Exceptional Units and Numbers of Small Mahler Measure

Joseph H. Silverman

CONTENTS

Let α be a unit of degree d in an algebraic number field, and assume that α is not a root of unity. We conduct a numerical investigation that suggests that if α has small Mahler measure, there are many values of n for which $1-\alpha^n$ is a unit and also many values of m for which $\Phi_m(\alpha)$ is a unit, where Φ_m is the m-th cyclotomic polynomial. We prove that the number of such values of n and m is bounded above by $O(d^{1+0.7/\log\log d})$, and we describe a construction of Boyd that gives a lower bound of $\Omega(d^{0.6/\log\log d})$.

INTRODUCTION

An algebraic number a is an *exceptional unit* if both a and $1-a$ are algebraic units. Siegel proved that there are only finitely many exceptional units in any number field, and there is a large literature devoted to proving quantitative and effective bounds for the set of exceptional units. For example, Evertse [1984] has proved that a number field of degree d has at most $3\cdot 7^{3d}$ exceptional units.

Let α be an algebraic unit of degree d that is not a root of unity. In this article we investigate how many powers of α can be exceptional units. Thus we will be looking at solutions of the unit equation

$$u+v=1 \qquad (u \text{ and } v \text{ units}) \tag{0.1}$$

in which the variable u is chosen from a cyclic subgroup of the group of units.

Let $E(\alpha)$ be the number of values of $n\geq 1$ such that α^n is an exceptional unit. Our main theoretical result will imply that there is an absolute and effectively computable constant c such that

$$E(\alpha)\leq cd^{1+0.7/\log\log d}. \tag{0.2}$$

Research partially supported by NSF grant DMS-9121727.

Keywords: units, Mahler measure, unit equation

AMS Subject Classification: 11R27, 11D61, 11R06, 11J68

© A K Peters, Ltd.
1058-6458/96 $0.50 per page

So in this special situation we are able to reduce Evertse's exponential bound to a bound that grows only a little faster than linearly.

A power α^n is an exceptional unit if and only if $1-\alpha^n$ is a unit. Now $1-\alpha^n$ factors as

$$1-\alpha^n = \prod_{m|n} \Phi_m(\alpha), \tag{0.3}$$

where Φ_m is the m-th cyclotomic polynomial. Thus $1-\alpha^n$ is an exceptional unit if and only if $\Phi_m(\alpha)$ is a unit for all $m|n$. Let $U(\alpha)$ denote the number of values of $m \geq 1$ such that $\Phi_m(\alpha)$ is a unit. Stewart [1977] has shown that, if $\Phi_m(\alpha)$ is a unit, then $m \leq e^{452}d^{67}$. This clearly gives $O(d^{67})$ as an upper bound for $U(\alpha)$. The following result gives a bound of the form $O(d^{1+o(1)})$ (see Theorem 4.1 for something stronger).

Theorem 0.1. *Let α be an algebraic unit of degree $d \geq 2$ that is not a root of unity. There is an absolute and effectively computable constant c such that*

$$U(\alpha) \leq cd^{1+0.7/\log\log d}.$$

We now briefly describe the contents of this article. We begin in Section 1 with some motivation for why one might be interested in studying the set of n such that α^n is an exceptional unit and the set of m such that $\Phi_m(\alpha)$ is a unit. We also use a method from [Blanksby and Montgomery 1971] to show heuristically why one might expect these sets to be large if α is a number of small Mahler measure. We follow up this observation in Section 2 with a numerical investigation of some specific α's catalogued by David Boyd [1977; 1978; 1990]. For example, we will exhibit an α of degree 18 such that $E(\alpha) \geq 25$, an α of degree 28 such that $U(\alpha) \geq 77$, and an α of degree 26 such that α^n is an exceptional unit for all $n = 1, \ldots, 10$.

The data we collect will suggest a possible general upper bound of the form

$$U(\alpha) \geq A\frac{\log d}{\log M(\alpha)} + B,$$

where A and B are absolute constants and $M(\alpha)$ is the Mahler measure of α. Unfortunately, this guess turns out to be much too ambitious. David Boyd (in a private communication) has pointed out that $U(\alpha)$ can grow more rapidly than any power of $\log d$, and further that no upper bound of the form $o(d)/\log M(\alpha) + O(1)$ is possible. We will describe Boyd's constructions in Section 5.

After the numerical results of Section 2, we turn in Section 3 to some preliminary inequalities that are needed for the proof of our main result. In Section 4 we prove something stronger than Theorem 0.1. The proof uses an upper bound for m essentially found in [Stewart 1977], a lower bound for the Mahler measure [Dobrowolski 1979], an elementary but involved estimate for values of cyclotomic polynomials (Proposition 3.3), and a sort of "supergap principle" (Lemma 4.3) that may be of some independent interest. Finally, in Section 5 we present Boyd's results.

We close this introduction with two remarks. First, our estimate (0.2) is really a bound for the number of solutions of (0.1) in which the variable u is chosen from a group $\Gamma \subset \mathbb{C}^*$ of rank 1. More generally, one can ask for a bound for the number of solutions of (0.1) with $u \in \Gamma_1$ and $v \in \Gamma_2$, where $\Gamma_1, \Gamma_2 \subset \mathbb{C}^*$ are groups of ranks r_1 and r_2. See, for example, the recent work [Bombieri et al. 1994] in which the authors use a supergap (or cluster) principle to prove their bounds. If $r_1 = 1$ and Γ_2 is the full unit group, (0.2) gives a bound that is almost linear in r_2. However, if $r_1 \geq 2$, the best known bounds are exponential in r_2. One possible explanation for this difference is the existence of the factorization (0.3) of $1-\alpha^n$ in the rank-1 case. Unfortunately, there is no analogous factorization of $1-\alpha_1^{n_1}\alpha_2^{n_2}$.

Second, there are natural elliptic analogues to the questions studied in this article. The analogue of an exceptional unit is an integral point on an elliptic curve. Thus let E/K be an elliptic curve defined over a number field, and let $P \in E(\bar{K})$ be a nontorsion point of degree d. One can ask for an upper bound for the number of integers n such that nP is an integral point on E, and similarly one can ask for the number of values of m such that $\Phi_{E,m}(P)$ is an S-unit, where $\Phi_{E,m}$ is the m-division polynomial of E and S is the set of primes of bad reduction for E. There are bounds known in terms of various quantities associated to K, E, and d (see [Hindry and Silverman 1988], for example), but all of them are worse than polynomial in d. It seems likely that the methods of this article will give a bound of the form $c(E/K)d^{3+o(1)}$. The exponent 3 reflects the best result currently known for the elliptic Lehmer conjecture [Masser 1989]. For elliptic curves with nonintegral j-invariant, the improved estimate in [Hindry and Silverman 1990] would probably yield an upper bound of the form $c(E/K)d^{2+o(1)}$. Similarly, for elliptic curves with complex multiplication, [Laurent 1983] could probably be used to reduce this to $c(E/K)d^{1+o(1)}$. We will not deal with the elliptic case in this article.

1. Exceptional units in cyclic groups

The *Mahler measure* $M(\alpha)$ of an algebraic integer α is defined by

$$M(\alpha)=\prod\max\{|\alpha|,1\},$$

where the product is over all embeddings of $\mathbb{Q}(\alpha)$ into $\mathbb{C}$. Clearly $M(\alpha)\geq 1$ for all α. An elementary result of Kronecker [Kronecker 1857] says that $M(\alpha)=1$ if and only if α is a root of unity. Equivalently, if α is not a root of unity, at least one of its conjugates must lie outside the unit circle.

Now consider the following dubious piece of logic suggested by Kronecker's theorem.

1. If $M(\alpha)$ is close to 1, then α should look like a root of unity.
2. If ζ is a root of unity, $1-\zeta^n$ tends to be a unit (or at least a p-unit, if ζ is a p^k-th root of unity).
3. Ergo, if $M(\alpha)$ is close to 1, then $1-\alpha^n$ should be a unit for many values of n.

This suspicious reasoning will be numerically vindicated in the next section. For example, a root γ_1 of the polynomial

$$\begin{aligned}x^{18}&+x^{17}+x^{16}+x^{15}-x^{12}-x^{11}-x^{10}\\&-x^9-x^8-x^7-x^6+x^3+x^2+x+1\end{aligned}\tag{1.1}$$

has Mahler measure approximately 1.188368147, and there are at least 25 values of n for which $1-\gamma_1^n$ is a unit.

A famous question of Lehmer [1933] is whether there exists an absolute constant $\varepsilon>0$ such that α is a root of unity whenever $M(\alpha)<1+\varepsilon$. For partial results on this problem, see [Blanksby and Montgomery 1971; Dobrowolski 1979; Mignotte 1977; Silverman 1994; Smyth 1971; Stewart 1978]. In particular, [Blanksby and Montgomery 1971] can be used to establish a connection between Lehmer's question and the powers of α that are exceptional units. This was our original motivation for studying this question. We briefly sketch the argument.

Let α be an algebraic unit of degree d with conjugates $\alpha_1,\dots,\alpha_d$. For each $1\leq i\leq d$, let $\beta_i=\alpha_i$ if $|\alpha_i|\leq 1$, and let $\beta_i=\alpha_i^{-1}$ otherwise. Let $N:\mathbb{Q}(\alpha)\to\mathbb{Q}$ denote the norm. For each integer $K\geq 1$ we consider the sum

$$\begin{aligned}S(\alpha,K)&:=\sum_{k=1}^{K}\Bigl(1-\frac{k}{K+1}\Bigr)\log\bigl|N(1-\alpha^k)\bigr|\\&=\sum_{k=1}^{K}\Bigl(1-\frac{k}{K+1}\Bigr)\sum_{j=1}^{d}\log|1-\alpha_j^k|\\&=\sum_{k=1}^{K}\Bigl(1-\frac{k}{K+1}\Bigr)\\&\qquad\times\Biggl(\sum_{j=1}^{d}\log\max\bigl\{|\alpha_j^k|,1\bigr\}+\log|1-\beta_j^k|\Biggr)\\&=K^2\log M(\alpha)\\&\qquad+\sum_{j=1}^{d}\sum_{k=1}^{K}\Bigl(1-\frac{k}{K+1}\Bigr)\log|1-\beta_j^k|.\end{aligned}$$

The final inner sums can be bounded using the Fourier averaging technique described in [Blanksby and Montgomery 1971]. One ends up with an estimate of the form

$$\log M(\alpha) \geq \frac{1}{K^2}\bigl(S(\alpha,K) - \tfrac{1}{2}d(\log(K+1)+1)\bigr) \tag{1.2}$$

[Silverman 1994, Proposition 2.3].

Thus one approach to answering Lehmer's question is to find a (small) value of K with the property that $S(\alpha,K)$ is large [Silverman 1994]. But $S(\alpha,K)$ will be large precisely when $N(1-\alpha^k)$ is large for many values of $1 \leq k \leq K$, so the worst possible case is when many of the α^k are exceptional units. Conversely, inequality (1.2) says that $S(\alpha,K)$ cannot be too large if $M(\alpha)$ is close to 1, and this in turn suggests that many of the α_k are exceptional units. So (1.2) helps to justify our earlier piece of dubious logic. Unfortunately, it does not appear that (1.2) by itself is strong enough to actually prove the existence of many exceptional units.

As mentioned in the introduction, the factorization of $1-x^n$ as a product of cyclotomic polynomials means that it is more natural to look at values of m for which $\Phi_m(\alpha)$ is a unit. Thus, if $1-\alpha$ is not a unit, $1-\alpha^n$ will never be a unit. But one might hope that the chances of the $\Phi_m(\alpha)$ being units are independent events in some (admittedly vague) probabilistic sense. As a numerical example, consider a root γ_2 of the polynomial

$$x^{18} - x^{12} - x^{11} - x^{10} - x^9 - x^8 - x^7 - x^6 + 1.$$

The Mahler measure of γ_2 is 1.2527759374, approximately. Evaluating the polynomial at $x=1$ shows that $N(1-\gamma_2) = -5$, so $1-\gamma_2^n$ is never a unit. On the other hand, there are at least 58 values of m for which $\Phi_m(\gamma_2)$ is a unit.

We also note that rearranging the sum defining $S(\alpha,K)$ gives

$$S(\alpha,K) = \sum_{m=1}^{M}\biggl(\sum_{k=1}^{[M/m]}\Bigl(1-\frac{km}{M+1}\Bigr)\biggr)\log\bigl|N\Phi_m(\alpha)\bigr|,$$

where the inner sum is approximately $K/2m$. So (1.2) also suggests that many values of $\Phi_m(\alpha)$ will be units if $M(\alpha)$ is close to 1.

As a final motivation for studying the question of how many values of m give units $\Phi_m(\alpha)$, we mention the recent article [Cohen et al. 1992], where the authors consider the largest real root γ_3 of the polynomial

$$x^{10} + x^9 - x^7 - x^6 - x^5 - x^4 - x^3 + x + 1. \tag{1.3}$$

(This γ_3 has the smallest known Mahler measure greater than 1—approximately 1.1762808.) They find that there are 66 values of $m < 1000$ for which $\Phi_m(\gamma_3)$ is a unit. They use these values, together with a few additional multiplicative relations among the other $\Phi_m(\gamma_3)$'s, as the starting point in climbing a polylogarithm ladder. The existence of so many units allows them to discover and numerically verify several relations among polylogarithms of order sixteen. They suggest that this "is quite possibly the highest order occurring for any algebraic number" because γ_3 probably has the smallest Mahler measure strictly greater than 1. However, we observe that for a root γ_1 of (1.1), there are at least 75 values of m for which $\Phi_m(\gamma_1)$ is a unit, so it might be worthwhile investigating polylogarithm ladders for γ_1.

2. NUMERICAL RESULTS

In this section we investigate some specific numbers of small Mahler measure. David Boyd [1977; 1978; 1990] has computed tables of such numbers. We begin with his list of small Salem numbers [Boyd 1977; 1978]. (A number α is a *Salem number* if $M(\alpha) = \alpha$ and if some conjugate of α lies on the unit circle.) The relevant data are given in Table 1.

Boyd [1990] gives for each even degree $4 \leq d \leq 40$ the number with smallest known Mahler measure. (For $d \leq 20$, he verifies that the number listed is actually the smallest.) In Table 2 we reproduce Boyd's list, together with the largest value of m and the number of values of m such that $\Phi_m(\alpha)$ is a unit, where we check all $m \leq 500$.

k	d	$\alpha = M(\alpha)$	A	B	C	D
1	10	1.176281	74	22	286	65
2	18	1.188368	74	25	210	72
3	14	1.200027	74	20	260	68
4	14	1.202617	74	20	300	64
5	10	1.216392	43	16	294	53
6	18	1.219721	91	19	294	64
7	10	1.230391	39	11	186	48
8	20	1.232614	73	13	300	64
9	22	1.235665	91	22	240	63
10	16	1.236318	67	14	210	56
11	26	1.237505	98	17	290	66
12	12	1.240726	47	11	240	49
13	18	1.252776	—	—	228	58
14	20	1.253331	—	—	252	56
15	14	1.255094	41	16	192	50
16	18	1.256221	47	15	294	54
17	24	1.260104	27	9	204	56
18	22	1.260284	61	16	270	57
19	10	1.261231	46	13	156	39
20	26	1.263038	74	19	250	59
21	14	1.267296	59	13	264	47
22	8	1.280638	23	8	140	35
23	26	1.281691	—	—	300	57
24	20	1.282496	41	10	210	55
25	18	1.284617	—	—	248	50
26	26	1.284747	91	16	280	54
27	30	1.285099	—	—	280	55
28	30	1.285122	85	12	266	62
29	30	1.285186	—	—	252	59
30	26	1.285197	46	14	294	58
31	44	1.285199	127	11	300	62
32	30	1.285235	83	16	264	57
33	34	1.285409	98	16	246	54
34	18	1.286396	73	14	180	47
35	26	1.286730	74	14	234	54
36	24	1.291741	—	—	162	46
37	20	1.292039	—	—	300	49
38	10	1.293486	39	11	210	36
39	18	1.295675	61	13	240	46
40	22	1.296421	61	13	210	49
41	28	1.296821	—	—	276	52
42	26	1.299745	53	16	168	52

TABLE 1. Small Salem numbers and their degrees d, from [Boyd 1977; 1978]. For each number we have also computed the largest $n \leq 300$ with $1-\alpha^n$ a unit (column A), the number of $n \leq 300$ with $1-\alpha^n$ a unit (a lower bound for $E(\alpha)$; column B), the largest $m \leq 300$ with $\Phi_m(\alpha)$ a unit (column C), and the number of $m \leq 300$ with $\Phi_m(\alpha)$ a unit (a lower bound for $U(\alpha)$; column D).

d	$M(\alpha)$	A	B
4	1.722084	22	6
6	1.401268	84	18
8	1.280638	210	35
10	1.176281	360	66
12	1.227786	170	49
14	1.200027	260	68
16	1.224279	420	57
18	1.188368	290	75
20	1.212824	396	67
22	1.205020	390	70
24	1.218855	408	70
26	1.223777	280	67
28	1.207950	330	77
30	1.225620	450	71
32	1.236198	480	65
34	1.229999	280	73
36	1.229483	462	73
38	1.223447	360	76
40	1.236250	360	70

TABLE 2. Small reciprocal numbers and their degrees d, from [Boyd 1980]. For each number we have computed the largest $m \leq 500$ with $\Phi_m(\alpha)$ a unit (column A) and the number of $m \leq 500$ with $\Phi_m(\alpha)$ a unit (a lower bound for $U(\alpha)$; column B).

The data in columns A–D of Tables 1 and 2 were calculated using PARI [Batut et al. 1993]. Note that $1-\alpha^n$ and $\Phi_m(\alpha)$ are units if and only if their norms equal ± 1. If $f(X)$ is the minimal polynomial of α over $\mathbb{Q}$, these norms can be computed as the resultants $\operatorname{Res}(f(X), 1 - X^n)$ and $\operatorname{Res}(f(X), \Phi_m(X))$. PARI is well suited to perform these computations, although for large values of m it turns out to be slightly quicker to compute

$$\operatorname{Res}(f(X), \Phi_m(X)) = \prod_{n|m} \operatorname{Res}(f(X), 1-X^n)^{\mu(m/n)},$$

where μ is the Möbius function. The advantage of this formula is that one can compute $X^n \bmod f(X)$ quite rapidly by using successive squaring.

Before we begin analyzing the data in our tables, we want to point out a few individual entries. The most famous, of course, is the first entry in Table 1, which also appears as the degree-10 entry in Table 2. It is widely believed that this number has the smallest Mahler measure strictly greater than 1. For this number there are at least 66 values of m for which $\Phi_m(\alpha)$ is a unit. (Note that Table 1, which only refers to values of m up to 300, missed the value $m = 360$ included in Table 2.)

The large number of units $\Phi_m(\alpha)$ for this particular α was exploited in [Cohen et al. 1992] to produce relations between polylogarithms of order 16. It might be interesting to perform similar computations using entry $k = 2$ in Table 1 with its 72 unit values, or using the entries of degrees 18 and 28 in Table 2, which have 75 and 77 unit values. These α's might allow the construction of polylogarithm relations of even higher order.

Another interesting entry is $k = 20$ in Table 1, which has the property that $1 - \alpha^n$ is a unit for all $1 \le n \le 10$. This is currently the longest known string of consecutive powers being exceptional units.

We are now going to try to interpret the data in our tables, especially the question of how the last column is related to the degree and Mahler measure of the number. For any number α, we will denote by $d(\alpha) = [\mathbb{Q}(\alpha) : \mathbb{Q}]$ the degree of α. In the course of proving our main theorem in Section 4, we will prove an inequality slightly weaker than

$$U(\alpha) \le c_1 d(\alpha) + c_2 \frac{d(\alpha)}{\log M(\alpha)} + c_3, \tag{2.1}$$

and it seems possible that our method is capable of producing exactly this estimate. See the proof of Theorem 4.1, especially (4.9).

At first glance this inequality seems reasonable for the data in Tables 1 and 2, since increasing the degree leads to additional units, and increasing the Mahler measure leads to fewer units. However, a second look makes it clear that a linear dependence in $d(\alpha)$ grows much too rapidly. In fact, the growth levels out quickly enough to suggest that $\log d(\alpha)$ might be more appropriate. Unfortunately, as shown by the constructions of Boyd described in Section 5, our data for α's of small Mahler measure turns out to be misleading. There do not exist bounds for $U(\alpha)$ of the form

$$U(\alpha) \le c_4 (\log d(\alpha))^N$$

or of the form

$$U(\alpha) \le c_5 \frac{d(\alpha)\psi\bigl(d(\alpha)\bigr)}{\log M(\alpha)}$$

for any exponent N and function $\psi(d)$ that tends to 0 as $d \to \infty$.

On the other hand, we would certainly expect that, for a fixed degree, the size of $U(\alpha)$ should decrease as the Mahler measure $M(\alpha)$ increases. This leads us to ask the following question.

Question 2.1. Are there absolute constants A and B such that

$$U(\alpha) \le A \frac{d}{\log M(\alpha)} + B$$

for all $d \ge 1$ and all algebraic units α of degree d that are not roots of unity?

Notice that one consequence of the inequality in this question is that, if $M(\alpha) \ge e^d$, there is a bound C for $U(\alpha)$ that is completely independent of α. But for any given d, there are only finitely many α's of degree d and $M(\alpha) \le e^d$, so for any given d there would be only finitely many α's with $U(\alpha) > C$.

3. PRELIMINARY ESTIMATES

In this section we will prove some preliminary estimates needed for the proof of our main theorem. We set the following (mostly) standard notation. See [Apostol 1976] for further details.

$\mu(n)$ Möbius μ function, equal to $(-1)^k$ if n is a product of k distinct primes, otherwise equal to 0.

$\sigma_0(m)$ the number of divisors of m.

$I(n)$ the identity function for Dirichlet multiplication, equal to 1 if $n = 1$, otherwise equal to 0.

$\boldsymbol{\mu}_n$ the set of n-th roots of unity.

$\boldsymbol{\mu}^*_m$ the set of primitive m-th roots of unity.

$\Phi_m(z)$ the m-th cyclotomic polynomial, equal to $\prod_{\zeta\in\boldsymbol{\mu}^*_m}(z-\zeta) = \prod_{n|m}(z^n-1)^{\mu(m/n)}$.

$\varphi(m)$ Euler's totient function, equal to $\deg\Phi_m$.

Lemma 3.1. *Let $n \ge 1$ be an integer, and let $w, \zeta \in \mathbb{C}$ satisfy $|w| \le 1$ and $|\zeta| = 1$. Then*

$$|w-\zeta| \ge \frac{1}{2}\left|\frac{w}{|w|}-\zeta\right|.$$

Proof. Replacing w by ζw and canceling $|\zeta|=1$, we may assume that $\zeta = 1$. Write $z = x+iy = re^{i\theta}$ with $|\theta| \le \pi$. If $|\theta| \ge \pi/2$, then $x<0$, so $|z-1| \ge |x-1| \ge 1$. Since we always have

$$\left|\frac{z}{|z|}-1\right| = |e^{i\theta}-1| \le 2,$$

this gives the desired result in this case.

Next suppose that $|\theta| \le \pi/2$. Then $|w-1| \ge |\sin\theta|$, since $|\sin\theta|$ is the distance from 1 to the ray determined by w. Hence

$$\begin{aligned}|w-1| &\ge |\sin\theta| = \tfrac{1}{2}|e^{i\theta}-e^{-i\theta}| \\ &= \tfrac{1}{2}|e^{2i\theta}-1| = \tfrac{1}{2}|e^{i\theta}-1|\,|e^{i\theta}+1| \\ &\ge \tfrac{1}{2}|e^{i\theta}-1| = \left|\frac{w}{|w|}-1\right|. \qquad \square\end{aligned}$$

Lemma 3.2. *Let $\alpha \in \mathbb{C}$ with $|\alpha| \le 1$ and α not a root of unity. Then, for all $n \ge 1$,*

$$\frac{1+|\alpha|^n}{\sqrt{(1-|\alpha|)^2+4|\alpha|\sin^2\dfrac{\pi}{2n}}} \le \frac{|\alpha^n-1|}{\min\limits_{\zeta\in\boldsymbol{\mu}_n}|\alpha-\zeta|} \le \sum_{j=0}^{n-1}|\alpha|^j. \tag{3.1}$$

In particular,

$$\tfrac{3}{10} \le \frac{|\alpha^n-1|}{\min\limits_{\zeta\in\boldsymbol{\mu}_n}|\alpha-\zeta|} \le n. \tag{3.2}$$

Remark. If we put $|\alpha| = 1$ in (3.1) and use the estimate $|\sin x| \ge 2x/\pi$, we obtain the following interesting inequality:

$$\frac{2n}{\pi} \le \frac{|\alpha^n-1|}{\min\limits_{\zeta\in\boldsymbol{\mu}_n}|\alpha-\zeta|} \le n$$

for all $\alpha\in\mathbb{C}$ with $|\alpha| = 1$. It is not hard to see that this is best possible.

Proof of Lemma 3.2. Replacing α by $\zeta\alpha$ for some $\zeta \in \boldsymbol{\mu}_n$, we may assume that

$$|\alpha-1| = \min_{\zeta\in\boldsymbol{\mu}_n}|\alpha-\zeta|.$$

This means that we can write $\alpha = re^{2\pi\theta}$ with $r \le 1$ and $|\theta| \le 1/2n$.

The upper bound in (3.1) follows trivially from the triangle inequality:

$$\left|\frac{\alpha^n-1}{\alpha-1}\right| = \left|\sum_{j=0}^{n-1}\alpha^j\right| \le \sum_{j=0}^{n-1}|\alpha|^j.$$

Then the upper bound in (3.2) is immediate from the assumption that $|\alpha| \le 1$.

To prove the lower bounds, we define

$$R(m) = \sum_{j=0}^{m-1} r^{2j} = \begin{cases}\dfrac{1-r^{2m}}{1-r^2} & \text{if } r<1,\\ m & \text{if } r=1.\end{cases}$$

We expand and regroup:

$$\begin{aligned}\left|\frac{\alpha^n-1}{\alpha-1}\right|^2 &= \left|\sum_{j=0}^{n-1} r^k e^{2\pi ij\theta}\right|^2 \\ &= \sum_{j=0}^{n-1}\sum_{k=0}^{n-1} r^{j+k}e^{2\pi i(j-k)\theta} \\ &= \sum_{u=-(n-1)}^{-1}\ \sum_{j=0}^{n-1+u} r^{-u+2j}e^{2\pi iu\theta} \\ &\qquad + \sum_{j=0}^{n-1} r^{2j} + \sum_{u=1}^{n-1}\ \sum_{k=0}^{n-1-u} r^{u+2k}e^{2\pi iu\theta} \\ &= R(n) + \sum_{u=1}^{n-1} R(n-u)(e^{2\pi iu\theta}+e^{-2\pi iu\theta}) \\ &= R(n) + 2\sum_{u=1}^{n-1} R(n-u)\cos(2\pi u\theta).\end{aligned}$$

Now we need only observe that if $1 \le u \le n-1$, the function $\cos(2\pi u\theta)$ on the interval $|\theta| \le 1/2n$ attains its absolute minimum at the endpoints. This means that, for a fixed modulus $|\alpha| = r$, the minimum value occurs at $\theta = 1/2n$, so

$$\begin{aligned}\frac{|\alpha^n - 1|}{\displaystyle\min_{\zeta\in\boldsymbol{\mu}_n} |\alpha-\zeta|} &\ge \frac{|(re^{2\pi i/2n})^n - 1|}{|re^{2\pi i/2n} - 1|}\\ &= \frac{1+r^n}{\sqrt{1+r^2-2r\cos(\pi/n)}}\\ &= \frac{1+r^n}{\sqrt{(1-r)^2+4r\sin^2(\pi/2n)}}.\end{aligned}$$

This completes the proof of (3.1).

In order to prove the lower bound in (3.2), we consider two cases, the first being $n \ge 1/(1-r)$. Then

$$\begin{aligned}(1-r)^2 + 4r\sin^2(\pi/2n) &\le (1-r)^2 + r\pi^2/n^2\\ &\le (1-r)^2(1+r\pi^2)\\ &\le (1-r)^2(1+\pi^2)\end{aligned}$$

(the first inequality because $|\sin x| \le x$, the second because of the assumption on n, and the last because $r \le 1$). Substituting this into the lower bound in (3.1), we get

$$\frac{1+r^n}{(1-r)\sqrt{1+\pi^2}} \ge \frac{3}{10(1-r)} \ge \tfrac{3}{10}$$

since $r \ge 0$. This proves the desired lower bound in this case. If, on the other hand, $n \le 1/(1-r)$, we have

$$\begin{aligned}(1-r)^2 + 4r\sin^2(\pi/2n) &\le 1/n^2 + 4r\sin^2(\pi/2n)\\ &\le (1+\pi^2)/n^2\end{aligned}$$

(the first inequality because $r \ge 1 - 1/n$, and the second because $r \le 1$ and $|\sin x| \le x$). Substituting this into the lower bound in (3.1), we get

$$\frac{1+r^n}{n^{-1}\sqrt{1+\pi^2}} \ge \tfrac{3}{10}n \ge \tfrac{3}{10},$$

which completes the proof of the lemma. □

Proposition 3.3. *For all $\alpha \in \mathbb{C}$ not a root of unity and satisfying $|\alpha| \le 1$, and for all integers $m \ge 1$,*

$$\frac{|\Phi_m(\alpha)|}{\displaystyle\min_{\zeta\in\boldsymbol{\mu}_m^*} |\alpha-\zeta|} \ge (118m)^{-3\sigma_0(m)/2}. \tag{3.3}$$

Proof. Write $\alpha = re^{2\pi i\theta}$ and choose an integer a satisfying

$$\left|\theta - \frac{a}{m}\right| \le \frac{1}{2m}.$$

Setting $\zeta_a = e^{2\pi i a/m}$, we have

$$\min_{\zeta\in\boldsymbol{\mu}_m} |\alpha-\zeta| = |\alpha - \zeta_a|.$$

Note, however, that this is the minimum over all m-th roots of unity, not just the primitive ones, since $\gcd(a,m)$ may be greater than one. So we write $a/m = A/M$ with $\gcd(A,M) = 1$, and then we have $\zeta_a \in \boldsymbol{\mu}_M^*$.

From [Apostol 1976, Theorem 2.1] we have

$$\sum_{\substack{n\mid m\\ n\equiv 0 \bmod M}} \mu\Big(\frac{m}{n}\Big) = \sum_{k\mid(m/M)} \mu\Big(\frac{m/M}{k}\Big) = I\left(\frac{m}{M}\right).$$

This allows us to write

$$\begin{aligned}\frac{|\Phi_m(\alpha)|}{\displaystyle\min_{\zeta\in\boldsymbol{\mu}_m^*} |\alpha-\zeta|} &= \frac{1}{\displaystyle\min_{\zeta\in\boldsymbol{\mu}_m^*} |\alpha-\zeta|} \prod_{n\mid m} |\alpha^n - 1|^{\mu(m/n)}\\ &= P_1\, P_2\, P_3,\end{aligned}$$

where

$$\begin{aligned}P_1 &= \frac{|\alpha-\zeta_a|^{I(m/M)}}{\displaystyle\min_{\zeta\in\boldsymbol{\mu}_m^*} |\alpha-\zeta|},\\ P_2 &= \prod_{\substack{n\mid m\\ n\equiv 0 \bmod M}} \left|\frac{\alpha^n-1}{\alpha-\zeta_a}\right|^{\mu(m/n)},\\ P_3 &= \prod_{\substack{n\mid m\\ n\not\equiv 0 \bmod M}} |\alpha^n - 1|^{\mu(m/n)}.\end{aligned}$$

We will treat each factor individually.

The first factor is easy. If $m = M$, then $\zeta_a \in \boldsymbol{\mu}_m^*$, so $P_1 = 1$; and if $m \ne M$, we can use the trivial estimate $|\alpha - \zeta| \le 2$ to get $P_1 \ge \frac{1}{2}$.

Next consider P_2. In this product every $n \equiv 0 \bmod M$, so we have $\zeta_a \in \boldsymbol{\mu}_n$. Further, our choice of a ensures that this is the n-th root of unity closest to α, so (3.2) gives the upper and lower bounds

$$n \ge \left|\frac{\alpha^n-1}{\alpha-\zeta_a}\right| \ge \tfrac{3}{10}.$$

Substituting these bounds into the definition of P_2 gives

$$\begin{aligned}\log P_2 &= \sum_{\substack{n\mid m\\ n\equiv 0 \bmod M}} \mu\left(\frac{m}{n}\right)\log\left|\frac{\alpha^n-1}{\alpha-\zeta_a}\right|\\ &\ge \sum_{\substack{n\mid m,\, \mu(m/n)=1\\ n\equiv 0 \bmod M}} \log\tfrac{3}{10} - \sum_{\substack{n\mid m,\, \mu(m/n)=-1\\ n\equiv 0 \bmod M}} \log n\\ &\ge \sum_{n\mid m}(-\log\tfrac{10}{3}-\log n)\\ &= -\sigma_0(m)\log\left(\frac{10\sqrt{m}}{3}\right).\end{aligned}$$

Exponentiating gives

$$P_2 \ge \left(10\sqrt{m}/3\right)^{-\sigma_0(m)}. \tag{3.4}$$

It remains to deal with P_3. Let $n|m$ be an integer with $n \not\equiv 0 \bmod M$, and choose $\xi \in \boldsymbol{\mu}_n$ so that

$$\min_{\zeta\in\boldsymbol{\mu}_n}|\alpha-\zeta| = |\alpha-\xi|.$$

Then, by Lemma 3.1, we have

$$\begin{aligned}\min_{\zeta\in\boldsymbol{\mu}_n}|\alpha-\zeta| &\ge \tfrac{1}{2}|e^{2\pi i\theta}-\xi| = \left|\sin\pi\left(\theta-\frac{b}{n}\right)\right|\\ &\ge 2\left|\theta-\frac{b}{n}\right|,\end{aligned} \tag{3.5}$$

where b is defined by $\xi = e^{2\pi i b/n}$ and where the last inequality used the fact that $|\sin(t)| \ge (2/\pi)\,|t|$ for $|t| \le \pi/2$.

Now we note that $\xi \ne \zeta_a$, since ζ_a is a primitive M-th root of unity and $M\nmid n$. This trivial but crucial observation implies that $b/n \ne a/m$, so

$$\left|\frac{b}{n}-\frac{a}{m}\right| = \left|\frac{bm/n-a}{m}\right| \ge \frac{1}{m}.$$

We also note that

$$\left|\theta-\frac{a}{m}\right| \le \frac{1}{2m},$$

since ζ_a is the closest m-th root of unity to $\alpha = re^{2\pi i\theta}$. These two estimates, combined with (3.5), yield

$$\min_{\zeta\in\boldsymbol{\mu}_n}|\alpha-\zeta| \ge 2\left(\left|\frac{b}{n}-\frac{a}{m}\right| - \left|\frac{a}{m}-\theta\right|\right) \ge \frac{1}{m}.$$

Hence, using the fact that $|\alpha| \le 1$ and Lemma 3.2, we get

$$\begin{aligned}2 \ge |\alpha^n-1| &= \frac{|\alpha^n-1|}{\min\limits_{\zeta\in\boldsymbol{\mu}_n}|\alpha-\zeta|}\cdot\min_{\zeta\in\boldsymbol{\mu}_n}|\alpha-\zeta|\\ &\ge \frac{3}{10m}.\end{aligned}$$

Using this, we are finally able to estimate P_3 as

$$\begin{aligned}P_3 &= \prod_{\substack{n\mid m\\ n\not\equiv 0 \bmod M}} |\alpha^n-1|^{\mu(m/n)} \ge \prod_{\substack{n\mid m\\ n\not\equiv 0 \bmod M}} \frac{3}{10m}\\ &\ge \prod_{n\mid m}\frac{3}{10m} = \left(\frac{10m}{3}\right)^{-\sigma_0(m)}.\end{aligned}$$

Combining this with the trivial estimate $P_1 \ge \frac{1}{2}$ and with (3.4), we get

$$\frac{\left|\Phi_m(\alpha)\right|}{\min\limits_{\zeta\in\boldsymbol{\mu}_m^*}|\alpha-\zeta|} = P_1\cdot P_2\cdot P_3 \ge \frac{1}{2}\left(\frac{100m^{3/2}}{9}\right)^{-\sigma_0(m)}.$$

Since $\sigma_0(m) \ge 2$ for all $m \ge 2$, this is stronger than (3.3). This completes the proof of Proposition 3.3. □

4. UPPER BOUNDS FOR UNITS $\Phi_m(\alpha)$

In this section we will prove the following bound:

Theorem 4.1. *Let $\varepsilon > 0$ and $\kappa \ge 1$. There is an effectively computable constant $c = c(\varepsilon,\kappa)$, depending only on ε and κ, such that any algebraic unit α of degree d that is not a root of unity satisfies*

$$\#\left\{m \ge 1 : |N\Phi_m(\alpha)| \le \kappa^d\right\} \le cd^{1+\frac{(1+\varepsilon)\log 2}{\log\log m}}. \tag{4.1}$$

Note that $\Phi_m(\alpha)$ is a unit if and only if the norm of $\Phi_m(\alpha)$ has absolute value 1, so this theorem is stronger than Theorem 0.1.

We begin with some preliminary calculations. The following result, which is essentially due to Stewart [1977], says that the largest m appearing in (4.1) satisfies $m \ll \max\{d^{265}, \log^{5/3}\kappa\}$. Thus the main feature of interest in Theorem 4.1 is the fact that the exponent of d in the upper bound is only slightly larger than 1.

Proposition 4.2 (after [Stewart 1977]). *Let α be an algebraic integer of degree $d \geq 2$ that is not a root of unity. If $m \geq (1000d)^{265}$, then*

$$\log|\Phi_m(\alpha)| > (1000d)^{50} m^{3/5}. \tag{4.2}$$

Proof. This result is contained in the proof of Theorem 1 of [Stewart 1977], so we just give a brief sketch. We begin with [Stewart 1977, eq. (12)], which in our notation says that

$$\begin{aligned} d\log|\Phi_m(\alpha)| > \varphi(m)\log M(\alpha) \\ - Cd\bigl(d+\log M(\alpha)\bigr)q(m)\log m, \end{aligned}$$

where $C = 2^{436}(3d)^{49}$ and $\log_2 q(m)$ is the number of distinct prime divisors of m. Next, [Stewart 1977, eq. (15)] gives

$$\log M(\alpha) \geq \frac{1}{1+52d\log 6d} \geq \frac{1}{100d^2}.$$

Combining these two estimates and doing a little algebra we get something stronger than

$$\log\bigl|\Phi_m(\alpha)\bigr| > \frac{1}{100d^3}\,\frac{\varphi(m)}{q(m)\log m} - 10^{155}d^{50}.$$

Next we use the fact that

$$\frac{\varphi(m)}{q(m)\log m} > m^{4/5}$$

[Stewart 1977, p. 88] to get

$$\begin{aligned} \log\bigl|\Phi_m(\alpha)\bigr| &> \frac{m^{4/5}}{100d^3} - 2^{437}3^{49}d^{50} \\ &= (1000d)^{50}m^{3/5}\Bigl(\frac{m^{1/5}}{10^{153}d^{53}} - \frac{10^5}{m^{3/5}}\Bigr). \end{aligned}$$

Finally, our assumption that $m > (1000d)^{265}$ gives something stronger than the desired result. □

Proof of Theorem 4.1. Unless otherwise indicated, the constants c_i appearing in this proof are effectively computable constants that depend only on ε and κ. We denote the conjugates of α by $\alpha_1, \dots, \alpha_d \in \mathbb{C}$ and define $\beta_1, \dots, \beta_d$ to be

$$\beta_i = \begin{cases} \alpha_i & \text{if } |\alpha_i| \leq 1, \\ \alpha_i^{-1} & \text{if } |\alpha_i| > 1. \end{cases}$$

Thus $|\beta_i| \leq 1$ for all i. We will need the estimates

$$\sigma_0(m) \leq c_6 m^{\frac{(1+\varepsilon)\log 2}{\log\log m}} \tag{4.3}$$

and

$$\varphi(m) \geq \frac{c_7 m}{\log\log m} \tag{4.4}$$

[Apostol 1976, Theorems 13.12(a) and 13.14(a)]. As we will see below, our argument hinges on the fact that $\varphi(m)$ grows much faster than $\sigma_0(m)$.

We take m to be in the set on the left-hand side of (4.1), so we can write

$$c_8 d \geq \log|N\Phi_m(\alpha)|.$$

This gives

$$\begin{aligned} c_8 d &\geq \sum_{i=1}^{d} \log|\Phi_m(\alpha_i)| \\ &= \sum_{\substack{i=1 \\ |\alpha_i|\geq 1}}^{d} \log|\alpha_i^{\varphi(m)}| + \sum_{i=1}^{d}\log|\Phi_m(\beta_i)|, \end{aligned}$$

since $\Phi_m(x) = \pm x^{\varphi(m)}\Phi_m(x^{-1})$. By the definition of the Mahler measure, the first sum on the right-hand side is $\varphi(m)\log M(\alpha)$, so Proposition 3.3 now gives

$$\begin{aligned} c_8 d &\geq \varphi(m)\log M(\alpha) \\ &\quad + \sum_{i=1}^{d}\Bigl(\log\min_{\zeta\in\boldsymbol{\mu}_m^*}|\beta_i-\zeta| - \tfrac{3}{2}\sigma_0(m)\log(118\sqrt{m})\Bigr) \\ &\geq \frac{c_7 m}{\log\log m}\log M(\alpha) \\ &\qquad -d\cdot c_6 m^{\frac{(1+\varepsilon)\log 2}{\log\log m}}\cdot\tfrac{3}{2}\log(118\sqrt{m}) \\ &\qquad\qquad + \sum_{i=1}^{d}\log\min_{\zeta\in\boldsymbol{\mu}_m^*}|\beta_i-\zeta| \end{aligned}$$

by (4.3) and (4.4). Rearranging the terms and adjusting the constants as necessary, we find

$$-\sum_{i=1}^{d} \log \min_{\zeta\in\boldsymbol{\mu}_m^*} |\beta_i - \zeta| \geq \frac{c_9 m \log M(\alpha)}{\log\log m} - c_{10} d m^{\frac{(1+\varepsilon)\log 2}{\log\log m}},$$

and hence

$$-d \log \min_{\substack{1\leq i\leq d\\ \zeta\in\boldsymbol{\mu}_m^*}} |\beta_i - \zeta| \geq \frac{c_9 m \log M(\alpha)}{\log\log m} - c_{10} d m^{\frac{(1+\varepsilon)\log 2}{\log\log m}}. \tag{4.5}$$

We are now going to assume that m satisfies

$$m^{1-\frac{(1+\varepsilon)\log 2}{\log\log m}} \geq \frac{d}{\log M(\alpha)} \quad \text{and} \quad m > d. \tag{4.6}$$

This means that at the end of the proof we will have to include all smaller values of m as possible elements of the set (4.1) whose size we are estimating. Now divide both sides of (4.5) by d and substitute in (4.6) to obtain

$$-\log \min_{\substack{1\leq i\leq d\\ \zeta\in\boldsymbol{\mu}_m^*}} |\beta_i - \zeta| \geq \frac{c_{11} m^{\frac{(1+2\varepsilon)\log 2}{\log\log m}}}{\log\log m} - c_{12} m^{\frac{(1+\varepsilon)\log 2}{\log\log m}}.$$

The function $m^{\varepsilon\log 2/\log\log m}$ grows faster than any power of $\log m$, so if we assume that $m > c_{13}$, we obtain the fundamental estimate

$$-\log \min_{\substack{1\leq i\leq d\\ \zeta\in\boldsymbol{\mu}_m^*}} |\beta_i - \zeta| \geq c_{14} m^{\frac{(1+\varepsilon)\log 2}{\log\log m}}.$$

Multiplying both sides by -1 and exponentiating yields the equivalent estimate

$$\min_{\substack{1\leq i\leq d\\ \zeta\in\boldsymbol{\mu}_m^*}} |\beta_i - \zeta| \leq \exp\Bigl(-c_{14} m^{\frac{(1+\varepsilon)\log 2}{\log\log m}}\Bigr). \tag{4.7}$$

The content of this inequality is that, if m is large and in the set (4.1), one of the β_i must be extremely close to some primitive m-th root of unity. In fact, the estimate (4.7) is so good that it implies that the m's in the set (4.1) satisfy a sort of "supergap principle", as described in the following result.

Lemma 4.3 (supergap principle). *Let $s > 0$ and $t \geq 1$ be fixed constants. There is a number $X_0(s,t)$ such that, for all $\beta \in \mathbb{C}$ and all $X \geq X_0(s,t)$, there is at most one m such that $X \leq m \leq X^t$ and*

$$\min_{\zeta\in\boldsymbol{\mu}_m^*} |\beta - \zeta| \leq \exp(-m^{s/\log\log m}).$$

Proof. Let $m_1 < m_2$ both satisfy this last inequality, and let $\zeta_1 \in \boldsymbol{\mu}_{m_1}^*$ and $\zeta_2 \in \boldsymbol{\mu}_{m_2}^*$ be the corresponding roots of unity closest to β. Then

$$\begin{aligned}|\zeta_1 - \zeta_2| &\leq |\beta - \zeta_1| + |\beta - \zeta_2| \\ &\leq \exp(-m_1^{s/\log\log m_1}) + \exp(-m_2^{s/\log\log m_2}).\end{aligned}$$

On the other hand, we know that $\zeta_1 \neq \zeta_2$, and clearly $\zeta_1\zeta_2^{-1}$ is an m_1m_2-th root of unity, so we have the trivial lower bound

$$\begin{aligned}|\zeta_1 - \zeta_2| = |\zeta_1\zeta_2^{-1} - 1| &\geq |e^{2\pi i/m_1m_2} - 1| \\ &= 2\,|\sin(\pi/m_1m_2)| \geq 4/m_1m_2.\end{aligned}$$

Combining the upper and lower bounds and using the assumption that $m_1 < m_2$, we find that

$$4/m_2^2 \leq 2\exp(-m_1^{s/\log\log m_1}).$$

If we further assume that m_1 is larger than some constant depending only on s, we find that

$$\log\log m_2 \geq \frac{s}{2}\,\frac{\log m_1}{\log\log m_1}.$$

But if $X \leq m_1 < m_2 \leq X^t$, this gives

$$\log(t\log X) \geq \frac{s}{2}\,\frac{\log X}{\log\log X},$$

which is a contradiction as soon as $X > X_0(s,t)$. □

We resume the proof of Theorem 4.1. It clearly suffices to bound the size of the set

$$\bigl\{m \geq d : |N\Phi_m(\alpha)| \leq \kappa^d\bigr\}. \tag{4.8}$$

Further, Stewart's result (4.2) says that any m in (4.8) satisfies

$$m \leq c_{15} d^{265}.$$

It thus suffices to prove Theorem 4.1 under the assumption that $d > c_{16}$, since the constant c in (4.1) can be adjusted to account for small values of d.

Now let $m_1 < m_2 < \cdots < m_N$ be the distinct elements in (4.8) that also satisfy the inequality (4.6). According to (4.7), we can assign to each $1 \le j \le N$ an index $i(j)$ so that

$$\min_{\zeta \in \boldsymbol{\mu}^*_{m_j}} |\beta_{i(j)} - \zeta| \le \exp(-c_{14} m_j^{(1+\varepsilon)\log 2/\log\log m_j}).$$

Since we are further assuming that $m \ge d > c_{16}$, we can absorb the constant into the power of m, so

$$\min_{\zeta \in \boldsymbol{\mu}^*_{m_j}} |\beta_{i(j)} - \zeta| \le \exp(-m_j^{c_{17}/\log\log m_j}).$$

On the other hand, Stewart's result (4.2) says that

$$m_j \le d^{c_{18}}.$$

Using the last two equations in the supergap principle (Lemma 4.3), we see that, for each β_i, there is at most one m_j with $i(j) = i$.

To summarize, we have shown that the set (4.8) contains no more than d elements satisfying the inequality (4.6). Hence (4.8) contains at most

$$d + d + c_{19}\Bigl(\frac{d}{\log M(\alpha)}\Bigr)^{1+\frac{(1+2\varepsilon)\log 2}{\log\log m}} \tag{4.9}$$

elements.

To complete the proof of the theorem, we apply Dobrowolski's theorem [1979], which says that

$$\log M(\alpha) \ge c_{20}\Bigl(\frac{\log\log d}{\log d}\Bigr)^3. \tag{4.10}$$

Substituting this into (4.9) and using

$$\frac{d}{\log M(\alpha)} \le c_{21}\Bigl(\frac{\log d}{\log\log d}\Bigr)^3 \le c_{22} d^{1+\varepsilon/\log\log d}$$

gives the desired result after adjusting the value of ε. This completes the proof of Theorem 4.1. □

Remark. Before Dobrowolski proved the estimate (4.10), Blanksby and Montgomery [1971] and Stewart [1978] had proved the weaker result

$$\log M(\alpha) \ge \frac{c_{23}}{d \log d}. \tag{4.11}$$

This estimate sufficed for Stewart [1977] to prove his polynomial upper bound for the largest value of m, but if we use (4.11) in place of (4.10), our upper bound (4.1) for the number of m's would look like $d^{2+o(1)}$ instead of $d^{1+o(1)}$. On the other hand, even if we knew Lehmer's conjecture that $\log M(\alpha) \ge c_{27}$, we would not be able to improve the upper bound in Theorem 4.1 unless we could also improve the lower bound in Proposition 3.3.

5. LOWER BOUNDS FOR UNITS $\Phi_m(\alpha)$

In this section we describe David Boyd's proof that the set $\mathcal{E}(\alpha)$ of values of $n \ge 1$ such that $1 - \alpha^n$ is a unit can be fairly large. We continue with the notations $d(\alpha)$ and $M(\alpha)$ for the degree and the Mahler measure of α. Further, we denote by $\mathcal{U}(\alpha)$ the set of $m \ge 1$ such that $\Phi_m(\alpha)$ is a unit. Notice that a lower bound for $E(\alpha)$ is automatically a lower bound for $U(\alpha)$. We begin with two elementary results.

Proposition 5.1. *Let p be a prime, and let $\beta = \alpha^{1/p}$ be any p-th root of α.*

(a) (after Rausch [1985]) *Let $K/\mathbb{Q}$ be a number field containing α. If α is not a p-th power in K^*, then $[K(\beta) : K] = p$.*

(b) (Boyd) *If $[\mathbb{Q}(\beta) : \mathbb{Q}(\alpha)] = p$, then*

$$\mathcal{E}(\beta) = p\,\mathcal{E}(\alpha) \cup (\mathcal{E}(\alpha) \setminus p\mathbb{Z}),$$
$$\mathcal{U}(\beta) = p\,\mathcal{U}(\alpha) \cup (\mathcal{U}(\alpha) \setminus p\mathbb{Z}).$$

In particular, $E(\beta) = 2E(\alpha)$ if every $n \in \mathcal{E}(\alpha)$ is prime to p, and $U(()\beta) = 2U(()\alpha)$ if every $n \in \mathcal{U}(\alpha)$ is prime to p.

Proof. (a) If $[K(\beta) : K] < p$, the polynomial $X^p - \alpha$ is reducible in $K[X]$, since it has a root $X = \beta$ of degree less than p. Factor $X^p - \alpha = g(X)h(X)$ with $g(X)$, $h(X) \in K[X]$ monic polynomials and $t = \deg g$ satisfying $1 \le t < p$. The polynomial $X^p - \alpha$ factors over $\bar{K}$ as

$$g(X)h(X) = X^p - \alpha = X^p - \beta^p = \prod_{\zeta \in \boldsymbol{\mu}_p} (X - \zeta\beta).$$

Comparing constant terms, we see that

$$g(0) = (-1)^t \zeta_1 \zeta_2 \cdots \zeta_t \beta^t$$

with $\zeta_1, \ldots, \zeta_t \in \boldsymbol{\mu}_p$. But $g(0) \in K$, so there is $\xi \in \boldsymbol{\mu}_p$ such that $\xi\beta^t \in K$. Further, $\beta^p = \alpha \in K$ and $\gcd(t,p) = 1$, so taking appropriate powers $(\xi\beta^t)^i(\beta^p)^j$ we find that there is a $\xi' \in \boldsymbol{\mu}_p$ such that $\xi'\beta \in K$. Then $\alpha = (\xi'\beta)^p$, contradicting the assumption that α is not a p-th power in K. Hence $X^p - \alpha$ is irreducible in $K[X]$, which proves that $[K(\beta) : K] = p$.

(b) We prove the result for $\mathcal{E}(\beta)$ and leave the similar proof for $\mathcal{U}(\beta)$ to the reader. The assumption that $[\mathbb{Q}(\beta) : \mathbb{Q}(\alpha)] = p$ means that the conjugates of β over $\mathbb{Q}(\alpha)$ are exactly the numbers $\zeta\beta$ with $\zeta \in \boldsymbol{\mu}_p$. So if we write N_K for the $K/\mathbb{Q}$ norm, we can compute

$$\begin{aligned} N_{\mathbb{Q}(\beta)}(1-\beta^n) &= N_{\mathbb{Q}(\alpha)}\Biggl(\prod_{\zeta\in\boldsymbol{\mu}_p}(1-(\zeta\beta)^n)\Biggr) \\ &= N_{\mathbb{Q}(\alpha)}\Biggl(\prod_{\zeta\in\boldsymbol{\mu}_p}\prod_{\eta\in\boldsymbol{\mu}_n}(1-\eta\zeta\beta)\Biggr) \\ &= N_{\mathbb{Q}(\alpha)}\Biggl(\prod_{\eta\in\boldsymbol{\mu}_n}(1-\eta^p\beta^p)\Biggr) \\ &= N_{\mathbb{Q}(\alpha)}\Biggl(\prod_{\eta\in\boldsymbol{\mu}_n}(1-\eta^p\alpha)\Biggr) \\ &= \begin{cases} N_{\mathbb{Q}(\alpha)}(1-\alpha^n) & \text{if } p\nmid n, \\ N_{\mathbb{Q}(\alpha)}(1-\alpha^{n/p})^p & \text{if } p\,|\,n. \end{cases} \end{aligned}$$

It follows that n is in $\mathcal{E}(\beta)$ if and only if $n/\gcd(n,p)$ is in $\mathcal{E}(\alpha)$, which is just another way to state the assertion in part (b) of the proposition. □

As Boyd points out, Proposition 5.1 can be used to find specific values of β for which $\mathcal{E}(\beta)$ and $\mathcal{U}(\beta)$ are large. For example, let $\alpha = \gamma_3$, a root of (1.3), be the number with smallest known Mahler measure greater than 1. One can check that the number of $m \in \mathcal{U}(\alpha)$ for which p divides m is

$$44, 29, 19, 11, 8, 6, 3, 3, 3, 1, 1, 2, 0, 1, 1, 0, 1$$

for $p = 2, 3, 5, \ldots, 59$, and is 0 for all other primes. It follows from Proposition 5.1 that

$$\mathcal{U}(\alpha^{1/2}) = 132 - 44 = 88,$$

and that $\mathcal{U}(\alpha^{1/3}) = 132 - 29 = 103$. Notice that these values are larger than the corresponding values in Table 2, column B, for degrees 20 and 30.

Proposition 5.1 can also be used to show that $\mathcal{E}(\alpha)$ and $\mathcal{U}(\alpha)$ may grow quite rapidly.

Corollary 5.2 (Boyd). (a) *Suppose that $\mathcal{E}(\alpha) \geq 1$, and fix $\varepsilon > 0$. For every $k \geq 1$, let $\beta_k = \alpha^{1/k}$ be a k-th root of α. Then there exists a sequence of numbers $k \to \infty$ such that $M(\beta_k) = M(\alpha)$ and*

$$E(\beta_k) > d(\beta_k)^{(1-\varepsilon)\log 2/\log\log d(\beta_k)}.$$

A similar result holds for $U(\beta_k)$.

(b) *Let $\psi(d)$ be any function such that $\psi(d) \to 0$ as $d \to \infty$. Then, for every $d \geq 1$ and every $C \geq 1$, there exists α satisfying $d(\alpha) \geq d$, $E(\alpha) \geq C$, and*

$$E(\alpha) \geq \frac{d(\alpha)\psi(d(\alpha))}{\log M(\alpha)}.$$

In particular, it is not possible to find absolute constants A and B such that

$$E(\alpha) \leq A + \frac{B\varphi(d(\alpha))}{\log M(\alpha)}$$

for all α.

Proof. (a) Let k be the product of all primes $p \leq t$ such that p does not divide any element of $\mathcal{E}(\alpha)$ and such that α is not a p-th power in $\mathbb{Q}(\alpha)^*$. Notice we have eliminated only finitely many primes, so $k \gg\ll e^t$.

Proposition 5.1(a) says that $[\mathbb{Q}(\alpha^{1/p}) : \mathbb{Q}(\alpha)] = p$ if p divides k. These degrees are relatively prime for different values of p, and $\mathbb{Q}(\beta_k)$ is the compositum of $\mathbb{Q}(\alpha^{1/p})$'s for $p|k$, so it follows that $[\mathbb{Q}(\beta_k) : \mathbb{Q}(\alpha)] = k$. In particular, this implies that $M(\beta_k) = M(\alpha)$.

Suppose p divides k. Proposition 5.2(b) tells us that $E(\alpha)^{1/p} = 2E(\alpha)$. But it tells us even more, since it says that $\mathcal{E}(\alpha^{1/p})$ is the union of $\mathcal{E}(\alpha)$ and $p\,\mathcal{E}(\alpha)$. Thus, if q is another prime dividing k, none of of the numbers in $\mathcal{E}(\alpha^{1/p})$ is divisible by q, so we can apply Proposition 5.2(b) to $\alpha^{1/p}$ and q to deduce that

$$E(\alpha^{1/pq}) = 2E(\alpha)^{1/p} = 4E(\alpha).$$

Continuing in this fashion, we find that

$$E(\beta_k) = E(\alpha^{1/k}) \geq 2^r E(\alpha) \geq 2^r,$$

where r is the number of primes dividing k. In particular, $r \geq \pi(t) + O(1) \geq (1-\varepsilon)t/\log t$ for all sufficiently large t. Combining the estimates $E(\beta_k) \geq 2^r$,

$$r \geq (1-\varepsilon)\frac{t}{\log t}$$

and $k \gg\ll e^t$, we obtain

$$E(\beta_k) \geq 2^{(1-\varepsilon)\log k/\log\log k}.$$

This gives the desired result, since

$$k = [\mathbb{Q}(\beta_k) : \mathbb{Q}(\alpha)] = d(\beta_k)/d(\alpha).$$

(b) We give only a sketch of the proof. Suppose that the assertion is false. For each $n \geq 1$, let α_n be a root of $f_n(x) = (x-1)^{2n} + x^n$. Assuming that f_n is irreducible, it follows from [Boyd 1980] that $M(\alpha_n) \approx c_{28}^n$, where $c_{28} = 1.90814\ldots$. One can show that, if n is a power of 2, then $f_n(x^k)$ is irreducible for all $k \geq 1$, so we restrict attention to values of n that are powers of 2. We also note that $N(1-\alpha_n) = \pm f_n(1) = \pm 1$, so $1 \in \mathcal{E}(\alpha_n)$. Just as in the proof of part (a), we now take k to be the product of the first t primes and consider $\beta_{n,k} = \alpha_n^{1/k}$. Again as in (a), Proposition 5.1 and our assumptions give $E(\beta_{n,k}) \geq 2^{(1-\varepsilon)t/\log t}$, $d(\beta_{n,k}) = kd(\alpha_n) = kn$, and $M(\beta_{n,k}) = M(\alpha_n) \approx c_{28}^n$. We can thus fix a value for t such that $E(\beta_{n,k}) \geq C$ and $d(\beta_{n,k}) \geq d$.

We are assuming that the assertion in (b) is false, so for every α we have

$$E(\alpha) \leq \frac{d(\alpha)\psi(d(\alpha))}{\log M(\alpha)}.$$

Using our estimates from above, we obtain

$$\begin{aligned} E(\beta_{n,k}) &\leq \frac{d(\beta_{n,k})\psi(d(\beta_{n,k}))}{\log M(\beta_{n,k})} \leq \frac{kn\psi(kn)}{\log c_{28}^n} \\ &= c_{29}k\psi(kn). \end{aligned}$$

Here the left-hand side goes to 0 as $n \to \infty$, while the right-hand side is greater than C. This contradiction completes the proof. □

ACKNOWLEDGEMENTS

I would like to thank A. Schinzel, and Cam Stewart for their helpful comments, David Bayer for some computational advice, and especially David Boyd for his many useful suggestions and for the constructions in Section 5 that punctured my original overoptimistic conjecture.

REFERENCES

[Apostol 1976] T. Apostol, *Introduction to Analytic Number Theory*, Undergraduate Texts in Mathematics, Springer, New York, 1976.

[Batut et al. 1993] C. Batut, D. Bernardi, H. Cohen and M. Olivier, *User's Guide to Pari-GP*. This manual is part of the program distribution, available by anonymous ftp from the host megrez.ceremab.u-bordeaux.fr.

[Blanksby and Montgomery 1971] P. E. Blanksby and H. L. Montgomery, "Algebraic integers near the unit circle", *Acta Arith.* **18** (1971), 355–369.

[Bombieri et al. 1994] E. Bombieri, J. Mueller and M. Poe, "S-unit equations and the cluster principle", in preparation.

[Boyd 1977] D. Boyd, "Small Salem numbers", *Duke Math. J.* **44** (1977), 315–328.

[Boyd 1978] D. Boyd, "Pisot and Salem numbers in intervals of the real line", *Math. Comp.* **32** (1978), 1244–1260.

[Boyd 1980] D. Boyd, "Reciprocal polynomials having small measure", *Math. Comp.* **35** (1980), 1361–1377; followup in *Math. Comp.* **53** (1989), 355–357.

[Cohen et al. 1992] H. Cohen, L. Lewin and D. Zagier, "A sixteenth order polylogarithm ladder", *Experimental Math.* **1** (1992), 25–34.

[Dobrowolski 1979] E. Dobrowolski, "On a question of Lehmer and the number of irreducible factors of a polynomial", *Acta Arith.* **34** (1979), 391–401.

[Evertse 1984] J.-H. Evertse, "On equations in S-units and the Thue–Mahler equation", *Invent. Math.* **75** (1984), 561–584.

[Hindry and Silverman 1988] M. Hindry and J. Silverman, "The canonical height and integral points on elliptic curves", *Invent. Math.* **93** (1988), 419–450.

[Hindry and Silverman 1990] M. Hindry and J. Silverman, "On Lehmer's conjecture for elliptic curves", pp. 103–116 in *Séminaire de Théorie des Nombres*, 1988–1989, Paris (edited by C. Goldstein), Prog. in Math. **91**, Birkhäuser, Boston, 1990.

[Kronecker 1857] L. Kronecker, "Zwei Sätze über Gleichungen mit ganzzahligen Coefficienten", *J. Reine Angew. Math.* **53** (1857), 133–175.

[Laurent 1983] M. Laurent, "Minoration de la hauteur de Néron–Tate", pp. 137–151 in *Séminaire de Théorie des Nombres*, 1981–1982, Paris (edited by C. Goldstein), Prog. in Math. **81**, Birkhäuser, Boston, 1983.

[Lehmer 1933] D. H. Lehmer, "Factorization of certain cyclotomic functions", *Annals of Math.* **34** (1933), 461–479.

[Masser 1989] D. Masser, "Counting points of small height on elliptic curves", *Bull. Soc. Math. France* **117** (1989), 247–265.

[Mignotte 1977] "Entiers algébriques dont les conjugués sont proches du cercle unité", *Séminaire Delange–Pisot–Poitou* **19** (1977/78), lecture 39.

[Rausch 1985] U. Rausch, "On a theorem of Dobrowolski about the product of conjugate numbers", *Colloquium Math.* **50** (1985), 137–142.

[Silverman 1994] J. H. Silverman, "Lehmer's conjecture and primes of small norm", preprint, Brown University, July 1994.

[Smyth 1971] C. J. Smyth, "On the product of conjugates outside the unit circle of an algebraic integer", *Bull. London Math. Soc.* **3** (1971), 169–175.

[Stewart 1977] C. L. Stewart, "Primitive divisors of Lucas and Lehmer numbers", pp. 79–92 in *Transcendence Theory: Advances and Applications* (edited by A. Baker and D. W. Masser), Academic Press, London, 1977.

[Stewart 1978] C. L. Stewart, "Algebraic integers whose conjugates lie near the unit circle", *Bul. Soc. Math. France* **196** (1978), 169–176.

Joseph H. Silverman, Mathematics Department, Box 1917, Brown University, Providence, RI 02912
(jhs@gauss.math.brown.edu)

Received November 29, 1994; accepted in revised form May 12, 1995