Claus Diem

Institut für Experimentelle Mathematik
Essen

On the ECDLP over Extension Fields

I thank the organizers and
B. Edixhoven, G. Frey, S. Galbraith, P. Gaudry,
J. Scholten, N. Thériault and E. Viehweg.

Claim. There exists a randomized algorithm which takes as input a tuple $(q, n, E/\mathbb{F}_{q^n}, P, Q)$, where q is a prime power, n a natural number, $E/\mathbb{F}_{q^n}$ an elliptic curve and $P, Q \in E(\mathbb{F}_{q^n})$ with $Q \in \langle P \rangle$ which computes the DLP with respect to P and Q and has the following property:

Let us fix $a, b \in \mathbb{R}$ with $0 < a < b$ and let us consider all instances with

$$a \log_2(q) \leq n \leq b \log_2(q).$$

Then restricted to these instances, the algorithm has an expected running time of

$$\mathcal{O}\Big(2^{D \cdot (n \cdot \log_2(q))^{3/4}}\Big)$$

bit operations for $D = \frac{4b+\epsilon}{a^{3/4}}$.

Please note.

1. I do not have a complete proof of this statement.

2. The algorithm is not practical.

The algorithm is a variant of the index calculus algorithm presented by Gaudry. The main difference is that we increase the factor base.

Let $k := \mathbb{F}_q$, $K := \mathbb{F}_{q^n}$.

Recall the basic features of Gaudry's basic algorithm.

The *factor base* is the set of points in $E(K)$ whose x-coordinates lie in a certain 1-dimensional subspace K_1 of K. It has "roughly" q elements.

The *relations*

$$\alpha P + \beta Q = R_1 + \cdots R_n$$

are found by solving certain systems of polynomial equations over k. These systems have n equations of degree $n \cdot 2^{n-2}$ in n variables. "Usually", the algebraic set they define is 0-dimensional.

Let us *assume* that the homogenizations of these systems define 0-dimensional (proj.) algebraic sets.

The complexity of solving these systems is

$$\mathcal{O}\Big((n \cdot 2^{n-2} \cdot e)^{3n} \cdot \log_2(q)^2\Big).$$

The time for finding the relations can be estimated as

$$\mathcal{O}\Big((n \cdot 2^{n-2} \cdot e)^{3n} \cdot \log_2(q)^2 \cdot n! \cdot q\Big).$$

The time for linear algebra is

$$\mathcal{O}\Big(q^2 \cdot (\log_2(q) \cdot n)^2\Big).$$

Let us for simplicity work with a total running time of

$$\mathcal{O}\Big(2^{3n^2} \cdot q^2\Big) = \mathcal{O}\Big(2^{3n^2+2\log_2(q)}\Big).$$

Let us consider all instances with

$$n \leq b\sqrt{\log_2(q)}$$

for some fixed $b > 0$.

Then we have

$$\mathcal{O}\Big(2^{3b^2\log_2(q)+2\log_2(q)}\Big) = \mathcal{O}\Big(2^{(3b^2+2)\log_2(q)}\Big).$$

Let us now consider all instances with

$$a\sqrt{\log_2(q)} \le n \le b^2\sqrt{\log_2(q)}$$

for fixed $0 < a < b$.

Then

$$\log_2(q) = (\sqrt{\log_2(q)} \cdot \log_2(q))^{2/3} \le \left(\frac{n}{a}\log_2(q)\right)^{2/3}$$

The total running time is thus

$$\mathcal{O}\Big(\exp_2(\frac{3b^2+2+\epsilon}{a^{2/3}} \cdot (n \cdot \log_2(q))^{2/3})\Big).$$

For larger n, the complexity can be improved by increasing the factor base and decreasing the size of the systems.

Recall: The factor base had $\approx q$ elements, and we tried to find relations

$$\alpha P + \beta Q = R_1 + \cdots + R_n.$$

Let $c \in [1, \ldots, n]$ (to be determined later) and let $m := [\frac{n}{c}]$.

Let K_m be a randomly chosen m-dimensional k-vector subspace of K. Let the factor base be the set of points in $E(K)$ whose x-coordinates lie in K_m. Then the factor base contains roughly q^m elements.

We try to find relations

$$\alpha P + \beta Q = R_1 + \cdots + R_c.$$

(Note that $n - mc \in [0, \ldots, c-1]$, but this difference can be made 1 or even 0.)

One can find such relations by solving certain systems with n variables in $mc \leq n$ unknowns of degree $c \cdot 2^{c-2}$ over k. One can expect that "usually" these systems define a zero-dimensional algebraic set. Let us again assume that "usually" the homogenizations also define a 0-dimensional (proj.) algebraic sets.

Then the complexity to solve these systems is

$$\mathcal{O}\Big(2^{3nc} \cdot \log_2(q)^2\Big),$$

and time to find enough relations is more-or-less

$$\mathcal{O}\Big(2^{3nc} \cdot q \cdot q^{m+1}\Big) = \mathcal{O}\Big(2^{3nc+(m+2)\log_2(q)}\Big)$$

(which is also the total running time).

This is "approximately"

$$\mathcal{O}\Big(2^{3nc+(\frac{n}{c}+2)\log_2(q)}\Big).$$

Let us set $c := [\sqrt{\log_2(q)}]$. Then we get

$$\mathcal{O}\Big(2^{(4n\sqrt{\log_2(q)}+3\log_2(q))}\Big).$$

Let us assume that

$$n \leq b\log_2(q).$$

Then we obtain a running time of

$$\mathcal{O}\Big(2^{(4b+\epsilon)\log_2(q)^{3/2}}\Big).$$

Let us assume that additionally

$$a\log_2(q) \leq n.$$

Then

$$\log_2(q)^{3/2} = (\log_2(q) \cdot \log_2(q))^{3/4} \leq \Big(\frac{n}{a} \cdot \log_2(q)\Big)^{3/4}$$

This gives a total running time of

$$\mathcal{O}\Big(\exp_2(\frac{4b+\epsilon}{a^{3/4}} \cdot (n\log_2(q))^{3/4})\Big).$$

On the heuristics.

- One can prove that a factor base with $\geq \frac{1}{2}q^m$ elements can be constructed in polynomial time.

- Using a further variation of the algorithm, one can prove in a certain sense that the systems "usually" define 0-dimensional algebraic sets.

Major open questions and tasks.

- Do the *homogenizations* of the systems really define 0-dimensional algebraic sets?

- Assume that $mc < n$. Is it then true that "usually" if there is at least one solution to the systems in k, there is exactly one?

- Make the algorithm (more) practical by replacing the summation polynomials!