Journal of Number Theory **92**, 348–375 (2002)
doi:10.1006/jnth.2001.2709, available online at http://www.idealibrary.com on IDEAL®

The Kernel of the Eisenstein Ideal

János A. Csirik[1]

AT&T Labs—Research, P.O. Box 971, 180 Park Avenue, Florham Park, New Jersey 07932
E-mail: janos@research.att.com

Communicated by D. Goss

Received September 2, 2000

Let N be a prime number, and let $J_0(N)$ be the Jacobian of the modular curve $X_0(N)$. Let $\mathbf{T}$ denote the endomorphism ring of $J_0(N)$. In a seminal 1977 article, B. Mazur introduced and studied an important ideal $I \subseteq \mathbf{T}$, the Eisenstein ideal. In this paper we give an explicit construction of the kernel $J_0(N)[I]$ of this ideal (the set of points in $J_0(N)$ that are annihilated by all elements of I). We use this construction to determine the action of the group $\mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q})$ on $J_0(N)[I]$. Our results were previously known in the special case where $N-1$ is not divisible by 16.

Key Words: modular curves; Eisenstein ideal.

1. INTRODUCTION

Let N be a prime number and let $J_0(N)$ denote the Jacobian of the modular curve $X_0(N)$. The variety $J_0(N)$ possesses certain naturally defined endomorphisms T_ℓ (for all primes $\ell \neq N$) and w. These endomorphisms together with $\mathbf{Z}$ (the multiplications by integers) generate the Hecke ring $\mathbf{T}_N$ of endomorphisms of $J_0(N)$. In his celebrated article "Modular curves and the Eisenstein ideal" [10], Mazur defined the Eisenstein ideal I in $\mathbf{T}_N$ as the ideal generated by $1+w$ and the $1+\ell-T_\ell$ and used it to identify the possible rational torsion subgroups of elliptic curves defined over the rational numbers. The Galois module $J_0(N)(\bar{\mathbf{Q}})[I]$ plays an important role in [10] and later studies of the arithmetic geometry of the curve $X_0(N)$.

Mazur proved that

$$J_0(N)[I] \cong \mathbf{Z}/n\mathbf{Z} \times \mathbf{Z}/n\mathbf{Z}$$

[1] This paper is based on the author's 1999 UC Berkeley thesis.

0022-314X/02 $35.00

as groups, for $n = (N-1)/\gcd(N-1, 12)$. In this paper we will study the action of the group $\mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q})$ on $J_0(N)[I]$. The group $J_0(N)[I]$ has two noteworthy Galois-invariant subgroups. The cuspidal subgroup C is generated by the divisor $c = 0 - \infty$ (the formal difference of the two cusps of $X_0(N)$). The group C is cyclic of order n and is pointwise fixed by $\mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q})$. The Shimura subgroup Σ is a finite flat subgroup scheme of $J_0(N)$ such that

$$\Sigma(\bar{\mathbf{Q}}) = \ker(\beta^*: J_0(N) \to J_1(N)),$$

where β^* is induced by the usual degeneracy map $\beta: X_1(N) \to X_0(N)$. The group Σ is also cyclic of order n, but is isomorphic to μ_n as a group scheme.

In this paper we shall give an explicit construction of $J_0(N)[I]$, the proof of which relies on an informed computation, and apply the construction in various ways. Mazur's paper [10] contains an explicit construction of $J_0(N)[I]$ only in the case $N \not\equiv 1 \pmod{16}$, although he remarks in a few places (e.g., [10, p. 130]) that a general description would be desirable. Our construction identifies the action of $\mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q})$ on $J_0(N)[I]$.

If n is odd (equivalently $N \not\equiv 1 \pmod 8$) then $C \cap \Sigma = 0$, so $J_0(N)[I] \cong C \oplus \Sigma$ and therefore we know the Galois action on $J_0(N)[I]$.

If n is even then $C \cap \Sigma \neq 0$ and more is needed to find the Galois action. In this case $C + \Sigma$ has index 2 in $J_0(N)[I]$. Therefore it suffices to find an "extra" point P in $J_0(N)[I]$ that is not in $C + \Sigma$. The knowledge of the Galois action on P, Σ and C then gives a description of the $\mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q})$-action of $J_0(N)[I]$. For the case $n \equiv 2 \pmod 4$ (or equivalently $N \equiv 9 \pmod{16}$), Mazur finds P by considering the Nebentypus covering $X_0^{\#}(N) \to X_0(N)$ of degree 2. Using a function constructed by Ogg and Ligozat, he obtains a divisor d on $X_0^{\#}(N)$ which turns out to be the pullback of a certain divisor on $X_0(N)$ that gives the extra point on $J_0(N)$.

This paper uses other coverings $X_0^{\#}(N) \to X_0(N)$ to generalize Mazur's construction and find extra points of $J_0(N)[I]$ for any $N \equiv 1 \pmod 8$. To find suitable divisors on our modular curves $X_0^{\#}(N)$, we use the theory of modular units: rational functions on a modular curve whose divisors are concentrated at the cusps. Our coverings $X_0^{\#}(N) \to X_0(N)$ are all intermediate to $X_1(N) \to X_0(N)$, enabling us to rely on the theory of modular units on $X_1(N)$. The units of $X(N)$ are treated in Kubert and Lang's [8]. We recall some of their results in Section 2. We then use the results of Section 2 to develop some results about the units of $X_1(N)$ in Section 3. (References [7, 8] also treat this case but restrict their attention to units whose divisors are supported at the rational cusps, and do not explicitly give the data necessary for the descent to $X_0^{\#}(N)$.) In Section 4, we construct a divisor on

$X_0^{\#}(N)$ and establish properties of the divisor that make our later arguments work. In Section 5, we prove that the extra point we obtain is in $J_0(N)[I]$ and use this fact to prove the following theorem, conjectured by Ribet. For any positive integer k, let χ_k denote the kth cyclotomic character $\mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q}) \to (\mathbf{Z}/k\mathbf{Z})^\times$ obtained via the identification $\mathrm{Gal}(\mathbf{Q}(\mu_k)/\mathbf{Q})) \cong (\mathbf{Z}/k\mathbf{Z})^\times$.

THEOREM 1.1. *$J_0(N)[I]$ has a basis e_1, e_2 over $\mathbf{Z}/n\mathbf{Z}$ such that*

(a) $c = e_1 + 2e_2$;

(b) *e_1 generates Σ;*

(c) *$\sigma \in \mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q})$ acts via left multiplication by*

$$\begin{pmatrix} \chi_n(\sigma) & (1-\chi_{2n}(\sigma))/2 \\ 0 & 1 \end{pmatrix}$$

with respect to the given basis $e_1 = \binom{1}{0}$, $e_2 = \binom{0}{1}$.

The results of this paper can also be used to clarify arguments in [15] and in [1]. Another application is determining the old subvariety of the modular Jacobian $J_0(NM)$, where N and M are distinct primes. This is described in [3].

2. NOTATION AND SETUP

For any non-zero rational number x, let $\mathrm{num}(x)$ denote the numerator of x, that is, the smallest positive integer n such that n/x is an integer.

We will now briefly summarize the relevant properties of the modular curves we will be using. The reader can find a thorough treatment of these in [5], as well as in the references cited below.

Let N be a positive integer. We shall consider the usual modular curves $X_0(N)$, $X_1(N)$ and $X(N)$, and their Jacobians $J_0(N)$, $J_1(N)$ and $J(N)$. These correspond to the moduli problems of classifying an elliptic curve with a cyclic subgroup of order N, an elliptic curve with a point of order N, and an elliptic curve with a symplectic identification of its N-torsion with $\mu_N \times \mathbf{Z}/N\mathbf{Z}$, respectively. These curves are all defined over $\mathbf{Q}$, as are the usual degeneracy maps (which are Galois coverings) $\alpha: X(N) \to X_1(N)$, $\beta: X_1(N) \to X_0(N)$ and $\gamma = \beta \circ \alpha$.

The curves $X_0(N)_{\mathbf{C}}$, $X_1(N)_{\mathbf{C}}$ and $X(N)_{\mathbf{C}}$ can also be regarded as compactified quotients of the complex upper half plane $\mathscr{H}^*/\Gamma_0(N)$, $\mathscr{H}/\Gamma_1(N)$ and $\mathscr{H}^*/\Gamma(N)$, respectively, where

$$\Gamma_0(N) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2\mathbf{Z} : c \equiv 0 \pmod{N} \right\},$$

$$\Gamma_1(N) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2\mathbf{Z} : c \equiv 0, d \equiv 1 \pmod{N} \right\},$$

$$\Gamma(N) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2\mathbf{Z} : a \equiv 1, b \equiv 0, c \equiv 0, d \equiv 1 \pmod{N} \right\},$$

and these subgroups of $\mathrm{SL}_2\mathbf{Z}$ act on the complex upper half plane $\mathscr{H}$ via fractional linear transformations. The points introduced during the compactification are called cusps.

Now let N be an odd prime number, and let

$$r = (N-1)/2.$$

The curve $X_0(N)$ has two cusps, denoted 0 and ∞. They are both defined over $\mathbf{Q}$ and are distinguished by the fact that under the natural map $X_0(N)_{\mathbf{C}} = \mathscr{H}^*/\Gamma_0(N) \to X(1)_{\mathbf{C}} = \mathscr{H}^*/\mathrm{SL}_2\mathbf{Z}$, the cusp 0 is ramified with index N and the cusp ∞ is unramified.

The curve $X_1(N)$ has $N-1$ cusps that come in two groups. We shall use Klimek's notation in [7] for them. The cusps $P_1, P_2, \ldots, P_r$ are defined over $\mathbf{Q}$ and are mapped to 0 under β: $X_1(N) \to X_0(N)$. The cusps $Q_1, Q_2, \ldots, Q_r$ are defined over $\mathbf{Q}(\mu_N)^+$ (the maximal totally real subfield of the Nth cyclotomic field) and are mapped to ∞ under β. All the cusps of $X_1(N)$ are unramified with respect to β.

The curve $X(N)$ has $(N^2-1)/2$ cusps and we use Shimura's notation in [13] to regard them as pairs $\pm\binom{x}{y}$ with $x, y \in \mathbf{F}_N$, not both equal to 0. In this representation, $\mathrm{Gal}(X(N)_{\mathbf{C}}/X(1)_{\mathbf{C}}) \cong \mathrm{PSL}_2\mathbf{F}_N$ acts naturally from the left. For $1 \leqslant i \leqslant r$, the cusps $\binom{*}{i}$ are all defined over $\mathbf{Q}(\mu_N)$ and map unramifiedly to P_i under α: $X(N) \to X_1(N)$. For $1 \leqslant i \leqslant r$, the cusps $\binom{i}{0}$ are all defined over $\mathbf{Q}(\mu_N)^+$ and map to Q_i under α with ramification index N.

Shimura's notation can be used to label the cusps of any modular curve. We shall now provide the translations to Shimura's system of all the names we use. On the curve $X_0(N)$, the cusps 0 and ∞ (respectively) are called $\binom{0}{1}$ and $\binom{1}{0}$ (respectively). On the curve $X_1(N)$, for any $1 \leqslant t \leqslant r$, our notation P_t corresponds to $\binom{0}{t}$, while Q_t corresponds to $\binom{t}{0}$.

Recall that a *unit* of a modular curve is a rational function on the curve that has its divisor concentrated at the cusps. (It is a unit of the ring of rational maps from the noncuspidal points of the curve to the affine line.) In [8], Kubert and Lang determined all the units of $X(N)$. We briefly recall their results here, using $\mathbf{Z}/N\mathbf{Z} \times \mathbf{Z}/N\mathbf{Z}$ as the indexing group instead of their $\frac{1}{N}\mathbf{Z}/\mathbf{Z} \times \frac{1}{N}\mathbf{Z}/\mathbf{Z}$. Let $e = (e_1, e_2)$ be a pair of integers such that not

both of e_1 and e_2 are divisible by N. One can use the classical Weierstrass σ and Dedekind η functions to define the Klein form $k_e(\tau)$ on $\mathscr{H}$. This form enjoys the properties

$$\forall \alpha = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2\mathbf{Z}, \qquad k_e(\alpha\tau) = (c\tau + d)^{-1} k_{e\alpha}(\tau) \tag{K1}$$

(where $e\alpha$ denotes usual matrix multiplication) and

$$\forall f = (f_1, f_2) \in N\mathbf{Z} \times N\mathbf{Z}, \qquad k_{e+f}(\tau) = \varepsilon(e, f)\, k_e(\tau), \tag{K2}$$

where

$$\varepsilon(e, f) = (-1)^{\frac{f_1 f_2}{N^2} + \frac{f_1}{N} + \frac{f_2}{N}} \exp\left(\frac{\pi i}{N^2}(e_1 f_2 - e_2 f_1)\right).$$

These Klein forms are then used to define for all $e = (e_1, e_2) \in (\mathbf{Z} \times \mathbf{Z}) \setminus (N\mathbf{Z} \times N\mathbf{Z})$ the *Siegel function*

$$g_e(\tau) = k_e(\tau)\, \eta^2(\tau).$$

Recall that

$$\forall \alpha = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2\mathbf{Z}, \qquad \eta^2(\alpha\tau) = \psi(\alpha)\, (c\tau + d)\, \eta^2(\tau), \tag{N}$$

where ψ is defined by its values on the two standard generators of $\mathrm{SL}_2\mathbf{Z}$ as

$$\psi\left(\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}\right) = \exp\left(\frac{\pi i}{6}\right), \qquad \psi\left(\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\right) = \exp\left(\frac{\pi i}{2}\right) = i.$$

Now [8, Chap. 4, Theorem 1.3] says

THEOREM 2.2. *The units of $X(N)_{\mathbf{C}}$ are exactly the functions of the form*

$$g = c \prod_{e \in E} g_e(\tau)^{m(e)},$$

for some constant c and some finite set $E \subseteq \mathbf{Z} \times \mathbf{Z}$, where the $m(e)$ satisfy the conditions

$$\sum_{e \in E} m(e) \equiv 0 \pmod{12}, \tag{U1}$$

$$\sum_{e = (e_1, e_2) \in E} e_1^2 m(e) \equiv 0 \pmod{N}, \tag{U2}$$

$$\sum_{e=(e_1,e_2)\in E} e_1 e_2 m(e) \equiv 0 \pmod{N}, \tag{U3}$$

$$\sum_{e=(e_1,e_2)\in E} e_2^2 m(e) \equiv 0 \pmod{N}. \tag{U4}$$

The order of such a g at a cusp $P=\binom{x}{y}$ of $X(N)$ is as follows. Pick $\alpha\in \mathrm{PSL}_2\mathbf{F}_N$ such that $\alpha\binom{1}{0}=P$, and let $\hat{\alpha}$ be a lift of α to $\mathrm{SL}_2\mathbf{Z}$. Let $(c_1(e), c_2(e))=e\hat{\alpha}$ be the components of the usual matrix product of e and $\hat{\alpha}$; thus, we may take $c_1(e)=xe_1+ye_2$. Then

$$\mathrm{ord}_P(g)=\sum_{e\in E} m(e)\frac{N}{2}B_2\left(\frac{c_1(e) \bmod N}{N}\right), \tag{1}$$

where $B_2(X)=X^2-X+1/6$ is the second Bernoulli polynomial, and we used $x \bmod N$ to denote the smallest non-negative residue of x modulo N. For details and a derivation using the q-expansion of g, see [8, Chap. 2, Sect. 3].

Let $e\in E$ and $\alpha\in \mathrm{SL}_2\mathbf{Z}$. From (K1) and (N), we conclude that

$$g_e(\alpha\tau)=\psi(\alpha)\, g_{e\alpha}(\tau), \tag{2}$$

so using (U1) and the fact that $\psi(\alpha)^{12}=1$ for any α, for $g(\tau)=c\prod_{e\in E} g_e(\tau)^{m(e)}$ we have

$$g(\alpha\tau)=c\prod_{e\in E} g_{e\alpha}(\tau)^{m(e)}. \tag{3}$$

By (K2), if $e\equiv e'(\mathrm{mod}\, N)$, then $g_e/g_{e'}$ is a root of unity. By (K1) with $\alpha=\left(\begin{smallmatrix}-1 & 0\\ 0 & -1\end{smallmatrix}\right)$, if $e+e'=(0,0)$, then $g_e/g_{e'}=-1$ (here we also used the fact that with this α, the ψ and the $c\tau+d$ factors in (N) are both -1, and so they multiply to 1). Hence all units of $X(N)_{\mathbf{C}}$ can be put into the form $g=c'\prod_{e\in E'} g_e(\tau)^{m(e)}$, with

$$\begin{aligned} E'=\quad \{\quad & (0,1),\ (0,2),\ \ldots,\ (0,r),\\ & (1,0),\ (1,1),\ (1,2),\ \ldots,\ (1,N-1),\\ & (2,0),\ (2,2),\ (2,2),\ \ldots,\ (2,N-1),\\ & \vdots\\ & (r,0),\ (r,1),\ (r,2),\ \ldots,\ (r,N-1)\}. \end{aligned}$$

Kubert and Lang [8, Chap. 5, Theorem 3.1] then prove that the degree zero divisors on $X(N)$ concentrated at the cusps span a finite subgroup of

the divisor class group (this was also proved in general for all modular curves by Manin and Drinfeld, see [9, footnote to Corollary 3.6]). The number of cusps on $X(N)$ is $(N^2-1)/2$, which is also the cardinality of the set E'. Since the divisor of a rational function has degree 0, this implies that the functions g_e with $e \in E'$ are independent except for a single relation. A simple calculation using (1) shows that $\prod_{e \in E'} g_e(\tau)$ is a constant, providing the sought-after relation.

To sum up, we have shown that

Fact 2.3. The function $g = c \prod_{e \in E'} g_e(\tau)^{m(e)}$ is constant if and only if the $m(e)$ are the same for all $e \in E'$.

3. THE UNITS OF $X_1(N)$

We now determine units of $X_1(N)_{\mathbf{C}}$. They can be identified with the units of $X(N)_{\mathbf{C}}$ that are invariant under $\mathrm{Gal}(X(N)_{\mathbf{C}}/X_1(N)_{\mathbf{C}})$, which is generated by $T = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right) \in \mathrm{Gal}(X(N)_{\mathbf{C}}/X(1)_{\mathbf{C}}) \cong \mathrm{PSL}_2\mathbf{F}_N$. Therefore the units of $X_1(N)$ can be determined from the knowledge of the units of $X(N)$ and their transformation properties under T.

DEFINITION 3.1. For all $1 \leqslant i \leqslant r$, let

$$g_i(\tau) = g_{(o,\,i)}(\tau)$$

and

$$s_i(\tau) = \prod_{j=0}^{N-1} g_{(i,\,j)}(\tau).$$

THEOREM 3.4. *The units of $X_1(N)_{\mathbf{C}}$ are exactly the functions of the form*

$$g(\tau) = c \prod_{i=1}^{r} g_i(\tau)^{c_i} s_i(\tau)^{d_i},$$

where c is a constant and the c_i and d_i satisfy

$$\sum_{i=1}^{r} c_i + N \sum_{i=1}^{r} d_i \equiv 0 \pmod{12}, \tag{V1}$$

$$\sum_{i=1}^{r} i^2 c_i \equiv 0 \pmod{N}, \tag{V2}$$

$$\sum_{i=1}^{r} i^2 d_i \equiv 0 \pmod{N}, \tag{V3}$$

Proof. Let $g(\tau) = c \prod_{e \in E'} g_e(\tau)^{m(e)}$ be a unit of $X(N)$. By determining the conditions the $m(e)$ must satisfy to be invariant under T, we can find a criterion for g to be a unit of $X_1(N)$. Assume then that g is invariant under the action of T. By (3) and Fact 2.3, $m: E' \to \mathbf{Z}$ must be constant on the orbits of T (acting on E' from the right). Since $(e_1, e_2)\, T = (e_1, e_1 + e_2)$, this shows that g can be written as a product of g_i and s_i as above, with $c_i = m((0, i))$ and $d_i = m((i, 0)) = m((i, 1)) = \cdots = m((i, N-1))$.

Condition (U1) translates immediately to (V1).

The condition (U2) translates as

$$\sum_{e \in E'} e_1^2 m(e) = N \sum_{i=1}^{r} i^2 d_i \equiv 0 \qquad (\mathrm{mod}\ N),$$

so it is necessarily satisfied. The condition (U3) translates as

$$\sum_{e \in E'} e_1 e_2 m(e) = \sum_{i=1}^{r} i \sum_{j=0}^{N-1} j d_i = \sum_{i=1}^{r} iNr d_i \equiv 0 \qquad (\mathrm{mod}\ N),$$

so it is also necessarily satisfied. Last, (U4) in our case is

$$\sum_{e \in E'} e_2^2 m(e)$$
$$= \sum_{i=0}^{r} i^2 c_i + \sum_{j=0}^{r} \sum_{i=0}^{N-1} i^2 d_j = \sum_{i=0}^{r} i^2 c_i + N \frac{(2N-1)\, r}{3} \sum_{j=0}^{r} d_j \equiv 0 \qquad (\mathrm{mod}\ N).$$

Since N is not divisible by 3, $r(2N-1)$ must be, so (U4) translates to (V2).

It remains to check that our function g is actually invariant under the action of T, as opposed to being invariant only up to multiplication by a constant. This is not automatic, despite the fact that (3) has no extra constant factors. Indeed, by restricting our indexing set to E' we sometimes have to convert some g_e (with $e \notin E'$) occurring in (3) to some $g_{e'}$ with $e' \in E'$, thereby introducing a factor of a root of 1.

Using (K2), we can see that when we are acting by T on our g, this factor will be

$$\prod_{j=1}^{r} \prod_{k=0}^{j-1} \varepsilon((j, k), (0, N))^{d_j} = \prod_{j=1}^{r} \prod_{k=0}^{j-1} (-1)^{d_j} \exp\left(\frac{\pi i}{N} j d_j\right)$$
$$= \prod_{j=1}^{r} (-1)^{j d_i} \exp\left(\frac{\pi i}{N} j^2 d_j\right).$$

Hence to ensure invariance, we need

$$N \sum_{j=1}^{r} jd_j + \sum_{j=1}^{r} j^2 d_j \equiv 0 \qquad (\mathrm{mod}\ 2N).$$

The condition mod 2 is automatic since N is odd and $j \equiv j^2 (\mathrm{mod}\ 2)$. The condition mod N is just (V3).

It is also clear from the above calculations that any such g as given in the statement of the theorem satisfies (U1)–(U4) and hence is actually a T-invariant unit of $X(N)$, so we have completed the proof. ▮

Remark. For any i, the Atkin–Lehner involution (associated to the matrix $\left(\begin{smallmatrix} 0 & 1 \\ -N & 0 \end{smallmatrix}\right)$) interchanges the functions g_i and s_i up to constants. (For more details, see the proof of Theorem 3.6.) Therefore, we expect the conditions (V1-3) to be invariant under exchanging the c_is and the d_is. This is clear for (V2) and (V3), but it is also true for (V1). Indeed, since $(N, 6) = 1$, we have $N^2 = 1 (\mathrm{mod}\ 12)$, and hence

$$12 \mid C + ND \Leftrightarrow 12 \mid N^2 C + ND \Leftrightarrow 12 \mid NC + D,$$

where we used the notation $C = \sum_{i=1}^{r} c_i$ and $D = \sum_{i=1}^{r} d_i$.

THEOREM 3.5. *The order of such a function g (mentioned in Theorem 3.4) at the cusps is*

$$\mathrm{ord}_{P_i}(g) = \sum_{i=1}^{r} \left(\frac{c_i N}{2} B_2\left(\frac{it \bmod N}{N} \right) + \frac{d_i}{12} \right)$$

$$\mathrm{ord}_{Q_i}(g) = \sum_{i=1}^{r} \left(\frac{c_i}{12} + \frac{d_i N}{2} B_2\left(\frac{it \bmod N}{N} \right) \right).$$

Proof. A straightforward calculation using (1) and the fact that $\sum_{i=1}^{r} B_2(i/N) = -r/(6N)$ gives the order of g at the cusps of $X(N)$. Then, using the ramification indices of the cusps of $X(N)$ over the cusps of $X_1(N)$, we obtain our result. ▮

Finally, we give the transformation properties of our functions under the Galois group of $X_1(N)_{\mathbf{C}}$ over $X_0(N)_{\mathbf{C}}$.

DEFINITION 3.2. As in [13, Sect. 2], for a odd positive integer u, let

$$\{\cdot\}_u : \mathbf{Z} \to \{0, 1, \ldots, (u-1)/2\}$$

be the function defined by

$$\{a\}_u \equiv \pm a \qquad (\mathrm{mod}\ u).$$

For brevity, let $\{\cdot\}$ denote $\{\cdot\}_N$.

THEOREM 3.6. *For any $b \in \{1, \ldots, r\}$ and $\alpha = \left(\begin{smallmatrix} s & f \\ Nh & t \end{smallmatrix}\right) \in \Gamma_0(N)$, we have*

$$g_b(\alpha\tau) = \psi(\alpha)\, \kappa(\alpha; b)\, g_{\{bt\}}(\tau),$$

with

$$\kappa(\alpha; b) = (-1)^{bh} \exp\left(\pi i \left(-\frac{b^2 ht}{N}\right)\right)(-1)^{\lfloor bt/N \rfloor}.$$

Although this fact will not be needed later, for reference we state that for any $a \in \{1, \ldots, r\}$ and a as above, we have

$$s_a(\alpha\tau) = \psi(\alpha)^N\, \kappa'(\alpha; a)\, s_{\{as\}}(\tau),$$

with

$$\kappa'(\alpha; a) = (-1)^{af} \exp\left(\pi i \left(\frac{a^2 sf}{N} + r\,\frac{a - \{as\}}{N}\right)\right)(-1)^{\lfloor as/N \rfloor}.$$

Proof. Using (2) we obtain

$$\begin{aligned} g_b(\alpha\tau) &= g_{(0,\,b)}(\alpha\tau) = \psi(\alpha)\, g_{(0,\,b)\,\alpha}(\tau) = \psi(\alpha)\, g_{(bhN,\,bt)}(\tau) \\ &= \psi(\alpha)\, \varepsilon((0, bt), (bhN, 0))\, g_{(0,\,bt)}(\tau) \\ &= \psi(\alpha)\, (-1)^{bh} \exp\left(\frac{\pi i}{N}(-b^2 ht)\right) g_{(0,\,bt)}(\tau). \end{aligned}$$

If $bt \equiv \{bt\} (\mathrm{mod}\ N)$ then

$$\begin{aligned} g_{(0,\,bt)}(\tau) &= \varepsilon((0, \{bt\}), (0, N\lfloor bt/N \rfloor))\, g_{(0,\,\{bt\})}(\tau) \\ &= (-1)^{\lfloor bt/N \rfloor}\, g_{\{bt\}}(\tau). \end{aligned}$$

If $bt \equiv -\{bt\} (\mathrm{mod}\ N)$ then

$$\begin{aligned} g_{(0,\,bt)}(\tau) &= \varepsilon((0, -\{bt\}), (0, N(\lfloor bt/N \rfloor + 1)))\, g_{(0,\,-\{bt\})}(\tau) \\ &= (-1)^{\lfloor bt/N \rfloor}\, g_{\{bt\}}(\tau). \end{aligned}$$

In either case, this completes the proof of the formula for $g_b(\alpha\tau)$.

For $s_a(\alpha\tau)$, a similar but more involved calculation can be used. Alternative, one might use the Atkin–Lehner involution $w_N = \left(\begin{smallmatrix} 0 & 1 \\ -N & 0 \end{smallmatrix}\right) \in \mathrm{SL}_2\mathbf{Z}$

and the q-expansion of $g_{(e_1, e_2)}$ in [8, Chap. 2, Sect. 1, K4] via the identity $w_N(\tau) = \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)(N\tau)$ to conclude that $g_a(w\tau)/s_a(\tau) = c \exp(\pi i r a/N)$, for some constant c that does not depend on a, and thereby reduce the calculation to the one done above. ▌

Remark. Theorems 3.4 and 3.5 allow us to determine the group of divisors supported at the cusps for any particular N. For example, consider the conjecture by Klimek in [7, p. 3]. Let $J_1^\infty(N)$ denote the group of divisors on $X_1(N)$ supported at the set $\{P_1, P_2, \ldots, P_r\}$ up to linear equivalence. Klimek proved that

$$\#J_1^\infty(N) = 4^{1-r} N \prod_{\chi \neq 1} B_{2,\chi},$$

(where χ runs over all non-trivial even characters of $(\mathbf{Z}/N\mathbf{Z})^\times$, and $B_{2,\chi}$ denotes the generalized Bernoulli numbers of Kubota and Leopoldt, see also [8, Chap. 6, Theorem 3.4] for another proof), and conjectured (presumably without the benefit of a computer) that the group $J_1^\infty(N)$ is always cyclic. He confirmed this conjecture for all $N \leqslant 23$. A simple (computer-aided) calculation using Theorems 3.4 and 3.5 shows that the conjecture is false for $N = 29$. We obtain

N	$J_1^\infty(N)$
2, 3, 5, 7	0
11	$\mathbf{Z}/5\mathbf{Z}$
13	$\mathbf{Z}/19\mathbf{Z}$
17	$\mathbf{Z}/584\mathbf{Z}$
19	$\mathbf{Z}/4383\mathbf{Z}$
23	$\mathbf{Z}/37181\mathbf{Z} \times \mathbf{Z}/11\mathbf{Z}$
29	$\mathbf{Z}/64427244\mathbf{Z} \times \mathbf{Z}/4\mathbf{Z} \times \mathbf{Z}/4\mathbf{Z}$
31	$\mathbf{Z}/1772833370\mathbf{Z} \times \mathbf{Z}/10$
⋮	⋮

4. SOME UNITS ON $X_0^\#(N)$

Assume from now on that

$$N \equiv 1 \pmod 8$$

(in particular $N \geqslant 17$).

DEFINITION 4.1. Let C_N denote the group $(\mathbf{Z}/N\mathbf{Z})^\times/\pm 1$.

Since $\beta\colon X_1(N)_{\mathbf{C}} \to X_0(N)_{\mathbf{C}}$ is a cyclic Galois covering of degree r, it has a unique intermediate covering of $X_0(N)_{\mathbf{C}}$ of any degree dividing r. Because of its uniqueness, any such curve is defined over $\mathbf{Q}$ (for a thorough treatment of these intermediate curves, see [4, IV, Sect. 3]). Letting $n = (N-1)/\gcd(N-1, 12)$, we know from [10, II, Sect. 2] that the intermediate curve $X_2(N)_{\mathbf{C}} \to X_0(N)_{\mathbf{C}}$ of degree n (the *Shimura covering*) is the largest étale covering of $X_0(N)_{\mathbf{C}}$ through which β factors. (As remarked before, the cusps of $X_0(N)$ are not branch points for β; it is the points with $j = 0$ and $j = 1728$ that ramify in β.)

DEFINITION 4.2. Write n as

$$n = 2^k v,$$

where 2^k is the largest power of 2 that divides n (and $v = n/2^k$ is an odd integer). Set $z = 3$ if $N \equiv 1 \pmod 3$ and $z = 1$ otherwise. For future use, we also set $q = 3/z$ here.

Let

$$\phi\colon X_0^{\#}(N) \to X_0(N)$$

be the unique covering of degree 2^k through which β factors. Let $J_0^{\#}(N) = \mathrm{Jac}(X_0^{\#}(N))$.

Observe that the definitions of k, v, z imply that

$$r = 2^{k+1} z v.$$

Since 2^k divides n, the Shimura covering factors through ϕ. This implies that ϕ is étale and that $\Sigma_0 = \ker(\phi^*\colon J_0(N) \to J_0^{\#}(N))$ is contained in $\Sigma = \ker(J_0(N) \to J_2(N))$.

The Galois group $X_1(N)_{\mathbf{C}}$ over $X_0(N)_{\mathbf{C}}$ is isomorphic to C_N, with $\left(\begin{smallmatrix} s & f \\ Nh & t \end{smallmatrix}\right)$ mapping to $\{t\}$. Let v be a generator of C_N. We will abuse notation to lighten it, and let the same v denote the generator of $\mathrm{Gal}(X_1(N)_{\mathbf{C}}/X_0(N)_{\mathbf{C}}) \cong C_N$ and the corresponding generator of the Galois group of the function field extension; so that $(vf)(\tau) = f(v\tau)$ for all functions f on $X_1(N)$. Let Ω denote the set of 2^kth powers in C_N. Then Ω is the Galois group of $X_1(N)$ over $X_0^{\#}(N)$ and

$$\#\Omega = 2zv.$$

By its uniqueness property, $X_0^{\#}(N)$ is Galois over $X_0(N)$.

The curve $X_0^{\#}(N)$ is the coarse moduli space for the problem of classifying elliptic curves with a point of order N, where (E, P) and (E', P') are to be considered equivalent if there is an isomorphism $\delta: E \to E'$ such that $\delta(P) = \pm b \cdot P'$ for some $b \in \Omega$.

We shall now construct some units of $X_0^{\#}(N)$. They will first be given as units of $X_1(N)$, and to check that they are actually units of $X_0^{\#}(N)$ we shall need the following lemmas. In these lemmas (and later), for any element $b \in C_N$ we let $\tilde{b}$ denote the representative for b in the set $\{1, 2, \ldots, r\}$. For example, $\sum_{b \in C_N} \tilde{b} = \sum_{i=1}^{r} i = r(r+1)/2$.

LEMMA 4.7. *For any coset Ω' of Ω,*

$$\sum_{b \in \Omega'} \tilde{b}^2 \equiv 0 \pmod{N}.$$

Proof. Let μ be a primitive root modulo N. Then a set of representatives for C_N in $\mathbf{Z}$ are $1, \mu, \ldots, \mu^{(N-3)/2}$. The representatives of a coset Ω' of Ω are $\mu^j, \mu^{j+2^k}, \mu^{j+2\cdot 2^k}, \ldots, \mu^{j+((N-1)/2^{k+1}-1)2^k}$ for some $0 \leqslant j < 2^k$. If

$$\tilde{b} \equiv \pm \mu^{j+t2^k} \pmod{N},$$

then

$$\tilde{b}^2 \equiv \mu^{2j+t2^{k+1}} \pmod{N},$$

so

$$\sum_{b \in \Omega'} \tilde{b}^2 \equiv \mu^{2j} \sum_{t=0}^{(N-1)/2^{k+1}-1} \mu^{2^{k+1}t} = \mu^{2j} \frac{\mu^{N-1}-1}{\mu^{2^{k+1}}-1} \equiv 0 \pmod{N}$$

by Fermat's Little Theorem. ∎

In the proofs of the next two lemmas, we shall use the following convention.

CONVENTION 4.8. For P a statement, let $[P]$ be 1 if P is true, 0 if P is false.

LEMMA 4.9. *Let t be an integer relatively prime to N. Then*

$$S = \sum_{b \in C_N} \lfloor \tilde{b}t/N \rfloor$$

is even if and only if t is a square modulo N.

Proof. Note that $\sum_{b\in C_N}\tilde{b} = r(r+1)/2$ is even (since r is divisible by 4), so

$$S = \sum_{b\in C_N}\lfloor \tilde{b}t/N\rfloor \equiv \sum_{b\in C_N}^{r}\tilde{b}t - \sum_{b\in C_N} N\lfloor \tilde{b}t/N\rfloor = \sum_{b\in C_N}(\tilde{b}t \bmod N) \pmod 2.$$

Now $\tilde{b}t \bmod N$ is either $\{\tilde{b}t\}$ or $N-\{\tilde{b}t\}$ depending on whether $\tilde{b}t \bmod N \leqslant r$ or not, respectively. Therefore, we can write S as

$$\begin{aligned} S &\equiv \sum_{b\in C_N}\{\tilde{b}t\} + \sum_{b\in C_N}[\tilde{b}t \bmod N > r]\,(-\{\tilde{b}t\} + N - \{\tilde{b}t\}) \\ &\equiv \sum_{b\in C_N}\tilde{b} + \sum_{b\in C_N}[\tilde{b}t \bmod N > r]\,(-\{\tilde{b}t\} + N - \{\tilde{b}t\}) \\ &\equiv \sum_{b\in C_N}[\tilde{b}t \bmod N > r]\,(N - 2\{\tilde{b}t\} \equiv \sum_{b\in C_N}[\tilde{b}t \bmod N > r] \pmod 2. \end{aligned}$$

Define m to equal $\sum_{b\in C_N}[\tilde{b}t \bmod N > r]$. Then

$$(-1)^m r! = \prod_{b\in C_N}\{\tilde{b}t\} \equiv \prod_{b\in C_N}(\tilde{b}t) = t^r r! \pmod N.$$

Since N does not divide $r!$, this implies that

$$(-1)^m \equiv t^r \pmod N.$$

Since $t^r \equiv 1 \pmod N$ exactly when t is a square modulo N, this proves our lemma. ∎

LEMMA 4.10. *Let Ω' be a coset of Ω and s, f, t, h integers with $st - Nfh = 1$ and $\{t\} \in \Omega$ and*

$$S = h(1-t)\sum_{b\in\Omega'}\tilde{b} + \sum_{b\in\Omega'}\lfloor \tilde{b}t/N\rfloor.$$

The parity of S does not depend on the choice of Ω'.

Proof. First observe that $st - Nfh = 1$ implies that if t is even then h must be odd, so in any case $h(1-t) \equiv t+1 \pmod 2$. Therefore,

$$S \equiv (t+1)\sum_{b\in\Omega'}\tilde{b} - N\sum_{b\in\Omega'}\lfloor \tilde{b}t/N\rfloor = \sum_{b\in\Omega'}\tilde{b} + \sum_{b\in\Omega'}(\tilde{b}t \bmod N) \pmod 2.$$

Since $t \in \Omega$, for b ranging over Ω the reductions of $\{\tilde{b}t\}$ to C_N just range over Ω', so we can once again use the method of the proof of Lemma 4.9. Accordingly,

$$S \equiv 2 \sum_{b \in \Omega'} \tilde{b} + \sum_{b \in \Omega'} [\tilde{b}t \bmod N > r]\,(-\{\tilde{b}t\} + N - \{\tilde{b}t\})$$

$$\equiv \sum_{b \in \Omega'} [\tilde{b}t \bmod N > r] \qquad (\text{mod } 2).$$

Let $m = \sum_{b \in \Omega'} [\tilde{b}t \bmod N > r]$. Then

$$(-1)^m \prod_{b \in \Omega'} \tilde{b} = \prod_{b \in \Omega'} \{\tilde{b}t\} \equiv \prod_{b \in \Omega'} (\tilde{b}t) = t^{\#\Omega} \prod_{b \in \Omega'} \tilde{b} \qquad (\text{mod } N)$$

Since N does not divide $\prod_{b \in \Omega'} \tilde{b}$,

$$(-1)^m \equiv t^{\#\Omega} \qquad (\text{mod } N),$$

and it is plain that the parity of m (which is the same as the parity of S) depends only on the choice of t and not on the choice of Ω'. ∎

Recall from Definition 4.2 that $q = 3/z$.

THEOREM 4.11. *Define the following three functions on* $X_1(N)$*:*

$$f(\tau) = \left(\prod_{b \in \Omega} g_{\tilde{b}}(\tau)^{2^k q} \right) \left(\prod_{b \in C_N} g_{\tilde{b}}(\tau)^{-q} \right),$$

$$g(\tau) = \left(\prod_{b \in \Omega} g_{\tilde{b}}(\tau)^{-q} \right) \left(\prod_{b \in \nu\Omega} g_{\tilde{b}}(\tau)^{q} \right),$$

$$h(\tau) = \left(\prod_{b \in \Omega} g_{\tilde{b}}(\tau)^{2q} \right).$$

Then the following are true:

(a) *the group* $\mathrm{Gal}(X_1(N)_{\mathbf{C}}/X_0^{\#}(N)_{\mathbf{C}})$ *is the subgroup of* $\mathrm{Gal}(X_1(N)_{\mathbf{C}}/X_0(N)_{\mathbf{C}})$ *that fixes the function* f;

(b) *the functions* g *and* h *are invariant under* $\mathrm{Gal}(X_1(N)_{\mathbf{C}}/X_0^{\#}(N)_{\mathbf{C}})$ *and can therefore be regarded as being defined on* $X_0^{\#}(N)_{\mathbf{C}}$*;*

(c) $vf = (-1)^{2^k} f$;

(d) $g(vg)(v^2 g)\cdots(v^{2^k-1} g) = -1$;

(e) *the divisor* $\mathrm{div}(f)$ *is divisible by* 2^k *in* $\mathrm{Div}^0(X_0^{\#}(N))$.

Remark. Henceforth we will regard f, g and h as functions on $X_0^{\#}(N)$. By (a) above, f descends to no smaller cover of $X_0(N)$.

Note. The function f above is the analogue in our situation of the Ogg–Ligozat function f_{OL} that was used in [10, II, Proposition (12.2)]. (In that paper, f_{OL} is called f.) Note however that if we restrict to the case of $N \equiv 9 \pmod{16}$ (equivalently $n \equiv 2 \pmod 4$) considered in that paper, our function f does not equal the function f_{OL}. Instead, the "correct" function f (the one we are using above) is equal to f_{OL}^q. However, since q is always equal to 1 or 3, and Mazur was constructing a point in a group of exponent two, f and f_{OL} worked equally well.

Proof. First we need to check (using Theorem 3.4) that f is actually a function on $X_1(N)$. Conditions (V1) and (V3) are clearly satisfied, since, in the notation of Theorem 3.4, each d_i is zero, and $\sum c_i$ is also zero. For (V2), note that we need that N divide

$$q\left(\sum_{b\in\Omega}\tilde{b}^2 2^k - \sum_{b\in C_N}\tilde{b}^2\right) = q\left(2^k\sum_{b\in\Omega}\tilde{b}^2 - \frac{r(r+1)}{6}N\right).$$

The first term is divisible by N by Lemma 4.7, the second is divisible by N since clearly $r(r+1)/6$ is an integer.

We have now confirmed that f is defined on $X_1(N)$. It remains to check that the largest subgroup Θ of C_N that fixes f is in fact Ω. Since the coefficients in f for those $g_{\tilde{b}}$ with $b \in \Omega$ are different from those for which $b \notin \Omega$, we must have $\Theta \subseteq \Omega$.

To check $\Theta = \Omega$ then, it remains to show that for any $\alpha = \left(\begin{smallmatrix} s & f \\ Nh & t \end{smallmatrix}\right) \in \Gamma_0(N)$ with $\{t\} \in \Omega$, we have $f(\alpha\tau) = f(\tau)$. Using Theorem 3.6 (and (V1) to get rid of the ψ factors), we need to confirm that

$$C = q\left(\sum_{b\in\Omega} 2^k(N\tilde{b}h - \tilde{b}^2 ht + N\lfloor \tilde{b}t/N\rfloor)\right)$$
$$- q\left(\sum_{b\in C_N}(N\tilde{b}h - \tilde{b}^2 ht + N\lfloor \tilde{b}t/N\rfloor)\right) \equiv 0 \pmod{2N}.$$

The expression C can be thought of as six separate sums, and it turns out that each of them is divisible by $2N$. This is obvious for the first, third, fourth and fifth; follows by Lemma 4.7 for the second; and follows by Lemma 4.9 for the sixth. Hence we have proved (a).

Similarly, we can check that g is defined on $X_0^{\#}(N)$. Again using Theorem 3.6, we need to confirm that

$$-q\left(\sum_{b\in\Omega}(N\tilde{b}h-\tilde{b}^2ht+N\lfloor\tilde{b}t/N\rfloor)\right)$$
$$+q\left(\sum_{b\in\Omega}(N\tilde{b}h-\tilde{b}^2ht+N\lfloor\tilde{b}t/N\rfloor)\right)\equiv 0 \qquad (\mathrm{mod}\ 2N).$$

The divisibility by N is immediate by Lemma 4.7, and for divisibility by 2 observe that for any coset Ω' of Ω we have

$$S=\sum_{b\in\Omega'}(N\tilde{b}h-\tilde{b}^2ht+N\lfloor\tilde{b}t/N\rfloor)$$
$$\equiv\sum_{b\in\Omega'}(\tilde{b}h-\tilde{b}ht+\lfloor\tilde{b}t/N\rfloor)\equiv h(1-t)\sum_{b\in\Omega'}\tilde{b}+\sum_{b\in\Omega'}\lfloor\tilde{b}t/N\rfloor \qquad (\mathrm{mod}\ 2).$$

Now an application of Lemma 4.10 completes the proof.

For $h(\tau)$, we need $12\,|\,2q\,(\#\Omega)$ to verify (V1). But $2q(\#\Omega)=2q\cdot 2zv=12v$ so this is clear. The rest of the proof is analogous to the proof for $g(\tau)$. This completes the proof of (b).

To prove (c), we calculate

$$(vf)(\tau)=f(v\tau)=\left(\prod_{b\in\Omega}g_{\tilde{b}}(v\tau)^{2^kq}\right)\left(\prod_{b\in C_N}g_{\tilde{b}}(v\tau)^{-q}\right)$$
$$=\left(\prod_{b\in\Omega}g_{\widetilde{vb}}(\tau)^{2^kq}\kappa(v;\tilde{b})^{2^kq}\right)\left(\prod_{b\in C_N}g_{\widetilde{vb}}(\tau)^{-q}\kappa(v;\tilde{b})^{-q}\right),$$

so (since $(-1)^q=-1$) it suffices to show that

$$\prod_{b\in\Omega}\kappa(v;\tilde{b})^{2^k}\prod_{b\in C_N}\kappa(v;\tilde{b})^{-1}=-1.$$

Pick some $\left(\begin{smallmatrix}s & f\\ Nh & t\end{smallmatrix}\right)\in\Gamma_0(N)$ that lifts v. Then t will necessarily generate $(\mathbf{Z}/N\mathbf{Z})^\times$, and in particular it will be a non-square modulo N. By Theorem 3.6, it suffices to show that

$$\exp\left(\frac{\pi i}{N}\left(2^k\sum_{b\in\Omega}(N\tilde{b}h-\tilde{b}^2ht+N\lfloor\tilde{b}t/N\rfloor)\right)\right)$$
$$\times\exp\left(\frac{\pi i}{N}\left(-\sum_{b\in C_N}(N\tilde{b}h-\tilde{b}^2ht+N\lfloor\tilde{b}t/N\rfloor)\right)\right)=-1.$$

The first exponential is clearly 1 by Lemma 4.7 and because 2^k is even. Clearly

$$\exp\left(\frac{\pi i}{N}\left(-\sum_{b\in C_N}(N\tilde{b}h-\tilde{b}^2 ht\right)\right)=1,$$

and we are done with (c) by Lemma 4.9.

For (d), note that

$$\begin{aligned}
g &= \left(\prod_{b\in\Omega} g_{\tilde{b}}^{-q}\prod_{b\in v\Omega} g_{\tilde{b}}^{q}\right)\\
vg &= \left(\prod_{b\in v\Omega} g_{\tilde{b}}^{-q}\prod_{b\in v^2\Omega} g_{\tilde{b}}^{q}\right)\left(\prod_{b\in\Omega}\kappa(v;\tilde{b})^{-q}\right)\left(\prod_{b\in v\Omega}\kappa(v;\tilde{b})^{q}\right)\\
v^2 g &= \left(\prod_{b\in v^2\Omega} g_{\tilde{b}}^{-q}\prod_{b\in v^3\Omega} g_{\tilde{b}}^{q}\right)\left(\prod_{b\in\Omega}\kappa(v;\tilde{b})^{-q}\right)\left(\prod_{b\in v\Omega}\kappa(v;\tilde{b})^{q}\right)\times\\
&\quad\times\left(\prod_{b\in v\Omega}\kappa(v;\tilde{b})^{-q}\right)\left(\prod_{b\in v^2\Omega}\kappa(v;\tilde{b})^{q}\right)\\
&= \left(\prod_{b\in v^2\Omega} g_{\tilde{b}}^{-q}\prod_{b\in v^3\Omega} g_{\tilde{b}}^{q}\right)\left(\prod_{b\in\Omega}\kappa(v;\tilde{b})^{-q}\right)\left(\prod_{b\in v^2\Omega}\kappa(v;\tilde{b})^{q}\right)\\
&\vdots\\
v^{2^k-1}g &= \left(\prod_{b\in v^{2^k-1}\Omega} g_{\tilde{b}}^{-q}\prod_{b\in v\Omega} g_{\tilde{b}}^{q}\right)\left(\prod_{b\in\Omega}\kappa(v;\tilde{b})^{-q}\right)\left(\prod_{b\in v^{2^k-1}\Omega}\kappa(v;\tilde{b})^{q}\right).
\end{aligned}$$

Therefore

$$g(vg)\,(v^2g)\cdots(v^{2^k-1}g)=\left(\prod_{b\in\Omega}\kappa(v;\tilde{b})^{-q}\right)^{2^k}\left(\prod_{b\in C_N}\kappa(v;\tilde{b})^{q}\right)=\frac{fg^{2^k}}{vf}=-1,$$

by (c). Thus the proof of (d) is complete.

We will use Theorem 3.5 to calculate $\operatorname{div}(f)$. It is immediately clear that $\operatorname{ord}_{Q_t}(f)=0$ for all the cusps Q_t. On the other hand, letting Ω' denote the coset of Ω containing the reduction of $\tilde{b}t$, and using the fact that $B_2(x)=B_2(1-x)$, we can calculate

$$\begin{aligned}
\operatorname{ord}_{P_t}(f) &= q\left(\sum_{b\in\Omega} 2^k\frac{N}{2}B_2\left(\frac{\tilde{b}t \bmod N}{N}\right)-\sum_{b\in C_N}\frac{N}{2}B_2\left(\frac{\tilde{b}t \bmod N}{N}\right)\right)\\
&= q\left(2^{k-1}N\sum_{b\in\Omega'}\left(\frac{\tilde{b}^2}{N^2}-\frac{\tilde{b}}{N}+\frac{1}{6}\right)-\frac{N}{2}\,\frac{(-r)}{6N}\right)\\
&= q2^{k-1}\sum_{b\in\Omega'}\left(\frac{\tilde{b}^2}{N}-\tilde{b}\right)+q\left(2^{k-1}N\frac{\#\Omega}{6}+\frac{r}{12}\right).
\end{aligned}$$

By Lemma 4.7, the first term is an integer and clearly it is divisible by 2^k. Using the identities $\#\Omega = 2zv$, $r = 2^{k+1}zv$, $qz = 3$, the latter term in the above sum reduces to

$$2^{k-1}(N+1)\,v,$$

which is divisible by 2^k since $N+1$ is even. So we have completed the proof of (e). ▮

We will need the following lemma. For a curve X defined over $\mathbf{Q}$ and a field K containing $\mathbf{Q}$, denote the function field of X over K by $K(X)$. It is well known that a finite abelian group equipped with a continuous action of $\mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q})$ is the same as a finite étale commutative group scheme over $\mathbf{Q}$. We will use this identification throughout the rest of this paper.

LEMMA 4.12. *Let $\phi: X \to Y$ be a finite étale reap of projective curves, with X, Y and ϕ defined over $\mathbf{Q}$. Assume that ϕ is Galois after some finite base extension $F/\mathbf{Q}$. Let $\Gamma = \mathrm{Gal}(\bar{\mathbf{Q}}(X)/\bar{\mathbf{Q}}(Y)) \cong \mathrm{Gal}(F(X)/F(Y))$ and assume that Γ is commutative. By the Picard functoriality of the Jacobians, we have an exact sequence*

$$0 \to K \to \mathrm{Jac}(Y)\,(\bar{\mathbf{Q}}) \xrightarrow{\phi} \mathrm{Jac}(X)\,(\bar{\mathbf{Q}})^{\Gamma},$$

where K denotes the finite $\mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q})$-module $\ker(\phi^*)$. Then

(a) *the group scheme K is isomorphic to Γ^D, the Cartier dual of Γ;*

(b) *if Γ is cyclic, then ϕ^* surjects onto $\mathrm{Jac}(X)\,(\bar{\mathbf{Q}})^{\Gamma}$.*

Remark. To make sense of Γ^D, we need to consider Γ as an étale group scheme over $\mathbf{Q}$. The group Γ is naturally acted upon by $\mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q})$ as follows: for any $\sigma \in \mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q})$ and $\tau \in \Gamma$, let $\sigma \cdot \tau = \tilde{\sigma}\tau\tilde{\sigma}^{-1}$ where $\sigma \in \mathrm{Gal}(\bar{\mathbf{Q}}(X)/\mathbf{Q}(X)) \cong \mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q})$ is any lift of σ.

Proof. Taking the exact sequence of low degree terms for the Hochschild–Serre spectral sequence of the étale cohomology of $\mathbf{G}_m$ over the base $\bar{\mathbf{Q}}$ as in [12, III, Theorem 2.20], we obtain

$$0 \to H^1(\Gamma, H^0(X_{et}, \mathbf{G}_m)) \to H^1(Y_{et}, \mathbf{G}_m) \to$$
$$H^0(\Gamma, H^1(X_{et}, \mathbf{G}_m)) \to H^2(\Gamma, H^0(X_{et}, \mathbf{G}_m))$$

which is

$$0 \to H^1(\Gamma, \bar{\mathbf{Q}}^{\times}) \to \mathrm{Pic}(Y) \to \mathrm{Pic}(X)^{\Gamma} \to H^2(\Gamma, \bar{\mathbf{Q}}^{\times})$$

(where the third map is ϕ^*). Now since Γ acts trivially on $\bar{\mathbf{Q}}^\times$, $H^1(\Gamma, \bar{\mathbf{Q}}^\times) \cong \mathrm{Hom}(\Gamma, \bar{\mathbf{Q}}^\times) \cong \Gamma^D$. The kernel of $\phi^*: \mathrm{Pic}(Y) \to \mathrm{Pic}(X)$ is contained in $\mathrm{Jac}(Y)$, so we have proved (a).

If Γ is cyclic, then by [16, VIII, Sect. 4], $H^2(\Gamma, \bar{\mathbf{Q}}^\times) \cong (\bar{\mathbf{Q}}^\times)^\Gamma / (\bar{\mathbf{Q}}^\times)^N = 0$. But if $\phi^*: \mathrm{Pic}(Y) \to \mathrm{Pic}(X)^\Gamma$ is surjective, then so is $\phi^*: \mathrm{Jac}(Y) \to \mathrm{Jac}(X)^\Gamma$, which proves (b). ∎

Now we are almost ready to find extra points in $J_0(N)[I]$. Recall that c denotes the divisor $0-\infty$ on $X_0(N)$.

THEOREM 4.13. *Let $d = (1/2^k)\,\mathrm{div}(f)$, considered as a point on $J_0^\#(N)$. Then*

(a) *the divisor d is rational over $\mathbf{Q}$;*

(b) *the divisor d is in the image of $\phi^*: J_0(N) \to J_0^\#(N)$;*

(c) $2d = \phi^*(v \cdot c)$.

Remark. In essence, we are trying to find "one half of c" in the group $J_0(N)[I]/\Sigma$. Assertion (c) in the above theorem shows that d is "one half of $v \cdot c$". Recall from Definition 4.2 that v is the odd part of n, so this is as good as finding half of c, but some calculations work out simpler this way. Assertion (b) will be used to show that our point pulls back to $J_0(N)$, and assertion (a) will be used to show that we are actually finding points in $J_0(N)[I]$.

Proof. As can be seen from the proof of Theorem 4.11(e), $\mathrm{div}(f)$ is concentrated at the cusps of $X_0^\#(N)$ that lie over the cusp 0 of $X_0(N)$. All of these cusps are rational over $\mathbf{Q}$, hence so is d, proving (a).

By Lemma 4.12(b), it suffices to check that d is fixed by v, the generator of the group $\mathrm{Gal}(X_0^\#(N)/X_0(N))$. By Theorem 4.11(c), $\mathrm{div}(f) - \mathrm{div}(vf) = -2k\,\mathrm{div}(g)$, so

$$d - vd = \frac{1}{2^k}\mathrm{div}(f) - \frac{1}{2^k}\mathrm{div}(vf) = \mathrm{div}(1/g),$$

which is a principal divisor, so $d = vd$ in $J_0^\#(N)$, concluding our proof of (b).

Let $d' = \mathrm{div}(f)/2^{k-1} - \mathrm{div}(h)$ be a divisor on $X_0^\#(N)$. Using Theorem 3.5, for any $1 \leqslant t \leqslant r$,

$$\mathrm{ord}_{Q_t}(d') = \frac{1}{2^{k-1}}\left(\sum_{b \in \Omega} \frac{2^k q}{12} - \sum_{b \in C_N} \frac{q}{12}\right) - \sum_{b \in \Omega} \frac{2q}{12} = 0 - 2zvq/6 = -v,$$

and

$$\operatorname{ord}_{P_t}(d') = \sum_{b \in \Omega} \frac{2^k qN}{2^k} B_2\left(\frac{\tilde{b}t \bmod N}{N}\right) - \sum_{b \in C_N} \frac{qN}{2^k} B_2\left(\frac{\tilde{b}t \bmod N}{N}\right)$$
$$- \sum_{b \in \Omega} \frac{2qN}{2} B_2\left(\frac{\tilde{b}t \bmod N}{N}\right) = -\frac{qN}{2^k} \frac{(-r)}{6N} = v.$$

Hence $d' = \phi^*(v \cdot c)$, so

$$2d - \operatorname{div}(h) = \phi^*(v \cdot c)$$

as divisors. But h is a function defined on $X_0^{\#}(N)$ by Theorem 4.11(b), so this proves (c). ∎

5. THE GALOIS STRUCTURE OF $J_0(N)[I]$

THEOREM 5.14. *Let $\mathscr{D}$ denote the group generated by d in $J_0^{\#}(N)$. Let $A = (\phi^*)^{-1}\mathscr{D}$. Then*

(a) *all the points of $\mathscr{D}$ are unramified at N;*

(b) *all the points of A are unramified at N;*

(c) *the group A is contained in $J_0(N)[I]$.*

Remark. Since $\#\ker(\phi) = \#\mathscr{D} = 2^k$, the group A has cardinality 2^{2k}. Therefore part (c) of the above theorem implies that A is the whole of the 2-primary component of $J_0(N)[I]$. Since the odd part of $J_0(N)[I]$ is the direct sum of the odd parts of C and Σ, we have now completed the concrete description of $J_0(N)[I]$ that we were aiming for.

Proof. Assertion (a) is immediate from Theorem 4.13(a), since the points of $\mathscr{D}$ are rational. (Note that since the action of $\mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q})$ on the cusps of $X_0^{\#}(N)$ factors through the cyclotomic character χ_N, the only way for a divisor supported at the cusps to be unramified at N is to be rational.)

Assertion (b) follows from [10, II, Lemma (16.5)]. Note that since the lemma just cited applies only to points of prime power order, we have to apply it separately to each of the primary components of the point of A in question.

Multiplication by 2^k annihilates d. Therefore $2^k A \subseteq \ker(\phi^*) \subseteq \Sigma$, so certainly all points in A are torsion points. By [15, Proposition 3.3], all torsion points of $J_0(N)$ that are unramified at N are in $J_0(N)[I]$, so we have proved $A \subseteq J_0(N)[I]$. ∎

Remark. For the reader's convenience we summarize another proof of part (c) of Theorem 5.14 that avoids invoking [15]. This proof also does not need the results of parts (a) and (b) of Theorem 5.14. We shall use the terminology and notation of [10]. Fix an embedding $\bar{\mathbf{Q}} \hookrightarrow \bar{\mathbf{Q}}_N$ and let J be the Néron model of $J_0(N)$ over $\mathbf{Z}_N$. Let $J_{/\mathbf{F}_N}$ denote the special fiber of J, and let $J^0_{/\mathbf{F}_N}$ denote the irreducible component of the identity in $J_{/\mathbf{F}_N}$. Let $\Sigma_{/\mathbf{F}_N}$ denote the reduction of Σ to $J_{/\mathbf{F}_N}$. Note that $\Sigma \cong \mu_n$, and so Σ is unramified at N. Therefore, by [17, Lemma 2], Σ reduces injectively to $\Sigma_{/\mathbf{F}_N}$. Then, by [10, II, Proposition (11.9)],

$$\Sigma_{/\mathbf{F}_N} \cap J^0_{/\mathbf{F}_N} = 0.$$

Thus, a point of Σ that reduces to a point in $J^0_{/\mathbf{F}_N}$ must be zero. We shall now use this observation to show that $A \subseteq J_0(N)[I]$.

It suffices to show that for an arbitrary point x of A and any element T of I, we have $Tx = 0$. The group of irreducible components of $J_{/\mathbf{F}_N}$ is Eisenstein, as can be seen from the title (and contents) of [6] (see also [14]). Therefore, the operator T sends the reduction of x into the identity component. In other words, Tx reduces into $J^0_{/\mathbf{F}_N}$.

On the other hand, we can use the formulae in [18, Sect. 2] to define actions of T_l (for $l \neq N$) and w on ($J_1(N)$ and therefore on) $J_0^{\#}(N)$ that are compatible with the actions defined on $J_0(N)$ via the map ϕ^*, and calculate (in the spirit of the proof of Theorem 4.13(c)) that $\mathscr{D}$ is annihilated by each $1+l-T_l$ and by $1+w$. Let T' be a lift of T to the ring $\mathbf{Z}[\ldots, T_l, \ldots, w]$ and let T'' be the image of T' in $\mathrm{End}(J_0^{\#}(N))$. Then we have a commutative diagram

$$\begin{array}{ccc} J_0(N) & \xrightarrow{\phi^*} & J_0^{\#}(N) \\ \downarrow{\scriptstyle T} & & \downarrow{\scriptstyle T''} \\ J_0(N) & \xrightarrow{\phi^*} & J_0^{\#}(N). \end{array}$$

Here x is mapped to $\phi^* x \in \mathscr{D}$ which is annihilated by T''. By the commutativity of the diagram we must have $Tx \in \ker(\phi^*) = \Sigma$. This completes our proof that $Tx = 0$.

Now that we established that $A \subseteq J_0(N)[I]$, we will determine the action of $\mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q})$ on A and then assemble what we know to find the action of $\mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q})$ on the whole of $J_0(N)[I]$.

DEFINITION 5.1. Let $\lambda: X_0^{\#\#}(N) \to X_0^{\#}(N)$ be a minimal covering of $X_0^{\#}(N)$ on which $f^{1/2^k}$ is defined.

By Theorem 4.11, parts (a) and (e), the degree of λ is 2^k and λ is étale. In fact, after base extension to $\mathbf{Q}(\mu_{2^{k+1}})$, $\lambda \circ \phi$ becomes a Galois covering with Galois group Γ. The group Γ can be regarded as a finite étale group scheme over $\mathbf{Q}$, and by Lemma 4.12(a), A will be its Cartier dual. This allows us to determine the action of $\mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q})$ on A.

Convention 5.15. Choose once and for all a primitive $2^{k+1}st$ root of unity $\zeta \in \bar{\mathbf{Q}}$. Then ζ^2 is the primitive 2^kth root of unity that we will use in explicit Cartier duality calculations.

THEOREM 5.16. *Let K denote the function field of $X_0(N)$ over $\mathbf{Q}$ and L the function field of $X_0^{\#}(N)$ over $\mathbf{Q}$, so that $L(f^{1/2^k})$ is the function field of $X_0^{\#\#}(N)$ over $\mathbf{Q}$.*

(a) *$L(f^{1/2^k}, \zeta)/K(\zeta)$ is a Galois extension with*

$$\Gamma = \mathrm{Gal}(L(f^{1/2^k}, \zeta)/K(\zeta)) \cong \mathbf{Z}/2^k\mathbf{Z} \times \mathbf{Z}/2^k\mathbf{Z}.$$

In terms of the basis described in the proof, any element σ of $\mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q})$ acts on Γ via the matrix

$$\begin{pmatrix} 1 & 0 \\ (\chi_{2^{k+1}}(\sigma)-1)/2 & \chi_{2^k}(\sigma) \end{pmatrix}.$$

(b) *The abelian group A is isomorphic to $\mathbf{Z}/2^k\mathbf{Z} \times \mathbf{Z}/2^k\mathbf{Z}$, with $\sigma \in \mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q})$ acting via*

$$\begin{pmatrix} \chi_{2^k}(\sigma) & (1-\chi_{2^{k+1}}(\sigma))/2 \\ 0 & 1 \end{pmatrix}.$$

Proof. We know that the field extension L/K is Galois of degree 2^k with cyclic Galois group generated by v, and this remains true for $L(\zeta)/K(\zeta)$. Clearly $L(f^{1/2^k}, \zeta)/L(\zeta)$ is also Galois (and cyclic) of degree 2^k. Since $K(\zeta)$ (and hence $L(\zeta)$) contains all 2^kth roots of unity, and by Theorem 4.11(c), $(vf)/f = (\zeta g)^{2^k}$, we can conclude that $L(f^{1/2^k}, \zeta) = L((vf)^{1/2^k}, \zeta)$. This way we obtain that $L(f^{1/2^k}, \zeta)$ contains all the 2^kth roots of $f, vf, \ldots$, and therefore that $L(f^{1/2^k}, \zeta)/K(\zeta)$ is a Galois extension.

To determine the group $\Gamma = \mathrm{Gal}(L(f^{1/2^k}, \zeta)/K(\zeta))$, observe that the field extension $L(f^{1/2^k}, \zeta)/K(\zeta)$ contains all the conjugates of its generator $f^{1/2^k}$. Therefore it is obtained as a splitting field of the polynomial F whose

roots are all the 2^kth roots of all conjugates of f. Since $vf = (-1)\, g^{2^k} f$, the 2^kth roots of vf are $\zeta g f^{1/2^k}, \zeta^3 g f^{1/2^k}, \ldots, \zeta^{2^{k+1}-1} g f^{1/2^k}$. Then

$$\begin{aligned} v^2 f &= v((-1)\, g^{2^k} f) = (-1)\, vg)^{2^k}\, (vf) \\ &= (-1)\, (vg)^{2^k}\, (-1)\, g^{2^k} f = (vg)^{2^k} g^{2^k} f, \end{aligned}$$

so the 2^kth roots of $v^2 f$ are $(vg)\, g f^{1/2^k}, \zeta^2 (vg)\, g f^{1/2^k}, \ldots, \zeta^{2^{k+1}-2} (vg)\, g f^{1/2^k}$. Hence it is clear that the roots of F are exactly the

$$\delta_{i,j} = \zeta^{2i+j} \left(\prod_{k=0}^{j-1} (v^k g) \right) f^{1/2^k},$$

where i and j range over the interval $[0, 2^k - 1]$.

To determine Γ, observe that it must act simply transitively on the set of all roots of F. Let $\varrho \in \Gamma$ be such that

$$\varrho\colon \delta_{0,0} = f^{1/2^k} \mapsto \delta_{1,0} = \zeta^2 f^{1/2^k}.$$

Taking 2^kth powers, we see that ϱ fixes f and hence all of $L(f^{1/2^k}, \zeta)$. So ϱ sends $\delta_{i,j}$ to $\delta_{i+1,j}$ (with $\delta_{2^k,j}$ to be interpreted as $\delta_{0,j}$).

Now consider the element $\bar{v} \in \Gamma$ for which

$$\bar{v}\colon \delta_{0,0} = f^{1/2^k} \mapsto \delta_{0,1} = \zeta g f^{1/2^k}.$$

Taking 2^kth powers again, we see that $\bar{v}$ sends f to vf, so it acts as v on $L(\zeta)$ (thereby justifying our choice of name for it). Note that

$$\bar{v}(\delta_{0,1}) = \bar{v}(\zeta g f^{1/2^k}) = \zeta(\bar{v}g)\, (\bar{v} f^{1/2^k}) = \zeta(vg)\, \zeta g f^{1/2^k} = \zeta^2 (vg)\, g f^{1/2^k} = \delta_{0,2},$$

similarly

$$\bar{v}(\delta_{0,2}) = \bar{v}(\zeta^2 (vg)\, g f^{1/2^k}) = \zeta^3 (v^2 g)\, (vg)\, g f^{1/2^k} = \delta_{0,3},$$

and so on. Finally, using Theorem 4.11(d) we obtain

$$\begin{aligned} \bar{v}(\delta_{0,2^k-1}) &= \bar{v}(\zeta^{2^k-1} (v^{2^k-2} g) \cdots (vg)\, g f^{1/2^k}) = \zeta^{2^k} (v^{2^k-1} g) \cdots (vg)\, g f^{1/2^k} \\ &= (-1)\, (-1)\, f^{1/2^k} = f^{1/2^k} = \delta_{0,0}. \end{aligned}$$

Hence $\bar{v}$ sends $\delta_{i,j}$ to $\delta_{i,j+1}$ (with $\delta_{i,2^k}$ to be interpreted as $\delta_{i,0}$).

This shows that Γ is generated by two commuting elements of order 2^k. In other words, we have $\Gamma \cong \mathbf{Z}/2^k\mathbf{Z} \times \mathbf{Z}/2^k\mathbf{Z}$, and we can represent elements of Γ as column vectors over $\mathbf{Z}/2^k\mathbf{Z}$, with $\bar{v}$ corresponding to $\binom{1}{0}$ and ϱ corresponding to $\binom{0}{1}$.

As for the action of $\mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q})$ on Γ, take some $\sigma \in \mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q})$ and consider its natural action on $L(f^{1/2^k}, \zeta)$ that leaves $L(f^{1/2^k})$ fixed. Both $\sigma\bar{v}\sigma^{-1}$ and $\varrho^{(\chi_{2^{k+1}}(\sigma)-1)/2}\,\bar{v}$ fix ζ and

$$\sigma\bar{v}\sigma^{-1}\colon f^{1/2^k} \mapsto f^{1/2^k} \mapsto \zeta g f^{1/2^k} \mapsto \zeta^{\chi_{2^{k+1}}(\sigma)} g f^{1/2^k}$$
$$\varrho^{(\chi_{2^{k+1}}(\sigma)-1)/2}\,\bar{v}\colon f^{1/2^k} \mapsto \zeta g f^{1/2^k} \mapsto \zeta^{\chi_{2^{k+1}}(\sigma)} g f^{1/2^k}.$$

Therefore $\sigma\bar{v}\sigma^{-1} = \varrho^{(\chi_{2^{k+1}}(\sigma)-1)/2}\,\bar{v}$. Similarly both $\sigma\varrho\sigma^{-1}$ and $\varrho^{\chi_{2^k}(\sigma)}$ fix ζ and

$$\sigma\varrho\sigma^{-1}\colon f^{1/2^k} \mapsto f^{1/2^k} \mapsto \zeta^2 f^{1/2^k} \mapsto \zeta^{2\chi_{2^{k+1}}(\sigma)} f^{1/2^k}$$
$$\varrho^{\chi_{2^k}(\sigma)}\colon f^{1/2^k} \mapsto \zeta^{2\chi_{2^k}(\sigma)} f^{1/2^k} = \zeta^{2\chi_{2^{k+1}}(\sigma)} f^{1/2^k}.$$

Therefore $\sigma\varrho\sigma^{-1} = \varrho^{\chi_{2^k}(\sigma)}$. Hence $\sigma \in \mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q})$ does act on the elements of Γ (represented by column vectors) as required. With this the proof of (a) is complete.

For (b), a simple calculation shows that if G is an étale group scheme over $\mathbf{Q}$ that is isomorphic to $\mathbf{Z}/m\mathbf{Z} \times \mathbf{Z}/m\mathbf{Z}$ with a Galois action described by

$$\begin{pmatrix} a(\sigma) & b(\sigma) \\ c(\sigma) & d(\sigma) \end{pmatrix}\colon \mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q}) \to \mathrm{GL}_2(\mathbf{Z}/m\mathbf{Z}),$$

then its Cartier dual G^D is also isomorphic to $\mathbf{Z}/m\mathbf{Z} \times \mathbf{Z}/m\mathbf{Z}$, but with a Galois action described in terms of the usual dual basis by

$$\begin{pmatrix} \chi_m(\sigma)\, a(\sigma^{-1}) & \chi_m(\sigma)\, c(\sigma^{-1}) \\ \chi_m(\sigma)\, b(\sigma^{-1}) & \chi_m(\sigma)\, d(\sigma^{-1}) \end{pmatrix}\colon \mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q}) \to \mathrm{Gl}_2(\mathbf{Z}/m\mathbf{Z}),$$

In our case this means that $A \cong \Gamma^D$ is isomorphic to $\mathbf{Z}/2^k\mathbf{Z} \times \mathbf{Z}/2^k\mathbf{Z}$ with the action of $\sigma \in \mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q})$ described by

$$\begin{pmatrix} \chi_{2^k}(\sigma)\, \chi_{2^k}(\sigma) & (\chi_{2^{k+1}}(\sigma^{-1})-1)/2 \\ 0 & 1 \end{pmatrix}$$

in terms of the basis $\bar{v}^D = \binom{1}{0}$, $\varrho^D = \binom{0}{1}$. But $\chi_{2^k}(\sigma)\,(\chi_{2^{k+1}}(\sigma^{-1})-1)/2 \equiv (1-\chi_{2^{k+1}}(\sigma))/2 \pmod{2^k}$ so we have completed the proof of this theorem. ∎

Proof of Theorem 1.1. Since the quotient group $\mathrm{Gal}(L(\zeta)/K(\zeta))$ of Γ is spanned by v, the dual subgroup $\Sigma_0 = \ker(J_0(N) \to J_0^{\#}(N))$ is spanned by $\bar{v}^D$ in A. One checks easily that $d \in \ker(J_0^{\#}(N) \to J_0^{\#\#}(N))$ corresponds to the image of ϱ in A/Σ_0 under Cartier duality, so we can see by

Theorem 4.13(c) that $v \cdot c \in A$ is represented by some vector $\binom{*}{2}$ in A. But since $v \cdot c \in A$ is $\mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q})$-invariant, we can use Theorem 5.16(b) to conclude that $v \cdot c = \binom{1}{2}$.

The odd part of $J_0(N)[I]$ is a direct product $\mu_v \times \mathbf{Z}/v\mathbf{Z}$. The constant part is generated by $2^k c$, so we can choose a basis g_1, $2^k c$ so that for any $\sigma \in \mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q})$, $\sigma(g_1) = \chi_v(\sigma)\, g_1$ and $\sigma(2^k c) = 2^k c$. Taking the basis consisting of g_1 and $g_2 = (2^k c - g_1)\,(v+1)/2$ instead, we have $2^k c = g_1 + 2g_2$ and σ acts via

$$\begin{pmatrix} \chi_v(\sigma) & (1-\chi_{2v}(\sigma))/2 \\ 0 & 1 \end{pmatrix}.$$

Now pick integers a, b such that $va + 2^k b = 1$. Then

$$e_1 = a\bar{v}^D + bg_1$$
$$e_2 = a\varrho^D + bg_2$$

is a basis of $J_0(N)[I]$ that clearly has all properties required in Theorem 1.1.

Finally, observe that if $N \not\equiv 1 \pmod 8$, then $n = v$ and the 2-primary part of $J_0(N)[I]$ is 0. So formally setting $\bar{v}^D = \varrho^D = 0$, we still have $v \cdot c = n \cdot c = 0 = \bar{v}^D + 2\varrho^D$ and $\varrho^D \in \Sigma$. The above argument about the prime-to-2 part works without a change, so we have proved Theorem 1.1 in this case too. ∎

Remark. As described in [15], H. W. Lenstra and K. Ribet proved a version of Theorem 1.1, where the expression $(1-\chi_{2n}(\sigma))/2$ in the statement of the theorem is replaced by a function

$$b\colon \mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q}) \to \mathbf{Z}/n\mathbf{Z},$$

satisfying the properties that for each $\sigma, \tau \in \mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q})$,

$$b(\sigma\tau) = b(\sigma) + \chi_n(\sigma)\, b(\tau),$$
$$2b(\sigma) = 1 - \chi_n(\sigma),$$

and that the kernel of b cuts out the $2n$th cyclotomic field. We shall show here that his result is strictly weaker than Theorem 1.1.

Indeed, let $b_0(\sigma) = (1-\chi_{2n}(\sigma))/2$, and let

$$\varepsilon\colon \mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q}) \to \mathbf{Z}/2\mathbf{Z}$$

be a *homomorphism* that factors through $\mathrm{Gal}(\mathbf{Q}(\mu_n)/\mathbf{Q})$. It is easy to check that for any such ε, the function

$$b(\sigma) = b_0(\sigma) + \frac{n}{2}\varepsilon(\sigma)$$

satisfies all of the above conditions. Since there is more than one choice for such a function ε when n divisible by 4, we have shown that the above result is weaker than Theorem 1.1.

ACKNOWLEDGMENTS

I thank Ken Ribet for many helpful conversations and for suggesting this problem to me. I also thank Arthur Ogus for his helpful advice. Some preliminary calculations were carred out using the computer program PARI-GP [2].

REFERENCES

1. M. H. Baker, Torsion points on modular curves, *Invent. Math.* **140**, No 3 (2000), 487–509.
2. C. Batut, D. Bernardi, H. Cohen, and M. Olivier, PARI-GP, computer software, 1995–2001.
3. J. A. Csirik, The old subvariety of $J_0(NM)$, submitted for publication.
4. P. Deligne and M. Rapoport, Les schémas de modules de courbes elliptiques, *in* "Modular Functions of One Variable II" (P. Deligne and W. Kuyk, Eds.), Lecture Notes in Mathematics, Vol. 349, pp. 143–316, Springer-Verlag, New York/Berlin, 1973.
5. F. Diamond and J. Im, Modular forms and modular curves, *in* "Seminar on Fermat's Last Theorem (Toronto, ON 1993–1994)," Canadian Mathematical Society Conference Proceedings, Vol. 17, pp. 39–133, Amer. Math. Soc., Providence, 1995.
6. B. Edixhoven, L'action de l'algèbre de Hecke sur les groupes de composantes des jacobiennes des courbes modulaires est "Eisenstein," *Astérisque* **196–197** (1992), 159–170.
7. P. E. Klimek, "Modular Functions for $\Gamma_1(N)$," Ph.D. dissertation, Berkeley, 1975.
8. D. S. Kubert and S. Lang, "Modular Units," Grundlehren der Mathematischen Wissenschaften, Vol. 244, Springer-Verlag, New York/Berlin, 1981.
9. Ju. I. Manin, Parabolic points and zeta functions of modular curves, *Izv. Akad. Nauk SSSR Ser. Math.* **36** (1972), 19–66. [In Russian]
10. B. Mazur, Modular curves and the Eisenstein ideal, *Inst. Hautes Études Sci. Publ. Math.* **47** (1977), 33–186; see also the Errata [11, pp. 187–188] for an important correction for line 18 of p. 105, also note that for some typos in Appendix I are corrected in [6, Sect. 4.4.1].
11. B. Mazur and A. Wiles, Class fields of abelian extensions of **Q**, *Invent. Math.* **76** (1984), 179–330.
12. J. S. Milne, "Étale Cohomology," Princeton Univ. Press, Princeton, NJ, 1980.
13. A. P. Ogg, Rational points on certain elliptic modular curves, *in* "Analytic Number Theory," Proc. Sympos. Pure Math., Vol. 24, pp. 221–231, Amer. Math. Soc., Providence, 1973.
14. K. A. Ribet, Irreducible Galois representations arising from component groups of Jacobians, *in* "Elliptic Curves, Modular Forms, and Fermat's Last Theorem (Hong Kong, 1993)," Ser. Number Theory, Vol. I, pp. 131–147, Internat. Press, Cambridge, MA, 1995.

15. K. A. Ribet, Torsion points on $J_0(N)$ and Galois representations, *in* "Arithmetic Theory of Elliptic Curves (Cetraro, 1997)," Lecture Notes in Mathematics, Vol. 1716, pp. 145–166, Springer-Verlag, New York/Berlin, 1999.
16. J.-P. Serre, "Local Fields," Graduate Texts in Mathematics, Vol. 67, Springer-Verlag, New York/Berlin, 1979.
17. J.-P. Serre and J. Tate, Good reduction of abelian varieties, *Ann. of Math. (2)* **88** (1968), 492–517.
18. A. Wiles, Modular curves and the class group of $\mathbf{Q}(\mu_p)$, *Invent. Math.* **58** (1980), 1–35.