PREDICTING THE ELLIPTIC CURVE CONGRUENTIAL GENERATOR

LÁSZLÓ MÉRAI

ABSTRACT. Let p be a prime and let $\mathbf{E}(\mathbb{F}_p)$ be an elliptic curve defined over the finite field $\mathbb{F}_p$ of p elements. For a point $G \in \mathbf{E}(\mathbb{F}_p)$ the elliptic curve congruential generator (with respect to the first coordinate) is a sequence (x_n) defined by the relation $x_n = x(W_n) = x(W_{n-1} \oplus G) = x(nG \oplus W_0)$, $n = 1, 2, \dots$, where $\oplus$ denotes the group operation in $\mathbf{E}(\mathbb{F}_p)$ and W_0 is an initial point. In this paper, we show that if some consecutive elements of the sequence (x_n) are given as integers, then one can recover the whole sequence, even if the prime and the elliptic curve are private.

1. Introduction

For a prime p, denote by $\mathbb{F}_p$ the field of p elements and always assume that it is represented by the first p-many non-negative integers $\{0, 1, \dots, p-1\}$.

Let $\mathbf{E}$ be an elliptic curve defined over $\mathbb{F}_p$ given by an *affine Weierstrass equation*, which for $\gcd(p, 6) = 1$ takes the form

$$y^2 = x^3 + Ax + B, \tag{1}$$

for some $A, B \in \mathbb{F}_p$ with non-zero discriminant $4A^3 + 27B^2 \neq 0$.

The $\mathbb{F}_p$-rational points $\mathbf{E}(\mathbb{F}_p)$ of $\mathbf{E}$ form an Abelian group (with respect to the usual addition denoted by $\oplus$) with the point at infinity $\mathcal{O}$ as the neutral element. For a point $P \in \mathbf{E}(\mathbb{F}_p)$, $P \neq \mathcal{O}$ we denote by $x(P)$ and $y(P)$ its affine components, $P = (x(P), y(P))$.

For a given point $G \in \mathbf{E}(\mathbb{F}_p)$ and *initial point* $W_0 \in \mathbf{E}(\mathbb{F}_p)$ the *elliptic curve congruential generator* is the sequence (W_n) of points of $\mathbf{E}(\mathbb{F}_p)$ satisfying the recurrence relation

$$W_n = G \oplus W_{n-1} = nG \oplus W_0, \quad n = 1, 2, \dots \tag{2}$$

We also define the *elliptic curve congruential generator with respect to the first coordinate* as the sequence (x_n) in $\mathbb{F}_p$ as

$$x_n = x(W_n) = x(nG \oplus W_0), \quad n = 1, 2, \dots \tag{3}$$

The elliptic curve congruential generator has been wildly studied, many positive results have been proven about its randomness, see [1–6, 8, 9, 11–15, 17] and see also the survey paper [16]. In particular, El Mahassni and Shparlinski [5] showed that (W_n), and so the sequence (x_n), is well-distributed. Hess and Shparlinski [8], and Topuzoğlu and Winterhof [17] provided lower bounds to the linear complexity profile of the sequence (x_n).

However, it is clear, that when the curve $\mathbf{E}$ are given, the sequence (W_n) is predictable from two consecutive points W_n, W_{n+1}. In [7], Gutierrez and Ibeas showed, that when the prime p and G are known, then the sequence (W_n) is predictable even if just an approximation of W_n, W_{n+1} are revealed (even if the curve $\mathbf{E}$ is private). These show that the security of the point sequence (W_n) is not well-established. However these attacks use the assumption that

2010 *Mathematics Subject Classification.* Primary 11Y50, 11Y55, 11T71, 14H52, 94A60.
Key words and phrases. elliptic curve, congruential generator, cryptography.

the prime and the curve (resp. the point G) is given which assumption is quite optimistic (in the viewpoint of the attacker).

In our cryptographic settings, all the parameters, as the initial point $W_0 = (x_0, y_0)$, the generator $G = (x, y)$, the parameters of the curve A, B and the prime p are assumed to be secret and just the output of the generator $x_1, x_2, \dots$ represented as non-negative integers are used. The main contribution of this paper is that all the secret parameters, and thus the whole sequence (x_n), can be computed in polynomial time (polynomial in $\log p$) if not too many consecutive elements are revealed.

The result suggest that for cryptographic application the elliptic curve congruential generator should be use with great care.

In Section 2 we summarize some basic fact about elliptic curves. In Section 3 we present the algorithm, and in Section 4 we discuss results of numerical tests.

Acknowledgements

The author would like to thank Harald Niederreiter for suggesting the problem. The author is partially supported by the Austrian Science Fund FWF Project F5511-N26 which is part of the Special Research Program "Quasi-Monte Carlo Methods: Theory and Applications" and by Hungarian National Foundation for Scientific Research, Grant No. K100291.

2. Background

In this section we summarize some basic facts about elliptic curves. First we recall the definition of the group operation $\oplus$ of $\mathbf{E}(k)$ with arbitrary field k. Then we extend the notion of elliptic curve when it is defined over a ring.

2.1. The group law on the curve.

Let k be a field of characteristic different from $2, 3$. Let $\mathbf{E}$ be an elliptic curve defined over k given by an affine Weierstrass equation (1) with $A, B \in k$, $4A^3 + 27B^2 \neq 0$. The group operation $\oplus$ in $\mathbf{E}(k)$ is defined in the following way.

Definition 1. The operation $\oplus$ is defined over $\mathbf{E}(k)$ as follows. If $P = (x_P, y_P)$ and $Q = (x_Q, y_Q)$ are in $\mathbf{E}(k)$, then

$$P \oplus Q = R = (x_R, y_R),$$

where

(i.) if $x_P \neq x_Q$, then

$$x_R = s^2 - x_P - x_Q, \quad y_R = s(x_P - x_R) - y_P, \quad \text{where } s = \frac{y_Q - y_P}{x_Q - x_P};$$

(ii.) if $x_P = x_Q$ but $y_P \neq y_Q$, then $P \oplus Q = \mathcal{O}$;

(iii.) if $P = Q$ and $y_P \neq 0$, then

$$x_R = s^2 - 2x_P, \quad y_R = s(x_P - x_R) - y_P, \quad \text{where } s = \frac{3x_P^2 + A}{2y_P};$$

(iv.) if $P = Q$ and $y_P = 0$, then $P \oplus Q = \mathcal{O}$.

2.2. **Elliptic curves over $\mathbb{Z}_m$.** If m is an odd composite integer, elliptic curve $\mathbf{E}$ can be also defined over $\mathbb{Z}_m$ via the *projective Weierstrass equation*

$$y^2 z = x^3 + Axz^2 + Bz^3 \tag{4}$$

with $A, B \in \mathbb{Z}_m$, $\gcd(4A^3+27B^2, m) = 1$. The $\mathbb{Z}_m$-rational points $\mathbf{E}(\mathbb{Z}_m)$ of $\mathbf{E}$ with projective coordinates can be represented as a triple $(x : y : z)$ such that $\gcd(m, x, y, z) = 1$ (but not necessarily $z = 0$ or 1).

As in the field case, group operation can be defined on $\mathbf{E}(\mathbb{Z}_m)$ (see, [10,18]) whose formulas correspond to Definition 1 if the divisor in (i.) or (ii.) is co-prime to m.

We remark that for odd integers m_1, m_2, $\gcd(m_1, m_2) = 1$ we have

$$\mathbf{E}(\mathbb{Z}_{m_1 m_2}) \cong \mathbf{E}(\mathbb{Z}_{m_1}) \otimes \mathbf{E}(\mathbb{Z}_{m_2})$$

as groups. Moreover, if $\mathbf{E}$ is an elliptic curve over $\mathbb{Q}$ defined by (1) with integers A, B, then for odd m with $\gcd(4A^3 + 27B^2, m) = 1$, the map

$$\begin{array}{ccc} \mathbf{E}(\mathbb{Q}) & \to & \mathbf{E}(\mathbb{Z}_m) \\ (x : y : z) & \mapsto & (x \bmod m : y \bmod m : z \bmod m) \end{array}$$

is a group homomorphism (where the representation $(x : y : z)$ is chosen as $x, y, z \in \mathbb{Z}$ and $\gcd(x, y, z) = 1$).

Finally, for arbitrary odd integers m_1, m_2 (not necessarily co-primes), the the map

$$\begin{array}{ccc} \mathbf{E}(\mathbb{Z}_{m_1 m_2}) & \to & \mathbf{E}(\mathbb{Z}_{m_1}) \\ (x : y : z) & \mapsto & (x \bmod m_2 : y \bmod m_2 : z \bmod m_2) \end{array}$$

is a group homomorphism.

3. Predicting the congruential generator on elliptic curve over $\mathbb{Z}_m$

Suppose we are given $x_1, \ldots, x_s$ an initial segment of a sequence (x_n) generated by an elliptic curve generator as non-negative integers. We would like to predict the remainder part of this sequence and specially, to compute the parameters of the generator, namely, the prime p, the parameters of the curve A, B and the points G, W_0.

If for two different generators with primes p and q, the revealed initial segments coincide, then the same initial segment is generated by an elliptic curve generator over $\mathbb{Z}_{p \cdot q}$. Clearly, in this case only the generator over the ring $\mathbb{Z}_{p \cdot q}$ is computable (without assuming the easiness of the integer factorization problem) and to recover the private parameters further revealed elements are needed.

On the other hand, if the curve $\mathbf{E}$ is defined over $\mathbb{Q}$ with non-negative integers A, B and $G, W_0 \in \mathbf{E}(\mathbb{Q})$ are points such that $x(iG \oplus W_0)$ are all integers for $i = 1, \ldots, s$, then there are infinitely many possible primes p (and generators) exist, namely all large enough primes are suitable.

Thus our aim is to determine the most general elliptic curve generator (possibly over a ring $\mathbb{Z}_m$ or over $\mathbb{Q}$) which generates the same initial segment.

The following theorem shows that if at least seven initial values are revealed, then it can be computed a curve $\mathbf{E}$ over $\mathbb{Q}$ or over $\mathbb{Z}_m$ with $p \mid m$ and points G, W_0 such that these revealed values are the initial segment of a sequence generated by an elliptic curve generator with $\mathbf{E}$, G and W_0. If more values are revealed, then better approximation can be given to the generator (i.e. to the prime p).

Theorem 1. *Let $p > 3$ be a prime number, $\mathbf{E} = \mathbf{E}_{A,B}$ be an elliptic curve over $\mathbb{F}_p$ given by the Weierstrass equation (1), $W_0, G \in \mathbf{E}(\mathbb{F}_p)$ such that $W_n \neq \pm G$ for $n = 1, \dots, 6$. If the sequence (x_n) is defined by (3) and $x_i \neq x_j$ ($1 < i < j \leq 7$), then there exists a polynomial time algorithm (polynomial in $\log p$) which, when given $x_1, \dots, x_7$ as integers, returns a curve $\mathbf{E}_{\widetilde{A},\widetilde{B}}$ over $\mathbb{Q}$ or over $\mathbb{Z}_m$ with $p \mid m$ and a sequence $\widetilde{x}_8, \widetilde{x}_9 \dots$ such that the sequence $x_1, \dots, x_7, \widetilde{x}_8, \widetilde{x}_9, \dots$ is generated by the elliptic curve congruential generator with $\mathbf{E}_{\widetilde{A},\widetilde{B}}$ and $\widetilde{x}_n \equiv x_n \mod p$ if $W_n \neq \mathcal{O}$.*

Remark 2. If we increase the number of revealed elements, then the algorithm of Theorem 1 can be extended which provides a better approximation m to p and hence a better approximation to (x_n), see Section 4. In practice, 8 elements contain enough information to learn the exact value of p.

Proof of Theorem 1. Let us assume, that the integers $x_1, \dots, x_7$ generated by an elliptic curve congruential generator (3) are given.

By (3), we can write

$$x_{i-1} = x(W_i \oplus (-G)), \quad x_i = x(W_i), \quad x_{i+1} = x(W_i \oplus G), \quad i = 2, \dots, 6,$$

where $-G$ is the (additive) inverse of G: $-G = (x, -y)$. By the addition law, by the assumption that $W_i \neq \pm G$ ($i = 1, \dots, 6$) and by (1) we have

$$x_{i-1} + x_{i+1} = \left(\frac{y_i + y}{x_i - x}\right)^2 - x_i - x + \left(\frac{y_i - y}{x_i - x}\right)^2 - x_i - x \tag{5}$$

$$= 2\frac{y_i^2 + y^2}{(x_i - x)^2} - 2(x_i + x) \tag{6}$$

$$= 2\frac{x_i^3 + Ax_i + B + y^2}{(x_i - x)^2} - 2(x_i + x), \quad i = 2, \dots, 6$$

in $\mathbb{F}_p$. Thus

$$(x_i - x)^2(x_{i-1} + x_{i+1}) \equiv 2(x_i^3 + Ax_i + B + y^2) - 2(x_i + x)(x_i - x)^2 \mod p, \quad i = 2, \dots, 6$$

i.e.,

$$(2x_i^2 + 2x_i(x_{i-1} + x_{i+1}))x + (2x_i - (x_{i-1} + x_{i+1}))x^2 + 2x_iA + 2B + 2y^2 - 2x^3 \equiv (x_{i-1} + x_{i+1})x_i^2 \mod p, \quad i = 2, \dots, 6. \tag{7}$$

Put

$$C = \begin{pmatrix} 2x_2^2 + 2x_2(x_1 + x_3) & 2x_2 - (x_1 + x_3) & 2x_2 & 2 & 2 & -2 \\ 2x_3^2 + 2x_3(x_2 + x_4) & 2x_3 - (x_2 + x_4) & 2x_3 & 2 & 2 & -2 \\ 2x_4^2 + 2x_4(x_3 + x_5) & 2x_4 - (x_3 + x_5) & 2x_4 & 2 & 2 & -2 \\ 2x_5^2 + 2x_5(x_4 + x_6) & 2x_5 - (x_4 + x_6) & 2x_5 & 2 & 2 & -2 \\ 2x_6^2 + 2x_6(x_5 + x_7) & 2x_6 - (x_5 + x_7) & 2x_6 & 2 & 2 & -2 \end{pmatrix} \in \mathbb{Q}^{5\times 6}$$

and

$$\mathbf{u} = \begin{pmatrix} (x_1 + x_3)x_2^2 \\ (x_2 + x_4)x_3^2 \\ (x_3 + x_5)x_4^2 \\ (x_4 + x_6)x_5^2 \\ (x_5 + x_7)x_6^2 \end{pmatrix} \in \mathbb{Q}^5.$$

Write $C = (\mathbf{c}_1, \dots, \mathbf{c}_6)$ with $\mathbf{c}_1, \dots, \mathbf{c}_6 \in \mathbb{Z}^5$. Then we have

Lemma 1. *Having the same assumption as in Theorem 1, the columns* $\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3, \mathbf{c}_4$ *are linearly independent over* $\mathbb{F}_p$.

Assuming Lemma 1 the matrix C has rank 4 over $\mathbb{F}_p$, and by (7) the congruence

$$C \cdot \mathbf{e} \equiv \mathbf{u} \mod p \tag{8}$$

has the solution $\mathbf{e} = (x, x^2, A, B, y^2, x^3)^T$.

In the next step we look for the maximal integer m such that the congruence (8) has a solution $\mathbf{e} = (e_1, \ldots, e_6)^T$ modulo m with the additional restriction $e_1^2 \equiv e_2 \mod m$. Clearly, we will have $p \mid m$.

During the algorithm, if m takes a finite value, then we can suppose that it is odd, since $p > 3$.

Since the congruence (8) has solutions, $\det(\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3, \mathbf{c}_4, \mathbf{u}) \equiv 0 \mod p$. First assume, that $\det(\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3, \mathbf{c}_4, \mathbf{u}) \neq 0$ and put $m = \det(\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3, \mathbf{c}_4, \mathbf{u})$. If $\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3, \mathbf{c}_4$ are linearly dependent modulo m, say, $\lambda_1\mathbf{c}_1 + \lambda_2\mathbf{c}_2 + \lambda_3\mathbf{c}_3 + \lambda_4\mathbf{c}_4 \equiv 0 \mod m$ with $(\lambda_1, \lambda_2, \lambda_3, \lambda_4) \neq (0, 0, 0, 0)$, then we have $p \mid \gcd(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$. Replacing m to $\gcd(m, \lambda_1, \lambda_2, \lambda_4, \lambda_5)$ we get that $\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3, \mathbf{c}_4$ are linearly independent modulo m. By the vanishing of the determinant, $\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3, \mathbf{c}_4, \mathbf{u}$ are linearly dependent mod m: $\lambda_1\mathbf{c}_1 + \lambda_2\mathbf{c}_2 + \lambda_3\mathbf{c}_3 + \lambda_4\mathbf{c}_4 \equiv \mu\mathbf{u} \mod m$. By the independence of $\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3, \mathbf{c}_4$, $\mu \not\equiv 0 \mod m$ and if μ is minimal and positive, the coefficients $\lambda_1, \lambda_2, \lambda_3, \lambda_4, \mu$ are unique. If $\gcd(m, \mu) > 1$, replacing m to $m/\gcd(m, \mu)$ we obtain $\widetilde{\lambda}_1\mathbf{c}_1 + \widetilde{\lambda}_2\mathbf{c}_2 + \widetilde{\lambda}_3\mathbf{c}_3 + \widetilde{\lambda}_4\mathbf{c}_4 \equiv \mathbf{u} \mod m$ with unique $\widetilde{\lambda}_1, \widetilde{\lambda}_2, \widetilde{\lambda}_3, \widetilde{\lambda}_4$. Finally, all solutions of the congruence

$$C \cdot \mathbf{e} \equiv \mathbf{u} \mod m \tag{9}$$

can be expressed as

$$e_1 = \widetilde{\lambda}_1, \quad e_2 = \widetilde{\lambda}_2, \quad e_3 = \widetilde{\lambda}_3, \quad e_4 + e_5 - e_6 = \widetilde{\lambda}_4. \tag{10}$$

If $\widetilde{\lambda}_1^2 \not\equiv \widetilde{\lambda}_2 \mod m$, we have to replace m to $\gcd(m, \widetilde{\lambda}_1^2 - \widetilde{\lambda}_2)$.

Next, consider the case when $\det(\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3, \mathbf{c}_4, \mathbf{u}) = 0$ (over $\mathbb{Q}$). Now, the equation

$$C \cdot \mathbf{e} = \mathbf{u}$$

has solutions over $\mathbb{Q}$. By Lemma 1, $\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3, \mathbf{c}_4$ are linearly independent over $\mathbb{F}_p$ and thus over $\mathbb{Q}$. As before

$$\lambda_1\mathbf{c}_1 + \lambda_2\mathbf{c}_2 + \lambda_3\mathbf{c}_3 + \lambda_4\mathbf{c}_4 = \mu\mathbf{u}, \quad \gcd(\lambda_1, \lambda_2, \lambda_3, \lambda_4, \mu) = 1, \quad \mu > 0 \tag{11}$$

with unique integers $\lambda_1, \lambda_2, \lambda_3, \lambda_4, \mu$. Moreover all the integer solution $(\lambda_1', \lambda_2', \lambda_3', \lambda_4', \mu')$ of (11) has the form $\gamma \cdot (\lambda_1, \lambda_2, \lambda_3, \lambda_4, \mu)$ with some $\gamma \in \mathbb{Z}$. Since $(x, x^2, A, B + y^2 - x^3, 1)$ is a solution modulo p we have for $\gcd(p, \mu) = 1$, that $(\lambda_1/\mu)^2 \equiv \lambda_2/\mu \mod p$. Put $m = (\lambda_1^2 - \lambda_2\mu)/\gcd(\lambda_1^2, \mu)$ so $p \mid m$. If $m \neq 0$, write the solutions of (9) in the same form as (10).

In both cases, if $m \neq 0$ write

$$\widetilde{x} = \widetilde{\lambda}_1, \quad \widetilde{(y^2)} = \widetilde{\lambda}_1^3 + \frac{\widetilde{\lambda}_4 + \widetilde{\lambda}_1\widetilde{\lambda}_3}{2}, \quad \widetilde{A} = \widetilde{\lambda}_3, \quad \widetilde{B} = \frac{\widetilde{\lambda}_4 - \widetilde{\lambda}_1\widetilde{\lambda}_3}{2} \quad \text{over } \mathbb{Z}_m.$$

Clearly, if $\widetilde{G} = (\widetilde{x}, \widetilde{y})$ with a $\widetilde{y} \in \mathbb{Z}_m[\zeta]/(\zeta^2 - \widetilde{(y^2)})$, $\widetilde{y}^2 = \widetilde{(y^2)}$, then $\widetilde{G} \in \mathbf{E}_{\widetilde{A},\widetilde{B}}$ over $\mathbb{Z}_m[\zeta]/(\zeta^2 - \widetilde{(y^2)})$. Moreover, the vector

$$\mathbf{e} = (\widetilde{x}, \widetilde{x}^2, \widetilde{A}, \widetilde{B}, \widetilde{(y^2)}, \widetilde{x}^3)^T \in \mathbb{Z}_m^6$$

is a solution of (9). If $\widetilde{y}_1 \in \mathbb{Z}_m[\xi]/(\xi^2 - x_1^3 - \widetilde{A}x_1 - \widetilde{B})$ is an element such that $\widetilde{W_1} = (x_1, \widetilde{y}_1)$ is on the curve $\mathbf{E}_{\widetilde{A},\widetilde{B}}$, then writing $\widetilde{W_0} = \widetilde{W_1} \oplus (-\widetilde{G})$, the sequence $(\widetilde{x}_n)$ generated by the elliptic curve congruence generator (with $\widetilde{W_0}$, $\widetilde{G}$) satisfies

$$\widetilde{x}_n = 2\frac{\widetilde{x}_{n-1}^3 + \widetilde{A}\widetilde{x}_{n-1} + \widetilde{B} + \widetilde{(y^2)}}{(\widetilde{x}_{n-1} - \widetilde{x})^2} - 2(\widetilde{x}_{n-1} + \widetilde{x}) - \widetilde{x}_{n-2}, \quad n = 2, 3, \dots, \tag{12}$$

so $\widetilde{x}_1, \widetilde{x}_2, \dots$ are in $\mathbb{Z}_m$ and thus $\widetilde{x}_i = x_i$ $(i = 1, 2, \dots, 7)$ as integers.

Finally, if $m = 0$, then write

$$\widetilde{x} = \frac{\lambda_1}{\mu}, \quad \widetilde{(y^2)} = \left(\frac{\lambda_1}{\mu}\right)^3 + \frac{\lambda_4}{2\mu} + \frac{\lambda_1\lambda_3}{2\mu^2}, \quad \widetilde{A} = \frac{\lambda_3}{\mu}, \quad \widetilde{B} = \frac{\lambda_4}{2\mu} - \frac{\lambda_1\lambda_3}{2\mu^2} \quad \text{over } \mathbb{Q}.$$

Then $\widetilde{G} = (\widetilde{x}, \widetilde{y}) \in \mathbf{E}_{\widetilde{A},\widetilde{B}}$ over $\mathbb{Q}\left(\sqrt{\widetilde{(y^2)}}\right)$ and the integer sequence $x_1 \dots, x_7$ is generated by $\widetilde{G}$. The set of possible primes p are those ones which $p \nmid \mu$ and $p > \max\{x_i : \ i = 1, \dots, 7\}$. □

Finally, it remains to prove Lemma 1.

Proof of Lemma 1. Clearly, it is enough to show that the vectors

$$\mathbf{v}_1 = \begin{pmatrix} x_2^2 + x_2(x_1 + x_3) \\ x_3^2 + x_3(x_2 + x_4) \\ x_4^2 + x_4(x_3 + x_5) \\ x_5^2 + x_5(x_4 + x_6) \\ x_6^2 + x_6(x_5 + x_7) \end{pmatrix}, \mathbf{v}_2 = \begin{pmatrix} x_1 + x_3 \\ x_2 + x_4 \\ x_3 + x_5 \\ x_4 + x_6 \\ x_5 + x_7 \end{pmatrix}, \mathbf{v}_3 = \begin{pmatrix} x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{pmatrix}, \mathbf{v}_4 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

are linearly independent over $\mathbb{F}_p$.

Suppose to the contrary that there are $\alpha_1, \dots, \alpha_4 \in \mathbb{F}_p$, $(\alpha_1, \dots, \alpha_4) \neq (0, \dots, 0)$ such that

$$\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \alpha_3\mathbf{v}_3 + \alpha_4\mathbf{v}_4 = 0.$$

i.e.,

$$\alpha_1(x_i^2 + x_i(x_{i-1} + x_{i+1})) + \alpha_2(x_{i-1} + x_{i+1}) + \alpha_3 x_i + \alpha_4 = 0, \quad i = 2, \dots, 6. \tag{13}$$

Substituting (5) to (13) we have

$$\alpha_1\left(x_i^2 + 2x_i\left(\frac{x_i^3 + Ax_i + B + y^2}{(x_i - x)^2} - (x_i + x)\right)\right) + 2\alpha_2\left(\frac{x_i^3 + Ax_i + B + y^2}{(x_i - x)^2} - (x_i + x)\right) + \alpha_3 x_i + \alpha_4 = 0, \quad i = 2, \dots, 6.$$

Clearing the denominator we get

$$\alpha_1\left(x_i^2(x_i - x)^2 + 2x_i\left(x_i^3 + Ax_i + B + y^2 - (x_i + x)(x_i - x)^2\right)\right) + 2\alpha_2\left(x_i^3 + Ax_i + B + y^2 - (x_i + x)(x_i - x)^2\right) + \alpha_3 x_i(x_i - x)^2 + \alpha_4(x_i - x)^2 = 0, \quad i = 2, \dots, 6,$$

which means that the polynomial $F(X) \in \mathbb{F}_p[X]$

$$F(X) = \alpha_1 \left(X^2(X-x)^2 + 2X\left(X^3 + AX + B + y^2 - (X+x)(X-x)^2\right)\right) \\ + 2\alpha_2 \left(X^3 + AX + B + y^2(X+x)(X-x)^2\right) + \alpha_3 X(X-x)^2 + \alpha_4(X-x)^2$$

has at least five zeros: x_i, $i = 2, \ldots 6$ (where $x = x(G)$ and $y = y(G)$ are fixed).

Write $F(X)$ into the following form

$$F(X) = \alpha_1 \left(X^4 + f_4(X)\right) + \alpha_3 \left(X^3 + f_3(X)\right) + (2\alpha_2 x + \alpha_4)X^2 \\ + (\alpha_2(2x^2 + 2A) - 2\alpha_4 x)X + \alpha_2(-2x^3 + 2B + 2y^2) + \alpha_4 x^2, \tag{14}$$

where $f_3, f_4 \in \mathbb{F}_p[X]$ with $\deg f_3 < 3$ and $\deg f_4 < 4$.

Since $\deg F \leq 4$, we must have that $F(X)$ is the zero polynomial. In this case we have

$$\alpha_1 = 0, \quad \alpha_3 = 0, \quad 2\alpha_2 x + \alpha_4 = 0. \tag{15}$$

Then the coefficients of X and 1 in (14) are

$$\alpha_2(3x^2 + A) = 0, \quad \alpha_2(2x^3 - B - y^2) = 0.$$

By (15) and $(\alpha_1, \ldots, \alpha_4) \neq (0, \ldots, 0)$ we also have $\alpha_2, \alpha_4 \neq 0$, thus

$$3x^2 + A = 0, \quad 2x^3 - B - y^2 = 0,$$

whence using (1) we get

$$2y^2 = x(3x^2 + A) - (x^3 + Ax + B - y^2) - (2x^3 - B - y^2) = 0,$$

thus

$$y = 0, \quad 3x^2 + A = 0.$$

Since $G = (x, y) \in \mathbf{E}(\mathbb{F}_p)$, we get that x is a multiple root of the right hand side of (1), which contradict that the discriminant of the curve is non-zero. □

4. Numerical tests

I have implemented the algorithm of Theorem 1 in SAGE. The algorithm have been tested for 1000 random examples of generators with 500-bit primes p. For seven revealed sequence elements, the algorithm computed the exact values of the parameters (p, A, B, G, W_0) in 95,2% of the cases. In the remainder cases, the algorithm provided a composite integer m and parameters $\widetilde{A}$, $\widetilde{B}$, $\widetilde{G}$, $\widetilde{W_0}$ such that $p \mid m$ and $\widetilde{A} \equiv A$, $\widetilde{B} \equiv B$, $\widetilde{G} \equiv G$, $\widetilde{W_0} \equiv W_0 \mod p$.

If the number of revealed sequence elements increases, then the algorithm can be modified to become more effective. Namely, if there are eight revealed sequence elements, then applying the algorithm for the first and the last seven elements, two approximation (m_1, A_1, B_1, G_1) and (m_2, A_2, B_2, G_2) are provided. Putting $m = \gcd(m_1, m_2, A_1 - A_2, B_1 - B_2, x(G_1) - x(G_2))$, m is a better approximation of p. Modifying the SAGE program in this way, the algorithm was successful in 100% of the cases.

5. Final remarks, open questions

In this paper we showed that the sequence $x_n = x(nG \oplus W_0)$ $(n = 1, 2, \dots)$ is highly predictable. In the literature, the distribution and the linear complexity profile of the general sequence $f(nG \oplus W_0)$ with $f \in \mathbb{F}_p(\mathbf{E})$ have been also studied (where $\mathbb{F}_p(\mathbf{E})$ is the function field of $\mathbf{E}(\mathbb{F}_p)$), see [1–6, 8, 9, 11–15]. The predictability of these sequence could be handle for individual functions f, but it is not clear whether there is a universal algorithm for all function f (or at least all function f with small degree).

An other possible question connected to the result is how much information about $x(nG \oplus W_0)$ we really need to recover all the private parameters. More precisely, if the curve is defined over a finite field $\mathbb{F}_{p^s}$ with degree $s > 1$, then one can define an integer sequence as $(x_1(nG \oplus W_0), \dots, x_r(nG \oplus W_0))$ $n = 1, 2, \dots$ where $r \leq s$ and $x_1(P), \dots, x_s(P)$ are the coordinates of $x(P)$ with respect to a fixed basis of $\mathbb{F}_{p^s}$ over $\mathbb{F}_p$. Is it possible to recover the whole sequence from an initial segment, at least when the degree s is fixed? Clearly, the most interesting case when $p = 2$ and the generator builds a binary sequence.

References

[1] P. H. T. Beelen, J. M. Doumen: Pseudorandom Sequences from Elliptic Curves, Finite Fields with Applications to Coding Theory, Cryptography and Related Areas, Springer-Verlag, Berlin-Heidelberg-New York, 37–52 (2002)

[2] Z. Chen, *Elliptic curve analogue of Legendre sequences*, Monatsh. Math. **154**, 1–10 (2008)

[3] Z. Chen, D. Gomez-Perez, G Pirsic, On lattice profile of the elliptic curve linear congruential generators. Period. Math. Hungar. 68 (2014), no. 1, 1–12.

[4] Chen, Z., Li, S., Xiao, G. Construction of pseudorandom binary sequences from elliptic curves by using discrete logarithm in Lecture Notes in Computer Science vol. 4086, pp. 285–294 Springer, Berlin, (2006)

[5] E. El Mahassni, I. Shparlinski, On the uniformity of distribution of congruential generators over elliptic curves. Sequences and their applications (Bergen, 2001), 257–264, Discrete Math. Theor. Comput. Sci. (Lond.), Springer, London, 2002.

[6] G. Gong,T. A. Berson, D. R. Stinson, Elliptic curve pseudorandom sequence generators. Selected areas in cryptography (Kingston, ON, 1999), 34–48, Lecture Notes in Comput. Sci., 1758, Springer, Berlin, 2000.

[7] J. Gutierrez, Á Ibeas, Inferring sequences produced by a linear congruential generator on elliptic curves missing high-order bits. Des. Codes Cryptogr. 45 (2007), no. 2, 199–212.

[8] F. Hess, I. E. Shparlinski, On the linear complexity and multidimensional distribution of congruential generators over elliptic curves. Des. Codes Cryptogr. **35** (2005), 111–117.

[9] H. Hu, L. Hu, D. Feng, On a class of pseudorandom sequences from elliptic curves over finite fields. IEEE Trans. Inform. Theory 53 (2007), no. 7, 2598–2605.

[10] H. W. Lenstra, Elliptic curves and number-theoretic algorithms. Proceedings of the International Congress of Mathematicians, Vol. 1, 2 (Berkeley, Calif., 1986), 99–120, Amer. Math. Soc., Providence, RI, 1987.

[11] H. Liu, A family of elliptic curve pseudorandom binary sequences.English summary) Des. Codes Cryptogr. 73 (2014), no. 1, 251–265.

[12] H. Liu, T. Zhan, X. Wang, Large families of elliptic curve pseudorandom binary sequences. Acta Arith. 140 (2009), no. 2, 135–144.

[13] L. Mérai, Construction of pseudorandom binary sequences over elliptic curves using multiplicative characters. Publ. Math. Debrecen 80 (2012), no. 1-2, 199–213.

[14] L. Mérai, Remarks on pseudorandom binary sequences over elliptic curves, Fund. Inform. 114 (2012) no. 3-4, 301–308,

[15] L. Mérai, A. Winterhof, On the linear complexity profile of some sequences derived from elliptic curves, submitted

[16] I. Shparlinski, Pseudorandom points on elliptic curves over finite fields, Algebraic geometry and its applications, 116–134, Ser. Number Theory Appl., 5, World Sci. Publ., Hackensack, NJ, 2008.

[17] A. Topuzoğlu, A. Winterhof, Pseudorandom sequences. Topics in Geometry, Coding Theory and Cryptography, 135–166, Algebr. Appl. **6**, Springer, Dordrecht, 2007.

[18] L. C. Washington, Elliptic Curves: Number Theory and Cryptography, Second edition. Chapman & Hall/CRC, Boca Raton, FL, 2008.

Johann Radon Institute for Computational and Applied Mathematics, Austrian Academy of Sciences Austrian Academy of Sciences
E-mail address: merai@cs.elte.hu