(N,n)=1} a_n q^n$, i.e., the sum

$$\sum_{n=1}^{\infty} b_n q^n, \qquad b_n = \begin{cases} a_n & \text{if } (n, N) = 1, \\ 0 & \text{otherwise.} \end{cases}$$

Then $b_n = 0$ whenever n is not a multiple of p; our aim is to show that $g = 0$. Now the point is that g may be considered as a weight-two mod p modular form on $\Gamma_0(N^2)$; this follows from [31, Prop. 3.64]. The hypothesis about the vanishing of the b_n means that g is annihilated by the operator $\theta = q\frac{d}{dq}$. Since $p > 2$, this forces $g = 0$ as desired, since θg has "filtration" $p + 3$ if g is non-zero. (This is Katz's generalization of the Serre-Swinnerton-Dyer theorem [11, p. 55].)

To complete the proof, we must show that $f = 0$. This follows from Proposition 4.10, Lemma 5.9 and Lemma 5.10 of [13, Ch. II]. □

Remark. The argument we have given is essentially that of [13], Ch. II, Prop. 14.13. The exploitation of $q\frac{d}{dq}$ to deal with the absence of T_p is a familiar ploy—it was used by the author in [22, Prop. 2] and by Wiles in [36, Lemma, p. 491].

Suppose now that $\mathfrak{m}$ is a maximal ideal of $\mathbf{T}$ of residue characteristic p. (We continue to assume $p \geq 5$.) Let $J[\mathfrak{m}]$ denote the kernel of $\mathfrak{m}$ on $J(\overline{\mathbf{Q}})$. By the results of [13], $J[\mathfrak{m}]$ defines the two-dimensional semisimple representation $\rho_{\mathfrak{m}}$ of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ which is associated with $\mathfrak{m}$; this representation is irreducible if and only if $\mathfrak{m}$ is not an Eisenstein prime of $\mathbf{T}$. When $\mathfrak{m}$ is non-Eisenstein, $\rho_{\mathfrak{m}}$ is a semistable representation: this follows from the results of Deligne-Rapoport [3] to the effect J has multiplicative reduction at N, together with Proposition 5.1.

Proposition 6.2. *The image of* $\rho_{\mathfrak{m}}\colon \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathbf{GL}(2, \mathbf{T}/\mathfrak{m})$ *is the group of elements of* $\mathbf{GL}(2, \mathbf{T}/\mathfrak{m})$ *having determinant in* $\mathbf{F}_p^*$.

Proof. This mild strengthening of [13, Ch. II, Prop. 14.12] may be derived directly from Theorem 2.5. Indeed, for each prime ℓ different from p and N, the image under $\rho_{\mathfrak{m}}$ of a Frobenius element for ℓ in $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ has trace T_ℓ mod $\mathfrak{m}$. It follows from the Lemma above that the T_ℓ mod $\mathfrak{m}$ generate $\mathbf{T}/\mathfrak{m}$. □

Remark. By using Theorem 2.6 in place of Theorem 2.5, we obtain the stronger statement that $\rho_{\mathfrak{m}}(X)$ is the group of matrices in $\mathbf{GL}(2, \mathbf{T}/\mathfrak{m})$ whose determinants lie in $\mathbf{F}_p^*$. A similar remark applies to the Proposition which follows.

Suppose next that $p \geq 5$ is such that *none* of the $\mathfrak{m}|p$ in $\mathbf{T}$ is an Eisenstein prime. This means simply that $N \not\equiv 1 \bmod p$. Let $\mathbf{F}$ be the $\mathbf{F}_p$-algebra $\prod_{m|p} \mathbf{T}/\mathfrak{m}$, and let

$$\rho\colon\ \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathbf{GL}(2, \mathbf{F})$$

be the product of the $\rho_{\mathfrak{m}}$.

Proposition 6.3. *The image of* ρ *is the group of matrices in* $\mathbf{GL}(2, \mathbf{F})$ *having determinant in* $\mathbf{F}_p^*$.

Proof. Indeed, the natural map $\mathbf{T}/p\mathbf{T} \to \mathbf{F}$ is surjective by the Chinese remainder theorem. Accordingly, the Lemma implies that $\mathbf{F}$ is generated by the images of the T_ℓ with ℓ prime to pN. In the language of Theorem 3.2, this means that $\mathbf{F}' = \mathbf{F}$. Applying that theorem, we find that the image of ρ is as stated. □

Continuing the discussion, we add the hypothesis that the surjection $\mathbf{T}/p\mathbf{T} \to \mathbf{F}$ is an isomorphism, i.e., that p is unramified in $\mathbf{T}$. Then ρ is the representation giving the action of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ on the group $J[p]$, viewed as a free $\mathbf{T}/p\mathbf{T}$-module of rank two. Proposition 6.3 thus furnishes a description of the Galois group of the field cut out by the p-division points of J. Further, as we have seen in Proposition 4.2, the result of Proposition 6.3 is equivalent to an analogous statement about the p-adic Galois representation defined by J. More precisely, because p in unramified in $\mathbf{T}$, $\mathbf{T}\otimes\mathbf{Z}_p$ is a product of discrete valuation rings. The Tate module $\mathrm{Ta}_p(J)$ of J is then automatically free of rank two over $\mathbf{T}\otimes\mathbf{Z}_p$. After choosing a basis of $\mathrm{Ta}_p(J)$, we may summarize the action of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ on the p-power division points of J by a homomorphism

$$\tilde{\rho}\colon\ \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathbf{GL}(2, \mathbf{T}\otimes\mathbf{Z}_p).$$

By Propositions 4.2 and 6.3, we find that the image of $\tilde{\rho}$ is "as large as possible":

Theorem 6.4. *Suppose that p is unramified in $\mathbf{T}$ and that p is prime to $6\cdot(N-1)$. Then $\tilde{\rho}\big(\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})\big) = \{\, M \in \mathbf{GL}(2, \mathbf{T}\otimes\mathbf{Z}_p) \mid \det M \in \mathbf{Z}_p^{*}\,\}$.*

Remark. We obtain the more precise equality

$$\tilde{\rho}(X) = \{\, M \in \mathbf{GL}(2, \mathbf{T}\otimes\mathbf{Z}_p) \mid \det M \in \mathbf{Z}_p^{*}\,\}$$

from Theorem 3.4 and Proposition 4.2.

7. Complements.

We continue our study of $J = J_0(N)$, where N is a prime number by studying products of p-adic representations attached to J. We are motivated by the discussions of [25, Ch. IV, §3] and [26, §4.4], whose tools serve very well in this context.

For each prime p, let $\mathrm{Ta}_p(J)$ be the $\mathbf{Z}_p$-adic Tate module of J and write ρ_p (rather than $\tilde{\rho}$, as above) for the p-adic representation of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ which is associated to J:

$$\tilde{\rho}\colon\ \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathrm{Aut}_{\mathbf{T}\otimes\mathbf{Z}_p}\left(\mathrm{Ta}_p(J)\right) \subset \mathrm{Aut}_{\mathbf{T}\otimes\mathbf{Q}_p}\left(\mathrm{Ta}_p(J)\otimes\mathbf{Q}\right).$$

Let G_p be the image of ρ_p. This group was determined in Theorem 6.4 for most prime numbers p. We shall describe the p-adic Lie algebra of G_p in general; thus we determine G_p "up to finite groups" even when (or especially when) p does not satisfy the hypothesis to Theorem 6.4.

We recall that Mazur proved [13, Ch. II, §15-§17] that $\mathrm{Ta}_p(J)$ is free of rank two over $\mathbf{T}\otimes\mathbf{Z}_p$ when p is odd, and also in many circumstances when $p = 2$. In this favorable situation, we may view G_p as a closed subgroup of $\mathbf{GL}(2, \mathbf{T} \otimes \mathbf{Z}_p)$. In any case, $\mathrm{Ta}_p(J) \otimes \mathbf{Q}_p$ is free of rank two over the $\mathbf{Q}_p$-algebra $\mathbf{T} \otimes \mathbf{Q}_p$ by [13, Ch. II, Lemma 7.7]. Hence we may always regard G_p as a closed subgroup of $\mathbf{GL}(2, E)$, where $E = \mathbf{T} \otimes \mathbf{Q}_p$. Note that E is a commutative semisimple $\mathbf{Q}_p$-algebra, i.e., the product of fields which are finite extensions of $\mathbf{Q}_p$. Because the determinant of ρ_p is the p-adic cyclotomic character, we have $G_p \subseteq H_p$, where

$$H_p := \{M \in \mathbf{GL}(2, E) \mid \det M \in \mathbf{Q}_p^*\}.$$

Proposition 7.1. *The group G_p is open in H_p.*

Proof. Let $\mathfrak{g}$ and $\mathfrak{h}$ be the p-adic Lie algebras of G_p and H_p, respectively. Thus

$$\mathfrak{g} \subseteq \mathfrak{h} = \mathbf{Q}_p \times \mathfrak{sl}_2(E),$$

where $\mathfrak{sl}_2(E) = [\mathfrak{h}, \mathfrak{h}]$ is the Lie algebra of two-by-two matrices over E with trace 0. Since the p-adic cyclotomic character has infinite order, g is not contained in $\mathfrak{sl}_2(E)$. The proposition states that $\mathfrak{g} = \mathfrak{h}$.

The equality $\mathfrak{g} = \mathfrak{h}$ is proved as [18, Th. 4.5.4] in the special case where $\mathbf{T} \otimes \mathbf{Q}$ is a field, i.e., where J is a simple abelian variety. (Note that $\mathrm{End}_{\overline{\mathbf{Q}}} J = \mathrm{End}_{\mathbf{Q}} J$, as was proved in [17]; hence J is simple over $\mathbf{Q}$ if and only if it is absolutely simple.) In the general case, J is isogenous over $\mathbf{Q}$ to a product $A_1 \times \cdots \times A_t$ of simple abelian varieties to which [18, Th. 4.5.4] applies. Thus $\mathfrak{g}$ and $\mathfrak{h}$ have equal images in $\mathrm{End}(\mathrm{Ta}_p(A_i) \otimes \mathbf{Q}_p)$ for each i. Moreover, one knows that $\mathrm{End}_{\mathfrak{g}}(\mathrm{Ta}_p(J)) = E = \mathrm{End}_{\mathfrak{h}}(\mathrm{Ta}_p(J))$. Indeed, the Tate conjecture for abelian varieties, which was proved in [5], implies that $\mathrm{End}_{\mathfrak{g}}(\mathrm{Ta}_p(J)) = (\mathrm{End}_{\overline{\mathbf{Q}}} J) \otimes \mathbf{Q}_p$. On the other hand, one knows by [17] that $(\mathrm{End}_{\overline{\mathbf{Q}}} J) \otimes \mathbf{Q} = \mathbf{T} \otimes \mathbf{Q}$; hence $(\mathrm{End}_{\overline{\mathbf{Q}}} J) \otimes \mathbf{Q}_p = E$, as was claimed. The proof of [18, Th. 4.4.10] now yields the required equality $\mathfrak{g} = \mathfrak{h}$. □

For each set of prime numbers S, let

$$\rho_S\colon\ \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \prod_{p \in S} \mathrm{Aut}(\mathrm{Ta}_p(J))$$

be the product of the ρ_p for p in S. Let G_S be the image of ρ_S, so that G_S is a closed subgroup of the product $\prod_{p \in S} G_p$.

Corollary 7.2. *If S is a finite set of primes, then G_S is open in $H_S := \prod_{p \in S} H_p$.*

Proof. The assertion to be proved follows from Proposition 7.1 and an argument due to Serre [25, Lemma 4, p. IV-24]:

Fix $p \in S$ for the moment, and ensure by a change of basis if necessary that G_p is a subgroup of $\mathbf{GL}(2, R)$, where R is the ring of integers of E. Let

$\mathbf{F}$ be the product of the residue fields of the factors of R, and let G'_p be the kernel of the composition

$$G_p \hookrightarrow \mathbf{GL}(2,R) \to \mathbf{GL}(2,F).$$

Clearly, G'_p has finite index in G_p, so that G'_p is open in H_p by Proposition 7.1. Further, G'_p is a pro-p group; in fact, it is a projective limit of nilpotent groups of p-power order.

Let G'_S be the inverse image of $\prod_{p\in S} G'_p$ in G_S. Since G'_S is a subgroup of $\prod_{p\in S} G'_p$, it is pro-nilpotent. Thus G'_S is the product of its Sylow subgroups. Now the p-Sylow subgroup of G'_S has finite index in G_p. Thus G'_S has finite index in $\prod_{p\in S} G_p$, a group which is open in H_S by Proposition 7.1. Hence G_S is open in H_S. □

Theorem 7.3 (Kaskel [10])**.** *Assume that N does not belong to S. Then $G_S = \prod_{p\in S} G_p$.*

Proof. We first consider the case where S is finite, arguing by induction on the size of S. The statement to be proved is evident if S has at most one element, so we may assume that $S = T \coprod \{p\}$ and that the statement is true with S replaced by T. We must show that the natural injection $G_S \hookrightarrow G_T \times G_p$ is an isomorphism, or equivalently that the Galois extensions of $\mathbf{Q}$ cut out by ρ_T and ρ_p are linearly disjoint. Let K be the intersection of these two fields. Then K is ramified only at N, since ρ_p is ramified only at N and at p, while ρ_T is unramified at p. Further, the inertia groups for the prime N in the image of ρ_p are pro-p groups since $J_0(N)$ is semistable at N. Similarly, the inertia groups for N in the image of ρ_T are profinite groups of order prime to p; indeed, p is not a member of T. Hence the inertia groups for N in $\mathrm{Gal}(K/\mathbf{Q})$ are trivial, so that K is an everywhere unramified extension of $\mathbf{Q}$. This gives $K = \mathbf{Q}$ and proves the linear disjointness.

The case where S is not necessarily finite now follows, since G_S is closed and dense in $\prod_{p\in S} G_p$. □

Corollary 7.4. *Let S be the set of prime numbers p which are prime both to $6(N-1)$ and to the discriminant of $\mathbf{T}$. Then*

$$G_S = \prod_{p\in S} \{M \in \mathbf{GL}(2, \mathbf{T}\otimes \mathbf{Z}_p) \mid \det M \in \mathbf{Z}_p^*\}.$$

Proof. We shall give two proofs of this result, the first in spirit of Kaskel's theorem. For each $p \in S$, we have $G_p = \{\, M \in \mathbf{GL}(2, \mathbf{T}\otimes\mathbf{Z}_p) \mid \det M \in \mathbf{Z}_p^* \,\}$ by Theorem 6.4. Let T be the set of primes $p \in S$ which are different from N; thus $S = T \coprod \{N\}$ if N is prime to the discriminant of $\mathbf{T}$ and $S = T$ if not.

We clearly have $G_T = \prod_{p \in T} \{ M \in \mathbf{GL}(2, \mathbf{T} \otimes \mathbf{Z}_p) \mid \det M \in \mathbf{Z}_p^* \}$ by Kaskel's theorem and Theorem 6.4. This gives what is wanted if $S = T$. In what follows, we suppose to the contrary that N belongs to S.

The idea now is to analyze the subgroup G_S of $G_T \times G_N$ as in the proof above. Let $K/\mathbf{Q}$ now be the obstruction to the equality $G_S = G_T \times G_N$. In other words, K is the intersection of the two extensions of $\mathbf{Q}$ whose Galois groups are the images of G_T and G_N. Clearly, K is ramified only at N. Moreover, I is an inertia group for N in $\mathrm{Gal}(K/\mathbf{Q})$, then the order of I (as a supernatural number) is divisible only by the primes in T. In particular, this order is prime to $N(N-1)$. On the other hand, it is easy to see that only N and primes dividing $N-1$ can intervene in the order of an inertia group for N in the image of ρ_N. Indeed, the restriction of ρ_N to an inertia group for N in $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ has the form $\begin{pmatrix} \chi & * \\ 0 & 1 \end{pmatrix}$, where χ is the N-adic cyclotomic character. Hence K is unramified at N, and we obtain the required equality $G_S = G_T \times G_N$ as in the proof of Kaskel's theorem.

To make a second proof of the equality

$$G_S = \prod_{p \in S} \{ M \in \mathbf{GL}(2, \mathbf{T} \otimes \mathbf{Z}_p) \mid \det M \in \mathbf{Z}_p^* \},$$

we fix a prime p in T and let X be the group which we considered in Theorem 2.6 and Theorem 3.4. We have $\rho_p(X) = \{ M \in \mathbf{GL}(2, \mathbf{T} \otimes \mathbf{Z}_p) \mid \det M \in \mathbf{Z}_p^* \}$ as explained in the remark at the end of the preceding section. On the other hand, if p' is an element of S which is different from p, $\rho_{p'}$ is unramified at p and therefore $\rho_{p'}(X) = \{1\}$. Thus $\rho_S(X)$ contains

$$\{ M \in \mathbf{GL}(2, \mathbf{T} \otimes \mathbf{Z}_p) \mid \det M \in \mathbf{Z}_p^* \},$$

viewed as a subgroup of the product. Hence we have

$$G_S \supseteq \prod_{p \in T} \{ M \in \mathbf{GL}(2, \mathbf{T} \otimes \mathbf{Z}_p) \mid \det M \in \mathbf{Z}_p^* \}.$$

This gives what is wanted if $S = T$, so we assume once again that N belongs to S. Then G_S is a subgroup of the product

$$\prod_{p \in S} \{ M \in \mathbf{GL}(2, \mathbf{T} \otimes \mathbf{Z}_p) \mid \det M \in \mathbf{Z}_p^* \}$$

which maps onto $\{ M \in \mathbf{GL}(2, \mathbf{T} \otimes \mathbf{Z}_N) \mid \det M \in \mathbf{Z}_N^* \}$ by Theorem 6.4. Moreover G_S contains the kernel of the natural projection

$$\prod_{p \in S} \{ M \in \mathbf{GL}(2, \mathbf{T} \otimes \mathbf{Z}_p) \mid \det M \in \mathbf{Z}_p^* \} \to \{ M \in \mathbf{GL}(2, \mathbf{T} \otimes \mathbf{Z}_N) \mid \det M \in \mathbf{Z}_N^* \}.$$

Therefore G_S is the full product. □

For our final result, we consider the product ρ_{f} of all of the representations ρ_p, i.e., the representation ρ_S where S is the set of all prime numbers. We view ρ_{f} as taking values in

$$\{M \in \mathbf{GL}(2, \mathbf{T}\otimes\mathbf{Q}_2) \mid \det M \in \mathbf{Q}_2^*\} \times \prod_{p\neq 2}\{M \in \mathbf{GL}(2, \mathbf{T}\otimes\mathbf{Z}_p) \mid \det M \in \mathbf{Z}_p^*\},$$

a group that we will call $\mathcal{A}$. (We separate 2 from the odd primes since it is not known that Ta_2 is free of rank two over $\mathbf{T} \otimes \mathbf{Z}_2$.)

Theorem 7.5. *The image of* ρ_{f} *is an open subgroup of* $\mathcal{A}$.

Proof. Let $\mathcal{A}_p$ be the pth component of $\mathcal{A}$, so that we have $\mathcal{A} = \prod \mathcal{A}_p$, with the product extended over all primes. Let S be the set of those p which are prime to $6(N-1)N$ and to the discriminant of $\mathbf{T}$. Fix $p \in S$ and let X be the subgroup of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ which we have considered repeatedly: the closed subgroup of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ generated by all inertia groups for p. As we have seen, $\rho_p(X) = \mathcal{A}_p$; on the other hand $\rho_{p'}(X) = \{1\}$ for $p' \neq p$. Hence $\rho_{\mathrm{f}}(X) = \mathcal{A}_p$, where $\mathcal{A}_p$ is considered as a subgroup of the product $\mathcal{A}$. On varying p, we find that $\rho_{\mathrm{f}}\left(\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})\right)$ contains the group $\prod_{p\in S}\mathcal{A}_p$, i.e., the kernel of the projection $\mathcal{A} \to \prod_{p\notin S}\mathcal{A}$. On the other hand, the image of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ in this finite product is open by Corollary 7.2. Hence $\rho_{\mathrm{f}}\left(\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})\right)$ is open in the full product. □

References

[1] C. Batut, D. Bernardi, H. Cohen and M. Olivier, *GP/PARI*, Available by anonymous ftp from `megrez.math.u-bordeaux.fr` or `math.ucla.edu`, in the directory `/pub/pari`.

[2] R. Coleman, B. Kaskel and K. Ribet, *Torsion points on* $X_0(N)$, Contemporary Math., to appear.

[3] P. Deligne and M. Rapoport, *Les schémas de modules de courbes elliptiques*, Lecture Notes in Math., Vol. 349, Springer-Verlag, Berlin and New York, (1973), 143-316.

[4] L.E. Dickson, *Linear groups with an exposition of the Galois field theory*, Dover, New York, 1958.

[5] G. Faltings, *Endlichkeitssätze für abelsche Varietäten über Zahlkörpern*, Invent. Math., **73** (1983), 349-366.

[6] G. Frey (Ed.) *On Artin's conjecture for odd* 2*-dimensional representations*, Lecture Notes in Math., Vol. 1585, Springer-Verlag, Berlin-Heidelberg-New York, 1994.

[7] D. Gorenstein, *Finite Groups*, Chelsea, New York, 1980.

[8] B.H. Gross, *A tameness criterion for Galois representations associated to modular forms* mod p, Duke Math. J., **61** (1990), 445-517.

[9] A. Grothendieck, *SGA*7 I*, Exposé* IX, Lecture Notes in Math., Vol. 288, Springer-Verlag, Berlin and New York, (1972), 313-523.

[10] B. Kaskel, *The adelic representation associated to* $X_0(37)$, PhD. thesis, UC Berkeley, May, 1996.

[11] N.M. Katz, *A result on modular forms in characteristic* p, Lecture Notes in Math., Vol. 601, Springer-Verlag, Berlin and New York, (1977), 53-61.

[12] S. Lang and H. Trotter, *Frobenius distributions in* $\mathbf{GL}_2$*-extensions*, Lecture Notes in Math., Vol. 504, Springer-Verlag, Berlin and New York, 1976.

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

[14] T. Miyake, *Modular forms*, Springer-Verlag, Berlin and New York, 1989.

[15] J. Oesterlé, *Travaux de Wiles* (et Taylor, ...), Partie II, Séminaire Bourbaki, **804** (1995).

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

[17] K.A. Ribet, *Endomorphisms of semi-stable abelian varieties over number fields*, Annals of Math., **101** (1975), 555-562.

[18] ———, *Galois action on division points of abelian varieties with many real multiplications*, Am. J. Math., **98** (1976), 751-804.

[19] ———, *On ℓ-adic representations attached to modular forms*, Invent. Math., **28** (1975), 245-275.

[20] ———, Mod p *Hecke operators and congruences between modular forms*, Invent. Math., **71** (1983), 193-205.

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

[22] ———, *Multiplicities of p-finite* mod p *Galois representations in* $J_o(Np)$, Boletim da Sociedade Brasileira de Matemática, Nova Série, **21** (1991), 177-188.

[23] ———, *Abelian varieties over* $\mathbf{Q}$ *and modular forms*, 1992 Proceedings of KAIST Mathematics Workshop, Taejon, Korea Advanced Institute of Science and Technology, (1992), 53-79.

[24] I. Schur, *Arithmetische Untersuchungen über endliche Gruppen linearer Substitutionen*, Sitz. Pr. Akad. Wiss., (1906), 164-184; Gesam. Abhl., **I**, 177-197, Springer-Verlag, Berlin-Heidelberg-New York-Tokyo, 1973.

[25] J.-P. Serre, *Abelian ℓ-adic representations and elliptic curves*, Addison-Wesley Publ. Co., Redding, Mass., 1989.

[26] ———, *Propriétés galoisiennes des points d'ordre fini des courbes elliptiques*, Invent. Math., **15** (1972), 259-331.

[27] ———, *Œuvres*, Springer-Verlag, Berlin-Heidelberg-New York-Tokyo, 1986.

[28] ———, *Sur les représentations modulaires de degré* 2 *de* $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$, Duke Math. J., **54** (1987), 179-230.

[29] ———, *Travaux de Wiles* (et Taylor, ...), Partie I, Séminaire Bourbaki, **803** (1995).

[30] G. Shimura, *A reciprocity law in non-solvable extensions*, Journal für die reine und angewandte Mathematik, **221** (1966), 209-220.

[31] ———, *Introduction to the arithmetic theory of automorphic functions*, Princeton University Press, Princeton, 1971.

[32] G. Shimura and Y. Taniyama, *Complex multiplication of abelian varieties and its applications to number theory*, Second printing corrected, Math. Society of Japan, Tokyo, 1975.

[33] J. Sturm, *On the congruence of modular forms*, Lecture Notes in Math., Vol. 1240, Springer-Verlag, Berlin and New York, (1987), 275-280.

[34] J. Tate, *The non-existence of certain Galois extensions of* **Q** *unramified outside* 2, Contemporary Mathematics, **174** (1994), 153-156.

[35] J.-L. Waldspurger, *Quelques propriétés arithmétiques de certaines formes automorphes sur* **GL**(2), Compositio Math., **54** (1985), 121-171.

[36] A. Wiles, *Modular elliptic curves and Fermat's Last Theorem*, Annals of Math., **141** (1995), 443-551.

This article was begun while the author was a research professor at the MSRI, where research was supported in part by NSF grant DMS-9022140. This work was further supported by the investigator's NSF Grants DMS 88-06815 and DMS 93-06898. The author thanks the referee for a number of interesting suggestions and questions and Ian Kiming and Jean-François Mestre for carrying out computations in connection with the example discussed at the end of §2.

University of California
Berkeley, CA 94720-3840 USA
E-mail address: ribet@math.berkeley.edu