Inventiones math. 34, 151–162 (1976)

Inventiones mathematicae

A Modular Construction of Unramified p-Extensions of $Q(\mu_p)$

Kenneth A. Ribet* (Princeton)

§ 1. Introduction

An odd prime p is called irregular if the class number of the field $\mathbf{Q}(\mu_p)$ is divisible by p (μ_p being, as usual, the group of p-th roots of unity). According to Kummer's criterion, p is irregular if and only if there exists an even integer k with $2 \leqq k \leqq p-3$ such that p divides (the numerator of) the k-th Bernoulli number B_k, given by the expansion

$$\frac{t}{e^t-1}+\frac{t}{2}-1=\sum_{n\geqq 2}\frac{B_n}{n!}t^n.$$

The purpose of this paper is to strengthen Kummer's criterion.

Let A be the ideal class group of $\mathbf{Q}(\mu_p)$, and let C be the $\mathbf{F}_p$-vector space A/A^p. The Galois group $\mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q})$ acts on C through its quotient $\Delta=\mathrm{Gal}(\mathbf{Q}(\mu_p)/\mathbf{Q})$. Since all characters of Δ with values in $\bar{\mathbf{F}}_p^*$ are powers of the standard character

$$\chi\colon \mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q}) \to \Delta \xrightarrow{\sim} \mathbf{F}_p^*$$

giving the action of $\mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q})$ on μ_p, the vector space C has a canonical decomposition

$$C=\bigoplus_{i \bmod (p-1)} C(\chi^i),$$

where

$$C(\chi^i)=\{c\in C \mid \sigma c=\chi^i(\sigma)c \text{ for all } \sigma\in\Delta\}.$$

(1.1) **Main Theorem.** *Let k be even, $2\leqq k\leqq p-3$. Then $p \mid B_k$ if and only if $C(\chi^{1-k})\neq 0$.*

In fact, the statement that $C(\chi^{1-k})\neq 0$ implies $p\mid B_k$ is well known [8, Th. 3]. Its converse is also familiar as a consequence of the conjecture that p is prime to the class number of the real subfield $\mathbf{Q}(\mu_p)^+$ of $\mathbf{Q}(\mu_p)$ [8, p. 434]. Thus the con-

---

* Sloan Fellow, and visitor at I.H.E.S.

tribution of this paper is to prove that $p|B_k$ implies $C(\chi^{1-k}) \neq 0$ *without* making a supplementary hypothesis.

By a "functoriality" formula for the Artin symbol [20, Th. 11.5, p. 199], this implication is equivalent to

(1.2) **Theorem.** *Suppose $p|B_k$. Then there exists a Galois extension $E/\mathbf{Q}$ containing $\mathbf{Q}(\mu_p)$ with the following properties:*

$$\left.H\left(\begin{array}{c} E \\ | \\ \mathbf{Q}(\mu_p) \end{array}\right.\begin{array}{c} \\ \\ \\ | \\ \mathbf{Q}\end{array}\right)G$$

(a) *The extension $E/\mathbf{Q}(\mu_p)$ is unramified.*
(b) *The group H is a non-zero abelian group of type $(p, \ldots, p)$, i.e., killed by p.*
(c) *If $\sigma \in G$ and $\tau \in H$, then*

$$\sigma\tau\sigma^{-1} = \chi(\sigma)^{1-k} \cdot \tau.$$

In fact, we shall prove (1.2) with $\mathbf{Q}(\mu_p)$ replaced by the unique subfield $\mathbf{Q}(\mu_p^{\otimes(1-k)})$ of $\mathbf{Q}(\mu_p)$ whose degree over $\mathbf{Q}$ is $(p-1)/(p-1, k-1)$. This subfield is the field corresponding to the kernel in $\mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q})$ of χ^{1-k}.

(1.3) **Theorem.** *Suppose $p|B_k$. Then there exists a finite field $\mathbf{F} \supseteq \mathbf{F}_p$ and a continuous representation*

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

with the properties:

(i) *$\bar{\rho}$ is unramified at all primes $l \neq p$.*

(ii) *The representation $\bar{\rho}$ is reducible (over $\mathbf{F}$) in such a way that $\bar{\rho}$ is isomorphic to a representation of the form*

$$\begin{pmatrix} 1 & * \\ 0 & \chi^{k-1} \end{pmatrix}.$$

That is, $\bar{\rho}$ is an extension of the 1-dimensional representation with character χ^{k-1} by the trivial 1-dimensional representation.

(iii) *The image of $\bar{\rho}$ has order divisible by p. In other words, $\bar{\rho}$ is not diagonalizable.*

(iv) *Let D be a decomposition group for p in $\mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q})$. Then $\bar{\rho}(D)$ has order prime to p, i.e., $\bar{\rho}|D$ is diagonalizable.*

Notice that (1.3) implies (1.2). Indeed, if $\bar{\rho}$ satisfies the above properties, then the image of $\bar{\rho}$ is the Galois group of an extension $E/\mathbf{Q}$ such that E is of type $(p, \ldots, p)$ over the field $\mathbf{Q}(\mu_p^{\otimes(1-k)})$. Now $E/\mathbf{Q}$ is unramified outside p by (i), and the $(p, \ldots, p)$ layer is a non-trivial extension by (iii). This $(p, \ldots, p)$ extension is unramified at (the unique prime over) p by (iv); hence it is *everywhere* unramified. Finally, the conjugation formula (c) of (1.2) follows from the matrix identity

$$\begin{pmatrix} a & b \\ 0 & d \end{pmatrix} \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a & b \\ 0 & d \end{pmatrix}^{-1} = \begin{pmatrix} 1 & ad^{-1}x \\ 0 & 1 \end{pmatrix}.$$

In proving (1.3) we begin by "finding" $\bar{\rho}$ in the p-adic representation associated with the modular variety $J_1(p)$ attached to forms of weight 2 on $\Gamma_1(p)$. Assuming that $p|B_k$, we construct a normalized eigenform $f=\sum a_n q^n$ in the space of such cusp forms which satisfies

$$a_l \equiv 1 + l^{k-1} \bmod \mathcal{M}$$

for all primes $l \neq p$, where $\mathcal{M}$ is a certain fixed ideal over p in the field generated by the coefficients a_n. This leads to our $\bar{\rho}$, and by the time we have constructed $\bar{\rho}$ we know from the construction that (i), (ii), and (iii) of (1.3) are satisfied by $\bar{\rho}$. It then remains to prove (iv). We then use the theorem of Deligne-Rapoport that the variety $J_1(p)/J_0(p)$ acquires everywhere good reduction over the real subfield $\mathbf{Q}(\mu_p)^+$ of $\mathbf{Q}(\mu_p)$ [5]. This implies that, locally at p, $\bar{\rho}|\mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q}(\mu_p)^+)$ is the representation attached to a finite flat commutative group scheme of type $(p, \ldots, p)$ over the integer ring of the completion $\mathbf{Q}(\mu_p)^+ \otimes \mathbf{Q}_p$. We note especially that the absolute ramification index of this completion is $(p-1)/2 < p-1$; this enables us to prove (iv) by applying results of Raynaud [15] on group schemes of type $(p, \ldots, p)$.

Our proof is motivated by two key ideas of Serre. The first idea (cf. [16]) is that the divisibility of B_k by p implies a congruence similar to the above one for some cusp form of weight k on $\mathbf{SL}(2, \mathbf{Z})$; hence a representation such as our $\bar{\rho}$ should be obtainable from the Deligne representation ρ_k attached to forms of weight k on $\mathbf{SL}(2, \mathbf{Z})$. Although our methods "find" in ρ_k a representation $\bar{\rho}$ which satisfies the first three properties of (1.3), a proof that this representation satisfies (iv) would seem to require unknown Galois-theoretic properties of étale cohomology. This leads to the second idea of Serre, that (mod p) representations coming from ρ_k ought to be visible (at least up to twist) on the Jacobian variety $J_1(p)$. (A similar idea is the starting point in a recent paper of Koike [10].) This is what led us to look at forms of weight 2.

We hope that our method will apply also to more general Kummer-like criteria, such as that given by Greenberg [7]. Some relevant computations have been made by Yamauchi [21].

The author wishes to acknowledge the invaluable assistance of N. Katz during the time this work was carried out. He is further indebted to J. Coates, P. Deligne, B. Mazur, and J-P. Serre for many helpful conversations.

§ 2. Reductions of Reducible Representations

Let K be a finite extension of $\mathbf{Q}_p$. Let $\mathcal{O}$ be its integer ring, $\mathbf{F}$ the residue field, and π a uniformizing parameter. Let V be a free module of rank 2 over K. A *lattice* in V is a free $\mathcal{O}$-module of rank 2 in V which generates V over K.

We suppose given a representation

$$\rho\colon\ G \to \mathbf{GL}(V)$$

of a group in V such that G leaves stable *some* lattices of V. (This latter condition is always satisfied if G is compact and ρ is continuous, for example.) If $T \subset V$ is stable by G; then G acts on $T/\pi T$, which is free of rank 2 over F. The associated map

$$\bar{\rho}\colon\ G \to \mathbf{GL}(T/\pi T)$$

will be called the *reduction* of ρ attached to T. It is known that the semi-simplification of $\bar{\rho}$ (as an $\mathbf{F}$-representation) is independent of the choice of T [4, 30.16], so that $\bar{\rho}$ is unique if one reduction (and hence every reduction) is simple.

We consider, however, the opposite situation, where the reductions are all reducible. Their semi-simplifications are then described by two characters φ_1, $\varphi_2\colon G \to \mathbf{F}^*$, which do not depend on the choice of T. A given reduction may be written matricially in one of the forms:

$$\begin{pmatrix} \varphi_1 & * \\ 0 & \varphi_2 \end{pmatrix}, \quad \begin{pmatrix} \varphi_1 & 0 \\ * & \varphi_2 \end{pmatrix}.$$

It is diagonalizable (i.e., semi-simple) if and only if its image has order prime to p.

(2.1) **Proposition.** *Suppose that the K-representation ρ is simple but that its reductions are reducible. Let φ_1 and φ_2 be the characters associated to the reductions of ρ. Then G leaves stable some lattice $L \subset V$ for which the associated reduction is of the form* $\begin{pmatrix} \varphi_1 & * \\ 0 & \varphi_2 \end{pmatrix}$ *but is not semi-simple.*

Proof. Choose a G-stable lattice of V together with an $\mathcal{O}$-basis of this lattice. Then ρ may be viewed as a map $G \to \mathbf{GL}(2, \mathcal{O})$. Any matrix $M \in \mathbf{GL}(2, K)$ such that $M\rho(G)M^{-1} \subseteq \mathbf{GL}(2, \mathcal{O})$ then defines another G-stable lattice together with a basis of it. The reduction attached to this new lattice is the map

$$G \to M\rho(G)M^{-1} \hookrightarrow \mathbf{GL}(2, \mathcal{O}) \to \mathbf{GL}(2, \mathbf{F}).$$

To prove the proposition, we do some calculations based on the formula

$$P \begin{pmatrix} a & \pi b \\ c & d \end{pmatrix} P^{-1} = \begin{pmatrix} a & b \\ \pi c & d \end{pmatrix},$$

where P is the matrix $\begin{pmatrix} 1 & 0 \\ 0 & \pi \end{pmatrix}$.

We first note that we may assume at the outset that the reduction of the given map $G \to \mathbf{GL}(2, \mathcal{O})$ is of the form $\begin{pmatrix} \varphi_1 & * \\ 0 & \varphi_2 \end{pmatrix}$ rather than the form $\begin{pmatrix} \varphi_1 & * \\ * & \varphi_2 \end{pmatrix}$, because if the latter occurs we can divide the upper-right corner entries by π and multiply the lower-left corner entries by π using the formula above. Let us make this assumption together with the following one: each reduction $\bar{\rho}$ of the form $\begin{pmatrix} \varphi_1 & * \\ 0 & \varphi_2 \end{pmatrix}$

is semi-simple. With these assumptions, we will show that ρ is itself reducible, and thus prove (2.1) by contradiction.

Set $M_0 = I$ (2×2 identity matrix). Inductively, we will define a converging sequence of matrices $M_i = \begin{pmatrix} 1 & t_i \\ 0 & 1 \end{pmatrix}$ such that $M_i \rho(G) M_i^{-1}$ consists of elements of $\mathbf{GL}(2, \mathcal{O})$ whose lower-left corner entries are divisible by π and whose upper-right corner entries are divisible by π^i. This will prove that ρ is reducible because the matrix $M = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}$ with $t = \operatorname{Lim} t_i$ will then be such that $M\rho(G)M^{-1}$ consists of matrices whose upper-right corner entries are 0.

According to the conjugation formula above, the induction assumption may be rephrased as follows: $P^i M_i \rho(G) M_i^{-1} P^{-i}$ consists of integral matrices whose lower-left corner entries are divisible by π^{i+1}. With this assumption, the representation $\sigma \mapsto P^i M_i \rho(\sigma) M_i^{-1} P^{-i} \pmod{\pi}$ is in the form $\begin{pmatrix} \varphi_1 & * \\ 0 & \varphi_2 \end{pmatrix}$ because $\sigma \mapsto \rho(\sigma) \pmod{\pi}$ is of this form. The representation in question is then by assumption semi-simple, so we may choose an element u of $\mathcal{O}$ such that the matrix $U = \begin{pmatrix} 1 & u \\ 0 & 1 \end{pmatrix}$ diagonalizes the (mod π) representation. That is, we can find a u in $\mathcal{O}$ so that

$$U P^i M_i \rho(G) M_i^{-1} P^{-i} U^{-1}$$

consists of matrices whose upper-right corner entries are divisible by π (and whose lower-left corner entries are still divisible by π^{i+1}: conjugation by U leaves unchanged the lower-left corner of any matrix). This gives that

$$(P^{-i} U P^i M_i)\, \rho(G) (P^{-i} U P^i M_i)^{-1}$$

consists of integral matrices whose lower-left corner entries are divisible by π and whose upper-right corner entries are divisible by π^{i+1}. Thus we may continue the induction by setting

$$M_{i+1} = P^{-i} U P^i M_i = \begin{pmatrix} 1 & t_i + \pi^i u \\ 0 & 1 \end{pmatrix}.$$

This formula makes visible the fact that $\{M_i\}$ converges.

§ 3. A Congruence between a Cusp Form and an Eisenstein Series

Let p be an odd prime and let μ_{p-1} be the group of complex $(p-1)$-st roots of unity. We consider modular forms of weights 1 and 2 on $\Gamma_1(p)$. For a character

$$\varepsilon\colon (\mathbf{Z}/p\mathbf{Z})^* \to \mu_{p-1}$$

(possibly the trivial one) we say that a form is of type ε if it satisfies the equation

$$f \left| \begin{pmatrix} a & b \\ c & d \end{pmatrix} \right. = \varepsilon(d) \cdot f$$

for all $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ in $\Gamma_0(p)$. (We lift ε as usual to a function on $\mathbf{Z}$.) A form of type ε is a cusp form if its q-expansion and that of $f \Big| \begin{pmatrix} 0 & -1 \\ p & 0 \end{pmatrix}$ both commence with 0; if the q-expansion of f commences with 0, then we say that f is a semi cusp form.

We will have need of the Eisenstein series. Let ε be a non-trivial even character. Then the two series

$$G_{2,\varepsilon} = L(-1, \varepsilon)/2 + \sum_{n \geqq 1} \sum_{d \mid n} \varepsilon(d)\, d q^n,$$

$$s_{2,\varepsilon} = \sum_{n \geqq 1} \sum_{d \mid n} \varepsilon(n/d)\, d q^n$$

are each of weight 2 and type ε. The space of modular forms of weight 2 and type ε is generated by the cusp forms and these two series, while the space of semi cusp forms of weight 2 and type ε is generated by $s_{2,\varepsilon}$ and the cusp forms. When ε is the trivial character, we still have an Eisenstein series $G_{2,\varepsilon}$ as above; it may be written

$$\frac{p-1}{24} + \sum_{n \geqq 1} \sum_{\substack{d \mid n \\ p \nmid d}} d q^n.$$

In weight 1 we use the series

$$G_{1,\varepsilon} = L(0, \varepsilon)/2 + \sum_{n \geqq 1} \sum_{d \mid n} \varepsilon(d)\, q^n$$

when ε is an *odd* character. The Eisenstein series are eigenforms for the Hecke operators $T(n)$, at least when n is prime to p.

Now fix a prime ideal $\mathfrak{p} \mid p$ of the field $\mathbf{Q}(\mu_{p-1})$. Then let $\omega\colon (\mathbf{Z}/p\mathbf{Z})^* \xrightarrow{\sim} \mu_{p-1}$ be the unique character which satisfies

$$\omega(d) \equiv d \pmod{\mathfrak{p}}$$

for all $d \in \mathbf{Z}$.

(3.1) **Lemma.** *Let k be even, $2 \leqq k \leqq p-3$. Then the modular forms $G_{2,\omega^{k-2}}$ and $G_{1,\omega^{k-1}}$ have $\mathfrak{p}$-integral q-expansions in $\mathbf{Q}(\mu_{p-1})$ which are congruent modulo $\mathfrak{p}$ to the q-expansion*

$$-B_k/2k + \sum_{n \geqq 1} \sum_{d \mid n} d^{k-1} q^n.$$

Proof. Aside from the constant terms of the series, the assertion follows immediately from the choice of ω. To prove the assertions about constant terms, we use the expresssions

$$L(0, \varepsilon) = \frac{-1}{p} \sum_{n=1}^{p-1} \varepsilon(n)(n - p/2),$$

$$L(-1, \varepsilon) = \frac{-1}{2p} \sum_{n=1}^{p-1} \varepsilon(n)(n^2 - pn + p^2/6)$$

of the L-values as generalized Bernoulli numbers, valid for *any* character $\varepsilon \pmod p$, cf. [11]. Using the congruence $\omega(n) \equiv n^p \pmod{\mathfrak{p}^2}$, we find

$$pL(0, \omega^{k-1}) \equiv -\sum_{n=1}^{p-1} n^{1+p(k-1)} \pmod{\mathfrak{p}^2},$$

$$pL(-1, \omega^{k-2}) \equiv \frac{-1}{2} \sum_{n=1}^{p-1} n^{2+p(k-2)} \pmod{\mathfrak{p}^2}.$$

On the other hand, if t is a positive even integer we have

$$pB_t \equiv \sum_{n=1}^{p-1} n^t \pmod{p^2}$$

according to [1, (8.8), p. 385]. The desired result follows by combining these facts with the Kummer congruence [1, Th. 5, p. 385].

(3.2) **Corollary.** *Let k be as above, and let n and m be even integers, $2 \leqq n, m \leqq p-3$, satisfying $n+m \equiv k \bmod (p-1)$. Then the product*

$$G_{1,\omega^{n-1}} G_{1,\omega^{m-1}}$$

is a modular form of weight 2 and type ω^{k-2} whose q-expansion coefficients are $\mathfrak{p}$-integers in $\mathbf{Q}(\mu_{p-1})$. Its constant term is a $\mathfrak{p}$-unit provided that neither B_n nor B_m is divisible by p.

Proof. Clear.

(3.3) **Theorem.** *Let k be as above. Then there exists a modular form g of weight 2 and type ω^{k-2} whose q-expansion coefficients are $\mathfrak{p}$-integers in $\mathbf{Q}(\mu_{p-1})$ and whose constant term is* 1.

Proof. It suffices to construct a g whose constant term is a $\mathfrak{p}$-unit. We first try the Eisenstein series $G_{2,\omega^{k-2}}$. By (3.1), this form will commence with a unit coefficient unless $p|B_k$. If this happens, we then try the products $G_{1,\omega^{n-1}} G_{1,\omega^{m-1}}$ as in (3.2). If none of these products works, then for every pair n, m as in (3.2) at least one of the two numbers B_n, B_m is divisible by p. Now let t be the number of even integers n, $2 \leqq n \leqq p-3$, such that p divides B_n. Then elementary reasoning shows that $t \geqq (p-1)/4$ if the theorem is false. However, we have $p^t | h_p^*$, where the integer h_p^* is the so-called first factor of the class number of $\mathbf{Q}(\mu_p)$ (see below). Hence to prove the theorem it will suffice to prove that

$$h_p^* < p^{(p-1)/4}.$$

According to Carlitz and Olson [3], we may write h_p^* in the form $\pm D/p^{(p-3)/2}$, where D is a certain determinant of dimension $(p-1)/2$ whose entries are integers between 1 and $p-1$. As Carlitz has pointed out [2], Hadamard's inequality then immediately gives

$$h_p^* < p^{(p+3)/4} 2^{-(p-1)/4}.$$

This implies the desired inequality because $h_p^* = 1$ for $p \leqq 19$ and $p \leqq 2^{(p-1)/4}$ for $p > 19$.

To prove that p^t divides h_p^* we use the expression

$$h_p^* = \alpha p \prod_{\substack{k=2 \\ k \text{ even}}}^{p-1} L(0, \omega^{k-1}),$$

where α is a certain power of 2 [7, p. 250]. It will be enough to show that $\mathfrak{p}^t$ divides h_p^* since $\mathfrak{p}$ is unramified. Now, by the $L(0, \varepsilon)$ formula given above, the quantity $p \cdot L(0, \omega^{p-2})$ is an algebraic integer. Thus what we want follows from (3.1): if $p | B_k$ with $2 \leqq k \leqq p-3$, then $\mathfrak{p}$ divides $L(0, \omega^{k-1})$.

Remarks. 1. Masley and Montgomery [13] give the bounds

$$(2\pi)^{-p/2} p^{(p-25)/4} \leqq h_p^* \leqq (2\pi)^{-p/2} p^{(p+31)/4}$$

for primes p bigger than 200. This shows that the elementary upper bound for h_p^* that we use is in fact reasonably sharp.

2. Theorem (3.3) may be proved more conceptually by methods of Mazur [14], using the Deligne-Rapoport study of the modular curve $X_1(p)$ at the prime p [5, p. DeRa–108]. One sees by Mazur's technique that g may be chosen so as to vanish at the cusp 0 of $X_1(p)$.

From this point on, we fix an even integer k $(2 \leqq k \leqq p-3)$ and make the assumption that $p | B_k$. We put $\varepsilon = \omega^{k-2}$. Since $B_2 = 1/6$, k is in fact at least 4; hence ε is a non-trivial even character. All modular forms will now be of weight 2 and type ε.

(3.4) **Proposition.** *There exists a semi cusp form $f = \sum_{n \geqq 1} a_n q^n$ such that the a_n are $\mathfrak{p}$-integers in $\mathbf{Q}(\mu_{p-1})$ and such that*

$$f \equiv G_k \equiv G_{2,\varepsilon} \bmod \mathfrak{p}$$

in q-expansions.

Proof. Take $f = G_{2,\varepsilon} - c \cdot g$, where c is the constant term of $G_{2,\varepsilon}$. Then f is a semi cusp form by construction, and we have $f \equiv G_{2,\varepsilon}$ because $\mathfrak{p} | c$ by (3.1) and the assumption $p | B_k$. Also $G_{2,\varepsilon} \equiv G_k$ by (3.1).

(3.5) **Proposition.** *There exists a non-zero cusp form f' of type ε which is an eigenform for all Hecke operators T_n with $(n, p) = 1$ and which has the property that for each prime $l \neq p$ the eigenvalue $\lambda(l)$ of $T(l)$ acting on f' satisfies*

$$\lambda(l) \equiv 1 + l^{k-1} \equiv 1 + \varepsilon(l)\, l \bmod \mathscr{M},$$

where $\mathscr{M}$ is a certain prime (independent of l) lying over $\mathfrak{p}$ in the field $\mathbf{Q}(\mu_{p-1}; \lambda(n))$ generated by the eigenvalues over $\mathbf{Q}(\mu_{p-1})$.

Proof (cf. Koike [9]). The semi cusp form f of (3.4) is a mod $\mathfrak{p}$-eigenform for the Hecke operators, because it is congruent to the eigenform $G_{2,\varepsilon}$. Its mod $\mathfrak{p}$-eigenvalues are congruent to those desired of f'. Hence we can apply the Deligne-Serre lemma [6, 6.11] to get a *semi* cusp form f' as in the statement of the proposition. We then must show that this f' is in fact a cusp form. But as remarked above, the space of semi cusp forms is generated by the space of cusp forms and the *eigenform* $s_{2,\varepsilon}$. Hence it suffices to show that f' cannot be $s_{2,\varepsilon}$. However the eigenvalue of $T(l)$ acting on $s_{2,\varepsilon}$ is $\varepsilon(l)+l$, and it is clear that we cannot have

$$\varepsilon(l)+l \equiv 1+l\varepsilon(l) \bmod \mathfrak{p}$$

unless $\varepsilon(l)=1$. Since ε is a non-trivial character, this gives what is wanted.

(3.6) **Proposition.** *Any form f' as in (3.5) is an eigenform for all Hecke operators $T(n)$ (including those for which $p|n$). Hence, after replacing f' by a multiple of f', we have*

$$f'=\sum_{n=1}^{\infty} \lambda(n)\, q^n$$

with $f'|T(n)=\lambda(n) f'$.

Proof. This follows directly from (3.5) and the theory of newforms (see, e.g., [12, Th. 3]) since there are no non-zero forms of weight 2 on $\mathbf{SL}(2, \mathbf{Z})$.

We restate what we have concluded from the hypothesis $p|B_k$:

(3.7) **Theorem.** *There exists a cusp form $f=\sum_{n\geqq 1} a_n q^n$ of weight 2 and some type ε which is a normalized ($a_1=1$) eigenform for all Hecke operators $T(n)$ and which satisfies*

$$a_l \equiv 1+l^{k-1} \equiv 1+\varepsilon(l)\, l \bmod \mathfrak{p}$$

for all primes $l \neq p$, where $\mathfrak{p}$ is a certain prime ideal over p in the field K generated by the coefficients of f, which does not depend on l.

Note that we may view ε as a (non-trivial) character with values in K^*, since formulas for the Hecke operators show that the values of ε lie in the field generated by the coefficients of f.

§ 4. Construction and Study of the (mod p) Representation

We retain the notations f, $\mathfrak{p}$, K of (3.7). In addition, we let $\mathcal{O}$ be the integer ring of K, $\mathcal{O}_{\mathfrak{p}}$ its completion at $\mathfrak{p}$, $K_{\mathfrak{p}}$ the completion of K at $\mathfrak{p}$, $\mathbf{F}$ the residue field of $\mathcal{O}_{\mathfrak{p}}$, $\pi \in \mathcal{O}_{\mathfrak{p}}$ a uniformizing parameter.

We let $A/\mathbf{Q}$ be the abelian variety attached to f by Shimura's construction [18, Th. 7.14]. We recall the following properties of A:

(i) The dimension of A is equal to the integer $[K:\mathbf{Q}]$, and K is included as a subring of the $\mathbf{Q}$-algebra $(\mathrm{End}_{\mathbf{Q}} A) \otimes \mathbf{Q}$ of endomorphisms of A defined over $\mathbf{Q}$. Thus the $\mathfrak{p}$-adic Tate module

$$V_{\mathfrak{p}} = V_p(A) \underset{K \otimes \mathbf{Q}_p}{\otimes} K_{\mathfrak{p}}$$

is a free $K_{\mathfrak{p}}$-module of rank 2 on which $\mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q})$ acts.

(ii) The variety A is a factor (over $\mathbf{Q}$) of the quotient of the modular variety $J_1(p)$ by the image in $J_1(p)$ of the variety $J_0(p)$. In particular, A has good reduction at all primes $l \neq p$ so that $V_{\mathfrak{p}}$ is unramified at all such primes. Furthermore, by a theorem of Deligne-Rapoport [5, Ex. 3.7 (i), p. DeRa-113], A acquires everywhere good reduction over the real cyclotomic field $\mathbf{Q}(\mu_p)^+$.

(iii) (Eichler-Shimura relation [19, Th. 1.4]). If $F_l \in \mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q})$ is a Frobenius element for a prime $l \neq p$, then the trace (resp., determinant) of its action on the $K_\mathfrak{p}$-vector space $V_\mathfrak{p}$ is a_l (resp., $l \cdot \varepsilon(l)$), regarded as an element of $K_\mathfrak{p}$.

Now we let $\rho\colon \mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q}) \to \mathrm{Aut}_{K_\mathfrak{p}} V_\mathfrak{p}$ be the map arising from the action of $\mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q})$ on $V_\mathfrak{p}$. From (iii) we deduce that the determinant of ρ is the product $\chi\varepsilon$, where we now regard ε as a character of $\mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q})$ and where χ is the standard cyclotomic character

$$\chi\colon \mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q}) \to \mathbf{Z}_p^* \subseteq K_\mathfrak{p}^*.$$

(4.1) **Proposition.** *The $K_\mathfrak{p}$ representation ρ is irreducible.*

Proof. Suppose otherwise. Then the semi-simplification of ρ, which is abelian, is described by two characters $\rho_1, \rho_2\colon \mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q}) \to K_\mathfrak{p}^*$. It is locally algebraic by [17, p. III-20] (or else because it comes from an abelian variety), so that each ρ_i may be written as an integral power χ^{n_i} of χ on an open subgroup of an inertia group for p in $\mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q})$. This implies that $\rho_i = \chi^{n_i}\varepsilon_i$, where ε_i is a character of finite order ramified only at p. Regarding the ε_i as Dirichlet characters, we have (for $l \neq p$) the equations

$$l^{n_1+n_2}\varepsilon_1(l)\,\varepsilon_2(l) = l\varepsilon(l),$$

$$a_l = \varepsilon_1(l)\, l^{n_1} + \varepsilon_2(l)\, l^{n_2}$$

because of (iii). From the first equation we get $n_1 + n_2 = 1$, so that one of the n_i, say n_1, is at least 1. Therefore $n_2 \leqq 0$. Looking at the second equation, we now see that $|a_l| \geqq l - 1$ for all $l \neq p$. When $l \geqq 7$, however, this contradicts the "Riemann hypothesis" $|a_l| \leqq 2\sqrt{l}$.

From now on, we use χ to denote the character "$\chi \bmod p$," namely the composition

$$\mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q}) \xrightarrow{\chi} \mathbf{Z}_p^* \to \mathbf{F}_p^* \hookrightarrow \mathbf{F}^*.$$ [1]

(4.2) **Proposition.** *There exists an $\mathcal{O}_\mathfrak{p}$-lattice $L \subset V_\mathfrak{p}$ invariant by* $\mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q})$ *for which the action of* $\mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q})$ *on $L/\pi L$ may be described matricially by*

$$\begin{pmatrix} 1 & * \\ 0 & \chi^{k-1} \end{pmatrix}$$

and is furthermore not semi-simple.

Proof. In view of (4.1) and (2.1) it suffices to show that there exists a lattice $T \subset V_\mathfrak{p}$ stable by $\mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q})$ for which the action of $\mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q})$ on $T/\pi T$ is reducible in such a way that its semi-simplification is given by the two characters 1 and χ^{k-1}. In fact, let T be any $\mathcal{O}_\mathfrak{p}$-lattice stable by $\mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q})$. By the Eichler-Shimura relation, if $l \neq p$ then a Frobenius element for l acts on $T/\pi T$ with trace $a_l \pmod{\pi}$ and determinant $l\varepsilon(l) \pmod{\pi}$. Because of (3.7) these numbers are respectively congruent to $l^{k-1}+1$ and $l^{k-1} \pmod{\pi}$. By the Čebotarev Density Theorem, the trace and determinant of the action of $\mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q})$ on $T/\pi T$ are respectively $1+\chi^{k-1}$ and χ^{k-1}.

---

[1] Thus we return to the notation used in the Introduction

According to the Brauer-Nesbitt Theorem [4, Th. 30.16], this implies the desired assertion about $T/\pi T$.

Let us set $M=L/\pi L$. This will be the representation space for the $\bar{\rho}$ of (1.3). In fact, property (ii) of this § together with (4.2) shows that the first three conditions of (1.3) are satisfied by the representation. It remains only to verify the fourth condition.

We consider the subgroup $\mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q}(\mu_p)^+)$ of $\mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q})$ corresponding to the real cyclotomic field $\mathbf{Q}(\mu_p)^+$. In this subgroup we consider a decomposition group D for the unique prime of $\mathbf{Q}(\mu_p)^+$ lying over p. Since $p \nmid [\mathbf{Q}(\mu_p)^+ : \mathbf{Q}]$, to verify the last condition of (1.3) it suffices to prove that the action of D on M is semi-simple, i.e. that the image of D in $\mathrm{Aut}\, M$ has order prime to p. It will be convenient to let E be the completion of the real cyclotomic field at p and to identify D with $\mathrm{Gal}(\bar{E}/E)$.

(4.3) **Proposition.** *The* $\mathrm{Gal}(\bar{E}/E)$*-module* M *is the Galois module attached to a finite flat commutative group scheme of type* $(p, \ldots, p)$ *over the integer ring* $\mathscr{R}$ *of* E.

Proof. After changing A by a $\mathbf{Q}$-isogeny we may assume that $\mathcal{O}$ operates on A and that M is isomorphic to the "kernel of $\mathfrak{p}$" on A. This makes M isomorphic to a submodule of the module of p-division points of A. By the Deligne-Rapoport theorem mentioned above, A acquires good reduction over E. Hence the module of p-division points has the property asserted of M: it is the Galois module attached to the scheme-theoretic kernel $\mathscr{A}_p$ of the map "multiplication by p" on the Neron model for A over $\mathscr{R}$. Then M for its part is the Galois module attached to the Zariski closure $\mathscr{M}$ of M in $\mathscr{A}_p$, cf. [15, § 2].

Before completing the proof that M is semi-simple as a D-module, we summarize the properties of M that we will use:

(a) It is free of rank 2 over $\mathbf{F}$,

(b) D acts trivially on a 1-dimensional subspace X of M and via the character $\chi\,(=\chi^{k-1})$ on the quotient $Y=M/X$.

(c) M is the module attached to a finite flat group scheme $\mathscr{M}$ of type $(p, \ldots, p)$ over $\mathscr{R}$.

(4.4) **Theorem.** *The image of* D *in* $\mathrm{Aut}\, M$ *has prime-to-p order.*

Proof. Let $\mathscr{X}$ be the Zariski closure of X in $\mathscr{M}$. The D-module attached to $\mathscr{X}$ is the trivial module X, and the absolute ramification index of E is $(p-1)/2 < p-1$. Hence $\mathscr{X}$ is a non-zero *constant* group scheme over $\mathscr{R}$ by the classification theorem of Raynaud [15, Th. (3.3.3)]. Hence $\mathscr{M}$ cannot be connected, since it has the étale subgroup $\mathscr{X}$.

Take the canonical exact sequence of D-modules

$$0 \to M^0 \to M \to M^{\mathrm{et}} \to 0,$$

where M^0 is associated with the largest connected subgroup of $\mathscr{M}$ and M^{et} with the largest étale quotient. Because M has a Galois-compatible $\mathbf{F}$-vector space structure, $\mathscr{M}$ is a "group scheme in $\mathbf{F}$-vector spaces" by the theorem of Raynaud mentioned above. In particular, the above exact sequence is a sequence of $\mathbf{F}$-vector spaces.

Now M^0 is not all of M because $\mathcal{M}$ is not connected. And $M^0 \neq 0$ because M^{et} is unramified but M is not (since it has the quotient Y). Thus M^0 is 1-dimensional. Further the fact that M^{et} is unramified and Y isn't shows that the image of M^0 in M is distinct from X. Hence D leaves stable both X and a line in M which is distinct from X. Since any element of order p in $\operatorname{Aut} M$ leaves stable a *unique* line, this proves what is wanted.

References

1. Borevich, Z.I., Shafarevich, I.R.: Number theory. New York: Academic Press 1966
2. Carlitz, L.: A generalization of Maillet's determinant and a bound for the first factor of the class number. Proc. A.M.S. **12**, 256–261 (1961)
3. Carlitz, L., Olson, F.R.: Maillet's determinant. Proc. A.M.S. **6**, 265–269 (1955)
4. Curtis, C., Reiner, I.: Representation theory of finite groups and associative algebras. New York: Interscience 1962
5. Deligne, P., Rapoport, M.: Les schémas de modules de courbes elliptiques. International Summer School on Modular Functions; Antwerp, 1972. Lecture Notes in Math. **349**, pp. 143–316. Berlin-Heidelberg-New York: Springer 1973
6. Deligne, P., Serre, J-P.: Formes modulaires de poids 1. Ann. Scient. Ec. Norm. Sup., 4ᵉ série, **7**, 507–530 (1974)
7. Greenberg, R.: A generalization of Kummer's criterion. Inventiones math. **21**, 247–254 (1973)
8. Herbrand, J.: Sur les classes des corps circulaires. J. Math. Pures et Appliquées, 9ᵉ série **11**, 417–441 (1932)
9. Koike, M.: On the congruences between Eisenstein series and cusp forms. US-Japan Number Theory Seminar; Ann Arbor, 1975. Photo-offset notes.
10. Koike, M.: Congruences between cusp forms of weight one and weight two and a remark on a theorem of Deligne and Serre. International Symposium on Algebraic Number Theory; Kyoto, 1976
11. Leopoldt, H.-W.: Eine Verallgemeinerung der Bernoullischen Zahlen. Abh. Math. Sem. Hamburg **22**, 131–140 (1958)
12. Li, W.-C.: Newforms and functional equations. Math. Ann. **212**, 285–315 (1975)
13. Masley, J.M., Montgomery, H.L.: Cyclotomic fields with unique factorization. Preprint.
14. Mazur, B.: Modular curves and the Eisenstein ideal. In preparation.
15. Raynaud, M.: Schémas en groupes de type $(p, \ldots, p)$. Bull. Soc. Math. France **102**, 241–280 (1974)
16. Serre, J-P.: Une interpretation des congruences relatives à la fonction τ de Ramanujan. Sém. Delange-Pisot-Poitou 1967–68, ex. 14
17. Serre, J-P.: Abelian l-adic representations and elliptic curves. New York: Benjamin 1968
18. Shimura, G.: Introduction to the arithmetic theory of automorphic functions. Publ. Math. Soc. Japan, nº 11, Tokyo-Princeton 1971
19. Shimura, G.: Class fields over real quadratic fields and Hecke operators. Ann. of Math. **95**, 130–190 (1972)
20. Tate, J.: Global class field theory. In: Algebraic number theory. Washington: Thompson 1967
21. Yamauchi, M.: On the fields generated by certain points of finite order on Shimura's elliptic curves. J. Math. Kyoto Univ. **14**, 243–255 (1974)

Received April 13, 1976

Kenneth A. Ribet
Fine Hall
Princeton, N.J. 08540
USA