Vanishing of Some Cohomology Groups and Bounds for the Shafarevich-Tate Groups of Elliptic Curves

by

Byungchul Cha

A dissertation submitted to the Johns Hopkins University in conformity with the requirements for the degree of Doctor of Philosophy.

Baltimore, Maryland
April 17, 2003.

ABSTRACT

Let E be an elliptic curve over $\mathbf{Q}$ and ℓ be an odd prime. Also, let K be a number field and assume that E has a semi-stable reduction at ℓ. Under certain assumptions, we prove the vanishing of the Galois cohomology group $H^1(\mathrm{Gal}(K(E[\ell^i])/K), E[\ell^i])$ for all $i \geq 1$. When K is an imaginary quadratic field with the usual Heegner assumption, this vanishing theorem enables us to extend a result of Kolyvagin, which finds a bound for the order of the ℓ-primary part of Shafarevich-Tate groups of E over K. This bound is consistent with the prediction of Birch and Swinnerton-Dyer conjecture.

Advisor : Cristian Popescu.

ACKNOWLEDGEMENT

I would like to express my gratitude to Professor Victor Kolyvagin, under whose guidance this work has been completed. I also thank Professor Cristian Popescu for his invaluable help and encouragement for the past few years.

I am very grateful to my parents for their constant support. Finally, special thanks go to my wife Kyoungsun Kim for her love and encouragement.

Contents

1 Introduction **1**

1.1 Historical background 1

1.2 Main theorem and its connection with Shafarevich-Tate groups 1

2 Vanishing of the cohomology groups $H^1(\mathcal{G}_i, E[\ell^i])$ **4**

3 The structure of $\mathcal{G}$ **7**

3.1 Subgroups of $\mathrm{GL}(V)$ 7

3.2 Conditions on $\mathcal{G}$ 9

3.3 The case where ℓ does not divide $|\mathcal{G}|$ 9

3.3.1 $\mathcal{G}$ cyclic 10

3.3.2 $\mathcal{G}$ dihedral 10

3.3.3 $\mathcal{G}$ is $\mathcal{A}_4, \mathcal{S}_4$ or $\mathcal{A}_5$ 13

3.4 The case where ℓ divides $|\mathcal{G}|$ 13

4 The exceptional case **17**

4.1 Vanishing of $H^1(\mathcal{G}_i, E[\ell^i])$ 17

4.2 An example 26

5 Application **27**

5.1 Extension of Kolyvagin's result on Ш(E/K) 27

5.2 Irreducible vs surjective 30

5.2.1 $\ell = 3$ 30

5.2.2 $\ell = 3$ or 5 31

5.2.3 $\ell > 5$ 31

1 Introduction

1.1 Historical background

An *elliptic curve* over $\mathbf{Q}$ is a smooth, projective curve of genus 1 defined over $\mathbf{Q}$, together with a distinguished rational point on it. Let E be an elliptic curve over $\mathbf{Q}$. The study of the finitely generated abelian group $E(\mathbf{Q})$, which consists of all the points with rational coordinates on E, is very important both for theoretical and practical reasons. However the algorithm of computing $E(\mathbf{Q})$ for any E is not, in general, guaranteed to terminate in finitely many steps. At the heart of this problem lies the finiteness of an abelian group, the, so called, *Shafarevich-Tate group* $\text{Ш}(E/\mathbf{Q})$. For any number field K, $\text{Ш}(E/K)$ is defined as the kernel of the restriction map $H^1(K, E) \longrightarrow \prod_v H^1(K_v, E)$, where v runs over all places of K. It has long been conjectured that $\text{Ш}(E/\mathbf{Q})$ should always be finite for any elliptic curve E. However, the first example of such an E was given in the 80's when Rubin [12] proved the finiteness of $\text{Ш}(E/\mathbf{Q})$ for the elliptic curve E with complex multiplication whose value at $s = 1$ of the Hasse-Weil L-function $L(E/\mathbf{Q}, s)$ is non-zero. Another fundamental advance in this direction was made soon thereafter by Kolyvagin [6], [7]. He proved the finiteness of $\text{Ш}(E/\mathbf{Q})$ for the elliptic curve E when the order of vanishing of $L(E/\mathbf{Q}, s)$ at $s = 1$ is less than or equal to 1. The arguments of both Rubin and Kolyvagin mentioned above can be described as particular instances of the application of the theory of *Euler systems*, which was developed by Kolyvagin in [7].

Kolyvagin's method of Euler Systems allowed him to further study the structure of Ш. For a certain class of prime numbers ℓ, he was able to determine a bound for the order of the ℓ-primary part of Ш. This result was particularly interesting because his bound is consistent with the prediction made in a celebrated conjecture of Birch and Swinnerton-Dyer. (See the next subsection for the precise statement.) This thesis originated as an attempt of making Kolyvagin's procedure applicable to a wider class of primes ℓ.

1.2 Main theorem and its connection with Shafarevich-Tate groups

First, we fix some notations. For a finite abelian group A, we will write $|A|$ for its order. And, "$\operatorname{ord}_\ell n$" will denote the maximal integer m such that ℓ^m divides the natural number n.

Let E be a (modular) elliptic curve over $\mathbf{Q}$ whose conductor is N. And let K be a finite extension of $\mathbf{Q}$. Fix an odd prime ℓ. For each natural number $i \geq 1$, $E[\ell^i]$ will denote the group of ℓ^i-torsion points of E. We let L_i be the smallest Galois extension of K over which $E[\ell^i]$ is defined, and $\mathcal{G}_i = \mathrm{Gal}(L_i/K)$ be its Galois group over K. In particular, we set $L := L_1 = K(E[\ell])$ and $\mathcal{G} := \mathcal{G}_1 = \mathrm{Gal}(L/K)$. Throughout this article, we will assume that ℓ satisfies the following.

Assumption 1.1. (a) There is a prime v of K over ℓ which is unramified in $K/\mathbf{Q}$, and E has either good reduction or multiplicative reduction over the completion K_v of K at v.

(b) $E(K)$ has no ℓ-torsion points.

Under this assumption, we prove

Main Theorem. $H^1(\mathcal{G}_i, E[\ell^i]) = 0$ *for all* $i \geq 1$ *unless* $\ell = 3$ *and* $\mathcal{G} \simeq G_{\mathrm{except}}$.

(See (1) below for the definition of G_{except}.) The proof consists of three steps. The first step consists of proving the main theorem when $\mathcal{G}$ contains a nontrivial homothety. If $\mathcal{G}$ does not contain a nontrivial homothety, we show in §3 that $\mathcal{G}$ is isomorphic to $G_{\mathrm{except}} \subseteq \mathrm{GL}_2(\mathbf{Z}/\ell\mathbf{Z})$, where G_{except} is defined as

$$G_{\mathrm{except}} = \left\{ \begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix} \,\middle|\, a \in (\mathbf{Z}/\ell\mathbf{Z})^* \quad \text{and} \quad b \in \mathbf{Z}/\ell\mathbf{Z} \right\}. \tag{1}$$

Finally, the exceptional case $\mathcal{G} \simeq G_{\mathrm{except}}$ is studied in §4. We prove the vanishing except the case where $\ell = 3$.

The motivation of this work is the following. Take $K = \mathbf{Q}(\sqrt{D})$ to be an imaginary quadratic extension with fundamental discriminant $D \neq -3, -4$ where all prime divisors of N split. We also let $y_K \in E(K)$ be the Heegner point associated with the maximal order in K. Kolyvagin [7] proves that, when y_K is of infinite order, $E(K)$ has rank one and the Shafarevich-Tate group $Ш(E/K)$ of E over K is finite. Let m be the largest integer such that $y_K \in \ell^m E(K)$ modulo ℓ-torsion points. In [8], Kolyvagin proves the following.

Theorem 1.2 (Kolyvagin). *Suppose that* y_K *is of infinite order. Assume that* ℓ *is an odd prime. If the Galois group* $\mathrm{Gal}(\mathbf{Q}(E[\ell])/\mathbf{Q})$ *is isomorphic to* $\mathrm{GL}_2(\mathbf{Z}/\ell\mathbf{Z})$*, then we have*

$$\mathrm{ord}_\ell \, |Ш(E/K)| \leq 2m.$$

This bound for the ℓ-part of $|\text{Ш}(E/K)|$ is consistent with the conjecture of Birch and Swinnerton-Dyer. In fact, Gross and Zagier [4] obtained a formula for the value of the derivative of the complex L-function of E over K in terms of the height of y_K. This formula, when combined with the conjecture of Birch and Swinnerton-Dyer, yields the following conjectural formula for the ℓ-order of $\text{Ш}(E/K)$.

Conjecture 1.3. *Suppose that y_K is of infinite order. Then $\text{Ш}(E/K)$ is finite and its ℓ-order is*

$$\operatorname{ord}_\ell |\text{Ш}(E/K)| = 2m + 2\operatorname{ord}_\ell \left(\frac{|E(K)_{\text{tor}}|}{c \cdot \Pi_{q|N} c_q} \right).$$

Here c_q is the number of connected components of the special fiber of the Néron model of E at q, and c is the Manin constant of a modular parametrization of E.

In view of conjecture 1.3, it is natural to expect that the assumption that $E(K)$ has no nontrival ℓ-torsion points should be sufficient to yield the same bound as in Theorem 1.2, even in the case where $\text{Gal}(\mathbf{Q}(E[\ell])/\mathbf{Q})$ is a proper subgroup of $\text{GL}_2(\mathbf{Z}/\ell\mathbf{Z})$. We are not proving this result in this thesis. Instead, under the condition that the mod ℓ Galois representation

$$\rho_{\mathbf{Q}} : \text{Gal}(\bar{\mathbf{Q}}/\mathbf{Q}) \longrightarrow \text{Aut}(E[\ell]) \simeq \text{GL}_2(\mathbf{Z}/\ell\mathbf{Z})$$

is *irreducible* over $\mathbf{Z}/\ell\mathbf{Z}$, we show that the same bound can be obtained (Theorem 5.1). See §5 for more detailed discussion in this direction.

2 Vanishing of the cohomology groups $H^1(\mathcal{G}_i, E[\ell^i])$

First, we investigate the natural maps between $H^1(\mathcal{G}_i, E[\ell^i])$.

Proposition 2.1. *For each $i \geq 1$, there is a natural injection*

$$H^1(\mathcal{G}_i, E[\ell^i]) \longrightarrow H^1(\mathcal{G}_{i+1}, E[\ell^{i+1}]). \tag{2}$$

Proof. There are two natural injections

$$H^1(\mathcal{G}_i, E[\ell^i]) \longrightarrow H^1(\mathcal{G}_{i+1}, E[\ell^i]) \tag{3}$$

and

$$H^1(\mathcal{G}_{i+1}, E[\ell^i]) \longrightarrow H^1(\mathcal{G}_{i+1}, E[\ell^{i+1}]). \tag{4}$$

Indeed, the map (3) is just the inflation in the exact sequence

$$0 \longrightarrow H^1(\mathcal{G}_i, E[\ell^i]) \xrightarrow{\text{Inf}} H^1(\mathcal{G}_{i+1}, E[\ell^i]) \xrightarrow{\text{Res}} H^1(\mathrm{Gal}(L_{i+1}/L_i), E[\ell^i])^{\mathcal{G}_i}. \tag{5}$$

Also, the map (4) is given as follows. The exact sequence

$$0 \longrightarrow E[\ell^i] \longrightarrow E[\ell^{i+1}] \xrightarrow{\ell^i} E[\ell] \longrightarrow 0$$

gives the $\mathcal{G}_{i+1}$-cohomology long exact sequence, part of which is

$$E[\ell]^{\mathcal{G}_{i+1}} \longrightarrow H^1(\mathcal{G}_{i+1}, E[\ell^i]) \longrightarrow H^1(\mathcal{G}_{i+1}, E[\ell^{i+1}]) \xrightarrow{(\ell^i)_*} H^1(\mathcal{G}_{i+1}, E[\ell]). \tag{6}$$

The group $E[\ell]^{\mathcal{G}_{i+1}}$ is zero by Assumption 1.1, (b). Therefore, the map $H^1(\mathcal{G}_{i+1}, E[\ell^i]) \longrightarrow H^1(\mathcal{G}_{i+1}, E[\ell^{i+1}])$ is injective. This is (4).

Finally, the composion of (3) and (4) gives (2). □

The following lemma tells us how to control the size of $H^1(\mathcal{G}_i, E[\ell^i])$ inductively.

Lemma 2.2. *If the restriction map*

$$\mathrm{Res} : H^1(\mathcal{G}_{i+1}, E[\ell^i]) \longrightarrow H^1(\mathrm{Gal}(L_{i+1}/L_i), E[\ell^i])^{\mathcal{G}_i}$$

in (5) is the zero map, then

$$\dim_{\mathbf{Z}/\ell\mathbf{Z}} \left(H^1(\mathcal{G}_i, E[\ell^i]) \otimes \mathbf{Z}/\ell\mathbf{Z}\right) = \dim_{\mathbf{Z}/\ell\mathbf{Z}} \left(H^1(\mathcal{G}_{i+1}, E[\ell^{i+1}]) \otimes \mathbf{Z}/\ell\mathbf{Z}\right).$$

In particular, the above equality is true if $H^1(\mathrm{Gal}(L_{i+1}/L_i), E[\ell^i])^{\mathcal{G}_i} = 0$.

Proof. Consider the short exact sequence

$$0 \longrightarrow E[\ell] \xrightarrow{\iota} E[\ell^{i+1}] \xrightarrow{\ell} E[\ell^i] \longrightarrow 0$$

of $\mathcal{G}_{i+1}$-modules. Its $\mathcal{G}_{i+1}$-cohomology long exact sequence shows that

$$(\iota)_* : H^1(\mathcal{G}_{i+1}, E[\ell^i]) \longrightarrow H^1(\mathcal{G}_{i+1}, E[\ell^{i+1}])$$

is injective. Therefore, the kernel of $(\ell^i)_*$ in (6) coincides with that of the endomorphism of multiplication by ℓ^i on $H^1(\mathcal{G}_{i+1}, E[\ell^{i+1}])$.

However, the sequence (5) says that $H^1(\mathcal{G}_i, E[\ell^i])$ is isomorphic to $H^1(\mathcal{G}_{i+1}, E[\ell^i])$. Now, from (6), $H^1(\mathcal{G}_{i+1}, E[\ell^i])$ is the kernel of the multiplication on $H^1(\mathcal{G}_{i+1}, E[\ell^{i+1}])$ by ℓ^i, so the lemma follows. □

We study the structure of $H^1(\mathrm{Gal}(L_{i+1}/L_i), E[\ell^i])^{\mathcal{G}_i} = \mathrm{Hom}_{\mathcal{G}_i}(\mathrm{Gal}(L_{i+1}/L_i), E[\ell^i])$ more closely.

Define $\mathcal{A}$ to be the additive group $M_2(\mathbf{Z}/\ell\mathbf{Z})$ of all 2×2 matrices with coefficients in $\mathbf{Z}/\ell\mathbf{Z}$, and turn it into a $\mathcal{G}_i$-module by first projecting $\mathcal{G}_i$ onto $\mathcal{G} = \mathcal{G}_1$ and then letting it act on $\mathcal{A}$ by conjugation. By definition, this action factors through $\mathcal{G}$.

Fix a basis for $E[\ell^{i+1}]$. Then, we can identify $\mathcal{G}_{i+1}$ with a subgroup of $\mathrm{GL}_2(\mathbf{Z}/\ell^{i+1}\mathbf{Z})$. An element of $\mathrm{Gal}(L_{i+1}/L_i)$ will be of the form $I_2 + \ell^i A$ for some matrix A with coefficients in $\mathbf{Z}/\ell^{i+1}\mathbf{Z}$, where I_2 is the 2×2 identity matrix in $\mathrm{GL}_2(\mathbf{Z}/\ell^{i+1}\mathbf{Z})$. Note that A modulo ℓ is uniquely determined, independent of the choice of A, hence defines an element of $\mathcal{A}$. Therefore the map

$$I_2 + \ell^i A \longmapsto A \bmod \ell$$

identifies $\mathrm{Gal}(L_{i+1}/L_i)$ with a $\mathcal{G}_i$-submodule of $\mathcal{A}$ which will be denoted by $\mathcal{C}_i$.

Let f be an element in $\mathrm{Hom}_{\mathcal{G}_i}(\mathrm{Gal}(L_{i+1}/L_i), E[\ell^i]) \simeq \mathrm{Hom}_{\mathcal{G}_i}(\mathcal{C}_i, E[\ell^i])$. Since $\mathcal{C}_i$ is of exponent ℓ, the image of f lies in $E[\ell] \subseteq E[\ell^i]$. Moreover, the action of $\mathcal{G}_i$ on $\mathcal{C}_i$ factors through $\mathcal{G} = \mathcal{G}_1$. Therefore, we have $\mathrm{Hom}_{\mathcal{G}_i}(\mathrm{Gal}(L_{i+1}/L_i), E[\ell^i]) \simeq \mathrm{Hom}_{\mathcal{G}}(\mathcal{C}_i, E[\ell])$. In summary, we obtain the isomorphism

$$H^1(\mathrm{Gal}(L_{i+1}/L_i), E[\ell^i])^{\mathcal{G}_i} \simeq \mathrm{Hom}_{\mathcal{G}}(\mathcal{C}_i, E[\ell]) \tag{7}$$

When $\mathrm{Hom}_{\mathcal{G}}(\mathcal{C}_i, E[\ell]) = 0$, one can control the rank of $H^1(\mathcal{G}_{i+1}, E[\ell^{i+1}])$ inductively. This is the case when $\mathcal{G}$ contains a *homothety*, that is, a $(\mathbf{Z}/\ell\mathbf{Z})^*$-multiple of the identity endomorphism of $E[\ell]$.

Theorem 2.3. *If $\mathcal{G}$ contains a nontrivial homothety, then $H^1(\mathcal{G}_i, E[\ell^i]) = 0$ for all $i \geq 1$.*

Proof. Let $\langle \eta \rangle$ be the cyclic subgroup of $\mathcal{G}$ generated by a nontrivial homothety η. Then obviously $E[\ell]^{\langle \eta \rangle} = 0$. Further the cohomology group $H^1(\langle \eta \rangle, E[\ell]) = 0$ since the order of $\langle \eta \rangle$ is prime to ℓ. Therefore, by the following Hochschild-Serre spectral sequence,

$$0 \longrightarrow H^1(\mathcal{G}/\langle \eta \rangle, E[\ell]^{\langle \eta \rangle}) \longrightarrow H^1(\mathcal{G}, E[\ell]) \longrightarrow H^1(\langle \eta \rangle, E[\ell])$$

we get $H^1(\mathcal{G}, E[\ell]) = 0$.

Now, assume that $H^1(\mathcal{G}_i, E[\ell^i]) = 0$ for some i. From Lemma 2.2 and (7), we only need to show that $\mathrm{Hom}_{\mathcal{G}}(\mathcal{C}_i, E[\ell]) = 0$. Let $f \in \mathrm{Hom}_{\mathcal{G}}(\mathcal{C}_i, E[\ell])$. Note that any homothety acts trivially on $\mathcal{A}$. So, for any $v \in \mathcal{C}_i$, we have

$$f(v) = f(v^{\eta}) = \eta f(v).$$

But, only the zero element of $E[\ell]$ can be fixed by η, hence $f(v) = 0$. Therefore $f \equiv 0$. □

3 The structure of $\mathcal{G}$

The main theorem in this section is

Theorem 3.1. *If $\mathcal{G}$ does not contain a nontrivial homothety, then $\mathcal{G}$ can be represented as*

$$G_{\text{except}} = \left\{ \begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix} \Big| \, a \in (\mathbf{Z}/\ell\mathbf{Z})^* \quad \text{and} \quad b \in \mathbf{Z}/\ell\mathbf{Z} \right\}$$

with respect to some basis for $E[\ell]$.

The proof of this theorem will be given throughout this section. The main tool used is a result of Serre [14, §§1–2]. Serre studies the image of the representation

$$\rho_K \; : \mathrm{Gal}(\bar{K}/K) \longrightarrow \mathrm{GL}(E[\ell])$$

restricted to the local Galois group. Together with a group theoretic argument, Serre's result is used to classify all the possible subgroups of $\mathrm{GL}_2(\mathbf{Z}/\ell\mathbf{Z})$ without homotheties that can occur as our Galois group $\mathcal{G}$. Our assumption that $E(K)$ has no ℓ-torsion points also helps us limit the possibilities.

3.1 Subgroups of $\mathrm{GL}(V)$

The definitions in this subsection are taken from [14, §§1–2]. We summarize what we need for our study of $\mathcal{G}$.

Let V be a two-dimensional vector space over $\mathbf{Z}/\ell\mathbf{Z}$. By $\mathrm{GL}(V)$, we mean the group of all linear automorphisms of V. For a 1-dimensional subspace V_1 of V, define $B(V_1) \subseteq \mathrm{GL}(V)$ to be the subgroup consisting of all $s \in \mathrm{GL}(V)$ such that $sV_1 = V_1$. Such a subgroup $B(V_1)$ is called a *Borel subgroup* of $\mathrm{GL}(V)$ defined by V_1. The subspace V_1 is the unique 1-dimensional subspace of V which is stable under $B(V_1)$. By choosing a basis for V appropriately, such a subgroup $B(V_1)$ can be represented by 2×2 matrices

$$B(V_1) = \left\{ \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} \Big| \, a, d \in (\mathbf{Z}/\ell\mathbf{Z})^* \quad \text{and} \quad b \in \mathbf{Z}/\ell\mathbf{Z} \right\}.$$

When V_1 and V_2 are two distinct 1-dimensional subspaces of V, we let $C(V_1, V_2) \subseteq \mathrm{GL}(V)$ be the set of all the elements $s \in \mathrm{GL}(V)$ such that $sV_1 = V_1$ and $sV_2 = V_2$. The subgroup $C(V_1, V_2)$ is called the *split Cartan subgroup* of $\mathrm{GL}(V)$ defined by V_1 and V_2. In the appropriate basis for V, $C(V_1, V_2)$ takes the form

$$C(V_1, V_2) = \left\{ \begin{pmatrix} a & 0 \\ 0 & c \end{pmatrix} \Big| a, c \in (\mathbf{Z}/\ell\mathbf{Z})^* \right\}.$$

Therefore $C(V_1, V_2)$ is isomorphic to a product of two cyclic groups of order $\ell - 1$. We also note that V_1 and V_2 are the only 1-dimensional subspaces of V which are stable under $C(V_1, V_2)$. Let C_1 be the subgroup of $C(V_1, V_2)$, consisting of all elements whose actions on V_1 are trivial. Similarly, one can define C_2 to be the subgroup of $C(V_1, V_2)$ which acts trivially on V_2. Then C_1 and C_2 can be represented by matrices of the from $\left(\begin{smallmatrix} 1 & 0 \\ 0 & * \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} * & 0 \\ 0 & 1 \end{smallmatrix}\right)$. Such subgroups C_1 and C_2 are called *semi-split Cartan subgroups* of $\mathrm{GL}(V)$.

Let $\mathbf{F}_{\ell^2}$ be the unique quadratic extension of the field $\mathbf{Z}/\ell\mathbf{Z}$. Then one can embed $\mathbf{F}_{\ell^2}^*$ into $\mathrm{GL}(V)$, by choosing a basis for $\mathbf{F}_{\ell^2}$ over $\mathbf{Z}/\ell\mathbf{Z}$ and by representing $\mathbf{F}_{\ell^2}^*$ in $\mathrm{GL}(V)$ via the regular representation with respect to the chosen basis for $\mathbf{F}_{\ell^2}$. A *non-split Cartan subgroup* of $\mathrm{GL}(V)$ is, by definition, a subgroup of $\mathrm{GL}(V)$ which is conjugate to the image of $\mathbf{F}_{\ell^2}^*$ under this embedding in $\mathrm{GL}(V)$. Any non-split Cartan subgroup is cyclic of order $\ell^2 - 1$. Relevant to our study are the facts that the subgroup $(\mathbf{Z}/\ell\mathbf{Z})^*$ in $\mathbf{F}_{\ell^2}^*$ maps onto the homotheties of $\mathrm{GL}(V)$ regardless of the choice of a basis for $\mathbf{F}_{\ell^2}$, and thus that any non-split Cartan subgroup of $\mathrm{GL}(V)$ contains all homotheties.

Finally, we define the *Cartan subgroups* of $\mathrm{PGL}(V) = \mathrm{GL}(V)/(\mathbf{Z}/\ell\mathbf{Z})^*$ to be the images in $\mathrm{PGL}(V)$ of the corresponding Cartan subgroups of $\mathrm{GL}(V)$. Clearly, a split and a non-split Cartan subgroup of $\mathrm{PGL}(V)$ are both cyclic and are of order $\ell - 1$ and $\ell + 1$ respectively.

We state a lemma which will be useful later.

Lemma 3.2. *If $s \in \mathrm{GL}(V)$ is of order prime to ℓ, then the cyclic subgroup generated by s is contained in a Cartan subgroup of $\mathrm{GL}(V)$.*

Proof. The element s is (absolutely) semisimple since its order is prime to ℓ. So, the cyclic group generated by s is a commutative semisimple subgroup of $\mathrm{GL}(V)$. However, every maximal commutative semisimple subgroup of $\mathrm{GL}(V)$ is a Cartan subgroup (See [10, Lemma 12.2, Chap 18.]), hence the lemma follows. □

3.2 Conditions on $\mathcal{G}$

Let v be the prime of K over ℓ as in Assumption (a) of 1.1, that is v is unramified in $K/\mathbf{Q}$ and E does not have an additive reduction over K_v. We fix a decomposition group $D = D_v$ of v in $\mathrm{Gal}(\bar{K}/K)$, and let $I = I_v$ be the inertia group of v in D_v.

Proposition 3.3. *Assume that $\mathcal{G}$ contains no nontrivial homothety. Then*

(a) *E has either ordinary or multiplicative reduction over K_v.*

(b) *$\mathcal{G}$ contains a semi-split Cartan subgroup of* $\mathrm{GL}(E[\ell])$*. In particular, $\mathcal{G}$ contains a cyclic subgroup of order $\ell - 1$.*

Proof. If E has a supersingular reduction over K_v, the subgroup $\rho_K(I) \subseteq \mathcal{G}$ is a non-split Cartan subgroup of $\mathrm{GL}(E[\ell])$ [14, Proposition 12] and it would contain all homotheties, which contradicts our assumption on $\mathcal{G}$. Therefore, we conclude that the reduction type of E over K_v is either ordinary or multiplicative. In either case, the subgroup $\rho_K(I) \subseteq \mathcal{G}$ contains a semi-split Cartan subgroup of $\mathrm{GL}(E[\ell])$. (See [14, Corollaire to Proposition 11] and [14, Corollaire to Proposition 13].) Also, see §§3.4 below. □

3.3 The case where ℓ does not divide $|\mathcal{G}|$

We investigate the case when ℓ does not divide $|\mathcal{G}|$.

As before, let V be a two-dimensional vector space over $\mathbf{Z}/\ell\mathbf{Z}$. The following classification result is [14, Proposition 16].

Proposition 3.4. *If H is a subgroup of* $\mathrm{PGL}(V)$ *whose order is not divisible by ℓ, then H is cyclic, dihedral, or isomorphic to one of the groups $\mathcal{A}_4, \mathcal{S}_4$ and $\mathcal{A}_5$.*

We claim that, if ℓ does not divide $|\mathcal{G}|$, then $\mathcal{G}$ must contain a nontrivial homothety.

The rest of this subsection will be devoted to the proof of this claim. From now on, we work under the assumption that the group $\mathcal{G}$ has no nontrivial homotheties. Propositions 3.4 and 3.3 will lead us into a case by case analysis and yield a contradiction for all cases.

Since $\mathcal{G}$ is assumed to have no homothety, its image $\tilde{\mathcal{G}}$ in $\mathrm{PGL}(E[\ell])$ is isomorphic to $\mathcal{G}$. By Proposition 3.4, there are three cases: $\mathcal{G}$ is cyclic, dihedral or isomorphic to one of the groups $\mathcal{A}_4, \mathcal{S}_4$ and $\mathcal{A}_5$.

3.3.1 $\mathcal{G}$ cyclic

By Lemma 3.2, $\mathcal{G}$ is contained in a Cartan subgroup S of $\mathrm{GL}(E[\ell])$. And, by Proposition 3.3, $\mathcal{G}$ contains a semi-split Cartan subgroup C of $\mathrm{GL}(E[\ell])$, so we have $C \subseteq \mathcal{G} \subseteq S$ as subgroups of $\mathrm{GL}(E[\ell])$.

We consider the case where S is non-split, so the order S is $\ell^2 - 1$. Recall that $\mathcal{G}$ maps isomorphically onto $\tilde{\mathcal{G}}$. Therefore, $\ell - 1$ divides $|\tilde{\mathcal{G}}|$, hence it also divides the order of the image $\tilde{S}$ of S in $\mathrm{PGL}(E[\ell])$, which is just $\ell + 1$. But, this is impossible unless $\ell = 3$. When $\ell = 3$, the group S is isomorphic to $\mathbf{F}_9^*$, and its subgroup consisting of all homotheties corresponds to $\mathbf{F}_3^*$ in $\mathbf{F}_9^*$. It is easy to check that every nontrivial subgroup of $\mathbf{F}_9^*$ contains $\mathbf{F}_3^*$. Therefore $\mathcal{G}$ must also contain a nontrivial homothety.

Next, we assume that S is split. From the inclusion $C \subseteq \mathcal{G} \subseteq S$, it follows that $\mathcal{G}$ should be equal to C, otherwise $\mathcal{G}$ would have a nontrivial homothety. But $C = \mathcal{G}$ is also impossible since it would violate the ℓ-torsion freeness of $E(K)$.

3.3.2 $\mathcal{G}$ dihedral

Next, we deal with the case where $\mathcal{G}$ is isomorphic to a dihedral group D_k of order $2k$ for some k.

First, let us assume $\ell > 3$. Again we denote by C a semi-split Cartan subgroup contained in $\mathcal{G}$, which is just a cyclic group of order $\ell - 1 \geq 4$. In particular, we have $k \geq 2$. But, if $k = 2$, then ℓ must be 5, and C is of order 4. However, D_2 cannot have such a subgroup. So, we have $k > 2$.

Lemma 3.5. *Let $D_k = \langle x, y \,|\, x^2 = 1, y^k = 1, xy^ix^{-1} = y^{-i} \quad \text{for all } i\rangle$ be the dihedral group with $k > 2$, generated by the elements x and y of order 2 and k respectively. If D_k contains a cyclic group C of order > 2, then C is a subgroup of $\langle y \rangle$.*

Proof. Any element of the form xy^i is of order 2, so no such element can generate C. □

Following the notation in the lemma, we let $x, y \in \mathcal{G}$ be the elements of order 2 and k respectively. Then, the lemma implies that $C \subseteq \langle y \rangle$. Fix a basis for $E[\ell]$ such that the subgroup C is represented by the matrices of the form $\left(\begin{smallmatrix} * & 0 \\ 0 & 1 \end{smallmatrix}\right)$. Let $x = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$. Then we have

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} s & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} s^{-1} & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

for all $s \in (\mathbf{Z}/\ell\mathbf{Z})^*$. Or equivalently

$$\begin{aligned} as &= s^{-1}a & \qquad b &= s^{-1}b \\ cs &= c & \qquad d &= d \end{aligned}$$

for all $s \in (\mathbf{Z}/\ell\mathbf{Z})^*$. Obviously, such $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{GL}_2(\mathbf{Z}/\ell\mathbf{Z})$ cannot exist.

Next, let us assume that $\ell = 3$. Again, we fix a basis for $\mathrm{GL}(E[3])$ so that the subgroup C is represented as $\{\left(\begin{smallmatrix} \pm 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)\}$. So, in particular, $\tau := \left(\begin{smallmatrix} -1 & 0 \\ 0 & 1 \end{smallmatrix}\right) \in \mathcal{G}$. Then, using the conditions that $\mathcal{G}$ has no homotheties and that 3 does not divide the order of $\mathcal{G}$, we will prove that $\left(\begin{smallmatrix} \pm 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$ are the only elements in $\mathcal{G}$. Then, this would be a contradiction to the assumption that $E(K)$ has no ℓ-torsion points.

The table in the following page summarizes this computation. The column labelled as "Comments" shows why $\mathcal{G}$ cannot contain the matrices σ, except the ones of Type 0. The matrices shown in the table exhaust all of $|GL_2(\mathbf{Z}/3\mathbf{Z})| = 48$ possibilities.

Type	The matrices σ	Comment
0	$\begin{pmatrix}1 & 0\\0 & 1\end{pmatrix}$ or $\begin{pmatrix}-1 & 0\\0 & 1\end{pmatrix}$	$\sigma \in C$
A	$\begin{pmatrix}-1 & 0\\0 & -1\end{pmatrix}$, $\begin{pmatrix}1 & 0\\0 & -1\end{pmatrix}$	σ or $\sigma\tau$ is a homothety.
B	$\pm\begin{pmatrix}0 & 1\\-1 & 0\end{pmatrix}$	σ^2 is of Type A.
B′	$\pm\begin{pmatrix}0 & 1\\1 & 0\end{pmatrix}$	$\sigma\tau$ is of Type B.
C	$\pm\begin{pmatrix}1 & 1\\0 & 1\end{pmatrix}, \pm\begin{pmatrix}1 & -1\\0 & 1\end{pmatrix}, \pm\begin{pmatrix}1 & 0\\1 & 1\end{pmatrix}, \pm\begin{pmatrix}1 & 0\\-1 & 1\end{pmatrix}$	3 divides the order of σ.
C′	$\pm\begin{pmatrix}-1 & 1\\0 & 1\end{pmatrix}, \pm\begin{pmatrix}1 & 1\\0 & -1\end{pmatrix}, \pm\begin{pmatrix}1 & 0\\1 & -1\end{pmatrix}, \pm\begin{pmatrix}-1 & 0\\1 & 1\end{pmatrix}$	$\sigma\tau$ or $\tau\sigma$ is of Type C.
D	$\pm\begin{pmatrix}1 & -1\\1 & 1\end{pmatrix}, \pm\begin{pmatrix}1 & 1\\-1 & 1\end{pmatrix}$	σ^2 is of Type B.
D′	$\pm\begin{pmatrix}-1 & 1\\1 & 1\end{pmatrix}, \pm\begin{pmatrix}1 & 1\\1 & -1\end{pmatrix}$	$\sigma\tau$ or $\tau\sigma$ is of Type D.
E	$\pm\begin{pmatrix}1 & 1\\1 & 0\end{pmatrix}$	σ^2 is of Type D'.
E′	$\pm\begin{pmatrix}1 & -1\\1 & 0\end{pmatrix}, \pm\begin{pmatrix}1 & 1\\-1 & 0\end{pmatrix}$	$\sigma\tau$ or $\tau\sigma$ is of Type E.
E″	$\pm\begin{pmatrix}-1 & 1\\1 & 0\end{pmatrix}$	$\sigma\tau$ is of Type E'.
F	$\pm\begin{pmatrix}0 & 1\\1 & 1\end{pmatrix}$	σ^2 is of Type D'.
F′	$\pm\begin{pmatrix}0 & -1\\1 & 1\end{pmatrix}, \pm\begin{pmatrix}0 & 1\\-1 & 1\end{pmatrix}$	$\sigma\tau$ or $\tau\sigma$ is of Type F.
F″	$\pm\begin{pmatrix}0 & 1\\1 & -1\end{pmatrix}$	$\sigma\tau$ is of Type F'.

3.3.3 $\mathcal{G}$ is $\mathcal{A}_4, \mathcal{S}_4$ or $\mathcal{A}_5$

Here ℓ cannot be 3, since 3 divides the orders of $\mathcal{A}_4, \mathcal{S}_4$ and $\mathcal{A}_5$. We again denote by C the subgroup of $\mathcal{G}$ which is cyclic of order $\ell - 1$ as in Proposition 3.3. Let's first assume that $\ell > 5$. Then, one of the groups $\mathcal{A}_4, \mathcal{S}_4$ and $\mathcal{A}_5$ must contain C, which is cyclic of order ≥ 6. This is impossible. We also note that 5 divides the order of $\mathcal{A}_5$. Therefore we have to do the case that $\ell = 5$ and $\mathcal{G}$ is isomorphic to either $\mathcal{A}_4$ or $\mathcal{S}_4$. But, the group $\mathcal{A}_4$ doesn't contain an element of order 4, that is, there is no 4-cycle in $\mathcal{A}_4$. The only case left is $\ell = 5$ and $\mathcal{G}$ isomorphic to $\mathcal{S}_4$.

Choose a basis for $\mathrm{GL}(E[5])$, so that C is of the form $\left(\begin{smallmatrix} * & 0 \\ 0 & 1 \end{smallmatrix}\right)$. Then, there are 2 generators $\left(\begin{smallmatrix} 2 & 0 \\ 0 & 1 \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} 3 & 0 \\ 0 & 1 \end{smallmatrix}\right)$ of C. Since their traces are different they are not conjugate to each other. However, the 4-cycles in $\mathcal{S}_4$ form a single conjugacy class, therefore $\mathcal{S}_4$ cannot be isomorphic to $\mathcal{G}$.

3.4 The case where ℓ divides $|\mathcal{G}|$

Now, we study the case when ℓ divides $|\mathcal{G}|$

Proposition 3.6. *If ℓ divides the order of the Galois group $\mathcal{G}$, then $\mathcal{G}$ is either isomorphic to the full group* $\mathrm{GL}(E[\ell])$ *or is contained in a Borel subgroup of* $\mathrm{GL}(E[\ell])$.

Proof. By [14, Proposition 15], either $\mathcal{G}$ contains $\mathrm{SL}(E[\ell])$ or $\mathcal{G}$ is contained in a Borel subgroup of $\mathrm{GL}(E[\ell])$.

Recall that v is assumed to be unramified in $K/\mathbf{Q}$. Therefore the extension $K/\mathbf{Q}$ is linearly disjoint with the cyclotomic extension $\mathbf{Q}(\mu_\ell)/\mathbf{Q}$. If $\mathcal{G}$ contains $\mathrm{SL}(E[\ell])$, then it must be equal to $\mathrm{GL}(E[\ell])$ since the determinant map

$$\det : \mathcal{G} \longrightarrow (\mathbf{Z}/\ell\mathbf{Z})^*$$

is surjective due to Weil pairing on $E[\ell]$. □

We keep the assumption that $\mathcal{G}$ has no homothety, and we further assume that ℓ divides the order of $\mathcal{G}$. We will finish the proof of Theorem 3.1.

Recall that v is a fixed prime of K over ℓ as in Assumption (a) of 1.1 and that $I = I_v$ denotes a inertia group of v. In Proposition 3.3, we proved that the reduction type of E over K_v is either ordinary or multiplicative. When E has ordinary reduction, $X_\ell \subseteq E[\ell]$

will denote the kernel of reduction modulo ℓ. In the case of multiplicative reduction, $E(\bar{K}_v)$ is isomorphic to $\bar{K}_v^*/q^{\mathbf{Z}}$ (over some unramified extension of K_v) for some q in the ring of integers of K_v. [15, chap V] Via this isomorphism, the group μ_ℓ of ℓ-th roots of unity maps into $E[\ell]$. We define X_ℓ to be the image of μ_ℓ in $\bar{K}_v^*/q^{\mathbf{Z}}$. In both cases, X_ℓ is 1-dimensional $\mathbf{Z}/\ell\mathbf{Z}$-subspace of $E[\ell]$. The following proposition is a simple consequence of Corollaire to Proposition 11 and Corollaire to Proposition 13 in [14].

Proposition 3.7. *Fix a nonzero $x \in X_\ell$ and $x' \notin X_\ell$. With respect to the basis $\{x, x'\}$ for $E[\ell]$, we have the following.*

(a) *If the wild part of I acts on $E[\ell]$ nontrivially, then $\rho_K(I)$ is equal to*

$$\left\{ \begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix} \;\middle|\; a \in (\mathbf{Z}/\ell\mathbf{Z})^* \quad \text{and} \quad b \in \mathbf{Z}/\ell\mathbf{Z} \right\}.$$

(b) *If the wild part of I acts on $E[\ell]$ trivially, then $\rho_K(I)$ is equal to*

$$\left\{ \begin{pmatrix} a & 0 \\ 0 & 1 \end{pmatrix} \;\middle|\; a \in (\mathbf{Z}/\ell\mathbf{Z})^* \right\}.$$

From Proposition 3.6, there is a Borel subgroup B of $\mathrm{GL}(E[\ell])$ containing $\mathcal{G}$. If there is another Borel subgroup B' of $\mathrm{GL}(E[\ell])$ containing $\mathcal{G}$, then $\mathcal{G}$ will leave stable two distinct 1-dimensional $\mathbf{Z}/\ell\mathbf{Z}$ subspaces of $E[\ell]$ defined by B and B'. Therefore $\mathcal{G}$ will be contained in a split Cartan subgroup, which is a contradiction to the assumption that ℓ divides $|\mathcal{G}|$. So, B is the unique Borel subgroup containing $\mathcal{G}$. We denote by V_1 the unique 1-dimensional subspace which is stable under the action of B.

We now claim that $V_1 = X_\ell$. Assume the contrary $V_1 \neq X_\ell$. Then we can take $\{x, v\}$ as a basis for $E[\ell]$ with $x \in X_\ell$ and $v \in V_1$. We will write the elements of $\mathrm{GL}(E[\ell])$ as 2×2 matrices with coefficients in $\mathbf{Z}/\ell\mathbf{Z}$ with respect to this basis.

Since V_1 is stable under the action of $\mathcal{G}$, any element in $\mathcal{G}$ is lower triangular. We let

$$\alpha, \delta \,:\, \mathrm{Gal}(\bar{K}/K) \longrightarrow (\mathbf{Z}/\ell\mathbf{Z})^*$$

be the group homomorphisms and

$$\gamma \,:\, \mathrm{Gal}(\bar{K}/K) \longrightarrow \mathbf{Z}/\ell\mathbf{Z}$$

be the function (not necessarily a homomorphism) such that

$$\rho_K(s) = \begin{pmatrix} \alpha(s) & 0 \\ \gamma(s) & \delta(s) \end{pmatrix}$$

for all $s \in \mathrm{Gal}(\bar{K}/K)$.

Fix an element $t \in \mathrm{Gal}(\bar{K}/K)$ with $\delta(t) \neq 1$. Such an element exists, otherwise $E(K)$ would contain a nonzero ℓ-torsion element.

- Case 1: $\alpha(t) = \delta(t)$ and $\gamma(t) = 0$.

  The element $\rho_K(t)$ is a nontrivial homothety in $\mathcal{G}$, so it is a contradiction.

- Case 2: $\gamma(t) = 0$ but $\alpha(t) \neq \delta(t)$.

  In view of our choice of the basis $\{x, v\}$, we see that $\gamma|_I = 0$ and $\delta|_I = 1$ by Lemma 3.7. It also follows that the wild part of I acts trivially on X_ℓ and $\alpha|_I : I \to (\mathbf{Z}/\ell\mathbf{Z})^*$ is surjective. So, we can find an element $t' \in I$ with $\alpha(t') = \alpha(t)^{-1}\delta(t)$. Then,

  $$\rho_K(t')\rho_K(t) = \begin{pmatrix} \alpha(t)^{-1}\delta(t) & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \alpha(t) & 0 \\ 0 & \delta(t) \end{pmatrix} = \begin{pmatrix} \delta(t) & 0 \\ 0 & \delta(t) \end{pmatrix}$$

  will be an element in $\mathcal{G}$. But this element is a nontrivial homothety.

- Case 3: $\gamma(t) \neq 0$.

  We fix $u \in \mathrm{Gal}(\bar{K}/K)$ with $\rho_K(u) = \left(\begin{smallmatrix} 1 & 0 \\ 1 & 1 \end{smallmatrix}\right)$, which is an element of order ℓ in $\mathcal{G}$. (Recall that we assume that ℓ divides the order of $\mathcal{G}$.) Let $k = -\gamma(t)\alpha(t)^{-1}$. Then,

  $$\rho_K(u^k)\rho_K(t) = \begin{pmatrix} 1 & 0 \\ k & 1 \end{pmatrix} \begin{pmatrix} \alpha(t) & 0 \\ \gamma(t) & \delta(t) \end{pmatrix} = \begin{pmatrix} \alpha(t) & 0 \\ 0 & \delta(t) \end{pmatrix}.$$

  So, we can replace our t with $u^k t$ and go back to Case 1 or Case 2. Again we obtain a contradiction. This finishes the proof of the claim that $V_1 = X_\ell$.

From now on, we fix a basis $\{x, y\}$ for $E[\ell]$ with $x \in X_\ell$ and $y \notin X_\ell$. Since the subspace $X_\ell = V_1$ is stable under the action of $\mathcal{G}$, all the elements in $\mathcal{G}$ are upper triangular with respect to this basis.

Now, We can complete the proof of Theorem 3.1. It will be an immediate corollary to the following group theoretic lemma.

Lemma 3.8. *Let G be a subgroup in $\mathrm{GL}_2(\mathbf{Z}/\ell\mathbf{Z})$ which is upper triangular. Assume that G satisfies*

(a) *ℓ divides $|G|$.*

(b) *G contains no nontrivial homothety.*

(c) *The subgroup*
$$\left\{ \begin{pmatrix} a & 0 \\ 0 & 1 \end{pmatrix} \Big| \, a \in (\mathbf{Z}/\ell\mathbf{Z})^* \right\}$$
is contained in G.

Then G is equal to the group

$$C = \left\{ \begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix} \Big| \, a \in (\mathbf{Z}/\ell\mathbf{Z})^* \quad b \in \mathbf{Z}/\ell\mathbf{Z} \right\}.$$

Proof. By (a), the matrix $\left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$ is in G. The condition (c) then says that C is contained in G. Consider the projection $\mathrm{GL}_2(\mathbf{Z}/\ell\mathbf{Z}) \to \mathrm{PGL}_2(\mathbf{Z}/\ell\mathbf{Z})$. Both G and C maps isomorphically onto their respective images in $\mathrm{PGL}_2(\mathbf{Z}/\ell\mathbf{Z})$ since they don't have any homotheties. But, it is easy to see that the image of C covers the image of all upper triangular matrices, in which G is contained. So, we have $C = G$. □

Taking $\{x, y\}$ as a basis for $E[\ell]$ as above, it is clear that $\mathcal{G}$ satisfies all the conditions in the lemma. The proof of Theorem 3.1 is completed.

4 The exceptional case

We prove the vanishing of $H^1(\mathcal{G}_i, E[\ell^i])$ when $\mathcal{G} \simeq G_{\text{except}}$. In this section, we assume that $\ell \neq 3$. This is necessary in proving the vanishing of $H^1(\mathcal{G}_i, E[\ell^i])$. However, the proof works well for $\ell = 3$ in some cases. See Remark 4.7.

and prove the vanishing.

4.1 Vanishing of $H^1(\mathcal{G}_i, E[\ell^i])$

We will work with a fixed system of compatible basis for $E[\ell^i]$ for all $i \geq 1$, or equivalently, a fixed basis for the Tate module $T_\ell(E)$ of E. This enables us to identify $\mathcal{G}_i$ with a subgroup of $\mathrm{GL}_2(\mathbf{Z}/\ell^i\mathbf{Z})$. In particular, we have the identification $\mathcal{G} = G_{\text{except}}$ at the first level $i = 1$.

We recall the following notations from §2; we let $\mathcal{G}_i$ act on $\mathcal{A} = M_2(\mathbf{Z}/\ell\mathbf{Z})$ by conjugation. The group $\mathrm{Gal}(L_{i+1}/L_i)$ is identified with a $\mathcal{G}_i$-submodule $\mathcal{C}_i$ of $\mathcal{A}$ via the identification

$$I_2 + \ell^i A \longmapsto A \bmod \ell. \tag{8}$$

From all this, we have that

$$H^1(\mathrm{Gal}(L_{i+1}/L_i), E[\ell^i])^{\mathcal{G}_i} \simeq \mathrm{Hom}_{\mathcal{G}}(\mathcal{C}_i, E[\ell]). \tag{9}$$

One can classify all the possible $\mathcal{G}$-submodules of $\mathcal{A}_0 \subseteq \mathcal{A}$, where $\mathcal{A}_0$ is defined by $\mathcal{A}_0 = \{A \in \mathcal{A} |\, \mathrm{Tr}\, A = 0\}$. Let $w = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right), u = \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$ and $v = \left(\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right)$ be elements of $\mathcal{A}_0$. And also let $\mathcal{W} = \langle w \rangle$ and $\mathcal{U} = \langle w, u \rangle$ be subspaces of $\mathcal{A}_0$.

Note that $\mathcal{G}$ is generated by $\tau := \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$ and $\sigma_a := \left(\begin{smallmatrix} a & 0 \\ 0 & 1 \end{smallmatrix}\right)$ for all $a \in (\mathbf{Z}/\ell\mathbf{Z})^*$.

Proposition 4.1. *The subspaces $\{0\}, \mathcal{W}, \mathcal{U}$ and $\mathcal{A}_0$ are the only $\mathcal{G}$-submodules of $\mathcal{A}_0$.*

Proof. One checks easily that $\mathcal{W}$ and $\mathcal{U}$ are invariant under the action of $\mathcal{G}$.

Take $\{w, u, v\}$ as a basis of $\mathcal{A}_0$. Then an elementary computation shows that the matrix

$$\begin{pmatrix} 1 & -2 & -1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

represents the action of $\tau \in \mathcal{G}$ on $\mathcal{A}_0$. So, the only subspaces invariant under the action of τ are $\{0\}, \mathcal{W}, \mathcal{U}$ and $\mathcal{A}_0$. □

Proposition 4.2. *We have the following*

(a) $\mathrm{Hom}_{\mathcal{G}}(\mathcal{A}_0, E[\ell]) = 0$.

(b) $\mathrm{Hom}_{\mathcal{G}}(\mathcal{U}, E[\ell]) \simeq \mathbf{Z}/\ell\mathbf{Z}$.

(c) $\mathrm{Hom}_{\mathcal{G}}(\mathcal{W}, E[\ell]) \simeq \mathbf{Z}/\ell\mathbf{Z}$.

Proof. With respect to the basis $\{w, u, v\}$, the action of $\sigma_a = \left(\begin{smallmatrix} a & 0 \\ 0 & 1 \end{smallmatrix}\right) \in \mathcal{G}$ on $\mathcal{A}_0$ is represented by

$$\begin{pmatrix} a & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & a^{-1} \end{pmatrix}.$$

Any map $f \in \mathrm{Hom}(\mathcal{A}_0, E[\ell])$ will be written as the matrix

$$f = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix}$$

with coefficients in $\mathbf{Z}/\ell\mathbf{Z}$. Then, f is $\mathcal{G}$-equivariant if and only if

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix} \begin{pmatrix} 1 & -2 & -1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix} \qquad \text{and}$$

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix} \begin{pmatrix} a & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & a^{-1} \end{pmatrix} = \begin{pmatrix} a & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix}$$

for all $a \in (\mathbf{Z}/\ell\mathbf{Z})^*$. Solving these linear conditions on a_{ij}, we get $a_{ij} = 0$ for all i and j, therefore, $f = 0$. We proved (a).

Similarly, the actions of τ and σ_a on $\mathcal{U}$, with respect to the basis $\{w, u\}$, are represented by the matrices

$$\begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix} \qquad \text{and} \qquad \begin{pmatrix} a & 0 \\ 0 & 1 \end{pmatrix}$$

respectively. Again, we write $f \in \mathrm{Hom}(\mathcal{U}, E[\ell])$ as

$$f = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}.$$

In this case, the same computation as above says that f is $\mathcal{G}$-equivariant when

$$f = a_{11}\begin{pmatrix} 1 & 0 \\ 0 & -2 \end{pmatrix}.$$

In particular, $\mathrm{Hom}_{\mathcal{G}}(\mathcal{U}, E[\ell])$ is isomorphic to $\mathbf{Z}/\ell\mathbf{Z}$ and is generated by the map which sends w and u to P_1 and $-2Q_1$ respectively.

For (c), the same argument is used. We omit the details, but we note that a generator of $\mathrm{Hom}_{\mathcal{G}}(\mathcal{W}, E[\ell]) \simeq \mathbf{Z}/\ell\mathbf{Z}$ can be chosen so as to send w to P_1. □

Corollary 4.3. *Let $\mathcal{S}$ be a $\mathcal{G}$-submodule of $\mathcal{A}_0$, and let $f \in \mathrm{Hom}_{\mathcal{G}}(\mathcal{S}, E[\ell])$. The function f is nonzero if and only if w is in $\mathcal{S}$ and $f(w) \neq 0$.*

Proof. In the two previous propositions, we computed $\mathrm{Hom}_{\mathcal{G}}(\mathcal{S}, E[\ell])$ for *any* $\mathcal{G}$-submodules $\mathcal{S}$ of $\mathcal{A}_0$. The corollary now follows from the description of generators of $\mathrm{Hom}_{\mathcal{G}}(\mathcal{S}, E[\ell])$. □

A similar result is needed for $\mathcal{G}$-submodules of $\mathcal{A}$, rather than those of $\mathcal{A}_0$. Let $\mathcal{H} = \{\left(\begin{smallmatrix} a & 0 \\ 0 & a \end{smallmatrix}\right) \in \mathcal{A} \mid a \in \mathbf{Z}/\ell\mathbf{Z}\}$. Then, $\mathcal{G}$ acts on $\mathcal{H}$ trivially and there is a decomposition $\mathcal{A} = \mathcal{A}_0 \oplus \mathcal{H}$ as $\mathcal{G}$ modules. Since $E[\ell]$ has no $\mathcal{G}$-invariant elements we have that $\mathrm{Hom}_{\mathcal{G}}(\mathcal{H}, E[\ell]) = 0$.

Proposition 4.4. *Let $\mathcal{X}$ be a $\mathcal{G}$-submodule of $\mathcal{A}$ and let $f \in \mathrm{Hom}_{\mathcal{G}}(\mathcal{X}, E[\ell])$. The function f is nonzero if and only if w is in $\mathcal{X}$ and $f(w) \neq 0$.*

Proof. If $\mathcal{H} \subseteq \mathcal{X}$, then $\mathcal{H}$ occurs as a direct summand of $\mathcal{X}$ as $\mathcal{G}$-modules, i.e. $\mathcal{X} = \mathcal{X}_0 \oplus \mathcal{H}$ with $\mathcal{X}_0 = \mathcal{X} \cap \mathcal{A}_0$. Then

$$\mathrm{Hom}_{\mathcal{G}}(\mathcal{X}, E[\ell]) = \mathrm{Hom}_{\mathcal{G}}(\mathcal{X}_0, E[\ell]) \oplus \mathrm{Hom}_{\mathcal{G}}(\mathcal{H}, E[\ell]) = \mathrm{Hom}_{\mathcal{G}}(\mathcal{X}_0, E[\ell]),$$

hence Corollary 4.3 gives the desired result.

When $\mathcal{H} \not\subseteq \mathcal{X}$ and $\mathcal{X} \neq 0$, we note that the map

$$i : \mathcal{X} \hookrightarrow \mathcal{A} \to \mathcal{A}/\mathcal{H} \simeq \mathcal{A}_0$$

is injective. Therefore, $i(\mathcal{X})$ is isomorphic to $\mathcal{W}$,$\mathcal{U}$ or $\mathcal{A}_0$ by Proposition 4.1. In particular, $\mathcal{X}$ must contain an element of the form $x = w+h$ for some $h \in \mathcal{H}$. Then for any $a \in (\mathbf{Z}/\ell\mathbf{Z})^*$, $\sigma_a x - x = (a-1)w \in \mathcal{X}$, or $w \in \mathcal{X}$. Since $\mathrm{Hom}_{\mathcal{G}}(\mathcal{X}, E[\ell]) = \mathrm{Hom}_{\mathcal{G}}(i(\mathcal{X}), E[\ell])$ the proof again follows from Corollary 4.3. □

We are now ready to prove

Theorem 4.5. *In the exceptional case $\mathcal{G} = G_{\text{except}}$, we have $H^1(\mathcal{G}_i, E[\ell^i]) = 0$ for all $i \geq 1$.*

Proof. First, we deal with the case $i = 1$. As before, let $\tau := \binom{1\,1}{0\,1}$ and $\sigma_a = \binom{a\,0}{0\,1}$ be in $\mathcal{G}$ for some $a \in (\mathbf{Z}/\ell\mathbf{Z})^*$. Consider the inflation-restriction sequence

$$0 \longrightarrow H^1(\mathcal{G}/\langle\tau\rangle, E[\ell]^{\langle\tau\rangle}) \longrightarrow H^1(\mathcal{G}, E[\ell]) \longrightarrow H^1(\langle\tau\rangle, E[\ell])^{\mathcal{G}/\langle\tau\rangle}.$$

The group $H^1(\mathcal{G}/\langle\tau\rangle, E[\ell]^{\langle\tau\rangle})$ is zero since $|\mathcal{G}/\langle\tau\rangle|$ is prime to ℓ. It remains to show the vanishing of $H^1(\langle\tau\rangle, E[\ell])^{\mathcal{G}/\langle\tau\rangle}$.

Let $P = \binom{1}{0}$ and $Q = \binom{0}{1}$ be the chosen basis of $E[\ell]$. If $f : \langle\tau\,\rangle \longrightarrow E[\ell]$ is a cocycle, representing a cohomology class $[f]$ in $H^1(\langle\tau\rangle, E[\ell])$, the association $[f] \mapsto f(\tau)$ defines an isomorphism

$$H^1(\langle\tau\rangle, E[\ell]) \simeq \frac{\{X \in E[\ell] \mid (1 + \tau + \cdots + \tau^{\ell-1})X = O\}}{(1-\tau)E[\ell]}.$$

Since $1 + \tau + \cdots + \tau^{\ell-1} = \binom{0\,0}{0\,0}$ and $(1-\tau)E[\ell] = \langle P\rangle$, we have

$$H^1(\langle\tau\rangle, E[\ell]) \simeq E[\ell]/\langle P\rangle \simeq \langle Q\rangle.$$

Now it is sufficient to prove that the cohomology class ϕ represented by the cocycle $f : \tau \mapsto Q$ is not fixed by the action of σ_a for some $a \in (\mathbf{Z}/\ell\mathbf{Z})^*$.

Note that $(\sigma_a)^{-1}\tau\sigma_a = \tau^{\bar{a}}$ for some $\bar{a} \in (\mathbf{Z}/\ell\mathbf{Z})^*$ with $a\bar{a} = 1$. The cohomlogy class ϕ^{σ_a} is represented by the cocycle f^{σ_a}, which sends τ to

$$\begin{aligned} f^{\sigma_a}(\tau) &= \sigma_a f(\tau^{\bar{a}}) = \sigma_a(1 + \tau + \cdots + \tau^{\bar{a}-1})f(\tau) \\ &= \begin{pmatrix} a & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \bar{a} & \bar{a}(\bar{a}-1)/2 \\ 0 & \bar{a} \end{pmatrix} f(\tau) \\ &= \begin{pmatrix} 1 & (\bar{a}-1)/2 \\ 0 & \bar{a} \end{pmatrix} f(\tau) \\ &= \frac{\bar{a}-1}{2}P + \bar{a}Q \equiv \bar{a}Q \bmod \langle P\rangle. \end{aligned}$$

Therefore, $\phi \neq \phi^{\sigma_a}$ if $a \neq 1$. We proved $H^1(\langle\tau\rangle, E[\ell])^{\mathcal{G}/\langle\tau\rangle} = 0$.

Now, let $i \geq 1$. Consider the restriction map

$$\mathrm{Res} : H^1(\mathcal{G}_{i+1}, E[\ell^i]) \longrightarrow H^1(\mathrm{Gal}(L_{i+1}/L_i), E[\ell^i])^{\mathcal{G}_i} \simeq \mathrm{Hom}_{\mathcal{G}}(\mathcal{C}_i, E[\ell])$$

which appeared in the exact sequence (5). We claim that this map is trivial. Once this claim is verified, the theorem will follow from Lemma 2.2.

Now, let g be a cocycle, representing a cohomology class in $H^1(\mathcal{G}_{i+1}, E[\ell^i])$ and let $f = \operatorname{Res}(g) \in \operatorname{Hom}_{\mathcal{G}}(\mathcal{C}_i, E[\ell])$. By Proposition 4.4, we only need to show that $f(w) = 0$. Via the identification (8), the element w corresponds to the matrix

$$\begin{pmatrix} 1 & \ell^i \\ 0 & 1 \end{pmatrix}.$$

Let $I_i := \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$ be the (multiplicative) identity element in the ring $M_2(\mathbf{Z}/\ell^{i+1}\mathbf{Z})$ of 2×2 matrices with coefficients in $\mathbf{Z}/\ell^{i+1}\mathbf{Z}$. We will show in Lemma 4.6 that there exists $A \in \mathcal{G}_{i+1}$ such that $A^{\ell^i} = \left(\begin{smallmatrix} 1 & \ell^i \\ 0 & 1 \end{smallmatrix}\right)$ and that

$$I_i + A + A^2 + \cdots A^{\ell^i - 1} = \ell^i \cdot M$$

for some $M \in M_2(\mathbf{Z}/\ell^{i+1}\mathbf{Z})$. Using this lemma, we compute

$$\begin{aligned} g \begin{pmatrix} 1 & \ell^i \\ 0 & 1 \end{pmatrix} &= g\left(A^{\ell^i}\right) \\ &= (I_i + A + A^2 + \cdots A^{\ell^i - 1}) g(A) \\ &= \ell^i \cdot M\, g(A) \end{aligned}$$

But, the cocycle g takes values in $E[\ell^i]$, so $g\left(\begin{smallmatrix} 1 & \ell^i \\ 0 & 1 \end{smallmatrix}\right) = 0$, and hence $f(w) = 0$. □

Lemma 4.6. *For each $i \geq 1$, there exists $A \in \mathcal{G}_{i+1}$ such that*

(a) $A^{\ell^i} = \left(\begin{smallmatrix} 1 & \ell^i \\ 0 & 1 \end{smallmatrix}\right)$.

(b) *Let $I_i := \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$ be in the ring $M_2(\mathbf{Z}/\ell^{i+1}\mathbf{Z})$ of 2×2 matrices with coefficients in $\mathbf{Z}/\ell^{i+1}\mathbf{Z}$. Then, in $M_2(\mathbf{Z}/\ell^{i+1}\mathbf{Z})$, we have*

$$I_i + A + A^2 + \cdots A^{\ell i - 1} = \ell^i \cdot M$$

for some $M \in M_2(\mathbf{Z}/\ell^{i+1}\mathbf{Z})$.

Proof. When $i = 1$, we let

$$A = \begin{pmatrix} 1 + \ell p & 1 + \ell q \\ \ell r & 1 + \ell s \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} + \ell \cdot \begin{pmatrix} p & q \\ r & s \end{pmatrix}$$

in $\mathcal{G}_2 \subseteq \operatorname{GL}_2(\mathbf{Z}/\ell^2\mathbf{Z})$ be any lift of τ for some integers p, q, r and s.

We will prove that, for any $n \geq 1$,

$$A^n = \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix} + \ell \cdot \begin{pmatrix} np + \frac{n(n-1)}{2}r & a_n p + b_n q + c_n r + d_n s \\ nr & \frac{n(n-1)}{2}r + ns \end{pmatrix} \tag{10}$$

where the sequences a_n, b_n, c_n and d_n are defined as

$$\begin{aligned} a_n &= n(n-1)/2 & b_n &= n \\ c_n &= n(n-1)(n-2)/6 & d_n &= n(n-1)/2. \end{aligned}$$

This formula is clear for $n = 1$. Now, we prove this for $n \geq 1$. Note that the following computation is in $\mathcal{G}_2$, so any multiple of ℓ^2 is replaced by 0.

$$\begin{aligned}
A^n \cdot A &= \left\{ \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix} + \ell \cdot \begin{pmatrix} np + \frac{n(n-1)}{2}r & a_n p + b_n q + c_n r + d_n s \\ nr & \frac{n(n-1)}{2}r + ns \end{pmatrix} \right\} \\
&\quad \times \left\{ \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} + \ell \cdot \begin{pmatrix} p & q \\ r & s \end{pmatrix} \right\} \\
&= \begin{pmatrix} 1 & n+1 \\ 0 & 1 \end{pmatrix} + \ell \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix} \begin{pmatrix} p & q \\ r & s \end{pmatrix} \\
&\quad + \ell \begin{pmatrix} np + \frac{n(n-1)}{2}r & a_n p + b_n q + c_n r + d_n s \\ nr & \frac{n(n-1)}{2}r + ns \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \\
&= \begin{pmatrix} 1 & n+1 \\ 0 & 1 \end{pmatrix} + \ell \begin{pmatrix} p + nr & q + ns \\ r & s \end{pmatrix} \\
&\quad + \ell \begin{pmatrix} np + \frac{n(n-1)}{2}r & (np + \frac{n(n-1)}{2}r) + (a_n p + b_n q + c_n r + d_n s) \\ nr & nr + \frac{n(n-1)}{2}r + ns \end{pmatrix} \\
&= \begin{pmatrix} 1 & n+1 \\ 0 & 1 \end{pmatrix} \\
&\quad + \ell \begin{pmatrix} (n+1)p + \frac{n(n+1)}{2}r & (np + q + \frac{n(n-1)}{2}r + ns) + (a_n p + b_n q + c_n r + d_n s) \\ (n+1)r & \frac{n(n+1)}{2}r + (n+1)s \end{pmatrix}.
\end{aligned}$$

So, the equation (10) is proved if the sequences a_n, b_n, c_n and d_n satisfy

$$\begin{aligned}
a_{n+1} &= n + a_n, & b_{n+1} &= 1 + b_n \\
c_{n+1} &= \frac{n(n-1)}{2} + c_n & d_{n+1} &= n + d_n.
\end{aligned}$$

This is immediate from the definitions, and (10) follows.

In particular, when $n = \ell$, all of a_ℓ, b_ℓ, c_ℓ and d_ℓ are divisible by ℓ. (We note here that this is the only place where the assumption $\ell \neq 3$ is needed.) Hence, from (10),

$$A^\ell = \begin{pmatrix} 1 & \ell \\ 0 & 1 \end{pmatrix}$$

in $\mathcal{G}_2$. For (b), we use (10) to compute

$$\begin{aligned} I_0 + A + \cdots + A^{\ell-1} &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \cdots + \begin{pmatrix} 1 & \ell-1 \\ 0 & 1 \end{pmatrix} + \ell M \\ &= \ell \begin{pmatrix} 1 & (\ell-1)/2 \\ 0 & 1 \end{pmatrix} + \ell M \end{aligned}$$

for some $M \in M_2(\mathbf{Z}/\ell^2\mathbf{Z})$. We proved (b) for $i = 1$.

Assume that $i \geq 2$. Let $A \in \mathcal{G}_i$ be such that

$$A^{\ell^{i-1}} = \begin{pmatrix} 1 & \ell^{i-1} \\ 0 & 1 \end{pmatrix}$$

in $\mathcal{G}_i$, and such that

$$I_{i-1} + A + \cdots + A^{\ell^{i-1}-1} = \ell^{i-1} M$$

in $M_2(\mathbf{Z}/\ell^i\mathbf{Z})$ for some $M \in M_2(\mathbf{Z}/\ell^i\mathbf{Z})$.

Choose any lift $\hat{A} \in \mathcal{G}_{i+1}$ of A. Let $T := (\hat{A})^{\ell^{i-1}}$ in $\mathcal{G}_{i+1}$. Then, the projection of T in $\mathcal{G}_i$ is equal to $A^{\ell^{i-1}}$. Therefore, we have

$$T = \begin{pmatrix} 1 & \ell^{i-1} \\ 0 & 1 \end{pmatrix} + \ell^i \begin{pmatrix} p & q \\ r & s \end{pmatrix}$$

for some integers p, q, r and s. For $n \geq 1$, we will prove the following formula inductively.

$$T^n = \begin{pmatrix} 1 & n\ell^{i-1} \\ 0 & 1 \end{pmatrix} + \ell^i \cdot n \begin{pmatrix} p & q \\ r & s \end{pmatrix} \tag{11}$$

The case $n = 1$ is clear. In the following computation, we note that any multiple of ℓ^{2i-1} can be replaced by zero, because the computation is in $\mathcal{G}_{i+1}$.

$$\begin{aligned} T^n \cdot T &= \left\{ \begin{pmatrix} 1 & n\ell^{i-1} \\ 0 & 1 \end{pmatrix} + \ell^i \cdot n \begin{pmatrix} p & q \\ r & s \end{pmatrix} \right\} \left\{ \begin{pmatrix} 1 & \ell^{i-1} \\ 0 & 1 \end{pmatrix} + \ell^i \begin{pmatrix} p & q \\ r & s \end{pmatrix} \right\} \\ &= \begin{pmatrix} 1 & (n+1)\ell^{i-1} \\ 0 & 1 \end{pmatrix} + \ell^i \cdot n \begin{pmatrix} p & q \\ r & s \end{pmatrix} \begin{pmatrix} 1 & \ell^{i-1} \\ 0 & 1 \end{pmatrix} + \ell^i \begin{pmatrix} 1 & n\ell^{i-1} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} p & q \\ r & s \end{pmatrix} \\ &= \begin{pmatrix} 1 & (n+1)\ell^{i-1} \\ 0 & 1 \end{pmatrix} + \ell^i \left\{ n \begin{pmatrix} p & q \\ r & s \end{pmatrix} + \begin{pmatrix} p & q \\ r & s \end{pmatrix} \right\} \\ &= \begin{pmatrix} 1 & (n+1)\ell^{i-1} \\ 0 & 1 \end{pmatrix} + \ell^i \cdot (n+1) \begin{pmatrix} p & q \\ r & s \end{pmatrix}. \end{aligned}$$

The equation (11) is proved.

Now, take $n = \ell$. Then, we have

$$(\hat{A})^{\ell^i} = T^\ell = \begin{pmatrix} 1 & \ell^i \\ 0 & 1 \end{pmatrix}$$

in $\mathcal{G}_{i+1}$. The part (a) is proved.

It remains to prove (b). First, we note that

$$I_i + \hat{A} + (\hat{A})^2 + \cdots + (\hat{A})^{\ell^{i-1}-1} = \ell^{i-1}\hat{M}$$

for some $\hat{M} \in M_2(\mathbf{Z}/\ell^{i+1}\mathbf{Z})$. From (11), we have

$$\begin{aligned} I_i + T + T^2 + \cdots T^{\ell-1} &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 1 & \ell^{i-1} \\ 0 & 1 \end{pmatrix} + \cdots + \begin{pmatrix} 1 & (\ell-1)\ell^{i-1} \\ 0 & 1 \end{pmatrix} + \ell\hat{N} \\ &= \ell \begin{pmatrix} 1 & \ell^{i-1}(\ell-1)/2 \\ 0 & 1 \end{pmatrix} + \ell\hat{N} \\ &= \ell\hat{N}' \end{aligned}$$

for some $\hat{N}, \hat{N}' \in M_2(\mathbf{Z}/\ell^{i+1}\mathbf{Z})$. Therefore,

$$\begin{aligned} I_i + \hat{A} + (\hat{A})^2 + \cdots + (\hat{A})^{\ell^i-1} &= (I_i + T + T^2 + \cdots T^{\ell-1})(I_i + \hat{A} + (\hat{A})^2 + \cdots + (\hat{A})^{\ell^{i-1}-1}) \\ &= (\ell\hat{N}')(\ell^{i-1}\hat{M}) = \ell^i(\hat{N}'\hat{M}'). \end{aligned}$$

The lemma is proved. □

Remark 4.7. The assumption $\ell \neq 3$ is needed only in the proof of Lemma 4.6. We investigate the case $\ell = 3$ more closely here.

As in the proof, let $A \in \mathcal{G}_2$ be a lift of τ with

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} + \ell \cdot \begin{pmatrix} p & q \\ r & s \end{pmatrix}.$$

When $\ell = 3$, we have $a_3 = 3$, $b_3 = 3$, $c_3 = 1$ and $d_3 = 3$. So, from the equation (10),

$$A^3 = \begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix} + 3 \cdot \begin{pmatrix} 0 & r \\ 0 & 0 \end{pmatrix}.$$

If $r \equiv 0 \bmod 3$, the proof in the lemma works without any change. If $r \equiv 1 \bmod 3$, then we can replace A by A^{-1} and the rest of the proof works again. If all the lifts A of τ in $\mathcal{G}_2$ are such that $r \equiv -1 \bmod 3$, then the proof does not work. And, this is the only case that we don't have the vanishing of $H^1(\mathcal{G}_i, E[\ell^i])$.

4.2 An example

Let A and B be the elliptic curves defined by the equations

$$\begin{aligned} A: &\quad y^2 + y = x^3 - x^2 - 10x - 20 \\ B: &\quad y^2 + y = x^3 - x^2 - 7820x - 263580 \end{aligned}$$

and fix $\ell = 5$. These curves are denoted by 11A1 and 11A2 respectively in Cremona's table [1]. They are also studied by Vélu in [16].

The group of rational torsion points $A(\mathbf{Q})_{\text{tors}}$ of the curve A is isomorphic to $\mathbf{Z}/5\mathbf{Z}$, generated by the point $P = (5, 5)$. And, the curve B has no rational torsion. There is an isogeny over $\mathbf{Q}$

$$f : A \longrightarrow B$$

of degree 5, whose kernel is generated by the point P.

Crucial is the fact that the Galois group $\mathrm{Gal}(\mathbf{Q}(A[\ell])/\mathbf{Q})$ can be expressed in matrix form as

$$\begin{pmatrix} 1 & 0 \\ 0 & * \end{pmatrix} \tag{12}$$

with respect to the basis $\{P, Q\}$ with some *non-rational* ℓ-torsion point Q of A [14, §§5.5.2]. Take $R = f(Q) \in B[\ell]$ and complete a basis for $B[\ell]$ by adding another point $S \in B[\ell]$. We prove that $\mathcal{G} = \mathrm{Gal}(\mathbf{Q}(B[\ell])/\mathbf{Q})$ is isomorphic to G_{except} with respect to the basis $\{R, S\}$.

The character which fills in the lower right coefficient in (12) is nothing but the mod ℓ cyclotomic character χ_ℓ because of Weil pairing. Also, note that the point R spans a proper $\mathcal{G}$-submodule of $B[\ell]$. Therefore, $\mathcal{G}$ will be upper-triangular. With respect to the basis $\{R, S\}$, The group $\mathcal{G}$ is represented as

$$\begin{pmatrix} \chi_\ell & \beta \\ 0 & 1 \end{pmatrix}.$$

The lower-right 1 is again due to Weil pairing. Further, β is nontrivial, otherwise B would have some rational ℓ-torsion points. So, $\mathcal{G}$ is isomorphic to G_{except}.

5 Application

For this section, our elliptic curve E is assumed to have no complex multiplication, unless stated otherwise.

5.1 Extension of Kolyvagin's result on Ш(E/K)

Let $K = \mathbf{Q}(\sqrt{D})$ be an imaginary quadratic extension with fundamental discriminant $D \neq -3, -4$ where all prime divisors of N split. The point $y_K \in E(K)$ will denote the Heegner point associated with the maximal order in K. When y_K is of infinite order, m is defined to be the largest integer such that $y_K \in \ell^m E(K)$ modulo ℓ-torsion points.

By means of our Main Theorem obtained in §2–§4, we will prove Theorem 1.2 under the weaker assumption "$\rho_{\mathbf{Q}}$ irreducible", instead of "$\rho_{\mathbf{Q}}$ surjective".

Theorem 5.1. *Suppose that y_K is of infinite order. Assume that ℓ does not divide D and that E has a good or multiplicative reduction at ℓ. If the Galois representation*

$$\rho_{\mathbf{Q}} : \mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q}) \longrightarrow \mathrm{Aut}(E[\ell])$$

is irreducible over $\mathbf{Z}/\ell\mathbf{Z}$, then

$$\mathrm{ord}_\ell\, |\text{Ш}(E/K)| \leq 2m.$$

Proof. The prime ℓ is unramified in $K/\mathbf{Q}$. Therefore, a ramification argument shows that $K/\mathbf{Q}$ is linearly disjoint with $\mathbf{Q}(E[\ell])/\mathbf{Q}$. Hence $\rho_{\mathbf{Q}}$ is irreducible, (resp. surjective) if and only if ρ_K is irreducible (resp. surjective). Note that the irreducibility of $\rho_{\mathbf{Q}}$ implies that $E(K)$ has no ℓ-torsion points. So, Assumption 1.1 is satisfied with the prime ℓ and K.

In [8], the surjectivity assumption is needed only for the proof of Proposition 2 in loc. cit. Therefore, it suffices to prove Proposition 2 only under the irreducibility assumption.

We will follow the notations in [8]. For any natural number n,

$$[\ ,\ ]_n : E[\ell^n] \times E[\ell^n] \longrightarrow \mu_{\ell^n}$$

is the Weil pairing on level ℓ^n with values in the group μ_{ℓ^n} of ℓ^n-th roots of unity. The group $E[\ell^n]$ admits the decomposition

$$E[\ell^n] = E[\ell^n]^+ \oplus E[\ell^n]^-$$

with respect to the action of a complex conjugation. We may and will choose the generators e_n^+ and e_n^- of $E[\ell^n]^+$ and $E[\ell^n]^-$ respectively in a compatible manner for all $n \geq 1$. That is, $\ell \cdot e_n^+ = e_{n-1}^+$ and $\ell \cdot e_n^- = e_{n-1}^-$.

Fix $n' > n$, and let $V = K(E[\ell^{n'}])$. For any $g \in \mathrm{Gal}(V/\mathbf{Q})$, we let $\alpha(g) = 1$ if g restricts to the identity on K, and $\alpha(g) = -1$ otherwise. Note that any g acts on $E[\ell^n]$ via its restriction to $\mathbf{Q}(E[\ell^n])$.

Lemma 5.2. *Let P and Q be in $E[\ell^n]$. If $[P, ge_n^-]_n = [Q, ge_n^+]_n^{-\alpha(g)}$ for all $g \in \mathrm{Gal}(V/\mathbf{Q})$, then $P = Q = O$.*

Proof of Lemma 5.2. Induction on n. When $n = 1$, we have

$$[P, ge_1^-]_1 = [Q, ge_1^+]_1^{-\alpha(g)} \tag{13}$$

for all $g \in \mathrm{Gal}(V/\mathbf{Q})$. Recall that the extensions $K/\mathbf{Q}$ and $\mathbf{Q}(E[\ell])/\mathbf{Q}$ are linearly disjoint. Therefore, each $\sigma \in \mathrm{Gal}(\mathbf{Q}(E[\ell])/\mathbf{Q})$ can lift to $\tilde{g_1}$ and $\tilde{g_2}$ in $\mathrm{Gal}(K(E[\ell])/\mathbf{Q})$ in such a way that $\tilde{g_1}$ restricts to the identity on K and $\tilde{g_2}$ restricts to the unique nontrivial element in $\mathrm{Gal}(K/\mathbf{Q})$. Further, $\tilde{g_1}$ and $\tilde{g_2}$ can be lifted to g_1 and g_2 in $\mathrm{Gal}(V/\mathbf{Q})$. By construction, $\alpha(g_1) = 1$ and $\alpha(g_2) = -1$. Applying g_1 and g_2 in (13), we get

$$[P, \sigma e_1^-]_1 = [Q, \sigma e_1^+]_1 = 1.$$

By the irreducibility assumption, it follows that $\{\sigma e_1^-\}_{\sigma \in \mathrm{Gal}(\mathbf{Q}(E[\ell])/\mathbf{Q})}$ generates $E[\ell]$, hence $P = O$. Similarly, $Q = O$.

Let $n > 1$. By raising the equation $[P, ge_n^-]_n = [Q, ge_n^+]_n^{-\alpha(g)}$ to its ℓ-th power, we get $[\ell P, g(\ell e_n^-)]_{n-1} = [\ell Q, g(\ell e_n^+)]_{n-1}^{-\alpha(g)}$. Equivalently, we have

$$[\ell P, ge_{n-1}^-]_{n-1} = [\ell Q, ge_{n-1}^+]_{n-1}^{-\alpha(g)}$$

for all $g \in \mathrm{Gal}(V/\mathbf{Q})$. By the induction hypothesis, $\ell P = \ell Q = O$. Therefore P and Q are in $E[\ell] \subseteq E[\ell^n]$. From the compatibility of Weil pairing, we have $[P, ge_n^-]_n = [P, ge_1^-]_1$ and $[Q, ge_n^+]_n = [Q, ge_1^+]_1$. We are reduced to the case $n = 1$, hence the lemma follows. □

We proceed to prove Proposition 2 in [8], keeping the same notations. The homomorphism $f : H^1(K, E[\ell^n]) \longrightarrow H^1(V, \mu_{\ell^n})$ in [8] is defined by, for all $z \in \mathrm{Gal}(\bar{\mathbf{Q}}/V)$,

$$f(h) : z \longmapsto [h^+(z), e_n^-]_n^2 [h^-(z), e_n^+]_n^2$$

where $h = h^+ + h^- \in H^1(K, E[\ell^n])$ is the decomposition with respect to the complex conjugation. In the proof of Proposition 2 in loc. cit. , the surjectivity assumption is needed (and nowhere else) to prove that f is injective.

The equation (18) in loc. cit. says that

$$[h^+(z), ge_n^-]_n = [h^-(z), ge_n^+]_n^{-\alpha(g)}$$

for all $g \in \mathrm{Gal}(V/\mathbf{Q})$. From Lemma 5.2, it follows that $h^+(z) = h^-(z) = 0$ for all $z \in \mathrm{Gal}(\bar{\mathbf{Q}}/V)$. Therefore h is in the kernel of the restriction map

$$H^1(K, E[\ell^n]) \longrightarrow H^1(V, E[\ell^n]).$$

However, the kernel is equal to the cohomology group $H^1(\mathcal{G}_{n'}, E[\ell^n])$. The following lemma is an easy corollary of our Main Theorem, and it will finish the proof of Theorem 5.1. □

Lemma 5.3. *$H^1(\mathcal{G}_{n'}, E[\ell^n]) = 0$ for all $n' > n$.*

Proof of Lemma 5.3. The short exact sequence

$$0 \longrightarrow E[\ell^n] \longrightarrow E[\ell^{n'}] \xrightarrow{\times \ell^n} E[\ell^{n'-n}] \longrightarrow 0$$

yields the long exact $\mathcal{G}_{n'}$-cohomology sequence, part of which is

$$E[\ell^{n'-n}]^{\mathcal{G}_{n'}} \longrightarrow H^1(\mathcal{G}_{n'}, E[\ell^n]) \longrightarrow H^1(\mathcal{G}_{n'}, E[\ell^{n'}]).$$

The irreducibility assumption implies that $E(K)$ has no ℓ-torsion points. Therefore, we have $E[\ell^{n'-n}]^{\mathcal{G}_{n'}} = 0$. And our Main Theorem tells us that $H^1(\mathcal{G}_{n'}, E[\ell^{n'}]) = 0$. □

Corollary 5.4. *Suppose that y_K, D and ℓ are as in Theorem 5.1. If $\ell > 37$ then*

$$\mathrm{ord}_\ell\, |\text{Ш}(E/K)| \leq 2m.$$

Proof. It is known by the work of Mazur [11] that, for an elliptic curve E over $\mathbf{Q}$ with no CM, the Galois representation $\rho_{\mathbf{Q}}$ is always irreducible for all $\ell > 37$. □

Remark 5.5. In [8], Kolyvagin not only finds the bound of $\mathrm{ord}_\ell\, |\text{Ш}(E/K)|$ but also determines the complete group structure of the ℓ-part of $\text{Ш}(E/K)$ in terms of the (higher) Heegner points of E. This result also carries over *mutatis mutandis* only if we assume the irreducibility of $\rho_{\mathbf{Q}}$.

5.2 Irreducible vs surjective

For a fixed elliptic curve E over $\mathbf{Q}$, the set of primes ℓ where the mod ℓ Galois representation $\rho_{\mathbf{Q}}$ is not surjective is usually small, (see [14] and [9]) and, in many cases, this set is empty [2], [3]. However, if we vary E, there is no *universal* bound for ℓ known yet for which $\rho_{E,\ell}$ is surjective for all E. Corollary 5.4 can therefore be regarded as an improvement of Theorem 1.2 from a computational point of view.

A natural question is then to look for those E and ℓ's such that the associated representation

$$\rho_{E,\ell} : \mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q}) \longrightarrow \mathrm{GL}_2(\mathbf{Z}/\ell\mathbf{Z})$$

is irreducible, but not surjective. The rest of the section will be devoted to how one can hope to find such examples.

5.2.1 $\ell = 3$

Following Serre [14, §5.3], we study the case $\ell = 3$ closely. Let

$$y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6$$

be the minimal Weierstrass equation of E over $\mathbf{Z}$. Define, as usual, the following constants;

$$b_2 = a_1^2 + 4a_2, \qquad b_4 = a_1a_3 + 2a_4, \qquad b_6 = a_3^2 + 4a_6,$$
$$b_8 = a_1^2a_6 - a_1a_4a_4 + 4a_2a_6 + a_2a_3^2 - a_4^2 = (b_2b_6 - b_4^2)/4$$
$$c_4 = b_2^2 - 24b_4, \qquad c_6 = 36b_2b_4 - b_2^3 - 216b_6,$$
$$\Delta = b_4^3 - 27b_6^2 + b_8(36b_4 - b_2^2) = (c_4^3 - c_6^2)/1728, \qquad j = c_4^3/\Delta$$

Let $x_i (i = 1, 2, 3, 4)$ be the $x-$coordinates of the nonzero 3-torsion points $\pm P_i (i = 1, 2, 3, 4)$ respectively. They form the zeroes of the polynomial

$$f(x) = 3x^4 + b_2x^3 + 3b_4x^2 + 3b_6x + b_8.$$

Proposition 5.6. *Suppose that Δ is a cube in $\mathbf{Q}^*$. If $f(x)$ has at most one rational zero, then $\rho_{E,\ell}$ is irreducible but not surjective.*

Proof. One knows (see [14, §5.3]) that the order of $G_3 := \rho_{E,3}(\mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q}))$ is not divisible by 3 if and only if Δ is a cube in $\mathbf{Q}^*$. When this happens, the group G_3 is contained in a normalizer of a Cartan subgroup C of $\mathrm{GL}_2(\mathbf{Z}/3\mathbf{Z})$. If C is non-split, G_3 is necessarily irreducible and not surjective. In the case that C is split, G_3 is equal to C or its normalizer. In the former case, we see that G_3 is isomorphic to one of the two groups

$$\begin{pmatrix} \pm 1 & 0 \\ 0 & \pm 1 \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} \pm 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Both of these groups project onto the same image in $\mathrm{GL}_2(\mathbf{Z}/3\mathbf{Z})/\{\pm 1\} \simeq \mathcal{S}_4$. It is a cyclic group of order 2, leaving two elements fixed and switching the other two. This implies that G_3 fixes two roots of $f(x) = 0$. Hence $f(x)$ has two rational zeroes.

When G_3 is equal to a normalizer of C, one can find an element from the normalizer which exchanges the two subspaces which are stable under the action of C. [14, §2.2] In particular, this shows that $\rho_{E,3}$ is irreducible. □

Example 5.7. The hypothesis in the proposition above can be checked easily. For example, take

$$y^2 + y = x^3 - 7x + 12.$$

This is the curve $245A1$ in Cremona's table. The discriminant $\Delta = -42875 = -5^3 7^3$ and the polynomial $f(x)$ is

$$f(x) = 3x^4 + 0x^3 + 3(-14)x^2 + 3 \cdot 49x + (-49) = 3x^4 - 42x^2 + 147x - 49.$$

One easily sees that $f(x)$ is irreducible over $\mathbf{Q}$, so the above proposition applies.

5.2.2 $\ell = 3$ or 5

If one has a single example of E with an irreducible, non-surjective representation $\rho_{E,\ell}$ with $\ell = 3$ or 5, we can generate many other examples of such representations using the parametrization given by Rubin and Silverberg [13]. The parametrization gives (isomorphism classes of) elliptic curve E_t, indexed by almost all rational number t, with $E_t[\ell] \simeq E[\ell]$ as $\mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q})$ modules. Note that a CM curve will always provide with such an example.

5.2.3 $\ell > 5$

The strategy in the previous paragraph – to start with one example E and then to construct other curves E' with $E'[\ell] \simeq E[\ell]$ as $\mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q})$ modules – fails when ℓ is larger than 5;

indeed it was a question of Mazur (cf. [11], p133) to determine all such E'. See [5] for the case $\ell = 7$. Of course, the larger ℓ is, the harder to find a non surjective $\rho_{E,\ell}$.

References

[1] J. Cremona. *Algorithms for modular elliptic curves.* Cambridge University Press, Cambridge, second edition, 1997.

[2] W. Duke. Elliptic curves with no exceptional primes. *C. R. Acad. Sci. Paris Sér. I Math.*, 325(8):813–818, 1997.

[3] D. Grant. A formula for the number of elliptic curves with exceptional primes. *Compositio Math.*, 122(2):151–164, 2000.

[4] B. Gross and D. Zagier. Heegner points and derivatives of L-series. *Invent. Math.*, 84(2):225–320, 1986.

[5] E. Halberstadt and A. Kraus. On the modular curves $Y_E(7)$. *Math. Comp.*, 69(231):1193–1206, 2000.

[6] V. Kolyvagin. Finiteness of $E(\mathbf{Q})$ and $\mathrm{SH}(E, \mathbf{Q})$ for a subclass of Weil curves. *Izv. Akad. Nauk SSSR Ser. Mat.*, 52(3):522–540, 670–671, 1988.

[7] V. Kolyvagin. Euler systems. In *The Grothendieck Festschrift, Vol. II*, pages 435–483. Birkhäuser Boston, Boston, MA, 1990.

[8] V. Kolyvagin. On the structure of Shafarevich-Tate groups. In *Algebraic geometry (Chicago, IL, 1989)*, pages 94–121. Springer, Berlin, 1991.

[9] A. Kraus. Une remarque sur les points de torsion des courbes elliptiques. *C. R. Acad. Sci. Paris Sér. I Math.*, 321(9):1143–1146, 1995.

[10] S. Lang. *Algebra.* Addison-Wesley, Reading, Massachusetts, 3rd edition, 1993.

[11] B. Mazur. Rational isogenies of prime degree (with an appendix by D. Goldfeld). *Invent. Math.*, 44(2):129–162, 1978.

[12] K. Rubin. Tate-Shafarevich groups and L-functions of elliptic curves with complex multiplication. *Invent. Math.*, 89(3):527–559, 1987.

[13] K. Rubin and A. Silverberg. Families of elliptic curves with constant mod p representations. In *Elliptic curves, modular forms, & Fermat's last theorem (Hong Kong, 1993)*, pages 148–161. Internat. Press, Cambridge, MA, 1995.

[14] J.-P. Serre. Propriétés galoisiennes des points d'ordre fini des courbes elliptiques. *Invent. Math.*, 15(4):259–331, 1972.

[15] J. Silverman. *Advanced topics in the arithmetic of elliptic curves*, volume 151 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1994.

[16] J. Vélu. Courbes elliptiques sur $\mathbf{q}$ ayant bonne réduction en dehors de $\{11\}$. *C. R. Acad. Sci. Paris Sér. A-B*, 273:A73–A75, 1971.

VITAE

BYUNGCHUL CHA

Byungchul Cha was born in Seoul, Korea in August, 1972. He received his Bachelor of Science degree in mathematics from Korea Advanced Institute of Science and Technology in Taejon, Korea in 1994. He started his graduate studies at the Johns Hopkins University in 1997.