Algebra & Number Theory

Volume 11
2017
No. 5

Collinear CM-points

Yuri Bilu, Florian Luca and David Masser

msp

ALGEBRA AND NUMBER THEORY 11:5 (2017)
dx.doi.org/10.2140/ant.2017.11.1047

Collinear CM-points

Yuri Bilu, Florian Luca and David Masser

André's celebrated theorem of 1998 implies that each complex straight line $Ax + By + C = 0$ (apart from obvious exceptions) contains at most finitely many points $(j(\tau), j(\tau'))$, where $\tau, \tau' \in \mathbb{H}$ are algebraic of degree 2. We show that there are only a finite number of such lines which contain more than two such points. As there is a line through any two complex points, this is the best possible result.

1. Introduction

André [1998] proved that a nonspecial irreducible plane curve in $\mathbb{C}^2$ may have at most finitely many CM-points. Here a *plane curve* is a curve defined by an irreducible equation $F(x, y) = 0$, where F is a polynomial with complex coefficients, and a *CM-point* (called also a *special point*) in $\mathbb{C}^2$ is a point whose coordinates are both singular moduli. Recall that a *singular modulus* is the invariant of an elliptic curve with complex multiplication; in other words, it is an algebraic number of the form $j(\tau)$, where j denotes the standard j-function on the upper half-plane $\mathbb{H}$ and

MSC2010: primary 11G15; secondary 11G18.
Keywords: CM points, André–Oort.

$\tau \in \mathbb{H}$ is an algebraic number of degree 2. Thus, a CM-point is a point of the form $(j(\tau), j(\tau'))$ with $\tau, \tau' \in \mathbb{H}$ algebraic of degree 2.

Special curves are those of the following types:

- "vertical lines" $x = j(\tau)$ and "horizontal lines" $y = j(\tau)$, where $j(\tau)$ is a singular modulus, and
- *modular curves* $Y_0(N)$, realized as the plane curves $\Phi_N(x, y) = 0$, where Φ_N is the modular polynomial of level N.

Recall that the polynomial $\Phi_N(X, Y) \in \mathbb{C}[X, Y]$ is the X-monic $\mathbb{C}$-irreducible polynomial satisfying $\Phi_N(j(z), j(Nz)) = 0$. It is known that actually $\Phi_N(X, Y) \in \mathbb{Z}[X, Y]$; this and other properties of Φ_N can be found, for instance, in [Cox 1989, Theorem 11.18].

Clearly, each special curve contains infinitely many CM-points, and André proved that special curves are characterized by this property.

André's result was the first nontrivial contribution to the celebrated André–Oort conjecture on the special subvarieties of Shimura varieties; see [Pila 2011] and the references therein.

Several other proofs (some conditional on the GRH) of André's theorem were suggested [Bilu et al. 2013; Breuer 2001; Edixhoven 1998; Kühne 2012; 2013; Pila 2009]. We specifically mention the argument of Pila [2009], based on an idea of Pila and Zannier [2008]. Pila [2011] extended it to higher dimensions, proving the André–Oort conjecture for subvarieties of $\mathbb{C}^n$. To state this result, one needs to introduce the notion of "special variety"; then Pila's theorem asserts that an algebraic subvariety of $\mathbb{C}^n$ has at most finitely many maximal special subvarieties. See Section 2 and Theorem 2.4 for the details.

Besides general results, some particular curves were considered. For instance, Kühne [2013, Theorem 5] proved that the straight line $x + y = 1$ has no CM-points,[1] and a similar result for the hyperbola $xy = 1$ was obtained in [Bilu et al. 2013]. The same conclusion was obtained in [Habegger et al. 2017] for the quartic curve

$$x^3y - 2x^2y^2 + xy^3 - 1728x^3 + 1216x^2y + 1216xy^2 - 1728y^3 + 3538944x^2$$
$$-2752512xy + 3538944y^2 - 2415919104x - 2415919104y + 549755813888 = 0;$$

this is equivalent to the fact that there are no complex $t \neq 0, 1, -1$ for which the two elliptic curves $Y^2 = X(X-1)(X-t)$ and $Y^2 = X(X-1)(X+t)$ both have complex multiplication.

One can ask about CM-points on general straight lines $Ax + By + C = 0$. One has to exclude from consideration the *special straight lines*: $x = j(\tau)$, $y = j(\tau)$ (where $j(\tau)$ is a singular modulus) and $x = y$, the latter being nothing else than the modular

[1] The same result was independently obtained in an earlier version of [Bilu et al. 2013] but did not appear in the final version.

curve $Y_0(1)$ (the modular polynomial Φ_1 is $X - Y$). According to the theorem of André, these are the only straight lines containing infinitely many CM-points.

In [Allombert et al. 2015] all CM-points lying on nonspecial straight lines defined over $\mathbb{Q}$ are listed. More generally, Kühne [2013, p. 5] remarks that, given a positive integer ν, at most finitely many CM-points belong to the union of all nonspecial straight lines defined over a number field of degree ν; moreover, for a fixed ν all these points can, in principle, be listed explicitly, though the implied calculation does not seem to be feasible.

Here we take a different point of view: instead of restricting the degree of field of definition, we study the (nonspecial) straight lines passing through at least three CM-points.

Such lines do exist [Allombert et al. 2015, Remark 5.3]: since

$$\det \begin{bmatrix} 1728 & -884736000 \\ 287496 & -147197952000 \end{bmatrix} = 0,$$

the three points $(0, 0)$, $(1728, 287496)$ and $(-884736000, -147197952000)$ belong to the same straight line $1331x = 8y$, and just as well for the points $(0, 0)$, $(1728, -884736000)$ and $(287496, -147197952000)$ on $512000x = -y$. Here

$$j\left(\frac{-1+\sqrt{-3}}{2}\right) = 0, \qquad j(\sqrt{-1}) = 1728, \qquad j(2\sqrt{-1}) = 287496,$$
$$j\left(\frac{-1+\sqrt{-43}}{2}\right) = -884736000, \qquad j\left(\frac{-1+\sqrt{-67}}{2}\right) = -147197952000.$$

Call an (unordered) triple $\{P_1, P_2, P_3\}$ of CM-points *collinear* if P_1, P_2, P_3 are pairwise distinct and belong to a nonspecial straight line.

In this paper we prove the following:

Theorem 1.1. *There exist at most finitely many collinear triples of CM-points.*

In particular, there exist at most finitely many nonspecial straight lines passing through three or more CM-points. This latter consequence looks formally weaker than Theorem 1.1, but in fact it is equivalent to it, due to the theorem of André.

Remark 1.2. The referee drew our attention to the phenomenon of *automatic uniformity*, discovered by Scanlon [2004]. Combining Theorem 4.2 from [Scanlon 2004] with Pila's Theorem 2.4 stated in the next section, one obtains the following "uniform" version of the theorem of André: there is a (noneffective) uniform upper bound c_d on the number of CM-points on an arbitrary nonspecial curve of geometric degree d (with an arbitrary field of definition). For every d, it is a widely open question what the optimal c_d actually is; moreover, even obtaining an effective upper bound for c_d seems to be quite difficult. It might be an easier question to ask for an

optimal bound c_d^* such that *all but finitely many* nonspecial curves of degree d contain at most c_d^* special points. In this language our Theorem 1.1 simply asserts that $c_1^* = 2$.

The idea of the proof of Theorem 1.1 is simple. Three points (x_i, y_i) lie on a line if and only if

$$\begin{vmatrix} 1 & 1 & 1 \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{vmatrix} = 0. \tag{1-1}$$

This defines a variety in $\mathbb{C}^6$ to which we can apply Pila's André–Oort result. This guarantees finiteness outside the special subvarieties of positive dimension. One easily detects "obvious" positive-dimensional special subvarieties: they correspond to the line being special in two dimensions or the three points not being distinct. The main difficulty is showing that there are no other positive-dimensional special subvarieties: this is the content of the "main lemma", whose proof occupies the overwhelming part of the article. Along the way we have to solve some auxiliary problems not only of André–Oort type but also of "mixed type" involving roots of unity.

It could be mentioned that, while the main lemma is completely effective, Theorem 1.1 is not because its deduction from the main lemma relies on Pila's Theorem 2.4, which is noneffective.

For analogous Diophantine assertions about lines proved also using "determinant varieties", the reader can consult the articles of Evertse, Győry, Stewart and Tijdeman [Evertse et al. 1988] about S-units or of Schlickewei and Wirsing [1997] about heights. In these papers, one is actually in the multiplicative group $\mathbb{G}_m^2$ and the appropriate special varieties are much easier to describe.

Plan of the article. In Section 2 we recall the general notion of special variety and state the already mentioned theorem of Pila, proving the André–Oort conjecture for subvarieties of $\mathbb{C}^n$.

In Section 3 we present the main lemma, which lists all maximal positive-dimensional special subvarieties of the "determinant variety" defined by (1-1), and we deduce Theorem 1.1 from the theorem of Pila and the main lemma.

In Sections 4, 5, 6 and 7 we obtain various auxiliary results used in the sequel. The proof of the main lemma occupies Sections 8 to 12. In Section 8 we collect some preliminary material and show how the proof of the main lemma splits into four cases. These cases are treated in Sections 9 to 12.

2. Special varieties and the theorem of Pila

We recall the definition of special varieties from [Pila 2011]. The referee pointed out that this is not the definition used in the standard formulation of the André–Oort

conjecture, and some work is required to show that the two are equivalent. However, this presents no issues for our purposes since the main result that we need, Pila's Theorem 2.4, proved in [Pila 2011], is stated therein in terms of this definition.

To begin, we define sets M in $\mathbb{C}^m$ (where $m \geq 1$) as follows. If $m = 1$, then $M = \mathbb{C}$, while if $m \geq 2$, then M is given by modular equations

$$\Phi_{N(i)}(x_1, x_i) = 0 \quad (i = 2, \ldots, m). \tag{2-1}$$

More generally for $\mathbb{C}^n$ (where $n \geq 1$), one takes a partition $n = l_0 + m_1 + \cdots + m_d$ (where $d \geq 0$) with $l_0 \geq 0$ and with $m_1 \geq 1, \ldots, m_d \geq 1$ (when $d \geq 1$) and defines sets K in $\mathbb{C}^n = \mathbb{C}^{l_0} \times \mathbb{C}^{m_1} \times \cdots \times \mathbb{C}^{m_d}$ as $L_0 \times M_1 \times \cdots \times M_d$, where L_0 (if $l_0 \geq 1$) is a single point whose coordinates are singular moduli and $M_1, \ldots, M_d$ (if $d \geq 1$) are as M above. Then any irreducible component $\widetilde{K}$ of K, which necessarily has the form

$$\widetilde{K} = L_0 \times \widetilde{M}_1 \times \cdots \times \widetilde{M}_d \tag{2-2}$$

with irreducible components $\widetilde{M}_1, \ldots, \widetilde{M}_d$ of $M_1, \ldots, M_d$, is an example of a special variety in the sense of Pila; and one gets all examples by permuting the coordinates. The dimension is d.

When $n = 2$ and $d = 1$, this agrees with the notion of special curve introduced in Section 1 because the polynomials Φ_N are irreducible.

The following property of special varieties is certainly known, but we could not find a suitable reference.

Proposition 2.1. *Let $0 \leq e \leq d \leq n$. Then every special variety of dimension d contains a Zariski-dense union of special varieties of dimension e.*

Proof. If $d = 0$, there is nothing to prove. Otherwise, by induction, it suffices to treat the case $e = d - 1$, with the special variety (2-2).

If $m_1 = 1$, then $\widetilde{M}_1 = \mathbb{C}$ and for each singular modulus ξ the variety $L_0 \times \{\xi\} \times \widetilde{M}_2 \times \cdots \times \widetilde{M}_d$ is special of dimension $d - 1$. As there are infinitely many singular moduli, the union is Zariski-dense in $\widetilde{K}$.

If $m_1 \geq 2$ (call it m), we note from (2-1) that x_1 is nonconstant on $\widetilde{M}_1$. Thus, the corresponding projection of $\widetilde{M}_1$ to $\mathbb{C}$ is dominant. We can therefore find infinitely many singular moduli ξ_1 for which some $(\xi_1, \xi_2, \ldots, \xi_m)$ lies in $\widetilde{M}_1$. As $\Phi_{N(i)}(\xi_1, \xi_i) = 0$ for $i = 2, \ldots, m$, it is clear that $\xi_2, \ldots, \xi_m$ are also singular moduli, and now the corresponding

$$L_0 \times \{(\xi_1, \xi_2, \ldots, \xi_m)\} \times \widetilde{M}_2 \times \cdots \times \widetilde{M}_d$$

do the trick. □

Special points are exactly those of the form $(\xi_1, \ldots, \xi_n)$, where each ξ_i is a singular modulus. To characterize the special curves in a similar way, it will be

convenient to use the language of "j-maps". A map $f : \mathbb{H} \to \mathbb{C}$ will be called a *j-map* if either $f(z) = j(\gamma z)$ for some $\gamma \in \mathrm{GL}_2^+(\mathbb{Q})$ (a *nonconstant j-map*) or $f(z) = j(\tau)$ with $\tau \in \mathbb{H}$ algebraic of degree 2 (a *constant j-map*). Here $\mathrm{GL}_2^+(\mathbb{Q})$ is the subgroup of $\mathrm{GL}_2(\mathbb{Q})$ consisting of matrices with positive determinants. We define a *j-set* to be of the form $\{(f_1(z), \dots, f_n(z)) : z \in \mathbb{H}\}$, where each f_k is a j-map and at least one of them is nonconstant.

Remark 2.2. It is worth noting that every j-map is $\Gamma(N)$-automorphic[2] for some positive integer N. This is trivially true for constant j-maps, and a nonconstant j-map $f = j \circ \gamma$ is $\gamma^{-1}\Gamma(1)\gamma$-automorphic. So it remains to note that $\gamma^{-1}\Gamma(1)\gamma$ contains $\Gamma(N)$ for a suitable N. Indeed, write $A \in \Gamma(N)$ as $I + NB$, where I is the identity matrix and B is a matrix with entries in $\mathbb{Z}$. Then the matrix $\gamma A \gamma^{-1} = I + N\gamma B\gamma^{-1}$ has entries in $\mathbb{Z}$ if N is divisible by the product of the denominators of the entries of γ and γ^{-1}.

It seems to be known (and even used in several places) that every special curve is a j-set and that the converse is also true. As we could not find a convincing reference, we provide here an argument. We thank the referee for many explanations on this topic.

Proposition 2.3. (1) *Any j-set is a Zariski-closed irreducible algebraic subset of $\mathbb{C}^n$.*

(2) *A subset of $\mathbb{C}^n$ is a j-set if and only if it is a special curve.*

Proof. In the proof of Part (1), we may restrict to the case when all $f_1, \dots, f_n$ are nonconstant j-maps. Denote by $Z \subset \mathbb{C}^n$ the j-set defined by these maps. According to Remark 2.2, the maps $f_1, \dots, f_n$ are $\Gamma(N)$-automorphic for some positive integer N. Hence, each f_i induces a regular map, also denoted by f_i, of the affine modular curve $Y(N) = \Gamma(N) \setminus \mathbb{H}$ to $\mathbb{C}$, and our Z is the image of the map $(f_1, \dots, f_n) : Y(N) \to \mathbb{C}^n$.

Furthermore, each f_i extends to a regular map $\bar{f}_i : X(N) \to \mathbb{P}^1(\mathbb{C})$ of projective curves, where $X(N)$ is the standard compactification of $Y(N)$, as explained, for instance, in [Diamond and Shurman 2005, §2.4]. The image $\bar{Z}$ of the map $(\bar{f}_1, \dots, \bar{f}_n) : X(N) \to \mathbb{P}^1(\mathbb{C})^n$ is Zariski-closed in $\mathbb{P}^1(\mathbb{C})^n$ and irreducible (being the image of an irreducible projective curve under a regular map). But for $x \in X(N)$, we have $\bar{f}_i(x) = \infty$ if and only if $x \in X(N) \setminus Y(N)$ (we write $\mathbb{P}^1(\mathbb{C}) = \mathbb{C} \cup \{\infty\}$ in the obvious sense). Hence, $Z = \bar{Z} \cap \mathbb{C}^n$, which shows that Z is Zariski-closed in $\mathbb{C}^n$ and irreducible. This proves Part (1).

[2]Recall that $\Gamma(N)$ is the kernel of the mod N reduction map $\mathrm{SL}_2(\mathbb{Z}) \to \mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})$, and "the function f is $\Gamma(N)$-automorphic" means $f \circ \eta = f$ for any $\eta \in \Gamma(N)$.

Part (2) is an easy consequence of Part (1). If f and g are two nonconstant j-maps, then there exists N such that $\Phi_N(f, g) = 0$. It follows that, up to coordinate permutations, any j-set is contained in $L_0 \times M$, where L_0 is a point whose coordinates are singular moduli and $M \subseteq \mathbb{C}^m$ is defined as in (2-1). Since our j-set is irreducible and Zariski-closed, it must be an irreducible component of $L_0 \times M$, that is, a special curve. In particular, a j-set is an irreducible one-dimensional algebraic set defined over $\overline{\mathbb{Q}}$.

Conversely, every special curve has (up to coordinate permutations) the shape $L_0 \times \widetilde{M}$, where $\widetilde{M}$ is an irreducible component of a set $M \subset \mathbb{C}^m$ defined as in (2-1). Recall that two complex numbers x, y satisfy $\Phi_N(x, y) = 0$ if and only if x and y are j-invariants of two elliptic curves linked by a cyclic N-isogeny. Now let $(\xi_1, \dots, \xi_m)$ be a transcendental point[3] of $\widetilde{M}$. Then the numbers $\xi_1, \dots, \xi_m$ are j-invariants of isogenous elliptic curves. Hence, if we write $\xi_1 = j(z)$ with some $z \in \mathbb{H}$, then there exist $\gamma_2, \dots, \gamma_m \in \mathrm{GL}_2^+(\mathbb{Q})$ such that $\xi_i = j(\gamma_i z)$ for $i = 2, \dots, m$.

Thus, $\widetilde{M}$ shares a transcendental point with the j-set defined by the j-maps $j, j \circ \gamma_2, \dots, j \circ \gamma_m$. Since both are Zariski-closed irreducible one-dimensional algebraic sets defined over $\overline{\mathbb{Q}}$, they must coincide. □

A similar "parametric" description can be given for higher dimensional special varieties. We do not go into this because we will not need it.

Pila [2011] generalized the theorem of André by proving the following:

Theorem 2.4 (Pila). *An algebraic set in $\mathbb{C}^n$ contains at most finitely many maximal special subvarieties.*

"Maximal" is understood here in the set-theoretic sense: let V be an algebraic set in $\mathbb{C}^n$ and $M \subseteq V$ a special variety; we call M a *maximal special subvariety* of V if for any special variety M' such that $M \subseteq M' \subseteq V$ we have $M = M'$.

If an algebraic curve is not special, then its only special subvarieties are special points, and we recover the theorem of André.

3. Main lemma and proof of Theorem 1.1

Theorem 1.1 is an easy consequence of Pila's Theorem 2.4 and the following lemma.

Lemma 3.1 (main lemma). *Let $f_1, f_2, f_3, g_1, g_2, g_3$ be j-maps, not all constant. Assume that the determinant*

$$\det \begin{bmatrix} 1 & 1 & 1 \\ f_1 & f_2 & f_3 \\ g_1 & g_2 & g_3 \end{bmatrix} \tag{3-1}$$

is identically 0. Then at least one of the following holds:

[3] "Transcendental" means here that the coordinates of this point are not all algebraic over $\mathbb{Q}$.

- $f_1 = f_2 = f_3$,
- $g_1 = g_2 = g_3$,
- *for some distinct* $k, \ell \in \{1, 2, 3\}$ *we have* $f_k = f_\ell$ *and* $g_k = g_\ell$,
- $f_k = g_k$ *for* $k = 1, 2, 3$.

In this section we prove Theorem 1.1 assuming the validity of the main lemma. Lemma 3.1 itself will be proved in the subsequent sections.

Consider the algebraic set in $\mathbb{C}^6$ consisting of the points $(x_1, x_2, x_3, y_1, y_2, y_3)$ satisfying

$$\begin{vmatrix} 1 & 1 & 1 \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{vmatrix} = 0. \tag{3-2}$$

Then Lemma 3.1 has the following consequence.

Corollary 3.2. *The algebraic set (3-2) has exactly six maximal special subvarieties of positive dimension*:

- *the subvariety* R_x, *defined in* $\mathbb{C}^6$ *by* $x_1 = x_2 = x_3$,
- *the subvariety* R_y, *defined in* $\mathbb{C}^6$ *by* $y_1 = y_2 = y_3$,
- *the three subvarieties* $S_{k,\ell}$, *defined in* $\mathbb{C}^6$ *by* $x_k = x_\ell$ *and* $y_k = y_\ell$, *where* $k, \ell \in \{1, 2, 3\}$ *are distinct, and*
- *the subvariety* T, *defined in* $\mathbb{C}^6$ *by* $x_k = y_k$ *for* $k = 1, 2, 3$.

Proof. Let $\widetilde{K}$ be a special variety in (3-2) of positive dimension. By Proposition 2.1 it contains a Zariski-dense union of special curves. By Proposition 2.3 each such curve is a j-set. By the main lemma, each j-set is contained in one of the subvarieties above. The latter are clearly irreducible and also special; for example with R_x we have $n = 6$, $d = 4$, and the partition with

$$l_0 = 0, \qquad m_1 = 3, \qquad m_2 = m_3 = m_4 = 1.$$

Taking closures we see that $\widetilde{K}$ itself is also contained in one of them. □

Now we are ready to prove Theorem 1.1. Let

$$P_k = (x_k, y_k) \quad (k = 1, 2, 3)$$

be three special points forming a collinear triple. Then the point $Q = (x_1, x_2, x_3, y_1, y_2, y_3)$ belongs to the algebraic set (3-2). Moreover, since our points are pairwise distinct, Q does not belong to any of $S_{k,\ell}$, and since the straight line passing through our points is not special, Q does not belong to any of R_x, R_y, T.

This shows that $\{Q\}$ is a zero-dimensional maximal special subvariety of the algebraic set (3-2), and we complete the proof by applying Theorem 2.4. □

The main lemma will be proved in Sections 8–12, after some preparations made in Sections 4–7.

4. Roots of unity

In this section we collect some facts about roots of unity used in the proof of the main lemma.

Lemma 4.1. *Let α be a sum of k roots of unity and N a nonzero integer. Assume that $N \mid \alpha$ (in the ring of algebraic integers). Then either $\alpha = 0$ or $k \geq |N|$.*

Proof. Assume $\alpha \neq 0$, and write $\alpha = N\beta$, where β is a nonzero algebraic integer. Then there exists an embedding $\mathbb{Q}(\alpha) \xrightarrow{\sigma} \mathbb{C}$ such that $|\beta^\sigma| \geq 1$. It follows that $|N| \leq |\alpha^\sigma|$. But since α is a sum of k roots of unity, we have $|\alpha^\sigma| \leq k$. □

Lemma 4.2. *Let a, b be nonzero rational numbers and η, θ roots of unity. Assume that $\alpha = a\eta + b\theta$ is of degree 1 or 2 over $\mathbb{Q}$. Then $\mathbb{Q}(\alpha)$ is one of the fields $\mathbb{Q}, \mathbb{Q}(i), \mathbb{Q}(\sqrt{-2}), \mathbb{Q}(\sqrt{-3}), \mathbb{Q}(\sqrt{2}), \mathbb{Q}(\sqrt{3}), \mathbb{Q}(\sqrt{5})$, and after a possible swapping of $a\eta$ and $b\theta$, and possible replacing of (a, η) by $(-a, -\eta)$ and/or (b, θ) by $(-b, -\theta)$, we have the following:*

(1) *If $\mathbb{Q}(\alpha) = \mathbb{Q}$, then*
   (a) *either both η and θ are ± 1 or*
   (b) *η is a primitive cubic root of unity, $\theta = \eta^{-1}$, and $a = b$, or*
   (c) *$\theta = -\eta$ and $a = b$.*

(2) *If $\mathbb{Q}(\alpha) = \mathbb{Q}(i)$, then*
   (a) *either $\eta = i$ and $\theta \in \{1, i\}$ or*
   (b) *η is a primitive 12th root of unity, $\theta = -\eta^{-1}$, and $a = b$.*

(3) *If $\mathbb{Q}(\alpha) = \mathbb{Q}(\sqrt{-3})$, then η is a primitive cubic root of unity, and θ is a cubic root of unity (primitive or not).*

(4) *If $\mathbb{Q}(\alpha) = \mathbb{Q}(\sqrt{-2})$, then η is a primitive 8th root of unity, $\theta = -\eta^{-1}$, and $a = b$.*

(5) *If $\mathbb{Q}(\alpha) = \mathbb{Q}(\sqrt{2})$, then η is a primitive 8th root of unity, $\theta = \eta^{-1}$, and $a = b$.*

(6) *If $\mathbb{Q}(\alpha) = \mathbb{Q}(\sqrt{3})$, then η is a primitive 12th root of unity, and*
   (a) *either $\theta = \eta^{-1}$ and $a = b$ or*
   (b) *$\theta = -\eta^3 (= \pm i)$ and $a = 2b$.*

(7) *If $\mathbb{Q}(\alpha) = \mathbb{Q}(\sqrt{5})$, then η is a primitive 5th root of unity, $\theta = \eta^{-1}$, and $a = b$.*

Proof. Without loss of generality, we may assume that a and b are coprime integers. Let N be the order of the multiplicative group generated by η and θ, and $L = \mathbb{Q}(\eta, \theta)$; then $[L : \mathbb{Q}] = \varphi(N)$, where φ is Euler's totient function.

If $\varphi(N) \leq 2$, then $N \in \{1, 2, 3, 4, 6\}$, and we have one of the options (1), (2a), or (3). If $\alpha = 0$, then we have option (1c).

From now on we assume that $\varphi(N) > 2$ and $\alpha \neq 0$. Since $\varphi(N) > 2$, there exists $\sigma \in \mathrm{Gal}(L/\mathbb{Q})$ such that $(\eta^\sigma, \theta^\sigma) \neq (\eta, \theta)$, but $\alpha^\sigma = \alpha$. We obtain

$$a(\eta - \eta^\sigma) = b(\theta^\sigma - \theta). \tag{4-1}$$

By our choice of σ, both sides of (4-1) are nonzero. Since a and b are coprime integers, we have $a \mid (\theta^\sigma - \theta)$, whence $|a| \leq 2$ by Lemma 4.1. Similarly, $|b| \leq 2$. It follows that $(a, b) \in \{(\pm 1, \pm 1), (\pm 1, \pm 2), (\pm 2, \pm 1)\}$. Swapping (if necessary) $a\eta$ and $b\theta$, and replacing (if necessary) (a, η) by $(-a, -\eta)$ and/or (b, θ) by $(-b, -\theta)$, we may assume that $a \in \{1, 2\}$ and $b = 1$. The rest of the proof splits into two cases.

The case $a = 2$ and $b = 1$. In this case (4-1) becomes $2(\eta - \eta^\sigma) = \theta^\sigma - \theta$. We must have $\theta^\sigma = -\theta$; otherwise all the conjugates of the nonzero algebraic integer $(\theta^\sigma - \theta)/2$ would be of absolute value strictly smaller than 1. Thus, we obtain $\eta - \eta^\sigma + \theta = 0$. Three roots of unity may sum up to 0 only if they are proportional to $(1, \zeta_3, \zeta_3^{-1})$, where ζ_3 is a primitive cubic of unity. We obtain $\theta/\eta = \zeta_3^{-1}$, and $\eta = \alpha(a + b\zeta_3^{-1})^{-1}$ is of degree at most 4 over $\mathbb{Q}$. Since $\theta = \eta^\sigma - \eta \in \mathbb{Q}(\eta)$, we obtain $L = \mathbb{Q}(\eta)$; in particular, η is a primitive N-th root of unity.

Thus, $\varphi(N) = [\mathbb{Q}(\eta) : \mathbb{Q}] \leq 4$, and in fact $\varphi(N) = 4$ because $\varphi(N) > 2$. Since $-\eta^\sigma/\eta = \zeta_3$, we must have $3 \mid N$. Together with $\varphi(N) = 4$, this implies that $N = 12$ and η is a primitive 12th root of unity. Hence, we have the option (6b).

The case $a = b = 1$. In this case $\eta - \eta^\sigma + \theta - \theta^\sigma = 0$. Four roots of unity may sum up to 0 only if two of them sum up to 0 (and the other two sum up to 0 as well). Since $\eta \neq \eta^\sigma$ and $\eta \neq -\theta$ (because $\alpha \neq 0$), we have $\eta = \theta^\sigma$ and $\eta^\sigma = \theta$. This implies that $L = \mathbb{Q}(\eta) = \mathbb{Q}(\theta)$, both η and θ are primitive N-th roots of unity, and $\sigma^2 = 1$.

We claim that the subgroup $H = \{1, \sigma\}$ is the stabilizer of $\mathbb{Q}(\alpha)$ in $G = \mathrm{Gal}(L/\mathbb{Q})$. Thus, let $\varsigma \in G$ satisfy $\alpha^\varsigma = \alpha$. Since $\eta + \eta^\sigma - \eta^\varsigma - \eta^{\sigma\varsigma} = 0$ and $\eta + \eta^\sigma \neq 0$, we must have either $\eta = \eta^\varsigma$ or $\eta = \eta^{\sigma\varsigma}$. Since $L = \mathbb{Q}(\eta)$, in the first case we have $\varsigma = 1$ and in the second case $\varsigma = \sigma^{-1} = \sigma$.

Thus, H is the stabilizer of $\mathbb{Q}(\alpha)$. Since $|H| = 2$ and $[G : H] = [\mathbb{Q}(\alpha) : \mathbb{Q}] = 2$, we obtain $\varphi(N) = |G| = 4$, which implies that $N \in \{5, 8, 10, 12\}$.

Now if $N = 5$, then we have option (7). If $N = 10$, then replacing (a, η) by $(-a, -\eta)$ and (b, θ) by $(-b, -\theta)$, we obtain option (7) as well. If $N = 8$, then we have one of the options (4) or (5). Finally, if $N = 12$, then we have one of the options (2b) or (6a). □

5. Singular moduli

In this section we collect miscellaneous properties of singular moduli used in the sequel. We start by recalling the notion of the *discriminant* of a singular modulus. Let $\tau \in \mathbb{H}$ be algebraic of degree 2; the endomorphism ring of the lattice $\mathbb{Z}\tau + \mathbb{Z}$ is an

Δ	-3	-4	-7	-8	-11	-12	-16	-19	-27
j	0	1728	-3375	8000	-32768	54000	287496	-884736	-12288000
Δ	-28		-43		-67		-163		
j	16581375		-884736000		-147197952000		-262537412640768000		

Table 1. Discriminants Δ with $h(\Delta) = 1$ and the corresponding singular moduli.

order in the imaginary quadratic field $\mathbb{Q}(\tau)$; the discriminant $\Delta = \Delta_\tau$ of this order will be called the *discriminant* of the singular modulus $j(\tau)$. This discriminant is a negative integer satisfying $\Delta \equiv 0, 1 \bmod 4$.

It is well-known (see, for instance, [Cox 1989, §11]) that

- any singular modulus of discriminant Δ is an algebraic integer of degree equal to the class number of Δ, denoted $h(\Delta)$, and
- the singular moduli of discriminant Δ are all conjugate over $\mathbb{Q}$; moreover, they form a complete set of $\mathbb{Q}$-conjugates.

A full description of singular moduli of given discriminant Δ is well-known as well. Denote by $T = T_\Delta$ the set of triples of integers (a, b, c) such that

$$\gcd(a, b, c) = 1, \quad \Delta = b^2 - 4ac, \quad \text{either } -a < b \le a < c \text{ or } 0 \le b \le a = c.$$

Then the map

$$(a, b, c) \mapsto j\left(\frac{b + \sqrt{\Delta}}{2a}\right) \tag{5-1}$$

defines a bijection from T_Δ onto the set of singular moduli of discriminant Δ. In particular, $h(\Delta) = |T_\Delta|$. The proof of this is a compilation of several classical facts, some of which go back to Gauss; see, for instance, [Bilu et al. 2016, §2.2] and the references therein.

It is crucial for us that the set T_Δ has only one triple (a, b, c) with $a = 1$. The corresponding singular modulus will be called the *principal* singular modulus of discriminant Δ. Note that the principal singular modulus is a real number; in particular,

$$\text{any singular modulus has a real } \mathbb{Q}\text{-conjugate.} \tag{5-2}$$

There exist exactly 13 discriminants Δ with $h(\Delta) = 1$. The corresponding singular moduli (and only they) are rational integers. The full list of the 13 rational singular moduli is well-known and reproduced in Table 1.

Finally, we use the inequality

$$|\,|j(\tau)| - e^{2\pi \operatorname{Im}\tau}\,| \le 2079, \tag{5-3}$$

which holds for every $\tau \in \mathbb{H}$ satisfying $\operatorname{Im}\tau \geq \sqrt{3}/2$ [Bilu et al. 2013, Lemma 1]. In particular, if $(a, b, c) \in T_\Delta$, then the number

$$\tau(a, b, c) = \frac{b + \sqrt{\Delta}}{2a}$$

satisfies $\operatorname{Im}\tau(a, b, c) \geq \sqrt{3}/2$ [Bilu et al. 2016, p. 403, (8)]. Hence, (5-3) applies with $\tau = \tau(a, b, c)$.

All the facts listed above will be repeatedly used in this section, sometimes without a special reference.

Lemma 5.1. *Let x be a singular modulus, and let x' be the principal singular modulus of the same discriminant. Then either $x = x'$ or $|x'| > |x| + 180000$.*

Proof. Let Δ be the common discriminant of x and x'. We may assume that $|\Delta| \geq 15$; otherwise, $h(\Delta) = 1$ and there is nothing to prove. We assume that $x \neq x'$ and will use (5-3) to estimate $|x|$ from above and $|x'|$ from below.

We have $x = j(\tau)$ and $x' = j(\tau')$, where $\tau = \tau(a, b, c)$ and $\tau' = \tau(a', b', c')$ for some $(a, b, c), (a', b', c') \in T_\Delta$. Since x' is principal, and x is not, we have $a' = 1$ and $a \geq 2$. Hence,

$$\operatorname{Im}\tau' = \pi|\Delta|^{1/2}, \qquad \operatorname{Im}\tau = \frac{\pi|\Delta|^{1/2}}{a} \leq \frac{\pi|\Delta|^{1/2}}{2}.$$

We obtain

$$|x'| \geq e^{\pi|\Delta|^{1/2}} - 2079, \qquad |x| \leq e^{\pi|\Delta|^{1/2}/2} + 2079,$$

which implies

$$|x'| - |x| \geq e^{\pi|\Delta|^{1/2}} - e^{\pi|\Delta|^{1/2}/2} - 4158 \geq e^{\pi\sqrt{15}} - e^{\pi\sqrt{15}/2} - 4158 > 180000,$$

as wanted. □

Lemma 5.2. *Let x, y be singular moduli, and let $a, b \in \mathbb{Z}$ be such that $|a|, |b| \leq 90000$. Assume that $y \neq b$ and that $(x - a)/(y - b)$ is a root of unity. Then either $x = y$ or $x, y \in \mathbb{Z}$. In particular, if x/y is a root of unity (with $y \neq 0$) or if $(x - 744)/(y - 744)$ is a root of unity, then $x = y$.*

Proof. Let x' and y' be the principal singular moduli of the same discriminants as x and y. We may assume that $|x'| \geq |y'|$. We may further assume, by conjugating, that $x = x'$. Then $y = y'$ as well since otherwise $|y| < |y'| - 180000$ by Lemma 5.1, and we obtain

$$|y| + 90000 \geq |y - b| = |x - a| = |x' - a| \geq |x'| - 90000 \geq |y'| - 90000 > |y| + 90000,$$

a contradiction. Thus, both x and y are principal singular moduli. In particular, both are real, which implies $x - a = \pm(y - b)$.

Now Theorem 1.2 of [Allombert et al. 2015] implies one of the following options:

(1) $x = y$ and $a = b$,

(2) $x, y \in \mathbb{Z}$, or

(3) x and y are distinct and of degree 2 over $\mathbb{Q}$.

We have to rule out option (3). Thus, assume that to be the case and let $f(T) = T^2 + AT + C$ and $g(T) = T^2 + BT + D$ be the $\mathbb{Q}$-minimal polynomials of x and y. Since x and y are both principal and distinct, they are not $\mathbb{Q}$-conjugate, which means that the polynomials F and G are distinct. We have either $x + y = a + b$ or $x - y = a - b$. Taking $\mathbb{Q}$-traces, we obtain $A + B = 2(a + b)$ or $A - B = 2(a - b)$. In particular, we have either $|A + B| \le 360000$ or $|A - B| \le 360000$.

However, our F and G are among the 29 Hilbert class polynomials associated to the imaginary quadratic orders of class number 2. The full list of such polynomials can be found in Table 2 of [Bilu et al. 2016]. A quick inspection of this table shows that, if A and B are middle coefficients of two distinct polynomials from this table, then $|A + B| > 360000$ and $|A - B| > 360000$. Hence, option (3) is impossible. This proves the first statement of the lemma.

In the special cases $a = b = 0$ or $a = b = 744$, we must have either $x = y$ or

$$x, y \in \mathbb{Z}, \quad x \ne y, \quad x + y \in \{0, 1488\}. \tag{5-4}$$

Inspecting Table 1, we find out that (5-4) is impossible. The lemma is proved. □

Lemma 5.3. *Let x and y be distinct principal singular moduli. Then $||x| - |y|| > 1600$.*

Proof. Denote by Δ_x and Δ_y the discriminants of x and y, respectively. We will assume that $|\Delta_x| > |\Delta_y|$. If $|\Delta_x| \le 12$, then $h(\Delta_x) = 1$, and the statement follows by inspection of Table 1. And if $|\Delta_x| \ge 15$, then

$$\begin{aligned} |x| - |y| &\ge (e^{\pi|\Delta_x|^{1/2}} - 2079) - (e^{\pi|\Delta_y|^{1/2}} + 2079) \\ &\ge e^{\pi|\Delta_x|^{1/2}} - e^{\pi|\Delta_x - 1|^{1/2}} - 4158 \\ &\ge e^{\pi\sqrt{15}} - e^{\pi\sqrt{14}} - 4158 \\ &> 60000, \end{aligned}$$

which is much stronger than needed. The lemma is proved. □

Lemma 5.4. *Let x be a singular modulus, and assume that the number field $\mathbb{Q}(x)$ is a Galois extension of $\mathbb{Q}$. Then the Galois group of $\mathbb{Q}(x)/\mathbb{Q}$ is 2-elementary, that is, isomorphic to $(\mathbb{Z}/2\mathbb{Z})^k$ for some k.*

Proof. This is well-known; see, for instance, Corollary 3.3 from [Allombert et al. 2015]. □

Lemma 5.5. *Let x, y be singular moduli and ε, η roots of unity. Then $\varepsilon(x - 744) + \eta(y - 744)$ is not a root of unity.*

Proof. We will assume that

$$\varepsilon(x-744)+\eta(y-744)=1$$

and derive a contradiction. We clearly have

$$||y|-|x||\le 1489. \tag{5-5}$$

We follow the same strategy as in the proof of Lemma 5.2. We denote by x' and y' the principal moduli of the same discriminants as x and y, respectively, and we may assume that $|x'|\ge|y'|$ and $x=x'$. We claim that $y=y'$ as well. Indeed, if $y\ne y'$, then Lemma 5.1 implies that

$$|y|+1489\ge|x|=|x'|\ge|y'|>|y|+180000,$$

a contradiction.

Thus, we may assume that both x and y are principal singular moduli. Lemma 5.3 and inequality (5-5) imply that $x=y$. Thus,

$$(\varepsilon+\eta)(x-744)=1.$$

In particular $0\ne\varepsilon+\eta\in\mathbb{R}$, which implies $\eta=\varepsilon^{-1}$.

Lemma 5.4 implies that the Galois group of the number field $\mathbb{Q}(x)=\mathbb{Q}(\varepsilon+\varepsilon^{-1})$ is 2-elementary. Since $\mathbb{Q}(\varepsilon+\varepsilon^{-1})$ is a subfield of degree at most 2 in $\mathbb{Q}(\varepsilon)$, the Galois group of $\mathbb{Q}(\varepsilon)/\mathbb{Q}$ is either 2-elementary or $\mathbb{Z}/4\mathbb{Z}$ times a 2-elementary group. But this group is $(\mathbb{Z}/n\mathbb{Z})^\times$, where n is the order of the root of unity ε. Using the well-known structure of the multiplicative group $(\mathbb{Z}/n\mathbb{Z})^\times$ (see, for instance, [Ireland and Rosen 1990, Theorem 3 in §4.1]), one easily finds out that any integer n with the property "the group $(\mathbb{Z}/n\mathbb{Z})^\times$ is either 2-elementary or $\mathbb{Z}/4\mathbb{Z}$ times a 2-elementary group" divides either 48 or 120. It follows that $|\varepsilon+\varepsilon^{-1}|\ge 2\sin(\pi/60)$ (recall that $\varepsilon+\varepsilon^{-1}=\varepsilon+\eta\ne 0$). Hence,

$$|x-744|\le\frac{1}{2\sin(\pi/60)}<10.$$

No principal singular modulus satisfies the latter inequality. □

Lemma 5.6. *The numbers* 744, 744 ± 1, 744 ± 2, 744 ± 196884, $744\pm1\pm196884$, $744\pm2\cdot196884$ *are not singular moduli.*

Proof. The proof is just by inspection of Table 1. □

Lemma 5.7. *Let θ be a root of unity. Then $744+\theta$ and $744+196884\theta$ are not singular moduli.*

Proof. If $744+\theta$ or $744+196884\theta$ is a singular modulus, then the cyclotomic field $\mathbb{Q}(\theta)$ has a real embedding by (5-2), which is possible only if $\theta=\pm1$. Now apply Lemma 5.6. □

Lemma 5.8. *Assume that a singular modulus of discriminant Δ is a sum of k roots of unity. Then*

$$|\Delta| \le \pi^{-2}(\log(k+2079))^2.$$

Proof. We may assume that our modulus (denote it by x) is principal and, as in the proof of Lemma 5.1, deduce from this that it satisfies $|x| \ge e^{\pi|\Delta|^{1/2}} - 2079$. On the other hand, since x is a sum of k roots of unity, we have $|x| \le k$, whence the result. □

Lemma 5.9. *Let η, θ be roots of unity, x a singular modulus, and $a, b, c \in \mathbb{Z}$. Assume that*

$$x = a\eta + b\theta + c, \quad a, b \ne 0, \quad |a| + |b| + |c| \le 3400000.$$

Then one of the following options holds:

- *We have $x \in \mathbb{Z}$.*
- *After possible replacing of (a, η) by $(-a, -\eta)$ and/or (b, θ) by $(-b, -\theta)$, we have the following: η is a primitive 5th root of unity, $\theta = \eta^{-1}$, $a = b$, and*

$$(a, c) \in \{(85995, -52515), (-85995, -138510), (565760, 914880), (-565760, 349120)\}. \quad (5\text{-}6)$$

Proof. Let Δ be the discriminant of the singular modulus x. Lemma 5.8 implies that

$$|\Delta| \le \pi^{-2}(\log(3400000+2079))^2 < 22.92. \quad (5\text{-}7)$$

Assume that $x \notin \mathbb{Z}$; then $h(\Delta) > 1$. Among negative quadratic discriminants satisfying (5-7), all but two have class number 1; these two are $\Delta = -15$ and $\Delta = -20$. In both cases $h(\Delta) = 2$ and $\mathbb{Q}(x) = \mathbb{Q}(\sqrt{5})$, so option (7) of Lemma 4.2 applies in both cases. After possible replacing of (a, η) by $(-a, -\eta)$ and/or (b, θ) by $(-b, -\theta)$, we obtain the following: η is a primitive 5th root of unity, $\theta = \eta^{-1}$, and $a = b$, so we have $x = a(\eta + \eta^{-1}) + c$.

The two singular moduli of discriminant $\Delta = -15$ are

$$\frac{-191025 \pm 85995\sqrt{5}}{2} = -\frac{191025}{2} \pm 85995(\tfrac{1}{2} + \eta + \eta^{-1}) = \begin{cases} \text{either} & 85995(\eta + \eta^{-1}) - 52515, \\ \text{or} & -85995(\eta + \eta^{-1}) - 138510, \end{cases}$$

which gives us the first two options in (5-6)

Similarly, the two singular moduli of discriminant $\Delta = -20$ are $632000 \pm 282880\sqrt{5}$, which gives the other two options. □

6. Rational matrices

In this section we obtain some elementary properties of $\mathbb{Q}$-matrices, which will be used in our study of j-maps in Section 7.

Recall that we denote by $\mathrm{GL}_2^+(\mathbb{Q})$ the subgroup of $\mathrm{GL}_2(\mathbb{Q})$ consisting of matrices of positive determinant. Unless the contrary is stated explicitly, in this section *matrix* refers to an element in $\mathrm{GL}_2^+(\mathbb{Q})$. We call two matrices A and A' *equivalent* (denoted $A \sim A'$) if there exists a matrix $B \in \mathrm{SL}_2(\mathbb{Z})$ and a scalar $\lambda \in \mathbb{Q}^\times$ such that $A' = \lambda BA$.

For $a, b \in \mathbb{Q}$ we define $\gcd(a, b)$ as the nonnegative $\delta \in \mathbb{Q}$ such that $a\mathbb{Z} + b\mathbb{Z} = \delta\mathbb{Z}$.

Given a matrix $A = \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]$, we define the *normalized left content* of A by

$$\mathrm{nlc}(A) = \frac{\gcd(a, c)^2}{\det A}.$$

Clearly, $\mathrm{nlc}(A) = \mathrm{nlc}(A')$ if $A \sim A'$.

Proposition 6.1. *Every matrix A is equivalent to an upper-triangular matrix of the form $\left[\begin{smallmatrix} a & b \\ 0 & 1 \end{smallmatrix}\right]$ with $a > 0$, where $a = \mathrm{nlc}(A)$. We have $\left[\begin{smallmatrix} a & b \\ 0 & 1 \end{smallmatrix}\right] \sim \left[\begin{smallmatrix} a' & b' \\ 0 & 1 \end{smallmatrix}\right]$ if and only if $a = a'$ and $b \equiv b' \bmod \mathbb{Z}$.*

Proof. It suffices to show that A is equivalent to an upper-triangular matrix; the rest is easy. Let $\binom{x}{y}$ be the left column of A and $\delta = \gcd(x, y)$. Then x/δ, $y/\delta \in \mathbb{Z}$, and there exist $u, v \in \mathbb{Z}$ such that $ux + vy = \delta$. Multiplying A on the left by the matrix $\left[\begin{smallmatrix} u & v \\ -y/\delta & x/\delta \end{smallmatrix}\right] \in \mathrm{SL}_2(\mathbb{Z})$, we obtain an upper-triangular matrix. □

Proposition 6.2. *Let A_1, A_2 be nonequivalent matrices. Then there exists a matrix B such that $\mathrm{nlc}(A_1B) \neq \mathrm{nlc}(A_2B)$.*

Proof. We may assume that $\mathrm{nlc}(A_1) = \mathrm{nlc}(A_2)$ (otherwise there is nothing to prove). Multiplying on the right by A_1^{-1}, we may assume that $A_1 = \left[\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right]$. We may further assume that $A_2 = \left[\begin{smallmatrix} a & b \\ 0 & 1 \end{smallmatrix}\right]$. Since $a = \mathrm{nlc}(A_2) = \mathrm{nlc}(A_1) = 1$, we have $A_2 = \left[\begin{smallmatrix} 1 & b \\ 0 & 1 \end{smallmatrix}\right]$, where $b \notin \mathbb{Z}$ since $A_2 \not\sim A_1$.

Now $B = \left[\begin{smallmatrix} 1 & 0 \\ -b^{-1} & 1 \end{smallmatrix}\right]$ would do. Indeed,

$$\mathrm{nlc}(A_1B) = \mathrm{nlc}(B) = \gcd(-b^{-1}, 1)^2, \quad \mathrm{nlc}(A_2B) = \mathrm{nlc}\begin{bmatrix} 0 & b \\ -b^{-1} & 1 \end{bmatrix} = b^{-2},$$

and we have to prove that $\gcd(-b^{-1}, 1) \neq |b|^{-1}$. This is equivalent to $\gcd(1, b) \neq 1$, which is true because $b \notin \mathbb{Z}$. □

One may wonder if the same statement holds true for more than two matrices: *given pairwise nonequivalent matrices $A_1, \ldots, A_n$, does there exist a matrix $B \in \mathrm{GL}_2^+(\mathbb{Q})$ such that $\mathrm{nlc}(A_1B), \ldots, \mathrm{nlc}(A_nB)$ are pairwise distinct*? The proof of the main lemma could have been drastically simplified if it were the case. Unfortunately, the answer is "no" already for three matrices, as the following example shows.

Example 6.3. Let

$$A_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \qquad A_2 = \begin{bmatrix} 1 & 1/2 \\ 0 & 1 \end{bmatrix}, \qquad A_3 = \begin{bmatrix} 4 & 0 \\ 0 & 1 \end{bmatrix}.$$

We claim that, for any matrix B, at least two of the numbers

$$\mathrm{nlc}(A_1B), \qquad \mathrm{nlc}(A_2B), \qquad \mathrm{nlc}(A_3B)$$

are equal. Indeed, write $B = \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]$. After multiplying by a suitable scalar, we may assume that $c = 2$. Now

$$\mathrm{nlc}(A_1B) = \frac{\gcd(a,2)^2}{\det B}, \quad \mathrm{nlc}(A_2B) = \frac{\gcd(a+1,2)^2}{\det B}, \quad \mathrm{nlc}(A_3B) = \frac{\gcd(4a,2)^2}{4\det B},$$

and we must show that among the three numbers

$$\gcd(a,2), \qquad \gcd(a+1,2), \qquad \tfrac{1}{2}\gcd(4a,2)$$

there are two equal. And this is indeed the case:

- if $\mathrm{ord}_2(a) > 0$, then $\gcd(a+1,2) = \frac{1}{2}\gcd(4a,2)$,
- if $\mathrm{ord}_2(a) = 0$, then $\gcd(a,2) = \frac{1}{2}\gcd(4a,2)$, and
- if $\mathrm{ord}_2(a) < 0$, then $\gcd(a,2) = \gcd(a+1,2)$.

Still, it is possible to prove something.

Proposition 6.4. *Let A_1, A_2, A_3 be pairwise nonequivalent matrices. Then there exists a matrix B such that among the numbers* $\mathrm{nlc}(A_1B), \mathrm{nlc}(A_2B), \mathrm{nlc}(A_3B)$ *one is strictly bigger than the two others.*

Proof. We may assume that $A_k = \left[\begin{smallmatrix} a_k & * \\ 0 & 1 \end{smallmatrix}\right]$ for $k = 1, 2, 3$. If the numbers a_k are pairwise distinct, then there is nothing to prove. Hence, we may assume that $a_1 = a_2$. Multiplying on the right by A_3^{-1} and afterwards by a suitable diagonal matrix, we may assume that

$$A_1 = \begin{bmatrix} 1 & b_1 \\ 0 & 1 \end{bmatrix}, \qquad A_2 = \begin{bmatrix} 1 & b_2 \\ 0 & 1 \end{bmatrix}, \qquad A_3 = \begin{bmatrix} a^{-1} & 0 \\ 0 & 1 \end{bmatrix},$$

where $a > 0$. Since $A_1 \not\sim A_2$, we have $b_1 \not\equiv b_2 \bmod \mathbb{Z}$, and we may assume $b_1 \notin \mathbb{Z}$.

Set $B = \left[\begin{smallmatrix} 1 & 0 \\ -b_1^{-1} & 1 \end{smallmatrix}\right]$. Then

$$\begin{aligned} \mathrm{nlc}(A_1B) &= b_1^{-2}, \\ \mathrm{nlc}(A_2B) &= \gcd(1 - b_1^{-1}b_2, b_1^{-1})^2, \\ \mathrm{nlc}(A_3B) &= a\gcd(a^{-1}, b_1^{-1})^2. \end{aligned} \tag{6-1}$$

Multiplying numbers (6-1) by ab_1^2, we must show that among the three numbers

$$a, \qquad a\gcd(b_1-b_2,1)^2, \qquad \gcd(b_1,a)^2 \tag{6-2}$$

one is strictly bigger than the others.

If the numbers in (6-2) are pairwise distinct, then there is nothing to prove. Now assume that two of them are equal. Since $b_1 \not\equiv b_2 \bmod \mathbb{Z}$, then $\gcd(b_1-b_2,1)<1$, and in particular, the first two of them are distinct.

Further, the equality $a=\gcd(b_1,a)^2$ is not possible either. Indeed, in this case for any prime number p we would have

$$\operatorname{ord}_p(a) = 2\min\{\operatorname{ord}_p(a), \operatorname{ord}_p(b_1)\},$$

which implies that either $\operatorname{ord}_p(a)=2\operatorname{ord}_p(b_1)>0$ or $\operatorname{ord}_p(b_1)\geq \operatorname{ord}_p(a)=0$. In particular, $\operatorname{ord}_p(b_1)\geq 0$ for any p, contradicting our assumption $b_1\notin\mathbb{Z}$.

Thus, the only possibility is $a\gcd(b_1-b_2,1)^2=\gcd(b_1,a)^2$, and we obtain

$$a > a\gcd(b_1-b_2,1)^2 = \gcd(b_1,a)^2. \qquad \square$$

7. Level, twist, and q-expansion of a j-map

In this section we collect some properties of j-maps used in the sequel.

Given $\gamma,\gamma'\in \mathrm{GL}_2^+(\mathbb{Q})$, we have $j(\gamma z)=j(\gamma' z)$ if and only if the matrices γ and γ' are *equivalent* in the sense of Section 6. Combined with Proposition 6.1, this gives the following:

Proposition 7.1. *Let f be a nonconstant j-map. Then there exist a unique positive number $m\in\mathbb{Q}$ and a unique modulo 1 number $\mu\in\mathbb{Q}$ such that $f(z)=j(mz+\mu)$.*

Note that $m=\mathrm{nlc}(\gamma)$ for any $\gamma\in\mathrm{GL}_2^+(\mathbb{Q})$ such that $f(z)=j(\gamma z)$.

Setting $q=e^{2\pi i z}$ and $\varepsilon=e^{2\pi i\mu}$, the map $f(z)=j(mz+\mu)$ admits the "q-expansion"

$$f(z)=\varepsilon^{-1}q^{-m}+744+196884\varepsilon q^m+21493760\varepsilon^2q^{2m}+o(q^{2m}), \tag{7-1}$$

where here and below we accept the following convention:

- $O(q^\ell)$ means "terms of q-degree ℓ or higher" and
- $o(q^\ell)$ means "terms of q-degree strictly higher than ℓ".

We call m and ε the *level* and the *twist* of the nonconstant j-map f. For a constant j-map, we set its level to be 0 and its twist undefined. The following property will be routinely used, usually without special reference:

two nonconstant j-maps coincide if and only if their levels and twists coincide. (7-2)

We will denote in the sequel $A = 196884$ and $B = 21493760$ so that (7-1) reads

$$f(z) = \varepsilon^{-1}q^{-m} + 744 + A\varepsilon q^m + B\varepsilon^2 q^{2m} + O(q^{2m}). \tag{7-3}$$

The following lemma will play an important role in Section 8.

Lemma 7.2. *Let f_1, f_2, f_3 be pairwise distinct j-maps, not all constant. Then there exists $\gamma \in \mathrm{GL}_2^+(\mathbb{Q})$ such that one of the maps $f_1 \circ \gamma$, $f_2 \circ \gamma$, $f_3 \circ \gamma$ has level strictly bigger than the two others.*

Proof. If only one of the maps f_k is nonconstant, then there is nothing to prove. If exactly two of them, say f_1 and f_2, are nonconstant, then Proposition 6.2 implies the existence of $\gamma \in \mathrm{GL}_2^+(\mathbb{Q})$ such that $f_1 \circ \gamma$ and $f_2 \circ \gamma$ have distinct levels, and we are done. Finally, if all the three are nonconstant, the result follows from Proposition 6.4. □

We conclude this section with a linear-independence property of nonconstant j-maps.

Lemma 7.3. *Let f, g be nonconstant j-maps satisfying a nontrivial linear relation $af + bg + c = 0$, where $(a, b, c) \in \mathbb{C}^3$ and $(a, b, c) \neq (0, 0, 0)$. Then $f = g$ and $a + b = c = 0$.*

Proof. Any two nonconstant j-maps parametrize the modular curve $Y_0(N)$ of a certain level N; in other words, we have $\Phi_N(f, g) = 0$, where $\Phi_N(x, y)$ is the N-th modular polynomial. If we also have $af + bg + c = 0$, then the polynomial $\Phi_N(x, y)$, being irreducible, must divide the linear polynomial $ax + by + c$. It follows that $N = 1$ since $\Phi_1(x, y) = x - y$ is the only modular polynomial of degree 1. □

8. Initializing the proof of the main lemma

In this section we start the proof of the main lemma. Thus, from now on, let $f_1, f_2, f_3, g_1, g_2, g_3$ be j-maps, not all constant and satisfying

$$\begin{vmatrix} 1 & 1 & 1 \\ f_1 & f_2 & f_3 \\ g_1 & g_2 & g_3 \end{vmatrix} = 0. \tag{8-1}$$

This can be rewritten as

$$(f_1 - f_2)(g_2 - g_3) = (f_2 - f_3)(g_1 - g_2). \tag{8-2}$$

If say $f_1 = f_2$, then we find from (8-2) that either $f_2 = f_3$, in which case $f_1 = f_2 = f_3$, or $g_1 = g_2$, in which case $f_1 = f_2$ and $g_1 = g_2$. Hence, we may assume in the sequel that

$$f_1, f_2, f_3 \text{ are pairwise distinct, and so are } g_1, g_2, g_3. \tag{8-3}$$

We will show that under this assumption

$$f_k = g_k \quad (k = 1, 2, 3). \tag{8-4}$$

Let m_k, n_k be the levels of f_k, g_k, respectively, for $k = 1, 2, 3$. If f_k and/or g_k is not constant, we denote the corresponding twists by $\varepsilon_k = e^{2\pi i \mu_k}$ and/or $\eta_k = e^{2\pi i \nu_k}$, respectively.

8A. ***Some relations for the levels.*** Since not all of our six maps are constant, we may assume that the three maps f_k are not all constant. Lemma 7.2 implies now that, after a suitable variable change, one of the numbers m_1, m_2, m_3 is strictly bigger than the others. After renumbering, we may assume that

$$m_1 > m_2, m_3.$$

We claim that

$$n_1 > n_2, n_3 \tag{8-5}$$

as well, and moreover,

$$m_1 - \max\{m_2, m_3\} = n_1 - \max\{n_2, n_3\}. \tag{8-6}$$

Indeed, assume that, say, $n_2 \geq n_1, n_3$. Then the leading terms of the q-expansion on the left and on the right of (8-2) are of the forms $cq^{-(m_1+n_2)}$ and $c'q^{-(\max\{m_2,m_3\}+n_2)}$ with some nonzero c and c'. (Precisely,

$$c = \begin{cases} \varepsilon_1^{-1}\eta_2^{-1}, & n_2 > n_3, \\ \varepsilon_1^{-1}(\eta_2^{-1} - \eta_3^{-1}), & n_2 = n_3 > 0, \\ \varepsilon_1^{-1}(g_2 - g_3), & n_2 = n_3 = 0, \end{cases}$$

and it follows from (8-3) that $c \neq 0$; in a similar way one shows that $c' \neq 0$.) And this is impossible because $m_1 + n_2 > \max\{m_2, m_3\} + n_2$. This proves that $n_1 > n_2, n_3$. In particular the three maps g_k are also not all constant. Again comparing the leading terms of the q-expansion on the left and on the right of (8-2), we obtain (8-6).

Swapping, if necessary, the functions f_k and g_k, we may assume that

$$m_1 \geq n_1, \tag{8-7}$$

and after renumbering, we may assume that

$$m_1 > m_2 \geq m_3. \tag{8-8}$$

Equation (8-6) now becomes

$$m_1 - m_2 = n_1 - \max\{n_2, n_3\}. \tag{8-9}$$

8B. ***One more lemma.*** Here is a less obvious property, which will be used in the proof several times.

Lemma 8.1. *In the above setup we cannot simultaneously have* $f_2 = g_3$ *and* $g_2 = f_3$.

Proof. If $f_2 = g_3$ and $g_2 = f_3$, then

$$0 = \begin{vmatrix} 1 & 1 & 1 \\ f_1 & f_2 & f_3 \\ g_1 & f_3 & f_2 \end{vmatrix} = (f_3 - f_2)(f_1 + g_1 - f_2 - f_3).$$

Since $f_2 \neq f_3$, this implies

$$f_1 + g_1 = f_2 + f_3. \tag{8-10}$$

We will see that this leads to a contradiction.

Observe first of all that $m_2 > 0$. Indeed, if $m_2 = 0$, then $m_3 = 0$ as well by (8-8). Hence, both f_2 and f_3 are constant, and (8-10) contradicts Lemma 7.3.

Next, we have $m_3 > 0$ as well. Indeed, if f_3 is constant, then comparing the constant terms in (8-10), we find $f_3 = 744$, contradicting Lemma 5.6.

Thus, we have $m_1 \geq n_1 > n_3 = m_2 \geq m_3 > 0$. Comparing the q-expansions

$$f_1 + g_1 = \begin{cases} \varepsilon_1^{-1} q^{-m_1} + \eta_1^{-1} q^{-n_1} + O(1), & m_1 > n_1, \\ (\varepsilon_1^{-1} + \eta_1^{-1}) q^{-m_1} + O(1), & m_1 = n_1,\ \varepsilon_1 \neq -\eta_1, \\ 1488 + 2B\varepsilon_1^2 q^{2m_1} + o(q^{2m_1}), & m_1 = n_1,\ \varepsilon_1 = -\eta_1, \end{cases}$$

$$f_2 + f_3 = \begin{cases} \varepsilon_2^{-1} q^{-m_2} + \varepsilon_3^{-1} q^{-m_3} + O(1), & m_2 > m_3, \\ (\varepsilon_2^{-1} + \varepsilon_3^{-1}) q^{-m_2} + O(1), & m_2 = m_3,\ \varepsilon_2 \neq -\varepsilon_3, \\ 1488 + 2B\varepsilon_2^2 q^{2m_2} + o(q^{2m_2}), & m_2 = m_3,\ \varepsilon_2 = -\varepsilon_3, \end{cases}$$

we immediately derive a contradiction. □

8C. ***The determinant*** $\mathcal{D}(q)$**.** We will study in the sequel a slightly modified version of the determinant from (8-1):

$$\mathcal{D}(q) = \begin{vmatrix} 1 & 1 & 1 \\ q^{m_1} f_1 & q^{m_1} f_2 & q^{m_1} f_3 \\ q^{n_1} g_1 & q^{n_1} g_2 & q^{n_1} g_3 \end{vmatrix}.$$

The advantage is that it has no negative powers of q. Equation (8-1) simply means that $\mathcal{D}(q)$ vanishes as a formal power series in q. It will be useful to write

$$\mathcal{D}(q) = \begin{vmatrix} 1 & 1 & 1 \\ q^{m_1}(f_1 - 744) & q^{m_1}(f_2 - 744) & q^{m_1}(f_3 - 744) \\ q^{n_1}(g_1 - 744) & q^{n_1}(g_2 - 744) & q^{n_1}(g_3 - 744) \end{vmatrix}. \tag{8-11}$$

This would allow us to eliminate the constant terms in the q-expansions of f_k and g_k.

It will be convenient to use the notation

$$\tilde{f}_k = \begin{cases} \varepsilon_k^{-1}, & m_k > 0, \\ f_k - 744, & m_k = 0, \end{cases} \qquad \tilde{g}_k = \begin{cases} \eta_k^{-1}, & n_k > 0, \\ g_k - 744, & n_k = 0 \end{cases} \tag{8-12}$$

so that

$$q^{m_1}(f_k - 744) = \tilde{f}_k q^{m_1 - m_k} + o(q^{m_1}), \qquad q^{n_1}(g_k - 744) = \tilde{g}_k q^{n_1 - n_k} + o(q^{n_1}).$$

Lemma 5.6 implies that

$$\tilde{f}_k, \tilde{g}_k \neq 0 \quad (k = 1, 2, 3), \tag{8-13}$$

which will be frequently used, usually without special references.

8D. ***The four cases.*** According to (8-5) and (8-8), there are four possible cases:

$$\begin{aligned} &m_2 = m_3, \\ &m_2 > m_3, \qquad n_2 > n_3, \\ &m_2 > m_3, \qquad n_2 = n_3, \\ &m_2 > m_3, \qquad n_3 > n_2. \end{aligned}$$

They are treated in the four subsequent sections, respectively. We will show that in the first two cases we have (8-4) and that the last two cases are impossible. The proofs in the four cases are similar in strategy but differ in technical details.

Most of our arguments are nothing more than careful manipulations with q-expansions. Still, they are quite technical, and to facilitate reading, we split proofs of each of the cases it into short logically complete steps.

9. The case $m_2 = m_3$

In this section we assume that

$$m_1 > m_2 = m_3.$$

We want to prove that in this case we have $f_k = g_k$ for $k = 1, 2, 3$.

Let us briefly describe the strategy of the proof. We already have (8-5), and after renumbering we may assume that

$$n_1 > n_2 \geq n_3.$$

Equation (8-9) now becomes

$$m_1 - m_2 = m_1 - m_3 = n_1 - n_2. \tag{9-1}$$

We start by proving that $n_2 = n_3$; see Section 9A. With this done, setting $m_2 = m_3 = m$ and $n_2 = n_3 = n$, we rewrite (9-1) as

$$m_1 - m = n_1 - n. \tag{9-2}$$

The next step is proving (see Section 9B) that $m_1 = n_1$. In view of (9-2) this would imply that $m = n$ as well. In particular, f_k and g_k are of the same level for every $k = 1, 2, 3$. After this, we will be ready to prove that $f_k = g_k$ for $k = 1, 2, 3$; see Section 9C.

9A. ***Proof of $n_2 = n_3$.*** In this subsection we prove that $n_2 = n_3$. Set

$$m_1 - m_2 = m_1 - m_3 = n_1 - n_2 = \lambda, \qquad n_1 - n_3 = \lambda' \geq \lambda.$$

We want to show that $\lambda' = \lambda$.

Assume that $\lambda' > \lambda$. Then by (8-7) all the m_k and n_k except perhaps n_3 are positive. We consider separately the cases $n_3 = 0$ and $n_3 > 0$.

The subcase $n_3 = 0$. If $n_3 = 0$, then using notation (8-12), we write $\tilde{g}_3 = g_3 - 744$ and

$$\begin{aligned}\mathcal{D}(q) &= \begin{vmatrix} 1 & 1 & 1 \\ \varepsilon_1^{-1} & \varepsilon_2^{-1} q^{\lambda} & \varepsilon_3^{-1} q^{\lambda} \\ \eta_1^{-1} & \eta_2^{-1} q^{\lambda} & \tilde{g}_3 q^{\lambda'} \end{vmatrix} + o(q^{n_1}) \\ &= (\varepsilon_1^{-1}\eta_2^{-1} - \varepsilon_2^{-1}\eta_1^{-1} + \varepsilon_3^{-1}\eta_1^{-1}) q^{\lambda} + \varepsilon_3^{-1}\eta_2^{-1} q^{2\lambda} + \varepsilon_1^{-1}\tilde{g}_3 q^{\lambda'} + o(q^{n_1}) + O(q^{\lambda+\lambda'}).\end{aligned}$$

The term with $q^{\lambda'}$ can be eliminated only if $\lambda' = 2\lambda$ and $\varepsilon_1^{-1}\tilde{g}_3 = \varepsilon_3^{-1}\eta_2^{-1}$, that is, $g_3 = 744 + \varepsilon_1\varepsilon_3^{-1}\eta_2^{-1}$, contradicting Lemma 5.7.

The subcase $n_3 > 0$. If $n_3 > 0$, then

$$\begin{aligned}\mathcal{D}(q) &= \begin{vmatrix} 1 & 1 & 1 \\ \varepsilon_1^{-1} & \varepsilon_2^{-1} q^{\lambda} & \varepsilon_3^{-1} q^{\lambda} \\ \eta_1^{-1} & \eta_2^{-1} q^{\lambda} & \eta_3^{-1} q^{\lambda'} + A\eta_3 q^{n_1+n_3} \end{vmatrix} + o(q^{n_1+n_3}) \\ &= (\varepsilon_1^{-1}\eta_2^{-1} - \varepsilon_2^{-1}\eta_1^{-1} + \varepsilon_3^{-1}\eta_1^{-1}) q^{\lambda} - \varepsilon_3^{-1}\eta_2^{-1} q^{2\lambda} - \varepsilon_1^{-1}\eta_3^{-1} q^{\lambda'} \\ &\qquad + \varepsilon_2^{-1}\eta_3^{-1} q^{\lambda+\lambda'} - A\varepsilon_1^{-1}\eta_3 q^{n_1+n_3} + o(q^{n_1+n_3}).\end{aligned}$$

As $n_1 + n_3 > \lambda'$, the term with $q^{n_1+n_3}$ can be eliminated only if either

$$\lambda < \lambda' < 2\lambda = n_1 + n_3 < \lambda + \lambda', \qquad \varepsilon_3^{-1}\eta_2^{-1} = -A\varepsilon_1^{-1}\eta_3,$$

which is impossible because A is not a root of unity, or

$$\lambda < \lambda', 2\lambda < n_1 + n_3 = \lambda + \lambda', \qquad \varepsilon_2^{-1}\eta_3^{-1} = A\varepsilon_1^{-1}\eta_3,$$

which is again impossible by the same reason.

Conclusion. Thus, we have proved that $n_2 = n_3$. Setting $m = m_2 = m_3$ and $n = n_2 = n_3$, we can summarize our knowledge as

$$m_1 > m_2 = m_3 = m, \qquad m_1 - m = n_1 - n = \lambda > 0,$$
$$n_1 > n_2 = n_3 = n, \qquad m_1 - n_1 = m - n \geq 0.$$

Together with (8-3) this implies that

$$\tilde{f}_2 \neq \tilde{f}_3, \qquad \tilde{g}_2 \neq \tilde{g}_3. \tag{9-3}$$

9B. ***Proof of $m_1 = n_1$.*** Now we want to prove that

$$m_1 = n_1. \tag{9-4}$$

Thus, assume that $m_1 > n_1$, in which case we also have $m > n$. We consider separately the subcases $n > 0$ and $n = 0$.

The subcase $n > 0$. If $n > 0$, then

$$\mathcal{D}(q) = \begin{vmatrix} 1 & 1 & 1 \\ \varepsilon_1^{-1} & \varepsilon_2^{-1} q^{\lambda} & \varepsilon_3^{-1} q^{\lambda} \\ \eta_1^{-1} & \eta_2^{-1} q^{\lambda} + A\eta_2 q^{n_1+n} & \eta_3^{-1} q^{\lambda} + A\eta_3 q^{n_1+n} \end{vmatrix} + o(q^{n_1+n})$$
$$= \begin{vmatrix} \varepsilon_1^{-1} & \varepsilon_2^{-1} - \varepsilon_3^{-1} \\ \eta_1^{-1} & \eta_2^{-1} - \eta_3^{-1} \end{vmatrix} q^{\lambda} + \begin{vmatrix} \varepsilon_2^{-1} & \varepsilon_3^{-1} \\ \eta_2^{-1} & \eta_3^{-1} \end{vmatrix} q^{2\lambda} + A\varepsilon_1^{-1}(\eta_2 - \eta_3) q^{n_1+n} + o(q^{n_1+n}) + o(q^{2\lambda}).$$

Here the coefficient of q^{λ} must vanish. If $2\lambda > n_1 + n$, then that of q^{n_1+n} must vanish too, but that would contradict (9-3). If $2\lambda < n_1 + n$, then the coefficient of $q^{2\lambda}$ must vanish and then that of q^{n_1+n}. It follows that $2\lambda = n_1 + n$ and

$$\begin{vmatrix} \varepsilon_2^{-1} & \varepsilon_3^{-1} \\ \eta_2^{-1} & \eta_3^{-1} \end{vmatrix} = A\varepsilon_1^{-1}(\eta_3 - \eta_2). \tag{9-5}$$

As noted, both sides of (9-5) are nonzero. Since the left-hand side is a sum of two roots of unity, Lemma 4.1 implies that $196884 = |A| \leq 2$, a contradiction. This completes the proof of (9-4) in the case $n > 0$.

The subcase $n = 0$. If $n = 0$, then g_2 and g_3 are distinct constants, and the other functions are nonconstant. Also, we have $\lambda = n_1$, and so

$$m_1 = m + n_1. \tag{9-6}$$

Now, using notation (8-12), we obtain

$$\begin{aligned}\mathcal{D}(q) &= \begin{vmatrix} 1 & 1 & 1 \\ \varepsilon_1^{-1} & \varepsilon_2^{-1}q^{n_1} + A\varepsilon_2 q^{m_1+m} & \varepsilon_3^{-1}q^{n_1} + A\varepsilon_3 q^{m_1+m} \\ \eta_1^{-1} + A\eta_1 q^{2n_1} & \tilde{g}_2 q^{n_1} & \tilde{g}_3 q^{n_1} \end{vmatrix} \\ &\qquad + o(q^{m_1+m}) + o(q^{2n_1}) \\ &= \begin{vmatrix} \varepsilon_1^{-1} & \varepsilon_2^{-1} - \varepsilon_3^{-1} \\ \eta_1^{-1} & \tilde{g}_2 - \tilde{g}_3 \end{vmatrix} q^{n_1} + \begin{vmatrix} \varepsilon_2^{-1} & \varepsilon_3^{-1} \\ \tilde{g}_2 & \tilde{g}_3 \end{vmatrix} q^{2n_1} + A\eta_1^{-1}(\varepsilon_3 - \varepsilon_2) q^{m_1+m} \\ &\qquad + o(q^{m_1+m}) + o(q^{2n_1}).\end{aligned}$$

As $\varepsilon_3 \neq \varepsilon_2$, the coefficient of q^{m_1+m} is nonzero; by Lemma 5.2 so is the coefficient of q^{2n_1}. This shows that $2n_1 = m_1 + m$. Together with (9-6) this implies $m_1 = 3m$ and $n_1 = 2m$; rescaling z, we may assume

$$m = 1, \qquad n_1 = 2, \qquad m_1 = 3.$$

Hence,

$$\begin{aligned}\mathcal{D}(q) &= \begin{vmatrix} 1 & 1 & 1 \\ \varepsilon_1^{-1} & \varepsilon_2^{-1}q^2 + A\varepsilon_2 q^4 + B\varepsilon_2^2 q^5 & \varepsilon_3^{-1}q^2 + A\varepsilon_3 q^4 + B\varepsilon_3^2 q^5 \\ \eta_1^{-1} + A\eta_1 q^4 & \tilde{g}_2 q^2 & \tilde{g}_3 q^2 \end{vmatrix} + O(q^6) \\ &= \begin{vmatrix} \varepsilon_1^{-1} & \varepsilon_2^{-1} - \varepsilon_3^{-1} \\ \eta_1^{-1} & \tilde{g}_2 - \tilde{g}_3 \end{vmatrix} q^2 + \left(\begin{vmatrix} \varepsilon_2^{-1} & \varepsilon_3^{-1} \\ \tilde{g}_2 & \tilde{g}_3 \end{vmatrix} + A\eta_1^{-1}(\varepsilon_3 - \varepsilon_2) \right) q^4 + B\eta_1^{-1}(\varepsilon_3^2 - \varepsilon_2^2) q^5 + O(q^6).\end{aligned}$$

Equating to 0 the coefficient of q^5, we obtain $\varepsilon_3 = \pm\varepsilon_2$, and (9-3) implies that $\varepsilon_3 = -\varepsilon_2$. Using this, and equating to 0 the coefficients of q^2 and q^4, we obtain

$$\varepsilon_1^{-1}(\tilde{g}_2 - \tilde{g}_3) = 2\varepsilon_2^{-1}\eta_1^{-1}, \qquad \varepsilon_2^{-1}(\tilde{g}_2 + \tilde{g}_3) = 2A\eta_1^{-1}\varepsilon_2,$$

from which we deduce $g_2 = \tilde{g}_2 + 744 = \varepsilon_1\varepsilon_2^{-1}\eta_1^{-1} + A\eta_1^{-1}\varepsilon_2^2 + 744$.

Now Lemma 5.9 implies that $g_2 \in \mathbb{Z}$, from which we deduce, using Lemma 4.2, that both roots of unity $\varepsilon_1\varepsilon_2^{-1}\eta_1^{-1}$ and $\eta_1^{-1}\varepsilon_2^2$ must be ± 1. Hence, g_2 is one of the four numbers $744 \pm 1 \pm A$, contradicting Lemma 5.6.

9C. ***Proof of $f_k = g_k$ for $k = 1, 2, 3$.*** In the previous subsection we proved that

$$m_1 = n_1 > m = n. \tag{9-7}$$

We want to now prove that

$$f_k = g_k \quad (k = 1, 2, 3). \tag{9-8}$$

We again distinguish the subcases $m = n > 0$ and $m = n = 0$. As before, we set $\lambda = m_1 - m = n_1 - n$.

The subcase $m = n > 0$. If $m = n > 0$ then

$$\mathcal{D}(q) = \begin{vmatrix} 1 & 1 & 1 \\ \varepsilon_1^{-1} & \varepsilon_2^{-1}q^{\lambda} + A\varepsilon_2 q^{\lambda+2m} & \varepsilon_3^{-1}q^{\lambda} + A\varepsilon_3 q^{\lambda+2m} \\ \eta_1^{-1} & \eta_2^{-1}q^{\lambda} + A\eta_2 q^{\lambda+2m} & \eta_3^{-1}q^{\lambda} + A\eta_3 q^{\lambda+2m} \end{vmatrix} + o(q^{\lambda+2m})$$

$$= \begin{vmatrix} \varepsilon_1^{-1} & \varepsilon_2^{-1} - \varepsilon_3^{-1} \\ \eta_1^{-1} & \eta_2^{-1} - \eta_3^{-1} \end{vmatrix} q^{\lambda} + \begin{vmatrix} \varepsilon_2^{-1} & \varepsilon_3^{-1} \\ \eta_2^{-1} & \eta_3^{-1} \end{vmatrix} q^{2\lambda} + A \begin{vmatrix} \varepsilon_1^{-1} & \varepsilon_2 - \varepsilon_3 \\ \eta_1^{-1} & \eta_2 - \eta_3 \end{vmatrix} q^{\lambda+2m} + o(q^{\lambda+2m}). \quad (9\text{-}9)$$

This implies the equations

$$\begin{vmatrix} \varepsilon_1^{-1} & \varepsilon_2^{-1} - \varepsilon_3^{-1} \\ \eta_1^{-1} & \eta_2^{-1} - \eta_3^{-1} \end{vmatrix} = 0, \qquad \begin{vmatrix} \varepsilon_2^{-1} & \varepsilon_3^{-1} \\ \eta_2^{-1} & \eta_3^{-1} \end{vmatrix} = 0, \qquad \begin{vmatrix} \varepsilon_1^{-1} & \varepsilon_2 - \varepsilon_3 \\ \eta_1^{-1} & \eta_2 - \eta_3 \end{vmatrix} = 0 \quad (9\text{-}10)$$

if $2\lambda \neq \lambda + 2m$ and the equations

$$\begin{vmatrix} \varepsilon_1^{-1} & \varepsilon_2^{-1} - \varepsilon_3^{-1} \\ \eta_1^{-1} & \eta_2^{-1} - \eta_3^{-1} \end{vmatrix} = 0,$$
$$\begin{vmatrix} \varepsilon_2^{-1} & \varepsilon_3^{-1} \\ \eta_2^{-1} & \eta_3^{-1} \end{vmatrix} = -A \begin{vmatrix} \varepsilon_1^{-1} & \varepsilon_2 - \varepsilon_3 \\ \eta_1^{-1} & \eta_2 - \eta_3 \end{vmatrix} \quad (9\text{-}11)$$

if $2\lambda = \lambda + 2m$. If both sides of (9-11) are nonzero, then Lemma 4.1 implies $196884 = |A| \leq 2$, a contradiction. Hence, in any case we have (9-10).

Resolving the first two equations from (9-10) in $\eta_1^{-1}, \eta_2^{-1}, \eta_3^{-1}$ and using (9-3), we obtain

$$(\eta_1, \eta_2, \eta_3) = \theta(\varepsilon_1, \varepsilon_2, \varepsilon_3)$$

for some $\theta \in \mathbb{C}$. Substituting this into the third equation in (9-10) and again using (9-3), we find $\theta = \pm 1$. If $\theta = -1$, then we get for $\mathcal{D}(q)$ the value

$$\begin{vmatrix} 1 & 1 & 1 \\ \varepsilon_1^{-1} + A\varepsilon_1 q^{2\lambda+2m} & \varepsilon_2^{-1}q^{\lambda} + A\varepsilon_2 q^{\lambda+2m} + B\varepsilon_2^2 q^{\lambda+3m} & \varepsilon_3^{-1}q^{\lambda} + A\varepsilon_3 q^{\lambda+2m} + B\varepsilon_3^2 q^{\lambda+3m} \\ -\varepsilon_1^{-1} - A\varepsilon_1 q^{2\lambda+2m} & -\varepsilon_2^{-1}q^{\lambda} - A\varepsilon_2 q^{\lambda+2m} + B\varepsilon_2^2 q^{\lambda+3m} & -\varepsilon_3^{-1}q^{\lambda} - A\varepsilon_3 q^{\lambda+2m} + B\varepsilon_3^2 q^{\lambda+3m} \end{vmatrix} + o(q^{\lambda+3m})$$

$$= \begin{vmatrix} 1 & 1 & 1 \\ \varepsilon_1^{-1} + A\varepsilon_1 q^{2\lambda+2m} & \varepsilon_2^{-1}q^{\lambda} + A\varepsilon_2 q^{\lambda+2m} + B\varepsilon_2^2 q^{\lambda+3m} & \varepsilon_3^{-1}q^{\lambda} + A\varepsilon_3 q^{\lambda+2m} + B\varepsilon_3^2 q^{\lambda+3m} \\ 0 & 2B\varepsilon_2^2 q^{\lambda+3m} & 2B\varepsilon_3^2 q^{\lambda+3m} \end{vmatrix} + o(q^{\lambda+3m})$$

$$= 2B\varepsilon_1^{-1}(\varepsilon_2^2 - \varepsilon_3^2)q^{\lambda+3m} + o(q^{\lambda+3m}),$$

which gives $\varepsilon_2 = \pm\varepsilon_3$, and $\varepsilon_2 = -\varepsilon_3$ by (9-3). Thus, we have $\varepsilon_2 = \eta_3 = -\varepsilon_3 = -\eta_2$, which implies that $f_2 = g_3$ and $g_2 = f_3$, contradicting Lemma 8.1.

The only remaining option is $\theta = 1$, which, together with (9-7), proves (9-8).

The subcase $m = n = 0$. This case can be easily settled using Lemma 7.3. Indeed, in the case $m = n = 0$ the functions f_1, g_1 are nonconstant, f_2, f_3, g_2, g_3 are constant, and

$$0 = \begin{vmatrix} 1 & 1 & 1 \\ f_1 & f_2 & f_3 \\ g_1 & g_2 & g_3 \end{vmatrix} = (g_2 - g_3) f_1 - (f_2 - f_3) g_1 + \begin{vmatrix} f_2 & f_3 \\ g_2 & g_3 \end{vmatrix}$$

is a nontrivial linear relation for f_1, g_1 (recall that $f_2 \neq f_3$ and $g_2 \neq g_3$ by (8-3)). By Lemma 7.3

$$f_1 = g_1, \quad f_2 - f_3 = g_2 - g_3, \quad \begin{vmatrix} f_2 & f_3 \\ g_2 & g_3 \end{vmatrix} = 0.$$

From the last two equations, one easily deduces that $f_2 = g_2$ and $f_3 = g_3$, proving (9-8).

10. The case $m_2 > m_3$ and $n_2 > n_3$

In this section we assume that

$$m_1 > m_2 > m_3, \qquad n_1 > n_2 > n_3. \tag{10-1}$$

As in the previous section, we will prove that in this case $f_k = g_k$ for $k = 1, 2, 3$.

The strategy of the proof is similar to that of the previous section. Equation (8-9) now reads

$$m_1 - m_2 = n_1 - n_2. \tag{10-2}$$

We start with proving that

$$m_1 - m_3 = n_1 - n_3; \tag{10-3}$$

see Section 10A. Then we prove, in Section 10B, that $m_1 = n_1$. Since, by this time, we will already know (10-2) and (10-3), this will imply that $m_k = n_k$ for every $k = 1, 2, 3$. After this, we prove that $f_k = g_k$ for $k = 1, 2, 3$ in Section 10C.

We set $m_1 - m_2 = n_1 - n_2 = \lambda$. We also have $m_1 \geq n_1$ by (8-7). Let us collect our knowledge:

$$m_1 > m_2 > m_3, \quad n_1 > n_2 > n_3, \quad m_1 - m_2 = n_1 - n_2 = \lambda > 0, \quad m_1 - n_1 = m_2 - n_2 \geq 0.$$

10A. ***Proof of $m_1 - m_3 = n_1 - n_3$.*** Now let us prove that $m_1 - m_3 = n_1 - n_3$. Using notation (8-12), we write

$$\begin{aligned} \mathcal{D}(q) &= \begin{vmatrix} 1 & 1 & 1 \\ \varepsilon_1^{-1} & \varepsilon_2^{-1} q^{\lambda} & \tilde{f}_3 q^{m_1 - m_3} \\ \eta_1^{-1} & \eta_2^{-1} q^{\lambda} & \tilde{g}_3 q^{n_1 - n_3} \end{vmatrix} + o(q^{n_1}) \\ &= \begin{vmatrix} \varepsilon_1^{-1} & \varepsilon_2^{-1} \\ \eta_1^{-1} & \eta_2^{-1} \end{vmatrix} q^{\lambda} + \tilde{f}_3 \eta_1^{-1} q^{m_1 - m_3} - \varepsilon_1^{-1} \tilde{g}_3 q^{n_1 - n_3} + o(q^{m_1 - m_3}) + o(q^{n_1 - n_3}). \end{aligned}$$

If $m_1 - m_3 \neq n_1 - n_3$, then we have one of the options

$$\lambda < m_1 - m_3 < n_1 - n_3, \qquad \lambda < n_1 - n_3 < m_1 - m_3.$$

In the first case $q^{m_1-m_3}$ cannot be eliminated, and in the second case $q^{n_1-n_3}$ cannot be eliminated. This proves that $m_1 - m_3 = n_1 - n_3$.

We set $m_1 - m_3 = n_1 - n_3 = \lambda'$. Thus,

$$\begin{gathered} m_1 > m_2 > m_3, \qquad n_1 > n_2 > n_3, \\ m_1 - m_2 = n_1 - n_2 = \lambda > 0, \qquad m_1 - m_3 = n_1 - n_3 = \lambda' > \lambda > 0, \\ m_1 - n_1 = m_2 - n_2 = m_3 - n_3 \geq 0. \end{gathered} \tag{10-4}$$

In addition to this, from

$$\mathscr{D}(q) = \begin{vmatrix} 1 & 1 & 1 \\ \varepsilon_1^{-1} & \varepsilon_2^{-1} q^{\lambda} & \tilde{f}_3 q^{\lambda'} \\ \eta_1^{-1} & \eta_2^{-1} q^{\lambda} & \tilde{g}_3 q^{\lambda'} \end{vmatrix} + o(q^{n_1}) = \begin{vmatrix} \varepsilon_1^{-1} & \varepsilon_2^{-1} \\ \eta_1^{-1} & \eta_2^{-1} \end{vmatrix} q^{\lambda} - \begin{vmatrix} \varepsilon_1^{-1} & \tilde{f}_3 \\ \eta_1^{-1} & \tilde{g}_3 \end{vmatrix} q^{\lambda'} + o(q^{\lambda'}),$$

we deduce that

$$\begin{vmatrix} \varepsilon_1^{-1} & \varepsilon_2^{-1} \\ \eta_1^{-1} & \eta_2^{-1} \end{vmatrix} = \begin{vmatrix} \varepsilon_1^{-1} & \tilde{f}_3 \\ \eta_1^{-1} & \tilde{g}_3 \end{vmatrix} = 0, \tag{10-5}$$

which means that

$$(\eta_1^{-1}, \eta_2^{-1}, \tilde{g}_3) = \theta(\varepsilon_1^{-1}, \varepsilon_2^{-1}, \tilde{f}_3) \tag{10-6}$$

with some root of unity θ.

10B. ***Proof of $m_1 = n_1$.*** In this subsection we show that $m_1 = n_1$. Thus, assume

$$m_1 > n_1, \tag{10-7}$$

in which case we also have

$$m_2 > n_2, \qquad m_3 > n_3. \tag{10-8}$$

We should also have

$$n_3 > 0. \tag{10-9}$$

Indeed, if $m_3 > n_3 = 0$, then the second equation in (10-5) reads $g_3 = 744 + \varepsilon_1 \varepsilon_3^{-1} \eta_1^{-1}$, which is impossible by Lemma 5.7.

Using (10-6), (10-7), (10-8), and (10-9), we obtain

$$\begin{aligned}\mathcal{D}(q) &= \begin{vmatrix} 1 & 1 & 1 \\ \varepsilon_1^{-1} & \varepsilon_2^{-1}q^{\lambda} & \varepsilon_3^{-1}q^{\lambda'} \\ \eta_1^{-1} & \eta_2^{-1}q^{\lambda} & \eta_3^{-1}q^{\lambda'} + A\eta_3 q^{n_1+n_3} \end{vmatrix} + o(q^{n_1+n_3}) \\ &= \begin{vmatrix} 1 & 1 & 1 \\ \varepsilon_1^{-1} & \varepsilon_2^{-1}q^{\lambda} & \varepsilon_3^{-1}q^{\lambda'} \\ 0 & 0 & A\eta_3 q^{n_1+n_3} \end{vmatrix} + o(q^{n_1+n_3}) \\ &= -A\varepsilon_1^{-1}\eta_3 q^{n_1+n_3} + o(q^{n_1+n_3}),\end{aligned}$$

a contradiction.

This proves that

$$m_k = n_k \quad (k = 1, 2, 3). \tag{10-10}$$

10C. ***Proof of $f_k = g_k$ for $k = 1, 2, 3$.*** To prove that $f_k = g_k$ for $k = 1, 2, 3$, we only need to show that

$$\theta = 1,$$

where θ is from (10-6). If $m_3 = n_3 = 0$, then rewriting the equality $\tilde{g}_3 = \theta \tilde{f}_3$ as $(g_3 - 744) = \theta(f_3 - 744)$, we deduce $\theta = 1$ from Lemma 5.2.

Now assume that $m_3 = n_3 > 0$. In this case

$$\begin{aligned}\mathcal{D}(q) &= \begin{vmatrix} 1 & 1 & 1 \\ \varepsilon_1^{-1} & \varepsilon_2^{-1}q^{\lambda} & \varepsilon_3^{-1}q^{\lambda'} + A\varepsilon_3 q^{m_1+m_3} \\ \eta_1^{-1} & \eta_2^{-1}q^{\lambda} & \eta_3^{-1}q^{\lambda'} + A\eta_3 q^{m_1+m_3} \end{vmatrix} + o(q^{m_1+m_3}) \\ &= \begin{vmatrix} 1 & 1 & 1 \\ \varepsilon_1^{-1} & \varepsilon_2^{-1}q^{\lambda} & \varepsilon_3^{-1}q^{\lambda'} + A\varepsilon_3 q^{m_1+m_3} \\ 0 & 0 & A\varepsilon_3(\theta^{-1} - \theta) q^{m_1+m_3} \end{vmatrix} + o(q^{m_1+m_3}) \\ &= -A\varepsilon_1^{-1}\varepsilon_3(\theta^{-1} - \theta) q^{m_1+m_3} + o(q^{m_1+m_3}),\end{aligned}$$

which implies $\theta = \pm 1$. If $\theta = -1$, then we get for $\mathcal{D}(q)$ the value

$$\begin{vmatrix} 1 & 1 & 1 \\ \varepsilon_1^{-1} + A\varepsilon_1 q^{2m_1} & \varepsilon_2^{-1}q^{\lambda} + A\varepsilon_2 q^{m_1+m_2} & \varepsilon_3^{-1}q^{\lambda'} + A\varepsilon_3 q^{m_1+m_3} + B\varepsilon_3^2 q^{m_1+2m_3} \\ -\varepsilon_1^{-1} - A\varepsilon_1 q^{2m_1} & -\varepsilon_2^{-1}q^{\lambda} - A\varepsilon_2 q^{m_1+m_2} & -\varepsilon_3^{-1}q^{\lambda'} - A\varepsilon_3 q^{m_1+m_3} + B\varepsilon_3^2 q^{m_1+2m_3} \end{vmatrix}$$
$$+ o(q^{m_1+2m_3})$$
$$\begin{aligned} &= \begin{vmatrix} 1 & 1 & 1 \\ \varepsilon_1^{-1} + A\varepsilon_1 q^{2m_1} & \varepsilon_2^{-1}q^{\lambda} + A\varepsilon_2 q^{m_1+m_2} & \varepsilon_3^{-1}q^{\lambda'} + A\varepsilon_3 q^{m_1+m_3} \\ 0 & 0 & 2B\varepsilon_3^2 q^{m_1+2m_3} \end{vmatrix} + o(q^{m_1+2m_3}) \\ &= -2B\varepsilon_1^{-1}\varepsilon_3^2 q^{m_1+2m_3} + o(q^{m_1+2m_3}),\end{aligned}$$

a contradiction.

Thus, in any case we have $\theta = 1$ in (10-6). Together with (10-10), this proves that $f_k = g_k$ for $k = 1, 2, 3$.

11. The case $m_2 > m_3$ and $n_2 = n_3$

In this section we assume that

$$m_1 > m_2 > m_3, \qquad n_1 > n_2 = n_3 \tag{11-1}$$

and will show that this is impossible.

Relation (8-9) now becomes $m_1 - m_2 = n_1 - n_2 = n_1 - n_3$. We set

$$m_1 - m_2 = n_1 - n_2 = n_1 - n_3 = \lambda. \tag{11-2}$$

Fist of all, let us rule out the case $n_2 = n_3 = 0$. In this case $n_1 = \lambda < m_1 - m_3$. Using notation (8-12), we write in this case

$$\mathcal{D}(q) = \begin{vmatrix} 1 & 1 & 1 \\ \varepsilon_1^{-1} & \varepsilon_2^{-1} q^{\lambda} & 0 \\ \eta_1^{-1} & \tilde{g}_2 q^{\lambda} & \tilde{g}_3 q^{\lambda} \end{vmatrix} + o(q^{\lambda}) = (\varepsilon_1^{-1}\tilde{g}_2 - \varepsilon_1^{-1}\tilde{g}_3 - \varepsilon_2^{-1}\eta_1^{-1}) q^{\lambda} + o(q^{\lambda}).$$

We obtain $\varepsilon_1^{-1}\tilde{g}_2 - \varepsilon_1^{-1}\tilde{g}_3 - \varepsilon_2^{-1}\eta_1^{-1} = 0$, which contradicts Lemma 5.5.

Thus, we may assume in the sequel that

$$n_2 = n_3 > 0. \tag{11-3}$$

Since $n_2 = n_3$, we have

$$\eta_2 \neq \eta_3, \tag{11-4}$$

which will be systematically used, sometimes without special reference.

Our principal objective will be to show that $m_3 = m_1 - 2\lambda$ and $n_1 = m_1 - \lambda/2$. The first of these two relations is proved already in Section 11A. The second one is more delicate and will be established in Section 11D, after some preparatory work done in the previous subsections. On the way, we will also prove certain inequalities relating the numbers m_k, n_k, and λ and certain relations for the twists. After all this is done, obtaining a contradiction will be relatively easy; see Section 11E.

11A. ***Proof of*** $2\lambda = m_1 - m_3 \leq n_1 + n_2$. Using notation (8-12), we write

$$\begin{aligned} \mathcal{D}(q) &= \begin{vmatrix} 1 & 1 & 1 \\ \varepsilon_1^{-1} & \varepsilon_2^{-1} q^{\lambda} & \tilde{f}_3 q^{m_1 - m_3} \\ \eta_1^{-1} & \eta_2^{-1} q^{\lambda} + A\eta_2 q^{n_1+n_2} & \eta_3^{-1} q^{\lambda} + A\eta_3 q^{n_1+n_2} \end{vmatrix} + o(q^{m_1}) + o(q^{n_1+n_2}) \\ &= (\varepsilon_1^{-1}\eta_2^{-1} - \varepsilon_1^{-1}\eta_3^{-1} - \varepsilon_2^{-1}\eta_1^{-1}) q^{\lambda} + \varepsilon_2^{-1}\eta_3^{-1} q^{2\lambda} + \eta_1^{-1}\tilde{f}_3 q^{m_1 - m_3} \\ &\qquad + A\varepsilon_1^{-1}(\eta_2 - \eta_3) q^{n_1+n_2} + o(q^{m_1 - m_3}) + o(q^{n_1+n_2}). \end{aligned} \tag{11-5}$$

First of all, this gives

$$\varepsilon_1^{-1}\eta_2^{-1} - \varepsilon_1^{-1}\eta_3^{-1} - \varepsilon_2^{-1}\eta_1^{-1} = 0. \tag{11-6}$$

A sum of three roots of unity can vanish only if they are proportional to the three distinct cubic roots of unity. In particular,

$$\eta_2/\eta_3 \text{ is a primitive 6th root of unity.} \tag{11-7}$$

We have $m_1 - m_3 \geq 2\lambda$. Indeed, if $2\lambda > m_1 - m_3$, then we must have

$$m_1 - m_3 = n_1 + n_2, \qquad \eta_1^{-1}\tilde{f}_3 = -A\varepsilon_1^{-1}(\eta_2 - \eta_3). \tag{11-8}$$

If $m_3 > 0$, this gives $\eta_1^{-1}\varepsilon_3^{-1} = -A\varepsilon_1^{-1}(\eta_2 - \eta_3)$, which is impossible because A does not divide a root of unity. And if $m_3 = 0$, then $f_3 = 744 - A\varepsilon_1^{-1}\eta_1(\eta_2 - \eta_3)$. Lemma 5.9 now implies that $f_3 \in \mathbb{Z}$, and we obtain $f_3 \in \{744 \pm 196884, 744 \pm 2 \cdot 196884\}$, contradicting Lemma 5.6.

We have $m_1 - m_3 \leq 2\lambda$. Indeed, if $2\lambda < m_1 - m_3$, then the term with $q^{2\lambda}$ cancels either a term in $o(q^{n_1+n_2})$ or the term with $q^{n_1+n_2}$. In the first situation the terms with $q^{m_1-m_3}$ and $q^{n_1+n_2}$ must cancel each other, and we are back to (11-8). In the second situation we must have

$$2\lambda = n_1 + n_2, \qquad \varepsilon_2^{-1}\eta_3^{-1} = -A\varepsilon_1^{-1}(\eta_2 - \eta_3),$$

which is impossible because $A = 196884$ does not divide a root of unity.

Thus, we proved that $m_1 - m_3 = 2\lambda$.

We have $n_1 + n_2 \geq 2\lambda$. Indeed, if $n_1 + n_2 < 2\lambda = m_1 - m_3$, then the nonzero term $A\varepsilon_1^{-1}(\eta_2 - \eta_3)q^{n_1+n_2}$ cannot be eliminated. (It is nonzero because of (11-4).)

Thus, we proved that

$$2\lambda = m_1 - m_3 \leq n_1 + n_2. \tag{11-9}$$

11B. ***Proof of $n_1 + n_2 > 2\lambda$.*** We want to show now that the inequality in (11-9) is strict. Thus, assume the contrary, that is,

$$2\lambda = m_1 - m_3 = n_1 + n_2. \tag{11-10}$$

Then (11-5) implies that

$$\varepsilon_2^{-1}\eta_3^{-1} + \eta_1^{-1}\tilde{f}_3 + A\varepsilon_1^{-1}(\eta_2 - \eta_3) = 0. \tag{11-11}$$

This implies that $m_3 = 0$. Indeed, if $m_3 > 0$, then (11-11) can be rewritten as

$$\varepsilon_2^{-1}\eta_3^{-1} + \eta_1^{-1}\varepsilon_3^{-1} = -A\varepsilon_1^{-1}(\eta_2 - \eta_3). \tag{11-12}$$

Both sides in (11-12) are nonzero by (11-4), and Lemma 4.1 implies that $2 \geq |A|$, a contradiction. Thus, we have $m_3 = 0$, which, together with (11-2) and (11-10), implies that

$$m_1 = 2\lambda, \qquad m_2 = \lambda, \qquad n_1 = \tfrac{3}{2}\lambda, \qquad n_2 = n_3 = \tfrac{1}{2}\lambda.$$

Rescaling, we may assume that $\lambda = 2$, which gives

$$m_1 = 4, \qquad m_2 = 2, \qquad m_3 = 0, \qquad n_1 = 3, \qquad n_2 = n_3 = 1.$$

Using (11-6) and (11-11), we obtain

$$\begin{aligned}\mathcal{D}(q) &= \begin{vmatrix} 1 & 1 & 1 \\ \varepsilon_1^{-1} & \varepsilon_2^{-1}q^2 & \tilde{f}_3 q^4 \\ \eta_1^{-1} & \eta_2^{-1}q^2 + A\eta_2 q^4 + B\eta_2^2 q^5 & \eta_3^{-1}q^2 + A\eta_3 q^4 + B\eta_3^2 q^5 \end{vmatrix} + O(q^6) \\ &= B\varepsilon_1^{-1}(\eta_2^2 - \eta_3^2)q^5 + O(q^6),\end{aligned}$$

which gives $\eta_2 = \pm\eta_3$, contradicting (11-7).

This proves that

$$2\lambda = m_1 - m_3 < n_1 + n_2. \tag{11-13}$$

11C. ***Proof of $m_3 > 0$.*** In addition to this, we have $m_3 > 0$. Indeed, equating to 0 the coefficient of $q^{2\lambda}$ in (11-5), we obtain

$$\varepsilon_2^{-1}\eta_3^{-1} + \eta_1^{-1}\tilde{f}_3 = 0. \tag{11-14}$$

If $m_3 = 0$, then this gives $f_3 = 744 - \varepsilon_2^{-1}\eta_3^{-1}\eta_1$, contradicting Lemma 5.7. This proves that

$$m_3 > 0, \tag{11-15}$$

and (11-14) becomes

$$\varepsilon_2^{-1}\eta_3^{-1} = -\varepsilon_3^{-1}\eta_1^{-1}. \tag{11-16}$$

11D. ***Proof of $m_1 + m_3 = n_1 + n_2 < 3\lambda$.*** Our next step is showing that $m_1 + m_3 = n_1 + n_2 < 3\lambda$. Using (11-6) and (11-16), we obtain

$$\begin{aligned}\mathcal{D}(q) &= \begin{vmatrix} 1 & 1 & 1 \\ \varepsilon_1^{-1} & \varepsilon_2^{-1}q^{\lambda} & \varepsilon_3^{-1}q^{2\lambda} + A\varepsilon_3 q^{m_1+m_3} \\ \eta_1^{-1} & \eta_2^{-1}q^{\lambda} + A\eta_2 q^{n_1+n_2} & \eta_3^{-1}q^{\lambda} + A\eta_3 q^{n_1+n_2} \end{vmatrix} + o(q^{m_1+m_3}) + o(q^{n_1+n_2}) \\ &= A\varepsilon_3\eta_1^{-1}q^{m_1+m_3} + A\varepsilon_1^{-1}(\eta_2 - \eta_3)q^{n_1+n_2} - \varepsilon_3^{-1}\eta_2^{-1}q^{3\lambda} \\ &\qquad + o(q^{m_1+m_3}) + o(q^{n_1+n_2}).\end{aligned} \tag{11-17}$$

We have $m_1 + m_3 \geq n_1 + n_2$. Indeed, if $m_1 + m_3 < n_1 + n_2$, then we must have $m_1 + m_3 = 3\lambda$ and $A\varepsilon_3\eta_1^{-1} = \varepsilon_3^{-1}\eta_2^{-1}$, which is impossible because A is not a root of unity.

We have $m_1+m_3 \le n_1+n_2$. Similarly, if $m_1+m_3 > n_1+n_2$, then we must have $n_1+n_2 = 3\lambda$ and $A\varepsilon_1^{-1}(\eta_2-\eta_3) = \varepsilon_3^{-1}\eta_2^{-1}$, which is impossible because A does not divide a root of unity.

We have $m_1+m_3 = n_1+n_2 < 3\lambda$. Indeed, if $m_1+m_3 = n_1+n_2 > 3\lambda$, then the $q^{3\lambda}$ cannot be eliminated. And if $m_1+m_3 = n_1+n_2 = 3\lambda$, then $A\varepsilon_3\eta_1^{-1} + A\varepsilon_1^{-1}(\eta_2-\eta_3) = \varepsilon_3^{-1}\eta_2^{-1}$, which is impossible because A does not divide a root of unity.

Thus, we proved that

$$m_1+m_3 = n_1+n_2 < 3\lambda. \tag{11-18}$$

Since $n_2 = n_1-\lambda$ and $m_3 = m_1-2\lambda$ (see (11-2) and (11-13)), this implies that

$$n_1 = m_1 - \tfrac{1}{2}\lambda. \tag{11-19}$$

Also, comparing the coefficients in (11-17), we obtain

$$\varepsilon_3\eta_1^{-1} + \varepsilon_1^{-1}\eta_2 - \varepsilon_1^{-1}\eta_3 = 0. \tag{11-20}$$

11E. ***Conclusion.*** We are almost done. Let us summarize the relations between the levels we already obtained. We deduce from (11-2), (11-15), (11-18), and (11-19)

$$m_2 = m_1-\lambda, \quad m_3 = m_1-2\lambda, \quad n_1 = m_1-\tfrac{1}{2}\lambda, \quad n_2 = n_3 = m_1-\tfrac{3}{2}\lambda, \quad 2\lambda < m_1 < \tfrac{5}{2}\lambda.$$

This implies the inequalities

$$2m_1 > m_1+m_2 = m_1+m_3+\lambda = n_1+n_2+\lambda > 3\lambda, \quad 2n_1 > 3\lambda, \quad n_1+2n_2 > m_1+2m_3.$$

It follows that

$$\begin{aligned}\mathcal{D}(q) = &\begin{vmatrix} 1 & 1 & 1 \\ \varepsilon_1^{-1} & \varepsilon_2^{-1}q^{\lambda} & \varepsilon_3^{-1}q^{2\lambda} + A\varepsilon_3 q^{m_1+m_3} + B\varepsilon_3^2 q^{m_1+2m_3} \\ \eta_1^{-1} & \eta_2^{-1}q^{\lambda} + A\eta_2 q^{n_1+n_2} & \eta_3^{-1}q^{\lambda} + A\eta_3 q^{n_1+n_2} \end{vmatrix} \\ &+ o(q^{m_1+2m_3}) + o(q^{3\lambda}) \\ = &-\varepsilon_3^{-1}\eta_2^{-1}q^{3\lambda} + B\varepsilon_3^2\eta_1^{-1}q^{m_1+2m_3} + o(q^{m_1+2m_3}) + o(q^{3\lambda}).\end{aligned}$$

We obtain $3\lambda = m_1+2m_3$ and $\varepsilon_3^{-1}\eta_2^{-1} = B\varepsilon_3^2\eta_1^{-1}$. But the last equation is impossible because B is not a root of unity. This proves that (11-1) is impossible in case (11-3).

12. The case $m_2 > m_3$ and $n_3 > n_2$

In this section we assume that

$$m_1 > m_2 > m_3, \qquad n_1 > n_3 > n_2 \tag{12-1}$$

(as usual with $m_1 \ge n_1$) and will, eventually, arrive at a contradiction. This is the nastiest case, and we beg for the reader's patience.

Relation (8-9) now becomes $m_1 - m_2 = n_1 - n_3$. We set $m_1 - m_2 = n_1 - n_3 = \lambda$. Using notation (8-12), we write

$$\begin{aligned}\mathscr{D}(q) &= \begin{vmatrix} 1 & 1 & 1 \\ \varepsilon_1^{-1} & \varepsilon_2^{-1} q^{\lambda} & \tilde{f}_3 q^{m_1-m_3} \\ \eta_1^{-1} & \tilde{g}_2 q^{n_1-n_2} & \eta_3^{-1} q^{\lambda} \end{vmatrix} + o(q^{n_1}) \\ &= \begin{vmatrix} \varepsilon_1^{-1} & \varepsilon_2^{-1} \\ \eta_1^{-1} & -\eta_3^{-1} \end{vmatrix} q^{\lambda} + \tilde{f}_3 \eta_1^{-1} q^{m_1-m_3} + \varepsilon_1^{-1} \tilde{g}_2 q^{n_1-n_2} + \varepsilon_2^{-1} \eta_3^{-1} q^{2\lambda} \\ &\qquad - \tilde{f}_3 \tilde{g}_2 q^{m_1-m_3+n_1-n_2} + o(q^{n_1}). \end{aligned} \tag{12-2}$$

Since $0 < \lambda < m_1 - m_3, n_1 - n_2$, this implies that

$$\begin{vmatrix} \varepsilon_1^{-1} & \varepsilon_2^{-1} \\ \eta_1^{-1} & -\eta_3^{-1} \end{vmatrix} = 0. \tag{12-3}$$

12A. ***Proof of $m_1 - m_3 = n_1 - n_2$.*** Let us start by proving that

$$m_1 - m_3 = n_1 - n_2. \tag{12-4}$$

Indeed, assume that $m_1 - m_3 \neq n_1 - n_2$. Then $q^{n_1-n_2}$ in (12-2) can be eliminated only if

$$n_1 - n_2 = 2\lambda, \qquad \varepsilon_1^{-1} \tilde{g}_2 = -\varepsilon_2^{-1} \eta_3^{-1}. \tag{12-5}$$

This implies also that $n_2 > 0$. Indeed, if $n_2 = 0$, then the second equality in (12-5) gives $g_2 = 744 - \varepsilon_1 \varepsilon_2^{-1} \eta_3^{-1}$, contradicting Lemma 5.7.

Using (12-3) and (12-5), we can now write

$$\begin{aligned}\mathscr{D}(q) &= \begin{vmatrix} 1 & 1 & 1 \\ \varepsilon_1^{-1} & \varepsilon_2^{-1} q^{\lambda} & \tilde{f}_3 q^{m_1-m_3} \\ \eta_1^{-1} & \eta_2^{-1} q^{2\lambda} + A\eta_2 q^{n_1+n_2} & \eta_3^{-1} q^{\lambda} \end{vmatrix} + o(q^{m_1}) + o(q^{n_1+n_2}) \\ &= \tilde{f}_3 \eta_1^{-1} q^{m_1-m_3} + A\varepsilon_1^{-1} \eta_2 q^{n_1+n_2} + o(q^{m_1-m_3}) + o(q^{n_1+n_2}). \end{aligned}$$

Here the term with $q^{m_1-m_3}$ cannot be eliminated by $o(q^{n_1+n_2})$ since then $m_1 - m_3 > n_1 + n_2$ and after elimination $q^{n_1+n_2}$ would still be standing. So

$$m_1 - m_3 = n_1 + n_2, \qquad \tilde{f}_3 \eta_1^{-1} = -A\varepsilon_1^{-1} \eta_2. \tag{12-6}$$

However, the second equality in (12-6) is impossible. Indeed, if $m_3 > 0$, then it becomes $\varepsilon_3^{-1} \eta_1^{-1} = -A\varepsilon_1^{-1} \eta_2$, which is clearly impossible because $A = 196884$ is not a root of unity. And if $m_3 = 0$, then it becomes $f_3 = 744 - A\varepsilon_1^{-1} \eta_1 \eta_2$, contradicting Lemma 5.7.

This proves (12-4). We set $m_1 - m_3 = n_1 - n_2 = \lambda'$. Since $m_1 \geq n_1$ by (8-7), we may summarize our present knowledge as

$$\begin{gathered} m_1 > m_2 > m_3, \qquad n_1 > n_3 > n_2, \\ m_1 - m_2 = n_1 - n_3 = \lambda > 0, \qquad m_1 - m_3 = n_1 - n_2 = \lambda' > \lambda, \\ m_1 - n_1 = m_2 - n_3 = m_3 - n_2 \geq 0. \end{gathered}$$

12B. ***Proof of $m_3 > 0$.*** In this subsection we prove that $m_3 > 0$. We will assume that $m_3 = 0$ and will arrive at a contradiction.

If $m_3 = 0$, then

$$m_1 = n_1 = \lambda', \qquad m_2 = n_3, \qquad m_3 = n_2 = 0. \tag{12-7}$$

Using (12-3), we obtain

$$\begin{aligned} \mathfrak{D}(q) &= \begin{vmatrix} 1 & 1 & 1 \\ \varepsilon_1^{-1} & \varepsilon_2^{-1} q^{\lambda} + A\varepsilon_2 q^{m_1+m_2} & \tilde{f}_3 q^{m_1} \\ \eta_1^{-1} & \tilde{g}_2 q^{m_1} & \eta_3^{-1} q^{\lambda} + A\eta_3 q^{m_1+m_2} \end{vmatrix} + o(q^{m_1+m_2}) \\ &= \begin{vmatrix} \varepsilon_1^{-1} & \tilde{f}_3 \\ -\eta_1^{-1} & \tilde{g}_2 \end{vmatrix} q^{m_1} + \varepsilon_2^{-1}\eta_3^{-1} q^{2\lambda} + A \begin{vmatrix} \varepsilon_1^{-1} & \varepsilon_2 \\ \eta_1^{-1} & -\eta_3 \end{vmatrix} q^{m_1+m_2} + o(q^{m_1+m_2}). \end{aligned} \tag{12-8}$$

The term with $q^{m_1+m_2}$ can be eliminated if either

$$\begin{vmatrix} \varepsilon_1^{-1} & \varepsilon_2 \\ \eta_1^{-1} & -\eta_3 \end{vmatrix} = 0, \tag{12-9}$$

or $m_1 + m_2 = 2\lambda$ and

$$A \begin{vmatrix} \varepsilon_1^{-1} & \varepsilon_2 \\ \eta_1^{-1} & -\eta_3 \end{vmatrix} = -\varepsilon_2^{-1}\eta_3^{-1}. \tag{12-10}$$

However, (12-10) is impossible because A does not divide a root of unity. Hence, we have (12-9). Together with (12-3), this implies that

$$(\varepsilon_1, \varepsilon_2) = \theta(\eta_1, -\eta_3), \quad \theta = \pm 1. \tag{12-11}$$

The rest of this subsection splits into three cases depending on the relation between m_2 and λ.

The case $m_2 > \lambda$. In this case $m_1 > 2\lambda$ and $q^{2\lambda}$ in (12-8) cannot be eliminated.

The case $m_2 < \lambda$. In this case $m_1 < 2\lambda$, and q^{m_1} in (12-8) can be eliminated only if $\varepsilon_1^{-1}\tilde{g}_2 + \eta_1^{-1}\tilde{f}_3 = 0$, which, combined with (12-11), gives $\tilde{g}_2 = -\theta\tilde{f}_3$. Lemma 5.2 implies that $\theta = -1$ and $\tilde{f}_3 = \tilde{g}_2$, that is, $f_3 = g_2$. Also, since $\theta = -1$, we obtain $\varepsilon_2 = \eta_3$, which, together with $m_2 = n_3$ (see (12-7)), implies that $f_2 = g_3$. This contradicts Lemma 8.1.

The case $m_2 = \lambda$. In this case $m_1 = 2\lambda < m_1 + m_2$ and $\varepsilon_1^{-1}\tilde{g}_2 + \eta_1^{-1}\tilde{f}_3 + \varepsilon_2^{-1}\eta_3^{-1} = 0$, which contradicts Lemma 5.5.

This completes the proof of impossibility of $m_3 = 0$.

12C. ***Proof of $n_2 > 0$.*** Thus, we have $m_3 > 0$. Let us now prove that $n_2 > 0$ as well. Indeed, if $n_2 = 0$, then

$$m_1 > n_1 = \lambda', \qquad m_2 > n_3, \qquad m_3 > n_2 = 0. \tag{12-12}$$

Using (12-3), we obtain

$$\begin{aligned}\mathcal{D}(q) &= \begin{vmatrix} 1 & 1 & 1 \\ \varepsilon_1^{-1} & \varepsilon_2^{-1}q^{\lambda} & \varepsilon_3^{-1}q^{n_1} \\ \eta_1^{-1} & \tilde{g}_2 q^{n_1} & \eta_3^{-1}q^{\lambda} \end{vmatrix} + o(q^{n_1}) \\ &= (\varepsilon_1^{-1}\tilde{g}_2 + \varepsilon_3^{-1}\eta_1^{-1})q^{n_1} + \varepsilon_2^{-1}\eta_3^{-1}q^{2\lambda} + o(q^{n_1}).\end{aligned}$$

Now to eliminate q^{n_1} we need to have one of the following:

$$\varepsilon_1^{-1}\tilde{g}_2 + \varepsilon_3^{-1}\eta_1^{-1} = 0, \tag{12-13}$$

$$\varepsilon_1^{-1}\tilde{g}_2 + \varepsilon_3^{-1}\eta_1^{-1} + \varepsilon_2^{-1}\eta_3^{-1} = 0. \tag{12-14}$$

However, since $\tilde{g}_2 = g_2 - 744$, (12-13) contradicts Lemma 5.7. Furthermore, applying Lemma 5.9 to (12-14), we obtain $g_2 \in \{744, 744\pm1, 744\pm2\}$, contradicting Lemma 5.6.

This proves that $n_2 > 0$. Let us summarize our present knowledge as

$$\begin{gathered} m_1 > m_2 > m_3 > 0, \qquad n_1 > n_3 > n_2 > 0, \\ m_1 - m_2 = n_1 - n_3 = \lambda > 0, \qquad m_1 - m_3 = n_1 - n_2 = \lambda' > \lambda, \\ m_1 - n_1 = m_2 - n_3 = m_3 - n_2 \geq 0. \end{gathered}$$

12D. ***Proof of $m_1 = n_1$.*** Next, we show that $m_1 = n_1$. Thus, assume that $m_1 > n_1$. Then we also have $m_2 > n_3$ and $m_3 > n_2$. Using (12-3), we write

$$\begin{aligned}\mathcal{D}(q) &= \begin{vmatrix} 1 & 1 & 1 \\ \varepsilon_1^{-1} & \varepsilon_2^{-1}q^{\lambda} & \varepsilon_3^{-1}q^{\lambda'} \\ \eta_1^{-1} & \eta_2^{-1}q^{\lambda'} + A\eta_2 q^{n_1+n_2} & \eta_3^{-1}q^{\lambda} \end{vmatrix} + o(q^{n_1+n_2}) \\ &= \begin{vmatrix} \varepsilon_1^{-1} & \varepsilon_3^{-1} \\ -\eta_1^{-1} & \eta_2^{-1} \end{vmatrix} q^{\lambda'} + \varepsilon_2^{-1}\eta_3^{-1}q^{2\lambda} - \varepsilon_3^{-1}\eta_2^{-1}q^{2\lambda'} + A\varepsilon_1^{-1}\eta_2 q^{n_1+n_2} \\ &\qquad + o(q^{n_1+n_2}). \end{aligned} \tag{12-15}$$

To eliminate $q^{n_1+n_2}$ we need one of the following to hold:

$$2\lambda = n_1 + n_2, \qquad \varepsilon_2^{-1}\eta_3^{-1} = -A\varepsilon_1^{-1}\eta_2, \tag{12-16}$$

$$2\lambda' = n_1 + n_2, \qquad \varepsilon_3^{-1}\eta_2^{-1} = A\varepsilon_1^{-1}\eta_2. \tag{12-17}$$

However, the second equations in both (12-16) and (12-17) cannot be true because A is not a root of unity.

This proves that $m_1 = n_1$. Moreover,

$$m_1 = n_1 > m_2 = n_3 > m_3 = n_2 > 0, \tag{12-18}$$

$$m_1 - m_2 = n_1 - n_3 = \lambda > 0, \qquad m_1 - m_3 = n_1 - n_2 = \lambda' > \lambda.$$

12E. ***Proof of*** $\boldsymbol{\lambda' = 2\lambda}$**.** Our next quest is proving that $\lambda' = 2\lambda$. Using (12-3) and (12-18), we obtain

$$\mathcal{D}(q) = \begin{vmatrix} 1 & 1 & 1 \\ \varepsilon_1^{-1} & \varepsilon_2^{-1}q^{\lambda} & \varepsilon_3^{-1}q^{\lambda'} \\ \eta_1^{-1} & \eta_2^{-1}q^{\lambda'} & \eta_3^{-1}q^{\lambda} \end{vmatrix} + o(q^{m_1}) = -\begin{vmatrix} \varepsilon_1^{-1} & \varepsilon_3^{-1} \\ \eta_1^{-1} & -\eta_2^{-1} \end{vmatrix} q^{\lambda'} + \varepsilon_2^{-1}\eta_3^{-1}q^{2\lambda} + o(q^{\lambda'}).$$

This already implies that $\lambda' \le 2\lambda$; otherwise $q^{2\lambda}$ cannot be eliminated.

The proof of the opposite inequality $\lambda' \ge 2\lambda$ is much more involved. Thus, assume that $\lambda' < 2\lambda$. Then we must have

$$\begin{vmatrix} \varepsilon_1^{-1} & \varepsilon_3^{-1} \\ \eta_1^{-1} & -\eta_2^{-1} \end{vmatrix} = 0.$$

Together with (12-3) this implies that

$$(\eta_1, -\eta_3, -\eta_2) = \theta(\varepsilon_1, \varepsilon_2, \varepsilon_3), \tag{12-19}$$

where θ is some root of unity. We obtain

$$\mathcal{D}(q) = \begin{vmatrix} 1 & 1 & 1 \\ \varepsilon_1^{-1} & \varepsilon_2^{-1}q^{\lambda} & \varepsilon_3^{-1}q^{\lambda'} + A\varepsilon_3 q^{m_1+m_3} \\ \theta^{-1}\varepsilon_1^{-1} & -\theta^{-1}\varepsilon_3^{-1}q^{\lambda'} - A\theta\varepsilon_3 q^{m_1+m_3} & -\theta^{-1}\varepsilon_2^{-1}q^{\lambda} \end{vmatrix} + o(q^{m_1+m_3})$$

$$= -\theta^{-1}\varepsilon_2^{-2}q^{2\lambda} + \theta^{-1}\varepsilon_3^{-2}q^{2\lambda'} + A\varepsilon_3\varepsilon_1^{-1}(\theta^{-1} - \theta)q^{m_1+m_3} + o(q^{m_1+m_3}).$$

To eliminate $q^{m_1+m_3}$ one of the following should be satisfied:

$$A\varepsilon_3\varepsilon_1^{-1}(\theta^{-1} - \theta) = \theta^{-1}\varepsilon_2^{-2}, \quad A\varepsilon_3\varepsilon_1^{-1}(\theta^{-1} - \theta) = -\theta^{-1}\varepsilon_3^{-2}, \quad A\varepsilon_3\varepsilon_1^{-1}(\theta^{-1} - \theta) = 0.$$

Since A does not divide a root of unity, only the third equation is possible, which implies $\theta = \pm 1$. If $\theta = -1$, then (12-18) and (12-19) imply that $f_2 = g_3$ and $f_3 = g_2$, contradicting Lemma 8.1. Thus, $\theta = 1$ and

$$(\eta_1, -\eta_3, -\eta_2) = (\varepsilon_1, \varepsilon_2, \varepsilon_3),$$

which gives us the relations

$$q^{m_1}(g_1-744)=q^{m_1}(f_1-744),$$
$$q^{m_1}(g_3-744)=-q^{m_1}(f_2-744)+O(q^{m_1+2m_2}),$$
$$q^{m_1}(g_2-744)=-q^{m_1}(f_3-744)+2B\varepsilon_3^2q^{m_1+2m_3}+o(q^{m_1+2m_3}).$$

Using this, and the identity

$$\begin{vmatrix}1 & 1 & 1\\ a & b & c\\ a & -c+x & -b\end{vmatrix}=c^2-b^2+x(a-c),$$

we obtain

$$\begin{aligned}\mathscr{D}(q)&=\begin{vmatrix}1 & 1 & 1\\ q^{m_1}(f_1-744) & q^{m_1}(f_2-744) & q^{m_1}(f_3-744)\\ q^{m_1}(f_1-744) & -q^{m_1}(f_3-744)+2B\varepsilon_3^2q^{m_1+2m_3} & -q^{m_1}(f_2-744)\end{vmatrix}\\ &\qquad+o(q^{m_1+2m_3})\\ &=2B\varepsilon_1^{-1}\varepsilon_3^2q^{m_1+2m_3}+(\varepsilon_3^{-1}q^{m_1-m_3}+A\varepsilon_3q^{m_1+m_3})^2\\ &\qquad-(\varepsilon_2^{-1}q^{m_1-m_2}+A\varepsilon_2q^{m_1+m_2})^2+o(q^{m_1+2m_3})\\ &=-\varepsilon_2^{-2}q^{2\lambda}+\varepsilon_3^{-2}q^{2\lambda'}+2B\varepsilon_1^{-1}\varepsilon_3^2q^{m_1+2m_3}+o(q^{m_1+2m_3})\end{aligned}$$

(recall that $\lambda=m_1-m_2$ and $\lambda'=m_1-m_3$). We see that to eliminate $q^{m_1+2m_3}$ we need to have either $2B\varepsilon_1^{-1}\varepsilon_3^2=\varepsilon_2^{-2}$ or $2B\varepsilon_1^{-1}\varepsilon_3^2=-\varepsilon_3^{-2}$; both are clearly impossible.

This proves that $\lambda'=2\lambda$. Thus,

$$m_1=n_1,\qquad m_2=n_3=m_1-\lambda,\qquad m_3=n_2=m_1-2\lambda>0. \tag{12-20}$$

12F. ***Proof of*** $2\lambda<m_1<3\lambda$**.** Now it is not difficult to show that

$$2\lambda<m_1<3\lambda. \tag{12-21}$$

In fact, $m_1>2\lambda$ is already in (12-20). Next, using (12-3), we obtain

$$\begin{aligned}\mathscr{D}(q)&=\begin{vmatrix}1 & 1 & 1\\ \varepsilon_1^{-1} & \varepsilon_2^{-1}q^{\lambda} & \varepsilon_3^{-1}q^{2\lambda}+A\varepsilon_3q^{m_1+m_3}\\ \eta_1^{-1} & \eta_2^{-1}q^{2\lambda}+A\eta_2q^{m_1+m_3} & \eta_3^{-1}q^{\lambda}\end{vmatrix}+o(q^{m_1+m_3})\\ &=(\varepsilon_1^{-1}\eta_2^{-1}+\varepsilon_3^{-1}\eta_1^{-1}+\varepsilon_2^{-1}\eta_3^{-1})q^{2\lambda}-\varepsilon_3^{-1}\eta_2^{-1}q^{4\lambda}-A\begin{vmatrix}\varepsilon_1^{-1} & \varepsilon_3\\ \eta_1^{-1} & -\eta_2\end{vmatrix}q^{m_1+m_3}\\ &\qquad+o(q^{m_1+m_3}).\end{aligned}$$

Since $m_1>2\lambda$, this gives

$$\varepsilon_1^{-1}\eta_2^{-1}+\varepsilon_3^{-1}\eta_1^{-1}+\varepsilon_2^{-1}\eta_3^{-1}=0. \tag{12-22}$$

Further, if $4\lambda < m_1 + m_3$, then $q^{4\lambda}$ cannot be eliminated. And if $4\lambda = m_1 + m_3$, then

$$-\varepsilon_3^{-1}\eta_2^{-1} = A\begin{vmatrix}\varepsilon_1^{-1} & \varepsilon_3\\ \eta_1^{-1} & -\eta_2\end{vmatrix},$$

which is impossible because A does not divide a root of unity.

Thus, we have $4\lambda > m_1 + m_3 = 2m_1 - 2\lambda$, that is, $m_1 < 3\lambda$, proving (12-21). In addition to this, to eliminate $q^{m_1+m_3}$ we need to have

$$\begin{vmatrix}\varepsilon_1^{-1} & \varepsilon_3\\ \eta_1^{-1} & -\eta_2\end{vmatrix} = 0.$$

Together with (12-3) this implies that

$$(\eta_1^{-1}, -\eta_3^{-1}, -\eta_2) = \theta(\varepsilon_1^{-1}, \varepsilon_2^{-1}, \varepsilon_3) \tag{12-23}$$

for some root of unity θ.

12G. ***Conclusion.*** It follows from (12-21) that $m_3 < \lambda$, whence

$$m_1 + 2m_3 < m_1 + m_3 + \lambda = m_1 + m_2 < 2m_1.$$

Using this, (12-3), (12-22), and (12-23), we obtain for $\mathcal{D}(q)$ the value

$$\begin{vmatrix}1 & 1 & 1\\ \varepsilon_1^{-1} & \varepsilon_2^{-1}q^{\lambda} & \varepsilon_3^{-1}q^{2\lambda} + A\varepsilon_3 q^{m_1+m_3} + B\varepsilon_3^2 q^{m_1+2m_3}\\ \eta_1^{-1} & \eta_2^{-1}q^{2\lambda} + A\eta_2 q^{m_1+m_3} + B\eta_2^2 q^{m_1+2m_3} & \eta_3^{-1}q^{\lambda}\end{vmatrix} + o(q^{m_1+2m_3})$$

$$= -\varepsilon_3^{-1}\eta_2^{-1}q^{4\lambda} - B\begin{vmatrix}\varepsilon_1^{-1} & \varepsilon_3^2\\ \eta_1^{-1} & -\eta_2^2\end{vmatrix} q^{m_1+2m_3} + o(q^{m_1+2m_3}).$$

Arguing as in Section 12F, we obtain from this $4\lambda > m_1 + 2m_3$ and

$$\begin{vmatrix}\varepsilon_1^{-1} & \varepsilon_3^2\\ \eta_1^{-1} & -\eta_2^2\end{vmatrix} = 0,$$

which, together with (12-23), implies that $\theta = -1$. It follows that $\eta_2 = \varepsilon_3$ and $\eta_3 = \varepsilon_2$; together with (12-18) this implies $g_2 = f_3$ and $g_3 = f_2$, contradicting Lemma 8.1.

This completes the proof of impossibility of (12-1). The main lemma is now fully proved.

Acknowledgments

Yuri Bilu was supported by the Agence National de la Recherche project "Hamot" (ANR 2010 BLAN-0115-01). We thank Bill Allombert, Qing Liu, Pierre Parent, Jonathan Pila, and Thomas Scanlon for useful discussions. We also thank the referee,

who did the hard job of verifying the proof, detected a number of inaccuracies, and made many helpful suggestions.

References

[Allombert et al. 2015] B. Allombert, Yu. Bilu, and A. Pizarro-Madariaga, "CM-points on straight lines", pp. 1–18 in *Analytic number theory*, edited by C. Pomerance and M. Th. Rassias, Springer, 2015. MR Zbl

[André 1998] Y. André, "Finitude des couples d'invariants modulaires singuliers sur une courbe algébrique plane non modulaire", *J. Reine Angew. Math.* **505** (1998), 203–208. MR Zbl

[Bilu et al. 2013] Yu. Bilu, D. Masser, and U. Zannier, "An effective 'Theorem of André' for CM-points on a plane curve", *Math. Proc. Cambridge Philos. Soc.* **154**:1 (2013), 145–152. MR Zbl

[Bilu et al. 2016] Yu. Bilu, F. Luca, and A. Pizarro-Madariaga, "Rational products of singular moduli", *J. Number Theory* **158** (2016), 397–410. MR Zbl

[Breuer 2001] F. Breuer, "Heights of CM points on complex affine curves", *Ramanujan J.* **5**:3 (2001), 311–317. MR Zbl

[Cox 1989] D. A. Cox, *Primes of the form $x^2 + ny^2$: Fermat, class field theory, and complex multiplication*, Wiley, 1989. MR Zbl

[Diamond and Shurman 2005] F. Diamond and J. Shurman, *A first course in modular forms*, Graduate Texts in Mathematics **228**, Springer, 2005. MR Zbl

[Edixhoven 1998] B. Edixhoven, "Special points on the product of two modular curves", *Compositio Math.* **114**:3 (1998), 315–328. MR Zbl

[Evertse et al. 1988] J.-H. Evertse, K. Győry, C. L. Stewart, and R. Tijdeman, "On S-unit equations in two unknowns", *Invent. Math.* **92**:3 (1988), 461–477. MR Zbl

[Habegger et al. 2017] P. Habegger, G. Jones, and D. Masser, "Six unlikely intersection problems in search of effectivity", *Math. Proc. Cambridge Philos. Soc.* **162**:3 (2017), 447–477. MR

[Ireland and Rosen 1990] K. Ireland and M. Rosen, *A classical introduction to modern number theory*, 2nd ed., Graduate Texts in Mathematics **84**, Springer, 1990. MR Zbl

[Kühne 2012] L. Kühne, "An effective result of André–Oort type", *Ann. of Math.* (2) **176**:1 (2012), 651–671. MR Zbl

[Kühne 2013] L. Kühne, "An effective result of André–Oort type, II", *Acta Arith.* **161** (2013), 1–19. MR Zbl

[Pila 2009] J. Pila, "Rational points of definable sets and results of André–Oort–Manin–Mumford type", *Int. Math. Res. Not.* **2009**:13 (2009), 2476–2507. MR Zbl

[Pila 2011] J. Pila, "O-minimality and the André–Oort conjecture for $\mathbb{C}^n$", *Ann. of Math.* (2) **173**:3 (2011), 1779–1840. MR Zbl

[Pila and Zannier 2008] J. Pila and U. Zannier, "Rational points in periodic analytic sets and the Manin–Mumford conjecture", *Atti Accad. Naz. Lincei Rend. Lincei Mat. Appl.* **19**:2 (2008), 149–162. MR Zbl

[Scanlon 2004] T. Scanlon, "Automatic uniformity", *Int. Math. Res. Not.* **2004**:62 (2004), 3317–3326. MR Zbl

[Schlickewei and Wirsing 1997] H. P. Schlickewei and E. Wirsing, "Lower bounds for the heights of solutions of linear equations", *Invent. Math.* **129**:1 (1997), 1–10. MR Zbl

Communicated by Jonathan Pila
Received 2016-01-02 Revised 2016-11-27 Accepted 2017-03-31

yuri@math.u-bordeaux.fr *Institut de Mathématiques de Bordeaux, Université de Bordeaux et CNRS, Talence, France*

florian.luca@wits.ac.za *School of Mathematics, University of the Witwatersrand, Johannesburg, South Africa*

Max Planck Institute for Mathematics, Bonn, Germany

david.masser@unibas.ch *Mathematisches Institut, Universität Basel, Basel, Switzerland*

Algebra & Number Theory

msp.org/ant

See inside back cover or msp.org/ant for submission instructions.

The subscription price for 2017 is US $325/year for the electronic version, and $520/year (+$55, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues and changes of subscriber address should be sent to MSP.

Algebra & Number Theory (ISSN 1944-7833 electronic, 1937-0652 printed) at Mathematical Sciences Publishers, 798 Evans Hall #3840, c/o University of California, Berkeley, CA 94720-3840 is published continuously online. Periodical rate postage paid at Berkeley, CA 94704, and additional mailing offices.

ANT peer review and production are managed by EditFLOW® from MSP.

PUBLISHED BY

mathematical sciences publishers

nonprofit scientific publishing

http://msp.org/

Algebra & Number Theory

Volume 11 No. 5 2017