Available online at www.sciencedirect.com
SCIENCE DIRECT®
Finite Fields and Their Applications 11 (2005) 200–229
FINITE FIELDS AND THEIR APPLICATIONS
http://www.elsevier.com/locate/ffa

Koblitz curve cryptosystems

Tanja Lange

Information-Security and Cryptography, Ruhr-University of Bochum, Universitätsstrasse 150, D-44780 Bochum, Germany

Received 20 February 2004; revised 26 May 2004

Available online 24 August 2004

Hyperelliptic curves over finite fields are used in cryptosystems. To reach better performance, Koblitz curves, i.e. subfield curves, have been proposed. We present fast scalar multiplication methods for Koblitz curve cryptosystems for hyperelliptic curves enhancing the techniques published so far. For hyperelliptic curves, this paper is the first to give a proof on the finiteness of the Frobenius-expansions involved, to deal with periodic expansions, and to give a sound complexity estimate.

As a second topic we consider a different, even faster set-up. The idea is to use a τ-adic expansion as the key instead of starting with an integer which is then expanded. We show that this approach has similar security and is especially suited for restricted devices as the requirements to perform the operations are reduced to a minimum.

Keywords: Cryptography; Discrete logarithm systems; Hyperelliptic curves; Koblitz curves; Frobenius expansions

1. Introduction

Many protocols for public key cryptography rely on the use of cyclic groups. In the Diffie–Hellman key exchange as well as in ElGamal's encryption and signature schemes the main operation is the computation of m times a group element. Thus a group is

E-mail address: lange@itsc.ruhr-uni-bochum.de (T. Lange),
URL: http://www.ruhr-uni-bochum.de/itsc/tanja

1071-5797/$ - see front matter
doi:10.1016/j.ffa.2004.07.001

suitable if this computation is fast, the group order can be determined efficiently, and—most importantly—the discrete logarithm problem, i.e. the problem of obtaining m from the knowledge of D and $E = mD$, is hard. Elliptic and hyperelliptic curves provide suitable groups—there are no currently known subexponential algorithms for solving the DLP on such curves of genus $g \leqslant 3$, except for curves of special classes. Furthermore, fast explicit formulae for addition and doubling exist, making the curves applicable in practice. The finite field the arithmetic is based on becomes smaller with increasing genus which might be advantageous for implementations. Compared to the common choice of the cyclic group as the multiplicative group of a finite field, the size of the finite field can be chosen much smaller on the cost of more complicated formulae to do arithmetic in the group.

If speed is an issue, cryptosystems based on curves can be speed-up considerably if one uses special curves. In this paper we investigate Koblitz curves; these are curves which are defined over a small finite field and are then considered over a large extension field. We show how to efficiently make use of the Frobenius endomorphism of the curve. To this end we detail the full generalization of Koblitz' ideas to hyperelliptic curves showing how to compute scalar multiples using the Frobenius endomorphism and give proofs on the properties of these expansions. We show that computing mD for $m \approx q^{gn}$ needs only $\approx n\frac{q^g-1}{q^g}$ group operations if $\lceil \frac{q^g-1}{2} \rceil$ elements can be precomputed and stored. One can trade-off storage for larger speed-up, e.g. if one is allowed to precompute and store $\lceil \frac{q^g(q^g-1)}{2} \rceil$ elements then one needs only $\approx n\frac{q^g-1}{2q^g}$ operations. As both q and g are assumed to be fairly small the storage requirements are low in any case.

Our main emphasis in this text is on hyperelliptic curves; from the properties we use elliptic curves are included as well. A generalization to arbitrary sub-field curves is obvious as the properties of the expansions depend only on the characteristic polynomial of the Frobenius endomorphism and not on special properties of the curve. To keep the mathematical background brief we do not mention more general curves, but all the results presented in the sequel apply to any Picard [11] or more generally any C_{ab} curve (see [1,2,12]).

Our approach is different from [5,43] as our expansions are shorter and are proven to be finite.

For elliptic curves, Koblitz [24] investigates using a Frobenius expansion as a secret instead of an integer which is then expanded. He credits the idea to H. Lenstra. This approach has the advantage that one saves the time (and more importantly the space for the code) needed for the expansion. In the case of $g = 1, q = 2$ Solinas [47] gives some heuristics that this approach should lead to uniformly distributed multipliers. The idea of using such random tuples instead of random integers m was pointed out to us by Schroeppel. We investigate the applications in protocols and consider attacks that might be possible due to this different choice.

The paper is organized as follows. We first recall the mathematical background needed in the following sections and sketch the development of Koblitz curves in cryptography. Then we present in detail the algorithms to obtain the fast method of computing m-folds, which is analyzed in the next section. The analysis contains a careful study on the length of the resulting expansions. In combination with the density

this allows to state the complexity of computing scalar multiples using the Frobenius endomorphism. For applications in restricted environments this might require too much computational overhead. We analyze the effects of a different set-up and deal with security concerns. Finally, we provide some examples to show the effects in practice and to give evidence that the asymptotical results obtained before already apply to the used setting.

2. Mathematical background

In this section we state results without proofs. For an introduction to hyperelliptic curves see the appendix by Menezes et al. [38], for more details and proofs we refer the interested reader to Lorenzini [33], Stichtenoth [49], and Frey and Lange [13].

2.1. Hyperelliptic curves and ideal class group

Let $q = p^r$ be a prime power and let $\mathbb{F}_q$ denote the finite field with q elements. The curves we consider can be defined via an equation of the type

$$C : y^2 + h(x)y = f(x), \quad f, h \in \mathbb{F}_q[x], \ \deg f = 2g+1, \deg h \leqslant g, \ f \text{ monic}, \tag{1}$$

where we require the curve to be nonsingular, i.e. no pair $(x, y) \in \overline{\mathbb{F}}_q^2$ satisfying the equation fulfills both partial derivative equations, where $\overline{\mathbb{F}}_q$ denotes the algebraic closure of $\mathbb{F}_q$. The curve C is called a *hyperelliptic curve of genus g*. In the case of odd q we may assume that $h = 0$.

The group one uses is the ideal class group of a maximal order of the function field $\mathbb{F}_{q^n}(x, y)/(y^2 + h(x)y - f(x))$, denoted by $\mathrm{Cl}(C/\mathbb{F}_{q^n})$. For applications, it is enough to keep the following routine in mind: take the polynomial ring $\mathbb{F}_{q^n}[x, y]$ and replace any occurrence of y^2 by $-h(x)y + f(x)$, thus every element is of the shape $a(x) + b(x)y$. The ideal class group is the factor group of the fractional ideals by the principal ideals.

For implementations, it is necessary to have a compact representation of the group elements. Each nontrivial ideal class can be represented via an ordered pair of polynomials $[u(x), v(x)]$, $u, v \in \mathbb{F}_{q^n}[x]$, $\deg v < \deg u \leqslant g$, u monic, that satisfy $u | f - v^2 - hv$. To unify notation we represent the class of the principal ideals by $[1, 0]$. Therefore, each class can be represented by at most $2g$ coefficients and if one considers classes in $\mathrm{Cl}(C/\mathbb{F}_{q^m})$ then the coefficients are in $\mathbb{F}_{q^m}$. The inverse of $[u, v]$ equals $[u, -h - v]$ where the second entry is reduced modulo u, hence, computing inverses can be performed efficiently. To need less storage for a class one can recover v from u and some additional information (see [21,48]). In any case the *key length* is $cng\log(q)$ for some small constant c depending on whether all users agree on the same curve or if the curve has to be included in the key as well. For the group size one has

$$|\mathrm{Cl}(C/\mathbb{F}_{q^n})| = q^{ng} + O(q^{n(g-1/2)}) \tag{2}$$

by the theorem of Hasse-Weil. Hence, the trade-off between group size and key length is optimal.

By a *Koblitz curve* we understand a curve defined over a small finite field which is considered over a large extension field. More requirements on the curve for cryptographic applications will be introduced later when the terminology is presented.

2.2. Frobenius endomorphism

In $\mathbb{F}_{q^n}$ the Frobenius automorphism maps x to x^q. This operation is inherited by the curve and by the ideal class group as well. The Frobenius endomorphism σ operates on the ideal classes via their representatives as $\sigma([u(x), v(x)]) = [\sigma(u(x)), \sigma(v(x))]$ for $u, v \in \overline{\mathbb{F}}_q[x]$, where $\sigma(\sum u_i x^i) = \sum u_i^q x^i$. It satisfies a characteristic polynomial in $\mathbb{Z}[T]$ of degree $2g$ of the form

$$P(T) = T^{2g} + a_1 T^{2g-1} + \cdots + a_g T^g + \cdots + a_1 q^{g-1} T + q^g. \tag{3}$$

From P one can easily obtain the group order of the ideal class group for any finite field extension. The complex roots τ_i of $P(T)$ have the following properties: $|\tau_i| = \sqrt{q}$, $\tau_{i+g} = \overline{\tau}_i$ for an appropriate ordering and the group order of the ideal class group over $\mathbb{F}_{q^n}$ is given by

$$|\mathrm{Cl}(C/\mathbb{F}_{q^n})| = \prod_{i=1}^{2g} (1 - \tau_i^n).$$

To compute $P(T)$ it is enough to know the number of points on the curve over $\mathbb{F}_q, \ldots, \mathbb{F}_{q^g}$ satisfying the defining equation of the curve. For $g = 1$ we simply have $a_1 = |C(\mathbb{F}_q)| - q - 1$ and for $g = 2$ it is $a_1 = |C(\mathbb{F}_q)| - q - 1, a_2 = (|C(\mathbb{F}_{q^2})| - q^2 - 1 + a_1^2)/2$. Hence, for curves defined over small finite fields, computing the group order poses no problem.

This is in contrast to the general case that for curves of genus >1 over fields of large characteristic it is still inefficient to determine the group order for randomly chosen curves. For genus two curves over prime fields the current record is held by Gaudry and Schost [18], but they need ≈ 1 week on a single machine to compute the group order for a single curve over $\mathbb{F}_p, \log_2 p = 80$.

2.3. Arithmetic in $\mathrm{Cl}(C/\mathbb{F}_{q^n})$

As usual we write the group additively. To compute scalar multiples of an element, doublings and general additions are needed. Cantor's algorithm [4,23] performs the group operations on the representatives $[u, v]$. Recently, very efficient explicit formulae for the most frequent cases of addition and doubling were published (cf. [29] and the references therein for $g = 2$ and [25,44] for $g = 3$). For elliptic curves such formulae have long been known. Using the standard affine representation, these formulae involve

field inversions in $\mathbb{F}_{q^n}$. For $g = 1$ and odd characteristic, an addition of ideal classes needs 1 inversion, 2 multiplications, and 1 squaring in $\mathbb{F}_{q^n}$ whereas a doubling needs one more squaring. For $g = 2$ we use 1 inversion, 3 squarings, and 22 multiplications for a generic addition and 2 more squarings for a doubling, both independent of the characteristic. Depending on the implementation environment it can be advantageous to trade-off the inversions for more multiplications using different coordinates.

Note that the size of the finite field decreases with increase of g if the group size q^{gn} remains fixed. For genus 3, q^n can be represented within 64 bits for common security requirements. This size of the finite field can be handled advantageously by some computers. To compare the effects for different genera one must take into account the costs of inversions relative to multiplications to find out for which system the arithmetic is fastest.

3. Background on Koblitz curves

3.1. Elliptic curves over $\mathbb{F}_2$

The first attempt to use the Frobenius endomorphism to speed up the computation on an elliptic curve was made by Menezes and Vanstone [37] using the curve $y^2 + y = x^3$ over $\mathbb{F}_{2^n}$. The characteristic polynomial of the Frobenius is $P(T) = T^2 + 2$, thus doubling is replaced by a two-fold application of the Frobenius endomorphism and taking the negative. However, these curves are supersingular and therefore weak [34]. As "the next best thing" Koblitz [24] suggested to use the remaining two nonsupersingular curves defined over $\mathbb{F}_2$, namely $y^2 + xy = x^3 + ax^2 + 1$, $a \in \{0, 1\}$. They are considered as curves over $\mathbb{F}_{2^n}$, where n is chosen large enough to achieve a group size of the desired bit length. The characteristic polynomial of the Frobenius endomorphism is $P(T) = T^2 + (-1)^a T + 2$.

The Frobenius endomorphism of the curve acts on a point $P = (x, y) \in \mathbb{F}_{2^n}^2$ of the curve C by mapping it to $\sigma(P) = (x^2, y^2)$. If the ground field is represented via a normal basis this operation is virtually for free as it is realized by a cyclic shift of the field elements. Also for polynomial basis representations a squaring of all coordinates is performed much faster than the whole addition formula (see [20] for a software implementation).

Let τ be a complex root of $P(T)$. To use the fast-to-compute endomorphism σ in computing mP for an integer m, one expands m to the base of τ using the relation $2 = -(-1)^a\tau - \tau^2$. Unfortunately this direct approach leads to expansions of twice the bit-length of m. Refinements have been obtained by Meier and Staffelbach [39] and Solinas [46]. A very detailed study can be found in Solinas [47]. To reduce the length of these expansions for a fixed extension field $\mathbb{F}_{2^n}$, one reduces m in $\mathbb{Z}[\tau]$ modulo $(\tau^n - 1)/(\tau - 1)$ and expands the resulting element. That is, one looks for an element $M \in \mathbb{Z}[\tau]$ that is equivalent to m modulo $(\tau^n - 1)/(\tau - 1)$ and which has a shorter expansion. Furthermore, Solinas suggests to use a signed digit τ-adic expansion achieving an expression of length n (the degree of extension) and density $\frac{1}{3}$.

3.2. Generalizations

For larger ground fields, such subfield curves have been studied by Müller [40] and Smart [45]. In any case the field of definition is small such that $P(T)$ can be computed easily. The process of expanding is as described above, however, their studies are not as detailed as Solinas'.

Already in his initial paper on hyperelliptic curves, Koblitz [23] suggested applying the Frobenius endomorphism in computations of 2^r-folds. Günther et al. [19] generalized the concept of Koblitz curves to larger genus curves and studied two curves of genus two over $\mathbb{F}_2$. In [26] it has been shown that this approach works for any genus and characteristic and this study has been detailed in [28].

4. Hyperelliptic Koblitz curves

The results of this section hold independently of the genus, characteristic, and size of the ground field. However, we suggest restricting to really small fields $\mathbb{F}_q$, $q \leqslant 7$ and large prime order extensions n. Additionally, we require $P(T)$ to be irreducible over $\mathbb{Z}$.

The size of the ground field needs to be kept small as the number of precomputations grows like q^g. The degree of extension should be prime to get an almost prime group order: due to $|\mathrm{Cl}(C/\mathbb{F}_{q^n})| = \prod_{i=1}^{2g}(1-\tau_i^n) = \prod_{i=1}^{2g}(1-\tau_i)(1+\tau_i+\cdots+\tau_i^{n-1}) = |\mathrm{Cl}(C/\mathbb{F}_q)| \prod_{i=1}^{2g}(1+\tau_i+\cdots+\tau_i^{n-1})$ we cannot avoid having a cofactor of size q^g, any divisor of n will lead to additional factors. Likewise a composite P gives rise to cofactors for any degree of extension. Furthermore, for composite or medium degree extensions, Weil descent attacks [16,17,36] have to be taken seriously. Therefore, we suggest to choose q and n prime for cryptographic applications. For this article we keep the arbitrary ground field $\mathbb{F}_q$ as the results are true in general.

Let $|\mathrm{Cl}(C/\mathbb{F}_{q^n})| = kl$ for a prime l. For cryptographic applications the cofactor k should not be significantly larger than the inevitable factor $|\mathrm{Cl}(C/\mathbb{F}_q)|$ from the ground field. From the Hasse–Weil bound (2) we can hope for $l \approx q^{g(n-1)}$. Furthermore, we assume that l is large such that $l^2 \nmid |\mathrm{Cl}(C/\mathbb{F}_{q^n})|$.

As supersingular curves are always weak under the Frey-Rück attack (cf. [14,15]) we suggest to avoid these curves for usual applications in DL systems. In any case one needs to check that for the minimal κ satisfying $l | q^{n\kappa} - 1$ we have $\kappa > \frac{2000}{n \log_2 q}$.

However, supersingular curves and—more generally—curves with small κ can be useful in pairing-based cryptosystems and the speed-up obtained from the Frobenius endomorphism can be exploited there as well.

Example 1. Over $\mathbb{F}_2$ we can classify up to isogenies the nine classes of hyperelliptic curves of genus 2 given by an equation of form (1) with irreducible $P(T)$, which are given in Table 1.

The first five examples were given in Koblitz [23]. Besides the first three classes these curves are nonsupersingular. The fourth and fifth case were studied by Günter et al. [19].

Table 1
Binary curves of genus 2

Equation of C	$P(T)$
$y^2+y=x^5+x^3$	$T^4+2T^3+2T^2+4T+4$
$y^2+y=x^5+x^3+1$	$T^4-2T^3+2T^2-4T+4$
$y^2+y=x^5+x^3+x$	T^4+2T^2+4
$y^2+xy=x^5+1$	T^4+T^3+2T+4
$y^2+xy=x^5+x^2+1$	T^4-T^3-2T+4
$y^2+(x^2+x)y=x^5+1$	T^4-T^2+4
$y^2+(x^2+x+1)y=x^5+1$	T^4+T^2+4
$y^2+(x^2+x+1)y=x^5+x$	$T^4+2T^3+3T^2+4T+4$
$y^2+(x^2+x+1)y=x^5+x+1$	$T^4-2T^3+3T^2-4T+4$

Group orders and characteristic polynomials $P(T)$ for all Koblitz curves of genus $\leqslant 4$ over $\mathbb{F}_q$ with $q\leqslant 7$ can be found in [27].

4.1. Expansions to the Base of τ

Like before let $P(T)$ denote the characteristic polynomial of the Frobenius endomorphism and let τ be one of its complex roots. To make use of the Frobenius endomorphism we need to be able to represent mD as a linear combination of the $\sigma^i(D)$ with bounded coefficients. This is equivalent to expanding m to the base of τ as $m=\sum_{i=0}^{l} r_i\tau^i$, where the $r_i\in R$ for a set of coefficients R to be defined later. If one precomputes rD for all occurring coefficients $r\in R$ then the computation of mD is realized by applications of the Frobenius endomorphism, table-look-ups and additions of ideal classes whenever the coefficient is nonzero.

The elements of $\mathbb{Z}[\tau]$ are of the form $c=c_0+c_1\tau+\cdots+c_{2g-1}\tau^{2g-1}$ with $c_i\in\mathbb{Z}$. By (3), τ satisfies a polynomial of degree $2g$ with constant term q^g. Thus one can replace the computation of q^gD by $q^gD=-(q^{g-1}a_1\sigma(D)+q^{g-2}a_2\sigma^2(D)+\cdots+a_g\sigma^g(D)+\cdots+a_1\sigma^{2g-1}(D)+\sigma^{2g}(D))$. But this need not be faster than computing q^gD by the usual method of double-and-add. Still it is the clue observation used in expanding an integer. To compute the expansion we need a division by τ with remainder.

Lemma 2. *$c=c_0+c_1\tau+\cdots+c_{2g-1}\tau^{2g-1}\in\mathbb{Z}[\tau]$ is divisible by τ if and only if $q^g|c_0$.*

Proof. $\tau|c\Leftrightarrow$

$$\begin{aligned} c &= \tau\tilde{c}=\tau(\tilde{c}_0+\tilde{c}_1\tau+\cdots+\tilde{c}_{2g-1}\tau^{2g-1})\\ &= \tilde{c}_0\tau+\tilde{c}_1\tau^2+\cdots+\tilde{c}_{2g-2}\tau^{2g-2}-\tilde{c}_{2g-1}(q^g+a_1q^{g-1}\tau+\cdots+a_1\tau^{2g-1})\\ &= -\tilde{c}_{2g-1}q^g+c_1\tau+\cdots+c_{2g-1}\tau^{2g-1} \qquad \Leftrightarrow q^g|c_0. \qquad \square \end{aligned}$$

Accordingly the set of coefficients R must include a complete set of remainders modulo q^g to allow an expansion. Since taking the negative of a class is essentially for free we will use $R = \{0, \pm 1, \pm 2, \ldots, \pm\lceil \frac{q^g-1}{2} \rceil\}$ as minimal set of remainders. Note that we would not need to include $-q^g/2$ in the case of even characteristic. But as we get it for free we will make use of it.

We now derive a τ-adic expansion of $m \in \mathbb{Z}$. Put $r_0 \equiv m \bmod q^g$ for $r_0 \in R$, $d_1 = (m - r_0)/q^g$, $r_1 \equiv -d_1 a_1 q^{g-1} \bmod q^g$ for $r_1 \in R$, and $d_2 = (-d_1 a_1 q^{g-1} - r_1)/q^g$. Then

$$\begin{aligned} m &= r_0 + m - r_0 = r_0 + d_1 q^g \\ &= r_0 - d_1(q^{g-1}a_1\tau + q^{g-2}a_2\tau^2 + \cdots + a_g\tau^g + \cdots + a_1\tau^{2g-1} + \tau^{2g}) \\ &= r_0 + \tau(-d_1 q^{g-1}a_1 - d_1 q^{g-2}a_2\tau - \cdots - d_1 a_g \tau^{g-1} - \cdots - d_1 a_1 \tau^{2g-2} - d_1\tau^{2g-1}) \\ &= r_0 + r_1\tau + \tau(d_2 q^g - d_1 q^{g-2}a_2\tau - \cdots - d_1 a_g \tau^{g-1} - \cdots - d_1 a_1 \tau^{2g-2} - d_1\tau^{2g-1}) \\ &= r_0 + r_1\tau + \tau^2(\ldots). \end{aligned}$$

The expansions derived by repeatedly applying this process with minimal remainders $|r_i| \leqslant \lceil \frac{q^g-1}{2} \rceil$ might become periodic in some cases. We study this question in Section 4.3. In the following algorithm we assume that R has been chosen to contain a complete set of remainders and some further coefficients if necessary. Furthermore, later on in the text we shall impose conditions to achieve a sparse representation and therefore we will use different choices of the set of coefficients R depending on the structure of $P(T)$.

Now we state the algorithm for expanding an element of $\mathbb{Z}[\tau]$ to the base of τ. Note that at the moment we would only need to represent integers, but in the further sections we will reduce the length of the representation. Thereby we stumble over this more general problem:

Algorithm 1.
INPUT: $c = c_0 + c_1\tau + \cdots + c_{2g-1}\tau^{2g-1}$, $P(T)$, *a suitable set* R.
OUTPUT: $r_0, \ldots, r_{\lambda-1}$ *with* $c = \sum_{i=0}^{\lambda-1} r_i \tau^i$, $r_i \in R$.

(1) *Put* $i := 0$;
(2) *While for any* $0 \leqslant j \leqslant 2g-1$ *there exists an* $c_j \neq 0$ *do*
  if $q^g | c_0$ *choose* $r_i := 0$;
  else choose $r_i \in R$ *with* $q^g | c_0 - r_i$;
    /**possibly taking into account further requirements*/
    /**in even characteristic choose* $r_i = c_0$ if $|c_0| = q^g/2$/*
  $d := (c_0 - r_i)/q^g$;
  for $0 \leqslant j \leqslant g-1$ *do*
    $c_j := c_{j+1} - a_{j+1} q^{g-j-1} d$;
  for $0 \leqslant j \leqslant g-2$ *do*
    $c_{g+j} := c_{g+j+1} - a_{g-j-1} d$;

$c_{2g-1} := -d;$
$i := i + 1;$
(3) *output* $(r_0, \ldots, r_{i-1})$.

4.2. On the finiteness of the representation

We now consider the finiteness of the τ-adic representations and establish the dependence of the length on an expression involving m in case of a finite representation. We show that for any curve the expansions are either finite or periodic and provide a way to find out what happens for a given individual curve and how to deal with periods.

For the original instance of elliptic Koblitz curves over $\mathbb{F}_2$, the ring $\mathbb{Z}[\tau]$ was Euclidean, this allowed an easy proof that the resulting expansion was finite. For elliptic curves over fields $\mathbb{F}_{2^r}$ with small r Müller [40] shows that the remainder of the expansion decreases in each step with respect to a certain norm and then shows that there are only finitely many elements of such a small norm and that they all allow a finite expansion. In our more general case the number theoretic norm as the product over all conjugates does not satisfy the Triangle inequality. Therefore, to investigate the finiteness we now consider a $2g$ dimensional lattice associated to the elements of $\mathbb{Z}[\tau]$.

Let $\tau_1, \ldots, \tau_g$ be the g independent roots of P and take the set of elements

$$\Lambda := \left\{ \left(\sum_{j=0}^{2g-1} c_j \tau_1^j, \ldots, \sum_{j=0}^{2g-1} c_j \tau_g^j \right) : c_j \in \mathbb{Z} \right\}.$$

These elements form a lattice in $\mathbb{C}^g$, since the sum of any two and integer multiples of the vectors are in Λ. Since the polynomial P is irreducible the lattice has full dimension $2g$. We now investigate the norm[1] of vectors in this lattice, where the norm is given by the usual Euclidean norm of $\mathbb{C}^g$

$$\mathcal{N} : (x_1, \ldots, x_g) \mapsto \sqrt{|x_1|^2 + \cdots + |x_g|^2},$$

where $|\ .\ |$ is the complex absolute value. We can also consider this lattice as a $2g$ dimensional lattice over $\mathbb{R}$ by the usual representation of $\mathbb{C}$ as $\mathbb{R}^2$.

By abuse of notation we write $\mathcal{N}(c)$ for $c = c_0 + c_1\tau + \cdots + c_{2g-1}\tau^{2g-1}$ and speak of the norm of c since these vectors are parameterized by the integers $c_0, \ldots, c_{2g-1}$.

[1] There are two notions of length—the length of the τ-adic expansion and the norm of the vector, which is often referred to as (Euclidean-)length in the literature. We hope not to confuse the reader and use norm in the second case.

Thus then $\mathcal{N}(c)$ reads

$$\mathcal{N}(c) = \sqrt{\sum_{i=1}^{g} \left| \sum_{j=0}^{2g-1} c_j \tau_i^j \right|^2}.$$

Now we study the behavior of the norm of the remainders during the expansion of c. Showing that the norm decreases down to a certain limit will be the important step to prove the following theorem:

Theorem 3. *Let C be a hyperelliptic curve of genus g and let τ be a root of the characteristic polynomial of the Frobenius endomorphism. Then the expansion of $c = c_0 + c_1\tau + \cdots + c_{2g-1}\tau^{2g-1} \in \mathbb{Z}[\tau]$ to the base of τ with coefficients in $R = \{0, \pm 1, \ldots, \pm\lceil \frac{q^g-1}{2} \rceil\}$ is either finite or becomes periodic.*

Proof. We first show that for elements of bounded norm the expansion cannot lead to a remainder with larger norm than that bound. Showing that the expansion of any element leads to a remainder of norm bounded by that constant concludes the proof.

Let $\mathcal{N}(c) < \sqrt{g}\frac{q^g}{2(\sqrt{q}-1)}$ (respectively $< \sqrt{g}\frac{q^g+1}{2(\sqrt{q}-1)}$ for even characteristic). Then using the Triangle inequality on $c = r + c - r =: r + c'\tau$, $c \equiv r \bmod q^g$, we get $\mathcal{N}(c'\tau) \leqslant \mathcal{N}(c) + \mathcal{N}(r) \leqslant \mathcal{N}(c) + \sqrt{g}(q^g - 1)/2$ (respectively $\mathcal{N}(c) + \sqrt{g}q^g/2$) and $\mathcal{N}(\tau c') = \sqrt{q}\mathcal{N}(c')$. Now direct calculation shows that $\mathcal{N}(c')$ is bounded by the same constant.

Since we consider a discrete lattice, the number of elements with bounded norm is finite. Thus the expansion of these elements of bounded norm either ends after hitting at most one time all these elements or runs into a cycle since the choice of the r—and therefore the next element c'—is unique for given c. Hence, for these elements the expansion is either periodic or finite.

The following two lemmata show that expanding an element c to the base of τ leads to an element c' with $\mathcal{N}(c') < \sqrt{g}\frac{q^g}{2(\sqrt{q}-1)}$ (or $< \sqrt{g}\frac{q^g+1}{2(\sqrt{q}-1)}$ in even characteristic) after at most $2\log_q \frac{2(\sqrt{q}-1)\mathcal{N}(m)}{\sqrt{g}} + 1$ steps concluding the proof. □

Later we shall refer to an algorithm to find these elements of small norm and show how to recognize periods and how to avoid them. Hence the problem is solved in practice.

Lemma 4. *Let q be odd. For every $m \in \mathbb{Z}[\tau]$ we have a unique expansion $m = \sum_{i=0}^{\lambda-1} r_i\tau^i + m'\tau^\lambda$, where $r_i \in \{0, \pm 1, \pm 2, \ldots, \pm\frac{q^g-1}{2}\}$,*

$$\mathcal{N}(m') < \sqrt{g}\frac{q^g}{2(\sqrt{q}-1)} \text{ and } \lambda \leqslant \left\lceil 2\log_q \frac{2(\sqrt{q}-1)\mathcal{N}(m)}{\sqrt{g}} \right\rceil.$$

Proof. Put $m_0 := m$. The expansion of m to the base of τ leads to

$$\begin{aligned} m_0 &= r_0 + m_1\tau = r_0 + r_1\tau + m_2\tau^2 \\ &= \sum_{i=0}^{j-1} r_i\tau^i + m_j\tau^j, \end{aligned}$$

where by Lemma 2 the $r_i \in \{0, \pm 1, \pm 2, \ldots, \pm\frac{q^g-1}{2}\}$ are uniquely determined.

The Triangle inequality for $\mathcal{N}$ leads to $\sqrt{q}\mathcal{N}(m_j) \leqslant \mathcal{N}(m_{j-1}) + \mathcal{N}(r_{j-1}) \leqslant \mathcal{N}(m_{j-1}) + \sqrt{g}\frac{q^g-1}{2}$. Hence,

$$\begin{aligned} \mathcal{N}(m_j) &\leqslant \frac{\mathcal{N}(m_0) + \sqrt{g}(q^g-1)/2\sum_{i=0}^{j-1} q^{i/2}}{q^{j/2}} \\ &< \frac{\mathcal{N}(m_0)}{q^{j/2}} + \sqrt{g}\frac{q^g-1}{2(\sqrt{q}-1)}. \end{aligned}$$

If we choose $j \geqslant 2\log_q \frac{2(\sqrt{q}-1)\mathcal{N}(m_0)}{\sqrt{g}}$, then $\frac{\mathcal{N}(m_0)}{q^{j/2}} \leqslant \frac{\sqrt{g}}{2(\sqrt{q}-1)}$ and the claim follows. □

For even characteristic we proceed similarly.

Lemma 5. *Let q be even. For every $m \in \mathbb{Z}[\tau]$ we have an expansion $m = \sum_{i=0}^{\lambda-1} r_i\tau^i + m'\tau^\lambda$, where $r_i \in \{0, \pm 1, \pm 2, \ldots, \pm\frac{q^g}{2}\}$,*

$$\mathcal{N}(m') < \sqrt{g}\frac{q^g+1}{2(\sqrt{q}-1)} \quad \text{and} \quad \lambda \leqslant \left\lceil 2\log_q \frac{2(\sqrt{q}-1)\mathcal{N}(m)}{\sqrt{g}} \right\rceil.$$

Proof. Put $m_0 := m$. The expansion of m to the base of τ leads to

$$\begin{aligned} m_0 &= r_0 + m_1\tau = r_0 + r_1\tau + m_2\tau^2 \\ &= \sum_{i=0}^{j-1} r_i\tau^i + m_j\tau^j, \end{aligned}$$

where the $r_i \in \{0, \pm 1, \pm 2, \ldots, \pm\frac{q^g}{2}\}$ are given like in Algorithm 1.

The Triangle inequality for $\mathcal{N}$ leads to $\sqrt{q}\mathcal{N}(m_j) \leqslant \mathcal{N}(m_{j-1}) + \mathcal{N}(r_{j-1}) \leqslant \mathcal{N}(m_{j-1}) + \sqrt{g}\frac{q^g}{2}$. Hence,

$$\begin{aligned}\mathcal{N}(m_j) &\leqslant \frac{\mathcal{N}(m_0) + \sqrt{g}q^g/2\sum_{i=0}^{j-1} q^{i/2}}{q^{j/2}} \\ &< \frac{\mathcal{N}(m_0)}{q^{j/2}} + \sqrt{g}\frac{q^g}{2(\sqrt{q}-1)}.\end{aligned}$$

If we choose $j \geqslant 2\log_q \frac{2(\sqrt{q}-1)\mathcal{N}(m_0)}{\sqrt{g}}$ then $\frac{\mathcal{N}(m_0)}{q^{j/2}} \leqslant \frac{\sqrt{g}}{2(\sqrt{q}-1)}$ and the claim follows. □

From the lemmata we see that the length of the expansion depends mainly on $\mathcal{N}(m)$. They leave open to study the length of expansions of elements of norm less than $K_{q,g} := \sqrt{g}\frac{\lfloor q^g/2 \rfloor + 1/2}{\sqrt{q}-1}$.

Remark 6. For elliptic curves, Müller [40] and Smart [45] followed this approach to give bounds on the length of the Frobenius expansions. There, the norm reads $\mathcal{N}(c)^2 = c_0^2 - a_1c_0c_1 + qc_1^2$ and the bounds on the remainders $\mathcal{N}(m') = \mathcal{N}(m'_0 + m'_1\tau)$ from Lemmas 4 and 5 can be translated to explicit bounds on the m'_i. After at most 3 further expansion steps the remainder is zero unless in the case of the pairs (q, a_1) equal to $(5, \pm 4)$ and $(7, \pm 5)$, where the expansion becomes periodic on input $\pm(q+1)/2$. The easy way out in these cases is to include $\pm(q+1)/2$ in the set of coefficients, using an additional precomputation and a little further space.

For arbitrary genus, one can state the norm $\mathcal{N}$ explicitly in the coefficients of the polynomial $P(T)$ and express it in terms of the coefficients $c_0, \ldots, c_{2g-1}$. This can be done using the symmetric functions in the τ_i and with the help of the formulae derived for computing the number of points on Koblitz curves in [28].

Example 7. For $g = 2$ we have for $c = c_0 + c_1\tau + c_2\tau^2 + c_3\tau^3$

$$\begin{aligned}\mathcal{N}(c)^2 = {} & 2c_0^2 - a_1c_0c_1 + (a_1^2 - 2a_2)c_0c_2 - (a_1^3 - 3(a_1a_2 - a_1q))c_0c_3 \\ & + 2qc_1^2 - a_1qc_1c_2 + (a_1^2 - 2a_2)qc_1c_3 \\ & + 2q^2c_2^2 - a_1q^2c_2c_3 \\ & + 2q^3c_3^2.\end{aligned}$$

For $g = 3$ we have for $c = c_0 + c_1\tau + c_2\tau^2 + c_3\tau^3 + c_4\tau^4 + c_5\tau^5$

$$\begin{aligned}\mathcal{N}(c)^2 = {} & 3c_0^2 - a_1c_0c_1 + (a_1^2 - 2a_2)c_0c_2 - (a_1^3 - 3(a_1a_2 - a_3))c_0c_3 \\ & + (a_1^4 - 4(a_1^2a_2 - a_1a_3 + a_2q) + 2a_2^2)c_0c_4\end{aligned}$$

$$
\begin{aligned}
&-(a_1^5 - 5(a_1^3a_2 - a_1^2a_3 - a_1a_2^2 + a_1a_2q + a_2a_3 - a_1q))c_0c_5\\
&\quad +3qc_1^2 - a_1qc_1c_2 + (a_1^2 - 2a_2)qc_1c_3 - (a_1^3 - 3(a_1a_2 - a_3))qc_1c_4\\
&\quad +(a_1^4 - 4(a_1^2a_2 - a_1a_3 + a_2q) + 2a_2^2)qc_1c_5 + 3q^2c_2^2 - a_1q^2c_2c_3\\
&\quad +(a_1^2 - 2a_2)q^2c_2c_4 - (a_1^3 - 3(a_1a_2 - a_3))q^2c_2c_5 + 3q^3c_3^2 - a_1q^3c_3c_4\\
&\quad +(a_1^2 - 2a_2)q^3c_3c_5 + 3q^4c_4^2 - a_1q^4c_4c_5 + 3q^5c_5^2.
\end{aligned}
$$

In general $\mathcal{N}(c)^2$ is a quadratic form in the $2g$ variables $c_0, \ldots, c_{2g-1}$. The coefficient of c_i^2 is gq^i and of $c_ic_j, i < j$ is $q^i(q^v + 1 - |C(\mathbb{F}_{q^v})|)$, where $v = j - i$. Due to its origin in the interpretation as Euclidean norm in a lattice, $\mathcal{N}^2$ is a positive definite quadratic form.

Experiments show that an element of norm bounded by $K_{q,g}$ has an expansion of length at most $2g + 1$ or becomes periodic. We did not succeed in proving this for arbitrary curves (see [28] for detailed study for $g = 2$). However, for each specific curve one can easily determine an upper bound on the length of the expansion:

Finke and Pohst [10] provide an algorithm for finding all vectors of norm bounded by a constant K in a lattice in $\mathbb{R}^s$, respectively for finding all arrays $(x_0, \ldots, x_{s-1})$ for which the value of the corresponding quadratic form in s variables is less than K. This allows to determine the complete set of elements of small norm. They prove the following upper bound on the number of elements of norm bounded by K:

$$
(2\lfloor K^{1/2}\rfloor + 1)\binom{\lfloor 4K\rfloor + s - 1}{\lfloor 4K\rfloor}.
$$

Thus for our constant $K_{q,g}$ we have at most $O\left((\sqrt{g}\frac{q^g}{\sqrt{q}-1})^{(4g-1)/2}\right)$ vectors of small norm. This bounds the maximal length of the expansion in the nonperiodic case, and also the length of the period.

We use the algorithm to find all elements of small norm for individual curves. For each of them we compute the expansion. These experiments show that for each such element $c = c_0 + \cdots + c_{2g-1}\tau^{2g-1}$ of small norm we have $c_i \in R$ for $1 \leqslant g \leqslant 2g - 1$ and $|c_0| \leqslant q^g$, and if $c_0 \notin R$ the other coefficients are fairly small. If no periods occur then every such element has an expansion of length at most $2g + 1$, thus either all $c_i \in R$ or the next remainder in the expansion has all coefficients in this set. Hence, the above bound is appropriate for the number of elements of small norm, however, the expansions are by far shorter than hitting each element once.

Together with Theorem 3 and Lemmas 4 or 5 respectively, this allows to sum up the result in the nonperiodic case.

Summary 1. *Let $P(T)$ be such that no periodic expansions occur. Then the length of the expansion of m with coefficients in R is bounded by*

$$\lceil 2\log_q \frac{2(\sqrt{q}-1)\mathcal{N}(m)}{\sqrt{g}} \rceil + k_{q,g},$$

where $k_{q,g}$ is the maximal length of an element of norm $< K_{q,g}$.

Our observations and experiments stress that $k_{q,g} = 2g+1$ is a good upper bound.

4.3. Periodic expansions

One argument that can be used in the proof of the finiteness in the elliptic curve case is that periods of length larger than one (except for a change of sign) cannot occur since otherwise the coefficients c_0 and c_1 would be larger than allowed. Now we investigate in which situations periods can occur at all. For the elliptic curve case the expansion can become cyclic only if $|a_1|-1 > (q-1)/2$ thus for $q < 14$. For odd characteristic these are just the cases where we included a further coefficient. For even characteristic it was shown in [40] by Müller that we always obtain a finite expansion if we use the set R as given above.

For curves of larger genus the situation is a bit different. First of all—although obvious from the experiments and motivated by the detailed example in [28] for the genus 2 case—we have no proof how large the coefficients of c with $\mathcal{N}(c) < K_{q,g}$ can get, but we can obtain some information as well, which makes it easy to check for periods for an individual curve.

In the following we assume that R consists of a complete set of remainders modulo q^g. For larger sets R' similar observations hold.

The first observation is that each element in the period must have norm less than $K_{q,g}$ as otherwise we know that the norm decreases.

4.3.1. Period length one

Assume that for a curve with $P(T) = T^{2g}+a_1T^{2g-1}+\cdots+a_gT^g+\cdots+a_1q^{g-1}T+q^g$ we have that

$$\begin{aligned} c &= c_0 + c_1\tau + \cdots + c_{2g-1}\tau^{2g-1} \\ &= r_0 \pm \tau(c_0 + c_1\tau + \cdots + c_{2g-1}\tau^{2g-1}) \end{aligned}$$

with $r_0 \in R$ and $\mathcal{N}(c)^2 < K_{q,g}^2$. Without loss of generality we assume that $c_0 > 0$ and therefore $c_0 > \lceil (q^g-1)/2 \rceil = r_{\max}$. Put $d = (c_0 - r_0)/q^g \in \mathbb{Z}_{>0}$.

The rules for expanding an element lead to a system of equations

$$\begin{aligned} \pm c_i &= c_{i+1} - da_{i+1}q^{g-i-1} && 0 \leqslant i \leqslant g-1, \\ \pm c_i &= c_{i+1} - da_{2g-1-i} && g \leqslant i \leqslant 2g-2, \\ \pm c_{2g-1} &= -d, \end{aligned}$$

where the signs are assumed simultaneously. If this system can be fulfilled for a curve with the positive sign for $(c_0, c_1, \ldots, c_{2g-1})$ then the equations hold for the quadratic twist of the curve with the opposite sign and the above coefficient vector with alternating signs. Thus we restrict ourselves to the case of positive sign. Inserting all equations in the one for c_0 yields

$$c_0 = -d - da_1 - \cdots - da_g - dag - 1q - \cdots - da_1 q^{g-1},$$

thus $c_0 = dq^g - d|\mathrm{Cl}(C/\mathbb{F}_q)|$. Inserting the definition of d we obtain

$$r_0 = -d|\mathrm{Cl}(C/\mathbb{F}_q)|.$$

Since both d and $|\mathrm{Cl}(C/\mathbb{F}_q)|$ are nonnegative and $r_0 \in R$ the crucial part to be fulfilled for either the curve or its twist is $\lceil (q^g - 1)/2 \rceil \geqslant d|\mathrm{Cl}(C/\mathbb{F}_q)|$. Since a lower bound on the class number is given by the Theorem of Hasse-Weil (2), q and d have to be such that $\lceil (q^g - 1)/2 \rceil \geqslant d(\sqrt{q} - 1)^{2g}$, i.e. q must be small enough.

We have just shown

Theorem 8. *Let C be a hyperelliptic curve over $\mathbb{F}_q$ of genus g. An expansion using the set of remainders R with maximal coefficient $r_{\max}$ can be periodic of period length* 1 *up to change of sign only if*

$$r_{\max} \geqslant |\mathrm{Cl}(\tilde{C}/\mathbb{F}_q)|,$$

where $\tilde{C}$ is either the curve or its quadratic twist. If in this case the period starts a some $c \in \mathbb{Z}[\tau]$, then $\mathcal{N}(c) < K_{q,g}$. One needs to include $\pm d(q^g - |\mathrm{Cl}(\tilde{C}/\mathbb{F}_q)|)$ into R for all $1 \leqslant d \leqslant r_{\max}/|\mathrm{Cl}(\tilde{C}/\mathbb{F}_q)|$ to guarantee finite expansions.

For a given curve it is again fairly easy to check whether the expansion can run into a cycle at all by applying the bound of Theorem 8. Furthermore, it shows which additional coefficients might have to be included in the set R. Using the algorithm of Finke and Pohst we can compute all elements of such a small norm and expand all these elements to the base of τ. This shows that not all the curves for which the inequality of the theorem holds lead to cyclic expansions. In case this happens, we just need to use the additionally included coefficients instead of the whole period that would follow. Thus if we choose such a curve for implementation we need to precompute and store d more elements. Since d and q are bounded by relatively small constants, the time for this further precomputation can be neglected.

Example 9. Put $g = 2, q = 3$. Among all the isogeny classes of curves with irreducible $P(T)$ only $P(T) = T^4 \pm 2T^3 + 2T^2 \pm 6T + 9$, $P(T) = T^4 \pm T^3 - 2T^2 \pm 3T + 9$, and $P(T) = T^4 \pm 3T^3 + 5T^2 \pm 9T + 9$ lead to periods. The coefficients to include are ± 5 in the first two cases and ± 6 in the last one.

Example 10. In the case of even characteristic the situation is even a bit more relaxed. If we choose coefficients from $\{0, \pm 1, \ldots, \pm q^g/2 - 1, q^g/2\}$ unless $c_0 = -q^g/2$ (cf. Algorithm 1) then for all classes of curves of genus two over $\mathbb{F}_2$ (see Tabular 1) the expansions are finite.

Even though we do not propose the ground field $\mathbb{F}_4$ we checked all possible curves. We run into a cycle only for $P(T) = T^4 \pm 4T^3 + 9T^2 \pm 16T + 16$. To deal with this we include ± 10 in the set of coefficients.

4.3.2. Longer periods

Here one can follow the same approach allowing more quotients d_i. A necessary condition to have a period of length 2 is that

$$-(d_0 + d_1)|\mathrm{Cl}(C/\mathbb{F}_q)| = r_0 + r_1$$

can be fulfilled for $r_0, r_1 \in R$.

For $d_0 = -d_1$ we get $r_0 = -r_1$, i.e. the case of period length one with a change of sign.

In the other cases we see as well, that d_1 and d_0 are of the same order and that both and q have to be reasonably small. On the other hand except for $d_0 = -d_1 = 1$ this did not occur in the experiments and the same holds for periods of higher order.

Again this can be explained by the bounds on the coefficients. If we have $|c_0| < \sqrt{q} r_{\max}$ and $|c_i| \in R$, $i \geqslant 1$, then $d_0 < (\sqrt{q} + 1)/2$ and $|d_1| < 1 + g + 1/\sqrt{q}$ in the worst case.

We did not get any longer periods.

4.3.3. Different strategies

A different way to proof the finiteness of such expansions can be extended from Lesage [32]. He investigates expansions to the base α, where α is a root of a quadratic polynomial over $\mathbb{Z}$ and the set of remainders is of cardinality $|\alpha|^2$, symmetric to 0. He uses difference equations to prove the finiteness and succeeds in general for the case of nonreal roots (except some cases where one obtains periods). For a special polynomial he computes the expected length of the expansion as well. The approach generalizes to the kind of polynomials considered here due to the symmetry of $P(T)$ but again the expressions for the general case involving the a_i cannot be handled. Like before it is possible to get bounds for an individual curve with explicit coefficients.

4.4. Reducing the length of the expansion

The strategy explained so far would lead to expansions of length $2 \log_q m \approx 2ng$—thus expansions that are $2g$ times as long as a q^g-adic expansion which mitigates the advantage of using the Frobenius. Thus, actually one does not expand m itself but looks for an element $M \in \mathbb{Z}[\tau]$ having a short expansion and $MD = mD$ for all $D \in \mathrm{Cl}(C/\mathbb{F}_{q^n})$. Once we decide to use such a curve we need to fix the field $\mathbb{F}_{q^n}$, i.e. the degree of extension. This gives us the additional equation $\tau^n = 1$. (Note that

for $a \in \mathbb{F}_{q^n}$ we have $a^{q^n} = a$ and the application of the Frobenius endomorphism on the ideal classes does nothing but raising field elements to the power of q.) Therefore, if $M \equiv m \bmod (\tau^n - 1)$ then $MD = mD$ for $D \in \mathrm{Cl}(C/\mathbb{F}_{q^n})$ and we can choose an equivalent M with a short expansion.

Remark 11. Note that for a fixed extension field, τ satisfies two equations. Since we consider only irreducible polynomials P and since the constant term of P is $q^g \neq \pm 1$, the polynomials $P(T)$ and $T^n - 1$ are co-prime. Thus their gcd over $\mathbb{Q}[T]$ is one. But we are working in $\mathbb{Z}[T]$. The ideal generated by these polynomials is a principal ideal generated by an integer (since the gcd over $\mathbb{Q}[T]$ is 1). In fact this number is equal to the cardinality of the ideal class group over $\mathbb{F}_{q^n}$.

Namely, write $P(T) = \prod_{i=1}^{2g}(T - \tau_i)$. Then in the ideal under consideration we have $T^n = 1$. Transforming $T \to T^n$ we have to evaluate

$$\prod_{i=1}^{2g}(T^n - \tau_i^n)_{|T^n=1} = \prod_{i=1}^{2g}(1 - \tau_i^n) = |\mathrm{Cl}(C/\mathbb{F}_{q^n})|,$$

which is indeed the class number.

To rephrase this, modulo $|\mathrm{Cl}(C/\mathbb{F}_{q^n})|$ these polynomials have a common linear factor. Hence, if we consider only the cyclic group of order l, the polynomials have a common factor $T - s$ in $\mathbb{F}_l[T]$, where l is a prime factor of $|\mathrm{Cl}(C/\mathbb{F}_{q^n})|$. This means that the operation of the Frobenius endomorphism on an ideal class corresponds to the multiplication of the ideal class by the integer s modulo l.

In the applications one usually restricts the computations to a subgroup of large prime order. Let $l \| |\mathrm{Cl}(C/\mathbb{F}_{q^n})|$ be a large prime such that $l^2 \nmid |\mathrm{Cl}(C/\mathbb{F}_{q^n})|$. Then we can even look for elements equivalent to m modulo $(\tau^n - 1)/(\tau - 1) = \tau^{n-1} + \tau^{n-2} + \cdots + \tau + 1$ as the Frobenius endomorphism cannot correspond to the identity.

In this section we prove the following theorem:

Theorem 12. *Let τ be a root of the characteristic polynomial $P(T)$ of the Frobenius endomorphism of the hyperelliptic curve C of genus g defined over $\mathbb{F}_q$. Consider the curve over $\mathbb{F}_{q^n}$ and let $m \in \mathbb{Z}$. There is an element $M \in \mathbb{Z}[\tau]$ such that*
(1) $m \equiv M \bmod (\tau^n - 1)/(\tau - 1)$
and
(2) $2\log_q \frac{2(\sqrt{q}-1)\mathcal{N}(M)}{\sqrt{g}} < n + 2g,$
where $\mathcal{N}$ denotes the norm defined in the previous section.

The proof is constructive, thus it provides a way to compute such an element M. Let us fix some notation which shall be useful for the proof and to state the algorithms. For an element $Q \in \mathbb{Q}$ let $z = \mathtt{nearest}(Q)$ be the nearest integer to Q, if ambiguity arises it is defined to be the integer with the least absolute value. This can be realized computationally by choosing $z = \lceil Q - 0.5 \rceil$ if $Q > 0$ and $z = \lfloor Q + 0.5 \rfloor$ else. We

will also use nearest(.) for elements of $\mathbb{Q}[\tau]$ where it is understood coefficient-wise.

Proof. In the field $\mathbb{Q}[\tau]$ one can invert elements. Thus, put $Q := m(\tau-1)/(\tau^n-1) \in \mathbb{Q}[\tau]$, so $Q = \sum_{i=0}^{2g-1} Q_i \tau^i$ where $Q_i \in \mathbb{Q}$. For $0 \leqslant i \leqslant 2g-1$ put $z_i = \texttt{nearest}\ (Q_i)$ and put

$$z := \sum_{i=0}^{2g-1} z_i \tau^i \quad \text{and} \quad M := m - z(\tau^n-1)/(\tau-1).$$

Thus it is easy to see that $m \equiv M \bmod (\tau^n-1)/(\tau-1)$. To compute the value

$$\mathcal{N}(M) = \mathcal{N}\left(m - \frac{z(\tau^n-1)}{(\tau-1)}\right) = \mathcal{N}\left(\frac{\tau^n-1}{\tau-1}\left(\frac{m(\tau-1)}{\tau^n-1} - z\right)\right)$$

we need an estimate on $\mathcal{N}(\frac{m(\tau-1)}{\tau^n-1} - z) = \mathcal{N}(Q-z)$.

$$\begin{aligned}\mathcal{N}(Q-z) &= \left(\sum_{j=1}^{g}\left|\sum_{i=0}^{2g-1}(Q_i - z_i)\tau_j^i\right|^2\right)^{\frac{1}{2}} \leqslant \left(\sum_{j=1}^{g}\left(\sum_{i=0}^{2g-1}|(Q_i-z_i)\tau_j^i|\right)^2\right)^{\frac{1}{2}} \\ &\leqslant \left(\sum_{j=1}^{g}\left(\frac{1}{2}\sum_{i=0}^{2g-1}\sqrt{q}^i\right)^2\right)^{\frac{1}{2}} = \left(\sum_{j=1}^{g}\left(\frac{\sqrt{q}^{2g}-1}{2(\sqrt{q}-1)}\right)^2\right)^{\frac{1}{2}} = \sqrt{g}\,\frac{q^g-1}{2(\sqrt{q}-1)}.\end{aligned}$$

Therefore we have

$$\begin{aligned}\mathcal{N}(M) &= \mathcal{N}\left(\frac{\tau^n-1}{\tau-1}\left(\frac{m(\tau-1)}{\tau^n-1} - z\right)\right) \leqslant \sum_{i=0}^{n-1}\mathcal{N}\left(\left(\frac{m(\tau-1)}{\tau^n-1} - z\right)\tau^i\right) \\ &= \sum_{i=0}^{n-1}\left(\sqrt{g}\,\frac{q^g-1}{2(\sqrt{q}-1)}q^{i/2}\right) = \sqrt{g}\,\frac{q^g-1}{2(\sqrt{q}-1)}\,\frac{\sqrt{q}^n-1}{\sqrt{q}-1}.\end{aligned}$$

It follows that

$$2\log_q \frac{2(\sqrt{q}-1)\mathcal{N}(M)}{\sqrt{g}} \leqslant 2\log_q\left(\frac{\sqrt{q}^n-1}{\sqrt{q}-1}\right) + 2\log_q(q^g-1) < n+2g. \qquad \square$$

Remark 13. The usage of `nearest` might not be the best choice, nevertheless it provides an efficient way to compute a length-reduced representation which works for every genus g, ground field $\mathbb{F}_q$, and degree of extension n. For the two binary elliptic curves, Solinas investigates in more detail an optimal way of reduction. Considering the lattice spanned by $\{1, \tau\}$ he shows that for each element of $\mathbb{Q}[\tau]$ there is a unique lattice point within distance less than $4/7$. For larger genera the computation of the nearest point is computationally hard to realize and we do not lose much choosing the "rounded" elements the way presented here.

Thus from the discussion of Section 4.2 we have the following result:

Theorem 14 (*Main result on the length*). *Let C be a hyperelliptic curve over $\mathbb{F}_q$ of genus g and with characteristic polynomial of the Frobenius endomorphism $P(T)$. Let P be such that the τ-adic expansion is not periodic and that for an element c of $\mathbb{Z}[\tau]$ of norm $< K_{q,g}$ the τ-adic expansion is no longer than $2g+1$. Then we have*:

For every element $m \in \mathbb{Z}$ we can compute a τ-adic expansion of length λ using coefficients in the set R only, where

$$\lambda \leqslant n + 4g + 1.$$

The remainder of this section is devoted to computational aspects. One first needs to compute $(\tau^n - 1)/(\tau - 1)$ and its inverse in $\mathbb{Q}[\tau]$, which is done *only once* for C and n. The computations are performed using recurrence sequences. To derive the inverse in $\mathbb{Q}[\tau]$ one uses the extended Euclidean algorithm. If C and n are system parameters these elements can be computed externally and stored on the device as they are independent of the chosen ideal classes. In Section 5 we state a way to circumvent this, see also Remark 15.

First of all one needs the representation of $(\tau^n - 1)/(\tau - 1)$ in $\mathbb{Z}[\tau]$.

Algorithm 2.
`INPUT:` $n \in \mathbb{N}$, $P(T)$.
`OUTPUT:` $e_0, \ldots, e_{2g-1} \in \mathbb{Z}$ *such that* $(\tau^n - 1)/(\tau - 1) = e_0 + e_1\tau + \cdots + e_{2g-1}\tau^{2g-1}$ *in* $\mathbb{Z}[\tau]$.

(1) *initialize*: $d_0 := 1$ *and* $d_i := 0$ *for* $1 \leqslant i \leqslant 2g-1$;
$e_0 := 1$ *and* $e_i := 0$ *for* $1 \leqslant i \leqslant 2g-1$;
(2) *for* $1 \leqslant j \leqslant n-1$ *do*
(a) $d_{\text{old}} := d_{2g-1}$;
(b) *for* $2g-1 \geqslant i \geqslant g$ *do*
$d_i := d_{i-1} - a_{2g-i} d_{\text{old}}$;
$e_i := e_i + d_i$;
(c) *for* $g-1 \geqslant i \geqslant 1$ *do*
$d_i := d_{i-1} - a_i q^{g-i} d_{\text{old}}$;
$e_i := e_i + d_i$;

(d) $d_0 := -q^g d_{\text{old}}$;
$e_0 := e_0 + d_0$;
(3) *output* $(e_0, e_1, ..., e_{2g-1})$.

Considering $e(T) = \sum e_i T^i$ we can invert e modulo $P(T)$ in $\mathbb{Q}[T]$ by the extended Euclidean Algorithm, as for $\gcd(e, P) = ee' + PP'$ one has $e' \equiv e^{-1} \bmod P$. For fixed genus and hence degree of the involved polynomials, this can be made explicit.

We now present the algorithm for computing scalar multiples as a whole.

Algorithm 3 (*Computation of m-folds using τ-adic expansions*).
INPUT: $m \in \mathbb{Z}$, $D = [u, v]$, $u, v \in \mathbb{F}_{q^n}[x]$, $P(T)$, R (*appropriate set of coefficients*), $e = (\tau^n - 1)/(\tau - 1)$, $e' \equiv e^{-1} \bmod P$.
OUTPUT: *mD represented by the reduced ideal* $H = [s, t]$, $s, t \in \mathbb{F}_{q^n}[x]$.
(1) /**precomputation*/
(a) *for* $i \in R, i > 0$ *compute*
$D(i) := iD$; /* *use double-and-add and previous computations*/
$D(-i) := -D(i)$;/* *for free, can also be computed from* $D(i)$ *when used*/
(2) /**compute m modulo* $(\tau^n - 1)/(\tau - 1)$ *using e and* e'*/
(a) *compute* $z(T) :=$`nearest`$(m \cdot e'(T))$;
(b) *let* $M = \sum_{i=0}^{2g-1} M_i T^i :\equiv m - e(T) \cdot z(T) \bmod P(T)$;
(3) /**compute the τ-adic representation of M* */
use Algorithm 1 *to compute expansion of* $M = \sum_{j=0}^{\lambda-1} r_j \tau^j$;
(4) /* *compute m-fold of D*;*/
(a) *initialize* $H := D(r_{\lambda-1})$;
(b) *for* $\lambda - 2 \geqslant i \geqslant 0$ *do*
(i) $H := \sigma(H)$; /* *this means cyclic shifting*/
(ii) *if* $r_i \neq 0$ *then*
$H := H + D(r_i)$; /* *one table-look-up, one addition*/
(c) *output* (H).

Step 1 needs to be performed only once per curve and base-point D, so in some applications one saves the precomputed points on the device and skips this step.

Remark 15. Arithmetic in $\mathbb{Q}$ has high system requirements. Therefore, for binary elliptic curves, Solinas [47] proposes *partial modular reduction*. Instead of computing $M \in \mathbb{Z}[\tau]$ of minimal norm he obtains an element $M' \equiv m \bmod (\tau^n - 1)/(\tau - 1)$ which might have a slightly longer expansion but the computations involve only truncated divisions by powers of 2 which can easily be realized in soft- and hardware. For the particular curves he considers one can explicitly state the group order as an expression of the degree of extension n and therefore find appropriate denominators giving a good approximation.

In our general case this is not possible, but one can work with an arbitrary good approximation $\tilde{e}'$ of e' in which all denominators are powers of two. The idea of Solinas

to use the number theoretic norm can be generalized to computing

$$\frac{(\tau_1-1)}{(\tau_1^n-1)}=\prod_{i=1}^{2g}\frac{(\tau_i-1)}{(\tau_i^n-1)}\prod_{i=2}^{2g}\frac{(\tau_i^n-1)}{(\tau_i-1)}=(kl)^{-1}\prod_{i=2}^{2g}\frac{(\tau_i^n-1)}{(\tau_i-1)}.$$

Thus, one can also precompute a Barrett-inversion of kl and get the inversion by multiplications and modular reductions.

Note that in any case the resulting $\tilde{M}$ will still be in the same class as m since $\tilde{M}=m-(\tau^n-1)/(\tau-1)\cdot\mathtt{nearest}\left(\sum_{i=0}^{2g-1}m\tilde{e}_i'\tau^i\right)\equiv m \bmod (\tau^n-1)/(\tau-1)$.

4.5. Complexity and comparison

The estimates for the complexity are given as number of group operations. Using precomputations as suggested one only needs to use additions. If the elements are represented with respect to a normal basis then $\sigma(D)$ can be computed for free. Thus we ignore these operations in the following.

By Theorem 14 the length of a τ-adic expansions is normally bounded by $n+4g+1$ in the nonperiodic case. Furthermore, the experiments show that even a bound of $n+4$ is sufficient for the range of q and g considered here.

The second important characteristic for the complexity is the density δ of the expansion. By density we mean the number of nonzero coefficients divided by the number of coefficients. Thus δ times the length gives the number of additions needed.

We first consider the minimal set $R=\{0,\pm1,\ldots,\pm\lfloor q^g/2\rfloor\}$. A zero-coefficient occurs with probability $1/q^g$, therefore the asymptotic density is $(q^g-1)/q^g<1$. Certainly all usual (signed) windowing methods carry through to τ-adic windows, thus if one precomputes all multiples $r_0D+r_1\sigma(D)+\cdots+r_{w-1}\sigma^{w-1}(D), r_i\in R, r_0\neq 0$ the density reduces to $(q^{wg}-1)/(wq^{wg})$ on average for fixed windows and to even less for sliding ones. Thus one can trade-off storage for larger speed-up. Depending on the curve one can also use other sets of coefficients, see [19,28] for details.

These numbers need to be compared to the usual arithmetic. Using binary double-and-add the number of operations is $\frac{3}{2}\log_2 m\sim\frac{3}{2}gn\log_2 q$ and using a NAF of m it still is $\sim\frac{4}{3}gn\log_2 q$.

Summary 2. *If we disregard storage and time for precomputations and assume a τ-adic expansion of length $\approx n+2g+1$, the speed-up factor is approximately*

$$\frac{3gq^g\log_2 q}{2(q^g-1)}>1.5g$$

compared to the binary expansion and

$$\frac{4gq^g\log_2 q}{3(q^g-1)}>1.3g$$

Table 2

g	Binary window	τ-adic, $w=1$	Speed-up factor	Binary window	τ-adic, $w=2$	Speed-up factor
2	$11/4n$	$3/4n$	11/3	$31/12n$	$3/7n$	$217/36 \sim 6$
3	$31/8n$	$7/8n$	31/7	$573/160n$	$7/15n$	$1719/224 \sim 7.6$
4	$79/16n$	$15/16n$	79/15	$1023/224n$	$15/31n$	$10571/1120 \sim 9.4$

Table 3

g	w_{bin}	Binary window	τ-adic, $w=1$	Speed-up factor	w_{bin}	Binary window	τ-adic, $w=2$	Speed-up factor
2	4	$47/8n$	$24/25n$	6	9	$6n$	$24/49n$	12
3	7	$511/64n$	n	8	13	$8n$	$124/249n$	16
4	9	$10n$	n	10	18	$10n$	$1/2n$	20

compared to the NAF expansion, both for the minimal set of coefficients and for n large compared to g and q.

Precomputations and signed digit expansions cannot be taken into account in a general formula, as it is a bit tricky to allow the same number of precomputations. We state $q = 2$ and 5 as examples allowing windows of length at most 2. Tables 2 and 3 list the average number of group operations to compute a scalar multiple using a signed digit windowing method and using the Frobenius endomorphism. For $q = 2$ and $w = 1$ the corresponding binary system is allowed to use a window of width g, for $w = 2$ a width of $2g - 1$ is more than fair.

For $q = 5$ we cannot directly express the width w_{bin} for the binary method as a function in g, thus we include this parameter in the table. Entries may be rounded to nearest integer.

4.6. Disadvantages

So far we have seen how to speed up the computation of m-folds on Koblitz curves. Certainly an attacker can also make use of the Frobenius endomorphism—first of all to speed up his computations. Furthermore, algorithms like Pollard's rho method allow to consider equivalence classes under the Frobenius endomorphism as "one element" (see [8,50] as the concepts generalize to hyperelliptic curves easily). This leads to a speed-up by a factor of $\sqrt{n}$ for Pollard's method. Thus allowing slightly larger field extensions n is enough to deal with this potential weakness.

The choice of Koblitz curves implies that one needs precomputations unless $g = 1, q = 2$ to obtain the described performance. Furthermore, to reduce the length of the

expansions we use polynomial arithmetic over $\mathbb{Q}$ which as well restricts the applications. We consider these two points in more detail in the next section.

5. Alternative set-up

For a cryptosystem or protocol based on Koblitz curves we now suggest to start with an expansion of fixed length and use this as the secret scalar—not caring to which integer it corresponds if at all. This implies that the device need not be able to perform polynomial arithmetic and to deal with arithmetic both in finite fields and in $\mathbb{Q}$.

If—as usual—we restrict ourselves to D of prime order l, we work in a cyclic group and σ operates as a group automorphism. Then for the action of the Frobenius we have $\sigma(D) = sD$, where s is an integer modulo l (see Remark 11). Hence, *any sum* $\sum_{i=0}^{\lambda-1} r_i \tau^i$ corresponds to an integer modulo l, namely to $\sum_{i=0}^{\lambda-1} r_i s^i \bmod l$, and one can replace the whole procedure to choose random integers and compute a reduced expansion described above by choosing a random λ-tuple of coefficients $r_i \in R$. The integer s is obtained via $\gcd(P(T), T^{n-1} + \cdots + T + 1) = (T - s)$ in $\mathbb{F}_l[T]$. This computation is done only once and s is included in the curve parameters. Here, we use the minimal set $R = \{0, \pm 1, \ldots, \pm\frac{q^g-1}{2}\}$ in odd characteristic and $R = \{0, \pm 1, \ldots, \pm(\frac{q^g}{2} - 1), \frac{q^g}{2}\}$ in even characteristic to avoid ambiguity. We assume that the chosen curve fulfills all requirements listed in the previous sections.

Remark 16. Likewise we can use the enlarged sets R of size $q^g(q^g - 1)$ and impose conditions on the density to obtain sequences $(r_0, \ldots, r_{\lambda-1})$ resembling outputs of the reduction and expansion procedure. Obviously, this leads to faster computations but it requires more precomputed points. We skip the details as most considerations are very similar.

By a random expansion of length λ we mean a tuple $r = (r_0, \ldots, r_{\lambda-1})$ with r_i chosen randomly in R along with the interpretation as $r_0 + r_1\tau + \cdots + r_{\lambda-1}\tau^{\lambda-1}$. We will show that a reasonable choice is $\lambda = n - 1$.

We first consider applications of this modified set-up and then investigate security issues.

5.1. Applications of the alternative set-up

We now care about the practicability of these new keys and show that we can still use the standard protocols:

In the *Diffie–Hellman key-exchange* [7] and in the *ElGamal cryptosystem* [9] all scalar multiplications can be performed using the random tuples. It is likewise possible that only one user applies the new idea whereas the other uses the standard Koblitz curve techniques or a binary expansion.

As *signature scheme* we choose an inversion-free scheme. However, the same considerations hold for any signature scheme. (For an overview of applicable schemes consider the Handbook of Applied Cryptography [35][Note 11.70].) Let $H(.)$ and $h(.)$

be hash functions from the message space resp. from the first polynomial of an ideal class to the integers modulo the group order l. The hash functions are public. **A** secretly chooses a and publishes $E_{\mathbf{A}} = aD$. To sign a message m, she chooses a nonce k and sends $\rho(k) = kD$ and $\mu(k, m) \equiv aH(m) + kh(kD) \bmod l$ together with the message m. To check the validity one compares μD and $H(m)E_{\mathbf{A}} + h(\rho)\rho$ and accepts upon equality. As one can see, the secret numbers a, k are not only used as multipliers but also as integers modulo l. As such they can be recovered using the correspondence of τ and s. To compute k as an integer $\tilde{k}$, we need at least $\lambda - 2$ multiplications modulo l plus some additions for the coefficients. Note, that we can compute $\tilde{k}$ and kD on the run as we need not store the coefficients and start from the highest power of s or σ respectively. Thus, we take λ random elements r of R and each time compute the intermediate results $\sigma(\tilde{\rho}) + rD$ and $s\tilde{k} + r \bmod l$ from the previous intermediate results $\tilde{\rho}$ and $\tilde{k}$. Usually this can be performed faster than computing the expansion as presented in Section 4. To obtain a we proceed the same way, but we save a together with its expansion.

5.2. Collisions

To apply the idea described in the previous section, we need to ensure that the corresponding multipliers occurring as integers modulo l are equally distributed. Respectively, we need to be aware of collisions. Since we know that $s^n \equiv 1 \bmod l$, because s corresponds to the Frobenius endomorphism on this restricted group, and $s \not\equiv 1 \bmod l$ the highest exponent of τ in the expansion should be less or equal to $n - 2$, to avoid multiple occurrences of a number. There can be other combinations of powers of s with bounded coefficients depending on the chosen curve, but here we try to exclude those polynomials that occur in any case. Namely, note that the known equivalences $1+s+\cdots+s^{n-1} \equiv 0 \bmod l$ and $s^{2g}+a_1 s^{2g-1}+\cdots+a_g s^g+\cdots+a_1 q^{g-1}s+q^g \equiv 0 \bmod l$ do not lead to such a representation, since in the first one the highest power is $n - 1$ and all powers $s^i \bmod l, 0 \leqslant i \leqslant n-2$ are different and also not equal to the negative of another power (n is an odd prime), the second one contains the coefficient $q^g \notin R$, and any combination of both still has the maximal power of $n - 1$ or too large coefficients. A mathematical sound study of collisions for these curves can be found in [31] showing that for l large enough collisions do not or only rarely occur for $\lambda = n - 1$.

5.3. Attacks

We now investigate extra weaknesses imposed by applying a random tuple instead of the expansion of a random integer. The obvious difference is that actual expansions are some τ-adic digits longer than the tuples. The standard low-storage square-root algorithms for computing discrete logarithms cannot make use of the fact that the last digits of the base τ expansion of the multiplier are zero. Clearly, a brute-force search throughout the key-space can make use of the reduced amount of possible keys, but this is far too inefficient to be threat. One can design a τ-adic baby-step-giant-step algorithm which slightly reduces the security but usually the storage requirements are prohibitively large. To play safe one should increase n by some bits.

If we consider digital signatures we have to pay more attention. We first outline what happens if parts of the *binary* expansion of the nonces are known and then show that this attack does not apply for our case. Building upon [3,22], Nguyen and Shparlinski [41] invented a way to reveal the secret signing key a if only some bits of the nonces k are known. They apply it to signature schemes based on the multiplicative group of a finite field and on elliptic curves. Although our notation and signature scheme differ from the one presented in [41], we now present essentially their ideas, however, in the new context of ideal class groups. Then we investigate to which extend this can be generalized to τ-adic bits. Let $\lfloor x \rfloor_l$ denote the unique integer $0 \leqslant \overline{x} < l$ with $x \equiv \overline{x} \bmod l$.

The task of computing a is transformed to a hidden number problem, which can be solved by lattice reduction. Assume that the highest j bits of k are known, i.e. one knows k' such that $0 \leqslant k - k' = \kappa < l/2^j$, where as before l denotes the prime group order. As $aH(m) \equiv \mu(k,m) - kh(kD) \equiv \mu(k,m) - (k' + \kappa)h(kD) \bmod l$, the attacker can compute

$$T(k,m) \equiv h(kD)^{-1}H(m) \bmod l, \quad U(k,m) \equiv -k' + \mu(k,m)h(kD)^{-1} \bmod l$$

from the publicly knows values and gets the problem of finding a such that

$$\lfloor U(k,m) - aT(k,m) \rfloor_l < l/2^j.$$

This hidden number problem can be solved given that one receives enough instances, i.e. different values of k and m for fixed a and equal most significant part of k, and that the nearest vector problem in the associated lattice can be solved (this is likely as the dimension is relatively low). Shparlinski and Nguyen verify this experimentally for elliptic curves and succeed even for a small number of known bits as $j = 3$. There are no reasons to expect a different behavior for larger genus curves. Hence, the attack has to be taken seriously.

Using our alternative scheme, the attacker knows that the "most significant τ-adic bits" of k are zero, respectively, as the corresponding integer s is easy to compute, that the highest powers of s do not occur. On the other hand we can bound s as an integer from below:

We have that $P(s) \equiv 0 \bmod l$. This equation cannot hold in the integers since we assume P to be irreducible. Hence, $s^{2g} + a_1 s^{2g-1} + \cdots + a_g s^g + \cdots + a_1 q^{g-1} s + q^g \geqslant l$. Neglecting lower order terms $s^{2g} + O(s^{2g-1}) \geqslant l \Rightarrow s \geqslant (1 + O(1))l^{1/2g} \approx q^{(n-1)/2}$. Therefore, s is large and in the expression $k = \sum_{i=0}^{n-2} r_i s^i$ one in fact computes modulo l. So, the attack does not carry through directly as one only knows that *modulo* l the highest coefficients are zero.

Techniques to solve subset sum problems [6,42] show that one can also deal with modular congruences by increasing the dimension by one and adding a further coordinate to stand for the unknown multiple of l. They show that for subset sums $\sum_{i=1}^{t} \alpha_i d_i$, $\alpha_i \in \{0,1\}$ out of t elements d_i modulo an integer l the secret coefficients can be determined for $t/\log_2 l < 0.94$. That setting allows only coefficients in $\{0,1\}$.

To model larger coefficients $\leqslant r_{\max}$ one includes $2d_i, 2^2 d_i, \ldots, 2^{\lfloor \log_2 r_{\max} \rfloor} d_i$ in the set for each d_i. If $r_{\max}$ is not a power of 2 one needs to take into account that not all linear combinations of these additional numbers are allowed but this does not affect the procedure. This enlarges t by a factor of $\approx \log_2 r_{\max}$. Allowing negative signs works just the same.

In our case l is the group order, i.e. $l \approx q^{g(n-1)}$. The d_i are the distinct powers s^i of s modulo l. Therefore the set contains $t = \lfloor 2(n-1)\log_2(\lfloor q^g/2 \rfloor) \rfloor \approx (n-1)g \log_2 q$ elements. The fraction $t/\log_2 l$ is very close to 1—for the approximations detailed above it even equals 1. In general, one should check that $t/\log_2 l > 0.94$ before applying a curve.

6. Example

In this section we present one example, however, further good instances are easy to get [27]. Consider the binary curve of genus 2 given by

$$C : y^2 + (x^2 + x + 1)y = x^5 + x + 1$$

with characteristic polynomial of the Frobenius endomorphism $P(T) = T^4 - 2T^3 + 3T^2 - 4T + 4$. For the extension of degree 89 the class number is almost prime

$$|\mathrm{Cl}(C/\mathbb{F}_{2^{89}})| = 2 \cdot 191561942608242456073498418252108663615312031512914969.$$

Let l be the large prime number. The operation of σ on the group of order l corresponds to the multiplication by

$$s = -109094763598619410884498554207763796660522627676801041 \bmod l$$

For a high-level comparison we provide two Magma programs. The program for this curve `FrobExample` and a program to play around with a user-defined curve `FrobSelf` can be obtained from [27]. A detailed paper about implementation of hyperelliptic Koblitz curves using normal and polynomial bases in comparison is in preparation [30]. It gives evidence that the theoretic and asymptotic results of this paper actually hold true in practice.

7. Conclusion

We gave details on the use of Koblitz curves and presented an alternative set-up in which the random integer m is replaced by a random $n-1$ tuple of elements from R. This alternative set-up allows to save the time needed to compute the expansion. Furthermore, in this case the mathematical features needed are reduced to a minimum,

e.g. no arithmetic in $\mathbb{Q}$ is used. Hence, this set-up is especially appropriate for memory-constrained environments like smart cards. The devices of the participants need only be able to perform addition, to execute σ, and to randomly choose elements from R. A little amount of storage is required to keep precomputed multiples.

The proposed alternative set-up can be applied to the usual protocols where in the case of a signature scheme one needs to compute the secret multiple as an integer as well. Concerning security issues, we considered generalizations of known attacks and dealt with collisions. To conclude one can say that using this modified system saves the time needed to compute the expansion without weakening the system.

An extremely careful user might feel better to use it only for ElGamal and Diffie–Hellman although to our knowledge signature schemes are just as well secure.

Remark 17. (1) In this paper we considered the effects of known τ-adic bits only in the section on the alternative set-up. The same considerations hold true for side-channel attacks where the leakage allows to obtain some τ-adic bits. Our analysis shows that Koblitz curve systems are not vulnerable to such attacks if the number of leaked bits is small, such that the parameter t in Section 5.3 is close to 1.

We thank the anonymous referee for pointing out this observation.

(2) One can restrict the key size even more by choosing a smaller set of coefficients for the τ-adic expansion. This reduces the storage requirements and the probability of collisions but for extreme choices—like $R' = \{0, \pm 1\}$, $g, q > 2$, thus without precomputations—one has to be aware of lattice based attacks on the subset sum problem [6,42]. If one tries to get around these by using longer keys of length $n + \varepsilon$, collisions get more likely since one has to deal with $1 + s + \cdots + s^{n-1} \equiv 0 \bmod l$. Then the zero element occurs at least $2\binom{\varepsilon + r'_{\max} - 1}{r'_{\max}} + 1$ times, where $r'_{\max}$ is the maximal coefficient of R'. Another idea is to consider only sparse representations to reduce the complexity. Although this reduces the size of the key-space as well, the implications are less dramatic.

(3) The use of reduced τ-expansions may help to improve any cryptographic method of key-exchange, signing and encryption based on the Jacobian of curves or other Abelian varieties which are defined over a smaller field than they are considered. Included are for example Jacobians of superelliptic and C_{ab}-curves and one might apply the construction to other efficiently computable endomorphisms with known characteristic polynomial.

(4) Unless $P(T) = T^{2g} + q^g$, the standard method as well as the alternative set-up can be applied to speed up pairing schemes based on supersingular curves, as pointed out by Stein.

Acknowledgments

This work evolved from my Ph.D. thesis, I express my deepest gratitude to my supervisor Gerhard Frey for everything he did. I would also like to thank Guillaume

Hanrot, Hendrik W. Lenstra, Phong Nguyen, Michael Pohst, René Schoof, and Igor Shparlinski for interesting discussions and their valuable suggestions.

References

[1] S. Arita, Algorithms for computations in jacobian group of C_{ab} curve and their application to discret-log-based public key cryptosystems, in: The Mathematics of Public Key Cryptography, Fields Institute, Toronto, 1999, pp. 1291–1299.
[2] A. Basiri, A. Enge, J.C. Faugère, N. Gürel, Implementing the arithmetic of $C_{3,4}$ curves, in: Algorithmic Number Theory Seminar ANTS-VI, Lecture Notes in Computer Science, vol. 3076, Springer, Berlin, 2004, pp. 87–101.
[3] D. Boneh, R. Venkatesan, Hardness of computing the most significant bits of secret keys in Diffie–Hellman and related schemes, in: Advances in Cryptology—Crypto '96, Lecture Notes in Computer Science, vol. 1109, Springer, Berlin, 1996, pp. 129–142.
[4] D.G. Cantor, Computing in the Jacobian of a hyperelliptic curve, Math. Comp. 48 (1987) 95–101.
[5] Y. Choie, J.W. Lee, Speeding up the scalar multiplication in the Jacobian of hyperelliptic curves using Frobenius map, in: Progress in Cryptology—Indocrypt 2002, Lecture Notes in Computer Science, vol. 2551, Springer, Berlin, 2002, pp. 285–295.
[6] M. Coster, A. Joux, B. LaMacchia, A. Odlyzko, C.-P. Schnorr, J. Stern, Improved low-density subset sum algorithms, Comp. Compl. 2 (1992) 111–128.
[7] W. Diffie, M.E. Hellman, New directions in cryptography, IEEE Trans. Inform. Theory 22 (6) (1976) 644–654.
[8] I. Duursma, P. Gaudry, F. Morain, Speeding up the discrete log computation on curves with automorphisms, in: Advances in cryptology—Asiacrypt'99, Lecture Notes in Computer Science, vol. 1716, Springer, Berlin, 1999, pp. 103–121.
[9] T. ElGamal, A public key cryptosystem and a signature scheme based on discrete logarithms, IEEE Trans. Inform. Theory 31 (4) (1985) 469–472.
[10] U. Finke, M. Pohst, Methods for calculating vectors of short length in a lattice, Math. Comput. 44 (1985) 463–482.
[11] S. Flon, R. Oyono, Fast arithmetic on Jacobians of Picard curves, in: Public Key Cryptography—PKC 2004, Lecture Notes in Computer Science, vol. 2947, Springer, Berlin, 2004, pp. 55–68.
[12] S. Flon, R. Oyono, C. Ritzenthaler, Fast addition on non-hyperelliptic genus 3 curves, cryptology ePrint Archive, Report 2004/118 (2004).
[13] G. Frey, T. Lange, Mathematical Background of Public Key Cryptography, Technical Report, vol. 10, IEM Essen, 2003.
[14] G. Frey, H.G. Rück, A remark concerning m-divisibility and the discrete logarithm problem in the divisor class group of curves, Math. Comp. 62 (1994) 865–874.
[15] S.D. Galbraith, Supersingular curves in cryptography, in: Advances in Cryptology—Asiacrypt 2001, Lecture Notes in Computer Science, vol. 2248, Springer, Berlin, 2001, pp. 495–513.
[16] S.D. Galbraith, Weil descent of Jacobians, in: D. Augot, C. Carlet (Eds.), WCC2001, Electronic Notes in Discrete Mathematics, vol. 6, Elsevier, Amsterdam, 2001, <http://www.elsevier.nl/gej-g/31/29/24/show/Products/notes/index.htt>.
[17] P. Gaudry F. Hess N.P. Smart, Constructive and destructive facets of Weil descent on elliptic curves, J. Cryptol. 15 (1) (2002) 19–46. <http://www.hpl.hp.com/techreports/2000/HPL-2000-10.html>.
[18] P. Gaudry, E. Schost, Construction of secure random curves of genus 2 over prime fields, in: Advances in Cryptology—Eurocrypt'2004, Lecture Notes in Computer Science, vol. 3027, Springer, Berlin, 2004, pp. 239–256.
[19] C. Günther, T. Lange, A. Stein, Speeding up the arithmetic on Koblitz curves of genus two, in: Selected Areas in Cryptography—SAC 2000, Lecture Notes in Computer Science, vol. 2012, Springer, Berlin, 2000, pp. 106–117.
[20] D. Hankerson, J. Hernandez, A. Menezes, Software implementation of elliptic curve cryptography over binary fields, in: Cryptographic Hardware and Embedded Systems CHES 2000, Lecture Notes in Computer Science, vol. 1965, Springer, Berlin, 2000, pp. 1–24.

[21] F. Hess, G. Seroussi, N.P. Smart, Two topics in hyperelliptic cryptography, in: Selected Areas in Cryptography—SAC 2001, Lecture Notes in Computer Science, vol. 2259, Springer, Berlin, 2001, pp. 181–189.
[22] N.G. Howgrave-Graham, N.P. Smart, Lattice attacks on digital signature schemes, Des. Codes Cryptography 23 (2001) 283–290.
[23] N. Koblitz, Hyperelliptic cryptosystems, J. Cryptology 1 (1989) 139–150.
[24] N. Koblitz, CM-curves with good cryptographic properties, in: Advances in Cryptology—Crypto'91, Lecture Notes in Computer Science, vol. 576, Springer, Berlin, 1992, pp. 279–287.
[25] J. Kuroki, M. Gonda, K. Matsuo, J. Chao, S. Tsuji, Fast genus three hyperelliptic curve cryptosystems, in: Proceedings of SCIS2002, IEICE, Japan, 2002, pp. 503–507.
[26] T. Lange, Efficient arithmetic on hyperelliptic Koblitz curves, Technical Report 2-2001, University Essen, 2001.
[27] T. Lange, Hyperelliptic curves allowing fast arithmetic, 2001, <http://www.itsc.ruhr.uni-bochum.de/tanja/KoblitzC.html>.
[28] T. Lange, Efficient arithmetic on hyperelliptic curves, Ph.D. Thesis, University Essen, 2001.
[29] T. Lange, Formulae for arithmetic on genus 2 hyperelliptic curves, http://www.itsc.ruhr-uni-bochum.de/tanja/preprints.html, J. AAECC (2004), to appear.
[30] T. Lange, M. Nöcker, M. Stevens, Optimal implementation of hyperelliptic Koblitz curves over $\mathbb{F}_{2^n}$, in preparation.
[31] T. Lange, I. Shparlinski, Collisions in fast generation of ideal classes and points on hyperelliptic and elliptic curves, J. AAECC (2004), to appear.
[32] J.-L. Lesage, Equations diophantiennes et corps quadratiques, Ph.D. Thesis, Université de Caen, 1998.
[33] D. Lorenzini, An invitation to Arithmetic Geometry, Graduate Studies in Mathematics, vol. 9, American Mathematical Society, Providence, RI, 1996.
[34] A. Menezes, T. Okamoto, S. Vanstone, Reducing elliptic curve logarithms to a finite field, IEEE Trans. Inform. Theory 39 (1993) 1639–1646.
[35] A. Menezes, P. van Oorschot, S. Vanstone, Handbook of Applied Cryptography, CRC Press, Boca Raton, FL, 1996.
[36] A. Menezes, M. Qu, Analysis of the Weil descent attack of Gaudry, Hess and Smart, preprint.
[37] A.J. Menezes, S. Vanstone, The implementation of elliptic curve cryptosystems, in: Advances in cryptology—AUSCRYPT '90, Lecture Notes in Computer Science, vol. 453, Springer, Berlin, 1990, pp. 2–13.
[38] A.J. Menezes, Y.-H. Wu, R. Zuccherato, An elementary introduction to hyperelliptic curves, in: N. Koblitz (Ed.), Algebraic Aspects of Cryptography, Springer, Berlin, 1998, pp. 155–178.
[39] W. Meier, O. Staffelbach, Efficient multiplication on certain nonsupersingular elliptic curves, in: Advances in Cryptology—Crypto'92, Lecture Notes in Computer Science, vol. 740, Springer, Berlin, 1993, pp. 333–344.
[40] V. Müller, Fast multiplication on elliptic curves over small fields of characteristic two, J. Cryptol. 11 (1998) 219–234.
[41] P.Q. Nguyen, I.E. Shparlinksi, The insecurity of the elliptic curve digital signature algorithm with partially known nonces, Des. Codes Cryptography 30 (2003) 201–217.
[42] P.Q. Nguyen, J. Stern, The hardness of the hidden subset sum problem and its cryptographic implications, in: Advances in Cryptology—Crypto '99, Lecture Notes in Computer Science, vol. 1666, Springer, Berlin, 1999, pp. 31–46.
[43] Y.-H. Park, S. Jeong, J. Lim, Speeding up point multiplication on hyperelliptic curves with efficiently-computable endomorphisms, in: Advances in Cryptology—Eurocrypt 2002, Lecture Notes in Computer Science, vol. 2332, Springer, Berlin, 2002, pp. 197–208.
[44] J. Pelzl, Fast hyperelliptic curve cryptosystems for embedded processors, Master's Thesis, Ruhr-University of Bochum, 2002.
[45] N.P. Smart, Elliptic curve cryptosystems over small fields of odd characteristic, J. Cryptol. 12 (1999) 141–151.
[46] J. Solinas, An improved algorithm for arithmetic on a family of elliptic curves, in: Advances in Cryptology—Crypto '97, Lecture Notes in Computer Science, vol. 1294, Springer, Berlin, 1997, pp. 371–375.

[47] J. Solinas, Efficient arithmetic on Koblitz curves, Designs Codes Cryptography 19 (2000) 195–249.
[48] C. Stahlke, Point compression on Jacobians of hyperelliptic curves over $\mathbb{F}_q$, cryptology ePrint Archive, Report 2004/030 (2004).
[49] H. Stichtenoth, Algebraic Function Fields and Codes, Springer, Berlin, 1993.
[50] M. Wiener, R. Zuccherato, Faster attacks on elliptic curve cryptosystems, in: Selected Areas in Cryptography—SAC'98, Lecture Notes in Computer Science, vol. 1556, Springer, Berlin, 1998, pp. 190–200.