Batch point compression in the context of advanced pairing-based protocols

Dmitrii Koshelev[0000−0002−4796−8989]
dimitri.koshelev@gmail.com

Computer sciences and networks department, Télécom Paris, France
https://www.telecom-paris.fr

Abstract. This paper continues previous ones about compression of points on elliptic curves $E_b\colon y^2 = x^3 + b$ (with j-invariant 0) over a finite field $\mathbb{F}_q$ of characteristic $p > 3$. More precisely, we will show in detail how any two (resp. three) points from $E_b(\mathbb{F}_q)$ can be quickly compressed to two (resp. three) elements of $\mathbb{F}_q$ (apart from a few auxiliary bits) in such a way that the corresponding decompression stage requires to extract only one cubic (resp. sextic) root in $\mathbb{F}_q$ (with several multiplications and without inversions). As a result, for many q occurring in practice the new compression-decompression methods are more efficient than the classical one with the two (resp. three) x or y coordinates of the points, which extracts two (resp. three) roots in $\mathbb{F}_q$. It will be explained why the new methods are useful in the context of modern real-world pairing-based protocols such as Groth16. As a by-product, when $q \equiv 2 \pmod 3$ (in particular, E_b is supersingular), we will obtain a two-dimensional analogue of Boneh–Franklin's encoding, that is a way to sample two "independent" $\mathbb{F}_q$-points on E_b at the cost of one cubic root in $\mathbb{F}_q$. Finally, the case of four and more points from $E_b(\mathbb{F}_q)$ will be commented on.

Keywords: batch point compression · Boneh–Franklin's encoding · conic bundle structure · cubic and sextic roots · elliptic curves of j-invariant 0 · Freeman's transformation · generalized Kummer varieties · high 2-adicity · rationality problems · recursive proof systems.

1 Introduction

Nowadays, pairing-based cryptography [12] can be certainly considered as an independent fruitful area of public-key cryptography, which is interesting from both mathematical and practical points of view. There are countless pairing-based protocols, many of which have found applications in the real world. It is worth noting protocols based on composite-order groups such as Boneh–Goh–Nissim's (BGN) *somewhat homomorphic encryption* [5] or Boneh–Sahai–Waters's *fully collusion resistant traitor tracing* [6]. It is also impossible not to mention *succinct non-interactive zero-knowledge (NIZK) proofs* among which one of the most popular is Groth16 [15]. And their *recursive compositions* are constructed via chains of elliptic curves as first suggested in [2].

Unfortunately, composite-order subgroups of $E_b(\mathbb{F}_q)$ must be very large to be protected against sub-exponential factorization algorithms. By virtue of Hasse's inequality (see, e.g., [12, Theorem 2.9]) we have $\#E_b(\mathbb{F}_q) \approx q$, hence pairing computation on E_b turns out to be very cumbersome as confirmed in [17]. Fortunately, with the help of so-called *Freeman's transformation* [14] (cf. [16, Sections 9-10]) we can almost always rewrite a protocol in the composite-order setting to the prime-order one operating with point vectors from $E_b^n(\mathbb{F}_q)$ for a smaller q and some $n \in \mathbb{N}$. In this case, an instance of the *subgroup decision problem* is a (prime-)order subgroup of $E_b^n(\mathbb{F}_q)$. For the majority of protocols it is sufficient to take $n = 2$, but there are some protocols (such as Katz–Sahai–Waters's *predicate encryption* [14, Section 7]) needing $n = 3$.

As said, e.g., in [11, Section 2.2] for the sake of efficiency of recursive proofs one needs to leverage pairing-friendly elliptic curves defined over *highly 2-adic fields* $\mathbb{F}_q$, that is the number $q - 1$ should be divisible by a non-small power 2^m, where $m \in \mathbb{N}$. More precisely, this allows to apply the fast Fourier transform (FFT) in order to speed up the polynomial arithmetic over $\mathbb{F}_q$. To be definite, we will suppose that high 2-adicity takes place if $m \geqslant 3$, but in practice usually $20 < m < 60$. Our choice follows from the fact that (as is known, e.g., from [12, Section 5.1.7]) for $q \equiv 1 \pmod 8$ it is problematic to express a square root in $\mathbb{F}_q$ via one exponentiation. Of course, we can always utilize Tonelli–Shanks's algorithm, namely [12, Algorithm 5.14] (cf. [32]), but it has a greater computational complexity.

Recall that curves E_b are ordinary (a.k.a. non-supersingular) if and only if the characteristic $p \equiv 1 \pmod 3$ or, equivalently, a primitive cubic root $\omega := \sqrt[3]{1}$ lies in $\mathbb{F}_p$. Since only curves E_b possess order 6 automorphism (of the form $[-\omega](x, y) := (\omega x, -y)$), according to [12, Section 3.2.5] such pairing-friendly ordinary curves are preferred in pairing-based cryptography. To the author's knowledge, at the moment, the most popular curves are BLS12-381 [31, Section 4.2.1] for a general use and BLS12-377 [11, Table 2] for one layer proof composition, where the numbers after the hyphen equal $\lceil \log_2(q) \rceil$. Moreover, the field $\mathbb{F}_q$ of the latter curve (in contrast to the former one) is highly 2-adic with $m = 46$. Among other things, the pages [19], [20] specify 2-cycles of curves of j-invariant 0 (over highly 2-adic fields) among which only one is pairing-friendly.

In compliance with [18, Examples IV.1.3.5-6] elliptic curves are not *rational*, i.e., they are not birationally isomorphic to the affine line $\mathbb{A}^1$. Therefore from the geometric point of view the most compact representation of them is on the affine plane $\mathbb{A}^2_{(x,y)}$, for example in the Weierstrass form. Consequently, any point from $E_b^n(\mathbb{F}_q) \subset \mathbb{F}_q^{2n}$ is obviously represented with the help of $2n\lceil \log_2(q) \rceil$ bits. In particular, for $n = 2$ (resp. $n = 3$) and $\log_2(q) \approx 380$ we obtain ≈ 1520 (resp. ≈ 2280) bits, which is quite a lot. For instance, two $\mathbb{F}_q$-points constitute a half of the proof in Groth16 [15, Table 1]. In comparison, with the same 128-bit security level classical (i.e., non-pairing-friendly) elliptic curves are defined over 256-bit fields $\mathbb{F}_q$. And many widespread cryptosystems on such curves (e.g., ECDH or ECDSA) don't require compressing several points at once, so it is sufficient to manipulate only 512 bits.

At the same time, by virtue of Hasse's inequality $\mathbb{F}_q$-points on E_b can be compressed to about half with regard to information theory. There is the classical compression-decompression method representing a point as its x (resp. y) coordinate in addition to one (resp. two) bits to uniquely recover the initial y (resp. x) coordinate via extracting in $\mathbb{F}_q$ the square (resp. cubic) root. In comparison with standard arithmetical operations in $\mathbb{F}_q$, the latter one is very costly, because even for $q \not\equiv 1 \pmod 8$ (resp. $q \not\equiv 1 \pmod{27}$) it consists in one exponentiation in $\mathbb{F}_q$ according to Lemma 3 (resp. 4). As a result, after compressing $\mathbb{F}_q$-point vectors of length $n = 2$ (resp. $n = 3$) we obtain ≈ 760 (resp. ≈ 1140) bits at the price of n exponentiations in the decompression stage.

1.1 Brief description of the new compression method and its relevance

Apart from $\tau_6 := [-\omega]$ there are on E_b the automorphisms

$$\tau_2 := \tau_6^3 \colon (x, y) \mapsto (x, -y), \qquad \tau_3 := \tau_6^4 \colon (x, y) \mapsto (\omega x, y)$$

of orders 2 and 3, respectively. For any $n \in \mathbb{N}$ and $m \in \{2, 3, 6\}$ consider the diagonal subgroup $G_{n,m} := \langle(\tau_m, \ldots, \tau_m)\rangle \simeq \mathbb{Z}/m$ of the automorphism group on E_b^n. Notice that it is Frobenius invariant even if $\omega \notin \mathbb{F}_q$. Further, introduce the $\mathbb{F}_q$-quotient $GK_{n,m} := E_b^n/G_{n,m}$, which is called *generalized Kummer variety* [34, Section 7], because for $m = 2$ this is a (usual) *Kummer variety* [34, Example 8.1]. Also, we need the notation of the quotient $\mathbb{F}_q$-cover $\varphi_{n,m} \colon E_b^n \to GK_{n,m}$, which, as usual [18, Theorem I.4.4], gives the function field extension $\mathbb{F}_q(GK_{n,m}) \hookrightarrow \mathbb{F}_q(E_b^n)$. Whenever $m = 2$ or $\omega \in \mathbb{F}_q$, by virtue of Artin's theorem (see, e.g., [27, Theorem VI.1.8]) $\varphi_{n,m}$ is a Galois cover whose Galois group equals $G_{n,m}$. Therefore $\varphi_{n,m}$ is a *Kummer cover* due to [27, Theorem VI.6.2]. All of the above is illustrated with the famous examples $\varphi_{1,2}(x, y) = x$ and $\varphi_{1,3}(x, y) = y$.

We see that $GK_{1,m}$ are obviously rational curves. More generally, there is the analogous notion of *(geometrically) rational variety* as defined in [18, Example II.8.20.1]. Rationality of the surfaces $GK_{2,3}$, $GK_{2,6}$ is a classical fact. According to [28, Section 2] the threefold $GK_{3,6}$ is also rational and there are [8, Questions 1.3, 1.4] about rationality of $GK_{4,6}$, $GK_{5,6}$. In turn, the varieties $GK_{n,m}$ are never rational for $n \geqslant m$ in accordance with [34, Example 8.10], [28, Remark 2.9]. In fact, we are interested in $\mathbb{F}_q$-rationality of $GK_{n,m}$. In a cryptographic context this concept [30, Definition 6.1] first arose in so-called *torus-based cryptography* for compressing $\mathbb{F}_q$-points of *algebraic tori*. By the way, since pairing values can be interpreted as such points, this compression technique is known to be useful in pairing-based cryptography.

For the Kummer covers $\varphi_{n,m}$ computing an inverse image $\varphi_{n,m}^{-1}(P)$ of a point $P \in \varphi_{n,m}\big(E_b^n(\mathbb{F}_q)\big)$ can be implemented by means of extracting in $\mathbb{F}_q$ some root of degree m. Suppose that $GK_{n,m}$ is an $\mathbb{F}_q$-rational variety and there are explicit formulas of a birational $\mathbb{F}_q$-isomorphism $\psi_{n,m} : GK_{n,m} \overset{\sim}{\dashrightarrow} \mathbb{A}^n$ and its inverse $\psi_{n,m}^{-1} : \mathbb{A}^n \overset{\sim}{\dashrightarrow} GK_{n,m}$. As is customary in algebraic geometry, the arrow $\dashrightarrow$ (resp. $\overset{\sim}{\dashrightarrow}$) means a (bi)rational map rather than an (iso)morphism, that is the

map may be undefined at some points. Treating them separately, we thus get a new compression-decompression method for all $\mathbb{F}_q$-points on E_b^n. Indeed, the compression (resp. decompression) stage consists in evaluating the map $\chi_{n,m} := \psi_{n,m} \circ \varphi_{n,m}$ at a general point $Q \in E_b^n(\mathbb{F}_q)$ (resp. finding $\chi_{n,m}^{-1}(R)$, where $R := \chi_{n,m}(Q)$).

For the surface $GK_{2,3}$ (resp. $GK_{2,6}$) $\mathbb{F}_q$-rationality is explicitly established in Section 2 (resp. [22, Sections 2-3]), although these results cannot be considered very important for pure mathematics because of their simplicity. Besides, it turns out that $\mathbb{F}_q$-formulas of $\psi_{3,6}^{\pm 1}$, derived in [28, Section 2] for $b = -1$, are still valid for any $b \in \mathbb{F}_q^*$. However if the field $\mathbb{F}_q$ is not highly 2-adic, to compress points from $E_b^2(\mathbb{F}_q)$ (resp. $E_b^3(\mathbb{F}_q)$) we will apply in Section 4 slightly another approach based on $\mathbb{F}_q$-rationality of $GK_{1,3}$ (resp. $GK_{2,3}$). Nevertheless, since the varieties $GK_{n,3}$ are not rational for $n > 2$, we can only hope for breakthroughs concerning $\mathbb{F}_q$-rationality of $GK_{4,6}$, $GK_{5,6}$. At the same time, cryptographers rarely come across protocols, obtained by Freeman's transformation, manipulating $\mathbb{F}_q$-point vectors of length greater than three.

We know that under the condition $q \equiv 2 \pmod 3$ a curve E_b is supersingular and every element of $\mathbb{F}_q$ has a unique cubic root in $\mathbb{F}_q$. Moreover, in accordance with [33, Theorem 3.3.15] the group $E_b(\mathbb{F}_q) \simeq \mathbb{Z}/(q+1)$. Although $\varphi_{n,3}$ are no longer Galois covers, we still can find the inverse image under $\varphi_{n,3}$ via extracting a cubic root in $\mathbb{F}_q$. In particular, $\varphi_{1,3}^{-1} : \mathbb{F}_q \to E_b(\mathbb{F}_q)$ and $\varphi_{2,3}^{-1} : GK_{2,3}(\mathbb{F}_q) \to E_b^2(\mathbb{F}_q)$ are true maps, that is they are correctly defined for each input argument. The former is widely known as *Boneh–Franklin's encoding* [12, Section 8.3.2]. The latter gives rise to the new encoding $\chi_{2,3}^{-1} \colon \mathbb{F}_q^2 \to E_b^2(\mathbb{F}_q)$, because points of a (possibly reducible) $\mathbb{F}_q$-curve, where $\psi_{2,3}^{-1}$ is not defined, as usual, can be easily processed independently. Thus $\chi_{2,3}^{-1}$ allows to generate two "independent" $\mathbb{F}_q$-points on E_b twice as efficient as $\varphi_{1,3}^{-1}$ applied two times. "Independence" means that the discrete logarithm between these points is unknown to anyone.

As far as the author knows, at the moment, supersingular curves are not preferable in the pairing context, because of their small embedding degrees ($\leqslant 3$ in the characteristic $p > 3$ according to [12, Section 4.3]). The only exception is a recent *verifiable delay function (VDF)* developed in [10], where pairings are combined with isogenies. However this and other isogeny-based protocols (such as CSIDH [7]) deal with many supersingular curves. Of course, CSIDH has one starting curve, which may be of j-invariant 0, but these protocols don't require to (often) sample points on it. Perhaps, in the near future advanced isogeny-based protocols will appear for which the task of efficient regular sampling on the starting curve is important.

An idea of batch compressing points on an elliptic $\mathbb{F}_q$-curve is not new. It has already arisen in [13] for any number $n \in \mathbb{N}$ of points (and not necessarily for j-invariant 0) under the name *multiple point compression* similarly to *double* and *triple* ones in [21]. The methods of these papers compress to $n+1$ elements of $\mathbb{F}_q$, i.e., the representation is not optimal. However their decompression stages don't need to find any roots in $\mathbb{F}_q$, but only one inverse element. In comparison

with root extraction, the inversion operation has much faster Euclidean-type constant-time implementations [4], [29].

Whenever n is large, the previous approach is expected to be the best trade-off between compactness and efficiency, even despite the fact that the number of necessary multiplications in $\mathbb{F}_q$ grows quite rapidly with increasing n. This tendency is already evident in the cases $n \in \{4, 5\}$ discussed in [13, Section 3]. For simplicity, let $k := n/3 \in \mathbb{N}$. Trivially, we can apply the new triple compression method separately k times. Unfortunately, in the corresponding decompression stage it is inevitable to extract k sextic roots, since there is no technique like Montgomery's trick for multiple inversions. Nevertheless, at least for $n \in \{2, 3\}$ the new approach is the best if bandwidth/memory is more critical than speed.

2 Derivation of formulas

By analogy with [23, Theorem 9], we have

Lemma 1. *There is (up to a birational $\mathbb{F}_q$-isomorphism) the affine model*

$$GK_{2,3} = (y_1^2 - b)t^3 - (y_0^2 - b) \quad \subset \quad \mathbb{A}^3_{(t,y_0,y_1)}$$

for which the corresponding quotient map has the form

$$\varphi_{2,3} \colon E_b^2 \dashrightarrow GK_{2,3} \qquad (x_0, y_0, x_1, y_1) \mapsto \Big(\frac{x_0}{x_1}, y_0, y_1\Big).$$

Proof. Clearly, $\mathbb{F}_q(GK_{2,3}) = \mathbb{F}_q(E_b^2)^{G_{2,3}}$, that is rational functions on $GK_{2,3}$ are $G_{2,3}$-invariant ones on E_b^2. Also, consider the field

$$F := \mathbb{F}_q(t, y_0, y_1) \subset \mathbb{F}_q(GK_{2,3}), \qquad \text{where} \qquad t := \frac{x_0}{x_1}.$$

Note that $F(x_1) = \mathbb{F}_q(E_b^2)$, because $x_0 = tx_1$. Since $x_1^3 = y_1^2 - b$, the extension degree $[\mathbb{F}_q(E_b^2) : F] \leqslant 3$. At the same time, $[\mathbb{F}_q(E_b^2) : \mathbb{F}_q(GK_{2,3})] = 3$ according to Artin's theorem. Thus $F = \mathbb{F}_q(GK_{2,3})$. Finally, by looking at the equalities $t^3 = x_0^3/x_1^3 = (y_0^2 - b)/(y_1^2 - b)$, we obtain the aforementioned equation for $GK_{2,3}$. There are no other dependencies between the coordinates t, y_0, y_1, because $GK_{2,3}$ is a surface in $\mathbb{A}^3$. It remains to apply [18, Corollary I.4.5].

Theorem 1. *The generalized Kummer surface $GK_{2,3}$ is $\mathbb{F}_q$-rational.*

Proof. Let's borrow the approach used for proving [23, Theorem 12]. It is based on the theory of *conic bundles* (see, e.g., [23, Section 1.4]), but the reader can verify the formulas below (e.g., in Magma [26]) without knowledge of this theory. There is the natural conic bundle structure

$$\pi \colon GK_{2,3} \to \mathbb{A}^1_t \qquad (t, y_0, y_1) \mapsto t.$$

In other words, $GK_{2,3}$ can be seen as an $\mathbb{F}_q(t)$-conic. In a diagonal form,

$$GK_{2,3} = -y_0^2 + t^3 y_1^2 + b(1-t^3).$$

Therefore the degenerate (i.e., reducible or, equivalently, singular) fibers of π lie over $t \in \{0,\infty\} \cup \{\omega^i\}_{i=0}^2$, where $\infty := (1:0) \in \mathbb{P}^1$. More precisely, for these t we see that $\pi^{-1}(t) = L_t^+ \cup L_t^-$, where

$$L_0^{\pm} := \begin{cases} t = 0, \\ y_0 = \pm\sqrt{b}, \end{cases} \qquad L_\infty^{\pm} := \begin{cases} t = \infty, \\ y_1 = \pm\sqrt{b}, \end{cases} \qquad L_{\omega^i}^{\pm} := \begin{cases} t = \omega^i, \\ y_1 = \pm y_0. \end{cases}$$

First, after the transformation

$$\tau := \begin{cases} z_0 := y_0, \\ z_1 := t y_1, \end{cases} \qquad \tau^{-1} = \begin{cases} y_0 := z_0, \\ y_1 := z_1/t \end{cases}$$

we obtain the cubic surface

$$GK'_{2,3} := \tau(GK_{2,3}) \;=\; -z_0^2 + t z_1^2 + b(1-t^3) \quad \subset \quad \mathbb{A}^3_{(t,z_0,z_1)}.$$

We then *blow down* [18, Section V.3] one of the components $\tau(L_1^{\pm})$ by means of the transformation

$$\theta := \begin{cases} y_0 := \dfrac{z_0 - z_1}{1-t}, \\ \\ y_1 := \dfrac{z_0 - t z_1}{1-t}, \end{cases} \qquad \theta^{-1} = \begin{cases} z_0 := -t y_0 + y_1, \\ z_1 := -y_0 + y_1, \end{cases}$$

coming to

$$S := \theta(GK'_{2,3}) \;=\; t y_0^2 - y_1^2 + b(t^2+t+1) \quad \subset \quad \mathbb{A}^3_{(t,y_0,y_1)}.$$

Further, blowing down simultaneously some pair of components over $t \in \{\omega, \omega^2\}$ has the form

$$\eta := \begin{cases} z_0 := \dfrac{(t+1)y_0 + y_1}{t^2+t+1}, \\ \\ z_1 := \dfrac{t y_0 + (t+1) y_1}{t^2+t+1}, \end{cases} \qquad \eta^{-1} = \begin{cases} y_0 := (t+1) z_0 - z_1, \\ y_1 := -t z_0 + (t+1) z_1, \end{cases}$$

which gives the simpler surface

$$T := \eta(S) \;=\; t z_0^2 - z_1^2 + b \quad \subset \quad \mathbb{A}^3_{(t,z_0,z_1)}.$$

Note that the maps τ, θ, η respect the conic bundle π, that is they can be seen as $\mathbb{F}_q(t)$-isomorphisms of conics. That's why we avoid the tautology $t := t$ in their description. Finally, the projection $pr\colon T \dashrightarrow \mathbb{A}^2_{(z_0,z_1)}$ is a desired map, because $t = (z_1^2 - b)/z_0^2$.

For the compositions $\psi_{2,3} := pr \circ \eta \circ \theta \circ \tau$ and $\chi_{2,3} := \psi_{2,3} \circ \varphi_{2,3}$ Magma [26] says that

$$\chi_{2,3}\colon E_b^2 \dashrightarrow \mathbb{A}^2_{(z_0,z_1)} \qquad \chi_{2,3} = \begin{cases} z_0 := \dfrac{x_1(2x_0^2y_1 - x_0x_1(y_0 - y_1) - 2y_0x_1^2)}{y_0^2 - y_1^2}, \\ z_1 := \dfrac{x_0^3y_1 + 2x_0x_1(x_0y_1 - y_0x_1) - y_0x_1^3}{y_0^2 - y_1^2}, \end{cases}$$

$$\psi_{2,3}^{-1}\colon \mathbb{A}^2_{(z_0,z_1)} \dashrightarrow GK_{2,3} \qquad \psi_{2,3}^{-1} = \begin{cases} t := \dfrac{z_1^2 - b}{z_0^2}, \\ y_0 := \dfrac{z_0^3z_1 - 2z_0(z_0 - z_1)(z_1^2 - b) - (z_1^2 - b)^2}{z_0^3}, \\ y_1 := -\dfrac{z_0^2(z_0 - 2z_1) + (2z_0 - z_1)(z_1^2 - b)}{z_1^2 - b}. \end{cases}$$

Let's consider the cases when the denominators equal zero. Obviously, $t \in \{0, \infty\} \Rightarrow x_0x_1 = 0$, and

$$y_0^2 - y_1^2 = 0 \quad \Leftrightarrow \quad \exists k \in \mathbb{Z}/6\colon (x_1, y_1) = [-\omega]^k(x_0, y_0). \tag{1}$$

In turn, it can easily be checked that $z_0 = 0$ (i.e., $z_1 = \pm\sqrt{b}$ under the condition $t \neq 0$) if and only if $(t, y_0, y_1) \in \mathrm{Im}(\varrho_\pm)$ for the sections of π given by

$$\varrho_\pm\colon \mathbb{A}^1_t \dashrightarrow GK_{2,3} \qquad \varrho_\pm := \begin{cases} y_0 := \pm\sqrt{b}(2t + 1), \\ y_1 := \dfrac{\pm\sqrt{b}(t + 2)}{t}. \end{cases}$$

It is readily seen that

$$t = \frac{y_0 \mp \sqrt{b}}{\pm 2\sqrt{b}} = \frac{\pm 2\sqrt{b}}{y_1 \mp \sqrt{b}}$$

and eventually we get the conics

$$C_{\pm 1} := \mathrm{Im}(\varrho'_{\pm 1}) \;=\; (y_0 \mp \sqrt{b})(y_1 \mp \sqrt{b}) - 4b \quad \subset \quad \mathbb{A}^2_{(y_0,y_1)},$$

where $\varrho'_{\pm 1} := pr \circ \varrho_\pm$.

3 New compression method for two points

We need the auxiliary sets

$$V' := \big\{(x, y) \in E_b \mid xy = 0\big\} \cup \big\{\mathcal{O}\big\} \quad \subset \quad E_b[2] \cup E_b[3],$$

$$V := E_b \times V' \;\cup\; V' \times E_b,$$

where $\mathcal{O} := (0:1:0)$. Formally, for two points $P_i = (x_i, y_i)$ from $E_b(\mathbb{F}_q) \setminus V'$ the new compression map has the form

$$\mathrm{com}_{2,3}\colon E_b^2(\mathbb{F}_q) \setminus V \quad \hookrightarrow \quad \mathbb{F}_q^2 \times [0,5] \times [0,2]$$

$$\mathrm{com}_{2,3}(P_0, P_1) := \begin{cases} (x_0, y_0, k, 0) & \text{if} \quad \exists k \in \mathbb{Z}/6\colon P_1 = [-\omega]^k(P_0), \\ (t, x_1, k, 1) & \text{if} \quad \exists k \in \mathbb{Z}/2\colon (y_0, y_1) \in C_{(-1)^k}, \\ (z_0, z_1, n, 2) & \text{otherwise.} \end{cases}$$

Here $(z_0, z_1) = \chi_{2,3}(P_0, P_1)$ and $n \in [0,2]$ is the position number of the element $x_1 \in \mathbb{F}_q^*$ in the set $\{\omega^i x_1\}_{i=0}^2 \cap \mathbb{F}_q^*$ with respect to some order in $\mathbb{F}_q^*$. For example, in the case of a prime q this can be the usual numerical one. It is worth noting that in the definition of $\mathrm{com}_{2,3}$ the condition (1) is successively checked by iterating over elements of $\mathbb{Z}/6$. The same strategy is applied for the second condition $\exists k \in \mathbb{Z}/2\colon (y_0, y_1) \in C_{(-1)^k}$. Further, the set $[0,5]\times[0,2]$ clearly requires 5 bits for representing its elements. Finally, since in discrete logarithm cryptography points of small orders don't occur, we omit the definition of the compression map on $V(\mathbb{F}_q)$ for the sake of simplicity, although it can be easily defined if desired.

The corresponding decompression map is given as follows:

$$\mathrm{com}_{2,3}^{-1}\colon \mathrm{Im}(\mathrm{com}_{2,3}) \quad \xrightarrow{\sim} \quad E_b^2(\mathbb{F}_q) \setminus V$$

$$\mathrm{com}_{2,3}^{-1}(z_0, z_1, m, \ell) = \begin{cases} (z_0, z_1, x_1, y_1) & \text{if} \quad \ell = 0 \text{ and } (x_1, y_1) = [-\omega]^m(z_0, z_1), \\ (z_0 z_1, y_0, z_1, y_1) & \text{if} \quad \ell = 1 \text{ and } (y_0, y_1) = \varrho'_{(-1)^m}(z_0), \\ (t x_1, y_0, x_1, y_1) & \text{if} \quad \ell = 2 \text{ and } (t, y_0, y_1) = \psi_{2,3}^{-1}(z_0, z_1), \end{cases}$$

where for $\ell = 2$ the initial $x_1 = \sqrt[3]{g_1}$ (for $g_1 := y_1^2 - b$) can be determined with the help of $m = n$.

According to the next lemma for $q \not\equiv 1 \pmod{27}$ the cubic root $\sqrt[3]{g_1}$ can be extracted at the cost of one exponentiation in $\mathbb{F}_q$ (in particular, without inverting the denominator of g_1).

Lemma 2. *Given $g = u/v \in (\mathbb{F}_q^*)^3$ such that u, $v \in \mathbb{F}_q^*$, we obtain:*

$$\sqrt[3]{g} = \begin{cases} u \cdot (u^2 v)^{(q-2)/3} & \text{if} \quad q \equiv 2 \pmod 3, \\ u^3 \cdot (u^8 v)^{(q-4)/9} & \text{if} \quad q \equiv 4 \pmod 9, \\ u v^5 \cdot (u v^8)^{(q-7)/9} & \text{if} \quad q \equiv 7 \pmod 9, \\ \zeta u v^8 \cdot (u^2 v^{25})^{(q-10)/27} & \text{if} \quad q \equiv 10 \pmod{27}, \\ \zeta u v^{17} \cdot (u v^{26})^{(q-19)/27} & \text{if} \quad q \equiv 19 \pmod{27} \end{cases}$$

for some $\zeta \in (\mathbb{F}_q^*)^{(q-1)/9}$.

Proof. Let's consider, e.g., the case $q \equiv 7 \pmod 9$, which is relevant for the curve BLS12-377. For $e := (q+2)/9 \in \mathbb{N}$ we have:

$$g^e = u^e \cdot v^{q-1-e} = u^e \cdot v^{(8q-11)/9} = uv^5 \cdot (uv^8)^{(q-7)/9},$$
$$(g^e)^3 = g^{(q+2)/3} = g^{(q-1)/3} \cdot g = g.$$

The cases $q \equiv 4 \pmod 9$ and $q \equiv 10 \pmod{27}$ are similarly processed in [25, Equalities (2), (3)]. The remaining cases $q \equiv 2 \pmod 3$ and $q \equiv 19 \pmod{27}$ are left to the reader.

Since the projective or *Jacobian coordinates* [12, Sections 2.3.2 and 10.7.9] are preferred in practice, the decompression stage doesn't require finding inverse elements at all. By definition, in these coordinates the curve E_b possesses the equations

$$\overline{E_b}\colon Y^2Z = X^3 + bZ^3, \qquad \overline{E_b}\colon Y^2 = X^3 + bZ^6,$$

respectively. And there are the birational isomorphisms

$$\sigma\colon \overline{E_b} \xrightarrow{\sim} E_b \qquad (X : Y : Z) \mapsto \left(\frac{X}{Z}, \frac{Y}{Z}\right), \qquad (X : Y : Z) \mapsto \left(\frac{X}{Z^2}, \frac{Y}{Z^3}\right),$$

respectively. By the way, in both cases,

$$\sigma^{-1}\colon E_b \xrightarrow{\sim} \overline{E_b} \qquad (x, y) \mapsto (x : y : 1).$$

If the compression stage starts from the projective or Jacobian coordinates, then even in the classical method it is necessary to compute one inverse in $\mathbb{F}_q$. Indeed, given two points $(X_i : Y_i : Z_i) \in \overline{E_b}(\mathbb{F}_q)$ with $Z_i \neq 0$, one needs the value $v := (Z_0Z_1)^{-1}$ in order to get $Z_0^{-1} = vZ_1$ and $Z_1^{-1} = vZ_0$. This famous trick, originally attributed to Montgomery, is clearly generalized to any number of inversions.

In the compression stage of the new method instead of the two inversions v, $(y_0^2 - y_1^2)^{-1}$ only one is also enough, because

$$\chi_{2,3} \circ \sigma^{\times 2} = \left(\frac{\mathrm{num}_0}{\mathrm{den}}, \frac{\mathrm{num}_1}{\mathrm{den}}\right)\colon \quad \overline{E_b}^2 \dashrightarrow \mathbb{A}^2_{(z_0,z_1)}$$

for some polynomials num_i, $\mathrm{den} \in \mathbb{F}_q[X_i, Y_i, Z_i]_{i=0}^1$ trivially obtained from the formulas of $\chi_{2,3}$. To determine the position number n one needs to know Z_1^{-1}, hence we should in fact invert $Z_1 \cdot \mathrm{den}$. It is worth emphasizing that all of the above is equally valid for the degenerate cases $\ell \in \{0, 1\}$. To sum up, an optimized implementation of $\mathrm{com}_{2,3}^{\pm 1}$ in Magma is represented in [26] for projective coordinates and $q \equiv 7 \pmod 9$.

4 Folklore compression method for two points and its variation for three ones

First, we put $f_i := x_i^3 + b$ and $g_i := y_i^2 - b$. Since the numbers 2, 3 are relatively prime, the roots $y_0 = \sqrt{f_0}$ and $x_1 = \sqrt[3]{g_1}$ can be extracted simultaneously, that

is at the cost of a sixth root in $\mathbb{F}_q$. Indeed, for $h := f_0^3 g_1^2$ it is sufficient to compute $\alpha := \sqrt[6]{h} = \sqrt{f_0}\sqrt[3]{g_1}$, because $\sqrt[3]{g_1} = f_0 g_1/\alpha^2$ and $\sqrt{f_0} = \alpha/\sqrt[3]{g_1}$. Moreover, by analogy with [22, Section 3], whenever $q \not\equiv 1 \pmod 8$, $q \not\equiv 1 \pmod{27}$, the value α can be expressed via one exponentiation in $\mathbb{F}_q$. This follows directly from the fact that $\alpha = \sqrt[3]{\sqrt{h}}$ and from the next easily verified lemmas.

Lemma 3. *Given a quadratic residue $a \in (\mathbb{F}_q^*)^2$, we have:*

$$\sqrt{a} = \begin{cases} a^{(q+1)/4} & \text{if} \quad q \equiv 3 \pmod 4, \\ \zeta \cdot a^{(q+3)/8} & \text{if} \quad q \equiv 5 \pmod 8 \end{cases}$$

for some $\zeta \in (\mathbb{F}_q^*)^{(q-1)/4}$.

Lemma 4. *Given a cubic residue $a \in (\mathbb{F}_q^*)^3$, we have:*

$$\sqrt[3]{a} = \begin{cases} a^{(2q-1)/3} & \text{if} \quad q \equiv 2 \pmod 3, \\ a^{(8q-5)/9} & \text{if} \quad q \equiv 4 \pmod 9, \\ a^{(q+2)/9} & \text{if} \quad q \equiv 7 \pmod 9, \\ \zeta \cdot a^{(2q+7)/27} & \text{if} \quad q \equiv 10 \pmod{27}, \\ \zeta \cdot a^{(q+8)/27} & \text{if} \quad q \equiv 19 \pmod{27} \end{cases}$$

for some $\zeta \in (\mathbb{F}_q^*)^{(q-1)/9}$.

In fact, the last lemma is nothing but Lemma 2 with $v = 1$.

Thus there is the compression map

$$E_b^2(\mathbb{F}_q) \setminus V \quad \hookrightarrow \quad \mathbb{F}_q^2 \times [0,5] \qquad\qquad (P_0, P_1) \mapsto (x_0, y_1, n),$$

where $n \in [0,5]$ is the position number of the element $y_0 x_1 \in \mathbb{F}_q^*$ in the set $\big\{(-1)^i \omega^j \cdot y_0 x_1\big\}_{i=0,j=0}^{1,2} \cap \mathbb{F}_q^*$ with respect to some order in $\mathbb{F}_q^*$. As above, n is used in the decompression stage for recovering the original y_0, x_1. Notice that at the heart of this method is $\mathbb{F}_q$-rationality of

$$E_b^2/G = E_b/G_{1,2} \times E_b/G_{1,3}, \qquad \text{where} \qquad G := G_{1,2} \times G_{1,3} \simeq \mathbb{Z}/6.$$

Let's call the method *folklore*, because it doesn't require an algebraic geometry technique, so perhaps someone already knows it. The significant drawback of the folklore method consists in the fact that (in contrast to $\mathrm{com}_{2,3}$) it doesn't work efficiently over highly 2-adic fields $\mathbb{F}_q$, that is Lemma 3 cannot be leveraged. The same drawback exists for the other method [22, Sections 2-3] based on $\mathbb{F}_q$-rationality of $GK_{2,6}$. Nevertheless, since the folklore one has a slightly simpler definition, we conclude that it is more preferred for use when possible.

Similarly, one can apply the folklore methodology to the new method with z_0, z_1 in order to compress three points $P_i = (x_i, y_i)$ from $E_b(\mathbb{F}_q) \setminus V'$. As earlier, consider the set

$$V := E_b^2 \times V' \ \cup \ E_b \times V' \times E_b \ \cup \ V' \times E_b^2.$$

It is about the compression map

$$E_b^3(\mathbb{F}_q)\backslash V \quad \hookrightarrow \quad \mathbb{F}_q^3\times[0,5]\times[0,2]\times[0,1] \qquad\qquad (P_0,P_1,P_2)\mapsto(z_0,z_1,x_2,n,\ell,s),$$

where $(z_0,z_1,m,\ell)=\mathrm{com}_{2,3}(P_0,P_1)$ and in the non-degenerate case $\ell=2$ the number $n\in[0,5]$ is the position of the element $x_1y_2\in\mathbb{F}_q^*$. In turn, for $\ell\in\{0,1\}$ we put $n:=m$ and the additional sign bit s is utilized to recover y_2 (regardless of P_0, P_1). Since for these ℓ the latter points are obtained without root computations, the overall complexity doesn't go beyond one exponentiation in $\mathbb{F}_q$.

Besides, pay attention that for $\ell=2$ the root $\sqrt[6]{h}$ (where $h:=g_1^2f_2^3$) can still be found at the cost of one exponentiation in $\mathbb{F}_q$ even if the inverse of the denominator of h (i.e., of g_1^2) is unknown. By analogy with $\sqrt{\cdot}$ (see, e.g., [3, Section 5]) and $\sqrt[3]{\cdot}$ (see Lemma 2), it is explained in [24, Section 2] how to do this for $q\equiv 3\pmod 4$, $q\equiv 2\pmod 3$, or, equivalently, $q\equiv 11\pmod{12}$. For self-completeness, let's repeat the reasoning. For $e:=(q+1)/12\in\mathbb{N}$ and $h=u/v\in(\mathbb{F}_q^*)^6$ such that u, $v\in\mathbb{F}_q^*$ we obtain:

$$h^e=u^e\cdot v^{q-1-e}=u^e\cdot v^{(11q-13)/12}=uv^9\cdot(uv^{11})^{(q-11)/12},$$
$$(h^e)^6=h^{(q+1)/2}=h^{(q-1)/2}\cdot h=h.$$

The author invites the reader to independently check that this trick is easily generalized to other $q\not\equiv 1\pmod 8$, $q\not\equiv 1\pmod{27}$.

Thus we completely justified Tables 1, 2, which contain a complexity comparison (all the operations are carried out in $\mathbb{F}_q$) of the compression-decompression methods for two and three points, respectively. As is customary, the addition, subtraction, and multiplication operations in $\mathbb{F}_q$ are omitted, because they are much cheaper. Taking this opportunity, the author emphasizes that arguments of the given paper, related to avoiding the inversion operation, are equally valid for the previous compression-decompression methods. In other words, the number of inversions in [23, Theorem 13] and [22, Tables 1, 2] can be actually reduced to only one in the compression stage (at the price of several multiplications). So, inter alia, one can resort to those methods to compress the second half of the proof in Groth16.

References

1. Aranha, D.F., Pagnin, E., Rodríguez-Henríquez, F.: LOVE a pairing. In: Longa, P., Ràfols, C. (eds.) Progress in Cryptology – LATINCRYPT 2021. LNCS, vol. 12912, pp. 320–340. Springer, Cham (2021)
2. Ben-Sasson, E., Chiesa, A., Tromer, E., Virza, M.: Scalable zero knowledge via cycles of elliptic curves. In: Garay, J.A., Gennaro, R. (eds.) Advances in Cryptology – CRYPTO 2014. LNCS, vol. 8617, pp. 276–294. Springer, Berlin, Heidelberg (2014)
3. Bernstein, D.J., Duif, N., Lange, T., Schwabe, P., Yang, B.Y.: High-speed high-security signatures. Journal of Cryptographic Engineering **2**(2), 77–89 (2012)

	Galois group	compression	decompression
classical method with x_0, x_1	$G_{1,2}^2$	one inversion	two $\sqrt{\cdot}$
classical method with y_0, y_1	$G_{1,3}^2$		two $\sqrt[3]{\cdot}$
folklore method with x_0, y_1	$G_{1,2} \times G_{1,3}$		one $\sqrt[6]{\cdot}$
new method with z_0, z_1	$G_{2,3}$		one $\sqrt[3]{\cdot}$
methods from [13, Sections 2.1 and 2.2]			for free

Table 1. Worst-case complexity for compressing $\overline{E_b}^2(\mathbb{F}_q)$ (with respect to the projective or Jacobian coordinates). The last method compresses to $\approx 3\lceil \log_2(q) \rceil$ bits and the other ones compress to $\approx 2\lceil \log_2(q) \rceil$ bits.

	Galois group	compression	decompression
classical method with x_0, x_1, x_2	$G_{1,2}^3$	one inversion	three $\sqrt{\cdot}$
classical method with y_0, y_1, y_2	$G_{1,3}^3$		three $\sqrt[3]{\cdot}$
folklore-classical method with x_0, x_1, y_2	$G_{1,2}^2 \times G_{1,3}$		one $\sqrt[6]{\cdot}$ and one $\sqrt{\cdot}$
folklore-classical method with x_0, y_1, y_2	$G_{1,2} \times G_{1,3}^2$		one $\sqrt[6]{\cdot}$ and one $\sqrt[3]{\cdot}$
new method with z_0, z_1, x_2	$G_{2,3} \times G_{1,2}$		one $\sqrt[6]{\cdot}$
method from [21, Section 3.2]			for free

Table 2. Worst-case complexity for compressing $\overline{E_b}^3(\mathbb{F}_q)$ (with respect to the projective or Jacobian coordinates). The last method compresses to $\approx 4\lceil \log_2(q) \rceil$ bits and the other ones compress to $\approx 3\lceil \log_2(q) \rceil$ bits.

4. Bernstein, D.J., Yang, B.Y.: Fast constant-time gcd computation and modular inversion. IACR Transactions on Cryptographic Hardware and Embedded Systems **2019**(3), 340–398 (2019)
5. Boneh, D., Goh, E.J., Nissim, K.: Evaluating 2-DNF formulas on ciphertexts. In: Kilian, J. (ed.) Theory of Cryptography. TCC 2005. LNCS, vol. 3378, pp. 325–341. Springer, Berlin, Heidelberg (2005)
6. Boneh, D., Sahai, A., Waters, B.: Fully collusion resistant traitor tracing with short ciphertexts and private keys. In: Vaudenay, S. (ed.) Advances in Cryptology – EUROCRYPT 2006. LNCS, vol. 4004, pp. 573–592. Springer, Berlin, Heidelberg (2006)
7. Castryck, W., Lange, T., Martindale, C., Panny, L., Renes, J.: CSIDH: an efficient post-quantum commutative group action. In: Peyrin, T., Galbraith, S. (eds.) Advances in Cryptology – ASIACRYPT 2018. LNCS, vol. 11274, pp. 395–427. Springer, Cham (2018)
8. Catanese, F., Oguiso, K., Verra, A.: On the unirationality of higher dimensional Ueno-type manifolds. Revue Roumaine de Mathématiques Pures et Appliquées **60**(3), 337–353 (2015)
9. Chatterjee, S., Hankerson, D., Menezes, A.: On the efficiency and security of pairing-based protocols in the type 1 and type 4 settings. In: Hasan, M.A., Helleseth, T. (eds.) Arithmetic of Finite Fields. WAIFI 2010. LNCS, vol. 6087, pp. 114–134. Springer, Berlin, Heidelberg (2010)
10. De Feo, L., Masson, S., Petit, C., Sanso, A.: Verifiable delay functions from supersingular isogenies and pairings. In: Galbraith, S., Moriai, S. (eds.) Advances in Cryptology – ASIACRYPT 2019. LNCS, vol. 11921, pp. 248–277. Springer, Cham (2019)
11. El Housni, Y., Guillevic, A.: Optimized and secure pairing-friendly elliptic curves suitable for one layer proof composition. In: Krenn, S., Shulman, H., Vaudenay, S. (eds.) Cryptology and Network Security. CANS 2020. LNCS, vol. 12579, pp. 259–279. Springer, Cham (2020)
12. El Mrabet, N., Joye, M.: Guide to pairing-based cryptography. Cryptography and Network Security Series, Chapman and Hall/CRC, New York (2017)
13. Fan, X., Otemissov, A., Sica, F., Sidorenko, A.: Multiple point compression on elliptic curves. Designs, Codes and Cryptography **83**(3), 565–588 (2017)
14. Freeman, D.M.: Converting pairing-based cryptosystems from composite-order groups to prime-order groups. In: Gilbert, H. (ed.) Advances in Cryptology – EUROCRYPT 2010. LNCS, vol. 6110, pp. 44–61. Springer, Berlin, Heidelberg (2010)
15. Groth, J.: On the size of pairing-based non-interactive arguments. In: Fischlin, M., Coron, J.S. (eds.) Advances in Cryptology – EUROCRYPT 2016. LNCS, vol. 9665, pp. 305–326. Springer, Berlin, Heidelberg (2016)
16. Groth, J., Sahai, A.: Efficient non-interactive proof systems for bilinear groups. In: Smart, N. (ed.) Advances in Cryptology – EUROCRYPT 2008. LNCS, vol. 4965, pp. 415–432. Springer, Berlin, Heidelberg (2008)
17. Guillevic, A.: Comparing the pairing efficiency over composite-order and prime-order elliptic curves. In: Jacobson, M., Locasto, M., Mohassel, P., Safavi-Naini, R. (eds.) Applied Cryptography and Network Security. ACNS 2013. LNCS, vol. 7954, pp. 357–372. Springer, Berlin, Heidelberg (2013)
18. Hartshorne, R.: Algebraic geometry, Graduate Texts in Mathematics, vol. 52. Springer, Berlin (1977)
19. Hopwood, D.: The pasta curves for Halo 2 and beyond (2020), `https://electriccoin.co/blog/the-pasta-curves-for-halo-2-and-beyond`

20. Hopwood, D.: Pluto/Eris supporting evidence (2021), https://github.com/daira/pluto-eris
21. Khabbazian, M., Gulliver, T.A., Bhargava, V.K.: Double point compression with applications to speeding up random point multiplication. IEEE Transactions on Computers **56**(3), 305–313 (2007)
22. Koshelev, D.: Faster point compression for elliptic curves of j-invariant 0. Mathematical Aspects of Cryptography **12**(4), 115–123 (2021)
23. Koshelev, D.: New point compression method for elliptic $\mathbb{F}_{q^2}$-curves of j-invariant 0. Finite Fields and Their Applications **69**, Article 101774 (2021)
24. Koshelev, D.: Some remarks on how to hash faster onto elliptic curves (2021), https://eprint.iacr.org/2021/1082
25. Koshelev, D.: Indifferentiable hashing to ordinary elliptic $\mathbb{F}_q$-curves of $j = 0$ with the cost of one exponentiation in $\mathbb{F}_q$. Designs, Codes and Cryptography **90**(3), 801–812 (2022)
26. Koshelev, D.: Magma code (2022), https://github.com/dishport/Batch-point-compression-in-the-context-of-advanced-pairing-based-protocols
27. Lang, S.: Algebra, Graduate Texts in Mathematics, vol. 211. Springer, New York (2002)
28. Oguiso, K., Truong, T.T.: Explicit examples of rational and Calabi–Yau threefolds with primitive automorphisms of positive entropy. Journal of Mathematical Sciences, the University of Tokyo **22**, 361–385 (2015)
29. Pornin, T.: Optimized binary GCD for modular inversion (2020), https://eprint.iacr.org/2020/972
30. Rubin, K., Silverberg, A.: Compression in finite fields and torus-based cryptography. SIAM Journal on Computing **37**(5), 1401–1428 (2008)
31. Sakemi, Y., Kobayashi, T., Saito, T., Wahby, R.S.: Pairing-friendly curves (2021), https://datatracker.ietf.org/doc/draft-irtf-cfrg-pairing-friendly-curves
32. Sarkar, P.: Computing square roots faster than the Tonelli–Shanks/Bernstein algorithm (2020), https://eprint.iacr.org/2020/1407
33. Tsfasman, M., Vlăduţ, S., Nogin, D.: Algebraic geometric codes: Basic notions, Mathematical Surveys and Monographs, vol. 139. American Mathematical Society, Providence (2007)
34. Ueno, K.: Classification of algebraic varieties, I. Compositio Mathematica **27**(3), 277–342 (1973)

Appendix A. Compressing $E_b(\mathbb{F}_{q^2}) \times E_{b_2}(\mathbb{F}_q)$

Throughout the current supplementary section, we will assume that $q \equiv 1 \pmod 3$ or, equivalently, $\omega \in \mathbb{F}_q$. Unlike the main part of the paper, here the opposite situation would be drastically different as it becomes clear below. Given $\gamma \in \mathbb{F}_q^* \setminus (\mathbb{F}_q^*)^2$, let $b = b_0 + b_1\sqrt{\gamma}$ and $b_0, b_1, b_2 \in \mathbb{F}_q$ such that $bb_2 \neq 0$. Our goal is to simultaneously compress points $(x, y) = (x_0 + x_1\sqrt{\gamma}, y_0 + y_1\sqrt{\gamma})$ and (x_2, y_2) from the sets $E_b(\mathbb{F}_{q^2})$, $E_{b_2}(\mathbb{F}_q)$, respectively (here $x_j, y_j \in \mathbb{F}_q$). This problem is relevant for *pairing delegation* [1] and *type* 4 *pairings* [9, Section 3] whenever the embedding degree of the curve E_{b_2} is equal to 12. In this popular case, E_b is a sextic twist of E_{b_2} over the field $\mathbb{F}_{q^2}$. See [12, Section 3.2.5] to understand the significance of twists in pairing-based cryptography.

For compressing $E_b(\mathbb{F}_{q^2})$ it is suggested to apply the method from [23]. The given method extracts a cubic root in $\mathbb{F}_q$ in the decompression stage. Therefore the concatenation of its result z_0, z_1 with x_2 gives rise to the compression method for $E_b(\mathbb{F}_{q^2})\times E_{b_2}(\mathbb{F}_q)$ with the cost of a sextic root in $\mathbb{F}_q$, by analogy with compressing three $\mathbb{F}_q$-points in Section 4.

Table 3 exhibits a complexity comparison (all the operations are carried out in $\mathbb{F}_q$) of the compression-decompression methods for points in the projective or Jacobian coordinates. As is customary, the addition, subtraction, and multiplication operations in $\mathbb{F}_q$ are omitted, because they are much cheaper. We use the fact (e.g., from [12, Section 5.2.1]) that an inverse element (resp. square root) in $\mathbb{F}_{q^2}$ can be expressed via an inverse element (resp. two square roots) in $\mathbb{F}_q$. However, to the author's knowledge, a cubic root in $\mathbb{F}_{q^2}$ is not computed through a few radicals in $\mathbb{F}_q$. As a result, in comparison with Table 2, the new table doesn't contain the very slow methods with the coordinates y_0, y_1, x_2 or y_0, y_1, y_2.

The method of [23] is similar to the one of Sections 2, 3. It is based on $\mathbb{F}_q$-rationality of the surface

$$GK_b := \alpha(t)(y_0^2 + \gamma y_1^2 - b_0) - \beta(t)(2y_0y_1 - b_1) \quad \subset \quad \mathbb{A}^3_{(t,y_0,y_1)},$$

where $\alpha(t) := 3t^2 + \gamma$ and $\beta(t) := t(t^2 + 3\gamma)$. The latter is nothing but the generalized Kummer surface $R_b/[\omega]_2$ (up to a birational $\mathbb{F}_q$-isomorphism). Here

$$R_b = \begin{cases} y_0^2 + \gamma y_1^2 = \rho_0 := x_0^3 + 3\gamma x_0 x_1^2 + b_0, \\ 2y_0y_1 = \rho_1 := \gamma x_1^3 + 3x_0^2 x_1 + b_1 \end{cases} \quad \subset \quad \mathbb{A}^4_{(x_0,x_1,y_0,y_1)}$$

is the *Weil restriction* (see, e.g., [30, Section 4]) of E_b, equipped with the $\mathbb{F}_q$-automorphism $[\omega]_2(x_0, x_1, y_0, y_1) := (\omega x_0, \omega x_1, y_0, y_1)$ of order 3. Notice that

$$t = \frac{x_0}{x_1}, \qquad x_1 = \sqrt[3]{\frac{2y_0y_1 - b_1}{\alpha(t)}} = \sqrt[3]{\frac{y_0^2 + \gamma y_1^2 - b_0}{\beta(t)}}.$$

	compression	decompression
classical method with x_0, x_1, x_2	one inversion	three $\sqrt{\cdot}$
folklore-classical method with x_0, x_1, y_2		one $\sqrt[6]{\cdot}$ and one $\sqrt{\cdot}$
new method with z_0, z_1, x_2		one $\sqrt[6]{\cdot}$

Table 3. Worst-case complexity for compressing $\overline{E_b}(\mathbb{F}_{q^2})\times\overline{E_{b_2}}(\mathbb{F}_q)$ (with respect to the projective or Jacobian coordinates). All the methods compress to $\approx 3\lceil\log_2(q)\rceil$ bits.

Although [23] doesn't deal with the case $q \equiv 1 \pmod 4$ (including the BLS12-377 curve), it is not difficult to generalize the results to the given case if desired. We are not going to do this, because our purpose is opposite, namely to specify

the $\mathbb{F}_q$-parametrization of GK_b as clearly as possible on the example of the BLS12-381 curve ($b_0 = b_1 = 4$ and $\gamma = -1$). That makes sense, since the description in [23, Section 3.1] is not sufficiently explicit.

First, $\sqrt{6} = \sqrt{-1}\cdot\sqrt{2}\cdot\sqrt{-3} \in \mathbb{F}_q$, because $\sqrt{-3} = 2\omega + 1 \in \mathbb{F}_q$, but $\sqrt{2} \notin \mathbb{F}_q$. Indeed, $4^2 \cdot 2$ is the norm of $b = 4(1 + \sqrt{-1})$ with respect to the quadratic extension $\mathbb{F}_{q^2}/\mathbb{F}_q$ and $\sqrt{b} \notin \mathbb{F}_{q^2}$ by virtue of [23, Remark 2]. Second, there is the birational $\mathbb{F}_q$-isomorphism

$$\tau\colon GK_b \dashrightarrow \mathbb{A}^2_{(z_0,z_1)} \qquad (t, y_0, y_1) \mapsto \Big(\frac{\mathrm{num}_{z_0}}{\mathrm{den}_z}, \frac{\mathrm{num}_{z_1}}{\mathrm{den}_z}\Big),$$

where

$$\mathrm{num}_{z_0} := f_0(t)y_0 + f_1(t)y_1, \qquad \mathrm{num}_{z_1} := -\sqrt{6}\cdot\alpha(t)(t^2 - 4t + 1),$$

$$\mathrm{den}_z := g_0(t)y_0 + g_1(t)y_1,$$

and

$$\begin{aligned}
f_0(t) &:= 6\big(\,(7\sqrt{6} - 13)t^3 - 13t^2 + (3\sqrt{6} - 1)t - 1\,\big),\\
f_1(t) &:= 3\sqrt{6}\cdot\alpha(t)\big(\,(\sqrt{6} - 3)t^2 + \sqrt{6}\cdot t - 1\,\big),\\
g_0(t) &:= 3\big(\,(\sqrt{6} + 2)t^4 + 2t^3 - 2(4\sqrt{6} - 5)t^2 + 10t - \sqrt{6}\,\big),\\
g_1(t) &:= 6\alpha(t)\big(\,(\sqrt{6} - 1)t - 1\,\big).
\end{aligned}$$

It turns out that

$$\tau^{-1}\colon \mathbb{A}^2_{(z_0,z_1)} \dashrightarrow GK_b \qquad (z_0, z_1) \mapsto \Big(\frac{\mathrm{num}_t}{\mathrm{den}_t}, \frac{\mathrm{num}_{y_0}}{\mathrm{den}_y}, \frac{\mathrm{num}_{y_1}}{\mathrm{den}_y}\Big),$$

where

$$\mathrm{num}_t := z_0^2 + 12z_1^2 - 1, \qquad \mathrm{den}_t := -2(z_0 + 6z_1^2), \qquad \mathrm{den}_y := -\sqrt{6}\cdot\alpha(t)(t^2 + 1),$$

$$\mathrm{num}_{y_0} := \alpha(t)\big(F_0(t)Z_0 + F_1(t)Z_1\big), \qquad \mathrm{num}_{y_1} := G_0(t)Z_0 + G_1(t)Z_1,$$

$$Z_0 := \frac{z_0\cdot\mathrm{den}_t + \mathrm{num}_t}{z_1\cdot\mathrm{den}_t}, \qquad Z_1 := \frac{1}{z_1}$$

and

$$\begin{aligned}
F_0(t) &:= 2\big(\,(\sqrt{6} - 1)t - 1\,\big), \qquad F_1(t) := (\sqrt{6} - 4)t^2 - 4t + \sqrt{6},\\
G_0(t) &:= -(\sqrt{6} + 2)t^4 - 2t^3 + 2(4\sqrt{6} - 5)t^2 - 10t + \sqrt{6},\\
G_1(t) &:= (\sqrt{6} + 2)t^5 + 2t^4 + 2(3\sqrt{6} - 8)t^3 - 16t^2 + (5\sqrt{6} - 2)t - 2.
\end{aligned}$$

All the written formulas are checked in Magma [26]. As usual, to compress any points from $E_b(\mathbb{F}_{q^2})$ it remains to process the degenerate cases when the denominators equal zero. In order not to complicate the text this is left as an elementary exercise.

Appendix B. Compressing $E_b(\mathbb{F}_{q^2})$ sub-optimally in such a way that decompressing is for free

Let's stick to the notation of the previous section. This one contains formulas obtained in the same way as in [21, Section 3.1] for compressing $E_b^2(\mathbb{F}_q)$ to $\approx 3\lceil\log_2(q)\rceil$ bits. The new formulas are very simple and important, but the author didn't find them anywhere else. So the appendix is a good place to write out them. Probably, the similar approach from [21, Section 3.2] in the 3-dimensional case may be also adapted for compressing $E_b(\mathbb{F}_{q^2})\times E_{b_2}(\mathbb{F}_q)$ to $\approx 4\lceil\log_2(q)\rceil$ bits.

Given a non-zero point $P=(x,y)\in E_b(\mathbb{F}_{q^2})$, consider the $\mathbb{F}_q$-elements

$$Y:=y_0+y_1,\qquad Y_1:=\frac{2(\rho_0-Y^2)+(\gamma+1)\rho_1}{2(\gamma-1)Y}.$$

Obviously, $\gamma\neq 1$. By looking at the defining equations of R_b, it is readily checked (see also [26]) that $y_1=Y_1$ whenever $Y\neq 0$. Therefore we get the compression map

$$\mathrm{com}\colon E_b(\mathbb{F}_{q^2})\setminus\{\mathcal{O}\}\quad\hookrightarrow\quad \mathbb{F}_q^3\times\{0,1\}$$

$$\mathrm{com}(P):=\begin{cases}(x_0,x_1,y_1,0) & \text{if}\quad Y=0,\\ (x_0,x_1,Y,1) & \text{otherwise.}\end{cases}$$

The corresponding decompression map has the form

$$\mathrm{com}^{-1}\colon \mathrm{Im}(\mathrm{com})\quad\twoheadrightarrow\quad E_b(\mathbb{F}_{q^2})\setminus\{\mathcal{O}\}$$

$$\mathrm{com}^{-1}(x_0,x_1,Y',\mathrm{bit})=\begin{cases}(x_0,x_1,-Y',Y') & \text{if}\quad \mathrm{bit}=0,\\ (x_0,x_1,Y'-Y_1,Y_1) & \text{if}\quad \mathrm{bit}=1.\end{cases}$$

Table 4 exhibits a complexity comparison (all the operations are carried out in $\mathbb{F}_q$) of the compression-decompression methods for $\mathbb{F}_{q^2}$-points on E_b. It is worth noting that all the remarks given for Table 3 still hold for the new one.

	compression	decompression
classical method with x_0, x_1	one inversion	two $\sqrt{\cdot}$
method from [23]		one $\sqrt[3]{\cdot}$
method from [22, Section 4] with $\frac{x_0}{x_1}$, $\frac{y_0}{y_1}$		one $\sqrt[6]{\cdot}$
new method with x_0, x_1, Y		for free

Table 4. Worst-case complexity for compressing $\overline{E_b}(\mathbb{F}_{q^2})$ (with respect to the projective or Jacobian coordinates). The last method compresses to $\approx 3\lceil\log_2(q)\rceil$ bits and the other ones compress to $\approx 2\lceil\log_2(q)\rceil$ bits.