Eisenstein ideals and λ-adic representations

By Elisabeth PAPIER and Kenneth A. RIBET*)

In memory of Takuro Shintani

The object of this note is to apply the techniques of Swinnerton-Dyer [5] to the study of certain 2-dimensional λ-adic representations of the Galois group $\mathrm{Gal}(\bar{\boldsymbol{Q}}/\boldsymbol{Q})$, namely those which are unramified outside the residue characteristic of λ and which are reducible modulo λ. We have been guided by certain portions of Mazur's Eisenstein ideal paper [1]; in particular, we introduce the analogue of Mazur's Hecke algebra $\boldsymbol{T}$, together with an ideal of $\boldsymbol{T}$ which we call the Eisenstein ideal. Making certain natural hypotheses, we show that this ideal is principal, giving a specific generator for it. We also determine (up to conjugation) the image of the given representation.

This work is an outgrowth of the first author's study of λ-adic representations attached to modular forms [2]. A subsequent article [3] will explore applications to such representations, including numerical examples.

1. Let l be an odd prime. Let $\bar{\boldsymbol{Q}}$ be an algebraic closure for $\boldsymbol{Q}$, and let $K_l \subset \bar{\boldsymbol{Q}}$ be the largest extension of $\boldsymbol{Q}$ which is unramified away from l and infinity. Let $G=\mathrm{Gal}(K_l/\boldsymbol{Q})$. Let E be a finite extension of $\boldsymbol{Q}_l$; let $\mathfrak{O}$, λ, and $\boldsymbol{F}$ be respectively the integer ring, the maximal ideal, and the residue field of E. Let

$$\rho : G \longrightarrow \boldsymbol{GL}(2, E)$$

be a continuous homomorphism. Thus ρ is a 2-dimensional λ-adic representation of $\mathrm{Gal}(\bar{\boldsymbol{Q}}/\boldsymbol{Q})$, unramified outside l. We shall write tr and det for the trace and the determinant of ρ, *a priori* functions on G with values in E. In fact, since G is compact, ρ is conjugate to a representation with values in $\boldsymbol{GL}(2, \mathfrak{O})$. Therefore, tr and det are $\mathfrak{O}$-valued.

Replacing ρ by such a conjugate $M\rho M^{-1}$, and composing it with the natural map

$$\boldsymbol{GL}(2, \mathfrak{O}) \longrightarrow \boldsymbol{GL}(2, \boldsymbol{F}),$$

we obtain a homomorphism

$$\bar{\rho} : G \longrightarrow GL(2, \boldsymbol{F}).$$

*) This author partially supported by the National Science Foundation under grant MCS 80-02317.

As is well known, this homomorphism may depend on the choice of M. However, the semisimplification of $\bar{\rho}$ depends only on G. We shall consider only the situation in which $\bar{\rho}$ is reducible, so that its semisimplification is described by 2 characters

$$\alpha,\ \beta : G \longrightarrow \boldsymbol{F}^*.$$

Let $\chi: G \to \boldsymbol{Z}_l^*$ be the l-adic cyclotomic character, and let

$$\omega : G \longrightarrow \boldsymbol{F}_l^*$$

be the reduction of χ modulo l. (Thus ω gives the action of G on the group $\mu_l \subset \bar{K}_l^*$ of l^{th} roots of unity.) Any continuous homomorphism $\varphi: G \to A$, where A is a profinite abelian group, must factor through χ; moreover, if the l-primary part of A is trivial, then φ must factor through ω. Since the only maps from $\boldsymbol{F}_l^*$ to $\boldsymbol{F}^*$ are powers of the natural inclusion $\boldsymbol{F}_l^* \overset{i}{\hookrightarrow} \boldsymbol{F}^*$, we may conclude that α and β are each the composition of i with some power of ω. We will write simply

$$\alpha = \omega^n, \qquad \beta = \omega^m,$$

with $n, m \in \boldsymbol{Z}/(l-1)\boldsymbol{Z}$. As our last general hypothesis, we will suppose that n and m are distinct, so that the two ratios $\alpha\beta^{-1}$, $\beta\alpha^{-1}$ are non-trivial characters. This hypothesis always holds if the character *det* is an *odd* character; in particular, it holds if ρ is the λ-adic representation attached to a holomorphic modular form of l-power level.

2. Our "Hecke algebra" $\boldsymbol{T}$ is simply the $\boldsymbol{Z}_l$-subalgebra of $\mathfrak{O}$ generated by the various quantities $tr(g)$, with $g \in G$. It is clear that $\boldsymbol{T}$ is a local $\boldsymbol{Z}_l$-algebra with maximal ideal $\mathfrak{M} = \boldsymbol{T} \cap \lambda$. The residue field $\boldsymbol{T}/\mathfrak{M}$ is the prime field $\boldsymbol{F}_l$, since for each $g \in G$ we have the mod λ congruence

$$\mathrm{tr}(g) \equiv \omega^n(g) + \omega^m(g) \in \boldsymbol{F}_l.$$

As a $\boldsymbol{Z}_l$-module, $\boldsymbol{T}$ is free of finite rank; it is therefore complete and separated with respect to its (l)-adic topology. Using the Artin-Rees lemma, one sees that this topology on $\boldsymbol{T}$ coincides with the $\mathfrak{M}$-adic topology on $\boldsymbol{T}$, cf. Bourbaki, *Alg. Comm.*, III, § 3, n°3, Prop. 7 (iii). We therefore have

$$\boldsymbol{T} \xrightarrow{\sim} \varprojlim_i \boldsymbol{T}/\mathfrak{M}^i,$$

which permits application of Hensel's lemma in $\boldsymbol{T}$. Because l is odd, the identity

$$2 \cdot \det(g) = \mathrm{tr}(g)^2 - \mathrm{tr}(g^2)$$

shows that the values of det are contained in $\boldsymbol{T}$.

Our definition of the "Eisenstein ideal" I of $\boldsymbol{T}$ is somewhat indirect. We observe that $\chi(G)$ is the profinite cyclic group $\boldsymbol{Z}_l^*$. Choose an element g_0 of G such that $\chi(g_0)$ generates $\boldsymbol{Z}_l^*$. Because n and m are distinct, the quadratic polynomial

$$X^2-\mathrm{tr}\,(g_0)X+\det\,(g_0)$$

has distinct roots modulo $\mathfrak{M}$. By Hensel's lemma, it factors over $\boldsymbol{T}$. Let r and s be its roots, ordered so that we have

$$r\equiv\omega^n(g_0)\,,\qquad s\equiv\omega^m(g_0)\qquad \mathrm{mod}\,\mathfrak{M}\,.$$

(2.1) LEMMA. *There exist unique characters* $\varphi,\psi: G\to\boldsymbol{T}^*$ *satisfying*

$$\varphi(g_0)=r\,,\qquad \psi(g_0)=s\,.$$

The product of these characters is det.

PROOF. Any character $G\to\boldsymbol{T}^*$ is the composition of χ with a unique map

$$\theta:\boldsymbol{Z}_l^*\longrightarrow\boldsymbol{T}^*\,.$$

Moreover, θ will be determined by its value on the generator $x=\chi(g_0)$ of $\boldsymbol{Z}_l^*$. The key point is that $\theta(x)$ can be selected arbitrarily: given $t\in\boldsymbol{T}^*$, we have $\theta(x)=t$ for some θ. This assertion follows easily from the fact that the residue field of $\boldsymbol{T}$ is the prime field $\boldsymbol{F}_l$, so that $\boldsymbol{T}^*$ is the product of the pro-l group $1+\mathfrak{M}$ and a cyclic group of order $l-1$.

We now define $\eta: G\to\boldsymbol{T}$ to be the function $tr-\varphi-\psi$ and define I to be the ideal of $\boldsymbol{T}$ generated by all quantities $\eta(g)$, for $g\in G$. The congruences

$$\mathrm{tr}\equiv\omega^n+\omega^m\equiv\varphi+\psi\qquad\mathrm{mod}\,\mathfrak{M}$$

show that I is contained in $\mathfrak{M}$. It is easily seen that the ideal I is intrinsic, although the characters φ and ψ obviously depend on g_0. More precisely, we have the following result.

(2.2) PROPOSITION. *Let* α *and* β *be characters* $G\to\boldsymbol{T}^*$, *and let* J *be an ideal of* $\boldsymbol{T}$. *Suppose that we have the congruence*

$$\mathrm{tr}\equiv\alpha+\beta\qquad\mathrm{mod}\,J\,. \tag{2.3}$$

Then I *is contained in* J. *Moreover, after permuting* α *and* β *if necessary, we have* $\alpha\equiv\varphi$ *and* $\beta\equiv\psi$ *modulo* J.

PROOF. We may assume that J is a proper ideal of $\boldsymbol{T}$, so that J is contained in $\mathfrak{M}$. The congruence (2.3) implies the congruence

$$\alpha\beta \equiv \det \quad \bmod J.$$

Specializing to g_0, we obtain the two congruences

$$\alpha(g_0)\cdot\beta(g_0) \equiv rs \quad \bmod J,$$

$$\alpha(g_0)+\beta(g_0) \equiv r+s \quad \bmod J.$$

Since r and s are incongruent mod $\mathfrak{M}$, it is clear that we have (possibly after permuting α and β):

$$\alpha(g_0) \equiv r, \qquad \beta(g_0) \equiv s \quad \bmod J.$$

This gives the last assertion of the proposition, i.e., the congruences $\alpha\equiv\varphi$, $\beta\equiv\psi$. We therefore have

$$\begin{aligned}\eta(g) &= \operatorname{tr}(g)-\varphi(g)-\psi(g)\\ &\equiv \operatorname{tr}(g)-\alpha(g)-\beta(g) \equiv 0 \quad \bmod J,\end{aligned}$$

for each $g\in G$. Thus $\eta(g)$ belongs to J for each g, so I is contained in J.

The following "numerical" variant of (2.2) shows how to establish congruences for all quantities $tr(g)$ by checking them for finitely many g. The idea of proving congruences in this way is one of the main themes of [5].

(2.4) *Let $g_1, \dots, g_t$ be elements of G for which I is generated by $\eta(g_1), \dots, \eta(g_t)$. Let J be an ideal of $\boldsymbol{T}$. Suppose that α and β are characters $G\to\boldsymbol{T}^*$ satisfying $\alpha\beta\equiv\det \pmod J$, together with the congruences*

$$\operatorname{tr}(g_i) \equiv \alpha(g_i)+\beta(g_i) \pmod J \tag{2.5}$$

for $i=0$ and for $i=1, \dots, t$. Then we have

$$\operatorname{tr}(g) \equiv \alpha(g)+\beta(g) \pmod J$$

for all $g\in G$.

PROOF. Again, we may suppose that J is a proper ideal. As before, we find that α and β coincide with φ and ψ (up to permutation) modulo J. Hence (2.5) shows that $\eta(g_i)\in J$ for $i=1, \dots, t$. By hypothesis, we have $I\subseteq J$, whence the tautologous congruence

$$\operatorname{tr} \equiv \varphi+\psi \pmod J.$$

The conclusion follows.

(2.6) PROPOSITION. *Let R be the $\boldsymbol{Z}_l$-subalgebra of $\boldsymbol{T}$ generated by all values of the character $\varphi\psi^{-1}: G\to\boldsymbol{T}^*$. Suppose that the character* det *is R^*-valued. Then the natural map*

$$R \longrightarrow \boldsymbol{T}/I$$

is surjective.

PROOF. We must show that the image of R in $\boldsymbol{T}/I$ contains the images modulo I of all $tr(g)$. It suffices to show that the image of R modulo I contains the images of all $\varphi(g)$ and $\psi(g)$. In fact, we will show that φ and ψ are already R^*-valued.

In view of the fact that R contains all values of $\varphi\psi^{-1}$ and of $det=\varphi\psi$, we know that R contains all quantities $\varphi(g)^2$, $\psi(g)^2$. Thus we are reduced to showing that a unit in R which "becomes" a square in $\boldsymbol{T}$ is already a square in R. This assertion is a consequence of Hensel's lemma (applied in R), together with the fact that the residue fields of R and of $\boldsymbol{T}$ coincide and have characteristic prime to 2.

3. We now begin study of the representation ρ. After replacing ρ by a conjugate $M\rho M^{-1}$, we may suppose that ρ takes values in $\boldsymbol{GL}(2, \mathfrak{O})$ and that its reduction $\bar{\rho}$ is given schematically by the matrix

$$\begin{pmatrix} \omega^n & * \\ 0 & \omega^m \end{pmatrix}.$$

In other words, letting $a, b, c, d: G \to \mathfrak{O}$ denote the matrix coefficients of ρ $\left(\text{so that } \rho=\begin{pmatrix} a & b \\ c & d \end{pmatrix}\right)$, we have

$$a \equiv \omega^n, \quad d \equiv \omega^m, \quad c \equiv 0 \quad (\text{mod } \lambda). \tag{3.1}$$

Since the eigenvalues r and s of $\rho(g_0)$ are distinct modulo λ, we may now find a matrix $N \in \boldsymbol{GL}(2, \mathfrak{O})$ such that $N\rho(g_0)N^{-1}=\begin{pmatrix} r & 0 \\ 0 & s \end{pmatrix}$. Making the replacement $\rho \mapsto N\rho N^{-1}$, we find that (3.1) is still satisfied and that $\rho(g_0)$ is the diagonal matrix $\begin{pmatrix} r & 0 \\ 0 & s \end{pmatrix}$.

(3.2) PROPOSITION. *For all $g \in G$, we have $a(g), d(g) \in \boldsymbol{T}$. For all pairs $g, g' \in G$, we have $b(g) \cdot c(g') \in \boldsymbol{T}$.*

PROOF. The first assertion follows from the fact that $tr(g)$ and $tr(gg_0)$ belong to $\boldsymbol{T}$ and that $r-s$ is a unit of $\boldsymbol{T}$. The second is then a consequence of the equation

$$b(g) \cdot c(g') = a(gg') - a(g) \cdot a(g'). \tag{3.3}$$

We now let $H=\mathrm{Gal}(K_l/\boldsymbol{Q}(\mu_{l^\infty}))$ be the kernel of the cyclotomic character χ. Let B be the $\boldsymbol{T}$-submodule of $\mathfrak{O}$ generated by all $b(g)$ with $g\in G$. Since the function b vanishes on the closure of the subgroup of G generated by g_0, we see that B is already generated by the $b(h)$ with $h\in H$. Similarly, we define C using the $c(g)$. We denote by BC the $\boldsymbol{T}$-submodule of $\mathfrak{O}$ generated by all products $\beta\cdot\gamma$ with $\beta\in B$, $\gamma\in C$. Then BC is generated by all products $b(g)c(g')$ so that, by (3.2), it is in fact an ideal of $\boldsymbol{T}$.

(3.4) PROPOSITION. *We have $I=BC$. Moreover, I is the ideal of $\boldsymbol{T}$ generated by the quantities $a(h)-1$ for $h\in H$, or alternately the ideal of $\boldsymbol{T}$ generated by the $d(h)-1$ for $h\in H$.*

PROOF. In view of the symmetry between a and d, we can prove the second assertion only relative to the $a(h)-1$. Let us temporarily denote by J the ideal of $\boldsymbol{T}$ generated by the $a(h)-1$. We will then establish the chain

$$BC\subseteq J\subseteq I\subseteq BC\,,$$

thus proving the proposition.

As a first step, we introduce the function "a mod BC" obtained by composing the coefficient function a with the canonical map $\boldsymbol{T}\to\boldsymbol{T}/BC$. Call this function $\bar{a}$. Using (3.3) again, we see that $\bar{a}$ is in fact a character $G\to(\boldsymbol{T}/BC)^*$. Since

$$a(g_0)=\varphi(g_0)\,,$$

we must have

$$\bar{a}\equiv\varphi \pmod{BC}\,,$$

i. e.,

$$a\equiv\varphi \pmod{BC}\,.$$

Similarly, we get

$$d\equiv\phi \pmod{BC}\,.$$

Adding these congruences, we find that $\eta(g)\in BC$ for all $g\in G$; therefore $I\subseteq BC$.

Now for each $h\in H$ we have $\varphi(h)=\phi(h)=1$. Therefore

$$[a(h)-1]+[d(h)-1]=\eta(h)\in I\,.$$

Similarly,

$$r(a(h)-1)+s(d(h)-1)=\eta(g_0h)\in I$$

Because $r-s$ is a unit of $\boldsymbol{T}$, we get

$$a(h)-1\,,\ d(h)-1\in I\,.$$

Therefore, $J\subseteq I$.

Finally, for $h, h' \in H$ we have

$$b(h)c(h') = a(hh') - a(h)a(h') \equiv 0 \qquad (\mathrm{mod}\, J).$$

This gives the inclusion $BC \subseteq J$.

(3.5) PROPOSITION. *Let* $g \in \mathrm{Gal}(K_l/\boldsymbol{Q}(\mu_l))$. *We have* $\varphi(g), \psi(g) \equiv 1 \ (\mathrm{mod}\, \mathfrak{M})$. *Further, we have*

$$\eta(g) \equiv b(g)c(g) \qquad \mathrm{mod}\, I\mathfrak{M}.$$

PROOF. Since the characters $\varphi \bmod \mathfrak{M}$ and $\psi \bmod \mathfrak{M}$ are powers of the mod l cyclotomic character, the first assertion is clear. Now $\varphi\psi = ad - bc$, so we have

$$b(g)c(g) - \eta(g) = u(\varphi(g)-1) + t(\psi(g)-1) + tu,$$

where

$$t = a(g) - \varphi(g) \in I$$

and

$$u = d(g) - \psi(g) \in I.$$

Since I is contained in $\mathfrak{M}$, the second assertion follows.

Now let M be the union of all finite abelian extensions of $\boldsymbol{Q}(\mu_l)$ in K_l which have l-power degree. The Galois group $X = \mathrm{Gal}(M/\boldsymbol{Q}(\mu_l))$ is a $\boldsymbol{Z}_l$-module on which $\Delta = \mathrm{Gal}(\boldsymbol{Q}(\mu_l)/\boldsymbol{Q})$ acts by conjugation. In other words, X is a module over the group ring $\boldsymbol{Z}_l[\Delta]$. As usual, X is the direct sum of the eigenspaces

$$X(\varepsilon) = \{x \in X \mid \delta \cdot x = \varepsilon(\delta) \cdot x \text{ for all } \delta \in \Delta\},$$

ε running over the group of $\boldsymbol{Z}_l^*$-valued characters of Δ. (In the above definition, $\varepsilon(\delta) \cdot x$ denotes the product of x and the "number" $\varepsilon(\delta) \in \boldsymbol{Z}_l^*$.) Notice that the various ε are the powers of the character obtained by composing the natural isomorphism

$$\Delta \xrightarrow{\sim} (\boldsymbol{Z}/l\boldsymbol{Z})^*$$

with the Teichmüller lifting

$$(\boldsymbol{Z}/l\boldsymbol{Z})^* \hookrightarrow \boldsymbol{Z}_l^*.$$

It is traditional to denote this character by ω. If we compose this new ω with the natural map $G \to \Delta$, we obtain a character, again denoted ω, which is just the Teichmüller lift of our original mod l cyclotomic character ω.

(3.6) THEOREM. *Suppose that l is prime to the class number of the maximal real subfield of $\boldsymbol{Q}(\mu_l)$. Then each $\boldsymbol{Z}_l$-module $X(\varepsilon)$ is cyclic.*

Recall that the hypothesis of (3.6) is the well known Vandiver conjecture for $\boldsymbol{Q}(\mu_l)$. It is true (at least) for all $l \leqq 125{,}000$ [6], and no counterexample is known.

PROOF. Since the assertion in question is essentially well known, we will give the proof rather rapidly. Let A be the l-primary part of the class group of $\boldsymbol{Q}(\mu_l)$; then by class field theory, A is given as a quotient of X. Let Y be the kernel of the natural map $X \to A$. Let $Y(\varepsilon)$ and $A(\varepsilon)$ be the eigenspaces analogous to the $X(\varepsilon)$ above.

It is easy to see that each eigenspace $Y(\varepsilon)$ is cyclic. Indeed, let U be the l-primary part of the group of units of the completion Φ of $\boldsymbol{Q}(\mu_l)$ at l, i.e., the group of units which are congruent to 1 modulo the maximal ideal of the ring of integers of Φ. Let $\mathcal{E}$ be the intersection (taken in Φ) of U and the group of units of $\boldsymbol{Q}(\mu_l)$. Using the l-adic logarithm map, one shows that the eigenspace $U(\varepsilon)$ is cyclic for each character $\varepsilon \neq \omega$, while $U(\omega)$ is the product of a cyclic $\boldsymbol{Z}_l$-module and the group μ_l. By class field theory, we have an isomorphism

$$Y \xleftarrow{\sim} U/\bar{\mathcal{E}},$$

where the $^{-}$ denotes "closure in the l-adic topology." The cyclicity then follows.

As a consequence, we obtain that $X(\varepsilon)$ is cyclic for each character ε such that $A(\varepsilon)$ vanishes. In view of the hypothesis, we may conclude that $X(\varepsilon)$ is cyclic for each even character ε.

To treat the other components, we introduce the odd part X^- of X, i.e., the direct sum of the $X(\varepsilon)$ with ε odd. Also, let $\mathcal{E}$ now be the group of "l-units" of $\boldsymbol{Q}(\mu_l)^+$, the maximal real subfield of $\boldsymbol{Q}(\mu_l)$. Thus $\mathcal{E}$ consists of all elements of $\boldsymbol{Q}(\mu_l)^+$ which are units locally at all non-archimedean primes of $\boldsymbol{Q}(\mu_l)^+$ except for the prime dividing l. As in [0, § 4], we see that the group $\mathcal{E}/\mathcal{E}^l$ is a cyclic $\boldsymbol{F}_l[\Delta]$-module. On the other hand, the hypothesis to (3.6) implies rather easily that the natural map

$$\mathcal{E}/\mathcal{E}^l \longrightarrow \operatorname{Hom}(X^-, \mu_l)$$

arising from Kummer theory, *a priori* an injection, is in fact an isomorphism. We may conclude that X^-/lX^- is a cyclic $\boldsymbol{F}_l[\Delta]$-module, and then by Nakayama's lemma that X^- is a cyclic $\boldsymbol{Z}_l[\Delta]$-module.

(3.7) THEOREM. *Suppose that each of the two eigenspaces* $X(\omega^{n-m})$ *and* $X(\omega^{m-n})$ *is cyclic. Then there exists a* $g \in \operatorname{Gal}(K_l/\boldsymbol{Q}(\mu_l))$ *for which*

$$B = \boldsymbol{T}\cdot b(g), \qquad C = \boldsymbol{T}\cdot c(g), \qquad I = \boldsymbol{T}\cdot \eta(g).$$

[The characters ω^{n-m} and ω^{m-n} are not assumed to be distinct.]

PROOF. We subject the function $b: G \to B$ to the following: we compose it with the projection $B \to B/\mathfrak{M}B$, and we restrict it to the subgroup $\operatorname{Gal}(K_l/\boldsymbol{Q}(\mu_l))$ of G. Let $\bar{b}$ be the new function that we obtain in this way. Since the values of a and of d on this subgroup are all congruent to 1 mod $\mathfrak{M}$, $\bar{b}$ is a homomorphism. Now $B/\mathfrak{M}B$ is an abelian l-group (in fact, an $\boldsymbol{F}_l$-vector space), so $\bar{b}$ must factor through X. A matrix calculation shows that

$$\bar{b}(\sigma\tau\sigma^{-1})=\omega^{n-m}(\sigma)\cdot\bar{b}(\tau)$$

for $\sigma\in G$, $\tau\in\mathrm{Gal}(K_l/\boldsymbol{Q}(\mu_l))$; thus, more precisely, $\bar{b}$ factors through the cyclic quotient $X(\omega^{n-m})$ of X. Therefore, if g is any element of $\mathrm{Gal}(K_l/\boldsymbol{Q}(\mu_l))$ whose image in $X(\omega^{n-m})$ generates $X(\omega^{n-m})$, then the image of $\bar{b}$ is the cyclic group generated by $\bar{b}(g)$. Thus $B/\mathfrak{M}B$ is generated as a $\boldsymbol{T}$-module (or as a group: the two notions coincide since $\boldsymbol{T}/\mathfrak{M}$ is the prime field $\boldsymbol{F}_l$) by $\bar{b}(g)$. By Nakayama's lemma, B is generated as a $\boldsymbol{T}$-module by $b(g)$.

Analogously, if g maps to a generator of $X(\omega^{m-n})$, then $C=\boldsymbol{T}\cdot c(g)$. Taking a g which maps to generators of both $X(\omega^{n-m})$ and $X(\omega^{m-n})$, we find that B is generated by $b(g)$ and C by $c(g)$. Hence $I=B\cdot C$ is generated by $b(g)c(g)$; by Nakayama's lemma, together with (3.5), it is generated alternately by $\eta(g)$.

REMARK. The above argument may be useful even when the $X(\omega^{\pm(n-m)})$ are not assumed to be cyclic. It provides a definite list of elements of I which generate I, the list reducing to a 1-element list in case of cyclicity.

(3.8) COROLLARY. *Suppose that Vandiver's conjecture is true for l and that I is non-zero. Then, after replacement of ρ by a conjugate $N\rho N^{-1}$ (with $N\in GL(2, E)$), the representation ρ takes values in $GL(2, \boldsymbol{T})$ and its matrix coefficients satisfy:*

$$a\equiv\varphi, \qquad d\equiv\psi, \qquad c\equiv 0 \qquad (\mathrm{mod}\, I). \tag{3.9}$$

PROOF. Let $\beta=b(g)$, $\gamma=c(g)$, with g as above. Then β is non-zero, since $I=(\beta\gamma)$ is non-zero. Taking $N=\begin{pmatrix}\beta^{-1} & 0\\ 0 & 1\end{pmatrix}$, we obtain a conjugate with the required properties.

REMARK. The hypothesis $I\neq 0$ is visibly satisfied whenever ρ is irreducible as a 2-dimensional representation of G. Conversely, if ρ is reducible, then (2.2) shows that I is 0.

4. In this § we will determine precisely the image of ρ, under the following four assumptions:

1) The characters ω^{n-m} and ω^{m-n} are distinct.
2) The ideal I is non-zero and is principal.
3) The representation ρ takes values in $\boldsymbol{GL}(2, \boldsymbol{T})$ and its coefficients satisfy (3.9).

4) The determinant of ρ, *det*, is $\boldsymbol{Z}_l^*$-valued.

Before beginning to do this, we should make comments about these axioms. The second and the third are obviously legacies of § 3. The fourth, or something like it, is needed to control the following phenomenon: if we replace ρ by the twist of ρ by a character of G, then $\boldsymbol{T}$ can change significantly, whereas the image of ρ is essentially unchanged. The first axiom means that ω^{n-m} is not *quadratic,* since we have already been assuming that it is non-trivial. The case where ω^{n-m} is quadratic is discussed by Swinnerton-Dyer in [5], and it is certain that his methods will give information in our more general setting. Finally, it might be worth noting that (1) excludes the case $l=3$.

From now on, we shall always assume that 1, 2, 3, and 4 above are true. We remind the reader that H denotes the Galois group $\mathrm{Gal}(K_l/\boldsymbol{Q}(\mu_{l^\infty}))$.

(4.1) THEOREM. *We have*

$$\rho(H)=\left\{\begin{pmatrix} a & b \\ c & d \end{pmatrix}\in \boldsymbol{SL}(2,\ \boldsymbol{T})\ \middle|\ a\equiv 1,\ d\equiv 1,\ c\equiv 0 \pmod{I}\right\}.$$

Let X denote the right-hand group. It is evident that $\rho(H)$ is contained in X, since φ and ϕ vanish on H. Our proof has two main steps: we first show that $\rho(H)$ maps onto a certain (rather modest) quotient of X, and we then show that any closed subgroup on X which maps onto this quotient must in fact be equal to X. In this sense, our theorem follows the pattern of results previously obtained by Serre [4, Lemma 3, p. IV-23] and by Swinnerton-Dyer [5, Th. 2, p. 75]. These two authors pass to larger and larger quotients of X by a technique involving formation of l^{th} powers. Here we do something a bit different: we pass to larger and larger quotients by taking commutators of pairs of elements. We learned this technique from an argument used by Mazur in proving a similar (unpublished) theorem; we will point out this argument when it appears below.

For $n\geqq 1$, let X_n be the image of X in $\boldsymbol{SL}(2,\ \boldsymbol{T}/I^n)$, namely the analogue of X with $\boldsymbol{T}$ replaced by the ring $\boldsymbol{T}/I^n$. It is enough to show that $\rho(H)$ maps onto each X_n. We show first that $\rho(H)$ maps onto X_2 and then that any subgroup of X_n ($n\geqq 3$) which maps onto X_{n-1} must in fact be equal to all of X_n.

(4.2) Let $\theta: X\to \boldsymbol{T}/I\times I/I^2$ be the map

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\longmapsto (b \bmod I,\ c \bmod I^2).$$

Then $\theta|_{\rho(H)}$ is surjective.

PROOF. As in § 3, let B and C be the ideals of $\boldsymbol{T}$ generated by the sets $b(H)$, $c(H)$. We have as before $B\subseteq T$, $C\subseteq I$, and $BC=I$. It follows that $C=I$, and then by Nakayama's lemma that $B=\boldsymbol{T}$. *A fortiori*, if we regard b as a map

$$b: H \longrightarrow \boldsymbol{T}/I$$

and c as a map

$$c: H \longrightarrow I/I^2,$$

we find that the targets in both cases are generated as $\boldsymbol{T}$-modules by the images of the maps. However, we can show that the images are $\boldsymbol{T}$-submodules of the targets, thereby proving that the maps are surjective.

For the sake of brevity, we will give the argument for this assertion only in the case of b. We note, first, that $b(H)$ is a subgroup of $\boldsymbol{T}/I$ because we have

$$a\equiv 1, \qquad d\equiv 1 \qquad (\mathrm{mod}\, I)$$

on H. On the other hand, we have in $\boldsymbol{T}/I$ the formula

$$b(\sigma\tau\sigma^{-1}) = (\varphi\psi^{-1})(\sigma)\cdot b(\tau), \tag{4.3}$$

which refines the formula used in the proof of (3.7); here σ is intended to be an element of G and τ to be an element of H. It shows that the set $b(H)$ is stable under multiplication by elements of the ring generated by the values of $\varphi\psi^{-1}$. Using the hypothesis that I is non-zero, we see that I has finite index in $\boldsymbol{T}$. Therefore, $b(H)$ is actually stable by the $\boldsymbol{Z}_l$-subalgebra R of $\boldsymbol{T}$ generated by the values of $\varphi\psi^{-1}$. As shown in (2.6), R maps onto $\boldsymbol{T}/I$. Thus, finally, $b(H)$ is stable under multiplication by elements of $\boldsymbol{T}$; it is therefore a $\boldsymbol{T}$-submodule of $\boldsymbol{T}/I$ and so is equal to $\boldsymbol{T}/I$.

To summarize, we have shown that b and c are surjective; we must now show that the product map (b, c) is surjective. We will refer to this map simply as θ. Suppose that $(\beta, \gamma)\in T/I\times I/I^2$ is in the image of θ. Choose $g\in G$ such that $u=(\varphi\psi^{-1})(g)$ is not congruent to $+1$ or -1 modulo $\mathfrak{M}$. Let $v\geqq 1$ be an integer congruent to u mod $\mathfrak{M}$. The image of θ contains $(v\beta, v\gamma)$ and also, because of (4.3), the couple $(u\beta, u^{-1}\gamma)$. Hence it contains

$$((u-v)\beta, (u^{-1}-v)\gamma).$$

Repeating the argument, we find that the image of θ contains

$$((u-v)^N\beta, (u^{-1}-v)^N\gamma)$$

for all integers $N\geqq 1$. For large N we have $(u-v)^N\in I$, since $u-v\in\mathfrak{M}$ and $\boldsymbol{T}/I$ is finite. On the other hand, $u^{-1}-v$ is a unit in $\boldsymbol{T}$, because of the way u was chosen. Thus, by the surjectivity of c, the image of θ contains all elements of

$T/I \times I/I^2$ of the form $(0, \gamma)$. This, together with the surjectivity of b, gives the surjectivity of θ.

(4.4) *The group $\rho(H)$ maps onto X_2.*

PROOF. Already this assertion will be a formal consequence of (4.2), i.e., a purely group theoretical statement having nothing to do with ρ. Namely, let Y be a subgroup of X_2 such that the map $\theta|_Y$ is surjective. Then we will show that Y coincides with X_2. We do this in two stages, each involving a commutator argument.

First, let

$$\Theta : X_2 \longrightarrow I/I^2 \times T/I \times I/I^2$$

be the map

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \longmapsto (a-1 \bmod I^2,\ b \bmod I,\ c \bmod I^2).$$

It becomes a homomorphism of groups when we give $I/I^2 \times T/I \times I/I^2$ the multiplication law

$$(\alpha, \beta, \gamma) * (\alpha', \beta', \gamma') = (\alpha + \alpha' + \beta\gamma',\ \beta + \beta',\ \gamma + \gamma'),$$

cf. [5, pp. 71-72]. Assuming that $\theta|_Y$ is surjective, we wish to see that $\Theta|_Y$ is surjective.

For this, it suffices to show that $\Theta(Y)$ contains all $(\alpha, 0, 0)$ with $\alpha \in I/I^2$. Given α, choose $y \in Y$ such that $\theta(y) = (0, \alpha)$ and $y' \in Y$ such that $\theta(y') = (1, 0)$. If y'' is the commutator of y and y', we find by a computation that $\Theta(y'') = (\alpha, 0, 0)$.

Now, assuming that $\Theta|_Y$ is surjective, we will show that $Y = X_2$ by showing that Y contains the kernel of Θ, which is the group

$$\left\{ \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix} \in SL(2, T/I^2) \,\middle|\, t \in I/I^2 \right\}.$$

(4.5) LEMMA. *Given $t \in I/I^2$, there exist $x, y \in I$ such that x and y are generators of the ideal I and such that*

$$t \equiv x - y \pmod{I^2}.$$

PROOF. Choose a representative for t in I, and denote this representative again by t. Let z be a generator of the ideal I. We have:

$$t = (t+z) - z\,; \qquad t = (t-z) - (-z).$$

It is easy to see that one of $(t \pm z)$ is a generator of I.

Indeed, suppose that $t=uz$ with $u\in \boldsymbol{T}$. Since l is an odd prime, u cannot be congruent both to $+1$ and to -1 modulo $\mathfrak{M}$. Thus one of $u\pm 1$ is a unit.

Now, given t, choose x and y as in the lemma. Let $b\in \boldsymbol{T}$ be the unit for which $y=bx$. Let $v=b^{-1}x$. Choose $M, N\in Y$ such that:

$$\Theta(M)=(0, b, x); \qquad \Theta(N)=(0, 1, v).$$

After some calculation, we find

$$MNM^{-1}N^{-1}=\begin{pmatrix}1 & t\\ 0 & 1\end{pmatrix},$$

thus completing the proof of (4.4).

We next consider briefly the $\boldsymbol{T}$-module $sl_2(\boldsymbol{T}/I)$ of 2×2 matrices over $\boldsymbol{T}/I$ which have trace 0. If β and β' are 2×2 matrices over $\boldsymbol{T}/I$, we set

$$[\beta, \beta']=\beta\beta'-\beta'\beta\in sl_2(\boldsymbol{T}/I).$$

(4.6) LEMMA. *Every matrix in $sl_2(\boldsymbol{T}/I)$ is a sum of elements of the form $[\beta, \beta']$.*

PROOF. Using that 2 is invertible in $\boldsymbol{T}/I$, one can prove this directly from the three formulas:

$$\left[\begin{pmatrix}0 & t\\ 0 & 0\end{pmatrix}, \begin{pmatrix}0 & 0\\ 1 & 0\end{pmatrix}\right]=\begin{pmatrix}t & 0\\ 0 & -t\end{pmatrix},$$

$$\left[\begin{pmatrix}a & 1\\ -1 & 0\end{pmatrix}, \begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}\right]=\begin{pmatrix}0 & a\\ a & 0\end{pmatrix},$$

$$\left[\begin{pmatrix}a & 1\\ 1 & 0\end{pmatrix}, \begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}\right]=\begin{pmatrix}0 & a\\ -a & 0\end{pmatrix}.$$

The lemma established, we will now complete the proof of (4.1) by using Mazur's argument which was alluded to above. Namely, we will show:

(4.7) *Suppose that Y is a subgroup of X_n (for $n\geqq 3$) which maps onto X_{n-1}. Then Y is equal to X_n.*

PROOF. We will show that Y contains the kernel of the natural map $X_n\to X_{n-1}$. A typical element of this kernel may be written

$$N=1+x^{n-1}M,$$

where x is a generator of the ideal I and M is a matrix with coefficients in $\boldsymbol{T}/I$. The condition $N\in \boldsymbol{SL}(2, \boldsymbol{T}/I^n)$ means precisely that M belongs to $sl_2(\boldsymbol{T}/I)$. By (4.6), we may assume that M is the commutator $[\beta, \beta']$. Supposing that this is so, we choose representatives for β and β' in $M(2, T)$ and denote these representatives again by β and β'. The determinants of $1+x\beta$ and $1+x^{n-2}\beta'$ are squares in $\boldsymbol{T}$, since they are congruent to 1 modulo $\mathfrak{M}$. Thus we may find $\alpha, \alpha' \in \boldsymbol{T}^*$ such that

$$\alpha(1+x\beta), \qquad \alpha'(1+x^{n-2}\beta') \in \boldsymbol{SL}(2, \boldsymbol{T}).$$

By induction, there exist $A, A' \in Y$ such that we have the mod I^{n-1} congruences:

$$A \equiv \alpha(1+x\beta)$$

$$A' \equiv \alpha'(1+x^{n-2}\beta').$$

Again, a computation gives

$$AA'A^{-1}A'^{-1} = 1+x^{n-1}M,$$

thus proving (4.7) and (4.1).

For the final results, it is convenient to introduce the following abuse of notation. We have already noted that each character $G \to \boldsymbol{T}^*$ is the composition of the cyclotomic character χ and a unique character $\boldsymbol{Z}_l^* \to \boldsymbol{T}^*$. Given a character of G, we will denote the corresponding character of $\boldsymbol{Z}_l^*$ by the *same* symbol. This abuse will be applied in the case of the three characters φ, ψ, and $det = \varphi\psi$.

Let $(\rho, \chi): G \to \boldsymbol{GL}(2, \boldsymbol{T}) \times \boldsymbol{Z}_l^*$ be the map given by

$$g \longmapsto (\rho(g), \chi(g)).$$

Then we have

(4.8) THEOREM. *The image of* (ρ, χ) *is the subgroup of* $\boldsymbol{GL}(2, \boldsymbol{T}) \times \boldsymbol{Z}_l^*$ *consisting of all pairs* $\left(\begin{pmatrix} a & b \\ c & d \end{pmatrix}, t\right)$ *which satisfy the conditions*:

$$\text{(4.9)} \qquad \begin{cases} ad-bc = \det(t) \\ a \equiv \varphi(t), \quad d \equiv \psi(t), \quad c \equiv 0 \quad (\mathrm{mod}\, I). \end{cases}$$

PROOF. It is clear that the image is contained in this group. Then the theorem follows immediately from (4.1) and the surjectivity of χ.

(4.10) COROLLARY. *The image of* ρ *is the group of matrices* $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ *such that* a, b, c, d *satisfy* (4.9) *for some* $t \in \boldsymbol{Z}_l^*$.

(4.11) COROLLARY. *The image of the map* (tr, det)$: G \to \boldsymbol{T} \times \boldsymbol{Z}_l^*$ *consists of all pairs* $(\alpha, \beta) \in \boldsymbol{T} \times \boldsymbol{Z}_l^*$ *satisfying*:

$$\text{(4.12)} \qquad \begin{cases} \beta = \det(t) \\ \alpha \equiv \varphi(t) + \psi(t) \quad (\mathrm{mod}\, I) \end{cases}$$

for some $t \in \boldsymbol{Z}_l^*$.

PROOF. Evidently, all pairs in the image satisfy (4.12). Conversely, suppose that (α, β) satisfies (4.12) with the element t of $\boldsymbol{Z}_l^*$. We put:

$$a=\varphi(t),$$

$$b=1,$$

$$c=\varphi(t)[\alpha-\varphi(t)-\psi(t)],$$

$$d=\alpha-\varphi(t).$$

The matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ has trace α and determinant β. By (4.11) we see that it lies in the image of ρ.

References

[0] Iwasawa, K., A note on cyclotomic fields, Invent. Math. **36** (1976), 115-123.
[1] Mazur, B., Modular curves and the Eisenstein ideal, Publ. Math. I.H.E.S. **47** (1977), 33-186.
[2] Papier, E., Thèse de 3e cycle, Paris, 1981.
[3] Papier, E., to appear.
[4] Serre, J-P., Abelian l-adic Representations and Elliptic Curves, New York, Benjamin, 1968.
[5] Swinnerton-Dyer, H.P.F., On l-adic representations and congruences for coefficients of modular forms (II), Lecture Notes in Math. **601**, Springer, 1977, 63-90.
[6] Wagstaff, S., The irregular primes to 125,000, Math. Comp. **32** (1978), 583-591.

(Received July 6, 1981)

Elisabeth Papier
Ecole Normale Supérieure
Mathématiques
1, rue Maurice Arnoux
92120 Montrouge
France

Kenneth A. Ribet
Mathematics Department
University of California
Berkeley, Ca. 94720
U.S.A.