Remarks on pseudorandom binary sequences over elliptic curves

László Mérai*

Abstract

In the paper the pseudorandomness of binary sequences defined over elliptic curves is studied and both the well-distribution and correlation measures are estimated. The paper is based on the Kohel-Shparlinski bound and the Erdős-Turán-Koksma inequality.

2000 Mathematics Subject Classification: Primary 11K45
Key words and phrases: pseudorandom, binary sequence, elliptic curve, character.

1 Introduction

Let \( p > 3 \) be a prime number and let us denote by \( \mathbb{F}_p \) the field of \( p \) elements which we represent by \( \{0, 1, \ldots, p-1\} \). Let \( \mathcal{E} \) be an elliptic curve over \( \mathbb{F}_p \) defined by

\[
y^2 = x^3 + Ax + B
\]

with non-zero discriminant (see [21]). The \( \mathbb{F}_p \)-rational points \( \mathcal{E}(\mathbb{F}_p) \) of \( \mathcal{E} \) form an Abelian group (with respect to the usual addition, see [21]) with the point in infinity \( O \) as the neutral element, where the group operation is denoted by \( \oplus \).

The elliptic curve congruential generator is defined by the relation

\[
P_n = G \oplus P_{n-1} = nG \oplus P_0
\]

---

*Research partially supported by Hungarian National Foundation for Scientific Research, Grant No. K67676 and by the Momentum fund of the Hungarian Academy of Sciences.
with the initial value $P_0$. Clearly (1) defined a cryptographically weak sequence, thus one may define a sequence by

$$s_n = f(P_n) = f(nG \oplus P_0)$$

with rational function $f \in \mathbb{F}_p(E)$ (see [2]).

The pseudorandom properties of sequences defined by (2) have already been studied in several aspect (see [2], [7], [9], [10], [11], [12], [14], [20], [22] and the reference therein.)

Here we study the pseudorandom properties of the binary sequences derived from (2).

First we prove an exponential sum estimate (which is based on the Kohel-Shparlinski bound) which implies a bound to the discrepancy of the sequence related to (2) using the Erdős-Turán-Koksma inequality. Then we derive bounds on the well-distribution measure $W(E_T)$ and the correlation measure $C_\ell(E_T)$ of order $\ell$ (see Section 3 below for the definition of the measures) of the binary sequence $E_T = \{e_1, e_2, \ldots, e_T\}$ defined by

$$e_n = \begin{cases} 
+1 & \text{if } f(nG) \in \{0, 1, \ldots, \frac{p-1}{2}\}, \\
-1 & \text{otherwise}.
\end{cases}$$

(3)

It was shown in [1] that for a “truly” random binary sequence $E_N \in \{+1, -1\}^N$ both pseudorandom measures $W(E_N)$ and $C_\ell(E_N)$ are “small” in terms of $N$ more precisely the sequence $E_N$ said to have “good” pseudorandom properties if both these measures $W(E_N)$ and $C_\ell(E_N)$ are greater than $N^{1/2}$ only by at most a power of $\log N$. In this paper we show that if $f$ satisfies certain conditions then the binary sequences $E_T$ defined by (3) has good pseudorandom properties.

Throughout the paper we use the notation $f(x) \ll g(x)$ or equivalently $f(x) = O(g(x))$ if $|f(x)| \leq cg(x)$ holds for some constant $c$ and for all $x$. For a real $\alpha \in \mathbb{R}$ we write $e(\alpha) = \exp(2\pi i \alpha)$ and $e_p(n) = e(n/p)$.

2 A character sum estimate

Let $\mathbb{F}_p(E)$ be the function field of $E$ over $\mathbb{F}_p$. For a rational function $f \in \mathbb{F}_p(E)$ and point $R \in E(\mathbb{F}_p)$, $R$ is a zero (resp. a pole) of $f$ if $f(R) = 0$ (resp.
Any function of \( \mathbb{F}_p(\mathcal{E}) \) has finitely many zeros and poles. The divisor of \( f \) is defined by

\[
\text{Div}(f) = \sum_{R \in \mathcal{E}(\mathbb{F}_p)} \text{ord}_R(f)[R],
\]

where \( \text{ord}_R(f) \) is the order of \( f \) in \( R \).

The degree of the pole divisor of \( f \) is denoted by \( \deg f \), for example \( \deg f = 2 \) if \( f(x, y) = x \) and \( \deg f = 3 \) if \( f(x, y) = y \).

The translation map by \( W \in \mathcal{E}(\mathbb{F}_p) \) on \( \mathcal{E}(\mathbb{F}_p) \) is defined by

\[
\tau_W : \mathcal{E}(\mathbb{F}_p) \to \mathcal{E}(\mathbb{F}_p), \quad P \mapsto P \oplus W.
\]

First we prove an estimate to exponential sums which is based on the Kohel-Shparlinski bound [13] (the incomplete version of the bound is studied in [5]):

**Lemma 1.** Let \( G \in \mathcal{E}(\mathbb{F}_p) \) be a point of order \( T \) and \( f \in \mathbb{F}_p(\mathcal{E}) \) be a non-constant rational function. Then for any fixed \( a, b, t \in \mathbb{N} \) with \( 1 \leq a \leq a + (t - 1)b < T \) the following holds:

\[
\left| \sum_{i=0}^{t-1} e_p(f((a + bi)G)) \right| \ll \deg f p^{1/2} \log T.
\]

**Theorem 1.** Let \( p > 3 \) be a prime number, \( G \in \mathcal{E}(\mathbb{F}_p) \) with order \( T \), and \( f \in \mathbb{F}_p(\mathcal{E}) \) be a non-constant rational function. Let \( p(T) \) be the least prime factor of \( T \). If one of the following condition holds:

1. \( \deg f < p(T) \) and \( \ell = 2 \);
2. \( \deg f < p(T) \) and \( (4 \deg f)^\ell < p(T) \)

then

\[
\sum_{\substack{1 \leq n \leq M \\ f(\xi \neq \infty)}} e_p(h_1f((n + d_1)G) + \cdots + h_\ell f((n + d_\ell)G)) \ll \ell \deg f p^{1/2} \log T
\]

uniformly in all integer vectors \((h_1, \ldots, h_\ell) \in \{0, 1, 2, \ldots, p - 1\}^\ell \setminus \{0\}\), and all distinct integers \( d_1, d_2, \ldots, d_\ell \in \{0, 1, \ldots, T - 1\} \).
We need the following lemma.

**Lemma 2.** Assume that $m, k, \ell \in \mathbb{N}$ and $k, \ell < p(m)$. Assume also that one of the following conditions holds:

1. $\ell \leq 2$,
2. $(4k)^\ell < p(m)$.

Then for all $A, B \subset \mathbb{Z}_m$ with $|A| = k$, $|B| = \ell$, there is a $c \in \mathbb{Z}_m$ so that the equation

$$a + b = c, \quad a \in A, \quad b \in B$$

has exactly one solution in $a, b$.

**Proof of Lemma 2.** This is a part of Theorem 3 in [19]. The original version of the lemma states that the number of solutions of (4) is not divisible by a given $d$ (where $A$ and $B$ are multisets), however if $d = 2$ (and $A$ and $B$ are only simple sets) then the theorem is proved in this sharper form.

**Proof of Theorem 1.** Let $h_{i_1} \leq \cdots \leq h_{i_r}$ denote the non-zero $h_i$’s. Then

$$\sum_{\substack{1 \leq n \leq M \\
f((n+d_i)G) \neq \infty \\
i=1,\ldots,\ell}} e_p(h_1 f((n+d_1)G) + \cdots + h_\ell f((n+d_\ell)G))$$

$$= \sum_{\substack{1 \leq n \leq M \\
f((n+d_{i_j})G) \neq \infty \\
j=1,\ldots,r}} e_p(h_{i_1} f((n+d_{i_1})G) + \cdots + h_{i_r} f((n+d_{i_r})G)) + O(\ell \deg f).$$

(5)

Write $F_{h_1,\ldots,h_\ell} = h_{i_1} f \circ \tau_{d_{i_1}G} + \cdots + h_{i_r} f \circ \tau_{d_{i_r}G}$. We show that $F_{h_1,\ldots,h_\ell}$ has at least one pole and therefore it cannot be a constant function.

We use the approach developed in [8] (see also [19]). Let $\mathcal{R}$ be a co-set of the group generated by $G$ in $E(\mathbb{F}_p)$. Note that $\mathcal{R}$ has the form

$$\mathcal{R} = \{S \oplus aG \mid a = 1, \ldots, T\}$$

with arbitrary $S \in \mathcal{R}$.

For a given co-set $\mathcal{R}$, which contains at least one pole of $f$, let $A$ be the set of the $a$’s such that $S \oplus aG$ is a pole of $f$ and let $B$ be the set of the
Clearly, \( S \oplus aG \oplus (-d_{ij})G \) (\( a \in A, -d_{ij} \in B \)) are poles of \( F_{h_1,\ldots,h_\ell} \), and no other pole belongs to \( R \). On the other hand

\[ |A| \leq \deg f \quad \text{and} \quad |B| = r \leq \ell. \]

By Lemma 2 there is a \( c \) so that it has exactly one representation in form (4), i.e., in form

\[ a - d_{ij} = c, \quad a \in A, -d_{ij} \in B. \]

Then, clearly for these \( a \) and \( d_{ij} \), \( S \oplus aG \oplus (-d_{ij})G \) is a pole of \( h_{ij}f \circ \tau_{d_{ij}}G \) but not of \( h_{ij}f \circ \tau_{d_{ij}}G \) for \( j \neq l \), thus it is a pole of \( F_{h_1,\ldots,h_\ell} \).

Finally the result follows from Lemma 1.

Using the exponential sum estimate proved in Theorem 1, the Erdős-Turán-Koksma inequality gives a non-trivial bound on the discrepancy of certain sequence. More precisely, for a given sequence \( \Gamma \) of \( N \) points

\[ \Gamma = \{(\gamma_{n,1},\ldots,\gamma_{n,r}) : n = 1,\ldots,N\} \quad (6) \]

in the \( r \)-dimensional unite cube \([0,1)^r\) the discrepancy \( \Delta(\Gamma) \) of \( \Gamma \) is defined by

\[ \Delta(\Gamma) \overset{\text{def}}{=} \sup_I \left| \frac{A_\Gamma(I)}{N} - |I| \right| \]

where \( I = \prod_{i=1}^r [u_i, v_i] \) is a subinterval of \([0,1)^r\) and \( A_\Gamma(I) \) is the number of points of \( \Gamma \) belongs to \( I \).

We use the Erdős-Turán-Koksma inequality in the following form (see [6, Theorem 1.21]):

**Lemma 3.** For any integer \( H > 1 \) and any sequence \( \Gamma \) of \( N \) points the discrepancy \( \Delta(\Gamma) \) satisfies

\[ \Delta(\Gamma) \leq \left( \frac{3}{2} \right)^r \left( \frac{2}{H+1} + \frac{1}{N} \sum_{h} \prod_{j=1}^{r} \frac{1}{\max\{|h_j|,1\}} \left| \sum_{n=1}^{N} e \left( \sum_{l=1}^{r} h_l \gamma_{n,l} \right) \right| \right) \]

where the outer sum is taken over all vectors \( h = (h_1,\ldots,h_r) \in \mathbb{Z}^r \setminus \{0\} \) such that \( |h_j| \leq H \) for all \( j = 1,\ldots,r \).
Theorem 1 and the Erdős-Turán-Koksma bound implies a bound of the discrepancy to the sequence of points

$$\Gamma(d, N, \ell) = \left\{ \left( \frac{f((n + d_1)G)}{p}, \ldots, \frac{f((n + d_\ell)G)}{p} \right) : n = 1, \ldots, N \right\}$$ (7)

(where we replace $f((n + d_i)G)/p$ by 0 if $(n + d_i)G$ is a pole of $f$).

**Corollary 4.** If $f$ defined as in Theorem 1, $\ell \geq 1$ is an integer, $d = (d_1, \ldots, d_\ell)$ is a vector of integers such that $d_1 < \cdots < d_\ell$, $N \in \mathbb{N}$ such that $N + d_\ell \leq T$ then

$$\Delta(\Gamma(d, N, \ell)) \ll \left( \frac{3}{2} \right)^\ell N^{-1} \deg f p^{1/2} (\log p)^{\ell} \log T.$$

### 3 Pseudorandom measures of binary sequences over elliptic curves

In order to study the pseudorandomness of finite binary sequences Mauduit and Sarközy introduced certain measures of pseudorandomness in [11]. For a given binary sequence

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

the well-distribution measure of $E_N$ 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 $a, b, t \in \mathbb{N}$ such that $1 \leq a \leq a + (t - 1)b \leq N$, and the correlation measure of order $\ell$ of $E_N$ is defined as

$$C_\ell(E_N) = \max_{M,d} \left| \sum_{n=1}^{M} e_{n+d_1} e_{n+d_2} \cdots e_{n+d_\ell} \right|,$$

where the maximum is taken over all $d = (d_1, \ldots, d_\ell)$ and $M$ such that $0 \leq d_1 < d_2 < \cdots < d_\ell \leq N - M$.

We remark that the pseudorandom measures of the sequence $E_T$ defined by (3) has already been studied in some special cases. For example Chen
and Xiao [5] studied the case when \( f(x, y) = x, y, y/2 \) or \( x/2 \). Later Liu, Zhan and Wang studied the case when \( f \) is a general “polynomial” (i.e. the only pole of \( f \) is \( O \)) or its multiplicative inverse [15]. However in the general case we need further criteria to the curve \( E(\mathbb{F}_p) \) and the function \( f \) as the following example shows.

**Example 3.1.** Let \( E \) defined by \( y^2 = x^3 - 2x \) over \( \mathbb{F}_{19} \). Then \( |E| = 20 \) and \( G = (2, 2) \) is a generator. If \( f(x, y) = 9x + \frac{1}{x} \) then the sequence \( E_{20} \) defined by (3) is

<table>
<thead>
<tr>
<th>( n )</th>
<th>( nG )</th>
<th>( f(nG) )</th>
<th>( e_n )</th>
<th>( n )</th>
<th>( nG )</th>
<th>( f(nG) )</th>
<th>( e_n )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>(2,2)</td>
<td>9</td>
<td>+1</td>
<td>11</td>
<td>(18,1)</td>
<td>9</td>
<td>+1</td>
</tr>
<tr>
<td>2</td>
<td>(7,14)</td>
<td>17</td>
<td>-1</td>
<td>12</td>
<td>(16,6)</td>
<td>17</td>
<td>-1</td>
</tr>
<tr>
<td>3</td>
<td>(15,1)</td>
<td>16</td>
<td>-1</td>
<td>13</td>
<td>(10,12)</td>
<td>16</td>
<td>-1</td>
</tr>
<tr>
<td>4</td>
<td>(11,6)</td>
<td>11</td>
<td>-1</td>
<td>14</td>
<td>(5,18)</td>
<td>11</td>
<td>-1</td>
</tr>
<tr>
<td>5</td>
<td>(13,10)</td>
<td>6</td>
<td>+1</td>
<td>15</td>
<td>(13,9)</td>
<td>6</td>
<td>+1</td>
</tr>
<tr>
<td>6</td>
<td>(5,1)</td>
<td>11</td>
<td>-1</td>
<td>16</td>
<td>(11,13)</td>
<td>11</td>
<td>-1</td>
</tr>
<tr>
<td>7</td>
<td>(10,7)</td>
<td>16</td>
<td>-1</td>
<td>17</td>
<td>(15,18)</td>
<td>16</td>
<td>-1</td>
</tr>
<tr>
<td>8</td>
<td>(16,13)</td>
<td>17</td>
<td>-1</td>
<td>18</td>
<td>(7,5)</td>
<td>17</td>
<td>-1</td>
</tr>
<tr>
<td>9</td>
<td>(18,18)</td>
<td>9</td>
<td>+1</td>
<td>19</td>
<td>(2,17)</td>
<td>9</td>
<td>+1</td>
</tr>
<tr>
<td>10</td>
<td>(0,0)</td>
<td>( \infty )</td>
<td>-1</td>
<td>20</td>
<td>( O )</td>
<td>( \infty )</td>
<td>-1</td>
</tr>
</tbody>
</table>

Here \( e_n = e_{n+10} \), thus \( C_2(E_{20}) \) is large. Then this binary sequence is not a pseudorandom sequence and thus the sequence \( s_n = x(nG) \) is not either.

We remark that in the case when the order of the elliptic curve group has small prime factors, many similar example of binary sequence could be constructed along the same line which are certainly not a random type.

**Theorem 2.** Let \( p > 3 \) be a prime number, let \( G \in E(\mathbb{F}_p) \) of order \( T \), \( f \in \mathbb{F}_p(E) \). If we define the sequence \( E_T = \{e_1, \ldots, e_T\} \) by (3) then we have

\[
W(E_T) \ll \deg f p^{1/2} \log p \log T. \tag{8}
\]

**Proof.** Let \( a, b, t \in \mathbb{N} \) such that \( 1 \leq a \leq a + (t - 1)b \leq T \). Then by [16, Theorem 2] we have

\[
\sum_{j=0}^{t-1} e_{a+jb} \ll t \Delta(\Gamma(t)),
\]

where

\[
\Gamma(t) = \left( \frac{f(aG)}{p}, \ldots, \frac{f((a+(t-1)b)G)}{p} \right).
\]
The Kohel-Shparlinski and the Erdős-Turán-Koksma bound give
\[
\Delta(\Gamma(t)) \ll t^{-1} \deg f \, p^{1/2} \log p \log T
\]
which proves the theorem.

\textbf{Theorem 3.} Let \( p, G, T, f, E_T \) be as in Theorem 2. Let \( p(T) \) be the least prime factor of \( T \). If one of the following condition holds:

1. \( \deg f < p(T) \) and \( \ell = 2 \);
2. \( \deg f < p(T) \) and \( (4 \deg f)^\ell < p(T) \)

then
\[
C_\ell(E_T) \ll \ell \deg f p^{1/2} (\log p)^\ell \log T. \tag{9}
\]

\textbf{Proof.} Let \( d = (d_1, \ldots, d_\ell) \) and \( M \) such that \( 0 \leq d_1 < \cdots < d_\ell < T - M \). Then [16, Theorem 1] and Corollary 4 give
\[
\left| \sum_{n=1}^M \varepsilon_{n+d_1} \cdots \varepsilon_{n+d_\ell} \right| \ll M \Delta(\Gamma(d, M, \ell)) \ll \ell \deg f p^{1/2} (\log p)^\ell \log T
\]
which proves the result.

\textbf{References}


[16] C. Mauduit, H. Niederreiter and A. Sárközy, On pseudorandom \([0,1)\)

[17] C. Mauduit, A. Sárközy, On finite pseudorandom binary sequences I:
Measures of pseudorandomness, the Legendre symbol, Acta Arith. \textbf{82}


[19] L. Mérai, Construction of pseudorandom binary sequences over elliptic

Recent trends in cryptography, 121–141, Contemp. Math., 477, Amer.


[22] A. Topuzoğlu, A. Winterhof, Pseudorandom sequences. Topics in geo-
metry, coding theory and cryptography, 135–166, Algebr. Appl., 6, Springer,

LÁSZLÓ MÉRAI, ALFRÉD RÉNYI INSTITUTE OF MATHEMATICS, BU-
DAPEST, PF. 127, H-1364 HUNGARY
E-mail address: merai@cs.elte.hu