Annales Henri Lebesgue
6 (2023) 95-116

ANNALES
HENRI LEBESGUE

MARCO STRENG

GENERATORS OF THE GROUP OF MODULAR UNITS FOR $\Gamma^1(N)$ OVER THE RATIONALS

GÉNÉRATEURS DU GROUPE DES UNITÉS MODULAIRES POUR $\Gamma^1(N)$ SUR LES RATIONNELS

Abstract. — We give two explicit sets of generators of the group of invertible regular functions over $\mathbf{Q}$ on the modular curve $Y^1(N)$.

The first set of generators is very surprising. It is essentially the set of defining equations of $Y^1(k)$ for $k \leqslant N/2$ when all these modular curves are simultaneously embedded into the affine plane, and this proves a conjecture of Derickx and Van Hoeij [DvH14]. This set of generators is an elliptic divisibility sequence in the sense that it satisfies the same recurrence relation as the elliptic division polynomials.

The second set of generators is explicit in terms of classical analytic functions known as Siegel functions. This is both a generalization and a converse of a result of Yang [Yan04, Yan09].

Résumé. — Nous donnons deux ensembles explicites de générateurs du groupe des fonctions régulières inversibles à coefficients rationnels sur la courbe modulaire $Y^1(N)$.

Le premier ensemble de générateurs est très surprenant. C'est essentiellement l'ensemble des équations qui définissent $Y^1(k)$ pour $k \leqslant N/2$ quand toutes ces courbes modulaires sont plongées simultanément dans le plan affine, ce qui prouve une conjecture de Derickx et Van

Keywords: modular units, modular functions, elliptic divisibility sequences, divsion polynomials.
2020 *Mathematics Subject Classification:* 11G16, 11B37, 11B39, 11F03, 14H52.
DOI: `https://doi.org/10.5802/ahl.160`

(*) The author would like to thank Peter Bruin, Maarten Derickx, Pınar Kılıçer and Mark van Hoeij for helpful discussions and the anonymous referee for helpful suggestions for improving the exposition.

Hoeij [DvH14]. Cet ensemble de générateurs est une suite de divisibilité elliptique dans le sens où il satisfait la même relation de récurrence que les polynômes de division elliptiques.

Le second ensemble de générateurs est explicite en termes de fonctions analytiques classiques appelées fonctions de Siegel. C'est à la fois une généralisation et une réciproque d'un résultat de Yang [Yan04, Yan09].

1. Introduction

Let $N \geqslant 1$ be an integer. The modular curve $Y^1(N)$ is a smooth, affine, geometrically irreducible algebraic curve over $\mathbf{Q}$, often also denoted by $Y_1(N)$. It has the following property: For every field K of characteristic zero, if $N \geqslant 4$ or K is algebraically closed, then we have

$$Y^1(N)(K) = \\ \{(E,P) : E \text{ is an elliptic curve over } K \text{ and } P \in E(K) \text{ has order } N\} \,/ \cong .$$

Here "=" denotes a functorial Galois-equivariant bijection, which we use to identify the left and right hand side; and we write $(E_1,P_1) \cong (E_2,P_2)$ when there is an isomorphism $\phi : E_1 \to E_2$ with $\phi(P_1) = P_2$.

Our object of study is the group of *modular units on* $Y^1(N)$, that is, the unit group $\mathcal{O}(Y^1(N))^*$ of the ring $\mathcal{O}(Y^1(N))$ of regular functions over $\mathbf{Q}$ on $Y^1(N)$. The curve $Y^1(N)$ has a smooth compactification $X^1(N)$, and the group $\mathcal{O}(Y^1(N))^*$ equals the group of meromorphic functions over $\mathbf{Q}$ on $X^1(N)$ with divisor supported on the set $X^1(N) \setminus Y^1(N)$ of *cusps*.

The *Tate normal form* (Section 2.1) gives an embedding $Y^1(N) \hookrightarrow \mathbf{A}^2$ for every $N \geqslant 4$, with the point $(B,C) \in \mathbf{A}^2$ corresponding to the curve

$$E : Y^2 + (1-C)XY - BY = X^3 - BX^2 \quad \text{and point} \quad P = (0,0). \tag{1.1}$$

Our first main result is as follows.

Theorem 1.1 (Conjecture 1 of Derickx and Van Hoeij [DvH14]). — *For all $k \geqslant 4$, let $F_k \in \mathbf{Q}[B,C]$ be the defining polynomial of $Y^1(k)$ inside $\mathbf{A}^2$. Then for all $N \geqslant 4$, the group $\mathcal{O}(Y^1(N))^*$ is $\mathbf{Q}^*$ times the free abelian group on B, D, F_4, F_5, …, $F_{\lfloor N/2 \rfloor + 1}$, where $D \in \mathbf{Q}[B,C]$ is the discriminant of* (1.1).

The functions F_k are given in terms of a recurrence relation, which we recall in Remark 2.9.

The theorem is interesting for a number of reasons. First of all, Derickx and Van Hoeij [DvH14] already used the functions in the theorem in order to compute the gonality of $Y^1(N)$ for all positive integers $N \leqslant 40$ and to give an upper bound on the gonality for $N \leqslant 250$. Our theorem helps explain why their method was successful.

Moreover, they found that the gonality is often achieved by functions from this set of generators. In particular, these functions are "small" functions in some sense, which we therefore hope are suitable for finding "small" models of modular curves $Y^1(N)$. Finding such small models directly in terms of another algebraic model has the advantage that no approximate numerics (such as floating point numbers or truncated power series) are needed in producing these models, as would be the case when using theta functions or Siegel functions directly or using modular forms.

Thirdly, as we will see in Section 2.1, the functions F_k are the primitive divisors of an *elliptic divisibility sequence (EDS)* $P_1, P_2, P_3, \ldots$ over the ring $\mathbf{Q}[B, C]$, which is in a way the *universal* EDS as it comes from the Tate normal form. In line with Ingram–Mahé–Silverman–Stange–Streng [IMS$^+$12] and Naskręcki [Nas16], all but finitely many terms P_k have a primitive divisor. In fact, we prove that all terms P_k with $k > 3$ have a *unique* primitive divisor F_k.

Finally, an explicit basis of the unit group could be useful for computing cuspidal divisor class groups similarly to [Yan09].

The proof proceeds by first linking the functions P_k to classical analytic Siegel functions, and then observing how a proof of Kubert and Lang for $Y(N)$ can be much simplified and strengthened when applying it to $Y^1(N)$. Our proof can be read without knowing the proof of Kubert and Lang, and can be seen as an introduction into their methods due to the disappearance of complications that arise in their proof.

We prove the main theorem using modular forms over $\mathbf{C}$. Let $\mathbf{H} \subset \mathbf{C}$ be the standard upper half plane, write $\mathbf{H}^* = \mathbf{H} \cup \mathbf{P}^1(\mathbf{Q}) \subset \mathbf{P}^1(\mathbf{C})$, and write

$$\Gamma^1(N) = \left\{ \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbf{Z}) : b \equiv 0, a \equiv d \equiv 1 \bmod N \right\}. \tag{1.2}$$

Recall the natural complex analytic isomorphism

$$\begin{aligned} \Gamma^1(N)\backslash\mathbf{H} &\longrightarrow Y^1(N)(\mathbf{C}) \\ \tau &\longmapsto (\mathbf{C}/\Lambda_\tau, \tau/N \bmod \Lambda_\tau), \end{aligned} \tag{1.3}$$

where $\Lambda_\tau = \tau\mathbf{Z} + \mathbf{Z}$. (If the reader is used to another parametrization, see Remarks 2.14 and 2.15 below.) The functions on $X^1(N)$ defined over $\mathbf{Q}$ correspond exactly to the meromorphic functions on $\Gamma^1(N)\backslash\mathbf{H}^*$ whose q-expansions at ∞ are rational, that is, in $\mathbf{Q}((q^{1/N}))$ with $q^a = \exp(2\pi i a\tau)$.

The group $\mathcal{O}(Y^1(N))^*$ therefore equals the group of meromorphic functions on $\Gamma^1(N)\backslash\mathbf{H}^*$ with rational q-expansion and divisor supported on $\mathbf{P}^1(\mathbf{Q})$.

Our second main result is as follows. For positive integers $k \leqslant N/2$, let H_k be the *Siegel function* given by (see also (2.7))

$$H_k(\tau) = i q^{\frac{1}{2}\left((k/N)^2 - k/N + \frac{1}{6}\right)} \left(1 - q^{k/N}\right) \prod_{n=1}^{\infty} \left(1 - q^{n+k/N}\right)\left(1 - q^{n-k/N}\right). \tag{1.4}$$

Theorem 1.2. — *Let*

$$S = \left\{ \prod_{k=1}^{\lfloor N/2 \rfloor} H_k^{e(k)} : \begin{array}{rcl} \forall_k \ e(k) & \in & \mathbf{Z}, \\ \sum_k e(k) & \in & 12\mathbf{Z}, \\ \sum_k k^2 e(k) & \in & \gcd(N,2)N\mathbf{Z} \end{array} \right\}.$$

Then S is free abelian of rank $\lfloor N/2 \rfloor$ and satisfies $\mathcal{O}(Y^1(N))^* = \mathbf{Q}^* \cdot S$.

Remark 1.3. — Kubert and Lang have results similar to Theorem 1.2 for the curve $Y(N)$ ([KL77, Theorems 1 and 2]; alternatively Theorems 1.1 and 1.2 in [KL81, Chapter 4]). Indeed, the results of loc. cit. can be combined into an analogue of our Theorem 1.2 for $\mathcal{O}(Y(N)_{\mathbf{C}})$, but for most N their result is only ‘up to power of two index'. For details, see Theorem 1.3 in [KL81, Chapter 4] and the text below it. See also Kubert [Kub81].

Remark 1.4. — Theorem 1.2 gives both a strengthening and a converse of Yang [Yan04, Corollary 3]. Indeed, loc. cit. gives the inclusion $S' \subset \mathbf{Q}(Y^1(N))$ if $S' \subset S$ is defined by the additional hypotheses $\sum_k ke(k) \in 2\mathbf{Z}$ and $\sum_k k^2e(k) \in 2N\mathbf{Z}$.

[Yan09, Theorems 1–5] give the analogue of Theorem 1.2 if one restricts to the functions with divisors supported on cusps $\frac{x}{y}$ for $\gcd(x, N) = 1$.

The dictionary between our functions and the functions of [Yan04, Yan09] is given in Remark 2.15 below. And in fact, with the conventions of Remark 2.15 the functions of [Yan09] are those with divisors supported on cusps $\frac{x}{Ny}$ for $\gcd(x, N) = 1$.

1.1. Overview and methods

Our proof consists of two parts. The first part is Section 3, which relates the functions of Theorems 1.1 and 1.2 via explicit expressions in both directions. We use formulas and techniques from the theory of elliptic divisibility sequences to relate division polynomials with the Weierstrass sigma function.

The second part is Section 4, in which we show that our functions indeed generate the full group. As in Kubert-Lang [KL77], one of the key ideas is to use the fact that every modular form with a rational q-expansion can be scaled to have an integer q-expansion. Together with Gauss' Lemma for power series with bounded denominators, this will show that if g^l is in our group for a modular function g, then so is g itself. We show that this idea works even better in the case of $\Gamma^1(N)$ over $\mathbf{Q}$ than in the case of [KL77], yielding results that are less general, but stronger, simpler and more elegant than the results of [KL77]. A detailed overview of this part of the proof is given at the beginning of Section 4.

Before we start the proof, Section 2 gives precise definitions of the functions appearing in Theorems 1.1 and 1.2.

After the proof is finished, we give two results that we get for free from our methods. In Section 5.1, we give generators of the *ring* $\mathcal{O}(Y^1(N))$ instead of generators of the unit group, and in Section 5.2, we express the generators of the unit group in terms of theta functions.

2. The functions appearing in the main results

2.1. The Tate normal form

Let E be an elliptic curve over a field K and $P \in E(K)$ a point of order > 3 (possibly non-torsion).

Lemma 2.1 (Tate normal form). — *Every pair (E, P) as above is isomorphic to a unique pair of the form*

$$E : Y^2 + (1 - C)XY - BY = X^3 - BX^2, \quad P = (0, 0) \tag{2.1}$$

for $B, C \in K$ with

$$D := B^3 \cdot \left(C^4 - 8BC^2 - 3C^3 + 16B^2 - 20BC + 3C^2 + B - C\right) \neq 0.$$

Conversely, for every pair $B, C \in K$ with $D \neq 0$, equation (2.1) *gives a pair* (E, P).

Proof. — Given (E, P), start with a general Weierstrass equation

$$Y^2 + A_1XY + A_3Y = X^3 + A_2X^2 + A_4X + A_6. \tag{2.2}$$

As P does not have order 1, it is affine, and we translate P to $(0,0)$ yielding $A_6 = 0$. As P does not have order 2, we have $A_3 \neq 0$, and we add $(A_4/A_3)X$ to Y to get $A_4 = 0$. As P does not have order 3, we get $A_2 \neq 0$, and we scale X and Y to get $A_2 = A_3$. Then we define $C = 1 - A_1$ and $B = -A_2 = -A_3$. This uses up all freedom for changing Weierstrass equations [Sil86, III.3.1(b)], so this form is uniquely defined. The quantity D is the discriminant of E, which is non-zero.

Conversely, if D is non-zero, then (E, P) defines an elliptic curve and a point on it, where the point does not have order 1, 2 or 3. □

For any elliptic curve E given by a general Weierstrass equation $y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6$ and any $k \in \mathbf{Z}$, the *k-division polynomial* ψ_k is given by

$$\psi_0 = 0, \qquad \psi_2 = 2y + a_1x + a_3,$$
$$\psi_k = t \cdot \prod_{Q \,\in\, (E[k]\setminus E[2])/\pm} (x - x(Q)), \qquad \text{where} \quad t = \begin{cases} k & \text{if } 2 \nmid k, \\ \frac{k}{2} \cdot \psi_2 & \text{if } 2 \mid k. \end{cases}$$

For any point P on E, we have $kP = 0$ if and only if $\psi_k(P) = 0$.

Let $P_k \in \mathbf{Z}[B, C]$ be the k-division polynomial ψ_k of the elliptic curve (2.1) evaluated in the point $P = (0,0)$. In particular, if $k \geqslant 4$ and (E, P) corresponds to $(B, C) \in K^2$ with $D \neq 0$, then P has order dividing k if and only if $P_k(B, C) = 0$.

Example 2.2. — For positive integers k, we compute the k-division polynomial with the SageMath [SageMath14] command

```
E.division_polynomial(k, two_torsion_multiplicity=1)
```

and obtain the following list.

$$\begin{aligned} P_1 &= 1 & P_5 &= -(C - B) \cdot B^8 \\ P_2 &= -B & P_6 &= -B^{12} \cdot \left(C^2 - B + C\right) \\ P_3 &= -B^3 & P_7 &= B^{16} \cdot \left(C^3 - B^2 + BC\right) \\ P_4 &= C \cdot B^5 & P_8 &= C \cdot B^{21} \cdot \left(BC^2 - 2B^2 + 3BC - C^2\right) \end{aligned}$$

For $k \geqslant 4$, let $F_k \in \mathbf{Q}[B, C]$ be P_k with all factors in common with D and P_d for $d < k$ removed (well-defined up to $\mathbf{Q}^*$). Following [DvH14], we let $F_3 = B \in \mathbf{Z}[B, C]$ and $F_2 = B^4/D \in \mathbf{Q}(B, C)$.

Example 2.3. — The rational functions F_k are as follows.

$$\begin{aligned}
F_2 &= B \cdot \left(C^4 - 8BC^2 - 3C^3 + 16B^2 - 20BC + 3C^2 + B - C\right)^{-1}\\
F_3 &= B\\
F_4 &= C\\
F_5 &= C - B\\
F_6 &= C^2 - B + C\\
F_7 &= C^3 - B^2 + BC\\
F_8 &= BC^2 - 2B^2 + 3BC - C^2
\end{aligned}$$

For $N \geqslant 4$, the point $P = (0, 0)$ on E is of order N if and only if $F_N = 0$. In particular, we get the following known model of $Y^1(N)$.

Proposition 2.4. — *Given $N \geqslant 4$, let $R = \mathbf{Q}[B, C, D^{-1}] \subset \mathbf{Q}(B, C)$ and let $Y = \operatorname{Spec}(R/F_N) \subset \operatorname{Spec}(R) \subset \mathbf{A}^2$. In other words, let Y be the curve over $\mathbf{Q}$ in the affine B, C-plane given by*

$$Y : F_N = 0, D \neq 0.$$

Then for all fields K of characteristic 0, we have $Y^1(N)(K) = Y(K)$. □

In fact, with a more careful analysis of the Tate normal form and division polynomials, one would get the following much stronger result, which we do not need for our main results, but which we give for completeness.

Proposition 2.5 (Jin [Jin13, Corollary 45]). — *Let $R' = \mathbf{Z}[B, C, D^{-1}, 1/N] \subset R$. The scheme $\operatorname{Spec}(R'/F_N)$ represents the "naive" $\Gamma_1(N)$ moduli problem of [Jin13] over $\mathbf{Z}[1/N]$.*

For every $k \geqslant 2$, the element $F_k \in \mathbf{Q}[B, C]$ now coincides with F_k of Derickx and Van Hoeij [DvH14]. It is irreducible in $\overline{\mathbf{Q}}[B, C]$ for $k \geqslant 4$ because the curve $Y^1(k)_{\mathbf{C}}$ is irreducible.

By taking B, C, D, F_k and P_k modulo F_N, we get modular functions b, c, d, f_k and p_k on $Y^1(N)$ for all $k, N \in \mathbf{Z}$ with $k \geqslant 2$, $N \geqslant 4$, and $N \nmid k$. Derickx and Van Hoeij show ([DvH14, Section 2]) that they are *modular units*, that is, functions with divisors supported at the cusps. Let $\mathcal{O}(Y^1(N))^* \subset \mathbf{Q}(X^1(N))^*$ be the group of all modular units. Our main result is the following.

Theorem 2.6 (Rephrasing of Theorem 1.1 above, [DvH14, Conjecture 1]). — *The group $\mathcal{O}(Y^1(N))^*/\mathbf{Q}^*$ is the free abelian group on $f_2, f_3, f_4, \ldots, f_{\lfloor N/2\rfloor+1}$.*

The first, small step of the proof is to rewrite the theorem in terms of p_k using the following lemma.

Lemma 2.7. — *For all $k \geqslant 3$, we have*

$$\langle F_2, F_3, \ldots, F_k\rangle \cdot \mathbf{Q}^* = \langle B, D, P_4, P_5, \ldots, P_k\rangle \cdot \mathbf{Q}^*.$$

Proof. — Let L_k be the left hand side and R_k the right. We prove by induction on k that we have $L_k = R_k$ and that all irreducible factors of both D and P_d for $d \leqslant k$ are elements of L_k.

For $k = 3$, we have $F_3 = B$ and $F_2 = B^4/D$ by definition, hence also $D = F_3^4/F_2$. As B and $D/B^3 = F_2^{-1}F_3$ are irreducible, the induction hypothesis follows for $k = 3$.

Suppose now that the induction hypothesis holds for $k = n - 1$. By definition F_n is P_n except for factors in common with D and P_d for $d < k$, but by the induction hypothesis all such factors are in $L_{n-1} = R_{n-1}$. In particular, we get $L_n = R_n$. The polynomial F_n is irreducible as mentioned below Proposition 2.5, hence the induction hypothesis also holds for $k = n$. □

By Lemma 2.7, we find that Theorem 2.6 is equivalent to the following.

Theorem 2.8. — *The group $\mathcal{O}(Y^1(N))^*/\mathbf{Q}^*$ is the free abelian group on b, d, p_4, $p_5, \ldots, p_{\lfloor N/2\rfloor+1}$.*

Remark 2.9. — The division polynomials ψ_k, and hence the polynomials P_k and the functions p_k, satisfy the following recurrence relation. For all $m, n, k \in \mathbf{Z}$, we have

$$\psi_{m+n}\psi_{m-n}\psi_k^2 = \psi_{m+k}\psi_{m-k}\psi_n^2 - \psi_{n+k}\psi_{n-k}\psi_m^2.$$

Taking $(k, m, n) = (1, l+1, l)$ or $(1, l+1, l-1)$, we get

$$\begin{aligned} \psi_{2l+1} &= \psi_{l+2}\psi_l^3 - \psi_{l+1}^3\psi_{l-1}, \\ \psi_{2l} &= \psi_2^{-1}\psi_l\left(\psi_{l+2}\psi_{l-1}^2 - \psi_{l-2}\psi_{l+1}^2\right), \end{aligned}$$

which gives p_k for all $k \geqslant 5$ starting from the initial terms p_1, p_2, p_3, p_4 of Example 2.2.

Example 2.10. — The curve $X^1(5)$ is defined by $0 = F_5 = C - B$, that is, by $B = C$. We compute

$$\begin{array}{rlcrl} p_1 = & 1 & \qquad & p_6 = & -c^{14} \\ p_2 = & -c & & p_7 = & c^{19} \\ p_3 = & -c^3 & & p_8 = & c^{25} \\ p_4 = & c^6 & & p_9 = & -c^{32} \\ p_5 = & 0 & & p_{10} = & 0 \\ d = & c^5 \cdot (c^2 - 11c - 1), & & & \end{array}$$

which, except for p_5 and p_{10}, all lie in the group generated by $b = c$ and d.

Example 2.11. — The curve $X^1(6)$ is defined by $0 = F_6 = C^2 - B + C$, that is, by $B = C(C+1)$. We compute

$$\begin{array}{rlcrl} p_1 = & 1 & \qquad & p_6 = & 0 \\ p_2 = & -c \cdot (c+1) & & p_7 = & -c^{20} \cdot (c+1)^{16} \\ p_3 = & -c^3 \cdot (c+1)^3 & & p_8 = & -c^{26} \cdot (c+1)^{21} \\ p_4 = & c^6 \cdot (c+1)^5 & & p_9 = & c^{33} \cdot (c+1)^{27} \\ p_5 = & c^{10} \cdot (c+1)^8 & & p_{10} = & c^{41} \cdot (c+1)^{33} \\ d = & c^6 \cdot (c+1)^3 \cdot (9c+1), & & & \end{array}$$

which indeed all, except for p_6, lie in the group generated by $b = c(c+1)$, d and p_4.

2.2. Siegel functions

This section defines the *Siegel functions* of Theorem 1.2 and recalls their transformation properties and q-expansions. Our main reference for this section is Fricke [Fri11]. We start by recalling the well-known Weierstrass sigma function and Dedekind eta function.

2.2.1. Lattices, sigma and eta

By a *lattice*, we will always mean a discrete subgroup $\Lambda \subset \mathbf{C}$ of rank 2. For example, for $\tau \in \mathbf{H}$, we have a lattice $\Lambda_\tau = \tau\mathbf{Z} + \mathbf{Z}$. For $\omega_1, \omega_2 \in \mathbf{C}$ with $\tau = \omega_1/\omega_2 \in \mathbf{H}$, we have a lattice $\omega_1\mathbf{Z} + \omega_2\mathbf{Z} = \omega_2\Lambda_\tau$.

We define the *Weierstrass sigma function* by ([Fri11, (1) on p. 258])

$$\sigma(z, \Lambda) = z \prod_{\substack{w \in \Lambda \\ \omega \neq 0}} \left(1 - \frac{z}{w}\right) \exp\left(\frac{z}{w} + \frac{1}{2}\left(\frac{z}{w}\right)^2\right)$$

for all $z \in \mathbf{C}$ and all lattices $\Lambda \subset \mathbf{C}$. We also define $\sigma(z, \tau) = \sigma(z, \Lambda_\tau)$.

Let $\zeta(z, \Lambda) = \frac{\frac{d}{dz}\sigma(z,\Lambda)}{\sigma(z,\Lambda)}$ be the logarithmic derivative of σ ([Fri11, (6) on p. 209]). It is quasi-periodic in the sense that we have

$$\zeta(z + \omega_i, \Lambda) = \zeta(z, \Lambda) + \eta_i,$$

for some $\eta_1, \eta_2 \in \mathbf{C}$, which we call the *basic quasi periods* associated to ω_1, ω_2 [Fri11, (4) on p. 196]. They satisfy the Legendre relation $\omega_1\eta_2 - \omega_2\eta_1 = 2\pi i$ ([Fri11, (6) on p. 160]).

Let η (not to be confused with η_1 and η_2) be the *Dedekind eta function*

$$\eta(\tau) = q^{1/24} \prod_{n=1}^{\infty} (1 - q^n) \qquad \text{where} \qquad q = \exp(2\pi i\tau)\,.$$

2.2.2. Klein forms and Siegel functions

For $a = (a_1, a_2) \in \mathbf{Q}^2 \setminus \mathbf{Z}^2$, we define the *Klein form* $\mathfrak{k}_a$ as a function of $\mathbf{R}$-linearly independent pairs $\omega_1, \omega_2 \in \mathbf{C}$ by

$$\mathfrak{k}_a(\omega_1, \omega_2) = \exp\left(-\tfrac{1}{2}\left(a_1\eta_1 + a_2\eta_2\right)\left(a_1\omega_1 + a_2\omega_2\right)\right) \sigma\left(a_1\omega_1 + a_2\omega_2, \omega_1\mathbf{Z} + \omega_2\mathbf{Z}\right).$$

There are many variants of the notation for Klein forms in the literature. Our Klein form $\mathfrak{k}_a$ equals $-\sigma_{gh}$ in the notation of [Fri11, (6) on p. 451] where $(g/N, h/N) = a$.

Define for $a = (a_1, a_2) \in \mathbf{Q}^2 \setminus \mathbf{Z}^2$ the function $\mathfrak{k}_a : \mathbf{H} \to \mathbf{C}$ by

$$\mathfrak{k}_a(\tau) = \omega_2^{-1}\mathfrak{k}_a(\omega_1, \omega_2), \tag{2.3}$$

for any $\omega_1, \omega_2 \in \mathbf{C}$ with $\omega_1/\omega_2 = \tau$. Indeed, by [Fri11, (7) on p. 452], this depends only on a and τ, not on ω_1 and ω_2. Our $\mathfrak{k}_a(\tau)$ is exactly $\mathfrak{k}_a\binom{\tau}{1}$ of Kubert and Lang [KL81, § 2.1, p. 27].

Define the *Siegel function*

$$h_a = 2\pi\eta^2\mathfrak{k}_a,$$

which is $-i$ times the function g_a of Kubert and Lang [KL81, § 2.1, p. 29].

Remark 2.12. — Our Klein forms and Siegel functions are the same as those in Kubert and Lang [KL75, KL77] up to multiplication by a constant and taking fractional powers. Kubert and Lang do not have the factor $\frac{1}{2}$ in the exponent in the definition of $\mathfrak{t}_a(\omega_1, \omega_2)$ ([KL75, p. 176]), but this is either due to a typo in [KL75] or due to different scaling conventions on e.g. ω_i and/or η_i. Indeed, the definition as we have given it satisfies [KL75, K2 on p. 177], and it would not have done so without the factor $\frac{1}{2}$.

The notation of Kubert and Lang varies a bit from paper to paper. For details of the relations between the functions, see the following equalities, where a superscript II refers to [KL75] and IV to [KL77]. Moreover, in the case of II, a positive integer N is understood to be fixed and we have $a = (r/N, s/N)$. Up to constant factors, we have

$$\begin{aligned} \mathfrak{t}_a &= \mathfrak{t}^{\mathrm{II}}_{r,s} = \mathfrak{t}^{\mathrm{IV}}_a, \\ h_a &= \left(g^{\mathrm{II}}_{r,s}\right)^{1/(12N)} = h^{\mathrm{IV}}_a = \begin{cases} g^{\mathrm{IV}}_a & \text{if } 2a \notin \mathbf{Z}^2, \\ (g^{\mathrm{IV}}_a)^2 & \text{if } 2a \in \mathbf{Z}^2. \end{cases} \end{aligned}$$

LEMMA 2.13. — *The Siegel functions h_a have the following expansions and transformation properties for all $a = (a_1, a_2) \in \mathbf{Q}^2 \setminus \mathbf{Z}$.*

(1) *Write $q^a = \exp(2\pi i a_2) \cdot q^{a_1} = \exp(2\pi i(a_1\tau + a_2))$. If $0 \leqslant a_1 \leqslant \frac{1}{2}$, then we have*

$$h_a = c(a) q^{\frac{1}{2}\left(a_1^2 - a_1 + \frac{1}{6}\right)} (1 - q^a) \prod_{n=1}^{\infty} (1 - q^n q^a)\left(1 - q^n q^{-a}\right), \tag{2.4}$$

where $c(a) = i \exp(\pi i a_2(a_1 - 1))$ is a constant.

(2) $h_{-a} = -h_a$.

(3) $h_{(a_1+n_1, a_2+n_2)} = (-1)^{n_1 n_2 + n_1 + n_2} e^{-\pi i(n_1 a_2 - n_2 a_1)} h_{(a_1, a_2)}$ *for all* $(n_1, n_2) \in \mathbf{Z}^2$.

(4) $h_{(a_1+1,0)} = -h_{(a_1,0)}$.

(5) *h_a up to multiplication by roots of unity depends only on the class of a in $(\mathbf{Q}^2/\mathbf{Z}^2)/\{\pm 1\}$.*

(6) *For all*

$$M = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \in \mathrm{SL}_2(\mathbf{Z}),$$

we have

$$h_a(M\tau) = \epsilon(M) h_{aM}(\tau), \tag{2.5}$$

where $\epsilon(M) \in \mathbf{C}^*$ is such that for all $\tau \in \mathbf{H}$,

$$\eta(M\tau)^2 = \epsilon(M)(\gamma\tau + \delta)\eta(\tau)^2. \tag{2.6}$$

(7) *The function ϵ from (2.6) satisfies $\epsilon(M)^{12} = 1$ and*

$$\epsilon\left(\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}\right) = \exp(2\pi i/12)^{-1}.$$

Proof. — The expansion in (1) is Fricke [Fri11, (7) on p. 452], but note that our q is the square of the q of Fricke. Equivalently, the expansion is $-i$ times Kubert and Lang's ([KL75, K5 on p. 178] or equivalently [KL81, K4 on p. 29]).

The identity $h_{-a} = -h_a$ of (2) follows from the anti-symmetry of σ as a function of z.

The identity of (3) is Fricke [Fri11, (4) on p. 451]. Identity (4) is a special case of (3).

Observation (5) follows immediately from (2) and (3).

As η^{24} has level 1, if we let $\epsilon(M, \tau) = \eta(M\tau)^2/((\gamma\tau + \delta)\eta(\tau)^2)$, then we get $\epsilon(M, \tau)^{12} = 1$, hence $\epsilon(M, \tau)$ is independent of τ, call it $\epsilon(M)$. A numerical evaluation yields the example value of (7), so it remains to prove equality (2.5) in (6).

First, [Fri11, (3) on p. 451] (equivalently [KL75, K1 on p. 177]) gives

$$\mathfrak{t}_a\left(M\begin{pmatrix}\omega_1\\ \omega_2\end{pmatrix}\right) = \mathfrak{t}_{aM}\left(\begin{pmatrix}\omega_1\\ \omega_2\end{pmatrix}\right).$$

In terms of $\tau = \omega_1/\omega_2$, this reads (by (2.3))

$$(\gamma\omega_1 + \delta\omega_2)\mathfrak{t}_a(M\tau) = \omega_2\mathfrak{t}_{aM}(\tau).$$

Now multiply this equality by 2π and (2.6) to get

$$(\gamma\omega_1 + \delta\omega_2)h_a(M\tau) = (\gamma\tau + \delta)\omega_2\epsilon(M)h_{aM}(\tau),$$

which proves (2.5). □

We use the shorthand notation

$$H_k = h_{(k/N,0)}, \tag{2.7}$$

which by Lemma 2.13(1) is the same as (1.4).

2.3. Remarks on the difference between Γ^1 and Γ_1

The curve that we denote by $Y^1(N)$ is often denoted $Y_1(N)$, mostly by authors who prefer to use the group $\Gamma_1(N)$ instead, which is defined as in (1.2) with $c \equiv 0$ instead of $b \equiv 0$. We now give two remarks for how to adapt Theorem 1.2 to that situation. We will not use these remarks in the rest of this article.

Remark 2.14. — There is a complex analytic isomorphism

$$\begin{aligned}\Gamma_1(N)\backslash\mathbf{H} &\longrightarrow Y^1(N)(\mathbf{C})\\ \tau &\longmapsto \left(\mathbf{C}/\Lambda_\tau, 1/N \bmod \Lambda_\tau\right).\end{aligned} \tag{2.8}$$

The field of functions on $X^1(N)$ defined over $\mathbf{Q}$ with that choice of parametrization is the field of meromorphic functions on $\Gamma_1(N)\backslash\mathbf{H}^*$ whose expansion at the cusp 0 is rational, that is, the functions in $\mathbf{Q}((\exp(-2\pi i\tau^{-1})))$.

The isomorphism $\Gamma_1(N)\backslash\mathbf{H} \to \Gamma^1(N)\backslash\mathbf{H}$ obtained by composing the two parametrizations (2.8) and (1.3) is given by

$$\left(\begin{smallmatrix}0 & 1\\ -1 & 0\end{smallmatrix}\right) : \tau \mapsto -\tau^{-1}.$$

In particular, if one uses the parametrization (2.8), then in Theorem 1.2 one should replace $H_k(\tau)$ by $H_k(-1/\tau)$.

Remark 2.15. — There is another complex analytic isomorphism, given by

$$\begin{aligned}\Gamma_1(N)\backslash\mathbf{H} &\longrightarrow Y^1(N)(\mathbf{C})\\ \tau &\longmapsto (\mathbf{C}/\Lambda_{N\tau}, \tau \bmod \Lambda_{N\tau})\,.\end{aligned}\tag{2.9}$$

The field of functions on $X^1(N)$ defined over $\mathbf{Q}$ with that choice of parametrization is the field of meromorphic functions on $\Gamma_1(N)\backslash\mathbf{H}^*$ whose expansion at the cusp ∞ is rational, that is, the function is in $\mathbf{Q}((\exp(2\pi i\tau)))$.

The isomorphism $\Gamma_1(N)\backslash\mathbf{H} \to \Gamma^1(N)\backslash\mathbf{H}$ obtained by composing the two parametrizations (2.9) and (1.3) is given by

$$\left(\begin{smallmatrix} N & 0\\ 0 & 1\end{smallmatrix}\right) : \tau \mapsto N\tau.$$

In particular, if one uses the parametrization (2.8), then in Theorem 1.2 one should replace $H_k(\tau)$ by $H_k(N\tau)$, which is denoted by $iE_k(\tau)$ in [Yan04, Yan09].

From now on, we only use the parametrization (1.3), and will not use Remark 2.14 or 2.15.

3. Relating the functions

We now give the first part of the proof of the main theorems: relating the groups given by the sets of generators of the theorems. We start by expressing the functions P_k and p_k of Section 2.1 in terms of the Weierstrass σ-function.

3.1. The Weierstrass sigma function

To any lattice $\Lambda \subset \mathbf{C}$ of rank two and any $z \in \mathbf{C}$, we associate an elliptic curve E with $E(\mathbf{C}) = \mathbf{C}/\Lambda$ and a point $P = (z \bmod \Lambda)$.

The curve E has a classical Weierstrass equation

$$W : y^2 = 4x^3 - g_2(\Lambda)x - g_3(\Lambda),\tag{3.1}$$

where $g_2(\Lambda) = 60\sum_{\omega\in\Lambda\backslash\{0\}}\omega^{-4}$ and $g_3(\Lambda) = 140\sum_{\omega\in\Lambda\backslash\{0\}}\omega^{-6}$. We let $\Delta = \Delta(\Lambda) = 16(g_2(\Lambda)^3 - 27g_3(\Lambda)^2)$ be the discriminant of the right hand side of (3.1).

After putting the pair (E, P) in Tate normal form, we get B and C as functions in z and Λ. In particular, we get expressions for P_k in terms of z and Λ. The following result gives these expressions.

Proposition 3.1. — *For any positive integer k, let*

$$\Phi_k = \frac{\sigma(kz,\Lambda)}{\sigma(z,\Lambda)^{k^2}} \quad \textit{and} \quad U = \frac{\Phi_3}{\Phi_2^3}.$$

Then we have

$$P_k = U^{k^2-1}\Phi_k \quad \textit{and} \quad D = U^{12}\Delta.$$

Proof. — Let $\wp(z,\Lambda)$ be the Weierstrass $\wp$-function and $\wp' = \frac{d}{dz}\wp$. Then for any $v \in \mathbf{C}$, we get a point $(x,y) = (\wp(v,\Lambda), \wp'(v,\Lambda))$ on (3.1).

We put the classical Weierstrass equation W in Tate normal form relative to the point $P = (x_0,y_0) = (\wp(z,\tau), \wp'(z,\tau))$. The transformation is of the form $X = u^2(x+t)$, $Y = \frac{1}{2}u^3(y+rx+s)$ with u,r,s,t functions of z and τ, where X and Y are the coordinate functions for the Tate normal form and x and y are the coordinate functions for the classical Weierstrass equation.

First, we compute the discriminant D of the Tate normal form. Completing the square to get an equation of the form $Y'^2 = X^3 + \cdots$ does not affect the discriminant or the X-coordinates of the two-torsion points. Note that the discriminant of a Weierstrass equation $(Y')^2 = X^3 + \cdots$ is 16 times the discriminant of the right hand side. Let Q_1, Q_2, Q_3 be the points of order 2 on E. Then

$$\begin{aligned} D &= 16 \cdot (X(Q_1) - X(Q_2))^2 \cdot (X(Q_2) - X(Q_3))^2 \cdot (X(Q_3) - X(Q_1))^2 \\ &= 16u^{12} \cdot (x(Q_1) - x(Q_2))^2 \cdot (x(Q_2) - x(Q_3))^2 \cdot (x(Q_3) - x(Q_1))^2 \\ &= u^{12}\Delta. \end{aligned}$$

Similarly, we have

$$P_k = k\sqrt{\prod_{Q\in E[k]\setminus\{0\}} (X - X(Q))} = u^{(k^2-1)}k\sqrt{\prod_{Q\in E[k]\setminus\{0\}} (x - x(Q))}, \tag{3.2}$$

where the square roots are is chosen to be monic polynomials times 1 or times $Y + \frac{1}{2}a_1X + \frac{1}{2}a_3$.

We use the classical identity

$$(-1)^{k+1}\ k\sqrt{\prod_{Q\in E[k]\setminus\{0\}} (x - x(Q))} = \frac{\sigma(kz,\Lambda)}{\sigma(z,\Lambda)^{k^2}}. \tag{3.3}$$

For a proof, see De Looij [dL10, Theorem 2.7]. The factor $(-1)^{k+1}$ does not appear in [dL10], but our choice of square root differs from the choice in loc. cit. by exactly that factor. The proof in [dL10] works by fixing the lattice Λ and showing that both sides are elliptic functions for that lattice with the same divisor and with equal leading terms in their power series.

Combining (3.2) and (3.3), we get

$$P_k = (-u)^{k^2-1}\frac{\sigma(kz,\Lambda)}{\sigma(z,\Lambda)^{k^2}} = (-u)^{k^2-1}\Phi_k, \tag{3.4}$$

so it suffices to prove $-u = U$.

Proving $-u = U$ could be done by a lengthy computation of the Tate normal form from W. Instead, simply note

$$1 = \frac{B^3}{B^3} = \frac{P_3}{P_2^3} = \frac{(-u)^{3^2-1}}{(-u)^{3(2^2-1)}}\frac{\Phi_3}{\Phi_2^3} = (-u)^{-1}U,$$

which finishes the proof. □

Next, we specialize to $\Lambda = \Lambda_\tau$ and $z = \tau/N$ consistently with the identification $\Gamma^1(N)\backslash\mathbf{H} \to Y^1(N)(\mathbf{C})$ of (1.3).

Corollary 3.2. — *For any integer $N \geqslant 4$ and any positive integer k with $N \nmid k$, let*

$$\phi_k = \frac{\sigma\left(\frac{k\tau}{N}, \tau\right)}{\sigma\left(\frac{\tau}{N}, \tau\right)^{k^2}} \quad \text{and} \quad u = \frac{\phi_3}{\phi_2^3}.$$

Then the following identities of meromorphic functions hold on $X^1(N)$:

$$p_k = u^{k^2-1}\phi_k \quad \text{and} \quad d = \left(2\pi\eta^2 u\right)^{12}.$$

Proof. — Take $\Lambda = \Lambda_\tau$ and $z = \tau/N$ in Proposition 3.1, and use the well known equality $\Delta(\Lambda_\tau) = (2\pi\eta(\tau)^2)^{12}$ (see [Fri11, (6) on p. 313]). □

3.2. The functions p_k in terms of the functions H_k

Now that we have expressed the functions p_k in terms of Weierstrass σ-functions, we use these expressions to express the p_k in terms of Siegel functions. Recall the functions H_k from (2.7).

Lemma 3.3. — *Let*

$$t = \frac{H_1^2 H_3}{H_2^3}.$$

Then for all integers $N \geqslant 4$ and $k \in \mathbf{Z} \setminus N\mathbf{Z}$ we have

$$p_k = t^{k^2-1}\frac{H_k}{H_1} \quad \text{and} \quad d = (tH_1)^{12}.$$

Proof. — In the notation of Corollary 3.2, we have

$$\begin{aligned} \phi_k &= \frac{\sigma(k\tau/N, \tau)}{\sigma(\tau/N, \tau)^{k^2}} = \frac{\mathfrak{t}_{(k/N,0)}}{\mathfrak{t}_{(1/N,0)}^{k^2}} = \frac{H_k}{H_1^{k^2}}\left(2\pi\eta^2\right)^{k^2-1} \\ &= \frac{H_k}{H_1}\left(\frac{H_1}{2\pi\eta^2}\right)^{1-k^2}, \\ u &= \phi_3\phi_2^{-3} = t \cdot \left(\frac{H_1}{2\pi\eta^2}\right), \\ p_k &= u^{k^2-1}\phi_k = t^{k^2-1}\frac{H_k}{H_1}, \\ d &= \left(2\pi\eta^2 u\right)^{12} = (tH_1)^{12}, \end{aligned}$$

so the result follows. □

Let $m = \lfloor N/2 \rfloor$. Next, we express p_{m+1} in terms of H_k with $1 \leqslant k \leqslant m$ using the periodicity and symmetry of H_k in k.

Lemma 3.4. — *Let t be as in Lemma 3.3, let $m = \lfloor N/2 \rfloor$, and let $v = t^{\gcd(2,N)N}$. Then we have*

$$p_{m+1} = \begin{cases} vp_m, & \text{if } N \text{ is odd,} \\ vp_{m-1}, & \text{if } N \text{ is even.} \end{cases}$$

Moreover, each of $d, p_2, p_4, p_5, p_6, \ldots, p_{m+1}$ *(including* $-b = p_2$*) is of the form*

$$f = \prod_{k=1}^{m} H_k^{e(k)},$$

where for every $k \in \{1, 2, \ldots, m\}$ *we have* $e(k) \in \mathbf{Z}$*, and where we have*

$$\sum_{k=1}^{m} e(k) \in 12\mathbf{Z} \qquad \text{and} \qquad \sum_{k=1}^{m} k^2 e(k) \in N\mathrm{gcd}(N, 2)\mathbf{Z}. \tag{3.5}$$

Proof. — Suppose first that N is odd, so $N = 2m + 1$. Lemma 3.3 gives

$$p_{m+1} = t^{(m+1)^2-1} H_{m+1}/H_1$$

and by Lemma 2.13 (parts (2) and (4)), we have $H_{m+1} = -H_{-(m+1)} = H_m$, hence

$$p_{m+1} = t^{2m+1} t^{m^2-1} H_m/H_1 = v p_m.$$

If N is even, then $N = 2m$ and $t^{(m+1)^2-1} = t^{4m} t^{(m-1)^2-1}$, so the same calculation gives $p_{m+1} = v p_{m-1}$.

A straightforward calculation verifies (3.5) for each expression in Lemma 3.3 or 3.4. Indeed, the value of $(\sum_k e(k), \sum_k k^2 e(k)) \in \mathbf{Z}^2$ is

$$\begin{array}{rl}
\left(1, n^2\right) & \text{for } H_n \text{ for all } n \in \mathbf{Z} \text{ with } N \nmid n, \\
(0, -1) & \text{for } t, \\
(12, 0) & \text{for } d, \\
(0, 0) & \text{for } p_n \text{ with } 1 \leqslant n \leqslant m \text{ (Lemma 3.3)}, \\
(0, -\mathrm{gcd}(N, 2)N) & \text{for } v \text{ and hence for } p_{m+1} \text{ (Lemma 3.4)}.
\end{array}$$

□

3.3. The functions H_k in terms of p_k

Now that we have expressions of p_k in terms of H_k, it is a matter of solving a system of linear equations to obtain the reverse expressions. These expressions are given in the following result.

Proposition 3.5. — *Let* $m = \lfloor N/2 \rfloor$*. Given* $e \in \mathbf{Z}^m$ *satisfying* (3.5) *and given*

$$f = \prod_{k=1}^{m} H_k^{e(k)},$$

let $\alpha = \frac{1}{12} \sum_k e(k)$ *and* $\beta = (N\mathrm{gcd}(2, N))^{-1} \sum_k k^2 e(k)$*. Then we have*

$$f = d^\alpha \left(p_{N-m-1} p_{m+1}^{-1}\right)^\beta \prod_{k=1}^{m} p_k^{e(k)}, \tag{3.6}$$

where $p_1 = 1$*,* $p_2 = -b$*,* $p_3 = -b^3$*, and* $N - m - 1 \in \{m - 1, m\}$*, so*

$$f \in \langle -b, d, p_4, p_5, \ldots, p_{m+1} \rangle \subset \mathcal{O}\left(Y^1(N)\right)^*.$$

Proof. — Note that Lemma 3.3 gives

$$\prod_{k=1}^{m} p_k^{e(k)} = t^{\sum_k k^2 e(k)^2} (tH_1)^{-\sum e(k)} \prod_{k=1}^{m} H_k^{e(k)} = v^{\beta} d^{-\alpha} \prod_{k=1}^{m} H_k^{e(k)}.$$

As Lemma 3.4 gives $v = p_{m+1} p_{N-m-1}^{-1}$, this proves (3.6). The formulas for p_1, p_2 and p_3 are in Example 2.2. □

The following result sums up in how far we have now proven the main theorems.

Proposition 3.6. — *Let S be the group of functions of the form $\prod_{k=1}^{m} H_k^{e(k)}$ satisfying* (3.5)*. If S has rank m and $\mathbf{Q}^* \cdot S$ contains $\mathcal{O}(Y^1(N))^*$, then all of Theorems 1.1, 1.2, 2.6 and 2.8 hold.*

Proof. — Let $T = \langle -b, d, p_4, p_5, \ldots, p_{m+1} \rangle \subset \mathcal{O}(Y^1(N))^*$. Lemma 3.3 and Proposition 3.5 show $S = T$, hence also $\mathbf{Q}^* \cdot S \subset \mathcal{O}(Y^1(N))^*$.

The leading coefficients of the q-expansions of the functions $H_k(\tau)$ are all i by (1.4), hence the leading coefficients of q-expansions of the elements of S are all 1, so $S \cap \mathbf{Q}^* = 1$. In particular, the rank of $(\mathbf{Q}^* \cdot S)/\mathbf{Q}^*$ equals the rank of S.

Under the assumption that this rank is m and that $\mathbf{Q}^* \cdot S$ contains $\mathcal{O}(Y^1(N))^*$, we get exactly Theorems 1.2 and 2.8.

By Lemma 2.7, Theorem 2.8 implies Theorems 1.1 and 2.6. □

4. q-expansions and Gauss' Lemma

Recall that S is the group of functions of the form $\prod_{k=1}^{m} H_k^{e(k)}$ satisfying (3.5), where $m = \lfloor N/2 \rfloor$. As stated in Proposition 3.6, it now suffices to prove that S has rank m and $\mathcal{O}(Y^1(N))^* \subset \mathbf{Q}^* \cdot S$.

Section 4.1 uses q-expansions to show that the Siegel functions H_k for $k = 1, 2, \ldots, m$ are multiplicatively independent. The group they generate then has the correct rank.

Section 4.2 combines this with Gauss' Lemma for power series to show that $\mathcal{O}(Y^1(N))^*$ is contained in $\mathbf{Q}^* \cdot \langle H_1, H_2, \ldots, H_{m-2}, H_{m-1}, H_m^{1/2} \rangle$.

Section 4.3 then uses explicit SL_2-actions to find restrictions on the exponent vectors, finishing the proof of $\mathcal{O}(Y^1(N))^* \subset \mathbf{Q}^* \cdot S$.

4.1. The rank

Proposition 4.1. — *The functions H_k for $k = 1, 2, \ldots, m$ are multiplicatively independent modulo $\mathbf{C}^*$. In other words, if*

$$\prod_{k=1}^{m} H_k^{e(k)} \in \mathbf{C}^*$$

with $e \in \mathbf{Z}^k$, then $e = 0$.

Proof. — We prove the result using q-expansions. Following [KL77], we define the *reduced form* f^* of a non-zero Laurent series f to be f divided by its lowest-degree term, so $f^* = 1 +$ higher order terms.

From (1.4), we have for $0 < k \leqslant N/2$:

$$\begin{aligned} H_k^* &= \left(1 - q^{k/N}\right) \prod_{n=1}^{\infty} \left(1 - q^{n+k/N}\right)\left(1 - q^{n-k/N}\right) \\ &= \begin{cases} 1 - q^{k/N} + O\left(q^{1-k/N}\right) & \text{if } 0 < k < N/2, \text{ and} \\ 1 - 2q^{1/2} + O\left(q^{3/2}\right) & \text{if } k = N/2. \end{cases} \end{aligned} \tag{4.1}$$

Suppose that we have $\prod_{k=1}^{m} H_k^{e(k)} \in \mathbf{C}^*$ for some $0 \neq e \in \mathbf{Z}^m$. Let k_0 be the smallest positive integer with $e(k_0) \neq 0$. Then (4.1) gives

$$1 = \prod_{k=k_0}^{m} \left(H_k^*\right)^{e(k)} = \begin{cases} 1 - e(k_0) q^{k_0/N} + O\left(q^{(k_0+1)/N}\right) & \text{if } 2k_0 \neq N, \text{ and,} \\ 1 - 2e(k_0) q^{k_0/N} + O\left(q^{(k_0+1)/N}\right) & \text{if } 2k_0 = N. \end{cases} \tag{4.2}$$

We get $e(k_0) = 0$, contradiction. □

Corollary 4.2. — *Let S be the group of functions $\prod H_k^{e(k)}$ satisfying* (3.5)*. Then the image of S in $\mathcal{O}(Y^1(N))^*/\mathbf{Q}^*$ has finite index.*

Proof. — Proposition 4.1 shows that S has rank m. We have

$$\operatorname{rk}\left(\mathcal{O}\left(Y^1(N)\right)^* / \mathbf{Q}^*\right) \leqslant \#\left(\left\{\text{cusps of } X^1(N)\right\} \big/ \operatorname{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})\right) - 1.$$

As the right hand side is m by [DvH14, Definition 1 in §2], this proves the result. □

We recover the following consequence of the Manin-Drinfeld theorem [Man72, Dri73], which states that the cuspidal parts of modular Jacobians are torsion.

Corollary 4.3. — *The group*

$$\frac{\operatorname{Div}^{0,\text{cusp}}\left(X^1(N)\right)}{\mathcal{O}\left(Y^1(N)\right)^* / \mathbf{Q}^*}$$

of cuspidal divisor classes of $X^1(N)$ is finite.

Proof. — As seen in the proof of Corollary 4.2, the two groups in the quotient both have rank m. □

4.2. Roots of power series

In Corollary 4.2, we have shown that every $f \in \mathcal{O}(Y^1(N))^*$ can be expressed as a product of powers $c \prod_{k=1}^{m} H_k^{e(k)}$ with $e \in \mathbf{Q}^m$, $c \in \mathbf{C}^*$ and $m = \lfloor N/2 \rfloor$. The current section is devoted to proving that the exponent $e(k)$ is an integer for $k \neq N/2$. The key idea, taken from Kubert and Lang [KL77] is to combine Gauss' lemma for power series with the fact that q-expansions of modular forms have bounded denominators.

We call a power series $f \in \mathbf{Z}[[x]]$ *primitive* if the ideal generated by its coefficients is (1). We then have the following variant of Gauss' lemma.

Lemma 4.4. — *Let $f, g \in \mathbf{Z}[[x]]$ be primitive power series. Then $fg \in \mathbf{Z}[[x]]$ is also primitive.*

Proof. — Given any prime number p, take the lowest-order terms of $(f \bmod p)$ and $(g \bmod p)$ (which exist by primitivity). Their product is a non-zero term of $(fg \bmod p)$, so $p \nmid fg$. □

We say that a Laurent series $f \in \mathbf{Q}((x))$ has *bounded denominators* if there is a non-zero $d \in \mathbf{Z}$ such that $df \in \mathbf{Z}((x))$.

Corollary 4.5. — *Let $f, g \in \mathbf{Q}[[x]]$ be power series with bounded denominators and constant term 1. If fg is in $\mathbf{Z}[[x]]$, then $f, g \in \mathbf{Z}[[x]]$.*

Proof. — Take $a, b \in \mathbf{Z}$ such that af and bg are primitive in $\mathbf{Z}[[x]]$. Then $(ab)(fg)$ is primitive by Lemma 4.4, hence $a, b \in \{\pm 1\}$. □

Proposition 4.6 (Special case of Kubert and Lang [KL77, Lemma 3.1]). — *Let f be a modular unit with rational q-expansion, that is, in $\mathbf{Q}((q^{1/M}))$ for some M. Then the q-expansion has bounded denominators.*

Proof. — See [KL77, Lemma 3.1] for the proof, of which we give a sketch here. After multiplying by a suitable power of η^{24}, the function becomes a cusp form. The vector space of cusp forms of given weight is generated by forms with *integer* Fourier expansions, hence the result follows. □

For a formal power series f with constant coefficient 1 and for $a, b \in \mathbf{Z} \setminus \{0\}$, we define $f^{a/b}$ to be the unique bth root of f^a with constant coefficient 1. For a holomorphic, non-vanishing function f on $\mathbf{H}$, we denote by $f^{a/b}$ any holomorphic bth root of f^a.

Proposition 4.7. — *Let f be a* modular *function of any level and suppose that we have*

$$f = c \prod_{k=1}^{m} H_k^{e(k)}$$

with $e \in \mathbf{Q}^m$, $c \in \mathbf{C}^*$ and $m = \lfloor N/2 \rfloor$. Then for all k we have $e(k) \in \mathbf{Z}$ if $2k \neq N$ and $e(k) \in \frac{1}{2}\mathbf{Z}$ if $2k = N$.

Proof. — Taking reduced forms (as defined in the proof of Proposition 4.1) on both sides, we get

$$(f^*)^n = \prod_{k=1}^{m} (H_k^*)^{n \cdot e(k)}$$

for some n with $ne \in \mathbf{Z}^m$. The right hand side has integer coefficients, so by Proposition 4.6 and Corollary 4.5, we find that f^* has integer coefficients.

We prove the result by induction on k. Suppose it is true for all $k < k_0$. We have

$$f^* \cdot \prod_{k=1}^{k_0 - 1} (H_k^*)^{-e(k)} = \prod_{k=k_0}^{m} (H_k^*)^{e(k)},$$

and the left hand side has integer coefficients. By (4.2), the right hand side has a coefficient $-e(k_0)$ if $2k_0 \neq N$ and $-2e(k_0)$ if $2k_0 = N$, hence the result follows. □

4.3. Using the action

Next, we use the action of SL_2. Recall $m = \lfloor N/2 \rfloor$.

THEOREM 4.8. — *Let $f \in \mathcal{O}(Y^1(N))^*$. Then $f = c\prod_{k=1}^{m} H_k^{e(k)}$, where $c \in \mathbf{Q}^*$ and $e \in \mathbf{Z}^m$ are uniquely determined by f. Moreover, the vector e satisfies* (3.5) *that is, it satisfies*

$$\sum_k e(k) \in 12\mathbf{Z} \qquad \textit{and} \qquad \sum_k k^2 e(k) \in N\gcd(N,2)\mathbf{Z}.$$

Proof. — By Corollary 4.2, we find that f can be written as $c\prod H_k^{e(k)}$ with $e(k) \in \mathbf{Q}$. Here e is uniquely determined by Proposition 4.1. Moreover, the numbers $e(1)$, $e(2)$, …, $e(m-1)$ are in $\mathbf{Z}$ by Proposition 4.7. Next, we prove (3.5), which also implies $e(m) \in \mathbf{Z}$.

Consider the matrix

$$M = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \in \Gamma^1(N).$$

Then we have $f(M\tau) = f(\tau)$, so we inspect the action of M on the functions H_k. Parts (6) and (7) of Lemma 2.13 give $H_k(M\tau) = \exp(2\pi i/12)^{-1} H_k(\tau)$ for this matrix M. In particular, we get $\sum_k e(k) \in 12\mathbf{Z}$.

Next, consider the matrix

$$M = \begin{pmatrix} 1 & N \\ 0 & 1 \end{pmatrix} \in \Gamma^1(N).$$

Again we have $f(M\tau) = f(\tau)$, that is, $f(\tau + N) = f(\tau)$, which shows that the q-expansion of f is in $\mathbf{C}((q^{1/N}))$. In the product expansion (2.4), we consider the leading term $-iq^{\frac{1}{2}(a_1^2 - a_1 + \frac{1}{6})}$ (with $a_1 = k/N$) of H_k. As the leading term of f is a constant times a power of $q^{1/N}$, we get

$$\frac{1}{12N^2} \sum_{k=1}^{m} e(k) \left(6k^2 - 6kN + N^2\right) \in \frac{1}{N}\mathbf{Z}.$$

As we already have $\sum e(k) \in 12\mathbf{Z}$, we get

$$\sum_{k=1}^{m} e(k) \left(k^2 - kN\right) \in 2N\mathbf{Z} \subset N\mathbf{Z},$$

hence in particular $\sum e(k)k^2 \in N\mathbf{Z}$. If N is odd, then this finishes the proof of (3.5). If N is even, then we get

$$\begin{aligned}(1-N)\sum_{k=1}^{m} e(k)k^2 &= \sum_{k=1}^{m} e(k)\left(k^2 - k^2 N\right) \\ &\equiv \sum_{k=1}^{m} e(k)\left(k^2 - kN\right) \equiv 0 \bmod 2N\mathbf{Z},\end{aligned}$$

and since $N-1$ is coprime to $2N$, this proves (3.5) and hence $e(m) \in \mathbf{Z}$.

It remains to prove $c \in \mathbf{Q}^*$. Let $g = f/c$, which is in $\mathcal{O}(Y^1(N))^*$ by Proposition 3.5. Then $c = f/g$ is a constant in $\mathcal{O}(Y^1(N))^*$, hence is in $\mathbf{Q}^*$. □

Proof of the main theorems. — Proposition 3.6 states exactly that Theorem 4.8 and the rank statement in Proposition 4.1 imply Theorems 1.1, 1.2, 2.6 and 2.8. □

Remark 4.9. — Results similar to Theorem 4.8, but assuming integral exponents $e(k)$ and working with $\Gamma(N)$, are already known. These results are insufficient for proving our main results as they assume that $e(k)$ is integral.

In the special case where N is coprime to 6, they can be used to an alternative proof of our Theorem 4.8 as follows. If N is odd, then Proposition 4.7 gives $e \in \mathbf{Z}^m$. For $e \in \mathbf{Z}^m$, Kubert and Lang [KL81, Theorem 5.2 and 5.3 on pp. 76–78 in Chapter 3] give conditions on e for f to be modular of level $\Gamma(N)$. The conditions are complicated, but if N is coprime to 6, then the conditions give exactly (3.5), which reproves Theorem 4.8 in that case.

5. Bonus section

There are two results that we get almost for free after all the work that was done towards the main theorem. We give them here.

5.1. Ring generators

In this section, we give complex analytic functions that generate the ring $\mathcal{O}(Y^1(N))$ itself, instead of its unit group.

Theorem 5.1. — *The ring $\mathcal{O}(Y^1(N))$ is generated as a $\mathbf{Q}$-algebra by the three functions*

$$b = -t^3\frac{H_2}{H_1}, \quad c = -H_4H_1^4H_2^{-5}, \quad d^{-1} = (tH_1)^{-12},$$

where

$$t = \frac{H_1^2H_3}{H_2^3}.$$

Proof. — By Proposition 2.5, we have $\mathcal{O}(Y^1(N)) = \mathbf{Q}[b, c, d^{-1}]$. We have $b = -p_2$ and $c = p_4/b^5$ by Example 2.2. Lemma 3.3 gives the formulas in terms of Siegel functions. □

Theorem 5.1 is comparable to the main result of Koo and Yoon [KY17]. Indeed, both give a set of complex analytic functions that generate the $\mathbf{Q}$-algebra $\mathcal{O}(Y^1(N))$, and through the isomorphism of Remark 2.15 also the $\mathbf{Q}$-algebra of holomorphic modular functions on $\Gamma_1(N)\backslash\mathbf{H}$ with rational $\mathbf{Q}$-expansion. The methods are however completely different.

As for the results themselves, they are different as well. First of all, the main result of [KY17] (that is, Theorems 4.5 and 5.2 and Corollary 5.3 of loc. cit.) are for $N = 2$, $N = 3$ and all N divisible by 4, 5, 6, 7 or 9, while our result is for all $N \geqslant 4$. Second, we give a uniform formula with three generators, while [KY17] has a few different cases, each with 2 to 6 generators.

5.2. Expressions in terms of the Jacobi theta function

In this section, we express the functions p_k in terms of the Jacobi theta function. This has two applications. First of all, this theta function can be numerically evaluated efficiently, as in Labrande [Lab18]. Second, it has a natural generalization to the moduli space of higher-dimensional abelian varieties (Riemann theta functions), potentially opening our results to future higher-dimensional generalisations.

For $c, d \in \mathbf{R}$, the *theta function* $\theta[c,d]$ *with characteristic* (c,d) is the function in $z \in \mathbf{C}$ and $\tau \in \mathbf{H}$ defined by

$$\begin{aligned}\theta[c,d](z,\tau) &= \sum_{n\in\mathbf{Z}} \exp\Big(\pi i(n+c)^2\tau + 2\pi i(n+c)(z+d)\Big)\\ &= e^{\pi i c^2\tau + 2\pi i c(z+d)} \cdot \theta[0,0](z + c\tau + d, \tau).\end{aligned}$$

We will use a special case, known as the *Jacobi theta function* $\theta_1 = \theta[\frac{1}{2},\frac{1}{2}] = \theta[-\frac{1}{2},-\frac{1}{2}]$, that is,

$$\theta_1(z,\tau) = i\sum_{n\in\mathbf{Z}} (-1)^n q^{\frac{1}{2}\left(n-\frac{1}{2}\right)^2} e^{\pi i(2n-1)z}.$$

Proposition 5.2. — *Consider the functions T_k given by*

$$T_k = \theta_1\left(\frac{k\tau}{N}, \tau\right).$$

Then we have for all integers k

$$p_k = \left(\frac{T_1^2T_3}{T_2^3}\right)^{k^2-1}\frac{T_k}{T_1} \quad \text{and} \quad d = \left(\frac{T_1^3T_3}{T_2^3\eta}\right)^{12}.$$

Proof. — Let $\theta_1'(z,\tau) = \frac{d}{dz}\theta_1(z,\tau)$. Let $\Lambda = 2\omega_1\mathbf{Z} + 2\omega_3\mathbf{Z}$ with $\tau = \omega_3/\omega_1 \in \mathbf{H}$. Then (6.22) in Theorem 6.5 on page 199 of [Mar67] states (note that our q is the square of the q in loc. cit.)

$$\sigma(z,\Lambda) = 2\omega_1\frac{\theta_1(z/(2\omega_1),\tau)}{\theta_1'(0,\tau)}\exp\Big(\eta_1 z^2/(2\omega_1)\Big).$$

We choose $\omega_1 = \frac{1}{2}$ and $\omega_3 = \frac{1}{2}\tau$ to get $\sigma(z,\tau) = c_1\exp(c_2z^2)\theta_1(z,\tau)$, where $c_1 = \theta_1'(0,\tau)^{-1}$ and $c_2 = \eta_1$ are functions of τ independent of z. We apply this to the formulas in Corollary 3.2 and get

$$\begin{aligned}\phi_k &:= \frac{\sigma\left(\frac{k\tau}{N},\tau\right)}{\sigma\left(\frac{\tau}{N},\tau\right)^{k^2}} = c_1^{1-k^2}\frac{T_k}{T_1^{k^2}},\\ u &:= \frac{\phi_3}{\phi_2^3} = c_1\frac{T_1^3T_3}{T_2^3},\\ p_k &= u^{k^2-1}\phi_k = \left(\frac{u}{c_1}\right)^{k^2-1}\frac{T_k}{T_1^{k^2}} = \left(\frac{T_1^2T_3}{T_2^3}\right)^{k^2-1}\frac{T_k}{T_1},\\ d &= (2\pi\eta^2u)^{12}.\end{aligned}$$

This proves the formula for p_k. To prove the formula for d, it suffices to prove $2\pi\eta^2 u = \pm u/(c_1\eta)$, or in other words, $2\pi\eta^3 = \pm\theta_1'(0,\tau)$. But that is exactly the formula for $\theta_1'(0,\tau)$ in the middle of page 210 of Markushevich [Mar67] together with (6.52) on page 211 of loc. cit. (In fact, reading further in [Mar67], we get that the sign is +, but we do not need this.) □

BIBLIOGRAPHY

[dL10] Rutger de Looij, *Elliptic divisibility sequences*, Master's thesis, Mathematical Sciences, Universiteit Utrecht, Netherlands, 2010, written under the supervision of Gunther Cornelissen, `https://studenttheses.uu.nl/handle/20.500.12932/7322`. ↑106

[Dri73] Vladimir G. Drinfeld, *Two theorems on modular curves*, Funkts. Anal. Prilozh. **7** (1973), no. 2, 83–84. ↑110

[DvH14] Maarten Derickx and Mark van Hoeij, *Gonality of the modular curve $X_1(N)$*, J. Algebra **417** (2014), 52–71. ↑95, 96, 99, 100, 110

[Fri11] Robert Fricke, *Die elliptischen Funktionen und ihre Anwendungen. Erster Teil. Die funktionentheoretischen und analytischen Grundlagen*, Springer, 2011, Reprint of the 1916 original. ↑102, 103, 104, 107

[IMS$^+$12] Patrick Ingram, Valéry Mahé, Joseph H. Silverman, Katherine E. Stange, and Marco Streng, *Algebraic divisibility sequences over function fields*, J. Aust. Math. Soc. **92** (2012), no. 1, 99–126. ↑97

[Jin13] Jinbi Jin, *Homogeneous division polynomials for Weierstrass elliptic curves*, `http://arxiv.org/abs/1303.4327v1`, 2013. ↑100

[KL75] Daniel S. Kubert and Serge Lang, *Units in the modular function field. II. A full set of units*, Math. Ann. **218** (1975), no. 2, 175–189. ↑103, 104

[KL77] ______, *Units in the modular function field. IV. The Siegel functions are generators*, Math. Ann. **227** (1977), no. 3, 223–242. ↑97, 98, 103, 110, 111

[KL81] ______, *Modular units*, Grundlehren der Mathematischen Wissenschaften, vol. 244, Springer, 1981. ↑97, 98, 102, 103, 113

[Kub81] Daniel S. Kubert, *The square root of the Siegel group*, Proc. Lond. Math. Soc. **43** (1981), no. 2, 193–226. ↑98

[KY17] Ja Kyung Koo and Dong Sung Yoon, *Generators of the ring of weakly holomorphic modular functions for* $\Gamma_1(N)$, Ramanujan J. **42** (2017), no. 3, 583–599. ↑113

[Lab18] Hugo Labrande, *Computing Jacobi's theta in quasi-linear time*, Math. Comput. **87** (2018), no. 311, 1479–1508. ↑114

[Man72] Yuri I. Manin, *Parabolic points and zeta functions of modular curves*, Izv. Akad. Nauk SSSR, Ser. Mat. **36** (1972), 19–66. ↑110

[Mar67] Alekseĭ I. Markushevich, *Theory of Functions of a Complex Variable. Vol. III*, Selected Russian Publications in the Mathematical Sciences, Chelsea Publishing; Prentice Hall, 1967, revised English edition translated and edited by Richard A. Silverman. ↑114, 115

[Nas16] Bartosz Naskręcki, *Divisibility sequences of polynomials and heights estimates*, New York J. Math. **22** (2016), 989–1020. ↑97

[SageMath14] The Sage Developers, *SageMath, the Sage Mathematics Software System (Version 6.2)*, 2014, `https://www.sagemath.org`. ↑99

[Sil86] Joseph H. Silverman, *The arithmetic of elliptic curves*, Graduate Texts in Mathematics, vol. 106, Springer, 1986. ↑99

[Yan04] Yifan Yang, *Transformation Formulas for Generalized Dedekind Eta Functions*, Bull. Lond. Math. Soc. **36** (2004), no. 5, 671–682. ↑95, 96, 98, 105

[Yan09] ______, *Modular units and cuspidal divisor class groups of $X_1(N)$*, J. Algebra **322** (2009), no. 2, 514–553. ↑95, 96, 97, 98, 105

Manuscript received on 24th September 2021,
revised on 7th October 2022,
accepted on 10th December 2022.

Recommended by Editors X. Caruso and B. Edixhoven.

eISSN: 2644–9463
This journal is a member of Centre Mersenne.

Marco STRENG
Universiteit Leiden,
Niels Bohrweg 1,
2333 CA Leiden (The Netherlands)
http://pub.math.leidenuniv.nl/~strengtc/
streng@math.leidenuniv.nl