Analysis of pseudo-random number generators based on elliptic curves

Tanja Lange

Igor Shparlinski

Technical University of Denmark

Denmark

Macquarie University

Australia

Overview

- Motivation
- Elliptic curves
- PRNGs from elliptic curves
  - Linear congruential generator
  - EC Power generator
  - Naor-Reingold generator
- Bounds for the power generator
- Koblitz curves

Motivation

- Pseudo-Random Number Generators are good to have
- Study helps to increase trust in EC-DLP
  bad performance with respect to some measure could be turned into attack.
- Some PRNGs related to

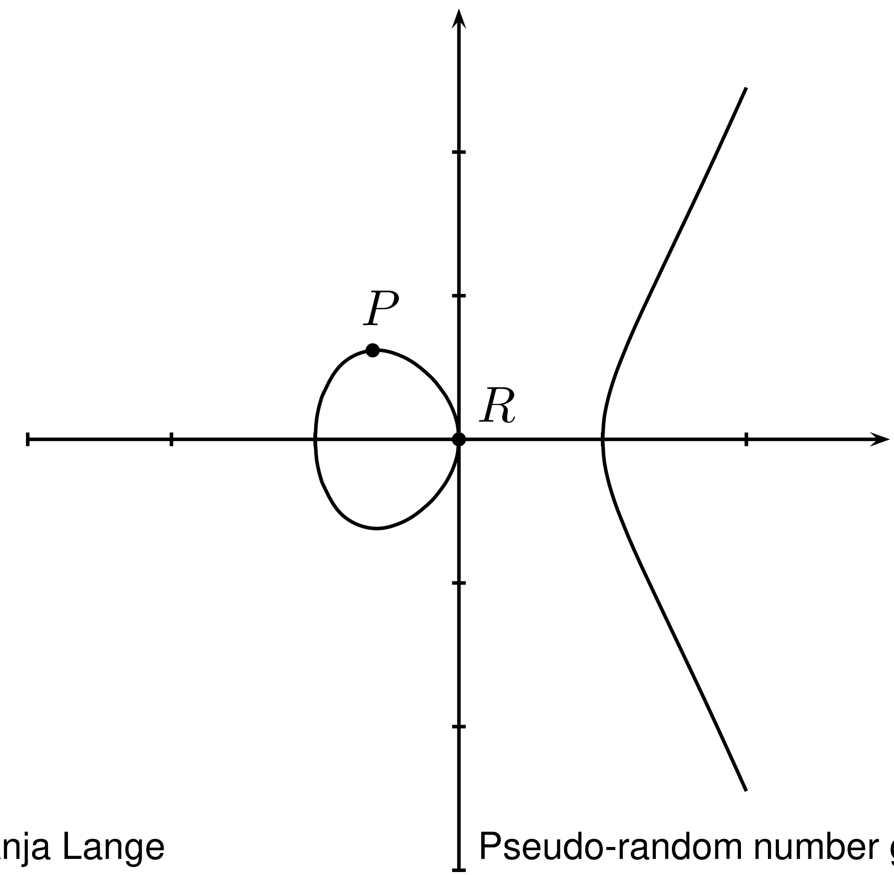
  Discrete Logarithm Problem (DLP)

  Given P and $Q = kP$ find k.
- Nice application of character sums, combinatorics, and curves :–)

Elliptic Curve

$$E : y^2 + \underbrace{(a_1 x + a_3)}_{h(x)}\, y = \underbrace{x^3 + a_2 x^2 + a_4 x + a_6}_{f(x)}, \;\; h, f \in \mathbb{F}_q[x]$$

Group: $E(\mathbb{F}_q) = \{\, (x, y) \in \mathbb{F}_q^2 \;:\; y^2 + h(x)y = f(x) \,\} \cup \{\, P_\infty \,\}$

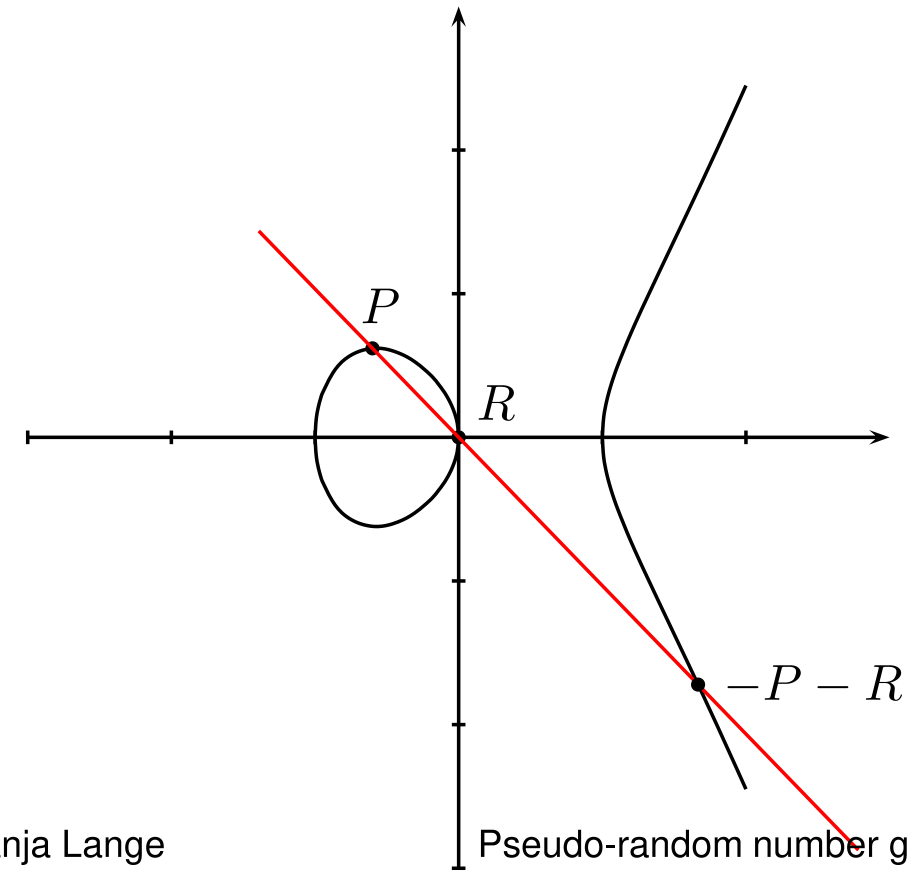


over $\mathbb{R}$

Elliptic Curve

$$E : y^2 + \underbrace{(a_1x + a_3)}_{h(x)}\, y = \underbrace{x^3 + a_2x^2 + a_4x + a_6}_{f(x)},\ \ h, f \in \mathbb{F}_q[x]$$

Group: $E(\mathbb{F}_q) = \{\,(x, y) \in \mathbb{F}_q^2 \,:\, y^2 + h(x)y = f(x)\,\} \cup \{\,P_\infty\,\}$

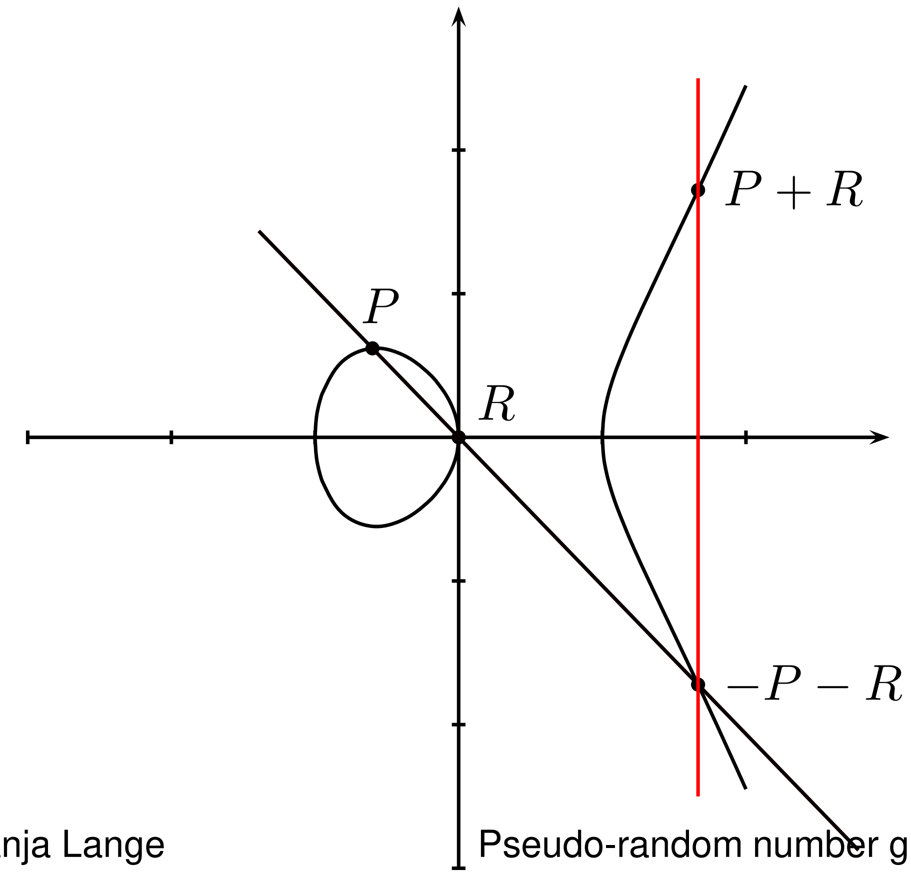


over $\mathbb{R}$

Elliptic Curve

$$E : y^2 + \underbrace{(a_1x + a_3)}_{h(x)}\, y = \underbrace{x^3 + a_2x^2 + a_4x + a_6}_{f(x)},\ \ h, f \in \mathbb{F}_q[x]$$

Group: $E(\mathbb{F}_q) = \{\, (x, y) \in \mathbb{F}_q^2 \,:\, y^2 + h(x)y = f(x) \,\} \cup \{\, P_\infty \,\}$

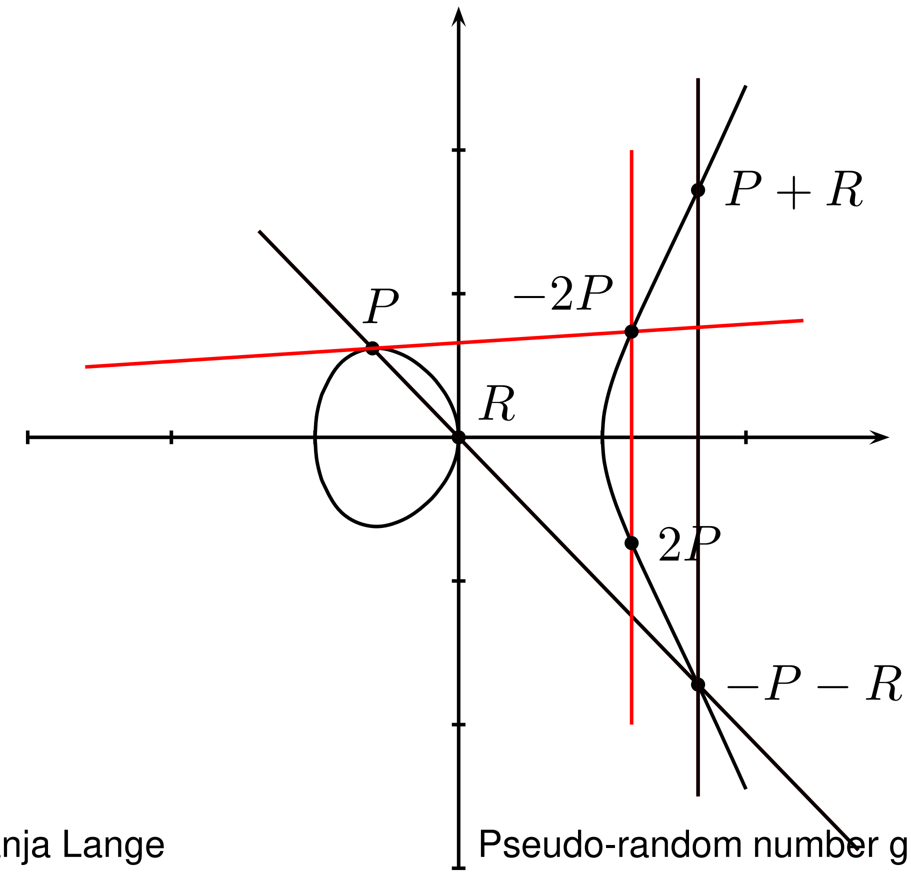


over $\mathbb{R}$

Elliptic Curve

$$E : y^2 + \underbrace{(a_1x + a_3)}_{h(x)}\, y = \underbrace{x^3 + a_2x^2 + a_4x + a_6}_{f(x)},\ \ h, f \in \mathbb{F}_q[x]$$

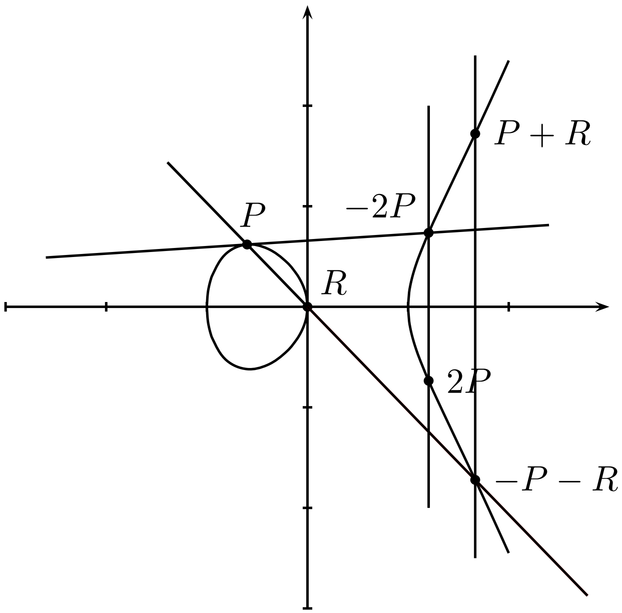
Group: $E(\mathbb{F}_q) = \{\, (x, y) \in \mathbb{F}_q^2 \ : \ y^2 + h(x)y = f(x)\,\} \cup \{\, P_\infty \,\}$



over $\mathbb{R}$

Group Law (q odd)

$$E: y^2 = x^3 + a_4 x + a_6,\ a_i \in \mathbb{F}_q$$



$$= (x_3, y_3) =$$
$$= (\lambda^2 - x_1 - x_2, \lambda(x_1 - x_3) - y_1),$$

where

$$\lambda = \begin{cases} (y_2 - y_1)/(x_2 - x_1) & \text{if } x_1 \neq x_2, \\ (3x_1^2 + a_4)/(2y_1) & \text{else.} \end{cases}$$

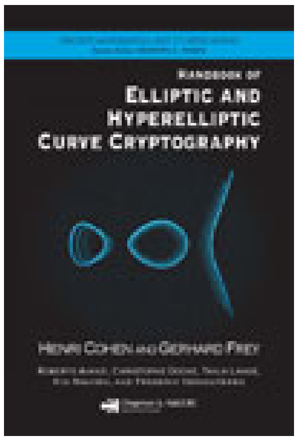
$$(x_1, y_1) + (x_2, y_2) =$$

$\Rightarrow$ consider sequence of points $P_n \in E(\mathbb{F}_q)$.

Let $P \in E(\mathbb{F}_q)$ be a point of order ℓ, often ℓ is assumed to be a prime.

Group Law (q odd)

$$E : y^2 = x^3 + a_4 x + a_6,\ a_i \in \mathbb{F}_q$$



$$= (x_3, y_3) =$$
$$= (\lambda^2 - x_1 - x_2, \lambda(x_1 - x_3) - y_1),$$

where

$$\lambda = \begin{cases} (y_2 - y_1)/(x_2 - x_1) & \text{if } x_1 \neq x_2, \\ (3x_1^2 + a_4)/(2y_1) & \text{else.} \end{cases}$$

$$(x_1, y_1) + (x_2, y_2) =$$

$\Rightarrow$ consider sequence of points $P_n \in E(\mathbb{F}_q)$.

Let $P \in E(\mathbb{F}_q)$ be a point of order ℓ, often ℓ is assumed to be a prime.

Sequences from points

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

Some applications use only x-coordinate or apply further maps to the coordinate values (e.g. hash functions or the trace map).

Elliptic curve PRNGs I

- EC Linear Congruential Generator, **EC-LCG**:
  For the "initial value" $Q_0 \in E(\mathbb{F}_q)$, consider the sequence:

  $$Q_k = P \oplus Q_{k-1} = kP \oplus Q_0, \qquad k = 1, 2, \dots .$$

  Easy to construct following element given two consecutive ones.
- Let $Q_k = (x_k, y_k)$. Use $(x_k)_{k=0}$, as sequence in $\mathbb{F}_q$ or normalize to $[0, 1)$ by using an enumeration of the field and dividing by q.
- Over $\mathbb{F}_{2^n}$ the sequence
  $(\mathrm{Tr}(x_0), \mathrm{Tr}(y_0), \mathrm{Tr}(x_1), \mathrm{Tr}(y_1), \mathrm{Tr}(x_2), \mathrm{Tr}(y_2), \mathrm{Tr}(x_3), \dots$
  was studied.
  Period length linked to the number of points on E.

Elliptic curve PRNGs II

- EC Power Generator, **EC-PG**: For an integer $e \geq 2$, consider the sequence (with $Q_0 = P$),

$$Q_k = eQ_{k-1} = e^k P, \qquad k = 1, 2, \ldots,$$

Determining e from Q_k, Q_{k-1} would be solving the DLP. Constructing the following sequence element given longer substrings related to generalized Diffie-Hellman problem.

Elliptic curve PRNGs III

- EC Naor-Reingold Generator, **EC-NRG**:
  Given an integer vector $(a_1, \ldots, a_n)$, consider the sequence:

$$Q_{\mathbf{a},k} = a_1^{k_1} \ldots a_n^{k_n} P, \qquad k = 1, 2, \ldots,$$

  where $k = k_1 \ldots k_n$ is the bit representation of k, $0 \leq k \leq 2^n - 1$.

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

$$f_{\mathbf{a},0} = 2^0 5^0 3^0 4^0 P = P, \qquad \ldots$$
$$f_{\mathbf{a},1} = 2^0 5^0 3^0 4^1 P = 4P, \quad f_{\mathbf{a},11} = 2^1 5^0 3^1 4^1 P = 24P = 5P$$
$$f_{\mathbf{a},2} = 2^0 5^0 3^1 4^0 P = 3P, \quad f_{\mathbf{a},15} = 2^1 5^1 3^1 4^1 P = 120P = 6P$$
$$f_{\mathbf{a},3} = 2^0 5^0 3^1 4^1 P = 12P$$

Character sums on elliptic curves

David Kohel, Igor Shparlinski, 2000:

Let $E/\mathbb{F}_p$ be an elliptic curve and $P \in H \subseteq E(\mathbb{F}_p)$ for some subgroup H. Then for any function f nonconstant on E we have

$$\sum_{P \in H} \exp\left(2\pi i \frac{f(P)}{p}\right) = O(p^{1/2})$$

Results

Florian Hess, T. L., Igor Shparlinski, 2002–2003:

> If the order ℓ of P is at least $p^{1/2+\varepsilon}$ then all above sequences (**EC-LCG**, **EC-PG**, **EC-NRG**) are reasonably well distributed.

E. g. EC power generator for $E/\mathbb{F}_q$ nonsupersingular, $q = p^r$, $T = \mathrm{ord}_\ell(e)$.

For any integer $\nu \geq 1$, the bound

$$\Delta_e(S^{\mathrm{PG}}) \ll T^{1-(3\nu+2)/2\nu(\nu+2)} \ell^{(\nu+1)/\nu(\nu+2)} q^{1/4(\nu+2)} (\ln p + 1)^r$$

holds for Δ the discrepancy of the sequence S^{PG} obtained from the power generator.

Schoenmakers & Sidorenko

Cryptanalysis of the Dual Elliptic Curve Generator, ePrint archive, 190/2006 (29th May).

- Analysis of a variant of the linear congruential generator.
- s_0 random, $Q = aP$, a secret
  $s_i = x(s_{i-1}P)$
  $r_i - \mathrm{lsb}_{240}(x(s_iQ))$
- r_i are not uniformly distributed – next bit is predictable without knowing a.
- Looks secure when less bits are extracted.

Koblitz curves over $\mathbb{F}_2$

Koblitz curves are subfield curves

$$E_a : Y^2 + XY = X^3 + aX^2 + 1, \qquad a = 0, 1.$$

Consider curve over extension field $\mathbb{F}_{2^n}$.

Frobenius automorphism φ of field extension extends to points: $\varphi((x, y)) = (x^2, y^2)$

$$P \in E(\mathbb{F}_{2^n}) \Rightarrow \varphi(P) \in E(\mathbb{F}_{2^n})$$

If $\ell > (\#E_a(\mathbb{F}_{2^n}))^{1/2}$ then $\varphi^j(P) \in \langle P \rangle$. Assume this case in the sequel.

Scalar multiplication

Neal Koblitz, Jerome Solinas 1992–1999:
Every integer m allows unique NAF (non-adjacent form) such that

$$mP = \sum \mu_i \varphi^i(P), \;\; \mu_i \in \{0, \pm 1\}, \mu_i \mu_{i-1} = 0.$$

Hence, obvious use in above method – just use expansion of e.
In cryptographic protocols often computation of mP for random m is required.
Faster:

choose random NAF $(\mu_0, \mu_1, \ldots, \mu_{k-1})$, $\mu_i \in \{0, \pm 1\}$, $\mu_i \mu_{i-1} = 0$ to save time to compute expansion.

General results on NAF

Bosma, 2001:

$$\mathcal{M}_k = \{(\mu_0, \ldots, \mu_{k-1}) \in \{0, \pm 1\}^k \mid \mu_j \mu_{j+1} = 0\}$$

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

Distribution properties I

For $k \le n$ we consider the points

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

Let $N_k(Q) = \#\{\mathbf{m} \in \mathcal{M}_k \mid P_{\mathbf{m}} = Q\}$.
For good distribution expect $N_k(Q) \sim \#\mathcal{M}_k/\ell$.

T. L., Igor Shparlinski, 2003:

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

Proof I

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

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

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 $|d_i| \le 2$, $i = 0, \dots, r-1$.

Proof II

The characteristic polynomial of 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,

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

Proof III

Let τ_1 and τ_2 be the roots of χ_a. Then

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

and $|\tau_1| = |\tau_2| = \sqrt{2}$.
Bounding $|F(\tau_j)|$, we get

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

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

Generalizations

Let σ be some endomorphism of E. Put

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

We consider the endomorphisms

- Doubling $\delta(P) = 2P$ on arbitrary curves E and fields $\mathbb{F}_q$.
- Frobenius endomorphism $\varphi(P) = (x^2, y^2)$ on the Koblitz curves $E_a/\mathbb{F}_{2^n}$.
- Complex multiplication $\psi(x,y) = \left(\frac{x^2-b}{b^2(x-c)}, \frac{y(x^2-2cx+b)}{b^3(x-c)^2}\right)$, where $b = (1+\sqrt{-7})/2$ and $c = (b-3)/4$, on
  $$E_{GLV} : y^2 = x^3 - \frac{3}{4}x^2 - 2x - 1$$
  over $\mathbb{F}_p$ with $p > 3$ and -7 quadratic residue modulo p.

Distribution properties II

The endomorphism ψ on E_{GLV} satisfies the same characteristic polynomial $T^2 - T + 2 = 0$ as the Frobenius endomorphism on E_1.

Thus L./Shparlinski 03 generalizes to:
Let $P \in E(\mathbb{F}_q)$ be of prime order ℓ. Then for any positive integer k and for every point $Q \in E(\mathbb{F}_q)$ the bound

$$N_{\sigma,k}(Q) \ll 2^k \ell^{-1} + 1$$

holds, where σ is one of the following endomorphisms:

- δ for an arbitrary curve E,
- φ for a Koblitz curve $E = E_a$, $a = 0, 1$,
- ψ for the GLV curve $E = E_{GLV}$.

Example: $\mathbb{F}_q = \mathbb{F}_{p^n}$ with $n > 1$

Fix basis $\omega_1, \ldots, \omega_n$ of $\mathbb{F}_{p^n}$ over $\mathbb{F}_p$ and a set $\mathcal{J} \subseteq \{1, \ldots, n\}$.

For fixed $\mathbf{c} = (c_j)_{j \in \mathcal{J}}$ with $c_j \in \mathbb{F}_p$ let $T_{\sigma,k}(\mathcal{J}, \mathbf{c})$ be the number of $\mathbf{m} \in \mathcal{M}_k$ for which

$$x(P_{\sigma,\mathbf{m}}) = \sum_{i=1}^{n} a_i \omega_i$$

and $a_j = c_j$ for every $j \in \mathcal{J}$.

I.e. $T_{\sigma,k}(\mathcal{J}, \mathbf{c})$ counts how often a fixed pattern is obtained in the x-coordinate of the points $P_{\sigma,\mathbf{m}}$.

Expect:

$$T_{\sigma,k}(\mathcal{J}, \mathbf{c}) \sim \frac{\#\mathcal{M}_k}{p^{\#\mathcal{J}}}.$$

Example: $\mathbb{F}_q = \mathbb{F}_{p^n}$ with $n > 1$

The bound

$$\max_{\mathcal{J},\mathbf{c}} \left| T_{\sigma,k}(\mathcal{J},\mathbf{c}) - \frac{\#\mathcal{M}_k}{p^{\#\mathcal{J}}} \right| \ll \#\mathcal{M}_k (p^{n/4\nu} \ell^{-1/2\nu} + 2^{-k/2\nu} p^{n(\nu+1)/4\nu^2})$$

holds with fixed integer $\nu \geq \frac{n \log p}{2k}$, where σ is one of the following endomorphisms:

- δ for an arbitrary curve E,
- φ for a Koblitz curve $E = E_a$, $a = 0, 1$.

Special case $\ell = q^{1+o(1)}$ and $k = \lfloor \log q \rfloor$:,
for any fixed $\alpha < 1/16$ the components on any $s \leq \alpha n$ positions are uniformly distributed.

The End!