# Estimating the Size of the Image of Deterministic Hash Functions to Elliptic Curves

Pierre-Alain Fouque and Mehdi Tibouchi*

École normale supérieure
Département d'informatique, Groupe de cryptographie
45, rue d'Ulm, F-75230 Paris CEDEX 05, France
{pierre-alain.fouque,mehdi.tibouchi}@ens.fr

**Abstract.** Let $E$ be a non-supersingular elliptic curve over a finite field $\mathbb{F}_q$. At CRYPTO 2009, Icart [5] introduced a deterministic function $\mathbb{F}_q \to E(\mathbb{F}_q)$ which can be computed efficiently, and allowed him and Coron [3] to define well-behaved hash functions with values in $E(\mathbb{F}_q)$. Some properties of this function rely on a conjecture which was left as an open problem in [5]. We prove this conjecture below as well as analogues for other hash functions. This allows us to prove that a related function is surjective, which is a first step towards efficient hashing to the whole set of points of elliptic curves.

**Keywords**: Elliptic Curves, Function Fields, Hash Functions.

## 1 Introduction

In cryptography, it has been an open problem for a long time to transform a random value in $\mathbb{F}_q$ into a random point on an elliptic curve in a deterministic and efficient manner. Such transformations $f$ are called hash functions since they have been used, say in the context of identity-based encryption, by first hashing an identity into a random value in $\mathbb{F}_q$ using a standard cryptographic hash function $h$ and then applying such a transformation to get a point on the curve: $H(m) = f(h(m))$. They have also applications in password-based authentication schemes. However, only probabilistic solutions were known before 2006.

The usual solution before 2006 was to take $x \in \mathbb{F}_q$ and check whether this value corresponds to a valid abscissa of a point on the elliptic curve. If not, try another abscissa until one of them works. Consequently, random bits are needed to perform this random search and the running time cannot be bounded and cannot be constant. The main drawback of this

* This research was carried out while the second author was visiting the Okamoto Research Laboratory at the NTT Information Sharing Platform (Tokyo, Japan).

approach is that for password-based authentication schemes, an adversary can perform timing attacks and off-line computations in exhaustive search attacks. Some passwords do not need to be tested if the number of iterations of the probabilistic process is not the correct one. Indeed, security proofs for password-based authentication schemes rely on the fact that only on-line attacks are possible and each try allows to remove a small constant number of passwords, ideally one. Other cryptographic solutions have been proposed to avoid the random process but they made the protocol more complex. One of these is to apply the protocol twice, once with the original curve and in parallel on one of the twisted curves of the original curve. Now, any value in $\mathbb{F}_q$ corresponds either to an abscissa of the original curve or of the associated twisted curve since the two curves represent a distinct union of $\mathbb{F}_q$. Finally, it is worth noticing that the function $h(m) \cdot G$ were $G$ is a generator of the point group of the curve is not a secure solution since the discrete log of the point is known and this makes most protocols insecure.

**Deterministic functions.** To construct such function, Shallue and van de Woestijne at ANTS 2006 [6] proposed a deterministic algorithm based on Skalba's inequality. The running time of this function is $O(\log^4 q)$. Later, a generalization for hyper-elliptic curve was proposed by Ulas [8]. At CRYPTO 2009, Icart [5] proposed another more efficient technique in $O(\log^3 q)$. Finally, Coron and Icart [3] propose another technique based on a variant of the Shallue-Woestijne-Ulas (SWU) function, and explain how to construct secure hash functions to elliptic curves based on Icart's function or SWU.

Ideally, it would be nice if the image of the hash function was the whole curve, and if the distribution on the points was statistically close to uniform. In order to prove such results, it is interesting to know how many points there are in the image. Icart showed a coarse bound for his function $f$: $q/4 \leq \#f(\mathbb{F}_q) \leq q$. He conjectured that

$$\left| \#f(\mathbb{F}_q) - \frac{5q}{8} \right| \leq \lambda \sqrt{q}$$

for some constant $\lambda$ but left this conjecture as an open problem. Similar statements can be formulated about the size of the image of other hash functions, such as the characteristic 2 version of Icart's function, or the simplified version of SWU proposed by Coron and Icart.

Here, we propose proofs of these conjectures by using number theoretic tools such as the Chebotarev density theorem. It is interesting to note

that, depending on the particular function we consider, the number of points in the image varies according to some Galois group associated with the function.

That way, we can give precise estimates for the number of points in the image, and since this number is large enough, it is easy to derive a surjective function: $F(u_1, u_2) = f(u_1) + f(u_2)$ from $\mathbb{F}_q^2$ to $E(\mathbb{F}_q)$ using a counting argument. This function $F$ was considered in [**?**], but this paper estabishes for the first time that it is surjective.

**Organization of the paper.** In section 2, we describe Icart's hash function and his conjecture. Then, we prove the conjecture for curves of odd characteristic, of characteristic 2 and finally for the variant of SWU.

## 2 Preliminaries

### 2.1 Icart's function

Let $\mathbb{F}_q$ be a finite field of characteristic $> 3$ and $E$ an elliptic curve over $\mathbb{F}_q$ that isn't supersingular. $E$ can be represented as the union of its neutral element $O$ and the set of points $(x, y)$ in the affine plane over $\mathbb{F}_q$ such that:

$$y^2 = x^3 + ax + b$$

for some suitable constants $a, b \in \mathbb{F}_q$ satisfying $4a^3 + 27b^2 \neq 0$ (non-singularity) and $a \neq 0$ (non-supersingularity).

When $q-1$ is not divisible by 3, Icart [5] defines the following function $f_{a,b} \colon \mathbb{F}_q \to E(\mathbb{F}_q)$. He sets $f_{a,b}(0) = O$ and for all $u \neq 0$, $f_{a,b}(u) = (x, y)$ with:

$$\begin{aligned} x &= \left(v^2 - b - \frac{u^6}{27}\right)^{1/3} + \frac{u^2}{3} \\ y &= ux + v \end{aligned}$$

where $v = (3a - u^4)/(6u)$. This function is shown to be well-defined and easily computed in deterministic polynomial time. Moreover, if $(x, y)$ is a point in $E(\mathbb{F}_q)$, then $f_{a,b}(u) = (x, y)$ if and only if $u$ satisfies the quartic equation

$$u^4 - 6xu^2 + 6yu - 3a = 0$$

### 2.2 Icart's conjecture

In [5], Icart conjectures that the image of $f_{a,b}$ contains $(5/8) \cdot \#E(\mathbb{F}_q) + O(q^{1/2})$ points of the curve. In view of the previous equation, and since the curve itself has $\#E(\mathbb{F}_q) = q + O(q^{1/2})$ points in $\mathbb{F}_q$, this conjecture can be stated as follows.

*Conjecture 1 (Icart).* Let $K = \mathbb{F}_q(x, y) = \mathbb{F}_q(x)[Y]/(Y^2 - x^3 - ax - b)$ be the function field of $E$, and $P$ the polynomial of $K[u]$ defined by $P(u) = u^4 - 6xu^2 + 6yu - 3a$. Let further $N$ be the number of points in $E(\mathbb{F}_q)$ at which the reduction of $P$ has a root in $\mathbb{F}_q$. Then

$$N = \frac{5}{8}q + O(q^{1/2})$$

The next section is devoted to the proof of this conjecture.

## 3 Proof of Icart's conjecture

### 3.1 Genericity of $P$

**Proposition 1.** *The polynomial $P(u) = u^4 - 6xu^2 + 6yu - 3a \in K[u]$ is irreducible over $K$, and its Galois group is $S_4$.*

*Proof.* Introduce the Ferrari resolvent of $P$:

$$\begin{aligned} C(u) &= u^3 + 12xu^2 + (36x^2 + 12a)u + 36y^2 \\ &= u^3 + 12xu^2 + (36x^2 + 12a)u + 36(x^3 + ax + b) \in \mathbb{F}_q(x) \end{aligned}$$

According to classical facts about the quartic equation (see e.g. [2, Theorem 13.1.1]), it suffices to prove that $P$ and $C$ are irreducible over $K$, and that their common discriminant

$$\Delta = -432(9x^6 + 18ax^4 + 90bx^3 - 39a^2x^2 - 54abx + 16a^3 + 81b^2)$$

is not a square in $K$. Moreover, we can prove these assertions after extending the field of scalars to $F = \bar{\mathbb{F}}_q$. Indeed, if they hold over $F$, they clearly hold *a fortiori* over $\mathbb{F}_q$. The following three lemmas conclude the proof.

**Lemma 1.** *The Ferrari resolvent cubic $C(u)$ is irreducible over $F(x, y)$.*

*Proof.* This amounts to showing that $C(u)$ has no root in $F(x,y)$. Note first that it is actually sufficient to prove it has no root in $F(x)$. Indeed, if it is irreducible in $F(x)$ but has a root in $F(x,y)$, the degree of the algebraic extension $F(x,y)/F(x)$ must be divisible by $\deg C(u) = 3$. But this extension is quadratic: hence a condraction.

Let then $f/g$ be a root of $C$ in $F(x)$, with $f$ and $g$ coprime polynomials. Multiplying the equation $C(f/g) = 0$ by $f^3$, we get

$$f^3 = -g \cdot \big( -12xf^2 - (36x^2 + 12a)fg - 36(x^3 + ax + b)g^2 \big)$$

Thus $g$ divides $f^3$, and since it is coprime to $f$, it must be constant. Without loss of generality, we thus have $g = 1$ and

$$f^3 + 12xf^2 + (36x^2 + 12a)f + 36(x^3 + ax + b) = 0$$

Let $m = \deg f$. Then the terms in the previous sum are of respective degrees $3m$, $2m+1$, $m+2$, 3. If $m \geq 2$, the sum is thus of degree $3m$, and if $m \leq 0$, it is of degree 3: in neither case can it be 0. The only possibility is thus $m = 1$ and $f = \alpha x + \beta$. We get

$$(\alpha^3 + 12\alpha^2 + 36\alpha + 36)x^3 + 3\beta(\alpha^2 + 8\alpha + 12)x^2 + \\ (3\alpha\beta^2 + 12a\alpha + 12\beta^2 + 36a)x + (\beta^3 + 12a\beta + 36b) = 0$$

in $F(x)$. Suppose $\beta \neq 0$. Since the coefficients of $x^3$ and $x^2$ must be zero, this gives $\alpha^3 + 12\alpha^2 + 36\alpha + 36 = \alpha^2 + 8\alpha + 12 = 0$, which is impossible, since the polynomials $X^3 + 12X^2 + 36X + 36$ and $X^2 + 8X + 12$ are coprime. Hence $\beta = 0$, and thus $\alpha^3 + 12\alpha^2 + 36\alpha + 36 = 12a(\alpha + 3) = 0$, which is similarly seen to be impossible (as $a \neq 0$). This completes the proof.

**Lemma 2.** *The discriminant $\Delta$ is not a square in $F(x,y)$.*

*Proof.* Again, we will show that it is sufficient to prove that $\Delta$ is not a square in $F(x)$. Indeed, suppose that $\Delta$ is not a square in $F(x)$ but becomes a square in $F(x,y)$. Since the extension is quadratic, this gives $F(x,y) = F(x, \sqrt{\Delta})$. In particular, if $\lambda$ is a root of $X^3 + aX + b$ in $F$, the extension $F(x, \sqrt{\Delta})/F(x)$ must be ramified at $(x - \lambda)$. In other words, if we specialize $\Delta(x)$ at $x = \lambda$, we must get 0. But

$$\begin{aligned}(\lambda - 3b/a)\Delta(\lambda) &= 16 \cdot 432(\lambda - 3b/a)(3a^2\lambda^2 + 9ab\lambda - a^3) \\ &= 3a^2(\lambda^3 + a\lambda + b) - (4a^3 + 27b^2)\lambda \neq 0\end{aligned}$$

hence a contradiction.

It remains to prove that $\Delta$ is not a square in $F(x)$, or equivalently in $F[x]$ (since $F[x]$ is integrally closed). A square root of $\Delta$ in $F[x]$ must have the form $S = \sqrt{-432} \cdot (3x^3 + rx^2 + sx + t)$. The coefficient of $x^5$ in $S^2$ must be 0, hence $r = 0$. The coefficient of $x^4$ must be $18a$, hence $s = 3a$. But then the coefficient of $x^2$ is equal to both $9a^2$ and $-39a^2$, which is a contradiction since $48a^2 \neq 0$. Hence the result.

**Lemma 3.** *The polynomial $P$ is irreducible over $F(x, y)$.*

*Proof.* Let $\sigma$ be the non trivial Galois automorphism of the extension $F(x, y)/F(x)$ ($\sigma(y) = -y$). If $P(u)$ decomposes as a product of non constant factors in $F(x, y)[u]$, then its norm $P_0(u) = P(u)P(u)^\sigma$ is reducible over $F(x)$. We will show that this is not the case. Note first that $P_0(u)$ can be written as $Q_0(u^2)$, where

$$Q_0(v) = v^4 - 12xv^3 + (36(x^3 + ax + b) - 6a)v^2 - 36(x^3 + b)v + 9a^2$$

Now $Q_0(v)$ is easily seen to be an irreducible polynomial of $F(x)[v]$. Indeed, if it had a root $f/g \in F(x)$, the rational function $f/g$ would be constant, which is clearly impossible. And if it decomposes as a product of degree 2 factors $Q_0 = (v^2 + rv + s)(v^2 + r'v + s')$, these factors are in $F[x]$ (integrally closed domain). Since $ss' = 9a^2$, both $s$ and $s'$ are constant. Then, since the coefficient of $v^2$, $rr' + s + s'$, is of degree 3, one must have $\deg(r + r') \geq 2$. But $r + r'$ is the coefficient of $v^3$ in $Q_0$, namely $-12x$, hence a contradiction.

Now let $w$ be a root of $P_0$ in the separable closure of $F(x)$, and let $L = F(x, w)$, $L' = F(x, w^2)$. $L'$ is a subfield of $L$, and a rupture field of $Q_0$. In particular $[L : F(x)] = [L : L'] \cdot [L' : F(x)] = 4[L : L']$. Since the polynomial $P_0$ is even, $-w$ is another root of $P_0$. As $w \notin F(x)$, $w \mapsto -w$ defines a non trivial $F(x)$-automorphism of $L$. This automorphism fixes $L'$, so $[L : L'] \geq 2$. This gives $[L : F(x)] \geq 8$, and thus $P_0$ must have an irreducible factor of degree $\geq 8$. In other words, $P_0$ is irreducible over $F(x)$ as required.

### 3.2 Applying Chebotarev

Now that Proposition 1 is established, Conjecture 1 readily follows from effective versions of the Chebotarev Density Theorem for function fields. One such version is [4, Proposition 6.4.8], from which one can easily deduce:

**Theorem 1 (Chebotarev).** *Let $K$ be an extension of $\mathbb{F}_q(x)$ of degree $d < \infty$ and $L$ a Galois extension of $K$ of degree $m < \infty$. Assume $\mathbb{F}_q$*

*is algebraically closed in $L$, and fix some subset $\mathscr{S}$ of* $\mathrm{Gal}(L/K)$ *stable under conjugation. Let $s = \#\mathscr{S}$ and $N(\mathscr{S})$ the number of places $v$ of $K$ of degree* 1*, unramified in $L$, such that the Artin symbol $\left(\frac{L/K}{v}\right)$ (defined up to conjugation) is in $\mathscr{S}$. Then*

$$\left|N(\mathscr{S}) - \frac{s}{m}q\right| \leq \frac{2s}{m}\big((m+g_L)\cdot q^{1/2} + m(2g_K+1)\cdot q^{1/4} + g_L + dm\big)$$

*Proof (of Conjecture 1).* In our case, $K$ is the function field of $E$ and $L$ the splitting field of $P(u)$. In particular, $d = 2$, $m = \#S_4 = 24$ and $g_K = 1$. We consider the subset $\mathscr{S} \subset \mathrm{Gal}(L/K) = S_4$ consisting of permutations with at least one fixed point—these are the conjugates of (1), (12) and (123), and there are $s = 1 + 6 + 8 = 15$ of them. Hence $s/m = 15/24 = 5/8$.

The places $v$ of $K$ of degree 1 correspond to points of $E(\mathbb{F}_q)$ (in the projective plane), and for a point $(x_0, y_0) \in E(\mathbb{F}_q)$ not at infinity, saying that $v = (x - x_0)$ has its Artin symbol in $\mathscr{S}$ means that the reduction of $P(u)$ at $(x_0, y_0)$ is a polynomial over $\mathbb{F}_q$ which decomposes into a products of factors at least one of which is of degree 1 (it splits completely if the symbol is (1), decomposes as two linear factors and a quadratic if it is (12) and a product of a linear factor and a cubic if it is (123) up to conjugation).

In other words, $N(\mathscr{S})$ is the same as $N$ in the statement of Conjecture 1 up to a constant number accounting for ramified places (at most 12 since $\Delta$ is a polynomial of degree 6 in $x$) and the point at infinity. We then get

$$\left|N - \frac{5}{8}q\right| \leq \frac{5}{4}\big((24+g_L)\cdot q^{1/2} + 72q^{1/4} + g_L + 48 + 13\big)$$

To bound $g_L$, note again that there are at most 12 ramified points, and the ramification index is at most $\deg P_0 = 4$ at each of them. The Riemann-Hurwitz formula thus gives

$$2 - 2g_L \geq 24(2 - 2g_K) - 12\cdot(4-1) \quad \text{i.e.} \quad g_L \leq 17$$

and thus

$$\left|N - \frac{5}{8}q\right| \leq \frac{5}{4}\big(41q^{1/2} + 72q^{1/4} + 78\big)$$

In particular, $N = (5/8)q + O(q^{1/2})$. Concretely, for all $q \geq 2^{29}$, we have

$$\left|N - \frac{5}{8}q\right| \leq 50q^{1/2} \tag{1}$$

## 4 Analogue in Characteristic 2

In [5], Icart also introduces a variant of his function for elliptic curves over finite fields $\mathbb{F}_q$ of even characteristic, i.e. $q = 2^n$. Such an elliptic curve has the form

$$y^2 + xy = x^3 + ax^2 + b$$

with $a, b \in \mathbb{F}_q$, $b \neq 0$. Icart's function for such a curve $E$ is defined when $n$ is odd as

$$\begin{aligned} f_{a,b} \colon \mathbb{F}_q &\to E(\mathbb{F}_q) \\ u &\mapsto (x, ux + v^2) \end{aligned}$$

where $v = a + u + u^2$ and $x = (v^4 + v^3 + b)^{1/3} + v$. It is shown that $u \in \mathbb{F}_q$ maps to $(x, y) \in E(\mathbb{F}_q)$ if and only if $P(u) = 0$, where $P \in K[u]$ is defined as

$$P(u) = u^4 + u^2 + xu + (a + y)$$

Using this result, we can prove the following analogue of Icart's conjecture.

**Proposition 2.** *The number of points $N$ in the image of $f_{a,b}$ satisfies:*

$$N = \frac{3}{4}q + O(q^{1/2})$$

*where the implied constant in the big-O is universal.*

The proof is identical to the one in §3.2. The only difference is that the Galois group of $P$ is $A_4$ instead of $S_4$, which leads to the constant $3/4$ instead of $5/8$ (as there are $1 + 8 = 9$ permutations out of 12 in $A_4$ which have at least one fixed point). Let us prove this fact now.

**Proposition 3.** *The polynomial $P(u) = u^4 + u^2 + xu + (a + y) \in K[u]$ is separable and irreducible over $K$, and its Galois group is $A_4$.*

*Proof.* Since $P' = x$ is a unit in $K[u]$, $P$ is certainly separable. Now, the relevant case of [2, Theorem 13.1.1] is easily seen to hold in any characteristic for separable polynomials, so it remains to prove that $P$ is irreducible, that its Ferrari resolvent $C$ is irreducible, and that their common discriminant $\Delta$ is a square in $K$.

Note first that $\Delta = x^4$, so the last point is obvious. Further, we have $C(u) = u^3 - u^2 - x^2$. If this polynomial had a root in $\mathbb{F}_q(x)$, it would be a polynomial of $\mathbb{F}_q[x]$ dividing $x^2$ by integral closure, which is clearly

impossible. Therefore, $C(u)$ is irreducible over $\mathbb{F}_q(x)$, and also over $K$ by the same degree argument as in the proof of Lemma 1.

Finally, let us prove that $P$ is irreducible. Let first $\sigma$ be the non-trivial Galois automorphism of $K/\mathbb{F}_q(x)$, namely $y \mapsto y + x$, and set $P_0 = PP^\sigma \in \mathbb{F}_q(x)$. It suffices to prove that $P_0$ is irreducible over $\mathbb{F}_q(x)$. We have

$$P_0 = (u^8+u^4)+x(u^4+u^2)+x^2(u^2+u)+(x^3+ax^2+ax+a^2+b) = Q_0(u^2+u)$$

where $Q_0(v) = v^4 + xv^2 + x^2v + (x^3 + ax^2 + ax + a^2 + b)$.

If $Q_0$ has a root over $\mathbb{F}_q(x)$, it is in fact in $\mathbb{F}_q[x]$, which is not possible by inspection of the degrees of the four terms in the sum. Similarly, if $Q_0$ can be written as a product of factors of degree 2, we have $Q_0 = (v^2+r+s)(v^2-r+s')$ with $r$, $s$ and $s'$ are all in $\mathbb{F}_q[x]$. We get $\deg(ss') = 3$, so the polynomials $s \pm s'$ must be of degree at least 2. Since $r(s-s') = -x^2$, this implies that $r$ is constant. But then the relation $s + s' - r^2 = x$ gives a contradiction. Therefore $Q_0$ is irreducible over $\mathbb{F}_q(x)$.

Then, let $w$ be a root of $P_0$ in the separable closure of $\mathbb{F}_q(x)$, and set $L = \mathbb{F}_q(x, w)$, $L' = \mathbb{F}_q(x, w + w^2)$. Like in the proof of Lemma 3, we have a tower of extensions $\mathbb{F}_q(x) \subset L' \subset L$, and $L'$ is a rupture field of $Q_0$, so $[L : \mathbb{F}_q(x)] = 4[L : L']$. Furthermore, since $P_0(u + 1) = P(u)$, $w \mapsto w + 1$ is a non-trivial $L'$-automorphism of $L$, which gives $[L : \mathbb{F}_q(x)] \geq 8$ and hence, $P_0$ is irreducible over $\mathbb{F}_q(x)$, which concludes the proof.

We can again give concrete bounds. With the notations of §3.2, we have $d = 2$, $m = 12$, $s = 8$, $g_K = 1$ and there is exactly one ramified point corresponding to $x = 0$. The Riemann-Hurwitz formula then gives $g_L \leq 2$, and thus:

$$\left| N - \frac{3}{4}q \right| \leq 21q^{1/2} + 54q^{1/4} + 42$$

In particular, for $q > 2^{16}$ we get

$$\left| N - \frac{3}{4}q \right| \leq 25q^{1/2} \tag{2}$$

## 5 Analogue for the simplified Shallue-Woestijne-Ulas algorithm

The first deterministic algorithm for hashing to elliptic curves was introduced by Shallue and Woestijne in [6]. It was later generalized and

simplified by Ulas in [8]. Coron and Icart [3] describe a further simplification of the Shallue-Woestijne-Ulas (SWU) algorithm for elliptic curves over fields $\mathbb{F}_q$ with $q \equiv 3 \pmod 4$, based on the following result.

**Theorem 2 ([3], Th. 5).** *Let $\mathbb{F}_q$ be a finite field and $g(x) := x^3 + ax + b$, where $ab \neq 0$. Consider the following rational functions.*

$$X_2(u) = -\frac{b}{a}\left(1 + \frac{1}{u^4 - u^2}\right), \quad X_3(u) = -u^2 X_2(u), \quad Z(u) = u^3 g\big(X_2(u)\big)$$

*Then we have $Z(u)^2 = -g\big(X_2(u)\big) \cdot g\big(X_3(u)\big)$.*

If $q \equiv 3 \pmod 4$, $-1$ is a quadratic non-residue in $\mathbb{F}_q$. Therefore, for each $u$, exactly one of $g\big(X_2(u)\big)$ and $g\big(X_3(u)\big)$ is a square. This leads to the following deterministic algorithm mapping elements in $\mathbb{F}_q$ to points on the curve $E_{a,b} : y^2 = x^3 + ax + b$.

**Simplified SWU algorithm.**
Input: $\mathbb{F}_q$ such that $q > 3$ and $q \equiv 3 \pmod 4$, parameters $a, b$ such that $ab \neq 0$, and input $u \in \mathbb{F}_q$.
Output: $(x, y) \in E_{a,b}(\mathbb{F}_q)$.

1. $\alpha \leftarrow -u^2$
2. $X_2 \leftarrow -\frac{b}{a}\left(1 + \frac{1}{\alpha^2 + \alpha}\right)$
3. $X_3 \leftarrow \alpha \cdot X_2$
4. $h_j \leftarrow X_j^3 + aX_j + b$, $j = 2, 3$
5. If $h_2$ is a square, return $(X_2, h_2^{(q+1)/4})$; otherwise, return $(X_3, -h_3^{(q+1)/4})$.

This algorithm is a slightly modified version of the one described in [3] §5.5. The only difference is the minus sign in $(X_3, -h_3^{(q+1)/4})$, which ensures that, up to three possible exceptions (points with a zero $x$-coordinate), the set of points obtained when $g\big(X_2(u)\big)$ is a square is disjoint from the set of points obtained when $g\big(X_3(u)\big)$ is a square (which improves the size of the image over the original version). Thus, the image of this function $\mathbb{F}_q \to E_{a,b}(\mathbb{F}_q)$ is the (almost disjoint) union of the sets $I_2$ and $I_3$ defined by

$$I_j = \left\{(x, y) \in E_{a,b}(\mathbb{F}_q) \mid \exists u \in \mathbb{F}_q,\ x = X_j(u) \text{ and } y = (-1)^j \sqrt{g(x)}\right\}$$

(where $\sqrt{\cdot}$ denotes the standard square root in $\mathbb{F}_q$, obtained by exponentiation by $(q+1)/4$). Again disregarding at most three points, $I_j$ consists of half the points on the curve with an $x$-coordinate of the form $X_j(u)$ for some $u$. Therefore, if $N$ is the number of points in the image of the

algorithm and $N_j$ denotes the number of points with an $x$-coordinate of the form $X_j(u)$, we get

$$N = \frac{N_2 + N_3}{2} + O(1)$$

and the implied constant is at most 6. We deduce the following result.

**Proposition 4.** *The number of points $N$ in the image of the simplified SWU algorithm satisfies:*

$$N = \frac{3}{8}q + O(q^{1/2})$$

*where the implied constant in the big-O is universal.*

*Proof.* The proof is again similar to the previous ones. What we actually show is that $N_j = (3/8)q + O(q^{1/2})$ for $j = 2, 3$, using the Chebotarev density theorem again. Note that for all $u \in \mathbb{F}_q \setminus \{-1, 0, 1\}$, we have

$$x = X_2(u) \iff u^4 - u^2 + \frac{1}{\omega} = 0$$
$$x = X_3(u) \iff u^4 - \omega u^2 + \omega = 0$$

where $\omega = \frac{a}{b}x + 1$. Hence, denoting by $K = \mathbb{F}_q(x, y)$ the function field of $E_{a,b}$, it suffices to prove that the polynomials $P_2(u) = u^4 - u^2 + 1/\omega$ and $P_3(u) = u^4 - \omega u^2 + \omega$ are irreducible and have Galois group $D_8$ (the 8-element dihedral group, viewed as a transitive subgroup of $S_4$) over $K$. Indeed, $D_8$ has 8 elements, 3 of which have a fixed point: the same technique as in §3.2 then gives the desired estimates for $N_2$ and $N_3$.

In view of [1, Theorems 2 and 3], a polynomial $P(u) = u^4 - ru^2 + s \in K[u]$ is irreducible with Galois group $D_8$ if and only if none of $s$, $\delta = r^2 - 4s$ or $s\delta$ are squares in $K$. For $P_2$, we have $(s, \delta, s\delta) = \frac{1}{\omega^2}(\omega, \omega(\omega - 4), \omega - 4)$, and for $P_3$, $(s, \delta, s\delta) = (\omega, \omega(\omega - 4), \omega^2(\omega - 4))$. Thus, all we have to prove is that $\omega$, $\omega - 4$ and $\omega(\omega - 4)$ are not squares in $K$. This is obvious in $\mathbb{F}_q(x)$ (since these are polynomials of $\mathbb{F}_q[x]$ which are not square), and extends to $K$ by a ramification argument as in the proof of Lemma 2.

## 6 Constructing surjective hash functions

In view of the previous results, none of the known functions $\mathbb{F}_q \to E(\mathbb{F}_q)$ are surjective. It is possible to use at least Icart's function to construct simple, efficient surjective hash functions, however, as explained in [5, Corollary 2].

Indeed, let $f$ be Icart's function on $\mathbb{F}_q$ (which can be a field of characteristic 2 or $> 3$) and consider

$$\begin{aligned} F : (\mathbb{F}_q)^2 &\to E(\mathbb{F}_q) \\ (u_1, u_2) &\mapsto f(u_1) + f(u_2) \end{aligned}$$

If we fix a point $P_0 \in E(\mathbb{F}_q)$, the sets $S_1 = \{f(u_1) \ / \ u_1 \in \mathbb{F}_q\}$ and $S_2 = \{P_0 - f(u_2) \ / \ u_2 \in \mathbb{F}_q\}$ both consist of $\alpha q + O(q^{1/2})$ points of $E$, with $\alpha = 5/8$ (resp. $3/4$) in characteristic $> 3$ (resp. 2). In any case, $\#S_1 + \#S_2 = 2\alpha q + O(q^{1/2})$, which is greater than $q + 2\sqrt{q} + 1 \geq \#E(\mathbb{F}_q)$ for large enough $q$. Hence, a pigeonhole argument ensures that $P_0$ is in the image of $F$ provided that $q$ is large enough.

More precisely, the conditions we have given for the explicit bounds (1) and (2) to hold are sufficient to ensure that the previous inequality is satisfied. Therefore, the function $F$ is surjective as soon as $q > 2^{29}$ in characteristic $> 3$ (resp. $q > 2^{16}$ in characteristic 2), which is always true in cryptographic applications.

This leads to a hash function construction that is subtantially more efficient than the one proposed in [3]. It is an interesting open problem, however, to determine whether this construction is equally secure.

## 7 Conclusion

In this paper, we provide a technique to analyze the image of some hash functions mapping elements of $\mathbb{F}_q$ to elliptic curves $E(\mathbb{F}_q)$. It relies on the Chebotarev density theorem in function fields, and in order to apply it, we need to prove the irreducibility of some related polynomial and compute its Galois group.

The same technique should apply similarly to any deterministic, algebraic hash function to curves of any genus. Depending on the particular hash function under consideration, the Galois group varies and the Chebotarev density theorem yields different results accordingly.

## Acknowledgements

We would like to thank Jean-Sébastien Coron and Thomas Icart for suggesting this problem to us, and both of them as well as Éric Brier and Go Yamamoto for helpful comments and discussions.

## References


1. L.C. Kappe, B. Warren, *An elementary test for the Galois group of a quartic polynomial*, Amer. Math. Monthly, vol. 96(2), Mathematical Association of America, 1989, pp. 133–137.
2. D.A. Cox, *Galois theory*, Series in Pure Mathematics, Wiley, 2004.
3. J.-S. Coron, T. Icart, *An indifferentiable hash function into elliptic curves*, preprint, 2009, `http://eprint.iacr.org/2009/340`.
4. M.D. Fried, M. Jarden, *Field arithmetic*, Second edition, Ergebnisse der Mathematik und ihre Grenzgebiete, vol. 11, Springer-Verlag, 2002.
5. T. Icart, *How to hash into elliptic curves*, Proceedings of Crypto 2009, LNCS, vol. 5677, Springer-Verlag, 2009, pp. 303–316.
6. A. Shallue and C. van de Woestjine, *Construction of rational points on elliptic curves over finite fields*, Proceedings of ANTS 2006, LNCS, vol. 4076, Springer-Verlag, 2006, pp. 510–524.
7. M. Skałba, *Points on elliptic curves over finite fields*, Acta Arith., vol. 117, IMPAN, 2005, pp. 293–301.
8. M. Ulas, *Rational points on certain hyperelliptic curves over finite fields*, Bull. Polish Acad. Sci. Math., vol. 55, IMPAN, 2007, pp. 97–104.