PERFECT POWERS WITH FEW BINARY DIGITS AND RELATED DIOPHANTINE PROBLEMS

MICHAEL A. BENNETT, YANN BUGEAUD AND MAURICE MIGNOTTE

Abstract. We prove that, for any fixed base $x \geq 2$ and sufficiently large prime q, no perfect q-th powers can be written with 3 or 4 digits 1 in base x. This is a particular instance of rather more general results, whose proofs follow from a combination of refined lower bounds for linear forms in Archimedean and non-Archimedean logarithms.

1. Introduction

A celebrated theorem of Mihăilescu [18] asserts that 8 and 9 are the only consecutive perfect powers of positive integers. One way of interpreting this result is that 9 is the only perfect power which can be written in the form $10\ldots01$ (with at least one digit 0) in some integer base. A (closely) related Diophantine problem is the search for perfect powers N having only digits 1 in some integer base other than N or $N-1$ (observe than N is written 11 in base $N-1$). We know precisely three numbers with this property, namely

$$3^4+3^3+3^2+3+1=11^2,\ \ 7^3+7^2+7+1=20^2 \ \text{ and } \ 18^2+18+1=7^3,$$

and it is widely believed that the corresponding Diophantine equation

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

commonly termed the Nagell–Ljunggren equation, has no solutions beyond the three listed above. Remarkably, with current technology, it is not even known whether this equation has finitely many solutions in the four variables. For a given fixed value of x, however, it is always possible to solve (1), at least in principle; this has been done for $x \leq 10^6$, see [6, 8, 9] and the survey [7]. The key

Date: 06 Oct. 2011.

2000 *Mathematics Subject Classification.* Primary: 11A63, Secondary: 11D61, 11J86.

Key words and phrases. Polynomial-exponential equations, digital problems.

Research supported by NSERC.

theoretical tools used to bound q in equation (1), for given x, are estimates for linear forms in two non-Archimedean logarithms. Refinements of this theory developed in [2, 3] enable one to solve equation (1) for large infinite families of x.

The starting point of the present paper is the following related problem. We fix an integer base $x \geq 2$ and aim to determine all the perfect powers whose representation in base x has relatively few non-zero digits. While an apparently simple question, for the situation where we allow more than three such digits, it does not appear to be a routine matter to determine whether the corresponding set of powers is finite or infinite (but see [12] for the case $x = 2$ with at most 4 digits). Since $(1 + x^\ell)^2 = 1 + 2x^\ell + x^{2\ell}$ for $\ell \geq 1$, there are infinitely many squares having only three non-zero digits in base x. From a result of Corvaja and Zannier [10], for x fixed, all but finitely many squares with this property can be classified by means of polynomial identities. The dependence in the proof of this fact on the Subspace Theorem, however, renders the finiteness statement ineffective (i.e. one may bound the cardinality of the finite exceptional set, but not the heights of its elements). It is still an open question to determine whether squares having only four non-zero digits in base $x > 2$ can be classified likewise.

In the case of base $x = 2$, a completely explicit characterization of odd squares with three binary digits has been obtained by Szalay [22]. Furthermore, Luca [16], extending work of Scott [19], showed that there are no squares of the form $p^a + p^b + 1$, where p is an odd prime and $a > b > 0$.

In the present paper, we establish a number of more general results along these lines. We prove that, for any fixed base $x \geq 2$ and sufficiently large prime q, no perfect q-th powers can be written with 3 or 4 digits 1 in base x. The key tool in our proof is the theory of linear forms in two logarithms, and especially a refinement in the non-Archimedean case obtained in [2, 3] (which, roughly speaking, allows one to replace the product of the heights occurring in the classical estimates by their sum, at least under certain technical hypotheses). This is the first application of this refinement which really utilizes its full strength. Additionally, our method applies to more general equations of the form $x_1^a + x_2^b + x_3^c + 1 = y^q$, where x_1, x_2, x_3 are fixed positive integers: under the assumption that x_1, x_2, x_3 have a common prime divisor $p \not\equiv 1 \pmod{q}$, we establish that q is effectively bounded in terms of x_1, x_2, x_3.

The same assumption of non-coprimality also appears in the work of Corvaja and Zannier [10], where, amongst other results, it is shown that, for any fixed prime q, the Diophantine equation

$$6^a + 2^b + 1 = y^q$$

has only finitely many solutions.

As noted by Corvaja and Zannier (see page 169 of [11]), the rather curious problems we consider here fit into a more general framework, as solutions to such polynomial-exponential equations correspond to S-integral points on certain projective varieties, for suitable sets of primes S. Viewed in this light, finiteness statements for these equations would follow from essentially the simplest open case of a deep conjecture of Lang-Vojta on Zariski denseness of S-integral points on varieties of log-general type (see e.g. page 486 of [13]).

The outline of this paper is as follows. The statements of our theorems are gathered in Section 2. Section 3 is devoted to collecting a number of lower bounds for linear forms in two logarithms from the existing literature. Our results themselves are proved in Section 4. We conclude the paper with a number of remarks and open problems.

2. Results

Throughout this paper, we denote by ϕ the Euler totient function. For a prime number p and a nonzero integer x, we write $\nu_p(x)$ for the largest power of p dividing x, and, for nonzero rational x/y, set $\nu_p(x/y) = \nu_p(x) - \nu_p(y)$. We begin with a result on the representation of perfect powers in base x with digits in $\{0, 1\}$:

Theorem 1. *There exists an absolute effective constant C with the following property. Let x be a positive integer and suppose that there exist integers a, b, y and a prime q such that*

$$x^a + x^b + 1 = y^q, \quad a > b > 0.$$

Then $q < C$ or q divides $\phi(x)$.

What is of interest here is that, subject to the coprimality of q and $\phi(x)$, the upper bound for q is independent of x. As will be apparent from our proofs, this coprimality condition can be replaced by the weaker assumption that y is congruent to 1 modulo each of the prime divisors (other than q) of x. A like condition appeared already in the study of the Nagell–Ljunggren equation (1); see Theorem 1 of [9].

In point of fact, the techniques we employ here, with additional arguments, enable us to treat rather more general equations, provided only that the "bases" have a common divisor.

Theorem 2. *There exists an absolute effective constant C with the following property. Let x_1 and x_2 be positive integers with $\gcd(x_1, x_2) > 1$. If there exist nonnegative integers a, b, y and an odd prime q, with q coprime to $(p-1)$ for some prime p dividing $\gcd(x_1, x_2)$, such that*

$$x_1^a + x_2^b + 1 = y^q, \tag{2}$$

then $q \leq C \log^2 \max\{x_1, x_2\}$.

Theorem 3. *There exists an absolute effective constant C with the following property. Let x_1, x_2 and x_3 be positive integers with $\gcd(x_1, x_2, x_3) > 1$. If there exist positive integers a, b, c, y and an odd prime q, with q coprime to $(p-1)$ for some prime p dividing $\gcd(x_1, x_2, x_3)$, such that*

$$x_1^a + x_2^b + x_3^c + 1 = y^q, \tag{3}$$

then $q \leq C \log^4 \max\{x_1, x_2, x_3\}$.

In the preceding two theorems, the exponents 2 and 4 present in the quantities

$$\log^2 \max\{x_1, x_2\}, \quad \log^4 \max\{x_1, x_2, x_3\}$$

are included to provide a flavour of what can be obtained. They are far from optimal and, as examination of our proofs suggest, may be readily reduced.

Our proofs demonstrate the power of the theory of linear forms in non-Archimedean logarithms; in some sense, they represent the first application of the techniques of [2, 3] to fully utilize the improvements inherent therein.

In a companion paper [1], we make a number of these results entirely explicit, solving completely, for example, the Diophantine equation $x^a + x^b + 1 = y^q$ for $x \in \{2, 3\}$ and $q \geq 2$, and showing that neither of the equations $6^a + 2^b + 1 = y^q$ and $2^a + 2^b + 2^c + 1 = y^q$ has a solution with $q \geq 5$ prime.

3. Linear forms in two logarithms

In this section, we gather estimates for linear forms in two logarithms, both in Archimedean and in non-Archimedean settings.

We begin with a very special version of a corollary obtained in [15]; here and henceforth, if r is a nonzero rational with $r = m/n$ for m and n coprime integers, we define the logarithmic height of r as $h(r) = \max\{\log|m|, \log|n|, 1\}$.

Theorem 4. *Let α_1 and α_2 be multiplicatively independent positive rational numbers, and b_1 and b_2 be positive integers. Define*

$$\Lambda = |b_2 \log \alpha_2 - b_1 \log \alpha_1|.$$

Then

$$\log \Lambda \geq -25.2 \left(\max\{\log b' + 0.38, 10\}\right)^2 h(\alpha_1)\, h(\alpha_2),$$

where

$$b' = \frac{b_1}{h(\alpha_2)} + \frac{b_2}{h(\alpha_1)}.$$

A more complicated statement, involving an additional parameter E, is given in [15]. This parameter allows us to considerably sharpen the lower estimate when α_1 and α_2 are very close to 1. Since this assumption is not fulfilled in the present work, we have chosen to quote this simplified statement.

We now turn our attention to non-Archimedean results. Suppose we are given two multiplicatively independent positive rational numbers α_1 and α_2. We define g to be the smallest positive integer such that both $\nu_p(\alpha_1^g - 1)$ and $\nu_p(\alpha_2^g - 1)$ are positive. Further, choose E so that

$$\nu_p(\alpha_i - 1) \geq E > \frac{1}{p-1}, \qquad \text{for } \ i = 1, 2.$$

As a special case of Theorem 2 of [2], we have

Theorem 5. *Let α_1 and α_2 be multiplicatively independent positive rational numbers, and b_1 and b_2 be positive integers. Consider the "linear form"*

$$\Lambda = \alpha_2^{b_2} - \alpha_1^{b_1}.$$

Then, for any fixed prime number p,

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

if p is odd or if $p = 2$ and $\nu_2(\alpha_2 - 1) \geq 2$, where

$$b' = \frac{b_1}{\log A_2} + \frac{b_2}{\log A_1} \quad \textit{and} \quad \log A_i \geq \max\{h(\alpha_i), E \log p\}.$$

If $p = 2$ and $\nu_2(\alpha_2 - 1) \leq 1$, then

$$\nu_p(\Lambda) \leq 208 \left(\max\{\log b' + 0.04, 10\}\right)^2 (\log A_1)\ (\log A_2)\,.$$

No parameter E occurs in the earlier bounds for the p-adic distance between two powers of rational numbers. By taking $E = 1$ in Theorem 5, we recover essentially an estimate proved in [5]. When the parameter E is large compared with the heights of α_1 and α_2, however, the results of [5] may be substantially improved. Indeed, if E is as large as $\min\{\log A_1, \log A_2\}$ (it cannot be much larger), then the quantity $(\log A_1)(\log A_2)/E$ becomes $\max\{\log A_1, \log A_2\}$. Thus, the product $(\log A_1)(\log A_2)$ arising in the classical estimates [5] is replaced by $\max\{\log A_1, \log A_2\}$.

Theorem 5 was subsequently generalized in [3] to treat simultaneously several non-Archimedean places. The assumptions of the next theorem (a special case of Theorem 3 of [3]) are very restrictive,

but are satisfied (and easy to check) in our context. Note that if $m = p_1^{j_1} \cdots p_k^{j_k}$ where the p_i's are distinct primes and $j_i \in \mathbb{N}$, we define, for a nonzero integer x,

$$\nu_m(x) = \min_{1 \le i \le k} \left[\frac{\nu_{p_i}(x)}{j_i} \right].$$

Theorem 6. *Let α_1 and α_2 be positive rational numbers with $\alpha_1 \neq 1$, b_1 and b_2 be positive integers and set*

$$\Lambda = \alpha_2^{b_2} - \alpha_1^{b_1}.$$

For any set of distinct primes $p_1, \ldots, p_k$ and positive integers $j_1, \ldots, j_k$, we set $m = p_1^{j_1} \cdots p_k^{j_k}$ and suppose that there exists a positive integer g such that for each i, we have

$$\nu_{p_i}(\alpha_1^g - 1) \ge j_i \quad \text{and} \quad \nu_{p_i}(\alpha_2^g - 1) \ge 1, \quad \text{if } p_i \ge 2,$$

and also

$$\nu_{p_i}(\alpha_1^g - 1) \ge 2 \quad \text{and} \quad \nu_{p_i}(\alpha_2^g - 1) \ge 2, \quad \text{if } p_i = 2.$$

Then, if m, b_1 and b_2 are relatively prime, we have

$$\nu_m(\Lambda) \le \frac{66.8\, g}{(\log m)^4} \big(\max\{\log b' + \log(\log m) + 0.64, 4 \log m\}\big)^2 (\log A_1)\, (\log A_2)\,,$$

where

$$b' = \frac{b_1}{\log A_2} + \frac{b_2}{\log A_1} \quad \text{and} \quad \log A_i \ge \max\{h(\alpha_i), \log m\}.$$

For an odd prime number p, the choice of $m = p^E$ illustrates that Theorem 6 is really a generalization of Theorem 5; we have included the latter result for sake of clarity. We stress that α_1 and α_2 need not be multiplicatively independent in Theorem 6; indeed, under this additional assumption, the constant 66.8 in the upper bound for $\nu_m(\Lambda)$ may be improved to 53.6.

4. Proofs

We first state a useful elementary lemma, which will be applied in several places.

Lemma 7. *Let $y \ge 2$ be an integer and q be a prime. Assume that $y^q \equiv 1 \pmod{q}$. Then $y \equiv 1 \pmod{q}$ and $y^q \equiv 1 \pmod{q^2}$. Moreover, if in addition $q \ge 3$, then $\nu_q(y-1) = \nu_q(y^q - 1) - 1$.*

To prove the lemma, observe that the binomial theorem gives

$$\begin{aligned} \frac{y^q - 1}{y - 1} = \frac{((y-1)+1)^q - 1}{y-1} &= (y-1)^{q-1} + \binom{q}{1}(y-1)^{q-2} + \ldots + \binom{q}{q-2}(y-1) + q \\ &= (y-1)^{q-1} + Kq(y-1) + q, \end{aligned}$$

for a suitable positive integer K.

Throughout our proofs, the constants implicit in the Vinogradov symbols $\ll$ and $\gg$ are absolute and positive. We assume that q is suitably large and show how to derive an upper bound for q.

4.1. **Proof of Theorem 2.** We will begin by proving the somewhat general Theorem 2. Let us suppose that there exist positive integers x_1, x_2, a, b, y and q satisfying equation (2), with $\gcd(x_1, x_2) = d > 1$. Without loss of generality, we may assume that $x_1 \geq x_2 \geq 2$, and that $q \gg 1$.

Suppose first that a and b satisfy $x_1^a > x_2^{2b}$. Then

$$0 < q \log y - a \log x_1 < 1 - x_1^a y^{-q} = y^{-q}(x_2^b + 1) < y^{-q}\left(x_1^{a/2} + 1\right) < y^{-q}\left(y^{q/2} + 1\right), \tag{5}$$

whereby, setting $\Lambda = |a \log x_1 - q \log y|$,

$$\log \Lambda \ll -q \log y. \tag{6}$$

We use this inequality, in conjunction with lower bounds for linear forms in two complex logarithms, to determine an upper bound for q.

Let us apply Theorem 4 with $b_1 = q, b_2 = a, \alpha_1 = y$ and $\alpha_2 = x_1$. Here, x_1 and y are multiplicatively independent since the fact that $\gcd(x_1, x_2) > 1$ implies the existence of a prime p dividing x_1 and x_2, but not y (via the equation $x_1^a + x_2^b + 1 = y^q$). Note that the inequality $a \log x_1 < q \log y$ implies that

$$b' \ll \frac{q}{\log x_1}.$$

Applying Theorem 4 to inequality (6), we thus have either

$$\log\left(\frac{q}{\log x_1}\right) \ll 1,$$

or

$$q \log y \ll \log^2\left(\frac{q}{\log x_1}\right) \log y \, \log x_1.$$

We thus conclude, in either case, that $q \ll \log x_1$, where the implied constant is absolute.

Similarly, if we assume that a and b are such that $x_2^b > x_1^{2a}$, we may apply Theorem 4 to the linear form $\Lambda = |b \log x_2 - q \log y|$ to reach a like conclusion. We may therefore suppose that a and b satisfy

$$x_1^{a/2} \leq x_2^b \leq x_1^{2a}. \tag{7}$$

Since $d > 1$, there exists a prime p dividing both x_1 and x_2. By assumption we may suppose that q is coprime to $(p-1)$. To deduce an upper bound upon q is this situation, we will appeal to lower bounds for linear forms in p–adic logarithms. Specifically, we apply Theorem 5 with $\alpha_2 = y, b_2 = q, b_1 = 1$ and either $\alpha_1 = x_1^a + 1$ or $\alpha_1 = x_2^b + 1$, chosen to guarantee that α_1 and y are multiplicatively independent. Note that if all three of $y, x_1^a + 1$ and $x_2^b + 1$ are pairwise multiplicatively dependent,

then, from the equation $x_1^a + x_2^b + 1 = y^q$, every prime dividing y necessarily divides both x_1 and x_2, an immediate contradiction. We thus have $g = 1$ and may assume that $ab \geq 2$. If $q = p$, then, Lemma 7 and $\nu_q(y^q - 1) \geq 1$ give that $\nu_q(y^q - 1) \geq 2$, whereby $\min\{a\,\nu_p(x_1), b\,\nu_p(x_2)\} \geq 2$ and hence we may take $E = \min\{a\,\nu_p(x_1), b\,\nu_p(x_2)\} - 1$. Otherwise, since q fails to divide $(p-1)$, we may choose $E = \min\{a\,\nu_p(x_1), b\,\nu_p(x_2)\}$. If $(p, E) \neq (2, 1)$, then applying Theorem 5 yields

$$E \ll \frac{M^2}{E^3 \log^4 p} \log y \log \alpha_1,$$

where

$$M = \max\left\{\log\left(\frac{1}{\log y} + \frac{q}{\log \alpha_1}\right) + \log(E \log p) + 0.4, 6E \log p, 5\right\}.$$

Notice that (7) implies $\alpha_1 > \frac{1}{2}\, y^{q/2}$, whereby $M = 6E \log p$ since $q \gg 1$, and so

$$E^2 \log^2 p \ll \log y \log \alpha_1 \ll q \log^2 y.$$

From inequality (7), we have

$$a \log x_1 \gg q \log y, \quad b \log x_2 \gg q \log y.$$

Recalling that $x_1 \geq x_2$,

$$E \geq \min\{a, b\} - 1 \gg \frac{q \log y}{\log x_1},$$

and hence $q \ll \log^2 x_1$, as desired. In case $(p, E) = (2, 1)$, then $\min\{a, b\} = 1$ and so (7) and the fact that $x_1 \geq x_2$ imply that $a = 1$, whereby $y^q \leq x_1^2 + x_1 + 1$ and $q \ll \log x_1$. This completes the proof of Theorem 2.

4.2. **Proof of Theorem 3.** Our approach is similar to that of the preceding subsection. Suppose that there exist positive integers $x_1, x_2, x_3, a, b, c, y$ and q satisfying equation (3), with $\gcd(x_1, x_2, x_3) > 1$, and fix a prime $p \mid \gcd(x_1, x_2, x_3)$ such that q does not divide $(p-1)$. Without loss of generality, we may assume that

$$a\,\nu_p(x_1) \geq b\,\nu_p(x_2) \geq c\,\nu_p(x_3), \tag{8}$$

and that $q \gg 1$. Suppose first that a, b and c are such that

$$x_1^a > \left(x_2^b + x_3^c\right)^2.$$

Then we have

$$|a \log x_1 - q \log y| < y^{-q}(x_2^b + x_3^c + 1) < y^{-q}\left(x_1^{a/2} + 1\right) < y^{-q}\left(y^{q/2} + 1\right) < 2\, y^{-q/2}.$$

Applying Theorem 4 and arguing as in Section 4.1 leads, again, to the conclusion that

$$q \ll \log^2 \max\{x_1, x_2, x_3\}.$$

Similarly, we reach an identical conclusion if either $x_2^b > (x_1^a + x_3^c)^2$ or $x_3^c > \left(x_1^a + x_2^b\right)^2$. We will thus suppose that

$$\max\left\{x_1^a, x_2^b, x_3^c\right\} \leq \min\left\{\left(x_1^a + x_2^b\right)^2, (x_1^a + x_3^c)^2, \left(x_2^b + x_3^c\right)^2\right\}. \tag{9}$$

Notice that inequality (9) implies that

$$\max\{b, c\} \gg \frac{q \log y}{\log \max\{x_1, x_2, x_3\}},$$

where the implicit constant is absolute, whereby, from (8),

$$b \gg \frac{q \log y}{\log^2 \max\{x_1, x_2, x_3\}}. \tag{10}$$

We turn now to consideration of non-Archimedean valuations. On the one hand, we have

$$\nu_p(y^q - (x_3^c + 1)) \geq b\, \nu_p(x_2) \geq b \gg \frac{q \log y}{\log^2 \max\{x_1, x_2, x_3\}}.$$

Let $\ell = \nu_p(x_3)$. By (10) we may assume that $b\nu_p(x_2) \geq 2$. If $p = q$, it then follows from Lemma 7 that $y^p \equiv 1 \pmod{p^2}$, whence $\ell c \geq 2$ and $y \equiv 1 \pmod{p^{\ell c - 1}}$. Otherwise, from the fact that q does not divide $(p-1)$, it follows that $y \equiv 1 \pmod{p^{\ell c}}$. Since we do not know *a priori* whether or not the integers y and $x_3^c + 1$ are multiplicatively independent, it is more convenient to apply Theorem 6 with $m = p^{\ell c - 1}$ and $g = 1$ or 2, than Theorem 5 with $E = \ell c - 1$. We conclude that

$$\nu_p(y^q - (x_3^c + 1)) \ll \ell c \nu_m(y^q - (x_3^c + 1)) \ll \frac{1}{\ell^3 c^2 \log^4 p} M^2 \log y \log x_3,$$

where

$$M = \max\left\{\log\left(\frac{1}{\log y} + \frac{q}{\ell c \log x_3}\right) + \log(\ell c \log p) + 0.64, 4\ell c \log p\right\} \ll \max\{\log q, \ell c \log p\}.$$

If $q > p^{\ell c}$, then we have

$$\frac{q}{\log^2 q} \ll \frac{\log x_3 \log^2 \max\{x_1, x_2, x_3\}}{\ell^3 c^2 \log^4 p} \ll \log^3 \max\{x_1, x_2, x_3\},$$

and so $q \ll \log^4 \max\{x_1, x_2, x_3\}$. If, on the other hand, $q \leq p^{\ell c}$, then

$$q \ll \frac{\log x_3 \log^2 \max\{x_1, x_2, x_3\}}{\ell \log^2 p} \ll \log^3 \max\{x_1, x_2, x_3\}.$$

This concludes the proof of Theorem 3.

4.3. **Proof of Theorem 1.** In this section, we will show that the upper bound upon q inherent in the equation $x_1^a + x_2^b + 1 = y^q$ can be made (essentially) absolute, in the case $x_1 = x_2 = x$. Here, we will appeal to lower bounds for linear forms in m-adic logarithms, due to the second author [3].

Since $a > b$, we have $y^q \equiv 1 \pmod{x^b}$. If q divides x, we deduce from Lemma 7 that $y^q \equiv 1 \pmod{q^2}$. Furthermore, denoting by q^ℓ the largest power of q dividing x, this implies that $b\ell = \nu_q(y^q - 1) \geq 2$ and that $y \equiv 1 \pmod{q^{b\ell - 1}}$, by Lemma 7 again and using that $q \geq 3$. Since we suppose q to be coprime to $\phi(x)$ (actually, assuming that q is coprime to $\phi(x/q^{\max\{0, \nu_q(x)-1\}})$ is sufficient), we derive that $y \equiv 1 \pmod{x^b/q}$ if q divides x, while, otherwise, $y \equiv 1 \pmod{x^b}$. In any case, we may conclude that $\log y > b \log x - \log q \geq (b \log x)/2$.

We apply Theorem 6 with $\alpha_1 = y, \alpha_2 = x^b + 1, b_1 = q, b_2 = 1$ and $m = x^b/q$ or $m = x^b$, depending on whether q divides x or not. Observe that $\log m \geq (b \log x)/2$ in either case. If x is odd or if $b \geq 2$, we may choose $g = 1$; otherwise, take $g = 2$. Since $\log y > (b \log x)/2$, Theorem 6 implies that

$$\nu_m \left(y^q - (x^b + 1)\right) \ll \frac{\log^2 \left(\frac{2q}{b \log x}\right) \log y}{b^3 \log^3 x} + \frac{\log y}{b \log x}.$$

Considering the cases where $\frac{q}{b \log x} < x^b$ and $\frac{q}{b \log x} \geq x^b$, separately, the inequality

$$\nu_m \left(y^q - (x^b + 1)\right) = \left[\frac{a}{b}\right] \gg \frac{q \log y}{b \log x}$$

implies, in either situation, that $q \ll 1$. This completes the proof of Theorem 1.

5. Concluding remarks

Remark 1. The techniques of the preceding sections may be readily modified to treat equations of the shape (by way of example)

$$d_1 x_1^a + d_2 x_2^b + d_3 x_3^c + 1 = y^q,$$

for where d_1, d_2 and d_3 are fixed integers and, as previously, x_1, x_2 and x_3 share a nontrivial common factor. To make such statements explicit requires recourse to lower bounds for linear forms in three logarithms; the corresponding theory is somewhat less well-developed than that for two logarithms.

Remark 2. We are unable to bound q in general, for the equation

$$x^a + x^b + x^c + x^d + 1 = y^q,$$

where $a > b > c > d > 0$. It is possible to do so, however, under the additional assumption that b is not too large compare to c, say $c \geq \kappa b$ for κ be a positive real number. Indeed, assuming that q is

coprime to $\phi(x)$, Theorem 4 again implies that $q \ll 1$ if $a \geq 2b$. Then, we may assume that $b \geq a/2$ and proceed exactly as in Section **??**, using that $y \equiv 1 \pmod{x^d}$. Since $y > x^d$, we find that

$$\left[\frac{c}{d}\right] \ll \frac{\log y}{d \log x} + \frac{\log^2\left(\frac{q}{d \log x}\right) \log y}{d^3 \log^3 x}.$$

From the inequality $c \geq \kappa b$, we thus obtain $\kappa q \log y \ll c \log x$ and hence conclude that q is bounded by a constant depending only on κ. The same conclusion remains true for the more general equation

$$x^{a_1} + x^{a_2} + \ldots + x^{a_{k-1}} + x^{a_k} + 1 = y^q,$$

where q is coprime to $\phi(x)$ and $a_1 > a_2 > \ldots > a_k > 0$ are such that $a_{k-1} \geq \kappa a_2$. The upper bound obtained again depends on κ.

Remark 3. We briefly discuss a related question. Let $x, y \geq 2$ be integers. In 1973, Senge and Straus [20] proved that the number of integers, the sum of whose digits in each of the bases x and y lies below a fixed bound, is finite if, and only if, x and y are multiplicatively independent. Their proof rests on the Thue–Siegel–Roth theorem and, hence, is ineffective. Using Baker's theory of linear forms in logarithms, in 1980 Stewart [21] succeeded in establishing an effective version of Senge and Straus' theorem. He showed that if x and y are multiplicatively independent, then, for every $c \geq 1$, each integer $m > 25$ whose sum of digits in base x as well as in base y is bounded by c satisfies

$$\frac{\log \log m}{\log \log \log m + c_1} < 2c + 1,$$

where c_1 is a positive constant which is effectively computable in terms of x and y only (see also [17]).

This means that a positive integer cannot have simultaneously very few digits in base x and in base y, when x and y are multiplicatively independent. Furthermore, for any given positive integer c, one is able, in principle, to determine the (finite) set of positive integers having no more than c digits in base x and in base y. Examples of complete resolution of the corresponding equations are given in, for instance, [4] and [23]. An open problem of a similar flavour is a conjecture of Erdős to the effect that there should be at most finitely many powers of 2 whose ternary representation contains only the digits 0 and 1. This appears to be well out of reach, at present; the interested reader is directed to the paper of Lagarias [14].

Acknowledgements: The authors would like to thank Umberto Zannier for providing us with his preprint [12]. Thanks are also due to the referee for careful reading and noticing several inaccuracies and obscure points.

References

[1] M. A. Bennett, Y. Bugeaud, and M. Mignotte, *Perfect powers with few binary digits and related Diophantine problems, II*. Preprint.

[2] Y. Bugeaud, *Linear forms in p-adic logarithms and the Diophantine equation* $(x^n - 1)/(x - 1) = y^q$, Math. Proc. Cambridge Philos. Soc. 127 (1999), 373–381.

[3] Y. Bugeaud, *Linear forms in two m-adic logarithms and applications to Diophantine problems*, Compositio Math. 132 (2002), 137–158.

[4] Y. Bugeaud, M. Cipu and M. Mignotte, *On the representation of Fibonacci and Lucas numbers in an integer base*, Ann. Sci. Math. Québec. To appear.

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

[6] Y. Bugeaud et M. Mignotte, *Sur l'équation diophantienne* $\frac{x^n-1}{x-1} = y^q$*, II*, C.R.A.S. Série I, t. 328 (1999).

[7] Y. Bugeaud et M. Mignotte, *L'équation de Nagell–Ljunggren* $\frac{x^n-1}{x-1} = y^q$, Enseign. Math. 48 (2002), 147–168.

[8] Y. Bugeaud and M. Mignotte, *On the Diophantine equation* $\frac{x^n-1}{x-1} = y^q$ *with negative* x. In: Number theory for the millennium, I (Urbana, IL, 2000), 145–151, A K Peters, Natick, MA, 2002.

[9] Y. Bugeaud, M. Mignotte and Y. Roy, *On the Diophantine equation* $(x^n - 1)/(x - 1) = y^q$, Pacific J. Math. 193 (2000), 257–268.

[10] P. Corvaja and U. Zannier, *On the Diophantine equation* $f(a^m, y) = b^n$, Acta Arith. 94 (2000), 25–40.

[11] P. Corvaja and U. Zannier, *Application of the Subspace theorem to certain Diophantine problems*. In: Diophantine Approximation, H. E Schlickewei et al., Editors, Springer-Verlag 2008, 161–174.

[12] P. Corvaja and U. Zannier, *Finiteness of odd perfect powers with four nonzero binary digits*. Preprint.

[13] M. Hindry and J. Silverman, Diophantine Geometry, An Introduction, Springer Verlag GTM 201, 2001.

[14] J. Lagarias, *Ternary expansions of powers of* 2, J. London Math. Soc. 79 (2009), 562–588.

[15] M. Laurent, *Linear forms in two logarithms and interpolation determinants. II.*, Acta Arith. 133 (2008), 325–348.

[16] F. Luca, *The Diophantine equation* $x^2 = p^a \pm p^b + 1$, Acta Arith. 112 (2004), 87–101.

[17] M. Mignotte, *Sur les entiers qui s'écrivent simplement en différentes bases*, European J. Combin. 9 (1988), 307–316.

[18] P. Mihăilescu, *Primary cyclotomic units and a proof of Catalan's conjecture*, J. Reine Angew. Math. 572 (2004), 167–195.

[19] R. Scott, *Elementary treatment of* $p^a \pm p^b + 1 = x^2$. Available online at the homepage of Robert Styer : http://www41.homepage.villanova.edu/robert.styer/ReeseScott/index.htm

[20] H. G. Senge and E. G. Straus, PV*-numbers and sets of multiplicity*, Period. Math. Hungar. 3 (1973), 93–100.

[21] C. L. Stewart, *On the representation of an integer in two different bases*, J. Reine Angew. Math. 319 (1980), 63–72.

[22] L. Szalay, *The equations* $2^n \pm 2^m \pm 2^l = z^2$, Indag. Math. (N.S.) 13 (2002), 131–142.

[23] T. Yamada, *On the Diophantine equation* $x^m = y^{n_1} + y^{n_2} + \cdots + y^{n_k}$, Glasg. Math. J. 51 (2009), 143–148.

Michael A. Bennett, Department of Mathematics, University of British Columbia, Vancouver, British Columbia, V6T 1Z2, CANADA

E-mail address: bennett@math.ubc.ca

Yann Bugeaud, Mathématiques, Université de Strasbourg, 7, rue René Descartes, 67084 Strasbourg, FRANCE

E-mail address: bugeaud@math.unistra.fr

Maurice Mignotte, Mathématiques, Université de Strasbourg, 7, rue René Descartes, 67084 Strasbourg, FRANCE

E-mail address: mignotte@math.unistra.fr