

COMPOSITIO MATHEMATICA

There are at most finitely many singular moduli that are S -units

Sebastián Herrero, Ricardo Menares  and Juan Rivera-Letelier

Compositio Math. **160** (2024), 732–770.

[doi:10.1112/S0010437X23007704](https://doi.org/10.1112/S0010437X23007704)



FOUNDATION
COMPOSITIO
MATHEMATICA




LONDON
MATHEMATICAL
SOCIETY
EST. 1865





There are at most finitely many singular moduli that are S -units

Sebastián Herrero, Ricardo Menares  and Juan Rivera-Letelier

ABSTRACT

We show that for every finite set of prime numbers S , there are at most finitely many singular moduli that are S -units. The key new ingredient is that for every prime number p , singular moduli are p -adically disperse. We prove analogous results for the Weber modular functions, the λ -invariants and the McKay–Thompson series associated with the elements of the monster group. Finally, we also obtain that a modular function that specializes to infinitely many algebraic units at quadratic imaginary numbers must be a weak modular unit.

1. Introduction

A *singular modulus* is the j -invariant of an elliptic curve with complex multiplication. These algebraic numbers lie at the heart of the theory of abelian extensions of imaginary quadratic fields, as they generate the ring class fields of quadratic imaginary orders. This was predicted by Kronecker and referred to by himself as his *liebsten Jugendtraum*.

A result going back at least to Weber, states that every singular modulus is an algebraic integer [Web08, §115, *Satz VI*]. Thus, the absolute norm of a singular modulus is a rational integer, and the same holds for a difference of singular moduli. Gross and Zagier gave an explicit formula for the factorization of the absolute norms of differences of singular moduli [GZ85]. Roughly speaking, this formula shows that these absolute norms are highly divisible numbers. In fact, Li showed recently that the absolute norm of every difference of singular moduli is divisible by at least one prime number [Li21]. Equivalently, that no difference of singular moduli is an algebraic unit. Li’s work extends previous results of Habegger [Hab15] and of Bilu, Habegger and Kühne [BHK20]. These results answered a question raised by Masser in 2011, which was motivated by results of André–Oort type.

In view of these results, one is naturally led to look at differences of singular moduli whose absolute norms are only divisible by a given set of prime numbers. To be precise, recall that for a set of prime numbers S , an algebraic integer is an S -unit if the only prime numbers dividing its absolute norm are in S . The following is our main result.

MAIN THEOREM. *Let S be a finite set of prime numbers and j_0 a singular modulus. Then, there are at most finitely many singular moduli j such that $j - j_0$ is an S -unit.*

To prove this result we follow Habegger’s original strategy in the case where $S = \emptyset$ in [Hab15]. The main new ingredient is that for every prime number p , singular moduli are p -adically disperse

Received 27 July 2021, accepted in final form 11 August 2023, published online 5 March 2024.

2020 Mathematics Subject Classification 11F03, 11G15 (primary), 11G16, 11J68, 37P45 (secondary).

Keywords: modular functions, complex multiplication, singular moduli, S -units.

© 2024 The Author(s). The publishing rights in this article are licensed to Foundation Compositio Mathematica under an exclusive licence.

(Theorem B in §1.2). We also prove analogous results for a more general class of modular functions that includes the Weber modular functions, the λ -invariants and the McKay–Thompson series associated with the elements of the monster group (Theorem A in §1.1). In the course of the proof of these results, we obtain that a modular function that specializes to infinitely many algebraic units at quadratic imaginary numbers must be a weak modular unit (Theorem D in §1.4).

We also propose a conjecture whose affirmative solution would yield a vast generalization of the Main Theorem. The conjecture is that for every prime number p , every algebraic number is p -adically badly approximable by singular moduli (Conjecture 1.3 in §1.3). We show that an affirmative solution to this conjecture, would imply a version of the Main Theorem for every nonconstant modular function f for a congruence or genus zero group and every algebraic value of f (Corollary 5.2 in §5).

1.1 Singular moduli that are S -units

Consider the usual action of $\mathrm{SL}(2, \mathbb{R})$ on the upper-half plane \mathbb{H} and consider the j -invariant as a holomorphic function defined on \mathbb{H} that is invariant under $\mathrm{SL}(2, \mathbb{Z})$. Moreover, denote by $\overline{\mathbb{Q}}$ the algebraic closure of \mathbb{Q} inside \mathbb{C} .

A subgroup Γ of $\mathrm{SL}(2, \mathbb{R})$ is *commensurable to* $\mathrm{SL}(2, \mathbb{Z})$, if the intersection $\Gamma \cap \mathrm{SL}(2, \mathbb{Z})$ has finite index in Γ and in $\mathrm{SL}(2, \mathbb{Z})$. For such a group, denote by $X(\Gamma)$ the Riemann surface obtained by compactifying the quotient $\Gamma \backslash \mathbb{H}$. The *genus of* Γ is the genus of $X(\Gamma)$. A *modular function for* Γ is a meromorphic function defined on \mathbb{H} that is obtained by lifting the restriction to $\Gamma \backslash \mathbb{H}$ of a meromorphic function defined on $X(\Gamma)$. A meromorphic function defined on \mathbb{H} is a modular function if and only if it is algebraically dependent with the j -invariant over \mathbb{C} (Proposition 2.1).

A modular function is *defined over* $\overline{\mathbb{Q}}$, if it is algebraically dependent with the j -invariant over $\overline{\mathbb{Q}}$. In this case, a *singular modulus of* f is a finite value that f takes at a quadratic imaginary number. Every singular modulus of f is in $\overline{\mathbb{Q}}$ (Proposition 2.3(i) in §2.2). We show that every modular function whose Fourier series expansion at $i\infty$ has coefficients in $\overline{\mathbb{Q}}$ is defined over $\overline{\mathbb{Q}}$ (Proposition A.1 in Appendix A).

Recall that for a set of prime numbers S , a number in $\overline{\mathbb{Q}}$ is an S -unit if the leading and constant coefficients of its minimal polynomial in $\mathbb{Z}[X]$ have all their prime factors in S .

THEOREM A. *Let f be a nonconstant modular function defined over $\overline{\mathbb{Q}}$ for a genus-zero group. Moreover, let f_0 be a singular modulus of f and let S be a finite set of prime numbers. Then, there are at most finitely many singular moduli \mathfrak{f} of f such that $\mathfrak{f} - f_0$ is an S -unit.*

Since $\mathrm{SL}(2, \mathbb{Z})$ is of genus zero and the j -invariant is a nonconstant modular function for $\mathrm{SL}(2, \mathbb{Z})$ defined over $\overline{\mathbb{Q}}$, the Main Theorem is Theorem A applied to the j -invariant. Theorem A also applies to the Weber modular functions, the λ -invariants and the McKay–Thompson series associated with the elements of the monster group. See §1.5 for details.

The following corollary is a direct consequence of Theorem A with f equal to the j -invariant and $f_0 = 0$, which is the j -invariant of every elliptic curve whose endomorphism ring is isomorphic to $\mathbb{Z}[(1 + \sqrt{3}i)/2]$.

COROLLARY 1.1. *For every finite set of prime numbers S , there are at most finitely many singular moduli of the j -invariant that are S -units.*

When restricted to the j -invariant and $S = \emptyset$, Theorem A is a particular instance of [Hab15, Theorem 2] and of [Li21, Corollary 1.3 with $m = 1$].¹ This last result extends the main result of [BHK20], that in the case where $S = \emptyset$ the set of singular moduli in Corollary 1.1 is empty. In contrast to these results, the proof of Theorem A, which follows the strategy of proof of [Hab15, Theorem 2] in the case where f is the j -invariant and $S = \emptyset$, does not give an effectively computable upper bound.

The number -2^{15} is an example of a singular modulus of the j -invariant that is a $\{2\}$ -unit. In fact, -2^{15} is the j -invariant of every elliptic curve whose endomorphism ring is isomorphic to $\mathbb{Z}[(1 + \sqrt{-11}i)/2]$. Numerical computations suggest an affirmative answer to the following question; see, e.g., [Sut].

Question 1.2. Is -2^{15} the unique singular modulus of the j -invariant that is a $\{2\}$ -unit?

For $j_0 = 0$ or 1728 and for the infinite set of prime numbers S for which every elliptic curve with j -invariant equal to j_0 has potential ordinary reduction, Campagna showed the following in [Cam21]: if j is a singular modulus of the j -invariant such that $j - j_0$ is an S -unit, then $j - j_0$ is, in fact, an algebraic unit. A combination of the Main Theorem and the arguments of Campagna shows that when $j_0 = 0$ or 1728 , the conclusion of the Main Theorem holds for some infinite sets of prime numbers S ; see § 1.5.

1.2 Singular moduli are disperse

Denote by $M_{\mathbb{Q}}$ the set of all prime numbers together with ∞ , put $\mathbb{C}_{\infty} := \mathbb{C}$ and denote by $|\cdot|_{\infty}$ the usual absolute value on \mathbb{C} . Moreover, for each prime number p let $(\mathbb{C}_p, |\cdot|_p)$ be a completion of an algebraic closure of the field of p -adic numbers \mathbb{Q}_p , and identify the algebraic closure of \mathbb{Q} inside \mathbb{C}_p with $\overline{\mathbb{Q}}$. For all v in $M_{\mathbb{Q}}$, α in \mathbb{C}_v and $r > 0$, put

$$\mathbf{D}_v(\alpha, r) := \{z \in \mathbb{C}_v : |z - \alpha|_v < r\}.$$

For a finite extension K of \mathbb{Q} inside $\overline{\mathbb{Q}}$, consider the Galois group $\text{Gal}(\overline{\mathbb{Q}}|K)$ and for each α in $\overline{\mathbb{Q}}$ denote by $O_K(\alpha)$ its orbit by $\text{Gal}(\overline{\mathbb{Q}}|K)$. The following result is stated for a modular function that is ‘defined over K ’ in the sense of Definition 2.2 in § 2.2. For a modular function to be defined over K , it is sufficient that its Fourier series expansion at $i\infty$ has coefficients in K (Proposition A.1 in Appendix A).

THEOREM B (Singular moduli are disperse). *Let K be a finite extension of \mathbb{Q} inside \mathbb{C} and let f be a nonconstant modular function defined over K . Then, for all v in $M_{\mathbb{Q}}$, α in \mathbb{C}_v and $\varepsilon > 0$, there is $r > 0$ such that the following property holds. For every singular modulus \mathfrak{f} of f such that $\# O_K(\mathfrak{f})$ is sufficiently large, we have*

$$\#(O_K(\mathfrak{f}) \cap \mathbf{D}_v(\alpha, r)) \leq \varepsilon \cdot \# O_K(\mathfrak{f}).$$

We first establish this result for the j -invariant, and then deduce the general case from this special case. The case where $v = \infty$ and f is the j -invariant is a direct consequence of the fact that the asymptotic distribution of the singular moduli of the j -invariant is given by a nonatomic measure [Duk88, CU04]. In the case where v is a prime number p , there are infinitely many measures that describe the p -adic asymptotic distribution of the singular moduli of the j -invariant. The main ingredient in the proof of Theorem B is that none of these measures has an atom in \mathbb{C}_p (Theorem 3.1 in § 3). We also prove an analogous result for the Hecke orbit of every point in \mathbb{C}_p (Theorem 3.2 in § 3.1). As a consequence, we obtain that a Hecke orbit cannot

¹ See Theorem D in § 1.4 for an extension of Habegger’s result to a general modular function defined over $\overline{\mathbb{Q}}$ and § 1.5 for further comments on Li’s result.

have a significant proportion of good approximations of a given point in \mathbb{C}_p (Corollary 3.4 in §3.1), thus improving a result of Charles in [Cha18]. The proofs of these results are based on the description of all the measures describing the p -adic asymptotic distribution of singular moduli and Hecke orbits, given in the companion papers [HMR20, HMR21].

1.3 Approximation by singular moduli

Let f be a nonconstant modular function, denote by Γ its stabilizer in $SL(2, \mathbb{R})$ and denote by f_0 the meromorphic function defined on $X(\Gamma)$ induced by f . A complex number is a *cuspidal value* of f if it is a value that f_0 takes at a cusp of $X(\Gamma)$. Note that the number of cuspidal values is finite. Moreover, f is a *Hauptmodul* if $X(\Gamma)$ is of genus zero and f_0 is a biholomorphism from $X(\Gamma)$ onto the Riemann sphere.

Let f be a nonconstant modular function defined over $\overline{\mathbb{Q}}$ and let v be in $M_{\mathbb{Q}}$. A number α in \mathbb{C}_v is *badly approximable in \mathbb{C}_v by the singular moduli of f* , if there are constants $A > 0$ and B such that for every singular modulus \mathfrak{f} of f different from α we have

$$-\log |\mathfrak{f} - \alpha|_v \leq A \log(\# O_{\mathbb{Q}}(\mathfrak{f})) + B.$$

If this property does not hold, then α is *well approximated in \mathbb{C}_v by the singular moduli of f* .

THEOREM C. *Let f be a nonconstant modular function defined over $\overline{\mathbb{Q}}$, let \mathfrak{f}_0 be a singular modulus of f and let v be in $M_{\mathbb{Q}}$. In the case where $v = \infty$, assume that \mathfrak{f}_0 is a non-cuspidal value of f and in the case where v is a prime number, assume that f is a Hauptmodul. Then, \mathfrak{f}_0 is badly approximable in \mathbb{C}_v by the singular moduli of f .*

In the case where $v = \infty$, the hypothesis that \mathfrak{f}_0 is a non-cuspidal value of f is necessary; see Proposition 2.7(i).

In the case where f is the j -invariant and $v = \infty$, the theorem above is a direct consequence of a result of Habegger [Hab15, Lemmas 5 and 8 and formula (11), or the proof of Lemma 6]. In fact, using results of David and Hirata-Kohno in [DH09], Habegger proved the stronger result that every algebraic number is badly approximable in \mathbb{C} by the singular moduli of the j -invariant. It is unclear to us whether the analogous result holds in the p -adic setting.

CONJECTURE 1.3. *Let p be a prime number. Then, every algebraic number is badly approximable in \mathbb{C}_p by the singular moduli of the j -invariant.*

We show that an affirmative solution to this conjecture would yield a version of Theorem A for a general congruence or genus zero group and a general algebraic value (Corollary 5.2 in §5).

1.4 Weak modular units are the only source of singular units

A *modular unit* is a modular function without zeros or poles in \mathbb{H} . A *weak modular unit* is a modular function u for which 0 is a cuspidal value of u and of $1/u$. Note that every nonconstant modular unit is a weak modular unit.

The singular moduli of modular units defined over $\overline{\mathbb{Q}}$ are a natural source of algebraic units; see, e.g., [KL81]. Roughly speaking, the following result asserts that among modular functions defined over $\overline{\mathbb{Q}}$, weak modular units are the only source of singular moduli that are algebraic units.

THEOREM D. *Let f be a nonconstant modular function defined over $\overline{\mathbb{Q}}$ that is not a weak modular unit. Then, there are at most finitely many singular moduli of f that are algebraic units.*

We also show that an affirmative solution to Conjecture 1.3 would imply a version of Theorem D for S -units (Corollary 5.1 in §5) and a version of Theorem D that holds under the weaker hypothesis that f is not a modular unit, but that is restricted to congruence or to genus-zero groups (Corollary 5.2 in §5). Note that for every modular unit f defined over $\overline{\mathbb{Q}}$, there is a finite set of prime numbers S such that every singular modulus of f is an S -unit; see Corollary 2.5(ii) in §2.3.

An *elliptic unit* is an algebraic unit that is the value of a modular unit defined over $\overline{\mathbb{Q}}$ at a quadratic imaginary number. A natural problem that arises from Theorem D is to determine those modular units that specialize to infinitely many elliptic units at quadratic imaginary numbers. Examples of such can be easily extracted from Weber’s book [Web08]. Recall that the *Weber modular functions* \mathbf{f} , \mathbf{f}_1 and \mathbf{f}_2 are given in terms of Dedekind’s η function by

$$\mathbf{f}(\tau) := \exp\left(-\frac{\pi i}{24}\right) \frac{\eta((\tau + 1)/2)}{\eta(\tau)}, \quad \mathbf{f}_1(\tau) := \frac{\eta(\tau/2)}{\eta(\tau)} \quad \text{and} \quad \mathbf{f}_2(\tau) := \sqrt{2} \frac{\eta(2\tau)}{\eta(\tau)};$$

see, e.g., [Web08, §34, (9)]. If p is a prime number satisfying $p \equiv -1 \pmod{8}$ and we put $\tau_p := \sqrt{pi}$, then the singular modulus $\mathbf{f}((\tau_p - 1)/(\tau_p + 1))$ of \mathbf{f} is equal to $\sqrt{2}/\mathbf{f}(\tau_p)$ by [Web08, §34, (18)] and it is an algebraic unit by [Web08, §142, p. 540]. Together with [Web08, §34, (13), (14)], this implies that the singular moduli $\mathbf{f}_1(-2/(\tau_p + 1))$ and $\mathbf{f}_2((\tau_p + 1)/2)$ of \mathbf{f}_1 and \mathbf{f}_2 are both algebraic units.

Other examples of modular units that specialize to infinitely many elliptic units can be found in [KL81]. We mention the λ -invariants or *modular λ functions*. These are six *Hauptmoduln* for the principal congruence group of level two, which can be defined as the roots of

$$256(1 - X + X^2)^3 - jX^2(1 - X)^2 = 0; \tag{1.1}$$

see, e.g., [Lan87, Chapter 18, §6]. Clearly, every singular modulus of a λ -invariant is a $\{2\}$ -unit. By, e.g., [Lan87, Chapter 12, §2, Corollary of Theorem 5] or the more recent results of Yang, Yin and Yu [YYY21, Theorem 1.1], each of the six λ -invariants has infinitely many singular moduli that are algebraic units.²

To prove Theorem D, we follow the strategy of proof of [Hab15, Theorem 2]. In particular, we use [Hab15, Lemmas 5 and 8 and formula (11)], whose proof is based on results of David and Hirata-Kohno in [DH09].

1.5 Notes and references

Theorem A applies to the Weber modular functions \mathbf{f} , \mathbf{f}_1 and \mathbf{f}_2 ; see §1.4 for the definition. In fact, \mathbf{f}_2 is a *Hauptmodul* by [YY16, Theorem 1.3(2)(a) and p. 19], and therefore so are \mathbf{f} and \mathbf{f}_1 by [Web08, §34, (13) and (14)]. On the other hand, each of these functions is defined over \mathbb{Q} in the sense of Definition 2.2 in §2.1 because it is a root of either

$$(X^{24} - 16)^3 - X^{24}j = 0 \quad \text{or} \quad (X^{24} + 16)^3 - X^{24}j = 0; \tag{1.2}$$

see, e.g., [YZ97] or [Web08, §126, (1)]. The singular moduli of Weber modular functions provide generators of ring class fields of quadratic imaginary orders; see [Web08, §126] and [Sch76, Satz 4.2]. In addition, the arithmetic complexity of these generators is sometimes significantly lower than that of the corresponding singular moduli of the j -invariant; see [YZ97] and [ES10] for a computational perspective. Since each of the functions \mathbf{f} , \mathbf{f}_1 and \mathbf{f}_2 is a root of one of

² Although these results only apply directly to one of the six λ -invariants, they automatically imply analogous results for each of the remaining five λ -invariants. Note that for every pair of λ -invariants λ_0 and λ_1 , there is γ in $\text{SL}(2, \mathbb{Z})$ such that $\lambda_1 = \lambda_0 \circ \gamma$.

the polynomials in (1.2), each of their singular moduli is an algebraic integer and a $\{2\}$ -unit. Moreover, as mentioned in § 1.4, the results of Weber imply that the singular moduli of \mathbf{f} , \mathbf{f}_1 and \mathbf{f}_2 are often algebraic units, in contrast to the singular moduli of the j -invariant. Note that 0 is a singular modulus of the j -invariant, but not of \mathbf{f} , \mathbf{f}_1 or \mathbf{f}_2 . In fact, \mathbf{f} , \mathbf{f}_1 and \mathbf{f}_2 are all modular units; see, e.g., Corollary 2.5(ii) in § 2.3.

Theorem A also applies to each of the six λ -invariants; see § 1.4 for the definition. This solves affirmatively a conjecture of Habegger (private communication, 2021). Note that each of these functions is defined over \mathbb{Q} , because it is a root of (1.1). As mentioned in § 1.4, the singular moduli of each of the six λ -invariants are often algebraic units, in contrast to those of the j -invariant. Note that each of the λ -invariants is a modular unit; see, e.g., Corollary 2.5(ii) in § 2.3.

The representation theory of the monster group provides a wealth of modular functions satisfying the hypotheses of Theorem A. In fact, by Borcherds' solution [Bor92, Theorem 1.1] of the monstrous moonshine conjecture of Conway and Norton [CN79], the McKay–Thompson series associated with a given element of the monster group is a *Hauptmodul* defined over \mathbb{Q} ; see Proposition A.1 in Appendix A. Moreover, the singular moduli of a fundamental McKay–Thompson series are often algebraic integers [CY96, Theorem I].

For distinct singular moduli j and j' of the j -invariant, Li gives in [Li21] an explicit lower bound for the absolute norm of $j - j'$ that implies that this algebraic integer is not an algebraic unit. When restricted to $j' = 0$, this is [BHK20, Theorem 1.1]. In fact, Li proves a stronger result for the values of modular polynomials at pairs of singular moduli of the j -invariant. Li's approach makes use of (extensions of) the work of Gross and Zagier in [GZ85], and it is different from those in [Hab15, BHK20]. Li does not treat the case of S -units in [Li21].

In the case where $j_0 = 0$ (respectively, 1728), the conclusion of the Main Theorem holds for certain classes of infinite sets of prime numbers S . In fact, if we put

$$S_0 := \{q: \text{prime number, } q \equiv 1 \pmod{3}\}$$

(respectively, $\{q: \text{prime number, } q \equiv 1 \pmod{4}\}$),

then the conclusion of the Main Theorem holds for every set of prime numbers S such that $S \setminus S_0$ is finite and does not contain $\{2, 3, 5\}$ (respectively, $\{2, 3, 7\}$). This is a direct consequence of the Main Theorem and the proof of [Cam21, Theorems 1.2 and 6.1].

1.6 Organization

In § 2 we establish general properties of modular functions (§ 2.1), their singular moduli (§ 2.2) and their cuspidal and omitted values (§ 2.3).

In § 3 we prove Theorem B. We first establish it in the special case of the j -invariant. The main ingredient in the proof of this special case, is that no measure describing the v -adic asymptotic distribution of the singular moduli of the j -invariant has an atom in \mathbb{C}_v . This follows from [CU04, Théorème 2.4] if $v = \infty$ and is stated as Theorem 3.1 in the case where v is a prime number. Together with [HMR21, Theorems A and B], this implies Theorem B in the case of the j -invariant as a direct consequence. The proof of Theorem 3.1 is based on the description of all these measures given in the companion papers [HMR20, HMR21]. We also use an analogous description for Hecke orbits given in [HMR20, HMR21]. We first establish a result analogous to Theorem 3.1 for Hecke orbits (Theorem 3.2) in § 3.1, and in § 3.2 we deduce Theorem 3.1 from this result. To prove Theorem 3.2, we first show that the images of a point under Hecke correspondences associated with different prime numbers are nearly disjoint (Lemma 3.5). We use this to show that an atom in \mathbb{C}_p of an accumulation measure of a Hecke orbit would

replicate indefinitely, thus creating infinite mass.³ In §3.3 we deduce Theorem B in the general case from the special case of the j -invariant, using the results about modular functions in §2.

In §4 we prove Theorem C. We first establish it in the special case of the j -invariant, which is stated in a slightly different form as Proposition 4.1. After a brief review of the work of Gross and Hopkins on deformation spaces of formal modules in §4.1, in §4.2 we give the proof of Proposition 4.1. First, we use that singular moduli are isolated in the ordinary reduction locus [HMR20, Corollary B], to restrict to the case where j and j_0 are both in the supersingular reduction locus. In the case where the conductors of D_j and D_{j_0} are both p -adic units, we use an idea in the proof of [Cha18, Proposition 5.11]. To extend this estimate to the general case, we use a formula in [HMR21] that shows how the canonical branch of the Hecke correspondence T_p relates CM points whose conductors differ by a power of p . In §4.3 we deduce Theorem C in the general case from the special case of the j -invariant, using the results about modular functions in §2.

In §5 we prove Theorems A and D. We follow Habegger's original strategy in the case of the j -invariant and $S = \emptyset$ in [Hab15], to prove a more general result that we state as Theorem A'. In particular, we use in a crucial way Colmez's bound [Col98, Théorème 1] in the form of [Hab15, Lemma 3]. The main new ingredient to implement Habegger's strategy is Theorem B. Theorem A' implies Theorem D as a direct consequence. Another direct consequence of Theorem A' is that an affirmative solution to Conjecture 1.3 would yield a version of Theorem D for S -units (Corollary 5.1) and a version of Theorem A for a general congruence or genus-zero group and a general algebraic value (Corollary 5.2). We prove Theorem A' in §5.1 and derive Theorem A and Corollary 5.2 from Theorem A' in §5.2.

2. Modular functions and their special values

In this section we prove general properties of modular functions, their singular moduli and their cuspidal and omitted values. In §2.1 we establish some general properties of modular functions (Proposition 2.1). In §2.2 we study arithmetic properties of singular moduli of modular functions defined over $\overline{\mathbb{Q}}$ (Proposition 2.3). Finally, in §2.3 we study cuspidal and omitted values of modular functions.

2.1 Modular functions

The goal of this section is to prove the following proposition.

PROPOSITION 2.1. *Every modular function is algebraically dependent with the j -invariant over \mathbb{C} . Conversely, let K be a subfield of \mathbb{C} and let f be a nonconstant meromorphic function defined on \mathbb{H} that is algebraically dependent with the j -invariant over K . Then, f is a modular function and there is a polynomial $\Phi(X, Y)$ with coefficients in a finite extension of K inside \mathbb{C} that is irreducible over \mathbb{C} and such that $\Phi(j, f)$ vanishes identically. Furthermore, $\Phi(X, Y)$ depends on both X and Y , and it satisfies the following properties.*

- (i) *For every (z, w) in the zero set of Φ in $\mathbb{C} \times \mathbb{C}$, there is τ in \mathbb{H} satisfying*

$$z = j(\tau) \quad \text{and} \quad w = f(\tau).$$

- (ii) *Up to a constant factor, Φ is the unique irreducible polynomial in $\mathbb{C}[X, Y]$ such that $\Phi(j, f)$ vanishes identically.*

³ See Remark 3.6 for a different strategy of proof.

In the proof of this proposition, which is given below, and in the rest of the paper, we use the following property: for every subfield K of \mathbb{C} , a polynomial in $K[X, Y]$ is irreducible over \mathbb{C} if and only if it is irreducible over an algebraic closure of K .

Definition 2.2. Let K be a subfield of \mathbb{C} . A modular function f is *defined over K* , if there is a polynomial $\Phi(X, Y)$ in $K[X, Y]$ that is irreducible over \mathbb{C} and such that $\Phi(j, f)$ vanishes identically. In this case, $\Phi(X, Y)$ is a *modular polynomial of f* .

In view of Proposition 2.1, in the case where $K = \overline{\mathbb{Q}}$ this definition coincides with that given in § 1.1. Note that if K is a subfield of \mathbb{C} , then every modular function having a modular polynomial in $K[X, Y]$ is defined over K . On the other hand, by Proposition 2.1 for every modular function f there is a modular polynomial of j and f in $\mathbb{C}[X, Y]$ and if, in addition, f is algebraically dependent with the j -invariant over K , then there is a modular polynomial of j and f with coefficients in a finite extension of K . Furthermore, a modular polynomial in $K[X, Y]$ of j and a modular function is unique up to a multiplicative constant in K^\times .

Proof of Proposition 2.1. Let f be a modular function. The case where f is constant being trivial, assume f is nonconstant. Let Γ be a subgroup of $\mathrm{SL}(2, \mathbb{R})$ that is commensurable to $\mathrm{SL}(2, \mathbb{Z})$ and such that f is invariant under Γ . Replacing Γ by $\Gamma \cap \mathrm{SL}(2, \mathbb{Z})$ if necessary, assume that Γ is a finite index subgroup of $\mathrm{SL}(2, \mathbb{Z})$. Then, the j -invariant and f induce meromorphic functions j_0 and f_0 defined on $X(\Gamma)$. Since the field of meromorphic functions defined on $X(\Gamma)$ has transcendence degree one over \mathbb{C} , there is a nonzero polynomial $\Phi(X, Y)$ in $\mathbb{C}[X, Y]$ such that $\Phi(j_0, f_0)$ vanishes identically. It follows that the function $\Phi(j, f)$ vanishes identically. This implies that the j -invariant and f are algebraically dependent over \mathbb{C} .

To prove the second assertion, let K be a subfield of \mathbb{C} and let f be a nonconstant meromorphic function defined on \mathbb{H} that is algebraically dependent with the j -invariant over K . Then, there is a polynomial $\Phi_0(X, Y)$ in $K[X, Y]$ such that $\Phi_0(j, f)$ vanishes identically. Suppose Φ_0 is not irreducible over \mathbb{C} . Then, we can find a finite extension \widehat{K} of K inside \mathbb{C} and a finite family $(\Phi_i)_{i \in I}$ of polynomials in $\widehat{K}[X, Y]$ that are irreducible over \mathbb{C} and whose product is equal to Φ_0 . It follows that at least one of the meromorphic functions in $\{\Phi_i(j, f) : i \in I\}$ vanishes identically. This proves that in all the cases there is a polynomial Φ with coefficients in a finite extension of K that is irreducible over \mathbb{C} and such that $\Phi(j, f)$ vanishes identically. Note also that, since the j -invariant is nonconstant and f is nonconstant by assumption, the polynomial Φ depends on both variables.

To prove that f is a modular function, denote by $\mathcal{M}(\mathbb{H})$ the field of all meromorphic functions defined on \mathbb{H} . Note that the polynomial $\Phi(j, Y)$ in $\mathcal{M}(\mathbb{H})[Y]$ is nonconstant and denote by \mathcal{Z} its finite number of zeros in $\mathcal{M}(\mathbb{H})$. The set \mathcal{Z} contains f and is invariant under the action of $\mathrm{SL}(2, \mathbb{Z})$ on $\mathcal{M}(\mathbb{H})$ given by $(\gamma, g) \mapsto g \circ \gamma$. Since \mathcal{Z} is finite, it follows that the stabilizer Γ of f in $\mathrm{SL}(2, \mathbb{R})$ is a finite index subgroup of $\mathrm{SL}(2, \mathbb{Z})$. Thus, to prove that f is a modular function, it is sufficient to prove that f is the lift of the restriction to $\Gamma \backslash \mathbb{H}$ of a meromorphic function defined on $X(\Gamma)$. To do this, it is sufficient to show that for every γ in $\mathrm{SL}(2, \mathbb{Z})$ the function $f \circ \gamma$ is meromorphic at $i\infty$; see, e.g., [Shi71, Proposition 1.30 and § 1.4]. Note that the functions $f \circ \gamma$ and $j \circ \gamma = j$ are algebraically dependent over \mathbb{C} . Thus, replacing $f \circ \gamma$ by f if necessary, it is sufficient to prove that f is meromorphic at $i\infty$. To do this, denote by δ the degree of $\Phi(X, Y)$ in Y and for each k in $\{0, \dots, \delta\}$ denote by $P_k(X)$ the coefficient of Y^k in $\Phi(X, Y)$ and by d_k the degree of P_k . Then, there are constants $R > 0$ and $C > 0$, such that for every z in \mathbb{C} satisfying $|z| \geq R$ we have $|P_\delta(z)| \geq C^{-1}|z|^{d_\delta}$, and such that every k in $\{0, \dots, \delta - 1\}$ we have $|P_k(z)| \leq C|z|^{d_k}$.

Thus, if we put

$$d := -d_\delta + \max\{d_1, \dots, d_{\delta-1}\},$$

then for every τ in \mathbb{H} satisfying

$$|j(\tau)| \geq R \quad \text{and} \quad |f(\tau)| \geq 1,$$

we have

$$|f(\tau)|^\delta = |P_\delta(j(\tau))|^{-1} \left| \sum_{k=0}^{\delta-1} P_k(j(\tau)) f(\tau)^k \right| \leq C^2 \delta |j(\tau)|^d |f(\tau)|^{\delta-1}.$$

Since the j -invariant has a pole at $i\infty$ (see, e.g., [Lan87, Chapter 4, §1]), we conclude that for every τ in \mathbb{H} whose imaginary part is sufficiently large we have $|f(\tau)| \leq C^2 \delta |j(\tau)|^d$. This implies that f is meromorphic at $i\infty$ and completes the proof that f is a modular function.

To prove item (i), let j_0 and f_0 be as above, note that the set of poles of j_0 is equal to the complement of $\Gamma \backslash \mathbb{H}$ in $X(\Gamma)$ and denote by P the set of poles of f_0 in $X(\Gamma)$. Moreover, denote by $\widehat{\Phi}$ the homogenization of Φ in $\mathbb{C}[X, Y, Z]$, so that $\widehat{\Phi}(X, Y, 1) = \Phi(X, Y)$, and let $Z(\widehat{\Phi})$ its zero set in $\mathbb{P}^2(\mathbb{C})$. Then,

$$\begin{aligned} (\Gamma \backslash \mathbb{H}) \setminus P &\rightarrow Z(\widehat{\Phi}) \\ x &\mapsto \iota(x) := [j_0(x) : f_0(x) : 1] \end{aligned}$$

has a unique continuous extension $\iota: X(\Gamma) \rightarrow Z(\widehat{\Phi})$ and this map is surjective; see, e.g., [Har77, Chapter II, Proposition 6.8]. Thus, for every (z, w) in $\mathbb{C} \times \mathbb{C}$ in the zero set of Φ there is x in $(\Gamma \backslash \mathbb{H}) \setminus P$ such that $\psi(x) = [z : w : 1]$. This implies item (i).

To prove item (ii), let $\check{\Phi}(X, Y)$ be an irreducible polynomial in $\mathbb{C}[X, Y]$ such that $\check{\Phi}(j, f)$ vanishes identically. Then, by item (i) the polynomial $\check{\Phi}$ vanishes on the zero set of Φ and, therefore, $\check{\Phi}$ is a multiple of Φ . Since, by assumption, $\check{\Phi}$ is irreducible, it follows that it is a constant multiple of Φ . This completes the proof of item (ii) and of the proposition. \square

2.2 Singular moduli of modular functions

The goal of this section is to establish some arithmetic properties of singular moduli of modular functions defined over $\overline{\mathbb{Q}}$. These are gathered in Proposition 2.3 below. To state it, we introduce some terminology.

For a finite extension K of \mathbb{Q} inside $\overline{\mathbb{Q}}$, consider the Galois group $\text{Gal}(\overline{\mathbb{Q}}|K)$ and for each α in \mathbb{C} denote by $O_K(\alpha)$ its orbit by $\text{Gal}(\overline{\mathbb{Q}}|K)$. Note that, with the notation introduced in § 1.2 we have $O(\alpha) = O_{\mathbb{Q}}(\alpha)$.

In this paper, a *discriminant* is the discriminant of an order in a quadratic imaginary extension of \mathbb{Q} . For every discriminant D denote by $h(D)$ the class number of the order of discriminant D in $\mathbb{Q}(\sqrt{D})$. For a singular modulus j , the discriminant of the endomorphism ring of an elliptic curve over \mathbb{C} whose j -invariant is equal to j only depends on j . Denote it by D_j . We use that for every singular modulus j of the j -invariant, we have

$$O_{\mathbb{Q}}(j) = \{\text{singular modulus } j' \text{ of the } j\text{-invariant with } D_{j'} = D_j\} \tag{2.1}$$

and

$$\# O_{\mathbb{Q}}(j) = h(D_j); \tag{2.2}$$

see, e.g., [Lan87, Chapter 10, Theorem 5].

PROPOSITION 2.3. *Let K be a finite extension of \mathbb{Q} inside \mathbb{C} . Then, for every nonconstant modular function f defined over K the following properties hold.*

- (i) Every singular modulus f_0 of f is in $\overline{\mathbb{Q}}$ and every element of $O_K(f_0)$ is also a singular modulus of f .
- (ii) There is a constant $C_0 > 0$ such that for every quadratic imaginary number τ_0 in \mathbb{H} that is not a pole of f , the singular modulus $f(\tau_0)$ of f satisfies

$$C_0^{-1} \cdot \# O_{\mathbb{Q}}(j(\tau_0)) \leq \# O_K(f(\tau_0)) \leq \# O_{\mathbb{Q}}(f(\tau_0)) \leq C_0 \cdot \# O_{\mathbb{Q}}(j(\tau_0)).$$

- (iii) For every $\varepsilon > 0$ there is a constant $C_1 > 0$ such that the following property holds. For every quadratic imaginary number τ in \mathbb{H} that is not a pole of f , we have

$$C_1^{-1} \cdot \# O_K(f(\tau))^{2-\varepsilon} \leq |D_{j(\tau)}| \leq C_1 \cdot \# O_K(f(\tau))^{2+\varepsilon}.$$

- (iv) For every $\varepsilon > 0$ there is a constant $C_2 > 0$, such that for every $R > 0$ we have

$$\#\{\text{singular modulus } \mathfrak{f} \text{ of } f \text{ such that } \# O_K(\mathfrak{f}) \leq R\} \leq C_2 R^{3+\varepsilon}. \tag{2.3}$$

Proof. Let $\Phi(X, Y)$ be a modular polynomial of f in $K[X, Y]$. Denote by δ_X (respectively, δ_Y) the degree of $\Phi(X, Y)$ in X (respectively, Y).

To prove items (i) and (ii), let f_0 be a singular modulus of f and let τ_0 be a quadratic imaginary number in \mathbb{H} such that $f_0 = f(\tau_0)$. Then, $j_0 := j(\tau_0)$ is a singular modulus of the j -invariant and, therefore, it is in $\overline{\mathbb{Q}}$. On the other hand, the polynomial $\Phi(j_0, Y)$ is nonzero because Φ is irreducible over \mathbb{C} and j is nonconstant. Since f_0 is a root of $\Phi(j_0, Y)$, it is in an extension of $K(j_0)$ of degree at most δ_Y . In particular, f_0 is in $\overline{\mathbb{Q}}$. To complete the proof of item (i), let σ in $\text{Gal}(\overline{\mathbb{Q}}|K)$ be given, and note that $\Phi(\sigma(j_0), \sigma(f_0)) = 0$. Since Φ is irreducible over \mathbb{C} , by Proposition 2.1 there is τ in \mathbb{H} such that

$$\sigma(j_0) = j(\tau) \quad \text{and} \quad \sigma(f_0) = f(\tau).$$

By (2.1), the number $\sigma(j_0)$ is a singular modulus of the j -invariant and, therefore, τ is a quadratic imaginary number. It follows that $\sigma(f_0)$ is a singular modulus of f . This completes the proof of item (i). To prove item (ii), note that by (2.2) and the fact that f_0 is in an extension of $K(j_0)$ of degree at most δ_Y we have

$$\# O_{\mathbb{Q}}(f_0) \leq \delta_Y [K(j_0) : \mathbb{Q}] \leq \delta_Y [K : \mathbb{Q}] \cdot \# O_{\mathbb{Q}}(j_0).$$

On the other hand, the polynomial $\Phi(X, f_0)$ is nonzero because Φ is irreducible over \mathbb{C} and f is nonconstant. Since j_0 is a root of $\Phi(X, f_0)$, it is in an extension of $K(f_0)$ of degree at most δ_X , and we have

$$\# O_{\mathbb{Q}}(j_0) \leq \delta_X [K(f_0) : \mathbb{Q}] = \delta_X [K : \mathbb{Q}] \cdot \# O_K(f_0).$$

This completes the proof of item (ii) with $C_0 = [K : \mathbb{Q}] \max\{\delta_X, \delta_Y\}$.

Item (iii) is a direct consequence of (2.1), (2.2), item (ii) and of the following estimate: for every $\varepsilon > 0$ there is $C > 0$ such that for every discriminant D , we have

$$C^{-1} |D|^{1/2-\varepsilon} \leq h(D) \leq C |D|^{1/2+\varepsilon}.$$

In the case where D is fundamental this is Siegel's estimate [Sie35, (1)]. To deduce the general case from the fundamental case; see, e.g., [Lan87, Chapter 8, § 1, Theorem 7] or [HMR21, (5.12) and Lemma 5.12].

To prove item (iv), let C_0 and C_1 be the constants given by items (ii) and (iii), respectively. Moreover, for each singular modulus \mathfrak{f} of f choose a quadratic imaginary number τ in \mathbb{H} such that $f(\tau) = \mathfrak{f}$, and put $j(\mathfrak{f}) := j(\tau)$. For every singular modulus \mathfrak{j} of the j -invariant there are at most δ_Y singular moduli \mathfrak{f} of f such that $j(\mathfrak{f}) = \mathfrak{j}$. Thus, by items (ii) and (iii) the left-hand side

of (2.3) is bounded from above by

$$\delta_Y \cdot \#\{\text{singular modulus } j \text{ of the } j\text{-invariant such that } \# O_{\mathbb{Q}}(j) \leq C_0 R \text{ and } |D_j| \leq C_1 R^{2+\varepsilon}\} \leq \delta_Y C_0 C_1 R^{3+\varepsilon}.$$

This proves item (iv), and completes the proof of the proposition. □

2.3 Cuspidal and omitted values of modular functions

Let f be a nonconstant modular function. A complex number α is a *value of f* if there is τ in \mathbb{H} such that $f(\tau) = \alpha$, and it is an *omitted value of f* if it is not a value of f .

For a modular function f , the following proposition gives a characterization of the cuspidal and omitted values of f . It shows, in particular, that every omitted value is cuspidal; see also Remark 2.6. Moreover, in Proposition 2.7 below we show that in the case where f is defined over $\overline{\mathbb{Q}}$, every cuspidal value of f is well approximated in \mathbb{C} by the singular moduli of f and that for every prime number p , every omitted value of f is badly approximable in \mathbb{C}_p by the singular moduli of f .

PROPOSITION 2.4. *Let f be a nonconstant modular function and let $\Phi(X, Y)$ be a modular polynomial of f . Then, for every complex number α the polynomial $\Phi(X, \alpha)$ is nonzero and the following properties hold.*

- (i) *The number α is an omitted value of f if and only if the polynomial $\Phi(X, \alpha)$ is constant.*
- (ii) *The number α is a cuspidal value of f if and only if the degree of the polynomial $\Phi(X, \alpha)$ is strictly smaller than the degree of $\Phi(X, Y)$ in X .*

In particular, every omitted value is cuspidal. Furthermore, if f is defined over a subfield K of \mathbb{C} , then every cuspidal value of f is in the algebraic closure of K inside \mathbb{C} .

The following corollary is an immediate consequence of this proposition. Note that a modular function f is holomorphic if and only if 0 is an omitted value of $1/f$, and f is a modular unit if and only if 0 is an omitted value of f and of $1/f$.

COROLLARY 2.5. *Let f and Φ be as in Proposition 2.4. If we consider $\Phi(X, Y)$ as a polynomial in Y with coefficients in $\mathbb{C}[X]$, then the following properties hold.*

- (i) *The modular function f is holomorphic if and only if the leading coefficient of $\Phi(X, Y)$ does not depend on X . In particular, for every holomorphic modular function f defined over $\overline{\mathbb{Q}}$, there is a finite set of prime numbers S such that every singular modulus of f is an S -integer.*
- (ii) *The modular function f is a modular unit if and only if neither the constant nor the leading coefficients of $\Phi(X, Y)$ depend on X . In particular, for every modular unit f defined over $\overline{\mathbb{Q}}$, there is a finite set of prime numbers S such that every singular modulus of f is an S -unit.*

Proof of Proposition 2.4. If the polynomial $\Phi(X, \alpha)$ were zero, then $\Phi(X, Y)$ would be divisible by $Y - \alpha$. This is impossible since $\Phi(X, Y)$ is irreducible in $\mathbb{C}[X, Y]$ and it depends on both variables (Proposition 2.1). This proves that $\Phi(X, \alpha)$ is nonzero.

To prove item (i), let α be a complex number such that $\Phi(X, \alpha)$ is nonconstant and let β be a root of this polynomial. Then, by Proposition 2.1(i) there is τ in \mathbb{H} such that $j(\tau) = \beta$ and $f(\tau) = \alpha$. In particular, α is a value of f and, therefore, it is not an omitted value of f . To prove the reverse implication, let τ in \mathbb{H} be such that $f(\tau)$ is finite. Then, the number $j(\tau)$ is a zero of the polynomial $\Phi(X, f(\tau))$. Since $\Phi(X, f(\tau))$ is nonzero, it follows that it is nonconstant. This completes the proof of item (i).

To prove item (ii), denote by d the degree of $\Phi(X, Y)$ in X and let $P(Y)$ be the coefficient of X^d in $\Phi(X, Y)$, seen as a polynomial in X with coefficients in $\mathbb{C}[Y]$. Furthermore, put

$$\Delta(X, Y) := P(Y)X^d - \Phi(X, Y) \tag{2.4}$$

and note that the degree in X of this polynomial is strictly less than d . Let Γ be the stabilizer of f in $\mathrm{SL}(2, \mathbb{R})$ and denote by f_0 the meromorphic function defined on $X(\Gamma)$ induced by f . Suppose that α is a cuspidal value of f . That is, α is a value that f_0 takes at a point in $\Gamma \backslash \mathbb{P}^1(\mathbb{Q})$. Since $\mathrm{SL}(2, \mathbb{Z})$ acts transitively on $\mathbb{P}^1(\mathbb{Q})$, there is γ in $\mathrm{SL}(2, \mathbb{Z})$ such that $f \circ \gamma(\tau) \rightarrow \alpha$ as $\Im(\tau) \rightarrow \infty$; see, e.g., [Shi71, Proposition 1.30]. Combined with (2.4), this implies

$$|P(\alpha)| = \lim_{\Im(\tau) \rightarrow \infty} |P((f \circ \gamma)(\tau))| = \lim_{\Im(\tau) \rightarrow \infty} \frac{|\Delta(j(\tau), (f \circ \gamma)(\tau))|}{|j(\tau)|^d} = 0.$$

This proves $P(\alpha) = 0$ and, therefore, that the degree of $\Phi(X, \alpha)$ is strictly less than d . To prove the reverse implication, suppose that α is a non-cuspidal value of f and let A be a finite subset of \mathbb{H} such that $f_0^{-1}(\alpha) = \Gamma \backslash (\Gamma \cdot A)$. Let $r > 0$ be sufficiently small so that there is a compact neighborhood N of A in \mathbb{H} such that

$$f_0^{-1}(\overline{\mathbf{D}_\infty(\alpha, r)}) = \Gamma \backslash (\Gamma \cdot N).$$

Reducing r if necessary, suppose that for every α' in $B(\alpha, r) \setminus \{\alpha\}$ we have $P(\alpha') \neq 0$ and let $(\alpha_i)_{i=1}^\infty$ be a sequence in $B(\alpha, r) \setminus \{\alpha\}$ converging to α . Then, for every i the polynomial $\Phi(X, \alpha_i)$ is of degree d and, therefore, by Proposition 2.1(i) there are $\tau_i^{(1)}, \dots, \tau_i^{(d)}$ in N such that

$$\Phi(X, \alpha_i) = P(\alpha_i) \prod_{\ell=1}^d (X - j(\tau_i^{(\ell)})). \tag{2.5}$$

Taking a subsequence if necessary, suppose that for every ℓ in $\{1, \dots, d\}$ the sequence $(\tau_i^{(\ell)})_{i=1}^\infty$ converges to an element τ_ℓ of N . Letting $i \rightarrow \infty$ in (2.5), we obtain

$$\Phi(X, \alpha) = P(\alpha) \prod_{\ell=1}^d (X - j(\tau^{(\ell)})).$$

Since $\Phi(X, \alpha)$ is nonzero, it follows that $P(\alpha)$ is nonzero and, therefore, that the degree of $\Phi(X, \alpha)$ is d . This completes the proof of item (ii).

To prove the remaining assertions, note that by combining items (i) and (ii) we obtain that every omitted value is cuspidal. On the other hand, by item (ii) the cuspidal values of f are precisely the zeros of $P(Y)$. In particular, there are at most finitely many cuspidal values of f . If f is defined over a subfield K of \mathbb{C} , then we can assume that the polynomial $\Phi(X, Y)$ is in $K[X, Y]$. This implies that $P(Y)$ is in $K[Y]$ and, therefore, that all of the cuspidal values of f are in the algebraic closure of K inside \mathbb{C} . □

Remark 2.6. The modular function

$$g := \frac{j}{j^2 - 1}$$

provides an example of a cuspidal value that is not omitted. In fact, this function is invariant under $\mathrm{SL}(2, \mathbb{Z})$ and the meromorphic function g_0 on $X(\mathrm{SL}(2, \mathbb{Z}))$ induced by g vanishes at the cusp $i\infty$. But 0 is not an omitted value of g , because $g((1 + \sqrt{3}i)/2) = 0$.

PROPOSITION 2.7. *For every nonconstant modular function f defined over $\overline{\mathbb{Q}}$, the following properties hold.*

- (i) Every cuspidal value of f is well approximated in \mathbb{C} by the singular moduli of f . In particular, every omitted value of f is well approximated in \mathbb{C} by the singular moduli of f .
- (ii) Let p be a prime number and let α be an omitted value of f . Then, there is $r > 0$ such that $\mathbf{D}_p(\alpha, r)$ contains no singular modulus of f . In particular, α is badly approximable in \mathbb{C}_p by the singular moduli of f .

The proof of this proposition is after the following lemma.

LEMMA 2.8. Let v be in $M_{\mathbb{Q}}$ and let $\Phi(X, Y)$ be an irreducible polynomial in $\mathbb{C}_v[X, Y]$ depending on both variables. Then, for every α in \mathbb{C}_v there are constants

$$C_3 > 1, \quad \theta > 0, \quad \eta > 0 \quad \text{and} \quad \eta' > 0$$

such that for every z in \mathbb{C}_v and every w in $\mathbb{C}_v \setminus \{\alpha\}$ sufficiently close to α and such that $\Phi(z, w) = 0$, exactly one of the following properties holds.

- (i) The polynomial $\Phi(X, \alpha)$ is nonconstant and, denoting by Z its finite set of zeros in \mathbb{C}_v , we have

$$\min\{|z - z_0|_v : z_0 \in Z\} < C_3|w - \alpha|_v^\theta. \tag{2.6}$$

- (ii) The degree of $\Phi(X, \alpha)$ is strictly smaller than that of $\Phi(X, Y)$ in X and we have

$$C_3^{-1}|w - \alpha|_v^{-\eta} < |z|_v < C_3|w - \alpha|_v^{-\eta'}. \tag{2.7}$$

Proof. Put $Q_0(X) := \Phi(X, \alpha)$ and note that our hypotheses that $\Phi(X, Y)$ is irreducible in $\mathbb{C}_v[X, Y]$ and that it depends on both variables, implies that $Q_0(X)$ is nonzero. Denote by ℓ_0 the degree of $Q_0(X)$ and let $R_0 > 1$ and M_0 in $]0, 1[$ be constants so that for every z in \mathbb{C}_v satisfying $|z|_v \geq R_0$, we have

$$M_0|z|_v^{\ell_0} \leq |Q_0(z)|_v \leq M_0^{-1}|z|_v^{\ell_0}. \tag{2.8}$$

Reducing M_0 if necessary, suppose that in the case where $Q_0(X)$ is constant we have

$$|Q_0(0)|_v \geq M_0, \tag{2.9}$$

and that in the case where $Q_0(X)$ is nonconstant for every z in \mathbb{C}_v we have

$$|Q_0(z)|_v \geq M_0 \min\{|z - z_0|_v : z_0 \in Z\}^{\ell_0}. \tag{2.10}$$

Note that $Y - \alpha$ divides $\Phi(X, Y) - Q_0(X)$. Let m_0 in $\mathbb{Z}_{>0}$ be the largest integer such that $(Y - \alpha)^{m_0}$ divides $\Phi(X, Y) - Q_0(X)$, and let $\Psi(X, Y)$ be the polynomial in $\mathbb{C}_v[X, Y]$ such that

$$\Phi(X, Y) - Q_0(X) = (Y - \alpha)^{m_0} \Psi(X, Y). \tag{2.11}$$

Denote by δ the degree of $\Psi(X, Y)$ in X . Regarding $\Psi(X, Y)$ as a polynomial in X with coefficients in $\mathbb{C}_v[Y]$, for each i in $\{0, \dots, \delta\}$ let $P_i(Y)$ be the coefficient of X^i in $\Psi(X, Y)$. Furthermore, denote by m_1 the order of $P_\delta(Y)$ at α . Then, there is a constant $M_1 > 1$ such that for every w in $\mathbb{C}_v \setminus \{\alpha\}$ that is sufficiently close to α , we have

$$|P_\delta(w)|_v > M_1^{-1}|w - \alpha|_v^{m_1}, \tag{2.12}$$

and such that for every i in $\{0, \dots, \delta\}$ we have $|P_i(w)|_v \leq M_1$. Thus, for every z in \mathbb{C}_v such that $\Phi(z, w) = 0$, we have

$$|\Psi(z, w)|_v \leq (\delta + 1)M_1 \max\{1, |z|_v\}^\delta \tag{2.13}$$

and

$$|\Psi(z, w) - P_\delta(w)z^\delta|_v \leq \delta M_1 \max\{1, |z|_v\}^{\delta-1}. \tag{2.14}$$

To prove the desired assertion, put

$$M_2 := \frac{M_0}{(\delta + 1)M_1},$$

and let w in $\mathbb{C}_v \setminus \{\alpha\}$ be sufficiently close to α so that (2.12), (2.13) and (2.14) hold and so that

$$|w - \alpha|_v^{m_0} < M_2 R_0^{-\delta}. \tag{2.15}$$

Furthermore, let z in \mathbb{C}_v be such that $\Phi(z, w) = 0$.

Case 1. $|z|_v < R_0$. If $Q_0(X)$ were constant, then by (2.9), (2.11) and (2.13) we would have

$$|w - \alpha|_v^{m_0} = \left| \frac{Q_0(z)}{\Psi(z, w)} \right|_v > M_2 R_0^{-\delta},$$

which contradicts (2.15) and proves that $Q_0(X)$ is nonconstant. Denoting by Z the nonempty set of zeros of $Q_0(X)$ in \mathbb{C}_v , by (2.10), (2.11) and (2.13) we have

$$|w - \alpha|_v^{m_0} = \left| \frac{Q_0(z)}{\Psi(z, w)} \right|_v > M_2 R_0^{-\delta} \min\{|z - z_0|_v : z_0 \in Z\}^{\ell_0}.$$

This proves (2.6) with $C_3 = M_2^{-(1/\ell_0)} R_0^{\delta/\ell_0}$ and $\theta = m_0/\ell_0$ and completes the proof that property (i) holds.

Case 2. $|z|_v \geq R_0$. By (2.8), (2.11) and (2.13), in this case we have

$$M_0 |z|_v^{\ell_0} \cdot |w - \alpha|_v^{-m_0} \leq |Q_0(z)| \cdot |w - \alpha|_v^{-m_0} = |\Psi(z, w)|_v \leq (\delta + 1)M_1 |z|_v^{\delta}. \tag{2.16}$$

If we had $\ell_0 \geq \delta$, then we would obtain $|w - \alpha|_v^{m_0} \geq M_2$. This contradicts (2.15) and proves that the degree ℓ_0 of $Q_0(X)$ is strictly less than the degree δ of $\Phi(X, Y)$ in X . Together with (2.16), this implies the first inequality in (2.7) with $C_3 = M_2^{1/(\delta-\ell_0)}$ and $\eta = m_0/(\delta - \ell_0)$. To prove the second inequality in (2.7), suppose

$$|z|_v \geq 2\delta M_1^2 |w - \alpha|_v^{-m_1}.$$

Then, by (2.12) and (2.14) we have

$$|P_\delta(w)z^\delta|_v \geq 2\delta M_1^2 |w - \alpha|_v^{-m_1} |P_\delta(w)z^{\delta-1}|_v > 2\delta M_1 |z|_v^{\delta-1} \geq 2|\Psi(z, w) - P_\delta(w)z^\delta|_v.$$

Together with the triangle inequality, (2.8), (2.11) and (2.12), this implies

$$M_0^{-1} |z|_v^{\ell_0} \cdot |w - \alpha|_v^{-m_0} \geq |Q_0(z)| \cdot |w - \alpha|_v^{-m_0} = |\Psi(z, w)|_v > \frac{1}{2} |P_\delta(w)z^\delta|_v \frac{1}{2} M_1^{-1} |w - \alpha|_v^{m_1} |z|_v^{\delta}.$$

Rearranging, we obtain the second inequality in (2.7) with

$$C_3 = \max\{2\delta M_1^2, (2M_0^{-1}M_1)^{1/(\delta-\ell_0)}\} \quad \text{and} \quad \eta' = \max\left\{m_1, \frac{m_0 + m_1}{\delta - \ell_0}\right\}.$$

This completes the proof that property (ii) holds.

Finally, note that for w in $\mathbb{C}_v \setminus \{\alpha\}$ that is sufficiently close to α , the inequality (2.6) and the first inequality in (2.7) cannot hold at the same time. This proves that properties (i) and (ii) cannot hold simultaneously and completes the proof of the lemma. \square

Proof of Proposition 2.7. Let $\Phi(X, Y)$ be a modular polynomial of f in $\overline{\mathbb{Q}}[X, Y]$ and note that for every v in $M_{\mathbb{Q}}$ the polynomial $\Phi(X, Y)$ is irreducible in $\mathbb{C}_v[X, Y]$.

To prove item (i), let α be a cuspidal value of f and let C_3 and η' be the constants given by Lemma 2.8 with $v = \infty$. That is, if we denote by Γ the stabilizer of f in $SL(2, \mathbb{R})$, then α is a value that f takes at a point in $\Gamma \backslash \mathbb{P}^1(\mathbb{Q})$. Since $SL(2, \mathbb{Z})$ acts transitively on $\mathbb{P}^1(\mathbb{Q})$, there is γ in $SL(2, \mathbb{Z})$ such that $f \circ \gamma(\tau) \rightarrow \alpha$ as $\Im(\tau) \rightarrow \infty$; see, e.g., [Shi71, Proposition 1.30]. Let $C > 0$ be a constant such that for every τ in \mathbb{H} such that $\Im(\tau)$ is sufficiently large, we have

$$|j(\tau)| \geq C \exp(2\pi\Im(\tau)); \tag{2.17}$$

see, e.g., [Lan87, Chapter 4, § 1]. Given a prime number p' satisfying $p' \equiv 1 \pmod{4}$, put

$$\tau_{p'} := i\sqrt{p'}, \quad j(p') := j(\tau_{p'}) \quad \text{and} \quad f(p') := f \circ \gamma(\tau_{p'}),$$

and note that $j(p')$ is a singular modulus of the j -invariant satisfying $D_{j(p')} = -4p'$ and that $f(p')$ is a singular modulus of f . If p' is sufficiently large, then by (2.17) with $\tau = \tau_{p'}$ property (i) in Lemma 2.8 cannot be satisfied with $z = j(p')$ and $w = f(p')$. Thus, property (ii) holds and we have

$$-\log |f(p') - \alpha| > \frac{1}{\eta'} \log |j(p')| - \frac{1}{\eta'} \log C_3 \geq \frac{\pi}{\eta'} \sqrt{|D_{j(p')}|} - \frac{1}{\eta'} \log \frac{C_3}{C}.$$

In view of Proposition 2.3(iii), this implies that α is well approximated in \mathbb{C} by the singular moduli of f . The second assertion of item (i) follows from the first and from the fact that every omitted value is cuspidal (Proposition 2.4).

To prove item (ii), let C_3 and η be the constants given by Lemma 2.8 with $v = p$ and put $r := C_3^{-1/\eta}$. By Proposition 2.4(ii), our hypothesis that α is an omitted value of f implies that the polynomial $\Phi(X, \alpha)$ is constant. Thus, if there were a quadratic imaginary number τ in \mathbb{H} such that $f(\tau)$ is sufficiently close to α in \mathbb{C}_p , then $f(\tau)$ would be in $\mathbf{D}_p(\alpha, r)$ and by Lemma 2.8 the singular modulus $j(\tau)$ of the j -invariant would satisfy

$$|j(\tau)|_p > C_3^{-1} |f(\tau) - \alpha|_p^{-\eta} > 1.$$

This is absurd, since $j(\tau)$ is an algebraic integer. This completes the proof of item (ii) and of the proposition. □

3. p -Adic limits of CM points

The goal of this section is to prove Theorem B. The main ingredient is Theorem 3.1 below. Together with [HMR21, Theorems A and B], which are summarized in Theorem 3.7 in § 3.2, Theorem 3.1 implies Theorem B in the case of the j -invariant as a direct consequence. The general case is deduced from this special case in § 3.3.

Throughout this section, fix a prime number p and let $(\mathbb{C}_p, |\cdot|_p)$ be as in the introduction. Denote by $Y(\mathbb{C}_p)$ the coarse moduli space of elliptic curves over \mathbb{C}_p . We consider $Y(\mathbb{C}_p)$ as a subspace of the Berkovich affine line $\mathbb{A}_{\text{Berk}}^1$ over \mathbb{C}_p , using the j -invariant to identify $Y(\mathbb{C}_p)$ with the subspace \mathbb{C}_p of $\mathbb{A}_{\text{Berk}}^1$. We endow the space of Borel measures on $\mathbb{A}_{\text{Berk}}^1$ with the weak topology with respect to the space of bounded and continuous real functions. Denote by x_{can} the ‘Gauss’ or ‘canonical’ point of $\mathbb{A}_{\text{Berk}}^1$. For x in $\mathbb{A}_{\text{Berk}}^1$ denote by δ_x the Dirac measure at x . An *atom* of a Borel measure ν on $\mathbb{A}_{\text{Berk}}^1$ is a point x in $\mathbb{A}_{\text{Berk}}^1$ such that $\nu(\{x\}) > 0$. A measure is *nonatomic* if it has no atoms.

The endomorphism ring of an elliptic curve over \mathbb{C}_p only depends on the corresponding class E in $Y(\mathbb{C}_p)$ of the elliptic curve. It is isomorphic to \mathbb{Z} or to an order in a quadratic imaginary

extension of \mathbb{Q} . In the latter case E is a *CM point* and its *discriminant* is the discriminant of its endomorphism ring. For every discriminant D , the set

$$\Lambda_D := \{E \in Y(\mathbb{C}_p) : \text{CM point of discriminant } D\}$$

is finite and nonempty. Denote by $\bar{\delta}_{D,p}$ the Borel probability measure on $Y(\mathbb{C}_p)$, defined by

$$\bar{\delta}_{D,p} := \frac{1}{\#\Lambda_D} \sum_{E \in \Lambda_D} \delta_E.$$

In contrast to the complex case, as the discriminant D tends to $-\infty$ the measure $\bar{\delta}_{D,p}$ does not converge in the weak topology. In fact, there are infinitely many different accumulation measures [HMR21, Corollary 1.1].

THEOREM 3.1. *Let p be a prime number. Then every accumulation measure of*

$$\{\bar{\delta}_{D,p} : D \text{ discriminant}\} \tag{3.1}$$

in the weak topology that is different from $\delta_{x_{\text{can}}}$ is nonatomic. In particular, no accumulation measure of (3.1) in the weak topology has an atom in $Y(\mathbb{C}_p)$.

One of the main ingredients in the proof of this result is the description of all accumulation measures of (3.1) given in the companion papers [HMR20, HMR21]. We also use an analogous description for Hecke orbits given in [HMR20, HMR21]. We first establish a result analogous to Theorem 3.1 for Hecke orbits (Theorem 3.2) in § 3.1, and in § 3.2 we deduce Theorem 3.1 from this result.

Denote by $\bar{\mathbb{Q}}_p$ the algebraic closure of \mathbb{Q}_p inside \mathbb{C}_p , and by \mathcal{O}_p and $\mathcal{O}_{\bar{\mathbb{Q}}_p}$ the ring of integers of \mathbb{C}_p and $\bar{\mathbb{Q}}_p$, respectively. For E in $Y(\mathbb{C}_p)$ represented by a Weierstrass equation with coefficients in $\mathcal{O}_{\bar{\mathbb{Q}}_p}$ having smooth reduction, denote by \mathcal{F}_E the formal group of E and by $\text{End}(\mathcal{F}_E)$ the ring of endomorphisms of \mathcal{F}_E that are defined over the ring of integers of a finite extension of \mathbb{Q}_p . Then $\text{End}(\mathcal{F}_E)$ is either isomorphic to \mathbb{Z}_p , or to a p -adic quadratic order; see, e.g., [Frö68, Chapter IV, § 1, Theorem 1(iii)]. In the latter case, E is said to have *formal complex multiplication* or to be a *formal CM point*.

An elliptic curve class E in $Y(\mathbb{C}_p)$ has *supersingular reduction*, if there is a representative elliptic curve over \mathcal{O}_p whose reduction is smooth and supersingular. Denote by $Y_{\text{sup}}(\mathbb{C}_p)$ the set of all elliptic curve classes in $Y(\mathbb{C}_p)$ with supersingular reduction.

3.1 On the limit measures of Hecke orbits

The goal of this section is to prove Theorem 3.2 below, which is the main ingredient in the proof of Theorem 3.1. To state it, we introduce some notation.

A *divisor on $Y(\mathbb{C}_p)$* is an element of the free abelian group

$$\text{Div}(Y(\mathbb{C}_p)) := \bigoplus_{E \in Y(\mathbb{C}_p)} \mathbb{Z}E.$$

For a divisor $\mathcal{D} = \sum_{E \in Y(\mathbb{C}_p)} n_E E$ in $\text{Div}(Y(\mathbb{C}_p))$, the *degree* and *support* of \mathcal{D} are defined by

$$\deg(\mathcal{D}) := \sum_{E \in Y(\mathbb{C}_p)} n_E \quad \text{and} \quad \text{supp}(\mathcal{D}) := \{E \in Y(\mathbb{C}_p) : n_E \neq 0\},$$

respectively. If, in addition, $\text{deg}(\mathcal{D}) \geq 1$ and for every E in $Y(\mathbb{C}_p)$ we have $n_E \geq 0$, then

$$\bar{\delta}_{\mathcal{D},p} := \frac{1}{\text{deg}(\mathcal{D})} \sum_{E \in Y(\mathbb{C}_p)} n_E \delta_E$$

is a Borel probability measure on $Y(\mathbb{C}_p)$.

For n in $\mathbb{Z}_{>0}$, the n th Hecke correspondence is the linear map

$$T_n: \text{Div}(Y(\mathbb{C}_p)) \rightarrow \text{Div}(Y(\mathbb{C}_p))$$

defined for E in $Y(\mathbb{C}_p)$ by

$$T_n(E) := \sum_{C \leq E \text{ of order } n} E/C,$$

where the sum runs over all subgroups C of E of order n . For background on Hecke correspondences, see [Shi71, §§ 7.2 and 7.3] for the general theory, or the survey [DI95, Part II].

THEOREM 3.2. *For each E in $Y(\mathbb{C}_p)$, every accumulation measure of $(\bar{\delta}_{T_n(E),p})_{n=1}^\infty$ in the weak topology that is different from $\delta_{x_{\text{can}}}$, is nonatomic. In particular, no accumulation measure of $(\bar{\delta}_{T_n(E),p})_{n=1}^\infty$ in the weak topology has an atom in $Y(\mathbb{C}_p)$.*

To prove Theorem 3.2, we first recall some results in [HMR21]. For E in $Y_{\text{supp}}(\mathbb{C}_p)$, define a subgroup \mathbf{Nr}_E of \mathbb{Z}_p^\times as follows. If E is not a formal CM point, then $\mathbf{Nr}_E := (\mathbb{Z}_p^\times)^2$. In the case where E is a formal CM point, denote by $\text{Aut}(\mathcal{F}_E)$ the group of isomorphisms of \mathcal{F}_E defined over $\mathcal{O}_{\mathbb{Q}_p}$, and by nr the norm map of the field of fractions of $\text{End}(\mathcal{F}_E)$ to \mathbb{Q}_p . Then,

$$\mathbf{Nr}_E := \{\text{nr}(\varphi) : \varphi \in \text{Aut}(\mathcal{F}_E)\}.$$

In all the cases \mathbf{Nr}_E is a multiplicative subgroup of \mathbb{Z}_p^\times containing $(\mathbb{Z}_p^\times)^2$. In particular, the index of \mathbf{Nr}_E in \mathbb{Z}_p^\times is at most two if p is odd, and at most four if $p = 2$.

For a coset \mathfrak{N} in $\mathbb{Q}_p^\times/\mathbf{Nr}_E$ contained in \mathbb{Z}_p , the partial Hecke orbit of E along \mathfrak{N} is

$$\text{Orb}_{\mathfrak{N}}(E) := \bigcup_{n \in \mathfrak{N} \cap \mathbb{Z}_{>0}} \text{supp}(T_n(E)).$$

In the following theorem we use the action of Hecke correspondences on compactly supported measures; see, e.g., [HMR21, § 2.8]. For n in $\mathbb{Z}_{>0}$, put

$$\sigma_1(n) := \sum_{d \geq 1, d|n} d.$$

THEOREM 3.3 [HMR21, Theorem C and Corollary 6.1]. *For every E in $Y_{\text{supp}}(\mathbb{C}_p)$ and all cosets \mathfrak{N} and \mathfrak{N}' in $\mathbb{Q}_p^\times/\mathbf{Nr}_E$ contained in \mathbb{Z}_p , the following properties hold.*

- (i) *The closure $\overline{\text{Orb}_{\mathfrak{N}}(E)}$ of $\text{Orb}_{\mathfrak{N}}(E)$ in $Y_{\text{supp}}(\mathbb{C}_p)$ is compact. Moreover, there is a Borel probability measure $\mu_{\mathfrak{N}}^E$ on $Y(\mathbb{C}_p)$ whose support is equal to $\overline{\text{Orb}_{\mathfrak{N}}(E)}$, and such that for every sequence $(n_j)_{j=1}^\infty$ in $\mathfrak{N} \cap \mathbb{Z}_{>0}$ tending to ∞ , we have the weak convergence of measures*

$$\bar{\delta}_{T_{n_j}(E),p} \rightarrow \mu_{\mathfrak{N}}^E \quad \text{as } j \rightarrow \infty.$$

- (ii) *For every E' in $\overline{\text{Orb}_{\mathbf{Nr}_E}(E)}$ and every n in $\mathfrak{N} \cap \mathbb{Z}_{>0}$, we have*

$$T_n(\overline{\text{Orb}_{\mathfrak{N}'}(E')}) = \overline{\text{Orb}_{\mathfrak{N}'}(E')} \quad \text{and} \quad \frac{1}{\sigma_1(n)} (T_n)_* \mu_{\mathfrak{N}'}^{E'} = \mu_{\mathfrak{N}'}^E.$$

The following corollary is an immediate consequence of Theorems 3.2 and 3.3 and [HMR20, Theorem C].

COROLLARY 3.4. *For every E in $Y(\mathbb{C}_p)$, α in \mathbb{C}_p , and $\varepsilon > 0$, there exists $r > 0$ such that the following set is finite:*

$$\{n \in \mathbb{Z}_{>0} : \deg(T_n(E)|_{\mathbf{D}_p(\alpha,r)}) \geq \varepsilon\sigma_1(n)\}.$$

Previously, Charles showed that the set above with $\mathbb{Z}_{>0}$ replaced by $\mathbb{Z}_{>0} \setminus p\mathbb{Z}_{>0}$ has zero density [Cha18, Proposition 3.2].

The proof of Theorem 3.2 is given after the following lemma.

LEMMA 3.5. *Let E_0 be in $Y(\mathbb{C}_p)$. If for distinct prime numbers q and q' we put*

$$I := \text{supp}(T_q(\widehat{E}_0)) \cap \text{supp}(T_{q'}(\widehat{E}_0)),$$

then we have $\deg(T_q(E_0)|_I) \leq 24$.

Proof. Given an elliptic curve \widehat{E} over \mathbb{C}_p , denote by $\text{End}(\widehat{E})$ and $\text{Aut}(\widehat{E})$ the set of all endomorphisms and the set of all automorphisms of \widehat{E} , respectively. Furthermore, for every elliptic curve \widehat{E}' over \mathbb{C}_p and every m in $\mathbb{Z}_{>0}$, denote by $\text{Hom}_m(\widehat{E}, \widehat{E}')$ the set of all isogenies from \widehat{E} to \widehat{E}' of degree m .

Choose an elliptic curve \widehat{E}_0 representing E_0 and for each E in I choose an elliptic curve \widehat{E} representing E and an isogeny ϕ_E in $\text{Hom}_{q'}(\widehat{E}, \widehat{E}_0)$. Let ϕ be in $\text{Hom}_q(\widehat{E}_0, \widehat{E})$ and set $\psi := \phi_E \circ \phi$. The isogeny ψ determines both E in I and ϕ . Indeed, suppose that there are E' in I and ϕ' in $\text{Hom}_q(\widehat{E}_0, \widehat{E}')$ with $\phi_{E'} \circ \phi' = \psi$. The group $\text{Ker}(\psi)$ has qq' elements, so it has a unique subgroup of order q . Since $\text{Ker}(\phi)$ and $\text{Ker}(\phi')$ are two such subgroups, we have $\text{Ker}(\phi) = \text{Ker}(\phi')$. Then $E = E'$ by [Sil09, Chapter III, Proposition 4.12], and from the equality $\phi_E \circ \phi = \phi_{E'} \circ \phi'$ we deduce $\phi = \phi'$. We thus have

$$\begin{aligned} \deg(T_q(E_0)|_I) &= \sum_{E \in I} \# \text{Hom}_q(\widehat{E}_0, \widehat{E}) / \# \text{Aut}(\widehat{E}) \\ &\leq \sum_{E \in I} \# \text{Hom}_q(\widehat{E}_0, \widehat{E}) \\ &\leq \sum_{E \in I} \#\{\phi_E \circ \phi : \phi \in \text{Hom}_q(\widehat{E}_0, \widehat{E})\} \\ &\leq \#\{\psi \in \text{End}(\widehat{E}_0) : \deg(\psi) = qq'\}. \end{aligned} \tag{3.2}$$

If E_0 is not a CM point, then this last number is equal to zero and the lemma follows in this case. Suppose E_0 is a CM point, so the field of fractions K of $\text{End}(\widehat{E}_0)$ is a quadratic imaginary extension of \mathbb{Q} . Denote by \mathcal{O}_K the ring of integers of K . Since each of the ideals $q\mathcal{O}_K$ and $q'\mathcal{O}_K$ is either prime or a product of two conjugate prime ideals, there are at most four ideals of \mathcal{O}_K of norm qq' . We thus have

$$\#\{\psi \in \text{End}(\widehat{E}_0) : \deg(\psi) = qq'\} \leq \#\{x \in \mathcal{O}_K : x\bar{x} = qq'\} \leq 4\#\mathcal{O}_K^\times \leq 24.$$

Together with (3.2) this completes the proof of the lemma. □

Proof of Theorem 3.2. By [HMR20, Theorem C], it is sufficient to assume that E is in $Y_{\text{supp}}(\mathbb{C}_p)$. Moreover, using Theorem 3.3(i) and [HMR20, Theorem C] again, it is sufficient to prove that for every coset \mathfrak{N} in $\mathbb{Q}_p^\times / \mathbf{Nr}_E$ contained in \mathbb{Z}_p the measure $\mu_{\mathfrak{N}}^E$ has no atom in $\overline{\text{Orb}_{\mathfrak{N}}(E)}$.

Fix E_0 in $\overline{\text{Orb}_{\mathfrak{N}}(E)}$ and let $N \geq 1$ be a given integer. Choose a set P of $2N$ prime numbers that are contained in $(\mathbb{Z}_p^\times)^2$ and that are larger than $100N$. Note that every q in P is a p -adic

square and that by Theorem 3.3(ii) we have

$$\frac{1}{\sigma_1(q)}(T_q)_*\mu_{\mathfrak{N}}^E = \mu_{\mathfrak{N}}^E.$$

Moreover, for all distinct q and q' in P denote by $I(q, q')$ the set I in Lemma 3.5 and put

$$S_q := \text{supp}(T_q(E_0)) \setminus \bigcup_{q' \in P, q' \neq q} I(q, q').$$

Then by the inequality $q > 100N$ and Lemma 3.5, we have

$$\begin{aligned} \deg(T_q(E_0)|_{S_q}) &\geq \deg(T_q(E_0)) - \sum_{q' \in P, q' \neq q} \deg(T_q(E_0)|_{I(q, q')}) \\ &\geq q + 1 - 24(\#P - 1) \\ &\geq \frac{q + 1}{2}. \end{aligned} \tag{3.3}$$

On the other hand, note that the signed measure $\mu_{\mathfrak{N}}^E - \mu_{\mathfrak{N}}^E(\{E_0\})\delta_{E_0}$ is nonnegative, thus the same holds for $(T_q)_*(\mu_{\mathfrak{N}}^E - \mu_{\mathfrak{N}}^E(\{E_0\})\delta_{E_0})$. Combined with Theorem 3.3(ii), this implies that for every E' in $Y(\mathbb{C}_p)$ we have

$$\begin{aligned} \mu_{\mathfrak{N}}^E(\{E'\}) &= \left(\frac{1}{q + 1}(T_q)_*\mu_{\mathfrak{N}}^E\right)(\{E'\}) \\ &\geq \frac{\mu_{\mathfrak{N}}^E(\{E_0\})}{q + 1}((T_q)_*\delta_{E_0})(\{E'\}) \\ &= \frac{\mu_{\mathfrak{N}}^E(\{E_0\})}{q + 1} \deg(T_q(E_0)|_{\{E'\}}). \end{aligned}$$

Together with (3.3) this implies

$$1 = \mu_{\mathfrak{N}}^E(\overline{\text{Orb}_{\mathfrak{N}}(E)}) \geq \sum_{q \in P} \mu_{\mathfrak{N}}^E(S_q) \geq \sum_{q \in P} \frac{\mu_{\mathfrak{N}}^E(\{E_0\})}{q + 1} \deg(T_q(E_0)|_{S_q}) \geq N\mu_{\mathfrak{N}}^E(\{E_0\}).$$

Since N is arbitrary, this implies that E_0 is not an atom of $\mu_{\mathfrak{N}}^E$ and completes the proof of the theorem. □

Remark 3.6. A different strategy to prove Theorem 3.2 is to use that for every E in $Y_{\text{sup}}(\mathbb{C}_p)$ and every coset \mathfrak{N} in $\mathbb{Q}_p^\times/\mathbf{Nr}_E$ contained in \mathbb{Z}_p , the measure $\mu_{\mathfrak{N}}^E$ is the projection of a certain homogeneous measure under an analytic map of finite degree. Theorem 3.2 then follows from the fact that the partial Hecke orbit $\text{Orb}_{\mathfrak{N}}(E)$ is infinite.

3.2 On the limit measures of CM points

The goal of this section is to prove Theorem 3.1. The proof is based on Theorem 3.2 and on the description of all accumulation measures of (3.1) given in the companion papers [HMR20, HMR21]. We start recalling some results in the latter.

Recall from § 2.2 that a discriminant is the discriminant of an order in a quadratic imaginary extension of \mathbb{Q} . A *fundamental discriminant* is the discriminant of the ring of integers of a quadratic imaginary extension of \mathbb{Q} . For each discriminant D , there is a unique fundamental discriminant d and a unique integer $f \geq 1$ such that $D = df^2$. In this case, d and f are the *fundamental discriminant* and *conductor of D* , respectively. A discriminant is *prime*, if it is fundamental and divisible by only one prime number. If d is a prime discriminant divisible

by p , then

$$p \equiv -1 \pmod{4} \text{ and } d = -p, \text{ or } p = 2 \text{ and } d = -4 \text{ or } d = -8.$$

A p -adic quadratic order is a \mathbb{Z}_p -order in a quadratic extension of \mathbb{Q}_p , and a p -adic discriminant is a set formed by the discriminants of all \mathbb{Z}_p -bases of a p -adic quadratic order. Every p -adic discriminant is thus a coset in $\mathbb{Q}_p^\times/(\mathbb{Z}_p^\times)^2$ contained in \mathbb{Z}_p .

The p -adic discriminant of a formal CM point E , is the p -adic discriminant of the p -adic quadratic order $\text{End}(\mathcal{F}_E)$. Given a p -adic discriminant \mathfrak{D} , put

$$\Lambda_{\mathfrak{D}} := \{E \in Y(\mathbb{C}_p) : \text{formal CM point of } p\text{-adic discriminant } \mathfrak{D}\}.$$

THEOREM 3.7 [HMR21, Theorems A and B]. *For every p -adic discriminant \mathfrak{D} , the following properties hold.*

- (i) *The set $\Lambda_{\mathfrak{D}}$ is a compact subset of $Y(\mathbb{C}_p)$, and there is a Borel probability measure $\nu_{\mathfrak{D}}$ on $Y(\mathbb{C}_p)$ whose support is equal to $\Lambda_{\mathfrak{D}}$ and such that the following equidistribution property holds. Let $(D_n)_{n=1}^\infty$ be a sequence of discriminants in \mathfrak{D} tending to $-\infty$, such that for every n , the fundamental discriminant of D_n is either not divisible by p , or not a prime discriminant. Then we have the weak convergence of measures*

$$\bar{\delta}_{D_n, p} \rightarrow \nu_{\mathfrak{D}} \quad \text{as } n \rightarrow \infty.$$

- (ii) *Suppose that there is a prime discriminant d divisible by p and an integer $m \geq 0$ such that $D := dp^{2m}$ is in \mathfrak{D} . Then there are Borel probability measures $\nu_{\mathfrak{D}}^+$ and $\nu_{\mathfrak{D}}^-$ on $Y(\mathbb{C}_p)$ such that the following equidistribution property holds. For every sequence $(f_n)_{n=0}^\infty$ in $\mathbb{Z}_{>0}$ tending to ∞ such that for every n we have $(d/f_n) = 1$ (respectively, $(d/f_n) = -1$), we have the weak convergence of measures*

$$\bar{\delta}_{D(f_n)^2, p} \rightarrow \nu_{\mathfrak{D}}^+ \quad (\text{respectively, } \bar{\delta}_{D(f_n)^2, p} \rightarrow \nu_{\mathfrak{D}}^-) \quad \text{as } n \rightarrow \infty.$$

The proof of Theorem 3.1 is given after the following proposition, in which we gather further properties of the limit measures in Theorem 3.7. To state it, we introduce some notation.

A p -adic discriminant is *fundamental*, if it is the p -adic discriminant of the ring of integers of a quadratic extension of \mathbb{Q}_p . Let \mathfrak{d} be a fundamental p -adic discriminant. For Δ in \mathfrak{d} , the field $\mathbb{Q}_p(\sqrt{\Delta})$ depends only on \mathfrak{d} , but not on Δ . Denote it by $\mathbb{Q}_p(\sqrt{\mathfrak{d}})$. Choose a formal CM point $E_{\mathfrak{d}}$ such that $\text{End}(\mathcal{F}_E)$ is isomorphic to the ring of integers of $\mathbb{Q}_p(\sqrt{\mathfrak{d}})$, as follows. If \mathfrak{d} does not contain a prime discriminant that is divisible by p , then choose an arbitrary formal CM point $E_{\mathfrak{d}}$ in $\Lambda_{\mathfrak{d}}$. In the case where \mathfrak{d} contains a prime discriminant d that is divisible by p , then d is the unique fundamental discriminant in \mathfrak{d} with this property and we choose $E_{\mathfrak{d}}$ in Λ_d . Note that if $\mathbb{Q}_p(\sqrt{\mathfrak{d}})$ is unramified over \mathbb{Q}_p , then $\mathbf{Nr}_{E_{\mathfrak{d}}} = \mathbb{Z}_p^\times$, and that if $\mathbb{Q}_p(\sqrt{\mathfrak{d}})$ is ramified over \mathbb{Q}_p , then $\mathbf{Nr}_{E_{\mathfrak{d}}}$ is a subgroup of \mathbb{Z}_p^\times of index two; see, e.g., [HMR21, Lemma 2.3].

Denote by v_p Katz's valuation on $Y_{\text{supers}}(\mathbb{C}_p)$, as defined in [HMR20, § 4.1] and put

$$N_p := \left\{ E \in Y_{\text{supers}}(\mathbb{C}_p) : v_p(E) < \frac{p}{p+1} \right\}.$$

For E in N_p , denote by $H(E)$ the canonical subgroup of E (see [Kat73, Theorem 3.10.7]). The *canonical branch of the Hecke correspondence* T_p is the map $\mathbf{t} : N_p \rightarrow Y_{\text{supers}}(\mathbb{C}_p)$ defined by $\mathbf{t}(E) := E/H(E)$. The map \mathbf{t} is analytic in the sense that it is given by a finite sum of Laurent series, each of which converges on all of N_p ; see, e.g., [HMR20, Theorem B.1].

Given a fundamental p -adic discriminant \mathfrak{d} and an integer $m \geq 0$, define the affinoid

$$A_{\mathfrak{d}p^{2m}} := \begin{cases} v_p^{-1} \left(\frac{1}{2} \cdot p^{-m} \right) & \text{if } \mathbb{Q}_p(\sqrt{\mathfrak{d}}) \text{ is ramified over } \mathbb{Q}_p; \\ v_p^{-1}([1, \infty]) & \text{if } \mathbb{Q}_p(\sqrt{\mathfrak{d}}) \text{ is unramified over } \mathbb{Q}_p \text{ and } m = 0; \\ v_p^{-1} \left(\frac{p}{p+1} \cdot p^{-m} \right) & \text{if } \mathbb{Q}_p(\sqrt{\mathfrak{d}}) \text{ is unramified over } \mathbb{Q}_p \text{ and } m \geq 1. \end{cases}$$

In the following proposition we summarize some of the results from [HMR21, Proposition 7.1, (7.13) and §§ 7.2 and 7.3].

PROPOSITION 3.8. *For every fundamental p -adic discriminant \mathfrak{d} we have*

$$\nu_{\mathfrak{d}} = \begin{cases} \mu_{\mathbb{Z}_p^\times}^{E_{\mathfrak{d}}} & \text{if } \mathbb{Q}_p(\sqrt{\mathfrak{d}}) \text{ is unramified over } \mathbb{Q}_p; \\ \frac{1}{2}(\mu_{\mathbf{Nr}_{\mathfrak{d}}}^{E_{\mathfrak{d}}} + \mu_{\mathbb{Z}_p^\times \setminus \mathbf{Nr}_{\mathfrak{d}}}^{E_{\mathfrak{d}}}) & \text{if } \mathbb{Q}_p(\sqrt{\mathfrak{d}}) \text{ is ramified over } \mathbb{Q}_p, \end{cases} \tag{3.4}$$

and for every integer $m \geq 1$ we have

$$\nu_{\mathfrak{d}p^{2m}} = \begin{cases} \frac{1}{p^{m-1}(p+1)} (\mathfrak{t}^m|_{A_{\mathfrak{d}p^{2m}}})^* \nu_{\mathfrak{d}} & \text{if } \mathbb{Q}_p(\sqrt{\mathfrak{d}}) \text{ is unramified over } \mathbb{Q}_p; \\ \frac{1}{p^m} (\mathfrak{t}^m|_{A_{\mathfrak{d}p^{2m}}})^* \nu_{\mathfrak{d}} & \text{if } \mathbb{Q}_p(\sqrt{\mathfrak{d}}) \text{ is ramified over } \mathbb{Q}_p. \end{cases} \tag{3.5}$$

If, in addition, \mathfrak{d} contains a prime discriminant, then we also have

$$\nu_{\mathfrak{d}}^+ = \mu_{\mathbf{Nr}_{\mathfrak{d}}}^{E_{\mathfrak{d}}} \quad \text{and} \quad \nu_{\mathfrak{d}}^- = \mu_{\mathbb{Z}_p^\times \setminus \mathbf{Nr}_{\mathfrak{d}}}^{E_{\mathfrak{d}}}, \tag{3.6}$$

and (3.5) holds for $\nu_{\mathfrak{d}p^{2m}}^+$ (respectively, $\nu_{\mathfrak{d}p^{2m}}^-$), with $\nu_{\mathfrak{d}}$ replaced by $\nu_{\mathfrak{d}}^+$ (respectively, $\nu_{\mathfrak{d}}^-$).

Proof of Theorem 3.1. Let $(D_n)_{n=1}^\infty$ be a sequence of discriminants tending to $-\infty$ such that the sequence of measures $(\bar{\delta}_{D_n,p})_{n=1}^\infty$ converges weakly to a measure different from $\delta_{x_{\text{can}}}$. By [HMR20, Theorem A], there is a constant $c > 0$ such that for every n we have $|D_n|_p > c$ and $\Lambda_{D_n} \subseteq Y_{\text{sup}}(\mathbb{C}_p)$. This implies that $(D_n)_{n=1}^\infty$ is contained in a finite union of p -adic discriminants; see, e.g., [HMR21, Lemmas 2.1 and A.1]. Taking a subsequence if necessary, assume that $(D_n)_{n=1}^\infty$ is contained in a p -adic discriminant \mathfrak{D} . Let \mathfrak{d} be the fundamental p -adic discriminant and $m \geq 0$ the integer such that $\mathfrak{D} = \mathfrak{d}p^{2m}$; see, e.g., [HMR21, Lemma A.1(i)].

Passing to a subsequence if necessary, there are two cases.

Case 1. For every n the fundamental discriminant of D_n is either not divisible by p , or not a prime discriminant. In this case the sequence $(\bar{\delta}_{D_n,p})_{n=1}^\infty$ converges to $\nu_{\mathfrak{D}}$ by Theorem 3.7(i). Then (3.4) in Proposition 3.8 and Theorem 3.2 imply that $\nu_{\mathfrak{d}}$ is nonatomic. This is the desired assertion in the case where $m = 0$. If $m \geq 1$, then the fact that $\nu_{\mathfrak{D}}$ is nonatomic follows from (3.5) in Proposition 3.8, together with the fact that $\nu_{\mathfrak{d}}$ is nonatomic and the analyticity of the canonical branch \mathfrak{t} of T_p .

Case 2. There is a prime discriminant d that is divisible by p and a sequence $(f_n)_{n=1}^\infty$ in $\mathbb{Z}_{>0}$ such that for every n we have $D_n = df_n^2$ and $(d/f_n) = 1$ (respectively, $(d/f_n) = -1$). In this case the sequence $(\bar{\delta}_{D_n,p})_{n=1}^\infty$ converges weakly to $\nu_{\mathfrak{d}}^+$ (respectively, $\nu_{\mathfrak{d}}^-$) by Theorem 3.7(ii). Then (3.6) in Proposition 3.8 and Theorem 3.2 imply that $\nu_{\mathfrak{d}}^+$ and $\nu_{\mathfrak{d}}^-$ are both nonatomic. This is the desired assertion in the case where $m = 0$. If $m \geq 1$, then that $\nu_{\mathfrak{d}}^+$ and $\nu_{\mathfrak{d}}^-$ are both nonatomic follows from the fact that $\nu_{\mathfrak{d}}^+$ and $\nu_{\mathfrak{d}}^-$ are both nonatomic, from the last assertion of Proposition 3.8 and from the fact that the canonical branch \mathfrak{t} of T_p is analytic. \square

3.3 Proof of Theorem B

In the case where f is the j -invariant, the desired estimate is a direct consequence of (2.2) and [CU04, Théorème 2.4] if $v = \infty$ and of Theorem 3.1 and [HMR21, Theorems A and B], which are summarized in Theorem 3.7 in §3.2, if v is a prime number.

To prove Theorem B in the general case, let K be a finite extension of \mathbb{Q} inside $\overline{\mathbb{Q}}$, let f be a nonconstant modular function defined over K and let $\Phi(X, Y)$ be a modular polynomial of f in $K[X, Y]$. Note that $\Phi(X, Y)$ is irreducible in $\mathbb{C}_v[X, Y]$. Let C_3, θ and η be given by Lemma 2.8 and denote by δ_X (respectively, δ_Y) the degree of $\Phi(X, Y)$ in X (respectively, Y). Furthermore, note that $\Phi(X, \alpha)$ is nonzero (Proposition 2.4) and denote by Z the (possibly empty) finite set of zeros of this polynomial in \mathbb{C}_v .

Let τ be a quadratic imaginary number in \mathbb{H} that is not a pole of f , and put

$$j := j(\tau) \quad \text{and} \quad f := f(\tau).$$

Then, j is a singular modulus of the j -invariant and f is a singular modulus of f . Noting that for every σ in $\text{Gal}(\overline{\mathbb{Q}}/K)$ we have $\Phi(\sigma(j), \sigma(f)) = 0$, by Lemma 2.8 there is $r_0 > 0$ independent of τ such that for every r in $]0, r_0[$ we have

$$\begin{aligned} \#(\mathcal{O}_K(f) \cap \mathbf{D}_v(\alpha, r)) &\leq \delta_Y \# \{j' \in \mathcal{O}_K(j) : |j'|_v \geq C_3^{-1} r^{-\eta}\} \\ &\quad + \delta_Y \sum_{z_0 \in Z} \#(\mathcal{O}_K(j) \cap \mathbf{D}_v(z_0, C_3 r^\theta)). \end{aligned} \tag{3.7}$$

In the case where v is a prime number, we have

$$\{j' \in \mathcal{O}_K(j) : |j'|_v > 1\} = \emptyset,$$

so the desired estimate for f follows from that for the j -invariant, together with (3.7) and Proposition 2.3(ii). To prove the theorem in the case where $v = \infty$, we use the fact that the limit measure μ_∞ in [CU04, Théorème 2.4], seen as a measure on $\mathbb{P}^1(\mathbb{C})$, is nonatomic. Thus, there is $R > 1$ such that

$$\mu_\infty(\{z \in \mathbb{C}_v : |z|_v > R\}) \leq \frac{\varepsilon}{2\delta_Y},$$

and, if $\Phi(X, \alpha)$ is nonconstant, such that for every z_0 in Z we have

$$\mu_\infty(\mathbf{D}_v(z_0, R^{-1})) \leq \frac{\varepsilon}{2\delta_X \delta_Y}.$$

Then, the desired estimate for f and $v = \infty$ follows from that for the j -invariant, together with (3.7) and Proposition 2.3(ii).

4. p -Adic approximation by singular moduli

The goal of this section is to prove the following proposition, from which we derive Theorem C. Throughout this section, fix a prime number p .

PROPOSITION 4.1. *Let j_0 be a singular modulus of the j -invariant. Then, there exists a constant $A > 0$ such that for every singular modulus j of the j -invariant that is different from j_0 we have*

$$-\log |j - j_0|_p \leq A \log |D_j|.$$

The archimedean counterpart of this estimate was shown by Habegger [Hab15, Lemmas 5 and 8 and formula (11)]. See also Conjecture 1.3 in §1.3.

After some preliminaries in §4.1, the proofs of Proposition 4.1 and Theorem C are given in §§4.2 and 4.3, respectively. To prove Proposition 4.1, we use that CM points outside $Y_{\text{sup}s}(\mathbb{C}_p)$

are isolated [HMR20, Corollary B] to restrict to the case where the CM points corresponding to j and j_0 are both in $Y_{\text{supers}}(\mathbb{C}_p)$. In the case where the conductors of D_j and D_{j_0} are both p -adic units, we use an idea in the proof of [Cha18, Proposition 5.11]. To deduce the general case from this particular case, we use a formula in [HMR21] showing how the canonical branch of T_p relates CM points whose conductors differ by a power of p (Theorem 4.5).

Denote by $Y_{\text{supers}}(\overline{\mathbb{F}}_p)$ the (finite) set of isomorphism classes of supersingular elliptic curves over $\overline{\mathbb{F}}_p$. For e in $Y_{\text{supers}}(\overline{\mathbb{F}}_p)$, denote by \mathbf{D}_e the set of all E in $Y(\mathbb{C}_p)$ having good reduction, and such that the reduced class is e . The set \mathbf{D}_e is a residue disc in $Y(\mathbb{C}_p)$.

4.1 Formal \mathbb{Z}_p -modules and elliptic curves

In this section, we briefly recall the work of Gross and Hopkins in [HG94], on deformation spaces of formal modules. See also [HMR21, §§ 2.4 to 2.7] for a more detailed account of the results needed here. For every e in $Y_{\text{supers}}(\overline{\mathbb{F}}_p)$, we describe an action of $(\text{End}(e) \otimes \mathbb{Z}_p)^\times$ on a certain ramified covering of \mathbf{D}_e . In the proof of Proposition 4.1 we use a relation between the metric on \mathbf{D}_e and the natural metric of the covering, which is stated as Theorem 4.2 below.

Fix e in $Y_{\text{supers}}(\overline{\mathbb{F}}_p)$ and a representative elliptic curve defined over \mathbb{F}_{p^2} that we also denote by e . Denote by \mathcal{F}_e the formal group of e endowed with its natural structure of formal \mathbb{Z}_p -module and set

$$\mathbf{B}_e := \text{End}_{\overline{\mathbb{F}}_p}(\mathcal{F}_e) \otimes \mathbb{Q}_p, \quad \mathbf{R}_e := \text{End}_{\overline{\mathbb{F}}_p}(\mathcal{F}_e) \quad \text{and} \quad \mathbf{G}_e := \text{Aut}_{\overline{\mathbb{F}}_p}(\mathcal{F}_e).$$

Then, \mathbf{B}_e is a division quaternion algebra over \mathbb{Q}_p and the sets \mathbf{R}_e and \mathbf{G}_e embed in \mathbf{B}_e as the maximal order and its group of units, respectively. Denote by $g \mapsto \bar{g}$ the involution of \mathbf{B}_e , and for g in \mathbf{B}_e denote by $\text{nr}(g) := g\bar{g}$ in \mathbb{Q}_p its *reduced norm*. On the other hand, the function $\text{ord}_{\mathbf{B}_e} : \mathbf{B}_e \rightarrow \mathbb{Z} \cup \{\infty\}$ defined for g in \mathbf{B}_e by $\text{ord}_{\mathbf{B}_e}(g) := \text{ord}_p(\text{nr}(g))$, is the unique valuation extending the valuation 2ord_p on \mathbb{Q}_p . Identifying \mathbf{R}_e and \mathbf{G}_e with their images in \mathbf{B}_e , we have

$$\mathbf{R}_e = \{g \in \mathbf{B}_e : \text{ord}_{\mathbf{B}_e}(g) \geq 0\} \quad \text{and} \quad \mathbf{G}_e = \{g \in \mathbf{B}_e : \text{ord}_{\mathbf{B}_e}(g) = 0\}.$$

The function $\text{dist}_{\mathbf{B}_e} : \mathbf{B}_e \times \mathbf{B}_e \rightarrow \mathbb{R}$ defined for g and g' in \mathbf{B}_e by

$$\text{dist}_{\mathbf{B}_e}(g, g') := p^{-(1/2) \text{ord}_{\mathbf{B}_e}(g-g')},$$

defines an ultrametric distance on \mathbf{B}_e that makes \mathbf{B}_e into a topological algebra over \mathbb{Q}_p .

Identify the residue field of \mathbb{C}_p with an algebraic closure $\overline{\mathbb{F}}_p$ of \mathbb{F}_p and denote by $\pi : \mathcal{O}_p \rightarrow \overline{\mathbb{F}}_p$ the reduction map. Moreover, denote by \mathbb{Q}_{p^2} the unique unramified quadratic extension of \mathbb{Q}_p inside \mathbb{C}_p , and by \mathbb{Z}_{p^2} its ring of integers.

Let R_0 be a complete, local, Noetherian \mathbb{Z}_p -algebra with maximal ideal \mathcal{M}_0 and residue field isomorphic to a subfield \mathbb{k}_0 of $\overline{\mathbb{F}}_p$ that contains \mathbb{F}_{p^2} . Fix a reduction map $R_0 \rightarrow \mathbb{k}_0$. We are mainly interested in the special case where R_0 the ring of integers of a finite extension of \mathbb{Q}_p contained in \mathbb{C}_p together with the restriction of π , or a quotient of such ring of integers together with the morphism induced by the restriction of π . We stick to the general case for convenience.

A *deformation of \mathcal{F}_e over R_0* is a pair (\mathcal{F}, α) , where \mathcal{F} is a formal \mathbb{Z}_p -module over R_0 and $\alpha : \tilde{\mathcal{F}} \rightarrow \mathcal{F}_e$ is an isomorphism of formal \mathbb{Z}_p -modules defined over \mathbb{k}_0 . Here, $\tilde{\mathcal{F}}$ is the formal group over \mathbb{k}_0 obtained as the base change of \mathcal{F} under the reduction map $R_0 \rightarrow \mathbb{k}_0$. Two such deformations (\mathcal{F}, α) and (\mathcal{F}', α') are *isomorphic*, if there exists an isomorphism φ in $\text{Iso}_{R_0}(\mathcal{F}, \mathcal{F}')$ with reduction $\tilde{\varphi}$ such that $\alpha' \circ \tilde{\varphi} = \alpha$. Denote by $\mathbf{X}_e(R_0)$ the set of isomorphism classes of deformations of \mathcal{F}_e over R_0 .

For the rest of this section, we further assume that our choice of the representative elliptic curve e is such that \mathcal{F}_e is isomorphic over \mathbb{F}_{p^2} to the specialization of a universal formal \mathbb{Z}_p -module of height two; see [HMR21, Lemma 2.5]. Then, a consequence of the work of Gross

and Hopkins is that there exists a bijection

$$\mathcal{M}_0 \rightarrow \mathbf{X}_e(R_0) \tag{4.1}$$

that is functorial in R_0 ; see [HG94, §12] and [HMR21, §2.5] for details. Moreover, we have the action

$$\begin{aligned} \text{Aut}_{\mathbb{k}_0}(\mathcal{F}_e) \times \mathbf{X}_e(R_0) &\rightarrow \mathbf{X}_e(R_0) \\ (\beta, (\mathcal{F}, \alpha)) &\mapsto \beta \cdot (\mathcal{F}, \alpha) := (\mathcal{F}, \beta \circ \alpha). \end{aligned}$$

Let \mathcal{K} be a finite extension of \mathbb{Q}_{p^2} inside \mathbb{C}_p , with ring of integers $\mathcal{O}_{\mathcal{K}}$ and residue field \mathbb{k} . Consider the reduction map $\mathcal{O}_{\mathcal{K}} \rightarrow \mathbb{k}$ obtained as the restriction of π to $\mathcal{O}_{\mathcal{K}}$. Denote by $\mathbf{Y}(e, \mathcal{O}_{\mathcal{K}})$ the space of isomorphism classes of pairs (E, α) formed by an elliptic curve E given by a Weierstrass equation with coefficients in $\mathcal{O}_{\mathcal{K}}$ and having smooth reduction, and an isomorphism $\alpha: \tilde{E} \rightarrow e$ defined over \mathbb{k} . Here, two pairs (E, α) and (E', α') are *isomorphic* if there exists an isomorphism $\psi: E \rightarrow E'$ defined over \mathbb{k} such that $\alpha' \circ \psi = \alpha$. Consider the natural map

$$\mathbf{Y}(e, \mathcal{O}_{\mathcal{K}}) \rightarrow \mathbf{X}_e(\mathcal{O}_{\mathcal{K}}) \tag{4.2}$$

mapping a class in $\mathbf{Y}(e, \mathcal{O}_{\mathcal{K}})$ represented by a pair (E, α) , to the class in $\mathbf{X}_e(\mathcal{O}_{\mathcal{K}})$ represented by the deformation $(\mathcal{F}_E, \hat{\alpha})$. Here, $\hat{\alpha}: \tilde{\mathcal{F}}_E \rightarrow \mathcal{F}_e$ is the isomorphism induced by α . This map is known to be a bijection; see [LST64, §6] or [MC10, Theorem 4.1]. We obtain a map

$$\Pi_{e, \mathcal{K}}: \mathbf{X}_e(\mathcal{O}_{\mathcal{K}}) \rightarrow Y_{\text{sup}}(\overline{\mathbb{Q}}_p) \cap \mathbf{D}_e, \tag{4.3}$$

by composing the inverse of (4.2) with the natural map from $\mathbf{Y}(e, \mathcal{O}_{\mathcal{K}})$ to $Y_{\text{sup}}(\overline{\mathbb{Q}}_p) \cap \mathbf{D}_e$.

Consider

$$\mathcal{K} := \{\text{finite extensions of } \mathbb{Q}_{p^2} \text{ inside } \mathbb{C}_p\}$$

as a directed set with respect to the inclusion. For each \mathcal{K} in \mathcal{K} , consider the parametrization (4.1) with $R_0 = \mathcal{O}_{\mathcal{K}}$. Taking a direct limit over \mathcal{K} and then a completion, we obtain a set $\hat{\mathbf{D}}_e$ that is parametrized by the maximal ideal of \mathcal{O}_p . The action of \mathbf{G}_e on the system $\{\mathbf{X}_e(\mathcal{O}_{\mathcal{K}}): \mathcal{K} \in \mathcal{K}\}$ extends to a continuous map $\mathbf{G}_e \times \hat{\mathbf{D}}_e \rightarrow \hat{\mathbf{D}}_e$ that is analytic in the second variable; see [HMR21, §2.6] for details.

In the following theorem, $\delta_e := \#\text{Aut}(e)/2$. Note that $\delta_e = 1$ if $j(e) \neq 0, 1728$ and that in all the cases we have $1 \leq \delta_e \leq 12$; see, e.g., [Sil09, Appendix A, Proposition 1.2(c)].

THEOREM 4.2 [HMR21, Theorem 2.7]. *Fix e in $Y_{\text{sup}}(\overline{\mathbb{F}}_p)$. Then, the system $\{\Pi_{e, \mathcal{K}}: \mathcal{K} \in \mathcal{K}\}$ given by (4.3) defines a ramified covering map*

$$\Pi_e: \hat{\mathbf{D}}_e \rightarrow \mathbf{D}_e,$$

such that for every x in $\hat{\mathbf{D}}_e$ and every E in \mathbf{D}_e we have

$$\begin{aligned} \min\{|x - x'|_p: x' \in \Pi_e^{-1}(E)\}^{\delta_e} &\leq |j(\Pi_e(x)) - j(E)|_p \\ &\leq \min\{|x - x'|_p: x' \in \Pi_e^{-1}(E)\}. \end{aligned}$$

4.2 Proof of Proposition 4.1

The proof of Proposition 4.1 is at the end of this section.

Let e be in $Y_{\text{sup}}(\overline{\mathbb{F}}_p)$. The set

$$L(e) := \{\phi \in \mathbb{Z} + 2\text{End}(e): \text{tr}(\phi) = 0\},$$

is a \mathbb{Z} -lattice of dimension three inside $\text{End}(e)$. Define for each integer $m \geq 1$,

$$V_m(e) := \{\phi \in L(e): \text{nr}(\phi) = m\}.$$

For each fundamental p -adic discriminant \mathfrak{d} and every discriminant D in \mathfrak{d} , the image of the set $V_{|D|}(e)$ by the natural map $\text{End}(e) \rightarrow \text{End}_{\mathbb{F}_p}(\mathcal{F}_e)$, denoted by $\phi \mapsto \widehat{\phi}$, is contained in

$$\mathbf{L}_{e,\mathfrak{d}} := \{\varphi \in \mathbb{Z}_p + 2\mathbf{R}_e : \text{tr}(\varphi) = 0, -\text{nr}(\varphi) \in \mathfrak{d}\};$$

see [HMR21, Lemma 2.1]. Let $U_{e,\mathfrak{d}}: \mathbf{L}_{e,\mathfrak{d}} \rightarrow \mathbf{G}_e$ be the function defined by

$$U_{e,\mathfrak{d}}(\varphi) := \begin{cases} \frac{\varphi^2 + \varphi}{2} & \text{if } \frac{\varphi^2 + \varphi}{2} \text{ belongs to } \mathbf{G}_e; \\ 1 + \frac{\varphi^2 + \varphi}{2} & \text{otherwise,} \end{cases}$$

and for each φ in $\mathbf{L}_{e,\mathfrak{d}}$ define

$$\text{Fix}_e(\varphi) := \{x \in \widehat{\mathbf{D}}_e : U_{e,\mathfrak{d}}(\varphi) \cdot x = x\}.$$

Given a fundamental p -adic discriminant \mathfrak{d} , denote by $\mathbb{Q}_{p^2}(\sqrt{\mathfrak{d}})$ the compositum of \mathbb{Q}_{p^2} and $\mathbb{Q}_p(\sqrt{\mathfrak{d}})$.

PROPOSITION 4.3 [HMR21, Lemmas 4.5(iv) and 4.15, and Proposition 5.6(i)]. *Fix e in $Y_{\text{supps}}(\mathbb{F}_p)$ and a fundamental p -adic discriminant \mathfrak{d} .*

- (i) *For φ and φ' in $\mathbf{L}_{e,\mathfrak{d}}$ the sets $\text{Fix}_e(\varphi')$ and $\text{Fix}_e(\varphi)$ coincide if φ' belongs to $\mathbb{Q}_p(\varphi)$ and they are disjoint if φ' is not in $\mathbb{Q}_p(\varphi)$.*
- (ii) *We have $\Pi_e^{-1}(\Lambda_{\mathfrak{d}} \cap \mathbf{D}_e) \subseteq \mathbf{X}_e(\mathcal{O}_{\mathbb{Q}_{p^2}(\sqrt{\mathfrak{d}})})$, and for every Δ in \mathfrak{d} we have*

$$\Pi_e^{-1}(\Lambda_{\mathfrak{d}} \cap \mathbf{D}_e) = \text{Fix}_e(\{\varphi \in \mathbf{L}_{e,\mathfrak{d}} : \text{nr}(\varphi) = -\Delta\}).$$

For a fundamental p -adic discriminant \mathfrak{d} , put $\varepsilon_{\mathfrak{d}} := \frac{1}{2}$ if $\mathbb{Q}_p(\sqrt{\mathfrak{d}})$ is ramified over \mathbb{Q}_p and $\varepsilon_{\mathfrak{d}} := 1$ if $\mathbb{Q}_p(\sqrt{\mathfrak{d}})$ is unramified over \mathbb{Q}_p .

LEMMA 4.4. *For every prime number p , every e in $Y_{\text{supps}}(\mathbb{F}_p)$ and every fundamental p -adic discriminant \mathfrak{d} , the following property holds. For all φ and $\check{\varphi}$ in $\mathbf{L}_{e,\mathfrak{d}}$ and all x in $\text{Fix}_e(\varphi)$ and \check{x} in $\text{Fix}_e(\check{\varphi})$, we have*

$$|x - \check{x}|_p \geq p^{-\varepsilon_{\mathfrak{d}}} \text{dist}_{\mathbf{B}_e}(\varphi\check{\varphi}, \check{\varphi}\varphi).$$

Proof. Let ϖ_0 and ϖ be uniformizers of $\mathcal{O}_{\mathbb{Q}_p(\sqrt{\mathfrak{d}})}$ and $\mathcal{O}_{\mathbb{Q}_p(\varphi)}$, respectively, and note that

$$\text{ord}_p(\varpi_0) = \varepsilon_{\mathfrak{d}} = \frac{1}{2} \text{ord}_{\mathbf{B}_e}(\varpi). \tag{4.4}$$

If $x = \check{x}$, then $\check{\varphi}$ is in $\mathbb{Q}_p(\varphi)$ by Proposition 4.3(i) and, therefore, $\varphi\check{\varphi} = \check{\varphi}\varphi$. Thus, the desired property holds in this case. Assume $x \neq \check{x}$. By Proposition 4.3(ii), x and \check{x} are both in $\Pi_e^{-1}(\Lambda_{\mathfrak{d}} \cap \mathbf{D}_e)$ and, therefore, in $\mathbf{X}_e(\mathcal{O}_{\mathbb{Q}_{p^2}(\sqrt{\mathfrak{d}})})$. In particular, there is an integer $N \geq 1$ such that $|x - \check{x}|_p = |\varpi_0|_p^N$. Let (\mathcal{F}, α) and $(\check{\mathcal{F}}, \check{\alpha})$ represent x and \check{x} , respectively, and denote by \mathcal{F}_N and $\check{\mathcal{F}}_N$ the base change of \mathcal{F} and $\check{\mathcal{F}}$ under the projection map $\mathcal{O}_{\mathbb{Q}_{p^2}(\sqrt{\mathfrak{d}})} \rightarrow R_0 := \mathcal{O}_{\mathbb{Q}_{p^2}(\sqrt{\mathfrak{d}})} / \varpi_0^N \mathcal{O}_{\mathbb{Q}_{p^2}(\sqrt{\mathfrak{d}})}$. Since (4.1) is a bijection, there is an isomorphism $\psi: \mathcal{F}_N \rightarrow \check{\mathcal{F}}_N$ defined over R_0 such that $\check{\alpha} = \alpha \circ \psi$. This implies that the maps

$$\text{End}_{R_0}(\mathcal{F}_N) \rightarrow \mathbf{R}_e \quad \text{and} \quad \text{End}_{R_0}(\check{\mathcal{F}}_N) \rightarrow \mathbf{R}_e,$$

given by

$$\phi \mapsto \alpha \circ \tilde{\phi} \circ \alpha^{-1} \quad \text{and} \quad \phi \mapsto \check{\alpha} \circ \tilde{\phi} \circ \check{\alpha}^{-1},$$

respectively, have the same image. Thus, by [Gro86, Proposition 3.3] we have

$$\mathcal{O}_{\mathbb{Q}_p(\varphi)} + \varpi^{N-1} \mathbf{R}_e = \mathcal{O}_{\mathbb{Q}_p(\check{\varphi})} + \varpi^{N-1} \mathbf{R}_e.$$

It follows that $\varphi\check{\varphi} - \check{\varphi}\varphi$ is in $\varpi^{N-1}\mathbf{R}_e$. Together with (4.4), this implies

$$\text{dist}_{\mathbf{B}_e}(\varphi\check{\varphi}, \check{\varphi}\varphi) \leq |\varpi_0|_p^{N-1} = p^{\varepsilon_0} |x - \check{x}|_p. \quad \square$$

In the following theorem we use the canonical branch \mathfrak{t} of T_p , recalled in § 3.2.

THEOREM 4.5 [HMR21, Theorem 4.6]. *Let d be a fundamental discriminant such that $\Lambda_d \subseteq Y_{\text{supps}}(\mathbb{C}_p)$. Then for every integer $r \geq 1$ and every integer $f \geq 1$ that is not divisible by p , we have*

$$\Lambda_{d(fpr)^2} = \begin{cases} \mathfrak{t}^{-1}(\Lambda_{df^2}) \cap v_p^{-1}\left(\frac{1}{2p}\right) & \text{if } r = 1 \text{ and } p \text{ ramifies in } \mathbb{Q}(\sqrt{d}); \\ \mathfrak{t}^{1-r}(\Lambda_{d(fp)^2}) & \text{if } r \geq 2 \text{ and } p \text{ ramifies in } \mathbb{Q}(\sqrt{d}); \\ \mathfrak{t}^{-r}(\Lambda_{df^2}) & \text{if } r \geq 1 \text{ and } p \text{ is inert in } \mathbb{Q}(\sqrt{d}). \end{cases}$$

The following lemma is [HMR20, Lemma 4.9]; see also [CM06, Lemma 4.8] and [Gro86, Proposition 5.3].

LEMMA 4.6. *Denote Katz’s valuation by v_p , as in § 3.2. Let D be a discriminant such that $\Lambda_D \subseteq Y_{\text{supps}}(\mathbb{C}_p)$ and let $m \geq 0$ be the largest integer such that p^m divides the conductor of D . Then for every E in $\text{supp}(\Lambda_D)$ we have*

$$\min \left\{ v_p(E), \frac{p}{p+1} \right\} = \begin{cases} \frac{1}{2} \cdot p^{-m} & \text{if } p \text{ ramifies in } \mathbb{Q}(\sqrt{D}); \\ \frac{p}{p+1} \cdot p^{-m} & \text{if } p \text{ is inert in } \mathbb{Q}(\sqrt{D}). \end{cases}$$

Proof of Proposition 4.1. Let E_0 be the CM point such that $j(E_0) = j_0$. Since CM points outside $Y_{\text{supps}}(\mathbb{C}_p)$ are isolated [HMR20, Corollary B], we can assume that E_0 is in $Y_{\text{supps}}(\mathbb{C}_p)$. Let e be the element of $Y_{\text{supps}}(\overline{\mathbb{F}}_p)$ such that E_0 is in \mathbf{D}_e .

Let \mathfrak{d} be the fundamental p -adic discriminant and $m \geq 0$ the integer such that E_0 is in $\Lambda_{\mathfrak{d}p^{2m}}$. Let j be a singular modulus different from j_0 and let E be the CM point satisfying $j(E) = j$. Without loss of generality, assume that $D_j \neq D_{j_0}$ and that E is in \mathbf{D}_e . In view of Lemma 4.6, we can also assume that there is a fundamental p -adic discriminant \mathfrak{d}' such that D_j is in $p^{2m}\mathfrak{d}'$; see also [HMR21, Lemma 2.1]. Since $\Lambda_{\mathfrak{d}p^{2m}}$ and $\Lambda_{\mathfrak{d}'p^{2m}}$ are both compact by Theorem 3.7(i) and they are disjoint if $\mathfrak{d}' \neq \mathfrak{d}$, we can also assume that $\mathfrak{d}' = \mathfrak{d}$. On the other hand, by Theorem 4.5 and the fact that the canonical branch \mathfrak{t} of T_p is analytic, it is sufficient to prove the lemma in the case where $m = 0$, so E_0 and E are both in $\Lambda_{\mathfrak{d}} \cap \mathbf{D}_e$.

Using $\delta_e \leq 12$ and Theorem 4.2, we can find x_0 in $\Pi_e^{-1}(E_0)$ and x in $\Pi_e^{-1}(E)$ such that

$$|j - j_0|_p \geq |x - x_0|_p^{12}. \quad (4.5)$$

On the other hand, by Proposition 4.3(ii) there is ϕ_0 in $\text{End}(e)$ such that $\widehat{\phi}_0$ satisfies the equation $X^2 - D_{j_0} = 0$, is in $\mathbf{L}_{e,\mathfrak{d}}$ and is such that x_0 is in $\text{Fix}_e(\widehat{\phi}_0)$. Similarly, we can find ϕ in $\text{End}(e)$ such that $\widehat{\phi}$ satisfies the equation $X^2 - D_j = 0$, is in $\mathbf{L}_{e,\mathfrak{d}}$ and is such that x is in $\text{Fix}_e(\widehat{\phi})$. Note that Proposition 4.3(i) and our assumption $D_j \neq D_{j_0}$, imply that $\phi_0\phi - \phi\phi_0$ is nonzero. Combined with the fact that deg is a positive-definite quadratic form on $\text{End}(e)$ and [Sil09, Chapter V, Lemma 1.2], this implies

$$\begin{aligned} \text{ord}_{\mathbf{B}_e}(\widehat{\phi}_0\widehat{\phi} - \widehat{\phi}\widehat{\phi}_0) &= \text{ord}_p(\text{nr}(\widehat{\phi}_0\widehat{\phi} - \widehat{\phi}\widehat{\phi}_0)) = \text{ord}_p(\text{deg}(\phi_0\phi - \phi\phi_0)) \\ &\leq \log_p(\text{deg}(\phi_0\phi - \phi\phi_0)) \leq \log_p(4 \text{deg}(\phi_0) \text{deg}(\phi)) \\ &= \log_p(4 \text{deg}(\phi_0)|D_j|). \end{aligned}$$

Together Lemma 4.4, (4.5) and the inequality $|D_j| \geq 2$, this implies the desired estimate. \square

4.3 Proof of Theorem C

In the case where $v = \infty$, Theorem C is a direct consequence of the following proposition, and in the case where v is a prime number, Theorem C is a direct consequence of Proposition 4.1, the following proposition and Lemma 4.8 below.

PROPOSITION 4.7. *Let f be a nonconstant modular function defined over $\overline{\mathbb{Q}}$, let $\Phi(X, Y)$ be a modular polynomial of f in $\overline{\mathbb{Q}}[X, Y]$ and let v be in $M_{\mathbb{Q}}$. Furthermore, let α in $\overline{\mathbb{Q}}$ be a non-cuspidal value of f if $v = \infty$, or such that every root of $\Phi(X, \alpha)$ is badly approximable in \mathbb{C}_v by the singular moduli of the j -invariant if v is a prime number. Then, α is badly approximable in \mathbb{C}_v by the singular moduli of f .*

In the case where $v = \infty$, the hypothesis that α is a non-cuspidal value of f is necessary by Proposition 2.7(i). In the case where v is a prime number and α is an omitted value of f , the hypothesis on α is automatically satisfied because the polynomial $\Phi(X, \alpha)$ is constant (Proposition 2.4(i)).

Proof of Proposition 4.7. Note that $\Phi(X, Y)$ is irreducible in $\mathbb{C}_v[X, Y]$ and that $\Phi(X, \alpha)$ is nonzero by Proposition 2.4. Denote by Z the (possibly empty) finite set of zeros of $\Phi(X, \alpha)$ in \mathbb{C}_v . Furthermore, let C_3, θ and η be given by Lemma 2.8.

Suppose $v = \infty$. By [Hab15, Lemmas 5 and 8 and formula (11)] there are constants $A > 0$ and B such that for every z_0 in Z and every singular modulus j of the j -invariant different from z_0 , we have

$$-\log |j - z_0|_v \leq A \log |D_j| + B. \tag{4.6}$$

On the other hand, Proposition 2.4(ii), Lemma 2.8 and our hypothesis that α is a non-cuspidal value of f imply that Z is nonempty and that for every quadratic imaginary number τ in \mathbb{H} such that $f(\tau)$ is sufficiently close to α , we have

$$\min\{|j(\tau) - z_0|_v : z_0 \in Z\} \leq C_3 |f(\tau) - \alpha|_v^\theta. \tag{4.7}$$

Together with (4.6) and Proposition 2.3(iii), this implies that α is badly approximable in \mathbb{C} by the singular moduli of f .

It remains to consider the case where v is a prime number p . Recall that C_3 and η are given by Lemma 2.8 and put $r := C_3^{-1/\eta}$. In the case where α is an omitted value of f , the desired assertion is given by Proposition 2.7(ii). Suppose that α is a value of f , so $\Phi(X, \alpha)$ is nonconstant by Proposition 2.4(i). In particular, Z is nonempty. Proposition 2.3(iii) with f replaced by j and our hypotheses imply that there are constants $A > 0$ and B such that for every z_0 in Z and every singular modulus j of the j -invariant different from z_0 we have (4.6). On the other hand, reducing r if necessary Lemma 2.8 implies that for every quadratic imaginary number τ in \mathbb{H} such that $f(\tau)$ is in $\mathbf{D}_p(\alpha, r)$ the singular modulus $j(\tau)$ satisfies either (4.7) or

$$|j(\tau)|_p > C_3^{-1} |f(\tau) - \alpha|_p^{-\eta} > 1.$$

This last chain of inequalities is impossible since $j(\tau)$ is an algebraic integer. We thus have (4.7). Together with (4.6) and Proposition 2.3(iii), this implies that α is badly approximable in \mathbb{C}_p by the singular moduli of f . □

LEMMA 4.8. *Let h be a Hauptmodul defined over $\overline{\mathbb{Q}}$ and let $\Phi(X, Y)$ be a modular polynomial of h in $\overline{\mathbb{Q}}[X, Y]$. Then, for every singular modulus \mathfrak{h}_0 of h , every root of $\Phi(X, \mathfrak{h}_0)$ is a singular modulus of the j -invariant.*

The proof of this lemma is given after the following one.

LEMMA 4.9. *Let γ be an element of $\mathrm{SL}(2, \mathbb{R})$ that is contained in a subgroup of $\mathrm{SL}(2, \mathbb{R})$ commensurable to $\mathrm{SL}(2, \mathbb{Z})$. Then, there are integers a, b, c and d such that $ad - bc > 0$ and such that for every τ in \mathbb{H} we have $\gamma(\tau) = (a\tau + b)/(c\tau + d)$. In particular, the image by γ of a quadratic imaginary number in \mathbb{H} is also quadratic imaginary.*

Proof. Let Γ be a subgroup of $\mathrm{SL}(2, \mathbb{R})$ commensurable to $\mathrm{SL}(2, \mathbb{Z})$ containing γ . Then, the set of cusps of Γ is equal to that of $\mathrm{SL}(2, \mathbb{Z})$, which is equal to $\mathbb{P}^1(\mathbb{Q})$ (see [Shi71, Proposition 1.30]). It follows that $\gamma(\mathbb{P}^1(\mathbb{Q})) = \mathbb{P}^1(\mathbb{Q})$ and, in particular, that $\gamma(\infty)$ is in $\mathbb{P}^1(\mathbb{Q})$. In the case where $\gamma(\infty) \neq \infty$, let $\hat{\gamma}$ be an element of $\mathrm{SL}(2, \mathbb{R})$ such that for every τ in \mathbb{H} we have

$$\hat{\gamma}(\tau) = \frac{1}{\gamma(\tau) - \gamma(\infty)}.$$

Otherwise, put $\hat{\gamma} := \gamma$. In all of the cases, we have $\hat{\gamma}(\mathbb{Q}) = \mathbb{Q}$ and, therefore, there is λ in \mathbb{Q} such that $\lambda > 0$ and such that for every τ in \mathbb{H} we have $\hat{\gamma}(\tau) = \lambda\tau + \hat{\gamma}(0)$. Since $\hat{\gamma}(0)$ is in \mathbb{Q} , this implies the desired assertion for $\hat{\gamma}$ and, therefore, for γ . □

Proof Lemma 4.8. Let τ_0 be a quadratic imaginary number in \mathbb{H} such that $h(\tau_0) = \mathfrak{h}_0$. Note that $\Phi(X, Y)$ is irreducible over \mathbb{C} , so by Proposition 2.1 for each root j of $\Phi(X, \mathfrak{h}_0)$ there is τ in \mathbb{H} such that

$$j = j(\tau) \quad \text{and} \quad \mathfrak{h}_0 = h(\tau).$$

Since h is a *Hauptmodul*, there is γ in the stabilizer of h in $\mathrm{SL}(2, \mathbb{R})$ such that $\gamma(\tau_0) = \tau$. By Lemma 4.9, τ is also a quadratic imaginary number and, therefore, j is a singular modulus of the j -invariant. □

5. Proof of Theorems A and D

In this section we prove the following theorem and we deduce from it Theorems A and D.

THEOREM A'. *Let f be a nonconstant modular function defined over $\overline{\mathbb{Q}}$ and $\Phi(X, Y)$ a modular polynomial of f in $\overline{\mathbb{Q}}[X, Y]$. Moreover, let α in $\overline{\mathbb{Q}}$ be a non-cuspidal value of f and let S be a finite set of prime numbers p such that every root of $\Phi(X, \alpha)$ is badly approximable in \mathbb{C}_p by the singular moduli of the j -invariant. Then, there are at most finitely many singular moduli \mathfrak{f} of f such that $\mathfrak{f} - \alpha$ is an S -unit.*

Theorem D is a direct consequence of Theorem A' with $S = \emptyset$ and $\alpha = 0$ applied to f and to $1/f$. Another direct consequence of Theorem A' is the following version of Theorem D for S -units, under the hypothesis that there is an affirmative solution to Conjecture 1.3.

COROLLARY 5.1. *Let S be a finite set of prime numbers p such that every algebraic number is badly approximable in \mathbb{C}_p by the singular moduli of the j -invariant. Moreover, let f be a nonconstant modular function defined over $\overline{\mathbb{Q}}$ that is not a weak modular unit. Then, there are at most a finite number of singular moduli of f that are S -units.*

The proof of Theorem A' is given in §5.1. In §5.2 we prove Theorem A and the following corollary of Theorem A'. To state it, recall that a subgroup of $\mathrm{SL}(2, \mathbb{R})$ is a *congruence group*, if for some N in $\mathbb{Z}_{>0}$ it contains

$$\left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}(2, \mathbb{Z}) : a, d \equiv 1 \pmod{N} \text{ and } b, c \equiv 0 \pmod{N} \right\}$$

as a finite index subgroup. The following corollary shows that an affirmative solution to Conjecture 1.3 would yield a version of Theorem A for a general congruence or genus zero group and a general algebraic value.

COROLLARY 5.2. *Let f be a nonconstant modular function defined over $\overline{\mathbb{Q}}$ for a congruence or a genus-zero group and let α in $\overline{\mathbb{Q}}$ be a value of f . Suppose that for every prime number p , every algebraic number is badly approximable in \mathbb{C}_p by the singular moduli of the j -invariant. Then, there are at most a finite number of singular moduli \mathfrak{f} of f such that $\mathfrak{f} - \alpha$ is an S -unit.*

Corollary 5.2 applied to f and to $1/f$ with $\alpha = 0$, shows that an affirmative solution to Conjecture 1.3 would yield a version Theorem D that holds under the weaker hypothesis that f is not a modular unit, but that is restricted to congruence or to genus-zero groups.

5.1 Proof of Theorem A'

Recall that $M_{\mathbb{Q}}$, and for each v in $M_{\mathbb{Q}}$, the norm field $(\mathbb{C}_v, |\cdot|_v)$, are defined in §1.2. Given a finite extension K of \mathbb{Q} inside $\overline{\mathbb{Q}}$, denote by M_K the set of all norms on K that for some v in $M_{\mathbb{Q}}$ coincide with $|\cdot|_v$ on \mathbb{Q} . For such w and v , write $w | v$, let $(\mathbb{C}_w, |\cdot|_w)$ be a completion of an algebraic closure of (K, w) and denote by K_w the closure of K inside \mathbb{C}_w . Note that $(\mathbb{C}_w, |\cdot|_w)$ and $(\mathbb{C}_v, |\cdot|_v)$ are isomorphic as normed fields and that $(K_w, |\cdot|_w)$ is a completion of (K, w) . Moreover, identify the algebraic closure of K inside \mathbb{C}_w with $\overline{\mathbb{Q}}$, put

$$\nu_w := \frac{[K_w : \mathbb{Q}_v]}{[K : \mathbb{Q}]} \quad \text{and} \quad \|\cdot\|_w := |\cdot|_w^{\nu_w},$$

and for all α in \mathbb{C}_v and $r > 0$ put

$$\mathbf{D}_w(\alpha, r) := \{z \in \mathbb{C}_v : |z - \alpha|_w < r\}.$$

Note that $\nu_w \leq 1$; see, e.g., [BG06, Corollary 1.3.2].

Let

$$\log^+ : [0, \infty[\rightarrow \mathbb{R} \quad \text{and} \quad \log^- : [0, \infty[\rightarrow \mathbb{R} \cup \{-\infty\}$$

be the functions defined by

$$\log^+(x) := \log \max\{1, x\} \quad \text{and} \quad \log^-(x) := \log \min\{1, x\}.$$

Denote by $h_W : \overline{\mathbb{Q}} \rightarrow \mathbb{R}$ the *Weil or naive height*, which for each finite extension K of \mathbb{Q} inside $\overline{\mathbb{Q}}$ and every α in K is given by

$$h_W(\alpha) = \sum_{w \in M_K} \log^+ \|\alpha\|_w.$$

In this formula, the right-hand side is independent of the finite extension K of \mathbb{Q} inside $\overline{\mathbb{Q}}$ containing α . Note that by the triangle inequality, for all α_0 and α in $\overline{\mathbb{Q}}$ we have

$$h_W(\alpha - \alpha_0) \geq h_W(\alpha) - h_W(\alpha_0) - \log 2. \tag{5.1}$$

The proof of Theorem A' is given after a couple of lemmas.

LEMMA 5.3. *Let α in $\overline{\mathbb{Q}}$ be given and let K be a finite extension of \mathbb{Q} inside $\overline{\mathbb{Q}}$ that does not necessarily contain α . Then, we have*

$$h_W(\alpha) = -\frac{1}{\#\mathcal{O}_K(\alpha)} \sum_{w \in M_K} \sum_{\alpha' \in \mathcal{O}_K(\alpha)} \log^- \|\alpha'\|_w.$$

Proof. Let \widehat{K} be a finite extension of K inside $\overline{\mathbb{Q}}$ containing $\mathcal{O}_K(\alpha)$, let

$$P(z) = z^d + a_{d-1}z^{d-1} + \dots + a_0$$

be the minimal polynomial of α in $K[x]$ and for every w in M_K put

$$\|P\|_w := \max\{\|a_j\|_w : j \in \{0, \dots, d-1\}\}.$$

For every archimedean w in M_K (respectively, \widehat{w} in $M_{\widehat{K}}$) put $S_w^1 := \{z \in \mathbb{C}_w : |z|_w = 1\}$ (respectively, $S_{\widehat{w}}^1 := \{z \in \mathbb{C}_{\widehat{w}} : |z|_{\widehat{w}} = 1\}$) and denote by λ_w (respectively, $\lambda_{\widehat{w}}$) the Haar measure of this group. Then, we have

$$\begin{aligned} \sum_{\alpha' \in \mathcal{O}_K(\alpha)} \log^+ \|\alpha'\|_w &= \int \log \|P(z)\|_w \, d\lambda_w(z) \\ &= \sum_{\substack{\widehat{w} \in M_{\widehat{K}} \\ \widehat{w}|_w}} \int \log \|P(z)\|_{\widehat{w}} \, d\lambda_{\widehat{w}}(z) \\ &= \sum_{\substack{\widehat{w} \in M_{\widehat{K}} \\ \widehat{w}|_w}} \sum_{\alpha' \in \mathcal{O}_K(\alpha)} \log^+ \|\alpha'\|_{\widehat{w}}; \end{aligned} \tag{5.2}$$

see, e.g., [BG06, Corollary 1.3.2 and Proposition 1.6.5]. Similarly, for every non-archimedean w in M_K we have

$$\begin{aligned} \sum_{\alpha' \in \mathcal{O}_K(\alpha)} \log^+ \|\alpha'\|_w &= \log \|P\|_w \\ &= \sum_{\substack{\widehat{w} \in M_{\widehat{K}} \\ \widehat{w}|_w}} \log \|P\|_{\widehat{w}} \\ &= \sum_{\substack{\widehat{w} \in M_{\widehat{K}} \\ \widehat{w}|_w}} \sum_{\alpha' \in \mathcal{O}_K(\alpha)} \log^+ \|\alpha'\|_{\widehat{w}}; \end{aligned}$$

see, e.g., [BG06, Corollary 1.3.2 and Lemma 1.6.3]. Combined with (5.2) and with the Galois invariance of the Weil height (see, e.g., [BG06, Proposition 1.5.17]), this implies

$$\begin{aligned} \#\mathcal{O}_K(\alpha) \, h_W(\alpha) &= \sum_{\widehat{w} \in M_{\widehat{K}}} \sum_{\alpha' \in \mathcal{O}_K(\alpha)} \log^+ \|\alpha'\|_{\widehat{w}} \\ &= \sum_{w \in M_K} \sum_{\alpha' \in \mathcal{O}_K(\alpha)} \log^+ \|\alpha'\|_w. \end{aligned} \tag{5.3}$$

On the other hand, by the product formula applied to the element $\prod_{\alpha' \in \mathcal{O}_K(\alpha)} \alpha'$ of K we have

$$\prod_{w \in M_K} \prod_{\alpha' \in \mathcal{O}_K(\alpha)} \|\alpha'\|_w = 1;$$

see, e.g., [BG06, Proposition 1.4.4]. Together with (5.3), this implies the desired identity. \square

The following lemma is an extension of [Hab15, Lemma 3] to the more general setting considered here.

LEMMA 5.4. *Let K be a finite extension of \mathbb{Q} inside $\overline{\mathbb{Q}}$ and let f be a nonconstant modular function defined over K . Then, for every singular modulus \mathfrak{f}_0 of f there are constants $A_0 > 0$*

and B_0 such that for every singular modulus \mathfrak{f} of f we have

$$h_W(\mathfrak{f} - \mathfrak{f}_0) \geq A_0 \log(\# O_K(\mathfrak{f})) + B_0.$$

Proof. Let $\Phi(X, Y)$ be a modular polynomial of f in $\overline{\mathbb{Q}}[X, Y]$ and denote by δ_X and δ_Y the degree of $\Phi(X, Y)$ in X and Y , respectively. For each k in $\{0, \dots, \delta_X\}$ let $P_k(Y)$ in $K[Y]$ be the coefficient of X^k in $\Phi(X, Y)$, and let $M_k > 0$ be such that for every α in $\overline{\mathbb{Q}}$ we have

$$h_W(P_k(\alpha)) \leq \deg(P_k) h_W(\alpha) + M_k;$$

see, e.g., [Sil09, Chapter VIII, Theorem 5.6]. Thus, if we put

$$\Delta := \sum_{k=0}^{\delta_X} \deg(P_k) \quad \text{and} \quad M := \sum_{k=0}^{\delta_X} M_k,$$

then for every quadratic imaginary number τ in \mathbb{H} that is not a pole of f we have

$$h_W(j(\tau)) - \delta_X \log 2 \leq \sum_{k=0}^{\delta_X} h_W(P_k(f(\tau))) \leq \Delta h_W(f(\tau)) + M;$$

see, e.g., [Sil09, Chapter VIII, Theorem 5.9]. Combined with (5.1), Proposition 2.3(iii) and [Hab15, Lemma 3], which is based on Colmez’s lower bound [Col98, Théorème 1], we obtain the desired estimate. \square

Proof of Theorem A’. Let K be a finite extension of \mathbb{Q} containing α and the coefficients of Φ and denote by S_0 the set of all w in M_K such that for some v in $S \cup \{\infty\}$ we have $w \mid v$. Let A_0 be the constant given by Lemma 5.4. By Propositions 2.3(ii) and 4.7, for every v in $M_{\mathbb{Q}}$ there are constants $A_v > 0$ and B_v such that for every singular modulus \mathfrak{f} of f we have

$$-\log |\mathfrak{f} - \alpha|_v \leq A_v \log(\# O_K(\mathfrak{f})) + B_v. \tag{5.4}$$

For every w in M_K such that $w \mid v$, put $A_w := A_v$ and $B_w := B_v$.

Suppose that there is a sequence of pairwise distinct singular moduli $(\mathfrak{f}_n)_{n=1}^{\infty}$ of f such that for every n the difference $\mathfrak{f}_n - \alpha$ is an S -unit. By Proposition 2.3(iv), we have

$$\# O_K(\mathfrak{f}_n) \rightarrow \infty \quad \text{as } n \rightarrow \infty. \tag{5.5}$$

Together with Theorem B applied to each v in $S \cup \{\infty\}$, this implies that there is r in $]0, 1[$ such that for every w in S_0 and every sufficiently large $n \geq 1$, we have

$$\#(O_K(\mathfrak{f}_n) \cap \mathbf{D}_w(\alpha, r)) \leq \frac{A_0}{2A_w(\#S_0 + 1)} \# O_K(\mathfrak{f}_n).$$

Thus, for every sufficiently large n we have

$$\frac{\#(O_K(\mathfrak{f}_n - \alpha) \cap \mathbf{D}_w(0, r))}{\# O_K(\mathfrak{f}_n - \alpha)} = \frac{\#(O_K(\mathfrak{f}_n) \cap \mathbf{D}_w(\alpha, r))}{\# O_K(\mathfrak{f}_n)} \leq \frac{A_0}{2A_w(\#S_0 + 1)}.$$

Combined with Lemma 5.3, (5.4), the fact that for every w in M_K we have $\nu_w \leq 1$ and our assumption that $\mathfrak{f}_n - \alpha$ is an S -unit, this implies that for some constant B independent of n

we have

$$\begin{aligned} h_W(f_n - \alpha) &= -\frac{1}{\#\mathcal{O}_K(f_n - \alpha)} \sum_{w \in S_0} \left(\sum_{\substack{\beta \in \mathcal{O}_K(f_n - \alpha) \\ |\beta|_w < r}} \log \|\beta\|_w + \sum_{\substack{\beta \in \mathcal{O}_K(f_n - \alpha) \\ r \leq |\beta|_w < 1}} \log \|\beta\|_w \right) \\ &\leq \sum_{w \in S_0} \frac{\#(\mathcal{O}_K(f_n - \alpha) \cap \mathbf{D}_v(0, r))}{\#\mathcal{O}_K(f_n - \alpha)} (A_w \log(\#\mathcal{O}_K(f_n)) + B_w) + (\#S_0 + 1) \log \frac{1}{r} \\ &\leq \frac{A_0}{2} \log(\#\mathcal{O}_K(f_n)) + B. \end{aligned}$$

In view of (5.5), letting $n \rightarrow \infty$ we obtain a contradiction with Lemma 5.4 that completes the proof of the theorem. \square

5.2 Proof of Theorem A and Corollary 5.2

The proofs are given after a few lemmas.

LEMMA 5.5. *Let V be a projective curve defined over $\overline{\mathbb{Q}}$ and let g_0 be a rational function defined on V over \mathbb{C} such that every zero and every pole of g_0 is in $V(\overline{\mathbb{Q}})$. If there exists z_0 in $V(\overline{\mathbb{Q}})$ such that $g_0(z_0) = 1$, then g_0 is defined over $\overline{\mathbb{Q}}$.*

Proof. If g_0 is constant, then it is equal to 1 and the result follows. Assume g_0 is nonconstant. Let σ in $\text{Aut}(\mathbb{C}|\overline{\mathbb{Q}})$ be given and denote by g_0^σ the image of g_0 under the action of σ on rational functions. The hypothesis that every zero and every pole of g_0 is in $V(\overline{\mathbb{Q}})$ implies that g_0 and g_0^σ have the same zeros and poles, and that the corresponding multiplicities are the same. This implies that g_0/g_0^σ is constant. Evaluating at z_0 and using

$$g_0^\sigma(z_0) = \sigma(g_0(z_0)) = \sigma(1) = 1 = g_0(z_0),$$

we conclude that $g_0 = g_0^\sigma$. Since σ is arbitrary, we get that g_0 is defined over $\overline{\mathbb{Q}}$. \square

LEMMA 5.6. *Let f be a nonconstant modular function defined over $\overline{\mathbb{Q}}$ for a genus-zero subgroup of $\text{SL}(2, \mathbb{R})$, such that 0 is a value of f . Then, there is a holomorphic Hauptmodul h defined over $\overline{\mathbb{Q}}$ and a nonconstant rational function $R(X)$ in $\overline{\mathbb{Q}}(X)$, such that 0 is a (non-cuspidal) value of h and we have*

$$R(0) = 0 \quad \text{and} \quad R(h) = f.$$

Proof. Our hypotheses imply that the stabilizer Γ of f in $\text{SL}(2, \mathbb{R})$ is of genus zero. Put $\widehat{\Gamma} := \Gamma \cap \text{SL}(2, \mathbb{Z})$ and note that $\widehat{\Gamma}$ has finite index in Γ and in $\text{SL}(2, \mathbb{Z})$. Denote by $\Pi: X(\widehat{\Gamma}) \rightarrow X(\Gamma)$ the map induced by the identity on \mathbb{H} and by \widehat{j} and \widehat{f} the meromorphic functions defined on $X(\widehat{\Gamma})$ that are induced by j and f , respectively. Note that the meromorphic function f_0 defined on $X(\Gamma)$ that is induced by f satisfies $\widehat{f} = f_0 \circ \Pi$.

First, we show that $X(\widehat{\Gamma})$ can be defined over $\overline{\mathbb{Q}}$ in such a way that \widehat{j} and \widehat{f} correspond to rational functions defined over $\overline{\mathbb{Q}}$. Let $\Phi(X, Y)$ in $\overline{\mathbb{Q}}[X, Y]$ be a modular polynomial for f and denote by $Z(\Phi)$ the zero set of Φ in \mathbb{C}^2 . Denote by P the finite subset of $X(\widehat{\Gamma})$ formed by the poles of \widehat{j} and those of \widehat{f} and let $\widehat{\varphi}$ be the function defined by

$$\begin{aligned} \widehat{\varphi}: X(\widehat{\Gamma}) \setminus P &\rightarrow Z(\Phi) \\ z &\mapsto (\widehat{j}(z), \widehat{f}(z)). \end{aligned}$$

By the definition of $\widehat{\Gamma}$, for every γ in $\text{SL}(2, \mathbb{Z})$ outside $\widehat{\Gamma}$ the meromorphic functions \widehat{f} and $\widehat{f} \circ \gamma$ are different. Thus, the set E_γ of all points of $X(\widehat{\Gamma})$ at which these functions agree is finite. Let \mathcal{R}

be a set of representatives of the right cosets of $\widehat{\Gamma}$ in $SL(2, \mathbb{Z})$ that are different from $\widehat{\Gamma}$ and put

$$E := P \cup \bigcup_{\gamma \in \mathcal{R}} E_\gamma.$$

Then, \mathcal{R} and E are both finite and the restriction of $\widehat{\varphi}$ to $X(\widehat{\Gamma}) \setminus E$ is injective. Thus, $\widehat{\varphi}$ induces a birational isomorphism between $X(\widehat{\Gamma})$ and $Z(\Phi)$. By [Har77, Chapter I, Corollary 6.11], there exist a smooth projective curve V defined over $\overline{\mathbb{Q}}$ and birational isomorphisms

$$\phi: X(\widehat{\Gamma}) \dashrightarrow V(\mathbb{C}) \quad \text{and} \quad \psi: V(\mathbb{C}) \dashrightarrow Z(\Phi),$$

such that ψ is defined over $\overline{\mathbb{Q}}$ and $\psi \circ \phi$ defines the same birational isomorphism as $\widehat{\varphi}$. Note that ϕ extends to an isomorphism $X(\widehat{\Gamma}) \rightarrow V(\mathbb{C})$; see, e.g., [Har77, Chapter I, Proposition 6.8]. Under this isomorphism, \widehat{j} and \widehat{f} correspond to the composition of ψ with the projections on the first and second coordinate on $Z(\Phi)$, respectively, both of which are defined over $\overline{\mathbb{Q}}$. Thus, ϕ and $V(\mathbb{C})$ induce an algebraic structure on $X(\widehat{\Gamma})$ over $\overline{\mathbb{Q}}$ for which \widehat{j} and \widehat{f} correspond to rational functions defined over $\overline{\mathbb{Q}}$. In what follows, we fix this algebraic structure on $X(\widehat{\Gamma})$.

Next, we show that $X(\Gamma)$ can be defined over $\overline{\mathbb{Q}}$ in such a way that Π corresponds to a rational function defined over $\overline{\mathbb{Q}}$. To do this, it is sufficient to show that there is a biholomorphic map $h_0: X(\Gamma) \rightarrow \mathbb{P}^1(\mathbb{C})$ for which the composition $h_0 \circ \Pi$ is defined over $\overline{\mathbb{Q}}$. Choose pairwise distinct numbers α_0, α_1 and α_∞ in $\overline{\mathbb{Q}}$ and choose z_0, z_1 and z_∞ in $f_0^{-1}(\alpha_0), f_0^{-1}(\alpha_1)$ and $f_0^{-1}(\alpha_\infty)$, respectively. Since $X(\Gamma)$ is of genus zero, there is a biholomorphic map $h_0: X(\Gamma) \rightarrow \mathbb{P}^1(\mathbb{C})$ mapping z_0, z_1 and z_∞ to $0, 1$ and ∞ , respectively. Thus, z_1 and every zero and every pole of $h_0 \circ \Pi$ is defined over $\overline{\mathbb{Q}}$ and, therefore, $h_0 \circ \Pi$ is defined over $\overline{\mathbb{Q}}$ by Lemma 5.5. We conclude that $X(\Gamma)$ and Π are both defined over $\overline{\mathbb{Q}}$ with respect to the algebraic structure induced by h_0 . In what follows, we fix this algebraic structure on $X(\Gamma)$. Note that f_0 is also defined over $\overline{\mathbb{Q}}$, because \widehat{f} is. Since the cuspidal values of f_0 are defined over $\overline{\mathbb{Q}}$ (Proposition 2.4), it follows that each cusp of $X(\Gamma)$ is also defined over $\overline{\mathbb{Q}}$.

To complete the proof of the lemma, note that our hypothesis that 0 is a value of f implies that there is τ_0 in \mathbb{H} such that $f(\tau_0) = 0$. The point z_0 of $X(\Gamma)$ defined by τ_0 is not a cusp of $X(\Gamma)$ and is defined over $\overline{\mathbb{Q}}$. It follows that there is a biholomorphic map $X(\Gamma) \rightarrow \mathbb{P}^1(\mathbb{C})$ defined over $\overline{\mathbb{Q}}$ mapping z_0 to 0 and the cusp of $X(\Gamma)$ defined by $i\infty$ to ∞ . The lift h to \mathbb{H} of this function is a holomorphic *Hauptmodul* for Γ that is defined over $\overline{\mathbb{Q}}$ and satisfies $h(\tau_0) = 0$. From the results proved in the previous paragraphs, it follows that there is $R(X)$ in $\overline{\mathbb{Q}}(X)$ such that $R(h) = f$. Evaluating at τ_0 , we conclude that $R(0) = 0$. Finally, note that since h is a *Hauptmodul* and 0 is a value of h , we have that 0 is a non-cuspidal value of h . □

LEMMA 5.7. *Let h be a holomorphic Hauptmodul defined over $\overline{\mathbb{Q}}$, let $R(X)$ in $\overline{\mathbb{Q}}(X)$ be non-constant and such that $R(0) = 0$ and put $f := R(h)$. Suppose that for every finite set of prime numbers S , there are at most a finite number of singular moduli of h that are S -units. Then, f is a nonconstant modular function defined over $\overline{\mathbb{Q}}$ and for every finite set of prime numbers S there are at most a finite number of singular moduli of f that are S -units.*

Proof. That f is a nonconstant modular function follows from the fact that h has the same property and that $R(X)$ is nonconstant. To show that f is defined over $\overline{\mathbb{Q}}$, note that h is algebraically dependent with the j -invariant over $\overline{\mathbb{Q}}$ because h is defined over $\overline{\mathbb{Q}}$. It follows that f is also algebraically dependent with the j -invariant over $\overline{\mathbb{Q}}$ and, therefore, that it is defined over $\overline{\mathbb{Q}}$.

Let S be a finite set of prime numbers. By Corollary 2.5(i) there is a finite set of prime numbers S_0 such that every singular modulus of h is an S_0 -integer. Our hypotheses

that $R(X)$ is nonconstant and $R(0) = 0$ imply that there are ℓ in $\mathbb{Z}_{>0}$, a in $\overline{\mathbb{Q}} \setminus \{0\}$ and monic polynomials $P(X)$ and $Q(X)$ in $\overline{\mathbb{Q}}[X]$, such that

$$P(0) \neq 0, \quad Q(0) \neq 0 \quad \text{and} \quad R(X) = aX^\ell \frac{P(X)}{Q(X)}. \tag{5.6}$$

Let S_1 be a finite set of prime numbers containing S and S_0 and such that each of the coefficients of $P(X)$ and of $Q(X)$ is an S_1 -integer and each of the numbers a , $P(0)$ and $Q(0)$ is an S_1 -unit.

By hypothesis, the set U of all those singular moduli of h that are S_1 -units is finite. Let \mathfrak{f} be a singular modulus of f outside the finite set $R(U)$, let τ be a quadratic imaginary number such that $f(\tau) = \mathfrak{f}$ and put $\mathfrak{h} := h(\tau)$. Then, $\mathfrak{f} = R(\mathfrak{h})$ and \mathfrak{h} is a singular modulus of h outside U . It follows that \mathfrak{h} is an S_1 -integer that is not an S_1 -unit. That is, there is a prime number p outside S_1 and σ in $\text{Gal}(\overline{\mathbb{Q}}|\mathbb{Q})$ such that $|\sigma(\mathfrak{h})|_p < 1$. Denote by $P^\sigma(X)$ and $Q^\sigma(X)$ the image of $P(X)$ and $Q(X)$ by the induced action of σ on $\overline{\mathbb{Q}}[X]$, respectively. In view of our choice of S_1 , we have

$$|\sigma(a)|_p = |P^\sigma(\sigma(\mathfrak{h}))|_p = |Q^\sigma(\sigma(\mathfrak{h}))|_p = 1.$$

Together with (5.6), this implies

$$|\sigma(\mathfrak{f})|_p = \left| \sigma(a)\sigma(\mathfrak{h})^\ell \frac{P^\sigma(\sigma(\mathfrak{h}))}{Q^\sigma(\sigma(\mathfrak{h}))} \right|_p = |\sigma(\mathfrak{h})|_p^\ell < 1.$$

This proves that \mathfrak{f} is not an S_1 -unit and, therefore, that it is not an S -unit. □

Proof of Theorem A. Put $f_0 := f - \mathfrak{f}_0$ and note that 0 is a value of f_0 . Let h and R be given by Lemma 5.6 with f replaced by f_0 . Combining Proposition 4.1, Lemma 4.8 and Theorem A' with f replaced by h and with $\alpha = 0$, we obtain that for every finite set of prime numbers S , there are at most a finite number of singular moduli of h that are S -units. Together with Lemma 5.7, this implies that f_0 has the same property. It follows that for every finite set of prime numbers S , there are at most a finite number of singular moduli \mathfrak{f} of f that $\mathfrak{f} - \mathfrak{f}_0$ is an S -unit. □

The proof of Corollary 5.2 is given after the following lemma.

LEMMA 5.8. *Let f be a nonconstant modular function defined over $\overline{\mathbb{Q}}$ for a congruence group Γ contained in $\text{SL}(2, \mathbb{Z})$. Then, for every cusp c of $X(\Gamma)$ there exists m in $\mathbb{Z}_{>0}$ and a modular unit g defined over $\overline{\mathbb{Q}}$ for Γ such that the following property holds. No cusp of $X(\Gamma)$ different from c is a zero or a pole of the meromorphic function defined on $X(\Gamma)$ induced by $f^m g$.*

Proof. Let j_0 and f_0 be the meromorphic functions defined on $X(\Gamma)$ induced by the j -invariant and by f , respectively. Moreover, let Z be the finite set of cusps of $X(\Gamma)$ and for each z in Z denote by n_z the order of f_0 at z . If for every z in $Z \setminus \{c\}$ we have $n_z = 0$, then the desired assertion holds with $m = 1$ and with g equal to the constant function equal to 1. Suppose this is not the case, so the divisor D on $X(\Gamma)$ defined by

$$D := \left(\sum_{z \in Z \setminus \{c\}} n_z \right) c - \sum_{z \in Z \setminus \{c\}} n_z z,$$

is nonzero. Note that the degree of D is zero. Applying the Manin–Drinfel’d theorem repeatedly [Dri73, Theorem 1], we obtain that there exists m in $\mathbb{Z}_{>0}$ and a nonconstant meromorphic function g_0 defined on $X(\Gamma)$ such that the divisor of zeros and poles of g_0 equals mD . It follows that the modular function g induced by g_0 is a modular unit for Γ such that $f_0^m g_0$ has no zeros or poles in $Z \setminus \{c\}$. To complete the proof of the lemma, it remains to show that there is a nonzero complex number s such that sg is defined over $\overline{\mathbb{Q}}$. To do this, note that the Riemann surface $X(\Gamma)$

has a structure of projective variety defined over $\overline{\mathbb{Q}}$ for which j_0 is given by a rational function defined over $\overline{\mathbb{Q}}$; see, e.g., [Shi71, Chapter 6.7]. In particular, each element of Z is defined over $\overline{\mathbb{Q}}$ with respect to this algebraic structure. Choose a point z_0 in $X(\Gamma) \setminus Z$ defined over $\overline{\mathbb{Q}}$, note that $g_0(z_0)$ is a nonzero complex number and put $s := g_0(z_0)^{-1}$. By Lemma 5.5 with g_0 replaced by sg_0 , the function sg_0 corresponds to a rational function on $X(\Gamma)$ defined over $\overline{\mathbb{Q}}$. Since this is also the case for j_0 , we have that j_0 and sg_0 are algebraically dependent over $\overline{\mathbb{Q}}$. This implies that sg is defined over $\overline{\mathbb{Q}}$ and completes the proof of the lemma. \square

Proof of Corollary 5.2. In the case where α is a non-cuspidal value of f , the desired assertion follows from Theorem A'. Suppose α is a cuspidal value of f and let Γ be the stabilizer of f in $\mathrm{SL}(2, \mathbb{R})$.

Suppose first that Γ is a congruence group, put $\widehat{\Gamma} := \Gamma \cap \mathrm{SL}(2, \mathbb{Z})$ and let c be a cusp of $X(\widehat{\Gamma})$ at which the meromorphic function f_0 defined on $X(\widehat{\Gamma})$ induced by f takes the value α . Let m and g be given by Lemma 5.8 with f replaced by $f - \alpha$ and with Γ replaced by $\widehat{\Gamma}$ and put $\widehat{f} := (f - \alpha)^m g$. Then, 0 is a non-cuspidal value of \widehat{f} or of $1/\widehat{f}$ and Theorem A' with α replaced by 0 implies that there are at most a finite number of singular moduli of \widehat{f} that are S -units. On the other hand, by Corollary 2.5 there is a finite set of prime numbers S_0 such that every singular modulus of g is an S_0 -unit. Putting $S_1 := S \cup S_0$, we conclude that there are at most a finite number of singular moduli \mathfrak{f} of f such that $\mathfrak{f} - \alpha$ is an S_1 -unit. Since S_1 contains S , this implies the desired assertion.

It remains to consider the case where Γ is of genus zero. Let h be the *Hauptmodul* given by Lemma 5.6 with f replaced by $f - \alpha$. Theorem A' with f replaced by h and with α replaced by 0, implies that for every finite set of prime numbers S there are at most a finite number of singular moduli of h that are S -units. Together with Lemma 5.7, this implies that $f - \alpha$ has the same property. \square

ACKNOWLEDGEMENTS

The authors would like to thank Shouwu Zhang for pointing out that a statement like the Main Theorem might be obtained from knowledge about the p -adic asymptotic behavior of CM points, and Philipp Habegger for sharing his conjecture on the λ -invariants that prompted Theorem A. Finally, we thank the referee for insightful comments. During the preparation of this work the first named author was partially supported by ANID/CONICYT FONDECYT Iniciación 11220567. The second named author was partially supported by ANID/CONICYT FONDECYT Regular 1211858. The third named author acknowledges partial support from NSF grant DMS-1700291. The authors would like to thank the Pontificia Universidad Católica de Valparaíso, the University of Rochester and Universitat de Barcelona for hospitality during the preparation of this work.

CONFLICTS OF INTEREST

None.

Appendix A. Fourier series expansion of modular functions

The goal of this section is to give conditions on a modular function to be defined over a given subfield of \mathbb{C} .

A meromorphic function f defined on \mathbb{H} is *periodic*, if there is h in $\mathbb{Z}_{>0}$ such that for every τ in \mathbb{H} we have $f(\tau + h) = f(\tau)$. The *period* of f is the least h satisfying this property. In this

case, f admits a Fourier series expansion at $i\infty$ of the form

$$f(\tau) = \sum_{n=-\infty}^{\infty} a_n \exp\left(\frac{2\pi in}{h}\tau\right).$$

The function f is *meromorphic* (respectively, *holomorphic*) at $i\infty$, if for every sufficiently large integer n (respectively, every n in $\mathbb{Z}_{>0}$) we have $a_{-n} = 0$.

Note that every modular function is periodic and therefore it admits a Fourier series expansion at $i\infty$. The goal of this appendix is to prove the following proposition.

PROPOSITION A.1. *Let f be a modular function whose Fourier series expansion at $i\infty$ has coefficients in a subfield K of \mathbb{C} . Then f is defined over K .*

The proof of this proposition is given after the following lemma.

LEMMA A.2. *Let K be a subfield of \mathbb{C} and let \mathcal{A} be a finite subset of \mathbb{C} that is not contained in K . Then, there is a field homomorphism $K(\mathcal{A}) \rightarrow \mathbb{C}$ that is the identity on K and that is different from the inclusion.*

Proof. Denote by \overline{K} the algebraic closure of K inside \mathbb{C} .

Suppose first that \mathcal{A} is contained in \overline{K} . By the primitive element theorem, there is α in \overline{K} such that $K(\mathcal{A}) = K(\alpha)$. Our assumption that \mathcal{A} is not contained in K implies that the minimal polynomial of α over K is of degree at least two. Thus, this polynomial has a root α' different from α . It follows that there is a field homomorphism $K(\alpha) \rightarrow \mathbb{C}$ that is the identity on K and that maps α to α' . It is thus different from the inclusion.

It remains to consider the case where \mathcal{A} is not contained in \overline{K} . In this case, there is a nonempty subset \mathcal{A}_0 of \mathcal{A} that is algebraically independent over \overline{K} . Increasing \mathcal{A}_0 if necessary, assume it is maximal with this property. Then, $\overline{K}(\mathcal{A})$ is a finite extension of $\overline{K}(\mathcal{A}_0)$. Since $\overline{K}(\mathcal{A}_0)$ is isomorphic to the field of rational functions with coefficients in \overline{K} in $\#\mathcal{A}_0$ variables, there is a field isomorphism $\sigma: \overline{K}(\mathcal{A}_0) \rightarrow \overline{K}(\mathcal{A}_0)$ that is the identity on \overline{K} and such that for some a_0 in \mathcal{A}_0 we have $\sigma(a_0) = 2a_0$. Since $\overline{K}(\mathcal{A})$ is a finite extension of $\overline{K}(\mathcal{A}_0)$ and \mathbb{C} is algebraically closed, σ extends to a field homomorphism $\overline{K}(\mathcal{A}) \rightarrow \mathbb{C}$. \square

Proof of Proposition A.1. We use that $1/j$ is holomorphic at $i\infty$; see, e.g., [Lan87, Chapter 4, §1]. Replacing f by $1/f$ if necessary, assume that f is also holomorphic at $i\infty$. Let $\Phi(X, Y)$ be a modular polynomial of f in $\mathbb{C}[X, Y]$ (Proposition 2.1). Replacing Φ by a constant multiple if necessary, assume that one of the coefficients of Φ is equal to 1. Denote by δ the degree of X in $\Phi(X, Y)$, and note that the polynomial

$$\Psi(X, Y) := X^\delta \Phi(1/X, Y)$$

in $\mathbb{C}[X, Y]$ is also irreducible.

For each pair of nonnegative integers (k, ℓ) , denote by $A_{k,\ell}$ the coefficient of $X^k Y^\ell$ in $\Psi(X, Y)$. Moreover, denote by I the set of all (k, ℓ) such that $A_{k,\ell} \neq 0$. By our normalization of Φ , there is (k_0, ℓ_0) such that $A_{k_0,\ell_0} = 1$. Suppose that $\Psi(X, Y)$ is not in $K[X, Y]$, so the set

$$\mathcal{A} := \{A_{k,\ell} : (k, \ell) \in I\}$$

is not contained in K . By Lemma A.2, there is a field homomorphism $\sigma: K(\mathcal{A}) \rightarrow \mathbb{C}$ that is the identity on K and that is different from the inclusion. It follows that for some (k', ℓ') in \mathcal{A} we have $\sigma(A_{k',\ell'}) \neq A_{k',\ell'}$.

For each integer $n \geq 0$, denote by $a_n^{k,\ell}$ the coefficient of $\exp((2\pi in/h)\tau)$ in the Fourier series expansion of $(1/j)^k f^\ell$. Since $1/j$ is holomorphic at $i\infty$ and its Fourier series expansion has

coefficients in \mathbb{Q} (see, e.g., [Lan87, Chapter 4, § 1]), our hypothesis implies that $a_n^{k,\ell}$ is in K . On the other hand, the fact that the function $\Psi(1/j, f)$ vanishes identically implies that for every integer $n \geq 0$ we have

$$\sum_{(k,\ell) \in I} A_{k,\ell} a_n^{k,\ell} = 0 \quad \text{and} \quad \sum_{(k,\ell) \in I} \sigma(A_{k,\ell}) a_n^{k,\ell} = 0.$$

It follows that the polynomial

$$\Psi_0(X, Y) := \sum_{(k,\ell) \in I} (\sigma(A_{k,\ell}) - A_{k,\ell}) X^k Y^\ell$$

in $\mathbb{C}[X, Y]$, is such that the function $\Psi_0(1/j, f)$ vanishes identically. Note also that Ψ is nonzero, because the coefficient of $X^{k'} Y^{\ell'}$ in $\Psi_0(X, Y)$ is nonzero by our choice of σ . Moreover, the coefficient of $X^{k_0} Y^{\ell_0}$ in $\Psi_0(X, Y)$ is zero, so Ψ_0 is not a scalar multiple of Ψ .

Consider the polynomial

$$\Phi_0(X, Y) := X^\delta \Psi_0(1/X, Y)$$

in $\mathbb{C}[X, Y]$. The functions $\Phi(j, f)$ and $\Phi_0(j, f)$ vanish identically. By Proposition 2.1(i), this implies that the polynomial Φ_0 vanishes on the zero set of Φ . Since Φ is irreducible over \mathbb{C} , we conclude that Φ divides Φ_0 . Since the degree of Φ_0 in X and in Y is less than or equal to the corresponding degree of Φ , we conclude Φ_0 is a scalar multiple of Φ . However, this would imply that Ψ_0 is a scalar multiple of Ψ , which is false. This contradiction proves that $\Phi(X, Y)$ is in $K[X, Y]$, and completes the proof of the proposition. \square

REFERENCES

- BHK20 Y. Bilu, P. Habegger and L. Kühne, *No singular modulus is a unit*, Int. Math. Res. Not. IMRN **2020** (2020), 10005–10041.
- BG06 E. Bombieri and W. Gubler, *Heights in Diophantine geometry*, New Mathematical Monographs, vol. 4 (Cambridge University Press, Cambridge, 2006).
- Bor92 R. E. Borcherds, *Monstrous moonshine and monstrous Lie superalgebras*, Invent. Math. **109** (1992), 405–444.
- Cam21 F. Campagna, *On singular moduli that are S -units*, Manuscripta Math. **166** (2021), 73–90.
- Cha18 F. Charles, *Exceptional isogenies between reductions of pairs of elliptic curves*, Duke Math. J. **167** (2018), 2039–2072.
- CY96 I. Chen and N. Yui, *Singular values of Thompson series*, in *Groups, difference sets, and the Monster (Columbus, OH, 1993)*, Ohio State University Mathematical Research Institute Publications, vol. 4 (de Gruyter, Berlin, 1996), 255–326.
- CU04 L. Clozel and E. Ullmo, *Équidistribution des points de Hecke*, in *Contributions to automorphic forms, geometry, and number theory* (Johns Hopkins University Press, Baltimore, MD, 2004), 193–254.
- CM06 R. Coleman and K. McMurdy, *Fake CM and the stable model of $X_0(Np^3)$* , Doc. Math. **Extra Vol.** (2006), 261–300.
- Col98 P. Colmez, *Sur la hauteur de Faltings des variétés abéliennes à multiplication complexe*, Compos. Math. **111** (1998), 359–368.
- CN79 J. H. Conway and S. P. Norton, *Monstrous moonshine*, Bull. Lond. Math. Soc. **11** (1979), 308–339.
- DH09 S. David and N. Hirata-Kohno, *Linear forms in elliptic logarithms*, J. Reine Angew. Math. **628** (2009), 37–89.

- DI95 F. Diamond and J. Im, *Modular forms and modular curves*, in *Seminar on Fermat's Last Theorem (Toronto, ON, 1993–1994)*, CMS Conference Proceedings, vol. 17 (American Mathematical Society, Providence, RI, 1995), 39–133.
- Dri73 V. G. Drinfel'd, *Two theorems on modular curves*, Funkcional. Anal. i Priložen. **7** (1973), 83–84.
- Duk88 W. Duke, *Hyperbolic distribution problems and half-integral weight Maass forms*, Invent. Math. **92** (1988), 73–90.
- ES10 A. Enge and A. V. Sutherland, *Class invariants by the CRT method*, in *Algorithmic number theory*, Lecture Notes in Computer Science, vol. 6197 (Springer, Berlin, 2010), 142–156.
- Frö68 A. Fröhlich, *Formal groups*, Lecture Notes in Mathematics, vol. 74 (Springer, Berlin–New York, 1968).
- Gro86 B. H. Gross, *On canonical and quasicanonical liftings*, Invent. Math. **84** (1986), 321–326.
- GZ85 B. H. Gross and D. B. Zagier, *On singular moduli*, J. Reine Angew. Math. **355** (1985), 191–220.
- Hab15 P. Habegger, *Singular moduli that are algebraic units*, Algebra Number Theory **9** (2015), 1515–1524.
- Har77 R. Hartshorne, *Algebraic geometry*, Graduate Texts in Mathematics, vol. 52 (Springer, New York–Heidelberg, 1977).
- HMR20 S. Herrero, R. Menares and J. Rivera-Letelier, *p-adic distribution of CM points and Hecke orbits I: Convergence towards the Gauss point*, Algebra Number Theory **14** (2020), 1239–1290.
- HMR21 S. Herrero, R. Menares and J. Rivera-Letelier, *p-Adic distribution of CM points and Hecke orbits. II: Linnik equidistribution on the supersingular locus*, Preprint (2021), [arXiv:2102.04865](https://arxiv.org/abs/2102.04865).
- HG94 M. J. Hopkins and B. H. Gross, *Equivariant vector bundles on the Lubin-Tate moduli space*, in *Topology and representation theory (Evanston, IL, 1992)*, Contemporary Mathematics, vol. 158 (American Mathematical Society, Providence, RI, 1994), 23–88.
- Kat73 N. M. Katz, *p-adic properties of modular schemes and modular forms*, Lecture Notes in Mathematics, vol. 350 (Springer, Berlin–New York, 1973), 69–190.
- KL81 D. S. Kubert and S. Lang, *Modular units*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 244 (Springer, New York–Berlin, 1981).
- Lan87 S. Lang, *Elliptic functions*, second edition, Graduate Texts in Mathematics, vol. 112 (Springer, New York, 1987), with an appendix by J. Tate.
- Li21 Y. Li, *Singular units and isogenies between CM elliptic curves*, Compos. Math. **157** (2021), 1022–1035.
- LST64 J. Lubin, J.-P. Serre and J. Tate, *Elliptic curves and formal groups*, in *Seminar at Woods Hole Institute on algebraic geometry (1964)*, <https://web.ma.utexas.edu/users/voloch/LST/lst.pdf>.
- MC10 K. McMurdy and R. Coleman, *Stable reduction of $X_0(p^3)$* , Algebra Number Theory **4** (2010), 357–431, with an appendix by Everett W. Howe.
- Sch76 R. Schertz, *Die singulären Werte der Weberschen Funktionen $f, f_1, f_2, \gamma_2, \gamma_3$* , J. Reine Angew. Math. **286** (1976), 46–74.
- Shi71 G. Shimura, *Introduction to the arithmetic theory of automorphic functions (Kanô Memorial Lectures, vol. 1)*, Publications of the Mathematical Society of Japan, vol. 11 (Princeton University Press, Princeton, NJ; Iwanami Shoten, Tokyo, 1971).
- Sie35 C. Siegel, *Über die Classenzahl quadratischer Zahlkörper*, Acta Arith. **1** (1935), 83–86.
- Sil09 J. H. Silverman, *The arithmetic of elliptic curves*, second edition, Graduate Texts in Mathematics, vol. 106 (Springer, Dordrecht, 2009).
- Sut A. Sutherland, *Norms of singular moduli for $|D| \leq 2000$* , <https://math.mit.edu/~drew/NormsOfSingularModuli2000.pdf>.
- Web08 H. Weber, *Lehrbuch der Algebra, volume III*, second edition (Braunschweig, 1908).
- YY16 T. Yang and H. Yin, *Some non-congruence subgroups and the associated modular curves*, J. Number Theory **161** (2016), 17–48.

THERE ARE AT MOST FINITELY MANY SINGULAR MODULI THAT ARE S -UNITS

- YYY21 T. Yang, H. Yin and P. Yu, *The lambda invariants at CM points*, Int. Math. Res. Not. IMRN **2021** (2021), 5542–5603.
- YZ97 N. Yui and D. Zagier, *On the singular values of Weber modular functions*, Math. Comp. **66** (1997), 1645–1662.

Sebastián Herrero sebastian.herrero.m@gmail.com

Departamento de Matemática y Ciencia de la Computación, Universidad de Santiago de Chile,
Av. Libertador Bernardo O'Higgins 3363, Santiago, Chile

Ricardo Menares rmenares.v@gmail.com

Facultad de Matemáticas, Pontificia Universidad Católica de Chile, Vicuña Mackenna 4860,
Santiago, Chile

Juan Rivera-Letelier riveraletelier@gmail.com

Department of Mathematics, University of Rochester, Hylan Building, Rochester, NY 14627,
USA

Compositio Mathematica is owned by the Foundation Compositio Mathematica and published by the London Mathematical Society in partnership with Cambridge University Press. All surplus income from the publication of *Compositio Mathematica* is returned to mathematics and higher education through the charitable activities of the Foundation, the London Mathematical Society and Cambridge University Press.