Acta Math. Hungar., **143 (1)** (2014), 192–200
DOI: 10.1007/s10474-014-0394-4
First published online February 9, 2014

ON SIMULTANEOUS DIGITAL EXPANSIONS OF POLYNOMIAL VALUES

M. MADRITSCH[1,2] and T. STOLL[1,2,*,†]

[1]Université de Lorraine, Institut Elie Cartan de Lorraine, UMR 7502, Vandoeuvre-lès-Nancy, F-54506, France

[2]CNRS, Institut Elie Cartan de Lorraine, UMR 7502, Vandoeuvre-lès-Nancy, F-54506, France
e-mails: manfred.madritsch@univ-lorraine.fr, thomas.stoll@univ-lorraine.fr

(Received June 30, 2013; revised September 6, 2013; accepted September 16, 2013)

Abstract. Let s_q denote the q-ary sum-of-digits function and let $P_1(X)$, $P_2(X) \in \mathbb{Z}[X]$ with $P_1(\mathbb{N}), P_2(\mathbb{N}) \subset \mathbb{N}$ be polynomials of degree $h, l \geqq 1$, $h \neq l$, respectively. In this note we show that $(s_q(P_1(n))/s_q(P_2(n)))_{n \geqq 1}$ is dense in $\mathbb{R}^+$. This extends work by Stolarsky [9] and Hare, Laishram and Stoll [6].

1. Introduction

Let $q \geqq 2$. Then we can express $n \in \mathbb{N}$ uniquely in base q as

$$n = \sum_{j \geqq 0} n_j q^j, \qquad n_j \in \{0, 1, \ldots, q-1\}. \tag{1}$$

Denote by $s_q(n) = \sum_{j \geqq 0} n_j$ the sum of digits of n in base q. The sum of digits of polynomial values has been at the center of interest in many works. We mention the (still open) conjecture of Gelfond [5] about the distribution of s_q of polynomial values in arithmetic progressions (see also [4,7,10]) and the fundamental work of Bassily and Kátai [1] on central limit theorems satisfied by s_q supported on polynomial values resp. polynomial values with prime arguments.

Stolarsky [9] examined the pointwise relationship between $s_q(n^h)$ and $s_q(n)$, where $h \geqq 2$ is a fixed integer. In particular, he used a result of Bose

---

* Corresponding author.

† The second author was supported by the Agence Nationale de la Recherche, grant ANR-10-BLAN 0103 MUNUM.

Key words and phrases: sum-of-digits, polynomial, digit, digital expansion, dense.
Mathematics Subject Classification: 11A63, 11B83, 11N60.

and Chowla [2] to prove that

$$\limsup_{n\to\infty} \frac{s_2(n^h)}{s_2(n)} = \infty. \tag{2}$$

Hare, Laishram and Stoll [6] generalized (2) to an arbitrary polynomial $P(X) \in \mathbb{Z}[X]$ of degree $h \geqq 2$ in place of X^h, and to base q in place of the binary base. Moreover, they showed that on the other side of the spectrum,

$$\liminf_{n\to\infty} \frac{s_q\bigl(P(n)\bigr)}{s_q(n)} = 0, \tag{3}$$

confirming a conjecture of Stolarsky.

From another point of view, not much is known about the pointwise relationship between the sum-of-digits of the values of two distinct fixed integer-valued polynomials $P_1(X)$, $P_2(X)$. Building up on work by Drmota [3], Steiner [8] showed that the distribution of the two-dimensional vector $\bigl(s_q\bigl(P_1(n)\bigr), s_q\bigl(P_2(n)\bigr)\bigr)_{n\geqq 1}$ obeys a central limit law (in fact, his results apply for general vectors and general q-additive functions). However, there are no local results available, such as an asymptotic formula or even a non-trivial lower estimate for

$$\#\bigl\{n < x : s_q\bigl(P_1(n)\bigr) = k_1,\ s_q\bigl(P_2(n)\bigr) = k_2\bigr\}, \tag{4}$$

where k_1 and k_2 are fixed positive integers.

The purpose of the present work is to extend both (2) and (3), and to make a first step towards understanding (4).

Our main result is as follows:

THEOREM 1.1. *Let $P_1(X), P_2(X) \in \mathbb{Z}[X]$ be polynomials of distinct degrees $h, l \geqq 1$ with $P_1(\mathbb{N}), P_2(\mathbb{N}) \subset \mathbb{N}$. Then*

$$\left(\frac{s_q\bigl(P_1(n)\bigr)}{s_q\bigl(P_2(n)\bigr)}\right)_{n\geqq 1}$$

is dense in $\mathbb{R}^+$.

REMARK 1. The proof extends to strictly q-additive functions in place of the sum-of-digits function s_q (we need, however, the condition that the weight attached to the digit $q-1$ is positive, cf. (16)). Recall that a strictly q-additive function f is a real-valued function f defined on the non-negative integers which satisfies $f(0) = 0$ and $f(n) = \sum_{j\geqq 0} f(n_j)$, where the n_j are the digits in the q-adic expansion (cf. (1)).

We first state some notation that is used throughout the paper. For integers a, b with $b < a$ we will write $[b, a]$ for the set of integers $\{b, b+1, \ldots, a\}$. For sets A and B, we write $mA = \{a_1 + \cdots + a_m : a_i \in A,\ 1 \leqq i \leqq m\}$ and $A + B = \{a + b : a \in A,\ b \in B\}$. For the sake of simplicity, we allow all constants to depend on q without further mentioning. Since we fix q already in the beginning there is not much harm to do so.

2. Proof of the main result

The proof of Theorem 1.1 will proceed in several steps. We first address the case $P_1(X) = X^h$, $P_2(X) = X^l$ which can be dealt with in a well arranged manner. The key idea in the proof is that

$$s_q(q^u) = 1, \qquad s_q(q^u - 1) = (q-1)u, \tag{5}$$

so that the first value is independent of u (and negligible, as $u \to \infty$) and the second one increases as u increases. In order to exploit this, we construct in Section 2.1 a polynomial $p(X)$ and determine the number of negative coefficients in $p(X)^t$ for $t \geqq 2$. In Section 2.2 we then show that, given a real number $r \in (0,1)$, we can choose the parameters of the polynomial in such a way that the ratios of the numbers of negative coefficients of $p(X)^h$ and $p(X)^l$ approximate r arbitrarily well. In Section 2.3 we link this ratio to the ratio of the sum-of-digits function under question and show that we obtain the same limit. The final section (Section 2.4) concerns the generalization to arbitrary polynomials $P_1(X), P_2(X)$.

2.1. Construction of the polynomial $p(X)$. In this section we construct the polynomial $p(X)$ which we will use later to approximate a given positive real ratio r. To clarify the underlying idea, we recall the method that has been used to prove (3) in [6] in the case of a monomial $P(X) = X^h$. First, the authors defined the polynomial

$$p(x) = q^k x^4 + q^k x^3 - x^2 + q^k x + q^k$$

for some (sufficiently large) fixed integer k. This polynomial has one negative coefficient and has the property that all coefficients in the expansion of $p(x)^h$ are positive for each fixed $h \geqq 2$ (each negative contribution to the coefficients is compensated by a large positive weight corresponding to the weights q^k). They then calculated $s_q\big(p(q^u)\big)$ and $s_q\big(p(q^u)^h\big)$. Since there is a negative coefficient in the expansion of the first value, this tends by (5) to infinity for $u \to \infty$. On the other hand, again by (5), the second value is bounded for $u \to \infty$, and the result follows.

For a general ratio r we have to make it possible to generate negative signs and to control them at the same time. In this section we prove the following lemma.

LEMMA 2.1. *Let $a, b, k, t \in \mathbb{N}$ with $a > b > a/2 > 0$, $t \geqq 2$ and*

$$\frac{a}{b} < \frac{t+1}{t}. \tag{6}$$

Set

$$A_1 = [0, a-b], \quad A_2 = [a-b+1, b-1], \quad A_3 = [b, a]$$

and define the polynomial $p(X) \in \mathbb{Z}[X]$ by

$$p(X) = q^k \sum_{i \in A_1} X^i - \sum_{i \in A_2} X^i + q^k \sum_{i \in A_3} X^i. \tag{7}$$

Denote by $C(p)$ the number of negative coefficients in the expansion of $p(X)$, and by $C(p^t)$ the number of negative coefficients in the expansion of $p(X)^t$. Then there exists $k_0 = k_0(t, a)$ such that for all $k \geqq k_0$,

$$C(p^t) = t\big((t+1)b - ta - 1\big).$$

Moreover, if $(t-1)a + b - 1 \geqq tb$ then all coefficients in the expansion of $p(X)^t$ are nonzero.

PROOF. We first determine the sign structure of the coefficients in the expansion of $p(X)^t$. For $0 \leqq i \leqq t$ set

$$Q_i = iA_1 + (t-i)A_3 = \big[(t-i)b, ta - ib\big]. \tag{8}$$

By (6) the sets Q_i, $1 \leqq i \leqq t$, are pairwise disjoint. We claim that there exists an integer $k_1(t, a) > 0$ such that for $k \geqq k_1(t, a)$ all the coefficients of the powers X^m with $m \in \bigcup_{i=0}^{t} Q_i$ in $p(X)^t$ are positive. To see this, we note that the positive coefficients of $p(X)$ that contribute to X^m have total weight at least q^{tk}, whereas the total contribution to X^m of terms that involve at least one negative coefficient of $p(X)$ is $O_{t,a}(q^{(t-1)k})$ (where the implied constant depends on t and a). Therefore, for each t and a there is $k_1(t, a) > 0$ such that for all $k \geqq k_1(t, a)$ the coefficients of $p(X)^t$ belonging to powers of the sets Q_i are positive.

For $0 \leqq i \leqq t-1$, we call

$$G_i = [1 + \max Q_{i+1}, -1 + \min Q_i] = \big[ta - (i+1)b + 1, (t-i)b - 1\big]$$

the *gap* between Q_{i+1} and Q_i. Each G_i contains

$$(t-i)b - 1 - (ta - (i+1)b + 1) + 1 = (t+1)b - ta - 1 \tag{9}$$

integers. Consider the coefficients of

$$X^m \quad \text{with} \quad m \in \bigcup_{i=0}^{t-1} G_i.$$

We claim that for sufficiently large k all these coefficients have negative sign. Since the Q_i's are disjoint, for each X^m with $m \in \bigcup_{i=0}^{t-1} G_i$ there must be a contributing term that involves at least one coefficient attached to some power with exponent in A_2. We use a similar argument as above: The total contribution from coefficients that involve $\geqq 2$ terms from A_2 is $O_{t,a}(q^{(t-2)k})$. On the other hand, the total weight of those contributions that involve exactly one coefficient from A_2 is (in modulus) at least $q^{(t-1)k}$. Therefore, there exists an integer $k_2(t,a) > 0$ such that for all $k \geqq k_2(t,a)$ the contributions of the terms that use exactly one term from A_2 are dominating.

The negative coefficients originate from terms that use one, three etc. factors with exponents in A_2. In order to pin down those terms with negative coefficients, we need only consider those negative coefficients that use exactly one factor from A_2. We define

$$N_i = iA_1 + A_2 + (t-1-i)A_3, \qquad 0 \leqq i \leqq t-1.$$

This simplifies to

$$N_i = \big[(a-b-1) + (t-1-i)b, i(a-b) + b - 1 + (t-1-i)a\big]. \tag{10}$$

In order to completely fill the gaps between two blocks of positive coefficients by negative ones, we suppose that for all i with $0 \leqq i \leqq t-1$,

$$\max N_i \geqq \min Q_i \quad \text{and} \quad \max Q_{i+1} \geqq \min N_i.$$

It is a straightforward calculation that both inequalities reduce to the same inequality, namely,

$$(t-1)a + b - 1 \geqq tb. \tag{11}$$

The statement of the lemma now follows from (11) and (9). $\square$

2.2. The number of negative coefficients of $p(X)^h$ and $p(X)^\ell$. The aim of this section is to prove the following result.

LEMMA 2.2. *Let $h, l \geqq 1$, $h \neq l$. Then the sequence*

$$\left(\frac{C(p^h)}{C(p^l)}\right)_{a>b>a/2>0}$$

is dense in $\mathbb{R}^+$.

PROOF. We first consider the case when $h > l \geqq 1$. Define the functions

$$f(a,b) = \frac{h}{l} \cdot \frac{(h+1) - h\frac{a}{b}}{(l+1) - l\frac{a}{b}}$$

and

$$f^*(a,b) = \frac{h}{l} \cdot \frac{(h+1) - h\frac{a}{b} - \frac{1}{b}}{(l+1) - l\frac{a}{b} - \frac{1}{b}}.$$

First suppose that $r = \frac{r_1}{r_2} \in (0,1)$ is a given rational number and set

$$a_0 = r_2 h(h+1) - r_1 l(l+1), \qquad b_0 = r_2 h^2 - r_1 l^2. \tag{12}$$

It is a straightforward calculation to check that a_0 and b_0 satisfy $a_0 > b_0 > a_0/2 > 0$,

$$\frac{a_0}{b_0} < \frac{h+1}{h} < \frac{l+1}{l} \tag{13}$$

and

$$b_0 - 1 \geqq \max\left(\frac{a_0}{2}, hb_0 - (h-1)a_0, lb_0 - (l-1)a_0\right). \tag{14}$$

The same is true for a_0 replaced by a_0N and b_0 by b_0N with $N \geqq 1$. Set $a = a_0$ and $b = b_0$ in Lemma 2.1 and write $p = p_0$ for the resulting polynomial. Then for sufficiently large k, all gaps in $p_0(X)^h$ and $p_0(X)^l$ are filled with powers having negative coefficients and

$$\frac{C(p_0^h)}{C(p_0^l)} = f^*(a_0, b_0).$$

Since $f(a_0, b_0) = \frac{r_1}{r_2}$ this implies that

$$\lim_{N\to\infty} f^*(a_0N, b_0N) = f(a_0, b_0) = \frac{r_1}{r_2},$$

which proves that $\left(C(p^h)/C(p^l)\right)$ is dense in $(0,1)$.

Now, suppose that $1 \leqq l < h$ and let $r = \frac{r_1}{r_2} \in \mathbb{Q}^+$. Denote by ν the minimal positive integer such that

$$\hat{r} := r\left(\frac{h+1}{l+1}\right)^{-\nu} \in (0,1). \tag{15}$$

Let a_0, b_0 the integers corresponding to $\hat{r}$ from the discussion before. Define

$$\hat{p}_0(X) = p_0(X), \quad \hat{p}_{i+1}(X) = \hat{p}_i(X)\left(1 + X^{h \deg \hat{p}_i + 1}\right), \quad 0 \leqq i \leqq \nu - 1.$$

From this definition we see that $C\big(\hat{p}_{i+1}^h\big) = (h+1)C\big(\hat{p}_i^h\big)$ and hence $C(\hat{p}_\nu^h) = (h+1)^\nu C\big(p_0^h\big)$. We therefore get

$$\frac{C\big(\hat{p}_\nu^h\big)}{C\big(\hat{p}_\nu^l\big)} = \frac{(h+1)^\nu C\big(p_0^h\big)}{(l+1)^\nu C\big(p_0^l\big)} = \frac{(h+1)^\nu}{(l+1)^\nu} f^*(a_0, b_0)$$

and, as before,

$$\lim_{N\to\infty} \frac{(h+1)^\nu}{(l+1)^\nu} f^*(a_0 N, b_0 N) = r,$$

as wanted.

We have now shown that $\big(C(p^h)/C(p^l)\big)$ is dense in $\mathbb{R}^+$ whenever $h > l \geqq 1$. This implies that $\big(C(p^l)/C(p^h)\big)$ is dense in $\mathbb{R}^+$ and this finishes the proof of the lemma. $\square$

2.3. The case of monomials. We now turn our attention to the ratio $s_q(n^h)/s_q(n^l)$.

Lemma 2.3. *Let $h, l \geqq 1$, $h \neq l$. Then the sequence*

$$\left(\frac{s_q(n^h)}{s_q(n^l)}\right)_{n\geqq 1}$$

is dense in $\mathbb{R}^+$.

Proof. We use the well known splitting property of the sum-of-digits function (see, for example, [6, Proposition 2.1] for a proof): For all $m_1, u \geqq 1$ and $1 \leqq m_2 < q^u$ we have

$$\begin{cases} s_q(m_1 q^u + m_2) = s_q(m_1) + s_q(m_2), \\ s_q(m_1 q^u - m_2) = s_q(m_1 - 1) + (q-1)u - s_q(m_2 - 1). \end{cases} \tag{16}$$

We apply (16) successively to the terms in the expansion of $p(q^u)^h$ and $p(q^u)^l$, respectively, and get

$$\lim_{u\to\infty} \frac{s_q\big(p(q^u)^h\big)}{s_q\big(p(q^u)^l\big)} = \lim_{u\to\infty} \left(\frac{s_q\big(p(q^u)^h\big)}{u} \cdot \frac{u}{s_q\big(p(q^u)^l\big)}\right)$$

$$= \frac{(q-1)C(p^h)}{(q-1)C(p^l)} = \frac{C(p^h)}{C(p^l)}.$$

Since by Lemma 2.2 this ratio is dense in $\mathbb{R}^+$, the result follows. $\square$

2.4. The case of general polynomials $P_1(X)$, $P_2(X)$. We are now ready to prove Theorem 1.1 in its generality.

Proof of Theorem 1.1. Without loss of generality, we can assume that both $P_1(X)$ and $P_2(X)$ have positive coefficients since otherwise there exists $n_0 = n_0(P_1, P_2)$ such that both $P_1(n+n_0)$ and $P_2(n+n_0)$ only have positive coefficients. We construct $p(X)$ as before with the monomials X^h, X^l in place of $P_1(X)$, $P_2(X)$. For $t \geqq 1$ consider

$$P(X) = \sum_{j=0}^{t} c_j X^j, \qquad c_j > 0, \quad 0 \leqq j \leqq t.$$

We claim that for sufficiently large k the polynomial $P\bigl(p(X)\bigr)$ has the same sign structure in its expansion as $p(X)^t$. First, we know that our construction fills up completely the gaps between $Q_{i+1}\bigl(p(X)^t\bigr)$ and $Q_i\bigl(p(X)^t\bigr)$ for any sufficiently large (fixed) k. Moreover, recall that we have shown that for $1 \leqq j \leqq t$ and $0 \leqq i \leqq j$, the total weight attached to each power X^m with $m \in Q_i\bigl(p(X)^j\bigr)$ (resp. $G_i\bigl(p(X)^j\bigr)$) is at least q^{jk} (resp. is $O_{a,j}(q^{(j-1)k})$).

Now, the relations (8), (10) and a comparison of the interval bounds imply that for all i, j and v with $0 \leqq i \leqq t-1$, $1 \leqq j \leqq t-1$ and $0 \leqq v \leqq j$,

$$G_i\bigl(p(X)^t\bigr) \cap Q_v\bigl(p(X)^j\bigr) = \emptyset.$$

This means that $P\bigl(p(X)\bigr)$ has at least the same number of powers with negative coefficients as $p(X)^t$. On the other hand, if

$$Q_i\bigl(p(X)^t\bigr) \cap G_v\bigl(p(X)^j\bigr) \neq \emptyset$$

then, as the weight associated to elements of $Q_i\bigl(p(X)^t\bigr)$ is dominant, we can find a sufficiently large k such that the coefficients to powers X^m for $m \in \bigcup_{0 \leqq i \leqq t} Q_i\bigl(p(X)\bigr)$ are positive. This shows, in particular, that for sufficiently large k the number of negative coefficients in the expansions of $P_1\bigl(p(X)\bigr)$ (resp. $P_2\bigl(p(X)\bigr)$) is $C(p^h)$ (resp. $C(p^l)$) and the same proof as in the case of monomials given before can be applied. $\square$

Acknowledgement. The second author wants to thank Kevin G. Hare (University of Waterloo) for several discussions on this problem. The authors are grateful to the referee for valuable comments that have led to an improved exposition of the result.

References

[1] N. L. Bassily and I. Kátai, Distribution of the values of q-additive functions on polynomial sequences, *Acta Math. Hungar.,* **68** (1995), 353–361.

[2] R. C. Bose and S. Chowla, Theorems in the additive theory of numbers, *Comm. Math. Helv.,* **37** (1962/63), 141–147.

[3] M. Drmota, The joint distribution of q-additive functions, *Acta Arith.,* **100** (2001), 17–39.

[4] M. Drmota, C. Mauduit and J. Rivat, The sum of digits function of polynomial sequence, *J. London Math. Soc.,* **84** (1) (2011), 81–102.

[5] A. O. Gelfond, Sur les nombres qui ont des propriétés additives et multiplicatives données, *Acta Arith.,* **13** (1967/1968), 259–265.

[6] K. G. Hare, S. Laishram and T. Stoll, Stolarsky's conjecture and the sum of digits of polynomial values, *Proc. Amer. Math. Soc.,* **139** (2011), 39–49.

[7] C. Mauduit and J. Rivat, La somme des chiffres des carrés, *Acta Math.,* **203** (2009), 107–148.

[8] W. Steiner, On the joint distribution of q-additive functions on polynomial sequences, *Theory Stoch. Process.,* **8** (2002), 336–357.

[9] K. B. Stolarsky, The binary digits of a power, *Proc. Amer. Math. Soc.,* **71** (1978), 1–5.

[10] T. Stoll, The sum of digits of polynomial values in arithmetic progressions, *Funct. Approx. Comment. Math.,* **47** (2012), 233–239.