Scalar Multiplication on Koblitz Curves using Double Bases

Roberto Avanzi[1] and Francesco Sica[2] *

[1] Institute for Cryptology and IT-Security
The Horst Görtz institute (HGI) of IT-Security
Ruhr-Universität Bochum
Universitätsstraße 150, D-44780 Bochum, Germany
roberto.avanzi@ruhr-uni-bochum.de
[2] Mount Allison University – AceCrypt
Department of Mathematics and Computer Science
67 York Street, Sackville, NB, E4L 1E6, Canada
fsica@mta.ca – http://www.acecrypt.com

Abstract. The paper is an examination of double-base decompositions of integers n, namely expansions loosely of the form

$$n = \sum_{i,j} A^i B^j$$

for some base $\{A, B\}$. This was examined in previous works [3, 4], in the case when A, B lie in $\mathbb{N}$.
On the positive side, we show how to extend the results of [3] to Koblitz curves over binary fields. Namely, we obtain a sublinear scalar algorithm to compute, given a generic positive integer n and an elliptic curve point P, the point nP in time $O\left(\frac{\log n}{\log\log n}\right)$ elliptic curve operations with essentially no storage, thus making the method asymptotically faster than any know scalar multiplication algorithm on Koblitz curves.
On the negative side, we analyze scalar multiplication using double base numbers and show that on a generic elliptic curve over a finite field, we cannot expect a sublinear algorithm. Finally, we show that all algorithms used hitherto need at least $\frac{\log n}{\log\log n}$ curve operations.

1 Introduction

In cryptographic algorithms making use of elliptic curves, the costliest part is the scalar multiplication nP, where P lies on the curve. In order to speed up this computation, it was proposed already at the very beginning of their use to adopt special kinds of elliptic curves where a large multiple of P can be computed very quickly. This is the case of endomorphism

* This work was partially supported by a NSERC Discovery Grant

curves [6] or Koblitz curves E_a (also called ABC-curves or anomalous curves) [8].

We will examine more closely this latter class of curves, defined over $\mathbb{F}_{2^{\mathbf{p}}}$. There exists a map (an endomorphism of the group of points $E_a(\mathbb{F}_{2^{\mathbf{p}}})$ of the curve), called the Frobenius endomorphism and denoted by τ, such that τP is a large multiple of P which can be computed in time $O(1)$ using normal bases or $O(\mathbf{p})$ using polynomial bases. Using τ, one can achieve good scalar multiplication algorithms, see Section 3. However, all these algorithms compute nP with[3] $\Omega(\log n)$ costly curve operations (such as a doubling or an addition). We call these algorithms linear (in the number of curve operations with respect to the bit size of the field), since also the number of curve operations is $O(\log n)$.

The novelty of our approach is to combine the use of τ with double bases, which were introduced in elliptic curve cryptography in [5]. We show how to find a decomposition

$$n = \sum_{i=1}^{\ell} (-1)^{e_i} \tau^{s_i} 3^{t_i}$$

with s_i, t_i nonnegative integers and $e_i \in \{0, 1\}$. The length ℓ of this expansion is $O(\log n / \log\log n)$. We then proceed similarly to [3] to reveal a scalar multiplication algorithm whose cost is $O(\log n / \log\log n)$ curve operations. We call such an algorithm sublinear, where the number of curve operations over the bit size of the field goes to zero.

This is a first instance of a sublinear scalar multiplication algorithm with very little precomputations (which depend only on $\mathbf{p}$, not the curve or the point P) or storage requirements ($O(\log \mathbf{p})$ bits).

2 Preliminaries

2.1 Koblitz Curves

A Koblitz curve E_a is an elliptic curve defined over $\mathbb{F}_{2^{\mathbf{p}}}$, with Weierstrass equation

$$E_a \quad : \quad y^2 + xy = x^3 + ax^2 + 1. \tag{1}$$

Here $a = 0$ or 1, and $\mathbf{p}$ is a prime chosen so to make the order of the group of points $E_a(\mathbb{F}_{2^{\mathbf{p}}})$ equal to twice if $a = 1$ (resp. four times if $a = 0$) a prime number, for at least one choice of a. A point $P \in E_a(\mathbb{F}_{2^{\mathbf{p}}})$ is then randomly chosen with order equal to that large prime. Note that in view

[3] We use the notation $\Omega(x)$ to mean $> cx$ for some positive c.

of Hasse's theorem, which states that $|\#E_a(\mathbb{F}_{2^{\mathbf{p}}}) - 2^{\mathbf{p}} - 1| < 2^{\frac{\mathbf{p}}{2}+1}$, $\operatorname{ord} P$ is very close to $2^{\mathbf{p}-1}$ if $a = 1$ and to $2^{\mathbf{p}-2}$ if $a = 0$.

Since E_a has coefficients in $\mathbb{F}_2$, the Frobenius map $\tau(x, y) = (x^2, y^2)$ is an endomorphism of $E_a(\mathbb{F}_{2^{\mathbf{p}}})$. Since squaring is a linear operation in characteristic two, computing τP is also linear and takes time at most $O(\mathbf{p})$. If normal bases are used to represent elements of $\mathbb{F}_{2^{\mathbf{p}}}$, then computing τP is much faster, since it amounts to making two rotations, which is essentially free.

We can view τ as a complex number of norm 2 satisfying the quadratic equation $\tau^2 - (-1)^{1-a}\tau + 2 = 0$, since for any P on the curve, $\tau^2 P + 2P = (-1)^{1-a}\tau P$. Explicitly,

$$\tau = \frac{(-1)^{1-a} + \sqrt{-7}}{2}.$$

It does not matter which "determination" of the square root we use, since it is the algebraic properties of τ that we need, hence we will fix $\mathrm{Im}\sqrt{-7} > 0$.

2.2 Continued Fractions

Continued fractions are a way to find very good rational approximations p_s/q_s (in terms of the maximum of the absolute values of p_s and q_s) to arbitrary real numbers, by an algorithmic process which generalizes the computation of the greatest common divisor (gcd) of two integers.

We list the properties of p_s/q_s, called the s-th convergent to α, relevant to this paper. There exists a sequence of positive integers $(a_s)_{s \geq 1}$ with

$$p_s = a_s p_{s-1} + p_{s-2} \quad \text{and} \quad q_s = a_s q_{s-1} + q_{s-2} \quad \text{for all } s \geq 1.$$

Therefore $q_s \geq q_{s-1} + q_{s-2}$ and similarly for p_s. These two sequences have at least a Fibonacci-like (exponential) growth. If $\alpha \notin \mathbb{Q}$, we have the following inequalities for all $s \geq 1$

$$0 < \alpha - \frac{p_{2s}}{q_{2s}} < \frac{1}{q_{2s}^2} \quad \text{and} \quad -\frac{1}{q_{2s-1}^2} < \alpha - \frac{p_{2s-1}}{q_{2s-1}} < 0 \ .$$

In particular, note that $\lim_{s\to\infty} p_s/q_s = \alpha$.

2.3 Measure of Irrationality

We begin with a famous result (usually proved with the "pigeon-hole" or box principle).

Theorem 1 (Dirichlet-Legendre). *Let $Q > 1$ and $\alpha \in \mathbb{R}$. There exist integers $0 < q < Q$ and $p \in \mathbb{Z}$ such that*

$$|q\alpha - p| < \frac{1}{Q}$$

The irrationality measure $\mu(\alpha)$ of $\alpha \in \mathbb{R} - \mathbb{Q}$ is defined as

$$\mu(\alpha) = \sup \left\{ r \in \mathbb{R} \colon \exists \infty \, (p,q) \in \mathbb{Z}^2 \text{ with } \left| \alpha - \frac{p}{q} \right| \leq \frac{1}{q^r} \right\} \ .$$

Notice that the convergents $\dfrac{p}{q}$ of the continued fraction expansion of α satisfy

$$\left| \alpha - \frac{p}{q} \right| \leq \frac{1}{q^2} \ ,$$

hence $\mu(\alpha) \geq 2$. It is known that the set of reals with irrationality measure greater than 2 has Lebesgue measure zero. Therefore, given α, we should conjecture that $\mu(\alpha) = 2$.

In the rest of the paper, we will then assume that the irrational numbers $\log_2 3$ and θ/π (see below for definition) have measure 2.

2.4 Double Bases

Following [4] we will call a $\{A, B\}$-integer a number which can be written as $A^i B^j$ for some nonnegative integers i, j. We will extend the definition to algebraic integers, more precisely, integers in $\mathbb{Z}[\tau]$. We will also allow $A, B \in \mathbb{Z}[\tau]$. We define a $\{A, B\}$-integer expansion of n as a decomposition of n into a sum of (possibly signed) $\{A, B\}$-integers.

3 Scalar Multiplication on Koblitz Curves

In this Section we are chiefly concerned with scalar multiplication techniques that do not make use of point precomputations or storage for on-the-fly computed tables of point multiples. Variants of these methods than can take advantage of such devices exist and in many cases have been extensively treated in the literature.

3.1 The τ-NAF

All facts here are stated without proofs: These are found in [12, 13].

Let us consider the Koblitz curve E_a defined over $\mathbb{F}_{2^{\mathbf{p}}}$ by equation (1), with base point P, and let τ denote the Frobenius endomorphism. We have seen that we can view $\tau(P)$ as *multiplication by* τ and let $\mathbb{Z}[\tau]$ operate on P, but in fact there exists an integer λ such that $\tau(P) = \lambda P$, and thus τ operates on the whole subgroup generated by P like multiplication by λ.

The τ-adic non-adjacent form (τ-NAF for short) of an integer $z \in \mathbb{Z}[\tau]$ is a decomposition $z = \sum_i z_i \tau^i$ where $z_i \in \{0, \pm 1\}$ with the *non-adjacency* property $z_j z_{j+1} = 0$, similarly to the classical NAF [10]. The average *density* (that is the average ratio of non-zero bits related to the total number of bits) of a τ-NAF is 1/3. Each integer z admits a unique τ-NAF.

The length of the τ-NAF expansion of a randomly chosen scalar n is $\approx 2\mathbf{p}$, whereas the bit length of n is $\approx \mathbf{p}$. But, for any point $P \in E_a(\mathbb{F}_{2^{\mathbf{p}}}) \setminus E_a(\mathbb{F}_2)$, $\tau^{\mathbf{p}} P = P$ and $\tau P \neq P$.

Since $\mathbb{Z}[\tau]$ is an Euclidian ring we can take the remainder ζ of n mod $(\tau^{\mathbf{p}} - 1)/(\tau - 1)$ and use it in place of n. This remainder will have smaller norm than that of $(\tau^{\mathbf{p}} - 1)/(\tau - 1)$, and thus it will have length at most $\mathbf{p}$. Its τ-NAF is called the *reduced* τ-NAF of n.

The double-and-add scalar multiplication algorithm is just a Horner scheme for the evaluation of nP using the binary expansion of $n = \sum_{i=0}^{\ell} n_i 2^i$ as $\sum_{i=0}^{\ell} n_i 2^i P$. In a similar way we can evaluate $zP = \sum_i z_i \tau^i(P)$ by a Horner scheme, and the the corresponding algorithm is called a τ-and-add algorithm. It is much faster than the double-and-add scheme on Koblitz curves because Frobenius evaluations are much faster than doublings.

3.2 Inserting a Doubling

Point halving [7, 11] is the inverse operation to point doubling and applies to *all* elliptic curves over binary fields, not only to Koblitz curves. Its evaluation is 2 to 3 times faster than that of a doubling and it is possible to rewrite the scalar multiplication algorithm using halving instead of doubling. The resulting method is very fast, but on Koblitz curves it is slower than the τ-and-add method.

In [1] it is proposed to insert a halving in the "τ-and-add" method to further speed up scalar multiplication. This approach brings a non-negligible speedup (on average 14% with suitable representations of the fields) with respect to the use of the τ-NAF, but it is not optimal. In [2] the method was refined in order to bring the speed-up to 25%, and the resulting method was proved to be optimal among the methods that do not require any precomputation.

The fundamental idea in these methods is that in order to perform $\sum_i e_i\tau^i(P) + \sum_i f_i\tau^i(Q)$ (with $e_i, f_i \in \{0, \pm 1\}$) where a linear relation $Q = f(P)$ exists, one can first compute $R = \sum_i f_i\tau^i(P)$, apply f to the result, and "resume" the τ-and-add loop corresponding to $\sum_i e_i\tau^i(P)$ starting however with $f(R)$ instead of setting the intermediate value to the zero point at the beginning.

In all these techniques the number of Frobenius applications is doubled with respect to the standard τ-and-add method, which is not a problem if the ground field is represented by a normal basis. In both papers it is shown that the performance is still improved if a polynomial basis is used: the expected speed-ups are around 12% and 21% for the methods in [1] and [2] respectively.

3.3 Other Development

Vuillaume, Okeya and Takagi in [9] generalize the approach of the previous Subsection applying it to expressions of the form $\sum_i e_{1,i}\tau^i(P) + \sum_i e_{3,i}\tau^i(f_3(P)) + \sum_i e_{5,i}\tau^i(f_5(P)) + \ldots$ where $f_3(1)$, $f_5(1)$, etc... form a complete set of residues modulo a suitable power of τ in the ring $\mathbb{Z}[\tau]$, and the $e_{j,i} \in \{0, \pm 1\}$. Such an expression can be easily obtained from a modified τ-adic windowing method, and if a window width w is used, then the τ-and-add loop must be resumed 2^{w-1} times. The method becomes then quickly impractical if a polynomial basis is used, because then Frobenius operations quickly become the dominant part of the computation, but the method has its merits if a normal basis is used. Some questions remain open: the relations f_j and their inverses must be described in an easy way, and the approch used in the paper works only for a few window sizes. Hence the authors cannot present the results in a completely general way. In the cases that have been described the reduction in memory consumption (or, equivalently, the speed-up with respect to other methods with no precomputations) are noteworthy.

4 Scalar Multiplication using Double Bases: Previous Work

Let p_s/q_s be the s-th convergent to $\log_3 2/2$ (this is a slight departure from [3]). Let $m = \mathbf{p}^{2/5}$. Fix s as the first odd index such that $p_s > m$. Then

$$0 < \frac{2}{m^{1+\epsilon}\log 2} < p_s\frac{2\log 3}{\log 2} - q_s < \frac{1}{m} < \frac{2}{m\log 2} \quad .$$

This shows the following lemma.

Lemma 1. *Using the above notations we have, as* $\mathbf{p} \to \infty$

$$\exp\left(\frac{1}{m^{1+\epsilon}}\right) < \frac{3^{p_s}}{2^{\frac{q_s}{2}}} < \exp\left(\frac{1}{m}\right) \quad . \tag{2}$$

The authors of [3] then use this lemma to prove the following "reduction" theorem (which we cite after fixing the value of m).

Theorem 2. *Let* n *be a large integer. There exists a* $\{2,3\}$*-integer* N *satisfying*

$$|n - N| < \frac{n}{\log^{\frac{1}{3}} n}$$

Repeated use of this theorem leads to an effective construction of a $\{2,3\}$-integer decomposition of n as in the following.

Theorem 3. *Every sufficiently large number* n *can be written as a sum*

$$n = \sum_{i=1}^{k} 2^{s_i} 3^{t_i}, \quad s_i, t_i \in \mathbb{N} \cup \{0\}$$

with $(s_i, t_i) \neq (s_j, t_j)$ *for* $i \neq j$ *and*

$$k \leq 3\,\frac{\log n}{\log\log n} + o\left(\frac{\log n}{\log\log n}\right) \quad .$$

Moreover, one can insure that $\max_i s_i \leq \log^{2/3+\epsilon} n$.

This last theorem allows to build a sublinear scalar multiplication algorithm, very similar to our Algorithm 3. That work forms the blueprint of our sublinear scalar multiplication algorithm for Koblitz curves.

5 Double Base $\{\tau, 3\}$ Expansions

The aim of the following sections is to produce an efficient decomposition of a scalar n as a signed sum of $\{\tau, 3\}$-integers

$$n = \sum_{i=1}^{\ell} (-1)^{e_i} \tau^{s_i} 3^{t_i}$$

with s_i, t_i nonnegative integers, $(s_i, t_i) \neq (s_j, t_j)$ for $i \neq j$ and $e_i \in \{0, 1\}$. Here $\ell = O(\log n / \log\log n)$.

As a first simplification we replace n its reduced τ-NAF, ζ, as defined in Section 3. This allows to cut by half the representation of n, since $nP = \zeta P$ on the curve.

Using a lexicographic order on powers of 3 and τ one can rewrite such a $\{\tau, 3\}$ expansion as

$$\zeta = \sum_{i=1}^{\mathcal{I}} 3^{t_i} \sum_{j=1}^{\mathcal{J}_i} (-1)^{e_{i,j}} \tau^{s_{i,j}} \tag{3}$$

where $e_{i,j} \in \{0,1\}$,

$$\sum_{i=1}^{\mathcal{I}} \sum_{j=1}^{\mathcal{J}_i} 1 = \ell \ , \quad t_i > t_{i+1} \quad \text{and} \quad s_{i,j} > s_{i,j+1} \ .$$

6 A New Scalar Multiplication

In this section we generalize the algorithms of [3] to ordinary Koblitz curves defined over $\mathbb{F}_{2^{\mathfrak{p}}}$. The main difference is that we have to view τ as a complex number, which requires controlling the argument of the numbers thus involved if we want to find "close" $\{\tau, 3\}$-numbers.

Let τ be the Frobenius endomorphism of E_a. Call $\theta = \arg(\tau)$. We first prove the following easy result.

Lemma 2. $\dfrac{\theta}{\pi} \notin \mathbb{Q}$.

Proof. We want to show that $\tau^a \notin \mathbb{R}$ for any $a \in \mathbb{Z}$. Let there otherwise be some such a. Let $M = \tau^a$. Taking norms, we get $M = \pm 2^{a/2}$ and a is even. But then $\tau^{a/2}\bar{\tau}^{a/2} = \pm\tau^a$ or $\bar{\tau}^{a/2} = \pm\tau^{a/2}$ which is impossible, since $\mathbb{Z}[\tau]$ is a unique factorization domain (it is Euclidean) and τ and $\bar{\tau}$ are two non-associated irreducibles. □

Theorem 4. *Let $\zeta \in \mathbb{Z}[\tau]$ be large. There exists a $\{\tau, 3\}$-number N satisfying either*

$$|\zeta - N| \le \frac{|\zeta|}{\log^{\frac{2}{25}} |\zeta|} \quad \mathit{or} \quad |\zeta + N| \le \frac{|\zeta|}{\log^{\frac{2}{25}} |\zeta|}$$

Proof. In view of (2) we have[4]

$$\left| \frac{3^{p_s}}{\tau^{q_s}} \right| \approx e^{\frac{1}{m}}$$

[4] We write something is $\approx f(m)$ for some function f to mean that it lies between $f(m^{1-\epsilon})$ and $f(m^{1+\epsilon})$. Similarly with $\mathfrak{m}$ instead of m. This will avoid notation cluttering, while giving enough indications for a complete technical proof.

We then take the largest power 2^ν less than or equal to $|\zeta|$. Define t as the largest integer such that

$$\left|\frac{3^{p_s}}{\tau^{q_s}}\right|^t \leq \frac{|\zeta|}{2^\nu} \tag{4}$$

and

$$q_s t < 2\nu \tag{5}$$

Then $\tilde{N} = \tau^{2\nu - tq_s} 3^{tp_s}$ satisfies

$$1 \leq \left|\frac{\zeta}{\tilde{N}}\right| \leq e^{\frac{1}{m}}$$

Unlike in the supersingular case we cannot conclude that $|\zeta - \tilde{N}|$ is small, because we need to adjust the argument of $\tilde{N}$. We will rely on the following result.

Lemma 3. *Let* ξ_1, ξ_2 *be two nonzero complex numbers and* $\mathfrak{m} \geq 3$ *such that* $1 \leq |\xi_1/\xi_2| \leq e^{\frac{1}{\mathfrak{m}}}$ *and* $\cos \arg(\xi_1/\xi_2) \geq e^{-\frac{1}{\mathfrak{m}}}$. *Then*

$$|\xi_1 - \xi_2| \leq \frac{2|\xi_2|}{\sqrt{\mathfrak{m}}}$$

Proof. See Appendix A.

We now find an integer $u \geq 0$ such that there exists an integer v with

$$|uq_s\theta - 2v\pi| < \frac{1}{\sqrt{\mathfrak{m}}}.$$

We can do this by looking at the continued fraction expansion of $q_s\theta/2\pi$ which is irrational by Lemma 2. The previous inequality becomes

$$\left|u\frac{q_s\theta}{2\pi} - v\right| < \frac{1}{2\pi\sqrt{\mathfrak{m}}}.$$

By the Dirichlet-Legendre theorem, v/u can be chosen as the convergent to $q_s\theta/2\pi$ with $u < 2\pi\sqrt{\mathfrak{m}}$ closest to this bound. By our assumption on irrationality measures, actually

$$u \approx \sqrt{\mathfrak{m}}$$

and

$$|uq_s\theta - 2v\pi| \approx \frac{1}{\sqrt{\mathfrak{m}}}. \tag{6}$$

This u can actually be precomputed, as it will depend only on the size $2^{\mathbf{p}}$ of the finite field (see below), not even on the curve. Define then

$$-k = \left\lfloor \frac{\arg_\pi \zeta - (2\nu - tq_s)\theta}{uq_s\theta - 2v\pi} \right\rceil \leq 0,$$

where $-\pi < \arg_\pi \zeta - (2\nu - tq_s)\theta < \pi$ is defined modulo π to make k non-negative. Then $k = O(\sqrt{\mathfrak{m}})$ and

$$|kuq_s\theta + \arg_\pi \zeta - (2\nu - tq_s)\theta - 2kv\pi| < \frac{1}{\sqrt{\mathfrak{m}}}.$$

Define now

$$N = \tilde{N} \left(\frac{3^{p_s}}{\tau^{q_s}} \right)^{ku} = \tau^{2\nu - (t+ku)q_s} 3^{(t+ku)p_s}. \tag{7}$$

Note that, if $\mathfrak{m}$ is small enough, N is a $\{\tau, 3\}$-integer. Also, either $|\arg(N/\zeta)| < \frac{1}{\sqrt{\mathfrak{m}}}$ or $|\arg(-N/\zeta)| < \frac{1}{\sqrt{\mathfrak{m}}}$ and thus we get $|\cos \arg(N/\zeta)| > e^{-\frac{1}{\mathfrak{m}}}$. Also,

$$1 \leq \left| \frac{N}{\zeta} \right| \leq \left| \frac{N}{\tilde{N}} \right| \leq e^{\frac{O(\mathfrak{m})}{m}} = e^{O\left(\frac{\mathfrak{m}}{m}\right)}$$

Thus choosing $\mathfrak{m} \leq m^{1/2-\epsilon}$, we can apply Lemma 3 to $\xi_1 = N$ or $\xi_1 = -N$ and $\xi_2 = \zeta$ to conclude that

$$|\zeta - N| \leq \frac{2|\zeta|}{m^{1/4-\epsilon}} \quad \text{or} \quad |\zeta + N| \leq \frac{2|\zeta|}{m^{1/4-\epsilon}}$$

hence Theorem 4, with $\mathfrak{m} = m^{2/5} = \mathbf{p}^{4/25}$.

Repeated applications of this theorem will give the next result, whose proof follows, *mutatis mutandis*, that of Theorem 3.

Theorem 5. *Every $\zeta \in \mathbb{Z}[\tau]$ with $\zeta\bar{\zeta} < \#E_a(\mathbb{F}_{2^{\mathbf{p}}})$ can be written as a sum*

$$\zeta = \sum_{i=1}^{\ell} (-1)^{e_i} \tau^{s_i} 3^{t_i}$$

with s_i, t_i nonnegative integers, $(s_i, t_i) \neq (s_j, t_j)$ for $i \neq j$ and $e_i \in \{0, 1\}$. Furthermore the length of the expansion is

$$\ell \leq 12.5 \frac{\mathbf{p}}{\log_2 \mathbf{p}} + o\left(\frac{\mathbf{p}}{\log \mathbf{p}} \right)$$

and one can insure that $\max_i t_i \leq \mathbf{p}^{4/5}$.

```
Input: An integer $\mathbf{p}$, the bit size of the ground field $\mathbb{F}_{2^{\mathbf{p}}}$.
Output: Three integers P_CONV, Q_CONV and U_CONV, and two floating point numbers MODULUS_RATIO and ANGLE_RATIO.

$m \leftarrow \mathbf{p}^{2/5}$
$s \leftarrow \min\{2j+1 \colon p_{2j+1} > m\}$
P_CONV $\leftarrow p_s$
Q_CONV $\leftarrow q_s$
MODULUS_RATIO $\leftarrow 2p_s \log_2 3 - q_s$
MODULUS_RATIO $\leftarrow 1/$MODULUS_RATIO
ANGLE_RATIO $\leftarrow uq_s\theta - 2v\pi$, as per $(6)$
ANGLE_RATIO $\leftarrow 1/$ANGLE_RATIO
U_CONV $\leftarrow u$
```

Algorithm 1. Precomputations (depending on $\mathbf{p}$)

In view of the fact that, by Hasse's theorem,

$$\mathbf{p} = \log_2 \#E_a(\mathbb{F}_{2^{\mathbf{p}}}) + O\left(\#E_a(\mathbb{F}_{2^{\mathbf{p}}})^{-1/2}\right) = \log_2 n + O(1)$$

on average for n, this is the analogue of Theorem 3 in our context. It is this constructive theorem which is responsible for the sublinear running time of Algorithm 3, with the same analysis as for Algorithm 2 in [3]. See also the next section for details.

Remark 1. We should note that Theorem 5 also holds for unsigned expansions, with the same constant. Hence we expect in practice a smaller value for the signed expansion, as suggested by practical examination. although at the moment is seems difficult to produce one. The reason to describe a signed expansion rather than an unsigned one is to have a significant speedup for even the smallest elliptic curves, such as K-163 and K-233.

7 Practical Estimates

Algorithms 1, 2 and 3 describe respectively the initial precomputation, the scalar recoding and the actual scalar multiplication. We draw some remarks concerning their application.

Input: An integer $\zeta \in \mathbb{Z}[\tau]$ with $2^{\mathbf{p}/4} < |\zeta| < 2^{\mathbf{p}}$, and constants P_CONV, Q_CONV, U_CONV and MODULUS_RATIO, ANGLE_RATIO.
Output: A set $\mathcal{S} = \{(e_1, s_1, t_1), \ldots, (e_\ell, s_\ell, t_\ell)\}$ with the property that $\zeta = \sum_{i=1}^{\ell} (-1)^{e_i} \tau^{s_i} 3^{t_i}$ with s_i, t_i nonnegative integers, $e_i \in \{0, 1\}$ and $\ell = O(\log n / \log\log n)$

```
$\mathcal{S} = \emptyset$
While $|\zeta| > 2^{\mathbf{p}^{4/5}}$ do
  $\epsilon \leftarrow 0$
  $\nu \leftarrow \lfloor \log_2 |\zeta| \rfloor$
  $t \leftarrow \lfloor 2\text{MODULUS\_RATIO}(\log_2 |\zeta| - \nu) \rfloor$
  $k_0 \leftarrow \operatorname{Arg}(\zeta / \tau^{2\nu - t\text{Q\_CONV}})$
  If $k_0\text{ANGLE\_RATIO} > 0$
    $k_0 \leftarrow k_0 - \operatorname{sign}(\text{ANGLE\_RATIO})\pi$
    $\epsilon \leftarrow 1$
  $k \leftarrow -\lfloor k_0\text{ANGLE\_RATIO} \rceil$
  $c \leftarrow t + k\text{U\_CONV}$
  $N \leftarrow \tau^{2\nu - c\text{Q\_CONV}} 3^{c\text{P\_CONV}}$
  $\mathcal{S} \leftarrow \mathcal{S} \cup \{(\epsilon, 2\nu - c\text{Q\_CONV}, c\text{P\_CONV})\}$
  $\zeta \leftarrow \zeta - (-1)^{\epsilon} N$
Find the $\tau$-NAF of $\zeta$, appending exponents and signs to $\mathcal{S}$.
Return $\mathcal{S}$
```

Algorithm 2. Binumber Scalar Decomposition

Algorithm 2 differs somewhat from its counterpart, Algorithm 1 in [3], in that we are always using the same MODULUS_RATIO, until we reach a sufficiently low stage, and then give up and use a τ-NAF. In fact, while $|\zeta| > 2^{\mathbf{p}^{4/5}}$, on applying Theorem 4 we keep dividing moduli by a quantity at least

$$\log_2^{\frac{2}{25}} 2^{\mathbf{p}^{4/5}} = \mathbf{p}^{\frac{8}{125}}.$$

Therefore we need less than

$$\frac{\mathbf{p}}{\log_2\left(\mathbf{p}^{\frac{8}{125}}\right)} = 15.625 \frac{\mathbf{p}}{\log_2 \mathbf{p}}$$

iterations to get down to $2^{\mathbf{p}^{4/5}}$. Below this threshold, a τ-NAF will have length $\leq \mathbf{p}^{4/5}$. Altogether, this gives a new bound for ℓ similar to the bound of Theorem 5 with 15.625 replacing 12.5.

There are indications, although at this point we cannot be more precise, that this algorithm is giving expansions shorter than the τ-NAF also

Input: A point P on the Koblitz curve E_a and a sequence of triplets of exponents $(e_{i,j}, s_{i,j}, t_i)$ as in (3).
Output: The point Q on the elliptic curve such that $Q = \zeta P$.

```
$Q \leftarrow \mathcal{O}$
For $i = 1$ to $\mathcal{I} - 1$
    $R \leftarrow (-1)^{e_{i,1}} P$
    For $j = 1$ to $\mathcal{J}_i$
        $R \leftarrow \tau^{s_{i,j} - s_{i,j+1}} R + (-1)^{e_{i,j+1}} P$
    $Q \leftarrow Q + R$
    $Q \leftarrow 3^{t_i - t_{i+1}} Q$
$R \leftarrow (-1)^{e_{\mathcal{I},1}} P$
For $j = 1$ to $\mathcal{J}_{\mathcal{I}}$
    $R \leftarrow \tau^{s_{\mathcal{I},j} - s_{\mathcal{I},j+1}} R + (-1)^{e_{\mathcal{I},j+1}} P$
$Q \leftarrow Q + R$
Return $Q$
```

Algorithm 3. Sublinear Multiplication

for cryptographically relevant curves, such as NIST curves K-163, K-233 etc. Indeed there is also a greedy algorithm for the τ-NAF expansion where at each step the intermediate variable is halved, whereas we divide it by a power of $\log |\zeta|$ (a small one, true, but for small values of $\mathbf{p}$ we use ad-hoc look-up which should at least find the smallest power of τ close to ζ). The constant in the bound of the expansion in practice seems much smaller than 15 (in [3], this is about 1, using only unsigned expansions, which means that the corresponding algorithm is 60% faster than the τ-and-add).

8 Impossibility Results

In this last part we want to show some limitations of the present method (which is not to say that double bases cannot be used more effectively in another way).

8.1 Why Double-Base Algorithms Are Sublinear Exclusively on Curves with Fast Endomorphisms

We now prove that the maximum of the exponents in any $\{2, 3\}$-integer expansion of n must be of order $\log n$. As a corollary we have that no such expansion can give rise to a sublinear scalar multiplication algorithm on a generic elliptic curve, where we can only hope to improve the scalar multiplication timings by a bounded factor.

Theorem 6. *Let*

$$n = \sum_{i=1}^{k} 2^{s_i} 3^{t_i}, \quad s_i, t_i \in \mathbb{N} \cup \{0\}$$

with $(s_i, t_i) \neq (s_j, t_j)$ *for* $i \neq j$*. Then, as* n *goes to infinity,*

$$\max_i (s_i, t_i) \geq \log_6 n + O(\log \log n) \ .$$

Proof. Let $s = \max_i (s_i, t_i)$. We have

$$n = \sum_{i=1}^{k} 2^{s_i} 3^{t_i} \leq k 6^s \ .$$

Since $k \leq \log^2 n / (\log 2 \log 3)$ we must have from $\frac{\log^2 n}{\log 2 \log 3} 6^s \geq n$ that

$$s \geq \frac{\log n}{\log 6} - \frac{2 \log \log n}{\log 2 \log 3 \log 6} \ .$$

□

Corollary 1. *A double base expansion of a generic scalar* n *cannot be converted into a sublinear scalar multiplication algorithm on a generic elliptic curve.*

Proof. Indeed, there are at least $\Omega(\log n)$ powers of 2 or 3, and on a generic elliptic curve these two operations are costly, hence Algorithm 3 computing nP on the elliptic curve will have to compute $\Omega(\log n)$ elliptic curve operations. □

Remark 2. The same goes of course for the $\{\tau, 3\}$-number algorithm described in this paper.

8.2 Limitations of Greedy-Type Algorithms

In this section, we take all logs to the base 2. We want to show the following theorem.

Theorem 7. *If we use a greedy algorithm to find a $\{\tau, 3\}$ expansion, then we must have*

$$\ell \geq \frac{\mathbf{p}}{\log \mathbf{p}} + o\left(\frac{\mathbf{p}}{\log \mathbf{p}}\right)$$

Remark 3. In particular, this shows that we cannot achieve a constant better than 1 in Theorem 5, at least with our method.

Proof. Since we are using a greedy algorithm to find all our $\{\tau, 3\}$-integers, we are restricting our pool of $\{\tau, 3\}$-integers to $\tau^s 3^t$ with $s \leq 2\mathbf{p}$ and $t \leq \mathbf{p}$, hence at most $2\mathbf{p}^2$ numbers. We know that the number of integers of norm less than $2^{\mathbf{p}}$ which can be represented by at most ℓ $\{\tau, 3\}$-numbers is upper bounded by

$$\sum_{i=1}^{\ell} 2^i \binom{2\mathbf{p}^2}{i} \leq \ell\, 2^\ell \binom{2\mathbf{p}^2}{\ell} = \frac{\Gamma(2\mathbf{p}^2+1)}{\Gamma(\ell+1)\Gamma(2\mathbf{p}^2-\ell)}\, \ell\, 2^\ell \ .$$

This is due to the fact that for any weight i, we can represent an integer by choosing i $\{\tau, 3\}$-integers among at most $2\mathbf{p}^2$ and each of them can have a positive or negative sign. The inequality follows from the ascertained fact that $\ell < \mathbf{p}$. Using Stirling's formula for $\Gamma(z)$, we arrive at the following asymptotic formula

$$\left(1 + \frac{\ell}{2\mathbf{p}^2 - \ell}\right)^{2\mathbf{p}^2 - \ell} \cdot 2^\ell \cdot \mathbf{p}^{2\ell} \cdot \frac{\ell^{3/2}}{\ell^\ell} \leq \frac{(2e)^\ell\ \mathbf{p}^{2\ell}\ \ell^{3/2}}{\ell^\ell} \ .$$

With $\ell = c\dfrac{\mathbf{p}}{\log \mathbf{p}}$ we transform the previous expression into

$$(2e)^{(c - c\log c)\frac{\mathbf{p}}{\log \mathbf{p}}} \cdot 2^{c\mathbf{p}} \cdot 2^{c\frac{\mathbf{p}\log\log \mathbf{p}}{\log \mathbf{p}}} \left(\frac{c\mathbf{p}}{\log \mathbf{p}}\right)^{3/2} < 2^{\mathbf{p}-3}$$

when $\mathbf{p} \to \infty$, as soon as $c < 1$. This contradicts the fact that we must find a representation of all the integers of norm less than $2^{\mathbf{p}}$, which are at least $2^{\mathbf{p}-2}$ (at least all remainders ζ of all possible n mod $(\tau^{\mathbf{p}} - 1)/(\tau - 1)$). □

Remark 4. The same theorem also holds for *any* unsigned $\{2, 3\}$ expansion as in [3, 4], or for those signed $\{2, 3\}$ expansion obtained with a greedy algorithm, since the only ingredient we need to make this cardinality-type argument work is an upper bound on the double exponents (a, b) in $2^a 3^b$, which we automatically have in theses cases.

It is not clear that the same holds in general (actually, it seems more plausible to have $o(\mathbf{p}/\log \mathbf{p})$ in signed expansions).

9 Conclusion

We have analyzed double-base expansions to the extent that we could generalize them to complex-valued bases $\{A, B\}$. We have then seen that since all present algorithms to compute them are greedy, they cannot achieve a better asymptotic bound than $\mathbf{p}/\log\mathbf{p}$.

This leaves many open questions, the first one being the existence of other non-greedy decomposition procedures. For instance, binary signed and unsigned expansions can be retrieved by a greedy (analytic) algorithm but also by a right-to-left "algebraic" algorithm. That the two algorithms yield the same result is assured by the uniqueness of such decompositions (NAF for the signed one). However, high redundancy in the case of double bases renders it more probable to achieve shorter expansions by non-analytic methods. In particular, it would be of utmost interest to find an algebraic "right-to-left" algorithm, since it could give rise to double-base chains as defined in [4], hence only one loop instead of the two in Algorithm 3.

References

1. R. Avanzi, M. Ciet, and F. Sica. Faster Scalar Multiplication on Koblitz Curves combining Point Halving with the Frobenius Endomorphism. In *(Proceedings of PKC 2004*, volume 2947 of *Lecture Notes in Computer Science*, pages 28–40. Springer, 2004.
2. R. Avanzi, C. Heuberger, and H. Prodinger. Minimality of the Hamming Weight of the τ-NAF for Koblitz Curves and Improved Combination with Point Halving. In *Proceedings of SAC 2005*, Lecture Notes in Computer Science. Springer 2006. to appear.
3. M. Ciet and F. Sica. An Analysis of Double Base Number Systems and a Sublinear Scalar Multiplication Algorithm. In E. Dawson and S. Vaudenay, editors, *Progress in Cryptology - Proceedings of Mycrypt 2005*, volume 3715 of *Lecture Notes in Computer Science*, pages 171–182. Springer, 2005.
4. V. S. Dimitrov, L. Imbert, and P. K. Mishra. Efficient and secure elliptic curve point multiplication using double-base chains. In *Advances in Cryptology - ASIACRYPT 2005*, volume 3788 of *Lecture Notes in Computer Science*, pages 59–78. Springer, 2005.
5. V. S. Dimitrov, G. A. Jullien, and W. C. Miller. An algorithm for modular exponentiation. *Information Processing Letters*, 66(3):155–159, 1998.
6. R. P. Gallant, J. L. Lambert, and S. A. Vanstone. Faster Point Multiplication on Elliptic Curves with Efficient Endomorphisms. In J. Kilian, editor, *Advances in Cryptology - Proceedings of CRYPTO 2001*, volume 2139 of *Lecture Notes in Computer Science*, pages 190–200. Springer, 2001.
7. E.W. Knudsen. Elliptic Scalar Multiplication Using Point Halving. In K.-Y.Lam, E. Okamoto, and C. Xing, editors, *Advances in Cryptography - Proceedings of ASIACRYPT 1999*, volume 1716 of *Lecture Notes in Computer Science*, pages 135–149. Springer, 1999.

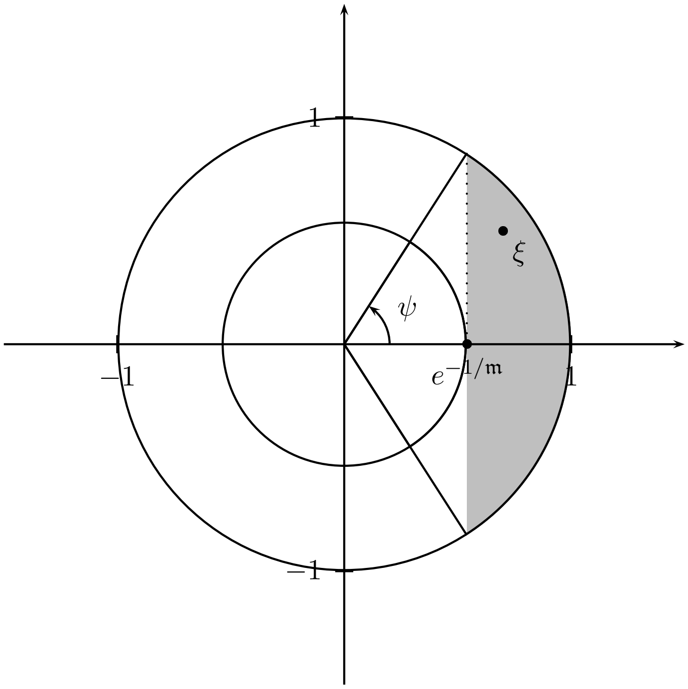
8. N. Koblitz. CM-curves with good cryptographic properties. In Joan Feigenbaum, editor, *Advances in Cryptology - Proceedings of CRYPTO 1991*, volume 576 of *Lecture Notes in Computer Science*, pages 279–287, Berlin, 1991. Springer.
9. K. Okeya, T. Takagi, and C. Vuillaume. Short Memory Scalar Multiplication on Koblitz Curves. In *Proceedings of CHES 2005*, volume 3659 of *Lecture Notes in Computer Science*, pages 91–105. Springer, 2005.
10. G.W. Reitwiesner. Binary arithmetic. *Advances in Computers*, 1:231–308, 1960.
11. R. Schroeppel. Elliptic curves: Twice as fast!, 2000. Presentation at the Crypto 2000 Rump Session.
12. J. Solinas. Efficient arithmetic on Koblitz curves. *Designs, Codes and Cryptography*, 19:195–249, 2000.
13. J. A. Solinas. An Improved Algorithm for Arithmetic on a Family of Elliptic Curves. In Burton S. Kaliski Jr., editor, *Advances in Cryptology - Proceedings of CRYPTO 1997*, volume 1294 of *Lecture Notes in Computer Science*, pages 357–371. Springer, 1997.

A Proof of Lemma 3

After rescaling, we may suppose that $\xi_2 = e^{-1/\mathfrak{m}}$ and $\xi_1 = \xi$ has modulus $e^{-1/\mathfrak{m}} \leq |\xi| \leq 1$, with $\cos\arg(\xi) \geq e^{-1/\mathfrak{m}}$. This means that ξ is in the grayed out sector in the figure. Let $0 < \psi < \pi/2$ be the angle such that $\cos\psi = e^{-1/\mathfrak{m}}$, as shown in the figure. Then the maximum of the distance between ξ and $e^{-1/\mathfrak{m}}$ is the dotted length. Analytically,

$$|\xi_1 - \xi_2| = \left|\xi - e^{-\frac{1}{\mathfrak{m}}}\right| \leq \sin\psi = \sqrt{1 - e^{-2/\mathfrak{m}}} \leq \frac{\sqrt{2}}{\sqrt{\mathfrak{m}}} = \frac{\sqrt{2}\, e^{1/\mathfrak{m}}}{\sqrt{\mathfrak{m}}} |\xi_2| \ .$$

Since $\sqrt{2}\, e^{1/\mathfrak{m}} < 2$ for $\mathfrak{m} \geq 3$ this concludes the proof.


1
ξ
ψ
-1
$e^{-1/\mathfrak{m}}$
1
-1