ALGEBRAIC NUMBERS CLOSE TO BOTH 0 AND 1

D. ZAGIER

Dedicated to the memory of D. H. Lehmer

ABSTRACT. A recent theorem of Zhang asserts that

$$H(\alpha) + H(1 - \alpha) \geq C$$

for all algebraic numbers $\alpha \neq 0$, 1, $(1 \pm \sqrt{-3})/2$, and some constant $C > 0$. An elementary proof of this, with a sharp value for the constant, is given (the optimal value of C is $\frac{1}{2} \log(\frac{1}{2}(1+\sqrt{5})) = 0,2406\ldots$, attained for eight values of α) and generalizations to other curves are discussed.

1. LEHMER'S CONJECTURE AND ZHANG'S THEOREM

The *Mahler measure* of an irreducible polynomial

$$(1) \quad f(x) = a_0 x^n + a_1 x^{n-1} + \cdots + a_n = a_0 \prod_{j=1}^{n} (x - \alpha_j) \qquad (a_i \in \mathbb{Z}, \ \gcd\{a_i\} = 1)$$

is defined as

$$(2) \quad M(f) = |a_0| \prod_{|\alpha_j| \geq 1} |\alpha_j| .$$

Clearly, $M(f) \geq 1$, with equality if and only if $f(x)$ is $\pm x$ or a cyclotomic polynomial, by Kronecker's theorem. In a famous article of 1933, Lehmer [4] asked whether in all other cases $M(f)$ is bounded away from 1, and specifically whether $M(f) \geq \alpha_0$, where $\alpha_0 = 1.1762808\ldots$ is the larger real root of the 10th-degree polynomial $x^{10} + x^9 - x^7 - x^6 - x^5 - x^4 - x^3 + x + 1$.

Lehmer's conjecture is still open today, though some progress has been made, both in getting lower estimates for $M(f)$ which are not too close to 1 for n large (cf. [1]) and in finding other particular polynomials with low Mahler measures [2]. The special number α_0 discovered by Lehmer has turned out to have other interesting properties. In particular, there are at least 71 multiplicatively independent multiplicative relations among the numbers $1 - \alpha_0^n$ ($n \in \mathbb{N}$), and there are so-called "ladder relations" among the values of polylogarithm functions at arguments α_0^n up to polylogarithms of at least the 16th order, both very possibly absolute records among all algebraic integers [3].

From a modern point of view, the Mahler measure is a version of the height. If K is an algebraic number field, then the (logarithmic) *height* of a number $\alpha \in K$ relative to K is defined by

$$(3) \qquad\qquad H_K(\alpha) = \sum_v \log\max(|\alpha|_v, 1),$$

where the sum is over all places v of K and the v-adic valuations are normalized as usual in such a way that the product of $|\alpha|_v$ over all Archimedean v is the absolute value of $N_{K/\mathbb{Q}}(\alpha)$, and the product of all $|\alpha|_v$ is equal to 1. It is an easy exercise—trivial if α is an algebraic integer—that $H_K(\alpha)$ is simply the logarithm of the Mahler measure of the minimal polynomial of α if $K = \mathbb{Q}(\alpha)$, so Lehmer's conjecture can be restated as the lower bound $H_K(\alpha) \geq \log\alpha_0 = 0.1623576\ldots$ for all nontorsion points in $K^\times$. The relative height $H_K(\alpha)$ depends on K, but in a very simple way: $H_K(\alpha)$ is simply multiplied by $[K':K]$ if K is replaced by a larger field K'. Hence, one can define a height $H(\alpha)$ independent of K by

$$H(\alpha) = \frac{1}{[K:\mathbb{Q}]}\, H_K(\alpha),$$

i.e., $H(\alpha) = \log M(f)^{1/n}$ if α is a root of the irreducible polynomial (1). For the absolute height, of course, Lehmer's conjecture is false (e.g., $H(\sqrt[n]{2}) = n^{-1}\log 2 \to 0$). Remarkably, however, there is a Lehmer-type statement for the *sum* of the absolute heights of α and $1-\alpha$, namely there is a universal positive lower bound

$$(4) \qquad\qquad H(\alpha) + H(1-\alpha) \geq C > 0$$

for all algebraic numbers except 0, 1, and $(1 \pm \sqrt{-3})/2$ (i.e., except for the four cases where both α and $1-\alpha$ are either 0 or roots of unity). This beautiful result was discovered recently by Shouwu Zhang as a special case of results on the self-intersections in the sense of Arakelov theory of admissible Hermitian line bundles over arithmetic surfaces (Theorems 6.3 and 6.5 of [7], applied to the curve $\{(x,y): x+y = 1, xy \neq 0\}$). The main purpose of this note is to give an elementary proof of (4) and at the same time to find the best possible value for C. Specifically, we shall show that the sum of the heights of α and $1-\alpha$, if nonzero, is universally bounded by the height of the golden ratio, with equality in exactly eight cases:

Theorem 1. *For all algebraic numbers $\alpha \neq 0$, 1, $(1 \pm \sqrt{-3})/2$, we have*

$$(5) \qquad H(\alpha) + H(1-\alpha) \geq \frac{1}{2}\log\frac{1+\sqrt{5}}{2} = 0.2406059\ldots,$$

with equality if and only if α or $1-\alpha$ is a primitive 10th root of unity.

Interestingly enough, the bound $\frac{1}{2}\log\frac{1}{2}(1+\sqrt{5})$ is exactly the same as the minimum of $H(\alpha)$ over all totally real numbers α, found by Schinzel twenty years ago [5].

We will also show how the method applies to give an equally elementary—but in general no longer sharp—proof of most cases of the more general result of Zhang that $H(\alpha)+H(\beta) \geq C(X) > 0$ for all but finitely many pairs of algebraic

numbers (α, β) lying on a curve $X \subset \mathbb{P}^1 \times \mathbb{P}^1$. (Equation (4) is the special case when X is the curve $\alpha + \beta = 1$.)

2. PROOF OF THEOREM 1

The basic idea, which we will illustrate first by proving (4) without optimizing C, is as follows. The absolute height of α, at least if α is an algebraic integer, is essentially the average of the logarithms of the absolute values of the conjugates of α outside the unit circle, so if this height is small, then most of the conjugates of α are within the disk $|z| \leq 1 + \varepsilon$ for some small ε. Since $H(\alpha^{-1}) = H(\alpha)$, the same is true for α^{-1} if this is also integral, so most of the conjugates of α must lie in an annulus $1 - \varepsilon \leq |z| \leq 1 + \varepsilon$. For the same reasons, if $H(1 - \alpha)$ is also small, then most of conjugates must lie in the annulus $1 - \varepsilon \leq |1 - z| \leq 1 + \varepsilon$ and hence must be near one of the intersection points $(1 \pm \sqrt{-3})/2$ of the two circles $|z| = 1$ and $|1 - z| = 1$, but this is impossible since then the norm of $\alpha^2 - \alpha + 1$ would be a rational integer of absolute value less than 1. To make this rigorous and get rid of the assumption that both α and $1 - \alpha$ are units, we observe that there is a universal constant $A \geq 1$ such that

$$(6) \qquad \log|\alpha^2 - \alpha + 1|_v + n_v \leq A \left(|\log|\alpha|_v| + |\log|1 - \alpha|_v| \right)$$

for all places v of K, where

$$(7) \qquad n_v = \begin{cases} 1 & \text{if } v \text{ is real,} \\ 2 & \text{if } v \text{ is complex,} \\ 0 & \text{if } v \text{ is non-Archimedean.} \end{cases}$$

This is obvious for v finite (if α is integral at v, then the right-hand side of (6) equals 0 and the left-hand side is ≤ 0; if not, then $|\alpha^2 - \alpha + 1|_v = |\alpha|_v |1 - \alpha|_v = |\alpha|_v^2 > 1$). If v is Archimedean, then $|\alpha|_v = |\sigma(\alpha)|^{n_v}$ for some embedding σ of K into $\mathbb{C}$, so the claim is that the function

$$z \mapsto \frac{\log|z^2 - z + 1| + 1}{|\log|z|| + |\log|z - 1||} \qquad (z \in \mathbb{C})$$

is bounded above by some constant $A \geq 1$, which is true because the function in question tends to 1 as $|z| \to \infty$, is negative near the two intersection points of the circles $|z| = 1$ and $|z - 1| = 1$, and is continuous elsewhere. Now summing (6) over all v and observing that

$$(8) \qquad \sum_v n_v = [K : \mathbb{Q}], \quad \sum_v \log|\beta|_v = 0, \quad \sum_v |\log|\beta|_v| = 2H_K(\beta)$$

$$(\forall \beta \in K)$$

(the last because $|x| = 2 \max(0, x) - x$), we obtain (4) with $C = 1/(2A)$.

We now give the proof of the sharp inequality (5).

Lemma. *For $z \in \mathbb{C}$, we have*

$$\max(0, \log|z|) + \max(0, \log|1 - z|)$$

$$\geq \frac{\sqrt{5} - 1}{2\sqrt{5}} \log|z^2 - z| + \frac{1}{2\sqrt{5}} \log|z^2 - z + 1| + \frac{1}{2} \log \frac{1 + \sqrt{5}}{2},$$

with equality if and only if z or $1 - z$ equals $e^{\pm \pi i/5}$ or $e^{\pm 3\pi i/5}$.

Proof. Let $f(z)$ denote the difference of the right and left sides of the proposed inequality. Clearly, $f(z) \to -\infty$ as z tends either to infinity or to one of

the points 0, 1, or $(1 \pm \sqrt{-3})/2$, and is continuous elsewhere, so it attains its maximum at some finite point(s). Since f is harmonic off the two circles $|z| = 1$ and $|1 - z| = 1$, the maximum principle tells us that the maxima occur only on these circles, and since the desired inequality is symmetric under $z \mapsto 1 - z$ or $z \mapsto \bar{z}$, we can assume $z = e^{i\theta}$, $0 \leq \theta \leq \pi$. Set $S = 4\sin^2 \theta/2$. We distinguish two cases. If $0 \leq \theta \leq \pi/3$ (i.e., $0 \leq S \leq 1$), then $|1 - e^{i\theta}| \leq 1$ and

$$f(e^{i\theta}) = \frac{\sqrt{5} - 1}{2\sqrt{5}} \log\left(2\sin\frac{\theta}{2}\right) + \frac{1}{2\sqrt{5}}\log(2\cos\theta - 1) + \frac{1}{2}\log\left(\frac{1 + \sqrt{5}}{2}\right)$$

$$= \frac{\sqrt{5} - 1}{4\sqrt{5}}\log(S) + \frac{1}{2\sqrt{5}}\log(1 - S) + \frac{1}{2}\log\left(\frac{1 + \sqrt{5}}{2}\right)$$

and differentiating this with respect to S, we see that the unique maximum on the interval $[0,1]$ is attained at $S = (3 - \sqrt{5})/2$ or $\theta = \pi/5$, where f vanishes. Similarly, in the interval $\pi/3 \leq \theta \leq \pi$ or $1 \leq S \leq 4$, we have $|1 - e^{i\theta}| \geq 1$ and hence

$$f(e^{i\theta}) = \frac{-\sqrt{5} - 1}{2\sqrt{5}}\log\left(2\sin\frac{\theta}{2}\right) + \frac{1}{2\sqrt{5}}\log(1 - 2\cos\theta) + \frac{1}{2}\log\left(\frac{1 + \sqrt{5}}{2}\right)$$

$$= \frac{-\sqrt{5} - 1}{4\sqrt{5}}\log(S) + \frac{1}{2\sqrt{5}}\log(S - 1) + \frac{1}{2}\log\left(\frac{1 + \sqrt{5}}{2}\right),$$

with the unique maximum value 0 at $S = (3 + \sqrt{5})/2$ or $\theta = 3\pi/5$. □

Proof of Theorem 1. The lemma immediately gives

$$\max(0, \log|z|_v) + \max(0, \log|1 - z|_v)$$
$$\geq \frac{\sqrt{5} - 1}{2\sqrt{5}}\log|z^2 - z|_v + \frac{1}{2\sqrt{5}}\log|z^2 - z + 1|_v + \frac{n_v}{2}\log\frac{1 + \sqrt{5}}{2},$$

(with n_v as in (7)) for v Archimedean, since then $|\alpha|_v = |\alpha|^{n_v}$. This inequality is trivial when v is non-Archimedean by the same argument as before (look separately at the cases $|\alpha|_v \leq 1$ and $|\alpha|_v > 1$). Now summing over all v, using (8), and dividing by $[K : \mathbb{Q}]$, we immediately get the inequality (5). □

3. COMPLEMENTS AND GENERALIZATIONS

A. **"Spectrum" of values of** $H(x) + H(1 - x)$. The proof of Theorem 1 can be improved to show that the real number on the right of (5) can be replaced by a strictly bigger one for all α except the eight values mentioned in the theorem. To prove (5), we showed that

(9)
$$\max(0, \log|z|_v) + \max(0, \log|1 - z|_v)$$
$$\geq C\, n_v + C_1 \log|z^2 - z|_v + C_2 \log|z^2 - z + 1|_v,$$

where the values of C_1 $(= (\sqrt{5} - 1)/2\sqrt{5})$ and C_2 $(= 1/2\sqrt{5})$ were chosen to maximize the value of C $(= \frac{1}{2}\log(1 + \sqrt{5})/2)$. Summing this over all v and dividing by n gave (4). If we add a term $C_3 \log|g(z)|_v$ to the right side of (9), where C_3 is a small positive constant and

$$g(z) = \phi_{10}(z)\phi_{10}(1 - z) = z^8 - 4z^7 + 8z^6 - 10z^5 + 11z^4 - 10z^3 + 7z^2 - 3z + 1$$

is the polynomial having the eight exceptional values of α as its roots, then the values of C, C_1, and C_2 can be reoptimized and the new value of C will be slightly larger, since the only places where the previously optimal value of C was attained or nearly attained were those near the eight roots of $g(z)$, where the expression $C_3 \log |g(z)|_v$ becomes very negative. This gives (4) with the new, larger value of C for all algebraic numbers α except for the twelve roots of $(z^2 - z)(z^2 - z + 1)g(z) = 0$. One could easily specify numerical values of the constants C_i and C which work. However, it is no longer clear whether the equality (after summing over places) is ever attained and hence whether we can repeat the process by adding yet another term $C_4 \log |h(z)|_v$ to (9), where h vanishes at the cases of equality. The question therefore remains open whether there is a whole "spectrum" of values

$$c_0 = 0 < c_1 = \frac{1}{2} \log \frac{1 + \sqrt{5}}{2} < c_2 < \cdots$$

such that $H(\alpha) + H(1 - \alpha) = c_j$ for some finite collection of algebraic numbers α and $H(\alpha) + H(1 - \alpha) > \limsup c_j$ for all other $\alpha \in \overline{\mathbb{Q}}$. For the related problem of low values of $H(\alpha)$ for α totally real, where as already mentioned the minimal value is attained and is the same as for our problem, an investigation of the spectrum was carried out, and its first four isolated points determined, by C. J. Smyth [6].

B. **Projective version of Theorem 1.** The quantity $H(\alpha)$ studied so far is in fact the height of α considered as a point of the projective line $\mathbb{P}^1$, so that Theorem 1 can be considered as a statement about the behavior of the height function $(x, y) \mapsto H(x) + H(y)$ of $\mathbb{P}^1 \times \mathbb{P}^1$ restricted to the curve $x + y = 1$ in $\mathbb{P}^1 \times \mathbb{P}^1$. It is perhaps more natural, and certainly more symmetric, to use instead the height function

$$H_{\mathbb{P}^2}(x : y : z) = \frac{1}{[K : \mathbb{Q}]} \sum_v \log \max(|x|_v, |y|_v, |z|_v)$$

$$((x : y : z) \in \mathbb{P}^2(K), \ K \subset \overline{\mathbb{Q}})$$

of the projective plane (note that this is independent of the choice of projective coordinates x, y, z by virtue of the product formula) and consider its restriction to the curve $x + y + z = 0$. The inequalities

$$\max(|x|, 1) + \max(|y|, 1) \geq \max(|x|, |y|, 1) \geq \frac{\max(|x|, 1) + \max(|y|, 1)}{2}$$

imply that

$$H(x) + H(y) \geq H_{\mathbb{P}^2}(x : y : 1) \geq \tfrac{1}{2}(H(x) + H(y))$$

and hence that the minimum of $H_{\mathbb{P}^2}$ on the curve $x + y + z = 0$ is somewhere between $C/2$ and C, where C is the optimal constant in equation (4). We can find the exact minimum by the same method as that used in §2. The analogue of the lemma there is the inequality

(10)
$$\log \max(|x|, |y|, |z|) \geq \frac{1}{2} \log \theta + \frac{1}{6\theta - 4} \log |xy + yz + zx|$$
$$+ \frac{\theta - 1}{3\theta - 2} \log |xyz|$$

for x, y, $z \in \mathbb{C}$ with $x + y + z = 0$, where $\theta = 1.46557\ldots$ is the real root of $\theta^3 - \theta^2 - 1 = 0$. (We omit the proof, which is similar to the one before.) This leads by the same argument as before to

Theorem 1′. *Let $C \subset \mathbb{P}^2$ be the curve $x + y + z = 0$. Then*

$$H_{\mathbb{P}^2}(P) \geq \tfrac{1}{2}\log\theta = 0.1911225\ldots \tag{11}$$

for all $P \in C(\overline{\mathbb{Q}})$ except for the five points

$$P = (1 : -1 : 0), \quad (1 : 0 : -1), \quad (0 : 1 : -1), \quad (1, \omega, \omega^2), \quad (1, \omega^2, \omega)$$

(ω = nontrivial cube root of unity) for which $H_{\mathbb{P}^2}(P)$ vanishes. The bound in (11) is sharp and is attained for exactly six values, namely at the points $P = (1 : \alpha - 1 : -\alpha)$, where α is a root of the equation $\alpha^6 - 3\alpha^5 + 7\alpha^4 - 9\alpha^3 + 7\alpha^2 - 3\alpha + 1 = 0$.

The six exceptional points of the theorem are the intersections of the curve $\sigma_1 = 0$ and $\sigma_2^3 = \sigma_3^2$ on $\mathbb{P}^2$, where σ_j denotes the jth elementary symmetric polynomial in x, y, and z. As in A above, we can improve the result by adding a term to the right-hand side of (10) which tends to $-\infty$ as $(x : y : z)$ approaches one of these six points, i.e., by replacing (10) by

$$\log\max(|x|, |y|, |z|) \geq C' + C'_1\log|\sigma_2| + C'_2\log|\sigma_3| + C'_3\log|\sigma_2^3 - \sigma_3^2| \qquad (\sigma_1 = 0)$$

for some small positive number C'_3 and appropriately chosen real numbers C', C'_1, and C'_2 with $2C'_1 + 3C'_2 + 6C'_3 = 1$. Numerical values which work are

$$C' = 0.2024850\ldots, \quad C'_1 = 0.1801634\ldots, \quad C'_2 = 0.1714179\ldots,$$
$$C'_3 = 0.0209032\ldots$$

(we do not describe the optimization process which produces these numbers), showing that the right-hand side of (11) can be replaced by the slightly larger real number C' for all α except for the six values mentioned in the theorem. Again, however, we do not know if there is a sequence of such best possible results leading to a "spectrum" of values of $H_{\mathbb{P}^2}(P)$, $P \in C(\overline{\mathbb{Q}})$.

C. Other curves. Finally, we can replace the special curve $\{x + y = 1\} \subset \mathbb{P}^1 \times \mathbb{P}^1$ or $\{x + y + z = 0\} \subset \mathbb{P}^2$ by an arbitrary curve X in $\mathbb{P}^1 \times \mathbb{P}^1$ or $\mathbb{P}^2$. We consider the case of $X \subset \mathbb{P}^1 \times \mathbb{P}^1$ given by the vanishing of a polynomial $F(x, y)$ with rational coefficients. We assume that the curve X^* given by $F(x^{-1}, y^{-1}) = 0$ has no component in common with X, and pick a polynomial $G(x, y) \in \mathbb{Z}[x, y]$ which vanishes at all the intersection points of X and X^* as well as at all the intersection points of X with $x = 0$, $y = 0$. For instance, if X is the curve $x + y = 1$, then X^* is the curve $x + y = xy$ and we can take $G(x, y) = xy(xy - 1)$. Now it is clear that $\log|G(x, y)|_v + n_v$ is bounded by $A(|\log|x|_v| + |\log|y|_v|)$ for all $(x, y) \in X(K)$ and all places v of K, if A is sufficiently large. (For v finite this is trivial because $n_v = 0$ and $|G(x, y)|_v$ is bounded by a polynomial in $|x|_v$ and $|y|_v$; and for v Archimedean it follows as before by noting that the quotient $(1 + \log|G|)/(|\log|x|| + |\log|y||)$ is bounded above for $|x| + |y| \to \infty$ because G is a polynomial, and is negative near all its discontinuities because $|G|$ is small whenever both $|x|$ and $|y|$ are near 1.) Summing over all places v and dividing by $[K : \mathbb{Q}]$ as usual, we obtain a universal lower bound $H(x) + H(y) \geq 1/(2A)$ for all $(x, y) \in X(\overline{\mathbb{Q}})$ except for

the finitely many common roots of F and G. It is clear that the method could also be applied to curves in $\mathbb{P}^2$ and in other situations.

4. ACKNOWLEDGMENT

I would like to thank Noam Elkies for several useful discussions and suggestions. In particular, he found the example which turned out to minimize $H(\alpha) + H(1 - \alpha)$ by computing $H(1 - \alpha)$ for roots of unity α of low order. The suggestion of looking at the projective plane version of Theorem 1, as carried out in B of §3, is also due to him. I also thank the referee for pointing out the connection with the results of [5] and [6].

BIBLIOGRAPHY

1. M. J. Bertin and M. Pathiaux-Delefosse, *Conjecture de Lehmer et petits nombres de Salem*, Queen's Papers in Pure and Appl. Math. **81** (1989), Queen's University, Kingston, Canada.

2. D. W. Boyd, *Reciprocal polynomials having small measure* I, II, Math. Comp. **35** (1980), 1361–1377 and **53** (1989), 355–357, S1–S6.

3. H. Cohen, L. Lewin, and D. Zagier, *A sixteenth order polylogarithm ladder*, Experiment. Math. **1** (1992).

4. D. H. Lehmer, *Factorization of certain cyclotomic functions*, Ann. of Math. **34** (1933), 461–479.

5. A. Schinzel, *On the product of the conjugates outside the unit circle of an algebraic number*, Acta Arith. **24** (1973), 385–399.

6. C. J. Smyth, *On the measure of totally real algebraic integers* I, II, J. Austral. Math. Soc. **30** (1980), 137-149 and Math. Comp. **37** (1981), 205–208.

7. S. Zhang, *Positive line bundles on arithmetic surfaces*, preprint, Princeton, 1992.

MAX-PLANCK-INSTITUT FÜR MATHEMATIK, GOTTFRIED-CLAREN-STRASSE 26, 5300 BONN 3, GERMANY
E-mail address: don@antigone.mpim-bonn.mpg.de