A survey on Cryptanalysis of Elliptic Curve Cryptography

Giordano Santilli

Università degli Studi di Trento

4 February 2020

Elliptic Curves

Weierstrass Equation

Let $\mathbb{K}$ be a field with characteristic different than 2 or 3 and A, $B \in \mathbb{K}$. A *Weierstrass Equation* f is an equation in $\mathbb{K}[x, y]$ of the form

$$f : y^2 = x^3 + Ax + B.$$

The quantity $\Delta = 4A^3 + 27B^2$ is called *discriminant* and if $\Delta \neq 0$, then f is said to be *non-singular*.

Elliptic Curves

Elliptic Curve

An *elliptic curve* E over $\mathbb{K}$ is the set of points $(x, y) \in \mathbb{K}^2$ which verifies a non-singular Weierstrass equation plus the *point at infinity* $\mathcal{O}$, that is

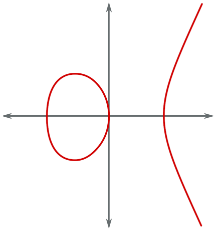
$$E(\mathbb{K}) = \left\{ (x, y) \in \mathbb{K}^2 \,\middle|\, y^2 = x^3 + Ax + B : A, B \in \mathbb{K} \right\} \cup \{\mathcal{O}\}.$$

Elliptic Curves

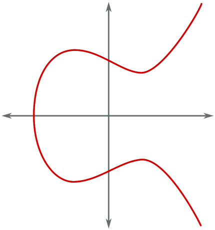
Elliptic Curve

An *elliptic curve* E over $\mathbb{K}$ is the set of points $(x, y) \in \mathbb{K}^2$ which verifies a non-singular Weierstrass equation plus the *point at infinity* $\mathcal{O}$, that is

$$E(\mathbb{K}) = \left\{ (x, y) \in \mathbb{K}^2 \,\middle|\, y^2 = x^3 + Ax + B : A, B \in \mathbb{K} \right\} \cup \{\mathcal{O}\}\,.$$

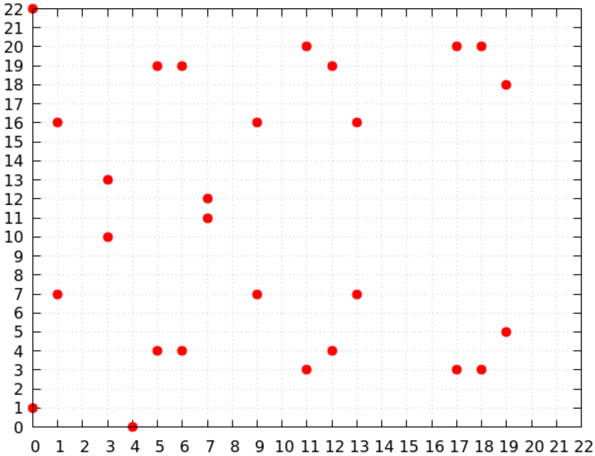


$y^2 = x^3 - x$



$y^2 = x^3 - x + 1$

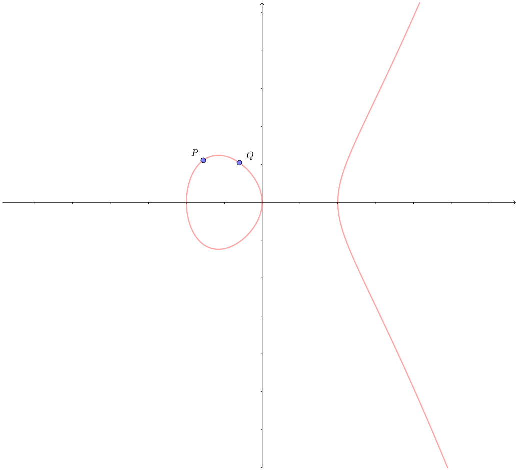
Elliptic Curves



$E : y^2 = x^3 + x + 1$ over the field $\mathbb{F}_{23}$

This curve has 28 points, including $\mathcal{O}$.

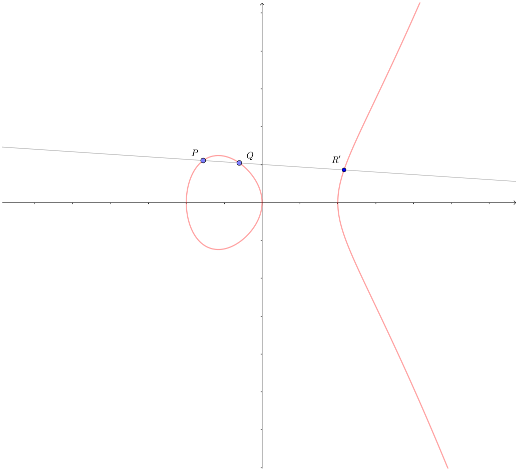
Group Structure in Elliptic Curves



Sum

Given two points P and Q over an elliptic curve E, the *sum* $P + Q$ is defined with the following algorithm:

Group Structure in Elliptic Curves

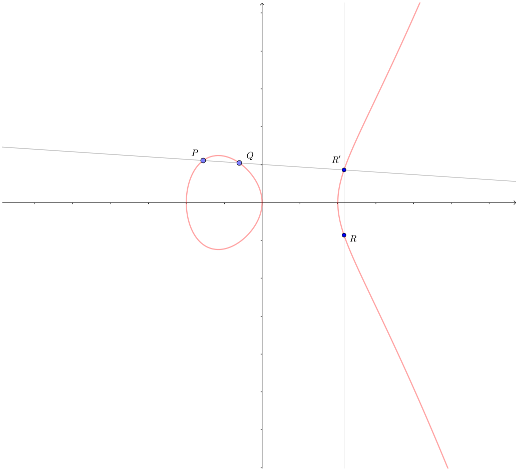


Sum

Given two points P and Q over an elliptic curve E, the *sum* $P + Q$ is defined with the following algorithm:

- Draw the line between P and Q. This line will intercept the curve E in a third point R'.

Group Structure in Elliptic Curves



Sum

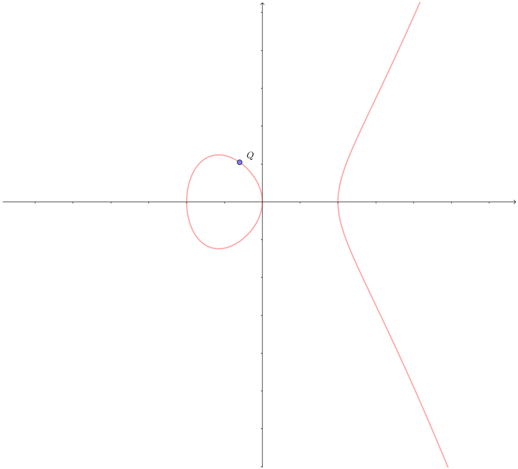
Given two points P and Q over an elliptic curve E, the *sum* $P + Q$ is defined with the following algorithm:

- Draw the line between P and Q. This line will intercept the curve E in a third point R'.
- Draw the symmetric point of R' with respect to the x-axis. This point is R and it is defined to be $R = P + Q$.

Group Structure in Elliptic Curves

Doubling

Given one point Q over an elliptic curve E, the *double* $2Q$ is defined with the following algorithm:

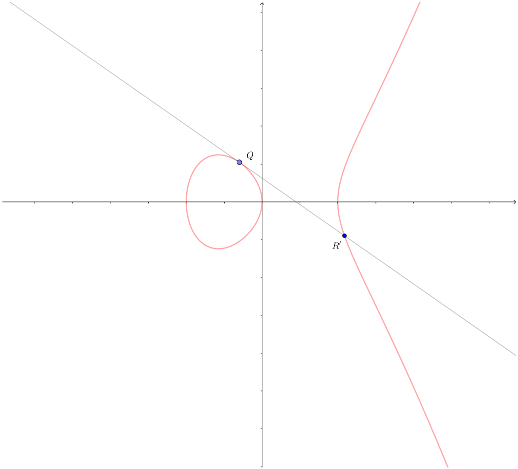


Group Structure in Elliptic Curves

Doubling

Given one point Q over an elliptic curve E, the *double* $2Q$ is defined with the following algorithm:

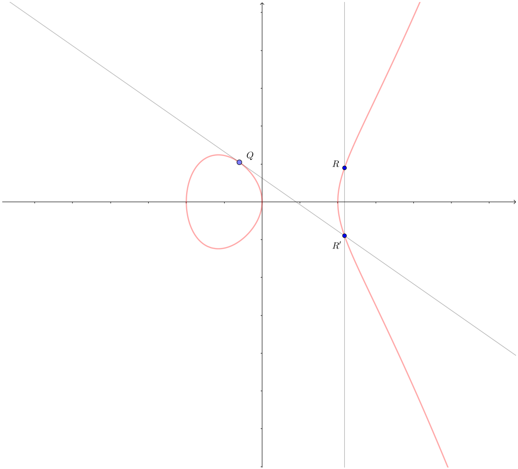
- Draw the tangent line through Q. This line will intercept the curve E in a second point R'.



Doubling

Given one point Q over an elliptic curve E, the *double* $2Q$ is defined with the following algorithm:

- Draw the tangent line through Q. This line will intercept the curve E in a second point R'.
- Draw the symmetric point of R' with respect to the x-axis. This point is R and it is defined to be $R = 2Q$.



Group Structure in Elliptic Curves

Theorem

Using this definition of point summation $(E, +)$ is an Abelian Group, i.e. given any $P, Q, R \in E$

- $(P + Q) + R = P + (Q + R)$ *(associativity)*,
- $P + \mathcal{O} = \mathcal{O} + P = P$ *(identity element)*,
- $-P \in E$ *such that* $P + (-P) = \mathcal{O}$ *(inverse element)*,
- $P + Q = Q + P$ *(commutativity)*.

Group Structure in Elliptic Curves

Theorem

Using this definition of point summation $(E, +)$ is an Abelian Group, i.e. given any $P, Q, R \in E$

- $(P + Q) + R = P + (Q + R)$ *(associativity)*,
- $P + \mathcal{O} = \mathcal{O} + P = P$ *(identity element)*,
- $-P \in E$ *such that* $P + (-P) = \mathcal{O}$ *(inverse element)*,
- $P + Q = Q + P$ *(commutativity)*.

Order

The *Order* of a point $P \in E$ is the smallest positive integer k such that

$$kP = \mathcal{O}.$$

Elliptic Curve Discrete Logarithm Problem (ECDLP)

ECDLP

Given an elliptic curve E defined over a finite field $\mathbb{F}_q$, a point $P \in E$ and another point Q which is a multiple of P, find $k \in \mathbb{N}^+$ such that $Q = kP$.

The number k is the *discrete logarithm* of Q to the base P and it is denoted as $\log_P Q = k$.

Given:

- $\mathbb{F}_q$, a finite field,
- E, an elliptic curve over $\mathbb{F}_q$,
- P, a base point of E (usually with big order),
- Q, a multiple of P.

$\longrightarrow$ Find $k \in \mathbb{N}^+$ such that $Q = k \cdot P$.

Public Parameters

- E, an elliptic curve over $\mathbb{F}_q$,
- P, a base point of E,
- N, the order of P.

Alice

- Generates a secret random number $a \in \mathbb{Z}_N$.
- Computes $\alpha = aP$ and sends it to Bob.
- Computes $\gamma = a\beta = abP$.

$$\xrightarrow{\alpha}$$
$$\xleftarrow[\beta]{}$$

Bob

- Generates a secret random number $b \in \mathbb{Z}_N$.
- Computes $\beta = bP$ and sends it to Alice.
- Computes $\gamma = b\alpha = baP$.

Suppose Bob wants to send a signed message to Alice.

Public Parameters

- E, an elliptic curve over $\mathbb{F}_q$,
- P, a base point of E,
- N, the order of P,
- m, the message to be signed,
- h, an hash function.

Key Generation - Bob

- Chooses a secret random number $k \in \mathbb{Z}_N$. This is the *secret key*.
- Computes $Q = kP = (x_1, x_2)$. This is the *public key*.

Suppose Bob wants to send a signed message to Alice.

Signature Generation - Bob

- Computes the hash of the message $e = h(m)$.
- Generates a random integer $t \in \mathbb{Z}_N$.
- Computes $r \equiv x_1 \bmod N$.
- Computes $s \equiv t^{-1}(e + rk) \bmod N$.
- The pair (r, s) is the *signature*.

Suppose Bob wants to send a signed message to Alice.

Signature Generation - Bob

- Computes the hash of the message $e = h(m)$.
- Generates a random integer $t \in \mathbb{Z}_N$.
- Computes $r \equiv x_1 \mod N$.
- Computes $s \equiv t^{-1}(e + rk) \mod N$.
- The pair (r, s) is the *signature*.

Signature Verification - Alice

- Computes the hash of the message $e = h(m)$.
- Computes $u \equiv es^{-1} \mod N$ and $v \equiv rs^{-1} \mod N$.
- Computes the point $(x_2, y_2) = uP + vQ$.
- If $r \equiv x_2 \mod N$, then the signature is *valid*.

Idea

If we fix $m \geq \lceil \sqrt{N} \rceil$, we can write $k = j_0 m + i_0$, with $0 \leq i_0, j_0 < m$. So $Q = kP = (j_0 m + i_0)P = j_0 mP + i_0 P$, therefore

$$Q - j_0 mP = i_0 P.$$

Idea

If we fix $m \geq \lceil \sqrt{N} \rceil$, we can write $k = j_0 m + i_0$, with $0 \leq i_0, j_0 < m$. So $Q = kP = (j_0 m + i_0)P = j_0 mP + i_0 P$, therefore

$$Q - j_0 mP = i_0 P.$$

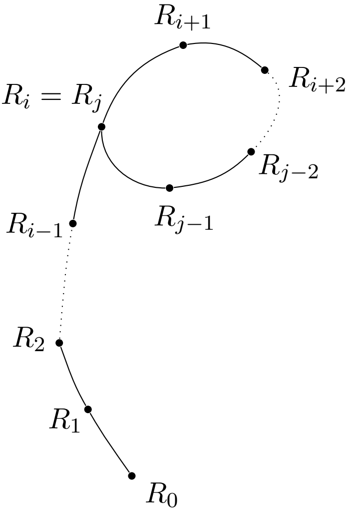
Algorithm - BSGS

- Fix an integer $m \geq \lceil \sqrt{N} \rceil$.
- Store a list of iP for $0 \leq i < m$ (*Baby Step*).
- Compute $Q - jmP$ for $0 \leq j < m$ until one of them matches an element of the stored list (*Giant Step*).
- If $i_0 P = Q - j_0 mP$, then $k \equiv i_0 + j_0 m \bmod N$.

Idea

Suppose there exists a pseudorandom function $f : E \to E$ and define an initial random point $R_0 = a_0 P + b_0 Q \in E$. It is possible to define the sequence $f(R_i) = R_{i+1}$.
Since E is a finite group, there exist $i < j$ such that $R_i = R_j$. Therefore, the period of the sequence R_i is a divisor of $j - i$.
So if $a_i P + b_i Q = R_i = R_j = a_j P + b_j Q$, then

$$k \equiv \frac{a_j - a_i}{b_i - b_j} \bmod N.$$



Algorithm - Pollard's ρ

- Define a random point $R_0 \in E$.
- Compute R_i and R_{2i} for $i = 1, 2, \ldots$.
- If $R_i = R_{2i}$, then $k = \gcd(N, i)$.

Remark

To compute the points for the algorithm, it is enough to store just one pair of the shape (R_i, R_{2i}) *at each iteration, since* $R_{i+1} = f(R_i)$ *and* $R_{2(i+1)} = f\left(f\left(R_{2i}\right)\right)$.

Idea

If N is a composite number of the shape

$$N = \prod_i q_i^{e_i},$$

with q_i prime numbers and $e_i \in \mathbb{N}^+$, it is possible to solve DLP for each $q_i^{e_i}$ and then combine the results together to find a solution for DLP modulo N.

For each q^e dividing N, k can be written as $k \equiv k_0 + k_1 q + \ldots + k_{e-1} q^{e-1} \bmod q^e$. The aim of the algorithm is therefore to recover k_0, k_1, ..., k_{e-1} for each q^e.

Semaev's Summation Polynomials

Summation Polynomials

Let $E(\mathbb{F}_q)$ be an elliptic curve. For any $n \geq 2$, the *n-th summation polynomial* $f_n(X_1, \ldots, X_n)$ is defined such that given x_1, x_2, ..., $x_n \in \overline{\mathbb{F}_q}$ (the algebraic closure of $\mathbb{F}_q$), then $f_n(x_1, x_2, \ldots, x_n) = 0$ if and only if there exist y_1, y_2, ..., $y_n \in \overline{\mathbb{F}_q}$ such that $(x_i, y_i) \in E(\overline{\mathbb{F}_q})$ and $(x_1, y_1) + (x_2, y_2) + \ldots + (x_n, y_n) = \mathcal{O}$.

Semaev's Summation Polynomials

Summation Polynomials

Let $E(\mathbb{F}_q)$ be an elliptic curve. For any $n \geq 2$, the *n-th summation polynomial* $f_n(X_1, \ldots, X_n)$ is defined such that given x_1, x_2, …, $x_n \in \overline{\mathbb{F}_q}$ (the algebraic closure of $\mathbb{F}_q$), then $f_n(x_1, x_2, \ldots, x_n) = 0$ if and only if there exist y_1, y_2, …, $y_n \in \overline{\mathbb{F}_q}$ such that $(x_i, y_i) \in E(\overline{\mathbb{F}_q})$ and $(x_1, y_1) + (x_2, y_2) + \ldots + (x_n, y_n) = \mathcal{O}$.

Theorem

- *The* 2*-nd summation polynomial is* $f_2(X_1, X_2) = X_1 - X_2$*.*
- *The* 3*-rd summation polynomial is* $f_3(X_1, X_2, X_3) = (X_1 - X_2)^2 X_3^2 - 2\left((X_1 + X_2)(X_1X_2 + A) + 2B\right) X_3 + \left((X_1X_2 - A)^2 - 4B(X_1 + X_2)\right)$*.*
- *For any* $n \geq 4$ *and* $n - 3 \geq k \geq 1$*, the* n*-th summation polynomial is* $f_n(X_1, X_2, \ldots, X_n) = \operatorname{Res}_X\left(f_{n-k}(X_1, \ldots, X_{n-k-1}, X), f_{k+2}(X_{n-k}, \ldots, X_n, X)\right)$*.*

Anomalous Curves

Anomalous Curve

An elliptic curve over $\mathbb{F}_p$ is called *anomalous* if $|E(\mathbb{F}_p)| = p$.

In 1998 Satoh, Araki and in 1999 Smart showed an algebraic attack to ECDLP over anomalous curves which involves the use of p-adic fields $\mathbb{Q}_p$.

Idea

Since p is a prime number, the group $E(\mathbb{F}_p)$ is isomorphic to $\mathbb{Z}_p$ and it is possible to define explicitly the isomorphism $\psi : E(\mathbb{F}_p) \to \mathbb{Z}_p$. The ECDLP over $E(\mathbb{F}_p)$ becomes: let $a, b \in \mathbb{Z}_p$ be such that b is a multiple of a modulo p. The problem is to find $k \equiv ba^{-1} \bmod p$, which is easy to compute using Euclid's extended algorithm.

In 1998 Semaev proved independently the same result, from a geometric point of view.

Addition Formulas

Let $P = (x_P, y_P)$ and $Q = (x_Q, y_Q)$ be two points on the elliptic curve $E : y^2 = x^3 + Ax + B$, then $P + Q = (\overline{x}, \overline{y})$, where

$$\overline{x} = \lambda^2 - x_P - x_Q \qquad \overline{y} = \lambda(x_P - \overline{x}) - y_P,$$

with

$$\lambda = \begin{cases} \frac{y_P - y_Q}{x_P - x_Q} & \text{if } P \neq Q, \\ \frac{3x_P^2 + A}{2y_P} & \text{if } P = Q. \end{cases}$$

Remark

The sum does NOT depend on B.

Idea

Suppose a protocol involves an elliptic curve $E : y^2 = x^3 + Ax + B$. Let $P = (x_P, y_P)$ be a point on another curve $E' : y^2 = x^3 + Ax + C$, with $C \neq B$, such that the order t of P is small on E' in order to compute easily the ECDLP instances in $< P >$. Then if P is used as base point for the protocol, the public key will be $Q = kP$, but since the ECDLP is easy, it is possible to recover $k \bmod t$.

Faulty implementation of the algorithm

Fault usage of random parameters

Suppose that in ECDSA t is fixed for each message. So given two messages m_1 and m_2, their corresponding sign is (r, s_1) and (r, s_2), where

$$\begin{cases} s_1 \equiv t^{-1}(h(m_1) + rk) \bmod N \\ s_2 \equiv t^{-1}(h(m_2) + rk) \bmod N. \end{cases}$$

Then, $t \equiv \frac{h(m_1)-h(m_2)}{s_1-s_2} \bmod N$ and from this it is possible to retrieve

$$k \equiv \frac{s_1 t - h(m_1)}{r} \bmod N.$$

Recap of the attacks

Attack	Complexity	Countermeasure
Brute-force	$O(N)$	Choose a large N.
Baby step-Giant step	$O(\sqrt{N})$ operations and memory usage	Choose a large N.
Pollard's ρ	$O(\sqrt{N})$	Choose a large N.
Silver-Pohlig-Hellman	If $N = \prod_i p_i^{e_i}$, $O\left(\sum_i e_i(\log N + \sqrt{p_i})\right)$	Choose a large and prime N.
Summation Polynomials	If E is defined over $\mathbb{F}_q$, then $O(q^2)$	Choose a large q.
Anomalous Curve Attack	If $\lvert E(\mathbb{F}_p)\rvert = p$, then $O(\log p)$	Choose a non-anomalous curve.
Chosen Input Point Attack	$O(\log^2 N)$	Check if $P \in E$.

THANK YOU
FOR THE ATTENTION!