



HAL
open science

The sum of digits of polynomial values in arithmetic progressions

Thomas Stoll

► **To cite this version:**

Thomas Stoll. The sum of digits of polynomial values in arithmetic progressions. *Functiones et Approximatio Commentarii Mathematici*, 2012, 47 (2), pp.233-239. 10.7169/facm/2012.47.2.7 . hal-01278713

HAL Id: hal-01278713

<https://hal.science/hal-01278713>

Submitted on 24 Feb 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

The sum of digits of polynomial values in arithmetic progressions

Thomas Stoll

Institut de Mathématiques de Luminy,
Université de la Méditerranée,
13288 Marseille Cedex 9, France
E-mail: stoll@iml.univ-mrs.fr

Abstract

Let $q, m \geq 2$ be integers with $(m, q - 1) = 1$. Denote by $s_q(n)$ the sum of digits of n in the q -ary digital expansion. Further let $p(x) \in \mathbb{Z}[x]$ be a polynomial of degree $h \geq 3$ with $p(\mathbb{N}) \subset \mathbb{N}$. We show that there exist $C = C(q, m, p) > 0$ and $N_0 = N_0(q, m, p) \geq 1$, such that for all $g \in \mathbb{Z}$ and all $N \geq N_0$,

$$\#\{0 \leq n < N : s_q(p(n)) \equiv g \pmod{m}\} \geq CN^{4/(3h+1)}.$$

This is an improvement over the general lower bound given by Dartyge and Tenenbaum (2006), which is $CN^{2/h!}$.

1 Introduction

Let $q, m \geq 2$ be integers and denote by $s_q(n)$ the sum of digits of n in the q -ary digital expansion of integers. In 1967/68, Gelfond [1] proved that for nonnegative integers a_1, a_0 with $a_1 \neq 0$, the sequence $(s_q(a_1n + a_0))_{n \in \mathbb{N}}$ is well distributed in arithmetic progressions mod m , provided $(m, q - 1) = 1$. At the end of his paper, he posed the problem of finding the distribution of s_q in arithmetic progressions where the argument is restricted to values of polynomials of degree ≥ 2 . Recently, Mauduit and Rivat [8] answered Gelfond's question in the case of squares.

2010 *Mathematics Subject Classification*: Primary 11A63; Secondary 11N37, 11N69.
Key words and phrases: sum of digits, polynomials, Gelfond's problem.

Theorem 1.1 (Mauduit & Rivat (2009)). *For any $q, m \geq 2$ there exists $\sigma_{q,m} > 0$ such that for any $g \in \mathbb{Z}$, as $N \rightarrow \infty$,*

$$\#\{0 \leq n < N : s_q(n^2) \equiv g \pmod{m}\} = \frac{N}{m} Q(g, d) + O_{q,m}(N^{1-\sigma_{q,m}}),$$

where $d = (m, q - 1)$ and

$$Q(g, d) = \#\{0 \leq n < d : n^2 \equiv g \pmod{d}\}.$$

The proof can be adapted to values of general quadratic polynomial instead of squares. We refer the reader to [7] and [8] for detailed references and further historical remarks. The case of polynomials of higher degree remains elusive so far. The Fourier-analytic approach, as put forward in [7] and [8], seems not to yield results of the above strength. In a recent paper, Drmota, Mauduit and Rivat [4] applied the Fourier-analytic method to show that well distribution in arithmetic progressions is obtained whenever q is sufficiently large.

In the sequel, and unless otherwise stated, we write

$$p(x) = a_h x^h + \cdots + a_0$$

for an arbitrary, but fixed polynomial $p(x) \in \mathbb{Z}[x]$ of degree $h \geq 3$ with $p(\mathbb{N}) \subset \mathbb{N}$.

Theorem 1.2 (Drmota, Mauduit & Rivat (2011)). *Let*

$$q \geq \exp(67h^3(\log h)^2)$$

be a sufficiently large prime number and suppose $(a_h, q) = 1$. Then there exists $\sigma_{q,m} > 0$ such that for any $g \in \mathbb{Z}$, as $N \rightarrow \infty$,

$$\#\{0 \leq n < N : s_q(p(n)) \equiv g \pmod{m}\} = \frac{N}{m} Q^*(g, d) + O_{q,m,p}(N^{1-\sigma_{q,m}}),$$

where $d = (m, q - 1)$ and

$$Q^*(g, d) = \#\{0 \leq n < d : p(n) \equiv g \pmod{d}\}.$$

It seems impossible to even find a single “nice” polynomial of degree 3, say, that allows to conclude for well distribution in arithmetic progressions for small bases, let alone that the binary case $q = 2$ is an emblematic case. Another line of attack to Gelfond’s problem is to find lower bounds that are valid for all $q \geq 2$. Dartyge and Tenenbaum [3] provided such a general lower bound by a method of descent on the degree of the polynomial and the estimations obtained in [2].

Theorem 1.3 (Dartyge & Tenenbaum (2006)). *Let $q, m \geq 2$ with $(m, q - 1) = 1$. Then there exist $C = C(q, m, p) > 0$ and $N_0 = N_0(q, m, p) \geq 1$, such that for all $g \in \mathbb{Z}$ and all $N \geq N_0$,*

$$\#\{0 \leq n < N : s_q(p(n)) \equiv g \pmod{m}\} \geq CN^{2/h!}.$$

The aim of the present work is to improve this lower bound for all $h \geq 3$. More importantly, we get a substantial improvement of the bound as a function of h . The main result is as follows.¹

Theorem 1.4. *Let $q, m \geq 2$ with $(m, q - 1) = 1$. Then there exist $C = C(q, m, p) > 0$ and $N_0 = N_0(q, m, p) \geq 1$, such that for all $g \in \mathbb{Z}$ and all $N \geq N_0$,*

$$\#\{0 \leq n < N : s_q(p(n)) \equiv g \pmod{m}\} \geq CN^{4/(3h+1)}.$$

Moreover, for monomials $p(x) = x^h$, $h \geq 3$, we can take

$$N_0 = q^{3(2h+m)} \left(2hq^2 (6q)^h\right)^{3h+1},$$

$$C = \left(16hq^5 (6q)^h \cdot q^{(24h+12m)/(3h+1)}\right)^{-1}.$$

The proof is inspired from the constructions used in [5] and [6] that were helpful in the proof of a conjecture of Stolarsky [9] concerning the pointwise distribution of $s_q(p(n))$ versus $s_q(n)$. As a drawback of the method of proof, however, it seems impossible to completely eliminate the dependency on h in the lower bound.

2 Proof of Theorem 1.4

Consider the polynomial

$$t(x) = m_3x^3 + m_2x^2 - m_1x + m_0, \quad (2.1)$$

where the parameters m_0, m_1, m_2, m_3 are positive real numbers that will be chosen later on in a suitable way. For all integers $l \geq 1$ we write

$$T_l(x) = t(x)^l = \sum_{i=0}^{3l} c_i x^i \quad (2.2)$$

¹Gelfond's work and Theorem 1.1 give precise answers for linear and quadratic polynomials, so we do not include the cases $h = 1, 2$ in our statement though our approach works without change.

to denote its l -th power. (For the sake of simplicity we omit to mark the dependency on l of the coefficients c_i .) The following technical result is the key in the proof of Theorem 1.4. It shows that, within a certain degree of uniformity in the parameters m_i , all coefficients *but one* of $T_l(x)$ are positive.

Lemma 2.1. *For all integers $q \geq 2$, $l \geq 1$ and $m_0, m_1, m_2, m_3 \in \mathbb{R}^+$ with*

$$1 \leq m_0, m_2, m_3 < q, \quad 0 < m_1 < l^{-1}(6q)^{-l}$$

we have that $c_i > 0$ for $i = 0, 2, 3, \dots, 3l$ and $c_i < 0$ for $i = 1$. Moreover, for all i ,

$$|c_i| \leq (4q)^l. \quad (2.3)$$

Proof. The coefficients of $T_l(x)$ in (2.2) are clearly bounded above in absolute value by the corresponding coefficients of the polynomial $(qx^3 + qx^2 + qx + q)^l$. Since the sum of all coefficients of this polynomial is $(4q)^l$ and all coefficients are positive, each individual coefficient is bounded by $(4q)^l$. This proves (2.3). We now show the first part. To begin with, observe that $c_0 = m_0^l > 0$ and $c_1 = -lm_1m_0^{l-1}$ which is negative for all $m_1 > 0$. Suppose now that $2 \leq i \leq 3l$ and consider the coefficient of x^i in

$$T_l(x) = (m_3x^3 + m_2x^2 + m_0)^l + r(x), \quad (2.4)$$

where

$$\begin{aligned} r(x) &= \sum_{j=1}^l \binom{l}{j} (-m_1x)^j (m_3x^3 + m_2x^2 + m_0)^{l-j} \\ &= \sum_{j=1}^{3l-2} d_j x^j. \end{aligned}$$

First, consider the first summand in (2.4). Since $m_0, m_2, m_3 \geq 1$ the coefficient of x^i in the expansion of $(m_3x^3 + m_2x^2 + m_0)^l$ is ≥ 1 . Note also that all the powers x^2, x^3, \dots, x^{3l} appear in the expansion of this term due to the fact that every $i \geq 2$ allows at least one representation as $i = 3i_1 + 2i_2$ with non-negative integers i_1, i_2 . We now want to show that for sufficiently small $m_1 > 0$ the coefficient of x^i in the first summand in (2.4) is dominant. To this end, we assume $m_1 < 1$ so that $m_1 > m_1^j$ for $2 \leq j \leq l$. Using $\binom{l}{j} < 2^l$ and a similar reasoning as above we get that

$$|d_j| < l2^l m_1 (3q)^l = l(6q)^l m_1, \quad 1 \leq j \leq 3l - 2.$$

This means that if $m_1 < l^{-1}(6q)^{-l}$ then the powers x^2, \dots, x^{3l} in the polynomial $T_l(x)$ indeed have positive coefficients. This finishes the proof. \square

To proceed we recall the following splitting formulas for s_q which are simple consequences of the q -additivity of the function s_q (see [5] for the proofs).

Proposition 2.2. *For $1 \leq b < q^k$ and $a, k \geq 1$, we have*

$$\begin{aligned} s_q(aq^k + b) &= s_q(a) + s_q(b), \\ s_q(aq^k - b) &= s_q(a - 1) + k(q - 1) - s_q(b - 1). \end{aligned}$$

We now turn to the proof of Theorem 1.4. To clarify the construction we consider first the simpler case of monomials,

$$p(x) = x^h, \quad h \geq 1.$$

(We here include the cases $h = 1$ and $h = 2$ because we will need them to deal with general polynomials with linear and quadratic terms.) Let $u \geq 1$ and multiply $t(x)$ in (2.1) by q^{u-1} . Lemma 2.1 then shows that for all *integers* m_0, m_1, m_2, m_3 with

$$q^{u-1} \leq m_0, m_2, m_3 < q^u, \quad 1 \leq m_1 < q^u / (hq(6q)^h), \quad (2.5)$$

the polynomial $T_h(x) = (t(x))^h = p(t(x))$ has all positive (*integral*) coefficients with the only exception of the coefficient of x^1 which is negative. Let u be an integer such that

$$q^u \geq 2hq(6q)^h \quad (2.6)$$

and let $k \in \mathbb{Z}$ be such that

$$k > hu + 2h. \quad (2.7)$$

For all u with (2.6) the interval for m_1 in (2.5) is non-empty. Furthermore, relation (2.7) implies by (2.3) that

$$q^k > q^{hu} \cdot q^{2h} \geq (4q^u)^h > |c_i|, \quad \text{for all } i = 0, 1, \dots, 3h,$$

where c_i here denotes the coefficient of x^i in $T_h(x)$. Roughly speaking, the use of a large power of q (i.e. q^k with k that satisfies (2.7)) is motivated by the simple wish to split the digital structure of the h -power according to Proposition 2.2. By doing so, we avoid to have to deal with carries when adding terms in the expansion in base q since the appearing terms will not interfere. We also remark that this is the point where we get the dependency of h in the lower bound of Theorem 1.4.

Now, by $c_2, |c_1| \geq 1$ and the successive use of Proposition 2.2 we get

$$\begin{aligned}
s_q(t(q^k)^h) &= s_q \left(\sum_{i=3}^{3h} c_i q^{ik} + c_2 q^{2k} - |c_1| q^k + c_0 \right) \\
&= s_q \left(\sum_{i=3}^{3h} c_i q^{(i-1)k} + c_2 q^k - |c_1| \right) + s_q(c_0) \\
&= s_q \left(\sum_{i=3}^{3h} c_i q^{(i-3)k} \right) + s_q(c_2 - 1) + k(q - 1) - s_q(|c_1| - 1) + s_q(c_0) \\
&= \sum_{i=3}^{3h} s_q(c_i) + s_q(c_2 - 1) + k(q - 1) - s_q(|c_1| - 1) + s_q(c_0) \\
&= k(q - 1) + M, \tag{2.8}
\end{aligned}$$

where we write

$$M = \sum_{i=3}^{3h} s_q(c_i) + s_q(c_2 - 1) - s_q(|c_1| - 1) + s_q(c_0).$$

Note that M is an integer that depends (in some rather obscure way) on the quantities m_0, m_1, m_2, m_3 . Once we fix a quadruple (m_0, m_1, m_2, m_3) in the ranges (2.5), the quantity M does not depend on k and is constant whenever k satisfies (2.7). We now exploit the appearance of the single summand $k(q - 1)$ in (2.8). Since by assumption $(m, q - 1) = 1$, we find that

$$s_q(t(q^k)^h), \quad \text{for } k = hu + 2h + 1, hu + 2h + 2, \dots, hu + 2h + m, \tag{2.9}$$

runs through a complete set of residues mod m . Hence, in any case, we hit a fixed arithmetic progression mod m (which might be altered by M) for some k with $hu + 2h + 1 \leq k \leq hu + 2h + m$.

Summing up, for u with (2.6) and by (2.5) we find at least

$$(q^u - q^{u-1})^3 (q^u / (hq(6q)^h) - 1) \geq \frac{(1 - 1/q)^3}{2hq(6q)^h} q^{4u} \tag{2.10}$$

integers n that in turn by (2.1), (2.5), (2.7) and (2.9) are all smaller than

$$q^u \cdot q^{3(hu+2h+m)} = q^{3(2h+m)} \cdot q^{u(3h+1)}$$

and satisfy $s_q(n^h) \equiv g \pmod{m}$ for fixed g and m . By our construction and by choosing $k > hu + 2h > u$ all these integers are distinct. We denote

$$N_0 = N_0(q, m, p) = q^{3(2h+m)} \cdot q^{u_0(3h+1)},$$

where

$$u_0 = \lceil \log_q (2hq(6q)^h) \rceil \leq \log_q (2hq^2(6q)^h).$$

Then for all $N \geq N_0$ we find $u \geq u_0$ with

$$q^{3(2h+m)} \cdot q^{u(3h+1)} \leq N < q^{3(2h+m)} \cdot q^{(u+1)(3h+1)}. \quad (2.11)$$

By (2.10) and (2.11), and using $(1 - 1/q)^3 \geq 1/8$ for $q \geq 2$, we find at least

$$\frac{(1 - 1/q)^3}{2hq(6q)^h} q^{4u} \geq \left(16hq^5 (6q)^h \cdot q^{(24h+12m)/(3h+1)} \right)^{-1} N^{4/(3h+1)}$$

integers n with $0 \leq n < N$ and $s_q(n^h) \equiv g \pmod{m}$. We therefore get the statement of Theorem 1.4 for the case of monomials $p(x) = x^h$ with $h \geq 3$. The estimates are also valid for $h = 1$ and $h = 2$.

The general case of a polynomial $p(x) = a_h x^h + \dots + a_0$ of degree $h \geq 3$ (or, more generally, of degree $h \geq 1$) follows easily from what we have already proven. Without loss of generality we may assume that all coefficients a_i , $0 \leq i \leq h$, are positive, since otherwise there exists $e = e(p)$ depending only on p such that $p(x + e)$ has all positive coefficients. Note that a finite translation can be dealt with choosing C and N_0 appropriately in the statement. Since Lemma 2.1 holds for all $l \geq 1$ and all negative coefficients are found at the same power x^1 , we have that the polynomial $p(t(x))$ has again all positive coefficients *but one* where the negative coefficient again corresponds to the power x^1 . It is then sufficient to suppose that

$$k > hu + 2h + \log_q \max_{0 \leq i \leq h} a_i$$

in order to split the digital structure of $p(t(q^k))$. In fact, this implies that

$$q^k > \left(\max_{0 \leq i \leq h} a_i \right) \cdot (4q^u)^h,$$

and exactly the same reasoning as before yields $\gg_{q,p} q^{4u}$ distinct positive integers that are $\ll_{q,m,p} q^{u(3h+1)}$ and satisfy $s_q(p(n)) \equiv g \pmod{m}$. This completes the proof of Theorem 1.4.

Acknowledgements

This research was supported by the Agence Nationale de la Recherche, grant ANR-10-BLAN 0103 MUNUM.

References

- [1] A. O. Gelfond, Sur les nombres qui ont des propriétés additives et multiplicatives données, *Acta Arith.* **13** (1967/1968), 259–265.
- [2] C. Dartyge, G. Tenenbaum, Sommes de chiffres de multiples d’entiers, *Ann. Inst. Fourier (Grenoble)* **55** (2005), no. 7, 2423–2474.
- [3] C. Dartyge, G. Tenenbaum, Congruences de sommes de chiffres de valeurs polynomiales, *Bull. London Math. Soc.* **38** (2006), no. 1, 61–69.
- [4] M. Drmota, C. Mauduit, J. Rivat, The sum of digits function of polynomial sequences, *J. London Math. Soc.* **84** (2011), 81–102.
- [5] K. G. Hare, S. Laishram, T. Stoll, Stolarsky’s conjecture and the sum of digits of polynomial values, *Proc. Amer. Math. Soc.* **139** (2011), 39–49.
- [6] K. G. Hare, S. Laishram, T. Stoll, The sum of digits of n and n^2 , *Int. J. Number Theory* (2011), to appear, doi:10.1142/S1793042111004319.
- [7] C. Mauduit, J. Rivat, Sur un problème de Gelfond: la somme des chiffres des nombres premiers, *Ann. of Math.* **171** (2010), 1591–1646.
- [8] C. Mauduit, J. Rivat, La somme des chiffres des carrés, *Acta Math.* **203** (2009), 107–148.
- [9] K. B. Stolarsky, The binary digits of a power, *Proc. Amer. Math. Soc.* **71** (1978), 1–5.