ACTA ARITHMETICA
135.2 (2008)

Pseudorandom binary sequences and lattices

by

Katalin Gyarmati (Budapest), Christian Mauduit (Marseille)
and András Sárközy (Budapest)

1. Introduction. Pseudorandom binary sequences play a role of basic importance in applications, in particular, in cryptography. The notion of pseudorandomness is usually defined in terms of computational complexity (see, e.g., [8]). As this approach has certain weak points, Mauduit and Sárközy [4] initiated another, more constructive approach (see [10] for a survey of the related work and for a comparison of a two approaches).

In the applications (e.g., in connection with image or bit map encryption) one also needs the multidimensional extension of this theory. Therefore Hubert, Mauduit and Sárközy [3], [5], [6] extended the constructive theory of pseudorandom binary sequences to the multidimensional situation by studying pseudorandom binary lattices. It turns out that the multidimensional case is much more difficult than the one-dimensional case; it takes a considerable effort to generalize the one-dimensional methods, results and constructions, and in most cases only much weaker partial results are achieved. Thus it is a natural question to ask: does one really need the multidimensional theory? Couldn't one utilize the simpler and more effective one-dimensional theory in the multidimensional case? Aren't there any simple and cheap but at the same time satisfactory ways to convert the one-dimensional results and constructions into multidimensional ones? In general, what is the connection between the one-dimensional and multidimensional cases? In this paper we study these questions. More precisely, since the two-dimensional case is simpler and more important than the three- or higher-dimensional ones, we will restrict ourselves to the study of the links between the one-dimensional and two-dimensional cases. How-

2000 *Mathematics Subject Classification*: Primary 11K45.

Key words and phrases: pseudorandom, binary lattice, concatenation.

Research of A. Sárközy partially supported by Hungarian National Foundation for Scientific Research, Grants K49693, K67676 and K72264, French–Hungarian exchange program F-06/48, and the János Bolyai Research Fellowship.

ever, with a little work our results and constructions could be extended to higher dimensions.

2. Some basic definitions and results in one, resp. n dimensions. In [4] Mauduit and Sárközy studied finite binary sequences

$$E_N = \{e_1, \dots, e_N\} \in \{-1, +1\}^N. \tag{2.1}$$

They introduced the following measures of pseudorandomness of such sequences: the *well-distribution measure* of the sequence (2.1) is 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 $a, b, t \in \mathbb{N}$ with $1 \le a \le a+(t-1)b \le N$, and the *correlation measure of order* k of E_N is defined as

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

where the maximum is taken over all $\underline{D} = (d_1, \dots, d_k)$ and M such that $0 \le d_1 < \cdots < d_k \le N - M$. The *combined* (well-distribution-correlation) pseudorandom measure of order k was also introduced:

$$Q_k(E_N) = \max_{a,t,\underline{D}} \left| \sum_{j=0}^{t} e_{ja+d_1} \cdots e_{ja+d_k} \right|$$

where the maximum is taken over all a, t and $\underline{D} = (d_1, \dots, d_k)$ with $d_1 < \cdots < d_k$ such that all the subscripts $ja + d_l$ belong to $\{1, \dots, N\}$. (Note that clearly $Q_1(E_N) = W(E_N)$.) The sequence E_N is considered to be a "good" pseudorandom sequence if $W(E_N)$ and, for "small" k, both $C_k(E_N)$ and $Q_k(E_N)$ are "small" in terms of N (in particular, both are $o(N)$ as $N \to \infty$). Indeed, later Cassaigne, Mauduit and Sárközy [2] showed that this terminology is justified since for fixed k for almost all $E_N \in \{-1, +1\}^N$, the measures $W(E_N), C_k(E_N)$ and $Q_k(E_N)$ are less than $N^{1/2}(\log N)^c$, where the constant c depends on k (see also [1]). It was also shown in [4] that the Legendre symbol forms a "good" pseudorandom binary sequence:

THEOREM A. *There is a number p_0 such that if $p > p_0$ is a prime number*, $k \in \mathbb{N}$, $k < p$ *and if we write*

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

(*where* $\left(\frac{n}{p}\right)$ *denotes the Legendre symbol*), *then*

$$Q_k(E_{p-1}) \le 9kp^{1/2} \log p.$$

The crucial tool in the proof of this theorem was the following consequence of Weil's theorem [12]:

Lemma 1. *Suppose that p is a prime number, χ is a non-principal character modulo p of order d (so that $d \mid p-1$), and the polynomial $f(x) \in \mathbb{F}_p[x]$ ($\mathbb{F}_p$ being the field of modulo p residue classes) has degree k and factorization $f(x) = b(x-x_1)^{d_1}\cdots(x-x_s)^{d_s}$ (where $x_i \neq x_j$ for $i \neq j$) in $\overline{\mathbb{F}}_p$ (the algebraic closure of $\mathbb{F}_p$) with*

$$(d, d_1, \ldots, d_s) = 1. \tag{2.2}$$

Let X, Y be real numbers with $0 < Y \le p$. Then

$$\Big|\sum_{X<n\le X+Y} \chi(f(n))\Big| < 9kp^{1/2}\log p.$$

Note that the same conclusion also holds if assumption (2.2) on $f(x)$ is replaced by

$$f(x) \text{ is not of the form } cg(x)^d \text{ with } c \in \mathbb{F}_p,\ g(x) \in \mathbb{F}_p[x] \tag{2.3}$$

(see [7], [11]).

In [3] Hubert, Mauduit and Sárközy extended this constructive theory of pseudorandomness to n dimensions (see also [5], [6]). Let I_N^n denote the set of n-dimensional vectors all of whose coordinates are $\{0, 1, \ldots, N-1\}$:

$$I_N^n = \{\underline{x} = (x_1, \ldots, x_n) : x_1, \ldots, x_n \in \{0, 1, \ldots, N-1\}\}.$$

We call this set the *n-dimensional N-lattice* or briefly (if n is fixed) the *N-lattice*. A function of the type

$$\eta(\underline{x}) :\ I_N^n \to \{-1, +1\} \tag{2.4}$$

is called an *n-dimensional binary N-lattice* or briefly a *binary lattice*. (Note that in the special case of $n = 1$ these functions are binary sequences $E_N \in \{-1, +1\}^N$.) In [3] the following measures of pseudorandomness of binary lattices were proposed: if $\eta = \eta(\underline{x})$ is an n-dimensional binary N-lattice of the form (2.4), $k \in \mathbb{N}$, and $\underline{u}_i$ ($i = 1, \ldots, n$) denotes the n-dimensional unit vector whose ith coordinate is 1 and the other coordinates are 0, then write

$$Q_k(\eta) = \max_{\underline{B}, \underline{d}_1, \ldots, \underline{d}_k, \underline{T}} \Big| \sum_{j_1=0}^{t_1} \cdots \sum_{j_n=0}^{t_n} \eta(j_1 b_1 \underline{u}_1 + \cdots + j_n b_n \underline{u}_n + \underline{d}_1) \cdots \eta(j_1 b_1 \underline{u}_1 + \cdots + j_n b_n \underline{u}_n + \underline{d}_k) \Big|,$$

where the maximum is taken over all n-dimensional vectors $\underline{B} = (b_1, \ldots, b_n)$, $\underline{d}_1, \ldots, \underline{d}_k$, and $\underline{T} = (t_1, \ldots, t_n)$ such that their coordinates are non-negative integers, $b_1, \ldots, b_n$ are non-zero, $\underline{d}_1, \ldots, \underline{d}_k$ are distinct, and all the points $j_1 b_1 \underline{u}_1 + \cdots + j_n b_n \underline{u}_n + \underline{d}_i$ occurring in the multiple sum belong to the

n-dimensional N-lattice I_N^n. Then $Q_k(\eta)$ is called the *pseudorandom* (briefly PR) measure of order k of η. (Note that in the one-dimensional case $Q_k(\eta)$ is the combined PR-measure Q_k of order k.)

It was proved in [3] that for a fixed $k \in \mathbb{N}$ and for a truly random n-dimensional binary N-lattice $\eta(\underline{x})$ we have

$$N^{n/2} \ll Q_k(\eta) \ll N^{n/2}(\log N^n)^{1/2}$$

with probability $> 1 - \varepsilon$, while the trivial upper bound for $Q_k(\eta)$ is N^n. Thus an n-dimensional binary N-lattice η can be considered as a "good" pseudorandom lattice if the PR measure of order k of η is "small" in terms of N (in particular, $Q_k(\eta) = o(N^n)$ for fixed n and $N \to \infty$) for small k.

Moreover, in [3] an example was given (by using the quadratic character of a finite field) of a "good" n-dimensional binary lattice (for any n).

In the rest of the paper we will restrict ourselves to the special case of $n = 2$, i.e., to two-dimensional binary lattices.

3. Binary lattices whose rows are "good" PR binary sequences. Suppose we want to construct a "good" PR two-dimensional lattice. As mentioned earlier, it is easier to construct binary sequences than binary lattices. Thus one might wish to construct a binary lattice by combining binary sequences. More precisely, assume that a sequence of "good" PR binary sequences $E_N^{(1)}, \dots, E_N^{(j)} = (e_1^{(j)}, \dots, e_N^{(j)}), \dots$ is given; then it is a natural idea to consider the two-dimensional binary lattice η whose jth row is the vector $E_N^{(j)}$, i.e.,

$$\eta((i, j-1)) = e_{i+1}^{(j)} \quad \text{for } j = 1, \dots, N,\ i = 0, 1, \dots, N-1. \tag{3.1}$$

If, say, $E_N^{(1)} = \dots = E_N^{(N)}$, then η is certainly not of PR type. Thus to ensure the pseudorandomness of η one needs an assumption on the connection between the sequences $E_N^{(j)}$. A natural assumption of this type is that the vectors $E_N^{(j)}$ are nearly orthogonal, i.e., the scalar products $(E_N^{(i)}, E_N^{(j)})$ are "small":

$$|(E_N^{(i)}, E_N^{(j)})| = |e_1^{(i)}e_1^{(j)} + e_2^{(i)}e_2^{(j)} + \dots + e_N^{(i)}e_N^{(j)}| \text{ is "small"} \quad \text{for } 1 \le i < j \le N. \tag{3.2}$$

So the question is: if $E_N^{(1)}, \dots, E_N^{(N)}$ are "good" PR binary sequences, and (3.2) holds, does this imply that the lattice η in (3.1) is a "good" PR binary lattice? We will show by an example that the answer is negative. This example shows that from "good" PR binary sequences we cannot construct a "good" lattice in this manner.

THEOREM 1. *Let p be a prime number, and for $j = 1, \dots, p$ define the binary sequence $E_p^{(j)} = (e_1^{(j)}, \dots, e_p^{(j)})$ by*

$$e_i^{(j)} = \begin{cases} \left(\dfrac{i+j}{p}\right) & \text{for } p \nmid i+j, \\ +1 & \text{for } p \mid i+j. \end{cases}$$

Define the binary lattice η by (3.1) (*with p in place of N*) *so that, for $(x, y) \in \{0, 1, \dots, p-1\}^2$,*

$$\eta((x,y)) = e_{x+1}^{(y+1)} = \begin{cases} \left(\dfrac{x+y+2}{p}\right) & \text{for } p \nmid x+y+2, \\ +1 & \text{for } p \mid x+y+2. \end{cases}$$

Then for $k \in \mathbb{N}$, $k < p$, $j = 1, \dots, p$ we have

$$Q_k(E_p^{(j)}) < 10kp^{1/2} \log p \tag{3.3}$$

(*so that $E_p^{(1)}, \dots, E_p^{(p)}$ are "good" PR binary sequences*) *and*

$$|(E_p^{(i)}, E_p^{(j)})| < 4p^{1/2} \quad \text{for } 1 \le i < j \le p \tag{3.4}$$

(*so that* (3.2) *also holds*), *but*

$$Q_2(\eta) \ge (p-1)^2. \tag{3.5}$$

Proof. Denote the quadratic character of $\mathbb{F}_p$ by χ^*:

$$\chi^*(n) = \begin{cases} \left(\dfrac{n}{p}\right) & \text{for } p \nmid n, \\ 0 & \text{for } p \mid n. \end{cases}$$

Then

$$\begin{aligned} Q_k(E_p^{(j)}) &= \max_{a,t,\underline{D}} \Big| \sum_{i=0}^{t} e_{ia+d_1}^{(j)} \cdots e_{ia+d_k}^{(j)} \Big| \\ &\le \max_{a,t,\underline{D}} \Bigg(\Bigg| \sum_{\substack{0 \le i \le t \\ p \nmid (j+ia+d_1)\cdots(j+ia+d_k)}} \left(\frac{(j+ia+d_1)\cdots(j+ia+d_k)}{p} \right) \Bigg| \\ &\qquad + \sum_{\substack{0 \le i \le t \\ p \mid (j+ia+d_1)\cdots(j+ia+d_k)}} 1 \Bigg) \\ &\le \max_{a,t,\underline{D}} \Big(\Big| \sum_{i=0}^{t} \chi^*((j+ia+d_1)\cdots(j+ia+d_k)) \Big| + k \Big), \end{aligned}$$

whence, by Lemma 1, (3.3) follows.

Moreover, for $1 \le i < j \le p$ we have

$$\begin{aligned}(3.6)\qquad |(E_p^{(i)}, E_p^{(j)})| &= \Big|\sum_{l=1}^{p} e_l^{(i)} e_l^{(j)}\Big| \\ &\le \Big|\sum_{\substack{1\le l\le p\\ p\nmid (l+i)(l+j)}} \left(\frac{(l+i)(l+j)}{p}\right)\Big| + \sum_{\substack{1\le l\le p\\ p\mid (l+i)(l+j)}} 1 \\ &\le \Big|\sum_{l=1}^{p} \chi^*((l+i)(l+j))\Big| + 2.\end{aligned}$$

It follows from Weil's theorem [12] (see also Lemma 2C in [11]) that the last sum is $\le 2p^{1/2}$. Thus (3.4) follows from (3.6).

Finally, it follows from the definition of $Q_k(\eta)$ that

$$\begin{aligned}(3.7)\qquad Q_2(\eta) &\ge \Big|\sum_{j_1=0}^{p-2}\sum_{j_2=1}^{p-1} \eta((j_1,j_2)+(0,0))\eta((j_1,j_2)+(+1,-1))\Big| \\ &= \Big|\sum_{j_1=0}^{p-2}\sum_{j_2=1}^{p-1} \eta((j_1,j_2))\eta((j_1+1,j_2-1))\Big|.\end{aligned}$$

We have

$$\begin{aligned}\eta((j_1,j_2))\eta((j_1+1,j_2-1)) &= \left(\frac{j_1+j_2+2}{p}\right)\left(\frac{j_1+j_2+2}{p}\right) \\ &= +1 \quad \text{for } p\nmid j_1+j_2+2\end{aligned}$$

and

$$\eta((j_1,j_2))\eta((j_1+1,j_2-1)) = (+1)(+1) = +1 \quad \text{for } p\,|\,j_1+j_2+2,$$

so that, from (3.7),

$$Q_2(\eta) \ge \sum_{j_1=0}^{p-2}\sum_{j_2=1}^{p-1} 1 = (p-1)(p-1) = (p-1)^2,$$

which proves (3.5) and completes the proof of Theorem 1.

Remark 1. We note that the construction of Theorem 1 is a special case of a more general construction: Let $E_N^{(1)} = \{e_1^{(1)}, \dots, e_N^{(1)}\} \in \{-1,+1\}^N$ be a truly random binary sequence, and for $2 \le j \le n$ let $E_N^{(j)}$ be a shifted version of $E_N^{(1)}$, so $E_N^{(j)} = \{e_1^{(j)}, \dots, e_N^{(j)}\} = \{e_j^{(1)}, e_{j+1}^{(1)}, \dots, e_N^{(1)}, e_1^{(1)}, e_2^{(1)}, \dots, e_{j-1}^{(1)}\}$. Then the $E_N^{(j)}$'s satisfy inequalities of type (3.3) and (3.4) (with N in place of p and with upper bounds $O(N^{1/2}(\log N)^c)$) with probability 1. Define a lattice η by

$$\eta(x,y) = e_{x+1}^{(y+1)} = e_{r_N(x+y+1)}^{(1)} \qquad (\text{for } (x,y) \in \{0,1,\dots,p-1\}^2)$$

where $r_N(x+y+1)$ denotes the least positive residue of $x+y+1$ modulo N. Similarly to (3.7) we easily get

$$Q_2(\eta) \geq (N-1)^2.$$

In Theorem 1 we presented a special case of the above construction, where $E_N^{(1)}$ was defined by the Legendre symbol, and then indeed (3.3) and (3.4) hold.

4. Trying to reduce the two-dimensional case to the one-dimensional one: the PR measures of order 1. The simplest and more natural way to reduce the two-dimensional case to the one-dimensional one is the following:

To any two-dimensional binary N-lattice

$$\eta(\underline{x}):\ I_N^2 \to \{-1,+1\} \tag{4.1}$$

we may assign a unique binary sequence $E_{N^2} = E_{N^2}(\eta) = (e_1, \dots, e_{N^2}) \in \{-1,+1\}^N$ by taking the first (from the bottom) row of the lattice (4.1), then the second row, etc.; in general, we set

$$e_{iN+j} = \eta((j-1,i)) \quad \text{for } i = 0,1,\dots,N-1,\ j = 1,\dots,N. \tag{4.2}$$

It is natural to ask: is it true that if $E_{N^2}(\eta)$ is a "good" PR binary *sequence* then η is a "good" PR two-dimensional lattice? Then "good" PR binary *sequences* would generate "good" PR-binary lattices automatically, so it would be sufficient to study binary sequences, and there would be no need for developing a theory of pseudorandomness of binary lattices. Unfortunately, the answer to this question is negative; we will show in Sections 4 and 5 that it may occur that the PR measures of the sequence $E_{N^2}(\eta)$ are small, but the corresponding PR-measures of the lattice η are large.

We will denote the PR measures of $E_{N^2}(\eta)$ by W, C_k, Q_k, while we write $\overline{Q}_k$ for the pseudorandom measure of order k of η. First we will compare the PR measures of order 1, i.e., $Q_1 = W$ and $\overline{Q}_1$.

Theorem 2. *For every even number $N = 2R \in \mathbb{N}$ there is a binary lattice η such that $Q_1(E_{N^2}(\eta))$ is "small":*

$$Q_1(E_{N^2}(\eta)) = W(E_{N^2}(\eta)) < 4N, \tag{4.3}$$

but $\overline{Q}_1(\eta)$ is large:

$$\overline{Q}_1(\eta) > \frac{1}{2} N^2. \tag{4.4}$$

Proof. Define an N-lattice of type (4.1) by

$$\eta((i,j)) = \begin{cases} +1 & \text{for } i = 0,1,\dots,R-1 \text{ and } j = 0,1,\dots,N-1, \\ -1 & \text{for } i = R, R+1,\dots,N-1 \text{ and } j = 0,1,\dots,N-1. \end{cases}$$

We will show that η satisfies (4.3) and (4.4).

By the definition of W and Q_1 we have

$$Q_1(E_{N^2}(\eta)) = W(E_{N^2}(\eta)) = \max_{a,b,t} \Big|\sum_{j=0}^{t-1} e_{a+jb}\Big|$$

where the maximum is taken over all $a, b, t \in \mathbb{N}$ with $1 \le a \le a + (t-1)b \le N^2$. Take one of the sums $\sum_{j=0}^{t-1} e_{a+jb}$ considered here. There are unique integers u, v with

$$\begin{aligned} &0 \le u \le v \le N-1,\\ &a \in (uN, uN+N],\\ &a+(t-1)b \in (vN, vN+N]. \end{aligned}$$

Then

$$\begin{aligned} \sum_{j=0}^{t-1} e_{a+jb} = &\sum_{\substack{0\le j\le t-1\\ a+jb\in(uN,(u+1)N]}} e_{a+jb} + \sum_{u<w<v} \sum_{\substack{0\le j\le t\\ a+jb\in(wN,(w+1)N]}} e_{a+jb}\\ &+ \sum_{\substack{0\le j\le t-1\\ a+jb\in(vN,(v+1)N]}} e_{a+jb}. \end{aligned} \tag{4.5}$$

Clearly,

$$\Big|\sum_{\substack{0\le j\le t-1\\ a+jb\in(uN,(u+1)N]}} e_{a+jb}\Big| \le \sum_{a+jb\in(uN,(u+1)N]} 1 \le N, \tag{4.6}$$

$$\Big|\sum_{\substack{0\le j\le t-1\\ a+jb\in(vN,(v+1)N]}} e_{a+jb}\Big| \le \sum_{a+jb\in(vN,(v+1)N]} 1 \le N, \tag{4.7}$$

and, for $u < w < v$, by the definition of η and E_{N^2},

$$\begin{aligned} \Big|\sum_{j:\, a+jb\in(wN,(w+1)N]} e_{a+jb}\Big| &= \Big|\sum_{j:\, a+jb\in(wN,wN+R]} \eta((a+jb-wN-1,w))\\ &\quad + \sum_{j:\, a+jb\in(wN+R,(w+1)N]} \eta((a+jb-wN-1,w))\Big|\\ &= \Big|\sum_{j:\, a+jb\in(wN,wN+R]} 1 - \sum_{j:\, a+jb\in(wN+R,(w+1)N]} 1\Big|\\ &= \big|\,|\{m : m\equiv a \pmod b,\ wN < m \le wN+R\}|\\ &\quad - |\{m : m\equiv a \pmod b,\ wN+R < m \le (w+1)N\}|\,\big|\\ &= |(|\{m : m\equiv a \pmod b,\ wN < m \le wN+R\}| - R/b)\\ &\quad - (|\{m : m\equiv a \pmod b,\ wN+R < m \le (w+1)N\}| - R/b)|\\ &\le 1+1=2. \end{aligned} \tag{4.8}$$

It follows from (4.5)–(4.8) that

$$\Big|\sum_{j=0}^{t-1} e_{a+jb}\Big| \le N + 2(v-u-1) + N < 4N,$$

which proves (4.3).

On the other hand, we have

$$\overline{Q}_1(\eta) \ge \Big|\sum_{j_1=0}^{R-1}\sum_{j_2=0}^{N-1} \eta((j_1,j_2))\Big| = \sum_{j_1=0}^{R-1}\sum_{j_2=0}^{N-1} 1 = RN = \frac{1}{2}\,N^2,$$

which proves (4.4).

Remark 2. It is easy to see that in the example above we have

$$Q_2(E_{N^2}(\eta)) \ge C_2(E_{N^2}(\eta)) \gg N^2.$$

One might like to give a construction where we also have $Q_2(E_{N^2}(\eta)) = o(N^2)$ or at least $C_2(E_{N^2}(\eta)) = o(N^2)$. We have not be able to give such a construction. So we arrive at the following natural question:

Problem 1. *Is it true that*

$$C_2(E_{N^2}(\eta)) = o(N^2) \;\Rightarrow\; \overline{Q}_1(\eta) = o(N^2)?$$

5. Trying to reduce the two-dimensional case to the one-dimensional case: the PR measures of order 2. One might wish to save the above idea on reducing the two-dimensional case to the one-dimensional one by also considering the PR measures of order 2. So one may ask: is it true that if $W(E_{N^2}(\eta))$ and $C_2(E_{N^2}(\eta))$ are small, then η must be a "good" PR binary lattice? Again, the answer is negative:

Theorem 3. *For every even number $N = 2R \in \mathbb{N}$ there is a binary lattice η such that $Q_1(E_{N^2}(\eta))$ and $C_2(E_{N^2}(\eta))$ are small*:

$$Q_1(E_{N^2}(\eta)) = W(E_{N^2}(\eta)) < 6N(\log N)^{1/2} \tag{5.1}$$

and

$$C_2(E_{N^2}(\eta)) < 12N(\log N)^{1/2}, \tag{5.2}$$

but $\overline{Q}_2(\eta)$ is large:

$$\overline{Q}_2(\eta) \ge \frac{1}{4}\,N^2. \tag{5.3}$$

Proof. We will present a probabilistic construction, more precisely we will consider all the binary N-lattices η satisfying certain conditions and chosen with equal probability, and then we will show that for $\varepsilon > 0$ and $N > N_0(\varepsilon)$, such a lattice η satisfies (5.1), resp. (5.2) with probability greater than $1-\varepsilon$, and all these lattices η also satisfy (5.3).

Define an N-lattice η so that

(i) for $0 \le x \le N-1$, $0 \le y \le R-1$ the numbers $\eta(x,y)$ are independent random variables with distribution

$$P(\eta(x,y)=+1) = P(\eta(x,y)=-1) = 1/2, \tag{5.4}$$

(ii) $\eta(x,y) = -\eta(x, y-R)$ for $R \le x \le N-1$, $R \le y \le N-1$,
(iii) $\eta(x,y) = \eta(x, y-R)$ for $0 \le x \le R-1$, $R \le y \le N-1$.

The structure of this binary lattice η is the following:

Y	$-Z$
Y	Z

Then defining the binary sequence $E_{N^2} = E_{N^2}(\eta) = (e_1, \dots, e_{N^2})$ by (4.2), it is easy to check that $e_1, \dots, e_{N^2}$ have the following properties:

(P1) For $n = 1, \dots, N^2$ the number e_n is a random variable with distribution

$$P(e_n = +1) = P(e_n = -1) = 1/2.$$

(P2) If $1 \le n < n+d \le N^2$ and $d \ne RN$, then the random variables e_n and e_{n+d} are independent.
(P3) If $1 \le n < n+d \le N^2$, $d = RN$, and we write n in the form $iN+j$ with $i \in \{0,1,\dots,R-1\}$, $j \in \{1,\dots,N\}$, then

$$e_{n+d} = \begin{cases} e_n & \text{for } 1 \le j \le R, \\ -e_n & \text{for } R < j \le N \ (=2R). \end{cases}$$

We will denote the mean value and standard deviation of the random variable ξ by $M(\xi)$ and $D(\xi)$, respectively. We will need Bernstein's inequality [9, Ch. 7]:

Lemma 2. *If $\xi_1, \dots, \xi_m$ are independent random variables with $M(\xi_k) = M_k$, $D(\xi_k) = D_k$ and $|\xi_k - M_k| \le K$ for $k = 1, \dots, m$, then, writing $\xi = \xi_1 + \dots + \xi_m$, $M = M_1 + \dots + M_m$ and $D = (D_1^2 + \dots + D_m^2)^{1/2}$, for any positive number $\mu \le D/K$ we have*

$$P(|\xi - M| \ge \mu D) \le 2 \exp\left(-\frac{\mu^2}{2(1 + \mu K/2D)^2} \right).$$

To estimate $W(E_{N^2}(\eta))$, fix positive integers a, b, t with $1 \le a \le a + (t-1)b \le N^2$, and consider the sum

$$S(a,b,t) = \sum_{j=0}^{t-1} e_{a+jb}.$$

Denote by t^* the largest integer for which

$$a + (t^* - 1)b < N^2/2.$$

Let

$$S_1(a,b,t)=\sum_{j=0}^{t^*-1} e_{a+jb}, \qquad S_2(a,b,t)=\sum_{j=t^*}^{t-1} e_{a+jb}.$$

Then

$$S(a,b,t)=S_1(a,b,t)+S_2(a,b,t).$$

By properties (P1) and (P2) we may use Lemma 2 with $e_{a+(k-1)b}$ in place of ξ_k for $k=1,\dots,t^*$ and for $k=t^*+1,\dots,t$, so that now $M_k=0$, $D_k=1/2$, $K=1/2$, $M=0$ and in the first case $D=\frac{1}{2}\,t^{*1/2}$ and in the latter case $D=\frac{1}{2}(t-t^*)^{1/2}$. Then using Lemma 2 with $\mu=12(\log N)^{1/2}$ we easily get

$$P(|S_1(a,b,t)|>6N(\log N)^{1/2})<\frac{1}{2N^8},$$
$$P(|S_2(a,b,t)|>6N(\log N)^{1/2})<\frac{1}{2N^8},$$

uniformly in a,b,t for $N>N_0$. From this and the triangle inequality we get

$$P(|S(a,b,t)|>12N(\log N)^{1/2})\le P(|S_1(a,b,t)|>6N(\log N)^{1/2}) + P(|S_2(a,b,t)|>6N(\log N)^{1/2})\le 1/N^8.$$

Thus we have

$$(5.5)\qquad P(W(E_{N^2})>12N(\log N)^{1/2})=P(\max_{a,b,t}|S(a,b,t)|>12N(\log N)^{1/2}) \le\sum_{a,b,t}P(|S(a,b,t)|>12N(\log N)^{1/2})\le\sum_{1\le a,b,t\le N^2}\frac{1}{N^8}=\frac{1}{N^2}.$$

Now we will estimate

$$(5.6)\qquad C_2(E_{N^2}(\eta))=\max_{L,d_1,d_2}\left|\sum_{n=1}^{L}e_{n+d_1}e_{n+d_2}\right|=\max_{U,V,d}\left|\sum_{n=U}^{V}e_ne_{n+d}\right|$$

where the maximum is taken over all U,V,d with $1\le U\le V<V+d\le N^2$. Consider one of these sums $\sum_{n=U}^{V}e_ne_{n+d}$. We have to distinguish two cases.

Case 1. Assume first that

$$(5.7)\qquad d\ne RN.$$

Define

$$\mathcal{A}_1=\{U,U+1,\dots,V\}\cap\bigcap_{k=0}^{+\infty}\{2kd+1,2kd+2,\dots,(2k+1)d\},$$
$$\mathcal{A}_2=\{U,U+1,\dots,V\}\cap\bigcap_{k=0}^{+\infty}\{(2k+1)d+1,(2k+1)d+2,\dots,(2k+2)d\},$$

so that

$$(5.8)\qquad \Big|\sum_{n=U}^{V} e_n e_{n+d}\Big| = \Big|\sum_{n\in\mathcal{A}_1} e_n e_{n+d} + \sum_{n\in\mathcal{A}_2} e_n e_{n+d}\Big| \le \Big|\sum_{n\in\mathcal{A}_1} e_n e_{n+d}\Big| + \Big|\sum_{n\in\mathcal{A}_2} e_n e_{n+d}\Big| = \Big|\sum\nolimits_1\Big| + \Big|\sum\nolimits_2\Big|.$$

It follows from (P1), (P2) and (5.7) that the terms of $\sum_1$ are independent random variables of distribution

$$P(e_n e_{n+d} = +1) = P(e_n e_{n+d} = -1) = 1/2 \qquad (\text{for } n \in \mathcal{A}_1).$$

Thus the terms of $\sum_1$ can be estimated by using Lemma 2 (in the same way as in the estimate of $W(E_{N^2})$). We obtain

$$(5.9)\qquad P\Big(\Big|\sum\nolimits_1\Big| > 6N(\log N)^{1/2}\Big) < \frac{1}{2N^8}$$

for $N > N_0$. In the same way we get

$$(5.10)\qquad P\Big(\Big|\sum\nolimits_2\Big| > 6N(\log N)^{1/2}\Big) < \frac{1}{2N^8}.$$

It follows from (5.8)–(5.10) that for all U, V and d (satisfying (5.7)) we have

$$P\Big(\Big|\sum_{n=U}^{V} e_n e_{n+d}\Big| > 12N(\log N)^{1/2}\Big) \le P\Big(\Big|\sum\nolimits_1\Big| > 6N(\log N)^{1/2}\Big) + P\Big(\Big|\sum\nolimits_2\Big| > 6N(\log N)^{1/2}\Big) < \frac{1}{2N^8} + \frac{1}{2N^8} = \frac{1}{N^8},$$

whence

$$(5.11)\qquad P\Big(\max_{U,V,d\neq RN}\Big|\sum_{n=U}^{V} e_n e_{n+d}\Big| > 12N(\log N)^{1/2}\Big) \le \sum_{U,V,d\neq RN} P\Big(\Big|\sum_{n=U}^{V} e_n e_{n+d}\Big| > 12N(\log N)^{1/2}\Big) < \sum_{U,V,d\neq RN}\frac{1}{N^8} \le (N^2)^3\,\frac{1}{N^8} = \frac{1}{N^2} \qquad (\text{for } N > N_0).$$

Case 2. Assume that

$$(5.12)\qquad d = RN.$$

Let K_1 and K_2 denote respectively the smallest and greatest integer K with

$$(KN, (K+1)N] \cap [U, V] \neq 0.$$

Then by (P3) and (5.12) we have

$$(5.13)\qquad \Big|\sum_{n=U}^{V} e_n e_{n+d}\Big|$$
$$= \Big|\sum_{n=U}^{(K_1+1)N} e_n e_{n+d} + \sum_{K=K_1+1}^{K_2-1}\sum_{n=KN+1}^{(K+1)N} e_n e_{n+d} + \sum_{n=K_2N+1}^{V} e_n e_{n+d}\Big|$$
$$\le \Big|\sum_{n=U}^{(K_1+1)N} 1\Big| + \sum_{K=K_1+1}^{K_2-1}\Big|\sum_{n=KN+1}^{KN+R} e_n e_{n+d} + \sum_{n=KN+R+1}^{(K+1)N} e_n e_{n+d}\Big| + \Big|\sum_{n=K_2N+1}^{V} 1\Big|$$
$$\le N + \sum_{K=K_1+1}^{K_2-1}\Big|\sum_{n=KN+1}^{KN+R} 1 + \sum_{n=KN+R+1}^{(K+1)N} (-1)\Big| + N$$
$$= 2N \qquad (\text{for } d = RN).$$

Finally, by (ii) we have

$$(5.14)\qquad \overline{Q}_2(\eta) \ge \Big|\sum_{j_1=0}^{R-1}\sum_{j_2=0}^{R-1} \eta((j_1,j_2)+(0,0))\eta((j_1,j_2)+(0,R))\Big|$$
$$= \Big|\sum_{j_1=0}^{R-1}\sum_{j_2=0}^{R-1} \eta((j_1,j_2))^2\Big| = \sum_{j_1=0}^{R-1}\sum_{j_2=0}^{R-1} 1 = R^2 = \frac{1}{4}N^2.$$

By (5.5), (5.11) and (5.13), for $N \ge N_0(\varepsilon)$ both (5.1) and (5.2) hold with probability greater than $1-\varepsilon$, and by (5.14), for all lattices η considered, (5.3) also holds; this completes the proof of Theorem 3.

Remark 3. In Theorem 3 we could have replaced $C_2(E_{N^2})$ by $Q_2(E_{N^2})$ but this would have made the argument lengthier, so we preferred to present this simpler version. It is easy to see that in the construction of Theorem 3 we have

$$Q_4(E_{N^2}(\eta)) \ge C_4(E_{N^2}(\eta)) \gg N^2.$$

Thus one might ask the following question:

Problem 2. *Is it true that* $Q_4(E_{N^2}(\eta)) = o(N^2)$ *implies* $\overline{Q}_2(\eta) = o(N^2)$*?*

Remark 4. Theorem 3 could be extended from $C_2(E_{N^2})$ to $C_k(E_{N^2})$ (and beyond, to $Q_k(E_{N^2})$) by using the following generalization of our construction: Let $N = 2kR$ where $k, R \in \mathbb{N}$. Define an N-lattice η so that

(i) for $0 \le x \le N-1$, $0 \le y \le (2k-2)R-1$ the numbers $\eta(x,y)$ are independent random variables with distribution

$$P(\eta(x,y) = +1) = P(\eta(x,y) = -1) = 1/2,$$

(ii) $\eta(x,y) = \prod_{i=1}^{k-1} \eta((x, y-2iR))$ for $0 \le x \le kR-1$, $(2k-2)R \le y \le N-1$,

(iii) $\eta(x,y) = -\prod_{i=1}^{k-1} \eta((x, y-2iR))$ for $kR \le x \le N-1$, $(2k-2)R \le y \le N-1$.

The structure of this lattice η is the following:

	$\overbrace{}^{kR}$	$\overbrace{}^{kR}$
$2R\{$	$\prod Y_i$	$-\prod Z_i$
$2R\{$	Y_{k-1}	Z_{k-1}
$\vdots$	$\vdots$	$\vdots$
$2R\{$	Y_3	Z_3
$2R\{$	Y_2	Z_2
$2R\{$	Y_1	Z_1

(Here $\prod Y_i$ means that the jth element in the lth row of this $2R \times kR$ matrix is the product of the corresponding elements of the matrices $Y_1, \dots, Y_{k-1}$; the meaning of $\prod Z_i$ is similar.)

It is easy to see that in this construction we have

$$Q_{2k}(E_{N^2}(\eta)) \ge C_{2k}(E_{N^2}(\eta)) \gg N^2.$$

This motivates the following question:

PROBLEM 3. *Is it true that if $Q_{2k}(E_{N^2}(\eta)) = o(N^2)$ for some fixed $k > 1$, then $\overline{Q}_k(\eta) = o(N^2)$?*

By Theorem 3 it may occur that $C_2(E_{N^2})$ is small but $\overline{Q}_2(\eta)$ is large. The opposite cannot occur:

THEOREM 4. *For every binary N-lattice η and $k \in \mathbb{N}$ we have*

$$Q_k(E_{N^2}(\eta)) \le 3N(\overline{Q}_k(\eta))^{1/2}.$$

Note that, as shown in [3], for a truly random two-dimensional N-lattice η the order of magnitude of $\overline{Q}_k(\eta)$ is N, so that the right hand side is $O(N^{3/2})$. Thus in general this theorem gives the nontrivial bound $O(N^{3/2})$ for $Q_k(E_{N^2}(\eta))$.

Proof. By the definition of $Q_k(E_{N^2}(\eta))$ there exist a, t and $D = (d_1, \dots, d_k)$ with $0 < d_1 < \dots < d_k$ such that

$$Q_k(E_{N^2}(\eta)) = \Big|\sum_{j=0}^{t} e_{ja+d_1} \cdots e_{ja+d_k}\Big|, \tag{5.15}$$

where all subscripts $ja+d_l$ belong to $\{1,\dots,N^2\}$. We split $\{1,\dots,N^2\}$ into several subsets. For $0\le i\le N-1$ the $(i+1)$-st subset is

$$I_i=\{iN+1,iN+2,\dots,(i+1)N\}.$$

For $0\le j\le t$ the minimum value of $ja+d_1$ is d_1. Write d_1 in the form

$$d_1=y_{\min}N+x_1\quad\text{ where } 0\le x_1\le N-1.$$

For $0\le j\le t$ the maximum value of $ja+d_1$ is $ta+d_1$. Write $ta+d_1$ in the form

$$ta+d_1=y_{\max}N+x_2\quad\text{ where } 0\le x_2\le N-1.$$

Then

$$\begin{aligned}
(5.16)\qquad Q_k(E_{N^2}(\eta)) &= \Big|\sum_{i=y_{\min}}^{y_{\max}}\sum_{\substack{0\le j\le t\\ ja+d_1\in I_i}} e_{ja+d_1}\cdots e_{ja+d_k}\Big|\\
&\le \Big|\sum_{\substack{0\le j\le t\\ ja+d_1\in I_{y_{\min}}\cup I_{y_{\max}}}} e_{ja+d_1}\cdots e_{ja+d_k}\Big|\\
&\quad+\Big|\sum_{i=y_{\min}+1}^{y_{\max}-1}\sum_{\substack{0\le j\le t\\ ja+d_1\in I_i}} e_{ja+d_1}\cdots e_{ja+d_k}\Big|\\
&\le 2N+\Big|\sum_{i=y_{\min}+1}^{y_{\max}-1}\sum_{\substack{0\le j\le t\\ ja+d_1\in I_i}} e_{ja+d_1}\cdots e_{ja+d_k}\Big|\\
&= 2N+\Big|\sum_{l=0}^{a-1}\sum_{\substack{i=y_{\min}+1\\ i\equiv l\ (\mathrm{mod}\ a)}}^{y_{\max}-1}\sum_{\substack{0\le j\le t\\ ja+d_1\in I_i}} e_{ja+d_1}\cdots e_{ja+d_k}\Big|\\
&\le 2N+\sum_{l=0}^{a-1}\Big|\sum_{\substack{i=y_{\min}+1\\ i\equiv l\ (\mathrm{mod}\ a)}}^{y_{\max}-1}\sum_{\substack{0\le j\le t\\ ja+d_1\in I_i}} e_{ja+d_1}\cdots e_{ja+d_k}\Big|.
\end{aligned}$$

It is easy to check that if $0\le l<a$ and

$$\begin{aligned}
(5.17)\qquad \{e_{ja+d_1}: ja+d_1\in I_l,\ j\in\mathbb{N}\}&\\
&=\{\eta(x_l,l),\eta(x_l+a,l),\dots,\eta(x_l+t_la,l)\},
\end{aligned}$$

then for $i\equiv l \pmod a$ we have

$$\{e_{ja+d_1}: ja+d_1\in I_i,\ j\in\mathbb{N}\}=\{\eta(x_l,i),\ \eta(x_l+a,i),\dots,\eta(x_l+t_la,i)\}.$$

In (5.16), j assumes values from the interval $[0,t]$. By the definition of $y_{\min}$ and $y_{\max}$, and (5.17), if $i\equiv l \pmod a$ and $y_{\min}+1\le i\le y_{\max}-1$, then

$$\{e_{ja+d_1}: ja+d_1\in I_i,\ 0\le j\le t\}=\{\eta(x_l,i),\ \eta(x_l+a,i),\dots,\eta(x_l+t_la,i)\}.$$

Write $d_i - d_1$ in the form

$$d_i - d_1 = d_{i,1}N + d_{i,2} \quad \text{with } 0 \le d_{i,2} \le N-1$$

and define

$$\underline{d}'_{i-1} = (d_{i,1}, d_{i,2}).$$

Then for $i \equiv l \pmod a$ and $y_{\min} + 1 \le i \le y_{\max} - 1$ we have

$$\sum_{\substack{0 \le j \le t \\ ja+d_1 \in I_i}} e_{ja+d_1} \cdots e_{ja+d_k} = \sum_{j=0}^{t_l} \eta((x_l + ja, i))\eta((x_l + ja, i) + \underline{d}'_1) \cdots \eta((x_l + ja, i) + \underline{d}'_{k-1}).$$

Let

$$\{i : i \equiv l \ (\mathrm{mod}\ a),\ y_{\min} + 1 \le i \le y_{\max} - 1\} = \{y_l, y_l + a, \dots, y_l + s_l a\}.$$

Then

$$\sum_{\substack{i=y_{\min}+1 \\ i \equiv l \,(\mathrm{mod}\, a)}}^{y_{\max}-1} \sum_{\substack{0 \le j \le t \\ ja+d_1 \in I_i}} e_{ja+d_1} \cdots e_{ja+d_k} = \sum_{i=0}^{s_l} \sum_{j=0}^{t_l} \eta((x_l + ja, y_l + ia))\eta((x_l + ja, y_l + ia) + \underline{d}'_1) \cdots \eta((x_l + ja, y_l + ia) + \underline{d}'_{k-1}).$$

By the definition of $\overline{Q}_k(\eta)$ we have

$$\left| \sum_{\substack{i=y_{\min}+1 \\ i \equiv l \,(\mathrm{mod}\, a)}}^{y_{\max}-1} \sum_{\substack{0 \le j \le t \\ ja+d_1 \in I_i}} e_{ja+d_1} \cdots e_{ja+d_k} \right| = \left| \sum_{i=0}^{s_l} \sum_{j=0}^{t_l} \eta((x_l + ja, y_l + ia))\eta((x_l + ja, y_l + ia) + \underline{d}'_1) \cdots \eta((x_l + ja, y_l + ia) + \underline{d}'_{k-1}) \right| \le \overline{Q}_k(\eta). \tag{5.18}$$

Using (5.16) and (5.18) we get

$$Q_k(E_{N^2}(\eta)) \le 2N + a\overline{Q}_k(\eta). \tag{5.19}$$

On the other hand, the number of terms in (5.15) is $t + 1 \le 2t \le 2N^2/a$, thus $Q_k(E_{N^2}(\eta)) \le 2N^2/a$. Therefore

$$a \le \frac{2N^2}{Q_k(E_{N^2}(\eta))}.$$

Using this and (5.19) yields

$$Q_k(E_{N^2}(\eta)) \le 2N + \frac{2N^2\overline{Q}_k(\eta)}{Q_k(E_{N^2}(\eta))},$$

$$Q_k(E_{N^2}(\eta))^2 \le 2NQ_k(E_{N^2}(\eta)) + 2N^2\overline{Q}_k(\eta),$$

$$\begin{aligned}(Q_k(E_{N^2}(\eta)) - N)^2 &\le N^2 + 2N^2\overline{Q}_k(\eta),\\ Q_k(E_{N^2}(\eta)) &\le N + N(1 + 2\overline{Q}_k(\eta))^{1/2},\\ Q_k(E_{N^2}(\eta)) &\le 3N(\overline{Q}_k(\eta))^{1/2}\end{aligned}$$

which was to be proved.

References

[1] N. Alon, Y. Kohayakawa, C. Mauduit, C. G. Moreira and V. Rödl, *Measures of pseudorandomness for finite sequences: typical values*, Proc. London Math. Soc. (3) 95 (2007), 778–812.

[2] J. Cassaigne, C. Mauduit and A. Sárközy, *On finite pseudorandom binary sequences VII: The measures of pseudorandomness*, Acta Arith. 103 (2002), 97–118.

[3] P. Hubert, C. Mauduit and A. Sárközy, *On pseudorandom binary lattices*, ibid. 125 (2006), 51–62.

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

[5] —, —, *On large families of pseudorandom binary lattices*, J. Uniform Distribution Theory 2 (2007), 23–37.

[6] —, —, *Construction of pseudorandom binary lattices by using the multiplicative inverse*, Monatsh. Math. 153 (2008), 217–231.

[7] —, —, *On finite pseudorandom sequences of k symbols*, Indag. Math. 13 (2002), 89–101.

[8] A. Menezes, P. van Oorschot and S. Vanstone, *Handbook of Applied Cryptography*, CRC Press, 1996.

[9] A. Rényi, *Probability Theory*, North-Holland Ser. Appl. Math. Mech. 10, North-Holland, Amsterdam, 1970.

[10] A. Sárközy, *On finite pseudorandom binary sequences and their applications in cryptography*, Tatra Mt. Math. Publ. 37 (2007), 123–136.

[11] W. M. Schmidt, *Equations over Finite Fields. An Elementary Approach*, Lecture Notes in Math. 536, Springer, New York, 1976.

[12] A. Weil, *Sur les courbes algébriques et les variétés qui s'en déduisent*, Act. Sci. Ind. 1041, Hermann, Paris, 1948.

Alfréd Rényi Institute of Mathematics
Reáltanoda utca 13-15
H-1053 Budapest, Hungary
E-mail: gykati@cs.elte.hu

Institut de Mathématiques de Luminy
CNRS, UMR 6206
163, avenue de Luminy, Case 907
F-13288 Marseille Cedex 9, France
E-mail: mauduit@iml.univ-mrs.fr

Department of Algebra and Number Theory
Eötvös Loránd University
Pázmány Péter sétány 1/c
H-1117 Budapest, Hungary
E-mail: sarkozy@cs.elte.hu

Received on 25.1.2008 (5623)