

On the digital representation of smooth numbers

Yann BUGEAUD and Hajime KANEKO

Abstract. *Let $b \geq 2$ be an integer. Among other results, we establish, in a quantitative form, that any sufficiently large integer which is not a multiple of b cannot simultaneously be divisible only by very small primes and have very few nonzero digits in its representation in base b .*

1. Introduction and results

Let a, b be positive, multiplicatively independent integers. Stewart [10] established that, for every sufficiently large integer n , the representation of a^n in base b has more than $(\log n)/(2 \log \log n)$ nonzero digits. His proof rests on a subtle application of Baker's theory of linear forms in complex logarithms of algebraic numbers. This result addresses a very special case of the following general (and left intentionally vague) question, which was introduced and discussed in [5]:

Do there exist arbitrarily large integers which have only small prime factors and, at the same time, few nonzero digits in their representation in some integer base?

The expected answer is *no* and a very modest step in this direction has been made in [5], by using a combination of estimates for linear forms in complex and p -adic logarithms. In the present work, we considerably extend Corollary 1.3 of [5] and, more generally, we show in a quantitative form that the maximum of the greatest prime factor of an integer n and the number of nonzero digits in its representation in a given integer base tends to infinity as n tends to infinity.

Throughout this note, b always denotes an integer at least equal to 2. Following [5], for an integer $k \geq 2$, we denote by $(u_j^{(k)})_{j \geq 1}$ the sequence, arranged in increasing order, of all positive integers which are not divisible by b and have at most k nonzero digits in their representation in base b . Said differently, $(u_j^{(k)})_{j \geq 1}$ is the ordered sequence composed of the integers $1, 2, \dots, b-1$ and those of the form

$$d_k b^{n_k} + \dots + d_2 b^{n_2} + d_1, \quad n_k > \dots > n_2 > 0, \quad d_1, \dots, d_k \in \{0, 1, \dots, b-1\}, \quad d_1 d_k \neq 0.$$

2010 *Mathematics Subject Classification* : 11A63, 11J86.

We stress that, for the questions investigated in the present note, it is natural to restrict our attention to integers not divisible by b . Obviously, the sequence $(u_j^{(k)})_{j \geq 1}$ depends on b , but, to shorten the notation, we have decided not to mention this dependence.

Theorem 1.1 of [5] implies that the greatest prime factor of $u_j^{(k)}$ tends to infinity as j tends to infinity. Its proof rests on the Schmidt Subspace Theorem and does not allow us to derive an estimate for the speed of divergence. Such an estimate has been established in [5], but only for $k \leq 3$. Following the proof of our main result (Theorem 1.2 below), we are able to extend this estimate to arbitrary integers k .

For a positive integer n , let us denote by $P[n]$ its greatest prime factor and by $\omega(n)$ the number of its distinct prime factors, with the convention that $P[1] = 1$. A positive real number B being given, a positive integer n is called B -smooth if $P[n] \leq B$.

Theorem 1.1. *Let $b \geq 2, k \geq 3$ be integers. Let ε be a positive real number. Then, there exists an effectively computable positive number j_0 , depending only on b, k , and ε , such that*

$$P[u_j^{(k)}] > \left(\frac{1}{k-2} - \varepsilon \right) \log \log u_j^{(k)} \frac{\log \log \log u_j^{(k)}}{\log \log \log \log u_j^{(k)}}, \quad \text{for } j > j_0.$$

In particular, there exists an effectively computable positive integer n_0 , depending only on b, k , and ε , such that any integer $n > n_0$ which is not divisible by b and is

$$\left(\frac{1}{k-2} - \varepsilon \right) (\log \log n) \frac{\log \log \log n}{\log \log \log \log n} \text{-smooth}$$

has at least $k + 1$ nonzero digits in its b -ary representation.

Our main result asserts that, given an integer $b \geq 2$, if the integer n is sufficiently large, then its greatest prime factor and the number of nonzero digits in its representation in base b cannot be simultaneously small.

Theorem 1.2. *Let $b \geq 2$ and $k \geq 2$ be integers. There exist an effectively computable real number c , depending at most on b , and an effectively computable, absolute real number C such that every sufficiently large positive integer n , which is not divisible by b and whose representation in base b has k nonzero digits, satisfies*

$$\frac{\log \log n}{k} \leq c + \log k + \omega(n)(C + \log \log P[n]) + \log \log(k \log P[n]).$$

Taking $k = 3$ in Theorem 1.2, we get the second assertion of Theorem 1.3 of [5]. The main ingredients for the proofs of both theorems are estimates for linear forms in complex and p -adic logarithms of algebraic numbers. The novelty in the present note is a more subtle use of estimates for linear forms in p -adic logarithms, where p is a prime divisor of the base b . With our new approach, the number k of nonzero digits need not to be fixed and can be allowed to depend on n , provided that it is rather small compared to n .

Several easy consequences of the proof of Theorem 1.2 are pointed out below. We extend the definition of the sequences $(u_j^{(k)})_{j \geq 1}$ as follows. For a positive real valued function f defined over the set of positive integers, we let $(u_j^{(f)})_{j \geq 1}$ be the sequence, arranged in increasing order, of all positive integers n which are not divisible by b and have at most $f(n)$ nonzero digits in their representation in base b .

Theorem 1.3. *Let $b \geq 2$ be an integer. Let f be a positive real valued function defined over the set of positive integers such that*

$$\lim_{u \rightarrow +\infty} f(u) = +\infty.$$

Assume that there exists a real number δ satisfying $0 < \delta < 1$ and

$$f(u) \leq (1 - \delta) \frac{\log \log u}{\log \log \log u}, \quad (1.1)$$

for any sufficiently large u , and set

$$\Psi_f(u) := \frac{\log \log u}{f(u)}, \quad \text{for } u \geq 3.$$

Then, for an arbitrary positive real number ε , we have

$$P[u_j^{(f)}] > (\delta_0 - \varepsilon) \Psi_f(u_j^{(f)}) \frac{\log \Psi_f(u_j^{(f)})}{\log \log \Psi_f(u_j^{(f)})}, \quad (1.2)$$

for any sufficiently large integer j , where

$$\delta_0 = \sup \left\{ \delta > 0 : f(u) \leq (1 - \delta) \frac{\log \log u}{\log \log \log u} \text{ for every large integer } u \right\}.$$

We gather in the next statement three immediate consequences of Theorem 1.3 applied with an appropriate function f .

Corollary 1.4. *Let $b \geq 2$ be an integer. There exists an effectively computable positive integer n_0 , depending only on b , such that any integer $n > n_0$ which is not divisible by b satisfies the following three assertions. If n is*

$$\frac{\log \log n}{2 \log \log \log n} \text{-smooth, then } n \text{ has at least } \log \log \log n$$

nonzero digits in its representation in base b . If n is

$$\sqrt{\log \log n \frac{\log \log \log n}{\log \log \log \log n}} \text{-smooth, then } n \text{ has at least } \frac{1}{3} \sqrt{\log \log n \frac{\log \log \log n}{\log \log \log \log n}}$$

nonzero digits in its representation in base b . If n is

$$\frac{1}{2} \log \log \log n \frac{\log \log \log \log n}{\log \log \log \log \log n} \text{-smooth, then } n \text{ has at least } \frac{\log \log n}{2 \log \log \log n}$$

nonzero digits in its representation in base b .

Let S be a finite, non-empty set of prime numbers. A rational integer is an integral S -unit if all its prime factors belong to S . We deduce from Theorem 1.2 lower bounds for the number of nonzero digits in the representation of integral S -units in an integer base.

Corollary 1.5. *Let $b \geq 2$ be an integer. Let S be a finite set of prime numbers. Then, for any positive real number ε , there exists an effectively computable positive integer n_0 , depending only on b, S , and ε , such that any integral S -unit $n \geq n_0$ which is not divisible by b has more than*

$$(1 - \varepsilon) \frac{\log \log n}{\log \log \log n}$$

nonzero digits in its representation in base b .

Let $a \geq 2, b \geq 2$ be coprime integers. By taking for S the set of prime divisors of a , Corollary 1.5 implies Stewart's result mentioned in the introduction (for the case where a and b are multiplicatively independent and not coprime, the proof of Corollary 1.5 can be easily adapted) and both proofs are different. Observe, however, that Stewart obtained in [10] a more general result, namely that, for any multiplicatively independent positive integers b and b' and any sufficiently large integer n , the number of nonzero digits in the representation of n in base b plus the number of nonzero digits in the representation of n in base b' exceeds $(\log \log n)/(2 \log \log \log n)$.

Our results are established in Section 3, by means of lower estimates for linear forms in logarithms gathered in Section 2. We postpone to Section 4 comments and remarks.

2. Lower estimates for linear forms in logarithms

The first assertion of Theorem 2.1 is an immediate consequence of a theorem of Matveev [8]. The second one is a slight simplification of the estimate given on page 190 of Yu's paper [12]. For a prime number p and a nonzero rational number z we denote by $v_p(z)$ the exponent of p in the decomposition of z in product of prime factors.

Theorem 2.1. *Let $n \geq 2$ be an integer. Let $x_1/y_1, \dots, x_n/y_n$ be nonzero rational numbers. Let b_1, \dots, b_n be integers such that $(x_1/y_1)^{b_1} \cdots (x_n/y_n)^{b_n} \neq 1$. Let A_1, \dots, A_n be real numbers with*

$$A_i \geq \max\{|x_i|, |y_i|, e\}, \quad 1 \leq i \leq n.$$

Set $B = \max\{3, |b_1|, \dots, |b_n|\}$. Then, we have

$$\log \left| \left(\frac{x_1}{y_1} \right)^{b_1} \cdots \left(\frac{x_n}{y_n} \right)^{b_n} - 1 \right| > -8 \times 30^{n+3} n^{9/2} \log(eB) \log A_1 \cdots \log A_n. \quad (2.1)$$

Let p be a prime number. Then, we have

$$v_p \left(\left(\frac{x_1}{y_1} \right)^{b_1} \cdots \left(\frac{x_n}{y_n} \right)^{b_n} - 1 \right) < (16e)^{2(n+1)} n^{5/2} (\log(2n))^2 \frac{p}{(\log p)^2} \log A_1 \cdots \log A_n \log B. \quad (2.2)$$

3. Proofs

Below, the constants c_1, c_2, \dots are effectively computable and depend at most on b and the constants C_1, C_2, \dots are absolute and effectively computable. Let N be a positive integer and k the number of nonzero digits in its representation in base b . We assume that b does not divide N , thus $k \geq 2$ and we write

$$N =: d_k b^{n_k} + \dots + d_2 b^{n_2} + d_1 b^{n_1},$$

where

$$n_k > \dots > n_2 > n_1 = 0, \quad d_1, \dots, d_k \in \{1, \dots, b-1\}.$$

Let q_1, \dots, q_s denote distinct prime numbers written in increasing order such that there exist non-negative integers r_1, \dots, r_s with

$$N = q_1^{r_1} \dots q_s^{r_s}.$$

Observe that

$$b^{n_k} \leq N < b^{n_k+1}. \quad (3.1)$$

Lemma 3.1. *Keep the above notation and set $k^* := \max\{k-2, 1\}$. If*

$$\log N \geq 2(\log b) \left(\frac{8 \log b}{\log 2} \right)^k, \quad (3.2)$$

then we have

$$n_k \leq \left(c_1 C_1^s k^* \left(\prod_{i=1}^s \log q_i \right) \log(k \log q_s) \right)^{k^*}. \quad (3.3)$$

Proof. First we assume that $n_k \geq 2n_{k-1}$. This covers the case $k = 2$. Since

$$\begin{aligned} \Lambda_a &:= \left| \left(\prod_{i=1}^s q_i^{r_i} \right) d_k^{-1} b^{-n_k} - 1 \right| = d_k^{-1} b^{-n_k} \sum_{h=1}^{k-1} d_h b^{n_h} \\ &\leq b^{1+n_{k-1}-n_k} \leq b^{-(n_k-2)/2}, \end{aligned}$$

we get

$$\log \Lambda_a \leq -\left(\frac{n_k}{2} - 1 \right) \log b. \quad (3.4)$$

Since $r_j \log q_j \leq (n_k + 1) \log b$ for $j = 1, \dots, s$, we deduce from (2.1) that

$$\log \Lambda_a \geq -c_2 C_2^s (\log q_1) \dots (\log q_s) (\log n_k). \quad (3.5)$$

Combining (3.4) and (3.5), we obtain

$$n_k \leq c_3 C_3^s \left(\prod_{i=1}^s \log q_i \right) (\log \log q_s),$$

which implies (3.3).

Now, we assume that $n_k < 2n_{k-1}$. In particular, we have $k \geq 3$. If there exists an integer j with $1 \leq j \leq k-3$ and $n_{1+j} \geq n_k^{j/(k-2)}$, then put

$$\ell := \min\{j : 1 \leq j \leq k-3, n_{1+j} \geq n_k^{j/(k-2)}\}.$$

Otherwise, set $\ell := k-2$. We see that

$$n_{\ell+1} \geq \frac{1}{2}n_k^{\ell/(k-2)} \quad \text{and} \quad n_\ell \leq n_k^{(\ell-1)/(k-2)}. \quad (3.6)$$

Let p be the smallest prime divisor of b . Put

$$\Lambda_u := \left(\prod_{i=1}^s q_i^{r_i} \right) \left(\sum_{h=1}^{\ell} d_h b^{n_h} \right)^{-1} - 1 = \left(\sum_{h=\ell+1}^k d_h b^{n_h} \right) \left(\sum_{h=1}^{\ell} d_h b^{n_h} \right)^{-1}.$$

We get by (3.6), (3.1), and (3.2) that

$$\begin{aligned} v_p(\Lambda_u) &\geq n_{\ell+1} - \frac{\log b^{1+n_\ell}}{\log p} \\ &\geq \frac{1}{2}n_k^{\ell/(k-2)} - \left(1 + n_k^{(\ell-1)/(k-2)}\right) \frac{\log b}{\log 2} \\ &\geq \frac{1}{2}n_k^{\ell/(k-2)} - 2n_k^{\ell/(k-2)} \frac{\log b}{n_k^{1/(k-2)} \log 2} \geq \frac{1}{4}n_k^{\ell/(k-2)}. \end{aligned} \quad (3.7)$$

We deduce from (2.2) and (3.6) that

$$v_p(\Lambda_u) \leq c_4 C_4^s (\log q_1) \cdots (\log q_s) n_k^{(\ell-1)/(k-2)} \log n_k. \quad (3.8)$$

By combining (3.7) and (3.8), we get

$$n_k^{1/(k-2)} \leq c_5 C_5^s (\log q_1) \cdots (\log q_s) (k-2) \log(n_k^{1/(k-2)}),$$

which implies (3.3) and completes the proof of Lemma 3.1. \square

Proof of Theorem 1.1.

We keep the above notation. In particular, N denotes an integer not divisible by b and with exactly k nonzero digits in its representation in base b . In view of [5], we assume that $k \geq 3$, thus $k^* = k-2$. Note that (3.2) holds if N is large enough. Then, we deduce from (3.1) and (3.3) that

$$\frac{\log \log N}{k-2} \leq c_6 + C_6 s + \log(k-2) + \sum_{i=1}^s \log \log q_i + \log \log(k \log q_s). \quad (3.9)$$

In particular, denoting by p_j the j -th prime number for $j \geq 1$ and defining s by $p_s = P[N]$, inequality (3.9) applied with $q_j = p_j$ for $j = 1, \dots, s$ shows that

$$\frac{\log \log N}{k-2} \leq c_6 + C_6 s + \log(k-2) + s \log \log P[N] + \log \log(k \log P[N]). \quad (3.10)$$

Let ε be a positive real number. By the Prime Number Theorem, there exists an effectively computable integer $s_0(\varepsilon)$, depending only on ε , such that, if $s \geq s_0(\varepsilon)$, then

$$s < (1 + \varepsilon) \frac{P[N]}{\log P[N]}. \quad (3.11)$$

For $s < s_0(\varepsilon/(2k-4))$, we derive from (3.10) an upper bound for N in terms of b, k , and ε . For $s \geq s_0(\varepsilon/(2k-4))$, it follows from (3.11) and the Prime Number Theorem that

$$\log \log N \leq (k-2 + \varepsilon) P[N] \frac{\log \log P[N]}{\log P[N]},$$

provided that N is sufficiently large in terms of b, k , and ε . This implies Theorem 1.1. \square

Proof of Theorem 1.2.

We assume that q_1, \dots, q_s are the prime divisors of N , thus in particular we have $s = \omega(N)$. If (3.2) is not satisfied, then $\log \log N \leq c_7 k$. Otherwise, by taking the logarithms of both sides of (3.3) and using (3.1), we get

$$\frac{\log \log N}{k} \leq c_8 + C_7 \omega(N) + \log k + \omega(N) \log \log P[N] + \log \log(k \log P[N]).$$

This establishes Theorem 1.2.

Proof of Theorem 1.3.

We argue as in the proof of Theorem 1.1. Let ε be a positive real number with $\varepsilon < \delta_0/3$. Suppose that j is sufficiently large and set $N := u_j^{(f)}$. It follows from (1.1) that (3.2) holds if N is large enough. Then, (3.3) holds with an integer k at most equal to $f(N)$. By using (3.10), (3.11) and the Prime Number Theorem, we get

$$\log \log N \leq (1 + \varepsilon) f(N) \left(\log f(N) + P[N] \frac{\log \log P[N]}{\log P[N]} \right).$$

It then follows from the definition of the positive real number δ_0 that

$$\log \log N < (1 - \delta_0 + 3\varepsilon) \log \log N + (1 + \varepsilon) f(N) P[N] \frac{\log \log P[N]}{\log P[N]},$$

hence

$$(\delta_0 - 3\varepsilon) \Psi_f(N) < (1 + \varepsilon) P[N] \frac{\log \log P[N]}{\log P[N]}.$$

Therefore, we have proved (1.2). \square

4. Additional remarks

Remark 4.1. Arguing as Stewart did in [11], we can derive a lower bound for $Q[u_j^{(k)}]$, where $Q[n]$ denotes the greatest square-free divisor of a positive integer n , similar to the lower bound for $Q[u_j^{(3)}]$ given in Theorem 4.1 of [5].

Remark 4.2. Let a, b be integers such that $a > b > 1$ and $\gcd(a, b) \geq 2$. Perfect powers in the double sequence $(a^m + b^n + 1)_{m, n \geq 1}$ have been considered in [2, 3, 5, 7]. The method of the proof of Theorem 1.2 allows us to establish the following extension of Theorem 4.3 of [5].

Theorem 4.1. *Let $k \geq 2$ and a_1, \dots, a_k be positive integers with $\gcd(a_1, \dots, a_k) \geq 2$. Let $\mathbf{v} = (v_j)_{j \geq 1}$ denote the increasing sequence composed of all the integers of the form $a_1^{n_1} + \dots + a_k^{n_k} + 1$, with $n_1, \dots, n_k \geq 1$. Then, for every positive ε , we have*

$$P[v_j] > \left(\frac{1}{k-1} - \varepsilon \right) \log \log v_j \frac{\log \log \log v_j}{\log \log \log \log v_j},$$

when j exceeds some effectively computable constant depending only on a_1, \dots, a_k , and ε .

Remark 4.3. Let n be a positive integer. Let $S = \{q_1, \dots, q_s\}$ be a finite, non-empty set of distinct prime numbers. Write $n = q_1^{r_1} \cdots q_s^{r_s} M$, where r_1, \dots, r_s are non-negative integers and M is an integer relatively prime to $q_1 \cdots q_s$. We define the S -part $[n]_S$ of n by

$$[n]_S := q_1^{r_1} \cdots q_s^{r_s}.$$

Theorem 1.1 of [5] asserts that, for every $k \geq 2$ and every positive real number ε , we have

$$[u_j^{(k)}]_S < (u_j^{(k)})^\varepsilon,$$

for every sufficiently large integer j . This implies that (and is a much stronger statement than) the greatest prime factor of $u_j^{(k)}$ tends to infinity as j tends to infinity. The proof uses the Schmidt Subspace Theorem and it is here essential that k is fixed. Moreover, this is an ineffective result.

The main goal of [5] was to establish an effective improvement of the trivial estimate $[u_j^{(3)}]_S \leq u_j^{(3)}$ of the form $[u_j^{(3)}]_S \leq (u_j^{(3)})^{1-\delta}$, for a small positive real number δ and for j sufficiently large. A key tool was a stronger version of Theorem 2.1 in the special case where $|b_n|$ is small. Unfortunately, for $k > 3$, the method of the proof of Theorem 1.2 does not seem to combine well with this stronger version of Theorem 2.1 to get an analogous result. We are only able to establish that, for any fixed integer $k \geq 4$ and any given positive real number ε , the upper bound

$$[u_j^{(k)}]_S \leq u_j^{(k)} \exp(-(\log u_j^{(k)})^{(1-\varepsilon)/(k-2)})$$

holds for every sufficiently large integer j .

Remark 4.4. Instead of considering the number of nonzero digits in the representation of an integer in an integer base, we can focus on the number of blocks composed of the same digit in this representation, a quantity introduced by Blecksmith, Filaseta, and Nicol [4]; see also [1, 6]. A straightforward adaptation of our proofs shows that analogous versions of Theorems 1.1 to 1.3 hold with ‘number of nonzero digits’ replaced by ‘number of blocks’. We omit the details.

Remark 4.5. In the opposite direction of our results, it does not seem to be easy to confirm the existence of arbitrarily large integers with few digits in their representation in some integer base and only small prime divisors. A construction given in Theorem 13 of [9] and based on cyclotomic polynomials shows that there exist an absolute, positive real number c and arbitrarily large integers N of the form $2^n + 1$ such that

$$P[N] \leq N^{c/\log \log \log N}.$$

Acknowledgements. The second author was supported by JSPS KAKENHI Grant Number 15K17505.

References

- [1] G. Barat, R. F. Tichy, and R. Tijdeman, *Digital blocks in linear numeration systems*. In: Number theory in progress, Vol. 2 (Zakopane-Kościełisko, 1997), 607–631, de Gruyter, Berlin, 1999.
- [2] M. A. Bennett, Y. Bugeaud, and M. Mignotte, *Perfect powers with few binary digits and related Diophantine problems, II*, Math. Proc. Cambridge Philos. Soc. 153 (2012), 525–540.
- [3] M. A. Bennett, Y. Bugeaud, and M. Mignotte, *Perfect powers with few binary digits and related Diophantine problems*, Ann. Sc. Norm. Super. Pisa Cl. Sci. 12 (2013), 525–540.
- [4] R. Blecksmith, M. Filaseta, and C. Nicol, *A result on the digits of a^n* , Acta Arith. 64 (1993), 331–339.
- [5] Y. Bugeaud, *On the digital representation of integers with bounded prime factors*, Osaka J. Math. To appear.
- [6] Y. Bugeaud, M. Cipu, and M. Mignotte, *On the representation of Fibonacci and Lucas numbers in an integer base*, Ann. Math. Qué. 37 (2013), 31–43.
- [7] P. Corvaja and U. Zannier, *S-unit points on analytic hypersurfaces*, Ann. Sci. École Norm. Sup. 38 (2005), 76–92.

- [8] E. M. Matveev, *An explicit lower bound for a homogeneous rational linear form in logarithms of algebraic numbers. II*, Izv. Ross. Acad. Nauk Ser. Mat. 64 (2000), 125–180 (in Russian); English translation in Izv. Math. 64 (2000), 1217–1269.
- [9] A. Schinzel, *On two theorems of Gelfond and some of their applications*, Acta Arith. 13 (1967), 177–236.
- [10] C. L. Stewart, *On the representation of an integer in two different bases*, J. reine angew. Math. 319 (1980), 63–72.
- [11] C. L. Stewart, *On the greatest square-free factor of terms of a linear recurrence sequence*. In: Diophantine equations, 257–264, Tata Inst. Fund. Res. Stud. Math., 20, Tata Inst. Fund. Res., Mumbai, 2008.
- [12] K. Yu, *p -adic logarithmic forms and group varieties. III*, Forum Math. 19 (2007), 187–280.

Yann Bugeaud
 Institut de Recherche Mathématique Avancée, U.M.R. 7501
 Université de Strasbourg et C.N.R.S.
 7, rue René Descartes
 67084 Strasbourg, FRANCE
 e-mail : bugeaud@math.unistra.fr

Hajime Kaneko
 Institute of Mathematics, University of Tsukuba, 1-1-1, Tennodai
 Tsukuba, Ibaraki, 305-8571, JAPAN
 Center for Integrated Research in Fundamental Science and Engineering (CiRfSE)
 University of Tsukuba, 1-1-1, Tennodai
 Tsukuba, Ibaraki, 305-8571, JAPAN
 e-mail: kanekoha@math.tsukuba.ac.jp