TORSION POINTS ON $J_0(N)$ AND GALOIS REPRESENTATIONS

Kenneth A. Ribet

University of California, Berkeley

To Barry Mazur, for his 60th birthday

Suppose that N is a prime number greater than 19 and that P is a point on the modular curve $X_0(N)$ whose image in $J_0(N)$ (under the standard embedding $\iota\colon X_0(N) \hookrightarrow J_0(N)$) has finite order. In [2], Coleman-Kaskel-Ribet conjecture that either P is a hyperelliptic branch point of $X_0(N)$ (so that $N \in \{ 23, 29, 31, 41, 47, 59, 71 \}$) or else that $\iota(P)$ lies in the cuspidal subgroup C of $J_0(N)$. That article suggests a strategy for the proof: assuming that P is not a hyperelliptic branch point of $X_0(N)$, one should show for each prime number ℓ that the ℓ-primary part of $\iota(P)$ lies in C. In [2], the strategy is implemented under a variety of hypotheses but little is proved for the primes $\ell = 2$ and $\ell = 3$. Here I prove the desired statement for $\ell = 2$ whenever N is prime to the discriminant of the ring $\operatorname{End} J_0(N)$. This supplementary hypothesis, while annoying, seems to be a mild one; according to W. A. Stein of Berkeley, California, in the range $N < 5021$, it false only in case $N = 389$.

1. Introduction

At the C.I.M.E. conference on the arithmetic of elliptic curves, I lectured on interrelated questions with a common underlying theme: the action of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ on torsion points of semistable abelian varieties over $\mathbf{Q}$. In this written record of my lectures, I focus on the modular curve $X_0(N)$ and its Jacobian $J_0(N)$ when N is a prime number. In this special case, $X_0(N)$ and $J_0(N)$ were studied intensively by B. Mazur in [9] and [10], so that we have a wealth of arithmetic information at our disposal.

The main theorem of this article complements the results of Coleman-Kaskel-Ribet [2] on the "cuspidal torsion packet" of $X_0(N)$. Recall that $X_0(N)$ has two cusps, customarily denoted 0 and ∞. Selecting the latter cusp as the more "standard" of the two, we use it to map $X_0(N)$ to $J_0(N)$, via the Albanese mapping ι which takes a point P of the curve to the class of the divisor $(P) - (\infty)$. This map is injective if the genus of $X_0(N)$ is non-zero.

Let g be the genus of $X_0(N)$. For the remainder of this preliminary discussion, make the hypothesis $g \geq 2$. (This hypothesis is satisfied if and only if $N \geq 23$.) Then ι

The author's research was partially supported by National Science Foundation contract #DMS 96 22801. The author thanks M. Baker, J. A. Csirik and H. W. Lenstra, Jr. for helpful conversations and suggestions.

identifies $X_0(N)$ with a subvariety of $J_0(N)$ of positive codimension. The torsion packet in question is the set Ω of points of $X_0(N)$ whose images in $J_0(N)$ have finite order. According to the Manin-Mumford conjecture, first proved by Raynaud in 1983 [13], Ω is a finite set.

The article [2] introduces a strategy for identifying Ω precisely. Clearly, Ω contains the two cusps 0 and ∞ of $X_0(N)$, whose images under ι have order $n := \operatorname{num}\left(\frac{N-1}{12}\right)$ and 1, respectively [9, p. 98]. Further, in the special case when $X_0(N)$ is hyperelliptic, we note in [2] that the hyperelliptic branch points of $X_0(N)$ belong to Ω if and only if N is different from 37. (Results of Ogg [11, 12] show that $X_0(N)$ is hyperelliptic if and only if N lies in the set $\{\,23, 29, 31, 37, 41, 47, 59, 71\,\}$.) In fact, suppose that $X_0(N)$ is hyperelliptic and that P is a hyperelliptic branch point on $X_0(N)$. Then $2\iota(P) = \iota(0)$ if $N \neq 37$, but P has infinite order when $N = 37$.

In [2], we advance the idea that Ω might contain only of the points we have just catalogued:

Guess 1.1. *Suppose that P is a point on $X_0(N)$ whose image in $J_0(N)$ has finite order. Then either P is one of the two cusps of $X_0(N)$, or $X_0(N)$ is a hyperelliptic curve and P is a hyperelliptic branch point of $X_0(N)$.*

In the latter case, (i.e., $X_0(N)$ hyperelliptic and P a hyperelliptic branch point with finite order in $J_0(N)$), it follows automatically that N is different from 37.

A reformulation of Guess 1.1 involves the cuspidal subgroup C of $J_0(N)$, i.e., the group generated by the point $\iota 0$. As we point out in [2], the results of [10] imply that the intersection of $X_0(N)$ and C (computed in $J_0(N)$) consists of the two cusps 0 and ∞. In words, to prove that a torsion point P of $X_0(N)$ is a cusp is to prove that it lies in the group C. For this, it is useful to decompose P into its primary parts: If P is a torsion point P of $J_0(N)$ and ℓ is a prime number, we let P_ℓ be the ℓ-primary part of P. Thus $P = \sum P_\ell$, the sum being extended over all primes, and we have $P \in C$ if and only if $P_\ell \in C$ for all primes ℓ.

Consider the following two statements (in both, we regard $X_0(N)$ as embedded in its Jacobian via ι):

Statement 1.2. *Suppose that P is an element of Ω and that ℓ is an odd prime. Then we have $P_\ell \in C$.*

Statement 1.3. *Suppose that P is an element of Ω and that $P_2 \notin C$. Then P is a hyperelliptic branch point of $X_0(N)$.*

It is clear that Guess 1.1 is equivalent to the conjunction of Statements 1.2 and 1.3. Indeed, suppose first that (1.1) is correct and that P is an element of Ω. If P is a cuspidal point (i.e., one of 0, ∞), then one has $P_\ell \in C$ for all primes ℓ. If P is not a cuspidal point, then P is a hyperelliptic branch point and $N \neq 37$; we then have $2P \in C$, so that $P_\ell \in C$ for all $\ell > 2$. Conversely, suppose that Statements 1.2 and 1.3 are true and that P is an element of Ω. If P_2 is not in C, then P is a hyperelliptic branch point (and is thus accounted for by the guess). If P_2 lies in C, then P_ℓ is in C for all primes ℓ, so that P is a point of C. As was mentioned above, this implies that P is one of the two cuspidal points on $X_0(N)$.

Our article [2] proves a number of results in the spirit of (1.2). For example, suppose that P is an element of Ω and ℓ is an odd prime different from N. Let g again be the genus of $X_0(N)$. Then $P_\ell \in C$ if ℓ is greater than $2g$ or if ℓ satisfies $5 \leq \ell < 2g$ and at least one of a number of supplementary conditions.

These notes prove a theorem in the direction of (1.3). This theorem requires an auxiliary hypothesis concerning the discriminant of the subring $\mathbf{T}$ of $\operatorname{End} J_0(N)$ which is generated by the Hecke operators T_m (with $m \geq 1$) on $J_0(N)$. (Many authors write the Hecke operator T_N as U_N.) According to [9, Prop. 9.5, p. 95], the Hecke ring $\mathbf{T}$ is in fact the full endomorphism ring of $J_0(N)$. Concerning the structure of $\mathbf{T}$, it is known that $\mathbf{T}$ is an order in a product $E = \prod E_t$ of totally real number fields. The discriminant $\operatorname{disc}(\mathbf{T})$ is the product of the discriminants of the number fields E_i, multiplied by the square of the index of $\mathbf{T}$ in its normalization. Our auxiliary hypothesis is the following statement:

Hypothesis 1.4. The discriminant of $\mathbf{T}$ is prime to N.

According to William Arthur Stein of Berkeley, California, Hypothesis 1.4 is false when $N = 389$ and true for all other primes $N \leq 5011$.

Theorem 1.5. *Suppose that P lies in Ω and that P_2 does not belong to C. In addition, suppose either that the order of P is prime to N or that Hypothesis 1.4 holds. Then $X_0(N)$ is hyperelliptic, and P is a hyperelliptic branch point of $X_0(N)$.*

Theorem 1.5 is a direct consequence of a Galois-theoretic statement which we prove in §7. Since this latter theorem is the main technical result of these notes, we state it now and then show how it implies Theorem 1.5.

Theorem 1.6. *Let N be a prime number, and let $J = J_0(N)$. Let ℓ be a prime different from N. Suppose that P is a point of finite order on $J_0(N)$ whose ℓ-primary component P_ℓ is not defined over $\mathbf{Q}$. Assume that at least one of the following hypotheses holds: (1) N is prime to the order of P; (2) ℓ is prime to $N-1$; (3) N is prime to the discriminant of $\mathbf{T}$. Then there is a $\sigma \in \operatorname{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ such that $\sigma P - P$ has order ℓ.*

Note that the hypothesis $g \geq 2$ is not needed for Theorem 1.6.

Proof that (1.6) implies (1.5). Let P be as in Theorem 1.5. Because P_2 does not lie in C, P_2 is not a rational point of $J_0(N)$ [9, Ch. III, Th. 1.2]. We apply Theorem 1.6 in this situation, taking $\ell = 2$. The theorem shows that there is a $\sigma \in \operatorname{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ such that the divisor $(\sigma P) - (P)$ on $X_0(N)$ has order 2 in $J_0(N)$. Accordingly, the points P and σP are distinct, and there is a rational function f on the curve $X_0(N)$ whose divisor is $2\big((\sigma P) - (P)\big)$. The function f has a double zero at σP, a double pole at P, and no other zeros or poles. It follows that the covering $X \to \mathbf{P}^1$ defined by f is of degree two and that P is ramified in the covering. Since the genus of X is at least 2, it follows that X is hyperelliptic and P is a hyperelliptic branch point. ■

We conclude this discussion with a second statement which will be proved only below. For this statement and for most of what follows, we again allow N be an arbitrary prime; i.e., we have no need of the assumption that $J = J_0(N)$ has dimension > 1. As in [9], we consider the Eisenstein ideal $\mathscr{I} \subseteq \mathbf{T}$ and form the kernel $J[\mathscr{I}] \subseteq J(\overline{\mathbf{Q}})$. Let $K = \mathbf{Q}(J[\mathscr{I}])$ be the field generated by the coordinates of the points in $J[\mathscr{I}]$. Recall that $n = \operatorname{num}\left(\frac{N-1}{12}\right)$. Then we have:

Theorem 1.7. *The field K is the field of $2n^{\text{th}}$ roots of unity.*

Theorem 1.7 is an essential ingredient in our proof of Theorem 1.6 in the crucial case where $\ell = 2$. Readers who are familiar with Mazur's article [9] will recognize that Theorem 1.7 follows directly from the results of that article if n is not divisible by 4. Moreover, as H. W. Lenstra, Jr. has pointed out, Theorem 1.7 may be proved rather easily by elementary arguments if n is divisible by 8. The most difficult case is therefore that for which n is divisible by 4 but not by 8; this case occurs precisely when $N \equiv 17 \bmod 32$. We will discuss Lenstra's observations in §4 and then prove Theorem 1.7 in the general case in §5 by exploiting Mazur's "congruence formula for the modular symbol" [9, Ch. II, §18]. An alternative proof of Theorem 1.7 was given recently by J. A. Csirik [3]. Csirik provides a complete concrete description of $J_0(N)[\mathscr{I}]$ which yields Theorem 1.7 as a corollary.

2. A local study at N

For the rest of this article, we take N to be a prime number and let $J = J_0(N)$. The assumption of §1 concerning the genus of $X_0(N)$ is no longer required.

We remind the reader that the results of Deligne and Rapoport [4] imply that J has purely multiplicative reduction at N. As explained in the Mazur-Rapoport appendix to [9], the fiber over $\mathbf{F}_N$ of the Néron model of J is the product of a cyclic component group Φ and a torus $J^0_{/\mathbf{F}_N}$.

The character group of this torus,

$$\mathscr{X} := \operatorname{Hom}_{\overline{\mathbf{F}}_N}\left(J^0_{/\mathbf{F}_N}, \mathbf{G}_{\mathrm{m}}\right),$$

is a free $\mathbf{Z}$-module of rank $\dim J$ which is furnished with compatible actions of $\mathbf{T}$ and the Galois group $\operatorname{Gal}(\overline{\mathbf{F}}_N/\mathbf{F}_N)$. Here, $\overline{\mathbf{F}}_N$ is of course an algebraic closure of the prime field $\mathbf{F}_N$. It will be convenient to choose a prime dividing N in $\overline{\mathbf{Q}}$ and to let $\overline{\mathbf{F}}_N$ be the residue field of this prime. Then if $D \subset \operatorname{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ is the decomposition group corresponding to the chosen prime, $\operatorname{Gal}(\overline{\mathbf{F}}_N/\mathbf{F}_N)$ is the quotient of D by its inertia subgroup I. Using the quotient map $D \to \operatorname{Gal}(\overline{\mathbf{F}}_N/\mathbf{F}_N)$, we view $\mathscr{X}$ as an unramified representation of D. As one knows, this action is "nearly" trivial: the generator $x \mapsto x^N$ of $\operatorname{Gal}(\overline{\mathbf{F}}_N/\mathbf{F}_N)$ acts on $\mathscr{X}$ as an automorphism of order 1 or 2, so that the group $\operatorname{Gal}(\overline{\mathbf{F}}_N/\mathbf{F}_{N^2})$ acts trivially on $\mathscr{X}$. (The group $\mathscr{X}$ is discussed in [14, §3] in the more general case where N is replaced by the product of a prime q and a positive integer which is prime to q.)

As far as the Hecke action goes, the group $\mathscr{X}$ is a free $\mathbf{Z}$-module whose rank is the same as that of $\mathbf{T}$, namely the dimension of J. Because $\mathbf{T}$ acts faithfully on $\mathscr{X}$, it is clear that $\mathscr{X} \otimes \mathbf{Q}$ is free of rank 1 over $\mathbf{T} \otimes \mathbf{Q}$. Thus $\mathscr{X}$ is a "$\mathbf{T}$-module of rank 1" in the sense of [9, Ch. II, §8]. (In fact, in [9, Ch. II, Prop. 8.3], Mazur notes in effect that $\mathscr{X} \otimes \mathbf{Q}_p$ is free of rank 1 over $\mathbf{T} \otimes \mathbf{Q}_p$ for each prime $p \neq N$.) It is natural to ask whether $\mathscr{X}$ is locally free of rank 1 over $\mathbf{T}$. In this section, we will answer the question affirmatively, except perhaps for certain primes (meaning: maximal ideals) of $\mathbf{T}$ which divide 2.

In what follows, we consider a maximal ideal $\mathfrak{m}$ of $\mathbf{T}$. Let p be the characteristic of the finite field $\mathbf{T}/\mathfrak{m}$. As in [9, Ch. II, §7], we let $\mathbf{T}_\mathfrak{m} = \varprojlim_\nu \mathbf{T}/\mathfrak{m}^\nu$ be the completion of $\mathbf{T}$ at $\mathfrak{m}$. As usual, we say that $\mathfrak{m}$ is ordinary if T_p is non-zero mod $\mathfrak{m}$ and supersingular otherwise.

Also, we recall that $\mathfrak{m}$ is *Eisenstein* if it divides (i.e., contains) the Eisenstein ideal $\mathscr{I}$ of $\mathbf{T}$. This latter ideal is defined (on p. 95 of [9]) as the ideal generated by the difference $T_N - 1$ and by the quantities $\eta_\ell := 1 + \ell - T_\ell$ as ℓ ranges over the set of primes different from N. The natural map $\mathbf{Z} \to \mathbf{T}/\mathscr{I}$ induces an isomorphism $\mathbf{Z}/n\mathbf{Z} \xrightarrow{\sim} \mathbf{T}/\mathscr{I}$, where n is the numerator of $\frac{N-1}{12}$. Thus the Eisenstein primes of $\mathbf{T}$ are in 1-1 correspondence with the prime ideals of $\mathbf{Z}/n\mathbf{Z}$ and therefore with the prime numbers which divide n.

Next, we write $J[\mathfrak{m}]$ for the group of points in $J(\overline{\mathbf{Q}})$ which are killed by all elements of $\mathfrak{m}$ (cf. [9, p. 91]). This group is a $\mathbf{T}/\mathfrak{m}$-vector space which is furnished with an action of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$. Recall the following key result of [9]:

Theorem 2.1. *Let $\mathfrak{m}$ be a maximal ideal of $\mathbf{T}$. If $\mathfrak{m}$ divides 2, suppose that $\mathfrak{m}$ is either Eisenstein or supersingular. Then $J[\mathfrak{m}]$ is of dimension two.*

Theorem 2.1 is proved in [9, Ch. II]. Note, however, that the discussions for $\mathfrak{m}$ Eisenstein and $\mathfrak{m}$ non-Eisenstein occur in different sections: one may consult Proposition 14.2 if $\mathfrak{m}$ is non-Eisenstein and (16.3) if $\mathfrak{m}$ is Eisenstein. (See also (17.9) if $\mathfrak{m}$ is Eisenstein and $\mathfrak{m}$ divides 2.)

When $\mathfrak{m}$ is Non-Eisenstein, Theorem 2.1 relates $J[\mathfrak{m}]$ and the standard representation $\rho_\mathfrak{m}$ of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ which is attached to $\mathfrak{m}$. By definition, $\rho_\mathfrak{m}$ is the unique (up to isomorphism) continuous semisimple representation $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathbf{GL}(2, \mathbf{T}/\mathfrak{m})$ satisfying: (i) $\det \rho_\mathfrak{m}$ is the mod p cyclotomic character; (ii) for each prime ℓ prime to pN, $\rho_\mathfrak{m}$ is unramified at ℓ and $\rho_\mathfrak{m}(\mathrm{Frob}_\ell)$ has trace T_ℓ mod $\mathfrak{m}$. (The existence and uniqueness of $\rho_\mathfrak{m}$ are discussed, for instance, in [14, §5].) The representation $\rho_\mathfrak{m}$ is irreducible if and only if $\mathfrak{m}$ is non-Eisenstein [9, Ch. II, Prop. 14.1 and Prop. 14.2]. The relation between $J[\mathfrak{m}]$ and $\rho_\mathfrak{m}$ is that the former representation *is* (i.e., defines or affords) the latter representation whenever $J[\mathfrak{m}]$ is irreducible and 2-dimensional [9, Ch. II, §14]. In particular, if $\mathfrak{m}$ is non-Eisenstein, then $J[\mathfrak{m}]$ affords the representation $\rho_\mathfrak{m}$ if either p is odd or $\mathfrak{m}$ is supersingular.

Suppose that $\rho_{\mathfrak{m}}$ is irreducible. Following [18, p. 189], we define $\rho_{\mathfrak{m}}$ to be finite at N if there is a finite flat $\mathbf{T}/\mathfrak{m}$-vector space scheme $\mathscr{V}$ of rank 2 over $\mathbf{Z}_N$ such that the restriction of $\rho_{\mathfrak{m}}$ to $D = \operatorname{Gal}(\overline{\mathbf{Q}}_N/\mathbf{Q}_N)$ is isomorphic to the two-dimensional representation $\mathscr{V}(\overline{\mathbf{Q}}_N)$. The following result is obtained by combining a 1973 theorem of Tate with the author's level-lowering result.

Proposition 2.2. *Let $\mathfrak{m}$ be a non-Eisenstein prime of $\mathbf{T}$. Then the two-dimensional Galois representation $\rho_{\mathfrak{m}}$ is not finite at the prime N.*

Proof. Suppose first that $\mathfrak{m}$ does not divide 2. Assume that $\rho_{\mathfrak{m}}$ is finite at N. Then [14, Th. 1.1] shows that $\rho_{\mathfrak{m}}$ is modular of level 1. (In applying [14, Th. 1.1], we take $N = N$, $p = N$, and $\ell = p$. Note that condition 2 of the theorem is satisfied except when $\mathfrak{m}$ divides N. In this case however, condition 1 of the theorem holds since we do not have $N \equiv 1 \bmod N$.) This is a contradiction, since there are no non-zero weight-2 cusp forms on $\Gamma_0(1)$.

Assume now that $\mathfrak{m}$ does divide 2. Suppose again that $\rho_{\mathfrak{m}}$ is finite at N. Then $\rho_{\mathfrak{m}}$ is an irreducible mod 2 two-dimensional representation of $\operatorname{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ which is unramified outside of the prime 2. An important theorem of Tate [20] proves, however, that there is no such representation. ■

Note that when p is different from N, $\rho_{\mathfrak{m}}$ is finite at N if and only if $\rho_{\mathfrak{m}}$ is unramified at N. Thus Proposition 2.2 shows, in particular, that $\rho_{\mathfrak{m}}$ is ramified at N for all $\mathfrak{m}$ such that $\rho_{\mathfrak{m}}$ is irreducible.

Theorem 2.3. *Let $\mathfrak{m}$ be a maximal ideal of $\mathbf{T}$. If $\mathfrak{m}$ divides 2, suppose that $\mathfrak{m}$ is either Eisenstein or supersingular. Then $\mathscr{X} \otimes_{\mathbf{T}} \mathbf{T}_{\mathfrak{m}}$ is free of rank 1 over $\mathbf{T}_{\mathfrak{m}}$.*

Proof. Since $\mathscr{X}$ is of rank 1, $\mathscr{X} \otimes \mathbf{T}_{\mathfrak{m}}$ is free of rank 1 if and only if it is cyclic. By Nakayama's lemma, the cyclicity amounts to the statement that $\mathscr{X}/\mathfrak{m}\mathscr{X}$ has dimension ≤ 1 over the field $\mathbf{T}/\mathfrak{m}$.

To prove this latter statement, i.e., the cyclicity of $\mathscr{X}/\mathfrak{m}\mathscr{X}$, we exploit the relation between $\mathscr{X}$ and torsion points of J. In the following discussion, for each integer $m \geq 1$, we let $J[m]$ be the group of points of J with values in $\overline{\mathbf{Q}}$ which have order dividing m. Thus $J[m]$ is a $\mathbf{T}[\operatorname{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})]$-module. Especially, we shall view $J[m]$ locally at N, i.e., as a $\mathbf{T}[D]$-module. One obtains from [6, 11.6.6–11.6.7] a $\mathbf{T}[D]$-equivariant exact sequence

$$0 \to \operatorname{Hom}(\mathscr{X}/m\mathscr{X}, \mu_m) \to J[m] \to \mathscr{X}/m\mathscr{X} \to 0. \tag{2.4}$$

(See, e.g., [15, pp. 669–670] for a discussion of this exact sequence when m is a prime number.) Especially, there is a natural identification of $\operatorname{Hom}(\mathscr{X}/m\mathscr{X}, \mu_m)$ with a subgroup of $J[m]$.

In particular, we find an injection

$$j\colon \operatorname{Hom}(\mathscr{X}/\mathfrak{m}\mathscr{X}, \mu_p) \hookrightarrow J[\mathfrak{m}];$$

here, p is again the residue characteristic of $\mathfrak{m}$. By Theorem 2.1, j is an isomorphism if $\mathscr{X}/\mathfrak{m}\mathscr{X}$ is not cyclic.

On the other hand, it is clear that j cannot be an isomorphism. Indeed, the group $\mathrm{Hom}(\mathscr{X}/\mathfrak{m}\mathscr{X}, \mu_p)$ is finite at N in the sense of [18] (since μ_p is finite), and we have seen in Proposition 2.2 that $J[\mathfrak{m}]$ is not finite at N. ■

3. The kernel of the Eisenstein ideal

We turn now to a study of the action of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ on the Eisenstein kernel in the Jacobian $J = J_0(N)$. Let $\mathscr{I}$ again be the Eisenstein ideal of $\mathbf{T}$, and recall that $n = \mathrm{num}\,\dfrac{N-1}{12}$. By $J[\mathscr{I}]$ we mean the kernel of $\mathscr{I}$ on J, i.e., the group of points in $J(\overline{\mathbf{Q}})$ which are annihilated by all elements of $\mathscr{I}$. The analysis of [9, Ch. II, §§16–18] shows that $J[\mathscr{I}]$ is free of rank two over $\mathbf{T}/\mathscr{I} \approx \mathbf{Z}/n\mathbf{Z}$.

The group $J[\mathscr{I}]$ contains the cuspidal group C, which was mentioned above, and also the Shimura subgroup Σ of J [9, Ch. II, §11]. The two groups C and Σ are $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$-stable and cyclic of order n. The actions of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ on these two groups are respectively the trivial action and the cyclotomic action ($\Sigma \approx \mu_n$). Accordingly, the intersection of C and Σ is trivial if n is odd; in that case, the inclusions of C and Σ in $J[\mathscr{I}]$ induce an isomorphism of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$-modules $C \oplus \Sigma \xrightarrow{\sim} J[\mathscr{I}]$. If n is even, however, $C \cap \Sigma$ has order 2, and the sum $C + \Sigma$ in $J[\mathscr{I}]$ (which is no longer direct) has index 2 in $J[\mathscr{I}]$.

In much of what follows, the reader may wish to assume that n is even; when n is odd, almost everything that we prove may be deduced immediately from the decomposition $J[\mathscr{I}] \approx C \oplus \Sigma$.

Proposition 3.1. *The group $J[\mathscr{I}]$ is unramified at N.*

Proof. We regard $J[\mathscr{I}]$ as a D-module, where $D = \mathrm{Gal}(\overline{\mathbf{Q}}_N/\mathbf{Q}_N)$ as above. We have a natural injection (analogous to the map j above)

$$\mathrm{Hom}(\mathscr{X}/\mathscr{I}\mathscr{X}, \mu_n) \hookrightarrow J[\mathscr{I}],$$

where $\mathscr{X}$ is again the character group associated with the reduction of J mod N. By combining this injection with the inclusion of Σ in $J[\mathscr{I}]$, we obtain a map of D-modules

$$\theta\colon \Sigma \oplus \mathrm{Hom}(\mathscr{X}/\mathscr{I}\mathscr{X}, \mu_n) \longrightarrow J[\mathscr{I}].$$

This map is again injective, in view of Proposition 11.9 of [9, Ch. II].

Now by Theorem 2.3, $\mathscr{X}$ is free of rank 1 locally at each prime $\mathfrak{m}$ dividing $\mathscr{I}$. Hence $\mathscr{X}/\mathscr{I}\mathscr{X}$, and therefore $\mathrm{Hom}(\mathscr{X}/\mathscr{I}\mathscr{X}, \mu_n)$, has order n. Thus the source of θ has n^2 elements. Since the target of θ has the same cardinality, we conclude that θ is an isomorphism of D-modules. The group $\Sigma \oplus \mathrm{Hom}(\mathscr{X}/\mathscr{I}\mathscr{X}, \mu_n)$, however, is unramified; note that D acts on $\mathscr{X}$ through its quotient $\mathrm{Gal}(\overline{\mathbf{F}}_N/\mathbf{F}_N)$. ■

We continue our study of the action of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ on $J[\mathscr{I}]$:

Proposition 3.2. *The Galois group* $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ *acts trivially on* $J[\mathscr{I}]/\Sigma$.

Proof. It is clear that Jordan-Hölder constituents of the $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$-module $J[\mathscr{I}]$ are all of the form μ_p or $\mathbf{Z}/p\mathbf{Z}$, with p dividing n. Indeed, $J[\mathscr{I}]$ is an extension of a group whose order divides 2 by a quotient of $\Sigma \oplus C$, where the latter group has the indicated property. Because $J[\mathscr{I}]$ is unramified at N, it is finite at N in Serre's sense; it extends to a finite flat group scheme over $\mathbf{Z}$. In the language of Chapter I of [9], $J[\mathscr{I}]$ is thus an admissible group scheme over $\operatorname{Spec} \mathbf{Z}[\frac{1}{N}]$ which extends to a finite flat group scheme G over $\operatorname{Spec} \mathbf{Z}$.

To analyze G, we follow the proof of Proposition 4.5 in [9, Ch. I]. The last step in the proof of that Proposition uses a result above it (Proposition 4.1) which applies only to groups of odd order. However, Steps 1–3 are perfectly applicable; they show that G is an extension of a constant group scheme by a μ-type group (dual of a constant group) $H \subseteq G$.

In particular, there is a subgroup Σ' of $J[\mathscr{I}]$ with the property that the action of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ on Σ' is cyclotomic, whereas the action of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ on $J[\mathscr{I}]/\Sigma'$ is trivial. By [9, Ch. III, Th. 1.3], Σ' is contained in Σ. Hence the action of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ on the quotient $J[\mathscr{I}]/\Sigma$ is indeed trivial. ■

Before studying further the $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$-action on $J[\mathscr{I}]$, we pause to establish a converse to Proposition 3.1.

Proposition 3.3. *Let* $P \in J(\overline{\mathbf{Q}})$ *be a torsion point on* J *for which the finite extension* $\mathbf{Q}(P)/\mathbf{Q}$ *is unramified at* N. *Then* P *lies in* $J[\mathscr{I}]$.

Proof. Let M be smallest $\mathbf{T}[\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})]$-submodule of $J(\overline{\mathbf{Q}})$ which contains both P and $J[\mathscr{I}]$. We must prove that M is annihilated by $\mathscr{I}$. Clearly, M is finite; indeed, we have $M \subseteq J[mn]$ if m is the order of P. Consider the Jordan-Hölder constituents of M, regarded as a $\mathbf{T}[\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})]$-module. If V is such a constituent, then the annihilator of V is a maximal ideal $\mathfrak{m}$ of $\mathbf{T}$. It follows from the discussion of [9, Ch. II, §14] that V is 1-dimensional over $\mathbf{T}/\mathfrak{m}$ if and only if $\mathfrak{m}$ is Eisenstein. If $\mathfrak{m}$ is not Eisenstein, then V is isomorphic to the irreducible representation $\rho_{\mathfrak{m}}$. (This follows from the discussion on page 115 of [9]. In fact, the main result of [1] can be used to prove the more precise fact that $J[\mathfrak{m}]$ is a direct sum of copies of $\rho_{\mathfrak{m}}$ when $\mathfrak{m}$ is non-Eisenstein.) However, Proposition 2.2 shows that $\rho_{\mathfrak{m}}$ is ramified at N when $\mathfrak{m}$ is non-Eisenstein. We conclude that all constituents of M belong to Eisenstein primes of $\mathbf{T}$. These constituents therefore have the form μ_p or $\mathbf{Z}/p\mathbf{Z}$, with p dividing n.

Returning to the language of [9, Ch. I], we see that M is an admissible group. As explained in the proof of the proposition above, M must be an extension of a constant group Q by a μ-type group M_0. Since M contains $J[\mathscr{I}]$ and since Σ is the maximal μ-type group in $J(\overline{\mathbf{Q}})$, we have $M_0 = \Sigma$. Next, note that the extension of $\mathbf{T}$-modules

$$0 \to \Sigma \to M \to Q \to 0$$

splits. The splitting is obtained as in the argument on p. 142 of [9] which proves [9, Ch. III, Th. 1.3]. Namely, specialization to characteristic N provides a map $M \to \Phi$, where Φ is the component group of J in characteristic N. We get a splitting because the restriction of this map to Σ is an isomorphism $\Sigma \xrightarrow{\sim} \Phi$. It follows that $\mathscr{I}$ annihilates M if and only if $\mathscr{I}$ annihilates Q.

Since $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ acts trivially on Q, the Eichler-Shimura relation shows that Q is annihilated by the differences $\eta_\ell = 1 + \ell - T_\ell$. To deduce from this the apparently stronger fact that Q is annihilated by all of $\mathscr{I}$ (which includes the generator $T_N - 1$), write Q as the direct sum $\oplus_{\mathfrak{m}} Q_{\mathfrak{m}}$, where the sum runs over the set of Eisenstein primes of $\mathbf{T}$. Each summand $Q_{\mathfrak{m}}$ is a module over $\mathbf{T}/\mu^\nu$, where ν is a suitable positive integer. It follows from [9, Ch. II, Th. 18.10] that the image of $\mathscr{I}$ in $\mathbf{T}/\mu^\nu$ is generated by a single element of the form η_ℓ. Thus $Q_{\mathfrak{m}}$ is annihilated by $\mathscr{I}$. Since this statement is true for each $\mathfrak{m}$, Q is annihilated by $\mathscr{I}$. ■

Our next goal is to study $J[\mathscr{I}]$ sufficiently closely to permit identification of the field $\mathbf{Q}(J[\mathscr{I}])$, i.e., to prove Theorem 1.7. For an alternative proof of Theorem 1.7, the reader may consult Csirik's forthcoming article [3], which determines $J[\mathscr{I}]$ completely by a method generalizing that of [9, Ch. II, §12–§13].

Recall that the cuspidal group C is provided with a natural generator, namely the image of the cusp 0 in J. We select generators for certain other cyclic groups by making use of the place over N that we have chosen in $\overline{\mathbf{Q}}$. As explained in §11 of [9, Ch. II], reduction to characteristic N induces isomorphisms among C, Σ and the group of components of $J_{/\mathbf{F}_N}$. In particular, we have a distinguished isomorphism $C \approx \Sigma$. Since C is provided with a generator, we obtain a basis of Σ. (See [5] for a comparison of the isomorphism $C \approx \Sigma$ with a second natural one.)

Since Σ and $J[\mathscr{I}]$ are free of ranks 1 and 2 over $\mathbf{Z}/n\mathbf{Z}$, the group $\mathbf{Q} := J[\mathscr{I}]/\Sigma$ is cyclic of order n. The intersection $C \cap \Sigma$ has order $\gcd(2, n)$ [9, Ch. II, Prop. 11.11]. The image of C in Q has order $n/\gcd(2, n)$. Choose a generator g of Q such that $2g$ is the image in Q of the chosen generator of C. Finally, note as above that reduction to characteristic N provides us with a splitting of the tautological exact sequence which displays Q as a quotient of $J[\mathscr{I}]$. This splitting writes $J[\mathscr{I}]$ as the direct sum $\Sigma \oplus Q$. (Said differently, $J[\mathscr{I}]$ is the direct sum of Σ and the "toric part" $\mathrm{Hom}(\mathscr{X}/\mathscr{I}\mathscr{X}, \mu_n)$ of $J[\mathscr{I}]$. The natural map $\mathrm{Hom}(\mathscr{X}/\mathscr{I}\mathscr{X}, \mu_n) \to Q$ is an isomorphism.)

Using the chosen generators of Σ and Q, we write $J[\mathscr{I}] = (\mathbf{Z}/n\mathbf{Z})^2$. In this model of $J[\mathscr{I}]$, Σ is the group generated by $(1, 0)$ and C is the group generated by $(1, 2)$. Since $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ preserves Σ and operates on Σ as the mod n cyclotomic character χ, and since $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ operates trivially on Q, the action of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ on $J[\mathscr{I}]$ is given in matrix terms by a map

$$\sigma \longmapsto \rho(\sigma) := \begin{pmatrix} \chi & b(\sigma) \\ 0 & 1 \end{pmatrix}.$$

Here, the map $\sigma \mapsto b(\sigma) \in \mathbf{Z}/n\mathbf{Z}$ is clearly a 1-cocycle: it verifies the identity

$$b(\sigma\tau) = b(\sigma) + \chi(\sigma)b(\tau)$$

for $\sigma, \tau \in \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$.

4. Lenstra's input

The contents of this section were suggested to the author by H. W. Lenstra, Jr. The author thanks him heartily for his help.

Lemma 4.1. *For all $\sigma \in \operatorname{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$, we have $2b(\sigma) = 1 - \chi(\sigma)$.*

Proof. For each σ, $\rho(\sigma)$ fixes the vector $\begin{pmatrix} 1 \\ 2 \end{pmatrix} \in C$. The lemma follows immediately. ■

Proposition 4.2. *The field $\mathbf{Q}(J[\mathscr{I}])$ is an abelian extension of $\mathbf{Q}$ which contains $\mathbf{Q}(\mu_n)$ and has degree 1 or 2 over $\mathbf{Q}(\mu_n)$.*

Proof. To say that $\mathbf{Q}(J[\mathscr{I}])$ is abelian over $\mathbf{Q}$ is to say that the image of ρ is abelian. This amounts to the identity $b(\sigma\tau) \overset{?}{=} b(\tau\sigma)$ for $\sigma, \tau \in \operatorname{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$. By the cocycle identity, the two sides of the equation are respectively $b(\sigma) + \chi(\sigma)b(\tau)$ and $b(\tau) + \chi(\tau)b(\sigma)$. These expressions are indeed equal, in view of the lemma above.

It is clear that the field $\mathbf{Q}(J[\mathscr{I}])$ contains $\mathbf{Q}(\mu_n)$ because the kernel of ρ is contained in the kernel of χ. Let H be this latter kernel; i.e., $H = \operatorname{Gal}\big(\overline{\mathbf{Q}}/\mathbf{Q}(\mu_n)\big)$. On H, $\chi = 1$; hence we have $2b = 0$ in $\mathbf{Z}/n\mathbf{Z}$. In other words, the group $\rho(H)$ is a subgroup of the group of matrices $\begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix}$ with $2x = 0$. Since this group has order $\gcd(2, n)$, the extension of $\mathbf{Q}$ cut out by ρ is an extension of $\mathbf{Q}(\mu_n)$ of degree 1 or 2. ■

The proof of Proposition 4.2 (or, alternatively, the decomposition $J[\mathscr{I}] = \Sigma \oplus C$) shows that $\mathbf{Q}(J[\mathscr{I}]) = \mathbf{Q}(\mu_n)$ if n is odd. Suppose now that n is even; write $n = 2^k n_{\mathrm{o}}$, where n_{o} is the "odd part" and $2^k \geq 2$ is the largest power of 2 dividing n. Then ρ is the direct sum of representations

$$\rho_2\colon \operatorname{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathbf{GL}(2, \mathbf{Z}/2^k\mathbf{Z}), \qquad \rho_{\mathrm{o}}\colon \operatorname{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathbf{GL}(2, \mathbf{Z}/n_{\mathrm{o}}\mathbf{Z}),$$

which are defined by the actions of $\operatorname{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ on the 2-primary part and the odd part of $J[\mathscr{I}]$, respectively. It is evident that the latter representation cuts out $\mathbf{Q}(\mu_{n_{\mathrm{o}}})$ and that the kernel of the former representation corresponds to an abelian extension K of $\mathbf{Q}$ which contains $\mathbf{Q}(\mu_{2^k})$ and has degree 1 or 2 over this cyclotomic field. Since ρ_2 is defined by the action of $\operatorname{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ on a group of 2-power division points of J, this representation can be ramified only at 2 and at N. We have seen, however, that ρ is unramified at N (Proposition 3.1). Hence $K/\mathbf{Q}$ is an abelian extension of $\mathbf{Q}$ which is ramified only at 2; it follows (e.g., from the proof that the "local Kronecker-Weber theorem" implies the usual, global one [21, Ch. 14]) that K is contained in the cyclotomic field $\mathbf{Q}(\mu_{2^\infty})$. Hence we have either $K = \mathbf{Q}(\mu_{2^k})$ or $K = \mathbf{Q}(\mu_{2^{k+1}})$. Accordingly, we have

$$\mathbf{Q}(\mu_n) \subseteq \mathbf{Q}(J[\mathscr{I}]) \subseteq \mathbf{Q}(\mu_{2n}).$$

In summary, the displayed inclusions hold both in the case when n is odd and when n is even. In the former case, the two cyclotomic fields are equal, and they coincide with $\mathbf{Q}(J[\mathscr{I}])$. In the latter case, there remains an ambiguity which will be resolved by the proof of Theorem 1.7.

Before turning to this proof in the general case, we present a *simple proof of Theorem 1.7* in the case where k is different from 2. To prove the Theorem is to show that ρ_2 cuts out the field $\mathbf{Q}(\mu_{2^{k+1}})$. This is perfectly clear if $k=0$, in which case ρ_2 is the trivial representation: the field $K=\mathbf{Q}$ is indeed the field of second roots of 1. If $k=1$, ρ_2 gives the action of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ on the group D which is described in [9, Ch. II, §12]; Lemma 12.4 of that section states that the field K is the field of fourth roots of unity.

Suppose now that k is at least 3, and choose $\sigma \in \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ so that $\chi(\sigma) \equiv 1+2^{k-1} \bmod 2^k$ and $\chi(\sigma) \equiv 1 \bmod n_{\mathrm{o}}$. It is evident that $\chi(\sigma^2)=1$; we will show, however, that $\rho_2(\sigma^2) \neq 1$. These two pieces of information imply that K is not contained in $\mathbf{Q}(\mu_n)$, which is precisely the information that we seek. To prove that $\rho_2(\sigma^2)$ is different from 1 is to show that $b(\sigma^2) \not\equiv 0 \bmod 2^k$. We have

$$b(\sigma^2) = (1+\chi(\sigma)b(\sigma) \equiv 2(1+2^{k-2})b(\sigma) \bmod 2^k$$

by the cocycle identity and the choice of σ. Since k is at least 3, the factor $(1+2^{k-2})$ is odd. Now $2b(\sigma) = 1-\chi(\sigma) \equiv -2^{k-1} \bmod 2^k$ in view of Lemma 4.1. Thus $b(\sigma)$ is divisible by 2^{k-2} but not by 2^{k-1}. It follows that $b(\sigma^2)$ is divisible by 2^{k-1} but not by 2^k. ■

5. Proof of Theorem 1.7

We return to the discussion of the general case, removing the assumption $k \geq 3$. Recall that ρ is the representation of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ giving the action of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ on $J[\mathscr{I}]$ and that $\eta_\ell = 1+\ell-T_\ell$ for each $\ell \neq N$.

Lemma 5.1. *Let ℓ be a prime number prime to nN. Suppose that $\rho(\mathrm{Frob}_\ell)=1$. Then η_ℓ belongs to $\mathscr{I}^2$.*

Proof. One has $\mathbf{T}/\mathscr{I}^2 = \bigoplus \mathbf{T}_\mathfrak{m}/\mathscr{I}^2\mathbf{T}_\mathfrak{m}$, where the sum is taken over the Eisenstein primes $\mathfrak{m}$ of $\mathbf{T}$. We must show that the image of η in $\mathbf{T}_\mathfrak{m}/\mathscr{I}^2\mathbf{T}_\mathfrak{m}$ is 0 for each such $\mathfrak{m}$. Fix $\mathfrak{m}$, and let p be the corresponding prime divisor of n. Consider the p-divisible group $J_\mathfrak{m} = \bigcup_\nu J[\mathfrak{m}^\nu]$ and its Tate module $\mathrm{Ta}_\mathfrak{m} := \mathrm{Hom}(\mathbf{Q}_p/\mathbf{Z}_p, J_\mathfrak{m})$. Let

$$\mathrm{Ta}_\mathfrak{m}^* := \mathrm{Hom}_{\mathbf{Z}_p}(\mathrm{Ta}_\mathfrak{m}, \mathbf{Z}_p) = \mathrm{Hom}(J_\mathfrak{m}, \mathbf{Q}_p/\mathbf{Z}_p);$$

the latter description of $\mathrm{Ta}_\mathfrak{m}^*$ presents this Tate module as the Pontryagin dual of $J_\mathfrak{m}$. Note that $\mathrm{Ta}_\mathfrak{m}$ and $\mathrm{Ta}_\mathfrak{m}^*$ have been shown to be free of rank 2 over $\mathbf{T}_\mathfrak{m}$ [9, Ch. II, Cor. 16.3]. The Tate pairing $\mathrm{Ta}_p(J) \times \mathrm{Ta}_p(J) \to \mathbf{Z}_p(1)$ may be viewed as an isomorphism $\mathrm{Ta}_\mathfrak{m} \approx \mathrm{Ta}_\mathfrak{m}^*(1)$ which is compatible with the natural actions of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ and $\mathbf{T}$ on the two modules.

Let $F = \mathrm{Frob}_\ell$. Since $1-F$ annihilates $J[\mathscr{I}]$, $1-F$ annihilates the Shimura subgroup $\Sigma \approx \mu_n$ of J, which is contained in $J[\mathscr{I}]$. Hence $\ell \equiv 1 \bmod n$. Accordingly, F acts as the identity on $\mathrm{Hom}(J[\mathscr{I}], \mu_n)$ and its p-primary subgroup $\mathrm{Hom}(J_\mathfrak{m}[\mathscr{I}], \mathbf{Q}_p/\mathbf{Z}_p)(1)$.

We may view this dual as the quotient $\mathrm{Ta}_{\mathfrak{m}}^*(1)/\mathscr{I}\,\mathrm{Ta}_{\mathfrak{m}}^*(1) \approx \mathrm{Ta}_{\mathfrak{m}}/\mathscr{I}\,\mathrm{Ta}_{\mathfrak{m}}$. Hence we have

$$(1-F)(\mathrm{Ta}_{\mathfrak{m}}) \subseteq \mathscr{I}\cdot\mathrm{Ta}_{\mathfrak{m}}\,.$$

Since $\mathrm{Ta}_{\mathfrak{m}}$ is free of rank 2 over $\mathbf{T}_{\mathfrak{m}}$, we obtain

$$\det\nolimits_{\mathbf{T}_{\mathfrak{m}}}\left(1-F \;\middle|\; \mathrm{Ta}_{\mathfrak{m}}\right) \in \mathscr{I}^2\mathbf{T}_{\mathfrak{m}}.$$

This proves what is needed, since the determinant we have calculated is nothing but η_ℓ; indeed, the the determinant and trace of F acting on $\mathrm{Ta}_{\mathfrak{m}}$ are ℓ and T_ℓ, respectively. ■

Theorem 5.2. *Assume that n is even. Let $\ell \neq N$ be a prime number which satisfies the congruence $\ell \equiv 1$ mod n but not the congruence $\ell \equiv 1$ mod $2n$. Assume further that the image of ℓ in $(\mathbf{Z}/N\mathbf{Z})^*$ is a generator of this cyclic group. Then $\rho(\mathrm{Frob}_\ell) \neq 1$.*

Proof. Let Δ be the unique quotient of $(\mathbf{Z}/N\mathbf{Z})^*$ of order n. To prove our result, we refer to §18 of [9, Ch. II]. In that section, one finds a homomorphism $\epsilon^+ \colon \mathscr{I}/\mathscr{I}^2 \to H^+/\mathscr{I}H^+$ and a map $\varphi\colon \Delta \to H^+/\mathscr{I}H^+$, both of which prove to be isomorphisms. The map $\kappa := \varphi^{-1}\circ\epsilon^+$ is an isomorphism $\mathscr{I}/\mathscr{I}^2 \xrightarrow{\sim} \Delta$. The *congruence formula for the winding homomorphism* yields

$$\kappa(\eta_\ell) = \tfrac{\ell-1}{2}\cdot\overline{\ell}.$$

Here, $\overline{\ell}$ is the image of $\ell \in (\mathbf{Z}/N\mathbf{Z})^*$ in Δ, and the operator $\frac{\ell-1}{2}$ is an exponent. (One is viewing the multiplicative abelian group Δ as a $\mathbf{Z}$-module.) Under our hypotheses, it is clear that $\frac{\ell-1}{2}\cdot\overline{\ell}$ has order 2 in Δ. Thus, by the congruence formula, η_ℓ is nonzero in $\mathscr{I}/\mathscr{I}^2$. Using Lemma 5.1, we deduce the required conclusion that $\rho(\mathrm{Frob}_\ell)$ is different from 1. ■

We now prove Theorem 1.7, i.e., the statement that $\mathbf{Q}(J[\mathscr{I}])$ coincides with the cyclotomic field $\mathbf{Q}(\mu_{2n})$.

As was explained above, the statement to be proved follows from the decomposition $J[\mathscr{I}] = \Sigma \oplus C$ when n is odd. Assume then that n is even. As we have discussed, the field $\mathbf{Q}(J[\mathscr{I}])$ is an extension of $\mathbf{Q}(\mu_n)$ of degree dividing 2. Moreover, if $\mathbf{Q}(J[\mathscr{I}])$ is indeed quadratic over $\mathbf{Q}(\mu_n)$, then $\mathbf{Q}(J[\mathscr{I}])$ has no choice but to be $\mathbf{Q}(\mu_{2n})$. To see that the extension $\mathbf{Q}(J[\mathscr{I}])/\mathbf{Q}(\mu_n)$ is non-trivial, we use the result above. Using the Chinese Remainder Theorem and Dirichlet's theorem on primes in an arithmetic progression, we may choose ℓ so as to satisfy the conditions of Theorem 5.2. A Frobenius element Frob_ℓ for ℓ in $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ then acts trivially on μ_n, but non-trivially on $J[\mathscr{I}]$. ■

6. Adelic representations

Let ℓ be a prime. As usual, we consider the ℓ-divisible group $J_\ell = \bigcup_\nu J[\ell^\nu]$ and its Tate modules $\mathrm{Ta}_\ell := \mathrm{Hom}(\mathbf{Q}_\ell/\mathbf{Z}_\ell, J_\ell)$ and $\mathrm{Ta}_\ell \otimes_{\mathbf{Z}_\ell} \mathbf{Q}_\ell$. The ℓ-adic representation of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ attached to J is the continuous homomorphism

$$\rho_\ell\colon\ \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathrm{Aut}(\mathrm{Ta}_\ell) \hookrightarrow \mathrm{Aut}\left(\mathrm{Ta}_\ell \otimes \mathbf{Q}_\ell\right))$$

which arises from the action of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ on Ta_ℓ.

This action is $\mathbf{T}$-linear, where $\mathbf{T}$ is the Hecke ring introduced above. Thus ρ_ℓ takes values, for example, in the group $\mathrm{Aut}_{\mathbf{T}_\ell}(\mathrm{Ta}_\ell)$, where $\mathbf{T}_\ell = \mathbf{T} \otimes \mathbf{Z}_\ell$. Note that the $\mathbf{Z}_\ell$-algebra $\mathbf{T}_\ell$ is the product of the completions $\mathbf{T}_\mathfrak{m}$ of $\mathbf{T}$ at the maximal ideals $\mathfrak{m}$ of $\mathbf{T}$ which divide ℓ. The corresponding decomposition of Ta_ℓ into a product of modules over the individual factors $\mathbf{T}_\mathfrak{m}$ of $\mathbf{T}_\ell$ is the natural decomposition of $\mathrm{Ta}_\ell = \prod_\mathfrak{m} \mathrm{Ta}_\mathfrak{m}$, where the $\mathrm{Ta}_\mathfrak{m}$ are the $\mathfrak{m}$-adic Tate modules which were introduced earlier.

As we have noted, Mazur proves in [9, Ch. II, §15–§18] that $\mathrm{Ta}_\mathfrak{m}$ is free of rank 2 over $\mathbf{T}_\mathfrak{m}$ for each maximal ideal $\mathfrak{m}$ of $\mathbf{T}$ which is not simultaneously ordinary, non-Eisenstein and of residue characteristic 2. Thus, after a choice of basis, $\mathrm{Aut}_{\mathbf{T}_\ell}(\mathrm{Ta}_\ell)$ becomes $\mathbf{GL}(2, \mathbf{T} \otimes \mathbf{Z}_\ell)$ for each prime $\ell > 2$. Thus, if ℓ is odd, ρ_ℓ may be viewed as a homomorphism

$$\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathbf{GL}(2, \mathbf{T} \otimes \mathbf{Z}_\ell).$$

Similarly, we may view ρ_2 as taking values in $\mathbf{GL}(2, \mathbf{T} \otimes \mathbf{Q}_2)$. Accordingly, the image G_ℓ of ρ_ℓ is a subgroup of $\mathbf{GL}(2, \mathbf{T} \otimes \mathbf{Q}_\ell)$ in all cases and a subgroup of $\mathbf{GL}(2, \mathbf{T} \otimes \mathbf{Z}_\ell)$ when ℓ is odd. The determinant of ρ_ℓ is the ℓ-adic cyclotomic character.

The group G_ℓ is studied in [16], where the following two results are obtained as Proposition 7.1 and Theorem 6.4, respectively:

Theorem 6.1. *The group G_ℓ is open in the matrix group*

$$\{\, M \in \mathbf{GL}(2, \mathbf{T} \otimes \mathbf{Q}_\ell) \mid \det M \in \mathbf{Q}_\ell^* \,\}.$$

Theorem 6.2. *Suppose that ℓ is at least 5 and is prime to the discriminant of* $\mathbf{T}$. *Suppose further that no maximal ideal $\mathfrak{m}|\ell$ is an Eisenstein ideal of* $\mathbf{T}$ *(i.e., that ℓ is prime to n). Then*

$$G_\ell = \{\, M \in \mathbf{GL}(2, \mathbf{T} \otimes \mathbf{Z}_\ell) \mid \det M \in {\mathbf{Z}_\ell}^* \,\}.$$

Consider next the adelic representation $\rho_\mathrm{f} := \prod_\ell \rho_\ell$, where the product is taken over the set of all prime numbers ℓ. The image G_f of ρ_f is a subgroup of the product $\prod_\ell G_\ell$, which in turn is contained in the group

$$\{\, M \in \mathbf{GL}(2, \mathbf{T} \otimes \mathbf{Q}_2) \mid \det M \in \mathbf{Q}_2^* \,\} \times \prod_{\ell \neq 2} \{\, M \in \mathbf{GL}(2, \mathbf{T} \otimes \mathbf{Z}_\ell) \mid \det M \in {\mathbf{Z}_\ell}^* \,\}.$$

According to [16, Th. 7.5], G_f is open in the latter product.

For each prime ℓ, let H_ℓ be the intersection of G_f with the group

$$1 \times \cdots \times 1 \times G_\ell \times 1 \times \cdots \times 1 \cdots ,$$

where G_ℓ is placed in the ℓth factor. Thus H_ℓ is a subgroup of G_ℓ which may be viewed as the image of the restriction of ρ_ℓ to the kernel of the representation $\prod_{\ell' \neq \ell} \rho_{\ell'}$.

Theorem 6.3. *Assume that ℓ satisfies the conditions of Theorem 6.2, i.e., that ℓ is prime to* $\operatorname{disc}\mathbf{T}$ *and distinct from 2 and 3. Assume further that ℓ is different from N. Then $H_\ell = G_\ell = \{\, M \in \mathbf{GL}(2, \mathbf{T} \otimes \mathbf{Z}_\ell) \mid \det M \in {\mathbf{Z}_\ell}^* \,\}$.*

Proof. The proof of this result is explained in the course of the proof of Theorem 7.5 of [16]: Fix ℓ, and let X be the smallest closed subgroup of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ which contains all inertia groups of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ for the prime ℓ. Since $\rho_{\ell'}(X) = \{1\}$ for all primes $\ell' \neq \ell$, $\rho_\ell(X)$ is a subgroup of H_ℓ, which in turn is contained in G_ℓ. As the author observed at the end of §6 of [16], the desired equality $\rho_\ell(X) = G_\ell$ follows from Theorem 3.4 and Proposition 4.2 of [16]. ■

We now present a variant of the result above for the prime $\ell = N$. For this, we let Γ be the subgroup $1 + N\mathbf{Z}_N$ of $\mathbf{Z}_N^*$, i.e., the N-Sylow subgroup of $\mathbf{Z}_N^*$.

Proposition 6.4. *Suppose that N is prime to the discriminant of* $\mathbf{T}$. *Then H_N contains the group $\{\, M \in \mathbf{GL}(2, \mathbf{T} \otimes \mathbf{Z}_N) \mid \det M \in \Gamma \,\}$.*

Proof. Let X now be the smallest closed subgroup of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ which contains the wild subgroups (i.e., N-Sylow subgroups) of all inertia groups for N in $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$. It follows from the exact sequence (2.4) that we have $\rho_\ell(X) = \{1\}$ for all $\ell \neq N$. (If $\ell \neq N$, inertia groups at N act unipotently in the ℓ-adic representations attached to J. Consequently, the image under ρ_ℓ of an inertia group at N is a pro-ℓ group.) Hence $\rho_N(X)$ is a subgroup of H_N, and it will suffice to show that $\rho_N(X) = \{\, M \in \mathbf{GL}(2, \mathbf{T} \otimes \mathbf{Z}_N) \mid \det M \in \Gamma \,\}$. We note that $\rho_N(X)$ is contained in this matrix group since the image of $\rho_N(X)$ under the determinant mapping $G_N \to \mathbf{Z}_N^*$ is a pro-N group. Since in fact the group $\det \rho_N(X)$ is all of Γ, the equality $\rho_N(X) = \{\, M \in \mathbf{GL}(2, \mathbf{T} \otimes \mathbf{Z}_N) \mid \det M \in \Gamma \,\}$ means that $\rho_N(X)$ contains $\mathbf{SL}(2, \mathbf{T} \otimes \mathbf{Z}_N)$.

Because $\mathbf{T}$ is unramified at N, [16, Prop. 4.2] implies that the inclusion

$$\rho_N(X) \supseteq \mathbf{SL}(2, \mathbf{T} \otimes \mathbf{Z}_N)$$

holds if and only if it holds "mod N" in the sense that the image of X in $\mathbf{GL}(2, \mathbf{T}/N\mathbf{T})$ contains $\mathbf{SL}(2, \mathbf{T}/N\mathbf{T})$. To say that this image contains $\mathbf{SL}(2, \mathbf{T}/N\mathbf{T})$ is in fact to say that the image coincides with $\mathbf{SL}(2, \mathbf{T}/N\mathbf{T})$; indeed, Γ maps to the trivial subgroup of $(\mathbf{Z}/N\mathbf{Z})^*$. The image in question is certainly a normal subgroup of $\mathbf{SL}(2, \mathbf{T}/N\mathbf{T})$ since X is normal in $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ and G_N contains $\mathbf{SL}(2, \mathbf{T} \otimes \mathbf{Z}_N)$. The ring $\mathbf{T}/N\mathbf{T}$ is a product of finite fields of characteristic N because $\mathbf{T}$ is unramified at N; intrinsically, $\mathbf{T}/N\mathbf{T} = \prod_{\mathfrak{m}} \mathbf{T}/\mathfrak{m}$, where $\mathfrak{m}$ runs over the maximal ideals of $\mathbf{T}$ which divide N.

Fix $\mathfrak{m}$ for the moment and let $\rho_{\mathfrak{m}} \colon \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathbf{GL}(2, \mathbf{T}/\mathfrak{m})$ be the mod $\mathfrak{m}$ reduction of the N-adic Galois representation ρ_N. This reduction is an irreducible two-dimensional representation because $\mathfrak{m}$ cannot be an Eisenstein prime; indeed, $\mathfrak{m}$ does not divide $N-1$. As we have seen in Proposition 2.2, $\rho_{\mathfrak{m}}$ cannot be "finite" (or *peu ramifiée*) in the sense of [18]; recall that the Main Theorem of [14] implies that $\rho_{\mathfrak{m}}$ would be modular of level 1 if it were finite. Thus $\rho_{\mathfrak{m}}$ is wildly ramified at N, so that the group $\rho_{\mathfrak{m}}(X)$ is non-trivial. But $\rho_{\mathfrak{m}}(X)$ is a normal subgroup of $\mathbf{SL}(2, \mathbf{T}/\mathfrak{m})$; we conclude that $\rho_{\mathfrak{m}}(X) = \mathbf{SL}(2, \mathbf{T}/\mathfrak{m})$.

Thus the image of $\rho_N(X)$ in $\mathbf{SL}(2, \mathbf{T}/N\mathbf{T}) = \prod_{\mathfrak{m}} \mathbf{SL}(2, \mathbf{T}/\mathfrak{m})$ is a normal subgroup of $\mathbf{SL}(2, \mathbf{T}/N\mathbf{T})$ which maps surjectively to each factor $\mathbf{SL}(2, \mathbf{T}/\mathfrak{m})$. By taking commutators with elements of the form $1 \times \cdots \times 1 \times g \times 1 \times \cdots \times 1$, we find that $\rho_N(X)$ maps surjectively to $\mathbf{SL}(2, \mathbf{T}/N\mathbf{T})$. Therefore, as was explained above, $\rho_N(X)$ contains $\mathbf{SL}(2, \mathbf{T} \otimes \mathbf{Z}_N)$. ■

Returning briefly to the group G_{f}, we note that we have

$$(\prod_\ell H_\ell) \subseteq G_{\mathrm{f}} \subseteq (\prod_\ell G_\ell),$$

where the products are taken over all prime numbers ℓ. A theorem of B. Kaskel [16, Th. 7.3] implies that the image of G_{f} in the group $G^N := \prod_{\ell \neq N} G_\ell$ is all of G^N. This suggests viewing the full product $\prod_\ell G_\ell$ as the binary product $G^N \times G_N$. Then G_{f} is a subgroup of this product which maps surjectively to each of the two factors. The group H_N may be viewed as the kernel of the projection map $G_{\mathrm{f}} \to G^N$; symmetrically, we let $H^N \subset G^N$ be the kernel of the second projection map. As is well known (see "Goursat's Lemma," an exercise in Bourbaki's Algèbre, Ch. I, §4), the projections from G_{f} onto G^N and G_N induce natural isomorphisms $G^N/H^N \approx G_{\mathrm{f}}/(H^N \times H_N)$ and $G_{\mathrm{f}}/(H^N \times H_N) \approx G_N/H_N$. We obtain as a consequence an isomorphism

$$\alpha \colon G^N/H^N \xrightarrow{\sim} G_N/H_N.$$

The group G_{f} contains $H^N \times H_N$ as a normal subgroup, and the image of G_{f} in

$$(G^N \times G_N)/(H^N \times H_N) = (G^N/H^N) \times (G_N/H_N)$$

is the graph of the isomorphism α.

It is worth remarking that G_{f} is open in $G^N \times G_N$ by [16, Th. 7.5]. Hence the groups H^N and H_N are open in G^N and G_N respectively. Thus the groups $G_f/(H^N \times H_N)$, G^N/H^N and G_N/H_N are finite groups which have the same order. The order of $(G^N \times G_N)/(H^N \times H_N)$ is the square of the orders of the three other groups. If N is prime to $\operatorname{disc} \mathbf{T}$, then the order of G_N/H_N is a divisor of $N-1$ by Prop. 6.4. Moreover as we will see below, the order of G_N/H_N is always divisible by Mazur's constant $n = \operatorname{num} \frac{N-1}{12}$.

Adopting a Galois-theoretic point of view, we let K be the subfield of $\overline{\mathbf{Q}}$ corresponding to the finite quotient $G_{\mathrm{f}}/(H^N \times H_N)$ of $\operatorname{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$. Let K_N be the extension of $\mathbf{Q}$ generated by the coordinates of the N-power torsion points on J and let K^N be the extension of $\mathbf{Q}$ which is defined similarly, using prime-to-N torsion points in place of N-power torsion points. Then the compositum $K_\infty = K^N K_N$ is the subfield of $\overline{\mathbf{Q}}$ corresponding to the quotient G_{f} of $\operatorname{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$, and it is clear that we have $\operatorname{Gal}(K_\infty/K_N) = H^N$ and $\operatorname{Gal}\left(K_\infty/K^N\right) = H_N$. Thus

$$G_{\mathrm{f}}/(H^N \times H_N) = G_N/H_N = G^N/H^N = \operatorname{Gal}(K/\mathbf{Q}).$$

What information do we have about K? We may restate Proposition 6.4 as follows: If $\mathbf{T}$ is unramified at N, then K is contained in the field of Nth roots of unity. Indeed, in that case, G_N/H_N is a quotient of

$$\{\, M \in \mathbf{GL}(2, \mathbf{T} \otimes \mathbf{Z}_N) \mid \det M \in \mathbf{Z}_N^* \,\}/\{\, M \in \mathbf{GL}(2, \mathbf{T} \otimes \mathbf{Z}_N) \mid \det M \in \Gamma \,\},$$

which corresponds (via the determinant) to the Galois group $\mathrm{Gal}\,(\mathbf{Q}(\mu_N)/\mathbf{Q})$. Without the assumption on $\operatorname{disc}\mathbf{T}$, we can remark, at least, that K is ramified only at N; it is a subfield of K_N, which is ramified only at N.

We now exhibit the lower bound for $[K : \mathbf{Q}]$ which was alluded to above, proving that K contains the unique subfield of $\mathbf{Q}(\mu_N)$ with degree n over $\mathbf{Q}$. (Since n is 1 only when $X_0(N)$ has genus 0, it follows that K is a non-trivial extension of $\mathbf{Q}$ whenever $J_0(N)$ is non-zero.) For this, we note first that K_N contains the field $\mathbf{Q}(\mu_N)$ of Nth roots of 1; indeed, K_N contains the field generated by the N-power roots of 1 in $\overline{\mathbf{Q}}$, since the determinant of ρ_N is the N-adic cyclotomic character. The Galois group $\mathrm{Gal}\,(\mathbf{Q}(\mu_N)/\mathbf{Q}) = (\mathbf{Z}/N\mathbf{Z})^*$ has a unique quotient of order n. As in the proof of Theorem 5.2, we refer to this quotient as Δ; field-theoretically, Δ corresponds to a Galois extension K_Δ of $\mathbf{Q}$ with

$$K_\Delta \subseteq \mathbf{Q}(\mu_N) \subset K_N.$$

Since $\mathrm{Gal}(K_\Delta/\mathbf{Q}) = \Delta$, $[K_\Delta : \mathbf{Q}] = n$.

Theorem 6.5. *The field K contains K_Δ.*

Proof. Let $\mathfrak{m}$ be an Eisenstein prime (i.e., maximal ideal) of $\mathbf{T}$; let ℓ be the corresponding divisor of n. The Tate module $\mathrm{Ta}_\mathfrak{m}$ which was introduced in the proof of Lemma 5.1 is free of rank two over $\mathbf{T}_\mathfrak{m}$, the completion of $\mathbf{T}$ at $\mathfrak{m}$. The action of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ on $\mathrm{Ta}_\mathfrak{m}$ is given by a representation

$$\rho_\mathfrak{m} \colon \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathbf{GL}(2, \mathbf{T}_\mathfrak{m})$$

whose determinant is the ℓ-adic cyclotomic character; if p is prime to ℓN, then the trace of $\rho_\mathfrak{m}(\mathrm{Frob}_p)$ is $T_p \in \mathbf{T}_\mathfrak{m}$, T_p being the pth Hecke operator. Taking the sum of the $\rho_\mathfrak{m}$ and then reducing mod $\mathscr{I}^2$, we obtain a representation

$$\rho \colon \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathbf{GL}(2, \mathbf{T}/\mathscr{I}^2)$$

with analogous properties. In particular, for each prime p prime to nN, the trace and determinant of $\rho(\mathrm{Frob}_p)$ are the images of T_p and p, respectively, in $\mathbf{T}/\mathscr{I}^2$.

Let $\eta\colon \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathbf{T}/\mathscr{I}^2$ be the function $1 + \det\rho - \operatorname{tr}\rho$. For p prime to nN, $\eta(\mathrm{Frob}_p)$ is the image in $\mathscr{I}/\mathscr{I}^2$ of the element $\eta_p = 1 + p - t_p$ of $\mathscr{I}$. In particular, the Cebotarev density theorem implies that η is a function $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathscr{I}/\mathscr{I}^2$.

As we recalled in the proof of Theorem 5.2, there is an isomorphism $\kappa\colon \mathscr{I}/\mathscr{I}^2 \xrightarrow{\sim} \Delta$ which satisfies the congruence formula

$$\kappa(\eta_p) = \tfrac{p-1}{2} \cdot \overline{p}$$

for all primes p not dividing nN. In this formula, $\overline{p}$ represents the image in Δ in the congruence class of p mod N. Let $\alpha\colon \operatorname{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathscr{I}/\mathscr{I}^2$ be the composite of: (1) the mod N cyclotomic character $\chi_N\colon \operatorname{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to (\mathbf{Z}/N\mathbf{Z})^*$; (2) the quotient map $(\mathbf{Z}/N\mathbf{Z})^* \to \Delta$; (3) the inverse of κ. Then we may write alternatively

$$\eta_p = \tfrac{p-1}{2} \cdot \alpha(p),$$

where the left-hand side is interpreted in $\mathscr{I}/\mathscr{I}^2$. If now $\chi = \chi_{2n}$ is the mod $2n$ cyclotomic character, then the formula for η_p and the Cebotarev density theorem imply the identity

$$\eta = \tfrac{\chi-1}{2} \cdot \alpha$$

of $\mathscr{I}/\mathscr{I}^2$-valued functions on $\operatorname{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$.

Let H be the kernel of $\rho \times \chi$. Then $\eta(hg) = \eta(g)$ for all $g \in \operatorname{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$, since $\rho(hg) = \rho(g)$ in that case. Let h be an element of H and take g to be a complex conjugation in $\operatorname{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$. Since $\chi(g) = -1$ and $\chi(h) = 1$, the equation $\eta(hg) = \eta(g)$ amounts to the identity $\alpha(hg) = \alpha(g)$. Since α is a homomorphism, we deduce that $\alpha(h) = 1$.

In other words, if $h \in \operatorname{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ is trivial under $\rho \times \chi$, then h is trivial in $\operatorname{Gal}(K_\Delta/\mathbf{Q})$. In particular, if $\rho^N(h) = 1$, then h fixes K_Δ. Accordingly, K_Δ is contained in the fixed field K^N of the kernel of ρ^N. Since, by construction, K_Δ is a subfield of K_N, K_Δ is contained in K. ■

Theorem 6.5, which will not be used in the proof of Theorem 1.6, suggests the problem of pinpointing K completely. According to Proposition 6.4 and Theorem 6.5, we have $K_\Delta \subseteq K \subseteq \mathbf{Q}(\mu_N)$ under the apparently mild assumption that N does not divide $\operatorname{disc}(\mathbf{T})$. Since $\operatorname{Gal}(\mathbf{Q}(\mu_N)/K_\Delta)$ is cyclic of order $(N-1)/n = \gcd(N-1, 12)$, to identify K under these circumstances is to calculate a divisor of $\gcd(N-1, 12)$, namely $[K : K_\Delta]$. In the cases where $X_0(N)$ has genus 0 (i.e., $N < 11$ and $N = 13$), we clearly have $K = \mathbf{Q} = K_\Delta$. In the case $N = 11$, K is constrained by our results to be either $\mathbf{Q}(\mu_{11})$ or the maximal real subfield of $\mathbf{Q}(\mu_{11})$. As was noted by Lang and Trotter [8] (see also [17, §5.3]), $K = \mathbf{Q}(\mu_{11})$ because the field generated by the 2-division points of $J_0(11)$ contains $\mathbf{Q}(\sqrt{-11})$. In the case $N = 37$, we have $\gcd(N-1, 12) = 12$, so that there are six a priori possibilities for K. In fact, Kaskel [7] shows that K is the maximal real subfield of $\mathbf{Q}(\mu_{37})$; the divisor in question is 6.

7. Proof of Theorem 1.6

We recall the statement to be proved: *Let P be a point of finite order on J whose ℓ-primary component P_ℓ is not rational point. Assume that at least one of the following statements is true: (1) N is prime to the order of P; (2) ℓ is prime to $N-1$; (3) N is prime to the discriminant of $\mathbf{T}$ (i.e., Hypothesis 1.4 holds). Then there is a $\sigma \in \operatorname{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ such that $\sigma P - P$ has order ℓ.*

In the proof that follows, we write P^ℓ for the sum of the p-primary components of P for primes different from ℓ. Thus $P = P_\ell + P^\ell$. Similarly, we put $P^N = P - P_N$.

Consider the extension $\mathbf{Q}(P_\ell)/\mathbf{Q}$, which is non-trivial by hypothesis. To orient the reader, we note that this extension can be ramified only at ℓ and at N, the latter prime being the unique prime of bad reduction of J. According to [2, Th. 2.2], $\mathbf{Q}(P_\ell)/\mathbf{Q}$ is automatically ramified at ℓ except perhaps when $\ell = 2$.

On the other hand, it is plausible that $\mathbf{Q}(P_\ell)/\mathbf{Q}$ is unramified at N. Let us first deal with this possibility, which turns out to be especially simple; here the hypotheses (1)–(3) are irrelevant. According to Proposition 3.3, P_ℓ lies in $J[\mathscr{I}]$. This latter group contains the Shimura subgroup Σ and the cuspidal group C of J. The source and target of the resulting natural map $\Sigma \oplus C \to J[\mathscr{I}]$ have order n; the kernel and cokernel of this map have order 1 if n is odd and order 2 if n is even.

To fix ideas, we assume for the moment that ℓ is an odd prime. Then P_ℓ lies in the ℓ-primary part of $J[\mathscr{I}]$, which is the direct sum of the ℓ-primary parts of Σ and C. Hence P_ℓ is the sum of a rational point of J and an element of ℓ-power order of $\Sigma \approx \mu_n$. Since P_ℓ is not rational, this latter element is non-trivial; its order may be written ℓ^a with $a \geq 1$. Let σ be an element of $\operatorname{Gal}\big(\overline{\mathbf{Q}}/\mathbf{Q}(\mu_{\ell^{a-1}})\big)$ which has non-trivial image in $\operatorname{Gal}\left(\mathbf{Q}(\mu_{\ell^a})/\mathbf{Q}(\mu_{\ell^{a-1}})\right)$. Then it is evident that $\sigma P_\ell - P_\ell$ has order ℓ on J. Indeed, this element is non-trivial since σ does not fix P_ℓ, but it is of order dividing ℓ since σ does fix ℓP_ℓ. Now the extension $\mathbf{Q}(\mu_{\ell^a})/\mathbf{Q}(\mu_{\ell^{a-1}})$ is ramified at ℓ; thus we may take σ to be in an inertia group for a prime of $\mathbf{Q}(\mu_{\ell^{a-1}})$ which lies over ℓ. This choice ensures that P^ℓ is fixed by σ. Then $\sigma P - P = \sigma P_\ell - P_\ell$ is a point of order ℓ, as desired.

Next, we suppose that $\ell = 2$; we continue to suppose that P_ℓ is unramified at N. Then $J[\mathscr{I}]$ has even order; i.e., n is even. If $P_\ell = P_2$ lies in $\Sigma + C$, then things proceed as in the case $\ell > 2$. However, as we recalled above, the sum $\Sigma + C$, which is not direct, represents a proper subgroup of $J[\mathscr{I}]$ (namely, one of index 2.) Hence we must discuss the case where P_2, which is a point in $J[\mathscr{I}]$, does not lie in the sum $\Sigma + C$.

In this case, the group $J[\mathscr{I}]$ is generated by its subgroup $\Sigma + C$ of index 2 together with the point P_2. Using Theorem 1.7, we find that

$$\mathbf{Q}(\mu_{2n}) = \mathbf{Q}(J[\mathscr{I}]) = K(P_2),$$

where $K = \mathbf{Q}(\Sigma + C) = \mathbf{Q}(\mu_n)$. The extension $\mathbf{Q}(\mu_{2n})/\mathbf{Q}(\mu_n)$ is a quadratic extension which is ramified at 2. We take σ in an inertia group for 2 which fixes K but not P_2. Since $2P_2$ lies in $\Sigma + C$, the difference $\sigma P_2 - P_2$ is of order 2. We have $\sigma P - P = \sigma P_2 - P_2$ in analogy with the situation already considered.

Having treated the relatively simple case where $\mathbf{Q}(P_\ell)/\mathbf{Q}$ is unramified at N, we assume from now on that P_ℓ is ramified at N. This assumption means that there is an inertia subgroup $I \subset \operatorname{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ for the prime N which acts non-trivially on P_ℓ. Hence there is a $\tau \in I$ such that the order of $\tau P - P$ is divisible by ℓ. We seek to construct a $\sigma \in I$ for which $\sigma P - P$ has order precisely ℓ.

Assume first that (1) holds, i.e., that the order of P is prime to N. Let m be this order, and let ℓd be the order of $\tau P - P$; thus, ℓd divides m. Recall the exact sequence of I-modules

$$0 \to \operatorname{Hom}(\mathscr{X}/m\mathscr{X}, \mu_m) \to J[m] \to \mathscr{X}/m\mathscr{X} \to 0. \tag{2.4}$$

Since m is prime to N, the two flanking groups are unramified. It follows, as is well known, that $A := \tau - 1$ acts on $J[m]$ as an endomorphism with square 0. By the binomial theorem, we find the equation $\tau^d = 1 + dA$ in $\operatorname{End} J[m]$. Therefore

$$\tau^d P - P = dAP = d(\tau - 1)P = d(\tau P - P)$$

is a point of order ℓ. We take $\sigma = \tau^d$.

Next, assume that (2) holds. Arguing as above, we may find an $s \in I$ such that $sP^N - P^N$ has order ℓ. Moreover, for each $i \geq 1$, we have $s^i P^N - P^N = i(sP^N - P^N)$. Consider again (2.4), with m replaced by m', the order of P_N. Let $j = \phi(m')$ (Euler ϕ-function). Then s^j acts trivially on the groups $\operatorname{Hom}(\mathscr{X}/m'\mathscr{X}, \mu'_m)$ and $\mathscr{X}/m'\mathscr{X}$ in (2.4), so that $s^{jm'}$ fixes P_N. By (2), j is prime to ℓ, and thus $i := jm'$ is prime to ℓ as well. Taking $\sigma = s^i$, we find that $\sigma P - P$ has order ℓ, as required.

We now turn to the most complicated case, that where (3) holds, but where (1) and (2) are no longer assumed. We change notation slightly, writing m (rather than m') for the order of P_N. Thus m is a power of N. Let s again be an element of I such that $sP^N - P^N$ has order ℓ.

We fix our attention once again on (2.4), which we view as a sequence of I-modules. Concerning the Hecke action, we note that the two groups

$$M := \operatorname{Hom}(\mathscr{X}/m\mathscr{X}, \mu_m), \quad M' := \mathscr{X}/m\mathscr{X}$$

are each free of rank 1 over $\mathbf{T}/m\mathbf{T}$ in view of Theorem 2.3 and the fact that $\mathbf{T}$ is Gorenstein away from the prime 2. The central group $J[m]$ is free of rank 2 over $\mathbf{T}/m\mathbf{T}$ because of [9, Ch. II, Cor. 15.2]. The inertia group I acts trivially on $\mathscr{X}$ and as the mod m cyclotomic character χ on μ_m. Thus M' is unramified, and M is ramified if m is different from 1.

We will be interested in the value of $\chi(s) \in (\mathbf{Z}/m\mathbf{Z})^*$. Let i be the prime-to-ℓ part of the order of $\chi(s)$, and replace s by s^i. After this replacement, the order of $\chi(s)$ is a power of ℓ. Also, as we have discussed, this replacement multiplies $sP^N - P^N$ by i. Since i is prime to ℓ, $sP^N - P^N$ remains of order ℓ.

If $\chi(s)$ is now 1, then the situation is similar to that which we just discussed. Namely, s^m is the identity on $J[m]$, and we may take $\sigma = s^m$.

Assume now that $\chi(s)$ is different from 1; thus $\chi(s)$ is a non-trivial ℓ-power root of 1. In this case, the $\mathbf{T}$-module $J[m]$ is the *direct sum* of two subspaces: the space where s acts as 1 and the space where s acts as $\chi(s)$ (which is not congruent to 1 mod N). Indeed, the endomorphism $\dfrac{s - \chi(s)}{1 - \chi(s)}$ of $J[m]$ is zero on $M = \operatorname{Hom}(\mathscr{X}/m\mathscr{X}, \mu_m)$ and the identity on $M' = \mathscr{X}/m\mathscr{X}$. It splits the exact sequence which is displayed above, giving us an isomorphism of $\mathbf{T}$-modules:

$$J[m] \approx M \oplus M'.$$

The module M', viewed as a submodule of $J[m]$, is the fixed part of $J[m]$ relative to the action of s.

We claim that there is an $h \in \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ such that $hP^N = P^N$ and such that $hP_N \in M'$. This claim will prove what is wanted, since the choice $\sigma = h^{-1}sh$ will guarantee that the difference $\sigma P - P$ is the ℓ-division point

$$h^{-1}(shP^N - hP^N) + h^{-1}(shP_N - hP_N) = h^{-1}(sP^N - P^N).$$

To find the desired h it suffices to produce an element of $\mathbf{SL}_{\mathbf{T}/m\mathbf{T}}J[m] \approx \mathbf{SL}(2, \mathbf{T}/m\mathbf{T})$ which maps P_N into M'. Indeed, Proposition 6.4 implies that all such elements arise from H_N, i.e., from elements of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ which fix torsion points of J with order prime to N.

To produce the required element of $\mathbf{SL}(2, \mathbf{T}/m\mathbf{T})$, we work explicitly. Choose $\mathbf{T}/m\mathbf{T}$-bases e' and e of the free rank 1 modules M' and M, and use $\{e', e\}$ as a basis of $J[m]$. Then M' is the span of the vector $(1,0)$ and M is the span of $(0,1)$. Let u and v be the coordinates of P_N relative to the chosen basis. We must exhibit a matrix in $\mathbf{SL}(2, \mathbf{T}/m\mathbf{T})$ which maps (u,v) to a vector with second component 0.

Because of the hypothesis that N is prime to $\operatorname{disc} \mathbf{T}$, $\mathbf{T} \otimes \mathbf{Z}_N$ is a finite product of rings of integers of finite unramified extensions of $\mathbf{Q}_N$. Thus $\mathbf{T}/m\mathbf{T}$ is a product of rings of the form $R = \mathscr{O}/m\mathscr{O}$, where $\mathscr{O}$ is the ring of integers of a finite unramified extension of $\mathbf{Q}_N$. It suffices to solve our problem factor by factor: given $(u,v) \in R^2$, we must find an element of $\mathbf{SL}(2, R)$ which maps (u,v) into the line generated by $(1,0)$. It is clear that we may write (u,v) in the form $N^t(u',v')$, where t is a non-negative integer and at least one of u', v' is a unit in R. Solving the problem for $(u'v')$ solves it for (u,v), so we may, and do, assume that either u of v is a unit.

If u is a unit, then

$$\begin{pmatrix} 1 & 0 \\ -vu^{-1} & 1 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} u \\ 0 \end{pmatrix}.$$

If v is a unit then $\begin{pmatrix} 0 & 1 \\ 1 & -uv^{-1} \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} v \\ 0 \end{pmatrix}$. ■

References

1. N. Boston, H. W. Lenstra, Jr., and K. A. Ribet, *Quotients of group rings arising from two-dimensional representations*, C. R. Acad. Sci. Paris Sér. I Math. **312** (1991), 323–328.

2. R. Coleman, B. Kaskel, and K. Ribet, *Torsion points on $X_0(N)$*, Contemporary Math. (to appear).

3. J. A. Csirik, *The Galois structure of $J_0(N)[I]$* (to appear).

4. P. Deligne and M. Rapoport, *Les schémas de modules de courbes elliptiques*, Lecture Notes in Math., vol. 349, Springer-Verlag, Berlin and New York, 1973, pp. 143–316.

5. E. De Shalit, *A note on the Shimura subgroup of $J_0(p)$*, Journal of Number Theory **46** (1994), 100–107.

6. A. Grothendieck, *SGA 7 I, Exposé IX*, Lecture Notes in Math., vol. 288, Springer-Verlag, Berlin and New York, 1972, pp. 313–523.

7. B. Kaskel, *The adelic representation associated to $X_0(37)$*, PhD. thesis, UC Berkeley, May, 1996.

8. S. Lang and H. Trotter, *Frobenius distributions in $\mathbf{GL}_2$-extensions*, Lecture Notes in Math., vol. 504, Springer-Verlag, Berlin and New York, 1976.

9. B. Mazur, *Modular curves and the Eisenstein ideal*, Publ. Math. IHES **47** (1977), 33–186.

10. ______, *Rational isogenies of prime degree*, Invent. Math. **44** (1978), 129–162.

11. A. Ogg, *Hyperelliptic modular curves*, Bull. Soc. Math. France **102** (1974), 449–462.

12. ______, *Automorphismes de courbes modulaires*, Sém. Delange-Pisot-Poitou (1974/1975, exposé 7).

13. M. Raynaud, *Courbes sur une variété abélienne et points de torsion*, Invent. Math. **71** (1983), 207–233.

14. K. A. Ribet, *On modular representations of* $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ *arising from modular forms*, Invent. Math. **100** (1990), 431–476.

15. ______, *Report on mod* ℓ *representations of* $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$, Proceedings of Symposia in Pure Mathematics **55 (2)** (1994), 639–676.

16. ______, *Images of semistable Galois representations*, Pacific Journal of Math. **81** (1997), 277–297.

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

18. ______, *Sur les représentations modulaires de degré 2 de* $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$, Duke Math. J. **54** (1987), 179–230.

19. G. Shimura, *A reciprocity law in non-solvable extensions*, Journal für die reine und angewandte Mathematik **221** (1966), 209–220.

20. J. T. Tate, *The non-existence of certain Galois extension of* $\mathbf{Q}$ *unramified outside 2*, Contemporary Mathematics **174** (1994), 153–156.

21. L. C. Washington, *Introduction to cyclotomic fields*, Graduate Texts in Math., vol. 83 (second edition), Springer-Verlag, Berlin-Heidelberg-New York, 1997.