isid/ms/2011/09
August 4, 2011
http://www.isid.ac.in/~statmath/eprints

The sum of digits of n and n^2

Kevin G. Hare
Shanta Laishram
Thomas Stoll

Indian Statistical Institute, Delhi Centre
7, SJSS Marg, New Delhi–110 016, India

THE SUM OF DIGITS OF n AND n^2

KEVIN G. HARE, SHANTA LAISHRAM, AND THOMAS STOLL

ABSTRACT. Let $s_q(n)$ denote the sum of the digits in the q-ary expansion of an integer n. In 2005, Melfi examined the structure of n such that $s_2(n) = s_2(n^2)$. We extend this study to the more general case of generic q and polynomials $p(n)$, and obtain, in particular, a refinement of Melfi's result. We also give a more detailed analysis of the special case $p(n) = n^2$, looking at the subsets of n where $s_q(n) = s_q(n^2) = k$ for fixed k.

1. INTRODUCTION

Let $q \geq 2$ and denote by $s_q(n)$ the sum of digits in the q-ary representation of an integer n. Recently, considerable progress has been made towards understanding the interplay between the sum-of-digits of some algebraically defined sequences, such as primes [5] and polynomials [1] or, in particular, squares [6]. In the latter, C. Mauduit and J. Rivat proved an asymptotic expansion of the sum of digits of squares [6] in arithmetic progressions. Their proof heavily relies on good estimates of quadratic Gauss sums. For the case of general polynomials $p(n)$ of degree $h > 2$ there is still a great lack of knowledge regarding their distribution with respect to digitally defined functionals [1].

Several authors studied the pointwise properties and relationships of $s_q(p(n))$, e.g., K. Stolarsky [8], B. Lindström [4], G. Melfi [7], and M. Drmota and J. Rivat [2]. In particular, a conjecture of Stolarsky [8] about some extremal distribution properties of the ratio $s_q(p(n))/s_q(n)$ has been recently settled by the authors [3]. Melfi [7] proposed to study the set of n's such that $s_2(n^2) = s_2(n)$, and he obtained that

$$\#\left\{n < N : \quad s_2(n^2) = s_2(n)\right\} \gg N^{1/40}. \tag{1}$$

Using heuristic arguments, Melfi conjectured the much stronger result

$$\#\left\{n < N : \quad s_2(n^2) = s_2(n)\right\} \approx \frac{N^{\beta}}{\log N} \tag{2}$$

and gave an explicit formula for $\beta \approx 0.75488\ldots$. The aim of the present paper is to provide a generalization to general $p(n)$ and base q of Melfi's result as well as to use the method of proof to sharpen Melfi's exponent in (1). Moreover, we provide a local analog, i.e., a lower bound for the number of n's such that $s_q(n^2) = s_q(n) = k$ for some fixed k.

K.G. Hare was partially supported by NSERC.

Computational support provided by CFI/OIT grant.

Th. Stoll was supported by the Agence Nationale de la Recherche, project "ANR-10-BLAN 0103".

Theorem 1.1. *Let $p(x) \in \mathbb{Z}[x]$ have degree at least 2, and positive leading coefficient. Then there exists an explicitly computable $\gamma > 0$, dependent only on q and $p(x)$, such that*

$$\#\left\{n < N,\ q \nmid n : \quad |s_q(p(n)) - s_q(n)| \leq \frac{q-1}{2}\right\} \gg N^{\gamma}, \tag{3}$$

where the implied constant depends only on q and $p(x)$.

This result is given in Section 2. In the general case of q-ary digits and polynomials $p(x)$, the bound $(q-1)/2$ in (3) cannot be improved. This is easily seen by recalling the well-known fact

$$s_q(n) \equiv n \bmod (q-1). \tag{4}$$

Indeed, if we set $p(x) = (q-1)x^2 + x + a$ for $a \in \mathbb{N}$ then we find that

$$s_q(p(n)) - s_q(n) \equiv p(n) - n \equiv a \bmod (q-1)$$

which could be any of $0, 1, \ldots, q-2$ depending only on the choice of a.

The method of proof of Theorem (1.1) allows to improve on Melfi's result (1).

Theorem 1.2.

$$\#\left\{n < N : \quad s_2(n^2) = s_2(n)\right\} \gg N^{1/19}. \tag{5}$$

Following on Melfi's paper [7], we examine the case when $p(n) = n^2$ and $q = 2$ in more detail. We consider the set of all n's such that $s_2(n) = s_2(n^2)$, and partition the set into the subsets dependent upon the value of $s_2(n)$. By noticing that $s_2(n) = s_2(2n)$ and $s_2(n^2) = s_2((2n)^2)$ we see that we can restrict our attention to odd n.

Theorem 1.3. *Let $k \leq 8$. Then*

$$\{n \text{ odd} : \quad s_2(n^2) = s_2(n) = k\}$$

is a finite set.

This was done by explicit computation of all such n which are given in Tables 1 and 2. A discussion of how these computations were made is given in Section 3.

Based on these initial small values of k, one might expect that this is always true. Let

$$n_2 = 1101111\underbrace{00\ldots00}_{r}1101111 \tag{6}$$

be written in base 2. Then $s_2(n) = s_2(n^2) = 12$ for all $r \geq 8$. This is in fact a special case of a more general property.

Theorem 1.4. *Let $k \geq 16$ or $k \in \{12, 13\}$. Then*

$$\{n < N,\ n \text{ odd} : \quad s_2(n^2) = s_2(n) = k\}$$

is an infinite set.

The proof of this result is given in Section 4. Despite great effort we are not able to decide the finiteness problem in the remaining cases $k \in \{9, 10, 11, 14, 15\}$. However, we will comment in Section 5 on some heuristic evidence making it seem unlikely that there are infinitely many solutions in the cases $k = 9$ and $k = 10$.

Somewhat surprisingly, a similar answer can be given if $q \geq 3$.

Theorem 1.5. *Let $q \geq 3$ and assume*

$$k \geq 94(q-1).$$

Then the equation

$$s_q(n^2) = s_q(n) = k \tag{7}$$

has infinitely many solutions in n with $q \nmid n$ if and only if

$$k(k-1) \equiv 0 \mod (q-1). \tag{8}$$

We show this result in Section 6.

2. Proof of Theorems 1.1 and 1.2

Following Lindström [4] we say that terms are *noninterfering* if we can use the following splitting formulæ:

Proposition 2.1. *For $1 \leq b < q^k$ and $a, k \geq 1$,*

$$s_q(aq^k + b) = s_q(a) + s_q(b), \tag{9}$$

$$s_q(aq^k - b) = s_q(a-1) + (q-1)k - s_q(b-1). \tag{10}$$

Proof. See [3]. □

Proof of Theorem 1.1: The proof uses a construction of a sequence with noninterfering terms which has already been used in [3]. However, to obtain the bound N^γ in (3) instead of a logarithmic bound, we have to make a delicate refinement. To begin with, define the polynomial

$$t_m(x) = mx^4 + mx^3 - x^2 + mx + m$$

where $m \in \mathbb{Z}$. Set $m = q^l - r$ with $1 \leq r \leq \lfloor q^{\alpha l} \rfloor$, $q \nmid r$ and $0 < \alpha < 1$. Obviously, for $\alpha < 1$ there exists $l_0(\alpha)$ such that for all $l > l_0(\alpha)$ we have $m \geq 3$. Furthermore let k be such that

$q^k > m$. By consecutively employing (9) and (10) we see that

$$
\begin{aligned}
s_q(t_m(q^k)) &= s_q(mq^{4k} + mq^{3k} - q^{2k} + mq^k + m) \\
&= s_q(mq^{4k}) + s_q(mq^{3k} - q^{2k}) + s_q(mq^k) + s_q(m) \\
&= s_q(m) + s_q(mq^k - 1) + s_q(m) + s_q(m) \\
&= (q-1)k + s_q(m-1) + 3s_q(m) \qquad (11)\\
&= (q-1)k + s_q(q^l - (r+1)) + 3s_q(q^l - r) \\
&= (q-1)k + (q-1)l - s_q(r) + 3((q-1)l - s_q(r-1)) \\
&= (q-1)k + 4(q-1)l - K(q,r)
\end{aligned}
$$

where K is depends only on q and r, and does not depend on k. First consider the easier case of monomials $p(n) = n^h$, $h \geq 2$ where we can give a somewhat more direct proof. We have

$$
\begin{aligned}
t_m(x)^h &= (mx^4 + mx^3 - x^2 + mx + m)^h \\
&= \sum_{j=0}^{4h} c_{j,h}(m) x^j \qquad (12)\\
&= m^h x^{4h} + hm^h x^{4h-1} + \left(\binom{h}{2} m^h - hm^{h-1}\right) x^{4h-2} \\
&\quad + \left(\left(h + \binom{h}{3}\right) m^h - 2\binom{h}{2} m^{h-1}\right) x^{4h-3} + \text{smaller powers.}
\end{aligned}
$$

From [3, Lemma 3.1] we have that $t_m(x)^h$ has only positive coefficients and $0 < c_{j,h}(m) \leq (2mh)^h$. This means that $s_q(t_m(q^k)^h)$ does not depend on k if k is sufficiently large (see (9)). More precisely, if $q^k > (2mh)^h$ (which is true if $q^k > (2h)^h q^{lh}$, or equivalently if $k > (h+1)l$ for sufficiently large l), then by a symmetry argument for the coefficients of $t_m(x)^h$,

$$
\begin{aligned}
s_q(t_m(q^k)^h) \geq 2 \Bigg(& s_q(m^h) + s_q(hm^h) + s_q\left(\binom{h}{2} m^h - hm^{h-1}\right) \\
& + s_q\left(\left(h + \binom{h}{3}\right) m^h - 2\binom{h}{2} m^{h-1}\right)\Bigg). \qquad (13)
\end{aligned}
$$

Consider the first summand $s_q(m^h)$ in (13). We have

$$
\begin{aligned}
m^h &= (q^l - r)^h = \sum_{j=0}^{h} \binom{h}{j} (-1)^{h-j} q^{jl} r^{h-j} \\
&= \sum_{j=0}^{h} (-1)^{h-j} d_j q^{jl} \qquad (14)
\end{aligned}
$$

which shows that m^h is a polynomial in q^l with coefficients of alternating signs. Now there are exactly $\lceil h/2 \rceil$ negative signs in this expansion. All coefficients in (14) are bounded in modulus by

$$0 < d_j \leq (2r)^h \leq (2q^{\alpha l})^h \leq q^{(\alpha l+1)h}, \tag{15}$$

and in turn their q-ary sum of digits is less than $s_q(d_j) \leq (q-1)(\alpha l+1)h$. (Note if equality is strict in (15) then $s_q(d_j) = 1$, otherwise it will have at most $(\alpha l + 1)h$ digits.) Therefore, by using (10), for $\lceil h/2 \rceil$ times, and observing that $s_q(d_j) \leq (q-1)(\alpha l+1)h$ we get that for fixed $\alpha < 1/h$ and sufficiently large l we have

$$\begin{aligned} s_q(m^h) &\geq \lceil h/2 \rceil (q-1)l - \lceil h/2 \rceil (q-1)(\alpha l+1)h \\ &\geq \frac{h}{2}(q-1)(l(1-\alpha h) - h). \end{aligned} \tag{16}$$

A similar argument can be applied to the other three summands in (13). This yields

$$s_q(t_m(q^k)^h) \geq 4h(q-1)(l(1-\alpha h) - h). \tag{17}$$

We recall that $s_q(t_m(q^k)^h)$ is independent of k (see discussion after (12), whereas from (11) we have $s_q(t_m(q^k))$ will increase by $q-1$ for each increase in k.

Take $\alpha = 1/(5h^2)$. Note that $h \geq 2$ and take k' and l sufficiently large so that

$$k' + 4h^2 \leq 4l \left(\frac{5h-6}{5} \right)$$

and

$$(h+1)l < k'$$

The second requirement is necessary for the validity of equation (17). This then implies that

$$\begin{array}{rrcl} & k' + 4h^2 & \leq & 4l\left(\frac{5h-6}{5}\right) \\ \Longrightarrow & k' & \leq & 4hl\left(1 - \frac{1}{5h} - \frac{1}{h}\right) - 4h^2 \\ \Longrightarrow & k' + 4l & \leq & 4hl(1-\alpha h) - 4h^2 \\ \Longrightarrow & (q-1)k' + 4(q-1)l - K(q,r) & \leq & 4h(q-1)(l(1-\alpha h) - h) \\ \Longrightarrow & s_q(t_m(q^{k'})) & \leq & s_q(t_m(q^{k'})^h) \end{array}$$

Recall that for each increase of k' by 1, the left hand side will increase by $q-1$, (by (11)), and the right hand side will remain fixed. Hence, for l sufficiently large, we can find a $k \geq k'$ such that

$$|s_q(t_m(q^k)^h) - s_q(t_m(q^k))| \leq \frac{q-1}{2}. \tag{18}$$

Summing up, we have obtained that for sufficiently large l we can find $\gg q^{\alpha l}$ values r where we in turn can provide a value k satisfying (18). In addition, each triple (l, r, k) gives rise to a different value of $t_m(q^k)$. We thus have (3).

Now consider the case of a general polynomial $p(x) = a_h x^h + a_{h-1}x^{h-1} + \cdots + a_0 \in \mathbb{Z}[x]$. There exist positive integers s_1 and s_2, both only depending on the polynomial $p(x)$ such that

$$p(q^{s_1}x + q^{s_2} + 1) = a'_h x^h + a'_{h-1}x^{h-1} + \cdots + a'_0$$

has only positive coefficients. With the notation of (12) we obtain

$$p(q^{s_1}t_m(x)+q^{s_2}+1) = \sum_{i=0}^{3} a'_h c_{4h-i,h}(m)\, x^{4h-i} + \sum_{i=4}^{7}\left(a'_h c_{4h-i,h}(m) + a'_{h-1}c_{4h-i,h-1}(m)\right) x^{4h-i} + \text{smaller powers}. \tag{19}$$

First suppose $h \geq 4$. By choosing s_1 sufficiently large (this choice again only depends on $p(x)$) we get that the coefficients of x^j in $p(q^{s_1}t_m(x)+q^{s_2}+1)$ with $4h-7 \leq j \leq 4h$ are polynomials in m of degree h since we can avoid unwanted cancellation for these coefficients. The coefficients of these terms (as polynomials in m) are alternating in sign, since for $h \geq 4$ and $i = 0, 1, \ldots, 2h-1$ we have

$$c_{i,h}(m) = c_{4h-i,h}(m) = \sum_{j=h-\lfloor i/2 \rfloor}^{h} d_{j,i,h} m^j \tag{20}$$

where $d_{j,i,h}d_{j+1,i,h} < 0$ for all j with $h - \lfloor i/2 \rfloor \leq j < h$. Setting $m = q^l - r$ we therefore can choose s_1, s_2 in the way that $a'_h c_{4h-i,h}(m) + a'_{h-1}c_{4h-i,h-1}(m)$ as a polynomial in q^l has $\lceil h/2 \rceil$ negative coefficients for each $= 0, 1, \ldots, 2h-1$. Now, for $q^{s_2}+1 < q^{s_1}$, we get by (11) that

$$s_q(q^{s_1}t_m(q^k) + q^{s_2} + 1) \leq (q-1)k + 4(q-1)l + 2.$$

In (19) we have therefore found eight summands sharing the property of the eight summands in the monomial case (see (13)). From this we proceed as in the case of monomials to get the statement.

It remains to deal with the cases of general quadratic and cubic polynomials, where we cannot directly resort to (20) (note that $8 > (2h-1)+1$ for $h = 2, 3$). We instead do a more direct calculation. Let $h = \deg p = 2$ which is the case of quadratic polynomials. By suitably shifting the argument $x \mapsto q^{s_1}x + q^{s_2} + 1$ we can arrange for a polynomial $p(q^{s_1}x+q^{s_2}+1) = a'_2x^2 + a'_1x + a'_0$ with $a'_2, a'_1, a'_0 > 0$ and $2a'_2 > a'_1$. Each coefficient of x^i in $p(q^{s_1}t_m(x)+q^{s_2}+1)$, $0 \leq i \leq 8$, is a function of m and of a'_2, a'_1 and a'_0. In a similar way as before (here we use 9 summands instead of the 8 in the case of $h \geq 4$) we obtain for sufficiently large l,

$$s_q(p(q^{s_1}t_m(q^k)+q^{s_2}+1)) > 8(q-1)l \geq 4h(q-1)l.$$

Now we can choose k suitably to get the assertion. Finally, for a cubic polynomial, we are able to achieve $p(q^{s_1}x+q^{s_2}+1) = a'_3x^3 + a'_2x^2 + a'_1x + a'_0$ with $a'_3, a'_2, a'_1, a'_0 > 0$ and $3a'_3 > a'_2$. Then, each coefficient of x^i in $p(q^{s_1}t_m(x)+q^{s_2}+1)$, $0 \leq i \leq 12$, is a function of m and a'_3, a'_2, a'_1, a'_0, and thus we get for sufficiently large l,

$$s_q(p(q^{s_1}t_m(q^k)+q^{s_2}+1)) > 12(q-1)l \geq 4h(q-1)l.$$

By choosing k suitably, we obtain the result. This completes the proof of Theorem 1.1. □

Proof of Theorem 1.2: We apply the method of proof of Theorem 1.1 to the special case $q = 2$ and $p(n) = n^2$. Instead of using the rather crude bounds, we here use exact values to get our result. To begin with, we observe that the largest coefficient (as $m \to \infty$) of $t_m(x)^2$ is the coefficient of x^4, namely $4m^2 + 1$. Therefore we get noninterfering terms when $2^k > 4m^2 + 1$. A sufficient condition for this is $2^k \geq 4 \cdot 2^{2l} = 2^{2l+2}$, or equivalently,

$$k \geq 2l + 2. \tag{21}$$

On the other hand, the coefficients of x^8 and x^7 (resp. x^1 and x^0) in $t_m(x)^2$ are m^2 and $2m^2$ which have the same binary sum of digits. Now assume $\alpha < 1/2$ and $l > l_0(\alpha)$ be sufficiently large. We then use Proposition 2.1 and set $m = 2^l - r$ with $1 \leq r \leq \lfloor 2^{\alpha l} \rfloor$ to obtain

$$\begin{aligned} s_2(t_m(2^k)^2) &\geq 4s_2(m^2) + s_2(4m^2+1) \\ &= 5s_2\left((2^{l-1} - r)2^{l+1} + r^2\right) + 1 \\ &\geq 5s_2(2^{l-1} - r) \\ &= 5\left((l-1) - s_2(r-1)\right) \\ &\geq 5(l-1) - 5\alpha l \\ &\geq (2+\varepsilon)l \end{aligned} \tag{22}$$

for any $0 < \varepsilon < 1/2$. This means that for any $\alpha < 1/2$ we have $\gg q^{\alpha l}$ values r where we in turn can provide a value k satisfying (18) which is due to

$$2l + 2 \leq k \leq (2+\varepsilon)l.$$

This yields

$$t_m(q^k) \leq 2q^{4k+l} \leq 2q^{4(2+\varepsilon)l+l} \leq q^{(9+5\varepsilon)l}.$$

Hence, letting $N = q^{(9+5\varepsilon)l}$ we note that we have

$$\gg q^{\alpha l} = \left(N^{\frac{1}{(9+5\varepsilon)l}}\right)^{\alpha l} = N^{\alpha/(9+5\varepsilon)} \geq N^{1/19}$$

solutions to (18). This finishes the proof. □

3. Proof of Theorem 1.3

The proof that there are only a finite number of odd n such that $s_2(n^2) = s_2(n) \leq 8$ is a strictly computational one. We discuss how our algorithm works. The code is available upon request from the first named author.

Consider

$$n = \sum_{i=1}^{k} 2^{r_i} = 2^{r_1} + 2^{r_2} + \cdots + 2^{r_k}$$

with $0 = r_1 < r_2 < r_3 < \cdots < r_k$. We have

$$n^2 = \sum_{i=1}^{k}\sum_{j=1}^{k} 2^{r_i+r_j} = \sum_{i=1}^{k} 2^{2r_i} + \sum_{i=1}^{k}\sum_{j=i+1}^{k} 2^{r_i+r_j+1}.$$

We therefore need to examine the exponents

$$\{2r_1,\ 2r_2,\ \ldots,\ 2r_k,\ r_1+r_2+1,\ r_1+r_3+1,\ \ldots,\ r_{k-1}+r_k+1\}$$

and the possible iterations between these exponents by carry propagation.

Clearly, $2r_1$ is the strict minimum within these exponents. Other relationships between exponents are not as clear. For example, $r_1 + r_3 + 1$ could be less than, equal to, or greater than $2r_2$ depending on the choices of r_3 and r_2. Each of these cases must be examined in turn. Numerous of these inequalities have implications for the order of other exponents in the binary expansion of n^2. So, once we make an assumption in our case by case analysis, this might rule out future possibilities. For example, if we assume that $2r_3 < 1 + r_1 + r_4$, then we have as a consequence that $1+r_2+r_3 < 1+r_1+r_4$ (by noticing that $r_2 < r_3$). In the case of equality we "group" terms. For example, if we assumed that $2r_3 = 1 + r_2 + r_4$, then we could, first, replace all occurrences of r_2 with $2r_3 - 1 - r_4$, and second replace $2^{2r_3} + 2^{1+r_2+r_4}$ by 2^{2r_3+1}.

Our algorithm occasionally finds a solution set with fractional or negative values for r_i, which is a contradiction. On the other hand, it is possible for the algorithm to find a solution, even if all of the exponents cannot be explicitly determined. This would happen if there is an infinite family of n with $s_2(n^2) = s_2(n) = k$ with some nice structure, (as is the case for $k = 12$, see (6)). The algorithm will detect, and report this. We used the method for k up to 8. For each of these values, there was only a finite number of n, and all of them are enumerated in Tables 1 and 2.

4. Proof of Theorem 1.4

For the proof of Theorem 1.4, we first state some auxiliary results. Denote by $(n)_2$ the binary representation of n, and $1^{(k)}$ a block of k binary 1. We begin with the following key observation.

Proposition 4.1. *If there exists u and v such that $s_2(u)+s_2(v) = s_2(u^2)+s_2(uv)+s_2(v^2) = k$, then for i sufficiently large, the numbers of the form $(n)_2 = u0^iv$ satisfy $s_2(n^2) = s_2(n) = k$.*

Proof. This follows at once from Proposition 2.1, relation (9). □

We use Proposition 4.1 to prove the following lemma.

Lemma 4.2. *Let $(u)_2 = 1^{(k_1)}01^{(n_1)}$ and $(v)_2 = 1^{(k_2)}01^{(n_2)}$. Assume that $n_1 \geq k_1+2$, $n_2 \geq k_2+2$ and $n_1 \geq n_2$. Then*

$$s_2(u^2) = n_1 \quad \text{and} \quad s_2(v^2) = n_2,$$

Base 10	Base 2	Base 10	Base 2
$\mathbf{s_2(n) = s_2(n^2) = 1}$		$\mathbf{s_2(n) = s_2(n^2) = 7}$	
1	1	127	1111111
		319	100111111
$\mathbf{s_2(n) = s_2(n^2) = 2}$		351	101011111
3	11	375	101110111
		379	101111011
$\mathbf{s_2(n) = s_2(n^2) = 3}$		445	110111101
7	111	575	1000111111
		637	1001111101
$\mathbf{s_2(n) = s_2(n^2) = 4}$		815	1100101111
15	1111	1087	10000111111
		1149	10001111101
$\mathbf{s_2(n) = s_2(n^2) = 5}$		1255	10011100111
31	11111	1815	11100010111
79	1001111	2159	100001101111
91	1011011	2173	100001111101
157	10011101	2297	100011111001
279	100010111	2921	101101101001
		4191	1000001011111
$\mathbf{s_2(n) = s_2(n^2) = 6}$		4207	1000001101111
63	111111	4345	1000011111001
159	10011111	6477	1100101001101
183	10110111	8689	10000111110001
187	10111011	10837	10101001010101
287	100011111	16701	100000100111101
317	100111101	18321	100011110010001
365	101101101	33839	1000010000101111
573	1000111101		
1071	10000101111		
1145	10001111001		
1449	10110101001		
4253	1000010011101		
4375	1000100010111		
4803	1001011000011		

TABLE 1. Odd n such that $s_2(n^2) = s_2(n) \leq 7$.

Base 10	Base 2	Base 10	Base 2
$s_2(n) = s_2(n^2) = 8$		**$s_2(n) = s_2(n^2) = 8$ (cont)**	
255	11111111	5811	1011010110011
639	1001111111	5865	1011011101001
703	1010111111	5911	1011100010111
735	1011011111	5971	1011101010011
751	1011101111	6479	1100101001111
759	1011110111	6557	1100110011101
763	1011111011	8415	10000011011111
893	1101111101	8445	10000011111101
975	1111001111	8697	10000111111001
1151	10001111111	10035	10011100110011
1215	10010111111	11591	10110101000111
1277	10011111101	11597	10110101001101
1455	10110101111	13233	11001110110001
1463	10110110111	13591	11010100010111
1495	10111010111	16575	100000010111111
1501	10111011101	16607	100000011011111
1599	11000111111	16889	100000111111001
1647	11001101111	17393	100001111110001
1661	11001111101	22807	101100100010111
2175	100001111111	23441	101101110010001
2301	100011111101	23575	101110000010111
2685	101001111101	25907	110010100110011
2919	101101100111	33777	1000001111110001
2987	101110101011	46377	1011010100101001
3259	110010111011	46881	1011011100100001
4223	1000001111111	51811	1100101001100011
4349	1000011111101	66173	10000001001111101
4601	1000111111001	67553	10000011111100001
4911	1001100101111	69521	10000111110010001
5069	1001111001101	133231	100000100001101111
5231	1010001101111	227393	110111100001000001
5799	1011010100111	266335	1000001000001011111

TABLE 2. Odd n such that $s_2(n^2) = s(n) = 8$.

and

$$s_2(uv) = \begin{cases} k_1+2 & \text{if } n_2 = k_1+1, n_1 = n_2+k_2+1 \\ n_2+1 & \text{if } n_2 > k_1+1, n_1 = n_2+k_2+1 \\ n_1+1 & \text{if } k_1 = k_2, n_1 > n_2. \end{cases}$$

Proof. Let $(U)_2 = 1^{(k)}01^{(n)}$ with $n \geq k+2$. Then $U = 2^n - 1 + 2^{n+1}(2^k-1)$ and we calculate

$$\begin{aligned} U^2 &= 2^{2n} - 2^{n+1} + 1 + 2^{n+2}(2^{n+k} - 2^n - 2^k + 1) + 2^{2n+2}(2^{2k} - 2^{k+1} + 1) \\ &= 1 + 2^{n+1} + 2^{2n} + 2^{n+k+2}(1 + 2 + 2^2 + \cdots + 2^{n+k-1}) - 2^{2n+k+2} \\ &= 1 + 2^{n+1} + 2^{n+k+2} + \cdots + 2^{2n-1} + 2^{2n+k+2} + 2^{2n+k+3} + \cdots + 2^{2n+2k+1}. \end{aligned}$$

Hence $s_2(U^2) = n$ and therefore $s_2(u^2) = n_1$ and $s_2(v^2) = n_2$.

Now, consider $s_2(uv)$. We have

$$\begin{aligned} uv =& 1 + 2^{n_1} + 2^{n_2} + 2^{n_1+n_2} - 2^{n_1+k_1+1} - 2^{n_2+k_2+1} - 2^{n_1+n_2+k_1+1} - \\ & 2^{n_1+n_2+k_2+1} + 2^{n_1+n_2+k_1+k_2+2}. \end{aligned}$$

We may assume that $k_1 \geq k_2$. Then

$$\begin{aligned} W &:= 2^{n_1+n_2+k_1+k_2+2} - 2^{n_1+n_2+k_2+1} - 2^{n_1+n_2+k_1+1} \\ &= 2^{n_1+n_2+k_2+1}(1 + 2 + \cdots + 2^{k_1-k_2-1} + 2^{k_1-k_2+1} + \cdots + 2^{k_1}) \end{aligned}$$

has $s_2(W) = k_1$. We distinguish three cases to conclude:

(1) Let $n_1 = n_2+k_2+1$ and $n_2 = k_1+1$. Then $uv = 1+2^{n_2}+W$ and hence $s_2(uv) = k_1+2$.
(2) Let $n_1 = n_2+k_2+1$ and $n_2 > k_1+1$. Then $uv = 1+2^{n_2}+W+2^{n_1+k_1+1}(2^{n_2-k_1-1}-1)$ and hence $s_2(uv) = 2+k_1+n_2-k_1-1 = n_2+1$.
(3) Let $k_1 = k_2 = k$ and $n_1 > n_2$. Then $uv = 1+2^{n_2}+2^{n_1}+W+2^{n_2+k+1}(2^{n_1-k-1}-1) - 2^{n_1+k+1}$ and hence $s_2(uv) = 3+k+n_1-k-2 = n_1+1$.

This finishes the proof. □

Proof of Theorem 1.4. Let n_1, n_2, k_1, k_2 be positive integers with $n_1 \geq k_1+2$, $n_2 \geq k_2+2$ and u, v be as in Lemma 4.2. Let $(N)_2 = u0^Rv$ be the binary representation of N where $R \geq n_1+n_2+k_1+k_2$. By Proposition 4.1 and Lemma 4.2 we have for any $R \geq n_1+n_2+k_1+k_2$,

$$\begin{aligned} s_2(N) &= s_2(u) + s_2(v) = n_1 + n_2 + k_1 + k_2, \\ s_2(N^2) &= s_2(u^2) + s_2(v^2) + s_2(uv) = n_1 + n_2 + s_2(uv). \end{aligned}$$

Let $k \geq 2$. Taking $k_1 = k_2 = k$ and $n_1 = n_2 = 2k$, we find from Lemma 4.2 and $2k \geq k+2$ that

$$s_2(N^2) = s_2(N) = 6k$$

implying there are infinite families of n such that $s_2(n) = s_2(n^2) = s$ for s of the form $6k$ with $k \geq 2$.

Let $k_2 = 2, k_1 \geq 3, n_2 = k_1 + 2$ and $n_1 = n_2 + k_2 + 1 = k_1 + 4$. Then $s_2(uv) = n_2 + 1$ by Lemma 4.2 implying $s_2(N^2) = s_2(N) = 3(k_1 + 2) + 1$. Hence there are infinite families of n such that $s_2(n) = s_2(n^2) = s$ for s of the form $3k + 1$ with $k \geq 5$.

Let $k_1 \geq k_2 \geq 3$ and $n_2 = k_1 + k_2 - 1, n_1 = n_2 + k_2 + 1$. Then $s_2(uv) = n_2 + 1 = k_1 + k_2$ from Lemma 4.2 implying $s_2(N^2) = s_2(N) = 3k_1 + 4k_2 - 1$. Let $k_2 = 3$. Then $s_2(N^2) = s_2(N) = 3(k_1 + 3) + 2$ for $k_1 \geq 3$ giving infinite families of n such that $s_2(n) = s_2(n^2) = s$ for s of the form $3k + 2$ with $k \geq 6$.

Let $k_2 = 4$. Then $s_2(N^2) = s_2(N) = 3(k_1 + 5)$ for $k_1 \geq 4$ giving infinite families of n such that $s_2(n) = s_2(n^2) = s$ for s of the form $3k$ with $k \geq 27$.

Summing up, we have infinite families of n with $s(n^2) = s(n) = s$ for all $s \geq 22$, respectively, $s \in \{12, 16, 18, 19, 20\}$. For $s \in \{13, 17, 21\}$, we take $(N)_2 = u0^R v$ with

$$s = 13 : u = 10111,\ v = 10110111111$$
$$s = 17 : u = 111011111,\ v = 10110111111$$
$$s = 21 : u = 11110111111,\ v = 111101111111.$$

This completes the proof of Theorem 1.4. □

5. Evidence that $s_2(n^2) = s_2(n) \leq 10$ is finite

All examples of infinite families with $s_2(n^2) = s_2(n) = k$ have the form given from Lemma 4.1. We show that there do not exist u and v satisfying Proposition 4.1, with $k \in \{9, 10\}$. We illustrate this method for $k = 8$, as it contains all of the key ideas without being overly cumbersome. The case of $k = 8$ is actually proved to be finite by the techniques of Section 3, but this does not detract from this example. The other two cases are similar.

Assume the contrary, that there exist u and v such that

$$s_2(u) + s_2(v) = s_2(u^2) + s_2(v^2) + s_2(uv) = 8$$

We easily see that $s_2(v), s_2(u) \geq 2$. Furthermore, as $s_2(u), s_2(v) \geq 2$, we see that $s_2(u^2), s_2(v^2) \geq 2$. Also, we have that $s_2(uv) \geq 2$. Therefore, we have $2 \leq s_2(u^2), s_2(v^2) \leq k - 4$. Lastly, we see that one of u or v must be "deficient", that $s_2(u^2) < s_2(u)$ or $s_2(v^2) < s_2(v)$.

Assume without loss of generality that $s_2(u^2) < s_2(u)$. Given the restrictions, we have that $2 \leq s_2(u) \leq 6$. Using the same algorithm as in Section 3, we can find all u such that $2 \leq s_2(u) \leq 6$ and $s_2(u^2) < s_2(u)$, $s_2(u^2) \leq 4$. These are the first three entries of Table 3.

Therefore, it suffices to show that there do not exist v for $u = 23, 47$ or 111 with $s_2(u) + s_2(v) = s_2(u^2) + s_2(v^2) + s_2(uv) = 8$.

(1) Let $u = 23 = 10111$. Given that $s_2(uv) \geq 2$ we have that $s_2(v) = 4$ and $s_2(v^2) \leq 3$. The only possible solution by Table 3 is $v = 23 = 10111$, but $s_2(uv) = 3$, a contradiction.
(2) Let $u = 47 = 101111$. Given that $s_2(uv) \geq 2$ we have that $s_2(v) = 3$ and $s_2(v^2) \leq 2$. There are no solutions by Table 3 for this, a contradiction.

Base 10	Base 2		
	u	$s_2(u)$	$s_2(u^2)$
23	10111	4	3
47	101111	5	4
111	1101111	6	4
95	1011111	6	5
5793	1011010100001	6	5
223	11011111	7	5
727	1011010111	7	5
191	10111111	7	6
367	101101111	7	6
415	110011111	7	6
1451	10110101011	7	6
46341	1011010100000101	7	6
479	111011111	8	5
447	110111111	8	6
887	1101110111	8	6

TABLE 3. $s_2(u) \le 8$, $s_2(u^2) < s_2(u)$ and $s_2(u^2) \le 6$.

(3) Let $u = 111 = 11101111$. Given that $s_2(uv) \ge 2$ we have that $s_2(v) = 2$ and $s_2(v^2) \le 2$. There is one possible solution to this by Table 3, namely $v = 3 = 11$. But then $s_2(uv) = 5$, a contradiction.

A similar, but more elaborate analysis can be done for $k = 9$ and $k = 10$ using the additional information in Table 3. Here we look at $2 \le s_2(u) \le 7$, $s_2(u^2) < s_2(u)$ and $s_2(u^2) \le 5$.

6. Proof of Theorem 1.5

The proof uses the strategy adopted for the case $q = 2$ (see Section 3). However, in order to handle more possible digits in the case of $q \ge 3$, the analysis is much more delicate. In the proof we will make frequent use of the fact (4) and of the splitting formulae of Proposition 2.1, which will apply if we have noninterfering terms at our disposal.

To begin with, the condition (8) is necessary, since (7) implies

$$s_q(n^2) - s_q(n) \equiv n^2 - n \equiv k^2 - k \equiv 0 \mod (q-1).$$

For the construction of an infinite family, we first prove a crucial lemma.

Lemma 6.1. *Let*

$$u = ((q-1)^k\, 0\, (q-1)^n e)_q$$

with $k \geq 2$*,* $n \geq k+2$ *and* $0 \leq e \leq q-2$*. Then*

$$s_q(u) = (q-1)(n+k) + e$$

and

$$s_q(u^2) = (q-1)(n+1) + f(q,e)$$

where

$$f(q,e) = s_q((q-e)^2) + s_q(2(q-1)(q-e)) - s_q(2(q-e)-1). \tag{23}$$

Proof. Since $u = e + (q^n - 1)q + (q^k - 1)q^{n+2}$, we get

$$\begin{aligned} u^2 &= (q-e)^2 + 2(q-1)(q-e)q^{n+1} - 2(q-e)q^{n+k+2} \\ &\quad + (q-1)^2 q^{2n+2} - 2(q-1)q^{2n+k+3} + q^{2n+2k+4}. \end{aligned} \tag{24}$$

By assumption that $n \geq k+2$ and $n, k \geq 2$, the terms in (24) are noninterfering. We therefore get

$$\begin{aligned} s_q(u^2) &= s_q((q-e)^2) + s_q(2(q-1)(q-e)) - s_q(2(q-e)-1) + (n-k)(q-1) \\ &\quad + s_q((q-1)^2 - 1) - s_q(2(q-1)-1) + (k+1)(q-1). \\ &= (n+1)(q-1) + s_q(q^2-2q) - s_q(2q-3) + f(q,e). \end{aligned}$$

The claimed value of $s_q(u^2)$ now follows by observing that $s_q(q^2-2q) = s_q(q-2) = q-2$ and $s_q(2q-3) = s_q(q+q-3) = 1+q-3 = q-2$. □

Now consider

$$\begin{aligned} u &= ((q-1)^{k_1}\ 0\ (q-1)^{n_1})_q, \\ v &= ((q-1)^{k_2}\ 0\ (q-1)^{n_2} e)_q \end{aligned}$$

where we suppose $k_1, n_1, k_2, n_2 \geq 2$ and $n_1 \geq k_1+2$, $n_2 \geq k_2+2$. Since $q \nmid n$ we further suppose that $e \neq 0$. We want to construct an infinite family of solutions to (7) of the form $n = (u0^{(i)}v)$, where i is a sufficiently large integer, such that terms will be noninterfering. Our task is to find an admissible set of parameters k_1, n_1, k_2, n_2 such that for sufficiently large $n_1 + n_2 + k_1 + k_2$ we have

$$\begin{aligned} s_q(u) + s_q(v) &= s_q(u^2) + s_q(2uv) + s_q(v^2) \\ &= e + (q-1)(n_1+n_2+k_1+k_2). \end{aligned} \tag{25}$$

First it is a straightforward calculation to show that $2uv = w_1 + w_2$ with

$$w_1 = 2q^{n_1+n_2+k_1+k_2+3} - 2(q-1)q^{n_1+n_2+k_1+2} - 2(q-1)q^{n_1+n_2+k_2+2} \tag{26}$$

and

$$\begin{aligned} w_2 &= 2(q-1)^2 q^{n_1+n_2+1} - 2(q-e)q^{n_1+k_1+1} - 2q^{n_2+k_2+2} \\ &\quad + 2(q-1)(q-e)q^{n_1} + 2(q-1)q^{n_2+1} + 2(q-e). \end{aligned} \tag{27}$$

Note that w_1 and w_2 are noninterfering because of $k_2 \geq 2$. Now, set

$$k_1 = n_2 \geq k_2 + 2, \qquad n_1 = 2k_2 - \alpha, \tag{28}$$

where we will later suitably choose $\alpha = \alpha(q, e)$ only depending on q and e. Then terms in (26) are again noninterfering and we get

$$\begin{aligned} s_q(w_1) &= s_q(2q^{k_1+1} - 2(q-1)q^{k_1-k_2} - 2(q-1)) \\ &= s_q(2q^{k_2+1} - 2q + 1) + (q-1)(k_1 - k_2) - s_q(2q-3) \\ &= 1 + k_2(q-1) + (q-1)(k_1 - k_2) - (q-2) \\ &= (k_1 - 1)(q-1) + 2. \end{aligned}$$

Next, by (28), we find that

$$\begin{aligned} w_2 = 2q^{k_1+2k_2-\alpha+1}&((q-1)^2 - (q-e)) - 2q^{k_1+k_2+2} \\ &+ 2(q-1)(q-e)q^{2k_2-\alpha} + 2(q-1)q^{k_1+1} + 2(q-e). \end{aligned} \tag{29}$$

In order to have terms noninterfering in (29), we impose the following inequalities on the parameters,

$$2 \leq k_1 + 1, \tag{30}$$

$$2 \leq (2k_2 - \alpha) - (k_1 + 1), \tag{31}$$

$$3 \leq (k_1 + k_2 + 2) - (2k_2 - \alpha) = k_1 - k_2 + 2 + \alpha, \tag{32}$$

$$1 \leq (k_1 + 2k_2 - \alpha + 1) - (k_1 + k_2 + 2) = k_2 - \alpha - 1. \tag{33}$$

Then we get

$$s(w_2) = (k_2 - \alpha - 1)(q-1) + g(q,e)$$

where

$$\begin{aligned} g(q,e) = s_q(2(q-e)) &+ s_q(2(q-1)) + s_q(2(q-1)(q-e)) \\ &+ s_q(2(q-1)^2 - (q-e) - 1) - 1. \end{aligned} \tag{34}$$

Summing up, we have

$$\begin{aligned} &s_q(u^2) + s_q(2uv) + s_q(v^2) \\ &\quad = (q-1)(n_1+1) + f(q,e) + (q-1)(n_2+1) + (k_1-1)(q-1) \\ &\qquad + 2 + (k_2 - \alpha - 1)(q-1) + g(q,e) \\ &\quad = (q-1)(2k_1 + 3k_2 - 2\alpha) + f(q,e) + g(q,e) + 2. \end{aligned}$$

Combining with (25) and (28) we therefore have

$$(q-1)(2k_1 + 3k_2 - 2\alpha) + f(q,e) + g(q,e) + 2 = (q-1)(2k_1 + 3k_2 - \alpha) + e \tag{35}$$

and

$$\alpha(q-1) = f(q,e) + g(q,e) - e + 2.$$

Rule (4) applied to (23) and (34) shows that the right hand side is indeed divisible by $q-1$ since $e^2 - e \equiv 0 \bmod (q-1)$ by assumption. Furthermore, we have by a crude estimation (using also (4)) that

$$0 \le \alpha \le 15. \tag{36}$$

Suppose $k_2 \ge 17$. Then (30) and (33) are satisfied. Rewriting (31) and (32) gives

$$1 + k_2 - \alpha \le k_1 \le 2k_2 - \alpha - 1. \tag{37}$$

Note that $k_1 \ge k_2 + 2$ is more restrictive than the first inequality in (37). On the other hand, since $k_2 \ge 2$, the interval given for k_1 in (37) has at least $(2 \cdot 17 - \alpha - 1) - (1 + 17 - \alpha) + 1 = 16$ terms. Therefore, $2k_1 + 3k_2$ hits all integers $\ge 2(1 + (k_2+1) - \alpha) + 3(k_2+1)$ for $k_2 \ge 17$. Thus, we find from (35) that all values

$$\begin{aligned}(q-1)(2k_1 + 3k_2 - \alpha) + e &\ge (q-1)(2 \cdot (19-0) + 3 \cdot 18) + (q-1) \\ &= 94(q-1)\end{aligned}$$

can be achieved. This completes the proof of Theorem 1.5.

References

[1] C. Dartyge, G. Tenenbaum, Congruences de sommes de chiffres de valeurs polynomiales, *Bull. London Math. Soc.* **38** (2006), no. 1, 61–69.

[2] M. Drmota, J. Rivat, The sum-of-digits function of squares, *J. London Math. Soc. (2)* **72** (2005), no. 2, 273–292.

[3] K. G. Hare, S. Laishram, T. Stoll, Stolarsky's conjecture and the sum of digits on polynomial values, *Proc. Amer. Math. Soc.* to appear.

[4] B. Lindström, On the binary digits of a power, *J. Number Theory* **65** (1997), 321–324.

[5] C. Mauduit, J. Rivat, Sur un probléme de Gelfond: la somme des chiffres des nombres premiers, *Annals of Mathematics*, **171** (2010), no. 3, 1591–1646.

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

[7] G. Melfi, On simultaneous binary expansions of n and n^2, *J. Number Theory* **111** (2005), no. 2, 248–256.

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

Department of Pure Mathematics, University of Waterloo, Waterloo, Ontario, Canada, N2L 3G1,

E-mail address: kghare@math.uwaterloo.ca

Stat Math Unit, Indian Statistical Institute (ISI), New Delhi, 110016, India,

E-mail address: shanta@isid.ac.in

Institut de Mathématiques de Luminy, Université de la Méditerranée, 13288 Marseille Cedex 9, France,

E-mail address: stoll@iml.univ-mrs.fr