On torsion in $J_1(N)$, II

by

S. Kamienny (Los Angeles, CA)

1. Introduction. In [5] we studied the primes that may occur as the order of a rational torsion point on $J_1(N)$ defined over a number field of degree $d$. In this sequel we continue the study of torsion from a different point of view. We use ideas introduced by Serre [12], [13] and later used by Ribet [10], to show that the image of the Gal($\mathbb{Q}/\mathbb{Q}$)-representation on the kernel of a non-Eisenstein maximal ideal of the Hecke algebra is usually quite large (see §5 for a precise statement). We then apply this result, using a variation of an idea of Boxall and Grant [2], to study the almost rational torsion in quotients of $J_1(N)$.

We would like to thank K. Ribet for many helpful discussions.

2. The modular curve, and its jacobian. Let $N$ be a prime $\geq 13$, and let $X_1(N)$ denote the non-singular projective curve over $\mathbb{Q}$ associated to the moduli problem of classifying, up to isomorphism, pairs $(E, P)$ consisting of an elliptic curve $E$ together with a point $P$ of $E$ of order $N$. As usual, we denote by $X_0(N)$ the non-singular projective curve over $\mathbb{Q}$ whose non-cuspidal points classify isomorphism classes of pairs $(E, C)$, where $E$ is an elliptic curve, and $C$ is a cyclic subgroup of $E$ of order $N$.

The curve $X_1(N)$ is a cyclic cover of $X_0(N)$ whose covering group $\Delta$ is isomorphic to $(\mathbb{Z}/N\mathbb{Z})^*/(\pm 1)$. The covering map $\pi : X_1(N) \to X_0(N)$ is given, on non-cuspidal points, by $\pi(E, P) = (E, C_P)$, where $C_P$ is the subgroup of $E$ generated by the point $P$. We denote by $\langle a \rangle$ the element of $\Delta$ whose action on non-cuspidal points is given by $\langle a \rangle(E, P) = (E, aP)$.

The curve $X_0(N)$ has two cusps 0 and $\infty$, each rational over $\mathbb{Q}$. The cusps are unramified in the cover $\pi : X_1(N) \to X_0(N)$, so there are $N - 1$ cusps on $X_1(N)$. One half of these cusps lie above the cusp $0 \in X_0(N)$. These are called the 0-cusps of $X_1(N)$. The other half of the cusps lie above

2000 Mathematics Subject Classification: 11G18, 14G35.

Key words and phrases: modular curve, jacobian, Galois representation, almost rational torsion.

© Instytut Matematyczny PAN, 2008
the cusp \( \infty \in X_0(N) \). We call these the \( \infty \)-cusp of \( X_1(N) \). We work with a model of \( X_1(N) \) in which the 0-cusps are \( \mathbb{Q} \)-rational, while the \( \infty \)-cusps are rational in \( \mathbb{Q}(\zeta_N)^+ \), the maximal totally real subfield of \( \mathbb{Q}(\zeta_N) \).

We denote by \( J_1(N) \) (respectively, \( J_0(N) \)) the jacobian of the modular curve \( X_1(N) \) (respectively, \( X_0(N) \)). The abelian variety \( J_0(N) \) is semi-stable over \( \mathbb{Q} \) with bad reduction only at the prime \( N \). The abelian variety \( J_1(N) \) also has good reduction away from \( N \), and the quotient abelian variety \( A = J_1(N)/\pi^*(J_0(N)) \) attains everywhere good reduction over the field \( \mathbb{Q}(\zeta_N)^+ \). We can actually do a bit better than this. If \( d > 1 \) is a divisor of \((N - 1)/2\), we let \( J_d \) denote the quotient (by a connected subvariety) of \( J_1(N) \) associated to weight two newforms on \( \Gamma_1(N) \) whose nebentypus character has order \( d \). Then \( J_d \) attains everywhere good reduction over the unique subfield \( \mathbb{Q}_d \) of \( \mathbb{Q}(\zeta_N)^+ \) whose degree over \( \mathbb{Q} \) is \( d \).

We embed \( X_1(N) \) into \( J_1(N) \), sending a 0-cusp to \( 0 \in J_1(N) \). The divisor classes supported only at the 0-cusps generate a finite subgroup \( C \) of \( J_1(N) \) of order \( M = N \cdot \prod (1/2) \cdot B_{2,\varepsilon} \) (see [6]), where the product is taken over all even characters \( \varepsilon \) of \((\mathbb{Z}/N\mathbb{Z})^* \). The prime-to-2 part of the group \( J_1(N)(\mathbb{Q})_{\text{tors}} \) has order equal to the largest odd divisor of \( M \) (see [5]). The divisor classes supported only at the \( \infty \)-cusps also generate a subgroup \( C^* \) of order \( M \). The points of this group are rational in \( \mathbb{Q}(\zeta_N)^+ \).

3. The Hecke operators. The standard Hecke operators \( T_\ell \) (\( \ell \) a prime \( \neq N \)) and \( U_N \) act as correspondences on the curve \( X_1(N) \). They thus induce endomorphisms of the jacobian \( J_1(N) \). We define the Hecke algebra \( \mathbb{T} \) to be the ring of endomorphisms of \( J_1(N) \) generated over \( \mathbb{Z} \) by the \( T_\ell, U_N, \) and \( \triangle \). It is a commutative ring of finite type over \( \mathbb{Z} \), and all of its elements are defined over \( \mathbb{Q} \). The Hecke algebra \( \mathbb{T} \) induces an algebra (again denoted by \( \mathbb{T} \)) of endomorphisms of the quotients \( J_d \).

Since \( J_1(N) \) and \( J_d \) have good reduction away from the prime \( N \), their Néron models \( J_{1/S} \) and \( J_{d/S} \) over \( S = \text{Spec} \mathbb{Z}[1/N] \) are abelian schemes; we denote their fibers at \( \ell \) by \( J_{1/F_\ell} \) and \( J_{d/F_\ell} \), respectively. The fibers \( J_{1/F_\ell} \) and \( J_{d/F_\ell} \) inherit an action of the appropriate Hecke algebra \( \mathbb{T} \) from the induced action of \( \mathbb{T} \) on the Néron models. The Eichler–Shimura relation (see [14])

\[
T_\ell = \text{Frob}_\ell + \ell(\ell)/\text{Frob}_\ell
\]

holds in \( \text{End}(J_{1/F_\ell}) \) (respectively, \( \text{End}(J_{d/F_\ell}) \)). We can, as usual, lift this relation to the \( p \)-divisible group \( J_p(\mathcal{O}) \) (respectively, \( (J_d)_p(\mathcal{O}) \)), where \( p \) is any prime \( \neq \ell, N \) as well as to any étale subgroup of \( J_\ell(\mathcal{O}) \). Of course, in the lifted relation, \( \text{Frob}_\ell \) is any \( \ell \)-Frobenius automorphism in \( \text{Gal}(\mathcal{O}/\mathbb{Q}) \).

4. Maximal ideals of the Hecke algebra. The Hecke algebra \( \mathbb{T} \) preserves the cuspidal groups \( C \) and \( C^* \). The Eisenstein ideal \( I \) (respectively, \( I^* \)
On torsion in $J_1(N)$, II

is the annihilator in $T$ of $C$ (respectively, $C^*$). It contains all elements of the form $T_\ell - (1 + \ell(\ell))$, for all $\ell \neq N$ (respectively, $T_\ell - (\ell + \langle \ell \rangle)$). The maximal ideals $\mathcal{M}$ of $T$ in the support of $I$ or $I^*$ are called Eisenstein primes. The residue characteristics of the Eisenstein primes are precisely the prime divisors of the order $M$ of the cuspidal group $C$. There are clearly only a finite number of such ideals, and they are easily distinguished from the non-Eisenstein primes. A consequence of [5] is that if $P$ is a $\mathbb{Q}$-rational torsion point in $J_1(N)$ of prime order $p > 2$ then $p$ is a divisor of $M$, and $P$ is cuspidal, i.e., $P$ is annihilated by an Eisenstein prime.

From now on we write $J$ for one of $J_1(N)/\mathbb{Q}$ or $J_d/\mathbb{Q}$. We will mostly be concerned with non-Eisenstein maximal ideals of the appropriate Hecke algebra $T$. If $\mathcal{M}$ is such an ideal we assume that $\mathcal{M}$ is unramified in $T$. The set of ramified maximal ideals is finite, and is easily computable. The next proposition is well known (see [4], for example).

**Proposition 4.1.** Let $\mathcal{M}$ be an unramified prime of $T$ of residue characteristic $p$, and let $\mathbb{F}$ be the residue field $T/\mathcal{M}$. Then the following hold:

1. The $\mathcal{M}$-adic Tate module $\text{Ta}(\mathcal{M})$ is free of rank two over the $\mathcal{M}$-adic completion $T_\mathcal{M}$.
2. The kernel $J[\mathcal{M}]$ is free of rank two over $\mathbb{F}$.
3. If all primes $N$ of $T$ of residue characteristic $p$ are unramified then $T \otimes \mathbb{Z}_p \cong \prod T_N$, where the product is taken over all maximal ideals $N | p$.
4. If all primes $N$ of $T$ of residue characteristic $p$ are unramified then $T/pT \cong \prod T/N$, where the product is taken over all maximal ideals $N | p$.

5. **Galois representations.** In the following we assume that $\mathcal{M}$ is an unramified, non-Eisenstein maximal ideal of $T$. We also assume that the residue characteristic $p$ of $\mathcal{M}$ is > 5, and $\neq N$. We write $J[\mathcal{M}]$ for the group of $\mathcal{M}$-torsion points of $J$, and $\varrho_{\mathcal{M}} : \text{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \text{GL}_2(\mathbb{F})$ for the representation giving the natural action of the absolute Galois group on the $\mathcal{M}$-torsion points of $J$.

**Proposition 5.1.** The above assumptions about $\mathcal{M}$ and $p$ imply that the representation $\varrho_{\mathcal{M}}$ is irreducible.

**Proof.** We may assume that $\mathcal{M}$ is not an ideal of the Hecke algebra $T$ associated to $J_0(N)$ since Mazur [7] has proved the irreducibility in this case. Now, if $J[\mathcal{M}]$ is reducible, we let $\mathcal{L}$ be a line fixed by $\text{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$. We let $\mathcal{O}$ be the ring of integers of $\mathbb{Q}(\zeta_N)^+$, and let $J/\mathcal{O}$ be the Néron model of $J/\mathbb{Q}(\zeta_N)^+$ over $\text{Spec} \mathcal{O}$. Let $\mathcal{G}$ be the Zariski closure of $\mathcal{L}$ in the kernel of $\mathcal{M}$ on $J/\mathcal{O}$. Then it follows from [8] that $\mathcal{G}$ is either $(\mathbb{Z}/p\mathbb{Z})_\mathcal{O}^f$ or $\langle \mu_p \rangle_\mathcal{O}^f$, where
$f$ is the residue class degree of $M$. In either case the arguments of [5] show that $M$ is Eisenstein, contrary to assumption.

The Eichler–Shimura relation shows that $\det(\varrho_M(\text{Frob}_\ell)) = \ell \cdot \varepsilon(\ell)$, where $\varepsilon$ is an even character of $(\mathbb{Z}/N\mathbb{Z})^*$ through which $\Delta$ acts on $J[M]$. It will be important to note that $\varepsilon$ is unramified outside of $N$. We may thus view the character $\det(\varrho_M) = \ell \cdot \varepsilon(\ell)$, where $\varepsilon$ is an even character of $(\mathbb{Z}/N\mathbb{Z})^*$ through which $\Delta$ acts on $J[M]$. It will be important to note that $\varepsilon$ is unramified outside of $N$. We may thus view the character $\det(\varrho_M) = \chi \cdot \varepsilon$, where $\chi$ is the $p$-cyclotomic character (which is, of course, unramified outside of $p$). Since $\chi$ is an odd character (i.e., $\chi(c) = -1$, where $c$ is complex conjugation), and $\varepsilon$ is even, we see that $\det(\varrho_M)$ is also odd.

Now let $I = I_\ell$ be a $\ell$-inertia subgroup of $\text{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$, and write $\varrho$ for $\varrho_M$. The semi-simplification of $\varrho|_I$ is described by a pair of characters $\phi, \phi^* : I \rightarrow F^*$. Since $\det(\varrho|_I) = \chi$, the cyclotomic character, we must have $\phi \cdot \phi^* = \chi$. Moreover, since the weight (see [13]) of the representation $\varrho$ is 2 it follows that either (1) $\phi$ or $\phi^*$ is $\chi$ (and the other one is trivial), or (2) $\phi, \phi^*$ are the fundamental characters of level two (see [13]). It follows, in either case, that $\varepsilon(\ell)$ is of order $p^\pm 1$.

**Proposition 5.2.** Suppose that the order of $\varepsilon$ is odd. Assume that $M$ is an unramified, non-Eisenstein maximal ideal of residue characteristic $p > 5$. Then the image of $\varrho$ has order divisible by $p$.

**Proof.** Let $G = \varrho(\text{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}))$, and assume that $p$ does not divide the order of $G$. Let $\overline{G}$ be the image of $G$ in $\text{PGL}_2(F)$. Since $p > 5$ we have the following possibilities (see [12]): $\overline{G}$ is either cyclic, dihedral, or one of the exceptional groups $S_4, A_4, A_5$.

If $\overline{G}$ is cyclic then $G$ is abelian, which contradicts the irreducibility of $\varrho$. Thus, the case where $\overline{G}$ is cyclic does not occur.

Let $\overline{I}$ be the image of $I$ in $\overline{G}$. Then $\overline{I}$ is cyclic since it may be viewed as the image of $\phi^* \cdot \phi^{-1}$ (see [13]), which is a finite subgroup of $\overline{F}^*$. Moreover, the order of $\overline{I}$ is $p \pm 1$. If $p \geq 7$ this rules out the exceptional groups $S_4, A_4,$ and $A_5$ since none of these groups has an element of order $p \pm 1$.

This leaves only the possibility that $\overline{G}$ is dihedral. Suppose that this is indeed the case. Let $\overline{C}$ be the large cyclic subgroup of $\overline{G}$. Then $\overline{I}$ is contained in $\overline{C}$ since $\overline{I}$ is cyclic, and of order $> 2$. The quadratic extension $L$ of $\mathbb{Q}$ corresponding to $\overline{C}$ is thus unramified at $p$. It follows that only the prime $N$ can ramify in $L$, so that $L$ must be the quadratic subfield of $\mathbb{Q}(\zeta_N)$. However, since the order of $\varepsilon$ is odd the ramification degree of $N$ in $L$ must also be odd. Indeed, let $d$ denote the order of $\varepsilon$. The module $J[M]$ may be realized as a module of torsion points on the quotient $J_d$, an abelian variety that attains everywhere good reduction over the field $\mathbb{Q}_d$. It follows immediately that the ramification degree of $N$ must be odd, as claimed. This shows that $L$ is an everywhere unramified extension of $\mathbb{Q}$, which is an obvious contradiction. Thus, $\overline{G}$ is not dihedral, and $p$ must divide the order of the image of $\varrho$, as desired.
On torsion in $J_1(N)$, II

Remark. (1) If the order of $\varepsilon$ is not divisible by 2 or 3 then we may also include $p = 5$ in Proposition 5.2 as we can then conclude that the exceptional groups $S_4$, $A_4$, and $A_5$ do not occur. To see this note that $\bar{G}$ is either $S_4$, or $A_4$ since its order is prime to 5. In fact, $\bar{G}$ must be $S_4$ since $I$ must be cyclic of order 4, and $A_4$ has no elements of order 4. We consider the $S_3$-extension $K$ of $\mathbb{Q}$ arising from the quotient $S_3$ of $S_4$. Only the primes 5 and $N$ can ramify in $K$, and the ramification degree of 5 must be 2. Since the order of $\varepsilon$ is not divisible by 2 or 3 we see that $N$ must be unramified in $K$. However, this means that $K$ is an everywhere unramified extension of $\mathbb{Q}(\sqrt{5})$, which is impossible since $\mathbb{Q}(\sqrt{5})$ has class number one.

(2) If the order of $\varepsilon$ is even then $N \equiv 1 \pmod{4}$. In that case the quadratic subfield of $\mathbb{Q}(\zeta_N)$ is a real quadratic field. If we could show that the action of complex conjugation on the quadratic field $L$ was non-trivial then we would again have a contradiction showing that $\bar{G}$ cannot be a dihedral group. We can sometimes do this by mimicking [12] as follows. If $G$ is dihedral then $\rho(G)$ is contained in the normalizer of a Cartan subgroup, but not in the Cartan subgroup itself. If the Cartan subgroup is non-split then complex conjugation $c$ must act non-trivially on $L$ since $\pm 1$ are the only involutions in a non-split Cartan subgroup (so the image of $c$ in $\bar{G}$ falls outside of the large cyclic subgroup $C$). If we can prove that $\rho(G)$ is never contained in the normalizer of a split Cartan subgroup then we can eliminate the hypothesis, in Proposition 5.2, that the order of $\varepsilon$ is odd.

Corollary 5.3. Let $\mathcal{M}$ be an unramified, non-Eisenstein maximal ideal of $\mathcal{T}$ of residue characteristic $p > 5$, and $\neq N$. Assume that the nebentypus character $\varepsilon$ associated to $\mathcal{M}$ has odd order. If $\rho$ is the representation of $\text{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ on the $\mathcal{M}$-torsion points of $J$ then $\text{Im}(\rho)$ contains a subgroup isomorphic to $\text{SL}_2(\mathbb{F}_p)$.

Proof. We closely follow Serre [12, 2.4], and Ribet [10, Corollary 2.3]. Let $G = \rho(\text{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}))$, and let $\sigma \in G$ be an element of order $p$. If $v$ is a non-zero vector in $\mathbb{F} \oplus \mathbb{F}$ that is fixed by $\sigma$ then there is a $\tau \in \bar{G}$ such that $v$ and $\tau v$ form a basis of $\mathbb{F} \oplus \mathbb{F}$ (since the irreducibility of $\rho$ means that $G$ cannot fix the one-dimensional subspace spanned by $v$). Then the matrix of $\rho$ with respect to the basis $\{v, \tau v\}$ is of the form $A(\alpha) = \begin{pmatrix} 1 & 0 \\ 0 & \alpha \end{pmatrix}$, and the matrix of $\tau \sigma \tau^{-1}$ is of the form $B = \begin{pmatrix} 1 & 0 \\ \beta & 1 \end{pmatrix}$. Multiplying by an appropriate scalar we may assume that $\alpha = 1$. It is well known that the group generated by $A(1)$ and $B$ contains $\text{SL}_2(\mathbb{F}_p)$.

We fix an odd divisor $d$ of $(N - 1)/2$, and work on the abelian variety $J_d$. We write $\mathcal{M}$ for a maximal ideal of $\mathcal{T}$, and $\mathcal{T}_\mathcal{M}$ for its completion. We continue to assume that $\mathcal{M}$ is unramified, so $\mathcal{T}_\mathcal{M} \otimes \mathbb{Q}$ is a finite unramified extension of $\mathbb{Q}_p$, and $\mathcal{T}_\mathcal{M}$ is a discrete valuation ring.
We write $\mathcal{R} = \mathcal{R}_M$ for the representation $\text{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \text{GL}_2(\mathbb{T}_M)$ giving the action of $\text{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ on the $M$-adic Tate module of $J_d$. It follows from the Eichler–Shimura relation that the determinant of $\mathcal{R}$ is $\chi \varepsilon$, where $\chi$ is the $p$-adic cyclotomic character $\chi : \text{Gal}(\mathbb{Q}/\mathbb{Q}) \to \mathbb{Z}_p^*$, and $\varepsilon$ is a character cutting out the field $\mathbb{Q}_d$ ($\varepsilon$ takes values in an unramified extension of $\mathbb{Z}_p$). Corollary 5.3 implies (as in Lemma 3 of [11, IV-23]) that the image of $\mathcal{R}$ contains $\text{SL}_2(\mathbb{Z}_p)$.

6. Almost rational torsion. Ribet (see [1] and [9]) has introduced the notion of almost rational torsion points on an abelian variety. He used this idea to give a new and beautiful proof of the Manin–Mumford conjecture. It also became immediately useful in proving the conjecture of Coleman, Kaskel, and Ribet (see [3]) that only the cusps and hyperelliptic branch points of $X_0(N)$ give rise to torsion points when the curve is embedded in its jacobian. We recall the definition and basic properties of almost rational torsion points here. Let $A$ be an abelian variety over a field $K$. A point $P$ in $A(\overline{K})$ is called almost rational over $K$ if, for all $\sigma, \tau \in \text{Gal}(\overline{K}/K)$, the equation $\sigma(P) + \tau(P) = 2P$ holds if and only if $P = \sigma(P) = \tau(P)$. Certainly, any rational point is almost rational, as is any Galois conjugate of an almost rational point. More important for us is the following.

**Lemma 6.1.**

(a) If $P$ is almost rational over $K$, and $\sigma \in \text{Gal}(\overline{K}/K)$ is such that $(\sigma - 1)^2 \cdot P = 0$, then $\sigma$ fixes $P$.

(b) If $L$ is an extension of $K$ contained in $\overline{K}$, and $P$ is almost rational over $K$, then $P$ is almost rational over $L$.

**Proof.** (a) To see this one calculates $(\sigma - 1)^2 \cdot P = \sigma^2(P) - 2\sigma(P) + P = 0$. Applying $\sigma^{-1}$ we see that $\sigma(P) + \sigma^{-1}(P) = 2P$. Since $P$ is almost rational we must have $\sigma(P) = \sigma^{-1}(P) = P$.

(b) is clear from the definition of almost rational.

We wish to describe the primes $p$ for which there exists an almost rational torsion point of order $p^\alpha$ on $J_d$. Let $S$ be the set of all primes $q$ such that if $M$ is a maximal ideal of $\mathbb{T}$ of residue characteristic $q$ then $\mathcal{R}_M(\text{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}))$ does not contain $\text{SL}_2(\mathbb{Z}_q)$. At worst $S$ contains the primes 2, 3, 5, $N$, the prime divisors of $M$, and the residue characteristics of those $M$ that are ramified in $\mathbb{T}$. Let $p \notin S$ be a prime, and suppose that there exists an almost rational point $P$ of order $p^\alpha$ on $J_d$. If we write $\mathcal{R}_p$ for the $\text{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$-representation on the $p$-adic Tate module, then, by the remarks at the end of §5, we know that $\mathcal{R}_p(\text{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}))$ contains a subgroup $G$ isomorphic to $\prod \text{SL}_2(\mathbb{Z}_p)$ (for us even $\prod \text{SL}_2(\mathbb{Z})$ will suffice), where the product is taken over all maximal ideals of $\mathbb{T}$ of residue characteristic $p$. We let $K$ be the extension of $\mathbb{Q}$ such that $\mathcal{R}_p(\text{Gal}(\overline{K}/K)) = G$. 
By Lemma 6.1(b) the point $P$ is almost rational over $K$. If $\sigma \in \text{Gal}(\overline{K}/K)$ is such that $R_p(\sigma)$ is an element all of whose components in $G \approx \prod \text{SL}_2(\mathbb{Z}_p)$ are transvections then $(\sigma - 1)^2 \cdot P = 0$. Since $P$ is almost rational, Lemma 6.1(a) tells us that $\sigma(P) = P$. Since $P$ is fixed by all such $\sigma$, we see that $P$ must be 0. We have thus proved the following.

**Theorem 6.2.** Let $N$ be a prime $\geq 13$, and let $d$ be an odd divisor of $(N-1)/2$. If there is an almost rational point $P$ on $J_d$ of prime power order $p^\alpha$ then $p \in \mathcal{S}$.

**References**


Department of Mathematics
University of Southern California
3620 South Vermont Ave.
Los Angeles, CA 90089-2532, U.S.A.
E-mail: kamienny@usc.edu

Received on 4.12.2007
and in revised form on 7.3.2008