On the Linear Complexity and Multidimensional Distribution of Congruential Generators over Elliptic Curves

Florian Hess
Department of Computer Science, University of Bristol
Bristol BS8 1UB, UK
florian@cs.bris.ac.uk

Igor E. Shparlinski
Department of Computing, Macquarie University
Sydney, NSW 2109, Australia
igor@ics.mq.edu.au

June 18, 2003

Abstract

We show that the elliptic curve analogue of the linear congruential generator produces sequences with high linear complexity and good multidimensional distribution.

1 Introduction

Let p be a prime and let $m \geq 1$ be an integer. We denote by $\mathbb{F}_p$ the field of p elements which we also identify with the set $\{0, 1, \ldots, p-1\}$.

Let w_0 and g be given elements of $\mathbb{F}_p$. We recall that the *linear congruential generator of pseudorandom numbers* is the sequence $w_1, w_2, \ldots$ of elements of $\mathbb{F}_p$ defined by the recurrence relation

$$w_n = gw_{n-1} = g^n w_0, \qquad n = 1, 2, \ldots, \tag{1}$$

with the *initial value* w_0.

Let $\mathbf{E}$ be an elliptic curve over $\mathbb{F}_p$, given by an affine *Weierstrass equation* of the form

$$y^2 + (a_1 x + a_3)y = x^3 + a_2 x^2 + a_4 x + a_6,$$

see [2, 22]. It is known, see [2, 22], that the set $\mathbf{E}(\mathbb{F}_p)$ of $\mathbb{F}_p$-rational points of $\mathbf{E}$ forms an Abelian group under an appropriate composition rule (which we denote by $\oplus$) and with the point at infinity $\mathcal{O}$ as the neutral element. We also recall that

$$|\#\mathbf{E}(\mathbb{F}_p) - p - 1| \leq 2p^{1/2},$$

where $\#\mathbf{E}(\mathbb{F}_p)$ is the number of $\mathbb{F}_p$-rational points, including the point at infinity $\mathcal{O}$. Furthermore, it will be convenient to work with places of $\mathbb{F}_p(E)$. A place H of degree d of $\mathbb{F}_p(\mathbf{E})$ corresponds to a Galois orbit of d points H_i in $\mathbf{E}(\mathbb{F}_{p^d})$. Let $P \in \mathbf{E}(\mathbb{F}_p)$. The points $H_i \oplus P$ form a Galois orbit and define a place which we denote by $H \oplus P$ (similarly with $\ominus$). Functions in $\mathbb{F}_p(\mathbf{E})$ are Galois invariant and hence have zeros and poles at places. We identify places of degree one and points.

Let $G \in \mathbf{E}(\mathbb{F}_p)$ be a point of order t, that is, t is the size of the cyclic group $\langle G \rangle$ generated by G. Let H be a place of degree d of $\mathbf{E}$ and let

$$\mathcal{F} = \{f_1, \ldots, f_r\} \tag{2}$$

be a set of $r \geq 1$ rational functions in $\mathbb{F}_p(\mathbf{E})$ with pole divisors of the form

$$(f_i)_\infty = (i + \delta)(H), \qquad 1 \leq i \leq r, \tag{3}$$

where

$$\delta = \begin{cases} 1, & \text{if } d = 1, \\ 0, & \text{if } d \geq 2. \end{cases}$$

Since $\mathbf{E}$ has genus one, such functions exist by the theorem of Riemann-Roch. We define $\rho = r + \delta$.

For $r = 2$ and $H = \mathcal{O}$ a natural example is given by $f_1(P) = x(P)$ and $f_2(P) = y(P)$, where $P = (x(P), y(P)) \neq \mathcal{O}$. In particular, $d = 1$, $\rho = 3$ for this example.

Generalising the above construction, for a given *initial value* $W_0 \in \mathbf{E}(\mathbb{F}_p)$, we define the *elliptic curve congruential generator of pseudorandom numbers with respect to* $\mathcal{F}$ as the sequence $(f_1(W_n), \ldots, f_r(W_n))$ of points in $\mathbb{F}_p^r$ where the W_n are defined by the recurrence relation

$$W_n = G \oplus W_{n-1} = nG \oplus W_0, \qquad n = 1, 2, \ldots, \tag{4}$$

with the *initial value* W_0 (see also [1, 4, 6, 7, 8, 14]). If $H \in \langle G \rangle \oplus W_0$ then the sequence is not defined for periodical values of n. In these cases we take any fixed vector (for example the zero vector) as the output of the generator. Obviously the sequence is purely periodic with period t, which is the order of G in the group of points of $\mathbf{E}(\mathbb{F}_p)$.

It is known that the linear congruential generator produces cryptographically weak sequences, see [5, 9, 10, 13], thus partially motivated by this fact an elliptic curve congruential generator has been introduced in [8] and then in [6] where some attractive properties of this generator and similar generators have been established. On the other hand, one of the advantages of the linear congruential generator (1) has been a variety of results about the distribution of its elements [11, 16, 17], In [4], using some recent bounds of exponential sums along subgroups of points on an elliptic curve from [12], a result was established about the distribution of points W_n in the square $[0, p-1]^2$. Here we consider the distribution of rs-tuples of points

$$\left(f_i\left(W_{n+j}\right),\ 1 \leq i \leq r,\ 1 \leq j \leq s\right), \quad n = 0, \ldots, t-1,$$

in an rs-dimensional cube $[0, p-1]^{rs}$. This question (for $r = 2$, $f_1(P) = x(P)$, $f_2(P) = y(P)$ and $s \geq 2$) has been posed in [4] and although it has been clear that the same technique should apply, this generalisation has not been obtained before due to lack of a certain linear independence result. Here we close this gap and prove the required linear independence of certain functions on elliptic curves which leads to a multidimensional analogue of [4]. Moreover, the same statement leads to a new lower bound on the linear complexity of the components of W_n. We recall, that exactly this reason, that is a very low linear complexity, has led to lattice reduction based attacks on the linear congruential generator of pseudorandom numbers (1), see [5, 9, 10, 13]. Thus we show that the pseudorandom numbers (4) are free of this disadvantage.

We remark that properties of elliptic curve analogues of some other pseudorandom sequences have been studied in [20, 21].

Throughout the paper, the implied constants in the symbols 'O' may sometimes depend on the integer parameters $d, r, s \geq 1$ and are absolute otherwise.

2 Preparations

Let $\mathcal{M}_{r,s}(\mathbb{F}_p)$ denote the set of all nonzero $r \times s$ matrices

$$C = (c_{i,j}, 1 \leq i \leq r, 1 \leq j \leq s)$$

over $\mathbb{F}_p$.

For a matrix $C \in \mathcal{M}_{r,s}(\mathbb{F}_p)$, a set $\mathcal{F}$ of functions (2) and a generic point Q on $\mathbf{E}$ we consider the function

$$\mathcal{L}_{C,\mathcal{F}}(Q) = \sum_{i=1}^{r} \sum_{j=1}^{s} c_{i,j} f_i \left(W_j \oplus Q\right),$$

as a function in the function field $\mathbb{F}_p(\mathbf{E})$.

Lemma 1. *For any matrix $C \in \mathcal{M}_{r,s}(\mathbb{F}_p)$ with $s \le t$ and any set $\mathcal{F}$ of functions (2) satisfying (3), $\mathcal{L}_{C,\mathcal{F}}(Q)$ is not constant. The subgroup $\langle G\rangle$ contains at most $sd\rho$ zeros of $\mathcal{L}_{C,\mathcal{F}}(Q)$. If $H \in \langle G\rangle \oplus W_0$ it contains at most s poles of $\mathcal{L}_{C,\mathcal{F}}(Q)$, which are of the form $H \ominus W_j$ for $1 \le j \le s$, and no poles otherwise.*

Proof. The function $f_i(W_j \oplus Q)$ of $\mathbb{F}_p(\mathbf{E})$ has the pole divisor

$$(f_i(W_j \oplus Q))_\infty = (i+\delta)(H \ominus W_j),$$

because $Q \mapsto W_j \oplus Q$ induces a translation automorphism of $\mathbb{F}_p(\mathbf{E})$. Since the $H \ominus W_j$ are different for $j = 1, \dots, s$, and the condition (3) and $C \neq 0$ hold, we obtain that $\mathcal{L}_{C,\mathcal{F}}(Q)$ has poles and is hence not constant. Furthermore, the pole divisor of $\mathcal{L}_{C,\mathcal{F}}(Q)$ has degree at most $sd\rho$ and support in $\{H \ominus W_j \,:\, 1 \le j \le s\}$. This latter set does not contain an element of $\langle G\rangle$ if $H \notin \langle G\rangle \oplus W_0$. The statements about the poles are hereby proven.

The bound for the pole divisor implies that $\mathcal{L}_{C,\mathcal{F}}(Q)$ has at most $sd\rho$ zeros in $\mathbf{E}(\overline{\mathbb{F}}_p)$, where $\overline{\mathbb{F}}_p$ is the algebraic closure of $\mathbb{F}_p$, hence (most likely even less) in G. This proves the statement about the zeros. □

Our other tool are bounds of exponential sums of the form

$$S(C, \mathcal{F}) = \sum_{n=1}^{t} \exp\left(2\pi i \mathcal{L}_{C,\mathcal{F}}(nG)/p\right),$$

where C is an $r \times s$ matrix over $\mathbb{F}_p$.

The following result is a partial case of Corollary 1 of [12].

Lemma 2. *For any set $\mathcal{F}$ of functions (2) satisfying (3), the bound*

$$\max_{C \in \mathcal{M}_{r,s}(\mathbb{F}_p)} |S(C, \mathcal{F})| = O\left(p^{1/2}\right)$$

holds.

Proof. For $t \le s$ the bound is trivial. Otherwise, by Lemma 1 we see that $\mathcal{L}_{C,\mathcal{F}}(Q)$ is not constant and thus Corollary 1 of [12] implies the result. □

3 Main Results

Let $W_n \in \mathbb{F}_p^2$ be a sequence generated by (4).

Given a set $\mathcal{F}$ of functions (2), we denote by $\Delta_s(\mathcal{F})$ the discrepancy of the following point set

$$\left(\frac{f_1(W_{n+1})}{p}, \ldots, \frac{f_r(W_{n+1})}{p}, \ldots, \frac{f_1(W_{n+s})}{p}, \ldots, \frac{f_r(W_{n+s})}{p} \right), \qquad (5)$$

where $n = 0, \ldots, t-1$, in the rs-dimensional unit cube. That is,

$$\Delta_{r,s} = \sup_{\mathcal{B} \subseteq [0,1]^{rs}} \left| \frac{T(\mathcal{B})}{t} - |\mathcal{B}| \right|,$$

where $T(\mathcal{B})$ is the number of points (5) which hit the box $\mathcal{B} = [\alpha_1, \beta_1] \times \ldots \times [\alpha_{rs}, \beta_{rs}] \subseteq [0,1]^{rs}$ of size $|\mathcal{B}| = (\beta_1 - \alpha_1) \ldots (\beta_{rs} - \alpha_{rs})$.

We recall that for at most s values of $n = 0, \ldots, t-1$ some of the above functions may not be defined, and we define the corresponding block of r coordinates in an arbitrary (but fixed) way (for example set it to zero).

Theorem 3. *For any set $\mathcal{F}$ of functions (2) satisfying (3), the bound*

$$\Delta_s(\mathcal{F}) = O(t^{-1} p^{1/2} \log^{rs} p)$$

holds.

Proof. The result follows immediately from Lemma 1 and a standard relation between the discrepancy and exponential sums, given by Corollary 3.11 in [17] for example. □

Generalising the corresponding definition for one dimensional sequences, for example, see [3, 15, 19], we define the *linear complexity* $L(N)$ of an r-dimensional sequence $(v_{1,n}, \ldots, v_{r,n})$, $n = 1, \ldots, N$ over $\mathbb{F}_p$ as the smallest s for which the following relations hold

$$\sum_{i=1}^{r} \sum_{j=1}^{s} c_{i,j} v_{i,n+j} = 0, \qquad 0 \le n \le N - s - 1, \qquad (6)$$

with some fixed $C = (c_{i,j}) \in \mathcal{M}_{r,s}(\mathbb{F}_p)$.

Theorem 4. *For any set $\mathcal{F}$ of functions (2) satisfying (3), the linear complexity $L(N)$ of the r-dimensional sequence*

$$(f_1(W_n), \ldots, f_r(W_n)), \qquad n = 1, \ldots, N,$$

satisfies

$$L(N) \geq \begin{cases} \min\{N/(d\rho+1), t/(d\rho+1)\}, & \text{if } H = W_0, \\ \min\{N/(d\rho+2), t/(d\rho+1)\}, & \text{if } H \in \langle G\rangle \oplus W_0, \\ \min\{N/(d\rho+1), t/(d\rho)\}, & \text{otherwise,} \end{cases}$$

for any N.

Proof. Let $s = L(N)$. Then $s \leq t$ and $\mathcal{L}_{C,\mathcal{F}}(nG) = 0$ whenever it is defined for $n = 0, \ldots, N-s-1$. Using Lemma 1 we see that $\mathcal{L}_{C,\mathcal{F}}(nG)$ is defined for at least $\min\{N-2s, t-s\}$ many distinct points nG in $\langle G\rangle$ if $H \in \langle G\rangle \oplus W_0$ and for at least $\min\{N-s, t\}$ distinct points otherwise. In the special case $H = W_0$ it is defined for at least $\min\{N-s, t-s\}$ many distinct points because the poles of $\mathcal{L}_{C,\mathcal{F}}(Q)$ are contained in $\{(t-1)G, \ldots, (t-s)G\}$. Furthermore, $\mathcal{L}_{C,\mathcal{F}}(nG) = 0$ for at most $sd\rho$ points in $\langle G\rangle$. Thus $\min\{N-s, t-s\} \leq sd\rho$, $\min\{N-2s, t-s\} \leq sd\rho$ and $\min\{N-s, t\} \leq sd\rho$ respectively. □

Let us choose an elliptic curve $\mathbf{E}$ with a cyclic point group, a generator $G \in \mathbf{E}(\mathbb{F}_p)$, thus $t = \#\mathbf{E}(\mathbb{F}_p) \sim p$, functions $f_i \in \mathbb{F}_p(\mathbf{E})$ with $(f_i)_\infty = (i+1)(\mathcal{O})$ for $1 \leq i \leq r$ and $W_0 = \mathcal{O}$. Then $d = 1$, $\rho = r+1$ and $\mathcal{O} = W_0$, and Theorem 4 gives

$$L(N) \geq \min\left\{\frac{N}{r+2}, \frac{t}{r+2}\right\}. \tag{7}$$

Let us now choose an elliptic curve $\mathbf{E}$ and $G \in \mathbf{E}(\mathbb{F}_p)$ such that $\#\mathbf{E}(\mathbb{F}_p) = 2t$ where t is the order of G. Then take $W_0 \in \mathbf{E}(\mathbb{F}_p)\backslash\langle G\rangle$ and choose functions f_i as above. Thus $d = 1$, $\rho = r+1$ and $\mathcal{O} \notin \langle G\rangle \oplus W_0$, and we obtain

$$L(N) \geq \min\left\{\frac{N}{r+2}, \frac{t}{r+1}\right\}. \tag{8}$$

Finally, we choose an elliptic curve with a cyclic point group over $\mathbb{F}_p$. Let G be a generator of this group, thus $t = \#\mathbf{E}(\mathbb{F}_p) \sim p$. We choose a place H of degree two and let $f_i \in \mathbb{F}_p(\mathbf{E})$ with $(f_i)_\infty = i(H)$. Then $d = 2$, $\rho = r$ and $H \notin \langle G \rangle \oplus W_0 = \mathbf{E}(\mathbb{F}_p)$, and by Theorem 4 we obtain

$$L(N) \geq \min\left\{\frac{N}{2r+1}, \frac{t}{2r}\right\}. \tag{9}$$

If we take for example $r = 1$, then from (9) we obtain a sequence of period about p with $L(N) \geq N/3$ from the above constructions by taking any function in $\mathbb{F}_p(\mathbf{E})$ which has precisely one simple pole at a place of degree two (this is smallest possible). The easiest examples of such functions are $f(P) = 1/(x(P)-a)$ where $a \in \mathbb{F}_p$ such that a is not equal to an x-coordinate of a point of $\mathbf{E}(\mathbb{F}_p)\backslash\{\mathcal{O}\}$.

Accordingly, the bound (8) with $r = 1$ leads to the same result but for sequences of period about $p/2$. We remark that for a random sequence of elements of $\mathbb{F}_p$ one should expect the linear complexity to be close to $N/2$, see [18].

4 Remarks

We observe that the implicit constant in the estimate of Theorem 3 can easily be evaluated.

We note that when $t < p^{1/2}\log p$ the result of Theorem 3 is trivial. On the other hand, using the bounds of Theorem 3.4 and Theorem 5.5 of [11] for the linear congruential generator (1), one can obtain nontrivial results for sequences of period $t \geq p^{1/3+\varepsilon}$ and $t \geq p^{\varepsilon}$ for all and almost all primes p, respectively. Obtaining similar improvements of Theorem 3 is a challenging problem.

It is natural to ask whether curves with special point groups, which lead to bounds (7), (8) and (9), exist and are common enough. It follows from

Corollary 6.2 in [23] that the majority of (isomorphism classes of) elliptic curves, namely about 75%, have indeed a cyclic point group, which is necessary for (7) and (9). For (8) we need a cyclic point group of even order. Among the curves $y^2 = f(x)$ with $f(x) = x^3 + ax + b \in \mathbb{F}_p[x]$ about 50% have precisely one point of order 2, corresponding to those $f(x)$ which have precisely one root in $\mathbb{F}_p$ ([2, p. 37]). Thus at least 25% of all (isomorphism classes of) elliptic curves do have cyclic point group of even order (heuristically we expect this to be more close to 50%). We remark that by Theorem 2.1(i) of [23] every cyclic group of order within the *Hasse-Weil interval* $[p + 1 - 2\sqrt{p}, p + 1 + 2\sqrt{p}]$ can be realized as the point group of an elliptic curve over $\mathbb{F}_p$ ($p \geq 5$). Several more results about elliptic curves with cyclic point groups can be found in [23, 24].

For supersingular elliptic curves more accurate results have been obtained in [6, 7, 14] but that technique cannot be extended to other curves. Moreover, these curves are usually considered as cryptographically weak.

Finally, one can also consider similar problems in extension fields $\mathbb{F}_q$ over $\mathbb{F}_p$ and study the distribution of traces $\mathrm{Tr}_{\mathbb{F}_q/\mathbb{F}_p}(f_i(W_n))$. Unfortunately we do not know how to establish an analogue of Lemma 1 for linear combinations of such traces.

References

[1] P. Beelen and J. Doumen, 'Pseudorandom sequences from elliptic curves', *Finite Fields with Applications to Coding Theory, Cryptography and Related Areas*, Springer-Verlag, Berlin, 2002, 37–52.

[2] I. Blake, G. Seroussi and N. Smart, *Elliptic curves in cryptography*, London Math. Soc., Lecture Note Series, **265**, Cambridge Univ. Press, 1999.

[3] T. W. Cusic, C. Ding and A. Renvall, *Stream ciphers and number theory*, Elsevier, Amsterdam, 1998.

[4] E. El Mahassni and I. E. Shparlinski, 'On the uniformity of distribution of congruential generators over elliptic curves', *Proc. Intern. Conf. on Sequences and their Applications, Bergen 2001*, Springer-Verlag, London, 2002, 257–264.

[5] A. M. Frieze, J. Håstad, R. Kannan, J. C. Lagarias and A. Shamir, 'Reconstructing truncated integer variables satisfying linear congruence', *SIAM J. Comp.*, **17** (1988), 262–280.

[6] G. Gong, T. A. Berson and D. A. Stinson, 'Elliptic curve pseudorandom sequence generators', *Lect. Notes in Comp. Sci.*, Springer-Verlag, Berlin, **1758** (2000), 34–49.

[7] G. Gong and C. C. Y. Lam, 'Linear recursive sequences over elliptic curves', *Proc. Intern. Conf. on Sequences and their Applications, Bergen 2001*, Springer-Verlag, London, 2002, 182–196.

[8] S. Hallgren, 'Linear congruential generators over elliptic curves', *Preprint CS-94-143*, Dept. of Comp. Sci., Cornegie Mellon Univ., 1994, 1–10.

[9] A. Joux and J. Stern, 'Lattice reduction: A toolbox for the cryptanalyst', *J. Cryptology* **11** (1998), 161–185.

[10] H. Krawczyk, 'How to predict congruential generators', *J. Algorithms*, **13** (1992), 527–545.

[11] S. V. Konyagin and I. Shparlinski, *Character sums with exponential functions and their applications*, Cambridge Univ. Press, Cambridge, 1999.

[12] D. R. Kohel and I. E. Shparlinski, 'Exponential sums and group generators for elliptic curves over finite fields', *Lect. Notes in Comp. Sci.*, Springer-Verlag, Berlin, **1838** (2000), 395–404.

[13] J. C. Lagarias, 'Pseudorandom number generators in cryptography and number theory', *Proc. Symp. in Appl. Math.*, Amer. Math. Soc., Providence, RI, **42** (1990), 115–143.

[14] C. C. Y. Lam and G. Gong, 'Randomness of elliptic curve sequences', *Research Report CORR 2002-18*, Faculty of Math., Univ. Waterloo, Waterloo, 2002, 1–11.

[15] A. J. Menezes, P. C. van Oorschot and S. A. Vanstone, *Handbook of applied cryptography*, CRC Press, Boca Raton, FL, 1996.

[16] H. Niederreiter, 'Quasi-Monte Carlo methods and pseudo-random numbers', *Bull. Amer. Math. Soc.*, **84** (1978), 957–1041.

[17] H. Niederreiter, *Random number generation and quasi–Monte Carlo methods*, SIAM, Philadelphia, 1992.

[18] H. Niederreiter and M. Vielhaber, 'Linear complexity profiles: Hausdorff dimensions for almost perfect profiles and measures for general profiles', *J. Compl.*, **13** (1997), 353–383.

[19] R. A. Rueppel, *Analysis and design of stream ciphers*, Springer-Verlag, Berlin, 1986.

[20] I. E. Shparlinski, 'On the Naor–Reingold pseudo-random number function from elliptic curves', *Appl. Algebra in Engin., Commun. and Computing*, **11** (2000), 27–34.

[21] I. E. Shparlinski and J. H. Silverman, 'On the linear complexity of the Naor–Reingold pseudo-random function from elliptic curves', *Designs, Codes and Cryprography*, **24** (2001), 279–289.

[22] J. H. Silverman, *The arithmetic of elliptic curves*, Springer-Verlag, Berlin, 1995.

[23] S. G. Vlăduţ, 'Cyclicity statistics for elliptic curves over finite fields', *Finite Fields and Their Appl.*, **5** (1999), 13–25.

[24] S. G. Vlăduţ, 'A note on the cyclicity of elliptic curves over finite field extensions', *Finite Fields and Their Appl.*, **5** (1999), 354–363.