Special Cohomology Classes for Modular Galois Representations

Benjamin Howard

Abstract. Building on ideas of Vatsal [23], Cornut [5] proved a conjecture of Mazur asserting the generic nonvanishing of Heegner points on an elliptic curve $E/\mathbb{Q}$ as one ascends the anticyclotomic $\mathbb{Z}_p$-extension of a quadratic imaginary extension $K/\mathbb{Q}$. In the present article Cornut’s result is extended by replacing the elliptic curve $E$ with the Galois cohomology of Deligne’s 2-dimensional $\ell$-adic representation attached to a modular form of weight $2k > 2$, and replacing the family of Heegner points with an analogous family of special cohomology classes.

0. Introduction

0.1. Statement of the main result. Let $f \in S_{2k}(\Gamma_0(N), \mathbb{C})$ be a normalized newform of weight $2k > 2$ and level $N \geq 4$. Fix a rational prime $\ell$ and embeddings of algebraic closures $\mathbb{Q}_\text{al} \hookrightarrow \mathbb{Q}_\ell$, $\mathbb{Q}_\text{al} \hookrightarrow \mathbb{C}$. Let $\Phi \subset \mathbb{Q}_\text{al}$ be a finite extension of $\mathbb{Q}_\ell$ containing all Fourier coefficients of $f$ and let $W_f$ be the 2-dimensional $\Phi$ vector space with Gal($\mathbb{Q}_\text{al}/\mathbb{Q}$)-action constructed by Deligne [6], so that the geometric Frobenius of a prime $q \nmid \ell N$ acts on $W_f$ with characteristic polynomial $X^2 - a_q(f) X + q^{2k-1}$. Let $K$ be a quadratic imaginary field satisfying the Heegner hypothesis that all prime divisors of $N$ are split in $K$, fix a prime $p \nmid N \cdot \text{disc}(K)$, and define $G = \text{Gal}(\mathbb{Q}_\text{al}/\mathbb{Q})$, so that $G_0 \subset G$ satisfies $G/G_0 \cong \mathbb{Z}_p$. In §5.1 we define for every $s \geq 0$ a subspace

$$\text{Heeg}_s(f) \subset H^1(H[p^s], W_f(k)).$$

This subspace is the higher weight analogue of the subspace generated by the Kummer images of Heegner points in the case $k = 1$, in which case

$$W_f(1) \cong \text{Hom}(A_f) \otimes \mathbb{Q}_\ell$$

for $A_f$ the modular abelian variety attached to $f$ by Eichler-Shimura theory. Such higher weight Heegner objects have been studied earlier by Brylinski [3], Nekovář [15, 16], and Zhang [25], and our construction of $\text{Heeg}_s(f)$ follows Nekovář’s [16] construction very closely. The main result (Theorem 5.1.1) extends the results of Cornut [5] and Vatsal [23] from the case $k = 1$, and is as follows:

Theorem A. Fix a character $\chi : G_0 \to \Phi^\times$ and let

$$\pi_\chi = \sum_{\sigma \in G_0} \chi(\sigma) \sigma \in \Phi[G_0].$$

Suppose $\ell \nmid p \cdot N \cdot \varphi(N) \cdot \text{disc}(K) \cdot (2k - 2)!$ ($\varphi$ is Euler’s function). As $s \to \infty$ the $\Phi$-dimension of $\pi_\chi \text{Heeg}_s(f)$ grows without bound.

This research was supported by an NSF postdoctoral fellowship.
Let \( X(N)/\mathbb{Q} \) be the usual (geometrically disconnected) moduli space of generalized elliptic curves over \( \mathbb{Q} \) with full level \( N \) structure, and let \( \mathcal{V}/\mathbb{Q} \rightarrow X(N)/\mathbb{Q} \) be the Kuga-Sato variety considered in \([20, 21]\). Thus \( \mathcal{V}/\mathbb{Q} \) is a desingularization of the \((2k - 2)\)-fold fiber product over \( X(N)/\mathbb{Q} \) of the universal generalized elliptic curve.

By work of Scholl \([21]\), Deligne’s \( \ell \)-adic representation \( W_f \) occurs as a summand of \( H^{2k-1}(\mathcal{V}/\mathbb{Q}^\text{et},\mathbb{Q}_\ell) \). Combining this with the \( \ell \)-adic Abel-Jacobi map of \([17]\) yields a map \([10, \S0.3]\)
\[
\Psi_f : \text{CH}^2(\mathcal{V}/F) \rightarrow H^{2k}(\mathcal{V}/F,\mathbb{Q}_\ell(k)) \rightarrow H^1(F,W_f(k))
\]
for any number field \( F \), where \( \text{CH}^2 \) denotes the Chow group of homologically trivial cycles of codimension \( k \), modulo rational equivalence. Nekovář \([16]\) shows that the image of \( \Psi_f \) is contained in the Bloch-Kato Selmer group
\[
\text{Sel}(F,W_f(k)) \subset H^1(F,W_f(k)).
\]
Taking \( F = H[p^s] \), the subspace \( \text{Heeg}_{n}(f) \) lies in the image of \( \Psi_f \).

As in \([23]\) we may write \( H[p^\infty] \) as the compositum of linearly disjoint (over \( K \)) fields \( F \) and \( K_\infty \) where \( F/K \) is tamely ramified at \( p \) with Galois group \( G_0 \), and \( K_\infty/K \) is the anticyclotomic \( \mathbb{Z}_p \)-extension. By Theorem \([A]\) (and under the hypotheses of that theorem), the dimension of the \( \chi \)-component of \( \text{Sel}(\mathbb{Q}^\text{ab}/H[p^s],W_f(k)) \) grows without bound. This provides some evidence for the standard conjecture predicting that for each character \( \chi \) of \( G_0 \)
\[
\dim_{\text{fil}} \pi_\chi \text{Sel}(H[p^s],W_f(k)) = \text{ord}_{s=k} \prod_{\psi} L(f \otimes K,\chi^{-1}\psi,s)
\]
where the product is over all characters \( \psi \) of \( \text{Gal}(K_\infty/K) \) of conductor \( \leq p^s \) and \( L(f \otimes K,\chi^{-1}\psi,s) \) is the twisted \( L \)-function defined as in \([10, \S0.5]\). Indeed, the Heegner hypothesis and the functional equation force \( L(f \otimes K,\chi^{-1}\psi,k) = 0 \) for each such \( \psi \), and so the right hand is \( \geq p^s \). One might hope to extend Kolyvagin’s theory of Euler systems so as to prove that the left hand side is \( p^s + O(1) \). Work of Nekovář \([15]\) and of Bertolini and Darmon \([2]\) give evidence that this is accessible.

It is conjectured that the kernel of \( \Psi_f \) is independent of the choice of prime \( \ell \). A proof would allow one to remove the undesirable hypothesis that \( \ell \neq p \) in Theorem \([A]\) leading to higher weight generalizations of the Iwasawa theoretic results of \([11, 18]\). It seems difficult to adapt the methods of the present article to treat the (most interesting) case \( \ell = p \); instead that case is treated in the forthcoming work \([9]\) using a completely different construction of Heegner cohomology classes in \( H^1(H[p^s],W_f(k)) \). The constructions and results of \([9]\) hold only for \( \ell = p \) and \( f \) ordinary at \( p \), but allow modular forms of odd weight (which seem inaccessible using the methods of the present article).

Zhang \([25]\) has proved a higher weight form of the Gross-Zagier theorem relating the height pairings of certain Heegner cycles in \( \text{CH}^k_0(\mathcal{V}/H[1]) \) to the derivatives \( L'(f \otimes K,\chi^{-1}\psi,k) \) for characters \( \psi \) of trivial conductor. The images of these Heegner cycles under \( \Psi_f \) generate our \( \text{Heeg}_0(f) \), and thus Theorem \([A]\) would yield nonvanishing results for \( L'(f \otimes K,\chi^{-1}\psi,k) \) if Zhang’s formula were extended to ramified characters, and (harder) if one knew the nondegeneracy of the height pairing on the Chow group \( \text{CH}^k_0(\mathcal{V}) \).

0.2. Notation and conventions. Throughout this article we use \( k, N, \) and \( M \) to denote positive integers with \( k > 1, N \geq 4, M \) squarefree, and \( (M,N) = 1 \). We
will be ultimately be concerned with the case $M = 1$, but must allow more general $M$ for technical reasons ($M$ will eventually be a divisor of $\text{disc}(K)$). We frequently abbreviate $N = NM$. The letters $\ell$ and $p$ denote rational primes with $(\ell p, N) = 1$

We allow $\ell = p$ unless stated otherwise (more precisely, we allow $\ell = p$ except in Sections 3 and 5).

The letter $\Lambda$ always denotes a $\mathbb{Z}_\ell$-algebra. If $S$ is a scheme on which $N$ is invertible we let $Y_0(N)/S$ (resp. $Y_1(N, M)/S$) be the coarse (resp. fine) moduli space of elliptic curves with $\Gamma_0(N)$ level structure (resp. $\Gamma_1(N, M) = \Gamma_1(N) \cap \Gamma_0(M)$ level structure). If $M = 1$ we omit it from the notation, and we sometimes omit $S$ if it is clear from the context. For a congruence subgroup $\Gamma \subset \text{SL}_2(\mathbb{Z})$ we will sometimes refer to a pair $(E, x)$ consisting of an elliptic curve $E$ together with a $\Gamma$ level structure $x$ on $E$ simply as a $\Gamma$ structure. Set $\Delta = (\mathbb{Z}/NM\mathbb{Z})^\times$ and let $\varphi$ be Euler’s function.

1. Augmented elliptic curves

Throughout $\mathfrak{F}$ $\Lambda = \mathbb{Z}/m\mathbb{Z}$ for some fixed $\ell$-power $m$, and $S = \text{Spec}(F)$ for $F$ a perfect field of characteristic prime to $\ell$.

1.1. Sheaves. If $L/F$ is an algebraic extension, let $\pi^{\text{univ}} : E^{\text{univ}} \to Y_1(N)/L$ be the universal elliptic curve, and define a locally constant constructible sheaf on $Y_1(N)/L$

$$\mathcal{L}_\Lambda = \text{Sym}^{2k-2}(R^1\pi^a_{univ}|_\Lambda).$$

The formation of this sheaf is compatible with base change in $L$, by the proper base change theorem. There are isomorphisms of étale sheaves on $Y_1(N)/L$

$$R^1\pi^a_{univ}|_\mu \cong \text{Hom}(E^{\text{univ}}[m], \mu_m) \cong E^{\text{univ}}[m]$$

where $E^{\text{univ}}[m]$ is the étale sheaf on $Y_1(N)/L$ associated to the group scheme $E^{\text{univ}}[m]$ and $\text{Hom}$ is sheaf Hom. Taking symmetric powers, there is a canonical isomorphism

$$\mathcal{L}_\Lambda(2k-2) \cong \text{Sym}^{2k-2}(E^{\text{univ}}[m]).$$

If we let $Y/L$ be a connected component of the open modular curve parameterizing elliptic curves over $L$ with $\Gamma_1(N) \cap \Gamma(m)$ level structure and fix a geometric point $\tilde{z} \to Y_1(N)/L$, then the forgetful covering map $Y/L \to Y_1(N)/L$ cuts out a quotient of the fundamental group $\pi_1 = \pi_1(Y_1(N)/L, \tilde{z})$. The group of $Y/L$-valued $m$-torsion points of the universal elliptic curve over $Y/L$ is canonically isomorphic to $\Lambda^2$ (via the universal $\Gamma(m)$ level structure), and the action of $\pi_1$ on this group identifies the aforementioned quotient with a subgroup of $\text{GL}_2(\Lambda)$ containing $\text{SL}_2(\Lambda)$. We thus obtain an action of $\pi_1$ on $\Lambda^2$ and so also on $\text{Sym}^{2k-2}\Lambda^2$. It is immediate from $\mathfrak{F}$ that the locally constant sheaf associated to this action is isomorphic to $\mathcal{L}_\Lambda(2k-2)$. From the discussion following $[7, \S 2\text{ Lemma 2}]$ we see that there is a perfect symmetric pairing of étale sheaves

$$\mathcal{L}_\Lambda(k-1) \otimes \mathcal{L}_\Lambda(k-1) \to \Delta.$$

For any étale sheaf $\mathcal{F}$ on $Y_1(N)/L$, define

$$\hat{H}^n(Y_1(N)/L, \mathcal{F}) = \text{Image}(H^n_{\text{c}}(Y_1(N)/L, \mathcal{F}) \to H^n(Y_1(N)/L, \mathcal{F})).$$
1.2. **Augmentations.** Let $L/F$ be an algebraic extension and let $\Gamma$ be any one of $\Gamma_0(N)$, $\Gamma_0(N)$, $\Gamma_1(N)$, or $\Gamma_1(N,M)$. If $E$ is an elliptic curve over $L$, define a $\text{Gal}(F^{\text{al}}/L)$-module

$$A_{\Lambda}(E) = (\text{Sym}^{2k-2} E[m])(1 - k)$$

(where $E[m] = E(F^{\text{al}})[m]$) and set $A_{\Lambda}^2(E) = A_{\Lambda}(E)^{\text{Gal}(F^{\text{al}}/L)}$. Note that $A_{\Lambda}$ and $A_{\Lambda}^2$ are naturally covariant functors on the category of elliptic curves over $L$. The construction of $A_{\Lambda}(E)$ depends on the embedding $L \hookrightarrow F^{\text{al}}$, but that of $A_{\Lambda}^2(E)$ does not, in the sense that the $\Lambda$-modules defined by two different choices are canonically isomorphic.

**Definition 1.2.1.** By a *\(\Lambda\)-augmented $\Gamma$ structure* over an algebraic extension $L/F$ we mean a triple $(E, x, \Theta)$ in which $(E, x)$ is an elliptic curve with $\Gamma$ structure over $L$ and $\Theta \in A_{\Lambda}^2(E)$.

Two $\Lambda$-augmented $\Gamma$ structures over $L$, $(E_0, x_0, \Theta_0)$ and $(E_1, x_1, \Theta_1)$, are isomorphic if there is an isomorphism (over $L$) of elliptic curves $\phi : E_0 \rightarrow E_1$ such that $\phi$ identifies $x_0$ with $x_1$ and $\phi(\Theta_0) = \Theta_1$. If $(E, x, \Theta)$ is a $\Lambda$-augmented $\Gamma$ structure over $F^{\text{al}}$ and $\sigma$ is an automorphism of $F^{\text{al}}$, there is an evident notion of the *conjugate $\Lambda$-augmented $\Gamma$ structure* $(E, x, \Theta)^\sigma = (E^{\sigma}, x^{\sigma}, \Theta^{\sigma})$.

**Definition 1.2.2.** Given a $\Lambda$-augmented $\Gamma$ structure $(E, x, \Theta)$ over $F^{\text{al}}$ the *field of moduli*, $L$, of $(E, x, \Theta)$ is the extension of $F$ characterized by the property that $\sigma \in \text{Gal}(F^{\text{al}}/F)$ fixes $L$ if and only if $(E, x, \Theta)^\sigma$ is isomorphic (over $F^{\text{al}}$) to $(E, x, \Theta)$.

**Remark 1.2.3.** We will often use $(E, C, \Theta)$ to denote a $\Lambda$-augmented $\Gamma_0(N)$ structure, and write $C \subset \mathbb{C}$ for the $\Gamma_0(N)$ structure obtained by forgetting the $\Gamma_0(1)$ structure.

If $z \in Y_1(N, M)/L$ is a closed point we let $E_z^{\text{univ}}$ be the pullback of the universal elliptic curve over $Y_1(N, M)/L$ to $k(z)$. Define the module of *$\Lambda$-augmented cycles on $Y_1(N, M)/L$*

$$A_{\Lambda}^2(Y_1(N, M)/L) = \bigoplus_z A_{\Lambda}^2(E_z^{\text{univ}}),$$

where the sum is over all closed points of $Y_1(N, M)/L$ (note that $A_{\Lambda}^2(E_z^{\text{univ}})$ means the points of $A_{\Lambda}(E_z^{\text{univ}})$ defined over the field of definition of $E_z^{\text{univ}}$, $k(z)$, not over $L$). For any set of closed points $Z \subset Y_1(N, M)/L$ define

$$A_{\Lambda}^2(Z; Y_1(N, M)/L)$$

in the same way, but with the sum restricted to $z \in Z$. We also define

$$A_{\Lambda}(\Gamma_1(N, M)) = \bigoplus_{(E,x)} A_{\Lambda}(E),$$

where the sum is over isomorphism classes of $\Gamma_1(N, M)$ structures over $F^{\text{al}}$. A $\Lambda$-augmented $\Gamma_0(N)$-structure $(E, x, \Theta)$ over $F^{\text{al}}$ defines an element of the modular $\mathbb{A}$, denoted the same way, by taking the element $\Theta$ in the summand attached to $(E, x)$ and 0 in the other summands. The module $A_{\Lambda}(\Gamma_1(N, M))$ has a natural action of $\text{Gal}(F^{\text{al}}/L)$, and

$$A_{\Lambda}^2(Y_1(N, M)/L) \cong A_{\Lambda}(\Gamma_1(N, M))^{\text{Gal}(F^{\text{al}}/L)}.$$

Indeed, a closed point $z \in Y_1(N, M)/L$ and a $\Theta \in A_{\Lambda}^2(E_z^{\text{univ}})$ determine a $\Lambda$-augmented $\Gamma_1(N, M)$ structure $(E_z^{\text{univ}}, x_z^{\text{univ}}, \Theta)$ over $k(z)$, where $(E_z^{\text{univ}}, x_z^{\text{univ}})$ is
the pullback to $k(z)$ of the universal $\Gamma_1(N,M)$ structure. Each embedding of $L$-algebras $k(z) \hookrightarrow F^\text{al}$ then determines a $\Lambda$-augmented $\Gamma_1(N,M)$ structure over $F^\text{al}$, and summing over all embeddings $k(z) \hookrightarrow F^\text{al}$ determines an element of the right hand side of (7). Extending linearly over all $z$ and $\Theta$ gives the desired map. The construction of the inverse is similar and easy.

1.3. A higher weight Kummer map. In this subsection $M = 1$. Let $L \subset F^\text{al}$ be an algebraic extension of $F$. Fix a closed point $z \in Y_1(N)/L$ and write $i_z$ for the closed immersion $\text{Spec}(k(z)) \to Y_1(N)/L$. Denote by $j : Y_1(N)/F^\text{al} \hookrightarrow X_1(N)/F^\text{al}$ the usual compactification.

**Lemma 1.3.1.** There are canonical isomorphisms

$$A_\Lambda^2(E^\text{univ}_z) \cong H^0(z,i_z^*\mathcal{L}_\Lambda(k-1)) \cong H^2_\Lambda(Y_1(N)/L,\mathcal{L}_\Lambda(k)).$$

**Proof.** The first isomorphism is induced by the isomorphism (3), and the second is a consequence of cohomological purity as in [13, Chapter VI §5].

**Lemma 1.3.2.** There is a canonical isomorphism

$$H^2(X_1(N)/L,j_*\mathcal{L}_\Lambda(k)) \cong H^1(F^\text{al}/L,\tilde{H}^1(Y_1(N)/F^\text{al},\mathcal{L}_\Lambda(k))).$$

**Proof.** One checks directly that $\text{Sym}^{2k-2}\Lambda^2$ has no $\text{SL}_2(\Lambda)$-invariants, and hence, by the discussion of [14], $H^0(Y_1(N)/F^\text{al},\mathcal{L}_\Lambda) = 0$. Using Poincaré duality we see also that the group

$$H^2(X_1(N)/F^\text{al},j_*\mathcal{L}) \cong H^2_\Lambda(Y_1(N)/F^\text{al},\mathcal{L})$$

is trivial, and so

$$H^i(X_1(N)/F^\text{al},j_*\mathcal{L}_\Lambda) = 0$$

for $i \neq 1$. Thus the Hochschild-Serre spectral sequence and the identification

$$H^1(X_1(N)/F^\text{al},j_*\mathcal{L}_\Lambda) \cong \tilde{H}^1(Y_1(N)/F^\text{al},\mathcal{L}_\Lambda)$$

yield the desired isomorphism. □

**Definition 1.3.3.** Combining Lemmas 1.3.1 and 1.3.2 with the homomorphism

$$H^2_\Lambda(Y_1(N)/L,\mathcal{L}_\Lambda(k)) \to H^2(X_1(N)/L,j_*\mathcal{L}_\Lambda(k))$$

we obtain a map

$$A_\Lambda^2(E^\text{univ}_z) \to H^1(F^\text{al}/L,\tilde{H}^1(Y_1(N)/F^\text{al},\mathcal{L}_\Lambda(k))).$$

for each closed point $z \in Y_1(N)/L$. This map extends linearly to define the $\Lambda$-augmented Kummer map

$$A_\Lambda^2(Y_1(N)/L) \to H^1(F^\text{al}/L,\tilde{H}^1(Y_1(N)/F^\text{al},\mathcal{L}_\Lambda(k))).$$

We now give an alternate definition of the $\Lambda$-augmented Kummer map. The proof of the equivalence of the two definitions requires only minor modification of [11, Lemma 9.4] and is omitted. Given a closed point $z \in Y_1(N)/L$, let $U = U/F^\text{al}$ be the open complement of $z \times_F F^\text{al}$ in $X_1(N)/F^\text{al}$. Excision and the relative cohomology sequence give the exact sequence

$$0 \to H^1(X_1(N)/F^\text{al},j_*\mathcal{L}_\Lambda) \to H^1(U,j_*\mathcal{L}_\Lambda) \to H^2_{z \times_F F^\text{al}}(Y_1(N)/F^\text{al},\mathcal{L}_\Lambda) \to 0$$
where the initial and terminating zeros are justified by cohomological purity and [8], respectively. Using Lemma [1.3.3] we may identify $\mathcal{A}_\Lambda(E^\text{univ}_w)$ with the $\text{Gal}(\text{F}^{\text{al}}/L)$-invariants of

$$
\bigoplus_{w \in \pi \times L^{\text{F}^{\text{al}}}} \mathcal{A}_\Lambda(E^\text{univ}_w) \cong H^2_{\pi \times L^{\text{F}^{\text{al}}}}(Y_1(N)/\text{F}^{\text{al}}, \mathcal{L}_\Lambda)(k),
$$

and the connecting homomorphism

$$
(11) \quad \mathcal{A}_\Lambda^0(E^\text{univ}_z) \to H^1(\text{F}^{\text{al}}/L, H^1(X_1(N)/\text{F}^{\text{al}}, j_* \mathcal{L}_\Lambda)(k))
$$

then agrees with Definition [1.3.3] using the identification of [9].

For $L/F$ any algebraic extension, the group $\Delta$ acts on $Y_1(N,M)/L$ through the diamond automorphisms. There is a similar action of $\Delta$ on $\mathcal{A}_\Lambda(\Gamma_1(N,M))$ commuting with the $\text{Gal}(\text{F}^{\text{al}}/L)$-action, and so $\Delta$ also acts on $\mathcal{A}_\Lambda^0(Y_1(N,M)/L)$ by [7], and on $\mathcal{A}_\Lambda^0(Z; Y_1(N,M)/L)$ for any subset $Z \subset Y_1(N,M)/L$ stable under $\Delta$. There is also a familiar action of $\Delta$ on the cohomology $H^i(Y_1(N)/L, \mathcal{L}_\Lambda(j))$ for any $i$ and $j$, on compactly supported cohomology, and on the cohomology supported on $Z$ for any closed set $Z \subset Y_1(N)/L$ stable under $\Delta$. The action of $\Delta$ is compatible with the $\Lambda$-augmented Kummer map of Definition [1.3.3].

### 1.4. Augmented $\Gamma_0(N)$ structures

Now fix a $\Lambda$-augmented $\Gamma_0(N)$ structure $(E, C, \Theta)$ over $\text{F}^{\text{al}}$ and suppose $L$ is a finite extension of $F$ containing the field of moduli of $(E, C, \Theta)$. In particular $L$ contains the field of moduli (in the usual sense) of the pair $(E, C)$, and so determines a closed point $y \in Y_0(N)/L$, with residue field $L$. Let $Z \subset Y_1(N,M)/L$ denote the set of closed points lying above $y$ under the forgetful degeneracy map $F_{N,M} : Y_1(N,M)/L \to Y_0(N)/L$.

Let $P_1, \ldots, P_{\varphi(N)}$ be the generators of $C$ (using the convention of Remark 1.2.3) and let $x_i$ be the $\Gamma_1(N,M)$ structure on $E$ determined by $P_i$ and the $\Gamma_0(M)$ structure underlying $C$. Define, using [7],

$$
(12) \quad F_{N,M}^*(E, C, \Theta) = \sum_{i=1}^{\varphi(N)} (E, x_i, \Theta) \in \mathcal{A}_\Lambda^0(Z; Y_1(N,M)/L)^\Delta.
$$

Taking $M = 1$ for the moment, we denote by

$$
(13) \quad \Omega_L(E, C, \Theta) \in H^1(\text{F}^{\text{al}}/L, H^1(Y_1(N)/\text{F}^{\text{al}}, j_* \mathcal{L}_\Lambda)(k))^\Delta
$$

the image of $F_{N,1}^*(E, C, \Theta)$ under the $\Lambda$-augmented Kummer map of Definition [1.3.3]. Allowing $L$ to vary over all finite extensions of $F$ containing the field of moduli of $(E, C, \Theta)$, the formation of $\Omega_L(E, C, \Theta)$ is compatible with the restriction maps on Galois cohomology.

### 1.5. Reduction and ramification

In this subsection we assume that $M = 1$ and $F$ is a finite extension of $\mathbb{Q}_q$ for some prime $q \nmid \ell N$. Let $\mathbb{F}^{\text{al}}$ and $\mathbb{F}$ denote the residue fields of $\text{F}^{\text{al}}$ and $F$, respectively, so that

$$
(14) \quad W_\Lambda \overset{\text{def}}{=} H^1(Y_1(N)/\text{F}^{\text{al}}, \mathcal{L}_\Lambda) \cong H^1(Y_1(N)/\mathbb{F}^{\text{al}}, \mathcal{L}_\Lambda)
$$

is an unramified $\text{Gal}(\text{F}^{\text{al}}/F)$-module. Let $(E, C, \Theta)$ be a $\Lambda$-augmented $\Gamma_0(N)$ structure over $\text{F}^{\text{al}}$, and assume that $E$ has good reduction. The reduction of $(E, C)$, a $\Gamma_0(N)$-structure over the field $\mathbb{F}^{\text{al}}$, is denoted $(\text{red}(E), \text{red}(C))$, and we identify...
$E(F_{\text{al}})[m]$ with $\text{red}(E) (F_{\text{al}})[m]$ as $\Lambda$-modules. This determines an isomorphism $\text{red} : A_\Lambda(E) \cong A_\Lambda(\text{red}(E))$, and so we obtain a $\Lambda$-augmented $\Gamma_0(N)$ structure

$$\text{red}(E, C, \Theta) = (\text{red}(E), \text{red}(C), \text{red}(\Theta))$$

over $F_{\text{al}}$. If $L \subseteq F_{\text{al}}$ is a finite extension of $F$ containing the field of moduli of $(E, C, \Theta)$ then the residue field of $L$, $\mathbb{L}$, contains the field of moduli of $\text{red}(E, C, \Theta)$, and so we may form

$$\Omega_L(\text{red}(E), \text{red}(C), \text{red}(\Theta)) \in H^1(F_{\text{al}}/L, W_A(k))^\Delta.$$

**Proposition 1.5.1.** Suppose $\ell \nmid \varphi(N)$. In the notation above, $\Omega_L(E, C, \Theta)$ is equal to the image of $\Omega_L(\text{red}(E), \text{red}(C), \text{red}(\Theta))$ under the inflation map

$$H^1(F_{\text{al}}/L, W_A(k)) \cong H^1(L_{\text{unr}}/L, W_A(k)) \rightarrow H^1(F_{\text{al}}/L, W_A(k)),$$

where $L_{\text{unr}} \subseteq F_{\text{al}}$ is the maximal unramified extension of $L$.

**Proof.** It is clear from the definition that the construction

$$(E, C, \Theta) \mapsto F^*_N(E, C, \Theta)$$

is compatible with reduction, so the proof of the proposition amounts to verifying that the constructions of $[11,3]$ extend across integral models. Let $Z \subseteq Y_1(N)/_{\text{al}}$ be as in $[14]$ and suppose for the moment that every $z \in Z$ has residue field $L$. Denoting the integer ring of $L$ by $\mathcal{O}_L$, each $z \in Y_1(N)/_{\text{al}}$ extends to a smooth section $\tilde{z} : \text{Spec}(\mathcal{O}_L) \rightarrow X_1(N)/_{\text{al}}$ of the canonical integral model of $X_1(N)$ over $\mathcal{O}_L$. Since $E^\text{univ}_z$ has potentially good reduction, this section does not meet the cusps in the special fiber, and so factors through the affine subscheme $Y_1(N)/_{\text{al}}$. The sequence $[10]$ extends across integral models over the integer ring of the maximal unramified extension of $L$, denoted $R$, to give the middle row of the commutative diagram

$$
\begin{array}{cccccc}
0 & \rightarrow & H^1(X_{/\text{al}}, j_* \mathcal{L}_A) & \rightarrow & H^1(U_{/\text{al}}, j_* \mathcal{L}_A) & \rightarrow & H^2_{\tilde{z} \times \text{al}}(Y_{/\text{al}}, \mathcal{L}_A) & \rightarrow & 0 \\
 & \uparrow & & \uparrow & & \uparrow & \downarrow & & \\
0 & \rightarrow & H^1(X_{/R}, j_* \mathcal{L}_A) & \rightarrow & H^1(U_{/R}, j_* \mathcal{L}_A) & \rightarrow & H^2_{\tilde{z} \times R}(Y_{/R}, \mathcal{L}_A) & \rightarrow & 0 \\
 & \downarrow & & \downarrow & & \downarrow & & & \\
0 & \rightarrow & H^1(X_{/\text{al}}, j_* \mathcal{L}_A) & \rightarrow & H^1(U_{/\text{al}}, j_* \mathcal{L}_A) & \rightarrow & H^2_{\tilde{z} \times \text{al}}(Y_{/\text{al}}, \mathcal{L}_A) & \rightarrow & 0 
\end{array}
$$

where we abbreviate $X = X_1(N)$ and $Y = Y_1(N)$, and write $U_{/R}$ for the complement of $\tilde{z} \times R$ in $X_{/R}$. Lemma $[13,1]$ implies that the rightmost vertical arrows are isomorphisms (the pullback of $\mathcal{L}_A$ to $\tilde{z} \times R$ is constant, so its global sections can be computed in either geometric fiber). The vertical arrows on the left are isomorphisms, by $[9]$ and the isomorphism of $[13]$. By the five lemma the arrows in the middle column are isomorphisms as well, and from this it follows that the map $[11]$ is compatible with reduction.

For the general case, choose a $z \in Z$ and an embedding $k(z) \hookrightarrow F_{\text{al}}$. Let $L'$ be the image of this embedding. It is easily seen that $L'$ does not depend on the point $z$ or the choice of embedding, that $L'/L$ is a Galois extension of degree dividing $\varphi(N)$, and that every point in $Z \times L L' \hookrightarrow Y_1(N)/_{L'}$ has residue field $L'$. The proposition follows from the bijectivity of the restriction map

$$H^1(F_{\text{al}}/L, W_A(k)) \cong H^1(F_{\text{al}}/L', W_A(k))^\text{Gal}(L'/L),$$
and of the analogous map on the level of residue fields, together with the special case considered above.

1.6. **Degeneracy maps.** Recall that $M$ is squarefree. Given a divisor $M' | M$ we define a degeneracy map

$$\alpha_{M'}^M : A_\Lambda(\Gamma_1(N, M)) \to A_\Lambda(\Gamma_1(N, M'))$$

as follows. Given a $\Lambda$-augmented $\Gamma_1(N, M)$ structure $(E, x, \Theta)$ over $F^{al}$, we define

$$\alpha_{M'}^M(E, x, \Theta) = (E, x', \Theta)$$

where $x'$ is the $\Gamma_1(N, M')$ structure on $E$ underlying $x$. Extend this $\Lambda$-linearly to a map on $\mathcal{A}_\Lambda(\Gamma_1(N, M))$. We also define a degeneracy map

$$\beta_{M'}^M : A_\Lambda(\Gamma_1(N, M)) \to A_\Lambda(\Gamma_1(N, M'))$$

as follows. Given a $\Lambda$-augmented $\Gamma_1(N, M)$ structure $(E, x, \Theta)$ over $F^{al}$ let $P$ and $D$ be the $\Gamma_1(N)$ and $\Gamma_0(M)$ structures underlying $x$. Let $D_0 \subset D$ be the subgroup of order $M/M'$, let $E' = E/D_0$, and let $P'$ and $D'$ be the images of $P$ and $D$ under $E \to E'$. Let $\Theta'$ be the image of $\Theta$ under $A_\Lambda(E) \to A_\Lambda(E')$. Write $x'$ for the $\Gamma_1(N, M')$ structure $(P', D')$ on $E'$. Now define

$$\beta_{M'}^M(E, x, \Theta) = (E', x', \Theta')$$

and again extend linearly. The maps $\alpha_{M'}^M$ and $\beta_{M'}^M$ respect the $\text{Gal}(F^{al}/F)$ action, and so induce maps

$$\alpha_{M', \beta_{M'}}^M : \mathcal{A}_\Lambda'(Y_1(N, M)/L) \to \mathcal{A}_\Lambda'(Y_1(N, M')/L)$$

for any algebraic extension $L/F$.

2. **Families of augmented CM points**

Let $\Lambda = \mathbb{Z}/m\mathbb{Z}$ for an $\ell$-power $m$. Fix a quadratic imaginary field $K \subset \mathbb{Q}^{al}$, assume that all prime divisors of $N$ are split in $K$, and fix an ideal $\mathfrak{m} \subset \mathcal{O}_K$ such that $\mathcal{O}_K/\mathfrak{m} \cong \mathbb{Z}/NZ$. Fix an elliptic curve $E_1$ over $\mathbb{Q}^{al}$ with complex multiplication by the maximal order $\mathcal{O}_K$ (there are #Pic($\mathcal{O}_K$) such curves). Let $j : \mathcal{O}_K \hookrightarrow \text{End}_{\mathbb{Q}^{al}}(E_1)$ be normalized so that pullback by $j(\alpha)$ acts as multiplication by $\alpha$ on the cotangent space of $E_1(\mathbb{C})$ for every $\alpha \in K$. Set $C_1 = E_1[\mathfrak{m}]$, a cyclic subgroup of order $N$.

**Definition 2.0.1.** An element $c \in \text{GL}_2(\mathbb{Q}_p)$ is cyclic if $(c^{-1} \mathbb{Z}_p^2)$ contains $\mathbb{Z}_p^2$ with cyclic quotient. The *degree* of a cyclic $c$ is

$$\deg(c) = [c^{-1} \mathbb{Z}_p^2 : \mathbb{Z}_p^2] = p^{\text{ord}_p(\det(c))}.$$ 

2.1. **A parametrized family of Heegner points.** A choice of isomorphism of $\mathbb{Z}_p$-modules $T_{x_1}(E_1) \cong \mathbb{Z}_p^2$ (which we now fix) determines a family of elliptic curves over $\mathbb{Q}^{al}$ parametrized by

$$\mathcal{T} = \mathbb{Q}_p^\times \text{GL}_2(\mathbb{Z}_p) \backslash \text{GL}_2(\mathbb{Q}_p)$$

as follows. For each cyclic subgroup $X \subset E_1[p^\infty] \cong (\mathbb{Q}_p/\mathbb{Z}_p)^2$ there is a cyclic $c_X \in \text{GL}_2(\mathbb{Q}_p)$ such that $X = (c_X^{-1} \mathbb{Z}_p^2)/\mathbb{Z}_p^2$. The assignment $X \mapsto c_X$ establishes a bijection between the set of such subgroups and the cyclic elements of $\text{GL}_2(\mathbb{Q}_p)$, modulo left multiplication by $\text{GL}_2(\mathbb{Z}_p)$. We denote the inverse by $c \mapsto X_c$. The projection map $\text{GL}_2(\mathbb{Z}_p) \backslash \text{GL}_2(\mathbb{Q}_p) \to \mathcal{T}$ establishes a bijection between the left $\text{GL}_2(\mathbb{Z}_p)$-orbits of cyclic elements and the set $\mathcal{T}$. To each $g \in \mathcal{T}$ we then define
Let $g$ be a cyclic $p$-power isogeny $f_g : E_1 \to E_g = E_1/X_g$ of degree $\deg(c)$. Define the degree of $g$ by $\deg(g) = \deg(c) = \deg(f_g)$. The elliptic curve $E_g$ over $\mathbb{Q}^{\text{al}}$ inherits a $\Gamma_0(N)$ structure $C_g = f_g(C_1) = E_g[\mathfrak{M} \cap \mathcal{O}_g]$, where $\mathcal{O}_g \subseteq \mathcal{O}_K$ is the largest order which leaves the subgroup $X_g \subset E_1[p^\infty]$ stable. The conductor of $\mathcal{O}_g$ is a power of $p$.

For each $g \in \mathcal{T}$ let $H_g$ be the ring class field of $\mathcal{O}_g$, thus $H_g/K$ is Galois with Galois group canonically identified with $\text{Pic}(\mathcal{O}_g)$. The Weil pairing on $E_1[m]$ provides a canonical (up to $\pm 1$) isomorphism between $(\text{Sym}^2 E_1[m])(-1)$ and the traceless $\Lambda$-module endomorphisms of $E_1[m]$. In particular there is a canonical (up to sign) element $\vartheta_1 \in (\text{Sym}^2 E_1[m])(-1)$ corresponding to $\sqrt{T} \in \text{End}_{\mathbb{Q}^{\text{al}}}(E_1)$. Let $\Theta_1 \in \mathcal{A}_\Lambda(E_1)$ be the image of $\vartheta_1^{k-1}$ under the natural projection

\[
\text{Sym}^{k-1}(\text{Sym}^2 E_1[m])(-1)) \to (\text{Sym}^{2k-2} E_1[m])(1-k),
\]

and define $\Theta_g = f_g(\Theta_1) \in \mathcal{A}_\Lambda(E_g)$. The data $K$, $E_1$, $\mathfrak{M}$, and $\text{Ta}_p(E_1) \cong \mathbb{Z}_2^2$ thus determine a family $g \mapsto (E_g, C_g, \Theta_g)$ of $\Lambda$-augmented $\Gamma_0(N)$ structures over $\mathbb{Q}^{\text{al}}$ parametrized by $\mathcal{T}$. This data is to remain fixed throughout the remainder of the article. Define a $\text{Gal}(\mathbb{Q}^{\text{al}}/K)$-module

\[
W_\Lambda = \check{H}^1(Y_1(N)/\mathbb{Q}^{\text{al}}, \mathcal{L}_\Lambda).
\]

Using the theory of complex multiplication it is easily seen that the field of moduli of $(E_g, C_g, \Theta_g)$ is $H_g$, and so the construction yields a family of cohomology classes parametrized by $g \in \mathcal{T}$

\[
g \mapsto \Omega_{H_g}(E_g, C_g, \Theta_g) \in H^1(\mathbb{Q}^{\text{al}}/H_g, W_\Lambda(k))^\Delta.
\]

Let $H[p^s]$ denote the ring class field of conductor $p^s$ of $K$. For each $s \geq 0$ define $\mathcal{T}_s \subset \mathcal{T}$ to be the subset consisting of all $g$ such that $H_g \subset H[p^s]$. For $g \in \mathcal{T}_s$ we let

\[
\Omega_s(g) = \Omega_{H[p^s]}(E_g, C_g, \Theta_g) \in H^1(\mathbb{Q}^{\text{al}}/H[p^s], W_\Lambda(k))^\Delta
\]

be the restriction of the cohomology class of $(16)$ to $\text{Gal}(\mathbb{Q}^{\text{al}}/H[p^s])$.

### 2.2. Level $M$ structure.

Suppose that the integer $M$ of $(12)$ is a divisor of $\text{disc}(K)$ and let $\mathfrak{M}$ be the unique $\mathcal{O}_K$-ideal of norm $M$. Although we assumed in $(2.1)$ that all prime divisors of $N$ are split in $K$, the constructions of $(2.1)$ work equally well with $N$ replaced by $\mathbf{N} = NM$ and $\mathfrak{M}$ replaced by $\mathfrak{M} \mathfrak{M}$. This has the effect of endowing each $(E_g, C_g)$ with the extra $\Gamma_0(N)$ structure $C_g = E_g[\mathfrak{M} \mathfrak{M} \cap C_g]$, so that $g \mapsto (E_g, C_g, \Theta_g)$ is a parametrized family of $\Lambda$-augmented $\Gamma_0(N)$ structures.

### 3. Reduction of the family.

In this section we prove Theorem $3.4.3$, which is our analogue of $[5$, Theorem 3.1$]$. Suppose $\ell \neq p$ and take $M$ to be a divisor of $\text{disc}(K)$ as in $(2.2)$. Assume $\ell \nmid \varphi(N)$ and $\ell > 2k - 2$. Let $\Lambda$ be a finite quotient of $\mathbb{Z}_\ell$. Let $\mathcal{Q}$ be a finite set of rational primes inert in $K$, all prime to $\ell p \mathfrak{N}$. For each $q \in \mathcal{Q}$, let $q$ denote the prime of $K$ above $q$ and fix an extension of $q$ to a place of $\mathbb{Q}^{\text{al}}$. We will abusively use $\mathcal{Q}$ to refer to the set of rational primes, the set of primes of $K$ above them, and also the set of chosen places of $\mathbb{Q}^{\text{al}}$. Let $\mathbb{F}_q$ and $\mathbb{F}_{\ell}^q$ denote the residue fields of $\mathbb{Q}^{\text{al}}$.
and $K$ at $q \in \Omega$, respectively, so that $\mathbb{F}_q$ has $q^2$ elements and $\mathbb{F}_q^{al}$ is algebraically closed. For each $q \in \Omega$ let

$$Z_0(N)q \subset Y_0(N)/\mathbb{F}_q \quad Z_1(N,M)q \subset Y_1(N,M)/\mathbb{F}_q$$

denote the subsets of supersingular points. The points of $Z_0(N)_q$ all have residue field $\mathbb{F}_q$, but this need not be true of $Z_1(N,M)_q$.

**Definition 3.0.1.** A subset $S \subset \text{Gal}(\mathbb{Q}^{al}/K)$ is chaotic if for any distinct $\sigma, \tau \in S$, the restriction of $\sigma \tau^{-1}$ to $\text{Gal}(H[p^\infty]/K)$ is not the Artin symbol of any idele with trivial $p$-component.

### 3.1 Simultaneous reduction

As $E_1$ has complex multiplication, any model of $E_1$ over a number field has everywhere potentially good reduction. Fix a finite Galois extension $F_1/K$ over which $E_1$ has a model with good reduction at every prime above every rational prime $q \in \Omega$, and fix such a model. All endomorphisms of $E_1$ are defined over $F_1$, and hence so is the subgroup $C_1 = E_1[9\mathbb{M}]$. For each $g \in T$ let $F_g$ be a finite extension of $F_1$, Galois over $K$, over which the subgroup $X_g$ is defined. We may then view $E_g, C_g,$ and the isogeny $f_g$ all as being defined over $F_g$. Fixing these choices, we may reduce everything at $w_q$ to obtain a family of $\Lambda$-augmented $\Gamma_0(N)$ structures over $\mathbb{F}_q^{al}$

$$\text{red}_q(E_g, C_g, \Theta_g) = (\text{red}_q(E_g), \text{red}_q(C_g), \text{red}_q(\Theta_g)).$$

We also denote by $\text{red}_q(f_g)$ the reduction of the isogeny $f_g$. Given an element $\sigma \in \text{Gal}(\mathbb{Q}^{al}/K)$ we may also form

$$\text{red}_q(E_g^\sigma, C_g^\sigma, \Theta_g^\sigma) = (\text{red}_q(E_g^\sigma), \text{red}_q(C_g^\sigma), \text{red}_q(\Theta_g^\sigma))$$

and

$$\text{red}_q(f_g^\sigma) : \text{red}_q(E_g^\sigma) \to \text{red}_q(E_g^\sigma).$$

To emphasize, we regard these as $\Lambda$-augmented $\Gamma_0(N)$ structures over $\mathbb{F}_q^{al}$, regardless of the residue field of $F_g$ at $q$. The field of moduli of $E_q$ is $\mathbb{F}_q$; a fact (Lemma 4.1.1) whose proof we postpone until the next section. Abbreviate

$$Z_q(M) = A_1^\Delta(Z_1(N,M)q; Y_1(N,M)/\mathbb{F}_q)^\Delta.$$  

Let $\Lambda[T]$ denote the free $\Lambda$-module on the set $T$. For each $\sigma \in \text{Gal}(\mathbb{Q}^{al}/K)$ and each $q \in \Omega$ define the reduction map $\text{Red}_{\sigma,q} : \Lambda[T] \to Z_q(M)$ by taking $L = \mathbb{F}_q$ in (12) and linearly extending

$$\text{Red}_{\sigma,q}(g) = F_{N,M}^\sigma(\text{red}_q(E_g^\sigma), \text{red}_q(C_g^\sigma), \text{red}_q(\Theta_g^\sigma)).$$

For any subset $S \subset \text{Gal}(\mathbb{Q}^{al}/K)$ define the simultaneous reduction map

$$\text{Red}_{S,\Omega} : \Lambda[T] \to \bigoplus_{(\sigma,q) \in S \times \Omega} Z_q(M)$$

by linearly extending $\text{Red}_{S,\Omega}(g) = \oplus_{\sigma,q} \text{Red}_{\sigma,q}(g)$. The reader may wish to skip directly to Theorem 3.4.3, the main result of 33.
3.2. Reduction at \( q \). Fix a \( q \in \mathbb{Q} \). Define \( S = \text{End}_{\mathbb{Q}(\mathbb{Z})}(\text{red}_q(E_1)) \) and \( B = S \otimes \mathbb{Q} \) so that \( B \) is a quaternion algebra ramified exactly at \( q \) and \( \infty \), and \( S \) is a maximal order in \( B \). Let \( R \subset S \) be the subring of endomorphisms which leave \( \text{red}_q(C_1) \) stable, so that \( R \) is a level \( N \)-Eichler order in \( B \). The embedding \( j : \mathcal{O}_K \to \text{End}_{F_1}(E_1) \) determines an embedding which we again denote by \( j \)

\[
j : K \cong \text{End}_{F_1}(E_1) \otimes \mathbb{Q} \to \text{End}_q(\text{red}_q(E_1)) \otimes \mathbb{Q} \cong B,
\]

with \( j(\mathcal{O}_K) \subset R \). For any rational prime \( r \) and any \( \mathbb{Z} \) (resp. \( \mathbb{Q} \)) algebra \( A \), set \( \Delta_r = A \otimes \mathbb{Z} \mathbb{Q}_r \) (resp. \( \Delta_r = A \otimes \mathbb{Q} \mathbb{Q}_r \)). Let \( \hat{B} \) be the restricted topological product \( \prod_r B_r \) with respect to the local orders \( R_r \subset B_r \), and define \( \hat{K}, \hat{R}, \ldots \) similarly. The embedding \( j \) induces embeddings \( \hat{K} \to \hat{B} \) and \( \hat{R} \to \hat{B} \) at every \( r \). We denote all of these again by \( j \).

Recall that we have fixed an isomorphism of \( \mathbb{Z}_p \)-modules \( \mathbb{T}_p(E_1) \cong \mathbb{Q}_p^2 \). As the \( p \)-adic Tate modules of \( E_1 \) and \( \text{red}_q(E_1) \) are canonically identified as \( \mathbb{Z}_p \)-modules (and \( R_p \subset B_p \) is a maximal order), this induces isomorphisms

\[
(21) \quad R_p \cong M_2(\mathbb{Z}_p) \quad B_p \cong M_2(\mathbb{Q}_p).
\]

We henceforth identify \( R_p^\times \cong \text{GL}_2(\mathbb{Z}_p) \) and \( B_p^\times \cong \text{GL}_2(\mathbb{Q}_p) \) using these isomorphisms, and in particular identify \( \mathbb{T} \) with \( \mathbb{Q}_p^r \mathbb{R}_p \). This gives a right action of \( B_p^\times \) (and hence also of \( B^\times \)) on \( \mathbb{T} \). The group \( R_p^\times \) acts on \( \mathbb{T}_p(\text{red}_q(E_1)) \) on the left, almost by definition, and we denote by \( \rho_q \) the action of \( R_p^\times \) on \( A_\Lambda(\text{red}_q(E_1)) \) obtained by taking symmetric powers, with the understanding that \( R_p^\times \) acts trivially on the twist \( \Lambda(1-k) \). Writing \( \text{det} \) for the reduced norm on \( B^\times \), \( B_p^\times \), and so on, we also define \( \rho_p^\times = \rho_q \circ \text{det}^1 \), and note that the center \( \mathbb{Z}_p^\times \subset B_p^\times \) acts trivially under \( \rho_p^\times \). The group \( \Gamma_q = R[1/p]^\times \) acts on \( A_\Lambda(\text{red}_q(E_1)) \) through \( \rho_q \) or \( \rho_p^\times \) by the inclusion \( \Gamma_q \hookrightarrow R_p^\times \).

Fix a \( \sigma \in \text{Gal}(\mathbb{Q}_K/K) \) whose restriction to \( K^\text{ab} \) (the maximal abelian extension of \( K \)) is equal to the Artin symbol of a finite idele \( \hat{\sigma} \in \hat{K}^\times \). Let \( b_{\sigma,q} \in B^\times \) be such that \( \hat{\sigma} b_{\sigma,q} \) lies in \( R_p^\times \) for all primes \( r \neq p \), and let \( \alpha_{\sigma,q} \in R_p^\times \) and \( \beta_{\sigma,q} \in B_p^\times \) be the \( \ell \) and \( p \) components, respectively, of \( \hat{\sigma} b_{\sigma,q} \in \hat{B}^\times \).

**Proposition 3.2.1.** Fix \( g, h \in \mathcal{T} \). There is a \( \gamma \in \Gamma_q \) such that \( g \beta_{\sigma,q} h = h \gamma \in \mathcal{T} \), if and only if there is an isomorphism of \( \Gamma_0(N) \) structures over \( \mathbb{F}_q \)

\[
\phi : \text{red}_q(E_g^\sigma, C_g^\sigma) \cong \text{red}_q(E_h, C_h).
\]

If these equivalent conditions hold then \( \phi \) may be chosen so that

\[
\phi(\text{red}_q(\Theta_g^\sigma)) = \omega_{\text{cycl}}^{-1}(\sigma) \cdot \text{red}_q(f_h)(\rho_q(\gamma \alpha_{\sigma,q}^{-1})\text{red}_q(\Theta_1)),
\]

where \( \omega_{\text{cycl}} \) is the \( \ell \)-adic cyclotomic character and \( \gamma \in \Gamma_q \) has the property that there are cyclic (in the sense of Definition 2.4.1) lifts \( c(g), c(h) \) of \( g \) and \( h \) satisfying \( c(g) \beta_{\sigma,q} = c(h) \gamma \) in \( R_p^\times \backslash B_p^\times \).

**Proof.** For any \( g \in \mathcal{T} \) there is an isomorphism of \( \Gamma_0(N) \) structures over \( \mathbb{F}_q^\text{al} \)

\[
(22) \quad \text{red}_q(E_g^\sigma, C_g^\sigma) \cong \text{red}_q(E_g \beta_{\sigma,q}, C_g \beta_{\sigma,q}).
\]

This is exactly the calculation performed in [5, \S 3.3]. On the other hand, by the parametrization of \( Z_0(N)_{\mathbb{Q}} \) given in [5, \S 2.3] there is an isomorphism

\[
(23) \quad \text{red}_q(E_g \beta_{\sigma,q}, C_g \beta_{\sigma,q}) \cong \text{red}_q(E_h, C_h)
\]
if and only if \(g \beta_{\sigma,q}\) and \(h\) lie in the same orbit under the right action of \(\Gamma_q\) on \(\mathcal{T}\). This proves the first claim. The proof of the second claim follows from an examination of the isomorphisms (22) and (23), and we give a sketch. The isomorphisms (22) and (23), disregarding the \(\Gamma_0(N)\) structure, arise from isomorphisms (again, see [5 §3.3])

\[
\text{red}_q(E_g^\sigma) \cong \text{Hom}_R(R \cdot c(g)j(\hat{\sigma}), \text{red}_q(E_1))
\]

(24)

\[
\text{red}_q(E_h) \cong \text{Hom}_R(R \cdot c(h), \text{red}_q(E_1))
\]

(25)

of functors on \(\mathbb{F}_q^a\)-schemes, where \(\text{Hom}_R\) means homomorphisms of left \(R\)-modules, and \(c(g)\) and \(c(h)\) are viewed as elements of \(\hat{\mathcal{B}}\) with trivial components away from \(p\). The map \(x \mapsto x\beta_{\sigma,q}\gamma^{-1}\) induces an isomorphism of left \(R\)-submodules of \(\hat{\mathcal{B}}\)

\[
R \cdot c(g)j(\hat{\sigma}) \xrightarrow{\cdot b_{\sigma,q}} R \cdot c(g)j(\hat{\sigma})b_{\sigma,q} = R \cdot c(g)\beta_{\sigma,q} = R \cdot c(h)\gamma^{-1} \xrightarrow{\cdot b_{\sigma,q}} R \cdot c(h)
\]

and so identifies \(\text{red}_q(E_g^\sigma) \cong \text{red}_q(E_h)\) and

\[
\text{Hom}_{R_\ell}(R_\ell \cdot j(\hat{\sigma})_\ell, \text{Ta}_\ell(\text{red}_q(E_1))) \cong \text{Hom}_{R_\ell}(R_\ell, \text{Ta}_\ell(\text{red}_q(E_1)))
\]

(26)

By the main theorem of complex multiplication, the isomorphism (24) may be chosen so that the induced isomorphism

\[
\text{Ta}_\ell(E_1) \xrightarrow{f_\Sigma} \text{Ta}_\ell(E_g^\sigma) \cong \text{Hom}_{R_\ell}(R_\ell \cdot j(\hat{\sigma})_\ell, \text{Ta}_\ell(\text{red}_q(E_1)))
\]

takes \(t \in \text{Ta}_\ell(E_1) \cong \text{Ta}_\ell(\text{red}_q(E_1))\) to the \(R_\ell\)-linear map determined by \(j(\hat{\sigma})_\ell \mapsto t\). The isomorphism (26) takes \(j(\hat{\sigma})_\ell \mapsto t\) to \(\alpha_{\sigma,q}\gamma^{-1} \mapsto t\). Under (25) this latter map corresponds to \(\text{red}_q(f_h)(\gamma\alpha_{\sigma,q}^{-1}t) \in \text{Ta}_\ell(\text{red}_q(E_h))\). This shows that the composition

\[
\text{Ta}_\ell(E_1) \xrightarrow{f_\Sigma} \text{Ta}_\ell(E_g^\sigma) \cong \text{Ta}_\ell(\text{red}_q(E_g^\sigma)) \cong \text{Ta}_\ell(\text{red}_q(E_h))
\]

is given by \(t \mapsto \text{red}_q(f_h)(\gamma\alpha_{\sigma,q}^{-1}t)\). The proposition now follows by taking symmetric powers and twisting by \(\Lambda(1-k)\).

\[\square\]

**Corollary 3.2.2.** Let \(\sigma\) and \(\beta_{\sigma,q}\) be as in Proposition 3.2.1. For each \(h \in \mathcal{T}\) there is a \(\varpi_{\sigma,q,h} \in A_{\Lambda}(\text{red}_q(E_1))\) with the property that for any \(g \in h\Gamma_q\beta_{\sigma,q}^{-1} \subset \mathcal{T}\) there exists an isomorphism of \(\Lambda\)-augmented \(\Gamma_0(N)\) structures over \(\mathbb{F}_q\)

\[
\text{red}_q(E_g^\sigma, C_g^\sigma, \text{deg}(g)^{1-k} \cdot \Theta_g^\sigma) \cong (\text{red}_q(E_h), \text{red}_q(C_h), \text{red}_q(f_h)(\rho_q^\ast(\gamma)\varpi_{\sigma,q,h}))
\]

(27)

where \(\gamma \in \Gamma_q\) is any element with \(g\beta_{\sigma,q} = h\gamma\) in \(\mathcal{T}\).

**Proof.** Suppose we have an equality \(g = h\gamma\beta_{\sigma,q}^{-1}\) in \(\mathcal{T}\) with \(g, h \in \mathcal{T}\) and \(\gamma \in \Gamma_q\). Fix cyclic lifts \(c(g)\) and \(c(h)\) of \(g\) and \(h\), respectively, to \(\hat{\mathcal{B}}\), and choose \(\gamma_0 \in \gamma \cdot \mathbb{Z}[1/p]\) so that \(c(g)\beta_{\sigma,q} = c(h)\gamma_0\) in \(\hat{\mathcal{B}}\). Using Proposition 3.2.1 and the fact that \(\rho_q^\ast(\gamma_0) = \rho_q^\ast(\gamma)\), one checks directly that (27) holds with

\[
\varpi_{\sigma,q,h} = \left(\text{deg}(g)\omega_{\text{cyc}}(\sigma)\det(\gamma_0^{-1})\det(\alpha_{\sigma,q})\right)^{1-k} \rho_q^\ast(\alpha_{\sigma,q}^{-1})\text{red}_q(\Theta).
\]

As \(\text{deg}(g)\det(\gamma_0)^{-1} = \text{deg}(h)p^{-\text{ord}_p \det(\beta_{\sigma,q})}\) depends on \(h\) but not on \(g\), the same is true of \(\varpi_{\sigma,q,h}\). \[\square\]
3.3. Vatsal’s lemma. Fix a subset $\mathcal{S} \subset \text{Gal}(\mathbb{Q}^{al}/K)$. For each $q \in \Omega$ and each $\sigma \in \mathcal{S}$ let $\beta_{\sigma,q} \in B_p^\times$ be as in Proposition 3.2.1. The quaternion algebra $B$ depends on $q$, but using the isomorphisms of [21] we identify $B_p^\times \cong \text{GL}_2(\mathbb{Q}_p)$ and view both $\beta_{\sigma,q}$ and $\Gamma_q$ as living in $\text{GL}_2(\mathbb{Q}_p)$ under this identification.

Lemma 3.3.1. For each $q \in \Omega$ there is a finite index subgroup $\Gamma_q^\ast \subset \Gamma_q$ containing $\mathbb{Z}[1/p]^{\times}$ such that $\det(\Gamma_q^\ast) = p^2$ and the restriction of $\rho_q^\ast$ to $\Gamma_q^\ast$ is trivial.

Proof. Define a subgroup $U = \prod U_r \subset \hat{B}^{\times}$ by

$$U_r = \begin{cases} \text{Ker}(\rho_q^\ast : R_\ell^\times \to \text{Aut}(A_{\Lambda}(\text{red}_q(E_1)))) & \text{if } r = \ell \\ B_p^{\times} & \text{if } r = p \\ R_r^\times & \text{else} \end{cases}$$

and let $\Gamma_q^\ast = B^{\times} \cap U \subset \hat{B}^{\times}$. Then $\Gamma_q^\ast \subset \Gamma_q$ is exactly the kernel of $\rho_q^\ast$ restricted to $\Gamma_q$. We must show that $\Gamma_q^\ast$ contains an element of norm $p$. By [21, Theoreme III.4.3] there is a $b_0 \in B^{\times}$ of norm $p$. Let $x = (x_r) \in U$ be an element of norm $p \in \hat{Q}^{\times}$. By strong approximation [21, Theoreme III.4.3] the norm one element $b_0^{-1}x \in \hat{B}^{\times}$ may be written in the form $b_1 y u = b_0^{-1}x$ for some norm one elements $b_1 \in B^{\times}$, $y \in B_p^{\times}$, and $u \in U$. Then $b_0 b_1$ has norm $p$ and is contained in $\Gamma_q^\ast$. □

Proposition 3.3.2. For each $q \in \Omega$ let $\Gamma_q^\ast$ be as in Lemma 3.3.1, and for each $(\sigma, q) \in \mathcal{S} \times \Omega$ set

$$\Gamma_{\sigma,q}^\ast = \beta_{\sigma,q} \Gamma_q^\ast \sigma_{\sigma,q} \subset \text{GL}_2(\mathbb{Q}_p).$$

If $\mathcal{S}$ is chaotic then the quotient map $T \to \prod_{(\sigma,q) \in \mathcal{S} \times \Omega} T / \Gamma_{\sigma,q}^\ast$ is surjective.

Proof. Let $\hat{\Gamma}_{\sigma,q}^\ast$ be the image of $\Gamma_{\sigma,q}^\ast$ in $\text{PGL}_2(\mathbb{Q}_p)$ and let $\hat{\Gamma}_{\sigma,q}^{\ast,1}$ the intersection of $\hat{\Gamma}_{\sigma,q}^\ast$ with $\text{PSL}_2(\mathbb{Q}_p)$ Then $\hat{\Gamma}_{\sigma,q}^{\ast,1}$ is discrete and cocompact by [21, p.104], and these subgroups are pairwise non-commensurable as $(\sigma,q)$ varies by [5, Proposition 3.7]. By Vatsal’s application of a theorem of Ratner (see [3, Proposition 3.11] or [23, Lemma 5.10]), the natural map

$$\text{PSL}_2(\mathbb{Q}_p) \to \prod_{(\sigma,q) \in \mathcal{S} \times \Omega} \text{PSL}_2(\mathbb{Z}_p) / \text{PSL}_2(\mathbb{Q}_p) / \hat{\Gamma}_{\sigma,q}^{\ast,1}$$

is surjective, and the proposition follows as in [5, Proposition 3.4]. □

3.4. Surjectivity of the reduction map. Assume $\mathcal{S} \subset \text{Gal}(\mathbb{Q}^{al}/K)$ is finite and chaotic.

Proposition 3.4.1. Fix $(\sigma', q') \in \mathcal{S} \times \Omega$, $\gamma_0, \gamma_1 \in \Gamma_q$, and $h \in T$. There exist $g_0, g_1 \in T$ such that

$$\deg(g_0)^{1-k} \text{Red}_{\mathcal{S}, \Omega}(g_0) - \deg(g_1)^{1-k} \text{Red}_{\mathcal{S}, \Omega}(g_1) \in \bigoplus_{(\sigma,q) \in \mathcal{S} \times \Omega} \mathbb{Z}_q(M)$$

has trivial components except at the summand $(\sigma,q) = (\sigma',q')$, at which the component is equal to

$$F_{N,M}^*(\text{red}_q(E_h), \text{red}_q(C_h), \text{red}_q(f_h)(\rho_q^*(\gamma_0) \omega_{\sigma,q,h} - \rho_q^*(\gamma_1) \omega_{\sigma,q,h})), $$

where $\omega_{\sigma,q,h} \in A_{\Lambda}(\text{red}_q(E_1))$ is the element of Corollary 3.4.2.
Proof. For each $i \in \{0, 1\}$ Proposition 3.3.2 allows us to choose a $g_i \in T$ such that the reduction map $T \to T/\Gamma_{g_i}^\sigma$ takes

$g_i \mapsto \begin{cases} h_i \beta^{-1}_{\sigma,q} \Gamma_{g_i}^\sigma = h_i \Gamma_{q}^\sigma \beta^{-1}_{\sigma,q} & \text{if } (\sigma,q) = (\sigma',q') \\ h_i \beta_{\sigma,q} \Gamma_{g_i}^\sigma = h_i \Gamma_{q}^\sigma \beta_{\sigma,q} & \text{if } (\sigma,q) \neq (\sigma',q') \end{cases}$

for every $(\sigma,q) \in S \times \Omega$. By Corollary 3.3.2, we have

$\text{red}_q(E_{g_0}^\sigma, C_{g_0}^\sigma, \deg(g_0)1^{-k}\Theta_{g_0}^\sigma) \cong \text{red}_q(E_{g_1}^\sigma, C_{g_1}^\sigma, \deg(g_1)1^{-k}\Theta_{g_1}^\sigma)$

as a $\Lambda$-augmented $\Gamma_0(N)$ structure over $\mathbb{F}_q$ whenever $(\sigma,q) \neq (\sigma',q')$, while

$\text{red}_q(E_{g_i}^\sigma, C_{g_i}^\sigma, \deg(g_i)1^{-k}\Theta_{g_i}^\sigma) \cong (\text{red}_q(E_h), \text{red}_q(C_h), \text{red}_q(f_h)(\rho_q^*(\gamma_i)\varpi_{\sigma,q,h}))$

if $(\sigma,q) = (\sigma',q')$. The proposition is now immediate from the definition (20) of $\text{Red}_{S,\Omega}$.

Lemma 3.4.2. Fix $q \in \Omega$ and suppose $\Lambda = \mathbb{Z}/\ell \mathbb{Z}$. Then $\mathcal{A}_\Lambda(\text{red}_q(E_1))$ has no proper, nonzero $\Lambda$-submodules which are stable under $\rho_q^*(\Gamma_q)$.

Proof. Fix a $\mathbb{Z}_\ell$-basis for the $\ell$-adic Tate module of $\text{red}_q(E_1)$, so that $R_\ell^\times$ is identified with $\text{GL}_2(\mathbb{Z}_\ell)$. Let $\Gamma_q^\times$, $R_\ell^\times$, $R_\ell^\times, 1$ denote the norm one elements of $\Gamma_q$ and $R_\ell^\times$, respectively. Then $\mathcal{A}_\Lambda(\text{red}_q(E_1))$ is identified with $\text{Sym}^{2k-2} \Lambda^2$ and the action of $\rho_q^*$ restricted to $\Gamma_q^\times$ is through

$\Gamma_q^\times \to R_\ell^\times, 1 \to \text{SL}_2(\mathbb{Z}_\ell) \to \text{SL}_2(\Lambda).$

Using strong approximation [24, Theoreme III.4.3] one may show that the first arrow has dense image, and so the composition is surjective. By the assumption $\ell > 2k-2$, $\text{Sym}^{2k-2} \Lambda^2$ has no proper, nonzero submodules stable under the action of $\text{SL}_2(\Lambda)$. □

Theorem 3.4.3. Let $S \subset \text{Gal}(\mathbb{Q}_\ell^\text{al}/K)$ be finite and chaotic, and suppose $\Lambda = \mathbb{Z}/\ell \mathbb{Z}$. Then the simultaneous reduction map (20) is surjective.

Proof. Fix $(\sigma',q') \in S \times \Omega$ and a supersingular point $z \in Z_0(\mathbb{N})_{q'}$. Let $Z \subset Z_1(\mathbb{N}, M)_{q'}$ be the set of closed points lying above $z$. We will show that the image of (20) contains the submodule

$(28) \quad \mathcal{A}_\Lambda^\sigma(Z; Y_1(N,M)_{/q'})^{\Lambda} \subset \bigoplus_{(\sigma,q) \in S \times \Omega} \mathcal{A}_\Lambda^\sigma(Z_1(N,M); Y_1(N,M)_{/q'})^{\Lambda}$

supported in the $(\sigma', q')$ component. The parametrization [5, §2.3] shows that the map $T \to Z_0(\mathbb{N})_{q'}$ defined by $h \mapsto \text{red}_{q'}(E_h, C_h)$ establishes a bijection $T/\Gamma_{q'} \cong Z_0(\mathbb{N})_{q'}$. Thus we may fix an $h \in T$ such that the supersingular $\Gamma_0(\mathbb{N})$ structure $\text{red}_{q'}(E_h, C_h)$ corresponds to the point $z$. For any $g \in T$, $\text{red}_{q'}(\Theta_g^{\sigma'}) \neq 0$ (from the construction one sees that $\Theta_1 \neq 0$, and $\ell \neq p$ implies that $f_\beta : \mathcal{A}(E_1) \to \mathcal{A}(E_g)$ is an isomorphism). It follows that $\varpi_{\sigma', q', h} \neq 0$. By Lemma 3.4.2, we may choose a $\gamma_1 \in \Gamma_q^\times$ such that

$\pi \overset{\text{def}}{=} \rho_q^*(\gamma_1)\varpi_{\sigma', q', h} \in \mathcal{A}_\Lambda(\text{red}_{q'}(E_1))$

is nonzero. Again by Lemma 3.4.2 choose $\gamma^{(0)}, \ldots, \gamma^{(n)} \in \Gamma_q^\times$ such that the elements $\rho_q^*(\gamma^{(i)})$, $0 \leq i \leq n$, generate $\mathcal{A}_\Lambda(\text{red}_{q'}(E_1))$. Set $\gamma^{(i)} = \gamma^{(i)}\gamma_1$ and let $g_0^{(i)}, g_1$ be as in Proposition 3.4.1 so that

$\deg(g_0^{(i)})1^{-k}\text{Red}_{S,\Omega}(g_0^{(i)}) - \deg(g_1)1^{-k}\text{Red}_{S,\Omega}(g_1)$
has trivial components except at the summand \((\sigma, q) = (\sigma', q')\), where the component is equal to
\[
F_{N,M}^{\gamma}(\text{red}_q(E_h), \text{red}_q(C_h), \text{red}_q(f_h)(\rho_q^* \gamma(i) \pi)).
\]
As \(i\) varies the elements \(\text{red}_q(f_h)(\rho_q^* \gamma(i) \pi)\) generate \(\mathcal{A}_A(\text{red}_q(E_h))\), and the elements \((29)\) generate the submodule \((28)\). \(\square\)

4. Augmented theorems of Deuring, Ihara, and Ribet

Let \(q \nmid N\) be a rational prime and let \(F = F\) when we refer to the notions of \([1]\) be a field of \(q^2\) elements with algebraic closure \(\overline{F}\). Unless specified otherwise, all geometric objects (e.g. \(Y_1(N), Y_1(N,M), \ldots\)) are defined over \(\text{Spec}(F)\). Let \(L = \mathbb{Z}/\ell\mathbb{Z}\) for some prime \(\ell\) and assume that \(\ell\) does not divide \(Nq\). Let \(L_\Lambda\) be the locally constant constructible sheaf on \(Y_1(N)\) defined by \([2]\). Denote by
\[
Z_1(N) \subset Y_1(N) \quad Z_1(N,M) \subset Y_1(N,M)
\]
the subsets of supersingular closed points.

4.1. Fields of moduli. We need a slight generalization of the well-known theorem of Deuring that all supersingular points on \(Y_0(N)\) have residue degree one.

**Lemma 4.1.1.** Let \(E\) be a supersingular elliptic curve over \(\overline{F}\), let \(C \subset E[N]\) be a cyclic subgroup of order \(N\), and let \(\Theta\) be any element of \(\mathcal{A}_A(E)\). The field of moduli of the \(\Lambda\)-augmented \(\Gamma_0(N)\) structure \((E, C, \Theta)\) is \(F\).

**Proof.** As \(E\) is supersingular, its \(j\)-invariant lies in \(\overline{F}\). Let \(A\) be an elliptic curve over \(F\) with the same \(j\)-invariant as \(E\), and let \(\text{Fr} \in \text{End}_F(A)\) be the degree \(q^2\) (relative) Frobenius. If \(\text{Fr} \in \mathbb{Z}\), then \(\text{Fr}\) commutes with all elements of \(\text{End}_{\overline{F}}(A)\), and so
\[
(30) \quad \text{End}_F(A) = \text{End}_{\overline{F}}(A).
\]
If \(\text{Fr} \notin \mathbb{Z}\) then \(\text{Fr}\) generates a quadratic imaginary subfield \(L\) of the definite quaternion algebra (ramified exactly at \(q\) and \(\infty\)) \(\text{End}_{\overline{F}}(A) \otimes \mathbb{Q}\), and \(q\) is nonsplit in \(L\). As \(\text{Fr}\) has degree \(q^2\) we must have \(\text{Fr} = \zeta^{-1} q\) for some root of unity \(\zeta \in L\), and in fact \(\zeta\) belongs to \(L \cap \text{End}_F(A)\) (this follows from the fact \([12, Corollary 12.3.5]\) that \(\text{Fr}\) and \([q]\) have the same scheme-theoretic kernel, and so there is a factorization \([q] = \zeta \circ \text{Fr}\) for some automorphism \(\zeta\) of \(A\). Replacing \(A\) by its twisted form corresponding to the cocycle sending the relative Frobenius \(\sigma \in \text{Gal}(\overline{F}/F)\) to \(\zeta \in \text{Aut}_F(E)\), a simple calculation shows that \((30)\) holds. Then \(\text{Fr}\) is a central element of \(\text{End}_{\overline{F}}(A)\), and so \(\text{Fr} = [\pm q]\).

With this choice of \(A\), \(\text{Gal}(\overline{F}/F)\) acts trivially on \(\mathcal{A}_A(A)\) and the triple \((A, C_A, \Theta_A)\) is defined over \(F\) for any cyclic order \(N\) subgroup \(C_A \subset A(\overline{F})\) and any \(\Theta_A \in \mathcal{A}_A(A)\). Over \(\overline{F}\) we may fix an isomorphism \(f : E \cong A\) and set \(C_A = f(C)\) and \(\Theta_A = f(\Theta)\). Then \((E, C, \Theta)\) and \((A, C_A, \Theta_A)\) are isomorphic (over \(\overline{F}\)) and so have the same field of moduli \(\overline{F}\). \(\square\)

4.2. Ihara’s theorem. We now recall a theorem of Ihara \([10]\) and derive some consequences; our exposition of Ihara’s theorem is influenced by the discussion of \([4, Chaptire 7]\). For each integer \(m\) prime to \(q\) set
\[
\mu^*_m = \text{Spec}(\overline{F}[X]/\Phi_m(X)) \quad \mu_m = \text{Spec}(\overline{F}[X]/(X^m - 1)),
\]
where $\Phi_m(X)$ is the $m^{th}$ cyclotomic polynomial. Let $Y(m)$ be the affine modular curve classifying “naive” level $m$ structures in the sense of [12] on elliptic curves over $\mathbb{F}$-schemes. Thus $Y(m)$ is a fine moduli space if $m > 2$, and for all $m$ (prime to $q$) the Weil pairing provides a canonical map $Y(m) \rightarrow \mu_m^*$ of $\mathbb{F}$-schemes. Fix a topological generator

$$\zeta = \lim_{(m,q)=1} \zeta_m \in \lim_{(m,q)=1} \mu_m^*(\mathbb{F}^\text{al}).$$

For each $m$ there is a map $\text{Spec}(\mathbb{F}[\zeta_m]) \rightarrow \mu_m^*$ determined by the map $X \mapsto \zeta_m$ on $\mathbb{F}$-algebras. Define

$$Y_\zeta(m) = Y(m) \times_{\mu_m^*} \text{Spec}(\mathbb{F}[\zeta_m]),$$

a smooth curve over $\mathbb{F}$ (geometrically disconnected unless $m \mid q^2 - 1$).

The subgroup $G_\zeta(m) \subset G(m) = \text{GL}_2(\mathbb{Z}/m\mathbb{Z})/\{\pm 1\}$ defined by

$$G_\zeta(m) = \{A \in G(m) \mid \det(A) \in q^{2z} \subset (\mathbb{Z}/m\mathbb{Z})^2\}$$

acts on both $Y_\zeta(m)$ and $\text{Spec}(\mathbb{F}[\zeta_m])$, and the actions are compatible with the structure map $Y_\zeta(m) \rightarrow \text{Spec}(\mathbb{F}[\zeta_m])$. Set $G^1(m) = \text{PSL}(\mathbb{Z}/m\mathbb{Z})$, let

$$\Gamma_0(m) = \left\{ \begin{pmatrix} * & * \\ 0 & * \end{pmatrix} \right\} \subset G(m) \quad \Gamma_1(m) = \left\{ \begin{pmatrix} 1 & * \\ 0 & * \end{pmatrix} \right\} \subset \Gamma_0(m)$$

be the habitual congruence subgroups, and let $\Gamma_{\text{fr}}(m) \subset G(m)$ be the center. For $* \in \{0, 1, \text{fr} \}$ let $Y_* (m)$ be the quotient of $Y_\zeta(m)$ by the action of

$$\Gamma_* (m) \cap G_\zeta(m).$$

The function field of the curve $Y_{\text{fr}}(m)$ is the field denoted $K_m$ in [10], and there is a canonical isomorphism of curves over $\mathbb{F}^\text{al}$.

$$Y_{\text{fr}}(m) \times_{\text{Spec}(\mathbb{F})} \text{Spec}(\mathbb{F}^\text{al}) \cong Y_\zeta(m) \times_{\text{Spec}(\mathbb{F}[\zeta_m])} \text{Spec}(\mathbb{F}^\text{al}).$$

Denote by $K_* (m)$ the function field of $Y_* (m)$ for $* \in \{0, 1, \text{fr} \}$ or for $*$ equal to the empty character, and view these as subfields of some fixed separable closure $K(1)^{\text{sep}}$. Define a $\Lambda$-vector space $L_\Lambda = \text{Sym}^{2k-2}\Lambda^{2}$ and endow $L_\Lambda$ with an action of $\text{Gal}(K(1)^{\text{sep}}/K(1))$ via

$$\text{Gal}(K(1)^{\text{sep}}/K(1)) \rightarrow \text{Gal}(K_\zeta(\ell)/K(1)) \cong G_\zeta(\ell) \subset \text{GL}_2(\Lambda)/\{\pm 1\}.$$

**Remark 4.2.1.** If the action of $\text{GL}_2(\Lambda)$ on $L_\Lambda$ is twisted by det, then the Galois action is twisted by the cyclotomic character. In particular $\Gamma_{\text{fr}}(\ell)$ acts trivially on $L_\Lambda \otimes \det^{1-k}$, and so the Galois action on $L_\Lambda(1-k)$ factors through

$$\text{Gal}(K(1)^{\text{sep}}/K(1)) \rightarrow \text{Gal}(K_{\text{fr}}(\ell)/K(1)) \rightarrow \text{Gal}(K_{\text{fr}}(\ell)/K(1)).$$

Under the bijection between locally constant étale sheaves on $Y_1(N)$ and modules for the absolute Galois group of $K_1(N)$ which are unramified outside of the cusps, $L_\Lambda(2k-2)$ corresponds to $L_\Lambda$.

**Definition 4.2.2.** If $M/K(1)$ is a separable extension, a cusp of $M$ is a place lying above the place $J = \infty$ of $K(1)$. A supersingular prime of $M$ is a place lying above a place $J = j$ of $K(1)$ with $j \in \mathbb{F}$ a supersingular $j$-invariant.

**Theorem 4.2.3.** (Hara) For any $m > 1$ with $(m,q) = 1$, $K_{\text{fr}}(m)$ has no non-trivial everywhere unramified extensions in which all supersingular primes are split completely. Furthermore, $K_{\text{fr}}(\infty) = \bigcup_{(r,q)=1} K_{\text{fr}}(r)$ is the maximal Galois extension of $K_{\text{fr}}(m)$ satisfying
(a) it is tamely ramified, and unramified outside the cusps of $K_{\text{ih}}(m)$,
(b) the supersingular primes of $K_{\text{ih}}(m)$ are split completely in $K_{\text{ih}}(\infty)$.

Proof. This is the main result of [10]. □

Corollary 4.2.4. Let $M_{\text{ih}}(N) \supset K_{\text{ih}}(\infty)$ be the maximal separable extension of $K_{\text{ih}}(N)$ unramified away from the cusps. The restriction map on Galois cohomology

$(31) H^1(M_{\text{ih}}(N)/K_{\text{ih}}(N), L_\Lambda(1-k)) \rightarrow$

$$
\left( \bigoplus_v H^1(K_{\text{ih}}(N)_v, L_\Lambda(1-k)) \right) \bigoplus \left( \bigoplus_w H^1(K_{\text{ih}}(N)^{\text{unr}}_w, L_\Lambda(1-k)) \right)
$$

is injective. Here the sum over $v$ is over all supersingular primes, the sum over $w$ is over all cusps, and the superscript $\text{unr}$ denotes maximal unramified extension.

Proof. First consider the restriction map (note Remark 4.2.1)

$(32) H^1(M_{\text{ih}}(N)/K_{\text{ih}}(N\ell), L_\Lambda(1-k)) \rightarrow \bigoplus_v \text{Hom}(H_v, L_\Lambda(1-k))$

where the sum is over all supersingular primes and all cusps, and $H_v \subset \text{Gal}(M_{\text{ih}}(N)/K_{\text{ih}}(N))$ is either the decomposition group or inertia group of a fixed place of $M_{\text{ih}}(N)$ above $v$, according as $v$ is supersingular or a cusp. Any homomorphism from $\text{Gal}(M_{\text{ih}}(N)/K_{\text{ih}}(N\ell))$ to $L_\Lambda(1-k)$ which vanishes on all $H_w$ factors through $\text{Gal}(\Phi/K_{\text{ih}}(N\ell))$ where $\Phi$ is the maximal Galois extension of $K_{\text{ih}}(N\ell)$ which is everywhere unramified and in which all supersingular primes split completely. By Theorem IX.2.4 $\Phi = K_{\text{ih}}(N\ell)$, and so the map (32) is injective. Thus any class in the kernel of (31) also lies in the kernel of restriction

$$H^1(M_{\text{ih}}(N)/K_{\text{ih}}(N), L_\Lambda(1-k)) \rightarrow H^1(M_{\text{ih}}(N)/K_{\text{ih}}(N\ell), L_\Lambda(1-k)),$$

and so is in the image of the inflation map

$(33) H^1(K_{\text{ih}}(N\ell)/K_{\text{ih}}(N), L_\Lambda(1-k)) \rightarrow H^1(M_{\text{ih}}(N)/K_{\text{ih}}(N), L_\Lambda(1-k))$

and is unramified at the cusps. The inertia subgroup in $\text{Gal}(K_{\text{ih}}(N\ell)/K_{\text{ih}}(N))$ of the cusp $\infty$ is an $\ell$-Sylow subgroup (this follows from [10 p. 167]), and so any element in the image of (33) which is unramified at the cusps is trivial by [22 Theorem IX.2.4]. □

Proposition 4.2.5. Let $j : Y_1(N) \hookrightarrow X_1(N)$ be the usual compactification and assume $\ell \nmid \varphi(N)$. The natural map

$(34) H^2_\zeta(N)(Y_1(N), \mathcal{L}_\Lambda(k))^{\Delta} \rightarrow H^2(X_1(N), j_* \mathcal{L}_\Lambda(k))^{\Delta}$

is surjective.

Proof. Let $i : C_1(N) \hookrightarrow X_1(N)$ denote the subscheme of cusps, i.e. the complement of $Y_1(N)$ in $X_1(N)$. From the exact sequence of sheaves on $X_1(N)$

$$0 \rightarrow j_! \mathcal{L}_\Lambda \rightarrow j_* \mathcal{L}_\Lambda \rightarrow i_* i^* j_* \mathcal{L}_\Lambda \rightarrow 0$$

and [13 Proposition II.2.3] we obtain the exact sequence

$$H^1(C_1(N), i^* j_* \mathcal{L}_\Lambda(k)) \rightarrow H^2_\zeta(N)(Y_1(N), \mathcal{L}_\Lambda(k)) \rightarrow H^2(X_1(N), j_* \mathcal{L}_\Lambda(k)) \rightarrow 0.$$
in which the terminating zero is justified by the observation that closed points on $X_1(N)$, having finite residue field, have cohomological dimension 1. It therefore suffices to prove the surjectivity of

$$H^2_{\text{cr}}(Y_1(N), \mathcal{L}_\Lambda(k))^\Lambda \oplus H^1(C_1(N), i^* j_* \mathcal{L}_\Lambda(k))^\Lambda \to H^2(Z_1(N), \mathcal{L}_\Lambda(k))^\Lambda.$$  

We take $\Lambda$-duals and translate the problem into the language of Galois cohomology. Let $M_1(N)$ denote the maximal extension of $K_1(N)$ unramified outside the cusps. By [14, Corollary II.4.13(c)] there is an isomorphism

$$H^2_{\text{cr}}(Y_1(N), \mathcal{L}_\Lambda(k)) \cong H^1(M_1(N)/K_1(N), L_\Lambda(1-k))^\vee$$

in which the superscript $\vee$ denotes $\Lambda$-dual. If we let $U$ denote the open complement of $Z_1(N)$ in $Y_1(N)$ then the pairing of [14, Corollary II.3.3] identifies the exact sequence [13, Proposition II.2.3(d)]

$$H^r_c(U, \mathcal{L}_\Lambda(k-1)) \to H^r_c(Y_1(N), \mathcal{L}_\Lambda(k-1)) \to \bigoplus_{z \in Z_1(N)} H^r(z, i^*_z \mathcal{L}_\Lambda(k-1))$$

with the dual of the relative cohomology sequence

$$H^{3-r}(U, \mathcal{L}_\Lambda(k)) \leftarrow H^{3-r}(Y_1(N), \mathcal{L}_\Lambda(k)) \leftarrow H^{3-r}_{Z_1(N)}(Y_1(N), \mathcal{L}_\Lambda(k)).$$

This gives the first isomorphism of

$$H^2_{Z_1(N)}(Y_1(N), \mathcal{L}_\Lambda(k))^\vee \cong \bigoplus_{z \in Z_1(N)} H^1(z, i^*_z \mathcal{L}_\Lambda(k-1))$$

$$\cong \bigoplus_v H^1(D_v/I_v, L_\Lambda(1-k)),$$

in which the second sum is over all supersingular primes and $I_v \subset D_v$ are the inertia and decomposition subgroups in $\text{Gal}(M_1(N)/K_1(N))$ of some choice of place above $v$. Finally local duality gives the second isomorphism of

$$H^1(C_1(N), i^* j_* \mathcal{L}_\Lambda(k))^\vee \cong \bigoplus_w H^1(D_w/I_w, L_\Lambda(2-k)^\Lambda)^\vee$$

$$\cong \bigoplus_w H^1(I_w, L_\Lambda(1-k))^D_w/I_w$$

where both sums are over all cusps. Thus the cokernel of (35) is isomorphic to the kernel of

$$H^1(M_1(N)/K_1(N), L_\Lambda(1-k))^\Lambda \to \bigoplus_v H^1(K_1(N)_v, L_\Lambda(1-k)) \bigoplus \bigoplus_w H^1(K_1(N)_{w, \text{unr}}, L_\Lambda(1-k))$$

where again the $v$’s range over supersingular primes and the $w$ range over cusps.

As we assume that $\ell$ is prime to $\varphi(N)$, the inflation-restriction sequence identifies the kernel of (36) with the kernel of

$$H^1(M_1(N)/K_0(N), L_\Lambda(1-k)) \to \bigoplus_v H^1(K_0(N)_v, L_\Lambda(1-k)) \bigoplus \bigoplus_w H^1(K_0(N)_{w, \text{unr}}, L_\Lambda(1-k))$$.
The fields $K_1(N)$ and $K_{th}(N)$ have a common extension which is unramified outside the cusps (namely $K_1(N)$) and so $M_1(N) = M_{th}(N)$. We may therefore consider the restriction map

$$H^1(M_1(N)/K_0(N), L_\Lambda(1 - k)) \to H^1(M_{th}(N)/K_{th}(N), L_\Lambda(1 - k)),$$

which is injective as $K_{th}(N)$ and $K_{th}(\ell)$ are linearly disjoint over $K(1)$, so that $L_\Lambda(1 - k)$ has no $\text{Gal}(K(1)^{\text{sep}}/K_{th}(N))$ invariants. The kernel of (37) therefore injects into the kernel of (31), which is trivial by Corollary 4.2.4. Thus (36) is injective and the proposition is proved. □

The following is our analogue of [5, Proposition 4.4].

**Corollary 4.2.6.** Assume $\ell \nmid \varphi(N)$. The $\Lambda$-augmented Kummer map

$$A^\Lambda_\alpha(Z_1(N); Y_1(N))^\Delta \to H^1(\mathbb{F}^{\text{al}}/\mathbb{F}, \tilde{H}^1(Y_1(N)/\mathbb{F}^{\text{al}}, \mathcal{L}_\Lambda)(k))^\Delta$$

of Definition 1.3.3 is surjective.

**Proof.** Lemma 1.3.1 gives isomorphisms

$$A^\Lambda_\alpha(Z_1(N); Y_1(N)) \cong \bigoplus_{z \in Z_1(N)} H^2_1(Y_1(N), \mathcal{L}_\Lambda(k)) \cong H^2_1(Y_1(N), \mathcal{L}_\Lambda(k))$$

which restrict to isomorphisms of $\Delta$-invariants. The claim is now immediate from Lemma 1.3.2 and Proposition 4.2.5. □

### 4.3. Degeneracy maps on supersingular points

Suppose $M = rM'$ for a prime $r$. The following theorem and its proof are based on work of Ribet [19, Theorem 3.15].

**Proposition 4.3.1.** Assume $\ell \nmid \varphi(N)$ and $\ell > 2k - 2$, and abbreviate

$$\mathcal{Z}(M) = A^\Lambda_\alpha(Z_1(N, M); Y_1(N, M))^\Delta$$

and similarly for $M'$. The sum of the degeneracy maps of (1.10)

$$\alpha^{M'}_M, \beta^{M'}_M : \mathcal{Z}(M) \to \mathcal{Z}(M') \oplus \mathcal{Z}(M')$$

is surjective.

**Proof.** Let $N_\Delta$ be the norm element in the group algebra $\Lambda[\Delta]$. Suppose we are given a $\Lambda$-augmented $\Gamma_1(N, M')$ structure over $\mathbb{F}^{\text{al}}$

$$(E, x, \Theta) \in A_\Lambda(\Gamma_1(N, M'))$$

with $E$ supersingular, and a degree $r^{2n}$ endomorphism $f : E \to E$ preserving the $\Gamma_0(NM')$ structure underlying $x$. Factor the endomorphism $f : E \to E$ as

$$E = E_0 \overset{h_0}{\to} E_1 \overset{h_1}{\to} \cdots \overset{h_{2n-1}}{\to} E_{2n-1} \overset{h_{2n}}{\to} E_{2n} = E$$

with each $h_i$ of degree $r$. Set $f_i = h_i \circ \cdots \circ h_1 : E \to E$, and let $x_i = f_i(x)$ be the induced $\Gamma_1(N, M')$ structure on $E_i$. For $i < 2n$ let $y_i$ be the $\Gamma_1(N, M)$ structure on $E_i$ obtained by adding the $\Gamma_0(r)$ structure $\ker(h_{i+1})$ to $x_i$, and for $i > 0$ let $y_i'$ be the $\Gamma_1(N, M)$ structure obtained by adding the $\Gamma_0(r)$ structure $\ker(h_{i})$. Define

$$\Theta_i = r^{i(1 - k)} f_i(\Theta) \in A_\Lambda(E_i).$$

A simple calculation of the degeneracy maps of (1.10) shows that the element

$$T = T_{E, x, f, \Theta} \in A_\Lambda(\Gamma_1(N, M))$$
defined by
\[
T = N_\Delta \left[ (E_0, y_0, \Theta_0) - (E_2, y_2^\nu, \Theta_2) + (E_2, y_2, \Theta_2) - (E_4, y_4^\nu, \Theta_4) + \ldots + (E_{2n-2}, y_{2n-2}, \Theta_{2n-2}) - (E_{2n}, y_{2n}^\nu, \Theta_{2n}) \right]
\]
satisfies $\beta_M^T(T) = 0$ and
\[
\alpha_M^T(T) = N_\Delta \cdot (E, x, \Theta - r^{2n(1-k)} f(\Theta)).
\]
It follows from Lemma 4.3.2 (with $N$ replaced by $N = NM$) that $T$ is fixed by the action of $\text{Gal}(\mathbb{F}^\text{al}/\mathbb{F})$, and so defines an element of $\mathcal{Z}(M)$.

We pause for a

**Lemma 4.3.2.** With $(E, x)$ as above, let $D$ denote the $\Gamma_0(NM')$ structure underlying the $\Gamma_1(N, M')$ structure $x$. The $\Lambda$-module $\mathcal{A}_\Lambda(E)$ has a set of generators $\mathcal{A}_{E,x}$ such that each $a \in \mathcal{A}_{E,x}$ has the form $a = \Theta_a - \text{deg}(f_a)^{(1-k)} f_a(\Theta_a)$ for some $\Theta_a \in \mathcal{A}_\Lambda(E)$ and some endomorphism $f_a : E \to E$ such that $f_a(D) = D$ and $\text{deg}(f_a)$ is an even power of $r$.

**Proof.** Set $R = \text{End}_{\mathbb{F}^\text{al}}(E, D)$, a level $N$ Eichler order in a quaternion algebra ramified exactly at $q$ and $\infty$, and let $\Gamma = R[1/r]^\times$. Let $\rho$ denote the natural action of $\Gamma$ on $\mathcal{A}_\Lambda(E)$, extend $\rho$ to an action of $\Gamma$ (recall $\ell \nmid M$ so that $r \neq \ell$, and let $\rho^* = \rho \otimes \det^{-1-k}$ be the twist such that $\mathbb{Z}[1/r]^\times \subset \Gamma$ acts trivially. All of this notation is exactly as in [3.2] with $p$ replaced by $r$. As in the proof of Lemma 3.4.2, $\mathcal{A}_\Lambda(E)$ has no submodules stable under the restriction of $\rho^*$ to the subgroup of norm one elements $\Gamma^1 \subset \Gamma$. As the set
\[
\mathcal{A}_{E,x} = \{ \Theta - \rho^* (\gamma) \Theta \mid \Theta \in \mathcal{A}_\Lambda(E), \gamma \in \Gamma^1 \}
\]
is stable under the action of $\rho^* (\Gamma^1)$, it must generate $\mathcal{A}_\Lambda(E)$. For each
\[
\Theta - \rho^* (\gamma) \Theta \in \mathcal{A}_{E,x},
\]
let $f = r^n \gamma$ for $n$ large enough that $r^n \gamma \in R$. Then $f$ has degree $r^{2n}$ and
\[
\Theta - \text{deg}(f)^{1-k} f(\Theta) = \Theta - \rho^* (f) \Theta = \Theta - \rho^* (\gamma) \Theta,
\]
so that $\mathcal{A}_{E,x}$ has the desired properties. \qed

If we let $E$ vary over all supersingular elliptic curves over $\mathbb{F}^\text{al}$, $x$ vary over all $\Gamma_1(N, M)$ structures on $E$, and $\Theta'$ vary over the set $\mathcal{A}_{E,x}$ of Lemma 4.3.2 the elements
\[
N_\Delta \cdot (E, x, \Theta') \in \mathcal{A}_\Lambda(\Gamma_1(N, M'))
\]
generate the submodule $\mathcal{Z}(M')$. Hence, by the construction of $T_{E,x,f,\Theta}$ above, there is a family $\{T_i\} \subset \mathcal{Z}(M)$ such that $\beta_{M'}^T(T_i) = 0$ for all $i$ and such that $\{\alpha_{M'}^T(T_i)\}$ generates $\mathcal{Z}(M')$. A construction similar to that of $T$ produces a family with the same properties but with the roles of $\alpha$ and $\beta$ reversed, completing the proof of Proposition 4.3.3. \qed

5. Nonvanishing of Heegner classes

Keep $K$, $E_1$, $\mathfrak{N}$, and $\mathcal{T}_\mathfrak{N}(E_1) \cong \mathbb{Z}^2$ as in [2.1] so that $K$ is an imaginary quadratic field in which the prime divisors of $\mathcal{N}$ are split, $\mathcal{O}_K/\mathfrak{N} \cong \mathbb{Z}/n\mathbb{Z}$, and $E_1$ is an elliptic curve over $\mathbb{Q}^\text{al}$ with complex multiplication by $\mathcal{O}_K$. Let $D = \text{disc}(K)$ and let $H[p^*]$, $\mathcal{G}$, and $\mathcal{G}_0$ be as in [3.1]. Let $f \in S_{2k}(\Gamma_0(N), \mathbb{C})$, $\Phi$, $\chi$, and $\pi_{\chi}$ also be as in [3.1]. Let $\mathbf{T}$ be the $\mathbb{Z}$-algebra generated by the Hecke operators $\{T_m : (m, N) = 1\}$ and
Heegner cohomology classes. For each finite quotient \( \Lambda \) of \( \mathbb{Z}_p \) we have the \( \text{Gal}(\mathbb{Q}^{\text{al}}/K) \)-module \( W_{\Lambda} \) of (13) and, for each \( s \geq 0 \), the family of cohomology classes \( \Omega_s(g) \) of (17) parametrized by \( g \in \mathcal{T}_s \subset \mathcal{T} \). If we set \( W_{\mathbb{Z}_E} = \varinjlim W_{\mathbb{Z}/E^\sigma} \), then the classes \( \Omega_s(g) \) are compatible as \( \Lambda = \mathbb{Z} / E^\sigma \) varies, and define classes
\[
\Omega_s(g) \in H^1(\mathbb{Q}^{\text{al}}/\mathbb{H}[p^s], W_{\mathbb{Z}_E}(k))^\Delta,
\]
and also classes (denoted the same way) in the cohomology of \( W_{\Lambda} = W_{\mathbb{Z}_E} \otimes \Lambda \) for any \( \mathbb{Z}_E \)-algebra \( \Lambda \). Other constructions made with \( \Lambda = \mathbb{Z} / E^\sigma \) extend to any \( \mathbb{Z}_E \)-algebra \( \Lambda \) in the same way. We denote by
\[
\text{Heeg}_s \subset H^1(\mathbb{Q}^{\text{al}}/\mathbb{H}[p^s], W_{\mathbb{Z}_E}(k))^\Delta
\]
the \( \Phi \)-submodule generated by the classes \( \Omega_s(g) \) as \( g \) ranges over \( \mathcal{T}_s \). By a well known theorem of Deligne, the Hecke algebra \( T \otimes \Phi \) acts on \( W_{\mathbb{Z}_E} \), and the Galois representation \( W_f = \pi_f W_{\mathbb{Z}_E} \) is a two dimensional \( \Phi \)-vector space. Set
\[
\text{Heeg}_s(f) = \pi_f \text{Heeg}_s.
\]

**Theorem 5.1.1.** Fix a character \( \chi : G_0 \to \Phi^\times \) and let \( \pi_\chi \) be as in (10.1). Suppose \( \ell \) does not divide \( p, N, \varphi(N), \text{disc}(K) \), or \( (2k - 2)! \) As \( s \) grows the \( \Phi \)-dimension of \( \pi_\chi \text{Heeg}_s(f) \) grows without bound.

**Proof.** Let \( r_1, r_2, \ldots \) be the prime divisors of \( D \) and let \( r_i \) denote the unique prime of \( K \) above \( r_i \). Let \( G_1 \subset \mathcal{G} \) be the subgroup generated by the Frobenius classes of the \( r_i \), so that \( G_1 \) has exponent 2, and in particular \( G_1 \subset G_0 \). Reordering the \( r_i \) if needed, choose \( n \) such that the Frobenius classes of \( r_1, \ldots, r_n \) form a basis for the \( \mathbb{Z}/2\mathbb{Z} \)-vector space \( G_1 \). Set \( M = r_1 \cdots r_n \), so that divisors of \( M \) are naturally in bijection with the elements of \( G_1 \). We denote this bijection by \( d \mapsto \sigma_d \). Set \( \mathfrak{M} = \mathfrak{m}_1 \cdots \mathfrak{m}_k \). For each \( \sigma \in G_1 \) fix once and for all an extension of \( \sigma \) to \( \text{Gal}(\mathbb{Q}^{\text{al}}/K) \), and let \( S_1 \) denote the set of extensions so chosen. Let \( S_0 \subset \text{Gal}(\mathbb{Q}^{\text{al}}/K) \) be chosen so that restriction to \( H[p^\infty] \) takes \( S_0 \) injectively into \( G_0 \) with image equal to a set of representatives for the cosets \( G_0 / G_1 \). Set \( \mathcal{S} = \{ \sigma \tau \mid \sigma \in S_1, \tau \in S_0 \} \).

As in (3) let \( \Omega \) be a finite set of rational primes, all inert in \( K \) and all prime to \( \ell p N \), and fix extensions of these places to \( \mathbb{Q}^{\text{al}} \). We will continue our practice of writing \( q \in \Omega \) to indicate that \( q \) is the prime of \( K \) above the rational prime \( q \in \mathcal{O} \). For each \( q \in \Omega \) define, using the notation (19), \( \lambda_d : Z_q(M) \to Z_q(1) \) by \( \lambda_d = \beta_1^d \circ \alpha_2^d \) where \( \alpha_2^d \) and \( \beta_1^d \) are the degeneracy maps of (17). Consider the composition
\[
(\mathbb{Z}/\ell \mathbb{Z})[\mathcal{T}] \to \bigoplus_{(\sigma, q) \in S_0 \times \Omega} Z_q(M) \mathcal{T}^\sigma \bigoplus_{(\sigma, q) \in S_0 \times \Omega} Z_q(1) \to \bigoplus_{(\sigma, q) \in S \times \Omega} Z_q(1)
\]
in which the first arrow is the map \( \text{Red}_{S_0 \times \Omega} \) of (10), and the final arrow rearranges the sum, taking the summand \( (\sigma, q, d) \) to the summand \( (\sigma_d \sigma, q) \).

**Lemma 5.1.2.** The composition (38) is surjective, and is equal to the simultaneous reduction map (20) defined with \( M = 1 \).
Proof. By [5, Lemma 4.5] the set $S_0$ is chaotic in the sense of Definition 3.0.1 and so Theorem 3.4.3 gives the surjectivity of $\text{Red}_{S_0, \mathcal{Q}}$. The surjectivity of $\oplus_{d | M} \lambda_d$ is an easy induction using Proposition 4.3.1.

Fix $g \in T$ and let $A$ be an elliptic curve over $\mathbb{Q}^\text{al}$ with complex multiplication by $\mathcal{O}_g \subset K$. If $d | M$, let $\mathfrak{d}$ be the unique $\mathcal{O}_g$-ideal of norm $d$ and set $A' = A/A[\mathfrak{d}]$. The main theorem of complex multiplication provides an isomorphism $A' \cong A'^{\sigma_d}$ such that the composition $A \to A' \cong A'^{\sigma_d}$ agrees with $P \mapsto P^{\sigma_d}$ for all torsion points $P \in A(\mathbb{Q}^\text{al})$ of order prime to $d$. Thus

$$\lambda_d(E^\sigma_g, C^\sigma_g, \Theta^\sigma_g) = (E^\sigma_g, C^\sigma_g, \Theta^\sigma_g)^{\sigma_d}$$

for any $\sigma \in \text{Gal}(\mathbb{Q}^\text{al}/K)$, and the lemma follows. \qed

For each $(\sigma, q) \in S \times \mathcal{Q}$ and each $g \in T_s$ the cohomology class $\Omega_s(g)$ is unramified at $q$, and, since the residue field of $H[p^s]$ at $q$ is $\mathbb{F}_q$, the localization of $\Omega_s(g)$ at $q$ defines a class

$$\text{loc}_{\sigma, q}(g) \in H^1(\mathbb{F}_q^\text{al}/\mathbb{F}_q, W_{Z_\ell}(k))^{\Delta}.$$

Summing over all $(\sigma, q) \in S \times \mathcal{Q}$ and extending linearly to the free $\mathbb{Z}_\ell$-module on $T_s$ defines

$$\text{loc}_{S, \mathcal{Q}} : \mathbb{Z}_\ell[T_s] \to \bigoplus_{(\sigma, q) \in S \times \mathcal{Q}} H^1(\mathbb{F}_q^\text{al}/\mathbb{F}_q, W_{Z_\ell}(k))^{\Delta}.$$

This map is compatible with the natural inclusions as $s$ varies. Proposition 1.5.1 gives the commutative diagram

$$(39) \quad \mathbb{Z}_\ell[T_s] \xrightarrow{\text{loc}_{S, \mathcal{Q}}} \bigoplus H^1(\mathbb{F}_q^\text{al}/\mathbb{F}_q, W_{Z_\ell}(k))^{\Delta} \xrightarrow{\text{Red}_{S, \mathcal{Q}}} \bigoplus A^\sigma_{Z/\mathbb{Z}_\ell}(Z_1(N)_q; Y_1(N)/\mathbb{F}_q)^{\Delta} \xrightarrow{\lambda_d} \bigoplus H^1(\mathbb{F}_q^\text{al}/\mathbb{F}_q, W_{Z/\mathbb{Z}_\ell}(k))^{\Delta}$$

where all sums are over $S \times \mathcal{Q}$, $\text{Red}_{S, \mathcal{Q}}$ is the restriction of the simultaneous reduction map (19), with $M = 1$, to $\mathbb{Z}_\ell[T_s]$, and the bottom horizontal arrow is the $\mathbb{Z}/\ell\mathbb{Z}$-augmented Kummer map of Definition 1.3.3. By Corollary 4.2.6 the bottom horizontal arrow is surjective, and by Lemma 5.1.2 the restriction of $\text{Red}_{S, \mathcal{Q}}$ to $\mathbb{Z}_\ell[T_s]$ is surjective for $s \gg 0$. The same argument as [14, Lemma 2.2] gives the exactness of

$$\begin{align*}
0 &\to W_{Z_\ell}(k) \xrightarrow{\ell} W_{Z_\ell}(k) \to W_{Z/\mathbb{Z}_\ell}(k) \to 0,
\end{align*}$$

and taking $\mathbb{F}_q^\text{al}/\mathbb{F}_q$ cohomology shows that the right vertical arrow is surjective with kernel equal to the image of multiplication by $\ell$. Applying Nakayama’s lemma, we have proved

Lemma 5.1.3. For $s \gg 0$ the restriction of $\text{loc}_{S, \mathcal{Q}}$ to $\mathbb{Z}_\ell[T_s]$ is surjective.

Let $R$ be the integer ring of $\Phi$, so that $W_R$ is an $R$ lattice in $W_\Phi$ and $\pi_f W_R$ is an $R$ lattice in $W_f = \pi_f W_\Phi$. Let

$$\text{Heeg}_{R,s} \subset H^1(\mathbb{Q}^\text{al}/H[p^s], W_R(k))^{\Delta}$$

be the $R$ submodule generated by the classes $\Omega_s(g)$ for $g \in T_s$, and abbreviate

$$T = \pi_f W_R(k) \subset W_f(k).$$
Lemma 5.1.4. For $s > 0$, the image of the composition

$$\text{Heeg}_{\mathbb{R}, s} \xrightarrow{\pi \times \pi_f} H^1(\mathbb{Q}^\text{an}/H[p^s], T) \xrightarrow{\otimes \text{loc}} \bigoplus_{q \in \Omega} H^1(\mathbb{Q}_q^\text{an}/K_q, T)$$

is $\bigoplus_{q \in \Omega} H^1(\mathbb{F}_q^\text{an}/\mathbb{F}_q, T)$, the submodule of unramified cohomology classes.

Proof. Using Proposition 1.5.1, we see that the image of the composition lies in the unramified cohomology, and is equal to the image of

$$R[T_s] \xrightarrow{\text{locx}_S, \Omega} \bigoplus_S \bigoplus_{q \in \Omega} H^1(\mathbb{F}_q^\text{an}/\mathbb{F}_q, W_R(k)) \Delta \xrightarrow{\chi} H^1(\mathbb{F}_q^\text{an}/\mathbb{F}_q, W_R(k)) \Delta \xrightarrow{\pi_f} H^1(\mathbb{F}_q^\text{an}/\mathbb{F}_q, T)$$

where arrow labeled $\chi$ takes the element $(x_\sigma)_{\sigma \in S}$ to $\sum_{\sigma \in S} \chi(\sigma)x_\sigma$. The first arrow is surjective for $s > 0$ by Lemma 5.1.3, the second is obviously surjective, and the third is surjective by the fact that $\text{Gal}(\mathbb{F}_q^\text{an}/\mathbb{F}_q)$ has cohomological dimension one.

Let $m$ denote the maximal ideal of $R$ and set $T = T \otimes_R R/m$. If $q \nmid \ell ND$ is a rational prime whose absolute Frobenius acts as complex conjugation on $K(T)$, the extension of $\mathbb{Q}$ cut out by the Galois action on $T = T \otimes_R R/m$, then clearly $q$ is inert in $K$ and the Frobenius of the unique prime $q$ of $K$ above $q$ acts trivially on $T$. By the Chebotarëv theorem we may choose $\Omega$ as large as we want and containing only primes of this form. For $s > 0$, Lemma 5.1.4 gives a surjection from $\pi_\chi \pi_f \text{Heeg}_{\mathbb{R}, s}$ to

$$\bigoplus_{q \in \Omega} H^1(\mathbb{F}_q^\text{an}/\mathbb{F}_q, T) \otimes R/m \cong H^1(\mathbb{F}_q^\text{an}/\mathbb{F}_q, T) \cong \bigoplus_{q \in \Omega} T/\text{(Frob}_q - 1)T \cong \bigoplus_{q \in \Omega} T.$$

Thus the $R/m$ dimension of $(\pi_\chi \pi_f \text{Heeg}_{\mathbb{R}, s}) \otimes_R R/m$ is at least $\#\Omega$ for $s > 0$. Enlarging $\Omega$, the $R/m$ dimension of $(\pi_\chi \pi_f \text{Heeg}_{\mathbb{R}, s}) \otimes_R R/m$ grows without bound as $s$ increases.

Lemma 5.1.5. The $R$-torsion submodule of $H^1(\mathbb{Q}^\text{an}/H[p^s], T)$ is finite and of bounded order as $s \to \infty$.

Proof. The $R$-torsion submodule of $H^1(\mathbb{Q}^\text{an}/H[p^s], T)$ is isomorphic to the quotient of

$$(40) H^0(\mathbb{Q}^\text{an}/H[p^s], T \otimes_{\mathbb{Z}_\ell} (\mathbb{Q}_\ell/\mathbb{Z}_\ell))$$

by its maximal divisible subgroup. Let $\ell \nmid pN$ be a rational prime which is inert in $K$, and let $\lambda$ be the prime of $K$ above $\ell$. Then $\lambda$ splits completely in $H[p^\infty]$ and, by Deligne’s proof of the Ramanujan conjecture, $\text{Frob}_\lambda = \text{Frob}_\ell^2$ acts on $W_f(k)$ with eigenvalues of (complex) absolute value $\ell^{2k-1}$. Hence $\text{Frob}_\lambda - 1$ is invertible on $W_f(k) \cong T \otimes_{\mathbb{Z}_\ell} \mathbb{Q}_\ell$. A snake lemma argument then shows that the order of (40) is bounded by the order of $T/\text{(Frob}_\lambda - 1)T$.

The $R$-torsion submodule of $\pi_\chi \pi_f \text{Heeg}_{\mathbb{R}, s}$ is contained in the torsion submodule of $H^1(\mathbb{Q}^\text{an}/H[p^s], T)$, and so is finite and bounded as $s \to \infty$ by Lemma 5.1.5. We have seen that the $R/m$ dimension of $(\pi_\chi \pi_f \text{Heeg}_{\mathbb{R}, s}) \otimes_R R/m$ increases without bound, and it now follows that the $R$-rank of $\pi_\chi \pi_f \text{Heeg}_{\mathbb{R}, s}$ also increases without bound. This complete the proof of Theorem 5.1.1.\qed
## References

1. M. Bertolini. Selmer groups and Heegner points in anticyclotomic $\mathbb{Z}_p$-extensions. *Compositio Math.*, 99(2):153–182, 1995.
2. M. Bertolini and H. Darmon. Kolyvagin’s descent and Mordell-Weil groups over ring class fields. *J. Reine Angew. Math.*, 412:63–74, 1990.
3. J.-L. Brylinski. Heights for local systems on curves. *Duke Math. J.*, 59(1):1–26, 1989.
4. C. Cornut. *Reduction de Familles de Points CM*. PhD thesis, U. Strasbourg I, 2000.
5. C. Cornut. Mazur’s conjecture on higher Heegner points. *Invent. Math.*, 148(3):495–523, 2002.
6. P. Deligne. Formes modulaires et représentations de GL(2). In *Modular functions of one variable, II* (Proc. Internat. Summer School, Univ. Antwerp, Antwerp, 1972), pages 55–105. Lecture Notes in Math., Vol. 349. Springer, Berlin, 1973.
7. F. Diamond and R. Taylor. Nonoptimal levels of mod $l$ modular representations. *Invent. Math.*, 115(3):435–462, 1994.
8. B. Howard. The Heegner point Kolyvagin system. *Compos. Math.*, 140(6):1439–1472, 2004.
9. B. Howard. Variation of Heegner points in Hida families. In preparation.
10. Y. Ihara. On modular curves over finite fields. In *Discrete Subgroups of Lie Groups and Applications to Moduli (Internat. Colloq., Bombay, 1973)*, pages 161–202. Oxford Univ. Press, Bombay, 1975.
11. U. Jannsen. *Mixed Motives and Algebraic K-theory*. Number 1400 in Lecture Notes in Mathematics. Springer-Verlag, New York, 1990.
12. N. Katz and B. Mazur. *Arithmetic Moduli of Elliptic Curves*, volume 108 of *Annals of Mathematics Studies*. Princeton University Press, Princeton, NJ, 1985.
13. J. S. Milne. *Étale Cohomology*, volume 33 of *Princeton Mathematical Series*. Princeton University Press, Princeton, N.J., 1980.
14. J. S. Milne. *Arithmetic Duality Theorems*, volume 1 of *Perspectives in Mathematics*. Academic Press Inc., Boston, MA, 1986.
15. J. Nekovář. Kolyvagin’s method for Chow groups of Kuga-Sato varieties. *Invent. Math.*, 107(1):99–125, 1992.
16. J. Nekovář. On the $p$-adic height of Heegner cycles. *Math. Ann.*, 302(4):609–686, 1995.
17. J. Nekovář. $p$-adic Abel-Jacobi maps and $p$-adic heights. In *The arithmetic and geometry of algebraic cycles (Banff, AB, 1998)*, volume 24 of *CRM Proc. Lecture Notes*, pages 367–379. Amer. Math. Soc., Providence, RI, 2000.
18. B. Perrin-Riou. Fonctions $L$ $p$-adiques, théorie d’Iwasawa et points de Heegner. *Bull. Soc. Math. France*, 115(4):399–456, 1987.
19. K. A. Ribet. On modular representations of $\text{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ arising from modular forms. *Invent. Math.*, 100(2):431–476, 1990.
20. A. J. Scholl. $L$-functions of Modular Forms and Higher R regulators. Book in preparation.
21. A. J. Scholl. Motives for modular forms. *Invent. Math.*, 100(2):419–430, 1990.
22. J.-P. Serre. *Local Fields*, volume 67 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1979. Translated from the French by Marvin Jay Greenberg.
23. V. Vatsal. Uniform distribution of Heegner points. *Invent. Math.*, 148(1):1–46, 2002.
24. M.-F. Vignéras. *Arithmétique des Algèbres de Quaternions*, volume 800 of *Lecture Notes in Mathematics*. Springer, Berlin, 1980.
25. S. Zhang. Heights of Heegner cycles and derivatives of $L$-series. *Invent. Math.*, 130(1):99–152, 1997.

Dept. of Mathematics, Harvard University, 1 Oxford St., Cambridge, MA. 02138.