# A Divisor Formula and a Bound on the $\mathbb{Q}$-gonality of the Modular Curve $X_1(N)$

MARK VAN HOEIJ AND HANSON SMITH

Abstract. We give a formula for divisors of modular units on $X_1(N)$ and use it to prove that the $\mathbb{Q}$-gonality of the modular curve $X_1(N)$ is bounded above by $\left[\frac{11N^2}{840}\right]$, where $[\bullet]$ denotes the nearest integer.

## 1. Introduction

The modular curve $X_1(N)$ parametrizes pairs $(E, \pm P)$ where $E$ is an elliptic curve and $P$ is a point of exact order $N$. As such it has been an object of interest for number theorists and arithmetic geometers. If $K$ is a field, then *$K$-gonality* of $X_1(N)$ is the minimum degree of a non-constant function $X_1(N) \to \mathbb{P}^1$ defined over $K$.

Table 1 in [5] gives the currently-best upper bounds the $\mathbb{Q}$-gonality of $X_1(N)$ for $N \leq 250$ and matching lower bounds for $N \leq 40$. Any non-constant function provides an upper bound for the gonality. The upper bounds in [5, Table 1] come from *modular units*. These are functions on $X_1(N)$ whose divisors are supported only on *cusps* (places on $X_1(N)$ where $E$ degenerates). In this note we prove a formula for the degree of a certain modular unit $F_7/F_8$. Its degree is a particularly good gonality bound when $N$ is prime, it is currently the best upper bound for all primes $N \leq 250$ except 31, 67, 101, where it is only one more.

A basis $F_2, F_3, \ldots$ of modular units was given in [5, Conjecture 1] which was proved in [16]. In order to quickly find the degree of any modular unit, a formula for the divisor of $F_k : X_1(N) \to \mathbb{P}^1$ was given at [22]. A proof for this formula was not given; the resulting degrees listed in [5, Table 1] were verified by other means. The main result in this paper is a proof for this formula (Theorem 4.2 in Section 4). As an application, Section 5 gives this bound

$$\text{Gonality}_{\mathbb{Q}}\left(X_1(N)\right) \;\leq\; \deg\left(\frac{F_7}{F_8} : X_1(N) \to \mathbb{P}^1\right) \;\leq\; \left[\frac{11N^2}{840}\right] \quad \text{if } N > 8.$$

Here $[\bullet]$ indicates rounding to the nearest integer. The second $\leq$ is an equality when $N$ is prime. The asymptotic growth $11N^2/840$ was already observed in [5, Section 2.1] and [17, page 11] (combine the factors 11/35 and 1/24) though a proof was not given.

The explicit divisors given in Theorem 4.2 have other applications as well, such as computing Galois representations for modular curves [6], computing the action of diamond operators [5], computing cuspidal class numbers of modular curves ([3], [4], [10, Chapters 5 and 6], [21], [24], [25], [26], [27]), computing optimized equations for $X_1(N)$ ([1], [18], [19]), and sporadic points on modular curves ([2], [11], [12], [15], [17]).

Section 2 reviews Puiseux expansions, elliptic curves, and division polynomials. Notation for modular units is given in Section 3. Section 4 gives the main theorem. In Section 5 we obtain the gonality bound as an application of the main theorem. Streng [16] used Siegel functions to prove [5, Conjecture 1]. This work implies another proof for Theorem 4.2, see Section 6 for details. Orders of Siegel functions are typically expressed in terms of Bernoulli polynomials. We observe that such expressions sum to piecewise linear functions (Section 6.3) when the corresponding product of Siegel functions is a modular unit.

To save the reader time, Appendix B tabulates the notations.

*Date*: May 3, 2020.

2010 *Mathematics Subject Classification.* Primary 11G16, Secondary 14H52, 11G05, 14G35, 11F03.

*Key words and phrases.* Modular curves, gonality, modular units, Siegel functions, torsion points on elliptic curves.

## 2. Preliminaries

2.1. **Places and Puiseux expansions.** If $f \in \mathbb{Q}(s)[x]$ is irreducible over $\overline{\mathbb{Q}}$, then $f$ defines an algebraic curve $C$ whose function field $\mathbb{Q}(C)$ is $\mathbb{Q}(s)[x]/(f)$.

We give a brief summary of Puiseux expansions, see [13, Chapter II] for more. A *Puiseux expansion* of $f$ at $s = 0$ is a root of $f$ in the algebraic closure of $\mathbb{Q}((s))$. This is contained in the algebraic closure of $\mathbb{C}((s))$, which is $\bigcup_{e=1}^{\infty} \mathbb{C}((s^{1/e}))$. The natural valuation

$$v_s : \mathbb{C}\left(\left(s^{1/e}\right)\right) \to \frac{1}{e}\mathbb{Z}\bigcup\{\infty\}$$

sends a non-zero series to its lowest exponent in $s$ and sends 0 to $\infty$.

For a Puiseux expansion $\mathbf{p}$, let $\mathbf{e_p}$ be the smallest $e$ for which $\mathbf{p} \in \mathbb{C}((s^{1/e}))$. From the embedding

$$\phi_{\mathbf{p}} : \mathbb{Q}(C) \to \mathbb{Q}((s))[\mathbf{p}] \subset \mathbb{C}\left(\left(s^{1/\mathbf{e_p}}\right)\right), \qquad \phi_{\mathbf{p}} : x \mapsto \mathbf{p}$$

we get a discrete valuation

$$v_{\mathbf{p}} : \mathbb{Q}(C) \to \mathbb{Z}\bigcup\{\infty\} \quad \text{given by} \quad v_{\mathbf{p}}(a) = \mathbf{e_p} \cdot v_s\left(\phi_{\mathbf{p}}(a)\right). \tag{1}$$

The factor $\mathbf{e_p}$ in (1) ensures that $v_{\mathbf{p}}(a)$ lands in $\mathbb{Z}\cup\{\infty\}$. Omitting this factor gives what we will call the *unweighted order* $v_s\left(\phi_{\mathbf{p}}(a)\right)$ of $a$, which is 1 at $a = s$ and $1/\mathbf{e_p}$ at a *local parameter*. The *residue field* $k_{\mathbf{p}}$ is defined as $\{a \in K_{\mathbf{p}} \,|\, v_{\mathbf{p}}(a) \geq 0\}$ modulo $\{a \in K_{\mathbf{p}} \,|\, v_{\mathbf{p}}(a) > 0\}$, where $K_{\mathbf{p}} = \mathbb{Q}((s))[\mathbf{p}]$.

A *place* on $C/\mathbb{Q}$ is a discrete valuation $v_P : \mathbb{Q}(C) \to \mathbb{Z}\bigcup\{\infty\}$. A place above $s = 0$ is a place with $v_P(s) > 0$. Puiseux expansions $\mathbf{p}$ and $\mathbf{p}_1$ are conjugate over $\mathbb{Q}((s))$ if and only if $v_{\mathbf{p}} = v_{\mathbf{p}_1}$, so *a place corresponds to a conjugacy class of Puiseux expansions.* A conjugacy class $\{\mathbf{p}, \ldots\}$ has $\mathbf{n_p} := \mathbf{e_p}\,\mathbf{f_p}$ elements, where $\mathbf{f_p} = [k_{\mathbf{p}} : \mathbb{Q}]$. A valuation $v_{\mathbf{p}} : \mathbb{Q}(C) \to \mathbb{Z}\bigcup\{\infty\}$ extends to $\mathbf{f_p}$ distinct valuations $\mathbb{C}(C) \to \mathbb{Z}\bigcup\{\infty\}$, so *one place on $C/\mathbb{Q}$ corresponds to $\mathbf{f_p}$ places on $C/\mathbb{C}$.*

*Example* 2.1. Let $\mathbf{p} = cs^{1/2} + \cdots$ where $c \neq 0$ and dots are terms of higher order. Then $v_s(\mathbf{p}) = 1/2$ so $\mathbf{p}^2/s = c^2 s^0 + \cdots$ has valuation 0 and hence $c^2 \in k_{\mathbf{p}}$. However, $c$ need not be in $k_{\mathbf{p}}$. In that case, to avoid constants not in $k_{\mathbf{p}}$, we rewrite $cs^{1/2}$ as $(\alpha s)^{1/2}$ where $\alpha = c^2 \in k_{\mathbf{p}}$.

**Definition 2.2.** Let $l_s(\mathbf{p})$ denote the *dominant term* (the term with lowest exponent) of a nonzero Puiseux expansion. We write $\mathbf{p}_1 \sim \mathbf{p}_2$ if and only if $l_s\left(\mathbf{p}_1\right) = l_s\left(\mathbf{p}_2\right)$. In general, $v_s\left(\mathbf{p}_1 - \mathbf{p}_2\right) \geq \min\left(v_s\left(\mathbf{p}_1\right), v_s\left(\mathbf{p}_2\right)\right)$ with equality if and only if $\mathbf{p}_1 \not\sim \mathbf{p}_2$.

Let $P$ be a place above $s = 0$ given by a Puiseux expansion $\mathbf{p} \in \mathbb{C}((s^{1/\mathbf{e_p}}))$ of $f$. Suppose we wish to compute $v_P(g)$ for some $g \in \mathbb{Q}(s)[x]$. Write $g = l\left(x - \mathbf{p}_1\right)\cdots\left(x - \mathbf{p}_n\right)$, where $l \in \mathbb{Q}(s)$ and the $\mathbf{p}_i \in \mathbb{C}((s^{1/\mathbf{e}_{\mathbf{p}_i}}))$ are the Puiseux expansions of $g$ at $s = 0$. Then $g(\mathbf{p}) = l\left(\mathbf{p} - \mathbf{p}_1\right)\cdots\left(\mathbf{p} - \mathbf{p}_n\right)$ and $v_P(g) = \mathbf{e_p} \cdot \left(v_s(l) + v_s\left(\mathbf{p} - \mathbf{p}_1\right) + \cdots + v\left(\mathbf{p} - \mathbf{p}_n\right)\right)$. If $\mathbf{p} \not\sim \mathbf{p}_i$ for each $i$, then:

$$v_P(g) = \mathbf{e_p} \cdot \left(\, v_s(l) + \min\left\{v_s(\mathbf{p}), v_s\left(\mathbf{p}_1\right)\right\} + \cdots + \min\left\{v_s(\mathbf{p}), v_s\left(\mathbf{p}_n\right)\right\}\,\right). \tag{2}$$

**Lemma 2.3.** *With $g$ and $\mathbf{p}_i$ as above, let $\mathbf{l}_i := l_s(\mathbf{p}_i)$. Suppose that $\mathbf{l}_1, \ldots, \mathbf{l}_n$ are distinct. Then $\mathbf{e}_{\mathbf{l}_i} = \mathbf{e}_{\mathbf{p}_i}$ and $k_{\mathbf{l}_i} = k_{\mathbf{p}_i}$.*

*Proof.* Note $\mathbf{e}_{\mathbf{l}_i} \leq \mathbf{e}_{\mathbf{p}_i}$ and $k_{\mathbf{l}_i} \subseteq k_{\mathbf{p}_i}$ because the ramification index and residue field of $\mathbf{p}_i$ must be at least as large as those of its dominant term $\mathbf{l}_i$. If at least one of those is not an equality, then $\mathbf{n}_{\mathbf{p}_i} > \mathbf{n}_{\mathbf{l}_i}$. In this case, $\mathbf{p}_i$ has more conjugates over $\mathbb{Q}((s))$ than $\mathbf{l}_i$, so there must be at least two conjugates with the same dominant term. Those conjugates are among $\mathbf{p}_1, \ldots, \mathbf{p}_n$ since $g \in \mathbb{Q}(s)[x]$, which implies that $\mathbf{l}_1, \ldots, \mathbf{l}_n$ are not distinct. □

2.2. **Elliptic curves, analytic viewpoint.** Let $0 < \epsilon \ll 1$ and consider the elliptic curve

$$E_\epsilon : \ y^2 = x(x-\epsilon)(x-1) \tag{3}$$

so $y = \sqrt{x(x-\epsilon)(x-1)}$. Let $E_\epsilon(\mathbb{C})$ denote the points on $E$ defined over $\mathbb{C}$. This is an additive group [14, Chapter VI], the identity $\mathcal{O}$ is the point at infinity. The period lattice is $\Lambda = \mathbb{Z}\omega_1 + \mathbb{Z}\omega_2$ where

$$\omega_1 = 2\int_1^\infty \frac{\mathrm{d}x}{y} = 2\int_0^\epsilon \frac{\mathrm{d}x}{y} = 4\,K\left(\sqrt{\epsilon}\right) = 2\pi\left(1 + \frac{1}{4}\epsilon + \frac{9}{64}\epsilon^2 + \frac{25}{256}\epsilon^3 + \cdots\right) \tag{4}$$

and

$$\omega_2 = 2\int_\epsilon^1 \frac{\mathrm{d}x}{y} = 2\int_{-\infty}^0 \frac{\mathrm{d}x}{y} = \frac{4}{i}\,K\left(\sqrt{1-\epsilon}\right) = \frac{\omega_1}{\pi i}\ln\left(\frac{16}{\epsilon} - 8 - \frac{5}{4}\epsilon + \cdots\right). \tag{5}$$

Here $K$ is the complete elliptic integral of the first kind [8, §19.2(ii)]

$$K(t) = \int_0^{\frac{\pi}{2}} \frac{d\theta}{\sqrt{1 - t^2\sin^2\theta}}.$$

In this section $\sim$ means that the $\epsilon$-dominant terms are the same, similar to Definition 2.2. For example

$$\omega_1 \sim 2\pi \quad \text{and} \quad \omega_2 \sim \frac{2}{i}\ln\left(\frac{16}{\epsilon}\right). \tag{6}$$

The notation $\approx$ will be used for approximations in intermediate steps, to indicate that they are sufficiently accurate to compute the main formulas (6),(10),(12),(13) up to $\sim$.

The Abel-Jacobi map is an isomorphism (as additive groups) from $E_\epsilon(\mathbb{C})$ to $\mathbb{C}/\Lambda$. Identify $E_\epsilon(\mathbb{C})/\pm$ with $\mathbb{P}^1(\mathbb{C})$ using $\pm P \mapsto x(P)$. Let $W := (\mathbb{C}/\Lambda)/\pm$. The Abel-Jacobi map (up to $\pm$) is a bijection:

$$\Psi : \mathbb{P}^1(\mathbb{C}) \to W, \quad \text{where} \quad \Psi(x_0) = \pm\left(\int_{x_0}^\infty \frac{\mathrm{d}x}{y} \ + \Lambda\right). \tag{7}$$

Its inverse is the Weierstrass $\wp$ function [9, 1,II,5,§1]. Each element of $W$ can be written uniquely as

$$\pm\left(r_1\omega_1 + r_2\omega_2 + \Lambda\right), \quad \text{with} \quad r_1 \in [0,1), \quad r_2 \in \left[0, \frac{1}{2}\right], \text{ and if } r_2 \in \left\{0, \frac{1}{2}\right\} \text{ then } r_1 \in \left[0, \frac{1}{2}\right]. \tag{8}$$

Although $W$ is not a group, it inherits the multiplication by $N$ map from $\mathbb{C}/\Lambda$. The order of the element (8) is $N$ if and only if $r_1, r_2 \in \mathbb{Q}$ and the least common multiple of their denominators is $N$. The image of $\mathbb{P}^1(\mathbb{R})$ under $\Psi$ is a rectangle in $W$ whose corners are the points of order 1 and 2.

Like in the modular description in Section 6.1, define the *Cartan* as $C(N) := \{0, \ldots, \lfloor N/2 \rfloor\}$. Let $W(N) \subset W$ be the set of elements of order $N$, and for each $c \in C(N)$ let $W_c(N) \subseteq W(N)$ be the subset where $r_2 = c/N \in [0, 1/2]$. Let $\mathbf{n}_c(N) := |W_c(N)|$ denote the cardinality of $W_c(N)$.

- Case $c = 0$. Then $\mathbf{n}_0(2) = 1$ and $\mathbf{n}_0(N) = \varphi(N)/2$ for $N > 2$.
- Case $0 < c < N/2$. Then $\mathbf{n}_c(N) = \varphi(d)N/d$, where $d = \gcd(c, N)$.
- Case $c = N/2$. Then $\mathbf{n}_1(2) = 2$ and $\mathbf{n}_{N/2}(N) = \varphi\left(\frac{N}{2}\right)$ for even $N > 2$.

For later use we define $\mathbf{e}_c(N), \mathbf{f}_c(N)$ with these formulas: $\mathbf{n}_c(N) = \mathbf{e}_c(N)\cdot\mathbf{f}_c(N)$ where $\mathbf{e}_2(4) := 1$ and $\mathbf{e}_c(N) := N/d$ otherwise. Define $C_c(N) := \Psi^{-1}(W_c(N)) \subset \mathbb{P}^1(\mathbb{C})$ so that

$$\bigcup_{c\in C} C_c(N) = \Psi^{-1}(W(N)) = \{\, x(P) \mid P \in E_\epsilon(\mathbb{C}) \text{ has exact order } N \,\}.$$

To prove the main theorem in Section 4 it suffices to compute these $x(P)$'s up to $\sim$. We find $C_0(2) = \{1\}$ and $C_1(2) = \{0, \epsilon\}$ from the definition. Next we compute $C_0(N)$ up to $\sim$ for $N \geq 2$.

For $C_0(N)$ we have $r_2 = 0$ and $x(P) \in [1, \infty)$. Let $y_1 = x\sqrt{x-1}$. If $\epsilon \ll |x|$ then $y \approx y_1$. For any $x_0 \in [1, \infty)$ we have

$$\Psi(x_0) = \int_{x_0}^{\infty} \frac{\mathrm{d}x}{y} \approx \int_{x_0}^{\infty} \frac{\mathrm{d}x}{y_1} = \pi - 2\arctan\left(\sqrt{x_0 - 1}\right). \tag{9}$$

Equating (9) to $r_1\omega_1 + 0\omega_2$ gives $x_0 \sim \sin(\pi r_1)^{-2} = 2/(1-\cos(2\pi r_1))$. Substituting $r_1 = a/N$ gives

$$C_0(N) \sim \left\{ \sin\left(\frac{a\pi}{N}\right)^{-2} \;\middle|\; 0 < a \leq \frac{N}{2},\ \gcd(a, N) = 1 \right\}. \tag{10}$$

For $C_{N/2}(N)$ we have $r_2 = 1/2$ and $x(P) \in [0, \epsilon]$. Let $y_0 = \sqrt{x(x-\epsilon)(-1)}$. If $|x| \ll 1$ then $y \approx y_0$. Let $x_0 \in [0, \epsilon]$. Working mod $\Lambda = \mathbb{Z}\omega_1 + \mathbb{Z}\omega_2$, see (7) and (4),(5), we have

$$\Psi(x_0) = \int_{-\infty}^{x_0} \frac{\mathrm{d}x}{y} = \frac{\omega_2}{2} + \int_0^{x_0} \frac{\mathrm{d}x}{y} \approx \frac{\omega_2}{2} + \int_0^{x_0} \frac{\mathrm{d}x}{y_0} = \frac{\pi + \omega_2}{2} - \arcsin\left(1 - \frac{2x_0}{\epsilon}\right). \tag{11}$$

Equating (11) to $r_1\omega_1 + \frac{1}{2}\omega_2$ gives $x_0 \sim \epsilon \cdot \sin(\pi r_1)^2$. Substituting $r_1 = a/N$ gives

$$C_{\frac{N}{2}}(N) \sim \left\{ \epsilon \cdot \sin\left(\pi \cdot \frac{a}{N}\right)^2 \;\middle|\; 0 \leq a \leq \frac{N}{2},\ \gcd\left(a, \frac{N}{2}\right) = 1 \right\}. \tag{12}$$

Now let $r_2 \in (0, 1/2)$ which corresponds to $\epsilon \ll |x_0| \ll 1$ under $\epsilon \to 0^+$. By equating the right-hand side of (9), or that of (11), to $r_1\omega_1 + r_2\omega_2$ and computing a series expansion we find

$$x_0 \sim -4e^{-2\pi i r_1} \left(\frac{\epsilon}{16}\right)^{2r_2}.$$

Substituting $r_1 = -a/N$ (the minus sign does not affect (13)) and $r_2 = c/N$ gives

$$C_c(N) \sim \left\{ -4\,\zeta_N^a \left(\frac{\epsilon}{16}\right)^{\frac{2c}{N}} \;\middle|\; 0 \leq a < N,\ \gcd(a, c, N) = 1 \right\}. \tag{13}$$

The $a$ and $c$ in (13) are the $a$ and $c$ appearing in the vectors in Subsection 6.1. After rewriting $\epsilon$ in terms of $s$ from Section 4, Equation (13) is enough to determine the Galois action.

2.3. **Division polynomials.** Let $K$ be a field of characteristic 0 and take $a, b \in K$ for which

$$E:\ y^2 = x^3 + ax + b \tag{14}$$

defines an elliptic curve over $K$. Following [14, Exercise 3.7], the division polynomials $Q_k \in \mathbb{Z}[x, y, a, b]$, $k = 1, 2, \ldots$ are defined by

$$Q_1 := 1, \quad Q_2 := 2y = 2\sqrt{x^3 + ax + b}, \quad Q_3 := 3x^4 + 6ax^2 + 12bx - a^2,$$
$$Q_4 := 4y(x^6 + 5ax^4 + 20bx^3 - 5a^2x^2 - 4abx - 8b^2 - a^3),$$

and the recursion relations

$$Q_{2k+1} = Q_{k+2}Q_k^3 - Q_{k-1}Q_{k+1}^3 \quad \text{for } k \geq 2 \tag{15}$$

$$Q_{2k} = \frac{(Q_{k+2}Q_{k-1}^2 - Q_{k-2}Q_{k+1}^2)Q_k}{Q_2} \quad \text{for } k \geq 3. \tag{16}$$

Recursively define $q_k$ to be $Q_k$ divided all $q_d$ with $d|k$ and $d < k$, so that $Q_k = \prod_{d|k} q_d$. One has $q_1 = 1$, $q_2 = Q_2$, $q_3 = Q_3$, $q_4 = Q_4/Q_2$, and so forth.

Division polynomials have the following properties:

(1) $Q_k$ is in $\mathbb{Z}[x, a, b]$ when $k$ is odd, and in $q_2 \cdot \mathbb{Z}[x, a, b]$ when $k$ is even.
(2) Let $\mathcal{O}$ be the identity in $(E(\overline{K}), +)$, let $E[k]$ be the points $P$ in $E(\overline{K})$ with $kP = \mathcal{O}$. Then $Q_k$ has one pole, of order $k^2 - 1$ at $\mathcal{O}$, and a root of order 1 at every $P \in E[k] - \{\mathcal{O}\}$. The roots of $q_k$ are the points of exact order $k$, denoted $E[=k] \subseteq E[k]$.

(3) The $\pm$ below means: choose only one element of each pair $\{P,-P\} \subset E[k]$ (this is not relevant for $k=2$ because $P=-P$ when $P \in E[2]$).

$$\text{If } k \text{ is odd}: \quad Q_k = k \prod_{P \in (E[k]-\{\mathcal{O}\})/\pm} (x - x(P)) \ \in \mathbb{Q}[x,a,b]. \tag{17}$$

$$\text{If } k \text{ is even}: \quad \frac{Q_k}{y} = k \prod_{P \in (E[k]-E[2])/\pm} (x - x(P)) \ \in \mathbb{Q}[x,a,b]. \tag{18}$$

$$\text{If } k > 2: \quad q_k = a_k \prod_{P \in E[=k]/\pm} (x - x(P)) \ \in \mathbb{Q}[x,a,b],$$

where $a_k = p$ if $k$ is a prime power and 1 otherwise.

$$\text{For } k = 2: \quad q_2^2 = 4 \prod_{P \in E[=2]} (x - x(P)) \ \in \mathbb{Q}[x,a,b].$$

The formulas imply that $Q_k$ is square-free, and if $d|k$, then $Q_d|Q_k$. Let $m_k$ denote the number of elements of $E[=k]$. We have $m_2 = 3$, $m_3 = 8$, and $12|m_k$ when $k > 3$. Note that $\deg_x(q_k) = m_k/2$.

## 3. Equations for $X_1(N)$ and modular units

The definitions of $Q_k$ and $q_k$ are not completely canonical; recurrence relations (15) and (16) are preserved under *scaling*. Scaling means multiplying $Q_k$ by $\alpha^{k^2-1}$, and $q_k$ by $\alpha^{m_k}$, for some fixed $\alpha \neq 0$. To obtain expressions that are independent of scaling, we take quotients

$$\tilde{Q}_k = \frac{Q_k}{q_2^{(k^2-1)/3}} \quad \text{and} \quad \tilde{q}_k = \frac{q_k}{q_2^{m_k/3}}. \tag{19}$$

As before we have $\tilde{Q}_k = \prod_{d|k} \tilde{q}_d$. Since $\tilde{Q}_k = \tilde{q}_k = 1$ for $k \in \{1,2\}$, we have $\tilde{Q}_k = \tilde{q}_k$ for $k < 6$. To avoid the fractional exponent in (19), we also introduce

$$F_3 = \tilde{q}_3^3 = \frac{q_3^3}{q_2^8} \quad \text{and} \quad F_k = \tilde{q}_k \text{ for } k > 3. \tag{20}$$

Let $\tilde{Q}_{k\backslash 3}$ be $\tilde{Q}_k/\tilde{q}_3$ if $3|k$ and $\tilde{Q}_k$ otherwise. Because $\tilde{Q}_k$ comes from $Q_k$ by scaling, it satisfies the recurrence relations. These relations inductively show that

$$\tilde{Q}_{k\backslash 3} = \prod_{3 \neq d|k} \tilde{q}_d \ \in \mathbb{Z}[F_3, F_4]. \tag{21}$$

Assuming that $F_3, F_4 \in \mathbb{Q}(x,a,b)$ are algebraically independent over $\mathbb{Q}$, the Appendix shows (21), and that $\tilde{Q}_{k\backslash 3}$ is *primitive* in $\mathbb{Z}[F_3,F_4]$, i.e. the gcd of the coefficients in $\mathbb{Z}$ is 1. The product in (21) is square-free since (17) and (18) are square-free. Then by induction $F_4, F_5, F_6, \ldots$ from (20) are primitive, co-prime, and square-free in $\mathbb{Z}[F_3,F_4]$.

**Henceforth**, $E$ will be the curve

$$E: y^2 = x^3 - 3j_0 x - 2j_0, \quad \text{where} \quad j_0 := \frac{j}{j-1728}. \tag{22}$$

Now $E$ is defined over $K = \mathbb{Q}(j)$, where $j$ is transcendental over $\mathbb{Q}$. So $a$ and $b$ from Section 2.3 will be $-3j_0$ and $-2j_0$ from here on. Now $q_3, q_4, \ldots$ are in $\mathbb{Z}[x, j_0]$. The $j$-invariant of $E$ is $j$, and the $j$-invariant of $E_\epsilon$ is

$$\frac{2^8(\epsilon^2-\epsilon+1)^3}{\epsilon^2(\epsilon-1)^2}. \tag{23}$$

In Section 4 we will equate (23) to $j$ in order to translate formulas given in terms of $\epsilon$ in Section 2.2 to similar formulas for $E$. Up to a simple transformation, $E$ is the universal elliptic curve $E_j$ from

Diamond and Shurman's book [7]. Sections 7.5 and 7.7 in [7] show that the modular curve $X_1(N)$ can be represented with the equation $q_N$ when $N > 2$. In particular, $q_N$ is irreducible in $\mathbb{Q}[x, j_0]$. Likewise $F_N$ is irreducible in $\mathbb{Z}[F_3, F_4]$. Although $q_2 \notin \mathbb{Z}[x, j_0]$, its square $4(x^3 - 3j_0x - 2j_0)$ is an equation for $X_1(2)$ that lies in $\mathbb{Z}[x, j_0]$.

## 4. The valuation of a division polynomial at a cusp

Recall from Section 2.1 that a *place* on $X_1(N)/\mathbb{Q}$ is a discrete valuation $v_P : \mathbb{Q}(X_1(N)) \to \mathbb{Z} \bigcup \{\infty\}$. Such a place is a *cusp over* $\mathbb{Q}$ when $v_P(j) < 0$. A function $g \in \mathbb{Q}(X_1(N))$ is called a *modular unit* if every place with $v_P(g) \neq 0$ is a cusp. If $k \neq N$ and $k > 2$, then $F_k$ is a *modular unit* in $X_1(N)$, see [5, Section 2]. However, to obtain a modular unit from $q_2$, it was necessary to take its $4^{\text{th}}$ power and scale it to

$$F_2 = \frac{q_2^4}{1728 j_0^2 (j_0 - 1)}. \tag{24}$$

Let $s = 1/j$, then *a cusp over* $\mathbb{Q}$ is a place above $s = 0$, which corresponds to a *conjugacy class of Puiseux expansions at* $s = 0$, see Section 2.1. Conjugation is always over $\mathbb{Q}((s))$ in this paper.

From (15),(16) one can compute $Q_2, Q_3, \ldots$ and then $q_2^2, q_3, q_4, \ldots \in \mathbb{Z}[x, j_0] \subset \mathbb{Q}(s)[x]$. We computed Puiseux expansions of $q_N$ (or $q_2^2$ if $N = 2$) at $s = 0$ for $N \leq 9$. Newton's algorithm gives arbitrarily many terms, but only dominant terms will be needed. Table 1 lists the dominant term of $\mathbf{p} + 1$ for *one* Puiseux expansion $\mathbf{p}$ *from each* conjugacy class $\{\mathbf{p}, \ldots\}$ (which has $\mathbf{n_p} := \mathbf{e_p}\,\mathbf{f_p}$ elements, see Section 2.1).

We use $\mathbf{p} + 1$ and $x + 1$ rather than $\mathbf{p}$ and $x$ because when $j \to \infty$ the curve $E$ in (22) becomes singular at $x = -1$. This is in contrast to $E_\epsilon$ which becomes singular at $x = 0$ when $\epsilon \to 0$.

| | $l_s(\mathbf{p}+1)$ with $\mathbf{p} \in C_0$ | $l_s(\mathbf{p}+1)$ $\mathbf{p} \in C_1$ | $l_s(\mathbf{p}+1)$ $\mathbf{p} \in C_2$ | $l_s(\mathbf{p}+1)$ $\mathbf{p} \in C_3$ | $l_s(\mathbf{p}+1)$ $\mathbf{p} \in C_4$ |
|---|---|---|---|---|---|
| $q_2^2$ | $3$ | $-24s^{1/2}$ | | | |
| $q_3$ | $4$ | $-12s^{1/3}$ | | | |
| $q_4$ | $6$ | $-12s^{1/4}$ | $0\,s^{1/2} - 672s$ | | |
| $q_5$ | $3\sin(\pi/5)^{-2}$ | $-12s^{1/5}$ | $-12s^{2/5}$ | | |
| $q_6$ | $12$ | $-12s^{1/6}$ | $-12(-s)^{1/3}$ | $-12s^{1/2}$ | |
| $q_7$ | $3\sin(\pi/7)^{-2}$ | $-12s^{1/7}$ | $-12s^{2/7}$ | $-12s^{3/7}$ | |
| $q_8$ | $3\sin(\pi/8)^{-2}$ | $-12s^{1/8}$ | $-12(-s)^{1/4}$ | $-12s^{3/8}$ | $-12(2s)^{1/2}$ |
| $q_9$ | $3\sin(\pi/9)^{-2}$ | $-12s^{1/9}$ | $-12s^{2/9}$ | $-12(\zeta_3 \cdot s)^{1/3}$ | $-12s^{4/9}$ |

Table 1. $l_s(\mathbf{p}+1)$ for one $\mathbf{p}$ from each conjugacy class $C_i$ over $\mathbb{Q}((s))$

The equation for $X_1(2)$ is $q_2^2 = 4(x^3 - 3j_0x - 2j_0)$, where $j_0 = j/(j - 1728) = 1/(1 - 1728s)$. To illustrate Table 1 for $N = 2$, factor $q_2^2 = 4(x - \mathbf{p}_0)(x - \mathbf{p}_{1a})(x - \mathbf{p}_{1b}) \in \overline{\mathbb{Q}((s))}[x]$. Row $q_2^2$ in Table 1 gives $l_s(\mathbf{p}_0 + 1) = 3$, $l_s(\mathbf{p}_{1a} + 1) = -24s^{1/2}$, and its conjugate $l_s(\mathbf{p}_{1b} + 1) = 24s^{1/2}$. This means

$$q_2^2 = 4((x+1) - 3 + \cdots)\left((x+1) + 24s^{1/2} + \cdots\right)\left((x+1) - 24s^{1/2} + \cdots\right), \tag{25}$$

where the dots indicate terms with higher powers of $s$. Likewise, for $N > 2$,

$$q_N = a_N \prod_{c=0}^{\lfloor N/2 \rfloor} (x - \mathbf{p}_{c,*}),$$

where $l_s(\mathbf{p}_{c,*} + 1)$ are the conjugates of the term listed in row $q_N$, column $C_c$.

*Example* 4.1. Counting conjugates, row $q_8$ in Table 1 gives two $\mathbf{p}$'s with $v_s(\mathbf{p}+1)=0$, eight $\mathbf{p}$'s with $v_s(\mathbf{p}+1)=1/8$, four with $1/4$, eight with $3/8$, and two with $1/2$. Indeed, $\deg_x(q_k)=m_8/2$ equals $2+8+4+8+2$.

Now take as an example the conjugacy class $C_1$ (a *cusp over* $\mathbb{Q}$) on $X_1(3)$. Row $q_3$, column $C_1$ gives $\mathbf{p}+1=-12s^{1/3}+\cdots$. Viewing $q_8$ as an element of $\mathbb{Q}(X_1(3))=\mathbb{Q}(s)[x]/(q_3)$, we can insert this data into Equation (2) to find

$$v_P(q_8)=3\cdot\left(2\min\left(\frac{1}{3},0\right)+8\min\left(\frac{1}{3},\frac{1}{8}\right)+4\min\left(\frac{1}{3},\frac{1}{4}\right)+8\min\left(\frac{1}{3},\frac{3}{8}\right)+2\min\left(\frac{1}{3},\frac{1}{2}\right)\right) \quad (26)$$

(the $+1$'s cancelled out). Omitting the factor of 3 gives the *unweighted order* from Section 2.1.

If $1<N$, $k\le 9$, and $N\neq k$, then Example 4.1 shows how one can use Table 1 to compute the valuation of $q_k$ (or $q_2^2$ if $k=2$) at any cusp of $X_1(N)$. To find a general formula, we will show that Observations (1)–(6) below, which hold in Table 1, hold for all $N>1$.

(1) $q_N$ ($q_2^2$ if $N=2$) has $\lfloor N/2\rfloor+1$ conjugacy classes (a.k.a. Galois orbits) $C_0,C_1,\ldots,C_{\lfloor N/2\rfloor}$ of Puiseux expansions at $s=0$. We number them so that if $\mathbf{p}\in C_c$ then $v_s(\mathbf{p}+1)=c/N$ except when $(N,c)=(4,2)$. This unique exceptional case is the irregular cusp of $X_1(4)$, where the $s^{c/N}$ term in Table 1 is $0s^{1/2}$ and $v_s(\mathbf{p}+1)=1$ instead.
(2) $C_0$ has $l_s(\mathbf{p}+1)=12/\left(2-\zeta_N-\zeta_N^{-1}\right)=3\sin(\pi/N)^{-2}$. The residue field is $\mathbb{Q}\left(\zeta_N+\zeta_N^{-1}\right)$.
(3) If $0<c<N/2$, then $\mathbf{p}\in C_c$ has $l_s(\mathbf{p}+1)=-12\left(\zeta_d\cdot s\right)^{c/N}$ (always up to conjugation) with $d=\gcd(c,N)$ and residue field $\mathbb{Q}(\zeta_d)$.
(4) If $N>4$ is even, then $\mathbf{p}\in C_{N/2}$ has

$$l_s(\mathbf{p}+1)=-24s^{1/2}+3\sin^2(\pi/N)\cdot 16s^{1/2}=-12(\beta\cdot s)^{1/2},$$

where $\beta:=\left(\zeta_N+\zeta_N^{-1}\right)^2$. The residue field is $\mathbb{Q}(\beta)$ (recall Example 2.1).
(5) $C_c\subset\mathbb{C}((s^{1/\mathbf{e}_c(N)}))$ has precisely $\mathbf{n}_c(N)$ elements, and the residue field has degree $\mathbf{f}_c(N)$ with $\mathbf{n}_c$, $\mathbf{e}_c$, $\mathbf{f}_c$ as in Section 2.2.
(6) Every $\mathbf{p}\in\bigcup_{N,c}C_c(N)$ has a unique $l_s(\mathbf{p}+1)$, so Equation (2) holds for all combinations. This implies that Example 4.1 generalizes to Theorem 4.2 below.

To see why Observations (1)–(6) hold, note that the curve $E_\epsilon$ in Section 2.2 differs from $E$ by the transformation

$$T:x\mapsto\mathbf{p}_{1a}+(\mathbf{p}_0-\mathbf{p}_{1a})x=(-1-24s^{1/2}+\cdots)+(3+24s^{1/2}+\cdots)x$$

that sends $0,\epsilon$, and 1 to $\mathbf{p}_{1a},\mathbf{p}_{1b}$, and $\mathbf{p}_0$, respectively. From $T(\epsilon)=\mathbf{p}_{1b}$ we find $\epsilon=16s^{1/2}+\cdots$ which can also be computed by equating $j=1/s$ to (23). Section 2.2 gives the $\epsilon$-dominant terms. Substituting $\epsilon\sim 16s^{1/2}$ and applying $T$ yields $l_s(\mathbf{p}+1)$ for every Puiseux expansion of $q_N$. Observation (6) immediately follows from this, but then Lemma 2.3 shows that $\mathbf{e}_\mathbf{p}$ and $k_\mathbf{p}$ can be read from $l_s(\mathbf{p}+1)$, and the remaining observations follow.

**Theorem 4.2. (MinFormula)**. *For $t\in[0,1/2]$, we define the following* unweighted order functions. *For $k=2$, define $v_2(t)=4t-1$, and for $k>2$, define*

$$v_k(t)=s_k\cdot\left(-\frac{m_k}{3}t+\sum_{c=1}^{\lfloor k/2\rfloor}\mathbf{n}_c(k)\min\left(t,\frac{c}{k}\right)\right), \quad (27)$$

*where $s_3=3$ and $s_k=1$ for $k>3$. Recall that $C_c(N)$ is a conjugacy class of Puiseux expansions, giving one cusp of $X_1(N)/\mathbb{Q}$, or a Galois orbit with $\mathbf{f}_c(N)$ cusps of $X_1(N)/\overline{\mathbb{Q}}$.*

*Let $2<N\neq k>1$ and $0\le c\le N/2$, then $F_k$, viewed as element of $\mathbb{Q}(s)[x]/(q_N)=\mathbb{Q}(X_1(N))$, has order $\mathbf{e}_c(N)\cdot v_k\left(\frac{c}{N}\right)$ at $C_c(N)$.*

If $N = 2$ we can not directly apply this formula to $F_k$ due to its denominator $q_2$, but the formula still holds for products where $q_2$ cancels out, such as $F_2^2 F_3$ and $F_2^{m_k/12} F_k$ for $k > 3$.

*Proof.* Observations (1)–(6) imply that the computation in Example 4.1 works in general, so

$$\mathbf{e}_c(N) \sum_{j=0}^{\lfloor k/2 \rfloor} \mathbf{n}_j(k) \min\left(\frac{c}{N}, \frac{j}{k}\right) \tag{28}$$

is the order of $q_k$ (or $q_2^2$ if $k = 2$) at $C_c(N)$ for any $N, k > 1$ with $N \neq k$. Theorem 4.2 follows by applying Equation (28) to $F_k$ in Equations (20) and (24), simplifying $\min(t, 0) = 0$ and $\min(t, 1/2) = t$, and noting that the denominator $1728 j_0^2 (j_0 - 1)$ in Equation (24) has a root of order 1 at $s = 0$. □

*Remark* 4.3. A cusp over $\mathbb{Q}$ corresponds to $\mathbf{f}_c(N)$ cusps over $\overline{\mathbb{Q}}$. Since the degree of the divisor of a function is zero,

$$\sum_{c=0}^{\lfloor N/2 \rfloor} \mathbf{f}_c(N)\, \mathbf{e}_c(N)\, v_k\left(\frac{c}{N}\right) = 0.$$

If $N$ is prime, then $\mathbf{e}_c(N)\, \mathbf{f}_c(N) = \mathbf{n}_c(N)$ is $N$ for $c > 0$, and $v_k(0) = 0$, so

$$\sum_{c=0}^{\lfloor N/2 \rfloor} N\, v_k\left(\frac{c}{N}\right) = 0.$$

Letting $N \to \infty$, we see $\int_{t=0}^{1/2} v_k(t) \mathrm{d}t = 0$. Since $\int_{t=0}^{1/2} \left(\min(t, \frac{c}{k}) - 4\frac{c}{k}(1 - \frac{c}{k})t\right) \mathrm{d}t$ equals 0 for any $\frac{c}{k} \in [0, 1/2]$, we do not need a formula for $m_k$, and can instead rewrite Equation (27) as

$$v_k(t) = s_k \sum_{0<c<k/2} \mathbf{n}_c(k) \left(\min\left(t, \frac{c}{k}\right) - 4\frac{c}{k}\left(1 - \frac{c}{k}\right) t\right), \qquad \text{for } k \geq 3. \tag{29}$$

This is the formula implemented in [22]. The sum (29) does not change if one replaces $0 < c < k/2$ by $0 \leq c \leq k/2$ because the summand vanishes at $c \in \{0, k/2\}$. Equation (29) with the factor $s_k$ removed gives the unweighted order function for $\tilde{q}_k$ (recall that $F_3 = \tilde{q}^3$ and $F_k = \tilde{q}_k$ if $k > 3$).

## 5. The degree of $F_7/F_8$ in $X_1(N)$

In this section we use Theorem 4.2 to prove an upper bound for the $\mathbb{Q}$-gonality of $X_1(N)$. Let $v(t) := v_7(t) - v_8(t)$, and let $m(t) := \max(0, v(t))$ as in Figure 1. Define

$$B_0(N) := \sum_{0<c<N/2} \mathbf{n}_c(N)\, m\left(\frac{c}{N}\right) \quad \text{and} \quad B_1(N) := \sum_{0<c<N/2} N\, m\left(\frac{c}{N}\right).$$

Theorem 4.2 gives

$$\operatorname{div}\left(\frac{F_7}{F_8}\right) = \sum_{0<c<N/2} \mathbf{e}_c(N)\, v\left(\frac{c}{N}\right) C_c, \qquad \text{so} \quad \deg\left(\frac{F_7}{F_8}\right) = B_0(N) \leq B_1(N). \tag{30}$$

We omit the terms $c = 0$ and $c = N/2$ in these sums because $v$ vanishes there. By Equation (27)

$$v(t) = 7\min\left(t, \frac{1}{7}\right) + 7\min\left(t, \frac{2}{7}\right) + 7\min\left(t, \frac{3}{7}\right) - 8\min\left(t, \frac{1}{8}\right) - 4\min\left(t, \frac{1}{4}\right) - 8\min\left(t, \frac{3}{8}\right) - 2t.$$

**Lemma 5.1.** *If $N$ is relatively prime to $420 = 3 \cdot 4 \cdot 5 \cdot 7$, then $B_1(N) = [11N^2/840]$. In general, $B_1(N) \leq [11N^2/840] + 2$ (where equality implies $7 \mid N$), and $B_0(N) \leq [11N^2/840]$.*

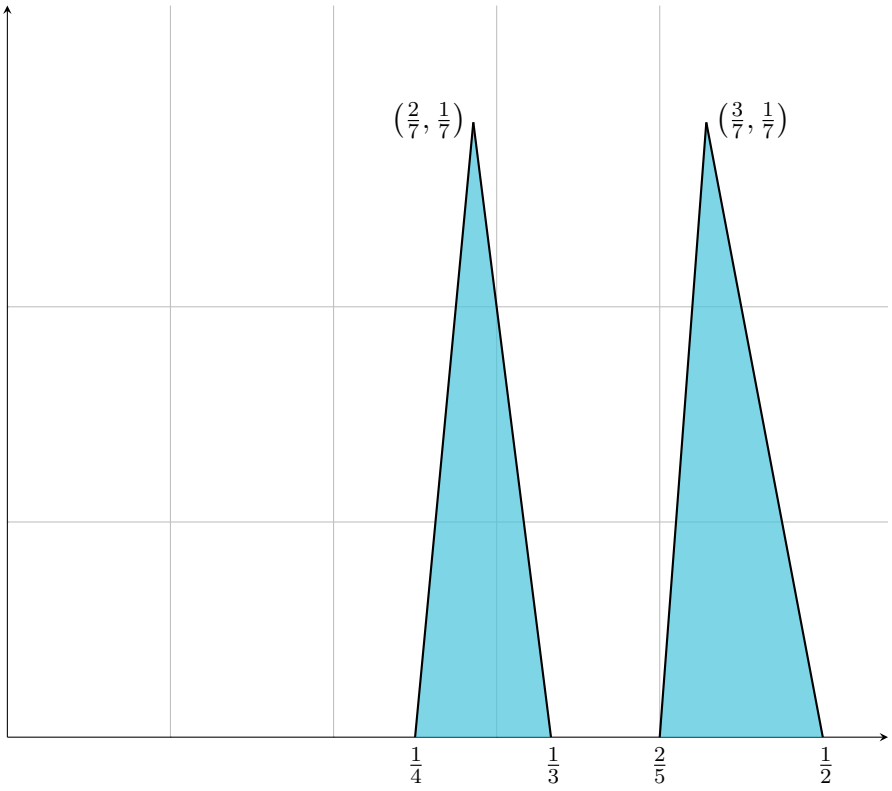


FIGURE 1. The function $m(t)$ graphed from 0 to $\frac{1}{2}$

*Proof.* Consider the intervals $I_1 := (1/4, 2/7]$, $I_2 := (2/7, 1/3)$, $I_3 := (2/5, 3/7]$, and $I_4 := (3/7, 1/2)$. These intervals partition the support of $m(t)$; see Figure 1. We define the functions $m_1(t) := 4t - 1$, $m_2(t) := 1 - 3t$, $m_3(t) := 5t - 2$, and $m_4(t) := 1 - 2t$. The graphs of the $m_j(t)$ over $I_j$ are exactly the line segments in Figure 1. We see $m(t) = m_j(t)$ if $t \in I_j$ and 0 otherwise.

Our goal is to bound

$$B_1(N) = \sum_{j=1}^{4} \sum_{\substack{c/N \in I_j \\ c \in \mathbb{Z}}} N\, m_j\Big(\frac{c}{N}\Big). \tag{31}$$

Since $N\, m_j(c/N) \in \mathbb{Z}$, we have $B_1(N) \in \mathbb{Z}$. Note that $B_1(N)$ is a Riemann sum of

$$N^2 \int_{t=0}^{1/2} m(t)\mathrm{d}t = N^2 \sum_{j=1}^{4} \int_{I_j} m_j(t)\mathrm{d}t = \frac{11}{840} N^2.$$

Since $m(t)$ is piece-wise linear, any error in $B_1(N)$, viewed as an approximation to this integral, must come from the corners:

$$\frac{1}{4},\ \frac{2}{7},\ \frac{1}{3},\ \frac{2}{5},\ \frac{3}{7},\ \text{and}\ \frac{1}{2}.$$

This error depends only on $N$ modulo $420 = 4 \cdot 7 \cdot 3 \cdot 5$. To demonstrate this, let $c_{1a}$ and $c_{1b}$, respectively, be the minimum and maximum integer $c$ with $c/N \in I_1$. Then $c_{1a}$ equals

$$\frac{N+4}{4}, \frac{N+3}{4}, \frac{N+2}{4},\ \text{or}\ \frac{N+1}{4}$$

depending on whether $N$ is, respectively, 0, 1, 2, or 3 mod 4. Likewise, the expression for $c_{1b}$ in terms of $N$ depends only on $N$ mod 7. Considering the intervals $I_2$, $I_3$, and $I_4$, we have a total of

$4 \cdot 7 \cdot 3 \cdot 5 = 420$ cases. Hence, the difference between the integral and its Riemann sum $B_1(N)$ depends only on $N$ mod 420.

As example we cover one of the 420 cases, namely $N \equiv 32 \bmod 420$. Here $c_{1a} = (N+4)/4$ and $c_{1b} = (2N-1)/7$, so

$$\sum_{\substack{c/N \in I_1 \\ c \in \mathbb{Z}}} N\, m_1\Big(\frac{c}{N}\Big)$$

has $n := c_{1b} - c_{1a} + 1 = N/28 - 1/7$ terms. The average of these $n$ terms is

$$\frac{1}{2}\Big(N\, m_1\Big(\frac{c_{1a}}{N}\Big) + N\, m_1\Big(\frac{c_{1b}}{N}\Big)\Big) = \frac{15N}{14} + \frac{5}{7},$$

so the $j = 1$ part of $B_1(N)$ in Equation (31) is $(N/28 - 1/7) \cdot (15N/14 + 5/7)$. Repeating this computation for $j = 2, 3, 4$ and summing, we find

$$B_1(N) = \frac{11N^2}{840} + \frac{43}{105}.$$

Since $|43/105| < 1/2$ we have $B_0(N) \le B_1(N) = [11N^2/840]$ for any $N \equiv 32 \bmod 420$.

In the same way we calculated the difference between $B_1(N)$ and $\big[11N^2/840\big]$ for all 420 cases, the programs are available at [23]. In all cases with $\gcd(N, 420) = 1$, we found $B_1(N) = \big[11N^2/840\big]$.

If $N$ is prime and $c > 0$, then the factor $\mathbf{n}_c(N)$ in the definition of $B_0(N)$ is $N$, and hence $B_0(N) = B_1(N)$. So for primes $N > 7$ we find

$$\deg\left(F_7/F_8\right) = B_0(N) = B_1(N) = \big[11N^2/840\big]\,. \tag{32}$$

For cases with $\gcd(N, 7) = 1 \ne \gcd(N, 2 \cdot 3 \cdot 5)$ the computation found $B_1(N) \le \big[11N^2/840\big] + 1$. Moreover, in these cases there is a $c/N$ in some $I_j$ with $\mathbf{n}_c(N) < N$. Then $B_0(N) < B_1(N)$, and so $B_0(N) \le \big[11N^2/840\big]$.

For the remaining cases $\gcd(N, 7) \ne 1$ we found $B_1(N) \le \big[11N^2/840\big] + 2$. The smallest $N$ for which that is sharp is $N = 49$. We check that a multiple of $7/N$ is in each of the intervals $(1/4, 1/3)$ and $(2/5, 1/2)$. These two multiples $c/N$ of $7/N$ each have $\mathbf{n}_c(N) < N$. So $B_0(N) \le B_1(N) - 2$. Hence we still have $B_0(49) \le \big[11 \cdot 49^2/840\big]$. The next $N$ with $B_1(N) = \big[11N^2/840\big] + 2$ is $N = 91$. For $N \ge 91$ the intervals $(1/4, 1/3)$ and $(2/5, 1/2)$ each have at least 7 consecutive $c/N$'s, and so $\gcd(c, N) > 1$ (which implies $\mathbf{n}_c(N) < N$) happens at least once in each of those intervals. Then the same argument shows that $B_0(N) \le \big[11N^2/840\big]$. □

From Lemma 5.1 we obtain:

**Theorem 5.2.** *For $N \ne 7, 8$ the modular unit*

$$\frac{F_7}{F_8} : X_1(N) \to \mathbb{P}^1$$

*has degree*

$$\deg\left(\frac{F_7}{F_8}\right) = B_0(N) \le \left[\frac{11N^2}{840}\right],$$

*with equality when $N$ is prime. If $N > 8$, then this is an upper bound for the gonality.*

*Proof.* We need $N \ne 7, 8$ to ensure $F_7, F_8 \ne 0$. If $N < 7$ then $B_0(N) = 0$ which means that $F_7/F_8$ is constant. The degree of a *non-constant* function is an upper bound for the gonality. It is easy to check that $B_0(N) > 0$ for $N > 8$. □

If $N$ is not prime, then the gonality is usually smaller than $B_0(N)$, see [5, Table 1]. If $N > 8$ is prime, then equation (32) gives an excellent gonality bound; the only primes $N < 250$ for which a sharper bound is known are 31, 67, 101 and for these cases, that bound is only one less.

## 6. Second proof for MinFormula

6.1. **Cusps: A modular interpretation.** Take the congruence subgroup

$$\Gamma_1(N) = \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathrm{SL}_2(\mathbb{Z}) \;\middle|\; \begin{bmatrix} a & b \\ c & d \end{bmatrix} \equiv \begin{bmatrix} 1 & * \\ 0 & 1 \end{bmatrix} \bmod N \right\}$$

where $*$ indicates the entry is unspecified. The extended complex upper half plane is

$$\overline{\mathcal{H}} = \mathcal{H} \cup \mathbb{Q} \cup \{\infty\},$$

where $\mathcal{H}$ is the usual complex upper half plane. The groups $\Gamma_1(N) \subseteq \mathrm{SL}_2(\mathbb{Z})$ act on the extended complex upper half plane $\overline{\mathcal{H}}$ by fractional linear transformations. The quotient is the modular curve $X_1(N)$.

Following [7, Chapter 3.8] and similar to Section 2.2, we represent cusps of $X_1(N)/\overline{\mathbb{Q}}$ with pairs of order $N$ vectors

$$\pm \begin{bmatrix} a \\ c \end{bmatrix} \in (\mathbb{Z}/N\mathbb{Z})^2 .$$

The Galois action on the cusps can be represented with matrices of the form

$$\pm \begin{bmatrix} y & z \\ 0 & 1 \end{bmatrix} \in \mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z})$$

on the order $N$ vectors in $(\mathbb{Z}/N\mathbb{Z})^2$, see [7, Sections 7.6-7.7]. Two vectors

$$\begin{bmatrix} a' \\ c' \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} a \\ c \end{bmatrix}$$

represent the same cusp when

$$\begin{bmatrix} a' \\ c' \end{bmatrix} = \pm \begin{bmatrix} a + jc \\ c \end{bmatrix}$$

for some $j \in \mathbb{Z}$. Two cusps represented this way are in the same Galois orbit if and only if $c = \pm c'$. So each Galois orbit is uniquely determined by $\pm c$, in other words, by an element of the *Cartan* $C(N) := (\mathbb{Z}/N\mathbb{Z})/\pm$, which is identified with $\{0, \ldots, \lfloor N/2 \rfloor\}$. We will denote such orbit by $C_c(N)$. Let $\mathbf{n}_c(N), \mathbf{e}_c(N), \mathbf{f}_c(N)$ be as in Section 2.2. There are $\mathbf{f}_c(N)$ cusps in $C_c(N)$, each of which is represented by $\mathbf{e}_c$ pairs of vectors in $(\mathbb{Z}/N\mathbb{Z})^2$, for a total of $\mathbf{n}_c = \mathbf{e}_c \mathbf{f}_c$ pairs.

The *width* of a cusp ([7, pages 59 and 60]) is defined as follows. Let $A \in \mathrm{SL}_2(\mathbb{Z})$ be such that $A \cdot \left[\begin{smallmatrix} a \\ c \end{smallmatrix}\right] = \infty = \left[\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right]$. The *width* $\mathbf{e}_{\left[\begin{smallmatrix} a \\ c \end{smallmatrix}\right]}(N)$ is the smallest positive integer for which

$$A \begin{bmatrix} 1 & \mathbf{e}_{\left[\begin{smallmatrix} a \\ c \end{smallmatrix}\right]}(N) \\ 0 & 1 \end{bmatrix} A^{-1} \in \Gamma_1(N).$$

A computation shows that this is $N/\gcd(c, N)$. So the *width* $\mathbf{e}_{\left[\begin{smallmatrix} a \\ c \end{smallmatrix}\right]}(N)$ is $N/\gcd(c, N)$, which equals the number $\mathbf{e}_c(N)$ from Sections 2.2 and 4 with *one exception*, namely $C_2(4)$. The cusp corresponding to $\left[\begin{smallmatrix} 1 \\ 2 \end{smallmatrix}\right]$ on $X_1(4)$ is the lone cusp in the orbit $C_2(4)$. It is the only irregular cusp for any modular curve $X_1(N)$, $X_0(N)$, or $X(N)$ ([7, page 75]). It has width 2, but it has 'order' 1. Throughout this paper $\mathbf{e}_c(N)$ denotes the width, except for the case $\mathbf{e}_2(4)$ where it denotes the 'order' 1.

6.2. **Siegel Functions.** We would like to define a class of functions on the complex upper half plane $\mathcal{H}$.

**Definition 6.1.** Let $(a_1, a_2) \in \mathbb{Q}^2 - \mathbb{Z}^2$. For $\tau \in \mathcal{H}$, define the *Siegel function* associated to $(a_1, a_2)$, denoted $g_{(a_1,a_2)}$, by the product

$$g_{(a_1,a_2)}(\tau) := -q^{\frac{1}{2}\mathbb{B}_2(a_1)} e^{2\pi i \frac{1}{2}(a_2(a_1-1))} (1 - e^{2\pi i a_2} q^{a_1}) \prod_{n=1}^{\infty} (1 - e^{2\pi i a_2} q^{n+a_1})(1 - e^{-2\pi i a_2} q^{n-a_1}),$$

where $q = e^{2\pi i \tau}$, and $\mathbb{B}_2(x) = x^2 - x + \frac{1}{6}$ is the second Bernoulli polynomial.

One can check that adding an integral vector to $(a_1, a_2)$ does not change the order of $g_{(a_1,a_2)}$, so we can interpret $(a_1, a_2)$ as a non-zero element of $(\mathbb{Q}/\mathbb{Z})^2$.

We are interested in the divisors of Siegel functions. From the $q$-expansion, we see that

$$\operatorname{ord}_\infty g_{(a_1,a_2)} = \mathbf{e}_\infty \cdot \frac{1}{2}\,\mathbb{B}_2(a_1).$$

Recall, $\infty$ denotes the standard prime at infinity given by the equivalence class of $\left[\begin{smallmatrix}1\\0\end{smallmatrix}\right]$ under the action of $\Gamma_1(N)$ and $\mathbf{e}_\infty = 1$ is its width. Consider another cusp of the modular curve $X_1(N)$ that corresponds to the orbit of $\left[\begin{smallmatrix}a\\c\end{smallmatrix}\right]$. Let $A \in \mathrm{SL}_2(\mathbb{Z})$ be a matrix such that

$$A \cdot \begin{bmatrix}1\\0\end{bmatrix} = \begin{bmatrix}a\\c\end{bmatrix}.$$

When $g_{(a_1,a_2)}$ is a function on $X_1(N)$, the *order of* $g_{(a_1,a_2)}$ at the cusp corresponding to $\left[\begin{smallmatrix}a\\c\end{smallmatrix}\right]$ is

$$\operatorname{ord}_{\left[\begin{smallmatrix}a\\c\end{smallmatrix}\right]}\left(g_{(a_1,a_2)}\right) = \mathbf{e}_c \cdot \frac{1}{2}\,\mathbb{B}_2\left(\left\{\left[(a_1,a_2)\cdot A\right]_1\right\}\right), \tag{33}$$

where $\{\bullet\} = \bullet - \lfloor\bullet\rfloor$ denotes the *fractional part* and $[\bullet]_1$ denotes the first entry of the vector. The paper [20] has a concise description of the above for an arbitrary modular curve, but [10, Chapter 2] has a more thorough exposition for $X(N)$; specifically, see the boxed equation on page 40. The reader should note that in [10], Kubert and Lang are considering the $q^{\frac{1}{N}}$ expansion. In the remainder of this paper, we will consider Siegel functions of the form $g_{(0,a)}$, with $a$ a nonzero element of $\mathbb{Q}/\mathbb{Z}$ of order dividing $N$. Following [16], we write

$$H_k := g_{\left(0,\frac{k}{N}\right)}, \quad \text{with} \quad k \in \mathbb{Z} - N\mathbb{Z}.$$

**Caution:** In [16], Streng considers the modular curve $X^1(N)$, while we have $X_1(N)$. The isomorphism $\Gamma^1(N) \setminus \mathcal{H} \to \Gamma_1(N) \setminus \mathcal{H}$ is given by $\begin{bmatrix}0 & -1\\1 & 0\end{bmatrix}$. This isomorphism sends $g_{(a,0)}$ to $g_{(0,-a)}$; however, $g_{(0,-a)} = -g_{(0,a)}$.

The *unweighted order of* $H_k$ at $c \in C(N)$ is

$$\operatorname{uord}_c\left(H_k\right) = \frac{1}{2}\,\mathbb{B}_2\left(\left\{c \cdot \frac{k}{N}\right\}\right). \tag{34}$$

Note that $x \mapsto \mathbb{B}_2(\{x\})$ is a continuous function even though $x \mapsto \{x\}$ is not.

6.3. **Generators of the Modular Units.** Describing the modular units on a given modular curve has long been a subject of interest. A significant motivation of Kubert and Lang's text [10] is to describe the units of $X(N)$ over $\mathbb{Q}(\zeta_N)$. They show that, with the exception of some 2-torsion elements when $N$ is even, the units are generated by the Siegel functions described above.

Motivated by [5, Conjecture 1], Streng [16] has used similar methods to describe all modular units on $X^1(N)$ over $\mathbb{Q}$. Before stating the result we introduce some of the relevant objects. We start with Tate normal form.

**Lemma 6.2.** *([16, Lemma 2.1]). If $E$ is an elliptic curve over a field $K$ of characteristic 0 (such as the elliptic curves in Equations (14) and (22)) and $P$ is a point on $E$ of order greater than 3 with $x(P) \in K$, then the pair $(E, \pm P)$ is isomorphic to a unique pair of the form*

$$E_T : Y^2 + (1-C)XY - BY = X^3 - BX^2, \qquad P = (0,0), \tag{35}$$

*where $B, C \in K$ and the discriminant*

$$D = B^3(16B^2 + (1 - 20C - 8C^2)B + C(C-1)^3) \neq 0.$$

*Further, each pair $B, C \in K$ with $D \neq 0$ satisfying (35) yields an elliptic curve and with a distinguished point $P$ of order greater than 3.*

This form $E_T$ is called *Tate normal form*. Let $K = \mathbb{Q}(j)$ and $E$ be as in Equation (22) and let $K_0 = K(x_0)$ where $x_0$ is transcendental over $K$. Let

$$P = \left(x_0, \sqrt{x_0^3 - 3j_0x_0 - 2j_0}\right).$$

Sending $P$ to $(0,0)$ and $E$ to Tate normal form with affine linear transformations results in expressions $B, C \in K_0$ (computation at [23]). Identify $x_0$ with $x$ so that $K_0$ becomes $\mathbb{Q}(x,j)$. Then $B, C \in \mathbb{Q}(x,j)$ are $B = -F_3$ and $C = -F_4$. Due to the uniqueness of the Tate normal form, it should also be possible to write $x, j$ in terms of $B, C$, and a computation [23] confirms that. Thus $\mathbb{Q}(B,C) = \mathbb{Q}(x,j)$.

A computation [23] shows $F_2 = B^4/D$. Conjecture 1 in [5], proved by Streng [16], says that for $N > 2$, the modular units in $\mathbb{Q}(X_1(N))$ modulo $\mathbb{Q}^*$ are freely generated by $F_2, \ldots, F_{\lfloor N/2 \rfloor + 1}$.

Considering the Tate normal form over $\mathbb{Z}[B,C]$, we can look at the $k^{\text{th}}$ division polynomial $\psi_{k,E_T}(x,y) \in \mathbb{Z}[B,C][x,y]$. As in [16, Example 2.2], evaluating $\psi_{k,E_T}$ at $(0,0)$ gives:

$$\begin{aligned}
&P_1 := \psi_{1,E_T}(0,0) = 1, \qquad P_2 := \psi_{2,E_T}(0,0) = -B, \qquad P_3 := \psi_{3,E_T}(0,0) = -B^3,\\
&P_4 := \psi_{4,E_T}(0,0) = CB^5, \qquad P_5 := \psi_{5,E_T}(0,0) = -(C-B)B^8,\\
&P_6 := \psi_{6,E_T}(0,0) = -B^{12}(C^2 - B + C), \qquad P_7 := \psi_{7,E_T}(0,0) = B^{16}(C^3 - B^2 + BC).
\end{aligned}$$

A computation shows $P_k = (q_3/q_2^3)^{k^2-1}Q_k$ for $k < 5$. This must then be true for all $k$ since both sequences $P_k$ and $Q_k$ satisfy the recurrence relations (15),(16), are preserved under scaling (defined in Section 3). From Equation (19),

$$P_k = \left(\frac{q_3}{q_2^3}\right)^{k^2-1} Q_k = \left(\frac{q_3}{q_2^{8/3}}\right)^{k^2-1} \tilde{Q}_k = (\tilde{q}_3)^{k^2-1} \prod_{d|k} \tilde{q}_d = F_3^{\lfloor k^2/3 \rfloor} \prod_{3<d|k} F_d. \tag{36}$$

In particular, the multiplicative group $\langle D, -B, P_4, \ldots, P_k \rangle$ equals $\langle F_2, \ldots, F_k \rangle$. Streng defined $F_k$ for $k > 3$ to be $P_k$ with all factors of $P_j$ with $j < k$ removed, Equation (36) makes this precise.

Since $\psi_{k,E_T}(P) = 0$ if and only if $P$ has order dividing $k$, we see $F_k(P) = 0$ if and only if $P$ has exact order $k$. As mentioned in Section 3, the polynomial $F_N$ is a model for the modular curve $X_1(N)$ for $N > 3$. The Tate normal form (35) is only defined for $N > 3$, so [5] used $x, j$ coordinates to construct $F_2$ and $F_3$. Rewritten in terms of $B, C$ they are $F_2 = B^4D^{-1}$ and $F_3 = -B$. We can now state the main result of [16]

**Theorem 6.3.** *[16, Theorem 1.1], [5, Conjecture 1]. The modular units of $X^1(N)$ are given by $\mathbb{Q}^*$ times the free abelian group on $B, D, F_4, F_5, \ldots, F_{\lfloor N/2 \rfloor + 1}$, or equivalently, $F_2, \ldots, F_{\lfloor N/2 \rfloor + 1}$.*

Streng gives $P_k$ explicitly in terms of Siegel functions.

**Lemma 6.4.** *[16, Lemma 3.3] For all $k \in \mathbb{Z} - N\mathbb{Z}$*

$$P_k = \left(\frac{H_1^2H_3}{H_2^3}\right)^{k^2-1} \frac{H_k}{H_1} \quad \text{and} \quad D = \left(\frac{H_1^2H_3}{H_2^3}\right)^{12} H_1^{12}.$$

Defining $\tilde{H}_k := H_k/H_1^{k^2}$, we get

$$P_k = \left(\frac{\tilde{H}_3}{\tilde{H}_2^3}\right)^{k^2-1} \tilde{H}_k, \quad F_3 = P_2 = \frac{\tilde{H}_3^3}{\tilde{H}_2^8}, \quad \text{and} \quad F_2 = \frac{P_2^4}{D} = \tilde{H}_2^4. \tag{37}$$

Setting $t = c/N$, equation (34) gives

$$\text{uord}_c\left(\tilde{H}_k\right) = \text{uord}_c\left(H_k\right) - k^2\text{uord}_c\left(H_1\right) = \frac{1}{2}\left(\mathbb{B}_2(\{kt\}) - k^2\mathbb{B}_2(\{t\})\right). \tag{38}$$

We say that a function $f : [0, 1/2] \to \mathbb{R}$ is *k-piecewise linear* if it is continuous and $f''(t) = 0$ for all $t \notin \frac{1}{k}\mathbb{Z}$. Two $k$-piecewise linear functions coincide if and only if they have: the same initial value $f(0)$, the same initial slope $f'(0^+)$, and the same change in slope at each $t \in \frac{1}{k}\mathbb{Z}$. These three conditions hold for the right-hand sides of (38) and (39) and thus:

$$\operatorname{uord}_c(\tilde{H}_k) = \frac{1}{2}\left((k^2-k)t - \frac{1}{6}(k^2-1)\right) + k\sum_{0<i<k/2}\left(\min\left(t, \frac{i}{k}\right) - t\right). \tag{39}$$

Applying (39) to $F_2$ and $F_3$ in (37) produces the unweighted order functions $v_2(t)$ and $v_3(t)$ in Theorem 4.2.

To verify $v_k(t)$ for the remaining $k > 3$, let $\tilde{v}_k(t)$ be the unweighted order function of $\tilde{q}_k$, i.e. $\tilde{v}_k(t)$ is the right hand side of equation (29) without the factor $s_k$. So $\tilde{v}_2(t) = 0$, $\tilde{v}_3(t) = \frac{1}{3}v_3(t)$ and $\tilde{v}_k(t) = v_k(t)$ for $k > 3$. The unweighted order function for $\tilde{Q}_k = \prod_{d|k}\tilde{q}_d$ according to Theorem 4.2 and Remark 4.3 is

$$\sum_{d|k}\tilde{v}_d(t) = \sum_{d|k}\sum_{0<c'<d/2}\mathbf{n}_{c'}(d)m_{c'/d}(t) = k\sum_{0<i<k/2}m_{i/k}(t), \tag{40}$$

where $m_a(t) = \min(t, a) - 4a(1-a)t$. We also used that $k$ is the sum of $\mathbf{n}_{c'}(d)$, taken over all $0 < c' < d/2$ with $d|k$ and $c'/d = i/k$.

Applying (39) to $\tilde{Q}_k = \tilde{H}_k/\tilde{H}_2^{(k^2-1)/3}$ gives the same result. To see this, note that $m_{i/k}(t)$, which contains $\min(t, i/k)$, appears in (40) with the same coefficient $k$ as the coefficient of $\min(t, i/k)$ in (39). The terms $\frac{1}{6}(k^2-1)$ from (39) cancel out for $\tilde{H}_k/\tilde{H}_2^{(k^2-1)/3}$ but then the remaining terms $(\cdots)t$ in (39),(40) must also match by the $\int_0^{1/2} = 0$ argument from Remark 4.3. This confirms $\tilde{v}_k(t)$ and thus $v_k(t)$ for the remaining $k$. This gives a second proof for most (except the case $N|k$, see lemma 6.4) of the reformulation of Theorem 4.2 given in Remark 4.3.

## Appendix A. Proof of Primitivity

**Proposition A.1.** *The polynomial $P_k$ is primitive in $\mathbb{Z}[B, C]$.*

*Proof.* Order the monomials lexicographically with the following rule

$$B^{n_1}C^{n_2} < B^{m_1}C^{m_2} \quad \text{when} \quad n_1 < m_1 \text{ or } (n_1 = m_1 \text{ and } n_2 < m_2).$$

If $R \in \mathbb{Q}[B, C]$, let $M(R)$ denote the smallest monomial of $R$. For example, if $R = 3B^2C^5 + B^3C$, then $M(R) = 3B^2C^5$. A key property is $M(R_1R_2) = M(R_1)M(R_2)$.

Let $c_k$ denote $\lceil k/3 \rceil$. It is enough to prove that

$$M(P_k) = (-1)^{c_k}(-B)^{\lfloor k^2/3 \rfloor}C^{c_k(c_k-1)/2} \tag{41}$$

since it shows that $P_k$ has at least one coefficient equal to $\pm 1$.

We will prove (41) by induction. First, a direct verification shows that (41) holds for $k = 1, 2, 3, 4$. Suppose now that $k$ is even, and write $l = \frac{k}{2}$. Recall the recursion relation (16)

$$P_k = \frac{P_l}{P_2}\left(P_{l+2}P_{l-1}^2 - P_{l-2}P_{l+1}^2\right).$$

The smallest monomial of the first summand $P_{l+2}P_{l-1}^2$ is

$$(-1)^{c_{l+2}+2c_{l-1}}(-B)^{\lfloor (l+2)^2/3 \rfloor + 2\lfloor (l-1)^2/3 \rfloor}C^{c_{l+2}(c_{l+2}-1)/2 + c_{l-1}(c_{l-1}-1)}.$$

For the second summand $-P_{l-2}P_{l+1}^2$ it is

$$(-1)^{c_{l-2}+2c_{l+1}}(-B)^{\lfloor (l-2)^2/3 \rfloor + 2\lfloor (l+1)^2/3 \rfloor}C^{c_{l-2}(c_{l-2}-1)/2 + c_{l+1}(c_{l+1}-1)}.$$

When $l \equiv 1 \bmod 3$, the second summand has the smallest monomial, and when $l \equiv 2 \bmod 3$, the first summand has the smallest monomial. When $3 \mid l$, we have to consider the exponent of $C$. In this case, the first summand is the smallest.

In each case, verifying Equation (41) is straightforward. For example, when $l \equiv 0 \bmod 3$ we have

$$\lfloor l^2/3 \rfloor + \lfloor (l+2)^2/3 \rfloor + 2\lfloor (l-1)^2/3 \rfloor = \frac{4l^2+3}{3} = \lfloor k^2/3 \rfloor + 1,$$

and

$$c_l(c_l-1)/2 + c_{l+2}(c_{l+2}-1)/2 + c_{l-1}(c_{l-1}-1) = \frac{l}{3}\left(\frac{2l-3}{3}\right) = \frac{c_k}{2}(c_k-1).$$

Now suppose $k$ is odd and write $k = 2l+1$. Recall the recursion relation (15)

$$P_k = P_{l+2}P_l^3 - P_{l-1}P_{l+1}^3.$$

For the first summand, the smallest monomial is

$$(-1)^{c_{l+2}+3c_l}(-B)^{\lfloor (l+2)^2/3 \rfloor + 3\lfloor l^2/3 \rfloor} C^{c_{l+2}(c_{l+2}-1)/2 + 3c_l(c_l-1)/2},$$

and for the second summand it is

$$(-1)^{c_{l-1}+3c_{l+1}}(-B)^{\lfloor (l-1)^2/3 \rfloor + 3\lfloor (l+1)^2/3 \rfloor} C^{c_{l-1}(c_{l-1}-1)/2 + 3c_{l+1}(c_{l+1}-1)/2}.$$

When $l \equiv 0 \bmod 3$, the first summand has the smaller monomial; when $l \equiv 1 \bmod 3$, considering the exponent of $C$ shows the second summand has the smaller monomial; and when $l \equiv 2 \bmod 3$, the first summand has the smaller monomial.

Verifying Equation (41) is again straightforward for each case. For example, when $l \equiv 1 \bmod 3$

$$\lfloor (l-1)^2/3 \rfloor + 3\lfloor (l+1)^2/3 \rfloor = \frac{4l^2+4l+1}{3} = \lfloor k^2/3 \rfloor,$$

and

$$c_{l-1}(c_{l-1}-1)/2 + 3c_{l+1}(c_{l+1}-1)/2 = \frac{4l^2-2l-2}{18} = \frac{c_k}{2}(c_k-1).$$

Repeating these computations for the remaining cases proves the proposition. □

Recall that $-B = F_3 = \tilde{q}_3^3$ and $-C = F_4$. From (36) we find that $\tilde{Q}_{k\setminus 3}$ from Section 3 is $P_k/(-B)^{\lfloor k^2/3 \rfloor}$ which is primitive in $\mathbb{Z}[B,C] = \mathbb{Z}[F_3,F_4]$ by equation (41).

## References


[1] H. Baaziz. Equations for the modular curve $X_1(N)$ and models of elliptic curves with torsion points. *Math. Comp.*, 79(272):2371–2386, 2010. ISSN 0025-5718. URL https://doi.org/10.1090/S0025-5718-10-02332-X.

[2] A. Bourdon, Ö. Ejder, Y. Liu, F. Odumodu, and B. Viray. On the level of modular curves that give rise to isolated j-invariants. *Advances in Mathematics*, 357:106824, 2019.

[3] P. Carlucci. Cuspidal divisor class groups of non-split Cartan modular curves. *Acta Arith.*, 187(4):301–327, 2019. ISSN 0065-1036. URL https://doi.org/10.4064/aa8516-6-2018.

[4] Y.-H. Chen. Cuspidal $\mathbb{Q}$-rational torsion subgroup of $J(\Gamma)$ of level $P$. *Taiwanese J. Math.*, 15 (3):1305–1323, 2011. ISSN 1027-5487. URL https://doi.org/10.11650/twjm/1500406301.

[5] M. Derickx and M. van Hoeij. Gonality of the modular curve $X_1(N)$. *J. Algebra*, 417:52–71, 2014. ISSN 0021-8693. URL https://doi.org/10.1016/j.jalgebra.2014.06.026.

[6] M. Derickx, M. van Hoeij, and J. Zeng. Computing Galois representations and equations for modular curves $X_H(l)$. *ArXiv e-prints*, Mar. 2014.

[7] F. Diamond and J. Shurman. *A first course in modular forms*, volume 228 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 2005. ISBN 0-387-23229-X.

[8] DLMF. *NIST Digital Library of Mathematical Functions*. http://dlmf.nist.gov/, Release 1.0.26 of 2020-03-15. URL http://dlmf.nist.gov/. F. W. J. Olver, A. B. Olde Daalhuis, D. W. Lozier, B. I. Schneider, R. F. Boisvert, C. W. Clark, B. R. Miller, B. V. Saunders, H. S. Cohl, and M. A. McClain, eds.

[9] A. Hurwitz and R. Courant. *Vorlesungen über allgemeine Funktionentheorie und elliptische Funktionen*. Interscience Publishers, Inc., New York, 1944.

[10] D. S. Kubert and S. Lang. *Modular units*, volume 244 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Science]*. Springer-Verlag, New York-Berlin, 1981. ISBN 0-387-90517-0.

[11] B. Mazur. Some comments on elliptic curves over general number fields and Brill-Noether modular varieties. 2014. URL http://people.math.harvard.edu/~mazur/papers/For.Maine.0718.2014(3).pdf.

[12] F. Najman. Torsion of rational elliptic curves over cubic fields and sporadic points on $X_1(n)$. *Math. Res. Lett.*, 23(1):245–272, 2016. ISSN 1073-2780. URL https://doi.org/10.4310/MRL.2016.v23.n1.a12.

[13] J.-P. Serre. *Local fields*, volume 67 of *Graduate Texts in Mathematics*. Springer-Verlag, New York-Berlin, 1979. ISBN 0-387-90424-7. Translated from the French by Marvin Jay Greenberg.

[14] J. H. Silverman. *The arithmetic of elliptic curves*, volume 106 of *Graduate Texts in Mathematics*. Springer, Dordrecht, second edition, 2009. ISBN 978-0-387-09493-9. URL https://doi.org/10.1007/978-0-387-09494-6.

[15] H. Smith. Ramification in the Division Fields of Elliptic Curves and an Application to Sporadic Points on Modular Curves. *arXiv e-prints*, art. arXiv:1810.04809, Oct. 2018.

[16] M. Streng. Generators of the group of modular units for $\Gamma^1(N)$ over $\mathbb{Q}$. *arXiv e-prints*, art. arXiv:1503.08127, Mar 2015.

[17] A. Sutherland. Torsion subgroups of elliptic curves over number fields. 2012. URL https://math.mit.edu/~drew/MazursTheoremSubsequentResults.pdf.

[18] A. Sutherland and M. van Hoeij. Defining equations for $X_1(N)$. 2014. URL https://math.mit.edu/~drew/X1_optcurves.html.

[19] A. V. Sutherland. Constructing elliptic curves over finite fields with prescribed torsion. *Mathematics of Computation*, 81(278):1131–1147, 2012. ISSN 00255718, 10886842. URL http://www.jstor.org/stable/23267989.

[20] A. V. Sutherland and D. Zywina. Modular curves of prime-power level with infinitely many rational points. *Algebra Number Theory*, 11(5):1199–1229, 2017. ISSN 1937-0652. URL https://doi.org/10.2140/ant.2017.11.1199.

[21] T. Takagi. Cuspidal class number formula for the modular curves $X_1(p)$. *J. Algebra*, 151(2):348–374, 1992. ISSN 0021-8693. URL https://doi.org/10.1016/0021-8693(92)90119-7.

[22] M. van Hoeij. Minformula (in file: cusp divisors program). 2013. URL www.math.fsu.edu/~hoeij/files/X1N.

[23] M. van Hoeij and H. Smith. Computations for gonality bound. 2020. URL http://math.colorado.edu/~hwsmith/code.html.

[24] Y. Yang. Modular units and cuspidal divisor class groups of $X_1(N)$. *J. Algebra*, 322(2):514–553, 2009. ISSN 0021-8693. URL https://doi.org/10.1016/j.jalgebra.2009.04.012.

[25] Y. Yang and J.-D. Yu. Structure of the cuspidal rational torsion subgroup of $J_1(p^n)$. *J. Lond. Math. Soc. (2)*, 82(1):203–228, 2010. ISSN 0024-6107. URL https://doi.org/10.1112/jlms/jdq013.

[26] H. Yoo. The rational cuspidal divisor class group of $X_0(N)$. *arXiv e-prints*, art. arXiv:1908.06411, Aug 2019.

[27] J. Yu. A cuspidal class number formula for the modular curves $X_1(N)$. *Math. Ann.*, 252(3):197–216, 1980. ISSN 0025-5831. URL https://doi.org/10.1007/BF01420083.

## Appendix B. Notation

| Notation | Brief Definition | References |
|---|---|---|
| $\mathbf{p}$ | a Puiseux expansion above $s=0$ | page 2 |
| $v_{\mathbf{p}}$ | discrete valuation associated to $\mathbf{p}$ | page 2 |
| $\mathbf{e_p}$ | smallest $e$ with $\mathbf{p} \in \mathbb{C}((s^{1/e}))$ $\quad (= v_{\mathbf{p}}(s))$ | page 2 |
| $k_{\mathbf{p}}$ | residue field associated to $\mathbf{p}$ | page 2 |
| $\mathbf{f_p}$ | $[k_{\mathbf{p}} : \mathbb{Q}]$ | page 2 |
| $\mathbf{n_p}$ | $\mathbf{e_p} \cdot \mathbf{f_p} = [\mathbb{Q}((s))[\mathbf{p}] : \mathbb{Q}((s))]$ | page 2 |
| $l_s(\mathbf{p})$ | dominant term of $\mathbf{p}$ | page 2 |
| $E_\epsilon$ | $y^2 = x(x-\epsilon)(x-1)$, with $0 < \epsilon \ll 1$ | page 3 |
| $\omega_1, \omega_2$ | periods of $E_\epsilon$ | page 3 |
| $W$ | $(\mathbb{C}/\Lambda)/\pm$ where $\Lambda = \mathbb{Z}\omega_1 + \mathbb{Z}\omega_2$ | page 3 |
| $W(N)$ | elements of order $N$ in $W$ | page 3 |
| $C(N)$ | the Cartan, $\{0, 1, \ldots, \lfloor N/2 \rfloor\}$ | page 3 |
| $W_i(N)$ | $i^{\text{th}}$ Galois orbit $\subset W(N)$ $\quad (i \in C(N))$ | page 3 |
| $C_i(N)$ | $i^{\text{th}}$ Galois orbit $\subset \{x(P) \,\vert\, P \text{ order } N\}$ | page 3 |
| $\mathbf{n}_i$ | $\lvert W_i(N)\rvert = \lvert C_i(N)\rvert = \mathbf{e}_i \cdot \mathbf{f}_i$ | page 3 |
| $E$ | $y^2 = x^3 + ax + b$, $a = -3j_0$, $b = -2j_0$ | pages 4, 5 |
| $j_0$, $j$ | $j_0 = j/(j-1728)$, $\;j = j$-invariant of $E$ | page 5 |
| $s$ | $s = 1/j$, $\;\text{roots}(s) = \{\text{cusps of } X_1(N)\}$ | page 6 |
| $E[=k]$ | $\{\text{points on } E \text{ of exact order } k\}$ | page 4 |
| $m_k$ | $m_k = \#$ points of exact order $k$ | page 5 |
| $Q_k$ | division polynomial of $E$ | pages 4, 5 |
| $q_k$ | $\text{roots}(q_k) = \{x(P) \mid P \text{ has order } k\}$, $\quad Q_k = \prod_{d\vert k} q_k$ | pages 4, 5 |
| $\tilde{Q}_k$ | rescaling of $Q_k$ to make it unique, $\;\tilde{Q}_k = Q_k/Q_2^{(k^2-1)/3}$ | page 5 |
| $\tilde{q}_k$ | rescaling of $q_k$, $\;\tilde{Q}_k = \prod_{d\vert k} \tilde{q}_d$ | page 5 |
| $\{F_k\}$ | basis of modular units, $F_k \in \mathbb{Q}(x, j_0) = \mathbb{Q}(x, j) = \mathbb{Q}(x, s)$ | [5], [16] |
| $F_2, F_3$ | $F_2 = q_2^4/\left(1728 j_0^2 (j_0 - 1)\right)$, $\;F_3 = \tilde{q}_3^3 = q_3^3/q_2^8$ | pages 6, 5 |
| $F_k$, $k > 3$ | $F_k = \tilde{q}_k = q_k/q_2^{m_k/3}$ | page 5 |
| $v_k(t)$ | piecewise linear function, gives $\text{div}\,(F_k)$ on any $X_1(N)$ | page 7 |

Table 2. Summary of notation for Sections 2-4

| Notation | Brief Definition | References |
|---|---|---|
| $v(t)$ | $v(t) = v_7(t) - v_8(t)$, gives div $(F_7/F_8)$ | page 8 |
| $m(t)$ | $\max(v(t), 0)$ | page 8 |
| $B_0(N)$ | degree of $F_7/F_8$ | page 8 |
| $B_1(N)$ | upper bound for $B_0(N)$ | page 8 |
| $I_i$ | intervals where $m(t)$ is linear | page 9 |
| $m_i(t)$ | linear function equal to $m(t)$ restricted to $I_i$ | page 9 |
| $\mathbf{e}_{\left[\begin{smallmatrix}a\\c\end{smallmatrix}\right]}(N)$ | width of cusp on $X_1(N)$ with representative $\left[\begin{smallmatrix}a\\c\end{smallmatrix}\right]$ | page 11 |
| $g_{(a_1,a_2)}$ | Siegel function associated to $(a_1, a_2) \in \mathbb{Q}^2 - \mathbb{Z}^2$ | page 11 |
| $\mathbb{B}_2$ | second Bernoulli polynomial, $x^2 - x + \frac{1}{6}$ | page 11 |
| $\{\bullet\}$ | fractional part, $\bullet - \lfloor \bullet \rfloor$ | page 12 |
| $[\bullet]_1$ | first entry in a vector | page 12 |
| $H_k$ | Siegel function on $X_1(N)$, $g_{(0,k/N)}$, $k \in \mathbb{Z} - N\mathbb{Z}$ | page 12 |
| uord$_c(H_k)$ | "unweighted order" of $H_k$ at $c \in C(N)$ | page 12 |
| $E_T$ | Tate normal form: $Y^2 + (1-C)XY - BY = X^3 - BX^2$ | page 12 |
| $B, C$ | $E \rightsquigarrow E_T$ gives $B = -F_3$, $C = -F_4$, $\mathbb{Q}(B, C) = \mathbb{Q}(x, j)$ | pages 12, 13 |
| $\Psi_{k,E_T}$ | $k^{\text{th}}$ division polynomial of $E_T$ | page 13 |
| $P_k$ | $P_k = \Psi_{k,E_T}(0,0) \in \mathbb{Z}[B, C]$, factors: $F_3$ and $\{F_d : 2 \neq d \mid k\}$ | page 13 |
| $\tilde{H}_k$ | scaled Siegel function, $H_k/H_1^{k^2}$ | page 13 |

TABLE 3. Summary of notation for Sections 5-6

DEPT. OF MATHEMATICS, FLORIDA STATE UNIVERSITY, TALLAHASSEE, FL 32306, USA
*Email address*: hoeij@math.fsu.edu

DEPT. OF MATHEMATICS, UNIVERSITY OF COLORADO, CAMPUS BOX 395, BOULDER, COLORADO 80309-0395
*Email address*: hanson.smith@colorado.edu or hansonsmith101@gmail.com