The sum-of-digits function on arithmetic progressions

Lukas Spiegelhofer and Thomas Stoll

The sum-of-digits function on arithmetic progressions

Lukas Spiegelhofer and Thomas Stoll

Let s_2 be the sum-of-digits function in base 2, which returns the number of nonzero binary digits of a nonnegative integer n. We study s_2 along arithmetic subsequences and show that—up to a shift—the set of m-tuples of integers that appear as an arithmetic subsequence of s_2 has full complexity.

1. Results

The binary sum-of-digits function s_2 is an elementary object studied in number theory. It is defined by the equation

$$s_2(\varepsilon_\nu 2^\nu + \cdots + \varepsilon_0 2^0) = \varepsilon_\nu + \cdots + \varepsilon_0,$$

where $\varepsilon_i \in \{0, 1\}$ for $0 \le i \le \nu$. Despite the simplicity of definition, the behaviour of s_2 on arithmetic progressions is not fully understood. Cusick's conjecture on the sum-of-digits function [Drmota et al. 2016; Spiegelhofer 2019] concerns this area of research: for an integer $t \ge 0$, we define the limit

$$c_t = \lim_{N \to \infty} \frac{1}{N} |\{n : 0 \le n < N, \ s_2(n+t) \ge s_2(n)\}|.$$

(The limit exists; see for example [Bésineau 1972]. In fact, the set in this definition is periodic with period 2^k for some k.) Cusick's conjecture states that

$$c_t > \tfrac{1}{2} \tag{1-1}$$

for all $t \ge 0$. Drmota, Kauers, and the first author [Drmota et al. 2016] proved that $c_t > \frac{1}{2}$ for *almost all t* in the sense of asymptotic density; we also wish to note [Emme and Prikhod'ko 2017; Emme and Hubert 2018; 2019], and the recent partial result [Spiegelhofer 2019].

For $t \ge 0$ and $k \in \mathbb{Z}$, we consider the quantity

$$\delta(k, t) = \lim_{N \to \infty} \frac{1}{N} |\{n : 0 \le n < N, \ s_2(n+t) - s_2(n) = k\}|$$

(as was the case for c_t, this asymptotic density exists [Bésineau 1972]). The family $\delta(\cdot, t)$ defines a probability distribution with mean 0. Clearly, Cusick's conjecture states that

$$\delta(0, t) + \delta(1, t) + \cdots > \tfrac{1}{2}. \tag{1-2}$$

Numerical computation, using the recurrence for the values $\delta(k, t)$ presented below, reveals that indeed $c_t > \frac{1}{2}$ for all $t < 2^{30}$. Moreover, since the sum (1-2) involves the summand $\delta(0, t)$, and the associated probability distribution is usually close to a normal distribution [Emme and Hubert 2019], the strict inequality $c_t > \frac{1}{2}$ is plausible. In fact the statement $\frac{1}{2}\delta(0, t) + \delta(1, t) + \delta(2, t) + \cdots > \frac{1}{2}$ is not true in general ($t = 27$ being the first counterexample); moreover the statement $\tilde{c}_t \leq \frac{1}{2}$, where $\tilde{c}_t = c_t - \delta(0, t)$, is true for $t < 2^{30}$ and for almost all t with respect to asymptotic density [Drmota et al. 2016].

In the current note, motivated by Cusick's conjecture, we are concerned with the $(m+1)$-tuple

$$(s_2(n), s_2(n+t), \ldots, s_2(n+mt)),$$

where $t \geq 0$ and $m \geq 1$ are integers. We aim to understand the set of tuples that can occur, as n and t run. In fact, our theorem states that, up to a shift, all tuples occur.

Theorem 1.1. *Assume that $k_1, \ldots, k_m \in \mathbb{Z}$. There exist n and t such that for $1 \leq \ell \leq m$,*

$$k_\ell = s_2(n + \ell t) - s_2(n).$$

This is a generalization of the statement that the Thue–Morse sequence **t** has full *arithmetic complexity*, meaning that every finite word $\omega \in \{0, 1\}^L$ occurs as an arithmetic subsequence of **t**. This was first proved in [Avgustinovich et al. 2003] and also follows from [Müllner and Spiegelhofer 2017; Konieczny 2019]. Our theorem should be compared to the results proved in [Sanna 2012], who considered pairwisely different values of $s_2(n + 0t), \ldots, s_2(n + \ell t)$.

Theorem 1.1 is not hard to prove for $m = 1$. We present three arguments leading to this fact:

(1) Assume first that $k \geq 0$. Set $n = 2^{k+1}$ and $t = 2^k - 1$. Then $s_2(n + t) = k + 1$ and $s_2(n) = 1$, yielding $k = s_2(n + t) - s_2(n)$. If $k < 0$, we set $n = 2^{-k+1} - 1$ and $t = 1$. Then $s_2(n) = -k + 1$ and $s_2(n + t) = 1$, which yields $s_2(n + t) - s_2(n) = k$.

Alternatively, we may also write, as in the case $m = 2$ presented below, $t = 2^c - 1$ and $n = 2^{c-1}(2^a - 1)$ for positive integers a and c. We obtain $s_2(n + t) = c$ and $s_2(n) = a$, and clearly the difference $c - a$ runs through all integers.

(2) We have $s_2(n + 1) - s_2(n) = 1 - v_2(n + 1) \leq 1$, where $v_2(m) = \max\{k \geq 0 : 2^k \mid m\}$ for $m \geq 1$ is the 2-adic valuation of m. This formula follows by considering the number of 1s with which the binary expansion of n ends. Since $s_2(2^\ell) = 1$ and $s_2(2^{\ell+1} - 1) = \ell + 1$, we obtain the fact that $s_2(n)$ attains all values in $\{1, \ldots, \ell + 1\}$ as n varies in $\{2^\ell, \ldots, 2^{\ell+1} - 1\}$. Let $k \in \mathbb{Z}$ be given and set $\ell = 2|k|$. Choose $n \in \{2^\ell, \ldots, 2^{\ell+1} - 1\}$ such that $s_2(n) = |k| + 1$ and $n' \in \{2^{\ell+1}, \ldots, 2^{\ell+2} - 1\}$ such that $s_2(n') = |k| + 1 + k$. Then $s_2(n') - s_2(n) = k$, which implies the statement.

(3) The quantities $\delta(k, t)$ satisfy the following recurrence [Drmota et al. 2016]:

$$\delta(k, 1) = \begin{cases} 2^{k-2}, & k \leq 1, \\ 0 & \text{otherwise,} \end{cases}$$

$$\delta(k, 2t) = \delta(k, t),$$

$$\delta(k, 2t + 1) = \tfrac{1}{2}\delta(k - 1, t) + \tfrac{1}{2}\delta(k + 1, t + 1).$$

From this, it is very easy to show that $\delta(k, t) > 0$ for all $k \leq s_2(t)$. For k given, choose t in such a way that $s_2(t) \geq k$; the positivity of the density $\delta(k, t)$ implies that there exists an n such that $s_2(n + t) - s_2(n) = k$.

For $m = 2$, it is also possible to obtain the statement by elementary considerations: consider integers $a, c \geq 1$, $b, d \geq 0$ and choose the integers n and t in such a way that the binary expansions look as follows:

$$
\begin{array}{l}
n: \overbrace{1\cdots 11}^{a} 0\cdots 0 \overbrace{1\cdots 1}^{b} 0\cdots 0 \\
t: \quad\; \underbrace{1 1\cdots 1}_{c} 0\cdots 0 \underbrace{1\cdots 1}_{d}.
\end{array}
$$

The sums of digits of n, $n + t$, and $n + 2t$ respectively are $a + b$, $b + c + d$, and $c + d$ respectively. By varying the variables, we can obtain the statement for all integers k_1 and k_2 such that $k_2 \leq k_1$. For the case $k_1 < k_2$, we use the following configuration of the integers n and t, where $a, d \geq 1$ and $c \geq 0$:

$$
\begin{array}{l}
n: \qquad\quad \overbrace{1\cdots 11}^{a} 0\cdots 0 \\
t: \underbrace{1\cdots 1}_{c} 0\cdots 0 \underbrace{1 1\cdots 1}_{d}.
\end{array}
$$

The sums of digits of n, $n + t$, and $n + 2t$ are a, d, and $c + d$ respectively, and we see that we obtain all pairs $(k_1, k_2) \in \mathbb{Z}^2$ such that $k_1 \leq k_2$.

However, the method quickly experiences difficulties, as multiplication by 3 is not a shift of the binary digits anymore. While we believe that the case $m = 3$ can be done with some effort, a general principle is not apparent. Therefore we choose a different approach.

We prove Theorem 1.1 by induction on m, the cases $m = 1, 2$ having been discussed above. Assume that $m \geq 3$ and let $k_1, \ldots, k_m \in \mathbb{Z}$ be given. By induction hypothesis, there exist t_0 and n_0 such that

$$
k_\ell = s_2(n_0 + \ell t_0) - s_2(n_0 + (\ell - 1)t_0) \quad \text{for } 1 \leq \ell < m.
$$

Set

$$
k'_m = s_2(n_0 + m t_0) - s_2(n_0 + (m - 1)t_0).
$$

We are going to show that we may vary k'_m by steps of ± 1, thus yielding the full statement.

By concatenation of binary expansions, it is sufficient to show the following statement:

There exist t_1, n_1 such that $s_2(n_1 + \ell t_1) - s_2(n_1 + (\ell - 1)t_1) = 0$ for $1 \leq \ell < m$

$$
\text{and } s_2(n_1 + m t_1) - s_2(n_1 + (m - 1)t_1) = \pm 1. \quad \text{(1-3)}
$$

This concatenation is straightforward and summarized in the following lemma, which we will also use again in a moment.

Lemma 1.2. *Let $\ell \geq 1$, $m \geq 1$, $n_0, \ldots, n_{k-1}$, and $t_0, \ldots, t_{k-1}$ be nonnegative integers. There exist nonnegative integers n and t such that*

$$
s_2(n + \ell t) - s_2(n + (\ell - 1)t) = \sum_{0 \leq j < k} (s_2(n_j + \ell t_j) - s_2(n_j + (\ell - 1)t_j))
$$

for $1 \leq \ell \leq m$.

Proof. The base case $k = 1$ is trivial; it is sufficient to prove the statement for $k = 2$, the general case following easily from repeated application of this case.

Let N be so large that $n_0 + mt_0 < 2^N$, and set $n = 2^N n_1 + n_0$ and $t = 2^N t_1 + t_0$. Since no carry propagation between the digits below and above N occurs, we can add up the contribution of the two blocks in order to yield the statement. $\qquad\square$

We reduce the problem further, using this block representation again: choose $t_j = 1$ for all $0 \le j < k$; it is sufficient to find a $k \ge 1$ and nonnegative integers n_j for $0 \le j < k$ such that

$$\sum_{0 \le j < k} (s_2(n_j + \ell) - s_2(n_j + \ell - 1)) = \begin{cases} 0 & \text{if } 1 \le \ell < m, \\ \pm 1 & \text{if } \ell = m. \end{cases} \tag{1-4}$$

In order to show (1-4), we use the telescoping sum

$$\sum_{a \le j < a + 2^L} g(j) = s_2(a + 2^L) - s_2(a) = g\left(\left\lfloor \frac{a}{2^L} \right\rfloor\right),$$

where $g(j) = s_2(j+1) - s_2(j)$. This representation yields for $1 \le \ell \le m$, where L is chosen such that $2^L \le m < 2^{L+1}$,

$$\sum_{2 \cdot 2^L - m + \ell \le j < 3 \cdot 2^L - m + \ell} g(j) = g\left(2 + \left\lfloor \frac{-m+\ell}{2^L} \right\rfloor\right) = \begin{cases} g(1) = 0 & \text{if } 1 \le \ell < m, \\ g(2) = 1 & \text{if } \ell = m, \end{cases}$$

$$\sum_{2^L - m + \ell \le j < 2 \cdot 2^L - m + \ell} g(j) = g\left(1 + \left\lfloor \frac{-m+\ell}{2^L} \right\rfloor\right) = \begin{cases} g(0) = 1 & \text{if } 1 \le \ell < m, \\ g(1) = 0 & \text{if } \ell = m, \end{cases}$$

$$\sum_{3 \cdot 2^{L+1} + \ell \le j < 4 \cdot 2^{L+1} + \ell} g(j) = g(3) = -1 \quad \text{for } 1 \le \ell \le m.$$

The first of these three identities yields the "+"-part of (1-4) by choosing $k = 2^L$ and $n_j = 2 \cdot 2^L - m + j$ for $0 \le j < k$.

The "−"-part is obtained from the second and third identities: by considering the disjoint union $J = [2^L - m, 2 \cdot 2^L - m) \cup [3 \cdot 2^{L+1}, 4 \cdot 2^{L+1})$, we have

$$\sum_{j \in J} (s_2(j + \ell) - s_2(j + \ell - 1)) = \begin{cases} 0 & \text{if } 1 \le \ell < m, \\ -1 & \text{if } \ell = m. \end{cases}$$

The statement follows by merging the two intervals and choosing n_j accordingly. This finishes the proof of our theorem.

2. Possible extensions

From our proof, it is possible to effectively construct integers n and t such that $s_2(n + \ell t) - s_2(n) = k_\ell$ for $1 \le \ell \le m$. In particular, this yields integers n and t such that $t_{n+\ell t} = \omega_\ell$ for $1 \le \ell \le m$, where $(\omega_1, \ldots, \omega_m) \in \{0, 1\}^m$ and t is the Thue–Morse sequence on $\{0, 1\}$. (Note that we also used $t(2^\lambda + n) = 1 - t(n)$ for $n < 2^\lambda$.) This gives a constructive result concerning the problem of full arithmetic complexity of the Thue–Morse sequence considered in [Avgustinovich et al. 2003; Konieczny 2019; Müllner and Spiegelhofer 2017].

As an extension of the presented line of research, we are interested in the proportion of cases in which $s_2(n + \ell t) - s_2(n) = k_\ell$ occurs (for $1 \leq \ell \leq m$). For this, we define more generally

$$\delta(\boldsymbol{k}, \boldsymbol{\varepsilon}, t) = \mathrm{dens}\{n \in \mathbb{N} : s_2(n + \ell t + \varepsilon_\ell) - s_2(n) = k_\ell \text{ for } 1 \leq \ell \leq m\},$$

where $\boldsymbol{k} = (k_1, \ldots, k_m) \in \mathbb{Z}^m$ and $\boldsymbol{\varepsilon} = (\varepsilon_1, \ldots, \varepsilon_m) \in \mathbb{N}^m$. This generalizes the array δ defined before. As in the one-dimensional case, the densities in this definition actually exist, and they satisfy the following recurrence relation:

$$\delta(\boldsymbol{k}, \boldsymbol{\varepsilon}, 2t) = \tfrac{1}{2}\delta(\boldsymbol{k}', \boldsymbol{\varepsilon}', t) + \tfrac{1}{2}\delta(\boldsymbol{k}'', \boldsymbol{\varepsilon}'', t),$$

where

$$\begin{aligned}
k'_\ell &= k_\ell - (\varepsilon_\ell \bmod 2), & \varepsilon'_\ell &= \left\lfloor \tfrac{1}{2}\varepsilon_\ell \right\rfloor, \\
k''_\ell &= k_\ell + 1 - ((\varepsilon_\ell + 1) \bmod 2), & \varepsilon''_\ell &= \left\lfloor \tfrac{1}{2}(\varepsilon_\ell + 1) \right\rfloor;
\end{aligned}$$

moreover,

$$\delta(\boldsymbol{k}, \boldsymbol{\varepsilon}, 2t + 1) = \tfrac{1}{2}\delta(\boldsymbol{k}', \boldsymbol{\varepsilon}', t) + \tfrac{1}{2}\delta(\boldsymbol{k}'', \boldsymbol{\varepsilon}'', t),$$

where

$$\begin{aligned}
k'_\ell &= k_\ell - ((\varepsilon_\ell + \ell) \bmod 2), & \varepsilon'_\ell &= \left\lfloor \tfrac{1}{2}(\varepsilon_\ell + \ell) \right\rfloor, \\
k''_\ell &= k_\ell + 1 - ((\varepsilon_\ell + \ell + 1) \bmod 2), & \varepsilon''_\ell &= \left\lfloor \tfrac{1}{2}(\varepsilon_\ell + \ell + 1) \right\rfloor.
\end{aligned}$$

This recurrence is the reason for the introduction of $\boldsymbol{\varepsilon}$ in the definition of $\delta(\boldsymbol{k}, \boldsymbol{\varepsilon}, t)$. To our knowledge, these identities are new; we give the idea of proof. We have

$$\begin{aligned}
\delta(\boldsymbol{k}, \boldsymbol{\varepsilon}, 2t) = \tfrac{1}{2} &\mathrm{dens}\{n \in \mathbb{N} : s(2n + 2\ell t + \varepsilon_\ell) - s(2n) = k_\ell \text{ for } 1 \leq \ell \leq m\} \\
+ \tfrac{1}{2} &\mathrm{dens}\{n \in \mathbb{N} : s(2n + 2\ell t + \varepsilon_\ell + 1) - s(2n + 1) = k_\ell \text{ for } 1 \leq \ell \leq m\},
\end{aligned}$$

and using the identities $s(2n) = s(n)$ and $s(2n + 1) = s(n) + 1$, we obtain

$$s(2n + 2\ell t + \varepsilon_\ell) = s\big(2\big(n + \ell t + \left\lfloor \tfrac{1}{2}\varepsilon_\ell \right\rfloor\big) + (\varepsilon_\ell \bmod 2)\big) = s(n + \ell t + \varepsilon'_\ell) + (\varepsilon_\ell \bmod 2);$$

therefore $s(2n + 2\ell t + \varepsilon_\ell) - s(2n) = k_\ell$ is equivalent to $s(n + \ell t + \varepsilon'_\ell) - s(n) = k'_\ell$. Analogous computations are valid for ε_ℓ replaced by $\varepsilon_\ell + 1$, $\varepsilon_\ell + \ell$, and $\varepsilon_\ell + \ell + 1$ respectively, which yields the claim.

This recurrence can be used to prove statements on the densities $\delta(\boldsymbol{k}, \boldsymbol{\varepsilon}, t)$. The general intuitive idea is that the differences $s(n + jt) - s(n)$, for $1 \leq j \leq m$, should be *almost independent for most t*; in light of this consideration, we consider generalizations of Cusick's conjecture and of the result of [Emme and Hubert 2019]. On the one hand, we may ask for multidimensional generalizations of (1-1), relating the relative sizes of the values $s_2(n), s_2(n + t), \ldots, s_2(n + mt)$ to one another. We propose the following conjecture, extending (1-1).

Conjecture 1. Assume that $m \geq 1$ is an integer. For an integer $t \geq 0$, define

$$c_t^{(m)} = \mathrm{dens}\{n \in \mathbb{N} : s(n) \leq s(n + jt) \text{ for } 1 \leq j \leq m\}.$$

Then for all $t \geq 0$,

$$c_t^{(m)} > \frac{1}{2^m}.$$

The statement is wrong for any larger constant in place of $1/2^m$. Also, define

$$C_t^{(m)} = \mathrm{dens}\{n \in \mathbb{N} : s(n) \leq s(n + t) \leq s(n + 2t) \leq \cdots \leq s(n + mt)\}.$$

Then for all $t \geq 0$,

$$C_t^{(m)} > \frac{1}{2^m m!}.$$

The constant $1/(2^m m!)$ is maximal.

On the other hand, we could ask for the overall shape of the m-dimensional probability distribution defined by $\delta(\,\cdot\,, \boldsymbol{\varepsilon}, t)$.

Problem 1. Prove a multidimensional generalization of the theorem in [Emme and Hubert 2019]: for most t, the densities dens$\{n \in \mathbb{N} : s_2(n + \ell t) - s_2(n) = k_\ell$ for $1 \leq \ell \leq m\}$ should define a probability distribution that is close to a multivariate Gaussian law.

We can now understand the intuition behind the constants in Conjecture 1: a bivariate normal distribution with mean $(0, 0)$ has one quarter of its total weight in the quadrant $\{(x, y) \in \mathbb{R}^2 : 0 \leq x, \, 0 \leq y\}$, and analogous considerations hold for higher dimensions. Concerning the values $C_t^{(2)}$, a bivariate normal distribution with mean $(0, 0)$ has one *eighth* of its total weight in the octant $\{(x, y) \in \mathbb{R}^2 : 0 \leq x \leq y\}$, which corresponds to the complex closed region $\left\{re^{2\pi i x} : r \geq 0, \, \frac{1}{8} \leq x \leq \frac{1}{4}\right\}$. Noting the fact that $s(n + 2t) \geq s(n + t)$ if and only if $s(n + 2t) - s(n) \geq s(n + t) - s(n)$, we see the link between the densities $\delta(\boldsymbol{k}, \boldsymbol{\varepsilon}, t)$ with $\boldsymbol{k}$ lying in this octant and the values $C_t^{(2)}$. Concerning higher dimensions, we note that the $m!$ sets

$$\{x \in \mathbb{R}^m : \|x\| \leq 1, \, 0 \leq x_{\sigma(1)} \leq x_{\sigma(2)} \leq \cdots \leq x_{\sigma(m)}\},$$

where σ is a permutation of $\{1, \ldots, m\}$, are unions of line segments $[0, z]$, where $\|z\| = 1$; they are pairwisely congruent (using a linear transformation that is a permutation matrix in the canonical base); an intersection of two distinct sets of this form has measure zero, and their union makes up a fraction $1/2^m$ of the unit ball in m dimensions.

We implemented the computation of the densities $\delta(\boldsymbol{k}, \boldsymbol{\varepsilon}, t)$ for two dimensions in the Sage computer algebra system [SageMath 2017], and the resulting Sage worksheet is available from the website of Spiegelhofer.[1] We obtain

$$c_t^{(2)} \geq c_{951}^{(2)} = \frac{94299}{262144} = 0.3597\ldots > \frac{1}{4} \quad \text{for } t \leq 2^{10}$$

and

$$C_t^{(2)} \geq c_{991}^{(2)} = \frac{43947}{262144} = 0.1676\ldots > \frac{1}{8} \quad \text{for } t \leq 2^{10}.$$

The implementation involves nine two-dimensional arrays of rational numbers (corresponding to the nine possibilities $(\varepsilon_1, \varepsilon_2) \in \{0, 1, 2\}^2$) and each of these calculations took about 5 minutes on a standard machine. We note that we did not optimize the Sage code, and certainly this computation can be sped up significantly.

Other conjectures similar to Conjecture 1 are conceivable: what about the other octants and quadrants A in the plane (including the borders)? Is it always true that $\frac{1}{8}$, resp. $\frac{1}{4}$, is a lower bound for the density dens$\{n \in \mathbb{N} : (s(n + t) - s(n), \, s(n + 2t) - s(n)) \in A\}$? We leave this question open for future discussion, but we note that the analogous problem in one dimension is true for almost all t: we also have

$$\tilde{c}_t = \text{dens}\{n \in \mathbb{N} : s(n + t) \leq s(n)\} > \frac{1}{2}$$

---

[1]https://dmg.tuwien.ac.at/spiegelhofer/

for t in a set having asymptotic density 1 [Drmota et al. 2016]. In other words, usually the median of the probability distribution defined by $\delta(\,\cdot\,, t)$ is very close to the mean value (which is 0). We believe that $\tilde{c}_t \geq \frac{1}{2}$ is true for all t, which complements Cusick's conjecture (1-1) (note that $\tilde{c}_1 = \frac{1}{2}$; thus there is no strict inequality).

We expect that nontrivial statements on both Conjecture 1 and Problem 1, at least for small m, can be obtained by extending the study of moments set forward in [Emme and Hubert 2019]. This is certainly not easy and will introduce technical difficulties that have to be surmounted.

References

[Avgustinovich et al. 2003] S. V. Avgustinovich, D. G. Fon-Der-Flaass, and A. E. Frid, "Arithmetical complexity of infinite words", pp. 51–62 in *Words, languages & combinatorics, III* (Kyoto, 2000), edited by M. Ito and T. Imaoka, World Sci. Publ., River Edge, NJ, 2003. MR

[Bésineau 1972] J. Bésineau, "Indépendance statistique d'ensembles liés à la fonction "somme des chiffres"", *Acta Arith.* **20** (1972), 401–416. MR Zbl

[Drmota et al. 2016] M. Drmota, M. Kauers, and L. Spiegelhofer, "On a conjecture of Cusick concerning the sum of digits of n and $n+t$", *SIAM J. Discrete Math.* **30**:2 (2016), 621–649. MR Zbl

[Emme and Hubert 2018] J. Emme and P. Hubert, "Normal distribution of correlation measures of binary sum-of-digits functions", preprint, 2018. arXiv

[Emme and Hubert 2019] J. Emme and P. Hubert, "Central limit theorem for probability measures defined by sum-of-digits function in base 2", *Ann. Sc. Norm. Super. Pisa Cl. Sci.* (5) **19**:2 (2019), 757–780. MR

[Emme and Prikhod'ko 2017] J. Emme and A. Prikhod'ko, "On the asymptotic behavior of density of sets defined by sum-of-digits function in base 2", *Integers* **17** (2017), art. id. A58. MR Zbl

[Konieczny 2019] J. Konieczny, "Gowers norms for the Thue–Morse and Rudin–Shapiro sequences", *Ann. Inst. Fourier* (*Grenoble*) **69**:4 (2019), 1897–1913. MR Zbl

[Müllner and Spiegelhofer 2017] C. Müllner and L. Spiegelhofer, "Normality of the Thue–Morse sequence along Piatetski–Shapiro sequences, II", *Israel J. Math.* **220**:2 (2017), 691–738. MR Zbl

[SageMath 2017] The Sage Developers, *SageMath, the Sage Mathematics Software System*, Version 8.1, 2017, available at http://www.sagemath.org.

[Sanna 2012] C. Sanna, "On arithmetic progressions of integers with a distinct sum of digits", *J. Integer Seq.* **15**:8 (2012), art. id. 12.8.1. MR Zbl

[Spiegelhofer 2019] L. Spiegelhofer, "Approaching Cusick's conjecture on the sum-of-digits function", *Integers* **19** (2019), art. id. A59. Zbl

LUKAS SPIEGELHOFER:

lukas.spiegelhofer@tuwien.ac.at
Institute of Discrete Mathematics and Geometry, Vienna University of Technology, Vienna, Austria

THOMAS STOLL:

thomas.stoll@univ-lorraine.fr
Institut Élie Cartan, Université de Lorraine, Vandœuvre-lès-Nancy, France

Moscow Journal of Combinatorics and Number Theory

See inside back cover or msp.org/moscow for submission instructions.