How to hash onto elliptic curves

Dimitri Koshelev

Parallel Computation Laboratory, École Normale Supérieure de Lyon, France

20/02/2023, CIRM, Luminy

Introduction

Let $\mathbb{F}_q$ be a finite field and $E = E_{a,b}$ be an elliptic $\mathbb{F}_q$-curve $y^2 = f(x) := x^3 + ax + b$.

Suppose there is a subgroup $G \subset E(\mathbb{F}_q)$ of a large prime order $r \mid N := \#E(\mathbb{F}_q)$. Besides, denote by $c := N/r$ the cofactor.

Many protocols of elliptic cryptography use a hash function $\mathcal{H} \colon \{0,1\}^* \to G$, namely signatures (e.g., Boneh–Lynn–Shacham), PAKE (Password-Authenticated Key Exchange) protocols, OPRFs (Oblivious Pseudorandom Functions), identity-based cryptography.

The vast majority of them assume that $\mathcal{H}$ is *indifferentiable from a random oracle* in the sense of Maurer et al.

A naive method consists in the scalar multiplication $m \mapsto [m]P$ for a fixed non-zero point $P \in G$. However, it is insecure due to dependency on the group structure of E.

Components of $\mathcal{H}$

Almost all known hash functions to G are the compositions $\mathcal{H} = [c] \circ h \circ \eta$.

Here $\eta\colon \{0,1\}^* \to S$ is a hash function to some finite set and $h\colon S \to E(\mathbb{F}_q)$ is just a map commonly called *encoding*.

The scalar multiplication $[c]$ on E is said to be *clearing cofactor*.

The set S is usually very simple, hence it is easy to combine η from existing hash functions $\{0,1\}^* \to \{0,1\}^\ell$ for $\ell \in \mathbb{N}$.

The most complicated component of $\mathcal{H}$ is no doubt h, because it is based on high-dimensional algebraic geometry.

Computation of h often has to be *constant-time* to guarantee protection against *timing attacks*.

In particular, iterating $x \in \mathbb{F}_q$ as long as $\sqrt{f(x)} \notin \mathbb{F}_q$ is unsafe.

Simplified Shallue–van de Woestijne–Ulas encoding

It consists in a parametrization $\varphi\colon \mathbb{A}^1_t \to C$ of a (possibly singular) rational $\mathbb{F}_q$-curve C lying on the algebraic surface

$$y^2 = df(x_1)f(x_2) \quad \subset \quad \mathbb{A}^3_{(x_1,x_2,y)},$$

where $d \in (\mathbb{F}_q)^* \setminus (\mathbb{F}_q^*)^2$.

Given $t \in \mathbb{F}_q$, for exactly one coordinate $a_i := x_i\big(\varphi(t)\big)$ the value $f_i := f(a_i)$ is a quadratic residue in $\mathbb{F}_q$.

Therefore, we get the point $P = \left(a_i, \sqrt{f_i}\right) \in E(\mathbb{F}_q)$.

The encoding requires extracting the square root in $\mathbb{F}_q$.

As is known, at least for $q \not\equiv 1 \pmod 8$ this is equivalent to raising f_i to some fixed power $n \in \mathbb{Z}/(q-1)$.

The denominators of P do not need to be inverted, because (weighted) projective coordinates on E are often preferred.

Statistical notions and statements

A map $f : S \to T$ between finite sets is ϵ-*regular* (for $\epsilon \in \mathbb{R}_{>0}$) if

$$\sum_{P \in T} \left| \frac{\#f^{-1}(P)}{\#S} - \frac{1}{\#T} \right| \leqslant \epsilon.$$

Further, f is ϵ-*samplable* whenever exists a randomized polynomial algorithm that for any $t \in T$ induces an ϵ-regular map $R \to f^{-1}(t)$, where R is a set of random values.

Finally, f is called ϵ-*admissible* if it is efficiently computable in constant time, ϵ-regular, and ϵ-samplable.

Theorem (Brier et al.)

Let $f : S \to T$ be an ϵ-admissible map for a negligible ϵ and $\eta : \{0,1\}^* \to S$ be an indifferentiable hash function. Then the composition $f \circ \eta$ is also indifferentiable.

A map $h\colon S \to G$ to a finite abelian group is called *B-well-distributed* (where $B \in \mathbb{R}_{>0}$) if for any non-trivial character $\chi\colon G \to \mathbb{C}^*$ we have

$$\left|\sum_{s\in S} \chi\big(h(s)\big)\right| \leqslant B.$$

Theorem (Tibouchi)

Let $h\colon S \to G$ and $g\colon T \to G$ be finite maps such that h is B-well-distributed and $\#g^{-1}(P) \leqslant n \in \mathbb{N}$ for all $P \in G$. Then the map

$$h \otimes g\colon S \times T \to G \qquad (s,t) \mapsto h(s) + g(t)$$

is ϵ-regular for $\epsilon = \dfrac{B}{\#S}\sqrt{n\dfrac{\#G}{\#T}}$.

For an encoding $h\colon \mathbb{F}_q \to E(\mathbb{F}_q)$ the image $\mathrm{Im}(h)$ is only a fraction of $\mathbb{F}_q$-points on E. So, h is not regular.

Instead, the square $h^{\otimes 2}$ is often considered. However, in this case, the running time is doubled.

Encodings of Skałba and Chávez-Saab et al.

Given a curve $E\colon y^2 = f(x)$, the threefold

$$T\colon y^2 = f(x_1)f(x_2)f(x_3) \quad \subset \quad \mathbb{A}^4_{(x_1,x_2,x_3,y)}$$

is at the core of the encodings.

We have the auxiliary map

$$h'\colon T(\mathbb{F}_q) \to E(\mathbb{F}_q) \qquad h' := \left(x_i, \sqrt{f(x_i)}\right) \quad \text{if} \quad \left(\frac{f(x_i)}{q}\right) \in \{0,1\}.$$

Skałba's encoding $h := h' \circ \varphi\colon \mathbb{F}_q^2 \to E(\mathbb{F}_q)$ is based on an $\mathbb{F}_q$-parametrization $\varphi\colon \mathbb{A}^2_{(t_1,t_2)} \dashrightarrow S \hookrightarrow T$ of the *Châtelet surface*

$$S\colon y_1^2 + 12ay_2^2 = f(x) \quad \subset \quad \mathbb{A}^3_{(x,y_1,y_2)}.$$

In turn, Chávez-Saab et al. provide a less cumbersome rational $\mathbb{F}_q$-map φ to another surface on T, namely

$$S = y_1^2 + (3x^2 + 4a)y_2^2 + f(x) \quad \subset \quad \mathbb{A}^3_{(x,y_1,y_2)}.$$

Year	Authors	Complexity	Conditions
2009	Icart	$\sqrt[3]{\cdot}$	$q \equiv 2 \pmod 3$
2010	Brier et al. (the simplified SWU map)	$\sqrt{\cdot}$	$ab \neq 0$
2019	Wahby, Boneh		$ab = 0$, E has a vertical $\mathbb{F}_q$-isogeny of small degree
2022	K.		$ab = 0$, the trace of E has a small divisor
2023		$\sqrt[7]{\cdot}$	$q \equiv 2, 4 \pmod 7$, j-invariant $-3^3 5^3$

Table: Taxonomy of (non-admissible) encodings $\mathbb{F}_q \to E(\mathbb{F}_q)$ to elliptic $\mathbb{F}_q$-curves $E : y^2 = x^3 + ax + b$

Year	Authors	Complexity	Conditions
2005	Skałba	$\sqrt{\cdot}$ and two $\left(\frac{\cdot}{q}\right)$	$a \neq 0$
2006	Shallue, van de Woestijne (modification)		
2022	Chávez-Saab et al.		three ones
2009-2010	Icart (combination with the simplified SWU map)	$\sqrt[6]{\cdot}$	$q \equiv 2 \pmod 3$, $ab \neq 0$
2022	K.	$\sqrt[3]{\cdot}$	$a = 0$, $\sqrt{b} \in \mathbb{F}_q$
		$\sqrt[4]{\cdot}$	$b = 0$
2023	K. (combination with the simplified SWU map)	$\sqrt[14]{\cdot}$	$q \equiv 2, 4 \pmod 7$, j-invariant $-3^3 5^3$

Table: Taxonomy of admissible encodings $\mathbb{F}_q^2 \to E(\mathbb{F}_q)$ to elliptic $\mathbb{F}_q$-curves $E : y^2 = x^3 + ax + b$

Conclusions

There are fast Euclidean-type algorithms finding $\left(\frac{\cdot}{q}\right)$.

However, it would be desirable to completely avoid this symbol, because its efficient constant-time implementation is less available in cryptographic libraries than the usual operations $+$, $-$, $*$ in the field.

Conjecture

For any elliptic $\mathbb{F}_q$-curve E there is an admissible encoding $\mathbb{F}_q^2 \to E(\mathbb{F}_q)$ at the cost of one radical $\sqrt[n]{\cdot}$ in $\mathbb{F}_q$ for some $n \in \mathbb{N}$ (in particular, without computing additional power residue symbols $\left(\frac{\cdot}{q}\right)_n$).

Besides, $\sqrt{\cdot}$ is unwanted in *highly 2-adic fields*, i.e., $2^\nu \mid q-1$ for a non-small $\nu \in \mathbb{N}$. Lots of modern curves are defined over such fields.

This allows to apply the fast Fourier transform to speed up the polynomial arithmetic in advanced cryptographic protocols.

Hashing to elliptic curves over highly 2-adic fields

Recently, K. constructed a new indifferentiable hash function $\mathcal{H}$ to almost any elliptic $\mathbb{F}_q$-curve E, having an $\mathbb{F}_q$-isogeny of degree 3.

Since $\mathcal{H}$ just has to compute a certain *Lucas sequence* element, its complexity always equals $O(\log(q))$ operations in $\mathbb{F}_q$ with a small constant hidden in O.

In comparison, whenever $q \equiv 1 \pmod 3$, almost all previous hash functions need to extract at least one square root in $\mathbb{F}_q$.

This amounts to $O(\log(q) + \nu^2)$ operations in $\mathbb{F}_q$ for the *Tonelli–Shanks algorithm* and $O(\log(q) + \nu^{3/2})$ ones for the novel *Sarkar algorithm*.

An analysis shows that $\mathcal{H}$ requires ≈ 8300 fewer multiplications in $\mathbb{F}_q$ than the earlier state-of-the-art hash function to the curve NIST P-224 (for which $\nu = 96$) from an American standard.

Thank you for your attention!

Encoding to $E_{0,b}\colon y^2 = x^3 + b$ such that $\sqrt{b} \in \mathbb{F}_q$

If a curve $E_{0,b}$ is ordinary, then $q \equiv 1 \pmod 3$, i.e., $\omega := \sqrt[3]{1} \in \mathbb{F}_q$, where $\omega \neq 1$. Pick $c \in \mathbb{F}_q^*$ such that $\left(\frac{c}{q}\right)_3 := c^{(q-1)/3} = \omega$. Let

$$g_i := y_i^2 - b, \qquad T\colon \begin{cases} g_1 = c g_0 t_1^3, \\ g_2 = c^2 g_0 t_2^3 \end{cases} \subset \mathbb{A}^5_{(y_0,y_1,y_2,t_1,t_2)},$$

$$h'\colon T(\mathbb{F}_q) \to E_{0,b}(\mathbb{F}_q) \qquad h' := \begin{cases} \left(\sqrt[3]{g_0},\ y_0\right) & \text{if} \quad \left(\frac{g_0}{q}\right)_3 \in \{0,1\}, \\ \left(\sqrt[3]{g_1},\ y_1\right) & \text{if} \quad \left(\frac{g_0}{q}\right)_3 = \omega^2, \\ \left(\sqrt[3]{g_2},\ y_2\right) & \text{if} \quad \left(\frac{g_0}{q}\right)_3 = \omega. \end{cases}$$

Given a parametrization $\varphi\colon \mathbb{A}^2_{(t_1,t_2)} \dashrightarrow S$ of a rational $\mathbb{F}_q$-surface $S \subset T$, we obtain the admissible encoding

$$h\colon \mathbb{F}_q^2 \to E_{0,b}(\mathbb{F}_q) \qquad h := \begin{cases} (0:1:0) & \text{if} \quad \varphi(t_1,t_2) = \infty, \\ (h' \circ \varphi)(t_1,t_2) & \text{otherwise.} \end{cases}$$

Main references on the topic of other authors

1. Brier E., Coron J.-S., Icart T., Madore D., Randriam H., Tibouchi M., *Efficient indifferentiable hashing into ordinary elliptic curves*, CRYPTO 2010.
2. Chávez-Saab J., Rodriguez-Henriquez F., Tibouchi M., *SwiftEC: Shallue–van de Woestijne indifferentiable function to elliptic curves. Faster indifferentiable hashing to most elliptic curves*, ASIACRYPT 2022.
3. Farashahi R. R., Fouque P.-A., Shparlinski I. E., Tibouchi M., Voloch J. F., *Indifferentiable deterministic hashing to elliptic and hyperelliptic curves*, Mathematics of Computation, 2013.
4. Faz-Hernandez A., Scott S., Sullivan N., Wahby R. S., Wood C. A., *Hashing to elliptic curves*, Internet-draft (CFRG), 2022.
5. Shallue A., van de Woestijne C. E., *Construction of rational points on elliptic curves over finite fields*, ANTS 2006.