Max-Planck-Institut für Mathematik
Bonn

Sequences of integers generated by two fixed primes

by

Alessandro Languasco
Florian Luca
Pieter Moree
Alain Togbé

Max-Planck-Institut für Mathematik
Preprint Series 2023 (34)

Date of submission: December 21, 2023

Sequences of integers generated by two fixed primes

by

Alessandro Languasco
Florian Luca
Pieter Moree
Alain Togbé

Max-Planck-Institut für Mathematik
Vivatsgasse 7
53111 Bonn
Germany

Dipartimento di Matematica
"Tullio Levi-Civita"
Università di Padova
Via Trieste 63
35121 Padova
Italy

School of Mathematics
University of the Witwatersrand
P.O. Box Wits 2050
South Africa

Department of Mathematics and Statistics
Purdue University Northwest
2200 169th Street
Hammond, IN 46323
USA

MPIM 23-34

SEQUENCES OF INTEGERS GENERATED BY TWO FIXED PRIMES

ALESSANDRO LANGUASCO, FLORIAN LUCA, PIETER MOREE, AND ALAIN TOGBÉ

In honorem Robert Tijdeman annos LXXX nati

Abstract. Let p and q be two distinct fixed prime numbers and $(n_i)_{i\geq 0}$ the sequence of consecutive integers of the form $p^a \cdot q^b$ with $a, b \geq 0$. Tijdeman gave a lower bound (1973) and an upper bound (1974) for the gap size $n_{i+1} - n_i$, with each bound containing an unspecified exponent and implicit constant. We will explicitly bound these four quantities. Earlier Langevin (1976) gave weaker estimates for (only) the exponents.

Given a real number $\alpha > 1$, there exists a smallest number m such that for every $n \geq m$, there exists an integer n_i in $[n, n\alpha)$. Our effective version of Tijdeman's result immediately implies an upper bound for m, which using the Koksma-Erdős-Turan inequality we will improve on. We present a fast algorithm to determine m when $\max\{p, q\}$ is not too large and demonstrate it with numerical material. In an appendix we explain, given n_i, how to efficiently determine both n_{i-1} and n_{i+1}, something closely related to work of Bérczes, Dujella and Hajdu.

1. Introduction

Given a set $S = \{p_1, \ldots, p_s\}$ of primes, the numbers $p_1^{e_1} \cdots p_s^{e_s}$ with non-negative exponents are called *S-units.* There are many Diophantine equations involving S-units, see for example Evertse et al. [5]. In this note we are interested in the distribution of S-units.

In case $p_1, \ldots, p_s$ are the first s prime numbers, the counting function of S-units up to x is denoted by $\psi(x, p_s)$, and the S-units are called p_s–friable or p_s–smooth. There is an extensive literature on these numbers, see Hildebrand and Tenenbaum [10] for a nice survey. The gaps between consecutive friable numbers were studied by Heath-Brown [9]. Tijdeman and Meijer [20] studied the distribution of the greatest common divisor of consecutive S-units (for general S). We will concentrate solely on the case where S has two elements.

Definition 1.1. Let p, q be two primes with $p < q$. We let $(n_i)_{i\geq 0}$ be the sequence of consecutive integers of the form $n = p^a \cdot q^b$ with $a, b \geq 0$.

Put $N_{p,q}(x) := \#\{n_i \leq x\}$. Note that $N_{p,q}(x) = \#\{(e, f) \in \mathbb{Z}^2_{\geq 0} : e \log p + f \log q \leq \log x\}$, that is it equals the number of lattice points inside a rectangular triangle with sides of length $\log x/ \log p$, respectively $\log x/ \log q$. The area of this triangle is

$$\frac{(\log x)^2}{2 \log p \cdot \log q}. \tag{1.1}$$

It is known that

$$N_{p,q}(x) = \frac{\log px \cdot \log qx}{2 \log p \cdot \log q} + o(\log x),$$

where the error term is worse if $\log px \cdot \log qx$ is replaced by $(\log x)^2$. Interestingly, Ramanujan in his famous first letter to Hardy (Jan. 16th, 1913), claims (in case $p = 2$ and $q = 3$) the latter main

Date: February 5, 2024.

2010 *Mathematics Subject Classification.* 11B83, 11J70, 11J86, 11N25.

Key words and phrases. Linear forms in logarithms, S-units, Bertrand's Postulate, Continued fractions, Discriminator.

term as an approximation, rather than the trivial (1.1). For more details and references the reader is referred to Moree [14].

Bérczes, Dujella, and Hajdu [2], for any term n_i determined n_{i+1}, at least in principle, without enumerating all terms of the sequence, and they gave an efficient algorithm to find n_{i+1} explicitly. They do so by analyzing the behavior of the continued fractions of $\log p/\log q$ (see Theorems 2.2 and 2.3 of [2]). In the appendix we present a shorter reproof and in addition show how to also efficiently find n_{i-1} explicitly.

As early as 1908, Thue [17] gave an ineffective proof that the gap size $n_{i+1}-n_i$ tends to infinity, which was made effective by Cassels [4] in 1960. Tijdeman [18, 19], half a century ago, gave bounds for the gap size. He derived these by making use of estimates of Baker [1] for linear forms in logarithms of algebraic numbers, which had become available a few years earlier.

Theorem 1.1. *Let $(n_i)_{i\ge 0}$ be as in Definition 1.1. There exist effective constants C_1 and C_2 such that*
$$\frac{n_i}{(\log n_i)^{C_1}} \ll_{p,q} n_{i+1}-n_i \ll_{p,q} \frac{n_i}{(\log n_i)^{C_2}}.$$
The constants C_1, C_2 and the two implicit constants all may depend on p and q.

Langevin [12], soon after Tijdeman's work appeared, gave effective bounds for the constants C_1 and C_2.

Theorem 1.2. *Under the conditions of Theorem 1.1, we have*
$$(2^{126}\log p\cdot\log q\cdot\log\log p)^{-1} < C_2 < C_1 < 2^{126}\log p\cdot\log q\cdot\log\log p.$$

We will improve on this result and also give explicit bounds for the implicit constants in Theorem 1.1.

Theorem 1.3. *Assuming $n_i\ge 3$, we have*
$$C_3\frac{n_i}{(\log n_i)^{C_1}} < n_{i+1}-n_i < C_4\frac{n_i}{(\log n_i)^{C_2}},$$
where $C_1 = 2\cdot 10^9\log p\cdot\log q$, $C_2 = C_1^{-1}$, $C_3 = (\log p)^{C_1}$, $C_4 = 8q$.

Langevin used a result of van der Poorten [21] to obtain Theorem 1.2. In Section 2 we will use a celebrated result of Matveev (Theorem 2.1) to prove Theorem 1.3. This helps to decrease the coefficient of the bound C_2 from a 38 digit number to a 9 digit one.

Very recently Stewart [16] derived an analogue of Theorem 1.1 for increasing sequences (n_i) such that the largest prime factor of n_i is at most $y(n_i)$, with $y(x)$ a non-decreasing function slowly tending to infinity.

In Section 3, we consider Bertrand's Postulate type results for our sequence.

Definition 1.2. Given any real number $\alpha > 1$, $n_{p,q}(\alpha)$ is the smallest integer such that every interval $[n, n\alpha)$ with $n \ge n_{p,q}(\alpha)$ contains an integer of the form $p^a\cdot q^b$ for every integer $n\ge n_{p,q}(\alpha)$.

Bertrand's Postulate type results arise on taking $\alpha = 2$. Theorem 1.3 gives right away that if $\alpha\in(1,p]$ and

$$n_i > \exp\Big(\Big(\frac{4q}{\alpha-1}\Big)^{2\cdot 10^9\log p\cdot\log q}\Big), \tag{1.2}$$

then $n_{i+1}\in(n_i,\alpha n_i)$. In Section 3 we will (slightly) improve on this lower bound and show that this conclusion already holds if

$$n_i > \exp\Big(2\log p\Big(\frac{60\log q}{\log\alpha}\Big)^{10^9\log p\cdot\log q}\Big).$$

In Section 4 we present an algorithm for efficiently determining $n_{p,q}(\alpha)$ and present some of its outputs. In particular, in Section 3.3 we use our algorithm to advance the understanding of the so-called *discriminator* of an infinite family of second-order recurrent sequences first studied by Faye, Luca and Moree [6] and more recently by Ferrari, Luca and Moree [7].

2. Proof of Theorem 1.3

We need linear forms in logarithms. For any non-zero algebraic number η of degree d over $\mathbb{Q}$, whose minimal polynomial over $\mathbb{Z}$ is $a\prod_{i=1}^{d}\bigl(X-\eta^{(i)}\bigr)$ (with $a>0$), we denote by

$$h(\eta)=\frac{1}{d}\Bigl(\log a+\sum_{i=1}^{d}\log\max\bigl(1,|\eta^{(i)}|\bigr)\Bigr)$$

the usual *absolute logarithmic height* of η. If η_1 and η_2 are algebraic numbers, then we have the basic properties

$$\begin{aligned}h(\eta_1\pm\eta_2)&\le h(\eta_1)+h(\eta_2)+\log 2,\\ h(\eta_1\eta_2^{\pm})&\le h(\eta_1)+h(\eta_2),\\ h(\eta_1^{j})&=|j|h(\eta_1),\end{aligned}$$

where j is any integer.

We recall Matveev's main theorem [13] in a version due to Bugeaud, Mignotte and Siksek [3, Thm. 9.4]. It applies to algebraic numbers, but we recall it here only for rational numbers. For a rational number $\gamma=r/s$ with coprime integers r and $s>0$, let $h(\gamma):=\max\{\log|r|,\log s\}$ be its *naive height.*

Theorem 2.1 (Matveev's theorem). *Let $\gamma_1,\ldots,\gamma_k$ be positive rational numbers, let $b_1,\ldots,b_k$ be non-zero integers, and assume that*

$$\Lambda:=\gamma_1^{b_1}\cdots\gamma_k^{b_k}-1, \tag{2.1}$$

is non-zero. For every real number $B\ge\max\{|b_1|,\ldots,|b_k|\}$ we have

$$\log|\Lambda|>-1.4\cdot 30^{k+3}\cdot k^{4.5}(1+\log B)\,h(\gamma_1)\cdots h(\gamma_k).$$

Note that $n_1=1,\ n_2=p$. Let $n_i=p^u\cdot q^v$. We assume that $n_i\ge 3$ (this holds for $i=2$ in all cases except if $p=2$, in which case we assume that $i\ge 3$). The lower bound on $n_{i+1}-n_i$ follows right–away from Theorem 2.1. Indeed, let $n_{i+1}=p^{u'}\cdot q^{v'}$. Note that $\max\{u,v,u',v'\}\le 2\max\{u,v\}$ (since certainly $n_i^2>n_i$ is a $\{p,q\}$-unit, clearly $n_i^2\ge n_{i+1}$). Then

$$n_{i+1}-n_i=n_i\bigl(p^{u'-u}q^{v'-v}-1\bigr).$$

By Theorem 2.1 the right–hand side is bounded below by

$$n_i\,\exp\bigl(-1.4\cdot 30^5\cdot 2^{4.5}(1+\log(2\max\{u,v\}))\log p\cdot\log q\bigr).$$

Since $\max\{u,v\}\ge 1$ and for $x\ge 2$ we have $1+\log x\le 2.5\log x$, we obtain

$$n_{i+1}-n_i>\frac{n_i}{(2\max\{u,v\})^{2\cdot 10^9\log p\cdot\log q}}.$$

Since $n_i=p^u\cdot q^v\le(pq)^{\max\{u,v\}}$, we get that $\max\{u,v\}\ge\log n_i/\log(pq)$. Thus,

$$n_{i+1}-n_i>C_3\frac{n_i}{(\log n_i)^{C_1}},$$

where

$$C_1=2\cdot 10^9\log p\cdot\log q,\qquad C_3=(0.5\log pq)^{C_1}>(\log p)^{C_1}.$$

In order to obtain an upper bound for $n_{i+1} - n_i$ we proceed as in Langevin [12] and make use of the sequence $(r_k/s_k)_{k\geq 0}$ of *convergents* of

$$\theta_{p,q} := \frac{\log p}{\log q}.$$

Suppose first that $p^u > q^v$. We assume that ℓ is the index verifying $s_\ell \leq u < s_{\ell+1}$. Since $p < q$, it follows that $a_0 = 0$ and $a_1 = \lfloor \log q / \log p \rfloor$. So, if $\ell = 0$, then we have $u < \log q / \log p$, so $p^u < q$. Thus, $v = 0$, $n_{i+1} \leq q$ and the inequalities hold provided that the multiplicative constant C_4 implied by the right Vinogradov symbol $\ll$ is taken to be at least $q(\log q)^{C_2}$.

Assume next that $\ell \geq 1$. One of the rational numbers r_ℓ/s_ℓ and $r_{\ell+1}/s_{\ell+1}$ is at least $\theta_{p,q}$. Choose $\ell' \in \{\ell, \ell+1\}$ such that $r_{\ell'}/s_{\ell'} > \theta_{p,q}$. By construction

$$p^{u-s_{\ell'}} \cdot q^{v+r_{\ell'}}$$

is integer larger than $p^u \cdot q^v$, and so it is at least n_{i+1}. Thus, we obtain

$$\log\Big(\frac{n_{i+1}}{n_i}\Big) \leq |\Lambda_{\ell'}|, \tag{2.2}$$

where

$$\Lambda_k := s_k \log p - r_k \log q.$$

By [8, Thm. 171] we have

$$|\Lambda_k| < \frac{\log q}{s_{k+1}}. \tag{2.3}$$

Combining (2.2) and (2.3) with $k = \ell' = \ell$ gives

$$\log\Big(\frac{n_{i+1}}{n_i}\Big) < \frac{\log q}{s_{\ell+1}}. \tag{2.4}$$

Next, assume $k = \ell' = \ell + 1$. Now (2.3) gives

$$\frac{s_{\ell+2}}{\log q} < \frac{1}{|\Lambda|}, \tag{2.5}$$

where for notational convenience we put $\Lambda = \Lambda_{\ell+1}$. We need to lower bound $|\Lambda|$. Note that $|p^{s_{\ell+1}} \cdot q^{-r_{\ell+1}} - 1| = |\exp(\Lambda) - 1|$. Either $|\Lambda| \geq 1/2$, or $|\Lambda| < 1/2$, in which case

$$|\exp(\Lambda) - 1| < 2|\Lambda|.$$

We lower bound the left–hand side above using Theorem 2.1. We get

$$2|\Lambda| > |p^{s_{\ell+1}} \cdot p^{-r_{\ell+1}} - 1| > \exp\big(-1.4 \cdot 30^5 \cdot 2^{4.5}\,(1 + \log\max\{s_{\ell+1}, r_{\ell+1}\})\,\log p \cdot \log q\big).$$

Note that we can assume that $s_{\ell+1} > r_{\ell+1}$. Indeed, otherwise we have $r_{\ell+1} \geq s_{\ell+1}$ and so

$$|\Lambda| = |(r_{\ell+1} - s_{\ell+1})\log q + s_{\ell+1}(\log q - \log p)| \geq \log\big(\frac{q}{p}\big) \geq \log\big(1 + \frac{1}{p}\big) > \frac{1}{2p},$$

which is a better inequality. In particular, $s_{\ell+1} \geq 2$. We thus get

$$\begin{aligned} |\Lambda| &> \exp\big(-\log 2 - 1.4 \cdot 30^5 \cdot 2^{4.5}\,(1 + \log s_{\ell+1})\log p \cdot \log q\big) \\ &> \exp\big(-2 \cdot 10^9\, \log s_{\ell+1}\, \log p \cdot \log q\big), \end{aligned}$$

where we have used the fact that $1 + \log s_{\ell+1} \leq 2.5 \log s_{\ell+1}$, for $s_{\ell+1} \geq 2$. Thus, inequality (2.5) gives

$$\frac{s_{\ell+2}}{\log q} < \frac{1}{|\Lambda|} < s_{\ell+1}^{C_1}.$$

Hence, we get

$$s_{\ell+1} > 0.5\, s_{\ell+2}^{C_2}.$$

Then, recalling (2.2)-(2.3), we obtain

$$\log\Big(\frac{n_{i+1}}{n_i}\Big) < \frac{\log q}{s_{\ell+2}} \le \frac{\log q}{s_{\ell+1}} < \frac{2\log q}{s_{\ell+2}^{C_2}}. \tag{2.6}$$

We thus get that in all cases ((2.4) and (2.6)) we have

$$\log\Big(\frac{n_{i+1}}{n_i}\Big) < \frac{\log q}{s_{\ell+1}} < \frac{2\log q}{s_{\ell+2}^{C_2}} < \frac{2\log q}{u^{C_2}}. \tag{2.7}$$

Since by assumption $p^u > q^v$ we have $p^u > n_i^{1/2}$ and hence $u > \frac{\log n_i}{2\log p}$ we obtain

$$\log\Big(\frac{n_{i+1}}{n_i}\Big) < \frac{2\,(2\log p)^{C_2}\,\log q}{(\log n_i)^{C_2}} < \frac{4\log q}{(\log n_i)^{C_2}}.$$

But $(n_{i+1}-n_i)/n_i \in (0, p-1]$. In this interval, the image of the function $\log(1+x)/x$ is in the interval $[\frac{\log p}{p-1}, 1]$. Writing $n_{i+1}/n_i = 1+x$, where $x = (n_{i+1}-n_i)/n_i$, we get that

$$n_{i+1} - n_i < \frac{4\,(p-1)\,n_i\,\log q}{(\log n_i)^{C_2}\,\log p},$$

and we can choose C_4 to be at least

$$\frac{4(p-1)\,\log q}{\log p}.$$

But, we also needed that this constant exceeds $q(\log q)^{C_2}$. Since $(\log q)^{C_2} < 2$, it follows that if we choose $C_4 \ge 4q$, then everything works out.

We proceed by a similar argument in the remaining case $q^v > p^u$. Note that $v \ge 1$. We take ℓ such that $r_\ell \le v < r_{\ell+1}$. We choose $\ell' \in \{\ell, \ell+1\}$ such that $r_{\ell'}/s_{\ell'}$ is smaller than $\theta_{p,q}$. Note that if $\ell = 0$, we then have $r_0/s_0 = 0 < \theta_{p,q}$. So, $\ell' \ge 0$ is well-defined even if $\ell = 0$. We consider the number

$$p^{u+s_{\ell'}} \cdot q^{v-r_{\ell'}},$$

which is an integer which is a $\{p,q\}$-unit exceeding n_i, so it is at least n_{i+1}. Assuming $s_{\ell+1} \ge 2$ and reasoning as before we deduce (2.7) and so

$$\log\Big(\frac{n_{i+1}}{n_i}\Big) < \frac{2\log q}{s_{\ell+2}^{C_2}}. \tag{2.8}$$

The above inequality assumes that $s_{\ell+1} \ge 2$ which holds for all $\ell \ge 0$ except when $\ell = 0$ and $q \in (p, p^2)$. But this is impossible since we must also have $0 = r_0 \le v < r_1 = 1$, so $v = 0$, therefore $i = 1$, which is false. We need to recast the above inequality in terms of v. By (2.3) we have

$$|r_{\ell+2}\log q - s_{\ell+2}\log p| < \frac{\log q}{s_{\ell+2}}.$$

Hence, using that $s_{\ell+2} \ge s_{\ell+1} \ge 2$ we obtain

$$s_{\ell+2}\log p \ge \Big(r_{\ell+2} - \frac{1}{s_{\ell+2}}\Big)\log q \ge 0.5\, r_{\ell+2}\log q.$$

Thus, we get

$$s_{\ell+2} \ge \frac{\log q}{2\log p} r_{\ell+2}, \quad s_{\ell+2}^{C_2} \ge \Big(\frac{\log q}{2\log p}\Big)^{C_2} r_{\ell+2}^{C_2} > \frac{r_{\ell+2}^{C_2}}{2}.$$

By (2.8) we get

$$\log\Big(\frac{n_{i+1}}{n_i}\Big) < \frac{4\log q}{r_{\ell+2}^{C_2}} < \frac{4\log q}{v^{C_2}}.$$

So, we get an upper bound in the right–hand side by a factor of at most 2 larger than in the case when $p^u > q^v$. Following along we obtain that in this case

$$n_{i+1} - n_i < C_4 \frac{n_i}{(\log n_i)^{C_2}},$$

we must have $C_4 > 8(p-1)\frac{\log q}{\log p}$, and so taking $C_4 = 8q$ suffices.

3. Bertrand's Postulate for the sequence $(n_i)_{i\geq 0}$

3.1. **The statement.** We will show a way to compute $n_{p,q}(\alpha)$. Tables 1 and 2 give some examples for $p = 2$ and $\alpha = 5/3$, respectively $3/2$, some of which are relevant for an application.

Basic results from the theory of Diophantine approximation (cf. Section 4.1), ensure the existence of integers e, f, g, and h such that

$$1 < \frac{p^f}{q^e} < \alpha \quad \text{and} \quad 1 < \frac{q^h}{p^g} < \alpha.$$

We claim that $n_{p,q}(\alpha) \leq p^g \cdot q^e$. In order to see this, one can observe that any integer $n := p^\ell \cdot q^k \geq p^g \cdot q^e$ satisfies either $k \geq e$, or $\ell \geq g$. In case $k \geq e$, we note that the number $p^{\ell+f} \cdot q^{k-e}$ lies in $[n, n\alpha)$. In case $\ell \geq g$, we have $p^{\ell-g} \cdot q^{k+h} \in [n, n\alpha)$. Next one tries to find an integer $n_{\text{new}} := p^\ell \cdot q^k \in [\lfloor p^g \cdot q^e/\alpha \rfloor, p^g \cdot q^e)$, where $\lfloor x \rfloor$ denotes the integral part of x. If successful, we continue until we fail, each time considering the interval $[\lfloor n_{\text{new}}/\alpha \rfloor, n_{\text{new}})$. In Section 4.1 we present a much more refined way of determining $n_{p,q}(\alpha)$.

3.2. **An upper bound for $n_{p,q}(\alpha)$.** Recall that the *discrepancy* D_N of a sequence $(a_m)_{m=1}^N$ of real numbers (not necessarily distinct) is defined as

$$D_N = \sup_{0\leq\gamma\leq 1} \left| \frac{\#\{m \leq N \ : \ \{a_m\} < \gamma\}}{N} - \gamma \right|,$$

where $\{x\}$ denotes the fractional part of a real number x. The Koksma-Erdős-Turán inequality (see, for example, Kuipers and Niederreiter [11, Lemma 3.2]) states that

$$D_N \leq \frac{3}{H} + \frac{3}{N} \sum_{m=1}^{H} \frac{1}{m\|a_m\|}, \tag{3.1}$$

where $\|x\|$ is the distance from x to the nearest integer and $H \leq N$ is an arbitrary positive integer. In this section we will improve on the bound (1.2) by applying this inequality to upper bound the discrepancy of the sequence $(j\theta_{p,q})_{j\geq 1}$, with $\theta_{p,q} := \log p/\log q$.

Theorem 3.1. *Let $p < q$ be primes and $\alpha \in (1, p]$. Put $C_5 = 10^9 \log p \cdot \log q$. There are positive integers f and g such that*

$$\left\{ f\frac{\log q}{\log p} \right\} \in \left(0, \frac{\log \alpha}{\log q}\right), \qquad \left\{ g\frac{\log p}{\log q} \right\} \in \left(1 - \frac{\log \alpha}{\log q}, 1\right), \tag{3.2}$$

and

$$\max\{f, g\} < \left(\frac{60 \log q}{\log \alpha}\right)^{C_5}. \tag{3.3}$$

In particular,

$$n_{p,q}(\alpha) \leq q^e \cdot p^g < p^{f+g} < \exp\left(2 \log p \left(\frac{60 \log q}{\log \alpha}\right)^{C_5}\right).$$

Proof. From the definition of D_N we see that the inequality

$$\#\{m \le N \ : \ \alpha \le \{a_m\} < \beta\} \ge (\beta - \alpha)N - 2D_N N \tag{3.4}$$

holds for all $0 \le \alpha \le \beta \le 1$. We will apply this with $a_m = m\,\theta_{p,q}$ for all $m = 1, \dots, N$. Writing

$$I = \Big(0, \frac{\log \alpha}{\log q}\Big), \qquad J = \Big(1 - \frac{\log \alpha}{\log q}, 1\Big),$$

both intervals of length $\log \alpha / \log q$, it follows from (3.4) that

$$\#\{m \le N : \{a_m\} \in I\} \ge |I|N - 2D_N N = \Big(\frac{\log \alpha}{\log q}\Big)N - 2D_N N, \tag{3.5}$$

and similarly with I replaced by J. In particular, if the right-hand side is positive, then there is $u \le N$ with $\{a_u\} \in I$. We will now upper bound D_N using (3.1). To bound $\|a_m\|$, note that

$$\|a_m\| = \Big|m\frac{\log p}{\log q} - t\Big| = \frac{|\Lambda|}{\log q}, \text{ with } \Lambda := m \log p - t \log q,$$

for some integer t with $t \le m(\log p)/(\log q) + 1 < 2m$. Since $\theta_{p,q}$ is irrational we have either $|\Lambda| \ge 1/2$, or $0 < |\Lambda| < 1/2$. If $0 < |\Lambda| < 1/2$, then

$$2|\Lambda| > |\exp(\Lambda) - 1|,$$

and we can apply Matveev's theorem to get a lower bound on $|\exp(\Lambda) - 1|$ and hence on $|\Lambda|$. In both cases we take in Matveev's theorem

$$k = 2, \quad \gamma_1 = p, \quad \gamma_2 = q, \quad b_1 = m, \quad b_2 = -t,$$

and, noting that we can set $B := 2m$, we get

$$2|\Lambda| > \exp\left(-c_1(\log p)(1 + \log(2m)) \log q\right), \tag{3.6}$$

where $c_1 = 1.4 \cdot 30^5 \cdot 2^{4.5} < 8 \cdot 10^8 - \log 2$. We get

$$\begin{aligned} |\Lambda| \ &> \ \exp(-8 \cdot 10^8(1 + \log(2m)) \log p \cdot \log q) \\ &= \ q^{-8 \cdot 10^8(1+\log(2m)) \log p} \quad \text{for} \quad m \ge 1. \end{aligned}$$

We thus obtain that, if $H \ge 60$ and $2m \le H$, then

$$1 + \log(2m) \le 1 + \log H \le 1.245 \log H \qquad (H \ge 60),$$

and so inequality (3.6) leads to

$$\begin{aligned} \frac{1}{\|a_m\|} \ &\le \ (\log q)\, q^{(8 \cdot 1.245)\cdot 10^8 \log H \log p} < q^{(10^9 - 3) \log H \log p} \\ &= \ H^{(10^9-3)\log p \cdot \log q} < H^{C_5 - 2}. \end{aligned}$$

Thus,

$$D_N \le 3\Big(\frac{1}{H} + \frac{H^{C_5-2}}{N}\sum_{m=1}^{H}\frac{1}{m}\Big) < 3\Big(\frac{1}{H} + \frac{H^{C_5-1}}{N}\Big),$$

where we trivially bounded the sum by H. Choosing $H := \lfloor N^{1/C_5} \rfloor$ we get, assuming still that $H \ge 60$ and therefore that

$$N^{1/C_5} \ge 60, \quad \text{which is equivalent to} \quad N \ge 60^{C_5}, \tag{3.7}$$

we obtain

$$D_N \le 3\Big(\frac{1}{H} + \frac{H^{C_5-1}}{N}\Big) \le 3\Big(\lfloor N^{1/C_5}\rfloor^{-1} + N^{-1/C_5}\Big) \le 7N^{-1/C_5}.$$

To derive the final inequality we used the trivial observation that if $x \geq 60$, then

$$\frac{1}{\lfloor x \rfloor} + \frac{1}{x} \leq \frac{1}{x}\Big(\frac{1}{1-\frac{1}{x}} + 1\Big) \leq \frac{119}{59} \cdot \frac{1}{x} < \frac{7}{3} \cdot \frac{1}{x}.$$

Turning now our attention to the inequality (3.5), we see that

$$N\Big(\frac{\log \alpha}{\log q} - 2D_N\Big) > N\Big(\frac{\log \alpha}{\log q} - 14N^{-1/C_5}\Big). \tag{3.8}$$

Thus, if $N \geq N_0$ with

$$N_0 := \Big(\frac{60 \log q}{\log \alpha}\Big)^{C_5}, \tag{3.9}$$

and hence $N_0 \geq 60^{C_5}$, the right–hand side of (3.8) is at least

$$\frac{23}{30} N \frac{\log \alpha}{\log q} \tag{3.10}$$

and hence positive. From what we have seen at the beginning of Section 3.1, we have $n_{p,q}(\alpha) \leq q^e \cdot p^g < p^{f+g}$. The proof is now completed on invoking (3.3). □

A similar proof also appears in Ferrari et al. [7] in the context of bounding the largest exceptional value of the discriminator of certain Lucas sequences.

3.3. **Application to the discriminator of Lucas type sequences.** Faye et al. [6] considered the sequence $\{U_n(k)\}_{n\geq 0}$ defined uniquely by

$$U_{n+2}(k) = (4k+2)U_{n+1}(k) - U_n(k), \quad U_0(k) = 0,\ U_1(k) = 1.$$

For $k = 1$, the sequence is

$$0, 1, 6, 35, 204, 1189, 6930, 40391, 235416, 1372105, \ldots,$$

which is A001109 in OEIS. On noting that

$$U_{n+2}(k) - U_{n+1}(k) = 4kU_{n+1}(k) + U_{n+1}(k) - U_n(k) \geq 1,$$

one sees that the sequence $U_n(k)$ consists of strictly increasing non-negative numbers. We can now define its *discriminator* $\mathcal{D}_k(n)$ as the smallest positive integer m such that $U_0(k), \ldots, U_{n-1}(k)$ are pairwise distinct modulo m.

Primes of the form $2^n - 1$ are called *Mersenne primes*. Note that n has to be a prime. The first few Mersenne primes are: $3, 7, 31, 127, 8191, 131071, 524287, \ldots$ The first few with exponent $n \equiv 1 \bmod 4$ are $31, 8191, 131071, 2305843009213693951, \ldots$ Choosing k to be a Mersenne prime, it turns out that $\mathcal{D}_k(n)$ is a $\{2, k\}$-unit for all n large enough.

Proposition 3.1. Let p be an odd prime such that also $q := 2^p - 1$ is a prime number. Let $(n_i)_{i\geq 0}$ be the sequence of consecutive S-units with $S = \{2, q\}$. Then

$$\mathcal{D}_q(n) \leq \min\{n_i \geq n\}, \tag{3.11}$$

with equality if the interval $[n, 3n/2)$ contains an S-unit n_j. In particular, if $n \geq n_{2,q}(3/2)$ we have equality in (3.11). If $p \equiv 1 \bmod 4$ and $p > 5$, then we have equality for $n \geq n_{2,q}(5/3)$ and $\mathcal{D}_q(n)$ is a $\{2, 5, q\}$-unit for every $n \geq 1$.

Proof. The first assertion follows on taking for k the Mersenne prime q in Theorem 3 of Faye et al. [6], the second assertion is a consequence of Theorem 3 of Ferrari et al. [6]. □

Numerical work by Matteo Ferrari never led to inequality in (3.11) in case $q = 2^p - 1$ with $p \equiv 3 \bmod 4$, but in case $p \equiv 1 \bmod 4$ he found examples, e.g.,

$$\mathcal{D}_{8191}(129) = 250, \quad \mathcal{D}_{131071}(129) = 250, \quad \mathcal{D}_{131071}(257) = 500.$$

q	$n_{2,q}(5/3)$	q	$n_{2,q}(5/3)$	q	$n_{2,q}(5/3)$
3	1	29	279	61	4358036
5	1	31	274839850	67	10066330
7	1	37	615	71	2458
11	1	41	20	73	2458
13	1	43	20	79	39
17	39322	47	20	83	39
19	154	53	20	89	39
23	10	59	66836	97	39

TABLE 1. $n_{2,q}(5/3)$; $3 \le q \le 100$.

q	$n_{2,q}(3/2)$	q	$n_{2,q}(3/2)$	q	$n_{2,q}(3/2)$
3	1	29	8971	61	18018054422
5	11	31	282022636380491	67	2932031007403
7	131	37	683	71	174763
11	6	41	683	73	174763
13	70	43	22	79	2731
17	699051	47	22	83	2731
19	171	53	1131	89	43
23	11	59	4381419	97	4139

TABLE 2. $n_{2,q}(3/2)$; $3 \le q \le 100$.

As predicted by Proposition 3.1, these values are indeed $\{2, 5, 2^p - 1\}$-units.

One finds that $n_{2,7}(3/2) = 131$ and computing $\mathcal{D}_7(n)$ for the integers $1 \le n \le 130$ shows that for $q = 7$ there is always equality in (3.11), in line with the observations of Ferrari. Unfortunately, we are not able to classify $\mathcal{D}_q(n)$ completely for any Mersenne prime $q > 7$, as then $n_{2,q}(5/3)$ appears to be very large (and hence $n_{2,q}(3/2)$ even more so). The Mersenne prime 127, for example, leads to a very large value of $n_{2,127}(3/2)$: the algorithm, which we will describe in Section 4.1, starts with the pair of exponents $(615, 1)$ and, after having generated 23 150 left neighbors, terminates with the pair $(6, 36)$. Hence, the final result is $n_{2,127}(3/2) = \lceil 2^6 \cdot 127^{36} \cdot 2/3 \rceil$, a number having 78 digits.

For $n_{2,8191}(5/3)$ the algorithm starts with the pair $(73\,800, 1)$ and, after having generated 195 078 401 left neighbors, terminates with the pair $(12, 1\,493)$. Hence, the final result is $n_{2,8191}(5/3) = \lceil 2^{12} \cdot 8191^{1493} \cdot 3/5 \rceil$, a number having 5 847 digits. The whole computation of this case took about 20 minutes and 30 seconds.

For $n_{2,131071}(5/3)$ the algorithm starts with the pair $(1\,544\,466, 1)$, so that n_{start} has 464 936 digits. After having generated 65 305 850 756 left neighbors (more than 65 billion neighbors ...), it terminates with the pair $(16, 23\,897)$. Hence, the final result is $n_{2,131071}(5/3) = \lceil 2^{16} \cdot 131071^{23897} \cdot 3/5 \rceil$, a number having 122 298 digits!! The whole computation of this case took about 5 days and 17 hours. The running times were obtained on a Dell Optiplex 3050, Intel i5-7500, 3.40GHz, 32GB of RAM running Ubuntu 22.04.03-LTS and PARI/GP 2.15.4.

Apart from having sharper estimates for n_{start}, the only hope to obtain $n_{2,q}(3/2)$ or $n_{2,q}(5/3)$ in case $q > 131071$ is a Mersenne prime, is to implement part of this algorithm using the C programming language. This is doable since, once one has gained the knowledge of a sufficiently large number of the $\theta_{2,q}$-convergents, the continued fraction part and the first trivial part can work on exponents only. On the other hand, for the second trivial part the use of multiprecision arithmetic is mandatory.

p	$q = 2^p - 1$	α	$n_{2,q}(\alpha)$	$\ell(n_{2,q}(\alpha))$
2	3	3/2	1	1
3	7	3/2	131	3
5	31	5/3	274 839 850	9
7	127	3/2	$\lceil 2^6 \cdot 127^{36} \cdot 2/3 \rceil$	78
13	8 191	5/3	$\lceil 2^{12} \cdot 8191^{1493} \cdot 3/5 \rceil$	5 847
17	131 071	5/3	$\lceil 2^{16} \cdot 131071^{23897} \cdot 3/5 \rceil$	122 298

TABLE 3. Known values of $n_{2,M_p}(\alpha)$ for Proposition 3.1; $\ell(n)$ is the number of decimals of n.

4. COMPUTATIONS

In this section, we will use heavy computations to search for $n_{p,q}(\alpha)$ and to study the extremal and average behaviors of the gaps $n_{i+1} - n_i$. In both cases, we will use the knowledge of the continued fractions convergents of $\theta_{p,q}$ to generate left or right neighbors of a given S-unit n_i.

The first computational problem we address is to search for $n_{p,q}(\alpha)$ since, as we will see later, its knowledge plays also a role in studying the extremal and average behaviors of the gaps $n_{i+1} - n_i$.

4.1. **Computation of $n_{p,q}(\alpha)$.** We now explain how the algorithm to compute $n_{p,q}(\alpha)$ works. First of all, thanks to Theorem 3.1, we know that it is possible to identify n_{start}, a suitable an upper bound for $n_{p,q}(\alpha)$. After having identified such a point, we will then generate a part of the n_i sequence by looking for S-units less than n_i.

We will now describe a way how to generate such a suitable starting point.

4.1.1. **Generating the starting point $n_{\mathbf{start}}$.** An easy argument, see Section 3.1, shows that if $f, e, g, h \geq 0$ are such that

$$1 < \frac{p^f}{q^e} < \alpha \quad \text{and} \quad 1 < \frac{q^h}{p^g} < \alpha \tag{4.1}$$

then $n_{p,q}(\alpha) \leq p^g \cdot q^e$. On taking logarithms these inequalities can be rewritten as

$$0 < \theta_{p,q} - \frac{e}{f} < \frac{\log \alpha}{f \log q} \qquad \text{and} \qquad 0 < \frac{h}{g} - \theta_{p,q} < \frac{\log \alpha}{g \log q}.$$

These type of inequalities appear in the theory of continued fractions and suggest to take

$$\frac{e}{f} = \frac{r_k}{s_k} \quad \text{and} \quad \frac{h}{g} = \frac{r_\ell}{s_\ell}, \tag{4.2}$$

where k is even and minimal, ℓ is odd and minimal and $(r_m/s_m)_{m\geq 0}$ is the sequence of convergents of $\theta_{p,q}$. Using the inequality

$$\left|\theta_{p,q} - \frac{r_m}{s_m}\right| < \frac{1}{s_m s_{m+1}},$$

see, e.g., [8, Thm. 171], we deduce that in both cases ($m = k$ and $m = \ell$) it suffices to require that

$$\frac{1}{s_m s_{m+1}} < \frac{\log \alpha}{s_m \log q},$$

that is

$$s_{m+1} > \frac{\log q}{\log \alpha}. \tag{4.3}$$

Clearly, the larger q will be, or the closer to 1 the value of α will be, the larger m will become. Let M be the minimal $m \geq 0$ such that (4.3) holds. In practice, to determine M we need a way to generate the continued fractions convergents of $\theta_{p,q}$. To do this, we heavily rely on the PARI/GP [15] internal functions. Letting $\varepsilon \in (0, 1)$ be the required accuracy for the computations, the `contfrac`(x) function of PARI/GP returns the list of the partial quotients $[a_0, \dots, a_n]$ of the continued fraction expansion of $x > 0$ so that $|x - (a_0 + 1/(a_1 + \cdots + 1/a_n))| < \varepsilon$. Using such a continued fraction expansion of x, the `contfracpnqn` function of PARI/GP returns two sorted lists containing $p_n \in \mathbb{N}^*$ and $q_n \in \mathbb{N}^*$, $(p_n, q_n) = 1$, $n \in \mathbb{N}$, the numerators and the denominators of the convergents p_n/q_n of x. In this way, it is relatively easy to obtain both the upper and lower convergents for each fraction $\theta_{p,q}$, $2 \leq p < q$ both primes, we have to work with. Our practical computations are performed with $\varepsilon = 10^{-19}$.

Having now such a sorted list of the denominators s_m of the convergents for $\theta_{p,q}$, a standard dyadic search procedure will quickly provide M. Nevertheless, it is possible to obtain some theoretical information about M; this might be useful in the case one has initially no access to the sorted list of the convergent denominators. Since $s_{m+1} \geq F_{m+2}$, where F_k is the k-th member of the Fibonacci sequence, it is enough to require that $F_{m+2} \geq \log q / \log \alpha$. Recalling that $F_{m+2} \geq \phi^m$, where $\phi = (1 + \sqrt{5})/2$, it suffices to take

$$m \geq \frac{\log \log q - \log \log \alpha}{\log \phi} > 0. \tag{4.4}$$

Hence $M \leq \lceil \frac{\log \log q - \log \log \alpha}{\log \phi} \rceil$. At the cost of complicating our inequality we can actually do a bit better. This improvement often turns out to make a real difference in numerical practice and so is worth the effort; this is due to the fact that both the numerators and the denominators of the continued fractions convergents are exponentially fast increasing sequences. As a consequence, being able to choose a smaller value for M ensures that much smaller values for the exponents in the definition of n_{start} can be chosen, see (4.5).

Namely, we will use the sharper inequality $F_k > \frac{\phi^k}{\sqrt{5}} - \frac{53}{500}$ for every $k \geq 3$ (this inequality is a consequence of Binet's formula $F_k = \frac{\phi^k - (-\phi)^{-k}}{\sqrt{5}}$ and the observation that for $k \geq 3$ one has $|\frac{(-\phi)^{-k}}{\sqrt{5}}| < \frac{53}{500}$). Hence if

$$m + 2 > \frac{\log\left(\frac{\log q}{\log \alpha} + \frac{53}{500}\right) + \frac{\log 5}{2}}{\log \phi}$$

for every $m \geq 1$, then $F_{m+2} \geq \log q / \log \alpha$. It is not hard to verify that $\frac{\phi^{m+2}}{\sqrt{5}} - \frac{53}{500} > \phi^m$ for every $m \geq 0$, and so the latter inequality gives a better lower bound for m than (4.4), namely

$$m \geq \left\lceil \frac{\log\left(\frac{\log q}{\log \alpha} + \frac{53}{500}\right) + \frac{\log 5}{2}}{\log \phi} \right\rceil - 2 \geq M.$$

Once M is determined with the dyadic search procedure or using the previously described estimates, we can take $\{k, \ell\} = \{M, M + 1\}$ in (4.2), and get that

$$n_{p,q}(\alpha) \leq p^g \cdot q^e \leq n_{\text{start}} := \min(p^{s_{M+1}} \cdot q^{r_M}; p^{s_M} \cdot q^{r_{M+1}}). \tag{4.5}$$

In (4.1) we want to find the solutions with p^g and q^e minimal. It follows from the basics of continued fractions that the approach with the convergents as described here is actually optimal. So in retrospect our choice in (4.2) was best possible.

4.1.2. **Generating left neighbors.** We take n_{start} as starting candidate and proceed as follows. Assume that we know $n_i = n_{\text{start}}$ and we want to obtain an S-unit $n^*_{i-1} \le n_{i-1}$ such that $n^*_{i-1} \ge \lfloor n_i/\alpha \rfloor$. Remark that the goal here is not to find n_{i-1}, the predecessor element in the n_i sequence, but just an S-unit less than n_i that verifies the Bertrand Postulate condition. For this reason in this procedure we will always choose, if possible, the S-unit n^*_{i-1} having the maximal distance from n_i compatible with the condition $n^*_{i-1} \ge \lfloor n_i/\alpha \rfloor$, because this reduces the total amount of computations to be performed to determine $n_{p,q}(\alpha)$.

In order to achieve this, we combine three different ways of searching for n^*_{i-1} in the following algorithm.

a) **Searching using continued fraction convergents.** Letting $n_i := p^a \cdot q^b$, we have two possible choices.

1) We search for ℓ such that $a \in [s_\ell, s_{\ell+1}]$, where $(r_m/s_m)_{m \ge 0}$ is the sequence of convergents to $\theta_{p,q}$. Choose l in $\{\ell, \ell-1\}$ such that $r_l/s_l < \theta_{p,q}$; this is equivalent to $x_1 := -s_l \log p + r_l \log q < 0$ and hence $q^{r_l}/p^{s_l} < 1$. Define $N_1 := n_i \exp(x_1) < n_i$.
2) We search for ℓ such that $b \in [r_\ell, r_{\ell+1}]$. Choose l in $\{\ell, \ell-1\}$ such that $r_l/s_l > \theta_{p,q}$; this is equivalent to $x_2 := s_l \log p - r_l \log q < 0$ and hence $p^{s_l}/q^{r_l} < 1$. Define $N_2 := n_i \exp(x_2) < n_i$.

Both N_1 and N_2 are less than n_i and thus candidates to be chosen as n^*_{i-1}. We need now select the best of them; i.e., the one whose distance from n_i is maximal, compatible with the condition $n^*_{i-1} \ge n_i/\alpha \ge \lfloor n_i/\alpha \rfloor$. We point out that we compare here with n_i/α, rather than with $\lfloor n_i/\alpha \rfloor$, since n_i/α allows us, by taking logarithms, to work on the exponents only; in other words, we try to avoid as long as possible the necessity of performing the costly computations of n_i and $\lfloor n_i/\alpha \rfloor$.

Let now $x_{\max} := \max(x_1; x_2)$ and $x_{\min} := \min(x_1; x_2)$. If $x_{\max} < -\log \alpha$, then N_1, N_2 are both $< n_i/\alpha$. In this case we terminate this step with n_i and we proceed with step b); we also remark that at this point we know that $n_{p,q}(\alpha) \le \lceil n_i/\alpha \rceil$.

Assume that $x_{\max} \ge -\log \alpha$; this means that at least one of x_1 and x_2 is $\ge -\log \alpha$. If both x_1 and x_2 are $\ge -\log \alpha$, the best choice is to select the smallest one, $x_{\min}$, since $n_i \exp(x_{\min})$ has the largest distance from n_i. Hence, if $x_{\min} \ge -\log \alpha$, we choose $n^*_{i-1} = n_i \exp(x_{\min}) = \min(N_1; N_2)$; otherwise we are in the case $x_{\min} < -\log \alpha \le x_{\max}$, and we are forced to choose $n^*_{i-1} = n_i \exp(x_{\max}) = \max(N_1; N_2)$. In both cases, we replace n_i with n^*_{i-1} and we repeat step a).

b) **Searching trivially: first part.** Assuming that step a) terminates with n_i, an improved approximation of $n_{p,q}(\alpha)$ is then obtained by trivially searching for a value of $x := a \log p + b \log q$ in the range $[\log(n_i/\alpha), \log n_i)$, where $a, b \in \mathbb{Z}$ run in the intervals $0 \le b \le \lfloor \log n_i / \log q \rfloor$ and $0 \le a \le \lfloor (\log n_i - b \log q)/\log p \rfloor$. Moreover, we can also exploit the fact that the search procedure of step a) produced $\overline{n} = \max(N_1; N_2) = p^{\overline{u}} \cdot q^{\overline{v}}$ such that $\overline{n} < n_i/\alpha$. This means that the value $\overline{x} := \overline{u} \log p + \overline{v} \log q$ is smaller than $\log(n_i/\alpha)$. Since we need to find $x \ge \log(n_i/\alpha) > \overline{x}$, we must have either $a > \overline{u}$ or $b > \overline{v}$. As a consequence, since $\log q > \log p$, the best strategy to determine x is to first search for $\overline{v} + 1 \le b \le \lfloor \log n_i / \log q \rfloor$ and $0 \le a \le \lfloor (\log n_i - b \log q)/\log p \rfloor$. If we have no success, we then work with $0 \le b \le \overline{v}$ and $\overline{u} + 1 \le a \le \lfloor (\log n_i - b \log q)/\log p \rfloor$.

If in one of the previously described procedures we find a solution $x \ge \log(n_i/\alpha)$, we have obtained an S-unit $n^*_{i-1} := \exp(x)$ such that $n_i/\alpha \le n^*_{i-1} < n_i$. In this case, we replace n_i with n^*_{i-1} and we start again the search described in step a). If we do not find any solution, we continue with step c).

c) **Searching trivially: second part.** After steps a)-b) are over, we have not yet determined $n_{p,q}(\alpha)$, since before, for efficiency reasons, we have replaced the condition $\log(\lfloor n_i/\alpha \rfloor) \le x$ with the sharper, but easier to compute, $\log(n_i/\alpha) \le x$. Hence we perform here another trivial search like the previous one, but using the the correct lower bound for x mentioned before. If we find such a solution x, we have obtained an S-unit $n^*_{i-1} := \exp(x)$ such that $\lfloor n_i/\alpha \rfloor \le n^*_{i-1} < n_i$. In this case we replace n_i with n^*_{i-1} and we start again the search described in step a). If we do not find

any solution x, this step has determined n_0, the smallest generated S-unit such that $n_{i-1} \geq \lfloor n_i/\alpha \rfloor$ holds for every $i \geq 1$. In this case, we continue with step d).
d) **Final computation.** In the final step we obtain $n_{p,q}(\alpha) = \lceil n_0/\alpha \rceil$.

The search in step a) is clearly the fastest one. Hence, the previously described procedure optimizes the computational cost by minimizing the number of times we are using the much slower trivial searches of steps b)-c). Moreover, except for step c), in which case the presence of the floor function forces us to compute n_i, the computations can be directly performed on the exponents a, b, rather than with the prime powers involved. This requires far less memory usage and, at the same time, as much smaller numbers are involved, improves the running time of the algorithm. We ran our program for $\alpha \in \{2, 7/4, 5/3, 3/2, 4/3\}$, $2 \leq p < q \leq 500$, on the cluster[1] located at the Dipartimento di Matematica "Tullio Levi-Civita" of the University of Padova; the running times were respectively 6 hours and 52 minutes ($\alpha = 2$), 9 hours and 39 minutes ($\alpha = 7/4$), 11 hours and 23 minutes ($\alpha = 5/3$), one day, 14 hours and 36 minutes ($\alpha = 3/2$), and 14 days, 20 hours ($\alpha = 4/3$).

To show the importance of working on the exponents only, we report here some data about the computations for $n_{3,83}(5/3)$. Here the starting value is $3^4 \cdot 83^{45}$, a number having 89 digits. After 2230 iterations our algorithm reached the pair $(100, 0)$ and hence

$$n_{3,83}(5/3) = \lceil 3^{100} \cdot 3/5 \rceil = 309226512439206798621876677859372763621264513201,$$

a number having 48 digits. Table 4 provides some data for further cases and makes manifestly clear that handling the problem directly - so, without working on the exponents only - would be infeasible.

p	q	α	a_{start}	b_{start}	a_{end}	b_{end}	iterations	$\ell(n_{\text{start}})$	$\ell(n_0)$	$n_{p,q}(\alpha)$
103	223	2	74 796	6	6	6 582	2 372 823 751	150 567	15 469	$\lceil 103^6 \cdot 223^{6582}/2 \rceil$
197	419	2	11 313	7	7	812	55 700 234	25 976	2 146	$\lceil 197^7 \cdot 419^{812}/2 \rceil$
13	89	7/4	44 875	4	6	3 266	566 222 486	49 997	6 374	$\lceil 13^6 \cdot 89^{3266} \cdot 4/7 \rceil$
137	311	7/4	5 174	6	6	1 410	10 339 605	11 071	3 528	$\lceil 137^6 \cdot 311^{1410} \cdot 4/7 \rceil$
89	479	5/3	4 415	8	10	290	7 066 379	8 628	797	$\lceil 89^{10} \cdot 479^{290} \cdot 3/5 \rceil$
293	491	5/3	5 221	11	11	55	12 553 989	12 910	176	$\lceil 293^{11} \cdot 491^{55} \cdot 3/5 \rceil$
79	293	3/2	5 170	10	12	296	10 276 001	9 836	753	$\lceil 79^{12} \cdot 293^{296} \cdot 2/3 \rceil$
313	487	3/2	14	9 749	880	12	8 056 800	26 236	2 229	$\lceil 313^{880} \cdot 487^{12} \cdot 2/3 \rceil$
101	293	4/3	16	14 616	3 412	12	40 708 783	36 088	6 869	$\lceil 101^{3412} \cdot 293^{12} \cdot 3/4 \rceil$
167	367	4/3	29 063	13	14	6 792	339 708 048	64 633	17 451	$\lceil 167^{14} \cdot 367^{6792} \cdot 3/4 \rceil$

TABLE 4. Some data obtained during the computations. In this table $n_0 = p^{a_{\text{end}}} \cdot q^{b_{\text{end}}}$ and $\ell(n)$ is the number of decimals of n.

We will now give a detailed description of the determination of $n_{13,89}(7/4)$.

Example 4.1. To explain why some cases are harder than others, we consider $p = 13$ and $q = 89$. The continued fractions convergents of $\theta_{13,89}$ have numerators and denominators respectively equal

[1]We used a machine having 2 x Intel(R) Xeon(R) CPU E5-2630L v31.80GHz and 192GB of RAM (but we used up to 128GB of RAM in our computations).

to:

$$0,\ 1,\ 1,\ 3,\ 4,\ 25643,\ 179505,\ 205148,\ldots$$
$$1,\ 1,\ 2,\ 5,\ 7,\ 44875,\ 314132,\ 359007,\ldots$$

Hence $r_0/s_0 = 0$, $r_1/s_1 = 1$, $r_2/s_2 = 1/2$, $r_3/s_3 = 3/5$, $r_4/s_4 = 4/7$, $r_5/s_5 = 25643/44875$ and so on. If $\alpha = 2$, we obtain $\lceil \frac{\log 89}{\log 2} \rceil = 7$ and hence we have $M = 3$, $s_{M+1} = 7$, $r_{M+1} = 4$, $s_M = 5$, $r_M = 3$ in (4.5). In this case the algorithm starts with the information that $n_{13,89}(2) \le 13^5 \cdot 89^4 = 23\,295\,754\,887\,613$, a number having 14 digits. On the other hand $\lceil \frac{\log 89}{\log(7/4)} \rceil = 9$, $M = 4$, $s_{M+1} = 44875$, $r_{M+1} = 25643$, $s_M = 7$, $r_M = 4$. So the algorithm starts just with the information that $n_{13,89}(7/4) \le 13^{44875} \cdot 89^4$, a number having 49 997 digits... The huge difference between the height $h(r_4/s_4)$ and $h(r_5/s_5)$ is responsible for $n_{13,89}(7/4)$ being much harder to compute than $n_{13,89}(2)$.

The situation for $p = 2$ and $q = 8191$ is similar: in this case $(r_2, s_2) = (1, 13)$, $(r_3, s_3) = (5677, 73800)$, $\lceil \frac{\log 8191}{\log 2} \rceil = 13$ and $\lceil \frac{\log 8191}{\log(5/3)} \rceil = 23$. Hence in order to establish that $n_{2,8191}(5/3) = \lceil 2^{12} \cdot 8191^{1493} \cdot 3/5 \rceil$, the algorithm starts with $n_{\text{start}} = 2^{73800} \cdot 8191$, a number having 22 220 digits, while, to obtain $n_{2,8191}(2) = 1$, it is sufficient to work with the much smaller starting number $2^{12} \cdot 8191 = 33\,550\,336$.

4.2. **Extremal behavior of the gaps.** We now show how to analyze the behavior of D_1 and D_2 implicitly defined in the inequalities

$$\frac{n_i}{(\log n_i)^{D_1}} < n_{i+1} - n_i < \frac{n_i}{(\log n_i)^{D_2}}, \tag{4.6}$$

where $n_i \ge 3$. We further define $\rho_i = \rho_i(p, q)$ implicitly by

$$n_{i+1} - n_i = \frac{n_i}{(\log n_i)^{\rho_i}}, \quad \textit{i.e.}, \quad \rho_i = -\frac{\log(\frac{n_{i+1}}{n_i} - 1)}{\log\log n_i} = \frac{\log(\frac{n_i}{n_{i+1} - n_i})}{\log\log n_i}, \tag{4.7}$$

so that, if $n_i \ge 3$, then $D_1 = D_1(p, q) = \max_i \rho_i(p, q)$ and $D_2 = D_2(p, q) = \min_i \rho_i(p, q)$. We note that if $n_{j+1} - n_j > n_j$ for some j (that is the Bertrand's Postulate property does not hold for n_j), then $\rho_j(p, q)$ is negative and hence $D_2(p, q) < 0$. As this occurs for at most finitely many j, we can allow ourselves to disregard these outliers. Hence we will evaluate $\rho_i(p, q)$ only for $n_i \ge n_{p,q}(2)$.

We remark that the smallest D_2 is zero and this value is reached when $n_{i+1} = 2n_i$; moreover, the largest D_1 are usually obtained with small powers of primes and, in fact, the maximal D_1 we got is for $n_i = 3, n_{i+1} = 4$, and is $\log 3/\log\log 3 \approx 11.681421\ldots$.

We also remark that $D_j = C_j - \log(C_{j+2})/\log\log n_i$, $j = 1, 2$, where C_1, C_2, C_3, C_4 are defined in Theorem 1.3. Recalling $C_3 = (\log p)^{C_1}$ and $C_4 = 8q$, this means that

$$C_1(p, q) = D_1(p, q)\frac{\log\log n_i}{\log\log n_i - \log\log p} \quad \text{and} \quad C_2(p, q) = D_2(p, q) + \frac{\log(8q)}{\log\log n_i}, \tag{4.8}$$

where, in the first case, we also have to assume $n_i \ne p$.

As we have just explained, in order to have meaningful results we need to work with $\rho_i \ge 0$, which is equivalent with $n_{i+1} \le 2n_i$. We will also require that $n_i := p^a \cdot q^b \ge 3$ and

$$n_i \ge N := \exp(\log p \cdot \log q)\ (= p^{\log q} = q^{\log p}).$$

Since sometimes we do not know the value of $n_{p,q}(2)$, the best strategy is then to start working with n_{start} as identified in Section 4.1.1 and use a modified form of the left neighbor search (explained in Section 4.1.2) to generate n_{i-1}, until we have produced L left neighbors, or we have reached $n_{p,q}(2)$. To do this, we need to slightly modify the search procedure in Section 4.1.2, since in this case the issue is to determine n_{i-1}, rather than to get as close as possible to $n_{p,q}(2)$. Hence in step a) of Section 4.1.2 we will always choose the value of $x_{\max}$ and in steps b)-c) of the same section we will search for the maximal value of the form $x := a \log p + b \log q$ in the ranges there defined, where $a, b \in \mathbb{Z}$ run in the intervals $0 \leq b \leq \lfloor \log n_i / \log q \rfloor$ and $0 \leq a \leq \lfloor (\log n_i - b \log q)/\log p \rfloor$. These changes to the search procedures increase their computational cost; unfortunately, there is no way to work differently since in this case we cannot halt the procedure as soon as we have found an S-unit less than n_i in the prescribed interval, but we need to be sure that such a point is the maximal S-unit less than n_i, or, in other words, that such a point is the left neighbor n_{i-1} of n_i.

If we have reached $n_{p,q}(2)$ without having generated L left neighbors, we start to generate right neighbors from n_{start} until we have obtained a total number of L neighbors. For the generation of the right neighbor n_{i+1} of n_i, we used the continued fraction approach described in [2, Theorem 2.2]. The theoretical justification for the neighbor search procedure is provided in the Appendix.

In each step of the previously described algorithm, as soon as we have determined one of the neighbors of n_i, we can evaluate ρ_i, defined in (4.7), the constants defined in (4.6) and (4.8).

The same remarks we made in Section 4.1 about using the exponents only in the computations apply here as well with a single exception: in evaluating $\log(n_{i+1}/n_i - 1)$ in (4.7), we are in fact forced either to generate both n_i and n_{i+1}, or to use

$$\log(n_{i+1}/n_i - 1) = \log\bigl(\exp(u \log p + v \log q) - 1\bigr),$$

in which we assume $u \log p + v \log q$, where $u, v \in \mathbb{Z}$ are such that $n_{i+1}/n_i = p^u \cdot q^v$, to be known. This is one of the most computationally costly steps of the whole procedure, but unfortunately there is no other way to compute ρ_i.

In this way we were able to collect, for every $2 \leq p < q$, with p and q both primes, all the values of $C_1, C_2, C_3, C_4, D_1, D_2$ and of their averaged values (defined in the next section). This is a heavy computation and the largest P we were able to work with was $P = 10^4$. The data in Table 5 were obtained as a part of computation having an accuracy of 19 decimal digits (but the results are here truncated at 10 digits) performed for every $2 \leq p < q < 10^4$, and using 10^4 neighbors, a computation which required about 2 days and 18 hours on the Dell Optiplex machine mentioned before.

4.3. **Average behavior of the gaps.** Define

$$\mu(p, q; k) := \frac{1}{k^+} \sum_{\substack{i=1 \\ \rho_i(p,q) \geq 0}}^{k} \rho_i(p, q), \quad \textit{with } k^+ := \#\{1 \leq i \leq k : \rho_i(p, q) \geq 0\}.$$

Lemma 4.1. *Let p, q be fixed and $\epsilon > 0$. Recall that $C_1 = 2 \cdot 10^9 \log p \cdot \log q$. There exists an integer $k_{p,q}(\epsilon)$ such that*

$$C_1^{-1} - \epsilon < \mu(p, q; k) < C_1 \quad \textit{for every } k \geq k_{p,q}(\epsilon).$$

Proof. By Theorem 1.3 we have

$$C_1^{-1} - \frac{\log C_4}{\log \log n_i} < \rho_i < C_1 - \frac{\log C_3}{\log \log n_i}.$$

The proof follows from these two inequalities, $C_3 > 0$ and the observation that $(\log \log n_i)^{-1}$ tends to zero. □

We cannot answer the following natural question.

Question 4.1. *Does $\lim_{k \to \infty} \mu(p, q; k)$ exist?*

The numerical work presented here suggests that the answer is **yes**. In this case we write $\mu(p, q)$ for the limit.

p	q	$D_1(p,q)$	$D_2(p,q)$	$\mu(p,q;k)$	p	q	$D_1(p,q)$	$D_2(p,q)$	$\mu(p,q;k)$	p	q	$D_1(p,q)$	$D_2(p,q)$	$\mu(p,q;k)$
2	3	11.6814212141	0.6261567724	1.0941489589	3	5	2.1605296213	0.1323896610	0.8603600118	5	7	1.1384784153	0.1109894596	0.7781281263
2	5	4.2441792885	0.3589187149	0.9474929648	3	7	1.1710669001	0.1383426171	0.7649429172	5	11	1.1887010049	0.0014698384	0.6590355370
2	7	2.9229728309	0.1201698368	0.8284546634	3	11	1.4012089272	0.2086747873	0.7566780823	5	13	1.3211241477	0.0310425989	0.7757032047
2	11	1.8178288541	0.2942087461	0.9858711527	3	13	2.6410842719	0.1285354272	0.7479745611	5	17	1.1035830903	0.1797045904	0.6901667574
2	13	1.7851594152	0.2088055868	0.7625506652	3	17	1.7461871466	0.0795647361	0.7255147576	5	19	1.1398389151	0.1578288483	0.6723537325
2	17	2.7188068070	0	0.8044481097	3	19	1.4225074511	0.2039287543	0.7949565378	5	23	0.9315775235	0.0005087729	0.6645526054
2	19	2.0907970943	0	0.7925235409	3	23	1.2065704232	0.0656360747	0.7683872665	5	29	1.0391485562	0.0813597681	0.6529045038
2	23	1.5215514442	0.2221675922	0.6835338899	3	29	0.9394181270	0.0151595599	0.7455240234	5	31	1.2305570961	0.1189302077	0.5509941851
2	29	1.8685980225	0.0635501877	0.6305104841	3	31	1.5149281514	0.0071479409	0.6700758602	5	37	1.8067753639	0.1479331658	0.7585371271
2	31	2.7834367088	0.0149561107	0.8398947566	3	37	1.4807635709	0.2296184840	0.6676913346	5	41	1.2902483645	0.0690378980	0.5759361463
2	37	1.4934915642	0	0.8004576906	3	41	1.0229032857	0.0138973896	0.7159407244	5	43	1.6660279429	0.0636126574	0.7032472651
2	41	1.3988710027	0	0.7933541465	3	43	0.9916232421	0.0692590381	0.7018910831	5	47	0.9651734788	0.0491064048	0.6754674325
2	43	1.3718840668	0	0.7791951977	3	47	2.2545776040	0.0535541917	0.7078101517	5	53	1.2873497260	0.0322254510	0.4291499629
2	47	1.4407350298	0	0.6686016472	3	53	0.9771181314	0.0200128231	0.7076989903	5	59	1.3174229255	0.0330207317	0.6428076307
2	53	1.4055386972	0	0.6391686515	3	59	1.5682950986	0.2189707656	0.8062334056	5	61	1.0939162195	0.1045174998	0.6276138163
2	59	1.7560379970	0.0474250925	0.7734676479	3	61	1.2885326430	0.1479792903	0.8006209830	5	67	1.0048592858	0.0534692090	0.6291197333
2	61	2.1308613919	0.0239511431	0.7720445935	3	67	0.9915469462	0.1316451029	0.7048711907	5	71	1.0184013877	0.1011594660	0.6371897583
2	67	2.1471869534	0.0460488967	0.6122509704	3	71	1.3780698058	0.0871486016	0.7826861503	5	73	2.1481873755	0.0470304122	0.5549093481
2	71	1.5526948207	0.1074029002	0.7469833108	3	73	0.9560032551	0.0070696424	0.6675886273	5	79	1.4575444151	0.0011909162	0.5340714572
2	73	1.3763643264	0.1382120440	0.7668223147	3	79	1.2434176789	0.0090970147	0.7737402410	5	83	1.2343218159	0.1386930533	0.5927933668
2	79	1.0720303269	0.2320129747	0.7474188012	3	83	1.3056800194	0.0042329888	0.6944828831	5	89	0.9266169776	0.0781941449	0.6043041547
2	83	2.3638631955	0.2964674293	0.8588722462	3	89	1.4320941282	0.0417272844	0.7999629578	5	97	1.0334458810	0.0680491434	0.5810298221
2	89	1.5378684405	0.2093917801	0.7841929598	3	97	1.2268141005	0.1029715740	0.6766541024					
2	97	2.5893808896	0.2590800786	0.6246448440										

TABLE 5. Computed values of $D_1(p,q), D_2(p,q), \mu(p,q;k)$ with $2 \le p \le 5$, $p < q \le 97$ and having generated 10^4 neighbors for each case.

Remark 4.1. Instead of $\mu(p,q;k)$ one can consider $\frac{1}{k}\sum_{i=1}^{k} \rho_i(p,q)$, which has also limit $\mu(p,q)$, if the limit exists. However, our preference is to work with $\mu(p,q;k)$, as numerically it seems to behave more regularly.

The programs and the results here described are available at the address: `www.math.unipd.it/~languasc/Sunits.html`.

`Acknowledgement.` We thank the authors of [2] for helpful correspondence.

A substantial part of the work described here, was carried out during the research stay of Togbé at the Max-Planck Institute for Mathematics in June-July 2023. A short visit of Luca also proved to be very helpful. Luca, Moree and Togbé thank the MPIM and its staff for the great working environment, the hospitality and the support.

The discriminator values reported on were kindly computed by Matteo Ferrari. The integers $n_{p,q}(\alpha)$ were computed on a machine of the cluster located at the Dipartimento di Matematica “Tullio Levi-Civita” of the University of Padova, see `https://hpc.math.unipd.it`. Languasco is grateful for having had such computing facilities at his disposal.

REFERENCES

[1] A. Baker, A sharpening of the bounds for linear forms in logarithms, *Acta Arith.* **21** (1972), 117–129.

[2] A. Bérczes, A. Dujella, L. Hajdu, Some Diophantine properties of the sequence of S-units, *J. Number Theory* **138** (2014), 48–68.

[3] Y. Bugeaud, M. Mignotte, S. Siksek, Classical and modular approaches to exponential Diophantine equations I. Fibonacci and Lucas powers, *Ann. of Math.* (2) **163** (2006), 969–1018.

[4] J.W.S. Cassels, On a class of exponential equations, *Ark. Mat.* **4** (1961), 231–233.

[5] J.-H. Evertse, K. Győry, C.L. Stewart, R. Tijdeman, S-unit equations and their applications, *New advances in transcendence theory* (Durham, 1986), 110–174, Cambridge Univ. Press, Cambridge, 1988.

[6] B. Faye, F. Luca, P. Moree, On the discriminator of Lucas sequences, *Ann. Math. Québec* **43** (2019), 51–71.

[7] M. Ferrari, F. Luca, P. Moree, On the discriminator of Lucas sequences II, in preparation.

[8] G. Hardy, E.M. Wright, *An introduction to the theory of numbers*, sixth edition, Oxford University Press, Oxford, 2008.

[9] D.R. Heath-Brown, The differences between consecutive smooth numbers, *Acta Arith.* **184** (2018), 267–285.

[10] A. Hildebrand and G. Tenenbaum, Integers without large prime factors, *J. Théor. Nombres Bordeaux* **5** (1993), 411–484.

[11] L. Kuipers and H. Niederreiter, *Uniform distribution of sequences*, Wiley-Interscience, New-York, 1974.

[12] M. Langevin, Quelques applications de nouveaux résultats de van der Poorten, *Séminaire Delange-Pisot-Poitou. Théorie des nombres*, tome 17, no 2 (1975-1976), exp. no G12, p. G1–G11.

[13] E.M. Matveev, An explicit lower bound for a homogeneous rational linear form in logarithms of algebraic numbers II, *Izv. Ross. Akad. Nauk. Ser. Math.* **64** (2000), 125–180 (in Russian); English translation *Izv. Math.* **64** (2000), 1217–1269.

[14] P. Moree, Integers without large prime factors: from Ramanujan to de Bruijn, *Integers* 14A (2014), A5, pp. 1–13.

[15] The PARI Group, *PARI/GP version 2.15.4*, Bordeaux, 2023. Available from `http://pari.math.u-bordeaux.fr/`.

[16] C.L. Stewart, On sequences of integers with small prime factors, preprint, `http://arxiv.org/abs/2308.02444`.

[17] A. Thue, Bermerkungen über gewisse Näherungsbrüche algebraischer Zahlen, *Christiania Vidensk. Selsk. Skr.* (1908), Nr. 3.

[18] R. Tijdeman, On integers with many small prime factors, *Comp. Math.* **26** (1973), 319–330.

[19] R. Tijdeman, On the maximal distance between integers composed of small primes. *Comp. Math.*, **28** (1974), 159–162.

[20] R. Tijdeman, H.G. Meijer, On integers generated by a finite number of fixed primes, *Compositio Math.* **29** (1974), 273–286.

[21] A.J. van der Poorten, Computing the lower bound for linear forms in logarithms, *Séminaire Delange-Pisot-Poitou. Théorie des nombres*, 17e année, 1975/76, no. 17, 9 pp.

Appendix A. On efficient left and right neighbors searches

Here we establish two theorems on determining the neighbors of n_i. We will just make use of [2, Lemma 3.1], which characterizes the best approximants of a real number θ in terms of the principal and intermediate convergents generated by its continued fraction. In line with [2] we will use the notation s_i instead of n_i. The sought for neighbors are then s_{i-1} and s_{i+1}.

Theorem A.1 (Left neighbor search). *Suppose $k \geq 1$ and that we are given $s_k = p^{c_k} \cdot q^{d_k}$, $(c_k, d_k) \in \mathbb{Z}^2_{\geq 0}$, $(c_k, d_k) \neq (0,0)$. Then we can compute its left neighbor s_{k-1} in the following way:*

i) Let $u_1/v_1 > 0$, $(u_1, v_1) \in \mathbb{Z}^2_{\geq 0}$, be the upper convergent of $\theta_{p,q}$ with maximal numerator for which $u_1 \leq d_k$ holds.

ii) Let $u_2/v_2 > 0$, $(u_2, v_2) \in \mathbb{Z}^2_{\geq 0}$, be the lower convergent of $\theta_{p,q}$ with maximal denominator for which $v_2 \leq c_k$ holds.

iii) Put $x := |v_1 \log p - u_1 \log q| - |v_2 \log p - u_2 \log q|$ and

$$c_{k-1} = \begin{cases} c_k + v_1 & \text{if } x < 0, \\ c_k - v_2 & \text{if } x > 0, \end{cases} \qquad d_{k-1} = \begin{cases} d_k - u_1 & \text{if } x < 0, \\ d_k + u_2 & \text{if } x > 0. \end{cases}$$

Then we have $s_{k-1} = p^{c_{k-1}} \cdot q^{d_{k-1}}$.

Proof. Let $x < 0$. Then we have $c_{k-1} = c_k + v_1$, $d_{k-1} = d_k - u_1$. Since u_1/v_1 is an upper convergent of $\theta_{p,q} = \log p / \log q$, we have $v_1 \log p - u_1 \log q < 0$, and we define $S := s_k p^{v_1} \cdot q^{-u_1} < s_k$. Now we prove that $S = s_{k-1}$.

By contradiction, if $S \neq s_{k-1}$ there exists $(a, b) \in \mathbb{Z}^2_{\geq 0}$, $(a,b) \neq (0,0)$, such that $S = p^{c_k + v_1} \cdot q^{d_k - u_1} < p^a \cdot q^b < p^{c_k} \cdot q^{d_k} = s_k$. This is equivalent to

$$0 < \frac{d_k - b}{\theta_{p,q}} - (a - c_k) < \frac{u_1}{\theta_{p,q}} - v_1. \tag{A.1}$$

If $b = d_k$, we must have $a < c_k$. Moreover, we will obtain that $u_1/(v_1 + c_k - a)$ is an upper convergent of $\theta_{p,q}$ having the same numerator of u_1/v_1. This implies $a = c_k$, which is impossible because $p^a \cdot q^b < s_k$ by definition.

So we have that $b \neq d_k$. Now, thanks to [2, Lemma 3.1], from (A.1) we obtain $d_k - b > u_1$. If $b > d_k$, it follows that $d_k > b + u_1 > d_k + u_1$ and hence $u_1 < 0$, which is a contradiction. Assume now $0 \leq b < d_k$. Then $\overline{u} := d_k - b > 0$ and $\overline{u} > u_1$. Again using (A.1), we have $\overline{v} := a - c_k > (\overline{u} - u_1)/\theta_{p,q} + v_1 > 0$ and $\overline{u}/\overline{v}$ is an upper convergent for $\theta_{p,q}$. But this is a contradiction, since u_1 is maximal between the numerators $\leq d_k$. This proves that $s_{k-1} = p^{c_k+v_1} \cdot q^{d_k-u_1}$. The case $x > 0$ can be proved analogously. □

In Theorem A.1 we used the same notation used in Theorem 2.2 in [2]; we think that this would help the reader to spot their differences more easily. Remark that the choices of the convergents of $\theta_{p,q}$ in this theorem are precisely the ones in Section 4.1.2, point a), and hence this s_{k-1} corresponds to the choice of $x_{\max}$ there.

We also include an alternative, and shorter, proof of Theorem 2.2 in [2].

Theorem A.2 (Right neighbor search)**.** *Suppose $k \geq 0$ and that we are given $s_k = p^{c_k} \cdot q^{d_k}$, $(c_k, d_k) \in \mathbb{Z}^2_{\geq 0}$. Then we can compute its right neighbor s_{k+1} in the following way:*

i) Let $u_1/v_1 > 0$, $(u_1, v_1) \in \mathbb{Z}^2_{\geq 0}$, be the upper convergent of $\theta_{p,q}$ with maximal denominator for which $v_1 \leq c_k$ holds.
ii) Let $u_2/v_2 > 0$, $(u_2, v_2) \in \mathbb{Z}^2_{\geq 0}$, be the lower convergent of $\theta_{p,q}$ with maximal numerator for which $u_2 \leq d_k$ holds.
iii) Put $x := |v_1 \log p - u_1 \log q| - |v_2 \log p - u_2 \log q|$ and

$$c_{k+1} = \begin{cases} c_k - v_1 & \text{if } x < 0, \\ c_k + v_2 & \text{if } x > 0, \end{cases} \qquad d_{k+1} = \begin{cases} d_k + u_1 & \text{if } x < 0, \\ d_k - u_2 & \text{if } x > 0. \end{cases}$$

Then we have $s_{k+1} = p^{c_{k+1}} \cdot q^{d_{k+1}}$.

Proof. Assume $x < 0$. Then $|v_1 \log p - u_1 \log q| < |v_2 \log p - u_2 \log q|$ and we choose $c_{k+1} = c_k - v_1$, $d_{k+1} = d_k + u_1$. Since u_1/v_1 is an upper convergent of $\theta_{p,q} = \log p/\log q$, we have $u_1 \log q - v_1 \log p > 0$, and we define $S := s_k p^{-v_1} \cdot q^{u_1} > s_k$. Now we prove that $S = s_{k+1}$.

By contradiction, if $S \neq s_{k+1}$ there exists $a, b \geq 0$, $(a, b) \neq (0, 0)$, such that $p^{c_k - v_1} \cdot q^{d_k + u_1} = S > p^a \cdot q^b > s_k = p^{c_k} \cdot q^{d_k}$. This is equivalent to

$$0 < (b - d_k) - (c_k - a)\theta_{p,q} < u_1 - v_1\theta_{p,q}. \tag{A.2}$$

If $a = c_k$, we must have $b > d_k$. Moreover, we will obtain that $(d_k - b + u_1)/v_1$ is an upper convergent of $\theta_{p,q}$ having the same denominator of u_1/v_1. This implies $b = d_k$, which is impossible because $p^a \cdot q^b > s_k$ by definition.

So we have that $a \neq c_k$. Now, thanks to [2, Lemma 3.1], from (A.2) we obtain $c_k > a + v_1$. If $a > c_k$, we obtain $c_k > a + v_1 > c_k + v_1$ and hence $v_1 < 0$, which is a contradiction. Assume now $0 \leq a < c_k$. Then $\overline{v} := c_k - a > 0$ and $\overline{v} > v_1$. Again using (A.2), we have $\overline{u} := b - d_k > \overline{v}\theta_{p,q} > 0$ and $\overline{u}/\overline{v}$ is an upper convergent for $\theta_{p,q}$. But this is a contradiction, since v_1 is maximal between the denominators $\leq c_k$. This proves that $s_{k+1} = p^{c_k - v_1} \cdot q^{d_k + u_1}$. The case $x > 0$ can be proved analogously. □

Dipartimento di Matematica "Tullio Levi-Civita", Università di Padova, Via Trieste, 63, 35121 Padova, Italy
Email address: alessandro.languasco@unipd.it

School of Mathematics, University of the Witwatersrand, P. O. Box Wits 2050, South Africa
Email address: florian.luca@wits.ac.za

Max-Planck-Institut für Mathematik, Vivatsgasse 7, D-53111 Bonn, Germany
Email address: moree@mpim-bonn.mpg.de

Department of Mathematics and Statistics, Purdue University Northwest, 2200 169th Street, Hammond, IN 46323, USA
Email address: atogbe@pnw.edu