

Effective bounds for singular units

Yu. Bilu, P. Habegger and L. Kühne

To David Masser

Abstract

A result of the second-named author states that there are only finitely many CM-elliptic curves over \mathbb{C} whose j -invariant is an algebraic unit. His proof depends on Duke's Equidistribution Theorem and is hence non-effective. In this article, we give a completely effective proof of this result. To be precise, we show that every singular modulus that is an algebraic unit is associated with a CM-elliptic curve whose endomorphism ring has discriminant less than 10^{15} . Through further refinements and computer-assisted computations, we eventually rule out all remaining cases, showing that no singular modulus is an algebraic unit. This allows us to exhibit classes of subvarieties in \mathbb{C}^n not containing any special points.

Contents

1	Introduction	1
2	An estimate for $\mathcal{C}_\varepsilon(\Delta)$	6
3	An upper bound for the height of a singular unit	10
4	Lower bounds for the height of a singular modulus	14
5	The estimate $\Delta < 10^{15}$	17
6	Handling the mid-range $10^{10} \leq \Delta < 10^{15}$	20
7	Handling the low-range $3 \cdot 10^5 \leq \Delta < 10^{10}$	26
8	The extra low-range	28

1 Introduction

Since the nineteenth century, j -invariants associated with elliptic curves having complex multiplication (CM), the so-called singular moduli, have been an object of study in number theory. A theorem of Weber [7, Theorem 11.1] states that every singular modulus is an algebraic integer. Under certain technical restrictions, Gross and Zagier [12] stated explicit formulas for the absolute norm of the difference between two singular moduli.

Motivated by effective results of André-Oort type [4, 17], David Masser raised in 2011 the question whether only finitely many singular moduli are algebraic units, that is, units of the ring of all algebraic integers.

Throughout this article, we call such hypothetical algebraic numbers *singular units*. Since there is no example of a singular unit in the literature, it seems legitimate to ask whether there are any singular units at all.

In [14], the second-named author answered Masser’s original question in the affirmative: There exist at most finitely many singular units. However, his proof is non-effective as it invokes Siegel’s lower bounds on the class number of imaginary quadratic fields [25] through Duke’s Equidistribution Theorem [10].

Here, we can give the following definite answer to Masser’s question as our main theorem.

Theorem 1.1. *There are no singular units.*

As already mentioned in [14], the explicit formulas of Gross and Zagier [12], which were extended by Dorman [9], can be used to compute the absolute norm of many singular moduli. However, it seems hard to extract a result like the above theorem from these formulas.

Theorem 1.1 is a formal consequence of our Theorems 5.1, 6.1, 7.1 and 8.1. Let us briefly sketch its proof. We say that a singular modulus is of discriminant Δ if it is the j -invariant of a CM-elliptic curve whose endomorphism ring is the imaginary quadratic order of discriminant Δ . We also write $\Delta = Df^2$ where D is the discriminant of the CM-field $\mathbb{Q}(\sqrt{\Delta})$, the fundamental discriminant, and f is the conductor of the endomorphism order. The singular moduli of a given discriminant Δ form a full Galois orbit over \mathbb{Q} of cardinality equal to the class number¹ $\mathcal{C}(\Delta)$.

Write ζ_3 (resp. ζ_6) for the third (resp. sixth) root of unity $e^{2\pi i/3}$ (resp. $e^{\pi i/3}$). Note that ζ_3 (resp. ζ_6) is the left (resp. right) vertex of the geodesic triangle enclosing the standard fundamental domain \mathcal{F} in the Poincaré upper half-plane. Given $\varepsilon \in (0, 1/3]$, denote by $\mathcal{C}_\varepsilon(\Delta)$ the number of singular moduli of discriminant Δ which can be written $j(\tau)$ where $\tau \in \mathcal{F}$ satisfies $|\tau - \zeta_3| < \varepsilon$ or $|\tau - \zeta_6| < \varepsilon$ and $j(\cdot)$ denotes Klein’s j -function. Since ζ_3 and ζ_6 are the only zeros of the j -function contained in the closure of \mathcal{F} , a pivotal ingredient in the proof of Theorem 1.1 is an upper bound on $\mathcal{C}_\varepsilon(\Delta)$. Indeed, a main point of the argument in [14] is the estimate

$$\mathcal{C}_\varepsilon(\Delta) \ll \mathcal{C}(\Delta)\varepsilon^2$$

which holds when $|\Delta|$ is sufficiently large (in terms of ε). Unfortunately, “sufficiently large” here is not effective; in fact, this is the place where Duke’s Equidistribution Theorem [10], generalized by Clozel and Ullmo [5] to arbitrary discriminants, is used.

Our main novelty is the following effective estimate (see Theorem 2.1):

$$\mathcal{C}_\varepsilon(\Delta) \ll F \left(\frac{\sigma_1(f)}{f} |\Delta|^{1/2} \varepsilon^2 + |\Delta|^{1/2} \varepsilon + \sigma_0(f) |\Delta|^{1/4} \varepsilon + 1 \right), \quad (1.1)$$

where

$$F = F(\Delta) = \max\{2^{\omega(a)} : a \leq |\Delta|^{1/2}\}.$$

¹We do not use the more traditional notation $h(\Delta)$ because of the risk of confusing it with the height $h(\cdot)$.

Here and in the sequel all implicit constants are effective, and we use the standard notation

$$\omega(n) = \sum_{p|n} 1, \quad \sigma_0(n) = \sum_{d|n} 1, \quad \sigma_1(n) = \sum_{d|n} d. \quad (1.2)$$

Using that

$$\omega(n) = o(\log n), \quad \log \sigma_0(n) = o(\log n), \quad \sigma_1(n) \ll n \log \log n, \quad (1.3)$$

we deduce from (1.1) that

$$\mathcal{C}_\varepsilon(\Delta) \leq |\Delta|^{1/2+o(1)} \varepsilon + |\Delta|^{o(1)} \quad (1.4)$$

as $|\Delta| \rightarrow \infty$.

The height of a singular unit α can be easily estimated in terms of $\mathcal{C}_\varepsilon(\Delta)$:

$$h(\alpha) \ll \frac{\mathcal{C}_\varepsilon(\Delta)}{\mathcal{C}(\Delta)} \log |\Delta| + \log(\varepsilon^{-1}), \quad (1.5)$$

see Theorem 3.1. Substituting (1.4), we obtain the following upper estimate:

$$h(\alpha) \leq \frac{|\Delta|^{1/2}}{\mathcal{C}(\Delta)} A \varepsilon + \frac{|\Delta|^{o(1)}}{\mathcal{C}(\Delta)} + O(\log(\varepsilon^{-1}))$$

where $A = A(\Delta) = |\Delta|^{o(1)}$ as $|\Delta| \rightarrow \infty$. Specifying

$$\varepsilon = \frac{\mathcal{C}(\Delta)}{A|\Delta|^{1/2}}$$

(which is a nearly optimal choice), we obtain the estimate

$$h(\alpha) \leq \frac{|\Delta|^{o(1)}}{\mathcal{C}(\Delta)} + o(\log |\Delta|) + O\left(\log^+ \frac{|\Delta|^{1/2}}{\mathcal{C}(\Delta)}\right)$$

where $\log^+(x) = \max\{\log x, 0\}$.

To obtain an upper bound on $|\Delta|$, we combine this bound with the following two lower estimates on $h(\alpha)$ (see Section 4)

$$h(\alpha) + 1 \gg \log |\Delta|, \quad (1.6)$$

$$h(\alpha) \gg \frac{|\Delta|^{1/2}}{\mathcal{C}(\Delta)}. \quad (1.7)$$

The bound (1.6) is rather deep and relies on work of Colmez [6] and Nakkajima-Taguchi [18]. On the contrary, (1.7) follows easily from the fact that one of the conjugates of our singular unit α is $j((\Delta + \sqrt{\Delta})/2)$. Nevertheless, (1.7) plays a crucial role when the class number is pathologically small so that it would contradict the Generalized Riemann Hypothesis (GRH). In fact, (1.7) becomes much stronger than (1.6) in these hypothetical cases.

Comparing upper and lower estimates, we obtain for large $|\Delta|$ that

$$\max\left\{\frac{|\Delta|^{1/2}}{\mathcal{C}(\Delta)}, \log |\Delta|\right\} \leq \frac{|\Delta|^{o(1)}}{\mathcal{C}(\Delta)} + o(\log |\Delta|) + O\left(\log^+ \frac{|\Delta|^{1/2}}{\mathcal{C}(\Delta)}\right),$$

which is clearly impossible.

To get an explicit bound on $|\Delta|$, we need to replace all implicit constants above with explicit ones. This relies in particular on a numerically sharp estimate for the arithmetic function $\omega(n)$ due to Robin [22]. In Section 5, we see that this leads to a bound $|\Delta| < 10^{15}$. While already effective, it is still not feasible to check directly by a computer-assisted proof that none of the singular moduli of discriminant $\Delta \in [-3, -10^{15})$ is an algebraic unit.

A refinement of our original arguments comes to our rescue. When $|\Delta| < 10^{15}$, we improve on the estimate (1.1) by bounding sums of the form $\sum_{n \in [a, b] \cap \mathbb{Z}} 2^{\omega(n)}$ in a more refined way. A natural idea is to use the Selberg-Delange method, which yields the asymptotic expansion

$$S(x) = \sum_{n \leq x} 2^{\omega(n)} = \lambda_0 x \log(x) + \lambda_1 x + O(e^{-c \log(x)}) \quad (1.8)$$

with explicit constants $\lambda_0, \lambda_1 \in \mathbb{C}$ and some positive constant $c > 0$ (see, for instance, [26, Theorem II.6.1]). There are two downsides of this method. First, the error term is suboptimal under assumption of the GRH. Second, it would need some effort to make the constant c actually explicit.

However, as $|\Delta| < 10^{15}$, we are only interested in the case where $[a, b]$ is a subinterval of $[1, 2 \cdot 10^7]$. In this range, a simple **SAGE** script using the **MPFI** library [21, 29] can be used to improve on (1.8) computationally (see Proposition 6.2). As a consequence, we obtain $|\Delta| < 10^{10}$ for any singular unit of discriminant Δ in Theorem 6.1.

This is still not sufficient to check all remaining cases, at least with modest computational means. The range is nevertheless small enough to use a counting algorithm in order to bound $\mathcal{C}_{10^{-3}}(\Delta)$ for all discriminants Δ satisfying $|\Delta| < 10^{10}$, see Lemma 7.2. This still needs a clever counting strategy, as determining $\mathcal{C}_{10^{-3}}(\Delta)$ for each discriminant is rather slow, comparable to computing separately each class number $\mathcal{C}(\Delta)$ in the same range. Our trick is to bound all $\mathcal{C}_{10^{-3}}(\Delta)$ simultaneously by running through a set containing all imaginary quadratic $\tau \in \mathcal{F}$ satisfying $|\tau - \zeta_3| < \varepsilon$ or $|\tau - \zeta_6| < \varepsilon$ and such that $j(\tau)$ is of discriminant Δ with $|\Delta| < 10^{10}$. For each τ encountered, we compute its discriminant $\Delta(\tau)$ after the fact and increment our counter for $\mathcal{C}_{10^{-3}}(\Delta(\tau))$. The thus obtained bounds for $\mathcal{C}_{10^{-3}}(\Delta)$ refine once again our previous inequalities, and allow us to conclude that $|\Delta| < 10^7$. Repeating this procedure once again, with a slightly changed ε , we achieve even $|\Delta| < 3 \cdot 10^5$ in Theorem 7.1. These remaining cases can now be dealt with directly, for which we use a **PARI** [28] program to prove Theorem 8.1, completing thereby the proof of Theorem 1.1.

It is very probable that our argument can be adapted to solve a more general problem: given an algebraic integer β , determine the singular moduli α such that $\alpha - \beta$ is a unit; or at least bound effectively the discriminants of such α . For instance, one may ask whether 0 is the only singular modulus α such that $\alpha - 1$ is a unit. In the general case, as explained in [14], this would require lower bounds for elliptic logarithmic

forms, but when β itself is a singular modulus, our argument extends almost without changes. One may go further and obtain an effective version of Theorem 2 from [14], which is an analogue of Siegel’s Finiteness Theorem for special points.

The already mentioned work of Gross-Zagier and Dorman [12, 9] inspires the following problem: determine all couples (α, β) of singular moduli such that $\alpha - \beta$ is a unit; presumably, there is none. As indicated above, when β is fixed and α varying, a version of our argument does the job, but if we let both α and β vary, the problem seems substantially more intricate.

Finally, let us discuss an application of Theorem 1.1 to effective results of André-Oort type. A point $(\alpha_1, \dots, \alpha_n) \in \mathbb{C}^n$ is called special if each α_i , $i \in \{1, \dots, n\}$, is a singular moduli. Since singular moduli are algebraic integers, the following statement is an immediate consequence of our main result.

Corollary 1.2. *For each polynomial P in unknowns X_2, \dots, X_n and coefficients that are algebraic integers in \mathbb{C} , the hypersurface defined by*

$$X_1 P(X_2, \dots, X_n) = 1$$

contains no special points.

In particular, $\alpha_1^{a_1} \cdots \alpha_n^{a_n} \neq 1$ for all special points $(\alpha_1, \dots, \alpha_n)$ and all integers $a_1 \geq 1, \dots, a_n \geq 1$. This corollary exhibits a rather general class of algebraic varieties of arbitrary dimension and degree for which the celebrated theorem of Pila [20] can be proved effectively and even explicitly. It is complementary to other effective results of André-Oort type recently obtained by the first- and third-named author [2].

Plan of the article In Section 2 we obtain an explicit version of the estimate (1.1). In Section 3 we obtain an upper estimate for the height of a singular unit. In Section 4 we obtain explicit versions of the lower estimates (1.6) and (1.7). In Section 5 we use all previous results to bound the discriminant of a singular unit as $|\Delta| < 10^{15}$. This bound is reduced to 10^{10} in Section 6 and to $3 \cdot 10^5$ in Section 7. Finally, in Section 8 we show that a discriminant of a singular unit satisfies $|\Delta| > 3 \cdot 10^5$.

Convention In this article we fix, once and for all, an embedding $\bar{\mathbb{Q}} \hookrightarrow \mathbb{C}$; this means that all algebraic numbers in this article are viewed as elements of \mathbb{C} .

Acknowledgments Yuri Bilu was partially supported by the University of Basel, the Fields Institute (Toronto), and the Xiamen University. Lars Kühne was supported by the Max-Planck Institute for Mathematics, the Fields Institute, and the Swiss National Science Foundation through an Ambizione grant. We thank Ricardo Menares and Amalia Pizarro for many useful conversations, Florian Luca and Aleksandar Ivic for helpful suggestions, Bill Allombert and Karim Belabas for a PARI tutorial, and Jean-Louis Nicolas and Cyril Mauvillain for helping to access Robin’s thesis [23].

2 An estimate for $\mathcal{C}_\varepsilon(\Delta)$

Let Δ be a negative integer satisfying $\Delta \equiv 0, 1 \pmod{4}$ and

$$\mathcal{O}_\Delta = \mathbb{Z}[(\Delta + \sqrt{\Delta})/2]$$

the imaginary quadratic order of discriminant Δ . Then $\Delta = Df^2$, where D is the discriminant of the imaginary quadratic field $\mathbb{Q}(\sqrt{\Delta})$ (the “fundamental discriminant”) and $f = [\mathcal{O}_D : \mathcal{O}_\Delta]$ is the conductor. We denote by $\mathcal{C}(\Delta)$ the class number of the order \mathcal{O}_Δ .

Up to \mathbb{C} -isomorphism there exist $\mathcal{C}(\Delta)$ elliptic curves with CM by \mathcal{O}_Δ . The j -invariants of these curves are called *singular moduli* of discriminant Δ . The singular moduli of discriminant Δ form a full Galois orbit over \mathbb{Q} of cardinality $\mathcal{C}(\Delta)$, see [7, Proposition 13.2].

Let \mathcal{F} be the standard fundamental domain in the Poincaré plane, that is, the open hyperbolic triangle with vertices $\zeta_3, \zeta_6, i\infty$, together with the geodesics $[i, \zeta_6]$ and $[\zeta_6, i\infty]$; here

$$\zeta_3 = e^{2\pi i/3} = \frac{-1 + \sqrt{-3}}{2}, \quad \zeta_6 = e^{\pi i/3} = \frac{1 + \sqrt{-3}}{2}.$$

Every singular modulus can be uniquely presented as $j(\tau)$, where $\tau \in \mathcal{F}$.

Now fix $\varepsilon \in (0, 1/3]$ and denote by $\mathcal{C}_\varepsilon(\Delta)$ the number of singular moduli of discriminant Δ that can be presented as $j(\tau)$ where $\tau \in \mathcal{F}$ satisfies

$$\min\{|\tau - \zeta_3|, |\tau - \zeta_6|\} < \varepsilon. \quad (2.1)$$

In this section we bound this quantity.

Define the *modified conductor* \tilde{f} by

$$\tilde{f} = \begin{cases} f, & D \equiv 1 \pmod{4}, \\ 2f, & D \equiv 0 \pmod{4}. \end{cases} \quad (2.2)$$

Then Δ/\tilde{f}^2 is a square-free integer.

Theorem 2.1. *For $\varepsilon \in (0, 1/3]$ we have*

$$\mathcal{C}_\varepsilon(\Delta) \leq F \left(\frac{16}{3} \frac{\sigma_1(\tilde{f})}{\tilde{f}} |\Delta|^{1/2} \varepsilon^2 + \frac{8}{3} |\Delta|^{1/2} \varepsilon + 8 |\Delta/3|^{1/4} \sigma_0(\tilde{f}) \varepsilon + 4 \right), \quad (2.3)$$

where

$$F = F(\Delta) = \max\{2^{\omega(a)} : a \leq |\Delta|^{1/2}\}. \quad (2.4)$$

Corollary 2.2. *In the set-up of Theorem 2.1 assume that $|\Delta| \geq 10^{14}$. Then*

$$\mathcal{C}_\varepsilon(\Delta) \leq F \left(9.83 |\Delta|^{1/2} \varepsilon^2 \log \log(|\Delta|^{1/2}) + 3.605 |\Delta|^{1/2} \varepsilon + 4 \right). \quad (2.5)$$

2.1 Some lemmas

We need some lemmas. For a prime number ℓ and a non-zero integer n we denote by $\text{ord}_\ell(n)$ the ℓ -adic order of n ; that is, $\ell^{\text{ord}_\ell(n)} \parallel n$.

Lemma 2.3. *Let ℓ be a prime number, $e \geq 1$ an integer, and Δ a non-zero integer with $\nu = \text{ord}_\ell \Delta$. Then the set of $b \in \mathbb{Z}$ satisfying $b^2 \equiv \Delta \pmod{\ell^e}$ is a union of at most 2 residue classes modulo $\ell^{e - \lfloor \min\{e, \nu\}/2 \rfloor}$ in all cases except when $\ell = 2$ and $e \geq 3$; in this latter case it is a union of most 4 such classes. Finally, the set of b equals a single residue class modulo $\ell^{e - \lfloor \min\{e, \nu\}/2 \rfloor}$ if $\nu \geq e$.*

Proof. We suppose first that $\nu = 0$, that is, $\ell \nmid \Delta$. In this case we have to count the number of elements in the multiplicative group $(\mathbb{Z}/\ell^e \mathbb{Z})^\times$ whose square is represented by Δ . If $\ell \geq 3$ or $\ell^e \in \{2, 4\}$, then $(\mathbb{Z}/\ell^e \mathbb{Z})^\times$ is a cyclic group. Then there are at most 2 square roots and this implies our claim. If $\ell = 2$ and $e \geq 3$, then $(\mathbb{Z}/2^e \mathbb{Z})^\times \cong \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2^{e-2}\mathbb{Z}$, and there are at most 4 square roots, as desired.

Now assume that $\nu < e$. Then $\text{ord}_\ell(b^2) = \nu$. So ν is even and we can write $b = \ell^{\nu/2} b'$, where $b' \in \mathbb{Z}$ is coprime to ℓ . Now $\Delta = \ell^\nu \Delta'$ with $\Delta' \in \mathbb{Z}$ coprime to ℓ , and $(b')^2 \equiv \Delta' \pmod{\ell^{e-\nu}}$. Above we already determined that, depending on the value of $\ell^{e-\nu}$, the set of possible b' consists of either at most 2 or at most 4 classes modulo $\ell^{e-\nu}$. Hence the set of possible $b = \ell^{\nu/2} b'$ consists of the same number of classes modulo $\ell^{e-\nu/2}$, as desired.

To prove the final claim assume that $\nu \geq e$. In this case $b^2 \equiv \Delta \pmod{\ell^e}$ is equivalent to $b \equiv 0 \pmod{\ell^{\lceil e/2 \rceil}}$. This means that the set of suitable b consists of exactly one class modulo $\ell^{\lceil e/2 \rceil} = \ell^{e - \lfloor e/2 \rfloor}$. \square

We say that $d \in \mathbb{Z}$ is a *quadratic divisor* of $n \in \mathbb{Z}$ if $d^2 \mid n$. We denote by $\text{gcd}_2(m, n)$ the greatest common quadratic divisor of m and n .

Lemma 2.4. *Let a be a positive integer and Δ a non-zero integer. Then the set of $b \in \mathbb{Z}$ satisfying $b^2 \equiv \Delta \pmod{a}$ consists of at most $2^{\omega(a/\text{gcd}_2(a, \Delta)) + 1}$ residue classes modulo $a/\text{gcd}_2(a, \Delta)$.*

Proof. For a prime power ℓ^e we only need the following simple consequence of Lemma 2.3 on the number of residue classes counted there. This number is at most $2^{\omega(\ell^e/\text{gcd}(\ell^e, \Delta))}$ if $\ell \geq 3$ and at most $2^{\omega(\ell^e/\text{gcd}(\ell^e, \Delta)) + 1}$ for $\ell = 2$. The current lemma follows from the Chinese Remainder Theorem. \square

The following lemma is trivial, but we state it here because it is our principal counting tool.

Lemma 2.5. *Let α and β be real numbers, $\alpha < \beta$, and m a positive integer. Then every residue class modulo m has at most $(\beta - \alpha)/m + 1$ elements in the interval $[\alpha, \beta]$.*

Given a negative integer $\Delta \equiv 0, 1 \pmod{4}$, denote by $T = T_\Delta$ the set of triples of integers (a, b, c) such that

$$\begin{aligned} \text{gcd}(a, b, c) &= 1, & \Delta &= b^2 - 4ac, \\ \text{either } -a < b \leq a < c & \text{ or } & 0 \leq b \leq a = c. \end{aligned} \tag{2.6}$$

For $(a, b, c) \in T_\Delta$ we set

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

Lemma 2.6. (i) For every $(a, b, c) \in T_\Delta$ the number $\tau(a, b, c)$ belongs to the standard fundamental domain.

(ii) For $(a, b, c) \in T_\Delta$ we have $0 < a \leq |\Delta/3|^{1/2}$, the equality being possible only if $\Delta = -3$ (and $a = b = c = 1$). We also have $c \geq |\Delta|^{1/2}/2$.

(iii) The map $(a, b, c) \mapsto j(\tau(a, b, c))$ defines a bijection from T_Δ onto the set of \mathbb{Q} -conjugates of $j(\tau)$. In particular, $\mathcal{C}(\Delta) = |T_\Delta|$.

Proof. For item (i) just note that (2.6) implies the inequalities

$$-\frac{1}{2} < \frac{b}{2a} \leq \frac{1}{2}, \quad \frac{b^2 + |\Delta|}{4a^2} \geq 1$$

and that the second one becomes equality only when $a = c$, in which case $b \geq 0$. For item (ii), since $|b| \leq a \leq c$, we have

$$4c^2 \geq |\Delta| = 4ac - b^2 \geq 4a^2 - a^2 = 3a^2.$$

with equality on the right only when $a = |b| = c$. Since $\gcd(a, b, c) = 1$, this is only possible when $a = b = c = 1$ and $\Delta = -3$.

Item (iii) is a combination of several classical results that can be found, for instance, in [7]. See [3, Proposition 2.5] for more details. \square

Lemma 2.7. Let $\varepsilon \in (0, 1/3]$ and let $(a, b, c) \in T_\Delta$ satisfy (2.1). Then

$$\frac{|\Delta|^{1/2}}{\sqrt{3} + 2\varepsilon} < a \leq \frac{|\Delta|^{1/2}}{\sqrt{3}}, \quad (2.7)$$

$$a(1 - 2\varepsilon) < |b| \leq a, \quad (2.8)$$

$$a \leq c < a(1 + \sqrt{3}\varepsilon + \varepsilon^2). \quad (2.9)$$

Note that (2.7) and (2.8) will be used already in Subsection 2.2, while (2.9) will be used only in Section 7.

Proof. For $\tau \in \mathcal{F}$ condition (2.1) implies that

$$\frac{\sqrt{3}}{2} \leq \operatorname{Im} \tau < \frac{\sqrt{3}}{2} + \varepsilon, \quad \frac{1}{2} - \varepsilon < |\operatorname{Re} \tau| \leq \frac{1}{2}.$$

Applying this for $\tau = \tau(a, b, c)$, we obtain (2.7) and (2.8). To prove (2.9), write

$$4ac = |\Delta| + b^2 < a^2(\sqrt{3} + 2\varepsilon)^2 + a^2 = 4a^2(1 + \sqrt{3}\varepsilon + \varepsilon^2),$$

and (2.9) follows. \square

2.2 Proof of Theorem 2.1

Note that, by definition,

$$\mathcal{C}_\varepsilon(\Delta) = \#\{(a, b, c) \in T_\Delta : \tau = \tau(a, b, c) \text{ satisfies (2.1)}\}.$$

Setting

$$I = \left(\frac{|\Delta|^{1/2}}{\sqrt{3} + 2\varepsilon}, \frac{|\Delta|^{1/2}}{\sqrt{3}} \right],$$

for $\tau = \tau(a, b, c)$ with $(a, b, c) \in T_\Delta$ we may re-write (2.7) and (2.8) as

$$a \in I, \quad b \in [-a, -a(1 - 2\varepsilon)) \cup (a(1 - 2\varepsilon), a]. \quad (2.10)$$

Since c is uniquely determined for given a, b and Δ , it suffices to bound the number of pairs (a, b) of integers satisfying $b^2 \equiv \Delta \pmod{a}$ and (2.10).

For every fixed a there are at most $(4\varepsilon \gcd_2(a, \Delta) + 2)2^{\omega(a)+1}$ suitable b , as follows from Lemmas 2.4 and 2.5; indeed, $\omega(a/\gcd(a, \Delta)) \leq \omega(a)$. Hence

$$\mathcal{C}_\varepsilon(\Delta) \leq 8\varepsilon \sum_{a \in I \cap \mathbb{Z}} \gcd_2(a, \Delta) 2^{\omega(a)} + 4 \sum_{a \in I \cap \mathbb{Z}} 2^{\omega(a)} \quad (2.11)$$

$$\leq 8\varepsilon F \sum_{a \in I \cap \mathbb{Z}} \gcd_2(a, \Delta) + 4F \#(I \cap \mathbb{Z}). \quad (2.12)$$

To estimate the sum, note that

$$\sum_{a \in I \cap \mathbb{Z}} \gcd_2(a, \Delta) \leq \sum_{d^2 | \Delta} d \cdot \#(I \cap d^2 \mathbb{Z}). \quad (2.13)$$

Since Δ/\tilde{f}^2 is a square-free integer, we have $d^2 | \Delta$ if and only if $d | \tilde{f}$. Also, since I is of length

$$|\Delta|^{1/2} \left(\frac{1}{\sqrt{3}} - \frac{1}{\sqrt{3} + 2\varepsilon} \right) < \frac{2}{3} |\Delta|^{1/2} \varepsilon,$$

we have, by Lemma 2.5,

$$\#(I \cap d^2 \mathbb{Z}) \leq \begin{cases} \frac{2}{3} \frac{|\Delta|^{1/2}}{d^2} \varepsilon + 1, & d \leq |\Delta/3|^{1/4}, \\ 0, & d > |\Delta/3|^{1/4}. \end{cases}$$

Hence

$$\begin{aligned} \sum_{d^2 | \Delta} d \cdot \#(I \cap d^2 \mathbb{Z}) &\leq \sum_{\substack{d | \tilde{f} \\ d \leq |\Delta/3|^{1/4}}} d \left(\frac{2}{3} \frac{|\Delta|^{1/2}}{d^2} \varepsilon + 1 \right) \\ &\leq \frac{2}{3} |\Delta|^{1/2} \varepsilon \sum_{d | \tilde{f}} d^{-1} + \sum_{\substack{d | \tilde{f} \\ d \leq |\Delta/3|^{1/4}}} d \\ &\leq \frac{2}{3} \frac{\sigma_1(\tilde{f})}{\tilde{f}} |\Delta|^{1/2} \varepsilon + |\Delta/3|^{1/4} \sigma_0(\tilde{f}). \end{aligned} \quad (2.14)$$

Finally, Lemma 2.5 implies that

$$\#(I \cap \mathbb{Z}) \leq \frac{2}{3} |\Delta|^{1/2} \varepsilon + 1. \quad (2.15)$$

Putting the estimates (2.12), (2.13), (2.14) and (2.15) together, we obtain (2.3). \square

2.3 Proof of Corollary 2.2

We need to estimate $\sigma_0(\tilde{f})$ and $\sigma_1(\tilde{f})$ in terms of $|\Delta|$. The following lemma uses a simple estimate for $\sigma_0(n)$ due to Nicolas and Robin [19]. Much sharper estimates can be found in Robin's thesis [23].

Lemma 2.8. *For $|\Delta| \geq 10^{14}$ we have*

$$\sigma_0(\tilde{f}) \leq |\Delta|^{0.192}, \quad (2.16)$$

$$\sigma_1(\tilde{f})/\tilde{f} \leq 1.842 \log \log(|\Delta|^{1/2}). \quad (2.17)$$

Proof. For proving (2.16) may assume that $\tilde{f} \geq 16$, otherwise there is nothing to prove. In [19] it is proved that for $n \geq 3$ we have

$$\frac{\log \sigma_0(n)}{\log 2} \leq 1.538 \frac{\log n}{\log \log n}.$$

The function $x \mapsto (\log x)/(\log \log x)$ is increasing for $x \geq 16$. Since

$$|\Delta| \geq 10^{14}, \quad 16 \leq \tilde{f} \leq |\Delta|^{1/2},$$

this gives

$$\begin{aligned} \log \sigma_0(\tilde{f}) &\leq 1.538 \log 2 \frac{\log(|\Delta|^{1/2})}{\log \log(|\Delta|^{1/2})} \\ &\leq \frac{1.538}{2} \log 2 \frac{\log |\Delta|}{\log \log(10^7)} \\ &< 0.192 \log |\Delta|, \end{aligned}$$

as wanted.

For proving (2.17) we use the estimate $\sigma_1(n) \leq 1.842n \log \log n$ which holds for $n \geq 121$, see [1, Theorem 1.3]. This proves (2.17) for $\tilde{f} \geq 121$. For $\tilde{f} \leq 120$ one can check directly that $\sigma_1(\tilde{f})/\tilde{f} \leq 3$ so that inequality (2.17) is also true in this case. \square

Proof of Corollary 2.2. If $|\Delta| \geq 10^{14}$ then Lemma 2.8 implies that

$$\begin{aligned} 8 \left| \frac{\Delta}{3} \right|^{1/4} \sigma_0(\tilde{f}) &\leq \frac{8}{3^{1/4}} |\Delta|^{0.442} \leq \frac{8}{3^{1/4} \cdot 10^{0.812}} |\Delta|^{1/2} \leq 0.938 |\Delta|^{1/2}, \\ \frac{16}{3} \frac{\sigma_1(\tilde{f})}{\tilde{f}} &\leq \frac{16}{3} \cdot 1.842 \log \log(|\Delta|^{1/2}) \leq 9.83 \log \log(|\Delta|^{1/2}). \end{aligned}$$

Substituting all this to (2.3), we obtain (2.5). \square

3 An upper bound for the height of a singular unit

In this section we obtain a fully explicit version of estimate (1.5). We use the notation $\mathcal{C}(\Delta)$, $\mathcal{C}_\varepsilon(\Delta)$, \mathcal{F} , ζ_3 , ζ_6 introduced in Section 2.

Let α be a complex algebraic number of degree m whose minimal polynomial over \mathbb{Z} is

$$P(x) = a_m x^m + \cdots + a_0 = a_m (x - \alpha_1) \cdots (x - \alpha_m) \in \mathbb{Z}[x].$$

Here $\gcd(a_0, a_1, \dots, a_m) = 1$ and $\alpha_1, \dots, \alpha_m \in \mathbb{C}$ are the conjugates of α over \mathbb{Q} . Then the height of α is defined by

$$h(\alpha) = \frac{1}{m} \left(\log |a_m| + \sum_{k=1}^m \log^+ |\alpha_k| \right),$$

where $\log^+(\cdot) = \log \max\{1, \cdot\}$. If α is an algebraic integer then

$$h(\alpha) = \frac{1}{m} \sum_{k=1}^m \log^+ |\alpha_k|.$$

It is known that $h(\alpha) = h(\alpha^{-1})$ when $\alpha \neq 0$.

Theorem 3.1. *Let α be a singular unit of discriminant Δ , and ε a real number satisfying $0 < \varepsilon \leq 4 \cdot 10^{-3}$. Then*

$$h(\alpha) \leq 3 \frac{\mathcal{C}_\varepsilon(\Delta)}{\mathcal{C}(\Delta)} \log |\Delta| + 3 \log(\varepsilon^{-1}) - 10.66. \quad (3.1)$$

Combining this with Corollary 2.2 and optimizing ε , we obtain the following consequence.

Corollary 3.2. *In the set-up of Theorem 3.1 assume that $|\Delta| \geq 10^{14}$. Then*

$$h(\alpha) \leq \frac{12A}{\mathcal{C}(\Delta)} + 3 \log \frac{A|\Delta|^{1/2}}{\mathcal{C}(\Delta)} - 3.77, \quad (3.2)$$

where $A = F \log |\Delta|$ and F is defined in (2.4).

3.1 Proof of Theorem 3.1

We start from some simple lemmas.

Lemma 3.3. *For $z \in \mathcal{F}$ we have*

$$|j(z)| \geq 42700 (\min\{|z - \zeta_3|, |z - \zeta_6|, 4 \cdot 10^{-3}\})^3.$$

Proof. This is an easy modification of Proposition 2.2 from [3]; just replace therein 10^{-3} by $4 \cdot 10^{-3}$. \square

In the next lemma we use the notation T_Δ and $\tau(a, b, c)$ introduced before Lemma 2.6.

Lemma 3.4. *Assume that $\Delta \neq -3$. Let $\tau = \tau(a, b, c)$, where $(a, b, c) \in T_\Delta$. Let ζ be one of the numbers ζ_3 or ζ_6 . Then*

$$|\tau - \zeta| \geq \frac{\sqrt{3}}{4|\Delta|}.$$

Proof. We have

$$|\tau - \zeta| \geq |\operatorname{Im} \tau - \operatorname{Im} \zeta| = \left| \frac{\sqrt{|\Delta|}}{2a} - \frac{\sqrt{3}}{2} \right| = \frac{||\Delta| - 3a^2|}{2a(\sqrt{|\Delta|} + a\sqrt{3})}.$$

Since $\Delta \neq -3$ we have $\Delta \neq -3a^2$, see item (ii) of Lemma 2.6. Hence

$$|\tau - \zeta| \geq \frac{1}{2a(\sqrt{|\Delta|} + a\sqrt{3})} \geq \frac{\sqrt{3}}{4|\Delta|},$$

the last inequality being again by item (ii) of Lemma 2.6. \square

Now we are ready to prove Theorem 3.1.

Proof of Theorem 3.1. Let $\alpha = \alpha_1, \alpha_2, \dots, \alpha_m \in \mathbb{C}$ be the conjugates of α over \mathbb{Q} . Then $m = \mathcal{C}(\Delta)$ and $\alpha_1, \dots, \alpha_m$ is the full list of singular moduli of discriminant Δ . Write them as $j(\tau_1), \dots, j(\tau_m)$, where $\tau_1, \dots, \tau_m \in \mathcal{F}$.

Since α is a unit, we have

$$h(\alpha) = h(\alpha^{-1}) = \frac{1}{m} \sum_{k=1}^m \log^+ |\alpha_k^{-1}|,$$

Hence

$$\begin{aligned} h(\alpha) &= \frac{1}{\mathcal{C}(\Delta)} \sum_{k=1}^m \log^+ |j(\tau_k)^{-1}| \\ &= \frac{1}{\mathcal{C}(\Delta)} \left(\sum_{\substack{1 \leq k \leq m \\ \min\{|\tau_k - \zeta_3|, |\tau_k - \zeta_6|\} < \varepsilon}} + \sum_{\substack{1 \leq k \leq m \\ \min\{|\tau_k - \zeta_3|, |\tau_k - \zeta_6|\} \geq \varepsilon}} \right) \log^+ |j(\tau_k)^{-1}|. \end{aligned} \quad (3.3)$$

We estimate each of the two sums separately.

Since $\varepsilon \leq 4 \cdot 10^{-3}$, Lemma 3.3 implies that each term in the second sum satisfies

$$\log^+ |j(\tau_k)^{-1}| \leq 3 \log(\varepsilon^{-1}) - \log 42700 \leq 3 \log(\varepsilon^{-1}) - 10.66.$$

Hence, using that $3 \log(\varepsilon^{-1}) > 10.66$, we obtain

$$\sum_{\substack{1 \leq k \leq m \\ \min\{|\tau_k - \zeta_3|, |\tau_k - \zeta_6|\} \geq \varepsilon}} \log^+ |j(\tau_k)^{-1}| \leq \mathcal{C}(\Delta) (3 \log(\varepsilon^{-1}) - 10.66). \quad (3.4)$$

As for the first sum, Lemmas 3.3 and 3.4 imply that each term in this sum satisfies

$$\log^+ |j(\tau_k)^{-1}| \leq \max \left\{ 0, 3 \log \frac{4|\Delta|}{\sqrt{3}} - \log 42700 \right\} \leq 3 \log |\Delta|.$$

Note that we may use here Lemma 3.4 because the only singular modulus of discriminant -3 is 0, which is not a unit.

Since the first sum has $\mathcal{C}_\varepsilon(\Delta)$ terms, this implies the estimate

$$\sum_{\substack{1 \leq k \leq m \\ \min\{|\tau_k - \zeta_3|, |\tau_k - \zeta_6|\} < \varepsilon}} \log^+ |j(\tau_k)^{-1}| \leq 3\mathcal{C}_\varepsilon(\Delta) \log |\Delta|. \quad (3.5)$$

Substituting (3.4) and (3.5) into (3.3), we obtain (3.1). \square

3.2 Proof of Corollary 3.2

To prove the corollary we need a lower bound for the quantity F defined in Theorem 2.1 and an upper bound for the class number $\mathcal{C}(\Delta)$.

Lemma 3.5. *Assume that $|\Delta| \geq 10^{14}$. Then $F \geq |\Delta|^{0.34/\log \log(|\Delta|^{1/2})}$ and $F \geq 18.54 \log \log(|\Delta|^{1/2})$.*

Proof. Define, as usual

$$\vartheta(x) = \sum_{p \leq x} \log p, \quad \pi(x) = \sum_{p \leq x} 1. \quad (3.6)$$

Then

$$\begin{aligned} \vartheta(x) &\leq 1.017x && (x > 0), \\ \pi(x) &\geq \frac{x}{\log x} && (x \geq 17), \end{aligned} \quad (3.7)$$

see [24], Theorem 9 on page 71 and Corollary 1 after Theorem 2 on page 69. Estimate (3.7) implies that

$$\pi(x) \geq 0.99995 \frac{x}{\log x} \quad (x \geq 13). \quad (3.8)$$

Setting here

$$x = \frac{\log(|\Delta|^{1/2})}{1.017}, \quad N = \prod_{p \leq x} p,$$

we obtain $N \leq |\Delta|^{1/2}$ and

$$\omega(N) = \pi(x) \geq \frac{0.99995 \log(|\Delta|^{1/2})}{1.017 \log \log(|\Delta|^{1/2})}.$$

Note that $x \geq (\log(10^7))/1.017 > 15$, so we are allowed to use (3.8). We obtain

$$F \geq 2^{\omega(N)} \geq |\Delta|^{\frac{0.99995 \log 2}{2 \cdot 1.017 \log \log(|\Delta|^{1/2})}} \geq |\Delta|^{0.34/\log \log(|\Delta|^{1/2})},$$

proving the first estimate.

To prove the second estimate, we deduce from the first estimate that

$$\log F - \log \log \log(|\Delta|^{1/2}) \geq 0.68 \frac{u}{\log u} - \log \log u, \quad (3.9)$$

where we set $u = \log(|\Delta|^{1/2})$. The right-hand side of (3.9), viewed as a function in u , is increasing for $u \geq \log(10^7)$. Hence

$$\log F - \log \log \log(|\Delta|^{1/2}) \geq 0.68 \frac{\log(10^7)}{\log \log(10^7)} - \log \log \log(10^7) \geq 2.92,$$

and $F \geq e^{2.92} \log \log(|\Delta|^{1/2}) \geq 18.54 \log \log(|\Delta|^{1/2})$. \square

Lemma 3.6. *For $\Delta \neq -3, -4$ we have*

$$\mathcal{C}(\Delta) \leq \pi^{-1} |\Delta|^{1/2} (2 + \log |\Delta|).$$

Proof. This follows from Theorems 10.1 and 14.3 in [16, Chapter 12]. Note that in [16] the right-hand side has an extra factor $\omega/2$, where ω is the number of roots of unity in the imaginary quadratic order of discriminant Δ . Since we assume that $\Delta \neq -3, -4$, we have $\omega = 2$, so we may omit this factor. \square

Proof of Corollary 3.2. Substituting the estimate for $C_\varepsilon(\Delta)$ from (2.5) into (3.1), we obtain the estimate

$$h(\alpha) \leq 3A \frac{9.83|\Delta|^{1/2}\varepsilon^2 \log \log(|\Delta|^{1/2}) + 3.605|\Delta|^{1/2}\varepsilon + 4}{C(\Delta)} + 3 \log(\varepsilon^{-1}) - 10.66$$

with $A = F \log |\Delta|$. Specifying

$$\varepsilon = 0.27 \frac{C(\Delta)}{A|\Delta|^{1/2}}$$

(this is a nearly optimal value, and it satisfies $\varepsilon \leq 4 \cdot 10^{-3}$ as verified below), we obtain, using Lemmas 3.5 and 3.6,

$$\begin{aligned} h(\alpha) &\leq 3 \cdot 9.83 \cdot (0.27)^2 \frac{\log \log(|\Delta|^{1/2})}{F} \frac{C(\Delta)}{|\Delta|^{1/2} \log |\Delta|} + 3 \cdot 3.605 \cdot 0.27 \\ &\quad + \frac{12A}{C(\Delta)} + 3 \log \frac{A|\Delta|^{1/2}}{C(\Delta)} - 3 \log 0.27 - 10.66 \\ &\leq \frac{12A}{C(\Delta)} + 3 \log \frac{A|\Delta|^{1/2}}{C(\Delta)} + \frac{3 \cdot 9.83 \cdot (0.27)^2 \cdot 0.34}{18.54} + 3 \cdot 3.605 \cdot 0.27 \\ &\quad - 3 \log 0.27 - 10.66 \\ &\leq \frac{12A}{C(\Delta)} + 3 \log \frac{A|\Delta|^{1/2}}{C(\Delta)} - 3.77, \end{aligned}$$

as wanted.

We only have to verify that $\varepsilon \leq 4 \cdot 10^{-3}$. We have $F \geq 256$ when $|\Delta| \geq 10^{14}$. Using Lemma 3.6, we obtain

$$\varepsilon = \frac{0.27C(\Delta)}{|\Delta|^{1/2} \log |\Delta|} \frac{1}{F} \leq 0.27\pi^{-1} \frac{2 + \log(10^{14})}{\log(10^{14})} \cdot \frac{1}{256} < 4 \cdot 10^{-4}.$$

The proof is complete. \square

4 Lower bounds for the height of a singular modulus

Now we establish explicit lower bounds of the form (1.6) and (1.7).

4.1 The “easy” bound

We start by proving a bound of the form (1.7).

Proposition 4.1. *Let α be a singular modulus of discriminant Δ . Assume that $|\Delta| \geq 16$. Then*

$$h(\alpha) \geq \frac{\pi|\Delta|^{1/2} - 0.01}{C(\Delta)}. \quad (4.1)$$

We need a simple lemma.

Lemma 4.2. *For $z \in \mathcal{F}$ with imaginary part y we have*

$$||j(z)| - e^{2\pi y}| \leq 2079.$$

If $y \geq 2$ then we also have $|j(z)| \geq 0.992e^{2\pi y}$.

Proof. The first statement is Lemma 1 of [4], and the second one is an immediate consequence. \square

Proof of Proposition 4.1. One of the conjugates of α over \mathbb{Q} is equal to $j((b + \sqrt{\Delta})/2)$, with $b = 1$ for Δ odd, and $b = 0$ for Δ even; it corresponds to the element $(1, b, (-\Delta + b^2)/4)$ of the set T_Δ . Hence

$$h(\alpha) \geq \frac{\log |j((b + \sqrt{\Delta})/2)|}{C(\Delta)}.$$

Using Lemma 4.2, we obtain

$$\log |j((b + \sqrt{\Delta})/2)| \geq \pi|\Delta|^{1/2} + \log 0.992 \geq \pi|\Delta|^{1/2} - 0.01.$$

Whence the result. \square

4.2 The “hard” bound

We are left with bound (1.6). We are going to prove the following.

Proposition 4.3. *Let α be a singular modulus of discriminant Δ . Then*

$$h(\alpha) \geq \frac{3}{\sqrt{5}} \log |\Delta| - 9.79. \quad (4.2)$$

The proof of Proposition 4.3 relies on the fact that it is possible to evaluate the Faltings height of an elliptic curve with complex multiplication precisely, due to the work of Colmez [6] and Nakajima-Taguchi [18]; for an exact statement see [13, Lemma 4.1].

Let E be an elliptic curve with CM by an order of discriminant Δ . We let $h_F(E)$ denote the stable Faltings height of E (using Deligne’s normalization [8]). The above-mentioned explicit formula for $h_F(E)$ is used in [15] to obtain the lower bound

$$h_F(E) \geq \frac{1}{4\sqrt{5}} \log |\Delta| - 5.93, \quad (4.3)$$

see Lemma 14(ii) therein. Unfortunately, this bound is numerically too weak for our purposes.

Proposition 4.3 will be deduced from the following numerical refinement of (4.3).

Proposition 4.4. *Let E be an elliptic curve with CM by an order of discriminant Δ . Then*

$$h_F(E) \geq \frac{1}{4\sqrt{5}} \log |\Delta| - \gamma - \frac{\log(2\pi)}{2} - \left(\frac{1}{2\sqrt{5}} - \frac{1}{6} \right) \log 2, \quad (4.4)$$

where $\gamma = 0.57721 \dots$ is the Euler constant.

Proof of Proposition 4.3 (assuming Proposition 4.4). Let E be an elliptic curve with $j(E) = \alpha$. We only need to relate $h_F(E)$ to $h(j(E))$. For this purpose we use Lemma 7.9 of Gaudron and Rémond [11]². In our notation they show that

$$h_F(E) \leq h(j(E))/12 - 0.72 \quad (4.5)$$

A quick calculation yields our claim. \square

To prove Proposition 4.4 we need a technical lemma. Set

$$\lambda = \frac{1}{2} - \frac{1}{2\sqrt{5}}, \quad (4.6)$$

and define the additive arithmetical functions $\beta(n)$ and $\delta(n)$ by

$$\beta(p^k) = \frac{\log p}{p+1} \frac{1-p^{-k}}{1-p^{-1}}, \quad \beta(n) = \sum_{p^k \parallel n} \beta(p^k), \quad \delta(n) = \lambda \log n - \beta(n). \quad (4.7)$$

Lemma 4.5. *For every positive integer n we have*

$$\delta(n) \geq \delta(2) = \left(\frac{1}{6} - \frac{1}{2\sqrt{5}} \right) \log 2.$$

Proof. Since $1/3 > \lambda > 1/4$, we have $\delta(2) < 0$ and $\delta(p) > 0$ for all primes $p \geq 3$. Also, for $k \geq 1$ and any prime p we have

$$\delta(p^{k+1}) - \delta(p^k) = \left(\lambda - \frac{1}{p^k(p+1)} \right) \log p > 0.$$

Since $\delta(4) > 0$, this proves that $\delta(p^k) > 0$ for every prime power $p^k \neq 2$, whence the result. \square

Proposition 4.4 is an immediate consequence of Lemma 4.5 and the following statement.

Proposition 4.6. *In the set-up of Proposition 4.4 we have*

$$h_F(E) \geq \frac{1}{4\sqrt{5}} \log |\Delta| + \lambda \log f - \beta(f) - \gamma - \frac{\log(2\pi)}{2}. \quad (4.8)$$

Since

$$\lambda \log f - \beta(f) \geq - \left(\frac{1}{2\sqrt{5}} - \frac{1}{6} \right) \log 2$$

by Lemma 4.5, this implies Proposition 4.4.

²The reader should be warned that our $h_F(E)$ is denoted $h(E)$ in [11].

Proof of Proposition 4.6. Write $\Delta = Df^2$ with D the fundamental discriminant and f the conductor. Define

$$e_f(p) = \frac{1 - \chi(p)}{p - \chi(p)} \frac{1 - p^{-\text{ord}_p(f)}}{1 - p^{-1}}, \quad c(f) = \frac{1}{2} \left(\sum_{p|f} e_f(p) \log p \right),$$

where $\chi(\cdot) = (D/\cdot)$ is Kronecker's symbol.

In the proof of Lemma 14 of [15]³, the stable Faltings height of E is estimated as

$$\begin{aligned} h_F(E) &\geq \frac{1}{4\sqrt{5}} \log |D| + \frac{1}{2} \log f - c(f) - \gamma - \frac{\log(2\pi)}{2}, \\ &= \frac{1}{4\sqrt{5}} \log |\Delta| + \lambda \log f - c(f) - \gamma - \frac{\log(2\pi)}{2}. \end{aligned}$$

Thus, to establish (4.8), we only have to prove that $c(f) \leq \beta(f)$. We have

$$\frac{1 - \chi(p)}{p - \chi(p)} = \begin{cases} 0, & \chi(p) = 1, \\ 1/p, & \chi(p) = 0, \\ 2/(p+1), & \chi(p) = -1. \end{cases}$$

Hence

$$\frac{1 - \chi(p)}{p - \chi(p)} \leq \frac{2}{p+1}$$

in any case. This implies that $c(f) \leq \beta(f)$. The proposition is proved. \square

5 The estimate $|\Delta| < 10^{15}$

In this section we obtain the first explicit upper bound for the discriminant of a singular unit.

Theorem 5.1. *Let Δ be the discriminant of a singular unit. Then $|\Delta| < 10^{15}$.*

Throughout this section Δ is the discriminant of a singular unit α , and we assume that $X = |\Delta| \geq 10^{14}$, as otherwise there is nothing to prove. Our principal tools will be the upper estimate (3.2) and the lower estimates (4.1), (4.2). We reproduce them here for convenience:

$$h(\alpha) \leq \frac{12A}{\mathcal{C}(\Delta)} + 3 \log \frac{AX^{1/2}}{\mathcal{C}(\Delta)} - 3.77, \quad (5.1)$$

$$h(\alpha) \geq \frac{\pi X^{1/2} - 0.01}{\mathcal{C}(\Delta)}, \quad (5.2)$$

$$h(\alpha) \geq \frac{3}{\sqrt{5}} \log X - 9.79. \quad (5.3)$$

Note that our assumption $X \geq 10^{14}$ implies that the right-hand side of (5.3) is positive.

³Note that our D is written Δ in [15].

5.1 The main inequality

Recall that $A = F \log X$. Minding 0.01 in (5.2) we deduce from (5.1), (5.2) and (5.3) the inequality

$$\frac{12A}{\mathcal{C}(\Delta)} + 3 \log \frac{AX^{1/2}}{\mathcal{C}(\Delta)} - 3.76 \geq \max \left\{ \frac{\pi X^{1/2}}{\mathcal{C}(\Delta)}, \frac{3}{\sqrt{5}} \log X - 9.78 \right\}.$$

Denoting

$$Y = \max \left\{ \frac{\pi X^{1/2}}{\mathcal{C}(\Delta)}, \frac{3}{\sqrt{5}} \log X - 9.78 \right\}, \quad (5.4)$$

we re-write this as

$$\frac{12A/\mathcal{C}(\Delta)}{Y} + \frac{3 \log A - 3.76}{Y} + \frac{\log(X^{1/2}/\mathcal{C}(\Delta))}{Y} \geq 1. \quad (5.5)$$

Note that $3 \log A - 3.76 > 0$, because $A \geq \log X \geq \log(10^{14}) > 20$. Hence we may replace Y by $\frac{3}{\sqrt{5}} \log X - 9.78$ in the middle term of the left-hand side in (5.5). Similarly, in the first term we may replace Y by $\pi X^{1/2}/\mathcal{C}(\Delta)$, and in the third term we may replace $X^{1/2}/\mathcal{C}(\Delta)$ by $\pi^{-1}Y$. We obtain

$$12\pi^{-1}AX^{-1/2} + \frac{3 \log A - 3.76}{\frac{3}{\sqrt{5}} \log X - 9.78} + 3 \frac{\log(\pi^{-1}Y)}{Y} \geq 1. \quad (5.6)$$

To show that (5.6) is not possible for $X \geq 10^{15}$, we will bound from above each of the three terms in its left-hand side. To begin with, we bound A .

5.2 Bounding F and A

Recall that $F = \max\{2^{\omega(a)} : a \leq X^{1/2}\}$ and $A = F \log X$.

Let $N_1 = 2 \cdot 3 \cdot 5 \cdots 1129$ be the product of the first 189 prime numbers. Define the real number c_1 from

$$\omega(N_1) = \frac{\log N_1}{\log \log N_1 - c_1}.$$

A calculation shows that $c_1 < 1.1713142$. Robin [22, Théorème 13] proved that

$$\omega(n) \leq \frac{\log n}{\log \log n - c_1}$$

for $n \geq 26$. This implies that

$$\frac{\log F}{\log 2} \leq \frac{1}{2} \frac{\log X}{\log \log X - c_1 - \log 2}, \quad (5.7)$$

$$\log A \leq \frac{\log 2}{2} \frac{\log X}{\log \log X - c_1 - \log 2} + \log \log X. \quad (5.8)$$

Indeed, the function

$$g(x) = \frac{\log x}{\log \log x - c_1}$$

is strictly increasing for $x \geq 6500$ and $g(6500) > 8$. If $a \leq X^{1/2}$ then either $a \leq 6500$ in which case $\omega(a) \leq 5 < g(6500) < g(X^{1/2})$ (recall that

$X \geq 10^{14}$), or $6500 < a \leq X^{1/2}$, in which case $\omega(a) \leq g(a) \leq g(X^{1/2})$. Thus, in any case we have

$$\omega(a) \leq g(X^{1/2}) = \frac{1}{2} \frac{\log X}{\log \log X - c_1 - \log 2},$$

which proves (5.7). The estimate (5.8) is an immediate consequence of (5.7).

5.3 Bounding the first term in (5.6)

Using (5.8), we estimate

$$\frac{\log(AX^{-1/2})}{\log X} \leq u_0(X),$$

where

$$u_0(x) = \frac{\log 2}{2} \frac{1}{\log \log x - c_1 - \log 2} + \frac{\log \log x}{\log x} - \frac{1}{2}.$$

The function $u_0(x)$ is decreasing for $x \geq 10^{10}$. Hence for $X \geq 10^{15}$ we have

$$\frac{\log(AX^{-1/2})}{\log X} \leq u_0(10^{15}) < -0.1908.$$

This proves the estimate

$$AX^{-1/2} < 10^{15 \cdot (-0.1908)} < 0.0014$$

for $X \geq 10^{15}$.

5.4 Bounding the second term in (5.6)

Using (5.8), we estimate

$$\frac{3 \log A - 3.76}{\frac{3}{\sqrt{5}} \log X - 9.78} \leq u_1(X)u_2(X),$$

where

$$u_1(x) = \frac{3 \log 2}{2} \frac{1}{\log \log x - c_1 - \log 2} + \frac{3 \log \log x - 3.76}{\log x},$$

$$u_2(x) = \left(\frac{3}{\sqrt{5}} - \frac{9.78}{\log x} \right)^{-1}.$$

Both functions $u_1(x)$ and $u_2(x)$ are decreasing for $x \geq 10^{10}$. Hence, for $X \geq 10^{15}$ we have

$$\frac{3 \log A - 3.76}{\frac{3}{\sqrt{5}} \log X - 9.78} \leq u_1(10^{15})u_2(10^{15}) < 0.7734.$$

5.5 Bounding the third term in (5.6)

The function $x \mapsto (\log x)/x$ is decreasing for $x \geq e$. Since for $X \geq 10^{14}$ we have

$$\pi^{-1}Y \geq \pi^{-1} \left(\frac{3}{\sqrt{5}} \log X - 9.78 \right) \geq e,$$

we have, for $X \geq 10^{14}$, the estimate

$$\frac{\log(\pi^{-1}Y)}{Y} \leq u_3(X),$$

where

$$u_3(x) = \frac{\log \left(\pi^{-1} \left(\frac{3}{\sqrt{5}} \log x - 9.78 \right) \right)}{\frac{3}{\sqrt{5}} \log x - 9.78}.$$

Moreover, the function $u_3(x)$ is decreasing for $x \geq 10^{14}$, which implies that

$$\frac{\log(\pi^{-1}Y)}{Y} \leq u_3(X) \leq u_3(10^{15}) < 0.0672$$

for $X \geq 10^{15}$.

5.6 Summing up

Now, when $X \geq 10^{15}$, we can combine the above estimates and bound the left-hand side of (5.6) by

$$12\pi^{-1} \cdot 0.0014 + 0.7734 + 3 \cdot 0.0672 < 0.981.$$

Hence, for $X \geq 10^{15}$ we cannot have (5.6). This contradiction completes the proof of Theorem 5.1.

6 Handling the mid-range $10^{10} \leq |\Delta| < 10^{15}$

In this section we rule out the existence of singular units with discriminants in the mid-range $[10^{10}, 10^{15})$, improving thereby the bound from the previous section.

Theorem 6.1. *Let Δ be the discriminant of a singular unit. Then $|\Delta| \notin [10^{10}, 10^{15})$.*

In Section 2 we estimate trivially $2^{\omega(a)} \leq F$. One might expect to do better by estimating the average order rather than the maximal order of the arithmetical function $2^{\omega(n)}$. This is accomplished in Subsection 6.1 and allows us to obtain, in Subsection 6.2, a new bound for $\mathcal{C}_\varepsilon(\Delta)$ in the range $10^{10} \leq |\Delta| < 10^{15}$. Using this, Theorem 6.1 is proved in Subsection 6.3 by an argument similar to the proof of Theorem 5.1.

Throughout this section n denotes a positive integer.

6.1 Average order of the function $2^{\omega(n)}$ on subintervals of $[0, 2 \cdot 10^7]$

For a positive real number x set

$$S(x) = \sum_{n \leq x} 2^{\omega(n)}.$$

We define $S(0) = 0$. As Theorem II.6.1 from [26] suggests, the function $S(x)$ can be well approximated by the function

$$g(x) = \lambda_0 x \log x + \lambda_1 x,$$

where

$$\begin{aligned} \lambda_0 &= \zeta(2)^{-1} = 0.607927101854026 \dots, \\ \lambda_1 &= -2 \frac{\zeta'(2)}{\zeta(2)^2} + \frac{2\gamma - 1}{\zeta(2)} = 0.786872460166245 \dots, \end{aligned}$$

and γ is the Euler constant. The function g is increasing on $[1, \infty)$.

As already mentioned in the introduction, the error term $|S(x) - g(x)|$ can be estimated by the Selberg-Delange method [26, Chapter II.5], but on our limited range it is more advantageous to obtain an optimal error term by a computer-assisted calculation.

Proposition 6.2. *For $2 \leq x \leq 2 \cdot 10^7$ we have*

$$g(x) - 1.010x^{1/2} \leq S(x) \leq g(x) + 0.712x^{1/2}, \quad (6.1)$$

and for $4 \cdot 10^4 \leq x \leq 2 \cdot 10^7$ we have

$$g(x) - 2.267 \frac{x^{1/2}}{\log x} \leq S(x) \leq g(x) + 2.598 \frac{x^{1/2}}{\log x}. \quad (6.2)$$

Proof. Set

$$\begin{aligned} c_1 &= \max_{2 \leq n \leq 2 \cdot 10^7} \frac{S(n) - g(n)}{\sqrt{n}}, \\ c_2 &= \max_{2 \leq n \leq 2 \cdot 10^7} \frac{g(n+1) - S(n)}{\sqrt{n}}, \\ c_3 &= \max_{4 \cdot 10^4 \leq n \leq 2 \cdot 10^7} \frac{(S(n) - g(n)) \log n}{\sqrt{n}}, \\ c_4 &= \max_{4 \cdot 10^4 \leq n \leq 2 \cdot 10^7} \frac{(g(n+1) - S(n)) \log n}{\sqrt{n}}. \end{aligned}$$

Then for $2 \leq n \leq 2 \cdot 10^7$ we have

$$S(n) \leq g(n) + c_1 \sqrt{n}, \quad S(n) \geq g(n+1) - c_2 \sqrt{n}.$$

Hence for $2 \leq x \leq 2 \cdot 10^7$ we have

$$\begin{aligned} S(x) &= S(\lfloor x \rfloor) \leq g(\lfloor x \rfloor) + c_1 \sqrt{\lfloor x \rfloor} \leq g(x) + c_1 \sqrt{x}, \\ S(x) &= S(\lfloor x \rfloor) \geq g(\lfloor x \rfloor + 1) - c_2 \sqrt{\lfloor x \rfloor} \geq g(x) - c_2 \sqrt{x}. \end{aligned}$$

In a similar way we show that for $4 \cdot 10^4 \leq x \leq 2 \cdot 10^7$

$$g(x) - c_4 \frac{\sqrt{x}}{\log x} \leq S(x) \leq g(x) + c_3 \frac{\sqrt{x}}{\log x}$$

having used that $x \mapsto \sqrt{x}/\log x$ is increasing on (e^2, ∞) . A computer-assisted calculation shows that

$$c_1 \leq 0.712, \quad c_2 \leq 1.010, \quad c_3 \leq 2.598, \quad c_4 \leq 2.267.$$

We verify this by means of a SAGE [29] script⁴ using the interval arithmetic MPFI package [21]. \square

Corollary 6.3. *Let A and B be positive real numbers satisfying*

$$0 < A < B \leq 2 \cdot 10^7, \quad B \geq 1.$$

Then

$$\sum_{A < n \leq B} 2^{\omega(n)} \leq \lambda_0(B - A)(1 + \log B) + \lambda_1(B - A) + 1.722B^{1/2} \quad (6.3)$$

If, in addition to this, $A \geq 4 \cdot 10^4$, then

$$\sum_{A < n \leq B} 2^{\omega(n)} \leq \lambda_0(B - A)(1 + \log B) + \lambda_1(B - A) + 4.865 \frac{B^{1/2}}{\log B}. \quad (6.4)$$

Proof. In general we have

$$S(B) - S(A) = \sum_{A < n \leq B} 2^{\omega(n)} = \sum_{n=[A]+1}^{\lfloor B \rfloor} 2^{\omega(n)} = S(\lfloor B \rfloor) - S(\lfloor A \rfloor).$$

Note that since $B > A > 0$ we remark that

$$B \log B - A \log A \leq (B - A)(1 + \log B). \quad (6.5)$$

When $A \geq 2$ estimate (6.3) follows immediately from (6.1) and (6.5).

Let us assume that $A < 2$, hence $\lfloor A \rfloor$ is 0 or 1 and $S(\lfloor A \rfloor) = 0$ or 1, respectively. For $B \geq 2$ we find

$$\sum_{A < n \leq B} 2^{\omega(n)} \leq \lambda_0 B \log B + \lambda_1 B + 0.712B^{1/2} - S(\lfloor A \rfloor),$$

and one easily verifies that

$$-S(\lfloor A \rfloor) + \lambda_0 A(1 + \log B) + \lambda_1 A - \lambda_0 B \leq 1.010B^{1/2}$$

by considering the cases $A \in (0, 1)$ and $A \in [1, 2)$ separately. This implies (6.3). And if $B < 2$, then $\sum_{A < n \leq B} 2^{\omega(n)} \leq 1 \leq 1.722B^{1/2}$. As $B \geq 1$, again we obtain (6.3).

Finally, for $4 \cdot 10^4 \leq A < B \leq 2 \cdot 10^7$ estimate (6.4) follows from (6.2) and (6.5). \square

⁴A link to the script `prop6_2.sage` is on the second named author's homepage. The running time is roughly 30 minutes on a regular desktop (Intel Xeon CPU E5-1620 v3, 3.50GHz, 32GB RAM).

6.2 Bounding $\mathcal{C}_\varepsilon(\Delta)$ for $10^{10} \leq |\Delta| < 10^{15}$

Now we can obtain a cardinal refinement of Theorem 2.1 for discriminants in the range $10^{10} \leq |\Delta| < 10^{15}$. We need a technical lemma using our notation (1.2).

Lemma 6.4. *Let n be an integer with $1 \leq n \leq 3.2 \cdot 10^7$.*

- (i) *We have $\sigma_1(n)/n \leq \sigma_1(21621600)/21621600 = 3472/715$.*
- (ii) *We have $\sigma_0(n) \leq 8.5n^{1/4}$.*

Proof. This can be proved in at least two ways. In our relatively small range, we can perform a quick computer-assisted calculation⁵. Alternatively, one can use the *On-Line Encyclopedia of Integer Sequences* [27]. Inspecting T. D. Noe and D. Kilminster's table in entry A004394, one can deduce (i). Similarly, (ii) follows from inspection of T. D. Noe's table in A002182. \square

Proposition 6.5. *In the set-up of Theorem 2.1 assume that*

$$10^{10} \leq |\Delta| < 10^{15}. \quad (6.6)$$

Then

$$\begin{aligned} \mathcal{C}_\varepsilon(\Delta) &\leq (8\varepsilon^2 + 0.811\varepsilon)|\Delta|^{1/2} \log |\Delta| + (28\varepsilon^2 + 2.829\varepsilon)|\Delta|^{1/2} \\ &\quad + 89\varepsilon|\Delta|^{3/8} + 31.06 \frac{|\Delta|^{1/4}}{\log |\Delta|}. \end{aligned}$$

Proof. As in the proof of Theorem 2.1, we set

$$I = \left(\frac{|\Delta|^{1/2}}{\sqrt{3} + 2\varepsilon}, \frac{|\Delta|^{1/2}}{\sqrt{3}} \right].$$

We again want to count pairs of integers (a, b) such that

$$a \in I, \quad b \in [-a, -a(1 - 2\varepsilon)) \cup (a(1 - 2\varepsilon), a], \quad b^2 \equiv \Delta \pmod{a}.$$

Lemmas 2.4 and 2.5 imply that, for every fixed a , the number of suitable b does not exceed $(4\varepsilon \gcd_2(a, \Delta) + 2)2^{\omega(a/\gcd(a, \Delta))+1}$. Hence

$$\mathcal{C}_\varepsilon(\Delta) \leq 8\varepsilon \sum_{a \in I \cap \mathbb{Z}} \gcd_2(a, \Delta) 2^{\omega(a/\gcd_2(a, \Delta)^2)} + 4 \sum_{a \in I \cap \mathbb{Z}} 2^{\omega(a)} \quad (6.7)$$

where we used that $\gcd_2(a, \Delta)^2$ divides $\gcd(a, \Delta)$. We estimate each of the terms separately.

⁵A link to the PARI script `lemma6_4.gp` is on the second-named author's homepage. The running time is under 2 minutes on a regular desktop (Intel Xeon CPU E5-1620 v3, 3.50GHz, 32GB RAM).

Estimating the first term in (6.7)

Recall that Δ/\tilde{f}^2 is a square-free integer. We have

$$\sum_{a \in I \cap \mathbb{Z}} \gcd_2(a, \Delta) 2^{\omega(a/\gcd_2(a, \Delta)^2)} \leq \sum_{d|\tilde{f}} d \sum_{a \in I \cap d^2 \mathbb{Z}} 2^{\omega(a/d^2)}. \quad (6.8)$$

To estimate the inner sum, write $a \in I \cap d^2 \mathbb{Z}$ as $a = d^2 a'$ with

$$a' \in d^{-2} I = \left(d^{-2} \frac{|\Delta|^{1/2}}{\sqrt{3} + 2\varepsilon}, d^{-2} \frac{|\Delta|^{1/2}}{\sqrt{3}} \right].$$

We estimate $\sum_{a' \in d^{-2} I \cap \mathbb{Z}} 2^{\omega(a')}$ using (6.3) with

$$A = d^{-2} \frac{|\Delta|^{1/2}}{\sqrt{3} + 2\varepsilon}, \quad B = d^{-2} \frac{|\Delta|^{1/2}}{\sqrt{3}}.$$

Since $|\Delta| \leq 10^{15}$, we have $B \leq 2 \cdot 10^7$. From

$$B - A \leq \frac{2}{3} d^{-2} |\Delta|^{1/2} \varepsilon, \quad \log B \leq \frac{1}{2} \log |\Delta| - \frac{1}{2} \log 3,$$

we obtain

$$\sum_{a' \in d^{-2} I \cap \mathbb{Z}} 2^{\omega(a')} \leq \frac{|\Delta|^{1/2}}{3d^2} (\lambda_0 \log |\Delta| + 2\lambda_0 + 2\lambda_1 - \lambda_0 \log 3) \varepsilon + \frac{1.722 |\Delta|^{1/4}}{\sqrt[4]{3} d}$$

as long as $B \geq 1$. If $B < 1$, then the sum on the left is 0 and the inequality remains valid as the right side is clearly positive. Hence the left-hand side of (6.8) is bounded by

$$\frac{|\Delta|^{1/2}}{3} (\lambda_0 \log |\Delta| + 2\lambda_0 + 2\lambda_1 - \lambda_0 \log 3) \varepsilon \frac{\sigma_1(\tilde{f})}{\tilde{f}} + \frac{1.722 |\Delta|^{1/4}}{\sqrt[4]{3}} \sigma_0(\tilde{f}),$$

where we use notation (1.2) and the identity $\sigma_1(\tilde{f})/\tilde{f} = \sum_{d|\tilde{f}} d^{-1}$.

Recall that $\tilde{f} \leq |\Delta|^{1/2} \leq 3.2 \cdot 10^7$. Hence Lemma 6.4 implies that

$$\sigma_1(\tilde{f})/\tilde{f} \leq 3472/715, \quad \sigma_0(\tilde{f}) \leq 8.5 \tilde{f}^{1/4} \leq 8.5 |\Delta|^{1/8}.$$

Taking into account the factor 8ε , the first term in (6.7) is thus at most

$$|\Delta|^{1/2} (8 \log |\Delta| + 28) \varepsilon^2 + 89 |\Delta|^{3/8} \varepsilon.$$

Estimating the second term in (6.7)

Set

$$A = \frac{|\Delta|^{1/2}}{\sqrt{3} + 2\varepsilon}, \quad B = \frac{|\Delta|^{1/2}}{\sqrt{3}}.$$

From (6.6) and $0 < \varepsilon \leq 1/3$ we deduce $A \geq 4 \cdot 10^4$ and $B \leq 2 \cdot 10^7$. This allows us to apply (6.4), and we obtain

$$\begin{aligned} \sum_{a \in I \cap \mathbb{Z}} 2^{\omega(a)} &\leq \frac{|\Delta|^{1/2}}{3} (\lambda_0 \log |\Delta| + 2\lambda_0 + 2\lambda_1 - \lambda_0 \log 3) \varepsilon \\ &\quad + \frac{4.865 |\Delta|^{1/4}}{\sqrt[4]{3} \log(|\Delta/3|^{1/2})} \\ &\leq \frac{|\Delta|^{1/2}}{3} (\lambda_0 \log |\Delta| + 2\lambda_0 + 2\lambda_1 - \lambda_0 \log 3) \varepsilon \\ &\quad + \frac{2 \cdot 4.865 |\Delta|^{1/4}}{\sqrt[4]{3} \log |\Delta|} \frac{\log 10^{10}}{\log(10^{10}/3)}. \end{aligned}$$

where for the last estimate we used the assumption $|\Delta| \geq 10^{10}$. Minding the factor 4 we find that the second term in (6.7) is at most

$$|\Delta|^{1/2} (0.811 \log |\Delta| + 2.829) \varepsilon + 31.06 \frac{|\Delta|^{1/4}}{\log |\Delta|}.$$

This concludes our proof of Proposition 6.5. \square

6.3 Proof of Theorem 6.1

Suppose that α is a singular unit of discriminant Δ and set $X = |\Delta|$. Assuming that $10^{10} \leq X < 10^{15}$, we arrive at a contradiction. As in Section 5 we use the estimates (5.2) and (5.3) which follow from Propositions 4.1 and 4.3, respectively.

We may no longer use Corollary 3.2, because its hypothesis $X \geq 10^{14}$ is not valid in our current range. Instead, we will apply Theorem 3.1 directly now. For this, let $\varepsilon \in (0, 4 \cdot 10^{-3}]$.

We define Y as in (5.4) and recall that $Y \leq 0.01 + h(\alpha)$. We find

$$Y \leq 3 \frac{\mathcal{C}_\varepsilon(\Delta)}{\mathcal{C}(\Delta)} \log X + 3 \log(\varepsilon^{-1}) - 10.65. \quad (6.9)$$

Using Proposition 6.5 we find

$$\begin{aligned} 1 &\leq 3(8\varepsilon^2 + 0.811\varepsilon) \frac{X^{1/2} (\log X)^2}{Y \mathcal{C}(\Delta)} + 3(28\varepsilon^2 + 2.829\varepsilon) \frac{X^{1/2} \log X}{Y \mathcal{C}(\Delta)} \\ &\quad + 3 \frac{89\varepsilon X^{3/8} \log X}{Y \mathcal{C}(\Delta)} + 3 \cdot 31.06 \frac{X^{1/4}}{Y \mathcal{C}(\Delta)} + \frac{3 \log(\varepsilon^{-1}) - 10.65}{Y}. \end{aligned}$$

For all but the final term on the right we use $Y \mathcal{C}(\Delta) \geq \pi X^{1/2}$ and for the remaining term we use $Y \geq \frac{3}{\sqrt{5}} \log X - 9.78 > 0$, as $X \geq 10^5$, to get

$$\begin{aligned} 1 &\leq \frac{3}{\pi} (8\varepsilon^2 + 0.811\varepsilon) (\log X)^2 + \frac{3}{\pi} (28\varepsilon^2 + 2.829\varepsilon) \log X \\ &\quad + \frac{267}{\pi} \varepsilon \frac{\log X}{X^{1/8}} + \frac{93.18}{\pi X^{1/4}} + \frac{3 \log(\varepsilon^{-1}) - 10.65}{\frac{3}{\sqrt{5}} \log X - 9.78}. \end{aligned} \quad (6.10)$$

Our choice is $\varepsilon = 10^{-4}$. The first two terms in the right-hand side of (6.10) are monotonously increasing, and the remaining three terms are

decreasing for $X \in [10^{10}, 10^{15}]$; note that $x \mapsto (\log x)/x^{1/8}$ is decreasing for $x \geq 3000 > e^8$. Using $X < 10^{15}$ for the first two terms and $X \geq 2 \cdot 10^{10}$ for the remaining three terms, we see that the right-hand side of (6.10) is strictly smaller than 0.962 if $X \in [2 \cdot 10^{10}, 10^{15}]$. Similarly, we infer that it is strictly smaller than 0.960 if $X \in [10^{10}, 2 \cdot 10^{10}]$. This completes the proof of Theorem 6.1. \square

7 Handling the low-range $3 \cdot 10^5 \leq |\Delta| < 10^{10}$

We now deal with the low-range $|\Delta| \in [3 \cdot 10^5, 10^{10}]$. For this range the upper bound on $\mathcal{C}_\varepsilon(\Delta)$ arises from a computer-assisted search algorithm.

We prove the following.

Theorem 7.1. *Let Δ be the discriminant of a singular unit. Then $|\Delta| \notin [3 \cdot 10^5, 10^{10}]$.*

The proof relies on the following lemma.

Lemma 7.2. *Let Δ be the discriminant of a singular modulus.*

(i) *If $10^7 \leq |\Delta| < 10^{10}$ and $\varepsilon = 10^{-3}$, then $\mathcal{C}_\varepsilon(\Delta) \leq 16$.*

(ii) *If $3 \cdot 10^5 \leq |\Delta| < 10^7$ and $\varepsilon = 4 \cdot 10^{-3}$, then $\mathcal{C}_\varepsilon(\Delta) \leq 6$.*

Proof. Let X_{\min} and X_{\max} be positive integers satisfying $X_{\min} < X_{\max}$, and let $\varepsilon \in (0, 1/3]$. We want to bound $\mathcal{C}_\varepsilon(\Delta)$ for all Δ in the interval $[-X_{\max}, -X_{\min}]$.

Recall that $\mathcal{C}_\varepsilon(\Delta)$ counts the triples (a, b, c) satisfying (2.6) such that

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

satisfies

$$\min\{|\tau - \zeta_3|, |\tau - \zeta_6|\} < \varepsilon. \quad (7.1)$$

Lemmas 2.6(ii) and 2.7 imply that such triples satisfy

$$\begin{aligned} \frac{|\Delta|^{1/2}}{2} &\leq c \leq \frac{|\Delta|^{1/2}}{\sqrt{3}}(1 + \sqrt{3}\varepsilon + \varepsilon^2), \\ c/(1 + \sqrt{3}\varepsilon + \varepsilon^2) &< a \leq c, \\ a(1 - 2\varepsilon) &< |b| \leq a. \end{aligned}$$

Note that, since $\varepsilon \in (0, 1/3]$, we have $b \neq 0$ and $(1 + \sqrt{3}\varepsilon + \varepsilon^2)/\sqrt{3} < 1$.

Hence, to bound $\mathcal{C}_\varepsilon(\Delta)$ on the interval $[-X_{\max}, -X_{\min}]$, it suffices, for every Δ in this interval, to count the triples (a, b, c) satisfying

$$\begin{aligned} X_{\min}^{1/2}/2 &\leq c \leq X_{\max}^{1/2}, \\ c/(1 + \sqrt{3}\varepsilon + \varepsilon^2) &< a \leq c, \\ a(1 - 2\varepsilon) &< b \leq a \end{aligned}$$

and $b^2 - 4ac = \Delta$, and multiply the maximal count by 2 (because we counted only triples with positive b). We phrase this counting procedure formally as Algorithm 1.

```

Input : Two positive integers  $X_{\min} < X_{\max}$  and  $\varepsilon > 0$ 
Output: an upper bound for  $\mathcal{C}_\varepsilon(\Delta)$  for all discriminants  $\Delta \in [-X_{\max}, -X_{\min}]$ 
1 counter  $\leftarrow$  pointer to array of length  $X_{\max} - X_{\min} + 1$  initialized to 0;
2 bound  $\leftarrow 0$ ;
3 for  $c \leftarrow \lfloor X_{\min}^{1/2}/2 \rfloor$  to  $\lfloor X_{\max}^{1/2} \rfloor$  do
4   for  $a \leftarrow \lfloor c/(1 + \sqrt{3}\varepsilon + \varepsilon^2) \rfloor$  to  $c$  do
5     for  $b \leftarrow \lfloor (1 - 2\varepsilon)a \rfloor$  to  $a$  do
6        $X \leftarrow 4ac - b^2$ ;
7       if  $X \geq X_{\min}$  and  $X \leq X_{\max}$  then
8          $pos \leftarrow X - X_{\min}$ ;
9          $counter[pos] \leftarrow counter[pos] + 2$ ;
10        if  $counter[pos] > bound$  then  $bound \leftarrow counter[pos]$ ;
11      end
12    end
13  end
14 end
15 return bound;

```

Algorithm 1: Compute an upper bound for $\mathcal{C}_\varepsilon(\Delta)$ in the range $\Delta \in [-X_{\max}, -X_{\min}]$

For a correct implementation, we have to avoid floating point arithmetic in determining the upper bounds on a and b used in the inner two for-loops. For this, we note that $\varepsilon = 10^{-3}$ implies that

$$0.998c \leq c/(1 + \sqrt{3}\varepsilon + \varepsilon^2), \quad 0.998a = a(1 - 2\varepsilon).$$

Similarly, $\varepsilon = 4 \cdot 10^{-3}$ implies that

$$0.993c \leq c/(1 + \sqrt{3}\varepsilon + \varepsilon^2), \quad 0.992a = a(1 - 2\varepsilon).$$

As we are only interested in an upper bound on $\mathcal{C}_\varepsilon(\Delta)$ for these two specific values of ε , we use these weaker rational bounds in our implementation of Algorithm 1 by means of a C-program⁶. It verifies directly the assertions of the lemma. \square

Proof of Theorem 7.1. Assume that α is a singular unit of discriminant $\Delta \in (-10^{-10}, -3 \cdot 10^5]$. Let $0 < \varepsilon \leq 4 \cdot 10^{-3}$. We set again Y as in (5.4). As in the proof of Theorem 6.1 we find (6.9). We infer that

$$\begin{aligned} 1 &\leq 3 \frac{\mathcal{C}_\varepsilon(\Delta)}{\mathcal{C}(\Delta)Y} \log X + \frac{3 \log(\varepsilon^{-1}) - 10.65}{Y} \\ &\leq \frac{3\mathcal{C}_\varepsilon(\Delta)}{\pi} \frac{\log X}{X^{1/2}} + \frac{3 \log(\varepsilon^{-1}) - 10.65}{\frac{3}{\sqrt{5}} \log X - 9.78} \end{aligned}$$

⁶A link to our program `algorithm1.c` is on the second-named author's homepage. The running time on a regular desktop (Intel Xeon CPU E5-1620 v3, 3.50GHz, 32GB RAM) was under a minute for item (i) and a few milliseconds for item (ii). Its memory usage for $[X_{\min}, X_{\max}] = [1, 10^{10}]$ is significant (5 GB) but this can be overcome by splitting $[1, 10^{10}]$ into subintervals and running the program separately for each interval. This feature is also implemented in our program through the macro `DISC_BLOCK_SIZE`.

where we use $\mathcal{C}(\Delta)Y \geq \pi X^{1/2}$ and $Y \geq \frac{3}{\sqrt{5}} \log X - 9.78$.

If $X \in [10^7, 10^{10})$ then we set $\varepsilon = 10^{-3}$ and use the estimate $\mathcal{C}_\varepsilon(\Delta) \leq 16$ from Lemma 7.2(i). Recall that $x \mapsto (\log x)/x^{1/2}$ is decreasing for $x \geq e^2$. So we find

$$1 \leq \frac{3 \cdot 16 \log(10^7)}{\pi \cdot 10^{7/2}} + \frac{3 \log(1000) - 10.65}{\frac{3}{\sqrt{5}} \log(10^7) - 9.78} < 0.929,$$

a contradiction.

When $X \in [3 \cdot 10^5, 10^7)$ we set $\varepsilon = 4 \cdot 10^3$. Then $\mathcal{C}_\varepsilon(\Delta) \leq 6$ by Lemma 7.2(ii). Using $X \geq 3 \cdot 10^5$ we find as before

$$1 \leq \frac{3 \cdot 6 \log(3 \cdot 10^5)}{\pi \cdot (3 \cdot 10^5)^{1/2}} + \frac{3 \log(250) - 10.65}{\frac{3}{\sqrt{5}} \log(3 \cdot 10^5) - 9.78} < 0.961,$$

another contradiction which completes this proof. \square

8 The extra low-range

The results of the three previous sections reduce the proof of Theorem 1.1 to the following assertion.

Theorem 8.1. *Let Δ be the discriminant of a singular unit. Then $|\Delta| \geq 3 \cdot 10^5$.*

Proof. Let α be a singular unit of discriminant Δ . We write $X = |\Delta|$. We may assume that $X \geq 4$ because the only singular modulus of discriminant -3 is $j(\zeta_3) = 0$, which is not an algebraic unit.

Recall from Section 2 that the Galois conjugates of α are precisely the singular moduli $j(\tau)$, where $\tau = \tau(a, b, c)$ with (a, b, c) as in (2.6). The imaginary part of such τ is $X^{1/2}/(2a)$ and $a \leq (X/3)^{1/2}$ by Lemma 2.6(ii). Lemma 4.2 implies that

$$|j(\tau)| \geq e^{2\pi X^{1/2}/(2a)} - 2079 = e^{\pi X^{1/2}/a} - 2079 > 23^{X^{1/2}/a} - 2079$$

as $e^\pi > 23$. Using Lemmas 3.3 and 3.4, we find that

$$|j(\tau)| \geq 42700 \min \left\{ \frac{\sqrt{3}}{4X}, 4 \cdot 10^{-3} \right\}^3.$$

These bounds together show that

$$|j(\tau)| \geq \max \left\{ 23^{\lfloor X^{1/2}/a \rfloor} - 2079, 42700 \min \left\{ \frac{2}{5X}, \frac{1}{250} \right\}^3 \right\}. \quad (8.1)$$

Based on this observation, Algorithm 2 prints a list of discriminants of potential singular units in the range $[-X_{\max}, -4]$. For this purpose, it computes a rational lower bound P for the absolute value of the $\mathbb{Q}(\alpha)/\mathbb{Q}$ -norm of each singular moduli in this range. Those singular moduli where $P \leq 1$ are then flagged as potential singular units.

```

Input : An integer  $X_{\max} \geq 1$ 
Output: Print a list containing all discriminants in  $[-X_{\max}, -4]$  that are attached to
a potential singular unit.
1 for  $X \leftarrow 4$  to  $X_{\max}$  do
2    $\Delta \leftarrow -X$ ;
3   if  $\Delta \equiv 2$  or  $3 \pmod{4}$  then next  $X$  ;
4    $P \leftarrow 1$ ;
5   for  $a \leftarrow 1$  to  $\lfloor \sqrt{X/3} \rfloor$  do
6      $n \leftarrow \lfloor X^{1/2}/a \rfloor$ ;
7     for  $b \leftarrow -a + 1$  to  $a$  do
8       if  $b^2 \not\equiv \Delta \pmod{4a}$  then next  $b$  ;
9        $c \leftarrow (b^2 - \Delta)/(4a)$ ;
10      if  $a > c$  then next  $b$ ;
11      if  $a = c$  and  $b < 0$  then next  $b$ ;
12      if  $\gcd(a, b, c) \neq 1$  then next  $b$ ;
13       $P \leftarrow P \cdot \max\{23^n - 2079, 42700 \min\{2/(5X), 1/250\}^3\}$ ;
14    end
15  end
16  if  $P \leq 1$  then print  $\Delta$ ;
17 end

```

Algorithm 2: Exclude singular units

We have implemented this algorithm as a PARI script⁷. The script flags only -4 , -7 and -8 as discriminants of potential singular units. The singular moduli of these discriminants are well-known [7, (12.20)]: they are 12^3 , -15^3 and 20^3 , respectively. None of them is a unit, which concludes the proof. \square

As indicated in the introduction, Theorem 1.1 is the combination of Theorems 5.1, 6.1, 7.1 and 8.1.

References

- [1] A. Akbary, Z. Friggstad, and R. Juricevic, *Explicit upper bounds for $\prod_{p \leq p_{\omega(n)}} p/(p-1)$* , *Contrib. Discrete Math.* **2** (2007), 153–160.
- [2] Yu. Bilu and L. Kühne, *Linear Equations in Singular Moduli*, ArXiv e-prints (2017).
- [3] Yu. Bilu, F. Luca, and A. Pizarro-Madariaga, *Rational products of singular moduli*, *J. Number Theory* **158** (2016), 397–410.
- [4] 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** (2013), no. 1, 145–152.
- [5] L. Clozel and E. Ullmo, *Équidistribution des points de Hecke*, *Contributions to automorphic forms, geometry, and number theory*, Johns Hopkins Univ. Press, Baltimore, MD, 2004, pp. 193–254.

⁷A link to our PARI script `algorithm2.gp` is on the second-named author’s homepage. The running time is about 23 minutes on a regular desktop computer (Intel Xeon CPU E5-1620 v3, 3.50GHz, 32GB RAM). The only floating point operation used approximates $X^{1/2}$ which leads to $n = \lfloor X^{1/2}/a \rfloor$. To rule out a rounding error in the floating point arithmetic we compare $(an)^2$ with X in our implementation.

- [6] P. Colmez, *Sur la hauteur de Faltings des variétés abéliennes à multiplication complexe*, Compositio Math. **111** (1998), no. 3, 359–368.
- [7] D. A. Cox, *Primes of the form $x^2 + ny^2$* , John Wiley & Sons, 1989.
- [8] Pierre Deligne, *Preuve des conjectures de Tate et de Shafarevitch (d’après G. Faltings)*, Astérisque (1985), no. 121-122, 25–41, Seminar Bourbaki, Vol. 1983/84. MR 768952
- [9] D. R. Dorman, *Special values of the elliptic modular function and factorization formulae*, J. Reine Angew. Math. **383** (1988), 207–220.
- [10] W. Duke, *Hyperbolic distribution problems and half-integral weight Maass forms*, Invent. Math. **92** (1988), no. 1, 73–90.
- [11] É. Gaudron and G. Rémond, *Théorème des périodes et degrés minimaux d’isogénies*, Comment. Math. Helv. **89** (2014), no. 2, 343–403.
- [12] B. H. Gross and D. B. Zagier, *On singular moduli*, J. Reine Angew. Math. **355** (1985), 191–220.
- [13] P. Habegger, *Weakly bounded height on modular curves*, Acta Math. Vietnam. **35** (2010), no. 1, 43–69.
- [14] P. Habegger, *Singular moduli that are algebraic units*, Algebra Number Theory **9** (2015), no. 7, 1515–1524.
- [15] P. Habegger, G. Jones, and D. Masser, *Six unlikely intersection problems in search of effectivity*, Mathematical Proceedings of the Cambridge Philosophical Society (2016), 1–31.
- [16] L.-K. Hua, *Introduction to Number Theory*, Springer-Verlag, 1982.
- [17] L. Kühne, *An effective result of André-Oort type*, Ann. Math. (2) **176** (2012), 651–671.
- [18] Y. Nakkajima and Y. Taguchi, *A generalization of the Chowla-Selberg formula*, J. Reine Angew. Math. **419** (1991), 119–124.
- [19] J.-L. Nicolas and G. Robin, *Majorations explicites pour le nombre de diviseurs de n* , Canad. Math. Bull. **26** (1983), 485–492.
- [20] J. Pila, *O-minimality and the André-Oort conjecture for \mathbb{C}^n* , Ann. of Math. (2) **173** (2011), no. 3, 1779–1840.
- [21] N. Revol and F. Rouillier, *The MPFI library*, <http://perso.ens-lyon.fr/nathalie.revol/software.html>, 2017.
- [22] G. Robin, *Estimation de la fonction de Tchebychef θ sur le k -ième nombre premier et grandes valeurs de la fonction $\omega(n)$ nombre de diviseurs premiers de n* , Acta Arith. **42** (1983), 367–389.
- [23] G. Robin, *Grandes valeurs de fonctions arithmétiques et problèmes d’optimisation en nombres entiers*, Ph.D. thesis, Université de Limoges, 1983.
- [24] J. B. Rosser and L. Schoenfeld, *Approximate formulas for some functions of prime numbers*, Illinois J. Math. **6** (1962), 64–94.
- [25] C. L. Siegel, *Über die Classenzahl quadratischer Zahlkörper*, Acta Arith. **1** (1935), no. 1, 83–86.
- [26] G. Tenenbaum, *Introduction to analytic and probabilistic number theory*, third ed., Graduate Studies in Mathematics, vol. 163, American Mathematical Society, Providence, RI, 2015, Translated from the 2008 French edition by Patrick D. F. Ion.
- [27] The OEIS Foundation Inc., *The On-Line Encyclopedia of Integer Sequences*, 2017, <https://oeis.org/>.
- [28] The PARI Group, *PARI/GP (Version 2.9.2)*, 2017, Bordeaux, <http://pari.math.u-bordeaux.fr/>.

[29] The Sage Developers, *Sagemath, the Sage Mathematics Software System (Version 7.3)*, 2016, <http://www.sagemath.org>.

Yuri Bilu, IMB, UNIVERSITÉ DE BORDEAUX AND CNRS, 351 COURS DE LA LIBÉRATION, 33405 TALENCE CEDEX, FRANCE
yuri@math.u-bordeaux.fr

Philipp Habegger, DEPARTMENT OF MATHEMATICS AND COMPUTER SCIENCE, UNIVERSITY OF BASEL, SPIEGELGASSE 1, 4051 BASEL, SWITZERLAND
philipp.habegger@unibas.ch

Lars Kühne, DEPARTMENT OF MATHEMATICS AND COMPUTER SCIENCE, UNIVERSITY OF BASEL, SPIEGELGASSE 1, 4051 BASEL, SWITZERLAND
lars.kuehne@unibas.ch