hexagon:2611.00478v2

A LOWER BOUND FOR THE ESSENTIAL MINIMUM OF THE ZHANG–ZAGIER HEIGHT

TEJASVI SINGH TOMAR

Abstract. For an algebraic number α let $h(\alpha)$ be its absolute logarithmic Weil height, and let $h_Z(\alpha) = h(\alpha) + h(1-\alpha)$. The essential minimum of h_Z is the supremum of the numbers H for which only finitely many algebraic numbers satisfy $h_Z(\alpha) \le H$. Zagier showed that it is at least $\frac{1}{2}\log\frac{1+\sqrt{5}}{2} = 0.2406059\ldots$, and the best lower bound so far is 0.24874, due to Flammang. We prove that the essential minimum is at least 0.24982933. The proof combines Smyth's method of auxiliary functions, here with five integer polynomials in $z(1-z)$, with a logarithmic energy term relative to a measure on the circles $|z| = 1$ and $|1-z| = 1$. Resultants and discriminants of minimal polynomials are nonzero integers. They bound the polynomial terms and the energy term, and a condition on the degrees absorbs the leading coefficients, so the argument applies to all algebraic numbers. A superharmonicity argument reduces the positivity of the auxiliary function to the unit circle, where two separately written programs verify it in ball arithmetic for a certificate with exact rational data. We also prove this lower bound in Lean.

1. Introduction

Let α be an algebraic number of degree n, and let $P(x) = a\prod_{i=1}^{n}(x-\alpha_i)$ be its minimal polynomial over $\mathbb{Z}$, primitive and with leading coefficient $a \ge 1$. The absolute logarithmic Weil height of α is

$$h(\alpha) = \frac{1}{n}\Big(\log a + \sum_{i=1}^{n}\log^+|\alpha_i|\Big), \qquad \log^+ x = \log\max(1,x).$$

The Zhang–Zagier height is $h_Z(\alpha) = h(\alpha) + h(1-\alpha)$. Its essential minimum is

$$\mu^{\mathrm{ess}}(h_Z) = \sup\big\{H \in \mathbb{R} :\ \{\alpha \in \overline{\mathbb{Q}} : h_Z(\alpha) \le H\} \text{ is finite}\big\}.$$

It is constant 82a in the collection of optimization constants of Davis, Ivanisvili, Tao and contributors [4]. Zhang [22, 23] proved that $\mu^{\mathrm{ess}}(h_Z) > 0$. Zagier [21] showed that $h_Z(\alpha) \ge \frac{1}{2}\log\frac{1+\sqrt{5}}{2} = 0.2406059\ldots$ for all $\alpha \notin \{0, 1, e^{\pm i\pi/3}\}$, with equality when α or $1-\alpha$ is a primitive tenth root of unity. Doche [5] raised the lower bound for the essential minimum to 0.24824, and Flammang [8] raised it to 0.24874. Both used Smyth's method of auxiliary functions [20]. From above, Doche's construction of algebraic numbers with small h_Z gives $\mu^{\mathrm{ess}}(h_Z) \le 0.25444$ [6], and Grinsztajn's computations based on it give $\mu^{\mathrm{ess}}(h_Z) \le 0.2536332$ [9]. Pinho [16] reports 0.2534449, using Doche's construction with factors from [9] and the Épi project [7] and further factors of his own; the Épi project reports 0.2535745. Burgos Gil, Menares, Qu and Sombra [2] proved a duality theorem for the essential minima of heights on $\overline{\mathbb{Q}}$ defined by Green functions, and h_Z is twice such a height, with Green function $\frac{1}{2}(\log^+|z| + \log^+|1-z|)$. For h_Z their theorem says that the supremum of the lower bounds from auxiliary functions, without energy terms, and the infimum of a dual linear program over measures are both equal to $\mu^{\mathrm{ess}}(h_Z)$. Their auxiliary functions use

2020 *Mathematics Subject Classification.* 11G50, 11R04; 11Y40, 31A05, 65G30.

Key words and phrases. Weil height, Zhang–Zagier height, essential minimum, auxiliary functions, logarithmic energy, discriminant, ball arithmetic, computer-assisted proof, formal verification, Lean 4.

polynomials in z; replacing $P(z)$ by $P(z)P(1-z)$ shows that polynomials in $z(1-z)$ give the same supremum.

Theorem 1.1. *Let* $m_0 = \frac{87900883718627874129}{351843720888320000000}$. *For every* $\varepsilon > 0$ *only finitely many algebraic numbers* α *satisfy* $h(\alpha) + h(1-\alpha) \le m_0 - \varepsilon$. *Consequently*

$$\mu^{\mathrm{ess}}(h_Z) \ge m_0 > 0.24982933.$$

The method. Smyth's method bounds h_Z from below by an inequality

$$\log^+ |z| + \log^+ |1-z| \;\ge\; m + \sum_j c_j \log |Q_j(z(1-z))| \qquad (z \in \mathbb{C})$$

with integer polynomials Q_j and positive weights c_j. Averaging over the conjugates of α turns the left side into $h_Z(\alpha)$, apart from leading coefficients. Each sum $\sum_i \log |Q_j(\alpha_i(1-\alpha_i))|$ is bounded below because a resultant is a nonzero integer. We add one further term. For a probability measure ν on $\mathbb{C}$ with potential $U_\nu(z) = \int \log|z-y|\, d\nu(y)$ and energy $I(\nu) = \int U_\nu \, d\nu$, the average of $2U_\nu - I(\nu)$ over the conjugates of α is at least the energy of a smoothed copy of their distribution, minus a small error. That energy is controlled by the discriminant of P, which is a nonzero integer. Orloski, Talebizadeh Sardari and Smith [15] combined Smyth's method with an energy condition of this kind, which comes from the integrality of discriminants, for the trace problem of Schur, Siegel and Smyth. With the energy term the inequality need only hold up to $\lambda(2U_\nu - I(\nu))$ for some $\lambda \ge 0$, and this extra freedom raises the bound. By the duality theorem of [2], auxiliary functions without energy terms come arbitrarily close to $\mu^{\mathrm{ess}}(h_Z)$, so the energy term does not move the limit of what auxiliary functions can give; it raises the bound obtained from a fixed finite set of polynomials (Section 6).

We take ν invariant under $z \mapsto 1-z$ and under complex conjugation, supported on the circles $|z| = 1$ and $|1-z| = 1$. These two circles cut the plane into four regions on each of which the auxiliary function is superharmonic, so its minimum lies on the circles, and by symmetry on the upper half of the unit circle. The certificate consists of five polynomials in $w = z(1-z)$, their weights, the knot values of a piecewise linear function from which ν is obtained by normalization and symmetrization, and the value m_0, all exact rationals. The nonnegativity of the auxiliary function on the half circle is proved in ball arithmetic by two programs, described in Section 4.

Section 2 proves the lower bound for all algebraic numbers from an auxiliary function with an energy term. Section 3 reduces the positivity to the unit circle, and Section 4 describes the certificate, its verification and the proof of Theorem 1.1. Section 5 describes the formal proof in Lean, and Section 6 collects numerical observations.

2. Auxiliary functions with an energy term

We write $f(z) = \log^+ |z| + \log^+ |1-z|$, $T(z) = z(1-z)$ and $\sigma(z) = 1-z$. For finite positive Borel measures θ, η on $\mathbb{C}$ with compact support let

$$U_\eta(z) = \int \log|z-y|\, d\eta(y), \qquad I(\theta,\eta) = \int U_\eta \, d\theta, \qquad I(\theta) = I(\theta,\theta).$$

Since $\log|z-y|$ is bounded above on compact sets, Tonelli's theorem gives $I(\theta,\eta) = I(\eta,\theta) \in [-\infty, \infty)$. If $I(\theta)$, $I(\eta)$ and $I(\theta,\eta)$ are finite, we put $I(\theta-\eta) = I(\theta) - 2I(\theta,\eta) + I(\eta)$. If moreover θ and η have the same total mass, then $I(\theta-\eta) \le 0$ by [19, Lemma I.1.8]. The lemma applies to the signed measure $\theta - \eta$, because its positive and negative parts are dominated by θ and η, so they have finite energy, and by bilinearity its energy is $I(\theta-\eta)$. (In [19] the energy is defined with $\log(1/|z-y|)$, so the inequality reads ≥ 0 there.) We write $B(y,r)$ for the open disc of radius r around y, $\sigma_*\eta$ for the image of η under σ, and $\bar\eta$ for its image under complex conjugation.

Lemma 2.1 (Leading coefficients)**.** *Let α have degree n, minimal polynomial $P(x) = a\prod_i(x-\alpha_i)$ as above. Then $h_Z(\alpha) = \frac{1}{n}\bigl(2\log a + \sum_{i=1}^n f(\alpha_i)\bigr)$.*

Proof. The polynomial $(-1)^n P(1-x)$ is irreducible, primitive, has leading coefficient a and the roots $1-\alpha_i$. So $h(1-\alpha) = \frac{1}{n}(\log a + \sum_i \log^+|1-\alpha_i|)$. □

Lemma 2.2 (Auxiliary function with an energy term)**.** *Let $Q_1,\dots,Q_\ell \in \mathbb{Z}[w]$ be nonzero, $c_1,\dots,c_\ell > 0$, $\lambda \ge 0$ and $m \in \mathbb{R}$. Let ν be a probability measure on $\mathbb{C}$ with compact support, and let $K > 0$ satisfy $\nu(B(y,r)) \le Kr$ for all $y \in \mathbb{C}$ and $0 < r \le 1$. (Then U_ν and $I(\nu)$ are finite, see the proof, so the function F in* (ii) *takes values in $(-\infty,+\infty]$.) Assume*

(i) $\sum_j c_j \deg Q_j + \lambda \le 1$;

(ii) $F(z) := f(z) - m - \sum_j c_j \log|Q_j(T(z))| - \lambda\bigl(2U_\nu(z) - I(\nu)\bigr) \ge 0$ *for all $z \in \mathbb{C}$, where $F = +\infty$ at the roots of the polynomials $Q_j(T(z))$.*

Let α be an algebraic number of degree $n \ge 1/(2K)$ that is not a root of any $Q_j(T(x))$. Then

$$h_Z(\alpha) \;\ge\; m - \frac{\lambda\,(1+\log(2Kn))}{n}. \tag{2.1}$$

Proof. Let $P(x) = a\prod_i(x-\alpha_i)$ be the minimal polynomial of α. The condition on ν implies that ν has no atoms. For $x \ne y$ we have $\max(\log|x-y|, \log r) = \log|x-y| + \int_0^r \mathbf{1}_{|x-y|<u}\,du/u$, so Tonelli's theorem and the condition on ν give, for every $y \in \mathbb{C}$ and $0 < r \le 1$,

$$\int \max\bigl(\log|x-y|, \log r\bigr)\,d\nu(x) = U_\nu(y) + \int_0^r \nu(B(y,u))\,\frac{du}{u} \;\le\; U_\nu(y) + Kr. \tag{2.2}$$

Taking $r = 1$, we get $U_\nu(y) \ge \int \log^+|x-y|\,d\nu(x) - K \ge -K$, and U_ν is bounded above on compact sets because ν has compact support. So U_ν is finite everywhere and $I(\nu)$ is finite.

Resultants. The polynomial $R_j(x) = Q_j(x(1-x))$ lies in $\mathbb{Z}[x]$ and has degree $2\deg Q_j$. It does not vanish at α, and P is irreducible, so R_j vanishes at no α_i. In particular every $F(\alpha_i)$ is a real number. Hence $\operatorname{Res}(P,R_j) = a^{2\deg Q_j}\prod_i R_j(\alpha_i)$ is a nonzero integer, and

$$\sum_{i=1}^n \log|Q_j(T(\alpha_i))| \;\ge\; -2\deg(Q_j)\log a. \tag{2.3}$$

Discriminant. The discriminant $\operatorname{disc}(P) = a^{2n-2}\prod_{i<k}(\alpha_i-\alpha_k)^2$ is a nonzero integer, so

$$\sum_{i\ne k} \log|\alpha_i - \alpha_k| \;\ge\; -(2n-2)\log a. \tag{2.4}$$

Smoothing. Let $\omega_{y,r}$ be the normalized arc length measure on the circle $|z-y| = r$, and let $\tau_r = \frac{1}{n}\sum_i \omega_{\alpha_i,r}$ for some $0 < r \le 1$. The potential of $\omega_{y,r}$ is $\max(\log|z-y|, \log r)$. Hence $I(\omega_{y,r}) = \log r$, and for $i \ne k$

$$I(\omega_{\alpha_i,r}, \omega_{\alpha_k,r}) = \int \max(\log|z-\alpha_k|, \log r)\,d\omega_{\alpha_i,r}(z) \;\ge\; \int \log|z-\alpha_k|\,d\omega_{\alpha_i,r}(z) \;\ge\; \log|\alpha_i-\alpha_k|,$$

the last step because $\log|z-\alpha_k|$ is subharmonic. With (2.4),

$$I(\tau_r) \;\ge\; \frac{1}{n^2}\sum_{i\ne k}\log|\alpha_i-\alpha_k| + \frac{\log r}{n} \;\ge\; -\frac{(2n-2)\log a}{n^2} + \frac{\log r}{n}. \tag{2.5}$$

Energy. The potential of τ_r is at least $\log r$, and the potential of ν is at least $-K$. Both potentials are bounded above on compact sets. So $I(\tau_r)$, $I(\nu)$ and $I(\tau_r,\nu)$ are finite, and since $I(\tau_r - \nu) \le 0$ we have $I(\tau_r) \le 2I(\tau_r,\nu) - I(\nu)$. By Fubini's theorem and (2.2),

$$I(\tau_r,\nu) = \frac{1}{n}\sum_i \int U_\nu\,d\omega_{\alpha_i,r} = \frac{1}{n}\sum_i \int \max\bigl(\log|x-\alpha_i|, \log r\bigr)\,d\nu(x) \;\le\; \frac{1}{n}\sum_i U_\nu(\alpha_i) + Kr.$$

Together with (2.5),

$$\frac{1}{n}\sum_{i=1}^{n}\big(2U_\nu(\alpha_i)-I(\nu)\big) \;\geq\; -\frac{(2n-2)\log a}{n^2}+\frac{\log r}{n}-2Kr. \tag{2.6}$$

Conclusion. We average (ii) over the conjugates, use (2.3) and (2.6), and add $\frac{2}{n}\log a$ as in Lemma 2.1:

$$h_Z(\alpha) \;\geq\; m+\frac{2\log a}{n}\Big(1-\sum_j c_j\deg Q_j-\lambda\,\frac{n-1}{n}\Big)-\lambda\Big(2Kr+\frac{\log(1/r)}{n}\Big).$$

The bracket is nonnegative by (i), and $\log a\geq 0$. Choosing $r=1/(2Kn)$, which is admissible since $n\geq 1/(2K)$, we obtain (2.1). □

Corollary 2.3. *Under the assumptions of Lemma 2.2, the set $\{\alpha\in\overline{\mathbb{Q}} : h_Z(\alpha)\leq m-\varepsilon\}$ is finite for every $\varepsilon>0$. In particular $\mu^{\mathrm{ess}}(h_Z)\geq m$.*

Proof. Let $\varepsilon>0$ and choose $N\geq 1/(2K)$ with $\lambda(1+\log(2Kn))/n<\varepsilon$ for all $n\geq N$. If $h_Z(\alpha)\leq m-\varepsilon$, then either α is one of the finitely many roots of the polynomials $Q_j(T(x))$, or α has degree less than N by (2.1). In the second case $h(\alpha)\leq h_Z(\alpha)\leq m$, and by Northcott's theorem [1, Theorem 1.6.8] there are only finitely many such α. Hence the set is finite, and $\mu^{\mathrm{ess}}(h_Z)\geq m-\varepsilon$ for every $\varepsilon>0$. □

Remark 2.4. Condition (i) follows from (ii), since $F(z)=2\big(1-\sum_j c_j\deg Q_j-\lambda\big)\log|z|+O(1)$ as $|z|\to\infty$. Lemma 2.2 applies to all algebraic numbers, not only to algebraic integers: the leading coefficient a enters through Lemma 2.1, (2.3) and (2.4), and condition (i) absorbs it in the last step of the proof.

3. Reduction to the unit circle

We call a measure *symmetric* if it is invariant under σ and under complex conjugation, and we write $C_1=\{|z|=1\}$ and $C_2=\{|1-z|=1\}=\sigma(C_1)$.

Lemma 3.1 (Reduction to the half circle)**.** *Let Q_j, c_j, λ, m, ν and K satisfy the assumptions of Lemma 2.2 that precede conditions* (i) *and* (ii)*, with ν symmetric, and assume the strict inequality $\sum_j c_j\deg Q_j+\lambda<1$. Then condition* (ii) *of Lemma 2.2 holds on $\mathbb{C}$ if $F(e^{it})\geq 0$ for all $t\in[0,\pi]$.*

Proof. The functions f, $|Q_j(T(z))|$ and U_ν are invariant under σ and under complex conjugation, and hence so is F. For $|Q_j(T(z))|$ this uses $T(1-z)=T(z)$ and the real coefficients of Q_j. For U_ν it uses $U_\nu(1-z)=U_{\sigma_*\nu}(z)$ and $U_\nu(\bar z)=U_{\bar\nu}(z)$. Since $F(e^{-it})=F(e^{it})$, the hypothesis implies $F\geq 0$ on C_1, and then $F\geq 0$ on C_2 because $F\circ\sigma=F$. The circles C_1 and C_2 cut $\mathbb{C}$ into four open connected regions. On each region f is harmonic, since it equals 0, $\log|z|$, $\log|1-z|$ or $\log|z(1-z)|$ there; across the circles f is only subharmonic. The functions $-c_j\log|Q_j(T(z))|$ and $-2\lambda U_\nu$ are superharmonic on $\mathbb{C}$. Hence F is superharmonic on each region. Since $\log|Q_j(T(z))|$ and U_ν are upper semicontinuous, F is lower semicontinuous on $\mathbb{C}$, with values in $(-\infty,+\infty]$. By the strict inequality,

$$F(z)=2\Big(1-\sum_j c_j\deg Q_j-\lambda\Big)\log|z|+O(1)\;\to\;+\infty\qquad(|z|\to\infty).$$

Since F is lower semicontinuous and $F\geq 0$ on $C_1\cup C_2$, we have $\liminf F(z)\geq 0$ as z tends to a boundary point of a region in $\mathbb{C}$, and $F\to+\infty$ at ∞. By the maximum principle for the subharmonic function $-F$ [18, Theorem 2.3.1], with boundaries taken in the Riemann sphere, $F\geq 0$ on each region, and hence on $\mathbb{C}$. □

4. The certificate

Data. The certificate uses the five polynomials of Table 1. The first two give $Q_1(T(z)) = z(1-z)$ and $Q_2(T(z)) = -(z^2 - z + 1)$, and $Q_4(T(z)) = \Phi_{10}(z)\Phi_{10}(1-z)$, where Φ_{10} is the tenth cyclotomic polynomial. These three polynomials vanish at the exceptional points of Zagier's theorem and at the points where his bound is attained. Flammang's auxiliary function [8, Table 1] contains the same five polynomials, with other weights; we return to it in Section 6.

Table 1. The polynomials $Q_j(w)$ and their weights c_j, rounded. The exact weights are dyadic rationals in the certificate.

j	$Q_j(w)$	c_j
1	w	0.22394147
2	$w-1$	0.17456579
3	w^3+w^2-2w+1	0.0000062581
4	$w^4-2w^3+4w^2-3w+1$	0.0046745171
5	$w^8-3w^7+8w^6-16w^5+26w^4-27w^3+17w^2-6w+1$	0.00079577628

The reference measure is built from a function G on $[0,\pi]$. It is continuous, nonnegative, and linear between the knots $t_k = k\pi/D$, $k = 0,\dots,D$, with $D = 2401$. Its knot values are rationals. Let ρ be the measure on the unit circle with density $G(|s|)$ with respect to ds at e^{is}, $s \in [-\pi,\pi]$, and let $\lambda = \rho(\mathbb{C}) = 2\int_0^\pi G$. This λ is the weight λ of Lemma 2.2. We put $\nu_1 = \rho/\lambda$ and $\nu = \frac12(\nu_1 + \sigma_*\nu_1)$. Then ν is a symmetric probability measure on $C_1 \cup C_2$. On the unit circle the energy term of Lemma 2.2 becomes

$$\lambda\big(2U_\nu(e^{it}) - I(\nu)\big) = U_\circ(t) + U_\times(t) - \frac{I(\rho)+J(\rho)}{2\lambda},$$

where

$$U_\circ(t) = \int_{-\pi}^{\pi} \log|e^{it}-e^{is}|\, G(|s|)\, ds, \qquad U_\times(t) = \int_{-\pi}^{\pi} \log|1-e^{it}-e^{is}|\, G(|s|)\, ds,$$

and $J(\rho) = \iint \log|1-z-y|\, d\rho(z)\, d\rho(y) = 2\int_0^\pi G(t)\, U_\times(t)\, dt$. Indeed $2U_\nu(z) = U_{\nu_1}(z) + U_{\nu_1}(1-z)$, with $\lambda U_{\nu_1}(e^{it}) = U_\circ(t)$ and $\lambda U_{\nu_1}(1-e^{it}) = U_\times(t)$. Since σ is an isometry, $I(\sigma_*\nu_1) = I(\nu_1)$, hence $I(\nu) = \frac12\big(I(\nu_1) + I(\nu_1,\sigma_*\nu_1)\big)$, and $\lambda^2 I(\nu_1) = I(\rho)$, $\lambda^2 I(\nu_1,\sigma_*\nu_1) = J(\rho)$. Finally $f(e^{it}) = \log^+|1-e^{it}| = \log^+(2\sin(t/2))$ for $t \in [0,\pi]$, and $T(e^{it}) = w(t)$ with $w(t) = e^{it} - e^{2it}$. So the function to be checked is

$$F(e^{it}) = \log^+\!\big(2\sin\tfrac{t}{2}\big) - m_0 - \sum_{j=1}^{5} c_j \log|Q_j(w(t))| - U_\circ(t) - U_\times(t) + \frac{I(\rho)+J(\rho)}{2\lambda}. \tag{4.1}$$

The density G vanishes outside $[225\pi/D, 695\pi/D] \cup [932\pi/D, \pi]$. Its maximum, computed directly from the knot values, is less than 1.034λ, so $\nu_1(B(y,r)) \le 1.034\,\pi r$ for $0 < r \le 1$, since the part of the unit circle inside a disc of radius $r \le 1$ is an arc of angle at most π whose chord has length at most $2r$, so its length is at most $2\arcsin r \le \pi r$. The same holds for $\sigma_*\nu_1$, and ν satisfies the hypothesis of Lemma 2.2 with $K = 3.25$.

Formulas on the circle. The Fourier series $\log|2\sin(x/2)| = -\sum_{k\ge1}\cos(kx)/k$ turns the integrals over the unit circle into Clausen-type sums. Let $\mathrm{Cl}_3(x) = \sum_{k\ge1}\cos(kx)/k^3$ and $\mathrm{Cl}_5(x) = \sum_{k\ge1}\cos(kx)/k^5$. Let $g(s) = G(|s|)$, extended 2π-periodically, let $x_p = p\pi/D$ for

$p = 0, \dots, 2D-1$, and let $A_p = g(x_p + \pi/D) - 2g(x_p) + g(x_p - \pi/D)$. Then

$$U_\circ(t) = \frac{D}{\pi} \sum_p A_p \operatorname{Cl}_3(t - x_p), \qquad I(\rho) = -\Big(\frac{D}{\pi}\Big)^2 \sum_{p,q} A_p A_q \operatorname{Cl}_5(x_p - x_q).$$

The second derivative of g in the sense of distributions is $\frac{D}{\pi}\sum_p A_p \delta_{x_p}$. Two integrations by parts give the Fourier coefficients of g, which yields the formula for $U_\circ$, and Parseval's formula gives $I(\rho)$. The first checker below evaluates $U_\times$ and $J(\rho)$ by quadrature with rigorous error bounds. The kernel $\log|1 - e^{it} - e^{is}|$ is singular only where $|1 - e^{is}| = 1$, that is at $s = \pm\pi/3$, and G vanishes near $\pi/3$.

Verification. We verify the following statement with two separately written programs in ball arithmetic. Both use Arb [10] through python-flint [17] and start from the certificate and the formula (4.1) for F; they share no code. The first checker works with a precision of 192 bits.

Proposition 4.1. *For the data of the certificate, $\sum_j c_j \deg Q_j + \lambda < 1$, and the function* (4.1) *satisfies $F(e^{it}) \geq 2.99 \cdot 10^{-9}$ for all $t \in [0, \pi]$.*

The first checker proves Proposition 4.1 as follows. All data are read exactly. The weighted degree sum is exact, λ is enclosed in a ball, and the upper end of the enclosure of the sum is less than 1. The quantities λ, $U_\circ$, $I(\rho)$, $U_\times$ and $J(\rho)$ are computed from the data.

- For $|x| \leq \pi$, Cl_3 and Cl_5 are evaluated from the series

$$\operatorname{Cl}_3(x) = \zeta(3) - \tfrac{3}{4}x^2 + \tfrac{1}{2}x^2 \log|x| - \sum_{n\geq 1} \frac{|B_{2n}|\, x^{2n+2}}{(2n)!\, 2n(2n+1)(2n+2)},$$

$$\operatorname{Cl}_5(x) = \zeta(5) - \tfrac{1}{2}\zeta(3)x^2 + \tfrac{25}{288}x^4 - \tfrac{1}{24}x^4 \log|x| + \sum_{n\geq 1} \frac{|B_{2n}|\, x^{2n+4}}{(2n)!\, 2n(2n+1)(2n+2)(2n+3)(2n+4)},$$

  where the B_{2n} are Bernoulli numbers, with rigorous tail bounds from $|B_{2n}|/(2n)! = 2\zeta(2n)/(2\pi)^{2n}$, after self-tests at points with known values. The double sum for $I(\rho)$ comes from an exact integer autocorrelation of the second differences.
- $U_\times$ is integrated in the form $U_\times(t) = \int_0^\pi (\log|1 - e^{it} - e^{is}| + \log|1 - e^{it} - e^{-is}|) G(s)\, ds$, with the three-point Gauss–Legendre rule on each grid interval where G is not identically zero. On an interval of length Δ the error of the rule is $\Delta^7 \varphi^{(6)}(\xi)/2016000$ for the integrand φ and some ξ in the interval. Since G is linear on the interval, $\varphi^{(6)}$ involves the fifth and sixth derivatives of the logarithmic kernel. Let $\kappa > 0$ be a lower bound for $|2\sin(s/2) - 1|$ on the interval, and let $\beta > 0$ satisfy $e^\beta - 1 < \kappa$. Since $|1 - e^{it} - e^{\pm is}| \geq |2\sin(s/2) - 1|$, the derivative $\partial_s \log(1 - e^{it} - e^{\pm is}) = \mp i e^{\pm is}/(1 - e^{it} - e^{\pm is})$ has modulus at most $e^\beta/(\kappa + 1 - e^\beta)$ on the disc of radius β around any point of the interval, and Cauchy estimates give

$$|\partial_s^k \log(1 - e^{it} - e^{\pm is})| \;\leq\; \frac{(k-1)!\, e^\beta}{\beta^{k-1}(\kappa + 1 - e^\beta)} \qquad (k \geq 1)$$

  for all real t. The same rule in t gives $J(\rho)$. Its error term uses the same estimate for the derivatives in t, at fixed s and with κ taken from the interval of s, and each value of $U_\times$ carries its own error ball.
- The zeros of $Q_j(w(t))$ on $[0, \pi]$ come from the cyclotomic factors of $Q_j(T(z))$, found exactly as common factors with the reciprocal polynomial. They are $t = 0$ for Q_1, $t = \pi/3$ for Q_2 and $t = \pi/5, 3\pi/5$ for Q_4, and a non-cyclotomic common factor would lead to rejection. Around each zero, F is bounded below on 8 balls that cover a window of at least 16 units of the grid $\pi/(D \cdot 2^{14})$ on each side, one-sided at $t = 0$, with an upper

bound for the vanishing factor $|Q_j|$ on each ball. The certified lower bounds of F on the windows at $t = 0$, $\pi/5$, $\pi/3$ and $3\pi/5$ are 2.76, 0.037, 1.99 and 0.046. The window at $\pi/3$ also covers the corner of $\log^+(2\sin(t/2))$ at that point.

- On the rest of $[0,\pi]$ we use the chord bound $F \ge \min(F(u),F(v)) - \frac{v-u}{4}\int_u^v (F'')_+$ on intervals $[u,v]$, with a rigorous upper bound for $(F'')_+$. If Λ is the chord of F over $[u,v]$, then $F-\Lambda = -\int_u^v \gamma(\cdot,s)F''(s)\,ds$ with the Green function $\gamma(x,s) = (\min(x,s)-u)(v-\max(x,s))/(v-u)$, and the bound follows from $0 \le \gamma \le (v-u)/4$. Here F' is absolutely continuous on each interval, since the corner of f and the zeros of the Q_j lie in windows and $U_\circ''(t) = \frac{D}{\pi}\sum_p A_p \log|2\sin\frac{t-x_p}{2}|$ has only integrable logarithmic singularities. These singularities sit at knots, and knots are never interior points of an interval. For a knot x_p at an end of an interval of length Δ, the positive part of the term $-\frac{D}{\pi}A_p\log|2\sin\frac{t-x_p}{2}|$ of F'' has integral at most $\frac{D}{\pi}A_p\Delta(1+\log\frac{\pi}{2\Delta})$ over the interval if $A_p > 0$, since $\sin\vartheta \ge 2\vartheta/\pi$ for $0 \le \vartheta \le \pi/2$, and the term is negative there if $A_p < 0$. The bound for $(F'')_+$ also drops the term $f'' \le 0$ and uses

$$|U_\times''(t)| \;\le\; \int_{-\pi}^{\pi} \frac{|1-e^{is}|}{(2\sin(|s|/2)-1)^2}\,G(|s|)\,ds \;\le\; 0.4491,$$

which follows from $\partial_t^2 \log(1-e^{it}-e^{is}) = e^{it}(1-e^{is})/(1-e^{it}-e^{is})^2$; the last bound is computed piecewise. Intervals are bisected, down to the grid $\pi/(D\cdot 2^{14})$, until the bound is positive. The windows and the initial intervals cover $[0,\pi]$, and bisection preserves the cover.

With 16 processes on a server with 384 threads the first check took 317 seconds.

The second checker was written independently of the first, without consulting its code, from the hypotheses of Lemmas 2.2 and 3.1 and the definition (4.1) of F, which is stored in the certificate. It works with a precision of 128 bits. It evaluates $U_\circ$ by the Clausen sums with its own series for Cl_3, and $U_\times$ by its Fourier series in powers of $\xi = 1-e^{it}$ or of $1/\xi$, with rigorous tails. Near $t = \pi/3$, where $|\xi|$ is close to 1, it uses only an upper bound, which suffices because $U_\times$ enters (4.1) with a minus sign: $U_\times(t) \le U_\circ(\pi/3) + \lambda\log(1+|\xi-e^{-i\pi/3}|/d_0)$, where d_0 is the distance from $e^{-i\pi/3}$ to the support of ρ. This bound follows from $U_\times(t) = U_\rho(\xi)$, $U_\rho(e^{-i\pi/3}) = U_\circ(\pi/3)$ and the triangle inequality $|\xi-y| \le |e^{-i\pi/3}-y|\,(1+|\xi-e^{-i\pi/3}|/d_0)$ for y in the support of ρ. It computes $I(\rho)$ by the Clausen sums, with Cl_5 evaluated through the polylogarithm and cross-checked by the Fourier series, and $J(\rho) = 2\int_0^\pi G\,U_\times$ from the same Fourier series, again with rigorous tails, where the moments $\int G(t)\xi^{\pm k}\,dt$ over each grid interval are computed in closed form by recursion. It factors each $Q_j(z-z^2)$ over $\mathbb{Z}$ into cyclotomic polynomials and a cofactor, and checks that the isolating balls of the complex roots of the cofactor do not meet the unit circle. Its breakpoints are the knots, the zeros on $[0,\pi]$ and the point $\pi/3$, all exact, and it bounds the term of a zero at an end of an interval from below by its value at the other end. On each interval it bounds the rest of F from below by a second-order Taylor expansion at the midpoint, with bisection. The lower bound for F'' on the interval comes from F'' at the midpoint and an enclosure of F'''; the contributions to F'' of terms whose singular point, a knot or a zero, lies within three interval widths are enclosed by their values at the two ends, where they are monotone, and the third derivatives of the logarithms of the polynomial cofactors are enclosed by Taylor expansions at the midpoint. For a knot x_p at an end of an interval of length $L \le 1$, the term $\frac{D}{\pi}A_p\,\mathrm{Cl}_3(t-x_p)$ of $U_\circ$ is kept in the value and the derivative at the midpoint, but its second derivative is left out of the lower bound for F''. Since $|x-s| \le L/2$ in the Taylor remainder, the omitted part of the remainder is at least $-\frac{L}{2}\cdot\frac{D}{\pi}|A_p|\int_0^L |\log(2\sin\frac{u}{2})|\,du \ge -\frac{L}{2}\cdot\frac{D}{\pi}|A_p|(L(1-\log L)+L^3/69)$, and the checker subtracts this amount when $A_p < 0$; when $A_p > 0$ the omitted part is nonnegative, since $2\sin(u/2) < 1$

for $u \le L \le 1$. With 16 processes on the same server it took 45 seconds. Table 2 lists the data and the results of both checks.

TABLE 2. The certificate and its two checks.

Bound m_0	$\frac{87900883718627874129}{351843720888320000000} = 0.2498293375\ldots$
$\sum_j c_j \deg Q_j$; λ; sum	$0.42359031\ldots$; $0.08660960\ldots$; $0.51019992\ldots < 1$
Grid for G	$D = 2401$ intervals, support $[225\pi/D, 695\pi/D] \cup [932\pi/D, \pi]$
$I(\rho)$; $J(\rho)$	$-1.3566155825\ldots \cdot 10^{-3}$; $1.3566156145\ldots \cdot 10^{-3}$
$(I(\rho) + J(\rho))/(2\lambda)$	$\ge 1.84 \cdot 10^{-10}$
Zeros of $Q_j(w(t))$ and the corner	$t = 0,\ \pi/5,\ \pi/3,\ 3\pi/5$
Initial intervals: first; second checker	2,404 (and 4 windows); 2,404
Interval checks: first; second checker	15,588; 2,592
Certified lower bound of F on $[0,\pi]$: first; second checker	$2.99 \cdot 10^{-9}$; $2.23 \cdot 10^{-8}$

Proof of Theorem 1.1. By Proposition 4.1, $\sum_j c_j \deg Q_j + \lambda < 1$ and the function (4.1) is nonnegative on $[0,\pi]$. The Q_j of Table 1 are nonzero polynomials in $\mathbb{Z}[w]$, all c_j are positive, $\lambda > 0$, and the measure ν is symmetric and satisfies the hypothesis of Lemma 2.2 with $K = 3.25$; the function (4.1) is $F(e^{it})$ for this ν. By Lemma 3.1, condition (ii) of Lemma 2.2 holds on $\mathbb{C}$ with $m = m_0$. The finiteness statement and $\mu^{\mathrm{ess}}(h_Z) \ge m_0$ then follow from Corollary 2.3. □

Remark 4.2. The proof is effective. With $\lambda < 0.0867$ and $K = 3.25$, inequality (2.1) gives

$$h_Z(\alpha) \ \ge\ m_0 - \frac{0.0867\,(2.872 + \log n)}{n}$$

for every algebraic number α of degree n that is not a root of the polynomials $Q_j(T(x))$; the condition $n \ge 1/(2K)$ of Lemma 2.2 holds for every $n \ge 1$. For instance $h_Z(\alpha) > 0.24898$ when $n \ge 1000$.

5. Formal verification

We also proved Theorem 1.1 in the Lean 4 proof assistant [12], with the Mathlib library [11].

The formal statement uses Mathlib's definitions. An algebraic number α enters through a polynomial $P \in \mathbb{Z}[x]$ of positive degree that is irreducible in $\mathbb{Z}[x]$ and vanishes at α, so that P is the minimal polynomial of α up to sign, and $h(\alpha) + h(1-\alpha)$ is written as $\big(\log M(P) + \log M(P(1-x))\big)/\deg P$, where M is Mahler's measure as defined in Mathlib. The theorem `ZZ82a.zz_main'` has no hypotheses. It states that for every $\varepsilon > 0$ only finitely many complex numbers α have such a polynomial with $h(\alpha) + h(1-\alpha) \le m_0 - \varepsilon$, where m_0 is the rational number of Theorem 1.1, and `ZZ82a.m0_gt` states that $m_0 > 0.24982933$.

The formal proof reduces the theorem, as in Sections 2 and 3, to the degree condition and to the nonnegativity of the function (4.1) on $[0,\pi]$. Both are verified by a checking program written in Lean without Mathlib, and a Lean proof shows that they hold for every input that the program accepts. The program works in fixed-point arithmetic with tables of values of the trigonometric functions, of the Clausen functions and of the logarithm, evaluates $U_\times$ by Fourier coefficients, and bounds F by chords and on windows around the zeros over a covering of $[0,\pi]$ by 2,783 pieces. Our scripts convert the certificate and the covering into Lean source files; these need not be trusted, since the correctness proof holds for every input that the program accepts.

The check is run in two ways. In `zz_main'` it is evaluated by `native_decide`, which executes compiled code. The trusted base then includes the Lean compiler, the C compiler that it uses and the Lean runtime, and the theorem depends, besides the standard axioms `propext`, `Classical.choice` and `Quot.sound`, on one auxiliary axiom, which states that the compiled check returns true. In `ZZ82a.zz_main_K`, which has the same statement, the Lean kernel evaluates the check, split into 286 modules, and the proof uses only the three standard axioms. In a separate check, `zz_main'` was rebuilt on another machine, and its exported proof was accepted by `leanchecker` and by two further implementations of the Lean kernel, nanoda [13] and lean4lean [3], with the auxiliary axiom admitted. The kernel-only proof was rebuilt from source, and Lean's kernel re-checked each of the 365 modules of its closure. Raising m_0 by $7.9 \cdot 10^{-9}$, the smallest margin over the pieces of the covering, makes the check fail. The Lean files, for Lean 4 version `v4.35.0-rc3` and Mathlib commit `c55e6e786f49`, are in the directory `lean/` of the repository given under Data availability.

6. Numerical observations

The statements in this section are numerical observations; we do not prove them. We took the twenty-four polynomials of Flammang's auxiliary function [8] and their weights from Neehan's transcription [14]; they agree with Table 1 of the version of [8] deposited in HAL (hal-03295880v1). Without the energy term, the linear program over these polynomials has the value 0.248746, which agrees with her bound. With the energy term we solved discretized convex programs on the side of measures, with a linear objective, linear constraints, the quadratic energy constraint and a slack variable dual to condition (i) of Lemma 2.2, which vanished at all computed optima. They range over measures $\theta = \frac{1}{2}(\theta_1 + \sigma_* \theta_1)$, where θ_1 has a piecewise linear density on the unit circle. A probability measure θ with compact support, $\int \log|Q_j(T(z))|\, d\theta \geq 0$ for all j and $I(\theta) \geq 0$ satisfies $\int f\, d\theta \geq m$ for every value m that Lemma 2.2 gives with the polynomials Q_j: integrating (ii) against θ and using $2I(\theta,\nu) - I(\nu) = I(\theta) - I(\theta-\nu) \geq I(\theta)$, we obtain $\int f\, d\theta \geq m + \sum_j c_j \int \log|Q_j(T(z))|\, d\theta + \lambda I(\theta) \geq m$. So the optimal values of these programs are, up to discretization error, upper bounds for every such m. They are about 0.24929 for the two polynomials w and $w-1$ alone and about 0.249833 for the five polynomials of Table 1. With the other nineteen polynomials of the list the value is the same, so these polynomials are not binding at the computed optimal measure.

The certificate comes from the program for the five polynomials in which θ_1 has a density that is linear between the knots $k\pi/2401$ and vanishes at $t = 0$, that is, 2400 hat functions at the interior knots and one at $t = \pi$: up to rounding, ν_1 is its optimal θ_1. With this reference measure fixed, the dual program is a linear program in m, the weights $c_j \geq 0$ and $\lambda \geq 0$, with the constraints $\sum_j c_j \deg Q_j + \lambda \leq 1$ and $F(e^{it}) \geq 0$ at grid points refined by cutting planes. Its value is 0.24982964 (rounded). The knot values of G are λ times those of the density of θ_1, the weights and knot values are the exact values of the double-precision solution, and m_0 is the value of the dual program minus $3 \cdot 10^{-7}$. This margin explains why floating-point evaluations of F have minimum about $2.9 \cdot 10^{-7}$. The certified lower bound $2.99 \cdot 10^{-9}$ is smaller because the bisection stops as soon as the chord bound is positive.

Two exhaustive searches tested polynomials Q against a computed optimal measure θ. The first covered the polynomials of degree d with positive leading coefficient, nonzero constant term and coefficients of absolute value at most $12, 8, 6, 4, 3, 2, 2$ for $d = 2, \ldots, 8$, with the measure computed for all twenty-four polynomials on a coarser grid. The second covered the same kind of polynomials with coefficients of absolute value at most $20, 10, 6, 4$ for $d = 3, \ldots, 6$, with the measure of the certificate. The sign of Q does not affect $\int \log|Q(T(z))|\, d\theta$, and a factor w changes it by less than 10^{-5}, since $\left|\int \log|T(z)|\, d\theta\right| < 10^{-5}$ at both measures. Apart from products of the

polynomials of Table 1, neither search found a polynomial with $\int \log|Q(T(z))|\,d\theta$ below 0.002 in the first search or below 0.001 in the second.

Data availability

The certificate, the two checkers, the Lean proof, and the programs that produced the certificate and the values of Section 6 are archived at doi:10.5281/zenodo.23552588 and available at `https://github.com/tejasvi/zhang-zagier-lower-bound-v3`. The certificate is the file `zz_certificate_D2401.json` of 80,144 bytes. The SHA-256 checksums of the certificate and of the two checkers `check_certificate.py` and `check_independent.py` are

`825ab2837911747b022b49f447b7decc63bc911476f46acded4bd262367eaca5`,
`89246136ab66f0964ad379c94f59c48e5121dfff3db72c1697a45e31cb889fbd`,
`b9574d06d939493fc3f338b3f05d2560d381187dc6886e85bb6e59da56971341`.

References

[1] E. Bombieri and W. Gubler, *Heights in Diophantine geometry*, New Math. Monogr. **4**, Cambridge Univ. Press, Cambridge, 2006.

[2] J. I. Burgos Gil, R. Menares, B. Qu and M. Sombra, *Closing the gap around the essential minimum of height functions with linear programming*, arXiv:2601.18978, 2026.

[3] M. Carneiro, *Lean4Lean: verifying a typechecker for Lean, in Lean*, arXiv:2403.14064, 2024.

[4] D. Davis, P. Ivanisvili, T. Tao and contributors, *Optimization constants in mathematics*, GitHub repository, `https://github.com/teorth/optimizationproblems`, 2026. Constant 82a: `https://teorth.github.io/optimizationproblems/constants/82a.html`, accessed October 2026.

[5] C. Doche, *On the spectrum of the Zhang–Zagier height*, Math. Comp. **70** (2001), 419–430.

[6] C. Doche, *Zhang–Zagier heights of perturbed polynomials*, J. Théor. Nombres Bordeaux **13** (2001), 103–110.

[7] Épi (BakeLab), *Zhang–Zagier C82 record*, Épi results repository, `https://github.com/BakeLab/Epi-results`, directory `records/zhang-zagier/zhang_zagier_c82`, commit `97fa6a4`, accessed October 2026.

[8] V. Flammang, *On the Zhang–Zagier measure*, Int. J. Number Theory **14** (2018), 2663–2671.

[9] M. Grinsztajn, *Code and certification data for an improved C82 Zhang–Zagier bound*, GitHub repository, `https://github.com/maaxgrin/zhang-zagier-c82-bound`, commit `cb88f92`, accessed October 2026.

[10] F. Johansson, *Arb: efficient arbitrary-precision midpoint-radius interval arithmetic*, IEEE Trans. Comput. **66** (2017), 1281–1292.

[11] The mathlib Community, *The Lean mathematical library*, in: Proc. 9th ACM SIGPLAN Int. Conf. Certified Programs and Proofs (CPP 2020), 2020, 367–381.

[12] L. de Moura and S. Ullrich, *The Lean 4 theorem prover and programming language*, in: Automated Deduction – CADE 28, Lecture Notes in Comput. Sci. **12699**, Springer, 2021, 625–635.

[13] `nanoda_lib`, a type checker for Lean 4, `https://github.com/ammkrn/nanoda_lib`.

[14] Neehan, *zhang-zagier-82a*, GitHub repository, `https://github.com/Neehan/zhang-zagier-82a`, file `certificate/flammang_table1.py`, commit `73e2929`, accessed October 2026.

[15] B. J. Orloski, N. Talebizadeh Sardari and A. Smith, *New lower bounds for the Schur–Siegel–Smyth trace problem*, Math. Comp. **94** (2025), 2005–2040; arXiv:2401.03252.

[16] R. Pinho, *Certificate and checkers for C82 $<$ 0.2534449*, GitHub repository, `https://github.com/rpinho/c82-height-certificate`, commit `69d8964`, accessed October 2026.

[17] python-flint: Python bindings for FLINT, version 0.9.0, `https://github.com/flintlib/python-flint`.

[18] T. Ransford, *Potential theory in the complex plane*, London Math. Soc. Student Texts **28**, Cambridge Univ. Press, Cambridge, 1995.

[19] E. B. Saff and V. Totik, *Logarithmic potentials with external fields*, Grundlehren Math. Wiss. **316**, Springer, Berlin, 1997.

[20] C. J. Smyth, *On the measure of totally real algebraic integers*, J. Austral. Math. Soc. Ser. A **30** (1980/81), 137–149.

[21] D. Zagier, *Algebraic numbers close to both* 0 *and* 1, Math. Comp. **61** (1993), 485–491.

[22] S. Zhang, *Positive line bundles on arithmetic surfaces*, Ann. of Math. (2) **136** (1992), 569–587.

[23] S. Zhang, *Positive line bundles on arithmetic varieties*, J. Amer. Math. Soc. **8** (1995), 187–221.

ORCID: `https://orcid.org/0000-0003-4668-3639`