Existence of primitive divisors of Lucas and Lehmer numbers

Yu. Bilu*, G. Hanrot and P.M. Voutier
(with an appendix by M. Mignotte)

Abstract

We prove that for $n > 30$, every n-th Lucas and Lehmer number has a primitive divisor. This allows us to list all Lucas and Lehmer numbers without a primitive divisor.

Whether the mathematicians like it or not, the computer is here to stay.
Folklore

Whether the computer likes it or not, mathematics is here to stay.
Beno Eckmann [32], p. xxiii

Contents

1 Introduction

A *Lucas pair* is a pair (α, β) of algebraic integers such that $\alpha + \beta$ and $\alpha\beta$ are non-zero coprime rational integers and α/β is not a root of unity. Given a Lucas pair (α, β), one defines the corresponding sequence of *Lucas numbers* by

$$u_n = u_n(\alpha, \beta) = \frac{\alpha^n - \beta^n}{\alpha - \beta} \quad (n = 0, 1, 2, \ldots) \tag{1}$$

*Partially supported by the CNPq (Brazil), Forschungsinstitut für Mathematik ETH Zürich, and by the Lise Meitner Fellowship (Austria), grant M00421-MAT.

1991 *Mathematics Subject Classification* primary 11B39; secondary 11Y50, 11J85, 11J86, 11Y65, 11D57

A *Lehmer pair* is a pair (α, β) of algebraic integers such that $(\alpha+\beta)^2$ and $\alpha\beta$ are non-zero coprime rational integers and α/β is not a root of unity. For a Lehmer pair (α, β), one defines the corresponding sequence of *Lehmer numbers* by

$$\tilde{u}_n = \tilde{u}_n(\alpha, \beta) = \begin{cases} \dfrac{\alpha^n - \beta^n}{\alpha - \beta} & \text{if } n \text{ is odd,} \\ \dfrac{\alpha^n - \beta^n}{\alpha^2 - \beta^2} & \text{if } n \text{ is even.} \end{cases} \tag{2}$$

Notice that every Lucas pair (α, β) is also a Lehmer pair, and

$$u_n = \begin{cases} \tilde{u}_n & \text{if } n \text{ is odd,} \\ (\alpha+\beta)\tilde{u}_n & \text{if } n \text{ is even.} \end{cases} \tag{3}$$

Lucas and Lehmer numbers are quite classical objects and were studied by many authors. See Ribenboim [33] for a comprehensive survey of results.

In the present paper we concentrate on one of the oldest problems about Lucas and Lehmer numbers: *the existence of primitive divisors.*

Let (α, β) be a Lucas pair. A prime number p is a *primitive divisor* of $u_n(\alpha, \beta)$ if p divides u_n but does not divide $(\alpha-\beta)^2 u_1 \cdots u_{n-1}$. (For instance, among the first several Fibonacci numbers

$$\boxed{1}, \boxed{1}, 2, 3, \boxed{5}, \boxed{8}, 13, 21, 34, 55, 89, \boxed{144}, 233, 377, 610, 987, 1597, \ldots$$

the framed ones have no primitive divisors.)

Similarly, let (α, β) be a Lehmer pair. A prime number p is a *primitive divisor* of $\tilde{u}_n(\alpha, \beta)$ if p divides $\tilde{u}_n$ but does not divide $(\alpha^2-\beta^2)^2 \tilde{u}_1 \cdots \tilde{u}_{n-1}$.

Remark 1.1 If (α, β) is a Lucas pair and $n \neq 2$ then, as follows from (3), a prime p is a primitive divisor of $u_n(\alpha, \beta)$ if and only if it is a primitive divisor of $\tilde{u}_n(\alpha, \beta)$.

The following problem goes back to the beginning of 20-th century (or even to earlier terms), though it does not seem to be ever formulated explicitly (perhaps, because one could hardly imagine it been solved in such generality):

Main problem *List all Lucas and Lehmer numbers without primitive divisors.*

More precisely: *classify all triples (α, β, n) such that that (α, β) is a Lucas (or Lehmer) pair, and $u_n(\alpha, \beta)$ (or $\tilde{u}_n(\alpha, \beta)$) has no primitive divisors.*

The first general result about the existence of primitive divisors dates back to as early as 1892, when Zsigmondy [46] proved that $u_n(\alpha, \beta)$ has a primitive divisor for $n > 6$ when $\alpha, \beta \in \mathbb{Z}$. (The particular case $\beta = 1$ was done even earlier [4].) Notice that this is best possible, since $2^6 - 1 = (2^2-1)^2(2^3-1)$. Independently, the same result was obtained by Birkhoff and Vandiver [9] in 1904.

In 1913 Carmichael [11] proved that if (α, β) is a *real* Lucas pair (that is, $\alpha, \beta \in \mathbb{R}$) then $u_n(\alpha, \beta)$ has a primitive divisor for $n > 12$. (Since the 12-th Fibonacci number has no primitive divisors, this is again best possible.) This was extended to the real Lehmer pairs by Ward [44], see [38], Lemma 8. (A Lehmer pair (α, β) is *real* if $\alpha^2, \beta^2 \in \mathbb{R}$.) See also Durst [14, 15].

It will be convenient to use the following terminology. A Lucas (respectively Lehmer) pair (α, β) such that $u_n(\alpha, \beta)$ (respectively $\tilde{u}_n(\alpha, \beta)$) has no primitive divisors will be called *n-defective Lucas* (respectively, *Lehmer*) *pair.*

Remark 1.2 If $n \neq 2$ and (α, β) is a Lucas pair, then, as follows from Remark 1.1, (α, β) is an n-defective Lucas pair if and only if it is an n-defective Lehmer pair.

In these terms, the previously mentioned results can be formulated as follows[1].

[1] In the introduction we number with letters (Theorem A, etc.) results that are only quoted but not proved in the present paper.

Table 1:

n	(a,b)
5	$(1,5),(1,-7),(2,-40),(1,-11),(1,-15),(12,-76),(12,-1364)$
7	$(1,-7),(1,-19)$
8	$(2,-24),(1,-7)$
10	$(2,-8),(5,-3),(5,-47)$
12	$(1,5),(1,-7),(1,-11),(2,-56),(1,-15),(1,-19)$
13	$(1,-7)$
18	$(1,-7)$
30	$(1,-7)$

Theorem A (Carmichael, Ward) *For $n>12$ there are no n-defective real Lucas and Lehmer pairs.*

The situation is much more complicated for non-real Lucas and Lehmer pairs. "Nothing appears to be known about the intrinsic divisors of Lucas and Lehmer numbers when α and β are complex,"— wrote Ward [44], p. 230, in 1955. Only in 1974 did Schinzel [36] prove the non-existence of n-defective Lucas and Lehmer pairs for n exceeding an effectively computable absolute constant n_0. (Earlier [34] he did it with n_0 depending on α and β; see also [35, 31] for more general results.) While Carmichael and Ward used skillful but, in principle, elementary arguments, Schinzel's proof relied upon deep tools of transcendence theory (Gelfond-Baker inequality).

Stewart [37] made Schinzel's result explicit by showing that $n_0 = e^{452}4^{67}$ would do. This was improved to $n_0 = 2\cdot 10^{10}$ in [41] and to $n_0 = 30030$ in [42]:

Theorem B [42] *For $n>30030$ there are no n-defective Lucas and Lehmer pairs.*

Let us say that an integer n is *totally non-defective* if no Lehmer pair is n-defective. (Trivially, for such n no Lucas pair is n-defective either.) In this terms Theorem B reads: *every integer $n>30030$ is totally non-defective.*

Stewart [37] pointed out that enumerating n-defective Lucas and Lehmer pairs for a fixed n reduces to solving an explicitly given binary norm Diophantine equation of degree $\varphi(n)/2$, which is either a linear equation, or a quadratic equation, or a Thue equation. Since such equations can be resolved effectively (for linear and quadratic equations this is a part of the folklore, for Thue equation this is due to Baker [1]), this means that Stewart reduced the main problem to a finite (though hopelessly long) computation.

In [40] the Thue equations corresponding to $n\le 30$ were resolved, using the method of Tzanakis and de Weger [39] (with some modifications). As a consequence, the following was proved.

Theorem C [40] *Let n satisfy $4<n\le 30$ and $n\ne 6$. Then, up to equivalence (see below), all n-defective Lucas pairs are of the form $\left((a-\sqrt{b})/2,(a+\sqrt{b})/2\right)$, where (a,b) are given in Table 1.*

Let n satisfy $6<n\le 30$ and $n\ne 8,10,12$. Then, up to equivalence, all n-defective Lehmer pairs are of the form $\left((\sqrt{a}-\sqrt{b})/2,(\sqrt{a}+\sqrt{b})/2\right)$, where (a,b) are given in Table 2.

Two Lucas pairs (α_1,β_1) and (α_2,β_2) are *equivalent* if $\alpha_1/\alpha_2=\beta_1/\beta_2=\pm1$. Two Lehmer pairs (α_1,β_1) and (α_2,β_2) are *equivalent* if $\alpha_1/\alpha_2=\beta_1/\beta_2\in\{\pm1,\pm\sqrt{-1}\}$. For equivalent Lucas pairs we have $u_n(\alpha_1,\beta_1)=\pm u_n(\alpha_2,\beta_2)$. Therefore one of them is n-defective if and only if the other is. The same holds for equivalent Lehmer pairs.

Though Theorem C misses some values of n, it is not difficult to classify defective Lucas and Lehmer pairs also for these values, following an idea of Stewart [37], Theorem 3.

Table 2:

n	(a,b)
7	$(1,-7),(1,-19),(3,-5),(5,-7),(13-3),(14,-22)$
9	$(5,-3),(7,-1),(7,-5)$
13	$(1,-7)$
14	$(3,-13),(5,-3),(7,-1),(7,-5),(19,-1),(22,-14)$
15	$(7,-1),(10,-2)$
18	$(1,-7),(3,-5),(5,-7)$
24	$(3,-5),(5,-3)$
26	$(7,-1)$
30	$(1,-7),(2,-10)$

Table 3:

n	(a,b)	
2	$(1,1-4q)$, $q\neq 1$	$(2^k,4^k-4q)$, $q\equiv 1 \bmod 2$, $(k,q)\neq(1,1)$
3	$(m,4\epsilon-3m^2)$, $m>1$	$(m,4\epsilon\cdot 3^k-3m^2)$, $m\not\equiv 0 \bmod 3$, $(k,m)\neq(1,2)$
4	$(m,2-\epsilon m^2)$, $m>1$, $m\equiv 1 \bmod 2$	$(m,4-\epsilon m^2)$, $m>2$, $m\equiv 0 \bmod 2$
6	$(m,(4-m^2)/3)$, $m\geq 4$, $m\not\equiv 0 \bmod 3$	$(m,(2^{k+2}-\epsilon m^2)/3)$, $m\equiv \pm 1 \bmod 6$, $k\equiv \epsilon' \bmod 2$
	$(m,4-\epsilon m^2/3)$, $m\equiv 0 \bmod 3$	$(m,2^{k+2}-\epsilon m^2/3)$, $m\equiv 3 \bmod 6$

Notation: q is a non-zero integer, k and m are positive integers, $\epsilon\in\{1,-1\}$, $\epsilon'=(1-\epsilon)/2$.

Theorem 1.3 *Any Lucas pair is* 1*-defective, and any Lehmer pair is* 1*- and* 2*-defective.*

For $n\in\{2,3,4,6\}$, *all (up to equivalence)* n*-defective Lucas pairs are of the form* $\left((a-\sqrt{b})/2,(a+\sqrt{b})/2\right)$, *where* (a,b) *are given in Table 3.*

For $n\in\{3,4,5,6,8,10,12\}$, *all (up to equivalence)* n*-defective Lehmer pairs are of the form* $\left((\sqrt{a}-\sqrt{b})/2,(\sqrt{a}+\sqrt{b})/2\right)$, *where* (a,b) *are given in Table 4.*

(See Section 3 for the proof.)

Motivated by extensive computations, one of us conjectured ([40], Conjecture 1) that

$$n_0=30, \tag{4}$$

which is best possible by Theorem C. In [41] the following result in favour of this conjecture was obtained.

Theorem D [41] *If a Lehmer pair* (α,β) *is defective for an integer* $n>30$*, then* $h(\beta/\alpha)>4$.

(Here $h(\cdot)$ stands for the absolute logarithmic height, see the end of this section.)

The main result of this paper confirms the conjecture (4).

Theorem 1.4 *Every integer* $n>30$ *is totally non-defective.*

Theorems C, 1.3 and 1.4 taken together **completely solve the main problem**.

The proof of Theorem 1.4 is long and involved. It equally relies upon heavy mathematics and heavy (rigorous) electronic computations (hence the epigraph!). The most crucial step is solving *many* Thue equations of *very high degree* (see Section 8), using methods developed in [6, 8, 18].

Now a short section-by-section overview of the paper. In Section 2, which relies mainly on the work of Stewart, we establish our basic arithmetic tool: the *cyclotomic criterion*. It has

Table 4:

n	(a,b)	
3	$(1+q, 1-3q)$, $q \neq 1$	$(3^k+q, 3^k-3q)$, $q \not\equiv 0 \bmod 3$, $(k,q) \neq (1,1)$, $k>0$
4	$(1+2q, 1-2q)$, $q \neq 1$	$(2^k+2q, 2^k-2q)$, $q \equiv 1 \bmod 2$, $(k,q) \neq (1,1),(2,1)$, $k>0$
5	$(\phi_{k-2\epsilon}, \phi_{k-2\epsilon}-4\phi_k)$, $k \geq 3$,	$(\psi_{k-2\epsilon}, \psi_{k-2\epsilon}-4\psi_k)$, $k \neq 1$
6	$(1+3q, 1-q)$, $q \neq 1$	$(3^l+3q, 3^l-q)$, $q \not\equiv 0 \bmod 3$, $l>0$
	$(2^k+3q, 2^k-q)$, $q \equiv 1 \bmod 2$, $k>0$	
	$(2^k3^l+3q, 2^k3^l-q)$, $q \equiv \pm 1 \bmod 6$, $k,l>0$	
8	$(\rho_{k-\epsilon}, \rho_{k-\epsilon}-4\pi_k)$, $k \geq 2$,	$(2\pi_{k-\epsilon}, 2\pi_{k-\epsilon}-4\rho_k)$, $k \geq 2$
10	$(\phi_{k-2\epsilon}-4\phi_k, \phi_{k-2\epsilon})$, $k \geq 3$,	$(\psi_{k-2\epsilon}-4\psi_k, \psi_{k-2\epsilon})$, $k \neq 1$
12	$(-\varsigma^{(i)}_{k-2\epsilon}, \varsigma^{(i)}_k)$, $(i,k) \neq (0,0),(0,1),(1,0),(2,0)$.	

Notation: q is a non-zero integer; k and l are non-negative integers; $\epsilon \in \{-1,1\}$; $i \in \{0,1,2,3\}$; $\{\phi_k\}$ is the Fibonacci sequence; $\{\psi_k\}$ is defined from $\psi_0=2$, $\psi_1=1$, $\psi_{k+1}=\psi_k+\psi_{k-1}$; $\{\pi_k\}$ and $\{\rho_k\}$ are defined from $\pi_0=0$, $\pi_1=1$, $\pi_{k+1}=2\pi_k+\pi_{k-1}$ and $\rho_0=\rho_1=1$, $\rho_{k+1}=2\rho_k+\rho_{k-1}$; $\left\{\varsigma^{(i)}_k\right\}$ are defined from $\varsigma^{(i)}_{k+1}=4\varsigma^{(i)}_k-\varsigma^{(i)}_{k-1}$ and the table

i	0	1	2	3
$\varsigma^{(i)}_0$	0	1	1	1
$\varsigma^{(i)}_1$	1	2	3	5

.

been (implicitly) known long ago and explicitly used in [40]–[42], but we found in the literature no complete proof of it. In the same section we show how the cyclotomic criterion reduces enumerating n-defective pairs to solving a norm equation.

In Section 3 we prove Theorem 1.3, which turns out to be a rather straightforward consequence of the cyclotomic criterion and other results of Section 2.

In Section 4 we show that it suffices to prove Theorem 1.4 only for odd square-free values of n.

In Section 5 we show that, for an n-defective Lehmer pair (α, β), the quotient $\gamma = \beta/\alpha$ is "very close" to a primitive n-th root of unity. The main tool is the cyclotomic criterion.

In Section 6 we obtain a non-trivial lower bound for $\arg \gamma^n$, as a consequence of a sharp lower estimate for linear forms in two logarithms due to Laurent, Mignotte, and Nesterenko [21] (in a slightly refined form contained in Mignotte's appendix). Comparing it with the result of the Section 5, and examining the continued fraction expansions of certain algebraic numbers, we prove Theorem 1.4 for many values of n, in particular, for all odd square-free $n > 2145$ and for all prime $n > 787$.

In Section 7 we deviate from our main subject to review the methods for numerical solution of Thue equations of high degree developed in [6, 8, 18].

In Section 8 we prove Theorem 1.4 for the remaining (odd square-free) n by resolving the corresponding Thue equations using the methods described in Section 7.

The appendix, due to Maurice Mignotte, contains the refinement of the result of Laurent-Mignotte-Nesterenko, used in Section 6.

Acknowledgments. We are indebted to Maurice Mignotte, who obtained, at our request, a refinement of the Laurent-Mignotte-Nesterenko theorem, and allowed us to include his manuscript as an appendix to our paper.

We are pleased to thank Alan Baker, Carl Pomerance, Paulo Ribenboim, Andrzej Rotkiewicz, Andrzej Schinzel and Cameron Stewart for stimulating discussions.

Our special gratitude is to Veikko Ennola, who carefully read the manuscript and corrected numerous inaccuracies.

1.1 Terminology, notation and conventions

We use $h(\alpha)$ for the absolute logarithmic height of the algebraic number α. Recall that

$$h(\alpha) := \frac{1}{[\mathbb{K}:\mathbb{Q}]} \sum_v [\mathbb{K}_v:\mathbb{Q}_v] \log\max(1, |\alpha|_v),$$

where $\mathbb{K}$ is any number field containing α and the sum runs over the valuations of the field $\mathbb{K}$ normalized to extend the standard infinite or p-adic valuations of $\mathbb{Q}$. It is well-known (and trivial) that the right-hand side is independent of the choice of $\mathbb{K}$. If $a_m x^m + \cdots + a_0 = a_m(x-\alpha_1)\cdots(x-\alpha_m) \in \mathbb{Z}[x]$ is the minimal polynomial of α over $\mathbb{Z}$ (so that $\gcd(a_0, \ldots, a_m) = 1$) then

$$h(\alpha) = \frac{1}{m}\left(\log|a_m| + \sum_{i=1}^{m} \log\max(1, |\alpha_i|)\right) \tag{5}$$

(see [20], end of Section 3.1). It follows immediately from the definition that

$$h(\alpha \pm \beta) \le h(\alpha) + h(\beta) + \log 2, \quad h(\alpha\beta^{\pm1}) \le h(\alpha) + h(\beta), \quad h(\alpha^n) = |n| h(\alpha).$$

These facts will be used throughout the paper without special reference.

We denote by $\arg z$ the principal value of the argument (that is, $-\pi < \arg z \le \pi$).

We denote by $\|\lambda\|$ the distance from $\lambda \in \mathbb{R}$ to the nearest integer.

We use $O_1(\cdot)$ as a quantitative version of the familiar $O(\cdot)$: $A = O_1(B)$ means $|A| \le B$.

We use the following arithmetical functions:

$P(n)$	the greatest prime divisor of n (with $P(1) = 1$);
$\varphi(n)$	Euler's function;
$\mu(n)$	Möbius function;
$\omega(n)$	the number of distinct prime divisors of n.

As usual, (a, b) stands for the greatest common divisor of a and b. Sometimes, to avoid confusion, we use the notation $\gcd(a, b)$.

We conclude this section by a very simple lemma which will be often used, sometimes without special reference.

Lemma 1.5 *Let p and q be rational integers, $\sqrt{p}$ an arbitrary square root of p, and α, β the roots of $X^2 - X\sqrt{p} + q$. Then:*

i. The algebraic number $\gamma = \beta/\alpha$ is of degree at most 2.

ii. If $\gcd(p, q) = 1$ then γ is a root of unity if and only if

$$(p, q) \in \{\pm(0,1), \pm(1,1), \pm(2,1), \pm(3,1), \pm(4,1)\}. \tag{6}$$

iii. If $\alpha^2, \beta^2 \notin \mathbb{R}$ then $\deg\gamma = 2$ and $|\alpha| = |\beta|$ (and henceforth $|\gamma| = 1$).

iv. If $\alpha^2, \beta^2 \notin \mathbb{R}$ and $\gcd(p, q) = 1$ then $h(\gamma) = \log|\alpha| = \log|\beta|$.

In particular, if (α, β) is a non-real Lehmer pair then $\deg\gamma = 2$, $|\gamma| = 1$ and $h(\gamma) = \log|\alpha| = \log|\beta|$.

Proof The numbers γ and γ^{-1} are the roots of $qX^2 - (p-2q)X + q$. This proves (i).

If $\gcd(p, q) = 1$ then γ can be a root of unity when and only when $q = \pm1$ and $|p - 2q| \le 2$, which is equivalent to (6).

The numbers α^2, β^2 are the roots of $X^2 - (p-2q)X + q^2$. If they are non-real then $\alpha^2 \ne \beta^2$, and α^2, β^2 are complex conjugate. Hence $|\alpha| = |\beta|$ and $|\gamma| = 1$. If in this case $\deg\gamma = 1$ then $\gamma = \pm1$, which implies $\alpha^2 = \beta^2$, a contradiction. Therefore $\deg\gamma = 2$.

Finally, if $\alpha^2, \beta^2 \notin \mathbb{R}$ and $\gcd(p,q) = 1$ then $qX^2 - (p-2q)X + q$ is the minimal polynomial of γ over $\mathbb{Z}$. By (5),

$$h(\gamma) = \frac{1}{2}\left(\log|q| + \max(0, \log|\gamma|) + \max(0, \log|\gamma^{-1}|)\right) = \frac{1}{2}\log|q| = \log|\alpha|.$$

The lemma is proved. ∎

2 Cyclotomic criterion and norm equation

In this section we establish our main arithmetical tool: the *cyclotomic criterion*. Though it has been known, at least implicitly, long ago, we found in the available literature no complete proof of it. The aim of the present section is to fill this gap.

Our argument relies upon the results of Stewart [38], though Stewart himself credits them to much earlier authors, notably Lucas [25], Carmichael [11] and Lehmer [22].

Proposition 2.1 *Let (α, β) be a Lehmer pair and $\{\tilde{u}_n\}$ the corresponding sequence of Lehmer numbers. Then:*

i. *For all positive integers n we have $(\alpha\beta, \tilde{u}_n) = 1$.*

ii. *If $d|n$ then $\tilde{u}_d|\tilde{u}_n$ and $(\tilde{u}_n/\tilde{u}_d, \tilde{u}_d)$ divides n/d.*

iii. *For all positive integers m and n we have $(\tilde{u}_m, \tilde{u}_n) = \tilde{u}_{(m,n)}$.*

iv. *If a prime p does not divide $\alpha\beta(\alpha^2-\beta^2)^2$ then p divides $\tilde{u}_{p-1}\tilde{u}_{p+1}$.*

v. *If a prime p divides $\tilde{u}_m$ then p divides $\tilde{u}_{mp}/\tilde{u}_m$.*

vi. *If in the previous item $p > 2$ then p exactly[2] divides $\tilde{u}_{mp}/\tilde{u}_m$.*

vii. *If $4|\tilde{u}_m$ then 2 exactly divides $\tilde{u}_{2m}/\tilde{u}_m$.*

viii. *If a prime $p > 2$ divides $(\alpha-\beta)^2$ then p divides $\tilde{u}_p$; if $p > 3$ then p exactly divides $\tilde{u}_p$.*

ix. *If a prime p divides $(\alpha+\beta)^2$ then p divides $\tilde{u}_{2p}$; if $p > 3$ then p exactly divides $\tilde{u}_{2p}$.*

Proof This is a summary of Lemmas 1–5 from [38]. Only assertion $\tilde{u}_d|\tilde{u}_n$ from item (ii) is not formally stated by Stewart but it is quite classical and easy to prove. (Indeed, if $n \equiv d \bmod 2$ then $\tilde{u}_n/\tilde{u}_d = (\alpha^n - \beta^n)/(\alpha^d - \beta^d) \in \mathbb{Z}$. This reduces the assertion to the case when d is odd and $n = 2d$, when $\tilde{u}_n/\tilde{u}_d = (\alpha^d + \beta^d)/(\alpha+\beta) \in \mathbb{Z}$.) Notice that Stewart's u_n corresponds to our $\tilde{u}_n$. ∎

Corollary 2.2 *Let (α, β) be a Lehmer pair and $\{\tilde{u}_n\}$ the corresponding sequence of Lehmer numbers. For any prime p not dividing $\alpha\beta$ there exists a positive integer m such that $p|\tilde{u}_m$. Let m_p be the smallest m with this property. Then*

$$p|\tilde{u}_m \Longleftrightarrow m_p|m, \tag{7}$$

and

$$\begin{aligned} &m_p = p && \text{if } p > 2 \text{ and } p|(\alpha-\beta)^2, && (8)\\ &m_p = 2p && \text{if } p|(\alpha+\beta)^2, && (9)\\ &m_p|(p-1) \text{ or } m_p|(p+1) && \text{otherwise.} && (10) \end{aligned}$$

In particular, if 2 does not divide $\alpha\beta$ then

$$m_2 = \begin{cases} 4, & \text{if } 2|(\alpha^2-\beta^2)^2, \\ 3, & \text{otherwise,} \end{cases} \tag{11}$$

[2] That is, p divides that number but p^2 does not.

and if 3 *does not divide* $\alpha\beta$ *then*

$$m_3 = \begin{cases} 3 \text{ or } 6, & \text{if } 3|(\alpha^2-\beta^2)^2, \\ 4, & \text{otherwise.} \end{cases} \tag{12}$$

Proof The existence of m such that $p|\widetilde{u}_m$ follows from items (iv), (viii) and (ix). The "$\Leftarrow$"-implication in (7) follows from item (ii) and the "$\Rightarrow$"-implication follows from item (iii). Further, (8) and (10) follow from (7) and items (viii) and (iv), respectively.

We are left with (9). It follows from (7) and item (ix) that $m_p = p$ or $m_p = 2p$ when $p|(\alpha+\beta)^2$. To establish (9), we have to prove that in this case p does not divide $\widetilde{u}_p$. For $p = 2$ this is obvious. Now assume that $p > 2$. By the binomial formula,

$$\alpha^p - \beta^p \equiv (\alpha-\beta)^p \bmod p. \tag{13}$$

Since $\gcd\left((\alpha+\beta)^2, (\alpha-\beta)^2\right)$ divides 4, and since $p|(\alpha+\beta)^2$, the algebraic integer $\alpha - \beta$ is prime to p. Therefore (13) yields

$$\widetilde{u}_p \equiv (\alpha-\beta)^{p-1} \bmod p. \tag{14}$$

Since the right-hand side of (14) is not divisible by p, so is the left-hand side. The corollary is proved. ∎

The following two polynomials will play crucial role in the sequel. The first one is $\Phi_n(X,Y)$, the homogeneous cyclotomic polynomial of order n:

$$\Phi_n(X,Y) = \prod_{\substack{1\le k<n \\ (k,n)=1}} \left(Y - e^{2\pi ik/n}X\right) \in \mathbb{Z}(X,Y).$$

The second one is $F_n(X,Y)$, the homogeneous *real* cyclotomic polynomial of order $n > 2$:

$$F_n(X,Y) = \prod_{\substack{1\le k<n/2 \\ (k,n)=1}} (Y - 2\cos(2\pi k/n)\cdot X) \in \mathbb{Z}(X,Y).$$

The two polynomials are related by the identity

$$\Phi_n(X,Y) = F_n(XY, X^2+Y^2). \tag{15}$$

If (α,β) is a Lehmer pair, then

$$x = \alpha\beta \quad \text{and} \quad y = \alpha^2+\beta^2 \tag{16}$$

are rational integers, which implies that $\Phi_n(\alpha,\beta) = F_n(x,y) \in \mathbb{Z}$ for $n > 2$.

Stewart [38], Lemma 6, showed that the prime divisors of the rational integer $\Phi_n(\alpha,\beta)$ satisfy very restrictive conditions. The following trivial observation is crucial for his argument[3]:

$$(n > 2,\ d < n,\ d|n) \Longrightarrow \Phi_n|(\widetilde{u}_n/\widetilde{u}_d). \tag{17}$$

We shall not need the full strength of Stewart's result, but only the following simple fact.

Proposition 2.3 (Stewart) *Let* (α,β) *be a Lehmer pair, and* p *a prime divisor of* $\Phi_n(\alpha,\beta)$, *where* $n > 2$. *Then* p *does not divide* $\alpha\beta$ *and* $n = m_p p^k$, *where* k *is a non-negative integer.*

Proof By (17), p divides $\widetilde{u}_n$. By Proposition 2.1 (i), p does not divide $\alpha\beta$. It then follows from (7) that $m_p|n$. Write $n = m_p t p^k$, where $(t,p) = 1$. If $t > 1$ then, again by (17), we have $p|\left(\widetilde{u}_n/\widetilde{u}_{n/t}\right)$. Also, $p|\widetilde{u}_{n/t}$, which again follows from (7). Now Proposition 2.1 (ii) implies that $p|t$, a contradiction. This proves that $t = 1$. The proposition is proved. ∎

Now we are ready to formulate and prove the cyclotomic criterion. For an integer $n > 3$ put $P'(n) = P(n/(n,3))$ (recall that $P(n)$ stands for the maximal prime divisor of n). In other words,

$$P'(n) = \begin{cases} 2, & \text{if } n \text{ is of the form } 2^k\cdot 3, \\ P(n), & \text{otherwise.} \end{cases}$$

[3] When the Lehmer pair (α,β) is fixed, we write sometimes Φ_n instead of $\Phi_n(\alpha,\beta)$.

Theorem 2.4 (cyclotomic criterion) *Let $n > 4$ be an integer distinct from 6 and 12. Then a Lehmer pair (α, β) is n-defective if and only if $\Phi_n(\alpha, \beta) \in \{\pm 1, \pm P'(n)\}$. Also, a Lehmer pair (α, β) is 12-defective if and only if $\Phi_{12}(\alpha, \beta) \in \{\pm 1, \pm 2, \pm 3, \pm 6\}$.*

Proof We start from the "**only if**" statement. Thus, let n be an a an integer satisfying

$$n > 4 \quad \text{and} \quad n \neq 6, \tag{18}$$

and (α, β) an n-defective Lehmer pair.

Let p be a prime divisor of Φ_n. By Proposition 2.3, p does not divide $\alpha\beta$ and $n = m_p p^k$, where $k \geq 0$. Since p is not a primitive divisor of $\widetilde{u}_n$, we have one of the following options:

$$n = m_p p^k \quad \text{with} \quad k \geq 1; \tag{19}$$

$$n = m_p \quad \text{and} \quad p|(\alpha^2 - \beta^2)^2. \tag{20}$$

It follows from Corollary 2.2 that, whichever of (19) and (20) holds, we have

$$p = \begin{cases} P'(n), & \text{if } n \neq 12, \\ 2 \text{ or } 3, & \text{if } n = 12. \end{cases} \tag{21}$$

To complete the proof of the "only if" assertion, it remains to establish the following:

$$p \text{ exactly divides } \Phi_n. \tag{22}$$

We consider two cases, which correspond to (19) and (20), respectively.

Case 1 $n = m_p p^k$ with $k \geq 1$

In this case $m_p|(n/p)$. Hence $p|\widetilde{u}_{n/p}$.

Subcase 1.a $p > 2$

Proposition 2.1 (vi) implies that p exactly divides $\widetilde{u}_n/\widetilde{u}_{n/p}$. It follows now from (17) that p exactly divides Φ_n.

Subcase 1.b $p = 2$ and 2 does not divide $(\alpha^2 - \beta^2)^2$

We have $m_2 = 3$ by Corollary 2.2, whence $n = 3 \cdot 2^k$ with $k \geq 2$. (If $k = 1$ then $n = 6$, which is contradicts the assumption.) Notice that $\Phi_3 = \widetilde{u}_3$ is even, and that $\Phi_3 - \Phi_6 = 2\alpha\beta \equiv 2 \bmod 4$. Hence 4 divides at least one of the numbers Φ_3 and Φ_6. Therefore 4 divides at least one of the numbers $\widetilde{u}_3$ and $\widetilde{u}_6$. By Proposition 2.1 (ii), $4|\widetilde{u}_6$. Since $6|(n/2)$, Proposition 2.1 (ii), again, implies that $4|\widetilde{u}_{n/2}$. Proposition 2.1 (vii) implies that 2 exactly divides $\widetilde{u}_n/\widetilde{u}_{n/2}$, and we complete the proof using (17).

Subcase 1.c $p = 2$ and $2|(\alpha^2 - \beta^2)^2$

In this case $m_2 = 4$, and $n = 2^k$ with $k \geq 3$. Using induction in k, we shall prove that 2 exactly divides Φ_{2^k} for $k \geq 3$. Indeed, $2|(\alpha^2 + \beta^2) = \Phi_4$, whence $4|(\alpha^2 + \beta^2)^2$. On the other hand, $2\alpha\beta \equiv 2 \bmod 4$, whence 2 exactly divides $\alpha^4 + \beta^4 = \Phi_8$.

Similarly, if $2|\left(\alpha^{2^{k-1}} + \beta^{2^{k-1}}\right) = \Phi_{2^k}$, then 2 exactly divides $\alpha^{2^k} + \beta^{2^k} = \Phi_{2^{k+1}}$. This proves (22) also in this subcase.

Case 2 $m_p = n$ and $p|(\alpha^2 - \beta^2)^2$

In this case $p > 3$, since otherwise $n \in \{3, 4, 6\}$, which is not allowed by the assumption. Hence p exactly divides $\widetilde{u}_n = \widetilde{u}_{m_p}$ by Proposition 2.1 (viii) and (ix). This proves the "only if" part.

Now establish the "**if**" statement. Let n be an integer satisfying (18) and (α, β) a Lehmer pair such that

$$\Phi_n \in \begin{cases} \{\pm 1, \pm P'(n)\}, & \text{if } n \neq 12, \\ \{\pm 1, \pm 2, \pm 3, \pm 6\}, & \text{if } n = 12. \end{cases} \tag{23}$$

We have to prove that (α, β) is n-defective.

Let p be a prime divisor of $\tilde{u}_n$. It follows from

$$\beta^n - \alpha^n = \prod_{d|n} \Phi_d(\alpha, \beta) \tag{24}$$

that $\tilde{u}_n$ belongs to the multiplicative group generated by the numbers Φ_m with $m > 2$ and $m|n$ and the numbers $(\alpha+\beta)^2$, $(\alpha^2-\beta^2)^2$. Hence p divides either one of those Φ_m or one of the numbers $(\alpha+\beta)^2$, $(\alpha^2-\beta^2)^2$.

If p divides Φ_m with $m < n$ then it cannot be a primitive divisor, as well as if it divides one of the $(\alpha+\beta)^2$, $(\alpha^2-\beta^2)^2$.

If p divides Φ_n then, by (23), it satisfies (21). In particular, $p|n$. If $m_p < n$ then p again is not a primitive divisor. If $m_p = n$ then $p|m_p$, which implies, by (10), that $p|(\alpha^2-\beta^2)^2$.

Thus, in no case can p be a primitive divisor of $\tilde{u}_n$. The theorem is proved. ■

Corollary 2.5 *An integer $n > 30$ is totally non-defective if the equation*

$$F_n(x, y) \in \{\pm 1, \pm P'(n)\} \tag{25}$$

has no solutions $(x, y) \in \mathbb{Z}^2$ with $|x| > e^8$.

Proof Let (α, β) be an n-defective Lehmer pair. Then $x = \alpha\beta$ and $y = \alpha^2 + \beta^2$ give a solution of (25). By Theorem A, $\alpha^2, \beta^2 \notin \mathbb{R}$. By Lemma (1.5), $|\alpha| = |\beta| = e^{h(\beta/\alpha)}$. By Theorem D, $h(\beta/\alpha) > 4$. Hence $|x| > e^8$. The corollary is proved. ■

3 Small n

In this section we prove Theorem 1.3, extending the argument outlined by Stewart in [37], Theorem 3. Obviously, any Lucas pair and any Lehmer pair is 1-defective, and any Lehmer pair is 2-defective. Also, it is easy to verify that if (a, b) is from Table 3, then $\left((a+\sqrt{b})/2, (a-\sqrt{b})/2\right)$ is an n-defective Lucas pair for the corresponding n, and if (a, b) from Table 4 then $\left((\sqrt{a}+\sqrt{b})/2, (\sqrt{a}-\sqrt{b})/2\right)$ is an n-defective Lehmer pair for the corresponding n.

It remains to show that for $n \in \{2, 3, 4, 6\}$ (respectively, $n \in \{3, 4, 5, 6, 8, 10, 12\}$) there are no (up to equivalence) n-defective Lucas (respectively, Lehmer) pairs other than those mentioned in the previous paragraph.

In this section (α, β) is a Lehmer pair. We use the notation $p = (\alpha+\beta)^2$, $q = \alpha\beta$, and we define $\sqrt{p}$ uniquely as $\alpha + \beta$, so that

$$\alpha, \beta = \frac{\sqrt{p} \pm \sqrt{p - 4q}}{2}.$$

We have

$$q \neq 0, \tag{26}$$

$$\gcd(\tilde{u}_n, q) = \gcd(\Phi_n, q) = \gcd(p, q) = 1, \tag{27}$$

$$(p, q) \notin \{\pm(1,1), \pm(2,1), \pm(3,1), \pm(4,1)\}. \tag{28}$$

If (α, β) is a Lucas pair then we put $m = \alpha + \beta = \sqrt{p}$, so that

$$\alpha, \beta = \frac{m \pm \sqrt{p - 4q}}{2}.$$

We may assume that

$$m > 0,$$

replacing (α, β) by an equivalent Lucas pair, and we have

$$\gcd(m, q) = 1, \tag{29}$$

$$(m, q) \notin \{(1,1), (2,1)\}. \tag{30}$$

$\underline{n=2}$

Let (α,β) be a 2-defective Lucas pair. Then every prime divisor of $u_2=m$ divides $(\alpha-\beta)^2=m^2-4q$. It follows from (29) that the only possible prime divisor of m is 2. Since $m>0$, we have $m=2^k$, where k is a non-negative integer. Again by (29), we have either $k=0$ or $q\equiv 1 \bmod 2$. Also, it follows from (30) that $(k,q)\notin\{(0,1),(1,1)\}$. Hence $(a,b)=(m,p-4q)$ is of one of the two types displayed in Table 3.

$\underline{n=3}$

Let (α,β) be a 3-defective Lehmer pair. Then every prime divisor of $\widetilde{u}_3=p-q$ divides $(\alpha^2-\beta^2)^2=(\widetilde{u}_3+q)(\widetilde{u}_3-3q)$. It follows from (27) that the only possible prime divisor of $\widetilde{u}_3$ is 3. Replacing (α,β) by an equivalent Lehmer pair, we may assume that $\widetilde{u}_3>0$. Hence $\widetilde{u}_3=3^k$, where k is a non-negative integer. Again by (27), we have either $k=0$ or $q\not\equiv 0 \bmod 3$. Since $p=3^k+q$, it follows from (28) that $(k,q)\notin\{(0,1),(1,1)\}$. Hence $(a,b)=(p,p-4q)$ is of one of the two types displayed in Table 4.

If (α,β) or $(\alpha i,\beta i)$ is a Lucas pair, then $p=\epsilon m^2$, where $\epsilon=1$ or -1 depending on whether we consider (α,β) or $(\alpha i,\beta i)$. It follows that $p-4q=4\cdot 3^k-3\epsilon m^2$. Hence $(a,b)=(m,(p-4q)\epsilon)$ is of one of the two types displayed in Table 3.

$\underline{n=4}$

Let (α,β) be a 4-defective Lehmer pair. Since $(\widetilde{u}_3,\widetilde{u}_4)=1$, every prime divisor of $\widetilde{u}_4=p-2q$ divides $(\alpha^2-\beta^2)^2=(\widetilde{u}_4+2q)(\widetilde{u}_4-2q)$. Hence the only possible prime divisor of $\widetilde{u}_4$ is 2. Replacing (α,β) by an equivalent Lehmer pair, we obtain $\widetilde{u}_4=2^k$, where either $k=0$ or $q\not\equiv 0 \bmod 2$. Since $p=2^k+2q$, it follows from (28) that $(k,q)\notin\{(0,1),(1,1),(2,1)\}$. Hence $(a,b)=(p,p-4q)$ is of one of the two types displayed in Table 4.

If (α,β) or $(\alpha i,\beta i)$ is a Lucas pair, then $p=2^k+2q=\epsilon m^2$. This is impossible when $k>1$, because q is odd. If $k=0$ then m is odd, and if $k=1$ then m is even. Since $p-4q=2^{k+1}-\epsilon m^2$, the pair $(a,b)=(m,(p-4q)\epsilon)$ is of one of the two types displayed in Table 3.

$\underline{n=5}$

Let (α,β) be a 5-defective Lehmer pair. By the cyclotomic criterion,

$$\Phi_5=(p-\eta^2q)(p-\overline{\eta}^2q)\in\{\pm1,\pm5\},$$

where $\eta=\frac{1+\sqrt5}{2}$ and $\overline{\eta}=\frac{1-\sqrt5}{2}$. Since η is the fundamental unit of $\mathbb{Q}(\sqrt5)$, we have either

$$p-\eta^2q=\epsilon_0\eta^{\epsilon k},\quad p-\overline{\eta}^2q=\epsilon_0\overline{\eta}^{\epsilon k}, \tag{31}$$

or

$$p-\eta^2q=\epsilon_0\sqrt5\eta^{\epsilon k},\quad p-\overline{\eta}^2q=-\epsilon_0\sqrt5\overline{\eta}^{\epsilon k}, \tag{32}$$

where $\epsilon_0,\epsilon\in\{1,-1\}$ and k is a non-negative integer. Replacing (α,β) by an equivalent Lehmer pair, we may assume that

$$\epsilon_0=-\epsilon^{k+1}\text{ in (31) and }\epsilon_0=-\epsilon^k\text{ in (32).} \tag{33}$$

Resolving the linear equations (31) and (32), and using (33), we obtain, respectively, either

$$\begin{aligned} p&=\epsilon^{k+1}\frac{\eta^{\epsilon k-2}-\overline{\eta}^{\epsilon k-2}}{\sqrt5}=\frac{\eta^{k-2\epsilon}-\overline{\eta}^{k-2\epsilon}}{\sqrt5}=\phi_{k-2\epsilon},\\ q&=\epsilon^{k+1}\frac{\eta^{\epsilon k}-\overline{\eta}^{\epsilon k}}{\sqrt5}=\frac{\eta^k-\overline{\eta}^k}{\sqrt5}=\phi_k,\end{aligned} \tag{34}$$

where $\{\phi_k\}$ is the Fibonacci sequence, or

$$\begin{aligned} p&=\epsilon^k\left(\eta^{\epsilon k-2}+\overline{\eta}^{\epsilon k-2}\right)=\eta^{k-2\epsilon}+\overline{\eta}^{k-2\epsilon}=\psi_{k-2\epsilon},\\ q&=\epsilon^k\left(\eta^{\epsilon k}+\overline{\eta}^{\epsilon k}\right)=\eta^k+\overline{\eta}^k=\psi_k,\end{aligned} \tag{35}$$

where $\{\psi_k\}$ is the classical *Lucas sequence* ([33], p. 43) defined from $\psi_0 = 2, \quad \psi_1 = 1$ and $\psi_{k+1} = \psi_k + \psi_{k-1}$.

By (28), we have $k \geq 3$ in the case (34), and $k \neq 1$ in the case (35). Hence $(a, b) = (p, p - 4q)$ is of one of the two types displayed in Table 4.

$\underline{n = 6}$

Let (α, β) be a 6-defective Lehmer pair. Then every prime divisor of Φ_6 divides $(\alpha^2 - \beta^2)^2 \tilde{u}_3 = (\Phi_6 + 3q)(\Phi_6 - q)(\Phi_6 + 2q)$. By (27), the only possible prime divisors of Φ_6 are 2 and 3. Replacing (α, β) by an equivalent Lehmer pair, we may assume that $\Phi_6 > 0$. We obtain $\Phi_6 = 2^k 3^l$, where k and l are non-negative integers. Again by (27), we have one of the four options:

$$\begin{array}{ll} k = l = 0; & k = 0 \text{ and } q \not\equiv 0 \bmod 3; \\ l = 0 \text{ and } q \not\equiv 0 \bmod 2; & q \equiv \pm 1 \bmod 6. \end{array}$$

Since $p = 2^k 3^l + 3q$, it follows from (28) that $(k, l, q) \neq (0, 0, 1)$. Hence $(a, b) = (p, p - 4q)$ is of one of the four types displayed in Table 4.

If (α, β) or $(\alpha i, \beta i)$ is a Lucas pair, then $p = 2^k 3^l + 3q = \epsilon m^2$. This is impossible when $l > 1$, because $q \not\equiv 0 \bmod 3$. When $l = 0$ we have $p = \epsilon m^2 = 2^k + 3q$, whence

$$\begin{array}{ll} \epsilon = 1 \text{ and } m \not\equiv 0 \bmod 3 & \text{when } k = 0, \\ m \equiv \pm 1 \bmod 6 \text{ and } k \equiv \epsilon' \bmod 2 & \text{when } k > 0. \end{array}$$

Here $\epsilon' = (1 - \epsilon)/2$. Since $p - 4q = (2^{k+2} - \epsilon m^2)/3$, this gives rise to the first two (depending on whether $k = 0$ or $k > 0$) types of $(a, b) = (m, (p - 4q)\epsilon)$ displayed in Table 3.

When $l = 1$ we have $p = \epsilon m^2 = 3 \cdot 2^k + 3q$, whence $m \equiv 0 \bmod 3$, and $m \equiv 3 \bmod 6$ if $k > 0$. We have $p - 4q = 2^{k+2} - \epsilon m^2/3$, which leads to the next two types.

$\underline{n = 8}$

Let (α, β) be an 8-defective Lehmer pair. By the cyclotomic criterion,

$$\Phi_8 = (p - q\eta\sqrt{2})(p + q\overline{\eta}\sqrt{2}) \in \{\pm 1, \pm 2\},$$

where $\eta = 1 + \sqrt{2}$ and $\overline{\eta} = 1 - \sqrt{2}$. Since η is the fundamental unit of $\mathbb{Q}(\sqrt{2})$, we have one of the two options

$$p - q\eta\sqrt{2} = \epsilon_0 \eta^{\epsilon k}, \qquad p + q\overline{\eta}\sqrt{2} = \epsilon_0 \overline{\eta}^{\epsilon k}, \tag{36}$$

$$p - q\eta\sqrt{2} = \epsilon_0 \sqrt{2}\eta^{\epsilon k}, \qquad p + q\overline{\eta}\sqrt{2} = -\epsilon_0 \sqrt{2}\overline{\eta}^{\epsilon k}, \tag{37}$$

where $\epsilon_0, \epsilon \in \{1, -1\}$ and k is a non-negative integer. Replacing (α, β) by an equivalent Lehmer pair, we may assume that

$$\epsilon_0 = -\epsilon^{k+1} \text{ in (36) and } \epsilon_0 = -\epsilon^k \text{ in (37)}. \tag{38}$$

Resolving the linear equations and using (38), we obtain $p = \rho_{k-\epsilon}, \quad q = \pi_k$ in case (36) and $p = 2\pi_{k-\epsilon}, \ q = \rho_k$ in case (37), where the sequences $\{\pi_k\}$ and $\{\rho_k\}$ are defined in Table 4. (Mention that $\{\pi_k\}$ is known as *Pell sequence*, and $\{2\rho_k\}$ is the *companion Pell sequence*, see [33], pp. 43–44.)

By (28), we have $k \geq 2$. Hence $(a, b) = (p, p - 4q)$ is of one of the two types displayed in Table 4.

$\underline{n = 10}$

One verifies immediately that $\Phi_{10}(\alpha, \beta) = \Phi_5(-\alpha, \beta)$. Hence (a, b) enters the line corresponding to $n = 10$ if and only if (b, a) enters the line corresponding to $n = 5$.

$\underline{n = 12}$

Let (α, β) be a 12-defective Lehmer pair. By the cyclotomic criterion,

$$\Phi_8 = (p - q\eta)(p - q\overline{\eta}) \in \{\pm 1, \pm 2, \pm 3, \pm 6\},$$

where $\eta = 2 + \sqrt{3}$ and $\overline{\eta} = 2 - \sqrt{3}$. Since η is the fundamental unit of $\mathbb{Q}(\sqrt{3})$, and since $2 = -\theta\overline{\theta} = -\theta^2\eta$, where $\theta = 1 + \sqrt{3}$ and $\overline{\theta} = 1 - \sqrt{3}$, we have one of the following four options:

$$p - q\eta = \epsilon_0 \eta^{\epsilon k}, \qquad p - q\overline{\eta} = \epsilon_0 \overline{\eta}^{\epsilon k}, \tag{39}$$

$$p - q\eta = \epsilon_0 \sqrt{3} \eta^{\epsilon k}, \qquad p - q\overline{\eta} = -\epsilon_0 \sqrt{3} \overline{\eta}^{\epsilon k}, \tag{40}$$

$$p - q\eta = \epsilon_0 \theta \eta^{\epsilon k}, \qquad p - q\overline{\eta} = \epsilon_0 \overline{\theta} \overline{\eta}^{\epsilon k}, \tag{41}$$

$$p - q\eta = \epsilon_0 \sqrt{3} \theta \eta^{\epsilon k}, \qquad p - q\overline{\eta} = -\epsilon_0 \sqrt{3} \overline{\theta} \overline{\eta}^{\epsilon k}, \tag{42}$$

where $\epsilon_0, \epsilon \in \{1, -1\}$ and k is a non-negative integer. Replacing (α, β) by an equivalent Lehmer pair, we may assume that

$$\epsilon_0 = -\epsilon \text{ in (39) and (41), and } \epsilon_0 = -1 \text{ in (40) and (42).} \tag{43}$$

Resolving the linear equations and using (43), we obtain $p = \varsigma_{k-\epsilon}^{(i)}$ and $q = \varsigma_k^{(i)}$, where $i = 0, \ldots, 3$ corresponds to (39)–(42), respectively, and the four sequences $\left\{\varsigma_k^{(i)}\right\}$ are defined in Table 4.

By (28), we have $(i, k) \neq (0, 0), (0, 1), (1, 0), (2, 0)$. Finally, $\varsigma_k^{(i)} - 4\varsigma_{k-\epsilon}^{(i)} = -\varsigma_{k-2\epsilon}^{(i)}$. Hence $(a, b) = (p, p - 4q)$ is as displayed in Table 4.

4 Reduction to odd square-free numbers

In this section we show that it suffices to prove Theorem 1.4 only for odd square-free numbers n.

It will be convenient to introduce a special partial ordering on the set of positive integers. Let n, n' be two positive integers. We say that n *dominates over* n' (notation: $n \succ n'$ or $n' \prec n$) if one of the following two conditions is satisfied.

i. $n = 2^k \cdot 3$ and $n' = 2^{k'} \cdot 3$ with $k' \geq k$;

ii.
- n is not of the form $2^k \cdot 3$ and
- the numbers n and n' have the same sets of odd prime divisors and $n | n'$; (in symbols: $n = 2^{k_0} p_1^{k_1} \cdots p_s^{k_s}$ and $n' = 2^{k'_0} p_1^{k'_1} \cdots p_s^{k'_s}$, where $p_1, \ldots, p_s$ are distinct odd primes ($s = 0$ allowed), and $k'_0 \geq k_0 \geq 0$, $k'_1 \geq k_1 > 0$, $\ldots$, $k'_s \geq k_s > 0$).

Thus, every positive integer is dominated either by an odd square-free integer or by 9.

Proposition 4.1 *Let $n \succ n'$, and put*

$$t = t(n, n') = \begin{cases} n'/n & \text{if } n \text{ is even or } n' \text{ is odd,} \\ n'/(2n) & \text{if } n \text{ is odd and } n' \text{ is even,} \end{cases} \tag{44}$$

$$\epsilon = \epsilon(n, n') = \begin{cases} 1 & \text{if } n \text{ is even or } n' \text{ is odd,} \\ -1 & \text{if } n \text{ is odd and } n' \text{ is even,} \end{cases} \tag{45}$$

Assume that $n > 4$ and $n \neq 6, 12$ and let (α, β) be an n'-defective Lehmer pair. Then $(\epsilon\alpha^t, \beta^t)$ is an n-defective Lehmer pair.

Proof It is easy to see that $n' \prec n$ yields $\Phi_{n'}(X, Y) = \Phi_n(\epsilon X^t, Y^t)$ and $P'(n') = P'(n)$. Hence the result follows from the cyclotomic criterion. ■

Corollary 4.2 *If $n' \prec n$ and n is totally non-defective, then so is n'.* ■

Corollary 4.3 *Assume that $n' \prec n$, $n > 4$, $n \neq 6, 12$ and $n' > 30$. If $h(\alpha/\beta) \leq 4$ for any n-defective Lehmer pair (α, β), then n' is totally non-defective.*

Proof Follows from Proposition 4.1 and Theorem D. ■

Proposition 4.4 *Let n be an integer satisfying $6 < n \leq 30$ and $n \neq 8, 10, 12$. Then for every n-defective Lehmer pair (α, β) we have $h(\alpha/\beta) \leq 4$.*

Proof Follows from Theorem C. ■

Theorem 4.5 *Theorem 1.4 is a consequence of the following formally weaker assertion: every odd square-free integer $n > 30$ is totally non-defective.*

Proof Assume that every odd square-free integer $n > 30$ is totally non-defective, and prove that this holds for every integer $n' > 30$.

If n' is dominated by an odd square-free integer $n > 30$, then n' is totally non-defective by Corollary 4.2.

Otherwise, the product of odd prime divisors of n' does not exceed 30. Therefore n' is dominated by a prime number not exceeding 30 or by one of the numbers 1, 9, 15, 21. Every $n' > 30$ dominated by 1 is dominated by 16, every $n' > 30$ dominated by 3 is dominated by 24, and every $n' > 30$ dominated by 5 is dominated by 20 or 25. Hence n' is dominated by one of the following numbers:

$$7,\ 9,\ 11,\ 13,\ 15,\ 16,\ 17,\ 19,\ 20,\ 21,\ 23,\ 24,\ 25,\ 29.$$

Therefore n' is totally non-defective by Corollary 4.3 and Proposition 4.4. ■

5 The quotient β/α is close to a root of unity

In this section we show that for an n-defective Lehmer pair (α, β), the quotient β/α is extremely close to a primitive n-th root of unity. The basic argument goes back to Schinzel [36] and Stewart [37], but the idea of using (47) and (49) seems to be new.

Theorem 5.1 *Let $n \geq 31$ be a square-free integer and (α, β) an n-defective Lehmer pair*[4]*. Put $\gamma = \beta/\alpha$. Then there exists a (single) primitive n-th root of unity ξ such that*

$$0 < \phi := |\arg(\gamma\xi^{-1})| < \min\left(\pi/n, c_1(n)|\alpha|^{-\varphi(n)}\right), \tag{46}$$

where

$$c_1(n) = \begin{cases} \pi & \text{if } n \text{ is prime,} \\ \pi n^{2^{\omega(n)-2}-1} P(n) & \text{if } n \text{ is composite.} \end{cases}$$

(Recall that $\varphi(n)$ is Euler's function, $P(n)$ is the maximal prime divisor of n, and $\omega(n)$ is the number of distinct prime divisors of n.)

Proof Since γ is not a root of unity, there exists a single n-th root of unity ξ such that $\phi := |\arg(\gamma\xi^{-1})| < \pi/n$. Let m be the divisor of n such that ξ is a primitive m-th root of unity. Then

$$\begin{array}{ll} |\arg\gamma^d| > d\pi/n & \text{if } d|n \text{ but } m \not| d, \\ |\arg\gamma^d| = d\phi < d\pi/n & \text{if } m|d|n. \end{array} \tag{47}$$

For any $z \in \mathbb{C}$ with $|z| = 1$ one has

$$(2/\pi)|\arg z| \leq |z-1| \leq |\arg z|. \tag{48}$$

Therefore

$$\begin{array}{rcccl} 2d/n & \leq & |\gamma^d - 1| & \leq & 2 \quad \text{if } d|n \text{ but } m \not| d, \\ (2/\pi)d\phi & \leq & |\gamma^d - 1| & \leq & d\phi \quad \text{if } m|d|n. \end{array} \tag{49}$$

[4] which is non-real by Theorem A

Also, we need the following arithmetical identities, the first two of them being well known, and the last two easy to prove:

$$\sum_{d|N}\mu(N/d)=\delta(N);\qquad \sum_{d|N}\mu(N/d)\log d=\Lambda(N);$$
$$\sum_{\substack{d|N\\ \mu(N/d)=1}}1=2^{\omega(N)-1}+\tfrac{\delta(N)}{2};\qquad \sum_{\substack{d|N\\ \mu(N/d)=1}}\log d=2^{\omega(N)-2}\log N+\tfrac{\Lambda(N)}{2}\quad (N \text{ is square-free}).$$

Here $\mu(N)$ is the Möbius function, $\Lambda(N)$ is the von Mangoldt function and

$$\delta(n)=\begin{cases}1, & \text{if } n=1,\\ 0, & \text{if } n\neq 1.\end{cases}$$

Using all this, we can estimate $|\Phi_n(1,\gamma)|$ from below:

$$\begin{aligned}
\log|\Phi_n(1,\gamma)| &= \sum_{d|n}\mu(n/d)\log\left|\gamma^d-1\right|\\
&\geq \sum_{\substack{m\not| d|n\\ \mu(n/d)=1}}\log\frac{2d}{n}-\sum_{\substack{m\not| d|n\\ \mu(n/d)=-1}}\log 2+\sum_{\substack{m|d|n\\ \mu(n/d)=1}}\log\frac{2}{\pi}+\sum_{m|d|n}\mu\left(\frac{n}{d}\right)\log(d\phi)\\
&= \sum_{\substack{d|n\\ \mu(n/d)=1}}\log\frac{2d}{n}-\sum_{\substack{d|n\\ \mu(n/d)=-1}}\log 2+\sum_{\substack{m|d|n\\ \mu(n/d)=1}}\log\frac{n}{\pi d}+\sum_{\substack{m|d|n\\ \mu(n/d)=-1}}\log 2+\\
&\quad \sum_{m|d|n}\mu\left(\frac{n}{d}\right)\log(d\phi)\\
&= \sum_{\substack{d|n\\ \mu(n/d)=1}}\log\frac{d}{n}+(\log 2)\sum_{d|n}\mu\left(\frac{n}{d}\right)+\sum_{\substack{m|d|n\\ \mu(n/d)=1}}\log\frac{2n}{\pi d}+\sum_{m|d|n}\mu\left(\frac{n}{d}\right)\log\frac{d\phi}{2}.
\end{aligned}$$

Since $n>1$, the second sum vanishes. We continue, replacing d by md in the last two sums:

$$\begin{aligned}
&= \sum_{\substack{d|n\\ \mu(n/d)=1}}\log d-(\log n)\sum_{\substack{d|n\\ \mu(n/d)=1}}1+\left(\log\frac{2n}{\pi m}\right)\sum_{\substack{d|(n/m)\\ \mu((n/m)/d)=1}}1-\sum_{\substack{d|(n/m)\\ \mu\left(\frac{n/m}{d}\right)=1}}\log d+\\
&\quad \sum_{d|(n/m)}\mu\left(\frac{n/m}{d}\right)\log d+\left(\log\frac{m\phi}{2}\right)\sum_{d|(n/m)}\mu\left(\frac{n/m}{d}\right)\\
&= 2^{\omega(n)-2}\log n+\Lambda(n)/2-2^{\omega(n)-1}\log n+\left(2^{\omega(n/m)-1}+\delta\left(n/m\right)/2\right)\log\left(2n/\pi m\right)\\
&\quad -2^{\omega(n/m)-2}\log(n/m)-\Lambda\left(n/m\right)/2+\Lambda\left(n/m\right)+\delta\left(n/m\right)\log(\phi m/2)\\
&= -2^{\omega(n)-2}\log n+\Lambda(n)/2+\left(2^{\omega(n/m)-1}+\delta\left(n/m\right)/2\right)\log(2/\pi)\\
&\quad +2^{\omega(n/m)-2}\log(n/m)+\Lambda\left(n/m\right)/2+\delta\left(n/m\right)\log(\phi m/2)\\
&\geq -2^{\omega(n)-2}\log n+\Lambda(n)/2+\log(2/\pi)+\delta\left(n/m\right)\log(\phi m/2).
\end{aligned}$$

Thus, we have proved that

$$\log|\Phi_n(1,\gamma)|\geq\begin{cases}-2^{\omega(n)-2}\log n-\log(\pi/2), & \text{if } m<n,\\ -\log\pi+\log n+\log\phi, & \text{if } m=n \text{ and } n \text{ is a prime number},\\ -\left(2^{\omega(n)-2}-1\right)\log n-\log\pi+\log\phi, & \text{if } m=n \text{ and } n \text{ is a composite number.}\end{cases}\tag{50}$$

On the other hand, one can easily estimate $|\Phi_n(1,\gamma)|$ from above using that $|\Phi_n(\alpha,\beta)| \le P(n)$ by the cyclotomic criterion:

$$\begin{aligned} \log|\Phi_n(1,\gamma)| &= \log|\Phi_n(\alpha,\beta)| - \varphi(n)\log|\alpha| \\ &\le \log P(n) - \varphi(n)h(\gamma) \qquad (51) \\ &\le \log n - 4\varphi(n), \qquad (52) \end{aligned}$$

because $h(\gamma) > 4$ by Theorem D.

If $m < n$ then $4\varphi(n) - \log n \le 2^{\omega(n)-2}\log n + \log(\pi/2)$. A simple computation shows that this inequality is contradictory when $31 \le n \le 30030$ (recall that $n \le 30030$ by Theorem B). Hence $m = n$, and the result is a direct consequence of (50) and (51). ■

Corollary 5.2 *In the set-up of Theorem 5.1 put $x = \alpha\beta$ and $y = \alpha^2 + \beta^2$. Then there exists an integer k satisfying*

$$0 < k < n/2, \quad \gcd(k,n) = 1, \qquad (53)$$

and

$$|x \cdot 2\cos(2\pi k/n) - y| \le 2c_1(n)|\alpha|^{2-\varphi(n)}. \qquad (54)$$

(Since (α,β) is a Lehmer pair, x and y are rational integers.)

Proof Using (46) and (48), we see that $|\gamma - \xi| = |\gamma\xi^{-1} - 1| \le c_1(n)|\alpha|^{-\varphi(n)}$, and similarly $|\gamma^{-1} - \xi^{-1}| \le c_1(n)|\alpha|^{-\varphi(n)}$. Since ξ is a primitive n-th root of unity, there exists an integer k satisfying (53) and such that $\xi + \xi^{-1} = 2\cos(2\pi k/n)$. Then

$$\begin{aligned} |x \cdot 2\cos(2\pi k/n) - y| &= |x|\left|\gamma + \gamma^{-1} - (\xi + \xi^{-1})\right| \\ &\le |\alpha|^2\left(|\gamma - \xi| + |\gamma^{-1} - \xi^{-1}|\right) \\ &\le 2c_1(n)|\alpha|^{2-\varphi(n)}, \end{aligned}$$

as wanted. ■

6 A lower estimate for $\arg(\beta/\alpha)^n$ and its consequences

In this section we show that an integer n is totally non-defective if it satisfies certain complicated, but rather mild inequality. In particular, every odd square-free $n > 2145$ and every prime $n > 787$ is totally non-defective.

If (α,β) is an n-defective Lehmer pair then Theorem 5.1 implies a sharp *upper* estimate for $|\arg\gamma^n|$, where $\gamma = \beta/\alpha$. On the other hand, Gelfond's theory of linear forms in two logarithms implies a sharp *lower* estimate for the same quantity. The main reference here is the recent paper [21] of Laurent, Mignotte and Nesterenko. In particular, Théorème 3 from [21] gives a lower estimate for the linear form $b_1\pi i - b_2\log\gamma$, where $|\gamma| = 1$, which is exactly what we need.

Unfortunately, the estimate from [21] becomes non-trivial only when b_1 and b_2 are sufficiently large, which makes it unsuitable for our purposes. At our request, Maurice Mignotte elaborated a more flexible version of this estimate, see Theorem A.1.3 from the appendix. As a consequence of it, we obtain the following.

Proposition 6.1 *Let γ be a complex algebraic number of degree 2, with $|\gamma| = 1$, but not a root of unity. Let $n \ge 234$ be an integer, λ a real number satisfying $1.8 \le \lambda \le 3$, and H a positive real number. Assume that $h(\gamma) > H$ and $|\arg\gamma| > \pi/n$. Then*

$$|\arg\gamma^n| \ge e^{-\left(c_3(\log n + c_2)^2 + 0.37\right)(c_4 + h(\gamma)) - 2\log(\log n + c_2) - c_5} \qquad (55)$$

where

$$\begin{aligned} c_2(\lambda,H) &= \log\left(\tfrac{1}{\pi\rho} + \tfrac{1}{\pi\rho+2H}\right) - \log\sqrt{k} + 0.856 + \tfrac{3\lambda}{2} + \tfrac{1}{k}\left(\tfrac{1}{6\rho\pi} + \tfrac{1}{1.5\pi\rho+3H}\right), \\ c_3(\lambda) &= 8\pi k\rho\lambda^{-1}, \quad c_4(\lambda) = 0.5\pi\rho, \quad c_5(\lambda) = 3\log 2 - 0.5\lambda - 2\log\lambda, \end{aligned}$$

and where the quantities ρ and k are defined as in Theorem A.1.3.

Proof We may assume that $\pi/n < \arg\gamma < \pi$, replacing γ by its complex conjugate if necessary. Put $b_2 = n$ and let b_1 be the nearest even integer to $n(\arg\gamma)/\pi$. Then $0 < b_1 \le n$, and $|\arg\gamma^n| = |b_1\pi i - b_2\log\gamma|$.

In the notation of Theorem A.1.3 we have $D = 1$ and $B = n \ge 234$. Hence

$$|\arg(\gamma^n)| = |b_1\pi i - b_2\log\gamma| \ge e^{-\left(c_3\mathcal{H}^2+0.37\right)(c_4+h(\lambda))-2\log\mathcal{H}-c_5},$$

where $\mathcal{H}$ is defined from (A.12). It remains to notice that $\mathcal{H} \le \log n + c_2$. ■

Corollary 6.2 *Let (α,β) be an n-defective Lehmer pair, where $n \ge 234$, and $\gamma = \beta/\alpha$. Let λ satisfy $1.8 \le \lambda \le 3$. Then:*

i. For any $H > 0$ either $h(\gamma) \le H$ or

$$\varphi(n) \le g(\lambda, n, H), \tag{56}$$

where

$$\begin{aligned} g(\lambda,n,H) \ = \ & \left(c_3\left(\log n + c_2(\lambda,H)\right)^2 + 0.37\right)\left(H^{-1}c_4(\lambda)+1\right) + \\ & H^{-1}\left(\log(nc_1(n)) + 2\log\left(\log n + c_2\right) + c_5\right) \end{aligned} \tag{57}$$

In particular, one always has

$$\varphi(n) \le g(\lambda, n, 4) \tag{58}$$

ii. For any $H > 0$ such that

$$\varphi(n) > c_3(\lambda)\left(\log n + c_2(\lambda,H)\right)^2 + 0.37 \tag{59}$$

we have

$$h(\gamma) \le \max\{H, f(\lambda,n,H)\} \tag{60}$$

where

$$f(\lambda,n,H) = \frac{c_4(\lambda)\left(c_3(\lambda)(\log n + c_2(\lambda,H))^2+0.37\right)+\log(nc_1(n))+2\log(\log n+c_2(\lambda,H))+c_5(\lambda)}{\varphi(n)-c_3(\lambda)(\log n+c_2(\lambda,H))^2-0.37} \tag{61}$$

In particular, if $\varphi(n) > c_3(\lambda)\left(\log n + c_2(\lambda,4)\right)^2 + 0.37$ then

$$h(\gamma) \le f(\lambda, n, 4). \tag{62}$$

Proof By Using (46) and (47) with $d = n$ we obtain

$$|\arg\gamma^n| \le nc_1(n)|\alpha|^{-\varphi(n)} = nc_1(n)e^{-\varphi(n)h(\gamma)} \tag{63}$$

(recall that $h(\gamma) = \log|\alpha|$). Also, using (47) with $d = 1$, we obtain $|\arg\gamma| > \pi/n$. Since $n \ge 234$, we may apply Proposition 6.1. Combining (55) and (63), we obtain

$$\begin{aligned} \varphi(n)h(\gamma) \ \le \ & \left(c_3\left(\log n + c_2\right)^2 + 0.37\right)\left(c_4 + h(\gamma)\right) + \\ & 2\log\left(\log n + c_2\right) + c_5 + \log(nc_1(n)) \end{aligned} \tag{64}$$

whenever $h(\gamma) \ge H$. This proves (56) and (60). Inequalities (58) and (62) follow if we recall that $h(\gamma) > 4$ by Theorem D. The proof is complete. ■

Corollary 6.2 suggests to consider the quantities

$$\min_{\substack{\lambda\in[1.8,3] \text{ sa-} \\ \text{tisfies (59)}}} f(\lambda,n,H) \quad \text{and} \quad \min_{\lambda\in[1.8,3]} g(\lambda,n,H).$$

To speed up the computations, we minimize f and g over a finite set of values of λ. For $i = 0, \dots, 12$ we put $\lambda_i = 1.8 + 0.1i$ and define

$$\tilde{f}(n,H) = \min_{\substack{i=0,\dots,12 \\ \lambda_i \text{ satisfies (59)}}} f(n,\lambda_i,H), \quad \widetilde{g}(n,H) = \min_{i=0,\dots,12} g(n,\lambda_i,H), \tag{65}$$

with the convention $\tilde{f}(n,H) = \infty$ if the minimum is taken over the empty set.

Further, for a given n we would like to find the optimal value of H. If $\varphi(n) > \widetilde{g}(n,\infty)$ then there exists a (unique) solution $H_0 = H_0(n)$ to the equation

$$\varphi(n) = \min_{i=0,\dots,12} c_3(\lambda)\left(\log n + c_2(\lambda,H)\right)^2 + 0.37. \tag{66}$$

Indeed, the right-hand side of (66) is a continuous decreasing function of H, which tends to ∞ when $H \to 0$ and tends to $\widetilde{g}(n,\infty)$ when $H \to \infty$.

Further, $\tilde{f}(n,H)$, considered as a function in H, is continuous and decreasing on the interval (H_0,∞), with $\tilde{f}(n,H_0) = \infty$ and $\tilde{f}(n,\infty) < \infty$. Therefore there exists a unique solution $H(n)$ to the equation $H = \tilde{f}(n,H)$.

A direct consequence of Corollary 6.2 is

Corollary 6.3 *Let (α,β) be an n-defective Lehmer pair, where $n \ge 234$, and $\gamma = \beta/\alpha$. Then:*

i. For any $H > 0$ either $h(\gamma) \le H$ or $\varphi(n) \le \widetilde{g}(n,H)$. In particular

$$\varphi(n) \le \widetilde{g}(n,4) \tag{67}$$

ii. If $\varphi(n) > \widetilde{g}(n,\infty)$ then $h(\gamma) \le H(n)$. ∎

Given a real number κ, we denote by $p_\nu(\kappa)/q_\nu(\kappa)$ the ν-th convergent of the continuous fraction expansion of κ. Given $\Omega > 0$, we put $q(\kappa,\Omega) = q_\nu(\kappa)$ and $q'(\kappa,\Omega) = q_{\nu+1}(\kappa)$, where ν is defined from $q_\nu(\kappa) \le \Omega < q_{\nu+1}(\kappa)$.

Lemma 6.4 *For any odd square-free n satisfying*

$$\widetilde{g}(n,1000) \le \varphi(n) \le \widetilde{g}(n,4) \tag{68}$$

and for any k satisfying (53) one has

$$q'\left(2\cos(2\pi k/n), \Omega(n)\right) \le (4c_1(n))^{-1} e^{4(\varphi(n)-2)}, \tag{69}$$

where $\Omega(n) = e^{2\min(1000,H(n))}$.

Proof This was done by a direct computation, using a C program and the PARI 2.0 programming library. There are 2563 numbers n satisfying (68), and in totality 5766966 numbers of the form $2\cos(2k\pi/n)$ were considered. Instead of $H(n)$ we used a good upper approximation $\widetilde{H}(n)$, satisfying $\widetilde{H}(n) - 0.1 \le f(n,\widetilde{H}(n)) \le \widetilde{H}(n)$. This approximation was computed by the following procedure: take a very large H_1, and compute the sequence $H_k = f(f(n,H_{k-1}))$, until the desired inequality holds. This always occurred after at most 6 steps (2 on average). The total computational time on a PC Pentium Pro 200 was roughly 15 days. ∎

Now we are ready to establish the main result of this section.

Theorem 6.5 *An odd square-free integer $n \ge 234$ satisfying*

$$\varphi(n) > \widetilde{g}(n,1000). \tag{70}$$

is totally non-defective.

Proof Assuming the contrary, fix an n-defective Lehmer pair (α, β). Then $h(\beta/\alpha) \le 1000$ (otherwise (70) contradicts Corollary 6.3 (i)).

On the other hand, since $\varphi(n) > \widetilde{g}(n, 1000) > \widetilde{g}(n, \infty)$, Corollary 6.3 (i) implies that $h(\beta/\alpha) \le H(n)$. This yields $\log|\alpha| = h(\beta/\alpha) \le (1/2)\log\Omega(n)$. In particular, for $x = \alpha\beta$ we have $|x| \le \Omega(n)$. Therefore, by the theory of continued fractions (see, for instance, [19], Theorems 13 and 16) one has

$$1/2q' \le \|q \cdot 2\cos(2\pi k/n)\| \le \|x \cdot 2\cos(2\pi k/n)\| \le 2c_1(n)|\alpha|^{2-\varphi(n)}, \tag{71}$$

where k is defined from Corollary 5.2 and

$$q = q\left(2\cos(2\pi k/n), \Omega(n)\right), \quad q' = q'\left(2\cos(2\pi k/n), \Omega(n)\right).$$

(Recall that $\|\cdot\|$ stands for the distance from the nearest integer.)

By Corollary 6.3 (i) the number n satisfies (67), which together with (70) gives (68). Since n is odd and square-free, we have (69). Combined with (71), this yields $\log|\alpha| = h(\beta/\alpha) \le 4$, which contradicts Theorem D. The proof is complete. ■

The following corollary (proved by the direct computation of the right-hand side of (70) for odd square-free $n \le 30030$) will not be used in the sequel, but it gives some idea about the strength of Theorem 6.5.

Corollary 6.6 *Every odd square-free integer n with $\omega(n) > 4$ is totally non-defective. If n is an odd square-free integer with $\omega(n) = k \le 4$ then n is totally non-defective whenever $n > N_1(k)$, where $N_1(k)$ is defined in the following table:*

k	1	2	3	4
$N_1(k)$	787	$1315 = 5 \cdot 263$	$1695 = 3 \cdot 5 \cdot 113$	$2145 = 3 \cdot 5 \cdot 11 \cdot 13$

In particular, every odd square-free $n > 2145$ is totally non-defective. ■

7 Numerical solution of Thue equations: an overview

In this section we deviate from our main subject to review methods for numerical solution of Thue equations of high degree. We shall use these methods for solving the equation (25).

The history of numerical solution of Diophantine equations dates back to 1969, when Baker and Davenport [2] completely solved a system of two Pell equations. They used the well-known fact that every "large" solution gives rise to a "very small" value of the linear form $\Lambda(b_1, b_2) = \log\alpha_0 + b_1\log\alpha_1 + b_2\log\alpha_2$ (where α_0, α_1 and α_2 are explicitly given algebraic numbers) at an integral point (b_1, b_2). Using Baker's theory of logarithmic forms, they obtained a huge (around 10^{400}) upper bound for $\max(|b_1|, |b_2|)$. After this, expanding $\log\alpha_2/\log\alpha_1$ into a continued fraction, they showed that $|\Lambda(b_1, b_2)|$ cannot be too small when b_1 and b_2 run the integers below the huge bound. Therefore the system cannot have "large" solutions, while "small" solutions can be easily enumerated.

This idea was developed in various directions by Pethő, Steiner, Tzanakis, de Weger, and many other authors. The subject became especially popular when Lenstra, Lenstra and Lovasz [23] suggested a polynomially quick algorithm for finding an "almost shortest" vector in a lattice (referred to as LLL-algorithm in the sequel). The LLL-algorithm made it possible to extend the idea of Baker and Davenport to logarithmic forms in three or more variables, when continued fractions can not be used anymore. See [39, 27] for a detailed description of the methods, history of the subject and extensive bibliography up to 1989.

As it was already indicated in the introduction, the method of Tzanakis and de Weger [39] for explicit resolution of Thue equations was used (with some modications) for the proof of Theorem C. However, the complete solution of the main problem requires solving Thue equations of very high degree, where the methods of [39, 27] are not efficient anymore. The main difficulties are:

First difficulty These methods require algebraic computations (in particular, computing fundamental units) in the associated number field. For fields of high degree, this problem is still far beyond the capacities of the modern computational number theory (see [12, 29, 30]).

Second difficulty The LLL-algorithm has to be applied to (many) lattices of very high dimention, which is very slow.

In [6] (see also [5]) it was shown that the second difficulty can be overcome: one can solve Thue equations using only continued fractions (as Baker and Davenport did), and without involving the LLL-algorithm.

Concerning the first difficulty, two different ideas were suggested in [8] and [18]:

First idea [8] If the associated number field contains a subfield of degree at least 3 (over $\mathbb{Q}$) then it is sufficient to compute units (and other algebraic stuff) in the subfield.

Second idea [18] To resolve a Thue equation, it suffices to compute a full rank system of undependent units (rather than a system of fundamental units) in the associated number field.

Of course, these ideas are not sufficient to eliminate the first difficulty in general, and finding a reasonable agorithm for numerical solution of arbitrary Thue equations of high degree remains a challenging problem.

Fortunately, the number field, associated to the equation (25), is the *real cyclotomic field* $\mathbb{L}^{(n)} = \mathbb{Q}(2\cos(2\pi/n))$, which has the following two properties:

First property The field $\mathbb{L}^{(n)}$ is abelian; in particular, if d divides its degree $\varphi(n)/2$ then $\mathbb{L}^{(n)}$ has a subfield of degree d over $\mathbb{Q}$.

Second property When n is a prime power, the field $\mathbb{L}^{(n)}$ has an *explicitly given* full rank system of independent units

$$\sin(k\pi/n)/\sin(\pi/n) \qquad (1 \le k < n/2, \quad (k,n)=1), \tag{72}$$

called *circular units.*

Therefore both the ideas can be sucessfully employed for the resolution of the equations (25). Moreover, the two ideas are quite independent and can be combined very efficiently.

Remark 7.1 Masley [26] (see also [24]) proved that circular units form a fundamental system for $n \le 67$. This allows one to solve the main problem for $n \le 67$ (see [17]), using only the method of [6].

What follows is a concise description of the method we used for the explicit resolution of Thue equations. It is combined from the methods of [8] and [18], with some additional ideas. In this section we found it methodologically better to consider a general Thue equation. In the next section we show how the method was adapted to the particular equations (25).

We give only a general overview of the method, omitting technical details like routine proofs, explicit expressions for constants, etc. All missing details can be found in [8], Section 3; in particular, the constants X_1 and c_8–c_{11}, as well as all the constants implied by $O(\cdot)$, $\ll$ or $\gg$ are explicitly displayed therein.

7.1 Preliminaries

We consider the Thue equation

$$F(x,y) = \mathcal{N}_{\mathbb{L}/\mathbb{Q}}(y - \alpha x) = a, \tag{73}$$

where $a = a_1/a_2$ is a rational number, α an algebraic number of degree $N \geq 3$, and $\mathbb{L} = \mathbb{Q}(\alpha)$.

Let $\mathbb{K}$ be a subfield of $\mathbb{L}$ of degree $[\mathbb{K}:\mathbb{Q}] = m \geq 3$. Denote by $\sigma_1, \ldots, \sigma_s$ the real embeddings of $\mathbb{K}$, and by $(\sigma_{s+1}, \sigma_{s+1+t}), \ldots, (\sigma_{s+t}, \sigma_{s+2t})$ the pairs of complex conjugate embeddings (so that $m = s + 2t$).

Put

$$\Phi(X,Y) = \mathcal{N}_{\mathbb{L}/\mathbb{K}}(Y - \alpha X) \in \mathbb{K}[X,Y], \quad \Phi_i(X,Y) = \sigma_i(\Phi(X,Y)), \tag{74}$$

so that $F(X,Y) = \Phi_1(X,Y) \cdots \Phi_m(X,Y)$.

7.1.1 Asymptotics for $\Phi_i(x,y)$

To begin with, recall that for any solution $(x,y) \in \mathbb{Z}^2$ of (73) with $|x| \geq X_1$ there exists a real conjugate α' to α (over $\mathbb{Q}$) such that

$$|y/x - \alpha'| \leq c_8 |x|^{-n}. \tag{75}$$

Here X_1 and c_8 are effectively computable positive constants. See, for instance, [39], Lemma 1.1, (reproduced in [8] as Lemma 3.1.1 (i)) where explicit expresssions for the constants X_1 and c_8 are given.

In concrete examples the constant X_1 is very small, and solutions satisfying $|x| \leq X_1$ can be easily enumerated.

From now on, we fix a real conjugate α' and consider solutions satisfying (75). (To find all the solutions with $|x| \geq X_1$, the procedure described below has to be repeated for all α'.)

The fixed conjugate α' satisfies $\Phi_{i_0}(1,\alpha') = 0$ for some $i_0 \in \{1, \ldots, s\}$. Arguing as in [8], Subsection 3.1, we obtain for every solution (x,y), satisfying (75), the relations

$$\Phi_i(x,y) = \Phi_i(1,\alpha') x^{N/m} e^{O\left(|x|^{-N}\right)} \quad (i \neq i_0). \tag{76}$$

7.1.2 The quantity Υ

Fix $i_1, i_2 \neq i_0$ and put

$$\Upsilon = \Upsilon(i_1, i_2) = \frac{\Phi_{i_1}(x,y)\Phi_{i_2}(1,\alpha')}{\Phi_{i_2}(x,y)\Phi_{i_1}(1,\alpha')}.$$

It follows from (76) that

$$\Upsilon = e^{O\left(|x|^{-N}\right)}. \tag{77}$$

Also, Proposition 3.1.2 of [8] guarantees that

$$\Upsilon \neq 1 \tag{78}$$

for a suitable choice of i_1 and i_2. We shall assume (78) in the sequel.

7.2 Fundamental units

Now let $\eta_1, \ldots, \eta_r$ be a system of fundamental units of the field $\mathbb{K}$. Since $\mathcal{N}_{\mathbb{K}/\mathbb{Q}} \Phi(x,y) = a$, there are only finitely many possibilities for the fractional ideal $(\Phi(x,y))$. It follows that

$$\Phi(x,y) = \mu \eta_1^{b_1} \cdots \eta_r^{b_r} \tag{79}$$

where $b_1, \ldots, b_r \in \mathbb{Z}$, and μ belongs to a finite effectively computable set $\mathrm{M} \subset \mathbb{K}$. Acting on (79) by the maps σ_i, and taking (the real part of) the logarithm, we obtain the following system of r linear equations in $r = s + t - 1$ variables $b_1, \ldots, b_r$:

$$\sum_{j=1}^{r} b_j \log |\sigma_i(\eta_j)| = -\log |\sigma_i(\mu)| + \log |\Phi_i(x,y)| \qquad (1 \leq i \leq s+t, \quad i \neq i_0). \tag{80}$$

Due to our numbering of σ_i, the linear system has a non-zero determinant. Resolving the system, and using (76), we obtain (see [8], equations (15) and (16))

$$b_i = \delta_i \log |x| + \lambda_i + O\left(|x|^{-N}\right) \quad (1 \le i \le r), \tag{81}$$

where δ_i and λ_i are real numbers, and $\delta_i \neq 0$ for some i.

Equalities (80) and (81) imply that

$$B \ll \log x \ll B, \tag{82}$$

where $B = \max(1, |b_1|, \ldots, |b_r|)$.

7.2.1 Baker's bound

It follows from (79) that

$$\Upsilon = \beta_0 \beta_1^{b_1} \cdots \beta_r^{b_r}, \tag{83}$$

where the algebraic numbers $\beta_0 = \beta_0^{(i_1,i_2)}, \beta_1 = \beta_1^{(i_1,i_2)}, \ldots, \beta_r = \beta_r^{(i_1,i_2)}$ can be easily expressed explicitly. Taking (the principal value of) the logarithm and using (77) and (78), we obtain $0 < |\Lambda| \ll |x|^{-N}$, where

$$\Lambda = \log \beta_0 + b_1 \log \beta_1 + \cdots b_r \log \beta_r + b_{r+1} \cdot \pi i,$$

and $b_{r+1} \in \mathbb{Z}$ satisfies $|b_{r+1}| \ll B$. Using (82), we derive from this that

$$0 < |\Lambda| \ll e^{-c_9 B}, \tag{84}$$

where c_9 is an effectively computable positive constant.

On the other hand, Baker's theory implies a lower estimate for $|\Lambda|$ of the form $|\Lambda| \geq B^{c_{10}}$, (provided $\Lambda \neq 0$) where $c_{10} > 0$ is effectively computable. The best known to us quantitative value for c_{10} is given by Baker and Wüstholz [3], page 20. For large B, the lower estimate for $|\Lambda|$ contradicts (84). This implies that $B \le B_0$, where B_0 is an effectively computable (large) positive number. In practical computations B_0 can range from 10^{30} to 10^{3000} (or even bigger).

See [8], Subsection 3.3, for further details.

7.2.2 Reduction

Our next task is to reduce the upper bound for B, using the continued fractions techniques. The idea goes back to Baker and Davenport [2]. We follow [8], Subsections 3.2 and 4.1. We shall assume that $r \geq 2$. If $r = 1$ then the reduction is similar and much simpler, see [8], Subsection 4.2.

Since $r \geq 2$, we may eliminate $\log |x|$ from (81). For instance, if $\delta_1 \neq 0$ then

$$|b_2 - \delta b_1 - \lambda| \le c_{11} |x|^{-N} \ll e^{-c_9 B}, \tag{85}$$

where δ and λ are real numbers and c_{11} a moderate constant. It follows that

$$B \ll c_{12} - c_9{}^{-1} \log \min_{|b_1|,|b_2| \le B_0} |b_2 - \delta b_1 - \lambda|, \tag{86}$$

where c_{12} is a small constant. The minimum in the right-hand side can be quickly computed (or, at least, minorated) by expanding δ into a continued fraction (see [7], Subsection 4.1, for the details). Heuristically, this minimum is of magnitude B_0^{-2}, which yields a new upper bound for B of magnitude $\log B_0$.

Iterating this procedure, one obtains, after two-three steps, a reasonable upper bound for B, of magnitude 10 or so. The possible $b_1, \ldots, b_r$ below this bound can be easily enumerated (see [8], Section 5).

Moreover, it turns out that, in most of the cases, enumerating small b_i is superfluous. Indeed, (85) yields an upper bound not only for B, but for $|x|$ as well, of the form

$$|x| \le {c_{11}}^{1/N} \left(\min_{|b_1|,|b_2| \le B_0'} |b_2 - \delta b_1 - \lambda| \right)^{-1/N}, \tag{87}$$

where B_0' is the reduced upper bound for B. Enumerating the solutions satisfying (87) can be done very quickly, because the right-hand side of (87) is quite moderate (typically, 10 or so), especially when N is large.

This method of reduction fails if, for instance, δ is a rational number. See [7], Subsection 4.6, and [17] for a detailed analysis of this and other "pathologies".

7.3 Non-fundamental units

It turns out that the method described above works, with slight modifications, if we have only a full rank system of independent units instead of fundamental units.

Thus, let now $\eta_1, \ldots, \eta_r$ be a full rank system of independent units of the field $\mathbb{K}$. Denote by b_0 the index of the subgroup generated by $\eta_1, \ldots, \eta_r$ in the full unit group. It is not easy to compute b_0, but one can majorate it using the estimate $R_{\mathbb{K}} \ge 0.2$ for the regulator of the field $\mathbb{K}$ (see [16, 28, 45]). Let $\tilde{b}_0$ be an upper bound for b_0 obtained this way. Typically, $\tilde{b}_0$ is of magnitude 10^{10}–10^{40}.

Instead of (79) we have now

$$\Phi(x,y)^{b_0} = \mu^{b_0} \eta_1^{b_1} \cdots \eta_r^{b_r}, \tag{79$'$}$$

which implies, instead of (81), the relation

$$b_i/b_0 = \delta_i \log|x| + \lambda_i + O\left(|x|^{-N}\right) \quad (1 \le i \le r). \tag{81$'$}$$

(Here and below the implicit constants in the equation (n') are equal to the correspondent constants in (n).) Further,

$$B/b_0 \ll \log x \ll B/b_0, \tag{82$'$}$$

where now $B = \max(b_0, |b_1|, \ldots, |b_r|)$.

7.3.1 Baker's bound

Instead of (83) we have

$$\Upsilon^{b_0} = \beta_0^{b_0} \beta_1^{b_1} \cdots \beta_r^{b_r}. \tag{83$'$}$$

If $\Upsilon^{b_0} \ne 1$, then, as before, we obtain the inequality

$$0 < |\Lambda| \ll \tilde{b}_0 e^{-c_9 B/\tilde{b}_0}, \tag{84$'$}$$

where now

$$\Lambda = b_0 \log \beta_0 + b_1 \log \beta_1 + \cdots b_r \log \beta_r + b_{r+1} \cdot \pi i.$$

Again, Baker's inequality implies that $B \le B_0'$, where B_0' is a large positive number.

If $\Upsilon^{b_0} = 1$, then $\arg \Upsilon \ge 2\pi/b_0$, which, together with (77) and (82), implies that $B \le B_0''$.

Thus, in any case $B \le B_0 := \max(B_0', B_0'')$. (In real computations, B_0'' is usually much smaller than B_0'.)

7.3.2 Reduction

Using

$$|b_2 - \delta b_1 - \lambda b_0| \le c_{11} \tilde{b}_0 |x|^{-N} \ll \tilde{b}_0 e^{-c_9 B/\tilde{b}_0}, \tag{85$'$}$$

we obtain

$$B/\tilde{b}_0 \ll c_{12} + c_9{}^{-1} \left(\log \tilde{b}_0 - \log \min_{\substack{1 \le b_0 \le \tilde{b}_0 \\ |b_1|, |b_2| \le B_0}} |b_2 - \delta b_1 - \lambda b_0| \right), \tag{86$'$}$$

The minimum can be quickly minorated using the 3-dimensional LLL, see [39] for the details. This implies a new upper bound for B, of magnitude $\tilde{b}_0 \log B_0$.

After a few (usually, two) iterations, one obtains $B \le B_0'$, where B_0' is of magnitude $\tilde{b}_0 \log \tilde{b}_0$.

Now we have the following upper bound for $|x|$:

$$|x| \le (c_{11}\tilde{b}_0)^{1/N} \left(\min_{\substack{1 \le b_0 \le \tilde{b}_0 \\ |b_1|, |b_2| \le B_0'}} |b_2 - \delta b_1 - \lambda b_0| \right)^{-1/N}. \tag{87$'$}$$

The right-hand side of (87$'$) is very moderate (typically, 30), and the solutions satisfying (87$'$) can be easily enumerated.

8 The final attack

Return to the proof of Theorem 1.4. In view of Theorem 6.5, it remains to prove the following.

Theorem 8.1 *Every odd square-free integer $n \ge 31$ satisfying*

$$n \le 233 \quad or \quad \varphi(n) \le \widetilde{g}(n, 1000) \tag{88}$$

is totally non-defective.

By Corollary 2.5, it is sufficient to show that, for every such n, the equation

$$\prod_{\substack{1 \le k < n/2 \\ (k,n)=1}} (Y - 2\cos(2\pi k/n) \cdot X) = \mathcal{N}_{\mathbb{L}/\mathbb{Q}}(Y - 2\cos(2\pi/n) \cdot X) \in \{\pm 1, \pm P(n)\} \tag{89}$$

(where $\mathbb{L} = \mathbb{L}^{(n)} = \mathbb{Q}(2\cos(2\pi/n)))$ has no solutions (x, y) with $|x| > e^8$. This was done using the method described in Section 7. Below we show how this method was adapted to the particular equation (89).

Thus, fix an odd square-free integer $n \ge 31$ satisfying (88). As indicated in Section 7, we have to choose a "small" subfield $\mathbb{K}$ of $\mathbb{L}$ of degree $m = [\mathbb{K}\colon \mathbb{Q}] \ge 3$. Note that, since $\mathbb{L}/\mathbb{Q}$ is abelian, for each m dividing $[\mathbb{L}\colon \mathbb{Q}] = \varphi(n)/2$, the field $\mathbb{L}$ has a subfield of degree m.

We defined the field $\mathbb{K}$ and organized the computations in two different ways depending on whether or not $\varphi(n)/2$ has a "small" divisor bigger than 2. To be precise, let

$$\{3, 5, 4, 7, 11, 13, 17, \ldots\} \tag{90}$$

be the set of all odd primes together with 4, ordered as indicated; that is, 5 precedes 4 and otherwise the ordering is natural. Let d be the smallest (with respect to this ordering) divisor of $\varphi(n)/2$ from the set (90). Then we have two cases: **Case A:** $d \le 11$ and **Case B:** $d \ge 13$.

8.1 Case A

We have three possibilities:

$$d \text{ divides } (p-1)/2 \text{ for some prime } p|n; \tag{91}$$

$$d = 4, \text{ no prime divisor of } n \text{ is } 1 \bmod 8, \text{ and at least one is } 5 \bmod 8; \tag{92}$$

$$d = 4 \text{ and all prime divisors of } n \text{ are } 3 \bmod 4. \tag{93}$$

In case (91) choose the smallest p with this property, and put $m = d$. The Galois group of the field $\mathbb{L}^{(p)} = \mathbb{Q}(\cos(2\pi/p))$ over $\mathbb{Q}$ is cyclic, which means that it has a single subfield of degree m over $\mathbb{Q}$. We let $\mathbb{K}$ be this subfield.

In case (92) let p be the smallest prime divisor of n satisfying $p \equiv 5 \bmod 8$, and q the smallest of the other prime divisors of n. If $q \equiv 5 \bmod 8$ then we put $\mathbb{K} = \mathbb{Q}\left(\sqrt{p}, \sqrt{q}\right)$. If $q \equiv 3 \bmod 4$ then 4 exactly divides $\varphi(pq)/2$, and we define $\mathbb{K}$ as the single subfield of $\mathbb{L}^{(pq)}$ of degree 4 over $\mathbb{Q}$.

Option (93) never occured in our computations.

Since the field $\mathbb{K}$ is reasonably "small", we computed the unit group, together with the class group, using the algorithm of Buchmann-Cohen-Diaz y Diaz-Olivier [10, 13], implemented as the PARI function `bnfinit`. It produces, in particular, a full rank system of independent units of the field $\mathbb{K}$, which are guaranteed to be fundamental if one assumes the generalized Riemann hypothesis. We then used the PARI function `bnfcertify` to verify the results of `bnfinit` unconditionnally (this was the hardest part of the computations). Note that though we are able to deal with non-fundamental systems of units, as described in Subsection 7.3, (we shall do this in Case B), it remains highly preferable to use fundamental systems, since the corresponding algorithm runs significantly faster.

The set M (see the beginning of Subsection 7.2) is computed by decomposing the ideal (P) into a product of prime ideals (PARI function `primedec`), enumerating all the ideals of norm $\pm P$, and testing them for being principal (function `bnfisprincipal`). Alternatively, all these steps are grouped in the function `bnfisintnorm`.

The use of PARI requires a defining equation of the field $\mathbb{K}$. It can be easily computed if one knows, with a sufficient precision, all conjugates of a generator of $\mathbb{K}$ over $\mathbb{Q}$. These can be found from Lemma 8.2 below. In the sequel, we put $\zeta := e^{2\pi/n}$ and identify the Galois group $\mathrm{Gal}\,\mathbb{Q}(\zeta)/\mathbb{Q}$ with $(\mathbb{Z}/n\mathbb{Z})^*$ by $a \mapsto (\zeta \mapsto \zeta^a)$.

Lemma 8.2 *Let G be a subgroup of $(\mathbb{Z}/n\mathbb{Z})^*$ of index m and $G_0, \ldots, G_{m-1}$ the cosets of G. Assume that $-1 \in G$. Then*

$$\xi_i = \sum_{a \in G_i} \cos(2a\pi/n) \quad (0 \le i \le m-1)$$

are algebraic integers, conjugate to ξ_0 over $\mathbb{Q}$. Put $\mathbb{K} = \mathbb{Q}(\zeta)^G$. Then $\mathbb{K} = \mathbb{Q}(\xi_i)$ for $0 \le i \le m-1$.

Proof Since $-1 \in G$, one has $\xi_i = \sum_{a \in G_i} \zeta^a$, In particular, ξ_i are algebraic integers.

Further, $\mathrm{Gal}\,\mathbb{Q}(\zeta)/\mathbb{Q} = (\mathbb{Z}/n\mathbb{Z})^*$ acts transitively on $\xi_0, \ldots, \xi_{m-1}$, and G stabilizes each of them. Hence $\xi_0, \ldots, \xi_{m-1}$ are conjugate over $\mathbb{Q}$ and belong to $\mathbb{K}$. Since $\mathrm{Gal}\,\mathbb{Q}(\zeta)/\mathbb{Q}$ is abelian, each of $\xi_0, \ldots, \xi_{m-1}$ generates the same subfield $\mathbb{K}_0$ of $\mathbb{K}$. In particular, $\mathbb{K}_0 = \mathbb{Q}(\xi_0, \ldots, \xi_{m-1})$.

To show that $\mathbb{K}_0 = \mathbb{K}$ we use the following general observation.

Let $k \subseteq K \subseteq L$ be a tower of fields of characteristic zero, and assume that $\alpha \in L$ generates L over k. Then the numbers

$$\mathrm{Tr}_{L/K}(\alpha^j) \quad (1 \le j \le [L\colon K]) \tag{94}$$

generate K over k.

(To prove this, notice that K is generated over k by the coefficients of the minimal polynomial of α over K, and these coefficients can be expressed as polynomials in the numbers (94) with integral coefficients.)

Now it is easy to complete the proof of the lemma. Since $\mathrm{Tr}_{\mathbb{Q}(\zeta)/\mathbb{K}}(\zeta^a) \in \{\xi_0, \ldots, \xi_{m-1}\}$ for any $a \in \mathbb{Z}$, we have $\mathbb{K} = \mathbb{Q}(\xi_0, \ldots, \xi_{m-1}) = \mathbb{K}_0$, as wanted. ∎

Thus, in Case A, we can easily compute the algebraic data needed for the method described in Section 7, which thereby can be readily applied (in [8], it was applied to much larger examples, up to $n = 5001$).

8.2 Case B

In this case we define m as the minimal odd prime divisor of $(P-1)/2$, where $P = P(n)$ is the maximal prime divisor of n. Notice that m need not be equal to d in the Case B. Further, we define $\mathbb{K}$ as the single subfield of $\mathbb{L}^{(P)}$ of degree m over $\mathbb{Q}$. This choice of $\mathbb{K}$ will be motivated in the last paragraph of Subsection 8.2.1.

We fix the isomorphism

$$\begin{array}{rcl} \mathcal{G} := (\mathbb{Z}/P\mathbb{Z})^*/\{\pm 1\} & \rightarrow & \mathrm{Gal}\left(\mathbb{L}^{(P)}/\mathbb{Q}\right) \\ a & \mapsto & \tau_a : \cos(2\pi/P) \mapsto \cos(2a\pi/P). \end{array} \tag{95}$$

The subgroup Γ of $\mathcal{G}$ of index m corresponds in (95) to $\mathrm{Gal}\left(\mathbb{L}^{(P)}/\mathbb{K}\right)$.

8.2.1 Unit group and set M

In the Case B, the field $\mathbb{K}$ is of large degree (at least 13) and of large discriminant[5]. This makes the computation of the class group and unit group by the Buchmann-Cohen-Diaz y Diaz-Olivier method very troublesome, and the certification almost hopeless.

However, even if the computation of a system of fundamental units seems intractable, we can compute very easily a system of units of full rank, obtained by taking the norm of the circular units of $\mathbb{L}^{(P)}$.

Lemma 8.3 *For every $a \in \mathcal{G}$ the algebraic number $\xi_a = |\sin(\pi a/P)/\sin(\pi/P)|$ is well-defined, and $(\xi_a)_{a\in\mathcal{G}\setminus\{1\}}$ is a full rank system of independent units of the field $\mathbb{L}^{(P)}$. Further, let $a_0 = 1, a_1, \ldots, a_{m-1} \in \mathcal{G}$ be a system of representatives modulo Γ, and $\eta_i = N_{\mathbb{L}^{(P)}/\mathbb{K}}(\xi_{a_i})$. Then $(\eta_1, \ldots, \eta_{m-1})$ is a full rank system of independent units of the field $\mathbb{K}$.*

Proof The independence of the ξ_a is a well-known fact, see for instance [43], Theorem 8.2. It implies that for any integers $\{b_a\}_{a\in\mathcal{G}}$

$$\left(\sum_{a\in\mathcal{G}} b_a \log|\sin(\pi a/P)| = 0, \quad \sum_{a\in\mathcal{G}} b_a = 0\right) \Rightarrow \left(\, b_a = 0 \quad (a \in \mathcal{G})\right) \tag{96}$$

It is easy to see that $\tau_a(\xi_b) = \pm\,|\sin(ab\pi/P)/\sin(a\pi/P)|$, which implies that

$$\eta_i = \pm\left(\prod_{a\in\Gamma} |\sin(aa_i\pi/P)|\right) \Big/ \left(\prod_{a\in\Gamma} |\sin(a\pi/P)|\right). \tag{97}$$

Now let $\eta_1^{b_1}\cdots\eta_{m-1}^{b_{m-1}} = 1$ be a dependence relation for the units $(\eta_1, \ldots, \eta_{m-1})$. Using (97), we rewrite this as

$$\sum_{l=0}^{m-1} b_l \sum_{a\in\Gamma} \log|\sin(aa_l\pi/P)| = 0,$$

where $b_0 = -b_1 - \cdots - b_{k-1}$. Hence $b_1 = \cdots = b_{k-1} = 0$ by (96). The lemma is proved. ∎

We also need to compute the set M. With our choice of the field $\mathbb{K}$, this problem is completely eliminated. Indeed, the prime ideal (P) totally ramifies in $\mathbb{L}^{(P)}$ as $(1-2\cos(2\pi/P))^{(P-1)/2}$. Hence it totally ramifies in $\mathbb{K}$ as $\left(N_{\mathbb{L}^{(P)}/\mathbb{K}}(1-2\cos(2\pi/P))\right)^m$. Thus, we can take as M the set $\{1, N_{\mathbb{L}^{(P)}/\mathbb{K}}(1-2\cos(2\pi/P))\}$. This is why in the Case B we choose $\mathbb{K}$ as a subfield of $\mathbb{L}^{(P)}$ (note that the construction of the set of independent units from Lemma 8.3 can be applied to any subfield of $\mathbb{L}^{(p)}$ for any prime p dividing n).

[5] more precisely, of large degree and of large *root-discriminant* $D_{\mathbb{K}}^{1/[\mathbb{K}:\mathbb{Q}]}$

8.2.2 Inverting a huge matrix, computing δ_i and λ_i.

The computationally limiting step of the algorithms of [6, 8, 18], once the algebraic number theory data is known, is inverting the matrix $[\log|\sigma_i(\eta_j)|]$, which is implicitly required to derive identity (81) from (80).

Because of the very high precision needed in this step[6] (we need to know δ_i and λ_i to a very high precision for the reduction step, up to several thousand digits in the worst case), and the extremely large dimension of the matrix in certain cases (up to 359), it is unrealistic to use straightforward Gauss elimination; we have to exploit the particular form of the matrix (coming from the fact that the field is abelian) to derive a computationally sound formula for the inverse matrix.

Lemma 8.4 *Let G be a finite abelian group, $\hat{G}$ its dual group, and f a $\mathbb{C}$ valued function on G. Put $M = [f(\sigma\tau^{-1}) - f(\sigma)]_{\sigma,\tau\neq 1}$. Then $\det M = \prod_{\chi\in\hat{G}\setminus\{1\}} \sum_\sigma \chi(\sigma)f(\sigma)$ and $M^{-1} = [a_{\sigma\tau}]$ where*

$$a_{\sigma\tau} = \sum_{\chi\neq 1} \frac{\chi(\tau^{-1})(\chi(\sigma)-1)}{|G|\sum_\alpha f(\alpha)\chi(\alpha)} \tag{98}$$

Proof We follow the proof presented in [43], Lemma 5.26, for the first part of the lemma. Let H be the vector space of $\mathbb{C}$-valued functions h on G such that $\sum_\sigma h(\sigma) = 0$. The group G acts on H by translation, that is, $(\sigma h)(X) = h(\sigma X)$. Define T as the linear transformation $T = \sum_\sigma f(\sigma)\sigma$.

We define two bases of H. Put

$$\psi_\tau(x) = \begin{cases} 1 - |G|^{-1} & \sigma = \tau, \\ -|G|^{-1} & \text{otherwise} \end{cases}.$$

Then $\Psi = \{\psi_\tau, \tau \neq 1\}$ form a basis for H. Note that $\psi_1 = -\sum_{\tau\neq 1}\psi_\tau$.

The non-trivial characters on G form a basis of eigenvectors of T:

$$T\chi(X) = \sum_\sigma f(\sigma)\chi(\sigma X) = (\sum_\sigma f(\sigma)\chi(\sigma))\chi(X),$$

so that the determinant and inverse of T are easy to compute in the basis $\hat{G}\setminus\{1\}$.

Now let us show that the matrix M is just the matrix of the endomorphism T in the basis Ψ. We compute $T\psi_\tau$:

$$\begin{aligned} T\psi_\tau(X) &= \sum_\sigma f(\sigma)\psi_\tau(\sigma X) = \sum_\sigma f(\sigma)\psi_{\sigma^{-1}\tau}(X) = \sum_\alpha f(\tau\alpha^{-1})\psi_\alpha(X) \\ &= f(\tau)\psi_1(X) + \sum_{\alpha\neq 1} f(\tau\alpha^{-1})\psi_\alpha(X) = \sum_{\alpha\neq 1}(f(\tau\alpha^{-1}) - f(\tau))\psi_\alpha(X) \end{aligned}$$

Together with the computation of the eigenvalues above, this proves the formula for the determinant.

One also readily verifies that $|G|\psi_\tau = \sum_{\chi\neq 1}\chi(\tau^{-1})\chi$ (recall that $|G| = |\hat{G}|$) and $\chi = \sum_{\tau\neq 1}(\chi(\tau)-1)\psi_\tau$. Hence

$$T^{-1}\psi_\tau = \sum_{\chi\neq 1}\frac{\chi(\tau^{-1})}{|G|}T^{-1}\chi = \sum_{\chi\neq 1}\frac{\chi(\tau^{-1})}{|G|\sum_\alpha f(\alpha)\chi(\alpha)}\chi = \sum_{\sigma\neq 1}\sum_{\chi\neq 1}\frac{\chi(\tau^{-1})(\chi(\sigma)-1)}{|G|\sum_\alpha f(\alpha)\chi(\alpha)}\psi_\sigma.$$

This concludes the proof. ∎

We used the lemma with $G = \mathcal{G}/\Gamma = \mathrm{Gal}\,(\mathbb{K}/\mathbb{Q})$ and with the function f defined on the cosets of Γ by $f(a\Gamma) = \sum_{\gamma\in\Gamma}\log|\sin(\pi a\gamma/P)|$. Applying Lemma 8.4, one can invert the matrix $[\log|\sigma_i(\eta_j)|]_{\substack{1\leq i\leq m,\ i\neq i_0 \\ 1\leq j\leq m-1}}$ and compute the numbers δ_i and λ_i. Unfortunately, even this computation is too long for the large m. The two main reasons for this are the following.

[6] For the control of the precision in Case A, see [6], Lemma 2.4.2.

First, the bigger is m, the poorer is Baker's bound B_0, and, consequently, the higher accuracy is required for δ_i and λ_i. In the worst case $n = 719$, when $m = 359$, we needed more than 8000 decimal digits for both δ_i and λ_i.

Second, as one can see from (98), the entries of the inverse matrix are given as sums of $m-1$ terms. Further, δ_i and λ_i themselves are sums of $m-1$ terms involving the entries of the inverse matrix (see [8], Equation (16)). Hence, in order to find δ_i or λ_i, one has to compute, with extremely high precision, a double-nested sum of $(m-1)^2$ terms.

Due to a special form of the summands, we managed to express δ_i as a simple sum of $m-1$ terms. (We omit the details, which are trivial but very technical.) Unfortunately, we found no way to simplify the expression for λ_i. Therefore, computing the numbers λ_i is the far slowest step of the algorithm.

In view of this, the following simple observation is really inevaluable: to perform the reduction, one needs to know δ_i and λ_i only for two distinct values of i, together with a reasonably sharp upper bound for the remaining values (which can be easily obtained). This trivial remark allowed us, for instance, to reduce the computation time for $n = 719$ by more than 99%.

8.3 Enumerating small solutions

This was unnecessary, since the upper bound for $|x|$ implied by (87) or (87$'$) was always much smaller than e^8.

8.4 Computational time

Apart from the 15 days needed for the continued fraction expansions of lemma (6.4), the total computational time was roughly one year-machine shared on a few SUN UltraSparc-1 and PC Pentiums of various clock speed.

The single value $n = 719$, for which we had to take $m = 359$, took one fifth of the time by itself.

The program was written in C, and used the PARI library. Note that in Case B this was probably not the right choice, since we did not need any specific PARI-function except the 3-dimensional LLL. Probaly, we could considerably speed up the computation (by 70–80%, according to our estimates), if, instead (or together with) PARI, we had used libraries having special routines for the operations on very large floating points numbers (for instance, fast multiplication routines), such as `gmp` or `ntl`. Unfortunately, we realized this only when a substantial part of the computations had already been complete. At that point, the time we could still have gained by using that libraries did not justify the additional effort required for rewriting the program.

Note that the algorithm readily distributes; even for a single value of n, different values of i_0 can be treated by different machines once two rows of the inverse matrix have been computed.

We acknowledge the support of GDR MEDICIS and Technische Universität Graz, who kindly allowed us to use their computer resources to complete the computations described in this section.

Appendix (by M. Mignotte)

A variant of a theorem of Laurent-Mignotte-Nesterenko[7]

A.1 Introduction.

Let α_1, α_2 be two non-zero algebraic numbers, and let $\log\alpha_1$ and $\log\alpha_2$ be any determinations of their logarithms. We consider the linear form

$$\Lambda = b_2 \log\alpha_2 - b_1 \log\alpha_1,$$

[7]The original manuscript of Mignotte was edited by Yu. Bilu in order to make it compatible with the notation, style and purposes of the present paper. Bilu accepts full responsibility for all remaining inaccuracies.

where b_1 and b_2 are positive integers. Without loss of generality, we suppose that $|\alpha_1| \geq 1$ and $|\alpha_2| \geq 1$. Put

$$D = [\mathbb{Q}(\alpha_1, \alpha_2) : \mathbb{Q}] \, / \, [\mathbb{R}(\alpha_1, \alpha_2) : \mathbb{R}].$$

The main result of [21] is:

Theorem A.1.1 *Let K, L, R_1, R_2, S_1, S_2 be positive integers with $K \geq 3$ and $L \geq 2$. Let $\rho > 1$ be a real number. Put $R = R_1 + R_2 - 1$, $S = S_1 + S_2 - 1$, $N = KL$,*

$$g = \frac{1}{4} - \frac{N}{12RS}, \qquad b = \frac{\Big((R-1)b_2 + (S-1)b_1\Big)}{2} \left(\prod_{k=1}^{K-1} k!\right)^{-2/(K^2-K)}.$$

Let a_1, a_2 be positive real numbers such that[8]

$$a_i \geq \rho\,|\log \alpha_i| - \log|\alpha_i| + 2Dh(\alpha_i),$$

for $i = 1, 2$. Suppose that

$$\mathrm{Card}\left\{\alpha_1^r\alpha_2^s : 0 \leq r < R_1,\ 0 \leq s < S_1\right\} \geq L, \tag{A.1}$$

$$\mathrm{Card}\left\{rb_2 + sb_1 : 0 \leq r < R_2,\ 0 \leq s < S_2\right\} > (K-1)L, \tag{A.2}$$

$$K(L-1)\log\rho - (D+1)\log N - D(K-1)\log b - gL\,(Ra_1 + Sa_2) > 0. \tag{A.3}$$

Then

$$|\Lambda'| \geq \rho^{-KL + (1/2)} \quad \textit{with} \quad \Lambda' = \Lambda \cdot \max\left\{\frac{LSe^{LS|\Lambda|/(2b_2)}}{2b_2},\ \frac{LRe^{LR|\Lambda|/(2b_1)}}{2b_1}\right\}. \tag{A.4}$$

Here we consider the "complex" case: $\alpha_1 = -1$ and α_2 is a complex algebraic number of modulus one. From Theorem A.1.1, we shall deduce the following slight sharpening of Théorème 3 of [21].

Theorem A.1.2 *Let α be a complex algebraic number with $|\alpha| = 1$, but not a root of unity, and $\log\alpha$ any determination of its logarithm. We put $D = [\mathbb{Q}(\alpha) : \mathbb{Q}]/2$. We consider the linear form*

$$\Lambda = b_1 i\pi - b_2\log\alpha, \tag{A.5}$$

where b_1 and b_2 are positive integers. Let ρ be a positive number with $1 < \rho < e^4$, and let a be a positive number satisfying

$$a \geq \min\Big\{0.5\pi\rho, 0.5\rho|\log\alpha| + Dh(\alpha)\Big\}. \tag{A.6}$$

Put

$$\lambda = \log\rho, \quad a_1 = \rho\pi, \quad a_2 = 2a, \quad b' = b_1/a_2 + b_2/a_1.$$

Let k be a positive number satisfying

$$\frac{\lambda k}{2} - \frac{\sqrt{k}}{3} - \frac{1}{6a_1} + \frac{1}{24a_1(1 + 2a_1\sqrt{k})} \geq 0. \tag{A.7}$$

Let h be a positive number, and define the following quantities:

$$L = \min\Big\{6, 2 + [2h/\lambda]\Big\}, \quad K = 1 + [kLa_1a_2], \tag{A.8}$$

$$\varepsilon = \frac{1}{L\sqrt{k}}\sqrt{\frac{K}{K-1}} \cdot \max\left\{\frac{\sqrt{(K-1)La_2/a_1}}{\sqrt{(K-1)La_2/a_1} - 1} \cdot \frac{1}{a_2}, \frac{L+1}{2L} \cdot \frac{1}{a_1}\right\},$$

$$\delta_1 = \frac{\lambda}{2} + \frac{1}{k}\left(\frac{1}{6\rho\pi} + \frac{1}{3a}\right) + \frac{1}{K}\left(\log KL + D\log\frac{L\sqrt{e}}{2\pi}\right),$$

$$\delta = \max\left\{0, \log b' + \log\left(\frac{1+\varepsilon}{2\sqrt{k}}\right) + \frac{3}{2} + f_1(K)\right\},$$

where

$$f_1(x) = \frac{1}{2}\log\frac{x}{x-1} + \frac{\log x}{6x(x-1)} + \frac{\log\frac{x}{x-1}}{x-1}.$$

Assume that

$$h \geq \delta D + \delta_1. \tag{A.9}$$

Then we have the lower bound

$$\log|\Lambda| \geq -(kL^2a_1a_2 + L - 1/2)\lambda - 2\log L - \log\Big(1 + \sqrt{k}\max\{a_1, a_2\}\Big) - D\log 2. \tag{A.10}$$

[8] Recall that $h(\alpha)$ stands for the absolute logarithmic height of the algebraic number α (see Subsection 1.1).

We also obtain the following more “user-friendly” result.

Theorem A.1.3 *Let α and D be as in Theorem A.1.2, $\log\alpha$ the principal value of the logarithm (that is, $-\pi < \operatorname{Im}\log\alpha \le \pi$) and Λ the linear form (A.5) with positive integral coefficients b_1 and b_2. Let λ be a positive number satisfying $1.8 \le \lambda \le 4$, and put*

$$\begin{aligned}
&\rho = e^{\lambda}, \quad a = 0.5\rho\pi + Dh(\alpha), \quad B = \max(13, b_1, b_2),\\
&t = \frac{1}{6\pi\rho} - \frac{1}{48\pi\rho\,(1+2\pi\rho/3\lambda)}, \quad k = \left(\frac{1/3+\sqrt{1/9+2\lambda t}}{\lambda}\right)^2,\\
&\mathcal{H} = \max\left\{3\lambda, D\left(\log B + \log\left(\frac{1}{\pi\rho}+\frac{1}{2a}\right) - \log\sqrt{k} + 0.833\right) + \frac{3\lambda}{2} + \frac{1}{k}\left(\frac{1}{6\rho\pi}+\frac{1}{3a}\right) + 0.023\right\}.
\end{aligned}$$

Then

$$\log|\Lambda| \ge -\left(8\pi k\rho\lambda^{-1}\mathcal{H}^2 + 0.37\right)a - 2\log\mathcal{H} + 0.5\lambda + 2\log\lambda - (D+2)\log 2. \tag{A.11}$$

If $\lambda \le 3$ and $B \ge 234$ then one can also put

$$\mathcal{H} = D\left(\log B + \log\left(\frac{1}{\pi\rho}+\frac{1}{2a}\right) - \log\sqrt{k} + 0.833\right) + \frac{3\lambda}{2} + \frac{1}{k}\left(\frac{1}{6\rho\pi}+\frac{1}{3a}\right) + 0.023. \tag{A.12}$$

Though this estimate still looks rather messy, it is very efficient for practical computations, since one can specyfy λ in the most optimal way. To give the reader some idea of which kind of estimate he can expect, notice that, with $\lambda = 2.8$ one obtains, after easy computations, the following:

$$\log|\Lambda| \ge -D^2\left((32h(\alpha)D + 813)\left((\log B + 5/D)^2 + 3\right) + 2201\right) + 12. \tag{A.13}$$

We omit the proof of (A.13) since this particular result is not used in the main text of the paper.

A.2 Proof of Theorem A.1.2

We shall apply Theorem A.1.1 with a suitable choice of the parameters. The proof follows closely the proof of Théorème 3 of [21].

A.2.1 The parameters of the proof

The parameters L and K are defined in (A.8). Now put

$$R_1 = 2, \quad S_1 = [(L+1)/2], \quad R_2 = 1 + \left[\sqrt{(K-1)La_2/a_1}\right], \quad S_2 = 1 + \left[\sqrt{(K-1)La_1/a_2}\right].$$

We need several simple properties of these parameters. First, $\lambda k/2 - \sqrt{k}/3 > 0$ yields

$$k > 4/(9\lambda^2). \tag{A.14}$$

Using this together with $L \ge 6$ and $a_2 \ge \pi\rho$, we minorate K as follows:

$$K \ge kLa_1a_2 \ge (8/3)\pi^2(\rho/\lambda)^2 \ge (8/3)\pi^2e^2 > 194.4,$$

which means that $K \ge 195$ and $K - 1 \ge (194/195)K$. Further, $L - 1 \ge 2h/\lambda$, which yields

$$\lambda(L-1) - h \ge \lambda(L-1)/2. \tag{A.15}$$

Finally,

$$S_2 > \sqrt{(K-1)La_1/a_2} > La_1\sqrt{\frac{194}{195}k} > \frac{2}{3}\pi\frac{\rho}{\lambda}L\sqrt{\frac{194}{195}} \ge \frac{2}{3}\pi eL\sqrt{\frac{194}{195}} > 5.6L > S_1. \tag{A.16}$$

A.2.2 Estimates for $\log b$ and $gL(Ra_1 + Sa_2)$

Define R, S, b and g as in Theorem A.1.1. Then

$$\log b \le \log b' + \log\left(\frac{1+\varepsilon}{2\sqrt{k}}\right) + \frac{3}{2} + f_1(K) - \frac{\log\left(2\pi K/\sqrt{e}\right)}{K-1} \le \delta - \frac{\log\left(2\pi K/\sqrt{e}\right)}{K-1}. \tag{A.17}$$

The proof is essentially the same as that of Lemma 10 of [21]. Notice that

$$\begin{aligned}
\frac{R_1-1}{R_2-1} &\le \frac{R_2}{R_2-1}\cdot\frac{1}{\sqrt{(K-1)La_2/a_1}} \le \frac{R_2}{R_2-1}\cdot\sqrt{\frac{K}{K-1}}\frac{1}{\sqrt{k}La_2} \le \varepsilon,\\
\frac{S_1-1}{S_2-1} &\le \frac{S_1}{S_2} \le \frac{L+1}{2\sqrt{(K-1)La_1/a_2}} \le \sqrt{\frac{K}{K-1}}\left(1+\frac{1}{L}\right)\frac{1}{2\sqrt{k}La_1} \le \varepsilon
\end{aligned}$$

(we used here the inequality $S_2 > S_1$, which follows from (A.16)). Now proceeding exactly as in [21], p. 315, we obtain (A.17).

From Lemma 11 of [21] we get

$$gL(Ra_1 + Sa_2) \le \left(\sqrt{k}/3 + 1/(6a_1) - \xi\right) a_1a_2L^2 + \sigma a_2 L, \tag{A.18}$$

where

$$\sigma = \frac{1}{6} + \frac{2a_1}{3a_2} \quad \text{and} \quad \xi = \frac{1}{24a_1\left(1+2a_1\sqrt{k}\right)}.$$

Notice that

$$\delta_1 = \frac{\lambda}{2} + \frac{\sigma}{ka_1} + \frac{1}{K}\left(\log KL + D\log\frac{L\sqrt{e}}{2\pi}\right) \tag{A.19}$$

A.2.3 Study of condition (A.3)

Put

$$\Phi_0 = K(L-1)\lambda - (D+1)\log(KL) - D(K-1)\log b - gL(Ra_1+Sa_2).$$

Using (A.17), (A.18), (A.9) and (A.15), we obtain

$$\begin{aligned}
\Phi_0 \quad &\ge \quad K(L-1)\lambda - (D+1)\log(KL) - (K-1)D\delta + D\log\left(2\pi K/\sqrt{e}\right) - \\
&\quad - \left(\sqrt{k}/3 + 1/(6a_1) - \xi\right)a_1a_2L^2 - \sigma a_2 L\\
&\ge \quad K(L-1)\lambda - \log K - (D+1)\log L - (K-1)h + (K-1)\delta_1 + \\
&\quad +D\log\left(2\pi/\sqrt{e}\right) - \left(\sqrt{k}/3 + 1/(6a_1) - \xi\right)a_1a_2L^2 - \sigma a_2 L\\
&\ge \quad K(L-1)\lambda/2 - \log K - (D+1)\log L + h + \delta_1(K-1) + \\
&\quad +D\log\left(2\pi/\sqrt{e}\right) - \left(\sqrt{k}/3 + 1/(6a_1) - \xi\right)a_1a_2L^2 - \sigma a_2 L.
\end{aligned}$$

This implies that $\Phi_0 \ge \Phi + \Theta$, where

$$\begin{aligned}
\Phi \quad &= \quad KL\lambda/2 - \left(\sqrt{k}/3 + 1/(6a_1) - \xi\right)a_1a_2L^2,\\
\Theta \quad &= \quad -\lambda K/2 + h - \sigma a_2 L - \log K - (D+1)\log L + D\log\left(2\pi/\sqrt{e}\right) + \delta_1(K-1).
\end{aligned}$$

Since $K > kLa_1a_2$, we get

$$\Phi > \left(\lambda k/2 - \sqrt{k}/3 - 1/(6a_1) + \xi\right)a_1a_2L^2 \ge 0, \tag{A.20}$$

by the hypothesis on k.

A.2.4 Proof of Theorem A.1.2

We consider two cases.

First case: *The integers*

$$rb_2 + sb_1 \quad (0 \le r < R_2, \quad 0 \le s < S_2) \tag{A.21}$$

are not pairwise distinct.
In this case there exist positive integers $r \le R_2 - 1$ and $s \le S_2 - 1$ such that $rb_2 = sb_1$. We have $s|\Lambda| = b_2|ri\pi - s\log\alpha|$, and by the Liouville inequality $\log|ri\pi - s\log\alpha| \ge -Dsh(\alpha) - D\log 2$. Hence

$$\begin{aligned} \log\Lambda &\ge -(S_2-1)Dh(\alpha) - \log(S_2-1) - D\log 2 \\ &\ge -(S_2-1)a_2 - e^{-1}(S_2-1) - D\log 2 \\ &> -1.4(S_2-1)a_2 - D\log 2. \end{aligned} \tag{A.22}$$

It is easy to see that

$$1.4(S_2-1)a_2 \le kL^2a_1a_2\lambda. \tag{A.23}$$

Indeed, $S_2 - 1 \le \sqrt{(K-1)La_1/a_2} \le \sqrt{k}La_1$. This reduces (A.23) to the inequality $1.4 \le \sqrt{k}\lambda L$, which is an immediate consequence of (A.14) and $L \ge 6$. This proves (A.23), which shows that (A.22) is sharper than the desired (A.10). This completes the proof in the first case.

Second case: *The integers (A.21) are pairwise distinct.*
In this case

$$\begin{aligned} &\mathrm{Card}\,\{\alpha_1^r\alpha_2^s : 0 \le r < R_1,\ 0 \le s < S_1\} = 2S_1 \ge L, \\ &\mathrm{Card}\,\{rb_2 + sb_1 : 0 \le r < R_2,\ 0 \le s < S_2\} = R_2S_2 > (K-1)L. \end{aligned}$$

Thus, conditions (A.1) and (A.2) of Theorem A.1.1 are satisfied. Concerning condition (A.3) it remains to prove, by (A.20), that $\Theta \ge 0$.

Since $L \ge 6$, we have $\log(L\sqrt{e}/(2\pi)) > 0$, which yields $h > \delta_1$. Using this and (A.19), we obtain

$$\begin{aligned} \Theta &= (\delta_1 - \lambda/2)K - \sigma a_2L + h - \delta_1 - \log KL - D\log L + D\log\left(2\pi/\sqrt{e}\right) \\ &> (\delta_1 - \lambda/2 - \sigma/(ka_1))\,K - \log KL - D\log\left(L\sqrt{e}/(2\pi)\right) \\ &= 0, \end{aligned}$$

as wanted. Thus, condition (A.3) is also verified.

By Theorem A.1.1 we have

$$\log|\Lambda'| \ge -\lambda KL + \lambda/2, \tag{A.24}$$

where Λ' is defined in (A.4). To obtain from this an estimate for Λ, notice that

$$R = R_1 + R_2 - 1 \le 2 + \sqrt{k}La_2 \le (1 + \sqrt{k}a_2)L, \tag{A.25}$$

$$S = S_1 + S_2 - 1 \le (L+1)/2 + \sqrt{k}La_1 \le (1 + \sqrt{k}a_1)L, \tag{A.26}$$

We may assume that

$$\max\{LS\Lambda/(2b_2), LR\Lambda/(2b_1)\} \le 1/2. \tag{A.27}$$

(Indeed, if (A.27) is false then by (A.25) and (A.26) we have $L^2\left(1 + \sqrt{k}\max\{a_1, a_2\}\right)\Lambda \ge 1$, which implies that $\log\Lambda \ge -2\log L - \log\left(1 + \sqrt{k}\max\{a_1, a_2\}\right)$, sharper than (A.10).) By (A.27)

$$\Lambda' \le 0.5\sqrt{e}\Lambda L\max\{R, S\} \le \Lambda L^2\left(1 + \sqrt{k}\max\{a_1, a_2\}\right)$$

which together with (A.24) yields (A.10). Theorem A.1.2 is proved. ∎

A.3 Proof of Theorem A.1.3

We use Theorem A.1.2 with $h = \mathcal{H} - \lambda$. Define a_1, a_2, L, K, ε, b', δ and δ_1 as required in Theorem A.1.2, and veryfy the conditions (A.6), (A.7) and (A.9).

A.3.1 Verification of (A.6) and (A.7)

Since we use the principal value of the logarithm, $|\log \alpha| \le \pi$. Hence (A.6) is satisfied.

To veryfy (A.7), notice that $t < 1/(6a_1)$, which yields

$$\sqrt{k} < \frac{1}{3\lambda} + \frac{1}{\lambda}\sqrt{\frac{1}{9} + \frac{\lambda}{3a_1}} < \frac{2}{3\lambda} + \frac{1}{2a_1}. \tag{A.28}$$

This implies the inequalities

$$\frac{1}{6a_1} - \frac{1}{24\left(1+2\sqrt{k}a_1\right)} < \frac{1}{6a_1} - \frac{1}{24\left(1+2a_1\left(2/(3\lambda)+1/(2a_1)\right)\right)} = t,$$

$$\frac{\lambda k}{2} - \frac{\sqrt{k}}{3} - \frac{1}{6a_1} + \frac{1}{24\left(1+2\sqrt{k}a_1\right)} > \frac{\lambda k}{2} - \frac{\sqrt{k}}{3} - t = 0,$$

as wanted.

A.3.2 Estimating L, K, $f_1(K)$ and ε

Since $h = \mathcal{H} - \lambda \ge 2\lambda$, we have

$$6 \le L \le 2\mathcal{H}/\lambda. \tag{A.29}$$

Further,

$$t = \frac{1}{6a_1}\left(1 - \frac{1}{8(1+2\pi e^{\lambda}/(3\lambda))}\right) \ge \frac{1}{6\pi\rho}\left(1 - \frac{1}{8\left(1+2\pi e^{1.8}/(3\cdot 1.8)\right)}\right) > \frac{0.052226}{\rho},$$

which yields

$$\sqrt{k}a_1 = \pi\left(\frac{\rho}{3\lambda} + \sqrt{\frac{\rho^2}{9\lambda^2} + \frac{2\rho^2 t}{\lambda}}\right) > \pi\left(\frac{e^{\lambda}}{3\lambda} + \sqrt{\left(\frac{e^{\lambda}}{3\lambda}\right)^2 + \frac{0.104452e^{\lambda}}{\lambda}}\right)\Bigg|_{\lambda=1.8} > 7.50099. \tag{A.30}$$

Since

$$a_2 > a_1, \tag{A.31}$$

we obtain

$$K > \left(\sqrt{k}a_1\right)^2 L > 56.26485L > 337.5 \tag{A.32}$$

and since K is an integer, we have

$$K \ge 338. \tag{A.33}$$

Hence

$$f_1(K) \le f_1(338) < 0.0015. \tag{A.34}$$

and

$$\varepsilon \le \frac{1}{\sqrt{k}a_1 L} \cdot \sqrt{\frac{K}{K-1}} \cdot \frac{\sqrt{(K-1)L}}{\sqrt{(K-1)L}-1} < \frac{1}{7.50099\cdot 6} \cdot \sqrt{\frac{338}{337}} \cdot \frac{\sqrt{337\cdot 6}}{\sqrt{337\cdot 6}-1} < 0.02276. \tag{A.35}$$

A.3.3 Estimating δ and δ_1 and verification of (A.9)

We have

$$\begin{aligned}\log b' + \log\left(\frac{1+\varepsilon}{2\sqrt{k}}\right) + \frac{3}{2} + f_1(K) &< \log B + \log\left(\frac{1}{a_1} + \frac{1}{a_2}\right) + \log 1.02276 - \\ &\quad - \log 2 - \log\sqrt{k} + 1.5 + 0.0015 \\ &< \log B + \log\left(\frac{1}{a_1} + \frac{1}{2a}\right) - \log\sqrt{k} + 0.83086 \end{aligned} \tag{A.36}$$

Using (A.28), we obtain

$$\begin{aligned}\log\left(1/a_1 + 1/(2a)\right) - \log\sqrt{k} &\ge \log(1/a_1) - \log\left(2/(3\lambda) + 1/(2a_1)\right) \\ &= -\log\left(2\pi e^{\lambda}/(3\lambda) + 1/2\right)\Big|_{\lambda=4} \\ &> -3.37.\end{aligned}$$

Since $B \geq 13$, in addition to (A.36) we have

$$\log B + \log\left(\frac{1}{a_1} + \frac{1}{2a}\right) - \log\sqrt{k} + 0.83086 > \log 13 - 3.37 + 0.83086 > 0.$$

Hence

$$\delta < \log B + \log\left(1/a_1 + 1/(2a)\right) - \log\sqrt{k} + 0.83086. \tag{A.37}$$

Further,

$$\frac{\log(KL)}{K} \leq \frac{\log 338}{338} + \frac{\log 6}{56.26485 \cdot 6} < 0.02254,$$

and

$$\frac{\log\left(L\sqrt{e}/(2\pi)\right)}{K} \leq \left.\frac{\log\left(L\sqrt{e}/(2\pi)\right)}{56.26485L}\right|_{L=2\pi\sqrt{e}} < 0.00172.$$

This shows that

$$\delta_1 < \frac{\lambda}{2} + \frac{1}{k}\left(\frac{1}{6\rho\pi} + \frac{1}{3a}\right) + 0.02254 + 0.00172D.$$

Hence

$$h = \mathcal{H} - \lambda > D\delta + \delta_1$$

by the very definition of $\mathcal{H}$. Condition (A.9) is verified.

A.3.4 Completing the proof

Thus, the assumption of Theorem A.1.2 are satisfied. Applying it, and using (A.29) and (A.31), we obtain

$$\begin{aligned}
\log|\Lambda| \quad &\geq \quad -(kL^2a_1a_2 + L - 1/2)\lambda - 2\log L - \log\left(1 + \sqrt{k}\max\{a_1, a_2\}\right) - D\log 2 \\
&\geq \quad -\left(8k\pi\rho\lambda^{-1}\mathcal{H}^2 + a^{-1}\log\left(1 + 2a\sqrt{k}\right)\right)a - 2\log\mathcal{H} + \lambda/2 + 2\log\lambda - \\
&\quad\quad -(D+2)\log 2.
\end{aligned} \tag{A.38}$$

The function $x^{-1}\log(1+x)$ decreases on the interval $[1, +\infty)$. Since $2a\sqrt{k} > a_1\sqrt{k} > 1$ by (A.30), we obtain, using (A.28) that

$$a^{-1}\log\left(1 + 2a\sqrt{k}\right) \leq \frac{2\log\left(1 + a_1\sqrt{k}\right)}{a_1} \leq \frac{\left.2\log\left(2\pi e^{\lambda}/(3\lambda) + 1/2\right)\right|_{\lambda=4}}{\left.\pi e^{\lambda}\right|_{\lambda=1.8}} < 0.37$$

Substituting this to (A.38), we obtain (A.11).

To complete the proof, it remains to show that, when $\lambda \leq 3$ and $B \geq 234$, one can define $\mathcal{H}$ as in (A.12). What we have to show is that under these assumptions, the right-hand side of (A.12) is greater than 3λ. Since

$$\log B + \log\left(\frac{1}{\pi\rho} + \frac{1}{2a}\right) - \log\sqrt{k} + 0.833 > \log B - \lambda - \log\pi - \log\sqrt{k} + 0.833 > 0,$$

it suffices to show that

$$\log B - \lambda - \log\pi - \log\sqrt{k} + 0.833 + 3\lambda/2 + 0.023 > 3\lambda,$$

or that

$$\log B > 5\lambda/2 + \log\sqrt{k} + 0.28873.$$

Using (A.28), we obtain

$$\frac{5\lambda}{2} + \log\sqrt{k} + 0.28873 \leq \log\left.\left(\frac{2}{3}\frac{e^{5\lambda/2}}{\lambda} + \frac{e^{3\lambda/2}}{2\pi}\right)\right|_{\lambda=3} + 0.28873 < 5.454 < \log 234 \leq \log B,$$

as wanted. This completes the proof of Theorem A.1.3. ∎

References

[1] *A. Baker*, Contribution to the theory of Diophantine equations I: On the representation of integers by binary forms, Phil. Trans. Roy. Soc. London **A263** (1968), 173–191.

[2] *A. Baker, H. Davenport*, The equations $3x^2 - 2 = y^2$ and $8x^2 - 7 = z^2$, Quat. J. Math. Oxford (2) **20** (1969), 129–137.

[3] *A. Baker and G. Wüstholz*, Logarithmic forms and group varieties, J. reine angew. Math. **442** (1993), 19–62.

[4] *A.S. Bang*, Taltheoretiske Undersøgelser, Tidsskrift for Mat. (5), **4** (1886), 70–80, 130–137.

[5] *Yu. Bilu*, Solving superelliptic Diophantine equations by the method of Gelfond–Baker, preprint 94-09, Mathématiques Stochastiques, Univ. Bordeaux 2, 1994.

[6] *Yu. Bilu, G. Hanrot*, Solving Thue Equations of High Degree, J. Number Th., **60** (1996), 373–392.

[7] *Yu. Bilu, G. Hanrot*, Solving superelliptic Diophantine equations by Baker's method, Compositio Math., **112** (1998), 273–312.

[8] *Yu. Bilu, G. Hanrot*, Thue equations with composite fields, Acta Arith., to appear.

[9] *G.D. Birkhoff, H.S. Vandiver*, On the integral divisors of $a^n - b^n$, Ann. Math. (2) **5** (1904), 173–180.

[10] *J. Buchmann*, Computing class groups and regulators in subexponential time, *in* Séminaire de Théorie des Nombres de Paris 1988-89, 27–39, Progr. Math. **91**, Birkhäuser, Boston, 1990.

[11] *P.D. Carmichael*, On the numerical factors of the arithmetic forms $\alpha^n \pm \beta^n$, Ann. Math. (2), **15** (1913), 30–70.

[12] *H. Cohen*, A Course in Computational Algebraic Number Theory, Graduate Texts in Math. **138**, Springer, 1993.

[13] *H. Cohen, F. Diaz y Diaz, M. Olivier*, Subexponential algorithms for class group and unit computations, J. Symbolic Comput., **24** (1997), 433–441.

[14] *L.K. Durst*, Exceptional real Lehmer sequences, Pacific J. Math., **9** (1959), 437–441.

[15] *L.K. Durst*, Exceptional real Lucas sequences, Pacific J. Math., **11** (1961), 489–494.

[16] *E. Friedman*, Analytic formulas for the regulator of a number field, Inv. Math., **98** (1989), 599–622.

[17] *G. Hanrot,* Résolution effective d'équations diophantiennes: algorithmes et applications, Thèse, Université Bordeaux 1, 1997.

[18] *G. Hanrot*, Solving Thue equations without the full unit group, Math. Comp., to appear.

[19] *A.Ya. Khintchine*, Continued fractions (in Russian), "Nauka", Moscow, 1978.

[20] *S. Lang*, Fundamentals of Diophantine Geometry, Springer, 1983.

[21] *M. Laurent, M. Mignotte, Y. Nesterenko*, Formes linéaires en deux logarithmes et déterminants d'interpolation, J. Number Theory **55** (1995), 285–321.

[22] *D.H. Lehmer*, An extended theory of Lucas' functions, Ann. Math. **31** (1930), 419–448.

[23] *A.K. Lenstra, H.W. Lenstra jr., L. Lovász,* Factoring polynomials with rational coefficients, Math. Ann. **261** (1982), 515–534.

[24] *F.J. van der Linden*, Class Number Computations of Real Abelian Number Fields, Math. Comp. **39** (1982), 693–707.

[25] *E. Lucas*, Théorie des fonctions numériques simplement périodiques, Amer. J. Math. **1** (1878), 184–240, 289–321.

[26] *J.M. Masley*, Class numbers of real cyclic number fields with small conductor, Compositio Math. **37** (1978), 297–319.

[27] *A. Pethő*, Computational Methods For the Resolution of Diophantine Equations, *in* R.A. Mollin (ed.), Number Theory: Proc. First Conf. Can. Number Th. Ass., Banff, 1988, de Gruyter, 1990, 477–492.

[28] *M.E. Pohst*, Eine Regulatorabschätzung, Abh. Math. Sem. Hamburg **47** (1978), 95–106.

[29] *M.E. Pohst*, Computational Algebraic Number Theory, DMV Seminar **21**, Birkhäuser, Basel, 1993.

[30] *M.E. Pohst, H. Zassenhaus*, Algorithmic Algebraic Number Theory, Cambridge Univ. Press, 1989.

[31] *L.P. Postnikova, A. Schinzel*, Primitive divisors of the expression $a^n - a^n$ in algebraic number fields (Russian), Mat. Sbornik **75** (1968), 171–177; Math. USSR Sbornik **4** (1968), 153–159.

[32] Proceedings of the International Congress of Mathematicians, Zürich, 1994, Volume **1**, Birkhäuser, 1995.

[33] *P. Ribenboim*, The Fibonacci numbers and the Arctic Ocean, *in* Behara, Fritsch and Lintz (eds.), Symposia Gaussiana, Conf. A, de Gruyter, 1995, 41–83.

[34] *A. Schinzel*, The intrinsic divisors of Lehmer numbers in the case of negative discriminant, Ark. Mat. **4** (1962), 413–416.

[35] *A. Schinzel*, On primitive prime factors of Lehmer numbers, Acta Arith. **8** (1962/63), 213–223, 251–257; **15** (1968), 49–70.

[36] *A. Schinzel*, Primitive divisors of the expression $A^n - B^n$ in algebraic number fields, J. reine angew. Math. **268/269** (1974), 27–33.

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

[38] *C. Stewart*, On divisors of Fermat, Fibonacci, Lucas and Lehmer numbers, Proc. London Math. Soc. (3) **35** (1977), 425-447.

[39] *N. Tzanakis, B.M.M. de Weger*, On the Practical Solution of the Thue Equation, J. Number Th. **31** (1989), 99–132.

[40] *P.M. Voutier*, Primitive divisors of Lucas and Lehmer sequences, Math. Comp. **64** (1995), 869–888.

[41] *P.M. Voutier*, Primitive divisors of Lucas and Lehmer sequences, II, J. Theor. Nombres Bordx., **8** (1996), 251–274.

[42] *P.M. Voutier*, Primitive divisors of Lucas and Lehmer sequences, III, Math. Proc. Cambridge Phil. Soc. **123** (1998), 407–419.

[43] *L.C. Washington*, Introduction to cyclotomic fields, Graduate Texts in Math. **83**, Springer, 1982.

[44] *M. Ward*, The intrinsic divisors of Lehmer numbers, Ann. Math. (2), **62** (1955), 230–236.

[45] *R. Zimmert*, Ideale kleiner Norm in Idealklassen und eine Regulatorabschätzung, Invent. Math., **62** (1981), 367–380.

[46] *K. Zsigmondy*, Zur Theorie der Potenzreste, Monatsh. Math. **3** (1892), 265–284.

Bilu: Mathematisches Institut, Universität Basel, Rheinsprung 21, 4051 Basel, Switzerland, e-mail yuri@math.unibas.ch
Hanrot: Projet POLKA, INRIA Lorraine, Technopole de Nancy-Brabois, 615, rue du Jardin Botanique, B.P. 101, F-54600 Villers-les-Nancy, France, e-mail: Guillaume.Hanrot@loria.fr
Voutier: 110 Hornsey Park Road, London, N8 0JY UK, e-mail: paul@optrak.co.uk
Mignotte: Institut de Mathematiques, Université Louis Pasteur, 7 rue René Descartes, 67087 Strasbourg, France, e-mail: mignotte@math.u-strasbg.fr