On the Diophantine equation $\frac{x^n-1}{x-1} = y^q$

Yann Bugeaud, Maurice Mignotte & Yves Roy (Strasbourg)

Abstract. We prove that if $(x, y, n, q) \neq (18, 7, 3, 3)$ is a solution of the Diophantine equation $(x^n - 1)/(x - 1) = y^q$ with q prime, then there exists a prime number p such that p divides x and q divides $p - 1$. This allows us to solve completely this Diophantine equation for infinitely many values of x. The proofs require several different methods in diophantine approximation together with some heavy computer calculations.

1991 Mathematics Subject Classification : 11D61.

1. Introduction

A longstanding conjecture claims that the Diophantine equation

$$\frac{x^n-1}{x-1} = y^q \quad \text{in integers } x > 1, y > 1, n > 2, q \geq 2 \tag{1}$$

has finitely many solutions, and, maybe, only those given by

$$\frac{3^5-1}{3-1} = 11^2, \ \frac{7^4-1}{7-1} = 20^2 \ \text{and} \ \frac{18^3-1}{18-1} = 7^3.$$

Among the known results, let us mention that Ljunggren [11] solved (1) completely when $q = 2$ and Ljunggren [11] and Nagell [12] when $3 \mid n$ and $4 \mid n$: they proved that in these cases there is no solution, except the previous ones. For more information and in particular for finiteness type results under some extra hypotheses, we refer the reader to Shorey & Tijdeman [17], [18] and to the survey of Shorey [16].

Very recently, some new results have been obtained by Bennett, Bugeaud, Mignotte, Roy, Saradha and Shorey, and, now, equation (1) is completely solved when x is a square (There is no solution in this case [15], [6], [1]), when x is a power of any integer in the interval $\{2, \ldots, 10\}$ (The only two solutions are listed above [5]) or, under hypothesis (H) below, when x is a power of a prime number [5]. In [5] and [6], the proofs require several different methods in diophantine approximation together with some heavy computer calculations, one of the main tools being a new lower bound for linear forms in two p-adic logarithms (see [3]), which applies very well to (1) and allows us to considerably reduce the time of computation.

In the present work, using the same methods, we show that (1) has no solution (x, y, n, q) with $y \equiv 1 \pmod{x}$. As a corollary, we answer a question left open by Edgar [7]. Further, we give an irrationnality statement for Mahler's numbers, which completes results of [15] and [3].

2. Statement of the results

First, we warn the reader that the proofs of the results obtained in Le [10] and in Yu & Le [19] are incorrect. Indeed, they all depend on Lemma 3 of [10], which is false, see the comment of Yuan Ping-Zhi [20]. The purpose of the present is to supply a correct proof of all their claims. However, we notice that our method is essentially different from theirs.

Theorem 1. *Equation (1) has no solution (x, y, n, q) where x and y satisfy the following hypothesis*

$$\textit{Every prime divisor of } x \textit{ also divides } y-1, \tag{H}$$

excepted $(18, 7, 3, 3)$. Consequently, for all other solutions (x, y, n, q) of (1) with q prime, there exists a prime number p such that p divides x and q divides $p - 1$.

The last assertion easily follows from the first one. Indeed, let (x, y, n, q) be a solution of (1), not satisfying (H). Let p be a prime factor of x such that p does not divide $y - 1$. Regarding (1) modulo p, we have $y^q \equiv 1 \pmod{p}$. However, $y \not\equiv 1 \pmod{p}$ and, since q is prime, q must divide $p - 1$. Consequently, equation (1) with $q \geq 3$ prime implies that $x > 2q$.

Remark : Saradha and Shorey [15] showed that equation (1) with (H) implies that $\max\{x, y, n, q\}$ is bounded by an effectively computable absolute constant. Further, Theorem 1 was proved for $x > 6 \times 10^{19}$ in [3] and for $x = z^t$ with $t \geq 1$ and $z = 6$, 10 or z prime in [5].

The main tool of the proof of Theorem 1 is a sharp lower bound for the p–adic distance between two rational numbers, obtained by Bugeaud [3]. Although this improvement of the estimates of Bugeaud & Laurent [4] seems to be very slight, it is of great interest for the problem investigated here. Indeed, it leads to considerably smaller numerical upper bounds, and allows us to do the numerical computation in a reasonable time. We should also mention that this tool can be used only when there is a prime factor of x dividing $y - 1$.

For fixed coprime rational positive integers a and b, we denote by $\mathrm{ord}_b\, a$ the least positive integer value of t for which $a^t \equiv 1 \pmod{b}$. It is known (*cf.* [7]) that every solution (x, y, n, q) of (1) with y an odd prime satisfies $n = \mathrm{ord}_y\, x$, and Edgar [7] asked whether we also have $q = \mathrm{ord}_x\, y$. Using Theorem 1, we are able to answer positively Edgar's question.

Theorem 2. *Every solution (x, y, n, q) of (1) satisfies $q = \mathrm{ord}_x\, y$.*

A problem arising in the theory of finite groups and strongly connected to Equation (1) is to find prime numbers P and Q and rational integers $n \geq 3$ and $a \geq 1$ such that $(Q^n - 1)/(Q - 1) = P^a$, see *e.g.* [8]. Our Theorem 1 allows us to prove that the latter equation with $a \geq 2$ is not solvable for $Q \in \{2, 3, 5, 7, 13, 17, 19, 37, 73, 97, \ldots\}$. In order to give a precise statement, we need first to introduce some notations. Let S_1 be the set of all positive integers greater than 1 and composed only by 2 and by the primes of the form $2^a + 1$, for $a \geq 1$. We remark that S_1 is much bigger than the set also denoted by S_1 in [15]. Let p be a prime number of the form $2^a\, 3^b + 1$, with integers $a \geq 0$, $b > 0$ and $p \not\equiv 55 \pmod{63}$. Thus $p \equiv 1, 4, 7 \pmod{9}$, $p \not\equiv 1 \pmod{7}$, and when $p \equiv 1 \pmod{9}$, we have $p \not\equiv 6 \pmod{7}$. Let f be any non-negative integer satisfying $f \not\equiv 1, 4 \pmod{6}$ if either $p \equiv 1 \pmod{9}$, $p \equiv 3, 4 \pmod{7}$ or $p \equiv 4 \pmod{9}$, and $f \not\equiv 2, 5 \pmod{6}$ if either $p \equiv 1 \pmod{9}$, $p \equiv 2, 5 \pmod{7}$ or $p \equiv 7 \pmod{9}$. Let S_2 be the set of all numbers of the form $2^f\, p$. Put $S_3 = S_1 \cup S_2$ and notice that S_3 is an infinite set. The next statement directly follows from Theorem 1. It completes Corollary 2 of [15] and Corollary 1 of [3]. For its proof, the reader is directed to [15] and [3].

Corollary 1. *Equation (1) has no solution (x, y, n, q), where $x = h^t$, with $h \in S_3$ and $t \geq 1$, other than $(h, t, y, n, q) = (3, 1, 11, 5, 2)$, $(7, 1, 20, 4, 2)$ and $(18, 1, 7, 3, 3)$.*

As already mentioned in details in [15], Theorem 1 can be applied to obtain irrationality statements. Let $g \geq 2$ and $h \geq 2$ be integers. For any integer $m \geq 1$, we define $(m)_h = a_1 \ldots a_r$ to be the sequence of digits of m written in basis h, *i.e.* $m = a_1\, h^{r-1} + \ldots + a_r$, with $a_1 > 0$ and $0 \leq a_i < h$ for $1 \leq i \leq r$. For a sequence $(n_i)_{i \geq 1}$ of non-negative integers, we put

$$a_h(g) = 0.(g^{n_1})_h\, (g^{n_2})_h \ldots$$

and we call Mahler's numbers the real numbers obtained in this way. It is known that $a_h(g)$ is irrational for any unbounded sequence $(n_i)_{i \geq 1}$; see the work of Sander [14] for an account of earlier results in this direction. Sander also considered the case when $(n_i)_{i \geq 1}$ is bounded with exactly two elements occurring infinitely many times, which are called limit points. As mentioned in [15], his paper contained an incorrect application of a result of Shorey & Tijdeman [17], hence his Theorem 3 remains unproved. Here, we extend Corollary 3 of [15] and Theorem 5 of [3] as follows.

Theorem 3. *Let $(n_i)_{i \geq 1}$ be a bounded sequence of non-negative integers which is not ultimately periodic and has exactly two limit points $N_1 < N_2$. Let $g \geq 2$ and $h \geq 2$ be integers such that $g \neq 1 + h + \ldots + h^{L-1}$ for every integer $L \geq 2$ if $(N_1, N_2) = (0, 1)$. Assume also that (N_1, N_2, g, h) is not equal to $(0, 2, 11, 3)$, $(0, 2, 20, 7)$, $(0, 3, 7, 18)$ or to $(1, 4, 7, 18)$ and that $g^{N_2 - N_1}$ is not equal to $1 + h$ whenever $g^{N_1} < h$. If $h \in S_3$, then $a_h(g)$ is irrational.*

Remark : We point out that all the assumptions imposed in Theorem 3 are necessary. Indeed, if (N_1, N_2, g, h) is equal to $(0, 2, 11, 3)$, $(0, 2, 20, 7)$, $(0, 3, 7, 18)$ or to $(1, 4, 7, 18)$, we see that all the digits of g^{N_1}

and g^{N_2} with respect to base h are identically equal to 1 in the first three cases and to 7 in the last case. Further, if $g^{N_1} < h$ and $g^{N_2-N_1} = 1+h$, then we write $g^{N_2} = g^{N_1} + g^{N_1} h$ to see that all the digits of g^{N_1} and g^{N_2} are identically equal to g^{N_1} with respect to base h. Observe that there are instances, for example $(N_1, N_2, g, h) = (0,2,2,3), (0,2,3,8), (0,3,2,7), (0,4,2,15), (0,2,3,8)$, when the relation $g^{N_2-N_1} = h+1$ with $g^{N_1} < h$ and $h \in S_3$ is satisfied. Finally, if $N_0 = 0$, $N_1 = 1$ and $g = 1+h+\ldots+h^{L-1}$ for an integer $L \geq 2$, then the digits of g^{N_1} and g^{N_2} are identically equal to 1. Thus $a_h(g)$ is rational in each of these cases.

3. Auxiliary results

Our main auxiliary result is a lower bound for the p–adic distance between two powers of algebraic numbers. Before stating it, we have to introduce some notation.

Let p be a prime number and denote by v_p the p–adic valuation normalized by $v_p(p) = 1$. Let x_1/y_1 and x_2/y_2 be two non-zero rational numbers and denote by g the smallest positive integer such that

$$v_p\big((x_1/y_1)^g - 1)\big) > 0 \quad \text{and} \quad v_p\big((x_2/y_2)^g - 1\big) > 0.$$

Assume that there exists a real number E such that $v_p\big((x_1/y_1)^g - 1)\big) \geq E > 1/(p-1)$. Theorem BU below provides explicit upper bounds for the p–adic valuation of

$$\Lambda = \left(\frac{x_1}{y_1}\right)^{b_1} - \left(\frac{x_2}{y_2}\right)^{b_2},$$

where b_1 and b_2 are positive integers. As in [3], we let $A_1 > 1, A_2 > 1$ be real numbers such that

$$\log A_i \geq \max\{\log|x_i|, \log|y_i|, E \log p\}, \;\; (i = 1, 2)$$

and we put

$$b' = \frac{b_1}{\log A_2} + \frac{b_2}{\log A_1}.$$

Theorem BU. *With the above notation, let x_1/y_1 and x_2/y_2 be multiplicatively independent and assume that either p is odd or $v_2(x_2/y_2 - 1) \geq 2$. Then we have the upper estimates*

$$v_p(\Lambda) \leq \frac{36.1\, g}{E^3\, (\log p)^4} \big(\max\{\log b' + \log(E\, \log p) + 0.4, 6\, E \log p, 5\}\big)^2 \log A_1 \log A_2$$

and

$$v_p(\Lambda) \leq \frac{53.8\, g}{E^3\, (\log p)^4} \big(\max\{\log b' + \log(E\, \log p) + 0.4, 4\, E \log p, 5\}\big)^2 \log A_1 \log A_2.$$

Proof : This is Theorem 2 of [3]. □

The following lemma is due to Saradha & Shorey [15] and originate in a work of Le [9]. Its proof uses Skolem's method.

Lemma 1. *Let (x, y, n, q) be a solution of equation (1) satisfying the hypothesis (H). Then we have*

$$x^{n+1-2\beta} \leq \left(\frac{n+3}{4}\right)^2 \left(2 + \frac{4}{x}\right)^{n-1} q^{\frac{\alpha\, q}{q-1}},$$

where $\alpha = n+1$ if q does not divide x, $\alpha = 2n$ if q divides x, and $\beta = \max\{1, n/q\}$.

Proof : This is Lemma 18 of [15]. □

Lemma 2. *Let (x, y, n, q) be a solution of equation (1) satisfying the hypothesis (H) and such that q does not divide x. Then we have $x \leq 2000$ if $q = 3$ and*

$$x \leq \max\{961, 2.1382\, q\},$$

if $q \geq 5$.

Proof : The case $q = 3$ follows from Lemma 1. Further, we deduce from the hypothesis (H) that $y \equiv 1 \pmod{x}$. Arguing as in the proof of Lemma 19 of [15], we get $n \geq q + 2$ and, for $q \geq 5$, we conclude exactly as in that lemma, and obtain the claimed upper bound. □

Lemma 3. *Equation (1) has no solution (x, y, n, q) with $y \leq 2n$.*

Proof : Let (x, y, n, q) be a solution of (1). Recall that a primitive prime divisor of $x^n - 1$ is congruent to 1 modulo n and that there exists a primitive prime divisor for every odd n (see [13], page 20), which, consequently, is greater or equal to $2n + 1$. If n is even, then $n = 2m$ with m odd, and we observe that $(x^m - 1)/(x - 1)$ and $x^m + 1$ are relatively prime, each having a primitive prime factor (see [13], page 20). Hence, the lemma is proved. □

An important tool of our proof is a corollary to the following very deep result of Bennett [1], which completes an earlier work of Bennett & de Weger [2].

Theorem BE. *If a, b and q are integers with $b > a \geq 1$ and $q \geq 3$, then the equation*

$$|ax^q - by^q| = 1$$

has at most one solution in positive integers (x, y).

Proof : This is Theorem 1.1 of [1]. □

Corollary BE. *Equation (1) has no solution (x, y, n, q) with $n \equiv 1 \pmod{q}$ and $q \geq 3$.*

Proof : This is Corollary 1.2 of [1], and this follows easily from Theorem BE. Indeed, let (x, y, n, q) be a solution of (1), and assume that for a rational integer ℓ we have $n = q\,\ell + 1$. Then we get $x\,(x^l)^q - 1 = (x-1)\,y^q$. In view of Theorem BE, this is impossible, since $(1, 1)$ is a solution of the equation $x\,X^q - (x-1)\,Y^q = 1$. □

4. Proof of Theorem 1

Let (x, y, n, q) be a solution of (1) satisfying the following assumption

$$\text{Every prime divisor of } x \text{ also divides } y - 1. \qquad (H)$$

In view of the results of [12], [11] and [5] stated in the beginning of section 1, we can suppose that $n \geq 5$ and $x \geq 11$. Moreover, we can restrict ourselves to the case when q is an odd prime number. Indeed, if ℓ is a prime divisor of q, we observe that $(x, y^{q/\ell}, n, \ell)$ is a solution of (1) such that every prime divisor of x also divides $y - 1$, hence divides $y^{q/\ell} - 1$.

- Sharp absolute upper bound for x and q under the assumption (H).

Our first goal is to obtain an absolute bound for q and for x, which improves Theorem 4 of [3]. More precisely, we distinguish the cases q divides x and q does not divide x, and we compute an upper estimate in each case. We proceed as follows : firstly, we apply Theorem BU in order to bound q by a polynomial in $\log x$ and, secondly, we deduce from Lemmas 1 and 2 that x is smaller than a polynomial in q.

Application of Theorem BU.

We put

$$\Lambda = (1-x) - \left(\frac{1}{y}\right)^q = -x^n\, y^{-q},$$

and we note that $1-x$ and $1/y$ are multiplicatively independent (for a proof, see [15], below inequality (51)). Let p be a prime factor of x and let $\alpha \geq 1$ be such that p^α divides x but $p^{\alpha+1}$ does not. We assume that $p^\alpha \neq 2$, and notice that, if 4 divides x, then $y \equiv 1 \bmod 4$ and Theorem BU applies with the prime $p = 2$.

Since p^α divides $(y-1) \cdot \frac{y^q-1}{y-1}$ and, in view of (H), p divides $y-1$, we get that p^α divides $y-1$ when $q \neq p$. If $q = p$, we infer from $\frac{y^p-1}{y-1} \equiv p \bmod p^2$ that $\max\{p, p^{\alpha-1}\}$ divides $y-1$. Thus, we deduce that, if $\alpha \geq 2$, we have $y \geq x^{(\alpha-1)/\alpha} \geq x^{1/2}$, whence, by $x^n > y^q$, we obtain the inequality

$$n \geq (q+1)/2, \tag{2}$$

which appears to be very useful. Applying Theorem BU and Lemma 3, we get $y \geq 11$ and

$$v_p(\Lambda) \leq \delta\, \frac{36.1}{\alpha^3\,(\log p)^4} \left(\max\left\{\log\left(\frac{q}{\log x} + 0.42\right) + 0.4, 6\,\alpha\,\log p, 5\right\}\right)^2 \log y\, \log(x-1)$$

and

$$v_p(\Lambda) \leq \delta\, \frac{53.8}{\alpha^3\,(\log p)^4} \left(\max\left\{\log\left(\frac{q}{\log x} + 0.42\right) + 0.4, 4\,\alpha\,\log p, 5\right\}\right)^2 \log y\, \log(x-1),$$

with $\delta = 1$ if $p \neq q$ or $\alpha = 1$ and $\delta = \alpha/(\alpha-1)$ if $p = q$ and $\alpha \geq 2$.

Further, $v_p(\Lambda) = n\alpha$ and $n \log x \geq q\, \log y$, thus we get

$$q \leq \delta\, \max\left\{\frac{36.1}{\alpha^2\,(\log p)^2}\left(\log\left(\frac{q}{\log x} + 0.42\right) + 0.4\right)^2, 36 \cdot 36.1, \frac{25 \cdot 36.1}{\alpha^2\,(\log p)^2}\right\} \frac{\log^2 x}{\alpha^2\,(\log p)^2} \tag{3}$$

and

$$q \leq \delta\, \max\left\{\frac{53.8}{\alpha^2\,(\log p)^2}\left(\log\left(\frac{q}{\log x} + 0.42\right) + 0.4\right)^2, 16 \cdot 53.8, \frac{25 \cdot 53.8}{\alpha^2\,(\log p)^2}\right\} \frac{\log^2 x}{\alpha^2\,(\log p)^2}. \tag{4}$$

Application of Lemma 1.

If q does not divide x, Lemma 2 provides the upper bound $x \leq 2000$ if $q = 3$ and, else,

$$x \leq \max\{961, 2.1382q\}. \tag{5}$$

Assume now that q divides x. If $q = 3$, an easy calculation leads to $x \leq 160000$. If $q \geq 5$ and $n \geq q$, it follows from Lemma 1 that

$$x \leq \left(\frac{n+3}{4}\right)^{\frac{2q}{n(q-2)+q}} \left(2 + \frac{4}{x}\right)^{\frac{q(n-1)}{n(q-2)+q}} q^{\frac{q^2}{q-1}\,\frac{2n}{n(q-2)+q}},$$

whence

$$x \leq \left(\frac{q+3}{4}\right)^{\frac{2}{q-1}} \left(2 + \frac{4}{x}\right)^{\frac{q}{q-2}} q^{\frac{2\,q^2}{(q-1)(q-2)}}, \tag{6}$$

and, assuming $x \geq 10^6$ and considering separately the cases $q \geq 800$ and $q < 800$, we get

$$x \leq \max\bigl\{1.4 \times 10^6, 2.05\, q^{2q^2/(q-1)(q-2)}\bigr\}. \tag{7}$$

If $q \geq 5$ and $n < q$, it follows from Lemma 1 that

$$x \leq \left(\frac{n+3}{4}\right)^{2/(n-1)} \left(2 + \frac{4}{x}\right) q^{\frac{q}{q-1}\,\frac{2n}{n-1}}. \tag{8}$$

Since the map $q \mapsto q^{q/(q-1)}$ is increasing, we infer from (2) that

$$x \leq \left(\frac{n+3}{4}\right)^{2/(n-1)} \left(2 + \frac{4}{x}\right) (2n-1)^{\frac{n(2n-1)}{(n-1)^2}},$$

and, assuming that $n \leq 400$, we get $x \leq 1.4 \times 10^6$. For $n > 400$, we deduce from (8) under the assumption $x \geq 10^6$ the bound

$$x \leq 2.05\, q^{\frac{2.005\, q}{(q-1)}}. \tag{9}$$

By (5), (7) and (9), we see that in all cases we have

$$x \leq \max\{1.4 \times 10^6, 2.05\, q^{2.005q/(q-1)}\}. \tag{10}$$

Absolute upper bound for q and for x.

Suppose first that q does not divide x and recall that we have assumed $x \geq 11$. If $x \neq 12$, then the primary part of x, *i.e.* the greatest prime power dividing x, is at least equal to 5. Using (3) with p^α replaced by 5 and noticing that $\delta = 1$, we get

$$q \leq \max\left\{5.39 \left(\log\left(\frac{q}{\log x} + 0.42\right) + 0.4\right)^2, 501.8\right\} \log^2 x. \tag{11}$$

We make a direct computation in the case $x = 12$ and combine (11) with (5) to obtain that, for all x, we have

$$q \leq 72000$$

and, consequently,

$$x \leq 154000. \tag{12}$$

Suppose now that q divides x, whence q^2 divides x, as easily seen. We then use (4) with the prime number $p = q$ and with $\alpha \geq 2$. Since $\delta \leq 2$ and $4\,\alpha\, \log 3 \geq 5$, we get

$$q \leq 2 \max\left\{\frac{53.8}{\alpha^2 (\log q)^2} \left(\log\left(\frac{q}{\log x} + 0.42\right) + 0.4\right)^2, 860.8\right\} \frac{\log^2 x}{\alpha^2 (\log q)^2}. \tag{13}$$

Combining (13) with (10) and replacing α by 2 yields

$$q \leq 1901$$

and

$$x \leq 7.76 \times 10^6. \tag{14}$$

- Strategy of the computational part of the proof.

In view of the above discussion, we are left to consider a finite number of pairs (x, q). It follows from (12) and (14) and from the results of [12], [11] and [5] that we have to prove that, for given integers $11 \leq x_0 \leq 7.76 \times 10^6$ and $q_0 \geq 3$, there is no solution (x_0, y, n, q_0) of (1) satisfying hypothesis (H), with $n \not\equiv 0 \pmod 3$ and $n \not\equiv 0 \pmod 4$. Furthermore, if such solution exists, known results on Catalan's equation imply that n must be odd. Indeed, if $n = 2m$, then $x_0^m - y_1^{q_0} = -1$ for some positive integer y_1 dividing y, and this is impossible since $x_0 < 10^{11}$, by a result of Hyyrö (see [13], pages 261 and 263).

We now describe our algorithm.

Firstly, using (3) or (4), we compute the bound on q obtained for x_0. If this bound is smaller than q_0 we have of course nothing to do (in this case, there is no solution), otherwise we have to work. And to work means :

– to consider the first prime numbers p with $p \equiv 1 \pmod{q_0}$ and to work modulo p,

– for each p, it leads to some conditions on the exponent n of equation (1), more precisely it implies that n belongs to some set modulo $p-1$,

– it appears that combining these conditions for several values of p and using that $n \not\equiv 0 \pmod 3$ and $n \not\equiv 0 \pmod 2$ yield $n \equiv 1 \pmod{q_0}$, which is excluded by Corollary BE, due to Bennett.

It remains us to describe how we proceed to treat all the pairs (x, q).

We first prove that there is no solution satisfying (H), $q \leq 97$, $n \not\equiv 0 \pmod 3$ and $n \not\equiv 0 \pmod 2$. Indeed, we prove a little more, namely that there is no solution with $q \leq 97$ and x less than the bound given by (6).

Secondly, assuming that $q \geq 101$ and q divides x, we treat the pairs (x, q) with

$$101 \leq q \leq 1901, \quad x \leq 7.76 \times 10^6 \quad \text{and} \quad q^2 \text{ divides } \quad x.$$

This can be done very quickly, since there are not many pairs to consider.

Lastly, we are left with pairs (x, q) such that q does not divide x. After some hundreds of hours of CPU with very fast computers, we could treat the full range $101 \leq q \leq 100000$ and x bounded by (5). It appeared that there is no solution. $\square$

5. Proof of Theorems 2 and 3

Proof of Theorem 2 : Let (x, y, n, q) be a solution of (1) and set $k = \mathrm{ord}_x y$. We have $y^q \equiv 1 \pmod x$ and k divides q. Letting $q' = q/k$ and $y' = y^k$ and assuming that $q' > 1$, we see that (x, y', n, q') is also a solution of (1), which satisfies $y' \equiv 1 \pmod x$. In view of Theorem 1, this is impossible. Thus $q' = 1$ and $k = q$, as asserted. $\square$

Proof of Theorem 3 : Sander ([14], Theorem 2) proved that $a_h(g)$ is irrational if and only if $g^{N_2 - N_1} \neq (h^{tL} - 1)/(h^t - 1)$ for every integer $L \geq 1$, where t is given by the inequalities $h^{t-1} \leq g^{N_1} < h^t$. As noticed in [15], we have $(N_1, N_2) = (0, 1)$ or $N_2 - N_1 \geq 2$. To the first case corresponds the first condition in the statement of Theorem 3. Now we assume that $N_2 - N_1 \geq 2$ and $L = 2$, *i.e.* $g^{N_2 - N_1} = h^t + 1$. We have $t \geq 2$ by an assumption of Theorem 3 and we observe that $(g, h, t) = (3, 2, 3)$ is excluded by the definition of t. Let p be a prime divisor of $N_2 - N_1$. In view of the results on Catalan equation, $N_2 - N_1$ and t are odd and we have $p \geq 5$. We re-write the equation as $h^t = G^p - 1$ with $G = g^{(N_2 - N_1)/p}$ to observe that h is divisible by a prime number of the form $1 + sp$. Thus $1 + sp = 2^a + 1$ or $1 + sp = 2^a 3^b + 1$ since $h \in S_3$. This is not possible for $p \geq 5$.

Consequently, we have $N_2 - N_1 \geq 2$ and $L \geq 3$, whence we deduce from Corollary 1 that $(h, t, g, L, N_2 - N_1)$ belongs to $\{(3, 1, 11, 5, 2), (7, 1, 20, 4, 2), (18, 1, 7, 3, 3)\}$, and we conclude by the definition of t. $\square$

Remark : Shorey has pointed out to us that an assumption $g^{N_2 - N_1} \neq h + 1$ whenever $g^{N_1} < h$ should be added in Corollary 3 of [15]. Indeed, observe that $a_h(g)$ is rational if $g^{N_2 - N_1} = h + 1$ with $g^{N_1} < h$.

Acknowledgement : The authors are very grateful to Tarlok Shorey. His numerous remarks helped us to considerably improve the presentation of the paper.

References

[1] M. Bennett. Rational approximation to algebraic number of small height : The diophantine equation $|\, ax^n - by^n \,| = 1$. Preprint.

[2] M. Bennett and B.M.M. de Weger. On the diophantine equation $|\, ax^n - by^n \,| = 1$. *Math. Comp.* **67** (1998), 413–438.

[3] Y. Bugeaud. Linear forms in p-adic logarithms and the Diophantine equation $(x^n - 1)/(x - 1) = y^q$. *Math. Proc. Cambridge Philos. Soc.* To appear.

[4] Y. Bugeaud and M. Laurent. Minoration effective de la distance p-adique entre puissances de nombres algébriques. *J. Number Th.* **61** (1996), 311–342.

[5] Y. Bugeaud, M. Mignotte. On integers with identical digits. Submitted.

[6] Y. Bugeaud, M. Mignotte, Y. Roy and T. N. Shorey. On the diophantine equation $(x^n - 1)/(x - 1) = y^q$. *Math. Proc. Cambridge Philos. Soc.* To appear.

[7] H. Edgar. Problems and some results concerning the Diophantine equation $1 + A + A^2 + \ldots + A^{x-1} = P^y$. *Rocky Mountain J. Math.* **15** (1985), 327–329.

[8] R. M. Guralnick. Subgroups of Prime Power Index in a Simple Group. *J. Algebra* **81** (1983), 304–311.

[9] Maohua Le. A note on the diophantine equation $(x^m - 1)/(x - 1) = y^n$. *Acta Arith.* **64** (1993), 19–28.

[10] Maohua Le. A note on perfect powers of the form $x^{m-1} + \ldots + x + 1$, *Acta Arith.* **69** (1995), 91–98.

[11] W. Ljunggren. Noen Setninger om ubestemte likninger av formen $(x^n - 1)/(x - 1) = y^q$. *Norsk. Mat. Tidsskr.* **25** (1943), 17–20.

[12] T. Nagell. Note sur l'équation indéterminée $(x^n - 1)/(x - 1) = y^q$. *Norsk. Mat. Tidsskr.* **2** (1920), 75–78.

[13] P. Ribenboim. Catalan's Conjecture. Academic Press, Boston (1994).

[14] J.W. Sander. Irrationality Criteria for Mahler's numbers. *J. Number Theory* **52** (1995), 145-156.

[15] N. Saradha and T. N. Shorey. The equation $(x^n - 1)/(x - 1) = y^q$ with x square. *Math. Proc. Cambridge Philos. Soc.* To appear.

[16] T. N. Shorey. Exponential diophantine equations involving product of consecutive integers and related equations. To appear.

[17] T. N. Shorey and R. Tijdeman. New applications of Diophantine approximations to Diophantine equations, *Math. Scand.* **39** (1976), 5–18.

[18] T. N. Shorey and R. Tijdeman. Exponential Diophantine equations, Cambridge Tracts in Mathematics 87 (1986), Cambridge University Press, Cambridge.

[19] Li Yu and Maohua Le. On the diophantine equation $(x^m - 1)/(x - 1) = y^n$, *Acta Arith.* **83** (1995), 363–366.

[20] Ping-Zhi Yuan. Comment : "A note on perfect powers of the form $x^{m-1} + \ldots + x + 1$", Acta Arith. 69 (1995), 91–98, by Maohua Le and "On the diophantine equation $(x^m - 1)/(x - 1) = y^n$", Acta Arith. 83 (1995), 363–366, by Li Yu & Maohua Le. *Acta Arith.* **83** (1998), 199.

Yann Bugeaud, Maurice Mignotte & Yves Roy
Université Louis Pasteur
U. F. R. de mathématiques
7, rue René Descartes
67084 STRASBOURG

e-mail : bugeaud@math.u-strasbg.fr   mignotte@math.u-strasbg.fr
yr@dpt-info.u-strasbg.fr