Practical Invalid Curve Attack Using Quadratic Twist

Pierre Chrétien

February 2025

1 Introduction

The so called *Invalid curve attack* is a real threat for cryptographic protocols based on elliptic curves. The attack has first been presented in [2] and the use of twists was described in [4]. OpenPGP.js prior to 4.2.0 was found to be vulnerable[1]. The Node.js secp256k1-node allows bindings to the "Bitcoin curve" `secp256k1` and was found to be vulnerable[2] to small subgroup attacks. Bluetooth was proved to be vulnerable to a "Fixed Coordinate" variant [3]. Edwards model has also been examined in [5]. The SafeCurves website and the associated paper [1] point out as *"An ECC implementor can stop an invalid-curve attack by checking whether the input point Q satisfies the correct curve equation; [...] But this creates a conflict between simplicity and security. An implementation that does not include this check is simpler and more likely to be produced, and will pass typical functionality tests."*

The rest of the paper is organized as follows. Section 2 recalls the basics mathematical concepts used in the sequel, we recall basics facts about discrete logarithm problem (DLP) and twists of elliptic curves. Section 3 presents the general setting of the attack and ways to exploit poor implementation and weak curves. Section 4 is a complete walktrhought an example. This paper has an expository role.

2 Background Material

Notations : We will denote by $\mathbb{F}_q$ the finite field with $q = p^n$ elements where $p \geq 5$ and $n \in \mathbb{N} - \{0\}$. We will denote by $E/\mathbb{F}_q$ an elliptic curve defined over $\mathbb{F}_q$. The reader is assumed to be familiar with basic theory of elliptic curves.
Short Weierstrass equations. Since $p \geq 5$, every elliptic curve $E/\mathbb{F}_q$ may be written as

$$E : y^2 = x^3 + ax + b, \;\; a, b \in \mathbb{F}_q.$$

This is a so called *short Weierstrass form* of the curve E *defined over* $\mathbb{F}_q$.

Remark 1. 1. The condition $p \geq 5$ is not a restriction in our context since p will usually be a large prime.

2. A short Weierstrass form is not unique. This will be completed in the subsection about twists.

Automorphisms. Let $E_1/\mathbb{F}_q$ and $E_2/\mathbb{F}_q$ be elliptic curves. These curves may be seen over $\overline{\mathbb{F}_q}$ that is, the coefficients of their equation may be seen as lying in $\overline{\mathbb{F}_q}$ instead of in $\mathbb{F}_q$. Every geometric isomorphism of elliptic curve ϕ from $E_1/\overline{\mathbb{F}_q}$ to $E_2/\overline{\mathbb{F}_q}$ has an affine part of the form

$$\phi(x, y) = (u^2x + r, u^3y + su^2x + t). \tag{1}$$

for $u \in \overline{\mathbb{F}_q}^*$, $r, s, t \in \overline{\mathbb{F}_q}$. We will denote geometric isomorphism as $\phi/\overline{\mathbb{F}_q}$. The isomorphism ψ is said to be *defined over* $\mathbb{F}_q$ or *rational* if $u, r, s, t \in \mathbb{F}_q$, we will denote it by $\psi/\mathbb{F}_q$.

Proposition 1. Let $E_i/\mathbb{F}_q$, $i \in \{1; 2\}$ be elliptic curves given by short Weierstrass equations.

$$E_i : y^2 = x^3 + a_ix + b_i, \;\; a_i, b_i \in \mathbb{F}_q.$$

A geometric isomorphism ϕ froom E_1 to E_2 is of the form

$$\phi(x, y) = (u^2x, u^3y).$$

Proof. This is included as a first step to fully understand isomorphisms in the quadratic twist case.
Let $(x, y) \in E_1$ and ϕ as given by (1). Applying ϕ to the equation of E_1 and expanding yields

$$\begin{aligned}
&y^2 = x^3 + a_1x + b_1 \\
&\Leftrightarrow (u^3y + su^2x + t)^2 = (u^2x + r)^3 + a_1(u^2x + r) + b_1 \\
&\Leftrightarrow u^6y^2 + s^2u^4x^2 + t^2 + 2u^5sxy + 2u^3ty + 2tsu^2x = u^6x^3 + 3ru^4x^2 + 3r^2u^2x + r^3 + a_1u^2x + a_1r + b_1 (*)
\end{aligned}$$

[1] https://www.cve.org/CVERecord?id=CVE-2019-9155
[2] https://nvd.nist.gov/vuln/detail/CVE-2024-48930

Identifying coefficients of xy and y with those of $y^2 = x^3 + a_2x + b_2$ yields $s = 0, t = 0$ (recall that $u \neq 0$ and $p \neq 2$).

$$(*) \Leftrightarrow u^6y^2 = u^6x^3 + 3ru^4x^2 + 3r^2u^2x + r^3 + a_1u^2x + a_1r + b_1$$

Then, identifying the coefficient of x^2 with the short equation of E_2 yields $r = 0$ (here we use $p \neq 3$). Thus $\phi(x, y) = (u^2x, u^3y)$. We conclude with the following computations that will be used in the sequel.

$$\begin{aligned} &u^6y^2 = u^6x^3 + a_1u^2x + b_1 \\ &\Leftrightarrow y^2 = x^3 + \frac{a_1}{u^4}x + \frac{b_1}{u^6} \\ &\Leftrightarrow \frac{a_1}{u^4} = a_2, \ \ \frac{b_1}{u^6} = b_2 (**) \end{aligned}$$

□

Proposition 2. Let $E/\mathbb{F}_q$ be an elliptic curves given by short Weierstrass equation $E : y^2 = x^3 + ax + b, \ \ a, b \in \mathbb{F}_q$. The *$j$-invariant of* E is defined to be

$$j(E) = 1728\frac{4a^3}{4a^3 + 27b^2}.$$

Let $E_1/\mathbb{F}_q, E_2/\mathbb{F}_q$ be elliptic curves, there exists an isomorphism $\phi/\overline{\mathbb{F}_q}$ from E_1 to E_2 if and only if $j(E_1) = j(E_2)$.

Remark 2.
1. We insist that the j-invariant classifies **geometric** isomorphism classes of elliptic curves over $\mathbb{F}_q$.
2. Thanks to $p \geq 5$, E has a short equation and $j(E)$ a special form in this case. Thus $j(E) \notin \{0, 1728\} \Leftrightarrow a, b \in \mathbb{F}_q^*$.

2.1 Twists of Elliptic Curves

Twists. Non trivial twists of $E/\mathbb{F}_q$ are elliptic curves $E'/\mathbb{F}_q$ being isomorphic to E when viewed over $\overline{\mathbb{F}_q}$ but not isomorphic to E when viewed over $\mathbb{F}_q$. Let $E/\mathbb{F}_q$ be an elliptic curve in short Weierstrass equation $y^2 = x^3 + ax + b$. Recall that $q = p^n$ and $p \geq 5$, so it is possible to write such an equation for E.

Definition 1. Le $E/\mathbb{F}_q$ be an elliptic curve. A *twist of* E is an elliptic curve $E_t/\mathbb{F}_q$ such that there is a geometric isomorphism $\phi/\overline{\mathbb{F}_q}$ of elliptic curves $\phi : E \simeq E_t$. A twist E_t of E is *trivial* if there exists an isomorphism ψ of elliptic curve defined over $\mathbb{F}_q$.

Definition 2. Let $d \in \mathbb{F}_q^*$. The *twist E_d of E by d* is the elliptic curve given in short Weierstrass equation

$$E_d : y^2 = x^3 + d^2ax + d^3b.$$

Remark 3. We did not specify that E_d is a non trivial. Actually, let δ be a square root of d in $\overline{\mathbb{F}_q}$ i.e. $\delta^2 = d$, then

$$\begin{aligned} &\phi : E \to E_d \\ &(x, y) \mapsto (\frac{x}{d}, \frac{y}{d\delta}) \end{aligned}$$

is a geometric isomorphism from E to E_d. It matches the relations (**) concluding proof of Proposition 1 with $a_1 = a, b_1 = b, a_2 = ad^2, b_2 = bd^3$ and $d = \frac{1}{u}$.

Proposition 3. Assume that $j(E) \neq 0, 1728$. The twist E_d is trivial if and only if $d \in (\mathbb{F}_q^*)^2$.

Proof. ($\Rightarrow$) Assume that there exists a rational isomorphism ψ from E to E_d. According to Proposition 1

$$\exists u \in \mathbb{F}_q^*, \ \psi(x, y) = (u^2x, u^3y)$$

According to (**), $\frac{a}{u^4} = ad^2$ and $\frac{b}{u^6} = bd^3$. Recall that since $p \geq 5$, the assumption about $j(E)$ is equivalent to $a, b \neq 0$. Thus $\frac{1}{u^4} = d^2$, $\frac{1}{u^6} = d^3$ and $d = \frac{d^3}{d^2} = \frac{1}{u^2} \in (\mathbb{F}_q^*)^2$.

($\Leftarrow$) Conversely, let $\delta \in \mathbb{F}_q^*$ such that $\delta^2 = d$. Then a rational isomorphism from E to E_d is

$$\psi(x, y) = (\frac{x}{d}, \frac{y}{d\delta})$$

□

Proposition 4. Assume that $E/\mathbb{F}_q$ has $j(E) \neq 0, 1728$. Then a twist $E_t/\mathbb{F}_q$ of $E/\mathbb{F}_q$ is either trivial or E_d for some $d \in (\mathbb{F}_q^*) \backslash (\mathbb{F}_q^*)^2$.

Proof. Assume that $E_t/\mathbb{F}_q$ is a non trivial twist of $E/\mathbb{F}_q$ with isomorphism $\phi : E \to E_t$ given by $\phi(x, y) = (u^2x, u^3y)$, $u \in \overline{\mathbb{F}_q}$, $u \notin \mathbb{F}_q$. Let $E_t : y^2 = x^3 + a_tx + b_t$, $a_t, b_t \in \mathbb{F}_q$, thus (**) yields

$$a_t = \frac{a}{u^4}, b_t = \frac{b}{u^6}$$

Then $u^2 = \frac{ba_t}{ab_t} \in \mathbb{F}_q$, i.e. $u \notin \mathbb{F}_q$ but $u^2 \in \mathbb{F}_q$. This means that $u \in \mathbb{F}_{q^2} \backslash \mathbb{F}_q$. Let $d := \frac{1}{u^2}$, then $a_t = d^2a, b_t = d^3b$ and $\phi(x, y) = \left(\frac{x}{d}, \frac{uy}{d}\right)$. □

Remark 4. Proposition 3 is wrong if $p < 5$. For example, let $p = q = 3$, $E : y^2 = x^3 + x + 1$ and $d = 2$. In this case, E is isomorphic to $E_d : y^2 = x^3 + x + 2$ with ψ given by constants $(u, r, s, t) = (2, 2, 0, 0)$. So E_d is a trivial twist of E but d is non square modulo 3. Note that the definition for the j-invariant when $p = 3$ gives here $j(E) = 0$.

Order of group of rational points. The group of rational points $E(\mathbb{F}_q)$ has order $\sharp E(\mathbb{F}_q) = q + 1 - t$ where t is the *Trace of Frobenius*. An extensive description of the Frobenius endomorphism is out scope for this paper, we only need some basics facts we recall below.

Proposition 5. 1. **(Hasse Bound)** One has $|t| \leq 2\sqrt{q}$.

2. One has $\sharp E_d(\mathbb{F}_q) = q + 1 + t$, thus $|\sharp E_d(\mathbb{F}_q) - \sharp E(\mathbb{F}_q)| = |2t| \leq 4\sqrt{q}$.

Proof. 1. The proof of Hasse Bound is technical, the interessted reader may refer to V Theorem 1.1 from [6]

2. We include this proof to give some insight into how points distribute over $E(\mathbb{F}_q)$, $E_d(\mathbb{F}_q)$. We will prove that

$$\sharp E(\mathbb{F}_q) + \sharp E_d(\mathbb{F}_q) = 2q + 2 \quad (\dagger)$$

substituting $\sharp E(\mathbb{F}_q) = q + 1 - t$ gives the result.

Let $f(x) = x^3 + ax + b$, $d \in \mathbb{F}_q^*$ be non square in $\mathbb{F}_q$ and $E_d : Y^2 = X^3 + d^2aX + d^3b$. The following change of variables on E_d (take care, this is not an isomorphism of E_d)

$$X = dx, Y = dy$$

yields $Y^2 = X^3 + d^2aX + d^3b \Leftrightarrow d^2y^2 = d^3x^3 + d^3ax + d^3b \Leftrightarrow y^2 = df(x)$.

- Let $x \in \mathbb{F}_q$ such that $f(x) = 0$, then $(x, 0) \in E(\mathbb{F}_q)$ and $(dx, 0) \in E_d(\mathbb{F}_q)$. Each curve gets one point.
- Let $x \in \mathbb{F}_q$ such that $f(x) \in (\mathbb{F}_q^*)^2$. Then $y^2 = f(x)$ has two solutions. Since d is non square, $df(x)$ is non square and $y^2 = df(x)$ has no solution in $\mathbb{F}_q$. So E gets two points and E_d zero.
- Let $x \in \mathbb{F}_q$ such that $f(x) \notin (\mathbb{F}_q^*)^2$. Then $df(x) \in (\mathbb{F}_q^*)^2$ and $y^2 = df(x)$ has two solution giving rise to two points on E_d. Here, E gets zero point.
- Count the point at infinty once for each curve.

Thus each $x \in \mathbb{F}_q$ contributes for 2 points in $E(\mathbb{F}_q) \cup E_d(\mathbb{F}_q)$ plus the points at infinity, yielding (†).

□

2.2 Discrete Logarithm Problem

Definition 3. Let G be a group in multiplicative notation. The **Discrete Logarithm Problem** (DLP) is : given $b, h \in G$ find $a \in \mathbb{Z}$ such that $h = b^a$.

Remark 5. 1. The group law on an elliptic curve being usually written in additive notation, DLP for elliptic curves is rephrased as : given $P, B \in E(\mathbb{F}_q)$ find $a \in \mathbb{Z}$ such that $P = aB$.

2. The DLP has a solution if and only if P is in the subgroup $<B>$ generated by B.

The following discussion gives the necessaries notions when dealing with generic methods to solve the DLP that will be used in the last section. The generic `discrete_log` method from Sage uses a combination of Pohlig-Hellman, Baby Step Giant Step (BSGS) , Pollard's kangaroo (i.e. Pollard's Lambda), and Pollard's Rho.

- The Pohlig-Helmann method. Let n be the order of a point $B \in E(\mathbb{F}_q)$ and $n = \prod_{i=1}^{m} p_i^{n_i}$ be its prime factorization. The subgroup H generated by B is cyclic of order n, thus has a unique cyclic subgroup H_i of order $p_i^{n_i}$ for each $i \in [|1; m|]$. By means of the Chinese Remainder Theorem (CRT), solving the DLP in H boils down to solve it in each H_i. The subgroups H_i have order a prime power, which may still be quite large. One can reduce the DLP from H_i to subgroups $\widetilde{H}_i$ of order **exactly** p_i. We restrict ourselves to solving the DLP in the $\widetilde{H}_i$'s.

- BSGS method is a collision finding algorithm to solve DLP that requires $\mathcal{O}(\sqrt{p_i})$ running time and $\mathcal{O}(\sqrt{p_i})$ storage. This gives an bound on the size of the p_i's for which we can hope to solve the DLP with this method.

- Pollard's Rho (resp. Lambda) algorithm is solving the DLP and has $\mathcal{O}(\sqrt{p_i})$ (resp. $\mathcal{O}(\sqrt{p_i})$) time complexity but $\mathcal{O}(1)$) (resp. $\mathcal{O}(\ln p_i)$) space complexity.

3 Invalid Curve Attack

3.1 General Setting

Invalid Cuve Attack is presented in [4]. Let $E/\mathbb{F}_q : y^2 = x^3 + ax + b$ be the public curve of the cryptosystem.

Main idea : The vulnerabilty lies in the fact that a malicious user may send a point that is **not** on $E/\mathbb{F}_q$ but on a weaker curve, say a quadratic twist, and then exploit it to solve an easier DLP. This is mainly due to b not beeing used for scalar multiplication on E.

1. Mallory may send an honest $P = (x, y)$ point on E and gets back some $\mathtt{k}.P$ from Bob. But Mallory may also send a malicious point $Q = (\tilde{x}, \tilde{y})$ on another curve $\tilde{E} : y^2 = x^3 + ax + \tilde{b}$ (note that the a coefficient remains unchanged). Since formulae for computing $\mathtt{k}.Q$ do not involve $b, \tilde{b}$, Bob will correctly compute $\mathtt{k}.Q$, believing he just made a computation on E.

- DLP peut etre trop dur sur E à cause de d'un ordre pas assez smooth.
- Si on peut faire calculer k*T pour T sur une autre courbe E' on peut trouver T modulo les premiers de l'ordre de E'.
- si le twist a un ordre avec d'autres facteurs premier que E alors on connait k modulo de nouveaux premiers.
- cela peut suffire à retrouver k (CRT)
- Expliquer que connaitre d modulo "suffisament" de premiers peut suffire, pas modulo "tous" les premiers.
- apport de montgomery qui limite le choix de E au twist (according to ...)

Remark 6. Even if $\sharp E(\mathbb{F}_q) = hp$ with p prime and h a small cofactor (a classical situation in cryptographic applications) then $\sharp E_d(\mathbb{F}_q)$ might have a prime factorization with many small primes (such numbers are called *smooth*). For example **brainpoolP256t1** curve has prime order but its quadratic twist E_t has a somehow smooth order since it factors in a product of 7 primes of which 6 have bit length less than 42 (the last prime factor has bit length 89).

4 Walkthrough example

References

[1] Daniel J. Bernstein and Tanja Lange. Safe curves for elliptic-curve cryptography. Cryptology ePrint Archive, Paper 2024/1265, 2024.

[2] Ingrid Biehl, Bernd Meyer, and Volker Müller. Differential fault attacks on elliptic curve cryptosystems. In Mihir Bellare, editor, *Advances in Cryptology — CRYPTO 2000*, pages 131–146, Berlin, Heidelberg, 2000. Springer Berlin Heidelberg.

[3] Eli Biham and Lior Neumann. Breaking the bluetooth pairing – the fixed coordinate invalid curve attack. Cryptology ePrint Archive, Paper 2019/1043, 2019.

[4] P.-A. Fouque, R. Lercier, D. Réal, and F. Valette. Fault Attack on Elliptic Curve with Montgomery Ladder Implementation. In *FDTC '08. 5th Workshop on Fault Diagnosis and Tolerance in Cryptography*, pages 92–98. IEEE-CS Press, August 2008.

[5] Samuel Neves and Mehdi Tibouchi. Degenerate curve attacks. Cryptology ePrint Archive, Paper 2015/1233, 2015.

[6] Joseph H Silverman. *The Arithmetic of Elliptic Curves*. Graduate texts in mathematics. Springer, Dordrecht, 2009.