Illinois Journal of Mathematics
Volume 46, Number 1, Spring 2002, Pages 185–197
S 0019-2082

ON THE PSEUDO-RANDOM PROPERTIES OF n^c

CHRISTIAN MAUDUIT, JOËL RIVAT, AND ANDRÁS SÁRKŐZY

Abstract. We estimate the well-distribution measure and correlation of order 2 of the binary sequence $E_N = \{e_1, \ldots, e_N\}$ defined by $e_n = +1$ if $0 \leqslant \{n^c\} < 1/2$ and $e_n = -1$ if $1/2 \leqslant \{n^c\} < 1$, where c is a real, non-integral number greater than 1 and $\{x\}$ denotes the fractional part of x. We also prove an upper bound for the well-distribution measure of an arbitrary binary sequence in terms of its generating function and show that there exists no upper bound of this type for the correlation. The proof is based on the Erdős-Turán inequality, which we establish with an improved constant.

1. Introduction

In a series of papers, J. Cassaigne, S. Ferenczi, C. Mauduit, J. Rivat and A. Sárközy studied finite pseudo-random binary sequences

$$E_N = \{e_1, \ldots, e_N\} \in \{-1, +1\}^N.$$

In particular, in [5] Mauduit and Sárközy first introduced the following measures of pseudo-randomness: the *well-distribution measure* of E_N, defined by

$$W(E_N) = \max_{a,b,t} \left| \sum_{j=0}^{t-1} e_{a+jb} \right|,$$

where the maximum is taken over all $a, b, t \in \mathbb{N}$ such that $1 \leqslant a \leqslant a+(t-1)b \leqslant N$; and the *correlation measure of order k* of E_N, defined as

$$C_k(E_N) = \max_{M,D} \left| \sum_{n=1}^{M} e_{n+d_1} e_{n+d_2} \cdots e_{n+d_k} \right|,$$

Received March 8, 2001; received in final form December 5, 2001.

2000 *Mathematics Subject Classification.* Primary 11K45, 11K36. Secondary 11K06, 11K38, 11L07.

The research of the third author was partially supported by the Hungarian National Foundation for Scientific Research, Grant No. T 029 759 and MKM fund FKFP-0139/1997.

where the maximum is taken over all $D = (d_1, \dots, d_k)$ and M such that $0 \leqslant d_1 < \cdots < d_k \leqslant N - M$. The sequence is considered a “good” pseudo-random sequence if the measures $W(E_N)$ and $C_k(E_N)$ are both “small” in terms of N, at least for “small” k (e.g., if the measures are $o(N)$ as $N \to \infty$).

In [5] it was shown that the Legendre symbol forms a “good” pseudo-random sequence. In [6, 1, 2, 7, 8, 11, 14] other special sequences were tested for pseudo-randomness. In [8] Mauduit and Sárközy posed the following problem of this type:

Denote the fractional part of α by $\{\alpha\}$, and write

$$\chi(x) = \begin{cases} +1 & \text{for } 0 \leqslant \{x\} < 1/2, \\ -1 & \text{for } 1/2 \leqslant \{x\} < 1. \end{cases}$$

Let $c > 0$, $c \notin \mathbb{N}$, and investigate the pseudo-random properties of the sequence

$$E_N = E_N(n^c) = \{e_1, \dots, e_N\}$$

defined by

$$e_n = \chi(n^c) \qquad (n = 1, 2, \dots, N).$$

Our main goal in this paper is to study this problem. Of course, the case when $0 < c < 1$ is trivial (since the correlation is trivially large); thus from now on we will restrict ourselves to the case $c > 1$. We will prove:

Theorem 1. *For $c > 1$, $c \notin \mathbb{N}$, $R = \lceil c \rceil$, $0 \leqslant b \leqslant a \leqslant (x-b)^{1-c/R}$, $x \to \infty$, we have*

$$\left| \sum_{an+b \leqslant x} e_{an+b} \right| \ll a^{-1+R/(2^R-1)} x^{1-(R-c)/(2^R-1)}. \tag{1}$$

Corollary 1. *For $c > 1$, $c \notin \mathbb{N}$, $R = \lceil c \rceil$, we have*

$$W(E_N) \ll N^{1-(R-c)/(2^R-1)}. \tag{2}$$

The proof of Theorem 1 will be based on the Erdős-Turán inequality, which we will prove here with constant 1, improving the value 3 given by Montgomery [9, Corollary 1.1]. This result may be of independent interest.

Our next result gives an estimate for the “short range” correlations of order 2 of E_N:

Theorem 2. *For $c > 1$, $c \notin \mathbb{N}$, $R = \lceil c \rceil$, $1 \leqslant d \leqslant N^{1-2(R-c)/(2^R-1)}$, $N \to \infty$, we have*

$$\left| \sum_{n \leqslant N} e_n e_{n+d} \right| \ll c N^{1-(R-c)/(2^R-1)} \log^2 N + \frac{N^{2-c}}{d}. \tag{3}$$

We expect that this result also applies to "long range" correlations, but we have not been able to prove this.

The above results, as well as the results in several earlier papers, give estimates for the well-distribution and correlation measures of certain special sequences. This raises the question if there are general inequalities for these measures, similar to the Erdős-Turán inequality, for the discrepancy measure. For the well-distribution measure, Theorem 3 below gives such an inequality.

For $N \in \mathbb{N}$, $k \in \mathbb{N}$, $1 \leqslant k \leqslant N$, $F_N = \{f_1, \dots, f_N\} \in \{-1, +1\}^N$, $\alpha \in \mathbb{R}$, write

$$\phi_k(F_N, \alpha) = \sum_{j=1}^{k} f_j \, \mathrm{e}(j\alpha), \tag{4}$$

where we use the standard notation $\mathrm{e}(\beta) = \exp(2i\pi\beta)$.

Theorem 3. *For all $N \in \mathbb{N}$, $F_N = \{f_1, \dots, f_N\} \in \{-1, +1\}^N$, we have*

$$W_N(F_N) \leqslant 2 \max_{1 \leqslant k \leqslant N} \max_{\alpha} |\phi_k(F_N, \alpha)|. \tag{5}$$

As an application we note, for example, that the upper bound (5.1) in [6] for the well-distribution measure of the Rudin-Shapiro sequence follows immediately from the estimate (4.10) in [6] and Theorem 3.

In contrast to the well-distribution measure, there is no similar inequality for the correlation measure; more precisely, the following theorem shows that there is no non-trivial upper bound for $C_2(F_N)$ in terms of the generating functions ϕ_k:

Theorem 4. *For $N \in \mathbb{N}$, $N \geqslant 4$, there is a binary sequence*

$$F_N = \{f_1, \dots, f_N\} \in \{-1, +1\}^N$$

such that for $k = 1, 2, \dots, N$ we have

$$\max_{\alpha} |\phi_k(F_N, \alpha)| < 2(2 + \sqrt{2})k^{1/2} \quad (\leqslant 2(2 + \sqrt{2})N^{1/2}), \tag{6}$$

but

$$C_2(F_N) > \frac{N}{4}. \tag{7}$$

We remark that in [6] we proved that the "truncated" Rudin-Shapiro sequence itself satisfies an inequality that is just slightly weaker than (7). However, the analysis of the connection between the generating functions and the pseudo-random measures given here is new. Moreover, to prove Theorem 4 we will use a slightly different construction, which is related to the Rudin-Shapiro sequence, but which better illustrates the underlying phenomenon.

2. Classical results on exponential sums

Throughout this paper, when we encounter a property $\mathfrak{P}$ which may or may not be satisfied according to the value of some parameters, we will make use of the following notation:

$$\mathbf{1}_{\mathfrak{P}} = \begin{cases} 1 & \text{whenever } \mathfrak{P} \text{ is satisfied,} \\ 0 & \text{otherwise.} \end{cases}$$

Definition 1. Let $\mathbf{x_1}, \dots, \mathbf{x_N}$ be a finite sequence of points in $\mathbb{R}^s$. The *discrepancy* of $\mathbf{x_1}, \dots, \mathbf{x_N}$ is defined by

$$D_N(\mathbf{x_1}, \dots, \mathbf{x_N}) = \sup_{I_1,\dots,I_s} \left| \frac{1}{N} \sum_{i=1}^{N} \mathbf{1}_{\{\mathbf{x_i}\} \in I_1 \times \cdots \times I_s} - \mu(I_1) \cdots \mu(I_s) \right|, \tag{8}$$

where the supremum is taken over all intervals I_1,. . . ,I_s contained in $[0, 1]$, $\mu(I)$ stands for the length of the interval I, and $\{\mathbf{y}\} = (\{y_1\}, \dots, \{y_s\})$, for $\mathbf{y} = (y_1, \dots, y_s) \in \mathbb{R}^s$,

Lemma 1 (Erdős-Turán). *For any integers $N > 0$, $H > 0$, and any real numbers $x_1, \dots, x_N$ we have*

$$D_N(x_1, \dots, x_N) \leqslant \frac{1}{H+1} + \sum_{h=1}^{H} \frac{1}{h} \left| \frac{1}{N} \sum_{n=1}^{N} \mathrm{e}(hx_n) \right|. \tag{9}$$

Remark 1. This is the well known Erdős-Turán inequality [3], except that we give a sharp explicit constant (equal to 1). This improves Corollary 1.1 of Montgomery [9], who had a factor 3 in front of the sum on the right hand side above. (This sharpening is not necessary in this paper; we present it here since it seems to be of independent interest.)

Proof. Vaaler [15, (2.28) and (2.29)] defines for $t \in [-1, 1]$

$$\hat{J}(t) = \pi|t|(1 - |t|)\cot(\pi|t|) + |t|, \quad \hat{K}(t) = (1 - |t|),$$

and notes [15, Theorem 6, p. 192] that $\hat{J}$ is even, nonnegative, continuously differentiable and strictly decreasing on $[0, 1]$. Following Vaaler, we set

$$\hat{J}_{H+1}(h) = \hat{J}(h/(H+1)), \quad \hat{K}_{H+1}(h) = \hat{K}(h/(H+1)).$$

Vaaler's inequality [15, (8.3)] implies[1]

$$D_N(x_1, \dots, x_N) \leqslant \frac{1}{H+1} + 2\sum_{h=1}^{H} \left(\frac{\hat{J}_{H+1}(h)}{\pi h} + \frac{\hat{K}_{H+1}(h)}{H+1} \right) \left| \frac{1}{N} \sum_{n=1}^{N} \mathrm{e}(hx_n) \right|. \tag{10}$$

[1] Vaaler uses *normalized* characteristic functions of intervals, but this condition can be removed: his proof is based on inequality (7.24) of his paper, which by continuity holds for any type of characteristic function of an interval.

Therefore it is sufficient to prove that for all $t \in [0,1]$,

$$\pi^{-1}\hat{J}(t) + t(1-t) \leqslant \frac{1}{2}. \tag{11}$$

By the monotonicity properties of the functions $t \mapsto \hat{J}(t)$ and $t \mapsto t(1-t)$ we get

$$\pi^{-1}\hat{J}(t) + t(1-t) \leqslant \begin{cases} \pi^{-1}\hat{J}(0) + 0.16 & \leqslant & 0.48 & \text{on } [0,0.2], \\ \pi^{-1}\hat{J}(0.2) + 0.21 & \leqslant & 0.4939 & \text{on } [0.2,0.3], \\ \pi^{-1}\hat{J}(0.3) + 0.25 & \leqslant & 0.4981 & \text{on } [0.3,1], \end{cases}$$

which completes the proof. □

Lemma 2 (Koksma-Szűsz). *Let $s > 0$ be an integer. For $\mathbf{h} = (h_1, \ldots, h_s) \in \mathbb{Z}^s$, write*

$$\varphi(\mathbf{h}) = \max_{j=1,\ldots,s} |h_j|, \quad r(\mathbf{h}) = \prod_{j=1}^{s} \max(|h_j|, 1).$$

Let $\mathbf{x_1}, \ldots, \mathbf{x_N}$ be a finite sequence of points of $\mathbb{R}^s$. For any integer $H > 0$ we have

$$D_N(\mathbf{x_1}, \ldots, \mathbf{x_N}) \ll_s \frac{1}{H} + \frac{1}{N} \sum_{0<\varphi(\mathbf{h})\leqslant H} \frac{1}{r(\mathbf{h})} \left| \sum_{n=1}^{N} \mathrm{e}(\mathbf{h} \cdot \mathbf{x_n}) \right|. \tag{12}$$

Lemma 3 (van der Corput). *Let R be an integer $\geqslant 2$. Suppose that f has R continuous derivatives on an interval $I \subseteq [N+1, 2N]$. Assume also that there is some constant F such that*

$$FN^{-r} \ll |f^{(r)}(x)| \ll FN^{-r}$$

for $x \in I$ and $r = 1, \ldots, R$. Then

$$\sum_{n\in I} \mathrm{e}(f(n)) \ll (FN^{-R})^{1/(2^R-2)} N + F^{-1}N. \tag{13}$$

Proof. See Theorem 2.9 of [4]. □

3. Well-distribution

Set $N = \lfloor (x-b)/a \rfloor$. We have

$$\sum_{n\leqslant N} e_{an+b} = \sum_{n\leqslant N} \left(\mathbf{1}_{\{(an+b)^c\}<1/2} - \frac{1}{2} \right) - \sum_{n\leqslant N} \left(\mathbf{1}_{\{(an+b)^c\}\geqslant 1/2} - \frac{1}{2} \right).$$

Hence, writing $x_n = (an+b)^c$, we obtain

$$\left| \sum_{n\leqslant N} e_{an+b} \right| \leqslant 2N \; D_N(x_1, \ldots, x_N).$$

In order to apply Lemma 1 we need the following exponential sum estimate:

Lemma 4. *For $N \geqslant 1$, $h \geqslant 1$, $0 \leqslant b \leqslant a$, $c > 1$, $R = \lceil c \rceil$, we have*

$$\sum_{n=1}^{N} \mathrm{e}(h(an+b)^c) \ll_c (ha^c)^{1/(2^R-2)} N^{1+(c-R)/(2^R-2)}.$$

Proof. We split the summation over n into $L = \lceil \log N / \log 2 \rceil$ intervals of the form $(N/2^\ell, N/2^{\ell-1}]$. We apply van der Corput's estimate (Lemma 3) to each such interval with

$$f(x) = ha^c \left(x + \frac{b}{a} \right)^c, \qquad F = ha^c (N2^{-\ell})^c,$$

so that

$$ha^c (N2^{-\ell})^{c-r} \ll f^{(r)}(x) \ll ha^c (N2^{-\ell})^{c-r}$$

for $r = 1, \dots, R$. We obtain

$$\sum_{n=1}^{N} \mathrm{e}(h(an+b)^c) \ll \sum_{1 \leqslant \ell \leqslant L} N2^{-\ell} \left((ha^c)^{1/(2^R-2)} (N2^{-\ell})^{(c-R)/(2^R-2)} + h^{-1} a^{-c} N^{-c} 2^{c\ell} \right).$$

Since $0 \leqslant R - c < 1$ and $R \geqslant 2$, we have $(R-c)/(2^R-2) \leqslant 1/2$. Thus

$$\sum_{1 \leqslant \ell \leqslant L} 2^{-\ell} 2^{-\ell(c-R)/(2^R-2)} \ll \sum_{\ell \geqslant 1} 2^{-\ell/2} \ll 1.$$

We also have

$$\sum_{1 \leqslant \ell \leqslant L} 2^{(c-1)\ell} \ll 2^{(c-1)L} \ll N^{c-1}.$$

Hence

$$\sum_{n=1}^{N} \mathrm{e}(h(an+b)^c) \ll N \left((ha^c)^{1/(2^R-2)} N^{(c-R)/(2^R-2)} + h^{-1} a^{-c} N^{-c} N^{c-1} \right),$$

$$\sum_{n=1}^{N} \mathrm{e}(h(an+b)^c) \ll (ha^c)^{1/(2^R-2)} N^{1+(c-R)/(2^R-2)} + h^{-1} a^{-c}.$$

We observe that $ha^c \geqslant 1$ and $1 + (c-R)/(2^R-2) \geqslant 1/2$, so the first term on the right hand side above dominates the second, and the result follows. □

Lemma 5. *For $N \geqslant 1$, $c > 1$, $R = \lceil c \rceil$, $0 \leqslant b \leqslant a \leqslant N^{(R/c)-1}$, we have*

$$D_N(x_1, \dots, x_n) \ll (a^c N^{c-R})^{1/(2^R-1)}.$$

Proof. By Lemma 1 it suffices to show the existence of an integer $H > 0$ such that

$$\frac{1}{H} + \frac{1}{N}\sum_{h=1}^{H}\frac{1}{h}\left|\sum_{n=1}^{N}\mathrm{e}(h(an+b)^c)\right| \ll (a^c N^{c-R})^{1/(2^R-1)}.$$

For any integer $H > 0$, the left hand side above can be estimated using Lemma 4 by

$$H^{-1} + (Ha^c)^{1/(2^R-2)} N^{(c-R)/(2^R-2)},$$

and the choice

$$H = \left\lfloor (a^c N^{c-R})^{-1/(2^R-1)} \right\rfloor \geqslant 1$$

gives the result. □

We have thus obtained the estimate

$$\left|\sum_{n\leqslant N} e_{an+b}\right| \ll N(a^c N^{c-R})^{1/(2^R-1)},$$

and replacing N by x/a this completes the proof of Theorem 1.

In order to obtain Corollary 1, let $x = N$. It suffices to consider the case not covered by Theorem 1, i.e., the case when $a \gg x^{1-c/R}$. In this case we just use the trivial upper bound x/a, which leads to

$$\left|\sum_{n\leqslant N} e_{an+b}\right| \ll x^{c/R}.$$

Then for $R \geqslant 2$ we have $R \leqslant 2^R - 1$, so that $\frac{R}{2^R-1}(1-\frac{c}{R}) \leqslant 1 - \frac{c}{R}$. Hence $c/R \leqslant 1 - (R-c)/(2^R-1)$, and Corollary 1 follows.

4. Correlation

We have

$$\begin{aligned}\sum_{n\leqslant N} e_n e_{n+d} = &\sum_{n\leqslant N}\left(\mathbf{1}_{\{n^c\}<1/2}\,\mathbf{1}_{\{(n+d)^c\}<1/2} - \frac{1}{4}\right)\\ &- \sum_{n\leqslant N}\left(\mathbf{1}_{\{n^c\}\geqslant 1/2}\,\mathbf{1}_{\{(n+d)^c\}<1/2} - \frac{1}{4}\right)\\ &- \sum_{n\leqslant N}\left(\mathbf{1}_{\{n^c\}<1/2}\,\mathbf{1}_{\{(n+d)^c\}\geqslant 1/2} - \frac{1}{4}\right)\\ &+ \sum_{n\leqslant N}\left(\mathbf{1}_{\{n^c\}\geqslant 1/2}\,\mathbf{1}_{\{(n+d)^c\}\geqslant 1/2} - \frac{1}{4}\right).\end{aligned}$$

Hence, writing $\mathbf{x_n} = (n^c, (n+d)^c)$, we obtain

$$\left| \sum_{n \leqslant N} e_n e_{n+d} \right| \leqslant 4N \ D_N(\mathbf{x_1}, \dots, \mathbf{x_N}).$$

For $c > 1$, $c \notin \mathbb{Z}$, $H \geqslant 2$, $|h|, |k| \leqslant H$, $2cN/H \leqslant n \leqslant N$, $1 \leqslant d \leqslant N/H^2$, we consider

$$f(n) = hn^c + k(n+d)^c.$$

We set $R = \lceil c \rceil$ and compute for $1 \leqslant r \leqslant R$,

$$f^{(r)}(n) = c(c-1)\cdots(c-r+1)n^{c-r}\left((h+k) + k\left((1+d/n)^{c-r} - 1\right)\right).$$

Since $|(c-r)\,kd/n| \leqslant 1/2$, we have

$$f^{(r)}(n) \asymp \begin{cases} c\,(c-1)\cdots(c-r+1)(h+k)n^{c-r} & \text{when } h+k \neq 0, \\ c\,(c-1)\cdots(c-r)kdn^{c-r-1} & \text{when } h+k = 0. \end{cases}$$

By Lemma 2,

$$\left| \sum_{n \leqslant N} e_n e_{n+d} \right| \ll \frac{N}{H} + S_1 + S_2 + S_3 + S_4,$$

where

$$S_1 = \sum_{\substack{1 \leqslant |h|, |k| \leqslant H \\ h+k \neq 0}} \frac{1}{|hk|} \left| \sum_{n=1}^{N} \mathrm{e}(hn^c + k(n+d)^c) \right| \quad \text{(i.e., } hk \neq 0,\ h+k \neq 0\text{)},$$

$$S_2 = \sum_{1 \leqslant |h| \leqslant H} \frac{1}{h^2} \left| \sum_{n=1}^{N} \mathrm{e}(hn^c - h(n+d)^c) \right| \quad \text{(i.e., } hk \neq 0,\ h+k = 0\text{)},$$

$$S_3 = \sum_{1 \leqslant |h| \leqslant H} \frac{1}{|h|} \left| \sum_{n=1}^{N} \mathrm{e}(hn^c) \right| \quad \text{(i.e., } k = 0\text{)},$$

$$S_4 = \sum_{1 \leqslant |k| \leqslant H} \frac{1}{|k|} \left| \sum_{n=1}^{N} \mathrm{e}(k(n+d)^c) \right| \quad \text{(i.e., } h = 0\text{)}.$$

Estimation of S_1: We have

$$S_1 \leqslant \sum_{\substack{1 \leqslant |h|, |k| \leqslant H \\ h+k \neq 0}} \frac{1}{|hk|} \left| \sum_{2cN/H \leqslant n \leqslant N} \mathrm{e}(hn^c + k(n+d)^c) \right| + O\left(c\frac{N}{H}\log^2 H\right).$$

We split the summation over n into $L = \lfloor \log(H/2c)/\log 2 \rfloor$ intervals of the form $(N/2^\ell, N/2^{\ell-1}]$ and apply van der Corput's estimate (Lemma 3) to each of these intervals. We obtain

$$\begin{aligned} S_1 &\ll \sum_{1\leqslant \ell\leqslant L}\ \sum_{\substack{1\leqslant |h|,|k|\leqslant H\\ h+k\neq 0}} \frac{1}{|hk|}\Biggl(|h+k|^{1/(2^R-2)}(N2^{-\ell})^{1+(c-R)/(2^R-2)} \\ &\qquad + \frac{(N2^{-\ell})^{1-c}}{|h+k|}\Biggr) + c\frac{N}{H}\log^2 H. \end{aligned}$$

Hence

$$\begin{aligned} S_1 &\ll \sum_{\substack{1\leqslant |h|,|k|\leqslant H\\ h+k\neq 0}} \frac{1}{|hk|}\Biggl((|h|^{1/(2^R-2)}+|k|^{1/(2^R-2)})N^{1+(c-R)/(2^R-2)} \\ &\qquad + \frac{N^{1-c}H^{c-1}}{|h+k|}\Biggr) + c\frac{N}{H}\log^2 H, \end{aligned}$$

so that

$$\begin{aligned} S_1 &\ll H^{1/(2^R-2)}N^{1+(c-R)/(2^R-2)}\log H + N^{1-c}H^{c-1}\log^2 H + c\frac{N}{H}\log^2 H \\ &\ll H^{1/(2^R-2)}N^{1+(c-R)/(2^R-2)}\log H + c\frac{N}{H}\log^2 H. \end{aligned}$$

Estimation of S_2: By similar arguments we obtain

$$S_2 \ll \sum_{1\leqslant h\leqslant H}\frac{1}{h^2}\left((hdN^{c-R-1})^{1/(2^R-2)}N + \frac{N^{2-c}}{hd}\right),$$

so that

$$S_2 \ll N^{1+(c-R)/(2^R-2)} + \frac{N^{2-c}}{d}.$$

Estimation of S_3 and S_4: Similarly we have

$$\begin{aligned} S_3 &\ll H^{1/(2^R-2)}N^{1+(c-R)/(2^R-2)}, \\ S_4 &\ll H^{1/(2^R-2)}N^{1+(c-R)/(2^R-2)} \end{aligned}$$

Estimation of the correlation: We have

$$\left|\sum_{n\leqslant N} e_n e_{n+d}\right| \ll H^{1/(2^R-2)}N^{1+(c-R)/(2^R-2)}\log H + c\frac{N}{H}\log^2 H + \frac{N^{2-c}}{d},$$

and the choice $H = \left\lfloor N^{(R-c)/(2^R-1)} \right\rfloor \geqslant 1$ gives

$$\left| \sum_{n \leqslant N} e_n e_{n+d} \right| \ll c N^{1+(c-R)/(2^R-1)} \log^2 N + \frac{N^{2-c}}{d}.$$

5. Proof of Theorem 3

For all $b \in \mathbb{N}$, $a \in \mathbb{Z}$, $k \in \mathbb{N}$, $k \leqslant N$, we have

$$\begin{aligned}
\left| \sum_{\substack{n \equiv a \bmod b \\ n \leqslant k}} f_n \right| &= \left| \sum_{n=1}^{k} f_n \frac{1}{b} \sum_{h=1}^{b} \mathrm{e}\left((n-a) \frac{h}{b} \right) \right| \\
&= \frac{1}{b} \left| \sum_{h=1}^{b} \mathrm{e}\left(-\frac{ah}{b} \right) \sum_{n=1}^{k} f_n \, \mathrm{e}\left(n \frac{h}{b} \right) \right| \\
&= \frac{1}{b} \left| \sum_{h=1}^{b} \mathrm{e}\left(-\frac{ah}{b} \right) \phi_k \left(F_N, \frac{h}{b} \right) \right| \\
&\leqslant \frac{1}{b} \sum_{h=1}^{b} \left| \phi_k \left(F_N, \frac{h}{b} \right) \right| \leqslant \max_{\alpha} |\phi_k(F_N, \alpha)|.
\end{aligned}$$

It follows that

$$\begin{aligned}
W(E_N) &= \max_{a,b,t} \left| \sum_{j=0}^{t-1} f_{a+jb} \right| = \max_{a,b,t} \left| \sum_{\substack{n \equiv a \bmod b \\ n \leqslant a+(t-1)b}} f_n - \sum_{\substack{n \equiv a \bmod b \\ n \leqslant a-1}} f_n \right| \\
&\leqslant \max_{a,b,t} \left(\left| \sum_{\substack{n \equiv a \bmod b \\ n \leqslant a+(t-1)b}} f_n \right| + \left| \sum_{\substack{n \equiv a \bmod b \\ n \leqslant a-1}} f_n \right| \right) \leqslant 2 \, \max_{\alpha} |\phi_k(F_N, \alpha)|.
\end{aligned}$$

Since this holds for all $k \leqslant N$, this proves (5).

6. Proof of Theorem 4

The construction is based on the Rudin-Shapiro sequence $\{r_0, r_1, \ldots\} \in \{-1, +1\}^\infty$ (see [12, 13]). This sequence can be defined by the recursion

$$\begin{aligned}
r_0 &= 1, \\
r_{2n} &= r_n \qquad (\text{for } n = 1, 2, \ldots), \\
r_{2n+1} &= (-1)^n r_n \qquad (\text{for } n = 0, 1, 2, \ldots);
\end{aligned}$$

see [10, p. 73]. Its most important feature is that the trigonometric polynomials $P_n(\alpha)$ defined by

$$P_n(\alpha) = \sum_{j=0}^{n-1} r_j \, \mathrm{e}(j\alpha) \qquad \text{for all } n \in \mathbb{N},$$

satisfy

$$\max_{\alpha} |P_n(\alpha)| \leqslant (2+\sqrt{2})\, n^{1/2}; \tag{14}$$

see [10, p. 166]. Since by Parseval's formula

$$\int_0^1 |P_n(\alpha)|^2 d\alpha = n,$$

(14) says that the maximum of the function $|P_n(\alpha)|$ exceeds its mean square by at most a constant factor.

Write $M = N - \lfloor N/2 \rfloor$, and define the sequence $F_N = \{f_1, \ldots, f_N\}$ by

$$f_j = r_{j-1} \qquad \text{for } j = 1, 2, \ldots, M, \tag{15}$$

$$f_{M+j} = r_{j-1} = f_j \qquad \text{for } j = 1, 2, \ldots, N-M. \tag{16}$$

Then the polynomial in (4) is

$$\begin{aligned} \phi_k(F_N, \alpha) &= \sum_{j=1}^{k} f_j \, \mathrm{e}(j\alpha) = \sum_{j=1}^{k} r_{j-1} \, \mathrm{e}(j\alpha) \\ &= \sum_{l=0}^{k-1} r_l \, \mathrm{e}((l+1)\alpha) = \mathrm{e}(\alpha) P_k(\alpha) \qquad \text{for } k \leqslant M \end{aligned} \tag{17}$$

and

$$\begin{aligned} \phi_k(F_N, \alpha) &= \sum_{j=1}^{M} f_j \, \mathrm{e}(j\alpha) + \sum_{j=M+1}^{N} f_j \, \mathrm{e}(j\alpha) \\ &= \sum_{j=1}^{M} r_{j-1} \, \mathrm{e}(j\alpha) + \sum_{j=M+1}^{N} r_{j-M-1} \, \mathrm{e}(j\alpha) \\ &= \sum_{l=0}^{M-1} r_l \, \mathrm{e}((l+1)\alpha) + \sum_{l=0}^{N-M-1} r_l \, \mathrm{e}((l+M+1)\alpha) \\ &= \mathrm{e}(\alpha) P_M(\alpha) + \mathrm{e}((M+1)\alpha) P_{n-M}(\alpha) \quad \text{for } M < k \leqslant N. \end{aligned} \tag{18}$$

If $M < k \leqslant N$, then it follows from (14) and (18) that

$$\begin{aligned} |\phi_k(F_N, \alpha)| &\leqslant |P_N(\alpha)| + |P_{N-M}(\alpha)| \\ &\leqslant (2+\sqrt{2})(M^{1/2} + (N-M)^{1/2}) \leqslant 2(2+\sqrt{2}) M^{1/2} \\ &\leqslant 2(2+\sqrt{2}) k^{1/2} \qquad \text{for } M < k \leqslant N, \end{aligned}$$

i.e., (6) holds. Similarly, for $k \leqslant M$, (6) follows from (14) and (17).

Moreover, by (15) and (16) we have

$$\begin{aligned} C_2(F_N) &\geqslant \left|\sum_{n=1}^{M-1} f_n f_{n+M}\right| = \left|\sum_{n=1}^{M-1} f_n^2\right| = M-1 = N - \lfloor N/2 \rfloor - 1 \\ &\geqslant N - \frac{N}{2} - \frac{N}{4} = \frac{N}{4}. \end{aligned}$$

This proves (7) and completes the proof of Theorem 4.

References

[1] J. Cassaigne, S. Ferenczi, C. Mauduit, J. Rivat, and A. Sárkőzy, *On finite pseudorandom binary sequences III (The Liouville function I)*, Acta Arith. **87** (1999), 367–390.

[2] ———, *On finite pseudorandom binary sequences IV (The Liouville function II)*, Acta Arith. **95** (2000), 343–359.

[3] P. Erdős and P. Turán, *On a problem in the theory of uniform distribution I, II*, Indag. Math. **10** (1948), 370–378; 406–413.

[4] S.W. Graham and G. Kolesnik, *Van der Corput's method of exponential sums*, London Mathematical Society Lecture Note Series, vol. 126, Cambridge University Press, Cambridge, 1991.

[5] C. Mauduit and A. Sárkőzy, *On finite pseudorandom binary sequences I: Measure of pseudorandomness, the Legendre symbol*, Acta Arith. **82** (1997), 365–377.

[6] ———, *On finite pseudorandom binary sequences II: the Champernowne, Rudin-Shapiro and Thue-Morse sequences. A further construction*, J. Number Theory **73** (1998), 256–276.

[7] ———, *On finite pseudorandom binary sequences, V. On $(n\alpha)$ and $(n^2\alpha)$ sequences*, Monatsh. Math. **129** (2000), 197–216.

[8] ———, *On finite pseudorandom binary sequences, VI. On $(n^k\alpha)$ sequences*, Monatsh. Math. **130** (2000), 281–298.

[9] H.L. Montgomery, *Ten lectures on the interface between analytic number theory and harmonic analysis*, CBMS Regional Conference Series in Mathemaatics, vol. 84, American Math. Soc., Providence, RI, 1994.

[10] M. Queffelec, *Substitution dynamical systems–spectral analysis*, Lecture Notes in Math., vol. 1294, Springer Verlag, New-York–Berlin, 1987.

[11] J. Rivat, *On pseudo-random properties of $P(n)$ and $P(n+1)$*, Period. Math. Hungar. **43** (2001), 121–136.

[12] W. Rudin, *Some theorems on Fourier coefficients*, Proc. Amer. Math. Soc. **10** (1959), 855–859.

[13] H. S. Shapiro, *Extremal problems for polynomials and power series*, Ph.D. thesis, M.I.T., 1951.

[14] A. Sárkőzy, *A finite pseudorandom binary sequence*, Studia Sci. Math. Hungar. **38** (2001), 377–384.

[15] J. D. Vaaler, *Some extremal functions in Fourier analysis*, Bull. Amer. Math. Soc. **12** (1985), 83–216.

Christian Mauduit, Institut de Mathématiques de Luminy, CNRS-UPR 9016, 163 avenue de Luminy, Case 907, 13288 Marseille Cedex 9, France
E-mail address: mauduit@iml.univ-mrs.fr

Joël Rivat, Institut Elie Cartan de Nancy, Université Henri Poincaré, B.P. 239, 54506 Vandœuvre-les-Nancy, France
E-mail address: rivat@iecn.u-nancy.fr

András Sárközy, Eőtvős Loránd University, Department of Algebra and Number Theory, Pazmany Peter setany 1/c, H-1117 Budapest, Hungary
E-mail address: sarkozy@cs.elte.hu