Uniform Distribution Theory **6** (2011), no.2, 143–156

uniform distribution theory

ON THE SYMMETRY OF FINITE PSEUDORANDOM BINARY SEQUENCES

Balázs Sziklai

ABSTRACT. K. Gyarmati introduced the symmetry measure in order to study pseudorandomness of finite binary sequences. This paper focuses on the generalization of this measure. We will give upper and lower bounds for the generalized measures. We will also give some examples which show that these generalizations are indeed useful.

Communicated by Christian Mauduit

1. Introduction

Recently many improvements have been made to generate and study finite binary sequences related to pseudorandomness. First in order to study pseudorandom properties certain statistical tests were introduced which classify sequences pseudorandom or not [7]. This method has the disadvantage that it only categorizes sequences, but it does not measure the quality of pseudorandomness. In 1996 C. Mauduit and A. Sárközy introduced new measures of pseudorandomness [8].

Throughout the paper we write

$$E_N = \{e_1, e_2, \dots, e_N\} \in \{-1, +1\}^N.$$

The *well-distribution measure* is defined by

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

2010 Mathematics Subject Classification: 11K06, 11K38.
Keywords: Pseudorandom, symmetry.
The author thanks the funding of the Hungarian Academy of Sciences under its Momentum Programme (LD-004/2010).

where the maximum is taken over all a, b, t such that

$$a \in \mathbb{Z}, b, t \in \mathbb{N} \quad \text{and} \quad 1 \leq a + b \leq a + tb \leq N.$$

While the *correlation measure* of order k is

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

where $D = (d_1, \dots, d_k) \in \mathbb{N}^k$, $0 < d_1 < \dots < d_k$ and the maximum is taken over all D and M such that $M + d_k \leq N$.

Beyond the above mentioned measures Mauduit and Sárközy also studied the Legendre symbol as a natural candidate for constructing pseudorandom sequences [8]. They introduced the following construction: for an arbitrary prime p write

$$e_n = \left(\frac{n}{p}\right), \quad E_{p-1} = (e_1, \dots, e_{p-1}). \tag{1}$$

It can be shown that both the well-distribution and correlation measure of E_{p-1} are small. This construction can be further extended. Goubin, Mauduit and Sárközy constructed large families of pseudorandom sequences by replacing $f(n)$ for $1 \leq n \leq p-1$ in place of n in (1) (see [4]). Still E_{p-1} has one bad feature which makes it unsuitable for some applications. Namely, if $p = 4k + 1$ for some integer k, then

$$\left(\frac{a}{p}\right) = \left(\frac{p-a}{p}\right)$$

and for $p = 4k + 3$,

$$\left(\frac{a}{p}\right) = -\left(\frac{p-a}{p}\right)$$

making E_{p-1} completely symmetric. To avoid this situation Gyarmati [5] introduced the symmetry measure:

$$S(E_N) = \max_{a<b} |H(E_N, a, b)| = \max_{a<b} \left|\sum_{j=0}^{[\frac{b-a}{2}]-1} e_{a+j} e_{b-j}\right|. \tag{2}$$

She also proved that the first $\frac{p-1}{2}$ elements of E_{p-1} have small symmetry measure. The symmetry property of finite binary sequences can be further studied. Every sequence $\{e_1, e_2, \dots, e_N\} \in \{-1, +1\}^N$ contains a large symmetrical subset since both $\{e_i : \ e_i = e_{N-i}\}$ and $\{e_i : \ e_i = -e_{N-i}\}$ are symmetrical and one of them is large. This implies that symmetrical patterns can occur quite frequently. Among the possible forms of these patterns only intervals have been studied previously. In general, we do not want to consider any binary sequence which shows some kind of symmetrical pattern as pseudorandom.

There are several ways to generalize the symmetry measure introduced by Gyarmati. The two most basic concepts - multiple symmetry centers and arithmetically symmetric subsequences - are presented in this paper.

Let

$$A = (a_1, a_2, \ldots, a_t) \in \mathbb{N}^t, \quad B = (b_1, b_2, \ldots, b_t) \in \mathbb{N}^t,$$

such that

$$1 \le a_i < b_i \le N \quad \text{for } i = 1, 2, \ldots, t \quad \text{and} \quad a_1 + b_1 < a_2 + b_2 < \cdots < a_t + b_t.$$

We write

$$SM(E_N, t) = \max_{A,B} \left\{ \sum_{i=1}^{t} \left| \sum_{j=0}^{\left\lceil \frac{b_i - a_i}{2} \right\rceil - 1} (e_{a_i+j} e_{b_i-j}) \right| \right\}, \tag{3}$$

where the maximum is taken over all possible A's and B's which satisfy the above conditions. In this way $SM(E_N, t)$ measures the sum of the t largest symmetrical subsequences with t *different* symmetry centers. Note that in case of an N long sequence there are $2N-3$ possible symmetry centers, therefore there is an upper limit on t, namely $t \le 2N - 3$.

We define $SD(E_N, d)$ by

$$SD(E_N, d) = \max_{a,b} \left| \sum_{j=0}^{\left\lceil \frac{b-a}{2d} \right\rceil - 1} e_{a+dj} e_{b-dj} \right|, \tag{4}$$

which counts the largest symmetric arithmetic progression of difference $d \in \mathbb{N}$ outgoing from a specific center (determined by a and b).

Note that

$$SM(E_N, 1) = SD(E_N, 1),$$

$$\left| SM(E_N, 1) - S(E_N) \right| = \left| SD(E_N, 1) - S(E_N) \right| \le 1$$

so indeed both measures are generalizations of the symmetry measure.

The main purpose of this paper is to give upper and lower bounds for the newly introduced measures. Gyarmati obtained analogue results for $S(E_N)$ in the paper [5]. Lower bounds for the measures $W(E_N)$ and $C_k(E_N)$ have been obtained in [1], [3], and [6] and average values of the measures $W(E_N)$ and $C_k(E_N)$ have been obtained in [2], [3], and [6].

Theorem 1. *There is an integer N_0 such that for all E_N for which $N > N_0$ we have:*

$$\text{a)} \qquad SM(E_N, t) > \frac{\sqrt{t}}{2}\left(\sqrt{\frac{N}{3}} - \sqrt{t}\right), \tag{5}$$

$$\text{b)} \qquad SD(E_N, d) > \frac{1}{2d}\left(\sqrt{\frac{N}{2}} - 1\right) \tag{6}$$

In the next theorem we will estimate

$$SM(E_N, t) \quad \text{and} \quad SD(E_N, d)$$

for 'random' binary sequences

$$E_N \in \{1, -1\}^N$$

(i.e., we are choosing each $E_N \in \{1, -1\}^N$ with probability $1/2^N$). In this way the upper bounds hold only for the majority of sequences.

Theorem 2. *For all $\epsilon > 0$ there exits $N_0(\epsilon)$ such that for $N > N_0(\epsilon)$ we have:*

$$\text{a)} \qquad P\Big(SM(E_N, t) \ < \ 3t\,(N \log\, N)^{1/2}\Big) > 1 - \epsilon, \tag{7}$$

$$\text{b)} \qquad P\left(SD(E_N, d) < \frac{3}{\sqrt{d}}(N \log N)^{1/2}\right) > 1 - \epsilon. \tag{8}$$

for $1 \le t \le 2N - 3$ and $1 \le d \le \sqrt{N}$.

2. Proofs

P r o o f o f T h e o r e m 1. We will use the same exponential sum in order to prove both parts of Theorem 1, namely, let

$$E_N = \{e_1, e_2, \ldots, e_N\}$$

be the usual N long binary sequence and

$$f(z) = \sum_{n=1}^{N} e_n z^n.$$

Throughout the proof $e(\alpha)$ will denote $\exp(2\pi i\alpha)$. By the Cauchy-Schwarz inequality and the Parseval formula we obtain:

$$K \stackrel{\text{def}}{=} \int_0^1 \Big|f\big(e(\alpha)\big)\Big|^4 d\alpha \ge \left(\int_0^1 \Big|f\big(e(\alpha)\big)\Big|^2 d\alpha\right)^2 = N^2. \tag{9}$$

On the other hand,

$$K = \int_0^1 \left| f^2\big(e(\alpha)\big) \right|^2 d\alpha = \int_0^1 \left| \sum_{n=1}^{N} \sum_{m=1}^{N} e_n e_m e\big((n+m)\alpha\big) \right|^2 d\alpha$$

$$= \int_0^1 \left| \sum_{k=2}^{2N} \left(\sum_{\max\{1,\ k-N\} \le n \le \min\{N,\ k-1\}} e_n e_{k-n} \right) e(k\alpha) \right|^2 d\alpha$$

$$= \sum_{k=2}^{2N} \left| \sum_{\max\{1,\ k-N\} \le n \le \min\{N,\ k-1\}} e_n e_{k-n} \right|^2 . \tag{10}$$

Now this sum can be estimated from above by both symmetry measures. First for $2 \le k_1, k_2, \ldots, \ k_t \le 2N$ such that $k_1, \ldots, k_t$ are different we have

$$\sum_{i=1}^{t} \left| \sum_{\max\{1,\ k_i-N\} \le n \le \min\{N,\ k_i-1\}} e_n e_{k_i-n} \right|^2$$

$$\le \left(\sum_{i=1}^{t} \left| \sum_{\max\{1,\ k_i-N\} \le n \le \min\{N,\ k_i-1\}} e_n e_{k_i-n} \right| \right)^2$$

$$\le \ (2SM(E_N, t) + t)^2,$$

where the second inequality follows from the definition of $SM(E_N, t)$. This yields

$$N^2 \le K = \sum_{k=2}^{2N} \left| \sum_{\max\{1,\ k-N\} \le n \le \min\{N,\ k-1\}} e_n e_{k-n} \right|^2$$

$$\le \left\lceil \frac{2N-1}{t} \right\rceil \big(2SM(E_N, t) + t\big)^2. \tag{11}$$

Finally, we can conclude

$$SM(E_N, t) \ge \frac{N}{2\sqrt{\left\lceil \frac{2N-1}{t} \right\rceil}} - \frac{t}{2} > \frac{N}{2\sqrt{\frac{2N+t}{t}}} - \frac{t}{2} > \frac{\sqrt{t}}{2} \left(\sqrt{\frac{N}{3}} - \sqrt{t} \right)$$

which completes the proof of the first part of Theorem 1 (5). Similarly, we can estimate (10) by using the definition of $SD(E_N, d)$:

$$\left| \sum_{\max\{1,\ k-N\} \le n \le \min\{N,\ k-1\}} e_n e_{k-n} \right|^2 \le (2dSD(E_N, d) + 1)^2.$$

So that

$$N^2 \le K \le \sum_{k=2}^{2N} \left| \sum_{\max\{1,\ k-N\} \le n \le \min\{N,\ k-1\}} e_n e_{k-n} \right|^2 \le (2N-1)(2dSD(E_N,d)+1)^2,$$

from which (6) follows. □

Proof of Theorem 2. The following lemma has been used in the paper of Cassaigne et al. [3], we shall need it as well. They omitted the proof, thus we prove it here.

Lemma 1. *Let k be a positive integer such that $k \le t^{2/3}$. If $t \to \infty$, then*

$$\binom{t}{[t/2]+k} = \binom{t}{[t/2]} \exp\left(\frac{-2k^2}{t} + O\left(\frac{k^3}{t^2}\right)\right).$$

Proof of Lemma 1. We denote by A the fraction of binomials. We discuss the case if t is even, the proof goes essentially the same if t is odd. We have

$$\begin{aligned} A &= \frac{\binom{t}{[t/2]+k}}{\binom{t}{[t/2]}} = \frac{((t/2)!)^2}{(t/2-k)!(t/2+k)!} \\ &= \frac{(t/2+1-k)(t/2+2-k)\dots(t/2+k-k)}{(t/2+1)(t/2+2)\dots(t/2+k)} \\ &= \left(1-\frac{k}{t/2+1}\right)\left(1-\frac{k}{t/2+2}\right)\dots\left(1-\frac{k}{t/2+k}\right). \end{aligned} \tag{12}$$

It is clear from (12) that

$$\left(1-\frac{2k}{t}\right)^k \le A \le \left(1-\frac{2k}{t+2k}\right)^k.$$

It is enough to prove that $A \cdot \exp\left(\frac{2k^2}{t}\right) = \exp\left(O\left(\frac{k^3}{t^2}\right)\right)$ when $t \to \infty$. We use the facts that $\left(1+\frac{x}{n}\right)^n \le e^x$ and $\left(1-\frac{x}{n+x}\right)^n \ge e^{-x}$ if $n \ge 1$ for all $x > -n$. The first estimate can be verified by taking the Taylor series of $e^{x/n}$ while the

second is the reciprocal of the first. We have

$$\begin{aligned} A \cdot e^{\frac{2k^2}{t}} &\geq \left(1 - \frac{2k}{t}\right)^k e^{\frac{2k^2}{t}} \\ &= \left(\left(1 - \frac{2k}{t-2k+2k}\right)^{t-2k}\right)^{\frac{k}{t-2k}} e^{\frac{2k^2}{t}} \\ &\geq e^{\frac{-2k^2}{t-2k}} e^{\frac{2k^2}{t}} = e^{2k^2\left(\frac{1}{t} - \frac{1}{t-2k}\right)} = e^{O(k^3/t^2)}. \end{aligned}$$

On the other hand,

$$\begin{aligned} A \cdot e^{\frac{2k^2}{t}} &\leq \left(1 - \frac{2k}{t+2k}\right)^k e^{\frac{2k^2}{t}} \\ &= \left(\left(1 - \frac{1}{(t+2k)/2k}\right)^{\frac{t+2k}{2k}}\right)^{2k^2/(t+2k)} e^{\frac{2k^2}{t}} \\ &\leq e^{\frac{-2k^2}{t+2k}} e^{\frac{2k^2}{t}} = e^{2k^2\left(\frac{1}{t} - \frac{1}{t+2k}\right)} = e^{O(k^3/t^2)}. \end{aligned}$$

Indeed, the above estimates together complete the proof of the lemma. □

Write $L = \frac{3}{\sqrt{d}}(N \log N)^{1/2}$. Then we have

$$\begin{aligned} P(SD(E_N, d) > L) =& P\left(\max_{a<b} \left| \sum_{j=0}^{\lceil (b-a)/2d \rceil - 1} e_{a+dj} e_{b-dj} \right| > L\right) \\ &\leq \sum_{a<b} P\left(\left| \sum_{j=0}^{\lceil (b-a)/2d \rceil - 1} e_{a+dj} e_{b-dj} \right| > L\right) \\ &\leq \binom{N}{2} \max_{a<b} P\left(\left| \sum_{j=0}^{\lceil (b-a)/2d \rceil - 1} e_{a+dj} e_{b-dj} \right| > L\right), \end{aligned}$$

where $a, b \in \mathbb{N}$ such that $1 \leq a < b \leq N$. It suffices to show that for all such a's and b's we have:

$$P\left(\left| \sum_{j=0}^{\lceil (b-a)/2d \rceil - 1} e_{a+dj} e_{b-dj} \right| > L\right) < \frac{2\epsilon}{N^2}. \tag{13}$$

Let $l = [(b-a)/2d]$. If $l < L$, then the probability in (13) is zero so we may assume that $l \geq L$. Write

$$M \stackrel{\text{def}}{=} 3\sqrt{2}(l \log l)^{1/2} \tag{14}$$

and

$$\big|\{j:\ 0\le j\le l-1,\ e_{a+dj}e_{b-dj}=-1\}\big|=h. \tag{15}$$

Then we have

$$\begin{aligned}\sum_{j=0}^{l-1}e_{a+dj}e_{b-dj} &= \big|\{j:\ 0\le j\le l-1,\ e_{a+dj}e_{b-dj}=1\}\big|\\ &\quad-\big|\{j:\ 0\le j\le l-1,\ e_{a+dj}e_{b-dj}=-1\}\big|\\ &= (l-h)-h=l-2h.\end{aligned}$$

The number of $2l$-long sequences for which (15) holds is

$$\binom{l}{h}2^h2^{l-h}=\binom{l}{h}2^l.$$

So the probability of (15) is $\frac{1}{2^l}\binom{l}{h}$. By the definition $l=[(b-a)/2d]\le N/2d$. It follows that $M\le L$, therefore, it is enough to consider

$$P\left(\left|\sum_{j=1}^{l-1}e_{a+dj}e_{b-dj}\right|>M\right)=\sum_{h:\ |l-2h|>M}\frac{1}{2^l}\binom{l}{h}=\frac{1}{2^l}\sum_{h:\ |h-l/2|>M/2}\binom{l}{h}.$$

If N is large enough, so is l since $l\ge L$. Using Lemma 1 we can estimate the above sum.

$$\begin{aligned}\sum_{h:\ |h-l/2|>M/2}\binom{l}{h} &= \sum_{h:\ |h-l/2|>\frac{3\sqrt{2}}{2}(l\log l)^{1/2}}\binom{l}{h}\\ &< l\binom{l}{[l/2]+\left[\frac{3\sqrt{2}}{2}(l\log l)^{1/2}\right]}\\ &\le l\binom{l}{[l/2]}\exp\left(-2\left(\frac{3\sqrt{2}}{2}(l\log l)^{1/2}\right)^2\frac{1}{l}+o(1)\right)\\ &< l\binom{l}{[l/2]}\exp\big(-9\log l+o(1)\big)<\frac{2^l}{l^8}.\end{aligned} \tag{16}$$

Finally, by (14) and (16) we conclude

$$\begin{aligned}P\left(\left|\sum_{j=0}^{l-1}e_{a+dj}e_{b-dj}\right|>L\right) &\le P\left(\left|\sum_{j=0}^{l-1}e_{a+dj}e_{b-dj}\right|>M\right)\\ &<\frac{1}{2^l}\frac{2^l}{l^8}\le\frac{1}{L^8}=\left(\frac{d}{9N\log N}\right)^4<\frac{2\epsilon}{N^2},\end{aligned}$$

which proves (13) and completes the second part of Theorem 2 (8). The first part immediately follows from the fact that

$$tdSD(E_N, d) \geq SM(E_N, t)$$

is true for every d. In particular, when $d = 1$

$$P\Big(SM(E_N, t) > 3t(N \log N)^{1/2}\Big) \leq P\Big(tSD(E_N, 1) > 3t(N \log N)^{1/2}\Big) < \epsilon. \quad \square$$

3. Constructions

To justify the newly introduced generalized symmetry measures, we now show a sequence which passes the usual statistical tests (i.e., it has small correlation, well-distribution and symmetry measure) even though it has a deeply symmetric structure.

Theorem 3. *There exists a sequence for which $S(E_N)$ is small, but $SD(E_N, 2)$ is big, that is*

$$S(E_N) \leq 54N^{1/2} \log N,$$
$$\left\lceil \frac{N-1}{4} \right\rceil \leq SD(E_N, 2). \tag{17}$$

Proof of Theorem 3. For the proof we will use two lemmas.

Lemma 2. *If p is a prime number, $f(x) \in F_p[x]$ is a polynomial of degree k such that it is not of the form $f(x) = b(g(x))^2$ with $b \in F_p$, $g(x) \in F_p[x]$, and X, Y are real numbers with $0 < Y \leq p$, then writing*

$$\chi_p^*(n) = \begin{cases} \left(\frac{n}{p}\right) & \text{for} \quad (n, p) = 1, \\ 0 & \text{for} \quad p|n, \end{cases}$$

we have

$$\left| \sum_{X < n \leq X+Y} \chi_p^*\big(f(n)\big) \right| < 9kp^{1/2} \log p\,.$$

For the proof see [8, p. 373]. Indeed there this result is deduced from Weil's theorem [9]. Also we need the following result of Gyarmati we mentioned previously.

Lemma 3. *The first* $\frac{p-1}{2}$ *elements of* E_{p-1} *have small symmetry measure. That is*

$$S(E_{\frac{p-1}{2}}) < 18p^{1/2}\log p,$$

where

$$E_{\frac{p-1}{2}} = \left(\left(\frac{1}{p}\right), \left(\frac{2}{p}\right), \ldots, \left(\frac{(p-1)/2}{p}\right)\right).$$

For the proof see [5].

Now we can construct a sequence which has the required property. We define

$$\widehat{E}_{p-1} = (e_1, e_2, \ldots, e_j, \ldots, e_{p-1}),$$

where

$$e_j = \begin{cases} \left(\frac{j}{p}\right) & \text{if } \ 1 \le j \le \frac{p-1}{2}, \\ \left(\frac{p-j}{p}\right) & \text{if } \ \frac{p+1}{2} \le j \le p-1, \quad j \ \text{ is even}, \\ -\left(\frac{p-j}{p}\right) & \text{if } \ \frac{p+1}{2} \le j \le p-1, \quad j \ \text{ is odd}. \end{cases}$$

In this way the second part of the sequence is the reflection of the first part except that we changed the sign of every second term. We will show that $\widehat{E}_{p-1}$ has small $S(\widehat{E}_{p-1})$ although $SD(\widehat{E}_{p-1}, 2)$ is huge. Indeed it is easy to check the second statement. Using (4) we obtain:

$$\left|\sum_{j=0}^{\lceil (p-2)/4\rceil - 1} e_{1+2j}e_{p-1-2j}\right| = \left\lceil \frac{p-2}{4} \right\rceil \le SD(E_{p-1}, 2). \tag{18}$$

To prove the first statement we have to check that $H(\widehat{E}_{p-1}, a, b)$ (defined by (2)) is small for every a and b such that $1 \le a < b \le p-1$.

I. If $b \le \frac{p-1}{2}$, we can directly apply Lemma 3 and obtain

$$\max_{a<b\le (p-1)/2} \left|H(\widehat{E}_{p-1}, a, b)\right| < 18p^{1/2}\log p. \tag{19}$$

II. If $a > \frac{p-1}{2}$, we have a very similar case, except here the sign of every second term was changed. We denote the subsequence of the first $(p-1)/2$ element of $\widehat{E}_{p-1}$ by $E_{\frac{p-1}{2}}$ and the subsequence of the remaining $(p-1)/2$ element by $\widehat{E}_{\frac{p-1}{2}}$. Observe that for every given c and d the magnitude of the sum

$$H(\widehat{E}_{\frac{p-1}{2}}, c, d) = \sum_{j=0}^{[(c-d)/2]-1} e_{c+j}e_{d-j}$$

is the same as $H(E_{\frac{p-1}{2}}, \frac{p-1}{2}+1-d, \frac{p-1}{2}+1-c)$ up to sign. Hence,

$$\left|H(\widehat{E}_{\frac{p-1}{2}}, c, d)\right| = \left|H(E_{\frac{p-1}{2}}, \frac{p-1}{2}+1-d, \frac{p-1}{2}+1-c)\right| < 18p^{1/2}\log p. \quad (20)$$

III. The remaining case is where $a \le \frac{p-1}{2}$ and $b > \frac{p-1}{2}$.

a) For $a = p-b$ we have $\max\limits_{a=p-b}\left|H(\widehat{E}_{p-1}, a, b)\right| \le 1$ by construction.

b) If $a > p-b$, then

$$\left|\sum_{j=0}^{[\frac{b-a}{2}]-1} e_{a+j}e_{b-j}\right| = \left|\sum_{j=0}^{\frac{p-1}{2}-a} e_{a+j}e_{b-j} + \sum_{j=\frac{p+1}{2}-a}^{[\frac{b-a}{2}]-1} e_{a+j}e_{b-j}\right|$$

$$\le \left|\sum_{j=0}^{\frac{p-1}{2}-a} e_{a+j}e_{b-j}\right| + 18p^{\frac{1}{2}}\log p. \quad (21)$$

In this way we split the sum into a first (outer) and a second (inner) part. The inner part of the sum can be estimated by (20), so we only need to deal with the outer part.

$$\left|\sum_{j=0}^{\frac{p-1}{2}-a} e_{a+j}e_{b-j}\right|$$

$$= \left|\sum_{j=0,\ \mathrm{j\ even}}^{\frac{p-1}{2}-a}\left(\frac{a+j}{p}\right)\left(\frac{p-b+j}{p}\right) - \sum_{j=1,\ \mathrm{j\ odd}}^{\frac{p-1}{2}-a}\left(\frac{a+j}{p}\right)\left(\frac{p-b+j}{p}\right)\right|$$

$$\le \left|\sum_{k=0}^{[\frac{p-1}{4}-\frac{a}{2}]}\left(\frac{a+2k}{p}\right)\left(\frac{p-b+2k}{p}\right)\right|$$

$$+ \left|\sum_{k=0}^{[\frac{p-1}{4}-\frac{a}{2}-\frac{1}{2}]}\left(\frac{a+2k+1}{p}\right)\left(\frac{p-b+2k+1}{p}\right)\right|$$

$$\le \left|\sum_{k=0}^{[\frac{p-1}{4}-\frac{a}{2}]}\left(\frac{(2k)^2+(a-b)2k-ab}{p}\right)\right|$$

$$+ \left|\sum_{k=0}^{[\frac{p-1}{4}-\frac{a}{2}-\frac{1}{2}]}\left(\frac{(2k+1)^2+(a-b)(2k+1)-ab}{p}\right)\right|.$$

Let

$$f_1(x) = 4x^2 + 2(a-b)x - ab \in \mathbb{F}_p[x]$$

and

$$f_2(x) = 4x^2 + 2(a-b+2)x - ab + a - b + 1 \in \mathbb{F}_p[x].$$

It is easy to check that both $f_1(x)$ and $f_2(x)$ can be written as

$$b_1(g_1(x))^2 \quad \text{and} \quad b_2(g_2(x))^2,$$

respectively, if and only if $a + b \equiv 0 \pmod{p}$. Now this is impossible since we assumed that $a > p - b$. Applying Lemma 2 twice we get

$$\begin{aligned}\left|\sum_{j=0}^{\frac{p-1}{2}-a} e_{a+j}e_{b-j}\right| &= \left|\sum_{k=0}^{[\frac{p-1}{4}-\frac{a}{2}]}\left(\frac{f_1(k)}{p}\right)\right| + \left|\sum_{k=0}^{[\frac{p-1}{4}-\frac{a}{2}-\frac{1}{2}]}\left(\frac{f_2(k)}{p}\right)\right| \\ &= \left|\sum_{X_1\le k\le X_1+Y_1}\chi_p^*(f_1(k))\right| + \left|\sum_{X_2\le k\le X_2+Y_2}\chi_p^*(f_2(k))\right| \\ &\le 18p^{1/2}\log p + 18p^{1/2}\log p = 36p^{1/2}\log p. \qquad (22)\end{aligned}$$

By (21) and (22) we obtain that $\max_{a>p-b}\left|H(\widehat{E}_{p-1}, a, b)\right| \le 54p^{1/2}\log p$.

c) For $a < p - b$ we can repeat the previous argument by symmetry. Again we split the sum. The inner part can be estimated by (19) and for the outer part we obtain the same result as in (22).

Now by (19), (20) and (22) we gather

$$S(\widehat{E}_{p-1}) \le 54p^{1/2}\log p,$$

which together with (18) completes the proof of Theorem 3. □

Now we show an example for a sequence E_N for which $SD(E_N, d)$ is small for every possible d.

Theorem 4. *If p is an odd prime and if we write*

$$E_{(p-1)/2} = \left(\left(\frac{1}{p}\right), \left(\frac{2}{p}\right), \ldots, \left(\frac{(p-1)/2}{p}\right)\right),$$

then we have

$$SD(E_{(p-1)/2}, d) \le 18p^{1/2}\log p$$

for every $1 \le d \le \frac{p-1}{2}$.

P r o o f o f T h e o r e m 4. By the definition

$$SD(E_{\frac{p-1}{2}},d)=\max_{a<b}\left|\sum_{j=0}^{\lceil (b-a)/2d\rceil -1} e_{a+dj}\ e_{b-dj}\right|.$$

We have to estimate the sum for all possible a's and b's.

$$\left|\sum_{j=0}^{\lceil (b-a)/2d\rceil -1} e_{a+dj}e_{b-dj}\right| = \sum_{j=0}^{\lceil (b-a)/2d\rceil -1}\left(\frac{a+dj}{p}\right)\left(\frac{b-dj}{p}\right)$$
$$= \sum_{j=0}^{\lceil (b-a)/2d\rceil -1}\left(\frac{-(dj)^2+(b-a)dj+ab}{p}\right). \tag{23}$$

Let $f(x)=-d^2x^2+(b-a)dx+ab\in\mathbb{F}_p[x]$. It is easy to see that $f(x)$ is the form of $b(g(x))^2$ if and only if $a+b\equiv 0 \pmod p$. In the present case this is impossible, since $1\le a<b\le (p-1)/2$. Applying Lemma 2 with 0 and $(b-a)/2d$ in place of X and Y we get

$$\left|\sum_{X\le j\le X+Y}\chi_p^*(f(j))\right| = \sum_{j=0}^{\lceil (b-a)/2d\rceil -1}\left(\frac{-(dj)^2+(b-a)dj+ab}{p}\right) < 18p^{1/2}\log p. \tag{24}$$

From (23) and (24) we obtain $S(E_{(p-1)/2})<18p^{1/2}\log p$ which gives the desired result. □

It seems plausible that similar constructions may work for $SM(E_N,t)$, however proving it is much harder. It remains to be shown that there exist sequences for which $S(E_N)$ is small, but $SM(E_N,t)$ is big.

Finally I would like to thank Katalin Gyarmati for the valuable discussions and András Bíró for the helpful advices.

REFERENCES

[1] ALON, N. – KOHAYAKAWA, Y.– MAUDUIT, C. – MOREIRA, C. G. – RÖDL, V.: *Measures of pseudorandomness for finite sequences: minimal values*, Combin. Probab. Comput. **15** (2006), no. 1–2, 1–29.

[2] ALON, N. – KOHAYAKAWA, Y. – MAUDUIT, C.– MOREIRA, C. G. – RÖDL, V.: *Measures of pseudorandomness for finite sequences: typical values*, Proc. Lond. Math. Soc. (3) **95** (2007), no. 3, 778–812.

[3] CASSAIGNE, J. – C. MAUDUIT, C. – SÁRKÖZY, A.: *On finite pseudorandom binary sequences VII: The measure of pseudorandomness*, Acta Arith. **103** (2002), no. 2, 97–118.
[4] GOUBIN, L. – MAUDUIT, C. – SÁRKÖZY, A.: *Construction of large families of pseudorandom binary sequences*, J. Number Theory **106** (2004), 56–69.
[5] GYARMATI, K.: *On a pseudorandom property of binary sequences*, Ramanujan J. **8** (2004), no. 3, 289–302.
[6] KOHAYAKAWA, Y. – MAUDUIT, C. – MOREIRA, C. G. – RÖDL, V.: *Measures of pseudorandomness for finite sequences: minimum and typical values.* In: *Proceedings of WORDS'03, TUCS Gen. Publ.* **27**, Turku Cent. Comput. Sci., Turku, 2003, pp. 159–169.
[7] KNUTH, D. E.: *Art of Computer Programming*, **2**, Addison-Wesley, Reading, Massachusetts 1997.
[8] MAUDUIT, C. – SÁRKÖZY, A.: *On finite pseudorandom binary sequences I: Measure of pseudorandomness, the Legendre symbol*, Acta Arith. **82** (1997), no. 4, 365–377.
[9] WEIL, A.: *Sur les courbes algébriques et les veriétés qui s'en déduisent*, Actualités Sci. Ind. **1041**, Publ. Inst. Math. Univ. Strasbourg **7** 1945, Hermann et Cie., Paris, 1948.

Received March 3, 2010
Accepted November 7, 2010

Balázs Sziklai
Institute of Economics
Hungarian Academy of Sciences
Budaörsi út 45.
H-1112, Budapest
HUNGARY
E-mail: sziklai@econ.core.hu.