ROCKY MOUNTAIN
JOURNAL OF MATHEMATICS
Volume 34, Number 4, Winter 2004

HOW SMOOTH IS $\phi(2^n + 3)$?

FLORIAN LUCA

Introduction. For any integer n let $P(n)$ denote the largest prime divisor of n with the convention that $P(\pm 1) = P(0) = 1$. We also let $\phi(n)$ denote the Euler function of n. In this paper, we show, among other things, that $P(\phi(2^n + 3))$ tends to infinity with n on a set of n of asymptotic density 1.

The example 2^n+3 that we chose is not incidental and is, in a certain sense, the smallest example of numbers of the form $2^n + b$, with a fixed integer b, and varying positive integers n, which is interesting for our type of problem. Suppose, let's say, that instead we look at the numbers 2^n+1. It is then well-known (see [**2**]), that for all sufficiently large n, the number 2^n+1 will have a prime divisor p which is congruent to 1 modulo n. In particular, n divides $\phi(2^n+1)$, and therefore $P(\phi(2^n+1)) \geq P(n)$. Since $P(n)$ tends to infinity on a set of n of asymptotic density 1 (see [**3**]), we get that $P(\phi(2^n+1))$ tends to infinity on a set of n of asymptotic density 1 as well. This example also hints as to why it is difficult to show that $P(\phi(2^n + 1))$ tends to infinity with n for *all* n. In fact, when $n = 2^k$ is a power of two, the number $F_k := 2^{2^k} + 1$ is what is known as a *Fermat number*. It is not known if there are infinitely many Fermat numbers which are primes, nor is it known if there are infinitely many Fermat numbers which are composite, but the standard believed conjecture is that there should be only finitely many Fermat numbers which are primes (see [**5**]). However, if this were not so, that is if infinitely many Fermat numbers F_k were primes, then for such k we would have $P(\phi(2^{2^k} + 1)) = 2$, which shows why it is difficult to prove that $P(\phi(2^n + 1))$ tends to infinity with n for *all* values of n. Since $\phi(2^n + 2) = \phi(2^{n-1} + 1)$, it follows that the same arguments as above apply to $P(\phi(2^n + 2))$, which is why we have chosen to look at the numbers of the form $2^n + 3$.

The numbers of the form $2^n + 3$ were chosen only as an example, and in this paper we will formulate and prove our main Theorem in

This work was supported in part by grants ECOS-ANUIES M02-M01, SEP-CONACYT 37259-E, and SEP-CONACYT 37260-E.

somewhat greater generality.

Let r and s be two non-zero integers with $r^2 + 4s \neq 0$. A *binary recurrent sequence of integers* $(u_n)_{n\geq 0}$ is a sequence of integers such that

$$u_{n+2} = ru_{n+1} + su_n \tag{1}$$

holds for all $n \geq 0$. It is well-known that there exist two constants $a,\ b$ such that the formula

$$u_n = a\alpha^n + b\beta^n \tag{2}$$

holds for all $n \geq 0$, where α and β are the two roots of the *characteristic equation*

$$x^2 - rx - s = 0. \tag{3}$$

The sequence $(u_n)_{n\geq 0}$ is said to be *non-degenerate* if $ab \neq 0$ and α/β is not a root of 1. Throughout this paper, we assume that $|\alpha| \geq |\beta|$.

Our main motivation in this paper is to give a lower bound of the type $P(\phi(|u_n|)) \geq f(n)$ with some increasing function f defined on the set of positive integers whose range is in the set of positive real numbers and which tends to infinity with n, and such that the above inequality holds for almost all positive integers n. Unfortunately, we cannot give such a lower bound *for all* binary recurrent sequences of integers $(u_n)_{n\geq 0}$, but we can do so for the sequences which satisfy certain technical assumptions. The assumptions in force throughout this paper are:

A1. *α is a rational number, and*

A2.1. *either β does not divide α, or*

A2.2. *β divides α but a/b is not an integer, or*

A.2.3 *β divides α, a/b is an integer, but a/b and β/α are multiplicatively dependent.*

Notice first of all that assumption A1 together with the fact that α and β are the roots of a monic quadratic equation with integer coefficients, imply that both α and β are integers, and therefore a and b are rational numbers. Thus, it makes sense to say that β does

not divide α (as in A2.1 above), or that a/b is not an integer (as in A2.2 above).

Our main result is the following:

Theorem. *Assume that* $(u_n)_{n\geq 0}$ *is a binary recurrent sequence of integers satisfying recurrence* (1)*, general formula* (2)*, and assumptions* A1 *and* A2 *above. Then, there exists an effectively computable positive constant* c_1 *such that the inequality*

$$P(\phi(|u_n|)) > c_1(\log n)^{1/6}(\log\log n)^{5/6} \tag{4}$$

holds for all positive integers n *except, perhaps, for a set of* n *of asymptotic density* 0*.*

Notice that the sequence of general term $u_n := 2^n + 3$ is binary recurrent, satisfies the general formula (2) with $(a,\ b,\ \alpha,\ \beta) = (1,\ 3,\ 2,\ 1)$, and so it also satisfies assumptions A1 and A2.3 of the above Theorem.

It would be of interest to remove the assumptions A1 and A2 and to prove that the above Theorem still holds. Unfortunately, this seems very hard to do. Assume first that α is an integer. Take, for example, $u_n := a2^n + 1$, where a is a fixed odd integer with $a > 1$. Notice that $(u_n)_{n\geq 0}$ does not satisfy any one of the assumptions from A2. Then obviously $P(\phi(u_n)) \leq P(a)$ provided that u_n is prime. In order to prove the Theorem for the sequence $(u_n)_{n\geq 0}$, we should be able first of all to infer that u_n is prime only on a set of positive integers n of asymptotic density zero. While this is probably true, the only result in this sense which is available to us is that this is indeed so if both the extended Riemann hypothesis and some other assumptions (such as Hypothesis 12 on page 112 of [**4**]) are assumed to hold. That is, there is no unconditional result claiming that indeed u_n is prime only on a set of n of asymptotic density zero.

The assumption A1 has been imposed just for technical reasons but our proof relies essentially on it. We shall describe where this assumption is needed in our argument in the next section.

We close this section by recalling that there are many papers in the literature dealing with finding non-trivial lower bounds for $P(u_n)$ in

terms of n. For example (see [**7**]), it is known that the inequality

$$P(u_n) > c_2\Big(\frac{n}{\log n}\Big)^{\frac{1}{d+1}} \tag{5}$$

holds for all positive integers n, where d is the degree of α over the rationals and c_2 is a computable constant depending only on $\alpha,\ \beta,\ a$, and b. If one wants a lower bound on $P(u_n)$ which holds not for *all positive integers* n, but only for *almost all positive integers* n, then it is known (see [**8**]) that the inequality

$$P(u_n) > \epsilon(n) n \log n \tag{6}$$

holds for all n, except for a set of n of asymptotic density zero, where $\epsilon(n)$ is any real valued function for which $\lim_{n\to\infty} \epsilon(n) = 0$. The interested reader should consult the survey paper [**9**] for a detailed account of the existing results concerning this type of problem.

A guide to the proof of the Theorem. The purpose of this section is twofold. First, we show that we can reduce the general problem to a slightly more particular one. Secondly, since the proof of the Theorem is rather involved, we give an informal account of the general principle behind the proof of the Theorem. The detailed proof appears in the next section.

Notations. Throughout the proof, we use $c,\ c_1,\ c_2, \dots$ to denote computable constants depending on $\alpha,\ \beta,\ a,\ b$, and we use the Vinogradov symbols $\ll$ and $\gg$ as well as the Landau symbols O and o with the meaning that they depend on the initial data $r,\ s,\ u_0$ and u_1. For every positive integer k and any large positive real number x we let $\log_k(x)$ be the $\log(\max(\log_{k-1}(x),\ e))$ for $k \geq 2$, and $\log_1(x) := \max(\log x,\ 1)$, where log denotes the natural logarithm function. Notice that for large x the function $\log_k(x)$ is nothing else but the composition of the natural logarithm function with itself k-times evaluated in x.

The reduction of the Problem. We shall suppose that $|\alpha| > |\beta|$. We show that we can assume that $a,\ b,\ \alpha,\ \beta$ are all integers, that $\gcd(a\alpha,\ b\beta) = 1$, that $\alpha > 0$ and that $a > 0$.

Indeed, if a and b are rational numbers which are not integers, then we may replace the sequence of general term u_n by the sequence of general term $\Delta u_n = (\Delta a)\alpha^n + (\Delta b)\beta^n$, where Δ is the greatest common denominator of the numbers a and b. In particular, we may assume that a and b are integers.

Set $l := \gcd(\alpha,\ \beta)$. If $l > 1$, then we may replace the sequence of general term u_n by the sequence of general term $u_n/l^n = a(\alpha/l)^n + b(\beta/l)^n$ and therefore we may assume that α and β are coprime. The same discussion applies to see that we may assume that a and b are coprime.

Now set $v_n := \gcd(a\alpha^n, b\beta^n) = \gcd(a, \beta^n) \cdot \gcd(b, \alpha^n)$. Suppose that there exists a positive integer n such that $v_n > 1$. In this case, there exists n_0 such that $v_n = v$ is constant for $n > n_0$, and by eliminating the first n_0 terms of the sequence $(u_n)_{n\geq 0}$, reindexing, and finally replacing the sequence of general term u_n by the sequence of general term $u_n/v = \big((a\alpha^{n_0})/v\big)\alpha^n + \big((b\beta^{n_0})/v\big)\beta^n$, it follows that we may indeed assume that $\gcd(a\alpha, b\beta) = 1$.

Finally, if $\alpha < 0$, we may replace the sequence of general term u_n by the sequence of general term $(-1)^n u_n = a(-\alpha)^n + b(-\beta)^n$ and thus assume that $\alpha > 0$. In particular, $\alpha > |\beta|$. Changing the sign of a, and/or eliminating the first few terms from the sequence $(u_n)_{n\geq 0}$, if needed, it follows that we may assume that $a > 0$ and that $u_n > 0$ for all $n \geq 0$.

Notice that all the transformations employed above do not affect the assumptions A1 and A2 of the Theorem, nor do they affect $P(\phi(u_n))$ once this is sufficiently large.

A sketch of the proof of the Theorem. We shall use $f(x)$ for some increasing function which is $O((\log x)^c)$, where $0 < c < 1/2$ is an absolute constant, and we shall try to determine such a function f such that the inequality $P(\phi(u_n)) < f(x)$ can hold only for a set of cardinality $o(x)$ of positive integers n in the interval $\big(x/\log x,\, x\big)$. Assume that n is one of these numbers. We start by writing $u_n := AB$, where A accounts for the multiplicative contribution of all prime numbers $p < f(x)$ in u_n and B accounts for the multiplicative contribution of the remaining primes in u_n. Using lower bounds for linear forms in logarithms, we show that A is small, thus B is large. Let p be any

prime divisor of B. For such p, we write $d := d(p)$ for the smallest positive integer k so that p divides the kth member of the Lucas sequence $(L_m)_{m\geq 0}$ of general term $L_m := (\alpha^m - \beta^m)/(\alpha - \beta)$ for all $m \geq 0$. We show that such a number $d(p)$ exists. Thus, p is a primitive divisor of $L_{d(p)}$. We then use an elementary counting argument to show that primes p dividing B such that $d(p)$ is "large" (say, at least as large as $\exp((\log x)^{1-2c})$) must exist, and we pick q to be the smallest one which satisfies the fact that $d(q)$ is as large as indicated above. We then show that for most n this number q is "small", namely at least as small as $\exp(\exp(3(\log x)^{1-2c}))$. This is the hardest part of the proof. It relies on rewriting the expression $u_n - \phi(u_n) = u_n - \phi(A) \cdot \phi(B)$ in such a way as to point out that the contribution of this "smallest" prime number q for which $d(q)$ satisfies the desired inequality can be more or less read off from an expression which is a difference of two numbers, each one of them being a multiplicative combination of rather small numbers appearing at rather large powers, such that the expression is suitable to deal with via lower bounds for linear forms in logarithms. If the expression we are looking at is non-zero, then we apply linear forms in logarithms to deal with this contribution of q. However, if the expression is zero, then linear forms in logarithms are useless, but we recognize an S-unit equation. To deal with this last case, we employ effective results on the number of solutions of such S-unit equations to infer that such equations can appear only for relatively few values of the positive integer n. Here is where assumption A2 is used. Having surpassed this point, we now notice that this number q is chosen in a unique way in terms of n, and moreover that if q is fixed, then n can run only in a certain arithmetical progression modulo $d(q)$. The assumption A1 now enters into the picture by saying that since q is primitive for $L_{d(q)}$, it follows that $q \equiv 1 \pmod{d(q)}$. Since $q - 1$ is $f(x)$-smooth (because $q - 1$ divides $\phi(u_n)$ which is assumed to be $f(x)$-smooth), we get that $d(q)$ is an $f(x)$-smooth number. If we were not assuming A1, then it is only known that if q is primitive for $L_{d(q)}$ then one of the congruences $q \equiv \pm 1 \pmod{d(q)}$ holds, but should q be congruent to -1 modulo $d(q)$ the fact that $q - 1$ is $f(x)$-smooth will tell us nothing about $d(q)$. Thus, under A1, q is a prime number such that $d(q)$ is an $f(x)$-smooth number which can run only in an interval bounded both from below and from above by some functions depending on x alone, and for each such q the positive integer $n < x$ sits in a unique arithmetic progression modulo $d(q)$. We then employ an elementary argument to

count how many such positive integers n there can be, and matching up all the bounds arising either from applications of lower bounds for linear forms in logarithms, or for the number of solutions of S-unit equations together with this last count, we finally get that the number of such positive integers $n < x$ satisfying $P(\phi(u_n)) < f(x)$ is $o(x)$ when $f(x)$ is chosen as indicated in the statement of the Theorem.

The proof of the Theorem. We use $f(x)$ to denote some function of x which is increasing for large values of x and tends to infinity with x. We shall try to find the best (i.e., "largest") such function f (which comes out of our arguments) and for which the inequality

$$P(\phi(u_n)) < f(x) \tag{7}$$

holds only on a set of cardinality $o(x)$ of positive integers n belonging to the interval $\big(x/\log x,\, x\big)$. As the conclusion of the Theorem suggests, our best f is a constant multiple of $(\log x)^{1/6} \cdot (\log_2 x)^{5/6}$.

For the moment, we shall work with an unknown such function f. Pick a large x_0 such that $f(x_0) > ab\alpha\beta(a+b)(\alpha-\beta)$, let $x > x_0$ be a large positive real number, and let n be a positive integer in the interval $x/\log x < n < x$ for which $P(\phi(u_n)) < f(x)$ holds. Assume that $p_1 < p_2 < \cdots < p_t < f(x)$ are all the prime numbers less that $f(x)$ and set $\mathcal{S} := \{p_1^{\alpha_1} \cdot \cdots \cdot p_t^{\alpha_t} \mid \alpha_i \geq 0\}$ to be the set of all positive integers m with $P(m) < f(x)$. Notice that for large x we have $t := \pi(f(x)) \leq \dfrac{2f(x)}{\log(f(x))}$. Since $\phi(u_n) \in \mathcal{S}$, we may write

$$u_n = \prod_{p^\alpha || n} p^\alpha = A \cdot B, \tag{8}$$

where

$$A := \prod_{\substack{p^\alpha || n \\ \alpha > 1}} p^\alpha \quad \text{and} \quad B := \prod_{p || n} p. \tag{9}$$

Clearly,

$$\phi(u_n) = \frac{A}{\mathrm{rad}(A)} \cdot \phi(\mathrm{rad}(A)) \cdot \phi(B), \tag{10}$$

where for a positive integer k we write $\mathrm{rad}(k) := \prod_{p|k} p$. Since $p \mid \phi(u_n)$ whenever $p \mid A$, it follows that $A \in \mathcal{S}$. We now bound the size of A. For every fixed prime $p < f(x)$, we have, by a standard application of linear forms in p-adic logarithms (see [**10**]), that if $p^{\alpha_p} || u_n$, then

$$\alpha_p \ll \frac{p}{\log p} \log n \le \frac{p}{\log p} \log x, \tag{11}$$

where the constant understood in $\ll$ above depends only on the initial data $r,\ s,\ u_0,\ u_1$ but not on p or x. Hence,

$$\log A = \sum_{p<f(x),\ p^{\alpha_p}||u_n} \alpha_p \log p \ll \log x \sum_{p<f(x)} p \ll f(x)^2 \log x. \tag{12}$$

Since certainly

$$\log u_n \gg n \gg \frac{x}{\log x},$$

it follows that

$$\log B = \log u_n - \log A \gg \frac{x}{\log x} - f(x)^2 \log x \ge \frac{x}{2 \log x}, \tag{13}$$

holds for any large enough real number x, provided that our function $f(x)$ satisfies $f(x) < (\log x)^{c_1}$ with some constant c_1, and large enough values of x. Since the function f shown at (4) certainly satisfies the above inequality with, say $c_1 := 1/5$, we get that inequality (13) holds for all large enough values of x. In particular, $B > 1$.

Now let $p \ge f(x)$ be any prime number which divides some member of the sequence $(u_n)_{n\ge 0}$. For this p, set $r := r(p)$ to be the minimal non-negative integer k for which $p \mid u_k$ and set $d := d(p)$ to be the minimal positive integer k for which p divides the kth term of the Lucas sequence $(L_n)_{n\ge 0}$ of general term

$$L_n := \frac{\alpha^n - \beta^n}{\alpha - \beta}, \quad \text{for } n \ge 0. \tag{14}$$

We claim that d exists, that $r < d$, and that $p \mid u_n$ if and only if $n \equiv r \pmod{d}$. To see this, notice that since $(u_n)_{n\ge 0}$ is periodic modulo p, infinitely many numbers n exist such that $p \mid u_n$. Pick

$m > n$ to be such that $p \mid u_n,\ p \mid u_m$ and the difference $m - n = k$ is minimal. In particular,

$$u_m = u_{n+k} = a\alpha^{n+k} + b\beta^{n+k} =$$

$$a\alpha^n(\alpha-\beta)\frac{\alpha^k-\beta^k}{\alpha-\beta} + \beta^k(a\alpha^n + b\beta^n) = a\alpha^n(\alpha-\beta)L_k + \beta^k u_n, \tag{15}$$

and since p divides both u_m and u_n and $p > a\alpha(\alpha-\beta)$, we read that $p \mid L_k$. From the well-known divisibility properties of Lucas sequence $(L_n)_{n\geq 0}$, it follows that $d \mid k$, and now the same argument as above shows that if $p \mid u_n$, then $p \mid u_{n+d}$ as well. Hence, by the minimality of k, we get $d = k$, and then further, by the minimality of r, we get that the number r is less than d, and that it is the unique non-negative integer $l < d$ for which $p \mid u_l$, and finally that any positive integer n for which $p \mid u_n$ must be congruent to r modulo d and conversely, if $n \equiv r \pmod{d}$, then $p \mid u_n$. We notice that $r > 0$, which follows from the fact that we are assuming that $p > a + b = u_0 > 0$.

We now pick q to be the smallest prime divisor of B for which $d(q) > x^{1/t^2}$. We show that this q exists and we find an upper bound on it. To show that q exists, let

$$C(x) := \prod_{p,\ d(p)\leq x^{1/t^2}} p.$$

Then certainly

$$C(x) \mid \prod_{1\leq d\leq x^{1/t^2}} L_d,$$

therefore

$$\begin{aligned}\log C(x) &\leq \log\Big(\prod_{1\leq d\leq x^{1/t^2}} L_d\Big) = \sum_{1\leq d\leq x^{1/t^2}} \log L_d \\ &\ll \sum_{1\leq d\leq x^{1/t^2}} d = O(x^{2/t^2}).\end{aligned} \tag{16}$$

Set

$$D := \gcd(B, C(x)), \tag{17}$$

and write

$$B = DE, \tag{18}$$

where obviously

$$E = \prod_{\substack{p \ge f(x),\ p \mid u_n \\ d(p) > x^{1/t^2}}} p. \tag{19}$$

By (16), we get

$$\log D \le \log C(x) \le c_2 x^{2/t^2}, \tag{20}$$

with some constant c_2, and by (13) and (20), we get that

$$\log E = \log B - \log D \ge \frac{x}{2\log x} - c_2 x^{2/t^2} > \frac{x}{3\log x} \tag{21}$$

holds for sufficiently large x. In particular, such a prime number q exists, and we write it as $q := q(n)$, and set $d(n) := d(q(n))$. To get an upper bound on q, write

$$E := q_1 q_2 \cdots q_k, \tag{22}$$

where $q = q_1 < q_2 < \cdots < q_k$ are distinct primes. Certainly, $E \le \phi(u_n) < u_n$, and therefore

$$2^k \le E \le u_n,$$

thus

$$k \log 2 \ll n < x,$$

and hence,

$$k \ll x. \tag{23}$$

In fact, using the fact that E is square-free, one can even infer that $k \ll x/(\log x)$ holds, but inequality (23) suffices for our purposes. Now write $F := AD$, therefore $u_n = EF$, with E and F coprime. Hence,

$$\frac{\phi(u_n)}{u_n} = \frac{\phi(F)}{F} \cdot \frac{\phi(E)}{E},$$

or

$$1-\frac{\phi(E)}{E}=1-\frac{F}{\phi(F)}\cdot\frac{\phi(u_n)}{u_n}=\frac{u_n-\frac{F}{\phi(F)}\phi(u_n)}{u_n}=\frac{(a\alpha^n-\frac{F}{\phi(F)}\cdot\phi(u_n))+b\beta^n}{a\alpha^n+b\beta^n}. \tag{24}$$

To get an upper bound on q, we use a lower bound depending on q on the left-hand side of (24) and an upper bound depending only on x (and $f(x)$) on the right-hand side of (24). For the left-hand side of (24) we write

$$1-\frac{\phi(E)}{E}=1-\prod_{i=1}^{k}\Big(1-\frac{1}{q_i}\Big). \tag{25}$$

In light of the inequality

$$1-\prod_{i=1}^{k}(1-x_i)\le\sum_{i=1}^{k}x_i \tag{26}$$

which holds for all $k\ge 1$ and all real numbers $x_i\in(0,1)$ for $i=1,\dots,k$, and which can be immediately proved by induction on k, we get, from (25), (26) and (23), that

$$1-\frac{\phi(E)}{E}=1-\prod_{i=1}^{k}\Big(1-\frac{1}{q_i}\Big)\le\sum_{i=1}^{k}\frac{1}{q_i}\ll\frac{x}{q}. \tag{27}$$

We now need a lower bound for the right-hand side. We look at the expression

$$a\alpha^n-\frac{F}{\phi(F)}\phi(u_n). \tag{28}$$

Assume first that the expression appearing at (28) is zero. In this case, we get

$$a\alpha^n=\frac{F}{\phi(F)}\phi(u_n)=F\phi(E),$$

therefore $F \mid a\alpha^n$. Since $F \mid (a\alpha^n + b\beta^n)$, we read that $F \mid b\beta^n$, and since $\gcd(a\alpha^n, b\beta^n) = 1$, we get $F = 1$. Thus, equation (28) becomes

$$a\alpha^n = \phi(E),$$

and formula (24) becomes

$$1 - \frac{\phi(E)}{E} = \frac{b\beta^n}{u_n} \ll \Big(\frac{\beta}{\alpha}\Big)^n.$$

On the other hand, since

$$1 - \frac{\phi(E)}{E} = 1 - \prod_{i=1}^{k}\Big(1 - \frac{1}{q_i}\Big) \geq \frac{1}{q_1},$$

we get

$$\frac{1}{q_1} \gg \Big(\frac{\beta}{\alpha}\Big)^n,$$

therefore the inequality

$$q_1 > \exp(c_3 n)$$

holds for all n sufficiently large with the constant $c_3 := \dfrac{\log(\alpha/\beta)}{2}$. So,

$$\alpha^n \gg u_n = E \geq q_1^k > \exp(kc_1 n),$$

which implies that $k < c_4$. We now write

$$b\beta^n = E - \phi(E) = q_1 q_2 \cdots q_k - \phi(E), \tag{29}$$

or

$$b\beta^n = \prod_{i=1}^{k}((q_i - 1) + 1) - \phi(E) = \sum_{\substack{I \subset \{1,2,\ldots,k\} \\ I \neq \{1,2,\ldots,k\}}} \prod_{i \in I}(q_i - 1). \tag{30}$$

Assume that $k = 1$. In this case, equation (30) becomes $b\beta^n = 1$, which shows that $b = \beta = 1$, so that u_n is of the form $u_n = a\alpha^n + 1$, and the only difficulty in carrying out the argument will be to investigate when $a\alpha^n + 1$ is a prime. As we have mentioned in the Introduction,

this is a very hard problem, and it is known, conditionally, that indeed u_n is a prime only for a set of n of asymptotic density zero. With our hypothesis however, we get that a and α are positive integers which are multiplicatively dependent. Thus, there exists a positive integer ρ, two non-negative integers μ and ν (with at least one of them positive) which are also coprime such that both $a = \rho^\mu$ and $\alpha = \rho^\nu$ hold. Therefore, $u_n = \rho^{\mu+\nu n}+1$ holds for all positive integers n. Replacing the sequence of general term u_n by the sequence of general term $u_n = \rho^n + 1$, it follows that it suffices to show that $\rho^n + 1$ can be a prime only on a set of n of asymptotic density zero. But the only chance that $\rho^n + 1$ can be a prime for infinitely many values of n is when ρ is even and n is a power of 2, and the set of powers of 2 is obviously of asymptotic density zero. This disposes of the only hard case of our problem, and it is only here where the assumptions A2 are really needed.

Assume now that $k > 1$. Since $k < c_4$, we may assume that k is fixed. Recalling that $q_i - 1 \in \mathcal{S}$, equation (30) is a particular case of an equation of the type

$$X = \sum_{\substack{I \subset \{1,2,\dots,k\} \\ I \neq \{1,2,\dots,k\}}} X_I$$

in $2^k \geq 4$ indeterminates $X := b\beta^n$ and $X_I := \prod_{i \in I}(q_i - 1)$ where I is a proper subset of $\{1, 2, \dots, k\}$ (including the subset $I = \varnothing$ for which $X_\varnothing := 1$). Since $X_I > 0$ for all I, it follows that the above equation is *non-degenerate* in the sense that no proper subsum of the form $X_{I_1} + \cdots + X_{I_j}$ vanishes. We now use a result of Schlickewei (see [**6**]) on the number of non-degenerate solutions of $\mathcal{S}$-equations. That is, if $\gamma_1, \dots, \gamma_m$ are fixed non-zero rational numbers, then there exists a number of at most

$$l \leq \exp\Big(2^{37m} t^6 \log(8t)\Big) \tag{31}$$

non-degenerate solutions $(X_1^{(j)}, \dots, X_n^{(j)})$ with $j = 1,\ 2, \dots, l$ for the equation

$$\sum_{i=1}^{m} \gamma_i X_i = 0, \quad \text{with } X_i \in \mathcal{S} \quad \text{for } i = 1, \dots, m, \tag{32}$$

so that for any other non-degenerate solution $(X_1, \dots, X_m)$ of equation (32) there exists a number $\rho \in \mathcal{S}$ and a number $s \le l$ such that $(X_1, \dots, X_m) = \rho(X_1^{(s)}, \dots, X_n^{(s)})$. Let us notice that from the above result it follows that equation (30) can have at most l solutions n, where l is bounded above as in (31) with $m = 2^k \le 2^{c_4}$. Indeed, we can label the indeterminates X_I for $I \subset \{1, \dots, k\}$ such that $1 = X_\varnothing = X_1$, such that if equation (30) has more than l solutions, then there must exist two solutions $(X_1, \dots, X_{2^k})$ and $(X_1', \dots, X_{2^k}')$ and a rational number $\rho \ne 1$ composed only from the primes $p_1, \dots, p_t$ such that $(X_1', \dots, X_{2^k}') = \rho(X_1, \dots, X_{2^k})$. In particular, $1 = X_1' = \rho X_1 = \rho$ forcing $\rho = 1$, which is a contradiction. This is for a fixed value of k, and now letting k run from 2 to c_4, we get that the number of solutions of (30) with $n < x$ is at most

$$c_4 \exp\Big(c_5 t^6 \log(8t)\Big) \tag{33}$$

with $c_5 := 2^{37 \cdot 2^{c_4}}$, and it is enough for our purposes to check that the number appearing at (33) is smaller than $x/(\log x)$. But this will be so provided that the inequality

$$c_5 t^6 \log(8t) < \log x - \log_2 x - \log c_4$$

holds, which will hold provided that the inequality

$$t^6 \log(8t) < c_6 \log x \tag{34}$$

holds, where one can take $c_6 := 1/(2c_5)$. The above inequality (34) is fulfilled provided that the inequality

$$t < c_7 \Big(\frac{\log x}{\log_2 x}\Big)^{1/6} \tag{35}$$

holds with some appropriate constant c_7, and in order for (35) to hold for large x it suffices that the inequality

$$\frac{f(x)}{\log f(x)} < c_8 \Big(\frac{\log x}{\log_2 x}\Big)^{1/6}, \tag{36}$$

is satisfied with, say $c_8 := c_7/2$. Clearly, inequality (36) holds provided that one chooses

$$f(x) := c_9 (\log x)^{1/6} (\log_2 x)^{5/6}, \tag{37}$$

with some appropriate constant c_9, as stated in the conclusion of the Theorem.

From now on, we may therefore assume that the expression (28) is non-zero. In this case, we can find a lower bound on the expression appearing at (28) by using a linear form in logarithms (see [**1**]). That is, write

$$\phi(u_n) := p_1^{\beta_1} \cdot \cdots \cdot p_t^{\beta_t}$$

and

$$\Big|a\alpha^n - \frac{F}{\phi(F)}\phi(u_n)\Big| = a\alpha^n\Big|1 - \Big(\frac{F}{a\phi(F)}\Big) \cdot p_1^{\beta_1} \cdot \cdots \cdot p_t^{\beta_t} \cdot \alpha^{-n}\Big|. \tag{38}$$

Since

$$\phi(u_n) < u_n \ll \alpha^n,$$

we get that

$$\max_{i=1}^{t}(\beta_i) \ll n. \tag{39}$$

For any rational number w let $H(w)$ be the maximum of the absolute values of its numerator and denumerator when written in reduced form. Since a is a constant, and $F/\phi(F)$ is a rational number which written in reduced form has its numerator greater than its denominator and the numerator is square-free and composed of primes less than x^{1/t^2}, we get that

$$\log\Big(H\Big(\frac{F}{a\phi(F)}\Big)\Big) \ll \sum_{p<x^{1/t^2}} \log p \ll x^{1/t^2}. \tag{40}$$

Let

$$\Omega := \prod_{i=1}^{t} \log p_i \tag{41}$$

and notice the following upper bound

$$\Omega \le (\log f(x))^t = \exp(t\log_2(f(x)) < \exp(2t\log_2 t), \tag{42}$$

which is valid for large values of x.

With the estimates (40), (42), and a classical lower bound for linear forms in complex logarithms (as in [**1**]), we deduce the existence of a constant $c_{10} > 1$ such that the inequality

$$\begin{aligned} &\Big|1-\Big(\frac{F}{a\phi(F)}\Big)\cdot p_1^{\beta_1}\cdot\cdots\cdot p_t^{\beta_t}\cdot\alpha^{-n}\Big| \\ &\qquad > \exp\Big(-t^{c_{10}t}\log\Big(H\Big(\frac{F}{a\phi(F)}\Big)\Big)\cdot\Omega\cdot\log n\Big) \\ &\qquad > \exp\Big(-x^{1/t^2}\log(n)\exp\Big(c_{10}t\log t+2t\log_2 t\Big)\Big) \end{aligned} \tag{43}$$

holds for large enough values of x. Let us observe that

$$\exp\Big(c_{10}t\log t+2t\log_2 t\Big) < x^{1/t^2} \tag{44}$$

holds for large enough values of x. Indeed, inequality (44) is implied by

$$c_{10}t^2(t\log t+2t\log_2 t) < \log x,$$

which obviously holds for large values of x because $t=\pi(f(x)) < f(x)$ and $f(x)$ is given by (37). Thus, with (43) and (44), we get that

$$\Big|1-\Big(\frac{F}{a\phi(F)}\Big)\cdot p_1^{\beta_1}\cdot\cdots\cdot p_t^{\beta_t}\cdot\alpha^{-n}\Big| > \exp(-x^{2/t^2}\log n). \tag{45}$$

We now show that the expression appearing at (28) is positive. Indeed, assuming that the expression appearing at (28) is negative, then from formula (24) and the fact that the left-hand side of (24) is positive, we get that $b\beta^n$ is positive and that

$$b\beta^n > \Big|a\alpha^n-\frac{F}{\phi(F)}\phi(u_n)\Big| > a\alpha^n\cdot\exp(-x^{2/t^2}\log n). \tag{46}$$

The above inequality implies, after rearranging it, taking logarithms, and recalling that $\alpha > |\beta|$, that

$$x^{2/t^2}\log x > x^{2/t^2}\log n \gg n \ge \frac{x}{\log x},$$

which is impossible for large enough values of x. This shows that for large enough values of x the expression appearing at (28) is positive. In particular, the numerator of the expression appearing in the right-hand side of (28) is at least as large as

$$a\alpha^n \cdot \exp(-x^{2/t^2}\log n) - |b\beta^n| > \frac{a}{2}\alpha^n \exp(-x^{2/t^2}\log n),$$

for large enough values of x. Since $u_n \ll \alpha^n$, it follows that

$$\frac{(a\alpha^n - \frac{F}{\phi(F)}\cdot\phi(u_n)) + b\beta^n}{a\alpha^n + b\beta^n} \gg \exp(-x^{2/t^2}\log n). \tag{47}$$

Combining estimate (47) with (27), we get

$$q = q_1 \ll x\exp(x^{2/t^2}\log x), \tag{48}$$

so that

$$q = q_1 < \exp(x^{3/t^2}) \tag{49}$$

holds for large enough values of x. Since $q-1 \in \mathcal{S}$, the number of numbers q that can fulfill (49) is certainly no more than

$$O(x^{3/t}) = o(x).$$

We now return to the values of n. From what we have said, but for $o(x)$ positive integers n in the interval $\big(x/(\log x),\, x\big)$ for which $\phi(u_n) \in \mathcal{S}$, a prime number $q := q(n)$ exists such that $q > f(x)$, $d(n) := d(q(n)) > x^{1/t^2}$, and q is minimal with this property. Moreover, this number q satisfies inequality (49) and the number of such numbers q is $o(x)$. Fix such a number q. Since $q \mid u_n$, this means that n is in the arithmetical progression $r(q) \pmod{d(q)}$. The number of such numbers $n < x$ is certainly at most

$$\Big\lfloor \frac{x}{d(q)} \Big\rfloor + 1 \le \begin{cases} \dfrac{2x}{d(q)}, & \text{if } d(q) < x, \\ 1, & \text{if } d(q) > x. \end{cases} \tag{50}$$

So, the total contributions when $d(q) > x$ are at most the number of such numbers q which, as we have seen, is $o(x)$. Thus, it remains to find an upper bound for

$$x \sum_{\substack{q-1 \in \mathcal{S} \\ x^{1/t^2} < d(q) < x}} \frac{1}{d(q)}. \tag{51}$$

In particular, the Theorem will be proved provided that we can show that

$$\sum_{\substack{q-1 \in \mathcal{S} \\ x^{1/t^2} < d(q) < x}} \frac{1}{d(q)} = o(1). \tag{52}$$

Set

$$\mathcal{D} := \{d \mid d = d(q) \text{ for some } q \text{ with } q - 1 \in \mathcal{S}\}. \tag{53}$$

In order to prove (52), we first need to understand an upper bound for the *multiplicity* of an element $d \in \mathcal{D}$. That is, given $d \in \mathcal{D}$, how many primes q with $q - 1 \in \mathcal{S}$ are there such that $d = d(q)$? Denote this number by $T(d)$. We shall show that

$$T(d) \le (3t)! d^{1 - \frac{1}{t+1}} \tag{54}$$

holds for large enough values of x and uniformly in d. Assume for the moment that we have proved inequality (54). Then we can bound the expression appearing in the left-hand side of (52) by saying that

$$\sum_{\substack{q-1 \in \mathcal{S} \\ x^{1/t^2} < d(q) < x}} \frac{1}{d(q)} \le (3t)! \sum_{\substack{d \in \mathcal{D} \\ x^{1/t^2} < d}} \frac{1}{d^{\frac{1}{t+1}}}. \tag{55}$$

Finally, let us notice that from the way the number $d := d(q)$ was defined, we get that for every prime number q which divides u_n for some n, q is a *primitive divisor* of $L_{d(q)}$. That is, $q \mid L_{d(q)}$ but q does not divide L_m for any positive integer $m < q$. From the well-known properties of the primitive divisors, we get that $d(q) \mid q - 1$. So, in

particular, when $q-1 \in \mathcal{S}$, we get that $d(q) \in \mathcal{S}$. In particular, we have

$$\sum_{\substack{q-1\in\mathcal{S}\\ x^{1/t^2}<d(q)<x}} \frac{1}{d(q)} \le \sum_{\substack{d\in\mathcal{S}\\ x^{1/t^2}<d}} \frac{1}{d^{\frac{1}{t+1}}}. \tag{56}$$

It remains to show that

$$\sum_{\substack{d\in\mathcal{S}\\ x^{1/t^2}<d}} \frac{1}{d^{\frac{1}{t+1}}} = o\Big(\frac{1}{(3t)!}\Big). \tag{57}$$

But obviously, the series

$$\sum_{d\in\mathcal{S}} \frac{1}{d^{\frac{1}{t+1}}} \tag{58}$$

is convergent, and the sum of the above series is precisely

$$g(t) := \prod_{i=1}^{t}\Big(1-\frac{1}{p_i^{\frac{1}{t+1}}}\Big)^{-1}. \tag{59}$$

We now find an upper bound on $g(t)$. Notice that with a fixed prime number p, we have

$$\Big(1-\frac{1}{p^{\frac{1}{t+1}}}\Big)^{-1} = \sum_{i\ge 0}\frac{1}{p^{\frac{i}{t+1}}} = \Big(\sum_{i=0}^{t}\frac{1}{p^{\frac{i}{t+1}}}\Big)\cdot\Big(\sum_{j\ge 0}\frac{1}{p^j}\Big) < (t+1)\Big(1-\frac{1}{p}\Big)^{-1},$$

so that

$$\begin{aligned} g(t) &< (t+1)^t \prod_{i=1}^{t}\Big(1-\frac{1}{p_i}\Big)^{-1} \\ &< \exp(t\log(t+1)+c_{11}\log_2 t) < \exp(2t\log t) \end{aligned} \tag{60}$$

holds for large enough values of x. To estimate the tail of the series (58) appearing in the left-hand side of formula (57), let $d \in \mathcal{S}$ be such that $d > x^{1/t^2}$, and assume that δ denotes the maximum of the

exponents at which the prime numbers dividing d can appear in the prime factorization of d. Then obviously

$$t \log f(x) \delta \geq \log d \geq \frac{\log x}{t^2},$$

so that

$$\delta \geq \frac{\log x}{t^3 \log f(x)} \geq \frac{\log x}{2t^3 \log t} \tag{61}$$

holds for all large enough values of x. By separating the prime power of maximal exponent δ from d, and then summing up over all the primes $p \in \mathcal{S}$ and over all the powers δ which are at least as large as shown in (61), we get that the sum appearing in the left hand side of (57) is bounded above by

$$\begin{aligned}
&\sum_{i=1}^{t} \sum_{j \geq \frac{\log x}{2t^3 \log t}} \frac{1}{p^{\frac{j}{t+1}}} \sum_{d \in S} \frac{1}{d^{\frac{1}{t+1}}} \\
&\quad \leq g(t) \sum_{i=1}^{t} \sum_{j \geq \frac{\log x}{2t^3 \log t}} \frac{1}{p^{\frac{j}{t+1}}} \\
&\quad \leq g(t) \sum_{i=1}^{t} \exp\Big(-\frac{\log x \log p_i}{2t^3(t+1)\log t}\Big)\Big(1 - \frac{1}{p_i^{\frac{1}{t+1}}}\Big)^{-1} \qquad (62) \\
&\quad \ll g(t) t \sum_{i=1}^{t} \exp\Big(-\frac{\log x \log p_i}{2t^3(t+1)\log t}\Big) \\
&\quad \leq g(t) t^2 \exp\Big(-\frac{\log x \log 2}{2t^3(t+1)\log t}\Big) \\
&\quad < \exp\Big(2t \log t + 2 \log t - \frac{\log x}{2t^3(t+1)\log t}\Big).
\end{aligned}$$

Since $1/(3t)^{3t} = o(1/(3t)!)$, it follows, with (62), that in order for (57) to hold it suffices that

$$\exp\Big(2t \log t + 2 \log t - \frac{\log x}{2t^3(t+1)\log t}\Big) < \frac{1}{(3t)^{3t}},$$

which is implied by

$$6t\log(3t) < \frac{\log x}{2t^3(t+1)\log t},$$

which is fulfilled provided that

$$12t^4(t+1)\log^2(3t) < \log x, \tag{63}$$

and (63) obviously holds for large values of x because $t = \pi(f(x)) < f(x)$ and $f(x)$ is given by formula (37).

Thus, the Theorem is proved once we are able to show that inequality (54) holds. To prove (54), assume that t is large, pick a number $d \in \mathcal{S}$, set $T := T(d)$, and write

$$\alpha^d \gg L_d = \prod_{i=1}^{T} q_i,$$

where $q_1 < q_2 < \cdots < q_T$ are distinct primes with $q_i - 1 \in \mathcal{S}$. Then certainly

$$\alpha^d \gg \prod_{i=1}^{T}(q_i - 1). \tag{64}$$

To get a large T, we have to assume that all the q_i's are as small as possible. But how small can we make the product on the left? Well, discarding the fact that q_i have to be primes, we will certainly want to first put $q - 1 = p_i$ for $i = 1, 2, \ldots, t$, then for the next numbers we will want to put $q - 1 = p_i p_j$ for $1 \le i \le j \le t$ and so on. The above argument shows that in order to get an upper bound on T, we should write

$$T = \binom{t}{t-1} + \binom{t+1}{t-1} + \cdots + \binom{t+u}{t-1} + N,$$

where u is the unique positive integer such that $0 \le N < \binom{t+u+1}{t-1}$ holds, and by (64) the maximal value of T will certainly be bounded above by those u and N for which the inequality

$$\binom{t}{1} + 2\binom{t+1}{2} + \cdots + (u+1)\binom{t+u}{t-1} + (u+2)N \le d \tag{65}$$

holds. Inequality (65) together with the obvious lower bound

$$\binom{t+i}{t-1} \geq \frac{(i+1)^{t-1}}{(t-1)!}$$

uniformly in i, shows that

$$\sum_{i=1}^{u+1} i^t \leq (t-1)!d,$$

and since

$$\sum_{i=1}^{u+1} i^t \gg \frac{u^{t+1}}{t+1}$$

holds uniformly in i and u, we get

$$u \ll (t+1)!^{\frac{1}{t+1}} d^{\frac{1}{t+1}}.$$

In particular, using now the fact that

$$\binom{t+i}{t-1} \leq (i+1)^{t-1}$$

holds uniformly in i, we get

$$\begin{aligned} T &\leq \sum_{i=0}^{u+1} \binom{t+i}{t-1} \leq \sum_{i=1}^{u+2} i^{t-1} \leq t(u+2)^{t-1} \\ &\ll t \cdot (t+1)!^{\frac{t-1}{t+1}} \cdot d^{\frac{t}{t+1}} \cdot \left(1+\frac{2}{u}\right)^{t-1} < (3t)! \cdot d^{1-\frac{1}{t+1}}, \end{aligned} \tag{66}$$

where in the last step of (66) we used the fact that

$$t((t+1)!)^{\frac{t-1}{t+1}} \cdot \left(1+\frac{2}{u}\right)^{t-1} \leq t \cdot ((t+1)!)^{\frac{t-1}{t+1}} \cdot 3^{t-1} = o((3t)!),$$

which holds for large t (for example, by Stirling's formula).

The Theorem is therefore proved.

Acknowledgment. I would like to thank the referee for comments and suggestions which improved the presentation of this paper.

REFERENCES

1. A. Baker, G. Wüstholz, *Logarithmic forms and group varieties*, J. Reine Angew. Math. **442** (1993), 19–62.

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

3. N. de Bruijn, *On the number of positive integers* $\leq x$ *and free of prime factors* $> y$, Nederl. Akad. Wetensch. Proc. Ser. A **54** (1951), 50–60.

4. C. Hooley, *Applications of sieve methods to the theory of numbers*, Cambridge Tracts in Math. **70**, Cambridge Univ. Press, Cambridge, 1976.

5. M. Křížek, F. Luca, L. Somer, *17 lectures on Fermat numbers. From number theory to geometry*, CMS Books in Mathematics/Ouvrages de Mathématiques de la SMC **9**, Springer-Verlag, Berlin, 2001.

6. H.P. Schlickewei, *S-unit equations over number fields*, Invent. Math. **102** (1990), 95–107.

7. T.N. Shorey, R. Tijdeman, *Exponential Diophantine equations*, Cambridge Tracts in Math. **87**, Cambridge Univ. Press, Cambridge, 1986.

8. C.L. Stewart, *On divisors of terms of linear recurrence sequences*, J. Reine Angew. Math. **333** (1982), 12–31.

9. C.L. Stewart, *On the greatest prime factor of terms of a linear recurrence sequence*, Number theory (Winnipeg, Man., 1983), Rocky Mountain J. Math. **15** (1985), 599–608.

10. K. Yu, *p-adic logarithmic forms and group varieties. II,* Acta Arith. **89** (1999), 337–378.

IMATE, UNAM, Ap. Postal 61-3 (Xangari), CP 58 089, Morelia, Michoacán, México
E-mail address: fluca@matmor.unam.mx