# ON THE DISTRIBUTION OF THE SUM OF DIGITS OF SUMS a + b

Christian Mauduit, Joel Rivat, András Sárközy

# ON THE DISTRIBUTION OF THE SUM OF DIGITS OF SUMS $a+b$

CHRISTIAN MAUDUIT, JOËL RIVAT, AND ANDRÁS SÁRKÖZY

Abstract. Let $\mathcal{A}$, $\mathcal{B}$ be large subsets of $\{1,\dots,N\}$. We study the distribution of the sum of binary digits of the sums $a+b$ with $(a,b)\in\mathcal{A}\times\mathcal{B}$.

## 1. Introduction

Throughout this paper we will use the following notations: $\mathbb{N}$, $\mathbb{N}_0$, $\mathbb{Z}$, $\mathbb{R}$ and $\mathbb{C}$ denote the set of positive integers, non-negative integers, integers, real numbers, resp. complex numbers, $\|x\|$ denotes the distance from $x$ to the nearest integer and we write $\mathrm{e}(\alpha)=e^{2i\pi\alpha}$. We will denote the sum of digits of an integer $n\geqslant 0$ written in base $g$ by $\mathrm{s}_g(n)$ and will write $\mathrm{s}_2(n)=\mathrm{s}(n)$.

Many papers have been written on the arithmetic properties of sumsets of "dense" sets of positive integers. A survey of the early work in this field is presented in [7]. In particular, in [5] the first and third author showed that if $\mathcal{A}$, $\mathcal{B}$ are "large" subsets of $\{1,2,\dots,N\}$, $g\in\mathbb{N}$ is fixed, $(m,g-1)=1$ and $m$ is "small", then the values of the sum of digit function $s_g(n)$ assumed over the sums $a+b$ (with $a\in\mathcal{A}$, $b\in\mathcal{B}$) are well-distributed modulo $m$:

**Theorem A.** *If $g\in\mathbb{N}$, $g\geqslant 2$, $m\in\mathbb{N}$, $(m,g-1)=1$, $r\in\mathbb{Z}$ and $\mathcal{A}$, $\mathcal{B}\subset\{1,2,\dots,N\}$, then we have*

$$\left|\{(a,b)\in\mathcal{A}\times\mathcal{B},\ s_g(a+b)\equiv r \bmod m\}-\frac{|\mathcal{A}|\,|\mathcal{B}|}{m}\right|\leqslant 2\,\gamma\,N^{\lambda}\,\left(|\mathcal{A}|\,|\mathcal{B}|\right)^{1/2},$$

*where $\lambda=\lambda(g,m)$ and $\gamma=\gamma(g,m)$ are defined by*

$$\lambda=\frac{1}{2\log g}\log\frac{g\sin(\pi/2m)}{\sin(\pi/2mg)}\ (<1),\ \ \gamma=\gamma(g,m)=\frac{g^2}{g^{\lambda}-1}.$$

In another paper [6] we formulated the conjecture that if $g\in\mathbb{N}$ is fixed and $\mathcal{A}$, $\mathcal{B}$ are "large" subsets of $\{1,2,\dots,N\}$, then there are $a\in\mathcal{A}$, $b\in\mathcal{B}$ such that their sum of digits $s_g(a+b)$ is equal to its expected value:

**Conjecture 1.** *If $\varepsilon>0$, $N>N_0(\varepsilon)$, $\mathcal{A}$, $\mathcal{B}\in\{1,2,\dots,N\}$ and $|\mathcal{A}|$, $|\mathcal{B}|>\varepsilon N$, then there are integers $a$, $b$ such that $a\in\mathcal{A}$, $b\in\mathcal{B}$ and*

$$s_g(a+b)=\lfloor(g-1)\nu/2\rfloor$$

*where $\nu=\nu(N)\in\mathbb{N}$ is defined by $g^{\nu}\leqslant N\leqslant g^{\nu+1}-1$.*

---

*Date*: August 28, 2017.

2010 *Mathematics Subject Classification.* 11A63, 11B13.

*Key words and phrases.* sumset, digits.

Research partially supported by the Hungarian National Research Development and Innovation Funds NK104183 and K119528, and the Agence Nationale de la Recherche project ANR-14-CE34-0009 MUDERA.

Later in [4] we proved this conjecture in the $g = 2$ special case in a slightly stronger form:

**Theorem B.** *For any $L > 0$ and $\varepsilon > 0$ there is a number $N_0 = N_0(L, \varepsilon)$ such that if $N \in \mathbb{N}$, $N > N_0$, $k \in \mathbb{N}$,*

$$\left| k - \frac{\log N}{2 \log 2} \right| < L (\log N)^{1/4}$$

*and*

$$\mathcal{A},\ \mathcal{B} \in \{1, 2, \dots, N\},$$

*then writing $\varrho = \left(\frac{\log 2}{8}\right)^{1/2}$, we have*

$$\left| \left|\{(a,b) \in \mathcal{A} \times \mathcal{B},\ s(a+b) = k\}\right| - \left(\frac{\log 4}{\pi}\right)^{1/2} \frac{|\mathcal{A}|\,|\mathcal{B}|}{(\log N)^{1/2}} \right|$$
$$< \frac{N}{(\log N)^{1/2} \exp((\varrho - \varepsilon)(\log\log N)^{1/2})} \left(|\mathcal{A}|\,|\mathcal{B}|\right)^{1/2}.$$

In this paper our goal is to extend the study of the distribution of the numbers $\mathrm{s}(a + b)$ from a small neighbourhood of the expedted value to a possibly large interval. We will prove the following theorem:

**Theorem 1.** *For $N \in \mathbb{N}$, $N > 2$, $\mathcal{A}$, $\mathcal{B} \in \{1, 2, \dots, N\}$, $0 < z < \frac{\log 2N}{\log 2}$ define $y_{2N} = y_{2N}(z)$ by*

$$z = \frac{\log 2N}{\log 4} + y_{2N} \cdot \frac{1}{2} \left(\frac{\log 2N}{\log 2}\right)^{1/2}.$$

*Then we have*

$$\text{(1)} \qquad \frac{1}{|\mathcal{A}|\,|\mathcal{B}|} \left|\{(a,b) \in \mathcal{A} \times \mathcal{B},\ s(a+b) < z\}\right|$$
$$= \Phi(y_{2N}) + O\left( \frac{N}{\sqrt{|\mathcal{A}|\,|\mathcal{B}|}} \frac{(\log\log N)(\log\log\log N)^{1/2}}{(\log N)^{1/4}} \right)$$

*uniformly in $z$ as $N$ tends to infinity, with $\Phi$ defined by*

$$\Phi(u) = (2\pi)^{-1/2} \int_{-\infty}^{u} e^{-t^2/2}\, dt.$$

This theorem shows that for

$$\text{(2)} \qquad \sqrt{|\mathcal{A}|\,|\mathcal{B}|} \Big/ \frac{N \log\log N}{(\log N)^{1/4}} \to +\infty$$

the numbers $\mathrm{s}(a + b)$ with $(a, b) \in \mathcal{A} \times \mathcal{B}$ are distributed in the same way as the numbers $\mathrm{s}(n)$ with $n \leqslant 2N$. (Indeed, compare this theorem with the statement of Lemma 2.) Probably condition (2) is not sharp. However, condition (2) cannot be replaced by

$$\text{(3)} \qquad \sqrt{|\mathcal{A}|\,|\mathcal{B}|} \gg \frac{N}{2^{\lfloor \sqrt{\log N} \rfloor}}$$

as the following example shows: let

$$\mathcal{A} = \mathcal{B} = \left\{ n \leqslant N :\ 2^{\lfloor \sqrt{\log N} \rfloor} \mid n \right\}.$$

Namely, for this sequences $\mathcal{A}$, $\mathcal{B}$ (3) holds, however, for $(a,b)\in\mathcal{A}\times\mathcal{B}$ the numbers $\mathrm{s}(a+b)$ tend to be smaller than expected by $c\sqrt{\log N}$ (since the last $\lfloor\sqrt{\log N}\rfloor$ digits of the sums $a+b$ do not contribute to $\mathrm{s}(a+b)$).

## 2. Sketch of the proof

We will use the circle method (in similar manner as in [2]). Let us write

$$G(\alpha)=\sum_{a\in\mathcal{A}}\mathrm{e}(a\alpha),\ \ H(\alpha)=\sum_{b\in\mathcal{B}}\mathrm{e}(b\alpha),$$

$$S_z(\alpha)=\sum_{\substack{n\leqslant 2N\\ \mathrm{s}(n)<z}}\mathrm{e}(n\alpha) \tag{4}$$

and

$$J=\int_0^1 G(\alpha)\,H(\alpha)\,S_z(-\alpha)\,d\alpha. \tag{5}$$

Then clearly we have

$$\begin{aligned} J&=\int_0^1\sum_{a\in\mathcal{A}}\sum_{b\in\mathcal{B}}\sum_{\substack{n\leqslant 2N\\ \mathrm{s}(n)<z}}\mathrm{e}((a+b-n)\alpha)\,d\alpha \\ &=\sum_{(a,b)\in\mathcal{A}\times\mathcal{B}}\ \sum_{\substack{n\leqslant 2N\\ \mathrm{s}(n)<z\\ a+b=n}}1=|\{(a,b)\in\mathcal{A}\times\mathcal{B},\ \mathrm{s}(a+b)<z\}|\end{aligned} \tag{6}$$

so that, indeed, in order to estimate the left hand side of (1) we have to estimate this integral $J$. The estimate of $J$ can be reduced to the estimate of the generating function $S_z(\alpha)$ which will be carried out in sections 3 and 4 and finally in section 5 we will complete the proof of the theorem by combining the lemmas proved in sections 3 and 4.

## 3. The estimate of $S_z(\alpha)$

For $x>0$ and $z>0$ let

$$\mathcal{R}_z(x)=\{n\in\mathbb{N}:\ n\leqslant x,\ \mathrm{s}(n)<z\} \tag{7}$$

and

$$R_z(x)=|\mathcal{R}_z(x)|=|\{n\in\mathbb{N}:\ n\leqslant x,\ \mathrm{s}(n)<z\}|\,. \tag{8}$$

Note that by using this notation the definition of $S_z(\alpha)$ in (4) can be rewritten as

$$S_z(\alpha)=\sum_{n\in\mathcal{R}_z(2N)}\mathrm{e}(n\alpha). \tag{9}$$

For $x>0$ we write

$$M_x=\frac{\log x}{\log 4},\ \ D_x=\frac{1}{2}\left(\frac{\log x}{\log 2}\right)^{1/2},$$

let

$$\mathcal{N}_x(y)=\{n:\ n\leqslant x,\ \mathrm{s}(n)<M_x+yD_x\},$$

and write

$$N_x(y)=|\mathcal{N}_x(y)|=|\{n:\ n\leqslant x,\ \mathrm{s}(n)<M_x+yD_x\}|\,.$$

The estimate of $S_z(\alpha)$ near 0 will be based on a theorem of Kátai and Mogyoródi [3]. Here we state their result in the special case when the base of the number system (denoted by them as $K$, and denoted above by us as $g$) is 2. They proved:

**Lemma 1.** *We have*

$$N_x(y) = x\,\Phi(y) + O\left(\frac{x \log\log x}{(\log x)^{1/2}}\right)$$

*uniformly in $y$ as $x$ tends to infinity.*

*Proof.* This is the $K = 2$ special case of Theorem 1 in [3]. □

**Lemma 2.** *For*

$$0 \leqslant z \leqslant \log 2N, \tag{10}$$

$$\frac{N}{(\log N)^{1/2}} < x \leqslant 2N \tag{11}$$

*we have*

$$R_z(x) = |\{n :\ n \leqslant x,\ \mathrm{s}(n) < z\}| = x\,\Phi(y_{2N}) + O\left(\frac{N \log\log N}{(\log N)^{1/2}}\right) \tag{12}$$

*where $y_{2N}$ is defined by*

$$y_{2N} = \frac{z - M_{2N}}{D_{2N}} \tag{13}$$

*Proof.* Writing

$$y_x = \frac{z - M_x}{D_x}, \tag{14}$$

by Lemma 1, (11) and (13), we have

$$\begin{aligned} R_z(x) &= |\{n :\ n \leqslant x,\ \mathrm{s}(n) < z\}| = |\{n :\ n \leqslant x,\ \mathrm{s}(n) < M_x + y_x D_x\}| \\ &= N_x(y_x) = x\Phi(y_x) + O\left(\frac{x \log\log x}{(\log x)^{1/2}}\right) = x\Phi(y_x) + O\left(\frac{N \log\log N}{(\log N)^{1/2}}\right). \end{aligned} \tag{15}$$

Observing by (11) that $\log x \gg \log N$,

$$\frac{1}{D_x} - \frac{1}{D_{2N}} = (\log 2)^{1/2} \int_{\log x}^{\log 2N} \frac{dt}{t^{3/2}} \ll \frac{\log(2N/x)}{(\log x)^{3/2}} \ll \frac{\log\log N}{(\log N)^{3/2}}$$

and

$$\frac{M_{2N}}{D_{2N}} - \frac{M_x}{D_x} = \frac{1}{(\log 2)^{1/2}} \int_{\log x}^{\log 2N} \frac{dt}{2\,t^{1/2}} \ll \frac{\log(2N/x)}{(\log x)^{1/2}} \ll \frac{\log\log N}{(\log N)^{1/2}},$$

it follows from (10), (13), (14) and (15) that

$$y_{2N} - y_x = z\left(\frac{1}{D_{2N}} - \frac{1}{D_x}\right) - \left(\frac{M_{2N}}{D_{2N}} - \frac{M_x}{D_x}\right) = O\left(\frac{\log\log N}{(\log N)^{1/2}}\right). \tag{16}$$

By the Lagrange mean value theorem there is a real number $\xi$ between $y_x$ and $y_{2N}$ such that

$$\left|\frac{\Phi(y_{2N}) - \Phi(y_x)}{y_{2N} - y_x}\right| = |\Phi'(\xi)| = \frac{1}{\sqrt{2\pi}} e^{-\xi^2/2} \leqslant \frac{1}{\sqrt{2\pi}}$$

whence, by (16),

$$|\Phi(y_{2N}) - \Phi(y_x)| \leqslant \frac{1}{\sqrt{2\pi}}\,|y_{2N} - y_x| = O\left(\frac{\log\log N}{(\log N)^{1/2}}\right)$$

so that

$$\Phi(y_{2N}) = \Phi(y_x) + O\left(\frac{\log\log N}{(\log N)^{1/2}}\right). \tag{17}$$

(12) follows from (15) and (17) which completes the proof of Lemma 2. □

Write

$$u = \frac{(\log N)^{1/4}(\log\log\log N)^{1/2}}{N}. \tag{18}$$

**Lemma 3.** *For*

$$0 \leqslant z \leqslant 2N \tag{19}$$

*and*

$$\|\alpha\| \leqslant u \tag{20}$$

*we have*

$$\left|S_z(\alpha) - \Phi(y_{2N})\sum_{n=1}^{2N} \mathrm{e}(n\alpha)\right| = O\left(\frac{u\,N^2\,\log\log N}{(\log N)^{1/2}}\right). \tag{21}$$

*Proof.* By (7), (8), (9) and partial summation we have, writing $N_0 = \left\lfloor 2N(\log N)^{-1/2}\right\rfloor$,

$$\begin{aligned} S_z(\alpha) = \sum_{n\in\mathcal{R}_z(2N)} \mathrm{e}(n\alpha) &= \sum_{n\in\mathcal{R}_z(N_0)} \mathrm{e}(n\alpha) + \sum_{n=N_0+1}^{2N} (R_z(n) - R_z(n-1))\,\mathrm{e}(n\alpha) \\ &= O(N_0) + \sum_{n=N_0}^{2N} R_z(n)\,(\mathrm{e}(n\alpha) - \mathrm{e}((n+1)\alpha)) + R_z(2N)\,\mathrm{e}((2N+1)\alpha). \end{aligned}$$

By (19), for $N_0 \leqslant n \leqslant 2N$ we can apply (12) and we get

$$S_z(\alpha) = O(N_0) + \Phi(y_{2N})\left(\sum_{n=N_0}^{2N} n\,(\mathrm{e}(n\alpha) - \mathrm{e}((n+1)\alpha)) + 2N\,\mathrm{e}((2N+1)\alpha)\right) + E_z(\alpha)$$

with

$$E_z(\alpha) = O\left(\frac{N\log\log N}{(\log N)^{1/2}}\right)\left(\sum_{n=N_0}^{2N} |1 - \mathrm{e}(\alpha)| + 1\right) = O\left(\frac{(\|\alpha\|\,N + 1)N\log\log N}{(\log N)^{1/2}}\right),$$

so that, since $0 \leqslant \Phi \leqslant 1$,

$$S_z(\alpha) = \Phi(y_{2N})\sum_{n=1}^{2N} \mathrm{e}(n\alpha) + O\left(\frac{N\log\log N}{(\log N)^{1/2}}\right) + O\left(\frac{u\,N^2\log\log N}{(\log N)^{1/2}}\right). \tag{22}$$

(21) follows from (22) and (18). □

## 4. The estimate of $S_z(\alpha)$ for large $\|\alpha\|$

For $(\theta,\alpha)\in\mathbb{R}^2$ and $\lambda\in\mathbb{Z}$ with $\lambda\geqslant 0$ let

$$F_\lambda(\theta,\alpha)=2^{-\lambda}\sum_{0\leqslant n<2^\lambda}\mathrm{e}(\mathrm{s}(n)\theta+n\alpha).$$

For $\lambda\geqslant 0$, we have the trivial upper bound

$$|F_\lambda(\theta,\alpha)|\leqslant 1 \tag{23}$$

and for $\lambda\geqslant 1$,

$$F_\lambda(\theta,\alpha)=2^{-\lambda}\prod_{j=0}^{\lambda-1}\left(1+\mathrm{e}(\theta+2^j\alpha)\right).$$

For $\alpha\in\mathbb{Z}$ we have $F_\lambda(0,\alpha)=1$ and for $\alpha\in\mathbb{R}\setminus\mathbb{Z}$ we have

$$|F_\lambda(0,\alpha)|=2^{-\lambda}\left|\frac{\sin\pi 2^\lambda\alpha}{\sin\pi\alpha}\right|\leqslant\left(2^\lambda\,\sin\pi\,\|\alpha\|\right)^{-1}. \tag{24}$$

For $0\leqslant\mu\leqslant\lambda$, $0\leqslant m<2^\mu$ and $n\geqslant 0$ we have $\mathrm{s}(m+2^\mu n)=\mathrm{s}(m)+\mathrm{s}(n)$ and

$$F_\lambda(\theta,\alpha)=F_\mu(\theta,\alpha)\,F_{\lambda-\mu}(\theta,2^\mu\alpha). \tag{25}$$

For $\lambda\geqslant 1$ we have

$$\frac{\partial F_\lambda}{\partial\theta}(\theta,\alpha)=2^{-\lambda}\sum_{i=0}^{\lambda-1}2i\pi\,\mathrm{e}(\theta+2^i\alpha)\prod_{\substack{j=0\\ j\neq i}}^{\lambda-1}\left(1+\mathrm{e}(\theta+2^j\alpha)\right),$$

so that for $(\theta,\alpha)\in\mathbb{R}^2$, $\lambda\geqslant 1$ we have the elementary upper bound

$$\left|\frac{\partial F_\lambda}{\partial\theta}(\theta,\alpha)\right|\leqslant\pi\lambda,$$

which, by the mean value theorem, gives (for $\lambda=0$ and $(\theta,\alpha)\in\mathbb{R}^2$ we have $F_\lambda(\theta,\alpha)=1$) for any integer $\lambda\geqslant 0$

$$|F_\lambda(\theta,\alpha)|\leqslant|F_\lambda(0,\alpha)|+\pi\theta\lambda \tag{26}$$

**Lemma 4.** *Let $\lambda\geqslant 0$ be an integer and $(\theta,\alpha)\in\mathbb{R}^2$. We have*

$$|F_\lambda(\theta,\alpha)|\leqslant e^{c/4}e^{-c\|\theta\|^2\lambda}, \tag{27}$$

*where*

$$c=\pi^2/20. \tag{28}$$

*Proof.* This is Lemma 3.2 of [4]. □

**Lemma 5.** *Let $\nu\geqslant 1$ be an integer and $\alpha\in\mathbb{R}$. For $0\leqslant\theta_0\leqslant\frac{1}{2}$, and $c$ defined by* (28) *we have*

$$\int_{\|\theta\|\geqslant\theta_0}|F_\nu(\theta,\alpha)|\,d\theta\leqslant\sqrt{\pi}\,e^{c/4}\frac{e^{-c\theta_0^2\nu}}{\sqrt{c\nu}}. \tag{29}$$

*Proof.* This is Lemma 3.3 of [4]. □

**Lemma 6.** *Let $\nu_1\geqslant 0$ be an integer and $\alpha\in\mathbb{R}$ such that $2^{-2-\nu_1}\leqslant\|\alpha\|\leqslant 2^{-1-\nu_1}$. For any $\theta\in\mathbb{R}$ and any integer $\nu\geqslant\nu_1$ we have*

$$|F_\nu(\theta,\alpha)|\leqslant\left(2^{\frac{1}{2}-\nu+\nu_1}+\pi\,|\theta|\,(\nu-\nu_1)\right)|F_{\nu_1}(\theta,\alpha)|. \tag{30}$$

*Proof.* By (25) we have

$$F_\nu(\theta,\alpha) = F_{\nu_1}(\theta,\alpha) F_{\nu-\nu_1}(\theta, 2^{\nu_1}\alpha),$$

and writing $\alpha = n+\beta$ with $2^{-2-\nu_1} \leqslant |\beta| \leqslant 2^{-1-\nu_1}$ we have $2^{\nu_1}\alpha = 2^{\nu_1} n + 2^{\nu_1}\beta$ with $2^{-2} \leqslant |2^{\nu_1}\beta| \leqslant 2^{-1}$ so that $\|2^{\nu_1}\alpha\| = |2^{\nu_1}\beta| \geqslant \frac{1}{4}$, and by (26) and (24) it follows

$$|F_{\nu-\nu_1}(\theta, 2^{\nu_1}\alpha)| \leqslant 2^{\frac{1}{2}-\nu+\nu_1} + \pi\theta(\nu-\nu_1)$$

and we get (30). □

**Lemma 7.** *For real numbers $U > 2$, $U^{-1} < \theta_1 \leqslant \frac{1}{2}$, integers $1 \leqslant \nu_1 \leqslant \nu$, and $\alpha \in \mathbb{R}$ such that $2^{-2-\nu_1} \leqslant \|\alpha\| \leqslant 2^{-1-\nu_1}$, we have*

$$(31) \qquad \int_{-1/2}^{1/2} \min\left(U, |\sin\pi\theta|^{-1}\right) |F_\nu(\theta,\alpha)|\, d\theta$$
$$\ll 2^{\nu_1-\nu}(1+\log(U\theta_1)) + (\nu-\nu_1)(U^{-1} + \nu_1^{-1/2}) + e^{-c\theta_1^2\nu} \log\frac{1}{2\theta_1}.$$

*where the implied constant is absolute.*

*Proof.* Let $\theta_0 = U^{-1}$, so that $0 < \theta_0 \leqslant \frac{1}{2}$. For $|\theta| \leqslant \theta_0$, since $\min\left(U, |\sin\pi\theta|^{-1}\right) \leqslant U = \theta_0^{-1}$, combining (30) with (23) applied with $\lambda = \nu_1$ and using parity we have

$$\int_{-\theta_0}^{\theta_0} \min\left(U, |\sin\pi\theta|^{-1}\right) |F_\nu(\theta,\alpha)|\, d\theta \leqslant 2U \int_0^{\theta_0} \left(2^{\frac{1}{2}-\nu+\nu_1} + \pi\theta(\nu-\nu_1)\right) d\theta$$

and by integration we get

$$\int_{-\theta_0}^{\theta_0} \min\left(U, |\sin\pi\theta|^{-1}\right) |F_\nu(\theta,\alpha)|\, d\theta \leqslant 2^{\frac{3}{2}+\nu_1-\nu} + \pi\, U^{-1}\,(\nu-\nu_1).$$

We have

$$\int_{\theta_0 \leqslant \|\theta\| \leqslant \theta_1} \min\left(U, |\sin\pi\theta|^{-1}\right) |F_\nu(\theta,\alpha)|\, d\theta \ll \int_{\theta_0 \leqslant \|\theta\| \leqslant \theta_1} |F_\nu(\theta,\alpha)| \frac{d\theta}{\theta}$$

and combining (30) and (23) the right hand side above is at most

$$2^{\frac{1}{2}-\nu+\nu_1} \int_{\theta_0 \leqslant \|\theta\| \leqslant \theta_1} \frac{d\theta}{\theta} + \pi(\nu-\nu_1) \int_{\theta_0 \leqslant \|\theta\| \leqslant \theta_1} |F_{\nu_1}(\theta,\alpha)|\, d\theta.$$

By (29) applied with $\nu_1$ in place of $\nu$ we obtain

$$\int_{\theta_0 \leqslant \|\theta\| \leqslant \theta_1} \min\left(U, |\sin\pi\theta|^{-1}\right) |F_\nu(\theta,\alpha)|\, d\theta \ll 2^{\nu_1-\nu} \log(\theta_1/\theta_0) + \frac{\nu-\nu_1}{\sqrt{\nu_1}}.$$

Observing that

$$\int_{\theta_1}^{1/2} e^{-c\theta^2\nu} \frac{d\theta}{\theta} \leqslant e^{-c\theta_1^2\nu} \int_{\theta_1}^{1/2} \frac{d\theta}{\theta} = e^{-c\theta_1^2\nu} \log\frac{1}{2\theta_1}.$$

by (27) we get

$$\int_{\theta_1 \leqslant \|\theta\| \leqslant 1/2} \min\left(U, |\sin\pi\theta|^{-1}\right) |F_\nu(\theta,\alpha)|\, d\theta \ll e^{-c\theta_1^2\nu} \log\frac{1}{2\theta_1}.$$

Combining these estimates leads to Lemma 7. □

**Lemma 8.** *For integers $0 \leqslant \nu_1 < \nu$, $(\theta, \alpha) \in \mathbb{R}^2$ such that $\|\theta\| < \frac{1}{4}$ and $2^{-2-\nu_1} \leqslant \|\alpha\| \leqslant 2^{-1-\nu_1}$ and $c$ defined by (28) we have*

$$|F_\nu(\theta,\alpha)| \ll \|\theta\|\, e^{-c\|\theta\|^2\nu} + 2^{\nu_1-\nu} + \exp\left(-\sigma(\theta)\sqrt{\nu-\nu_1}\right), \tag{32}$$

*where $\sigma(\theta) = \sqrt{-\frac{1}{2}(\log 2)\log\left(\sin\pi(\|\theta\|+\frac{1}{4})\right)} = \frac{\log 2}{2} + O(\|\theta\|)$.*

*Proof.* This is Lemma 3.4 of [4] with $\nu_1 + 2$ in place of $\nu_1$ and therefore a modified absolute implied constant. The range of $\|\alpha\|$ is extended by continuity. □

**Lemma 9.** *For integers $0 \leqslant \nu_1 < \nu$, $(\theta_0, \alpha) \in \mathbb{R}^2$ such that $0 < \theta_0 < \frac{1}{4}$ and $2^{-2-\nu_1} \leqslant \|\alpha\| \leqslant 2^{-1-\nu_1}$, and $c$ defined by (28), we have*

$$\int_{\|\theta\|\leqslant\theta_0} |F_\nu(\theta,\alpha)|\, d\theta \ll \frac{1-e^{-c\theta_0^2\nu}}{\nu} + \theta_0 2^{\nu_1-\nu} + \theta_0\exp\left(-\sigma(\theta_0)\sqrt{\nu-\nu_1}\right).$$

*Proof.* This is Lemma 3.5 of [4] with $\nu_1 + 2$ in place of $\nu_1$ and therefore a modified absolute implied constant. The range of $\|\alpha\|$ is extended by continuity. □

**Lemma 10.** *For integers $0 \leqslant \nu_1 < \nu$, $\alpha \in \mathbb{R}$ such that $2^{-2-\nu_1} \leqslant \|\alpha\| < 2^{-1-\nu_1}$ and $U \geqslant 5$ we have*

$$\int_{-1/2}^{1/2} \min\left(U, |\sin\pi\theta|^{-1}\right) |F_\nu(\theta,\alpha)|\, d\theta \ll \frac{1}{U} + \frac{1}{\nu^{1/2}} + \exp\left(-\sigma(1/5)\sqrt{\nu-\nu_1}\right)\log U. \tag{33}$$

*Proof.* Let $\theta_0 = U^{-1}$, so that $0 < \theta_0 \leqslant \frac{1}{5}$. Since $\min\left(U, |\sin\pi\theta|^{-1}\right) \leqslant U = \theta_0^{-1}$, by Lemma 9 we have

$$\int_{-\theta_0}^{\theta_0} \min\left(U, |\sin\pi\theta|^{-1}\right) |F_\nu(\theta,\alpha)|\, d\theta \ll \frac{1-e^{-c\theta_0^2\nu}}{\theta_0\nu} + 2^{\nu_1-\nu} + \exp\left(-\sigma(\theta_0)\sqrt{\nu-\nu_1}\right),$$

and by the mean value theorem and the monoticity of $\sigma$ we get

$$\int_{-\theta_0}^{\theta_0} \min\left(U, |\sin\pi\theta|^{-1}\right) |F_\nu(\theta,\alpha)|\, d\theta \ll \theta_0 + 2^{\nu_1-\nu} + \exp\left(-\sigma(1/5)\sqrt{\nu-\nu_1}\right).$$

By Lemma 8 we have

$$\int_{\theta_0\leqslant\|\theta\|\leqslant 1/5} \min\left(U, |\sin\pi\theta|^{-1}\right) |F_\nu(\theta,\alpha)|\, d\theta$$
$$\ll \int_{\theta_0}^{1/5} e^{-c\theta^2\nu} d\theta + \left(2^{\nu_1-\nu} + \exp\left(-\sigma(1/5)\sqrt{\nu-\nu_1}\right)\right)\int_{\theta_0}^{1/5} \theta^{-1} d\theta,$$

and writing $\theta = \theta_0 + t$ we have

$$\int_{\theta_0}^{1/5} e^{-c(\theta^2-\theta_0^2)\nu} d\theta \leqslant \int_0^{+\infty} e^{-c(t^2+2\theta_0 t)\nu} dt \leqslant \int_0^{+\infty} e^{-ct^2\nu} dt = \frac{\sqrt{\pi}}{2\sqrt{c\nu}},$$

so that the quantity above is

$$\ll \nu^{-1/2} e^{-c\theta_0^2\nu} + 2^{\nu_1-\nu}\log(\theta_0^{-1}) + \exp\left(-\sigma(1/5)\sqrt{\nu-\nu_1}\right)\log(\theta_0^{-1}).$$

By Lemma 5 we have

$$\int_{\|\theta\|>1/5} \min\left(U, |\sin\pi\theta|^{-1}\right) |F_\nu(\theta,\alpha)|\, d\theta \leqslant \left|\sin\frac{\pi}{5}\right|^{-1} \sqrt{\pi}\, e^{c/4} \frac{e^{-c\nu/25}}{\sqrt{c\nu}}.$$

Gathering the estimates above and observing that $2^{\nu_1-\nu} \ll \exp\left(-\sigma(1/5)\sqrt{\nu-\nu_1}\right)$ we get Lemma 10. □

We will need a special case of the so called "Chernoff bounds":

**Lemma 11.** *Let $\nu \geqslant 1$ be an integer and $X_1, \ldots, X_\nu$ be independent random variables such that $\mathbb{P}(X_j = 1) = \frac{1}{2}$ and $\mathbb{P}(X_j = 0) = \frac{1}{2}$ for $j = 1, \ldots, \nu$. Then for any $t > 0$ we have*

$$\mathbb{P}\left(\left|X_1 + \cdots + X_\nu - \tfrac{\nu}{2}\right| > t\right) < 2\exp(-2\,t^2/\nu).$$

*Proof.* E.g. apply Corollary A.1.2 of [1] to the random variables $1-2X_1, \ldots, 1-2X_\nu$ with $a = 2t$. □

**Lemma 12.** *Let $\nu \geqslant 1$ be an integer and $\xi_\nu > 0$. We have*

$$\operatorname{card}\left\{0 \leqslant n < 2^\nu,\ \left|\mathrm{s}(n) - \tfrac{\nu}{2}\right| > \xi_\nu\sqrt{\nu}\right\} < 2^{\nu+1}\exp\left(-2\,\xi_\nu^2\right).$$

*Proof.* Apply Lemma 11 with $t = \xi_\nu\sqrt{\nu}$. □

For any $\alpha \in \mathbb{R}$, $N \in \mathbb{N}$, and $z \in \mathbb{R}$ let us write

$$T(\alpha, N, z) = N^{-1} \sum_{\substack{0 \leqslant n < N \\ \mathrm{s}(n) < z}} \mathrm{e}(n\alpha). \tag{34}$$

**Lemma 13.** *Let $\nu \geqslant 3$ be an integer and $\alpha \in \mathbb{R} \setminus \mathbb{Z}$. For $z \in \mathbb{R}$ we have*

$$|T(\alpha, 2^\nu, z)| \ll \|\alpha\|^{-1}\,2^{-\nu}\log\log\nu + \nu^{-1/2}(\log\nu)^2, \tag{35}$$

*where the implied constant is absolute.*

*Proof.* Assume first that $z \leqslant \frac{\nu}{2} - \frac{1}{2}\sqrt{\nu\log\nu}$. By Lemma 12 with $\xi_\nu = \frac{1}{2}\sqrt{\log\nu}$ we have

$$|T(\alpha, 2^\nu, z)| \ll \exp(-\tfrac{1}{2}\log\nu) = \nu^{-1/2}. \tag{36}$$

Assume now that $z > \frac{\nu}{2} + \frac{1}{2}\sqrt{\nu\log\nu}$. Observing that

$$\left|\sum_{n<2^\nu} \mathrm{e}(n\alpha)\right| \leqslant \frac{1}{|\sin(\pi\alpha)|} = \frac{1}{\sin(\pi\|\alpha\|)} \leqslant (2\|\alpha\|)^{-1}$$

and writing

$$T(\alpha, 2^\nu, z) = 2^{-\nu} \sum_{0 \leqslant n < 2^\nu} \mathrm{e}(n\alpha) - 2^{-\nu} \sum_{\substack{0 \leqslant n < 2^\nu \\ \mathrm{s}(n) \geqslant z}} \mathrm{e}(n\alpha),$$

by Lemma 12 with $\xi_\nu = \frac{1}{2}\sqrt{\log\nu}$ we have

$$|T(\alpha, 2^\nu, z)| \ll \|\alpha\|^{-1}\,2^{-\nu} + \exp(-\tfrac{1}{2}\log\nu) = \|\alpha\|^{-1}\,2^{-\nu} + \nu^{-1/2}.$$

It remains to consider the case when

$$\frac{\nu}{2} - \frac{1}{2}\sqrt{\nu\log\nu} < z \leqslant \frac{\nu}{2} + \frac{1}{2}\sqrt{\nu\log\nu}.$$

Writing $k_\nu = \left\lfloor \frac{\nu}{2} - \frac{1}{2}\sqrt{\nu\log\nu} \right\rfloor$ we have

$$T(\alpha, 2^\nu, z) = \sum_{k_\nu \leqslant k < z} 2^{-\nu} \sum_{\substack{0 \leqslant n < 2^\nu \\ \mathrm{s}(n) = k}} \mathrm{e}(n\alpha) + T(\alpha, 2^\nu, k_\nu),$$

and

$$2^{-\nu} \sum_{\substack{0 \leqslant n < 2^\nu \\ s(n)=k}} \mathrm{e}(n\alpha) = \int_{-1/2}^{1/2} F_\nu(\theta,\alpha)\,\mathrm{e}(-k\theta)\,d\theta$$

so that

$$T(\alpha,2^\nu,z) = \int_{-1/2}^{1/2} \left( \sum_{k_\nu \leqslant k < z} \mathrm{e}(-k\theta) \right) F_\nu(\theta,\alpha)\,d\theta + T(\alpha,2^\nu,k_\nu),$$

hence using (36) with $k_\nu$ in place of $z$ we have $|T(\alpha,2^\nu,k_\nu)| = O(\nu^{-1/2})$ and we get

$$|T(\alpha,2^\nu,z)| \leqslant \int_{-1/2}^{1/2} \min(\lceil z \rceil - k_\nu, |\sin \pi\theta|^{-1})\ |F_\nu(\theta,\alpha)|\ d\theta + O(\nu^{-1/2}).$$

We choose $U = 2 + \sqrt{\nu \log \nu}$ and observe that $\lceil z \rceil - k_\nu \leqslant U$.

If $\|\alpha\| \leqslant 2^{-1-\nu}$ then (35) holds by the trivial estimate. Therefore we can assume that $\|\alpha\| > 2^{-1-\nu}$. Let $0 \leqslant \nu_1 < \nu$ be the unique integer such that $2^{-2-\nu_1} < \|\alpha\| \leqslant 2^{-1-\nu_1}$, and let us first assume that

$$\nu - \nu_1 \geqslant \left(\sigma\left(\tfrac{1}{5}\right)\right)^{-2} \left(\tfrac{1}{2}\log\nu + \log\log\nu\right)^2 \tag{37}$$

holds, so that

$$\exp\left(\sigma(1/5)\sqrt{\nu-\nu_1}\right) \geqslant \sqrt{\nu}\log\nu.$$

Applying (33) we get

$$|T(\alpha,2^\nu,z)| \ll \nu^{-1/2}.$$

If condition (37) does not hold, in particular we have

$$\nu - \nu_1 \ll (\log\nu)^2, \tag{38}$$

thus $\nu_1 \geqslant 1$, and taking

$$\theta_1 = \sqrt{\frac{\log\nu + 2\log\log\nu}{2c\nu}},$$

we observe that, using (28), we have $1 < \frac{\log 3}{\sqrt{2c}} < U\theta_1$ for $\nu \geqslant 3$, and $\theta_1 \leqslant \frac{1}{2}$ for $\nu \geqslant 22$. Moreover we have

$$U\theta_1 \ll \log\nu,\ U^{-1} \ll \nu^{-1/2},\ \nu_1^{-1/2} \ll \nu^{-1/2},\ \log\frac{1}{2\theta_1} \ll \log\nu,$$

and

$$e^{c\theta_1^2\nu} = \nu^{1/2}\log\nu$$

and we can apply (31) to get

$$|T(\alpha,2^\nu,z)| \ll 2^{\nu_1-\nu}\log\log\nu + (\nu-\nu_1)\nu^{-1/2},$$

and by (38) we get (35) for $\nu \geqslant 22$. Finally (35) holds for $3 \leqslant \nu < 22$ by using the trivial upper bound and modifying the implied absolute constant. □

**Lemma 14.** *Let $N \geqslant 16$ be an integer and $\alpha \in \mathbb{R} \setminus \mathbb{Z}$. For $z \in \mathbb{R}$ we have*

$$|T(\alpha,N,z)| \ll \frac{(\log\log N)(\log\log\log N)}{N\,\|\alpha\|} + \frac{(\log\log N)^2}{(\log N)^{1/2}}. \tag{39}$$

*where the implied constant is absolute.*

*Proof.* If $N$ is a power of 2 then (39) follows from (35). If $N$ is not a power of 2, let $r \geqslant 2$ be an integer and write $N = 2^{j_1} + \cdots + 2^{j_r}$ with $j_1 > \cdots > j_r$. If $\|\alpha\| \leqslant (2N)^{-1}$ (35) holds trivially so we may assume that $\|\alpha\| > (2N)^{-1}$. We have

$$\begin{aligned}\sum_{\substack{0\leqslant n<N\\ \mathrm{s}(n)<z}} \mathrm{e}(n\alpha) &= \sum_{\substack{0\leqslant n<2^{j_1}\\ \mathrm{s}(n)<z}} \mathrm{e}(n\alpha) + \sum_{\substack{0\leqslant n<N-2^{j_1}\\ \mathrm{s}(2^{j_1}+n)<z}} \mathrm{e}((2^{j_1}+n)\alpha)\\ &= \sum_{\substack{0\leqslant n<2^{j_1}\\ \mathrm{s}(n)<z}} \mathrm{e}(n\alpha) + \mathrm{e}(2^{j_1}\alpha) \sum_{\substack{0\leqslant n<N-2^{j_1}\\ \mathrm{s}(n)<z-1}} \mathrm{e}(n\alpha)\end{aligned}$$

hence

$$N\left|T(\alpha,N,z)\right| \leqslant 2^{j_1}\left|T(\alpha,2^{j_1},z)\right| + (N-2^{j_1})\left|T(\alpha,N-2^{j_1},z-1)\right|.$$

Iterating this process we get

$$N\left|T(\alpha,N,z)\right| \leqslant \sum_{i=1}^{r} 2^{j_i}\left|T(\alpha,2^{j_i},z-i)\right|.$$

It follows that

$$N\left|T(\alpha,N,z)\right| \leqslant \sum_{0\leqslant j<\frac{\log N}{\log 2}} 2^j \max_{z\in\mathbb{R}}\left|T(\alpha,2^j,z)\right|. \tag{40}$$

Let us first assume that assume that $\|\alpha\| > N^{-1}(\log N)^{3/2}$. By (35)

$$N\left|T(\alpha,N,z)\right| \leqslant 7 + \sum_{3\leqslant j<\frac{\log N}{\log 2}} \left(\|\alpha\|^{-1}\log\log j + 2^j j^{-1/2}(\log j)^2\right).$$

Observing that

$$\|\alpha\|^{-1}\sum_{3\leqslant j<\frac{\log N}{\log 2}} \log\log j \ll \frac{N}{(\log N)^{3/2}}(\log N)\log\log\log N = \frac{N\log\log\log N}{(\log N)^{1/2}},$$

and

$$\sum_{3\leqslant j<\frac{\log N}{\log 2}} 2^j j^{-1/2}(\log j)^2 \ll (\log\log N)^2 \sum_{3\leqslant j<\frac{\log N}{2\log 2}} 2^j + \frac{(\log\log N)^2}{(\log N)^{1/2}} \sum_{\frac{\log N}{2\log 2}\leqslant j<\frac{\log N}{\log 2}} 2^j$$

so that

$$\sum_{3\leqslant j<\frac{\log N}{\log 2}} 2^j j^{-1/2}(\log j)^2 \ll \frac{N(\log\log N)^2}{(\log N)^{1/2}},$$

and for $\|\alpha\| > N^{-1}(\log N)^{3/2}$ we get (39).

It remains to consider the case where

$$(2N)^{-1} < \|\alpha\| \leqslant N^{-1}(\log N)^{3/2}. \tag{41}$$

Let

$$J = \frac{-\log(\|\alpha\|)}{\log 2} > 0,\ \text{ i.e.: } 2^J = \|\alpha\|^{-1}.$$

In (40) we use the trivial estimate $\left|T(\alpha,2^j,z)\right| \leqslant 1$ for $0 \leqslant j \leqslant J$ and (35) for $J < j < \log N/\log 2$. This leads to

$$N\left|T(\alpha,N,z)\right| \ll \sum_{0\leqslant j\leqslant J} 2^j + \sum_{J<j<\frac{\log N}{\log 2}} \left(\|\alpha\|^{-1}\log\log j + 2^j j^{-1/2}(\log j)^2\right).$$

We have

$$\sum_{0\leqslant j\leqslant J} 2^j \ll 2^J = \|\alpha\|^{-1}$$

and as above

$$\sum_{J<j<\frac{\log N}{\log 2}} 2^j j^{-1/2}(\log j)^2 \ll \frac{N(\log\log N)^2}{(\log N)^{1/2}}.$$

Now

$$\sum_{J<j<\frac{\log N}{\log 2}} \log\log j \ll (\log\log\log N)\left(1+\frac{\log N}{\log 2}-J\right)$$

and using (41)

$$\frac{\log N}{\log 2}-J = \frac{\log(N\,\|\alpha\|)}{\log 2} \ll \log\log N$$

and this completes the proof of (39). □

**Lemma 15.** *For*

$$0 \leqslant z \leqslant \log(2N) \tag{42}$$

*and*

$$\|\alpha\| > u \tag{43}$$

*(where $u$ is defined by* (18)*) we have*

$$\left| S_z(\alpha) - \Phi(y_{2N}) \sum_{n=1}^{2N} \mathrm{e}(n\alpha) \right| \ll \frac{(\log\log N)\log\log\log N}{u} + \frac{N(\log\log N)^2}{(\log N)^{1/2}}. \tag{44}$$

*Proof.* By (18) and (43) it follows from 34 and 39 that

$$\begin{aligned} |S_z(\alpha)| = \left| \sum_{\substack{n\leqslant 2N\\ \mathrm{s}(n)<z}} \mathrm{e}(n\alpha) \right| &\ll \frac{(\log\log N)\log\log\log N}{\|\alpha\|} + \frac{N(\log\log N)^2}{(\log N)^{1/2}} \\ &\ll \frac{(\log\log N)\log\log\log N}{u} + \frac{N(\log\log N)^2}{(\log N)^{1/2}} \end{aligned} \tag{45}$$

and by (43) we have

$$\left| \Phi(y_{2N}) \sum_{n=1}^{2N} \mathrm{e}(n\alpha) \right| \leqslant \left| \frac{e(2N\alpha)-1}{e(\alpha-1)} \right| \ll \frac{1}{\|\alpha\|} \leqslant \frac{1}{u}. \tag{46}$$

(44) follows from (45) and (46). □

**Lemma 16.** *Uniformly for*

$$0 \leqslant z \leqslant \log 2N$$

*and all $\alpha\in\mathbb{R}$ we have*

$$\left| S_z(\alpha) - \Phi(y_{2N}) \sum_{n=1}^{2N} \mathrm{e}(n\alpha) \right| \ll \frac{N(\log\log N)(\log\log\log N)^{1/2}}{(\log N)^{1/4}}. \tag{47}$$

*Proof.* Combining Lemmas 3 and 15 we get for all $\alpha$ that

$$\begin{aligned}&\left| S_z(\alpha) - \Phi(y_{2N}) \sum_{n=1}^{2N} \mathrm{e}(n\alpha) \right| \\ &\qquad \ll \frac{uN^2 \log\log N}{(\log N)^{1/2}} + \frac{(\log\log N)\log\log\log N}{u} + \frac{N(\log\log N)^2}{(\log N)^{1/2}},\end{aligned}$$

and the choice of $u$ made by (18) gives (47). □

## 5. Completion of the proof of Theorem 1

We have

$$J = \int_0^1 G(\alpha)\, H(\alpha)\, S_z(-\alpha)\, d\alpha = \Phi(y_{2N})\, J_1 + J_2$$

where

$$J_1 = \int_0^1 G(\alpha) H(\alpha) \sum_{n=1}^{2N} \mathrm{e}(-n\alpha)\, d\alpha$$

and

$$\begin{aligned} J_2 &= \int_0^1 G(\alpha)H(\alpha) \left( S_z(-\alpha) - \Phi(y_{2N}) \sum_{n=1}^{2N} \mathrm{e}(-n\alpha) \right) d\alpha \\ &\leqslant \int_0^1 |G(\alpha)H(\alpha)| \left| S_z(-\alpha) - \Phi(y_{2N}) \sum_{n=1}^{2N} \mathrm{e}(-n\alpha) \right| d\alpha, \end{aligned}$$

whence by using (47)

$$J_2 \ll \frac{N(\log\log N)(\log\log\log N)^{1/2}}{(\log N)^{1/4}} \int_0^1 |G(\alpha)H(\alpha)|\; d\alpha,$$

by the Cauchy-Schwarz inequality and by Parseval identity

$$\int_0^1 |G(\alpha)H(\alpha)|\; d\alpha \leqslant \left( \int_0^1 |G(\alpha)|^2\, d\alpha \right)^{1/2} \left( \int_0^1 |H(\alpha)|^2\, d\alpha \right)^{1/2} = \sqrt{|\mathcal{A}|\,|\mathcal{B}|},$$

so that

$$J_2 \ll \frac{N(\log\log N)(\log\log\log N)^{1/2}}{(\log N)^{1/4}} \sqrt{|\mathcal{A}|\,|\mathcal{B}|}. \tag{48}$$

Clearly we have

$$J_1 = \int_0^1 \sum_{a\in\mathcal{A}} \sum_{b\in\mathcal{B}} \sum_{n=1}^{2N} \mathrm{e}((a+b-n)\alpha)\, d\alpha = \sum_{a\in\mathcal{A}} \sum_{b\in\mathcal{B}} 1 = |\mathcal{A}|\,|\mathcal{B}|\,. \tag{49}$$

By (48) and (49) we have

$$\left| J - \Phi(y_{2N})\, |\mathcal{A}|\, |\mathcal{B}| \right| \ll \frac{N(\log\log N)(\log\log\log N)^{1/2}}{(\log N)^{1/4}} \sqrt{|\mathcal{A}|\,|\mathcal{B}|}. \tag{50}$$

(1) follows from (6) and (50) and this completes the proof of the theorem.

## References

[1] N. Alon and J. H. Spencer, *The probabilistic method*, Wiley-Interscience Series in Discrete Mathematics and Optimization, John Wiley & Sons, Inc., Hoboken, NJ, third ed., 2008. With an appendix on the life and work of Paul Erdős.

[2] P. Erdős, H. Maier, and A. Sárközy, *On the distribution of the number of prime factors of sums* $a+b$, Trans. Amer. Math. Soc., 302 (1987), pp. 269–280.

[3] I. Kátai and J. Mogyoródi, *On the distribution of digits*, Publ. Math. Debrecen, 15 (1968), pp. 57–68.

[4] C. Mauduit, J. Rivat, and A. Sárközy, *On the digits of sumsets*, Canadian Journal of Mathematics, 69 (2017), pp. 595–612.

[5] C. Mauduit and A. Sárközy, *On the arithmetic structure of sets characterized by sum of digits properties*, J. Number Theory, 61 (1996), pp. 25–38.

[6] C. Mauduit and A. Sárközy, *On the arithmetic structure of the integers whose sum of digits is fixed*, Acta Arithmetica, 81 (1997), pp. 145–173.

[7] A. Sárközy, *Hybrid problems in number theory*, in Number Theory, New York 1985-88, vol. 1383 of Lecture Notes in Mathematics, Springer-Verlag, 1989, pp. 146–169.

Christian MAUDUIT, Université d'Aix-Marseille and Institut Universitaire de France, Institut de Mathématiques de Marseille, CNRS UMR 7373, 163, avenue de Luminy, Case 907, 13288 MARSEILLE Cedex 9, France

*E-mail address*: `mauduit@iml.univ-mrs.fr`

Joël RIVAT, Université d'Aix-Marseille, Institut de Mathématiques de Marseille, CNRS UMR 7373, 163 avenue de Luminy, Case 907, 13288 Marseille Cedex 9, France.

*E-mail address*: `rivat@iml.univ-mrs.fr`

András SÁRKÖZY, Eötvös Loránd University, Department of Algebra and Number Theory, H-1117 Budapest, Pázmány Péter sétány 1/c, Hungary.

*E-mail address*: `sarkozy@cs.elte.hu`