Algorithmen & mathematische Modellierung

Forschungsschwerpunkt

Algorithmen und mathematische Modellierung

TU Graz
Graz University of Technology

Redundant τ-adic Expansions II: Non-Optimality and Chaotic Behaviour

Clemens Heuberger

Project Area(s):
Analysis of Digital Expansions with Applications in Cryptography

Institut für Optimierung und Diskrete Mathematik (Math B)

Report 2008-4, February 2008

Redundant τ-adic Expansions II: Non-Optimality and Chaotic Behaviour

Clemens Heuberger

Abstract. When computing scalar multiples on Koblitz curves, the Frobenius endomorphism can be used to replace the usual doublings on the curve. This involves digital expansions of the scalar to the complex base $\tau = (\pm 1 \pm \sqrt{-7})/2$ instead of binary expansions. As in the binary case, this method can be sped up by enlarging the set of valid digits at the cost of precomputing some points on the curve. In the binary case, it is known that a simple syntactical condition (the so-called w-NAF-condition) on the expansion ensures that the number of curve additions is minimised. The purpose of this paper is to show that this is not longer true for the base τ and $w \in \{4, 5, 6\}$. Even worse, it is shown that there is no longer an online algorithm to compute an optimal expansion from the digits of some standard expansion from the least to the most significant digit, which can be interpreted as chaotic behaviour. The proofs heavily depend on symbolic computations involving transducer automata.

Mathematics Subject Classification (2000). 11A63, 68W13, 68Q45, 94A60, 90C27.

1. Introduction

The principle of elliptic curve cryptography is that scalar multiples of a point can be computed quickly, whereas the inverse operation, the discrete logarithm problem, is believed to be hard. It is a natural aim to optimise the scalar multiplication.

In [15], Koblitz discussed the curves (since then associated with his name)

$$E_a : Y^2 + XY = X^3 + aX^2 + 1, \qquad \text{with } a \in \{0, 1\},$$

which are defined over $\mathbb{F}_2$ and whose point group $E_a(\mathbb{F}_{2^n})$ over $\mathbb{F}_{2^n}$ is considered. The Frobenius automorphism $\tau : \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}$, which sends an element to its square, can be extended to an endomorphism of $E_a(\mathbb{F}_{2^n})$. It turns out [15] that τ satisfies the equation

$$\tau^2 - \mu\tau + 2 = 0 \tag{1}$$

where $\mu = (-1)^{1-a}$. This implies that τ can be identified with the complex number

$$\frac{\mu + \sqrt{-7}}{2}.$$

According to the classification of Kátai and Kovács [12], the imaginary quadratic number τ is the basis of a canonical number system (cf. Kátai and Szabó [13]) in $\mathbb{Z}[\tau]$, i.e., every element $z \in \mathbb{Z}[\tau]$ can be represented as $z = \sum_{j=0}^{\ell} \eta_j \tau^j$ for digits $\eta_j \in \{0, 1\}$. Using this representation, a scalar multiple $n \cdot P$ with $n \in \mathbb{Z}$ (or even $n \in \mathbb{Z}[\tau]$) and $P \in E_a(\mathbb{F}_{2^n})$ can be computed as

$$n \cdot P = \sum_{j=0}^{\ell} \eta_j \tau^j(P). \tag{2}$$

C. Heuberger is supported by the Austrian Science Foundation FWF, project S9606, that is part of the Austrian National Research Network "Analytic Combinatorics and Probabilistic Number Theory."

The latter sum can be evaluated efficiently using Horner's scheme. This is a generalisation of the double-and-add scheme (cf. Knuth [14]) on arbitrary elliptic curves, where a digit expansion to the base of 2 is used. The attractive feature is that an application of the Frobenius endomorphism is much cheaper (or even almost free when normal bases are used) than doubling.

The number of elliptic curve additions required to calculate $n \cdot P$ using (2) almost (i.e., one addition less is required) equals the Hamming weight of $(\eta_\ell \dots \eta_0)$, i.e., the number of nonzero digits η_j.

On an elliptic curve, subtraction of a point is (almost) as cheap as addition of a point. Therefore, Morain and Olivos [17] proposed (in the binary case) to allow digits -1, too. This can be carried over to Koblitz curves. Since there are many representations $z = \sum_{j=0}^{\ell} \eta_j \tau^j$ with digits $\eta_j \in \{0, 1, -1\}$, one can choose a representation that minimises the Hamming weight. It turns out that every $z \in \mathbb{Z}[\tau]$ has a unique representation such that $\eta_j \cdot \eta_{j+1} = 0$ for all j (cf. Solinas [23, 24]), called the τ-Non-Adjacent Form (or τ-NAF)[1], and that the τ-NAF minimises the Hamming weight (cf. Gordon [7] and Avanzi, Heuberger and Prodinger [3, 4]). In the binary case, the same has already be shown by Reitwiesner [22].

If one allows still larger digit sets, the Hamming weight can be further decreased. This comes at the cost of precomputing and storing $\eta_j \cdot P$ for all digits η_j. Solinas [24] proposed the following set of digits: Fix a positive integer w and for each residue class modulo τ^w coprime to τ, choose the element of minimal norm to be a digit. Furthermore, 0 is a digit. This construction is called the digit set of minimal norm representatives modulo τ^w. Solinas proved that every element $z \in \mathbb{Z}[\tau]$ admits a unique representation $z = \sum_{j=0}^{\ell} \eta_j \tau^j$ such that among every w consecutive digits, there is at most one non-zero digit. This representation is called the τ-w-NAF of z. For $w = 1$, this corresponds to the canonical number system representation with digits from $\{0, 1\}$. The τ-2-NAF is just the τ-NAF described above.

This paper deals with the question whether the τ-w-NAF of an element $z \in \mathbb{Z}[\tau]$ minimises the Hamming weight over all representations of z with the same digits, but without the non-adjacency condition. For $w = 1$, this is trivially true. As mentioned above, for $w = 2$, optimality has been shown in [7, 3, 4]. For $w = 3$, optimality has been shown in Avanzi, Heuberger and Prodinger [3, 4] in a slightly different language. In this paper, we show that the τ-w-NAF (with the minimal norm representatives as digits) is *not* optimal for $w \in \{4, 5, 6\}$. This is in sharp contrast to the binary case, where the digit set of minimal norm representatives simply consists of zero and all odd integers of absolute value less than 2^{w-1} and where w-NAFs with this digit set minimise the Hamming weight, cf. Avanzi [1], Muir and Stinson [19] and Phillips and Burgess [20].

This raises the question whether the concept of the τ-w-NAF is the "right" concept. Is it possible to choose another syntactic condition which also ensures uniqueness of the representation and, at the same time, minimises the Hamming weight? On of the attractive features of the τ-w-NAF is the fact that it can be computed by an online algorithm from any representation with a pre-specified set of digits from right to left, i.e., starting with the least significant digit. This is equivalent to the existence of a finite deterministic transducer automaton to compute the τ-w-NAF from right to left.

The—perhaps surprising—answer given in this paper is that it is impossible to construct an online algorithm (or, equivalently, a finite deterministic transducer automaton) to compute an optimal representation from right to left for $w \in \{4, 5, 6\}$. In particular, we will exhibit examples of pairs of integers which are congruent modulo arbitrarily high powers of τ, but whose least significant digits in their optimal expansions have to different. One would conjecture a similar behaviour for higher values of w, but at present, the required symbolic computations seem to be out of reach with current computers. We note that a similar behaviour has also been found by the author [9] for bases of canonical number systems in the Gaussian integers. One may interpret this as chaotic behaviour, since knowing the input arbitrarily precisely (in a τ-adic sense) does not allow to determine the least significant digit of optimal expansions.

[1] This name comes from the fact that $\eta_j \eta_{j+1} = 0$ implies that a τ-NAF does not have adjacent non-zeros.

The digit set of minimal norm representatives is not the only useful one, cf. Avanzi, Heuberger and Prodinger [5, 2]. We also provide examples of chaos for other digit sets considered in that paper.

In some cryptosystems, e.g., ECDSA, it is required to compute linear combinations $m\cdot P+n\cdot Q$ for integers m and n and points P, $Q\in E_a(\mathbb{F}_{2^n})$. As remarked by Straus [26], one can do better than simply adding the results of the scalar multiplications $m\cdot P$ and $n\cdot Q$: Using a *joint expansion* $\binom{m}{n}=\sum_{j=0}^{\ell}\boldsymbol{\eta}_j 2^j$ where $\boldsymbol{\eta}_j\in\{0,1\}^2$ are digit vectors and precomputing $P+Q$, one can compute the linear combination by ℓ doublings and $H(\boldsymbol{\eta}_\ell,\dots,\boldsymbol{\eta}_0)-1$ additions, where $H(\boldsymbol{\eta}_\ell,\dots,\boldsymbol{\eta}_0)$ denotes the joint Hamming weight, i.e., the number of nonzero digit vectors. This procedure is also known as Shamir's trick. As in the one-dimensional case, one can allow digits -1 to introduce redundancy. Solinas [25] introduced the *Joint Sparse Form* by syntactical conditions, which is unique and minimises the joint Hamming weight. Generalisations have been made in Grabner, Heuberger and Prodinger [8], Proos [21] and Heuberger and Muir [10].

For the case of base τ, the same syntactical condition can be imposed (cf. Ciet, Lange, Sica and Quisquater [6]). Uniqueness is preserved, but minimality is not. In fact, more complicated syntactical properties have been proposed by Zhu, Kuang and Zhang [27], where the average joint Hamming weight could be reduced, but which are still non-optimal. After the above announced results on chaotic behaviour of τ-w-NAF, one may wonder whether the same is true for joint expansions in base τ. The answer is affirmative and is proved in Theorem 3.

We now turn to the methods employed in this paper. At first glance, computing optimal expansions of an element seems to be a difficult task with exponential running time, because for each digit, one has a certain number of choices. A closer analysis, however, shows that only a finite number of carries can actually occur, which is encoded by the transducer automaton translating any expansion with the given digit set to the "canonical" representation, i.e., the τ-w-NAF or the τ-JSF. This results in an algorithm to compute optimal expansions which is linear in the length of the expansions—but beware, the implicit constant (depending on w and the digit set) is huge.

A computer search exhibits candidates for pairs of integers which are arbitrarily close, but whose least significant digits in their optimal expansions are different. The above mentioned linear time algorithm is then tweaked to deal with those candidates, which is possible due to the essentially periodic patterns in their canonical representations. This leads to heavy symbolic computations with transducer automata. These result in a shortest path calculation in a large directed graph. All shortest paths correspond to all optimal expansions of the given integers.

The remaining paper is structured as follows: In Section 2, we fix the notations and introduce the digit sets used in this paper. The following Section 3 collects all results in the one-dimensional case. The transducer automata forming the base of the remaining proofs are introduced in Section 4. Section 5 is devoted to the asymptotically efficient computation of optimal expansions of single elements of $\mathbb{Z}[\tau]$, whereas Section 6 deals with the computation of optimal expansions of families of integers given by essentially periodic $\mathcal{D}$-w-NAFs, which leads to the proof of the results in the one-dimensional case. Finally, Section 7 discusses the case of higher dimensions.

2. $\mathcal{D}$-expansions

Let $\mu\in\{\pm1\}$ and τ be a root of the equation (1). It is well known that $\mathbb{Z}[\tau]$ is an Euclidean domain and that τ is a prime element in this ring. For $w\geq 1$, the prime residue classes modulo τ^w are those residue classes modulo τ^w that are relatively prime to τ.

We now collect the basic definitions and notations used in the first part of this paper. We use standard notations for finite and infinite words, automata and transducer automata, cf. for instance Lothaire [16]. However, our infinite words will be left-infinite, i.e., $\dots\eta_3\eta_2\eta_1\eta_0$, and all automata will process their arguments from right to left. The length of a finite word will be denoted by $\mathsf{length}(\eta_{\ell-1}\dots\eta_0):=\ell$. In order to avoid any confusion, we write

$$w^{(\ell)}=\underbrace{w\dots w}_{\ell\text{ repetitions}}$$

for the ℓth power of the finite word w with respect to concatenation and reserve superscripts without parentheses for powers of complex numbers with respect to the usual multiplication in $\mathbb{C}$.

Definition 2.1. Let $\mathcal{D}$ be a (finite) subset of $\mathbb{Z}[\tau]$ containing 0. A *$\mathcal{D}$-expansion* of $z \in \mathbb{Z}[\tau]$ is a left infinite word $\boldsymbol{\eta} = \dots \eta_2\eta_1\eta_0 \in \mathcal{D}^\omega$ over the alphabet $\mathcal{D}$ such that

1. only a finite number of the digits η_j is nonzero,
2. $\mathsf{value}(\boldsymbol{\eta}) := \sum_{j\geq 0} \eta_j\tau^j = z$, i.e., $\boldsymbol{\eta}$ is indeed an expansion of z.

The *Hamming weight* $\mathsf{weight}(\boldsymbol{\eta})$ of $\boldsymbol{\eta}$ is the number of nonzero digits η_j.

The *length* of the expansion[2] $\boldsymbol{\eta}$ is defined as

$$\mathsf{length}(\boldsymbol{\eta}) := 1 + \max\{j : \eta_j \neq 0\}.$$

A $\mathcal{D}$-expansion $\boldsymbol{\eta}$ of z is called an *optimal $\mathcal{D}$-expansion* of z if its Hamming weight is minimum amongst all $\mathcal{D}$-expansions of z. The set of optimal $\mathcal{D}$-expansions of z is denoted by

$$\mathsf{opt}(z) := \{\boldsymbol{\eta} \in \mathcal{D}^\omega : \boldsymbol{\eta} \text{ is an optimal } \mathcal{D}\text{-expansion of } z\}.$$

Let $w \geq 1$ be an integer. A $\mathcal{D}$-expansion of z is called a *$\mathcal{D}$-w-Non-Adjacent-Form ($\mathcal{D}$-w-NAF) of z*, if

3. each factor $\eta_{j+w-1}\dots\eta_j$ of length w, i.e., each block of w consecutive digits, contains at most one nonzero digit η_k, $j \leq k \leq j+w-1$.

A $\{0, \pm 1\}$-2-NAF is also called a *τ-NAF*.

A set $\mathcal{D}$ which consists of zero and exactly one representative of every prime residue class modulo τ^w and such that each $z \in \mathbb{Z}[\tau]$ admits a $\mathcal{D}$-w-NAF is called a w-Non-Adjacent-Digit-Set (w-NADS).

It is easily seen that if $\mathcal{D}$ is a w-NADS, then each $z \in \mathbb{Z}[\tau]$ has a *unique* $\mathcal{D}$-w-NAF, which will be denoted by

$$\mathsf{NAF}(z).$$

Furthermore, if $z \equiv z' \pmod{\tau^{k+w}}$ for some integer k, then the k least significant digits of their $\mathcal{D}$-w-NAFs agree.

The following families of digit sets $\mathcal{D}$ will be considered in this paper.

2.1. Minimal Norm Representatives modulo τ^w

In Avanzi, Heuberger and Prodinger [5], it has been shown that every prime residue class modulo τ^w indeed contains exactly one element of minimal norm.

Definition 2.2. Let $w \geq 2$. Then the set $\mathsf{MNR}(w)$ consisting of 0 and the unique element of minimal norm for every prime residue class modulo τ^w is called the *set of minimal norm representatives modulo τ^w*.

Solinas [24] proved that for $w \geq 1$, $\mathsf{MNR}(w)$ is indeed a w-NADS.

2.2. Short τ-NAF Representatives

Definition 2.3. Let $w \geq 2$. Then the set $\mathsf{SNR}(w)$ is defined as

$$\mathsf{SNR}(w) = \{0\} \cup \{\mathsf{value}(\boldsymbol{\eta}) : \boldsymbol{\eta} \text{ is a } \tau\text{-NAF of length at most } w \text{ with } \eta_0 \neq 0 \text{ and } \eta_{w-1} \in \{0, \eta_0\}\}$$

and is called the *set of short τ-NAF representatives.*

In [2] it is shown that $\mathsf{SNR}(w)$ is a w-NADS for all $w \geq 2$. In fact, $\mathsf{SNR}(w) = \mathsf{MNR}(w)$ for $w \in \{2, 3\}$ as well as for $(w, \mu) = (4, 1)$. For $w \geq 4$, the rule $\eta_{w-1} \in \{0, \eta_0\}$ is somewhat arbitrary (cf. [2]), but in this paper, we stick to this definition.

[2] We use the same notation as for the length of finite words, where we also count leading zeros, which would be meaningless in the case of an infinite word.

2.3. Powers of $\bar{\tau}$

Definition 2.4. Let $w \geq 2$. Then the set $\mathsf{P}\bar{\tau}(w)$ is defined as

$$\mathsf{P}\bar{\tau}(w) = \{0\} \cup \{\pm\bar{\tau}^k : 0 \leq k < 2^{w-2}\}$$

and is called the *set of powers of* $\bar{\tau}$.

For $w \geq 2$, $\mathsf{P}\bar{\tau}(w)$ contains exactly one representative of every prime residue class modulo τ^w, cf. [2]. For $w \in \{2,3,4,5,6\}$, $\mathsf{P}\bar{\tau}(w)$ is even a w-NADS, for $w \in \{7,8,9,10,11,12\}$ it is not a w-NADS, cf. [2]. It turns out that $\mathsf{MNR}(w) = \mathsf{P}\bar{\tau}(w)$ for $w \in \{2,3,4\}$. As explained in [2], this digit set can be used for a sub-linear scalar multiplication algorithm on Koblitz curves, the key ingredient is a relation between multiplication by $\bar{\tau}$ and point halving on the curve.

3. Results in the one-dimensional case

Avanzi, Heuberger and Prodinger [3, 4] showed that for $w \in \{2,3\}$ a $\mathcal{D}$-w-NAF with $\mathcal{D}$ being the set of minimal norm representatives modulo τ^w is actually always an optimal $\mathcal{D}$-expansion. The same result is also trivially true for $w = 1$.

In the binary case (where τ is replaced by 2), it turns out that the analogous result is true for all positive w, cf. Avanzi [1] and Muir and Stinson [18]. So one might conjecture that the same is also true for our choice of τ.

We show that this conjecture is false by considering the following example.

Example 3.1. Consider $\mu = -1$, $w = 4$, and the set $\mathcal{D} = \mathsf{MNR}(4)$, i.e., $\mathcal{D} = \{0, \pm 1, \pm 1 \pm \tau, \pm(3+\tau)\}$ (all signs are independent). We note that

$$\mathsf{value}(0^\omega 1000(-1-\tau)000(1-\tau)) = -9 = \mathsf{value}(0^\omega(-3-\tau)00(-1)).$$

The first expansion is the $\mathcal{D}$-w-NAF and has Hamming weight 3, whereas the second expansion does not satisfy the $\mathcal{D}$-w-NAF-condition, has Hamming weight 2 and is even shorter.

Examples for other values of w and $\mathcal{D}$ can easily be extracted from Table 1.

Even worse, we exhibit chaotic behaviour in the following sense: for every positive integer ℓ, we exhibit a pair of numbers which are congruent modulo τ^ℓ, but whose optimal $\mathcal{D}$-expansions must differ even at the least significant position. Thus it is impossible to compute an optimal $\mathcal{D}$-expansion of z by a deterministic transducer automaton or an online algorithm.

We remark that such chaotic behaviour has previously been found to occur in $\{0, \pm 1\}$-expansions in $\mathbb{Z}[i]$, cf. Heuberger [9].

Theorem 1. *Let*

- $w = 4$ *and* $\mathcal{D} \in \{\mathsf{MNR}(4), \mathsf{SNR}(4), \mathsf{P}\bar{\tau}(4)\}$ *or*
- $w = 5$ *and* $\mathcal{D} \in \{\mathsf{MNR}(5), \mathsf{SNR}(5), \mathsf{P}\bar{\tau}(5)\}$ *or*
- $w = 6$ *and* $\mathcal{D} \in \{\mathsf{MNR}(6), \mathsf{SNR}(6)\}$.

For every positive integer ℓ, *there exist elements* z_ℓ, $z'_\ell \in \mathbb{Z}[\tau]$ *given in Table* 1 *with the following two properties:*

1. *The numbers* z_ℓ *and* z'_ℓ *are congruent modulo* τ^ℓ.
2. *For all optimal* $\mathcal{D}$*-expansions* $\boldsymbol{\eta}$ *and* $\boldsymbol{\eta}'$ *of* z_ℓ *and* z'_ℓ, *respectively, the least significant digits* η_0 *and* η'_0 *differ.*

For clarity, we state the result for $w = 4$ explicitly:

Example 3.2. Let $w = 4$ and $\mathcal{D} = \mathsf{MNR}(4) = \{0, \pm 1, \pm 1 \pm \tau, \pm(3 - \mu\tau)\}$. For every nonnegative integer ℓ, we define

$$\begin{aligned} z_\ell &:= \mathsf{value}\big(\quad 0^\omega(\mu-\tau)(000(-3\mu+\tau))^{(\ell)}0000(1-\mu\tau)000(-1)\big), \\ z'_\ell &:= \mathsf{value}\big(0^\omega(-\mu)000(\mu-\tau)(000(-3\mu+\tau))^{(\ell)}0000(1-\mu\tau)000(-1)\big), \end{aligned} \tag{3}$$

where $(000(-3\mu+\tau))^{(\ell)}$ means that this 4 digit block is repeated ℓ times.

Then $z_\ell \equiv z'_\ell \pmod{\tau^{4\ell+13}}$. All $\mathcal{D}$-optimal expansions of z_ℓ are given by

$$0^\omega(000(3-\mu\tau))^{(\ell_2)}00(\mu-\tau)(000(-3\mu+\tau))^{(\ell_1)}0000(1-\mu\tau)000(-1), \tag{4}$$

where ℓ_1 and ℓ_2 are nonnegative integers summing up to ℓ. There is only one $\mathcal{D}$-optimal expansion of z'_ℓ, it is given by

$$0^\omega(000(-3+\mu\tau))^{(\ell+1)}0000(-3\mu+\tau)00(1+\mu\tau). \tag{5}$$

In particular, the least significant digit of all optimal expansions of z_ℓ is -1, whereas the unique optimal expansion of z'_ℓ has least significant digit $(1+\mu\tau)$.

Note that the $\mathcal{D}$-optimal expansion of z'_ℓ has Hamming weight $\ell+3$, whereas the $\mathcal{D}$-w-NAF of z'_ℓ given in (3) has Hamming weight $\ell+4$.

w	μ	$\mathcal{D}$	
4	μ	MNR	$\mathsf{NAF}(z_\ell) = 0^\omega(\mu-\tau)\,(000(-3\mu+\tau))^{(\ell)}\,0000(1-\mu\tau)000(-1)$
	($\ell \geq 0$)		$\mathsf{opt}(z_\ell) = \{0^\omega\,(000(3-\mu\tau))^{(\ell_2)}\,00(\mu-\tau)\,(000(-3\mu+\tau))^{(\ell_1)}$ $0000(1-\mu\tau)000(-1) \mid \ell_1,\ell_2 \geq 0 \text{ and } \ell_1+\ell_2=\ell\}$
			$\mathsf{NAF}(z'_\ell) = 0^\omega(-\mu)000(\mu-\tau)\,(000(-3\mu+\tau))^{(\ell)}\,0000(1-\mu\tau)000(-1)$
			$\mathsf{opt}(z_\ell) = \{0^\omega\,(000(-3+\mu\tau))^{(\ell+1)}\,0000(-3\mu+\tau)00(1+\mu\tau)\}$
4	-1	SNR	$\mathsf{NAF}(z_\ell) = 0^\omega(-1)\,(0000(-3+\tau)0)^{(\ell)}\,00(3-\tau)$
	($\ell \geq 0$)		$\mathsf{opt}(z_\ell) = 0^\omega\,(00000(-3+\tau))^{(\ell)}\,001$
			$\mathsf{NAF}(z'_\ell) = 0^\omega\,(00000(-3+\tau))^{(\ell)}\,000(3-\tau)$
			$\mathsf{opt}(z_\ell) = 0^\omega\,(00000(-3+\tau))^{(\ell)}\,000(3-\tau)$
5	-1	MNR	$\mathsf{NAF}(z_\ell) = 0^\omega(1-2\tau)\,(00000(-3-\tau))^{(\ell)}\,0000(1+3\tau)$
	($\ell \geq 1$)		$\mathsf{opt}(z_\ell) = \{0^\omega(1-2\tau)\,(00000(-3-\tau))^{(\ell)}\,0000(1+3\tau)\}$
			$\mathsf{NAF}(z'_\ell) = 0^\omega(-1)\,(0000(-3-\tau)0)^{(\ell)}\,000(1+3\tau)$
			$\mathsf{opt}(z'_\ell) = \{0^\omega\,(00000(1+3\tau))^{(\ell)}\,000(-1)\}$
5	1	MNR	$\mathsf{NAF}(z_\ell) = 0^\omega(-1+2\tau)00\,(00000(3-\tau))^{(\ell)}\,0000(1-3\tau)$
	($\ell \geq 1$)		$\mathsf{opt}(z_\ell) = \{0^\omega(-1+2\tau)00\,(00000(3-\tau))^{(\ell)}\,0000(1-3\tau)\}$
			$\mathsf{NAF}(z'_\ell) = 0^\omega(-1)\,(0000(3-\tau)0)^{(\ell)}\,000(1-3\tau)$
			$\mathsf{opt}(z'_\ell) = \{0^\omega\,(00000(1-3\tau))^{(\ell)}\,000(-1)\}$
5	-1	SNR	$\mathsf{NAF}(z_\ell) = 0^\omega(-1-\tau)\,(0000(-5-4\tau)000000(-5-4\tau))^{(\ell)}$ $0000(-5-4\tau)0000(3+3\tau)$
	($\ell \geq 0$)		$\mathsf{opt}(z_\ell) = \{0^\omega\,(000000(-5-4\tau)0000(-5-4\tau))^{(\ell)}\,000000(-3-3\tau)0001\}$
			$\mathsf{NAF}(z'_\ell) = 0^\omega\,(0000(-5-4\tau)000000(-5-4\tau))^{(\ell)}\,0000(-5-4\tau)0000(3+3\tau)$
			$\mathsf{opt}(z'_\ell) = \{0^\omega\,(0000(-5-4\tau)000000(-5-4\tau))^{(\ell)}$ $0000(-5-4\tau)0000(3+3\tau)\}$
5	1	SNR	$\mathsf{NAF}(z_\ell) = 0^\omega 1\,(000000(5-4\tau)0000(-5+4\tau))^{(\ell)}\,0000(-3+\tau)0000(3-3\tau)$
	($\ell \geq 1$)		$\mathsf{opt}(z_\ell) = \{0^\omega\,(0000000(-5+4\tau)000(-5+4\tau))^{(\ell)}$ $0000000(-5+4\tau)00(3+\tau)\}$
			$\mathsf{NAF}(z'_\ell) = 0^\omega(-1+\tau)\,(0000(5-4\tau)0000(-5+4\tau)00)^{(\ell)}$ $00(-3+\tau)0000(3-3\tau)$
			$\mathsf{opt}(z'_\ell) = \{0^\omega(1-\tau)\,(0000000(-5+4\tau)000(-5+4\tau))^{(\ell)}$ $0000(3-3\tau)\}$
5	-1	$\mathsf{P}\bar{\tau}$	$\mathsf{NAF}(z_\ell) = 0^\omega(1+\tau)\,(00000(5-\tau))^{(\ell)}\,0000(-1-3\tau)$
	($\ell \geq 0$)		$\mathsf{opt}(z_\ell) = \{0^\omega\,(00000(-1-3\tau))^{(\ell_2)}\,000(1+\tau)\,(00000(5-\tau))^{(\ell_1)}$ $0000(-1-3\tau) \mid \ell_1,\ell_2 \geq 0 \text{ and } \ell_1+\ell_2=\ell\}$
			$\mathsf{NAF}(z'_\ell) = 0^\omega(1+\tau)0000(1+\tau)\,(00000(5-\tau))^{(\ell)}\,0000(-1-3\tau)$
			$\mathsf{opt}(z'_\ell) = \{0^\omega\,(00000(-3+7\tau))^{(\ell)}\,00000(-3+7\tau)00(-1+\tau)\}$

TABLE 1. Explicit elements z_ℓ and z'_ℓ for Theorem 1. For $w=4$, $\mu=1$ we have $\mathsf{SNR}(4) = \mathsf{MNR}(4)$. For $w=5$, $\mu=1$, $\mathcal{D} = \mathsf{P}\bar{\tau}(5)$, $\mathsf{opt}(z_\ell)$ is given by a regular expression, where "$\|$" denotes alternatives and * denotes the Kleene star.

w	μ	$\mathcal{D}$	
5	1	$\mathsf{P}\bar{\tau}$	$\mathsf{NAF}(z_\ell) = 0^\omega(-1)\,(000000(-7+5\tau))^{(\ell+1)}\,00000(-3+\tau)0000(-1+3\tau)$
	$(\ell \geq 0)$		$\mathsf{opt}(z_\ell) = \{\boldsymbol{\eta} \in 0^\omega\big(0000000000(5+\tau)00(3-\tau)$
			$\quad \| \, 000000000(3-\tau)000(-5-\tau)$
			$\quad \| \, 000000000000(-3+\tau)(-3-7\tau) \, \| \, 000000(-1+\tau)\big)^*$
			$0000000000000(-3-7\tau)00000(-3-7\tau)(-1)$
			$\mid \mathsf{length}(\boldsymbol{\eta}) = 23+7\ell\}$
			$\mathsf{NAF}(z'_\ell) = 0^\omega\,(000000(-7+5\tau))^{(\ell)}\,00000(-3+\tau)0000(-1+3\tau)$
			$\mathsf{opt}(z'_\ell) = \{0^\omega\,(000000(-7+5\tau))^{(\ell)}\,000000000(3+7\tau)(-3+\tau),$
			$0^\omega\,(000000(-7+5\tau))^{(\ell)}\,00000(-3+\tau)0000(-1+3\tau)\}$
6	-1	MNR	$\mathsf{NAF}(z_\ell) = 0^\omega 100000(1+3\tau)\,(00000(5+3\tau))^{(\ell)}\,00000(3+4\tau)$
	$(\ell \geq 1)$		$\mathsf{opt}(z_\ell) = \{0^\omega(3+4\tau)\,(00000(5+3\tau))^{(\ell)}\,0000(-1-2\tau)\}$
			$\mathsf{NAF}(z'_\ell) = 0^\omega(1+3\tau)\,(00000(5+3\tau))^{(\ell)}\,00000(3+4\tau)$
			$\mathsf{opt}(z'_\ell) = \{0^\omega(1+3\tau)\,(00000(5+3\tau))^{(\ell)}\,00000(3+4\tau)\}$
6	1	MNR	$\mathsf{NAF}(z_\ell) = 0^\omega(1-3\tau)\,(00000(5-3\tau))^{(\ell)}\,00000(3-4\tau)$
	$(\ell \geq 1)$		$\mathsf{opt}(z_\ell) = \{0^\omega(1-3\tau)\,(00000(5-3\tau))^{(\ell)}\,00000(3-4\tau)\}$
			$\mathsf{NAF}(z'_\ell) = 0^\omega 100000(1-3\tau)\,(00000(5-3\tau))^{(\ell)}\,00000(3-4\tau)$
			$\mathsf{opt}(z'_\ell) = \{0^\omega(-3+4\tau)\,(00000(-5+3\tau))^{(\ell)}\,0000(-1+2\tau)\}$
6	-1	SNR	$\mathsf{NAF}(z_\ell) = 0^\omega\,(000000(1-2\tau))^{(\ell)}\,00000(-5-\tau)$
	$(\ell \geq 1)$		$\mathsf{opt}(z_\ell) = \{0^\omega\,(000000(1-2\tau))^{(\ell)}\,00000(-5-\tau)\}$
			$\mathsf{NAF}(z'_\ell) = 0^\omega(-1)(00000(1-2\tau)0)^{(\ell)}0000(-5-\tau)$
			$\mathsf{opt}(z'_\ell) = \{0^\omega\,(000000(-5-4\tau))^{(\ell)}\,00001\}$
6	1	SNR	$\mathsf{NAF}(z_\ell) = 0^\omega(3-\tau)\,(00000009000000000(-9))^{(\ell)}$
	$(\ell \geq 1)$		$000000(1-3\tau)00000(7-\tau)$
			$\mathsf{opt}(z_\ell) = \{0^\omega(3-\tau)\,(00000009000000000(-9))^{(\ell)}$
			$000000(1-3\tau)00000(7-\tau)\}$
			$\mathsf{NAF}(z'_\ell) = 0^\omega 100000(1-3\tau)\,(0000009000000000(-9)0)^{(\ell+1)}$
			$00000(1-3\tau)00000(7-\tau)$
			$\mathsf{opt}(z'_\ell) = \{0^\omega(-9)\,(00000000(9-2\tau)00000000(-9+2\tau))^{(\ell)}$
			$00000000(9-2\tau)000000(3+\tau)0000000(-9+2\tau)000(-1+3\tau)\}$

TABLE 1. Explicit elements z_ℓ and z'_ℓ for Theorem 1 (continued).

4. Computing $\mathcal{D}$-w-NAFs by Transducer Automata

As an auxiliary result, we show that it is possible to compute a $\mathcal{D}$-w-NAF by a transducer automaton. This result is similar to Heuberger and Prodinger [11, Section 3].

Lemma 4.1. *Let $w \geq 1$ and $\mathcal{D}$ be a w-NADS. Then there is a transducer $\mathcal{T}$ on the alphabet $\mathcal{D}$ transforming an arbitrary $\mathcal{D}$-expansion $0^\omega\mathbf{d}$ to $\mathsf{NAF}(\mathsf{value}(0^\omega\mathbf{d}))$ from right to left.*

More precisely, there is a constant c depending on $\mathcal{D}$ such that for any finite words $\mathbf{d}$, $\boldsymbol{\eta}$ over the alphabet $\mathcal{D}$, the words $0^{(c)}\mathbf{d}$ and $\boldsymbol{\eta}$ are the input and output labels of a successful path in $\mathcal{T}$ if and only if $0^\omega\boldsymbol{\eta} = \mathsf{NAF}(\mathsf{value}(0^\omega\mathbf{d}))$.

The number of states for this transducer $\mathcal{T}$ in the case of the digits sets introduced in Section 2 are shown in Table 2.

Proof. We first define a possibly larger transducer $\widetilde{\mathcal{T}}$ and remove unnecessary states and transitions afterwards.

w	μ	$\#V(\mathcal{T})$ for $\mathcal{D}=\mathsf{MNR}(w)$	$\#V(\mathcal{T})$ for $\mathcal{D}=\mathsf{SNR}(w)$	$\#V(\mathcal{T})$ for $\mathcal{D}=\mathsf{P}\bar{\tau}(w)$
2	±1	13	13	13
3	±1	89	89	89
4	±1	575	575	575
5	±1	2469	4609	17051
6	-1	10191	15309	
6	1	10191	21159	

TABLE 2. Size of the transducer automaton constructed in Lemma 4.1

Set $M := \max\{|d| : d \in \mathcal{D}\}$ and

$$C := M\left(\frac{1}{\sqrt{2}-1}+\frac{1}{2^{w/2}-1}\right).$$

The sets $\widetilde{V}$ of states and $\widetilde{E}$ of transitions of $\widetilde{\mathcal{T}}$ are defined to be

$$\widetilde{V} := \left\{(z,\ell) : \ell \in \{0,\dots,w-1\}, z \in \mathbb{Z}[\tau], |z| \le C + M\frac{2^{\ell/2}-1}{\sqrt{2}-1}, \ell = 0 \text{ or } \tau \text{ does not divide } z\right\}$$

and $\widetilde{E} := \widetilde{E}_1 \cup \widetilde{E}_2 \cup \widetilde{E}_3 \cup \widetilde{E}_4$ with

$$\begin{aligned}
\widetilde{E}_1 &:= \{(z,0) \xrightarrow{d|0} ((z+d)/\tau, 0) : (z,0) \in \widetilde{V}, d \in \mathcal{D} \text{ and } \tau \text{ divides } (z+d)\},\\
\widetilde{E}_2 &:= \{(z,0) \xrightarrow{d|\varepsilon} (z+d,1) : (z,0) \in \widetilde{V}, d \in \mathcal{D}, \text{ and } \tau \text{ does not divide } (z+d)\},\\
\widetilde{E}_3 &:= \{(z,\ell) \xrightarrow{d|\varepsilon} (z+d\tau^\ell, \ell+1) : (z,\ell) \in \widetilde{V}, d \in \mathcal{D}, 0 < \ell < w-1\},\\
\widetilde{E}_4 &:= \{(z,w-1) \xrightarrow{d|0^{(w-1)}\eta} ((z+d\tau^{w-1}-\eta)/\tau^w, 0) : (z,w-1) \in \widetilde{V}, d \in \mathcal{D},\\
&\qquad \eta \in \mathcal{D} \text{ with } z + d\tau^{w-1} \equiv \eta \pmod{\tau^w}\}
\end{aligned} \tag{6}$$

respectively. Here, the symbol ε stands for the empty word. The set of initial states and the set of terminal states are both defined to consist of state $(0,0)$ only.

A routine verification shows that for all $(z,\ell) \xrightarrow{d|\boldsymbol{\eta}} (z',\ell') \in \widetilde{E}$, the pair (z',ℓ') is indeed an element of $\widetilde{V}$ and the invariants

$$\begin{aligned}
z + d\tau^\ell &= \mathsf{value}(\boldsymbol{\eta}) + z'\tau^{\mathsf{length}(\boldsymbol{\eta})},\\
\ell + 1 &= \mathsf{length}(\boldsymbol{\eta}) + \ell'
\end{aligned}$$

hold for finite words $\boldsymbol{\eta} \in \mathcal{D}^*$.

By induction, these invariants extend to paths in $\widetilde{\mathcal{T}}$, too: For a path from (z,ℓ) to (z',ℓ') with input and output labels $\mathbf{d}$ and $\boldsymbol{\eta}$, respectively, we have

$$z + \mathsf{value}(\mathbf{d})\tau^\ell = \mathsf{value}(\boldsymbol{\eta}) + z'\tau^{\mathsf{length}(\boldsymbol{\eta})}, \tag{7a}$$

$$\ell + \mathsf{length}(\mathbf{d}) = \mathsf{length}(\boldsymbol{\eta}) + \ell'. \tag{7b}$$

The labels of the path are concatenated from right to left. We note that by construction, $\boldsymbol{\eta}$ satisfies the w-NAF condition.

Consider a successful path (i.e., a path from the unique initial state $(0,0)$ to the unique terminal state $(0,0)$) with input and output labels $\mathbf{d}$ and $\boldsymbol{\eta}$, respectively. Then (7a) simply states that $0^\omega\boldsymbol{\eta} = \mathsf{NAF}(\mathsf{value}(0^\omega\mathbf{d}))$.

We also claim that for each $(z,\ell) \in \widetilde{V}$, there is a path from (z,ℓ) to the terminal state $(0,0)$ whose input label is a word consisting of zeros only. This can be proved by induction on the length[3] of the $\mathcal{D}$-w-NAF $\boldsymbol{\theta}$ of z which has been assumed to exist. The main fact is that transitions in $\widetilde{E}_1$ output the least significant digit of $\boldsymbol{\theta}$, whereas transitions in $\widetilde{E}_4$ output the w least significant digits of $\boldsymbol{\theta}$.

[3]The length in the sense of Definition 2.1.

Thus for any $\mathcal{D}$-expansion $0^\omega\mathbf{d}$ of some $z \in \mathbb{Z}[\tau]$, there is a successful path with input label $0^{(c)}\mathbf{d}$ for a suitable number c of leading zeros. The exact number of leading zeros is irrelevant since there is a transition $(0,0) \xrightarrow{0|0} (0,0)$. The output label is then—up to leading zeros—$\mathsf{NAF}(z)$.

Finally we define $\mathcal{T}$ with sets V and E of states and transitions, respectively, to be the sub-transducer of $\widetilde{\mathcal{T}}$ spanned by the states which are actually reachable from the initial state $(0,0)$. □

5. Computing Optimal $\mathcal{D}$-Expansions

It is conceptually easy to compute an optimal $\mathcal{D}$-expansion of some $y \in \mathbb{Z}[\tau]$ recursively: If τ divides y, the least significant digit of any $\mathcal{D}$-expansion of y equals 0. Otherwise, we consider an optimal expansion (calculated recursively) of $(y-d)/\tau$ for all $d \in \mathcal{D} \setminus \{0\}$ and choose $d \in \mathcal{D}$ such that the Hamming weight of an optimal expansion is minimum. Termination of this procedure can be enforced by pruning all expansions with Hamming weight larger than the Hamming weight of the corresponding $\mathcal{D}$-w-NAF. From the description of the algorithm, we expect it to have exponential running time.

However, this can be improved:

Theorem 2. *Let $w \geq 1$ and $\mathcal{D}$ be a w-NADS. Then there is an algorithm to compute a $\mathcal{D}$-optimal expansion of $y \in \mathbb{Z}[\tau]$ in $O(\log|y|)$ time, where the implicit constant depends on $\mathcal{D}$.*

We remark that the implicit O-constant depends on the size of the transducer $\mathcal{T}$ described in Lemma 4.1, so one cannot expect miracles from this result. However, the idea will be used to prove Theorem 1.

Proof. Let $y \in \mathbb{Z}[\tau]$, $\boldsymbol{\theta} = \mathsf{NAF}(y)$ and $K := \mathsf{length}(\boldsymbol{\theta})$. Note that $K \sim 2\log_2|y|$ by an estimate of Solinas [24, Equation (53)].

We construct a new transducer $\mathcal{T}_y$ from the transducer defined in Lemma 4.1 whose underlying input automaton only accepts $\mathcal{D}$-expansions of y. This can be done by restricting the output of $\mathcal{T}$: we only allow output which agrees with the $\mathcal{D}$-w-NAF $\boldsymbol{\theta}$ of y. This corresponds to the concatenation of $\mathcal{T}$ with the automaton accepting the word $\boldsymbol{\theta}$ only, i.e., the output of $\mathcal{T}$ is used as input to this second automaton. In order to achieve this explicitly, we must manage a pointer describing the number of output digits already verified.

More precisely, we define the transducer $\mathcal{T}_y$ as follows. The set of states V_y is defined to be

$$V_y := \{(z,\ell,k) : (z,\ell) \in V, 0 \leq k \leq K\}.$$

The only initial state is $(0,0,0)$, the only terminal state is $(0,0,K)$.

The set of transitions E_y is defined to be

$$E_y := \big\{(z,\ell,k) \xrightarrow{d|\eta_{m-1}\dots\eta_0} (z',\ell',\min\{k+m,K\}) : m \geq 0, 0 \leq k \leq K, \\ (z,\ell) \xrightarrow{d|\eta_{m-1}\dots\eta_0} (z',\ell') \in E, (\eta_{m-1},\dots,\eta_0) = (\theta_{k+m-1},\dots,\theta_k)\big\}. \quad (8)$$

The crucial invariant in this transducer is the following: There is a path from $(0,0,0)$ to (z,ℓ,k) in $\mathcal{T}_y$ with input and output labels $\mathbf{d}$ and $\boldsymbol{\eta}$, respectively, if and only if there is a path from $(0,0)$ to (z,ℓ) with the same labels in $\mathcal{T}$, $k = \min(K, \mathsf{length}(\boldsymbol{\eta}))$, and $\boldsymbol{\eta}$ is a suffix of $\boldsymbol{\theta}$, i.e., $\eta_j = \theta_j$ for $0 \leq j < \mathsf{length}(\boldsymbol{\eta})$. This can easily be proved by induction on the length of $\mathbf{d}$.

At this point, a comment on the rôle of K seems to be adequate: intuitively, the argument would be simpler if $\min\{k+m,K\}$ in the definition of E_y would be replaced with $k+m$. Then we would have $k = \mathsf{length}(\boldsymbol{\eta})$ in the above invariant. However, this would construct infinitely many states $(z,\ell,K+j)$, $j \geq 0$, which are all equivalent, since θ_{K+j} vanishes anyway. Therefore, the truncation at K has been chosen in order to avoid equivalent states and to obtain a finite transducer.

The above invariant states that there is a successful path in $\mathcal{T}_y$ with input label $0^{(c)}\mathbf{d}$ if and only if $\mathsf{value}(0^\omega\mathbf{d}) = \mathsf{value}(\boldsymbol{\theta}) = y$. Here, c is the constant from Lemma 4.1. In this case, the output label is a suffix of $\boldsymbol{\theta}$.

The cost of a transition is defined as the Hamming weight of its input label. Thus the optimal $\mathcal{D}$-expansions of y are exactly the input labels of shortest successful paths. From (8) and (6) we deduce that if there is a transition from (z, ℓ, k) to (z', ℓ', k') with $k < K$, the pair (k, ℓ) is lexicographically smaller than (k', ℓ'). This implies that the transducer is $(w \cdot K + 1)$-partite with node classes $V_{y,k,\ell} = \{(z, k, \ell) \in V_y\}$, $0 \le k < K$, $0 \le \ell < w$, and $V_{y,K} = \{(z, K, \ell) \in V_y\}$. Thus the shortest paths can be computed with a running time which is linear in K. For instance, a variant of the Ford-Bellman algorithm which processes the edges in lexicographically increasing order of (k, ℓ) for the start node (z, k, ℓ) $(k < K)$ does only need one loop for the transitions starting at a vertex (z, k, ℓ) with $k < K$. Only the final component $V_{y,K}$ (which is independent of y) requires a full shortest path search. □

6. Proof of Theorem 1

We present the details of the proof of Theorem 1 for the case $w = 4$ and $\mathcal{D} = \mathsf{MNR}(4) = \mathsf{P}\bar{\tau}(4)$. All other cases listed in Theorem 1 and Table 1 are proved analogously, cf. the remarks at the end of this section.

We first consider z_ℓ. We construct an auxiliary transducer which is similar to that in the proof of Theorem 2. The difference is that we deal with all values of ℓ simultaneously. So we are not storing a pointer k to the given $\mathcal{D}$-w-NAF, but we store the whole language which is still expected. This corresponds to the concatenation of $\mathcal{T}$ with the automaton accepting the regular language corresponding to (3).

Let $\mathcal{L}$ be the language given by the regular expression $0^*000(\mu-\tau)(000(-3\mu+\tau))^*0000(1-\mu\tau)000(-1)$ over the alphabet $\mathcal{D}$. Here, $(-3\mu+\tau)$ is a literal (as a digit in $\mathcal{D}$), and not an alternation. The Kleene star (finite repetition) is denoted by $(\dots)^*$, as usual. Obviously, the language $\mathcal{L}$ has been chosen to correspond with the $\mathcal{D}$-w-NAF of z_ℓ given in (3). We set

$$\begin{aligned}\mathcal{M} := \{&0^*000(\mu-\tau)(000(-3\mu+\tau))^*0000(1-\mu\tau)000(-1),\\ &0^*000(\mu-\tau)(000(-3\mu+\tau))^*0000(1-\mu\tau),\\ &0^*000(\mu-\tau)(000(-3\mu+\tau))^*0,\\ &0^*000(\mu-\tau)(000(-3\mu+\tau))^*,\\ &0^*\}.\end{aligned}$$

The reason for this choice is that output labels of transitions in $\mathcal{T}$ are either empty words, a single 0, or w-digit words with $(w-1)$ leading zeros.

The auxiliary transducer $\mathcal{T}_\mathcal{L}$ is defined by its set of states $V_\mathcal{L}$ with

$$V_\mathcal{L} := \{(z, \ell, M) : (z, \ell) \in \mathcal{T}, M \in \mathcal{M}\},$$

its set of transitions

$$\begin{aligned}E_\mathcal{L} := \big\{(z,\ell,M) \xrightarrow{d|\eta_{m-1}\dots\eta_0} (z',\ell',M') : m \ge 0, M, M' \in \mathcal{M},\\ (z,\ell) \xrightarrow{d|\eta_{m-1}\dots\eta_0} (z',\ell') \in E, M'\eta_{m-1}\dots\eta_0 \subseteq M\big\},\end{aligned}$$

its unique initial state $(0, 0, \mathcal{L})$ and its unique terminal state $(0, 0, 0^*)$.

The following invariant holds: There is a path from the initial state $(0, 0, \mathcal{L})$ to a state (z, ℓ, M) with input and output labels $\mathbf{d}$ and $\boldsymbol{\eta}$, respectively, if and only if there is path from $(0, 0)$ to (z, ℓ) with the same labels such that $M\boldsymbol{\eta} \subseteq \mathcal{L}$.

Thus there is a successful path in $\mathcal{T}_\mathcal{L}$ with input and output labels $0^{(c)}\mathbf{d}$ and $\boldsymbol{\eta}$ if and only if $\mathbf{d}$ is a $\mathcal{D}$-expansion of some z_ℓ which is given by its $\mathcal{D}$-w-NAF $\boldsymbol{\eta}$.

We computed the transducer $\mathcal{T}$ for $\mu = -1$ and for $\mu = 1$ separately; in both cases, there are 2003 states reachable from the initial state. From 608 of those, the terminal state is reachable.

We intend to compute shortest paths in $\mathcal{T}_\mathcal{L}$. In contrast to Theorem 2, we cannot use the same cost function, since this would mask out z_ℓ for $\ell > 0$. Therefore, we define the cost of a transition to be the Hamming weight of its input label minus the Hamming weight of its output label. Using the Ford-Bellman algorithm shows that the shortest path from the initial state to the

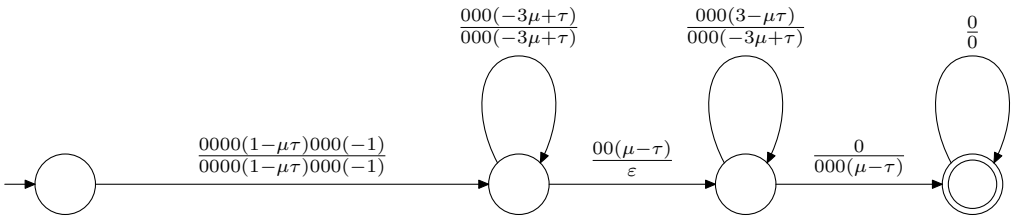

FIGURE 1. Transducer transforming all optimal $\mathcal{D}$-expansions of z_ℓ to its $\mathcal{D}$-w-NAF, where $w = 4$ and $\mathcal{D}$ is the set of minimal norm representatives modulo τ^w.

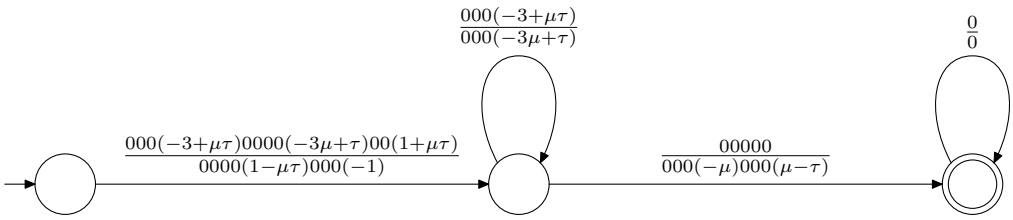

FIGURE 2. Transducer transforming all optimal $\mathcal{D}$-expansions of z'_ℓ to its $\mathcal{D}$-w-NAF, where $w = 4$ and $\mathcal{D}$ is the set of minimal norm representatives modulo τ^w.

terminal state has cost 0 (for both choices of μ). This means that there is no $\ell \geq 0$ such that z_ℓ admits a $\mathcal{D}$-expansion of Hamming weight less than the Hamming weight of its $\mathcal{D}$-w-NAF. This immediately shows that a $\mathcal{D}$-expansion of some z_ℓ is an optimal $\mathcal{D}$-expansion if and only if it has the same Hamming weight as the corresponding $\mathcal{D}$-w-NAF. This is the case if and only if there is a successful path in $\mathcal{T}_\mathcal{L}$ with input label $0^{(c)}\mathbf{d}$ of total cost 0.

Furthermore, the Ford-Bellman algorithm also yields the vertex potentials $\pi(s)$, $s \in V_\mathcal{L}$, defined to be the shortest distance between the initial state and s. We call a transition $s \xrightarrow{\mathbf{d}|\boldsymbol{\eta}} s'$ optimal if $\pi(s') = \pi(s) + \mathsf{weight}(\mathbf{d}) - \mathsf{weight}(\boldsymbol{\eta})$. A successful path is a shortest path in $\mathcal{T}_\mathcal{L}$ if and only if it only uses optimal transitions. We now drop all non-optimal transitions of $\mathcal{T}_\mathcal{L}$. Then, we consider the sub-transducer spanned by the states from which the terminal state is reachable. The result is shown in Figure 1, where independent paths have been contracted to transitions for graphical reasons. In fact, the results for $\mu = \pm 1$ are very similar and can be uniformly shown in the same figure. Thus the transducer in Figure 1 transforms all optimal $\mathcal{D}$-expansions of some z_ℓ to its $\mathcal{D}$-w-NAF. It is now easy to read off the optimal $\mathcal{D}$-expansions. These are indeed shown in (4).

Now, we turn to z'_ℓ. We use the same ideas as for z_ℓ. The corresponding regular language is denoted by $\mathcal{L}'$. The corresponding transducer $\mathcal{T}_{\mathcal{L}'}$ has 2495 states reachable from the initial state, from 855 of those, the terminal state is reachable. Using the Ford-Bellman algorithm shows that the shortest path in $\mathcal{T}_{\mathcal{L}'}$ from the initial to the terminal state has length -1. Thus there is a nonnegative ℓ such that z'_ℓ admits a $\mathcal{D}$-expansion whose Hamming weight is the Hamming weight of the corresponding $\mathcal{D}$-w-NAF decreased by 1. Since the $\mathcal{D}$-w-NAF has Hamming weight $\ell + 4$, the minimum Hamming weight of a $\mathcal{D}$-expansion of z_ℓ is at least $\ell + 3$.

An expansion $\mathbf{d}$ of z_ℓ has Hamming weight $\ell + 3$ if and only if the corresponding successful path with input label $\mathbf{d}$ in $\mathcal{T}_{\mathcal{L}'}$ has total cost -1. Thus we apply the above reduction process again. This yields the transducer in Figure 2. From this transducer, we immediately see that there is indeed a $\mathcal{D}$-expansion of z_ℓ of Hamming weight $\ell + 3$ for every nonnegative integer ℓ. Obviously, there is only one such expansion, and this expansion is shown in (5). This concludes the proof of the Theorem for $w = 4$.

As stated in the introductory remarks of this section, the proof of Theorem 1 for the other values of w and $\mathcal{D}$ is analogous. The number of states of the various transducers is given in Table 3. The computation of $\mathsf{opt}(z'_\ell)$ in the largest case ($w = 6$, $\mu = 1$, $\mathcal{D} = \mathsf{SNR}(6)$) took 65

w	μ	$\mathcal{D}$	$\#V(\mathcal{T})$	$\#V(\mathcal{T}_\mathcal{L})$	$V(\mathcal{T}^*_\mathcal{L})$	$V(\mathcal{T}_{\mathcal{L}'})$	$V(\mathcal{T}^*_{\mathcal{L}'})$
4	-1	$\mathsf{MNR} = \mathsf{P}\bar{\tau}$	575	2003	608	2495	855
4	-1	SNR	575	3465	1076	3465	767
4	1	$\mathsf{MNR} = \mathsf{SNR} = \mathsf{P}\bar{\tau}$	575	2003	608	2495	855
5	-1	MNR	2469	7841	1630	7770	1753
5	1	MNR	2469	12275	2031	7770	1753
5	-1	SNR	4609	22990	4555	23177	3783
5	1	SNR	4609	26570	5584	26581	5844
5	-1	$\mathsf{P}\bar{\tau}$	17051	51647	7616	67470	12596
5	1	$\mathsf{P}\bar{\tau}$	17051	79914	13142	80075	12162
6	-1	MNR	10191	34964	5531	25145	3657
6	1	MNR	10191	25145	3657	34964	5531
6	-1	SNR	15309	52344	4600	52203	6904
6	1	SNR	21159	214609	19278	235138	23785

TABLE 3. Number of states in the transducers used in the proof of Theorem 1

days on a Intel® Core™ 2 Duo CPU E6850 at 3.00 GHz running Mathematica® 5.2 under Linux 2.6.22. In most cases, the transducers describing optimal expansions were of the same shape as the transducer in Figure 2. In some cases, however, the transducers were slightly more complex. As representative examples, we show the transducers for $\mathsf{opt}(z'_\ell)$ for $(w, \mu, \mathcal{D}) = (6, 1, \mathsf{SNR}(6))$ in Figure 3 and for $\mathsf{opt}(z_\ell)$ for $(w, \mu, \mathcal{D}) = (5, 1, \mathsf{P}\bar{\tau}(5))$ in Figure 4. The latter case is the one leading to the regular expression in Table 1.

7. Joint τ-Expansions

In this second part of the paper, we turn our attention to joint τ-expansions of pairs of integers in $\mathbb{Z}[\tau]$.

Definition 7.1. A *joint expansion* of $\mathbf{z} \in \mathbb{Z}[\tau]^2$ is a left infinite word $\mathbf{H} = \dots \boldsymbol{\eta}_2\boldsymbol{\eta}_1\boldsymbol{\eta}_0$ over the alphabet $\{0, 1, -1\}^2$ such that

1. only a finite number of the digit vectors $\boldsymbol{\eta}_j$ is nonzero,
2. $\mathsf{value}(\mathbf{H}) := \sum_{j\ge 0} \boldsymbol{\eta}_j \tau^j = \mathbf{z}$, i.e., $\mathbf{H}$ is indeed an expansion of $\mathbf{z}$.

The digit vectors $\boldsymbol{\eta}_j$, $j \ge 0$ will also called the *columns* of $\mathbf{H}$.

The *joint Hamming weight* $\mathsf{weight}(\mathbf{H})$ of $\mathbf{H}$ is the number of nonzero digit vectors $\boldsymbol{\eta}_j$.

A joint expansion $\mathbf{H}$ of $\mathbf{z}$ is called an *optimal joint expansion* of $\mathbf{z}$ if its Hamming weight is minimum amongst all joint expansions of $\mathbf{z}$.

We again need a "canonical" joint expansion as point of reference for the transducer automata. We choose the generalisation of Solinas' [25] Joint Sparse Form (JSF) to base τ as proposed by Ciet, Lange, Sica and Quisquater [6].

Definition 7.2. An expansion $\mathbf{H}$ is called a τ-JSF if it fulfils the following conditions:

1. Among three consecutive columns, at least one is a zero column.
2. For all $j \ge 0$ and $i \in \{1, 2\}$, we have $\eta_{i,j+1} \cdot \eta_{i,j} \ne \mu$.
3. If $\eta_{i,j+1} \cdot \eta_{i,j} \ne 0$ for some $j \ge 0$ and some $i \in \{1, 2\}$, then $\eta_{i',j+1} \in \{\pm 1\}$ and $\eta_{i',j} = 0$, where $i' = 3 - i$ is the other row index.

Here, the components of the jth column $\boldsymbol{\eta}_j$ are denoted by $\binom{\eta_{1j}}{\eta_{2j}}$.

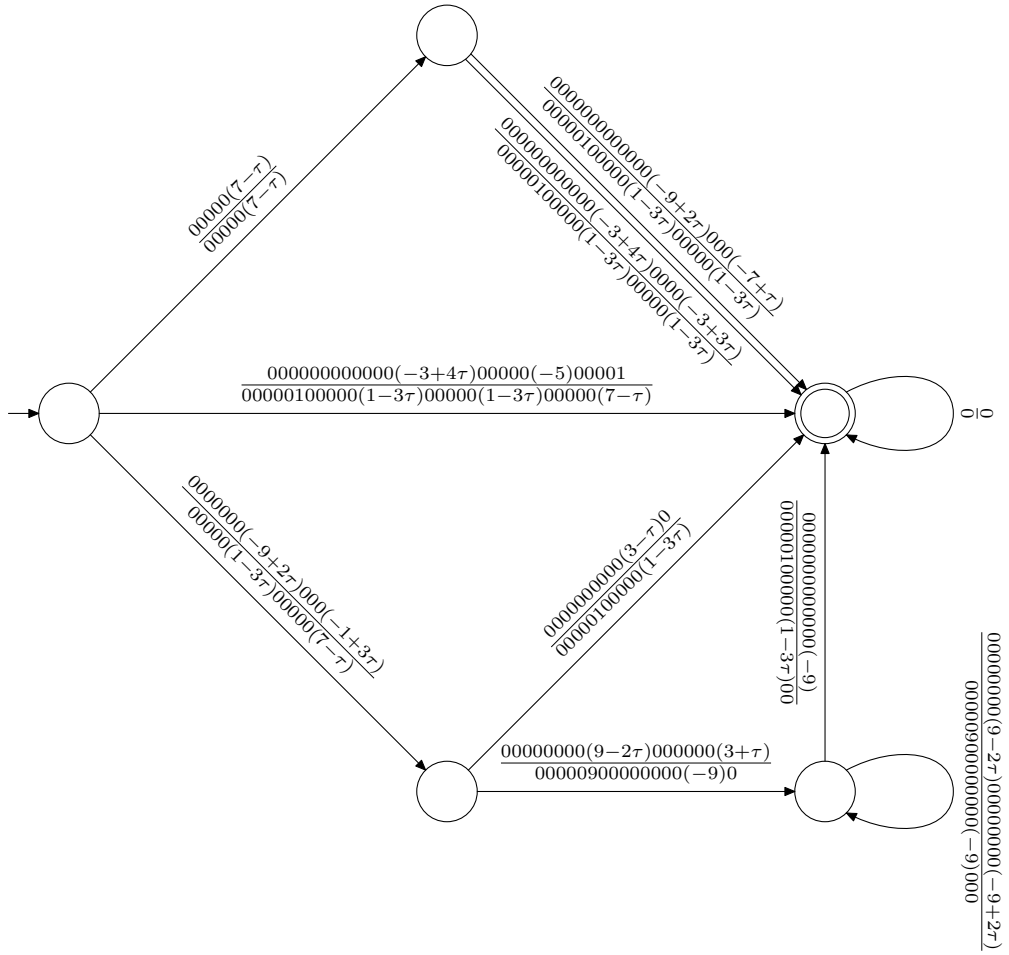

FIGURE 3. Transducer describing $\mathsf{opt}(z'_\ell)$ in the case $w = 6$, $\mu = 1$, $\mathcal{D} = \mathsf{SNR}(6)$.

As stated in [6], every $\mathbf{z} \in \mathbb{Z}[\tau]^2$ admits exactly one τ-JSF. The proof of this fact is promised to appear in the journal version of [6], which is not yet available at the time of this writing. However, it can be proved independently.

There is a transducer automaton translating joint expansions to the τ-JSF representing the same integer vector. This transducer has 289 states whence it is not shown here.

Theorem 3. *For every nonnegative integer ℓ, we consider the integer vectors $\mathbf{z}_\ell$, $\mathbf{z}'_\ell \in \mathbb{Z}[\tau]^2$ given by*

$$\mathbf{z}_\ell := \mathsf{value}\left(\begin{pmatrix}0\\0\end{pmatrix}^{\omega} 0 \begin{pmatrix}0&0&0&0&0&0\\0&\mu&0&0&\bar{1}&0\end{pmatrix}^{(\ell)} \begin{matrix}0&0&1&0&0\\0&\mu&0&0&\bar{1}\end{matrix}\right),$$

$$\mathbf{z}'_\ell := \mathsf{value}\left(\begin{pmatrix}0\\0\end{pmatrix}^{\omega} \begin{pmatrix}0&0&0&0&0&0\\0&0&\mu&0&0&\bar{1}\end{pmatrix}^{(\ell)} \begin{matrix}0&0&0&1&0&0\\0&0&\mu&0&0&\bar{1}\end{matrix}\right).$$

We have

$$\mathbf{z}_\ell - \mathbf{z}'_\ell = \begin{pmatrix}0\\(6-\mu\tau)\tau^{6\ell}\end{pmatrix},$$

in particular, $\mathbf{z}_\ell \equiv \mathbf{z}'_\ell \pmod{\tau^{6\ell}}$.

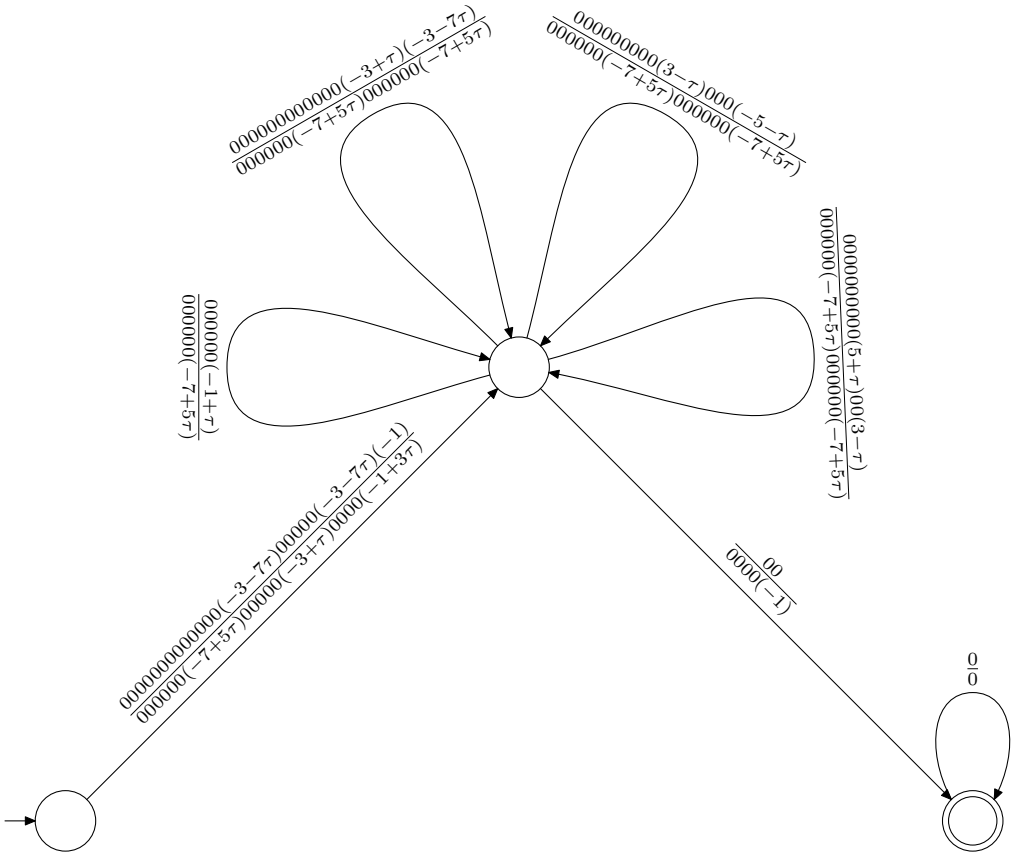

FIGURE 4. Transducer describing $\mathsf{opt}(z_\ell)$ in the case $w = 5$, $\mu = 1$, $\mathcal{D} = \mathsf{P}\bar{\tau}(5)$.

The only optimal joint expansions of $\mathbf{z}_\ell$ *and* $\mathbf{z}'_\ell$ *are given by*

$$\begin{pmatrix}0\\0\end{pmatrix}^{\omega}\begin{pmatrix}0&0&0&0&0&0\\0&0&\bar{1}&0&0&\mu\end{pmatrix}^{(\ell)}\begin{matrix}0&0&1&0&0\\0&0&\bar{1}&\bar{\mu}&1\end{matrix}$$

and

$$\begin{pmatrix}0\\0\end{pmatrix}^{\omega}\begin{pmatrix}0&0&0&0&0&0\\0&0&\mu&0&0&\bar{1}\end{pmatrix}^{(\ell)}\begin{matrix}0&0&0&1&0&0\\0&0&\mu&0&0&\bar{1}\end{matrix},$$

respectively. In particular, the least significant digit vector of the optimal joint expansions of $\mathbf{z}_\ell$ *and* $\mathbf{z}'_\ell$ *differ.*

Proof. The proof runs along the same lines as the proof of Theorem 1. The transducers $\mathcal{T}_{\mathcal{L}}$ and $\mathcal{T}_{\mathcal{L}'}$ have 1048 states in both cases for μ, the reduced transducers $\mathcal{T}^*_{\mathcal{L}}$ and $\mathcal{T}^*_{\mathcal{L}'}$ have 225 and 197 states, respectively. The transducers describing the optimal expansions of $\mathbf{z}'$ and $\mathbf{z}'_\ell$ are given in Figures 5 and 6, respectively. □

Corollary 7.3. *It is impossible to compute an optimal joint expansion of a digit vector* $\mathbf{z}$ *by a deterministic transducer automaton or an online algorithm from right to left.*

Remark 7.4. This result can immediately be generalised to higher dimensions $d > 2$ by filling up the additional rows by zeros.

Remark 7.5. Instead of using the τ-JSF, one could also use a generalisation of the Simple Joint Sparse Form (SJSF) proposed by Grabner, Heuberger and Prodinger [8] to the base of τ. The only difficulty is that the τ-SJSF of the two vectors $\mathbf{z}_\ell$ and $\mathbf{z}'_\ell$ given in Theorem 3 have a Hamming

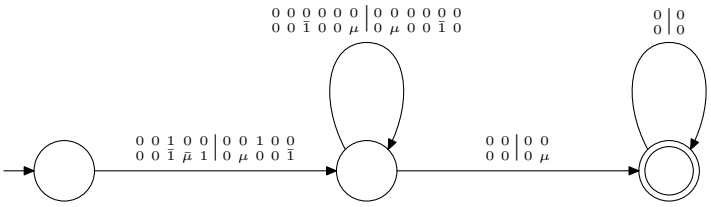

FIGURE 5. Transducer transforming all optimal joint τ-expansions of $\mathbf{z}_\ell$ to its τ-JSF.

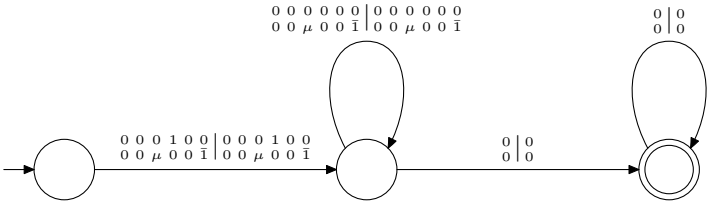

FIGURE 6. Transducer transforming all optimal joint τ-expansions of $\mathbf{z}'_\ell$ to its τ-JSF.

weight which exceeds that of their optimal expansion by an amount which is linear in ℓ. This complicates the argument in the proof somewhat and motivated our decision to take the τ-JSF as "standard-representation" in this case.

References

1. R. Avanzi, *A Note on the Signed Sliding Window Integer Recoding and a Left-to-Right Analogue*, Selected Areas in Cryptography: 11th International Workshop, SAC 2004, Waterloo, Canada, August 9-10, 2004, Revised Selected Papers, Lecture Notes in Comput. Sci., vol. 3357, Springer-Verlag, Berlin, 2004, pp. 130–143.
2. R. M. Avanzi, C. Heuberger, and H. Prodinger, *Redundant τ-adic expansions I: Non-adjacent digit sets and their applications to scalar multiplication*, Preprint.
3. ———, *Minimality of the Hamming weight of the τ-NAF for Koblitz curves and improved combination with point halving*, Selected Areas in Cryptography: 12th International Workshop, SAC 2005, Kingston, ON, Canada, August 11–12, 2005, Revised Selected Papers (Preneel B. and Tavares St., eds.), Lecture Notes in Comput. Sci., vol. 3897, Springer, Berlin, 2006, pp. 332–344.
4. ———, *Scalar multiplication on Koblitz curves. Using the Frobenius endomorphism and its combination with point halving: Extensions and mathematical analysis*, Algorithmica **46** (2006), 249–270.
5. ———, *On redundant τ-adic expansions and non-adjacent digit sets*, Selected Areas in Cryptography: 13th International Workshop, SAC 2006, Montreal, Canada, August 2006, Revised Selected Papers (E. Biham and A. Youssef, eds.), Lecture Notes in Comput. Sci., vol. 4356, Springer, Berlin, 2007, pp. 285–301.
6. M. Ciet, T. Lange, F. Sica, and J.-J. Quisquater, *Improved algorithms for efficient arithmetic on elliptic curves using fast endomorphisms*, Advances in cryptology — EUROCRYPT 2003. International conference on the theory and applications of cryptographic techniques, Warsaw, Poland, May 4–8, 2003. Proceedings (E. Biham, ed.), Lecture Notes in Comput. Sci., vol. 2656, Springer, Berlin, 2003, pp. 388–400.
7. D. M. Gordon, *A survey of fast exponentiation methods*, J. Algorithms **27** (1998), 129–146.
8. P. J. Grabner, C. Heuberger, and H. Prodinger, *Distribution results for low-weight binary representations for pairs of integers*, Theoret. Comput. Sci. **319** (2004), 307–331.

9. C. Heuberger, *Minimal redundant digit expansions in the Gaussian integers*, J. Théor. Nombres Bordeaux **14** (2002), 517–528.
10. C. Heuberger and J. Muir, *Minimal weight and colexicographically minimal integer representations*, J. Math. Cryptol. **1** (2007), 297–328.
11. C. Heuberger and H. Prodinger, *Analysis of alternative digit sets for nonadjacent representations*, Monatsh. Math. **147** (2006), 219–248.
12. I. Kátai and B. Kovács, *Canonical number systems in imaginary quadratic fields*, Acta Math. Hungar. **37** (1981), 159–164.
13. I. Kátai and J. Szabó, *Canonical number systems for complex integers*, Acta Sci. Math. (Szeged) **37** (1975), 255–260.
14. D. E. Knuth, *Seminumerical algorithms*, third ed., The Art of Computer Programming, vol. 2, Addison-Wesley, 1998.
15. N. Koblitz, *CM-curves with good cryptographic properties*, Advances in cryptology—CRYPTO '91 (Santa Barbara, CA, 1991), Lecture Notes in Comput. Sci., vol. 576, Springer, Berlin, 1992, pp. 279–287.
16. M. Lothaire, *Algebraic combinatorics on words*, Encyclopedia of Mathematics and its Applications, vol. 90, Cambridge University Press, Cambridge, 2002.
17. F. Morain and J. Olivos, *Speeding up the computations on an elliptic curve using addition-subtraction chains*, RAIRO Inform. Théor. Appl. **24** (1990), 531–543.
18. J. A. Muir and D. R. Stinson, *New minimal weight representations for left-to-right window methods*, Topics in Cryptology — CT-RSA 2005 The Cryptographers' Track at the RSA Conference 2005, San Francisco, CA, USA, February 14–18, 2005, Proceedings (A. J. Menezes, ed.), Lecture Notes in Comput. Sci., vol. 3376, Springer, Berlin, 2005, pp. 366–384.
19. ———, *Minimality and other properties of the width-w nonadjacent form*, Math. Comp. **75** (2006), 369–384.
20. B. Phillips and N. Burgess, *Minimal weight digit set conversions*, IEEE Trans. Comput. **53** (2004), 666–677.
21. J. Proos, *Joint sparse forms and generating zero columns when combing*, Tech. Report CORR 2003-23, Centre for Applied Cryptographic Research, 2003, available at `http://www.cacr.math.uwaterloo.ca/techreports/2003/`.
22. G. W. Reitwiesner, *Binary arithmetic*, Advances in computers, vol. 1, Academic Press, New York, 1960, pp. 231–308.
23. J. A. Solinas, *An improved algorithm for arithmetic on a family of elliptic curves*, Advances in Cryptology — CRYPTO '97. 17th annual international cryptology conference. Santa Barbara, CA, USA. August 17–21, 1997. Proceedings (B. S. Kaliski, jun., ed.), Lecture Notes in Comput. Sci., vol. 1294, Springer, Berlin, 1997, pp. 357–371.
24. ———, *Efficient arithmetic on Koblitz curves*, Des. Codes Cryptogr. **19** (2000), 195–249.
25. ———, *Low-weight binary representations for pairs of integers*, Tech. Report CORR 2001-41, Centre for Applied Cryptographic Research, University of Waterloo, 2001, available at `http://www.cacr.math.uwaterloo.ca/techreports/2001/corr2001-41.ps`.
26. E. Straus, *Addition chains of vectors (problem 5125)*, American Mathematical Monthly **71** (1964), 806–808.
27. YueFei Zhu, BaiJie Kuang, and YaJuan Zhang, *An improved algorithm for up + vq on a family of elliptic curves*, 19th IEEE International Parallel and Distributed Processing Symposium (IPDPS'05) — Workshop 17, IEEE Computer Society, Los Alamitos, CA, USA, 2005, p. 294.

Clemens Heuberger
Institut für Mathematik B
Technische Universität Graz
Austria
e-mail: `clemens.heuberger@tugraz.at`