Random Points on Elliptic Curves

Igor E. Shparlinski

Centre for Advanced Computing:
Algorithms and Cryptography

Macquarie University

Joint work with:

Florian Hess (U. of Bristol)

and

Tanja Lange (U. of Bochum)

What do we want?

Question: Given an elliptic curve $\mathbb{E}$ over a finite field $\mathbb{F}_q$, find a **random** point $P \in \mathbb{E}$?

Subquestion: How is $\mathbb{E}$ given?
– Probably by a *Weierstraß equation*

$$y^2 = x^3 + Ax + B$$

Subquestion: What is "a random point"?
– We need some measure of the uniformity of distribution

Subquestion: Why do we need this at all?
– It is mathematically interesting
– Cryptographers may like mathematicians a little more

By-product: Better (= more secure) sequences of random bits

Assume that $x^3 + Ax + B$ has a root $\alpha \in \mathbb{F}_q$. Then $P = (\alpha, 0)$ is as random as any other point

...even if it is clearly not a *random point*

Moral: Random points do not exist, but sequences (P_n) of random points might.

We want to come up with a computer algorithm

$$\Downarrow$$

Forget **random**, settle for **pseudorandom**: sequences of points which

- are generated by some deterministic algorithm (and it better be fast ...)

- satisfy some properties which are natural for random points on $\mathbb{E}$ (chosen independently and uniformly)

Quality Control

Discrepancy:

For a sequence of N points in the unit cube $[0,1)\times[0,1)$

$$\Gamma = \Big\{\Big(\gamma_{1,n},\gamma_{2,n}\Big) \in [0,1)\times[0,1), \quad 1 \le n \le N\Big\}$$

its *discrepancy* $\Delta(\Gamma)$ is

$$\Delta(\Gamma) = \sup_{B\subseteq[0,1)^2} \left| \frac{T_\Gamma(B)}{N} - |B| \right|,$$

where $T_\Gamma(B)$ is the number of points of Γ which hit the box

$$B = [\alpha_1,\beta_1)\times[\alpha_2,\beta_2) \subseteq [0,1)\times[0,1)$$

and the supremum is taken over all such boxes.

Distribution Requirement:

Map points $P_n = (x_n, y_n) \in \mathbb{F}_q^2$ into $[0,1) \times [0,1)$ and request that the discrepancy is small.

If $q = p$ prime then there is a natural map

$$P_n \to \left(\frac{x_n}{p}, \frac{y_n}{p}\right)$$

(think $\mathbb{F}_p = \{0, 1, \ldots, p-1\}$).

If $q \neq p$ then ... well, we will worry about this later.

We may also wish to study multidimesional distributions of the s-tuples of points $P_n, \ldots, P_{n+s-1}$:

$$\left(\frac{x_n}{p}, \frac{y_n}{p}, \ldots, \frac{x_{n+s-1}}{p}, \frac{y_{n+s-1}}{p}\right)$$

Many, **but not all**, results extend to this case as well.

Main Facts

Let $\mathbb{E}$ be given by a *Weierstraß equation* over $\mathbb{F}_q$

$$y^2 = x^3 + Ax + B$$

(if $\gcd(q, 6) = 1$).

- Hasse–Weil bound: $|\#\mathbb{E}(\mathbb{F}_q) - q - 1| \le 2q^{1/2}$

- $\mathbb{E}(\mathbb{F}_q)$ is an Abelian group, we denote by $\oplus$ the group operation

Sequences

$\boxed{\text{Fix a point } G \in \mathbb{E}(\mathbb{F}_p) \text{ of order } t}$

- EC Linear Congruential Generator, **EC-LCG**:

  For the "initial value" $U_0 \in \mathbb{E}(\mathbb{F}_q)$, consider the sequence:

$$U_k = G \oplus U_{k-1} = kG \oplus U_0, \qquad k = 1, 2, \ldots .$$

- EC Power Generator, **EC-PG**:

  For an integer $e \geq 2$, consider the sequence (with $W_0 = G$),

$$W_k = eW_{k-1} = e^k G, \qquad k = 1, 2, \ldots ,$$

- EC Naor-Reingold Generator, **EC-NRG**:

  Given an integer vector $(a_1, \dots a_n)$, consider the sequence:

$$f_{\mathbf{a}}(x) = a_1^{x_1} \dots a_n^{x_n} G, \qquad x = 1, 2, \dots,$$

  where $x = x_1 \dots x_n$ is the bit representation of x, $0 \le x \le 2^n - 1$.

  Example: $n = 4$, $t = 19$ and $\mathbf{a} = (2, 5, 3, 4)$

$$f_{\mathbf{a}}(0) = 2^0 5^0 3^0 4^0 G = G,$$

$$f_{\mathbf{a}}(1) = 2^0 5^0 3^0 4^1 G = 4G,$$

$$f_{\mathbf{a}}(2) = 2^0 5^0 3^1 4^0 G = 3G,$$

$$f_{\mathbf{a}}(3) = 2^0 5^0 3^1 4^1 G = 12G,$$

$$\dots$$

$$f_{\mathbf{a}}(11) = 2^1 5^0 3^1 4^1 G = 24G = 5G,$$

$$\dots$$

$$f_{\mathbf{a}}(15) = 2^1 5^1 3^1 4^1 G = 120G = 6G,$$

$\boxed{\text{They all have analogues in the group } \mathbb{F}_p^*}$

Sean Hallgren, **1994**: **EC-LCG**

Also by
Gong, Berson, Stinson, **2001**:
Beelen, Doumen, **2002**:
El Mahassni, Hess, I.S., **2001-2003**:

Tanja Lange, I.S., **2003**: **EC-PG**

Bill Banks, Frances Griffin, Daniel Lieman, Joe Silverman, I.S., **1999-2001**: **EC-NRG**

Burt Kaliski, **1986**: **EC-PRNG** in a complexity theory sense

Results

Florian Hess, Tanja Lange, I.S., **2002–2003**:

Theorem:
*If the order t of G is at least $p^{1/2+\varepsilon}$ then all above sequences (**EC-LCG**, **EC-PG**, **EC-NRG**) are reasonable well distributed*

Conjecture:
The above sequences are very well distributed

Proof ingredients:

- Bounds of exponential sums

  David Kohel, I.S., **2000**:
  $$\sum_{P \in \mathcal{H}} \exp\left(2\pi i \frac{f(P)}{p}\right) = O(p^{1/2})$$
  for any subgroup $\mathcal{H} \in \mathbb{E}(\mathbb{F}_p)$ and any function f which is not constant on $\mathbb{E}$.

- Results about not vanishing some functions over $\mathbb{E}$

Getting more technical . . .

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

$$C = \left(c_{i,j}, \quad 1 \le i \le r,\ 1 \le j \le s\right)$$

over $\mathbb{F}_p$.

Let H be a place of degree d of $\mathbb{E}$ and let

$$\mathcal{F} = \{f_1, \ldots, f_r\} \tag{1}$$

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

$$(f_i)_\infty = (i + \delta)(H), \qquad 1 \le i \le r, \tag{2}$$

where

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

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

Let (U_k) be the EC linear congruential generator (i.e. $U_k = G \oplus U_{k-1} = kG \oplus U_0$).

For a matrix $C \in \mathcal{M}_{r,s}(\mathbb{F}_p)$, a set $\mathcal{F}$ of functions (1) and a generic point Q on $\mathbb{E}$ we consider the function

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

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

Getting less technical . . .

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

Underlying results — Example

Lemma:
For any matrix $C \in \mathcal{M}_{r,s}(\mathbb{F}_p)$ with $s \le t$ and any set $\mathcal{F}$ of functions satisfying (2), $\mathcal{L}_{C,\mathcal{F}}(Q)$ is not constant. The subgroup $\langle G \rangle$ contains at most $sd\rho$ zeros of $\mathcal{L}_{C,\mathcal{F}}(Q)$. If $H \in \langle G \rangle \oplus U_0$ it contains at most s poles of $\mathcal{L}_{C,\mathcal{F}}(Q)$, which are of the form $H \ominus U_j$ for $1 \le j \le s$, and no poles otherwise.

Why Will This Work?

Question: Is it typical for a point $G \in \mathbb{E}(\mathbb{F}_q)$ to be of order $t \geq q^{1/2+\varepsilon}$?

$\boxed{\text{Elementary Number Theory:}}$

$$\Downarrow$$

If the *exponent* $\mathbf{e}_q(\mathbb{E})$ (the size of the largest cyclic subgroup of $\mathbb{E}(\mathbb{F}_q)$ is large, then the order of a "typical" point $G \in \mathbb{E}(\mathbb{F}_q)$ is large as well.

Facts: $E(\mathbb{F}_q)$ is

- either cyclic, thus $\mathbf{e}_q(\mathbb{E}) = \#E(\mathbb{F}_q) \sim q$ — as good as it gets;

- or is isomorphic to a product of two cyclic groups $\mathbb{Z}/M \times \mathbb{Z}/L$ with $L|M$, thus
$$\mathbf{e}_q(\mathbb{E}) = M \geq \#E(\mathbb{F}_q)^{1/2} \sim q^{1/2}$$
. . . falls below the threshold $q^{1/2+\varepsilon}$.

Better Bounds?

Rene Schoof, **1991**:

If $\mathbb{E}$ (defined over $\mathbb{Q}$) has no complex multiplication then

$$\mathbf{e}_p(\mathbb{E}) \geq p^{1/2} \frac{\log p}{\log \log p}$$

. . . still below the threshold $p^{1/2+\varepsilon}$.

Bill Duke, **2003**:

For almost all primes p and **all** curves $\mathbb{E}$ over $\mathbb{F}_p$

$$\mathbf{e}_p(\mathbb{E}) \geq p^{3/4-\varepsilon}$$

. . . comfortably above the threshold!!

Even Better Bounds?

The bounds on the discrepancy attain their full strength when $G \in \mathbb{E}(\mathbb{F}_q)$ is of order close to q.

Question: Is it typical for $\mathbf{e}_q(\mathbb{E})$ to be close to q?

Bill Duke, **2003**:
For any curve $\mathbb{E}$ (defined over $\mathbb{Q}$) and almost all primes p

$$\mathbf{e}_p(\mathbb{E}) \geq p^{1-\varepsilon}$$

unconditionally if E has complex multiplication and under the **Extended Riemann Hypothesis**, otherwise.

I.S., **2003**:
For any prime p and almost all curves $\mathbb{E}$ (defined over $\mathbb{F}_p$)

$$\mathbf{e}_p(\mathbb{E}) \geq p^{1-\varepsilon}.$$

Florian Luca and I.S., **2004**:
For any curve $\mathbb{E}$ (defined over $\mathbb{F}_q$) and almost all integers n

$$\mathbf{e}_{q^n}(\mathbb{E}) \geq q^{(1-\varepsilon)n}$$

The proof of the last bound is based of some deep facts of the theory of Diophantine Approximations

- Subspace Theorem;
- Lower bounds of linear forms of p-adic logarithms;
- Upper bounds on the number of zeros of linear recurrence sequences.

The proof also gives an upper bound on

$$d(q^n) = \gcd(\#\mathbb{E}(\mathbb{F}_{q^n}), q^n - 1)$$

which appears in the estimate of the complexity of the structure finding algorithm of *Victor Miller,* **1984-2004**:

How do we proceed?

Facts:

- *Weil Pairing:* If $E(\mathbb{F}_p) \cong \mathbf{Z}/M \times \mathbf{Z}/L$ with $L|M$, then $L|q-1$.

- *Hendrik Lenstra,* **1987**: For any N, the probability that $\#\mathbb{E}(\mathbb{F}_q) = N$ for a random curve $\mathbb{E}$ is $O\left(q^{-1/2}\log q(\log\log q)^2\right)$.

  Thus all values of $N \in [q-2q^{1/2}, q+2q^{1/2}]$ are taken about the same number of times.

The question about $e_q(\mathbb{E})$ is now reduced to studying how often $N \in [q-2q^{1/2}, q+2q^{1/2}]$ has a large common divisor with $q-1$.

Is Cyclicity Typical?

- Fix the field — Vary the curve:

  Sergei Vlăduţ, **1999**:
  At least 75% of elliptic curves over $\mathbb{F}_q$ are cyclic, but **not** 100%

- Fix the curve: — Vary the field:

  Related to the **Lang–Trotter Conjecture!**

  Alina Cojocaru, Ram Murty, Bill Duke, **2001-2003**: (a series of results)

  – under the **Extended Riemann Hypothesis** the set of primes for which $\mathbb{E}(\mathbb{F}_p)$ is cyclic is of positive density (depending on $\mathbb{E}$);

  – the smallest prime for which $\mathbb{E}(\mathbb{F}_p)$ is cyclic is not too large.

Anything Else?

Linear Complexity

We do not want any embedded linear structures in our sequences.

This **vague** request can be made rigorous via the notion of the **linear complexity**.

Florian Hess, Tanja Lange, Joe Silverman, I.S., **1999–2003**:

Theorem:
*The above sequences (***EC-LCG***,* **EC-PG***,* **EC-NRG***) are reasonably good w.r.t. this property.*

Conjecture:
The above sequences are very good w.r.t. this property

Predictability

<u>**Scenario**</u>:

- The <u>User</u> generates a sequence of points W_n but shows to the word only some partial information (e.g., discards some bits of the components $x(W_n)$ and $y(W_n)$ of W_n).

- The <u>Attacker</u> knows
  - the general rule how the points W_n are generated
  - some partial information about the chosen parameters (such as the initial values, coefficients of the interatively applied tranformation, the underlying field, etc.)

  and wants after polynomially many observations of the "truncated" values of W_n, continue to generate the same sequence.

Donald Knuth, **1985**:
Alain Frieze, Johan Håstad, Ravi Kannan, Jeff Lagarias and Adi Shamir, **1988**:
Joan Boyar, **1989**:
Hugo Krawczyk, **1992**:

— Attacks on the linear congruential generator over a finite field.

Simon Blackburn, Domingo Gomez-Perez, Jaime Gutierrez and I.S., **2003**:

— Attacks on non-linear congruential generator over a finite field:

$$w_k \equiv f(w_{k-1}) \pmod p, \quad k = 1, 2 \ldots,$$

where

$$f(X) \in \mathbb{F}_p[X], \quad \text{(polynomial generator)}$$

or

$$f(X) = aX^{-1} + b, \quad \text{(inversive generator)}$$

Question: Study these attacks for EC-PRNG's

Different Approach

Jerome Solinas, Tanja Lange, **1999–2003**:

Koblitz curves over $\mathbb{F}_2$:

$$\mathbb{E}_a : Y^2 + XY = X^3 + aX^2 + 1, \qquad a = 0, 1.$$

σ = Frobenius endomorphism: $\sigma((x, y)) = (x^2, y^2)$

P = point on $\mathbb{E}_a(\mathbb{F}_{2^n})$ of prime order $\ell | \#\mathbb{E}_a(\mathbb{F}_{2^n})$.

$\mathcal{M}_k = \{(\mu_1, \ldots, \mu_k) \in \{0, \pm 1\}^k \mid \mu_j \mu_{j+1} = 0\}$

$$\#\mathcal{M}_k = \frac{4}{3} 2^k + O(1).$$

For $k \leq n$ we consider the points

$$P_{\mathbf{m}} = \sum_{j=0}^{k-1} \mu_j \sigma^j(P), \qquad \mathbf{m} = (\mu_0, \ldots, \mu_{k-1}) \in \mathcal{M}_k.$$

$\boxed{\textbf{No doublings!}}$

Question: Can we say anything about the distribution of $P_{\mathbf{m}}$?

Let $N_k(Q) = \#\{\mathbf{m} \in \mathcal{M}_k \mid P_{\mathbf{m}} = Q\}$.

If $\ell > (\#\mathbb{E}_a(\mathbb{F}_{2^n}))^{1/2}$ then $\sigma^j(P) \in \langle P \rangle$, thus

$$P_{\mathbf{m}} \in \langle P \rangle$$

Subquestion: Is it true that "typically" $N_k(Q) \sim \#\mathcal{M}_k/\ell$.

Tanja Lange, I.S., **2003**:

Theorem:
$N_k(Q) = O\left(\#\mathcal{M}_k/\ell\right)$

Proof. Let $1 \le r \le k$ and $2^r \le \ell/8$.

Assume that $N_k(Q) \ge \#\mathcal{M}_{k-r} + 1$ solutions.

Then there are two distinct vectors

$$\mathbf{m}_\nu = (\mu_{\nu,0}, \ldots, \mu_{\nu,k-1}) \in \mathcal{M}_k, \qquad \nu = 1, 2,$$

with $P_{\mathbf{m}_1} = P_{\mathbf{m}_2} = Q$ and such that $\mathbf{m}_1$ and $\mathbf{m}_2$ agree at the last $k - r$ components.

Therefore,

$$\mathcal{O} = P_{\mathbf{m}_1} - P_{\mathbf{m}_2} = \sum_{i=0}^{r-1} d_i \sigma^i(P),$$

where $d_i = m_{1,i} - m_{2,i}$ and thus $|d_i| \leq 2$, $i = 0, \ldots, r-1$.

The characteristic polynomial of $\mathbb{E}_a$ (and of σ) is

$$\chi_a(T) = T^2 + (-1)^a T + 2.$$

Assume that $\chi_a(T)$ is irreducible (the other case is similar . . . but not too interesting).

Hence $\chi_a(T)$ and

$$F(T) = \sum_{i=0}^{r-1} d_i T^i$$

have a common root in the algebraic closure of $\mathbb{F}_\ell$.

Therefore,

$$\operatorname{Res}(\chi_a, F) \equiv 0 \pmod{\ell}.$$

Let τ_1 and τ_2 be the roots of χ_a. Then

$$\operatorname{Res}(\chi_a, F) = F(\tau_1)F(\tau_2)$$

and $|\tau_1| = |\tau_2| = \sqrt{2}$.

Bounding $|F(\tau_j)|$, we get

$$|\operatorname{Res}(\chi_a, F)| < 2^{r+3} \le \ell.$$

Hence, $\operatorname{Res}(\chi_a, F) = 0$. – Easy to rule out by looking at the constant coefficient of χ_a. $\square$

Open Questions

- Better discrepancy bounds?

- Better exponent bounds? Is $p^{3/4}$ the "right" answer (if we want it for all curves over $\mathbb{F}_p$)?

- Low degree polynomial structures embedded in the sequences of points? (We do not want them!)

- Other natural and easy ways to generate sequences on elliptic curves? Any use of the Frobenius automorphism? Weil and Tate pairings?

- Sequences associated with other nice algebraic geometry structures? E.g. Jacobians?