JOURNAL de Théorie des Nombres de BORDEAUX

anciennement Séminaire de Théorie des Nombres de Bordeaux

Clemens HEUBERGER et Daniel KRENN

Optimality of the Width-w Non-adjacent Form: General Characterisation and the Case of Imaginary Quadratic Bases

Tome 25, n° 2 (2013), p. 353-386.

<http://jtnb.cedram.org/item?id=JTNB_2013__25_2_353_0>

Journal de Théorie des Nombres
de Bordeaux **25** (2013), 353–386

Optimality of the Width-w Non-adjacent Form: General Characterisation and the Case of Imaginary Quadratic Bases

par Clemens HEUBERGER et Daniel KRENN

Résumé. On étudie des systèmes de numération $\sum_{j=0}^{\ell-1} \Phi^j(D_j)$ avec un endomorphisme Φ d'un groupe abélien. On démontre que dans un tel système la condition w-NAF (chaque bloc de w chiffres consécutifs contient au plus un chiffre non nul) minimise le poids de Hamming par rapport à toutes les représentations avec le même ensemble de chiffres si et seulement si la condition de sous-additivité est vérifiée (la somme de deux représentations de poids 1 admet une représentation optimale w-NAF).

Ce résultat est ensuite appliqué sur les bases entières quadratiques complexes, qui sont utilisées pour la multiplication par un scalaire dans la cryptographie sur les courbes elliptiques. On présente un critère algorithmique et des réponses génériques pour des cas différents. Des entiers quadratiques complexes de trace au moins 3 (en valeur absolue) admettent une w-NAF optimale pour $w \geq 4$. Il en va de même pour le cas particulier de la base $(\pm 3 \pm \sqrt{-3})/2$ (quatre cas) et $w \geq 2$, qui correspond à des courbes de Koblitz en caractéristique trois. Dans le cas de $\tau = \pm 1 \pm i$ (quatre cas), l'optimalité dépend de la parité de w. Des résultats numériques pour des traces plus petites sont présentés.

Abstract. We consider digit expansions $\sum_{j=0}^{\ell-1} \Phi^j(d_j)$ with an endomorphism Φ of an Abelian group. In such a numeral system, the w-NAF condition (each block of w consecutive digits contains at most one nonzero) is shown to minimise the Hamming weight over all expansions with the same digit set if and only if it fulfills the subadditivity condition (the sum of every two expansions of weight 1 admits an optimal w-NAF).

This result is then applied to imaginary quadratic bases, which are used for scalar multiplication in elliptic curve cryptography.

The authors are supported by the Austrian Science Fund (FWF): W1230, Doctoral Program "Discrete Mathematics".

Mots clefs. τ-adic expansions, width-w non-adjacent forms, redundant digit sets, elliptic curve cryptography, Koblitz curves, Frobenius endomorphism, scalar multiplication, Hamming weight, optimality, imaginary quadratic bases.

Classification math. 11A63, 94A60.

Both an algorithmic criterion and generic answers for various cases are given. Imaginary quadratic integers of trace at least 3 (in absolute value) have optimal w-NAFs for $w \geq 4$. The same holds for the special case of base $(\pm 3 \pm \sqrt{-3})/2$ (four cases) and $w \geq 2$, which corresponds to Koblitz curves in characteristic three. In the case of $\tau = \pm 1 \pm i$ (again four cases), optimality depends on the parity of w. Computational results for small trace are given.

1. Introduction

Let τ be an imaginary quadratic algebraic integer. We consider τ-adic (multi-)expansions for an element of $\mathbb{Z}[\tau]$ using a redundant digit set, i.e. our expansions need not be unique without any further constraints. The question that arises is how to find "good" representations. This problem comes from elliptic curve cryptography, where one is interested in expansions leading to efficient calculation schemes.

A scalar multiplication, one of the key operations in elliptic curve cryptosystems, can be carried out by a double-and-algorithm or by using the Frobenius endomorphism ("Frobenius-and-add"), cf. Koblitz [16] and Solinas [25, 26]: The Frobenius endomorphism φ fulfils a quadratic equation $\varphi^2 - p\varphi + q = 0$ for integers p and q depending on the curve. We identify φ with the complex root τ of the same equation. We represent a scalar n as $n = \sum_{j=0}^{\ell} \eta_j \tau^j$ for η_j from some suitable digit set $\mathcal{D}$. Then the scalar multiplication nP for some P on the curve can be computed as $nP = \sum_{j=0}^{\ell} \eta_j \varphi^j(P)$. The latter sum can be efficiently computed using Horner's scheme, where the ηP for $\eta \in \mathcal{D}$ have to be precomputed. The number of applications of φ (which is computationally cheap) corresponds to the length of the expansion, the number of additions corresponds to the weight of the expansion, i.e. the number of nonzero digits η_j.

Some of the results of this article do not depend on the setting in an imaginary quadratic number field. Those are valid in the following more general (abstract) setting and may be used in other situations as well.

A general numeral system is an Abelian group $\mathcal{A}$ together with a group endomorphism Φ and a digit set $\mathcal{D}$, which is a finite subset of $\mathcal{A}$ including 0. The endomorphism acts as base in our numeral system. A common choice is multiplication by a fixed element. In this general setting we consider *multi-expansions*, which are simply finite sums with summands $\Phi^k(d)$, where $k \in \mathbb{N}_0$ and d a nonzero digit in $\mathcal{D}$. The "multi" in the expression "multi-expansion" means that we allow several summands with the same k. If the k are pairwise distinct, we call the sum an *expansion*. Note that in the context of the Frobenius-and-add algorithm, multi-expansions are as good as expansions, as long as the weight is low.

A special expansion is the *width-w non-adjacent form*, or *w-NAF* for short. It will be the main concept throughout this article. In a w-NAF-expansion, the k in different summands differ at least by a fixed positive rational integer w, i.e. considering an expansion as a sequence $(\eta_j)_{j\in\mathbb{N}_0} \in \mathcal{D}^{\mathbb{N}_0}$ (or alternatively as finite word over the alphabet $\mathcal{D}$), each block $\eta_{j+w-1}\dots\eta_j$ of width w contains at most one nonzero digit. The term "non-adjacent form" at least goes back to Reitwiesner [24]. More details and precise definitions of those terms can be found in Section 2.

Obviously, a w-NAF has low weight and therefore leads to quite efficient scalar multiplication in the Frobenius-and-add method. The main question investigated in this article is: Does the w-NAF minimise the weight, i.e. the number of nonzero digits, among all possible representations (multi-expansions) with the same digit set? If the answer is affirmative, we call the w-NAF-expansion *optimal*.

We give conditions equivalent to optimality in Section 3 in the setting of general numeral systems. We show that each group element has an optimal w-NAF-expansion if and only if the digit set is "w-subadditive", which means that each multi-expansion with two summands has a w-NAF-expansion with weight at most 2. This condition can be verified algorithmically, since there are only finitely many nontrivial cases to check. More precisely, one has to consider the w-NAFs corresponding to $w(\#\mathcal{D}-1)^2$ multi-expansions. Another way to verify w-subadditivity is to use the geometry of the digit set. This is done in our imaginary quadratic setting, see below.

Now consider some special cases of numeral systems, where optimality or non-optimality of the non-adjacent form is already known. Here, multiplication by a base element is chosen as endomorphism Φ. In the case of 2-NAFs with digit set $\{-1,0,1\}$ and base 2, optimality is known, cf. Reitwiesner [24]. This was reproved in Jedwab and Mitchell [14] and in Gordon [9]. That result was generalised in Avanzi [4], Muir and Stinson [22] and in Phillips and Burgess [23]. There, the optimality of the w-NAFs with base 2 was shown. As digit set, zero and all odd numbers with absolute value less than 2^{w-1} were used. In this setting, there is also another optimal expansion, cf. Muir and Stinson [21]. Using base 2 and a digit set $\{0,1,x\}$ with $x\in\mathbb{Z}$, optimality of the 2-NAFs is answered in Heuberger and Prodinger [12]. Some of these results will be reproved and extended to arbitrary rational integer bases with our tools in Section 4. That proof will show the main idea how to use the geometry of the digit set to show w-subadditivity and therefore optimality.

We come back to our imaginary quadratic setting, so suppose that the imaginary quadratic base τ is a solution of $\tau^2-p\tau+q=0$, where p and q are rational integers with $q>p^2/4$. Here, $\mathbb{Z}[\tau]$ plays the rôle of the

group and multiplication by τ is taken as the endomorphism. We suppose that the digit set consists of 0 and one representative of minimal norm of every residue class modulo τ^w, which is not divisible by τ, and we call it a *minimal norm representatives digit set*. It can be shown that, taking such a digit set, every element of $\mathbb{Z}[\tau]$ admits a unique w-NAF, cf. Blake, Murty and Xu [7, 6, 5], Koblitz [17] and Solinas [25, 26] for some special cases or Heuberger and Krenn [11] for a general result. All those definitions and basics can be found in Section 6 in a precise formulation.

First, consider the cases $|p| = 1$ and $q = 2$, which comes from a Koblitz curve in characteristic 2, cf. Koblitz [16], Meier and Staffelbach [20], and Solinas [25, 26]. There optimality of the w-NAFs can be shown for $w \in \{2, 3\}$, cf. Avanzi, Heuberger and Prodinger [1, 2]. The case $w = 2$ can also be found in Gordon [9]. For the cases $w \in \{4, 5, 6\}$, non-optimality was shown, see Heuberger [10].

In the present paper we give a general result on the optimality of the w-NAFs with imaginary quadratic bases, namely when $|p| \geq 3$, as well as some results for special cases. So let $|p| \geq 3$. If $w \geq 4$, then optimality of the w-NAFs could be shown in all cases. If we restrict to $|p| \geq 5$, then the w-NAFs are already optimal for $w \geq 3$. Further, we give a condition — p and q have to fulfil a special inequality —, when 2-NAFs are optimal. All those results can be found in Section 7. There we show that the digit set in those cases is w-subadditive by using its geometry.

In the last four sections some special cases are examined. Important ones are the cases $|p| = 3$ and $q = 3$ coming from Koblitz curves in characteristic 3. In Kröll [18] optimality of the w-NAFs was shown for $w \in \{2, 3, 4, 5, 6, 7\}$ by using a transducer and some heavy symbolic computations. In this article we prove that the w-NAF-expansions are optimal for all $w \geq 2$, see Section 8. In Section 9 we look at the cases $|p| = 2$ and $q = 2$. There the w-NAF-expansions are optimal if and only if w is odd. In the cases $p = 0$ and $q \geq 2$, see Section 10, non-optimality of the w-NAFs with odd w could be shown.

2. Expansions and Numeral Systems

This section contains the abstract definition of numeral systems and the definition of expansions. Further, we specify the width-w non-adjacent form and notions related to it.

Abstract numeral systems can be found in van de Woestijne [28], which are generalisations of the numeral systems used, for example, in Germán and Kovács [8]. We use that concept to define w-NAF-numeral systems.

Definition 2.1. A *pre-numeral system* is a triple $(\mathcal{A}, \Phi, \mathcal{D})$ where $\mathcal{A}$ is an Abelian group, Φ an endomorphism of $\mathcal{A}$ and the *digit set* $\mathcal{D}$ is a subset of $\mathcal{A}$ such that $0 \in \mathcal{D}$ and each nonzero digit is not in the image of Φ.

Note that we can assume Φ is not surjective, because otherwise the digit set would only consist of 0.

Before we define expansions and multi-expansions, we give a short introduction on multisets. We take the notation used, for example, in Knuth [15].

Notation 2.2. A *multiset* is like a set, but identical elements are allowed to appear more than once. For a multiset A, its cardinality $\#A$ is the number of elements in the multiset. For multisets A and B, we define new multisets $A \uplus B$ and $A \setminus B$ in the following way: If an element occurs exactly a times in A and b times in B, then it occurs exactly $a + b$ times in $A \uplus B$ and it occurs exactly $\max(a - b, 0)$ times in $A \setminus B$.

Now a pre-numeral system (and multisets) can be used to define what expansions and multi-expansions are.

Definition 2.3 (Expansion)**.** Let $(\mathcal{A}, \Phi, \mathcal{D})$ be a pre-numeral system, and let $\boldsymbol{\mu}$ be a multiset with elements $(d, n) \in (\mathcal{D} \setminus \{0\}) \times \mathbb{N}_0$. We define the following:

(1) We set
$$\mathsf{weight}(\boldsymbol{\mu}) := \#\boldsymbol{\mu}$$
and call it the *Hamming-weight of* $\boldsymbol{\mu}$ or simply *weight of* $\boldsymbol{\mu}$. The multiset $\boldsymbol{\eta}$ is called *finite*, if its weight is finite.

(2) We call an element $(d, n) \in \boldsymbol{\mu}$ an *atom* and $\Phi^n(d)$ the *value of the atom* (d, n).

(3) Let $\boldsymbol{\mu}$ be finite. We call
$$\mathsf{value}(\boldsymbol{\mu}) := \sum_{(d,n)\in\boldsymbol{\mu}} \Phi^n(d)$$
the *value of* $\boldsymbol{\mu}$.

(4) Let $z \in \mathcal{A}$. A *multi-expansion of* z is a finite $\boldsymbol{\mu}$ with $\mathsf{value}(\boldsymbol{\mu}) = z$.

(5) Let $z \in \mathcal{A}$. An *expansion of* z is a multi-expansion $\boldsymbol{\mu}$ of z where all the n in $(d, n) \in \boldsymbol{\mu}$ are pairwise distinct.

We use the following conventions and notations. If necessary, we see an atom as a multi-expansion or an expansion of weight 1. We identify an expansion $\boldsymbol{\eta}$ with the sequence $(\eta_n)_{n\in\mathbb{N}_0} \in \mathcal{D}^{\mathbb{N}_0}$, where $\eta_n = d$ for $(d, n) \in \boldsymbol{\eta}$ and all other $\eta_n = 0$. For an expansion $\boldsymbol{\eta}$ (usually a bold, lower case Greek letter) we will use η_n (the same letter, but indexed and not bold) for the elements of the sequence. Further, we identify expansions (sequences) in $\mathcal{D}^{\mathbb{N}_0}$ with finite words over the alphabet $\mathcal{D}$ written from right (least significant digit) to left (most significant digit), except left-trailing zeros, which are usually skipped. Besides, we follow the terminology of Lothaire [19] for words.

Note, if $\boldsymbol{\eta}$ is an expansion, then the weight of $\boldsymbol{\eta}$ is
$$\mathsf{weight}(\boldsymbol{\eta}) = \#\{n \in \mathbb{N}_0 \,|\, \eta_n \neq 0\}$$

and the value of $\boldsymbol{\eta}$ is

$$\mathsf{value}(\boldsymbol{\eta}) = \sum_{n\in\mathbb{N}_0} \Phi^n(\eta_n)\,.$$

For the sake of completeness—although we do not need it in this paper—a pre-numeral system is called *numeral system* if each element of $\mathcal{A}$ has an expansion. We call the numeral system *non-redundant* if there is exactly one expansion for each element of $\mathcal{A}$, otherwise we call it *redundant*. We will modify this definition later for w-NAF numeral systems.

Before going any further, we want to see some simple examples for the given abstract definition of a numeral system. We use multiplication by an element τ as endomorphism Φ. This leads to values of the type

$$\mathsf{value}(\boldsymbol{\eta}) = \sum_{n\in\mathbb{N}_0} \eta_n\tau^n$$

for an expansion $\boldsymbol{\eta}$.

Example 2.4. The binary numeral system is the pre-numeral system

$$(\mathbb{N}_0, z\mapsto 2z, \{0,1\}).$$

It is a non-redundant numeral system, since each integer admits exactly one binary expansion. We can extend the binary numeral system to the pre-numeral system

$$(\mathbb{Z}, z\mapsto 2z, \{-1,0,1\}),$$

which is a redundant numeral system.

In order to get a non-redundant numeral system out of a redundant one, one can restrict the language, i.e. we forbid some special configurations in an expansion. There is one special kind of expansion, namely the non-adjacent form, where no adjacent nonzeros are allowed. A generalisation of it is defined here.

Definition 2.5 (Width-w Non-Adjacent Form)**.** Let w be a positive integer and $\mathcal{D}$ be a digit set (coming from a pre-numeral system). Let $\boldsymbol{\eta} = (\eta_j)_{j\in\mathbb{N}_0} \in \mathcal{D}^{\mathbb{N}_0}$. The sequence $\boldsymbol{\eta}$ is called a *width-w non-adjacent form*, or *w-NAF* for short, if each factor[1] $\eta_{j+w-1}\dots\eta_j$, i.e. each block of length w, contains at most one nonzero digit.

A w-NAF-expansion is an expansion that is also a w-NAF.

Note that a w-NAF-expansion is finite. With the previous definition we can now define what a w-NAF numeral system is.

[1]See Lothaire [19] for the used terminology on words.

Definition 2.6. For a positive integer w, a pre-numeral system $(\mathcal{A}, \Phi, \mathcal{D})$ is called a *w-NAF numeral system* if each element of $\mathcal{A}$ admits a w-NAF-expansion, i.e. for each $z \in \mathcal{A}$ there is a w-NAF $\boldsymbol{\eta} \in \mathcal{D}^{\mathbb{N}_0}$ with $\mathsf{value}(\boldsymbol{\eta}) = z$. We call a w-NAF numeral system *non-redundant* if each element of $\mathcal{A}$ has a unique w-NAF-expansion, otherwise we call it *redundant*.

Now we continue the example started above.

Example 2.7. The redundant numeral system

$$(\mathbb{Z}, z \mapsto 2z, \{-1, 0, 1\})$$

is a non-redundant 2-NAF numeral system. This fact has been shown in Reitwiesner [24]. More generally, for an integer w at least 2, the numeral system

$$(\mathbb{Z}, z \mapsto 2z, \mathcal{D}),$$

where the digit set $\mathcal{D}$ consists of 0 and all odd integers with absolute value smaller than 2^{w-1}, is a non-redundant w-NAF numeral system, cf. Solinas [25, 26] or Muir and Stinson [22].

Finally, since this paper deals with the optimality of expansions, we have to define the term "optimal". This is done in the following definition.

Definition 2.8 (Optimal Expansion)**.** Let $(\mathcal{A}, \Phi, \mathcal{D})$ be a pre-numeral system, and let $z \in \mathcal{A}$. A multi-expansion or an expansion $\boldsymbol{\mu}$ of z is called *optimal* if for any multi-expansion $\boldsymbol{\nu}$ of z we have

$$\mathsf{weight}(\boldsymbol{\mu}) \leq \mathsf{weight}(\boldsymbol{\nu}),$$

i.e. $\boldsymbol{\mu}$ minimises the Hamming-weight among all multi-expansions of z. Otherwise $\boldsymbol{\mu}$ is called *non-optimal*.

The "usual" definition of optimal, cf. [24, 14, 9, 4, 22, 23, 21, 12, 1, 2, 10], is more restrictive: An expansion of $z \in \mathcal{A}$ is optimal if it minimises the weight among all expansions of z. The difference is that in Definition 2.8 we minimise over all multi-expansions. The use of multi-expansions is motivated by applications: we want to do efficient operations. There it is no problem to take multi-expansions if they are "better", so it is more natural to minimise over all of them instead of just over all expansions.

3. The Optimality Result

This section contains our main theorem, the Optimality Theorem, Theorem 3.2. It contains four equivalences. One of them is a condition on the digit set and one is optimality of the w-NAF. We start with the definition of that condition on the digit set.

Definition 3.1. Let $(\mathcal{A}, \Phi, \mathcal{D})$ be a pre-numeral system, and let w be a positive integer. We say that the digit set $\mathcal{D}$ is *w-subadditive* if the sum of the values of two atoms has a w-NAF-expansion of weight at most 2.

In order to verify the w-subadditivity-condition it is enough to check atoms $(c, 0)$ and (d, n) with $n \in \{0, \dots, w-1\}$ and nonzero digits c and d. Therefore, one has to consider $w\,(\#\mathcal{D}-1)^2$ multi-expansions.

Theorem 3.2 (Optimality Theorem)**.** *Let $(\mathcal{A}, \Phi, \mathcal{D})$ be a pre-numeral system with*

$$\bigcap_{m\in\mathbb{N}_0} \Phi^m(\mathcal{A}) = \{0\}\,,$$

and let w be a positive integer. Then the following statements are equivalent:

(1) The digit set $\mathcal{D}$ is w-subadditive.

(2) For all multi-expansions $\boldsymbol{\mu}$ there is a w-NAF-expansion $\boldsymbol{\xi}$ such that

$$\mathsf{value}(\boldsymbol{\xi}) = \mathsf{value}(\boldsymbol{\mu})$$

and

$$\mathsf{weight}(\boldsymbol{\xi}) \le \mathsf{weight}(\boldsymbol{\mu})\,.$$

(3) For all w-NAF-expansions $\boldsymbol{\eta}$ and $\boldsymbol{\vartheta}$ there is a w-NAF-expansion $\boldsymbol{\xi}$ such that

$$\mathsf{value}(\boldsymbol{\xi}) = \mathsf{value}(\boldsymbol{\eta}) + \mathsf{value}(\boldsymbol{\vartheta})$$

and

$$\mathsf{weight}(\boldsymbol{\xi}) \le \mathsf{weight}(\boldsymbol{\eta}) + \mathsf{weight}(\boldsymbol{\vartheta})\,.$$

(4) If $z \in \mathcal{A}$ admits a multi-expansion, then z also admits an optimal w-NAF-expansion.

Note that if we assume that each element $\mathcal{A}$ has at least one expansion (e.g. by assuming that we have a w-NAF numeral system), then we have the equivalence of w-subadditivity of the digit set and the existence of an optimal w-NAF-expansion for each group element.

We will use the term "addition" in the following way: The addition of two group elements x and y means finding a w-NAF-expansion of the sum $x+y$. Addition of two multi-expansions shall mean addition of their values.

Proof of Theorem 3.2. For a nonzero $z \in \mathcal{A}$, we define

$$L(z) := \max\{m \in \mathbb{N}_0 \mid z \in \Phi^m(\mathcal{A})\}\,.$$

The function L is well-defined, because

$$\bigcap_{m\in\mathbb{N}_0} \Phi^m(\mathcal{A}) = \{0\}\,.$$

We show that *(1)* implies *(2)* by induction on the pair

$$(\mathsf{weight}(\boldsymbol{\mu})\,, L(\mathsf{value}(\boldsymbol{\mu})))$$

for the multi-expansion $\boldsymbol{\mu}$. The order on those pairs is lexicographic. In the case $\mathsf{value}(\boldsymbol{\mu}) = 0$, we choose $\boldsymbol{\xi} = 0$ and are finished. Further, if the multi-expansion $\boldsymbol{\mu}$ consists of less than two elements, then there is nothing to do, so we suppose $\mathsf{weight}(\boldsymbol{\eta}) \ge 2$.

We choose an atom $(d, n) \in \boldsymbol{\mu}$ (note that $d \in \mathcal{D} \setminus \{0\}$ and $n \in \mathbb{N}_0$) with minimal n. If $n > 0$, then we consider the multi-expansion $\boldsymbol{\mu}'$ arising from $\boldsymbol{\mu}$ by shifting all indices by n, use the induction hypothesis on $\boldsymbol{\mu}'$ and apply Φ^n. Note that $\boldsymbol{\mu}'$ and $\boldsymbol{\mu}$ have the same weight, but

$$L(\mathsf{value}(\boldsymbol{\mu}')) = L(\mathsf{value}(\boldsymbol{\mu})) - n < L(\mathsf{value}(\boldsymbol{\mu}))\,.$$

So we can assume $n = 0$. Set $\boldsymbol{\mu}^\star := \boldsymbol{\mu} \setminus \{(d, 0)\}$. Using the induction hypothesis, there is a w-NAF-expansion $\boldsymbol{\eta}$ of $\mathsf{value}(\boldsymbol{\mu}^\star)$ with weight strictly smaller than $\mathsf{weight}(\boldsymbol{\mu}) = \mathsf{weight}(\boldsymbol{\mu}^\star) + 1$.

Consider the addition of $\boldsymbol{\eta}$ and the digit d. If the digits η_ℓ are zero for all $\ell \in \{0, \ldots, w-1\}$, then the result follows by setting $\boldsymbol{\xi} = \ldots \eta_{w+1}\eta_w 0^{w-1} d$. So we can assume

$$\boldsymbol{\eta} = \boldsymbol{\beta} 0^{w-k-1} b 0^k$$

with a w-NAF $\boldsymbol{\beta}$, a digit $b \neq 0$ and $k \in \{0, \ldots, w-1\}$. Note that there are at least $w-1$ zeros on the left hand-side of b in $\boldsymbol{\eta}$, but for our purposes, it is sufficient (and more convenient) to consider only $w-k-1$ zeros. Since the digit set $\mathcal{D}$ is w-subadditive, there is a w-NAF $\boldsymbol{\gamma}$ of $\Phi^k(b) + d$ with weight at most 2. If the weight is strictly smaller than 2, we use the induction hypothesis on the multi-expansion $\boldsymbol{\beta} 0^w \uplus \boldsymbol{\gamma}$ to get a w-NAF $\boldsymbol{\xi}$ with the desired properties and are done. Otherwise, denoting by J the smallest index with $\gamma_J \neq 0$, we distinguish between two cases: $J = 0$ and $J > 0$.

First let $J = 0$. The w-NAF $\boldsymbol{\beta}$ (seen as multi-expansion) has a weight less than $\mathsf{weight}(\boldsymbol{\eta})$, so, by induction hypothesis, there is a w-NAF $\boldsymbol{\xi}'$ with

$$\mathsf{value}(\boldsymbol{\xi}') = \mathsf{value}(\boldsymbol{\beta}) + \mathsf{value}(\ldots \gamma_{w+1}\gamma_w)$$

and

$$\mathsf{weight}(\boldsymbol{\xi}') \leq \mathsf{weight}(\boldsymbol{\beta}) + \mathsf{weight}(\ldots \gamma_{w+1}\gamma_w)\,.$$

We set $\boldsymbol{\xi} = \boldsymbol{\xi}'\gamma_{w-1} \ldots \gamma_0$. Since $\boldsymbol{\xi}$ is a w-NAF-expansion we are finished, because

$$\begin{aligned}\mathsf{value}(\boldsymbol{\xi}) &= \Phi^w(\mathsf{value}(\boldsymbol{\beta})) + \Phi^w(\mathsf{value}(\ldots \gamma_{w+1}\gamma_w)) + \mathsf{value}(\gamma_{w-1} \ldots \gamma_0) \\ &= \Phi^w(\mathsf{value}(\boldsymbol{\beta})) + \Phi^k(b) + d = \mathsf{value}(\boldsymbol{\eta}) + d = \mathsf{value}(\boldsymbol{\mu}^\star) + d = \mathsf{value}(\boldsymbol{\mu})\end{aligned}$$

and

$$\begin{aligned}\mathsf{weight}(\boldsymbol{\xi}) = \mathsf{weight}(\boldsymbol{\xi}') + \mathsf{weight}(\gamma_{w-1} \ldots \gamma_0) &\leq \mathsf{weight}(\boldsymbol{\beta}) + \mathsf{weight}(\boldsymbol{\gamma}) \\ &\leq \mathsf{weight}(\boldsymbol{\eta}) + 1 \leq \mathsf{weight}(\boldsymbol{\mu}^\star) + 1 = \mathsf{weight}(\boldsymbol{\mu})\,.\end{aligned}$$

Now, in the case $J > 0$, we consider the multi-expansion $\boldsymbol{\nu} := \boldsymbol{\beta} 0^w \uplus \boldsymbol{\gamma}$. We use the induction hypothesis for $\boldsymbol{\nu}$ shifted by J (same weight, L decreased by J) and apply Φ^J on the result.

The proofs of the other implications of the four equivalences are simple. To show that *(2)* implies *(3)*, take $\boldsymbol{\mu} := \boldsymbol{\eta} \uplus \boldsymbol{\vartheta}$, and *(3)* implies *(1)* is the special case when $\boldsymbol{\eta}$ and $\boldsymbol{\vartheta}$ are atoms.

Further, for *(2)* implies *(4)* take an optimal multi-expansion $\boldsymbol{\mu}$ (which exists, since z admits at least one multi-expansion). We get a w-NAF-expansion $\boldsymbol{\xi}$ with $\mathsf{weight}(\boldsymbol{\xi}) \le \mathsf{weight}(\boldsymbol{\mu})$. Since $\boldsymbol{\mu}$ was optimal, equality is obtained in the previous inequality, and therefore $\boldsymbol{\xi}$ is optimal, too. The converse, *(4)* implies *(2)*, follows using $z = \mathsf{value}(\boldsymbol{\mu})$ and the property that optimal expansions minimise the weight. □

Let X and Y be subsets in an additively written semigroup. Then we write

$$X + Y := \{x + y \,|\, x \in X, y \in Y\}\,,$$

see, for example, Hungerford [13]. We use that notion from now on.

Proposition 3.3. *Let $(\mathcal{A}, \Phi, \mathcal{D})$ be a pre-numeral system with*

$$\bigcap_{m \in \mathbb{N}_0} \Phi^m(\mathcal{A}) = \{0\}\,,$$

and let w be a positive integer. We have the following sufficient condition: Suppose we have sets U and S such that $\mathcal{D} \subseteq U$, $-\mathcal{D} \subseteq U$, $U \subseteq \Phi(U)$ and all elements in S are atoms. If $\mathcal{D}$ contains a representative for each residue class modulo $\Phi^w(\mathcal{A})$ which is not contained in $\Phi(\mathcal{A})$ and

$$\Big(\Phi^{w-1}(U) + U + U\Big) \cap \Phi^w(\mathcal{A}) \subseteq S \cup \{0\}\,, \tag{3.1}$$

then the digit set $\mathcal{D}$ is w-subadditive.

Sometimes it is more convenient to use (3.1) of this proposition instead of the definition of w-subadditive. For example, in Section 4 all digits lie in an interval U and all nonzero integers in that interval have a w-NAF expansion with weight 1. The same technique is used in the optimality result of Section 7.

Proof of Proposition 3.3. Consider $y = \mathsf{value}((c, 0) \uplus (d, n))$, where $(c, 0)$ and (d, n) are atoms with $n \in \{0, \dots, w-1\}$. If $y = 0$, we have nothing to do, so we can assume $y \neq 0$. First suppose $y \notin \Phi(\mathcal{A})$. Because of our assumptions on $\mathcal{D}$ there is a digit a such that

$$z := \Phi^n(d) + c - a \in \Phi^w(\mathcal{A})\,.$$

If z is not zero, then, using our sufficient condition, there is an atom (b, m) with value z, and we have $m \ge w$. The w-NAF-expansion $b0^{m-1}a$ does what we want.

Now suppose $y \in \Phi^k(\mathcal{A})$ with a positive integer k, which is chosen maximally. That case can only happen when $n = 0$. Since $y \neq 0$ and our

assumptions on $\mathcal{D}$ there is an w-NAF-expansion of y with an atom (a,k) as least significant digit. If $k \in \{0,\dots,w-1\}$, then

$$z := d + c - \Phi^k(a) \in \Phi^{w+k}(\mathcal{A})\,.$$

If z is nonzero it is a value of an atom (b,m), $m \geq w+k$, because of (3.1), and we obtain a w-NAF-expansion of y with atoms (b,m) and (a,k). If $k \geq w$, then

$$z := d + c \in \Phi^w(\mathcal{A})$$

and z is the value of an atom (b,m) by (3.1). We get a w-NAF-expansion $b0^m$. □

Sometimes the w-subadditivity-condition is a bit too strong, so we do not get optimal w-NAFs. In that case one can check whether $(w-1)$-NAFs are optimal. This is stated in the following remark, where the w-subadditive-condition is weakened.

Remark 3.4. Suppose that we have the same setting as in Theorem 3.2. We call the digit set *w-weak-subadditive* if the sum of the values of two atoms (c,m) and (d,n) with $|m-n| \neq w-1$ has a w-NAF-expansion with weight at most 2.

We get the following result: If the digit set $\mathcal{D}$ is w-weak-subadditive, then each element of $\mathcal{A}$, which has at least one multi-expansion, has an optimal $(w-1)$-NAF-expansion. The proof is similar to the proof of Theorem 3.2, except that a "rewriting" only happens when we have a $(w-1)$-NAF-violation.

4. Optimality for Integer Bases

In this section we give a first application of the abstract optimality theorem of the previous section. We reprove the optimality of the w-NAFs with a minimal norm digit set and base 2. But the result is more general: We prove optimality for all integer bases (with absolute value at least 2). This demonstrates one basic idea how to check whether a digit set is w-subadditive or not.

Let b be an integer with $|b| \geq 2$ and w be an integer with $w \geq 2$. Consider the non-redundant w-NAF numeral system

$$(\mathbb{Z}, z \mapsto bz, \mathcal{D})$$

where the digit set $\mathcal{D}$ consists of 0 and all integers with absolute value strictly smaller than $\frac{1}{2}\,|b|^w$ and not divisible by b. We mentioned the special case base 2 of that numeral system in Example 2.7. See also Reitwiesner [24] and Solinas [26].

The following optimality result can be shown. For proofs of the base 2 setting cf. Reitwiesner [24], Jedwab and Mitchell [14], Gordon [9], Avanzi [4], Muir and Stinson [22], and Phillips and Burgess [23].

Theorem 4.1. *With the setting above, the w-NAF-expansion for each integer is optimal.*

Proof. We show that the digit set $\mathcal{D}$ is w-subadditive by verifying the sufficient condition of Proposition 3.3. Then optimality follows from Theorem 3.2. First, note that the w-NAF-expansion of each integer with absolute value at most $\frac{1}{2}|b|^{w-1}$ has weight at most 1, because either the integer is already a digit, or one can divide by a power of b to get a digit. Further, we have $\mathcal{D} = -\mathcal{D}$. Set $U = \left[-\frac{1}{2}|b|^w, \frac{1}{2}|b|^w\right]$ and $S = b^w(U \cap \mathbb{Z} \setminus \{0\})$. We have to show

$$\left(b^{w-1}U + U + U\right) \cap b^w\mathbb{Z} \subseteq S \cup \{0\} = b^wU \cap b^w\mathbb{Z}.$$

If we can show

$$b^{-w}\left(b^{w-1}U + U + U\right) \subseteq \left[-\tfrac{1}{2}|b|^w, \tfrac{1}{2}|b|^w\right],$$

the inclusion above follows by multiplying with b^w and taking the intersection with $b^w\mathbb{Z}$.

So let

$$b^wz = b^{w-1}c + a + d$$

for some digits a, c and d. A digit has absolute value less than $\frac{1}{2}|b|^w$, so

$$|z| < |b|^{-w}\left(|b|^{w-1} + 2\right)\tfrac{1}{2}|b|^w = \left(|b|^{-1} + 2|b|^{-w}\right)\tfrac{1}{2}|b|^w \leq \tfrac{1}{2}|b|^w,$$

where we also used the assumptions $|b| \geq 2$ and $w \geq 2$. Thus, the desired inclusion is shown. □

5. Voronoi Cells

We first start to define Voronoi cells. Let $\tau \in \mathbb{C}$ be an algebraic integer that is imaginary quadratic, i.e. τ is solution of an equation $\tau^2 - p\tau + q = 0$ with $p, q \in \mathbb{Z}$ and such that $q - p^2/4 > 0$.

Definition 5.1 (Voronoi Cell)**.** We set

$$V := \{z \in \mathbb{C} \,|\, \forall y \in \mathbb{Z}[\tau]\colon\ |z| \leq |z - y|\}$$

and call it the *Voronoi cell for* 0 corresponding to the set $\mathbb{Z}[\tau]$. Let $u \in \mathbb{Z}[\tau]$. We define the *Voronoi cell for* u as

$$V_u := u + V = \{u + z \,|\, z \in V\} = \{z \in \mathbb{C} \,|\, \forall y \in \mathbb{Z}[\tau]\colon\ |z - u| \leq |z - y|\}.$$

The point u is called *centre of the Voronoi cell* or *lattice point corresponding to the Voronoi cell*.

An example of a Voronoi cell in a lattice $\mathbb{Z}[\tau]$ is shown in Figure 5.1. Two neighbouring Voronoi cells have at most a subset of their boundary in common. This can be a problem, when we tile the plane with Voronoi cells and want that each point is in exactly one cell. To fix this problem we define

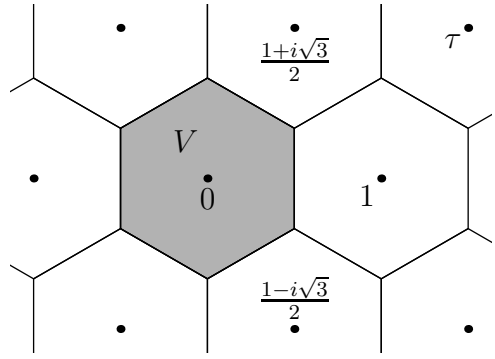


FIGURE 5.1. Voronoi cell V for 0 corresponding to the set $\mathbb{Z}[\tau]$ with $\tau = \frac{3}{2} + \frac{i}{2}\sqrt{3}$.

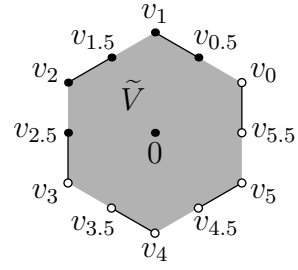


FIGURE 5.2. Restricted Voronoi cell $\widetilde{V}$ for 0 corresponding to the set $\mathbb{Z}[\tau]$ with $\tau = \frac{3}{2} + \frac{i}{2}\sqrt{3}$.

a restricted version of V. This is very similar to the construction used in Avanzi, Heuberger and Prodinger [3] and in Heuberger and Krenn [11].

Definition 5.2 (Restricted Voronoi Cell)**.** Let V_u be a Voronoi cell with its centre u as above. Let $v_0, \dots, v_{m-1}$ with appropriate $m \in \mathbb{N}$ be the vertices of V_u. We denote the midpoint of the line segment from v_k to v_{k+1} by $v_{k+1/2}$, and we use the convention that the indices are meant modulo m.

The *restricted Voronoi cell* $\widetilde{V}_u$ consists of

- the interior of V_u,
- the line segments from $v_{k+1/2}$ (excluded) to v_{k+1} (excluded) for all k,
- the points $v_{k+1/2}$ for $k \in \{0, \dots, \lfloor \frac{m}{2} \rfloor - 1\}$, and
- the points v_k for $k \in \{1, \dots, \lfloor \frac{m}{3} \rfloor\}$.

Again we set $\widetilde{V} := \widetilde{V}_0$.

In Figure 5.2 the restricted Voronoi cell of 0 is shown for $\tau = \frac{3}{2}+\frac{i}{2}\sqrt{3}$. The second condition in the definition is used because it benefits symmetries. The third condition is just to make the midpoints unique. Obviously, other rules[2] could have been used to define the restricted Voronoi cell.

The statements (including proofs) of the following lemma can be found in Heuberger and Krenn [11]. We use the notation $\mathcal{B}(z,r)$ for an open ball with centre z and radius r and $\overline{\mathcal{B}}(z,r)$ for a closed ball.

Lemma 5.3 (Properties of Voronoi Cells)**.** *We have the following properties:*

(a) The vertices of V are given explicitly by

$$v_0 = 1/2 + \frac{i}{2\operatorname{Im}(\tau)}\left(\operatorname{Im}(\tau)^2 + \{\operatorname{Re}(\tau)\}^2 - \{\operatorname{Re}(\tau)\}\right),$$

$$v_1 = \{\operatorname{Re}(\tau)\} - \frac{1}{2} + \frac{i}{2\operatorname{Im}(\tau)}\left(\operatorname{Im}(\tau)^2 - \{\operatorname{Re}(\tau)\}^2 + \{\operatorname{Re}(\tau)\}\right),$$

$$v_2 = -1/2 + \frac{i}{2\operatorname{Im}(\tau)}\left(\operatorname{Im}(\tau)^2 + \{\operatorname{Re}(\tau)\}^2 - \{\operatorname{Re}(\tau)\}\right) = v_0 - 1,$$

$$v_3 = -v_0,$$

$$v_4 = -v_1$$

and

$$v_5 = -v_2.$$

All vertices have the same absolute value. If $\operatorname{Re}(\tau) \in \mathbb{Z}$, then $v_1 = v_2$ and $v_4 = v_5$, i.e. the hexagon degenerates to a rectangle.

(b) The Voronoi cell V is convex.

(c) We get $\overline{\mathcal{B}}\left(0, \frac{1}{2}\right) \subseteq V$.

(d) The inclusion $\tau^{-1}V \subseteq V$ holds.

6. Digit Sets for Imaginary Quadratic Bases

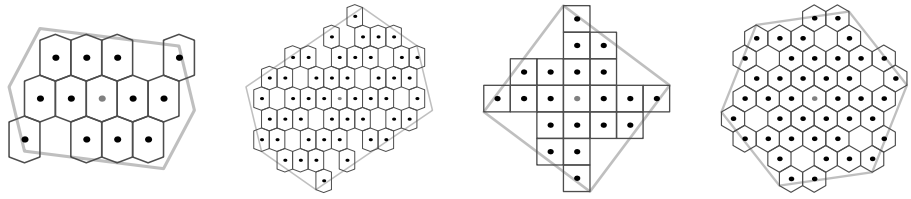
In this section we assume that $\tau \in \mathbb{C}$ is an imaginary quadratic algebraic integer, i.e. τ is solution of an equation $\tau^2 - p\tau + q = 0$ with $p, q \in \mathbb{Z}$ and such that $q - p^2/4 > 0$. By V we denote the Voronoi cell of 0 of the lattice $\mathbb{Z}[\tau]$, by $\widetilde{V}$ the corresponding restricted Voronoi cell, cf. Section 5.

We consider w-NAF numeral systems

$$(\mathbb{Z}[\tau], z \mapsto \tau z, \mathcal{D}),$$

where the digit set $\mathcal{D}$ is the so called "minimal norm representatives digit set". The following definition specifies that digit set, cf. Solinas [25, 26],

[2]The rule has to make sure that the complex plane can be covered entirely and with no overlaps by restricted Voronoi cells, i.e. the condition $\mathbb{C} = \biguplus_{z\in\mathbb{Z}[\tau]} \widetilde{V}_z$ has to be fulfilled.



(a) Digit set for $\tau = \frac{3}{2} + \frac{i}{2}\sqrt{7}$ and $w = 2$. (b) Digit set for $\tau = \frac{3}{2} + \frac{i}{2}\sqrt{7}$ and $w = 3$. (c) Digit set for $\tau = 2 + i$ and $w = 2$. (d) Digit set for $\tau = \frac{5}{2} + \frac{i}{2}\sqrt{3}$ and $w = 2$.

FIGURE 6.1. Minimal norm representatives digit sets modulo τ^w. For each digit η, the corresponding Voronoi cell V_η is drawn. The large scaled Voronoi cell is $\tau^w V$.

Blake, Murty and Xu [5] or Heuberger and Krenn [11]. It is used throughout this article, whenever we have the setting (imaginary quadratic base) mentioned above.

Definition 6.1 (Minimal Norm Representatives Digit Set)**.** Let w be an integer with $w \geq 2$ and $\mathcal{D} \subseteq \mathbb{Z}[\tau]$ consist of 0 and exactly one representative of each residue class of $\mathbb{Z}[\tau]$ modulo τ^w that is not divisible by τ. If all such representatives $\eta \in \mathcal{D}$ fulfil $\eta \in \tau^w \widetilde{V}$, then $\mathcal{D}$ is called the *minimal norm representatives digit set modulo* τ^w.

The previous definition uses the restricted Voronoi cell $\widetilde{V}$ for the point 0, see Definition 5.2, to choose a representative with minimal norm. Note that by construction of $\widetilde{V}$, there is only one such choice for the digit set. Some examples of such digit sets are shown in Figures 6.1, 8.1, 9.1 and 10.1.

Remark 6.2. The definition of a minimal norm representative digit set, Definition 6.1, depends on the definition of the restricted Voronoi cell $\widetilde{V}$, Definition 5.2. There we had some freedom in choosing which part of the boundary is included in $\widetilde{V}$, cf. the remarks after Definition 5.2. We point out that all results given here for imaginary quadratic bases are valid for any admissible configuration of the restricted Voronoi cell, although only the case corresponding to Definition 5.2 will be presented.

Using a minimal norm representatives digit set, each element of $\mathbb{Z}[\tau]$ corresponds to a unique w-NAF, i.e. the pre-numeral system given at the beginning of this section is indeed a w-NAF numeral system. This is stated in the following theorem, which can be found in Heuberger and Krenn [11].

Theorem 6.3 (Existence and Uniqueness Theorem)**.** *Let w be an integer with $w \geq 2$. Then the pre-numeral system*

$$(\mathbb{Z}[\tau], z \mapsto \tau z, \mathcal{D}),$$

where $\mathcal{D}$ is the minimal norm representatives digit set modulo τ^w, is a non-redundant w-NAF numeral system, i.e. each lattice point $z \in \mathbb{Z}[\tau]$ has a unique w-NAF-expansion $\boldsymbol{\eta} \in \mathcal{D}^{\mathbb{N}_0}$ with $z = \mathsf{value}(\boldsymbol{\eta})$.

7. Optimality for Imaginary Quadratic Bases

In this section we assume that $\tau \in \mathbb{C}$ is an imaginary quadratic algebraic integer, i.e. τ is solution of an equation $\tau^2 - p\tau + q = 0$ with $p, q \in \mathbb{Z}$ and such that $q - p^2/4 > 0$. Further let w be an integer with $w \geq 2$ and let

$$(\mathbb{Z}[\tau], z \mapsto \tau z, \mathcal{D})$$

be the non-redundant w-NAF numeral system with minimal norm representatives digit set modulo τ^w, cf. Section 6.

Our main question in this section, as well as for the remaining part of this article, is the following: For which bases and which w is the width-w non-adjacent form optimal? To answer this, we use the result from Section 3. If we can show that the digit set $\mathcal{D}$ is w-subadditive, then optimality follows. This is done in the lemma below. The result will then be formulated in Corollary 7.2, which, eventually, contains the optimality result for our mentioned configuration.

Lemma 7.1. *Suppose that one of the following conditions hold:*

(i) $w \geq 4$ *and* $|p| \geq 3$,
(ii) $w = 3$ *and* $|p| \geq 5$,
(iii) $w = 3$, $|p| = 4$ *and* $5 \leq q \leq 9$,
(iv) $w = 2$, p *even, and*

$$\left(\frac{1}{\sqrt{q}} + \frac{2}{q}\right)^2 \left(q - \frac{p^2}{4} + 1\right) < 1$$

or equivalently

$$|p| > 2\sqrt{q + 1 - \frac{q^2}{\left(2 + \sqrt{q}\right)^2}},$$

(v) $w = 2$, p *odd and*

$$\left(\frac{1}{\sqrt{q}} + \frac{2}{q}\right)^2 \left(q - \frac{p^2}{4} + \frac{1}{4}\right)^2 \left(q - \frac{p^2}{4}\right)^{-1} < 1.$$

Then the digit set $\mathcal{D}$ is w-subadditive.

The conditions (iv) and (v) of Lemma 7.1, i.e. the case $w = 2$, are illustrated graphically in Figure 7.1.

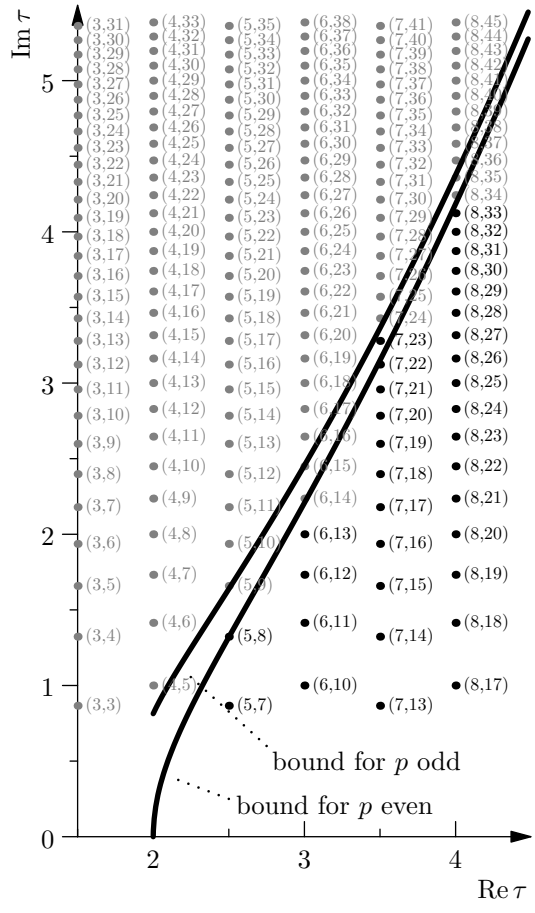


FIGURE 7.1. Bounds for the optimality of 2-NAFs. The two curves correspond to the conditions (iv) and (v) of Lemma 7.1. A dot corresponds to a valid τ. If the dot is black, then the 2-NAFs of that τ are optimal, gray means not decidable with this method. Each dot is labelled with $(|p|, q)$.

Proof. We denote the interior of V by $\operatorname{int}(V)$. If

$$\tau^{w-1}V + V + V \subseteq \tau^w \operatorname{int}(V)$$

holds, then the digit set $\mathcal{D}$ is w-subadditive since $\mathcal{D} \subseteq \tau^w V$, $-\mathcal{D} \subseteq \tau^w V$, $V \subseteq \tau V$ and $z \in \tau^w \operatorname{int}(V) \cap \mathbb{Z}[\tau]$ implies that there is an integer $\ell \geq 0$ with $z \in \tau^\ell \mathcal{D}$. The sufficient condition of Proposition 3.3 was used with $U = \tau^w V$ and $S = \tau^w \operatorname{int}(V) \setminus \{0\}$.

Since V is convex, it is sufficient to show that

$$\tau^{w-1}V + 2V \subseteq \tau^w \operatorname{int}(V)\,.$$

This will be done by showing

$$\left(|\tau|^{-1} + 2\,|\tau|^{-w}\right)|V| < \tfrac{1}{2},$$

where $|V|$ denotes the radius of the smallest closed disc with centre 0 containing V. By setting

$$T(p,q,w) := 2\left(|\tau|^{-1} + 2\,|\tau|^{-w}\right)|V|\,,$$

we have to show that

$$T(p,q,w) < 1.$$

Note that $T(p,q,w) > 0$, so it is sufficient to show

$$T^2(p,q,w) < 1.$$

For each of the different conditions given, we will check that the inequality holds for special values of p, q and w and then use a monotonicity argument to get the result for other values of p, q and w. In the following we distinguish between even and odd p.

Let p be even, first. Then $\frac{1}{2} + \frac{i}{2}\operatorname{Im}(\tau)$ is a vertex of the Voronoi cell V. This means $|V| = \frac{1}{2}\sqrt{1+q-p^2/4}$. Inserting that and $|\tau| = \sqrt{q}$ in the asserted inequality yields

$$T^2(p,q,w) = \left(\frac{1}{\sqrt{q}} + 2q^{-w/2}\right)^2 \left(1 + q - \frac{p^2}{4}\right) < 1.$$

It is easy to see that the left hand side of this inequality is monotonically decreasing in $|p|$ (as long as the condition $q > p^2/4$ is fulfilled) and monotonically decreasing in w. We assume $p \geq 0$.

If we set $p = 4$ and $w = 4$, we get

$$T^2(4,q,4) = -\frac{12}{q^4} + \frac{4}{q^3} - \frac{12}{q^{5/2}} + \frac{4}{q^{3/2}} - \frac{3}{q} + 1,$$

which is strictly monotonically increasing for $q \geq 5$. Further we get

$$\lim_{q\to\infty} T^2(4,q,4) = 1.$$

This means $T^2(4,q,4) < 1$ for all $q \geq 5$. Since $p \geq 4$ implies $q \geq 5$ and because of the monotonicity mentioned before, the case (i) for the even p is completed.

If we set $p = 6$ and $w = 3$, we get

$$T^2(6,q,3) = -\frac{32}{q^3} - \frac{28}{q^2} - \frac{4}{q} + 1,$$

which is obviously less than 1. Therefore, again by monotonicity, the case (ii) is done for the even p.

If we set $p = 4$ and $w = 3$, we obtain

$$T^2(4, q, 3) = -\frac{12}{q^3} - \frac{8}{q^2} + \frac{1}{q} + 1,$$

which is monotonically increasing for $5 \leq q \leq 18$. Further we get

$$T^2(4, 9, 3) = \frac{242}{243} < 1$$

and $T^2(4, 10, 3) > 1$. This means $T^2(4, q, 3) < 1$ for all q with $5 \leq q \leq 9$. So case (iii) is completed.

The condition given in (iv) is exactly

$$T^2(p, q, 2) < 1$$

for even p, so the result follows immediately.

Now, let p be odd. Then $\frac{i}{2\,\mathrm{Im}(\tau)} \left(\mathrm{Im}(\tau)^2 + \frac{1}{4}\right)$ is a vertex of the Voronoi cell V. This means

$$|V| = \frac{1}{2} \left(q - \frac{p^2}{4}\right)^{-1/2} \left(q - \frac{p^2}{4} + \frac{1}{4}\right).$$

Inserting that in the asserted inequality yields

$$T^2(p, q, w) = \left(q - \frac{p^2}{4}\right)^{-1} \left(q - \frac{p^2}{4} + \frac{1}{4}\right)^2 \left(2q^{-w/2} + q^{-1/2}\right)^2 < 1.$$

Again, it is easy to verify that the left hand side of this inequality is monotonically decreasing in p (as long as the condition $q \geq p^2/4 + 1/4$ is fulfilled) and monotonically decreasing in w. We assume $p \geq 0$.

If we set $p = 3$ and $w = 4$, we get

$$T^2(3, q, 4) = \frac{4(q-2)^2 \left(q^{3/2} + 2\right)^2}{q^4(4q - 9)}$$

which is strictly monotonically increasing for $q \geq 3$. Further we get

$$\lim_{q \to \infty} T^2(3, q, 4) = 1.$$

This means $T^2(3, q, 4) < 1$ for all $q \geq 3$. Since $p \geq 3$ implies $q \geq 3$ and because of the monotonicity mentioned before, the case (i) for the odd p is finished.

If we set $p = 5$ and $w = 3$, we get

$$T^2(5, q, 3) = \frac{4(q-6)^2(q+2)^2}{q^3(4q - 25)}$$

which is strictly monotonically increasing for $q \geq 7$. Further we get

$$\lim_{q \to \infty} T^2(5, q, 3) = 1.$$

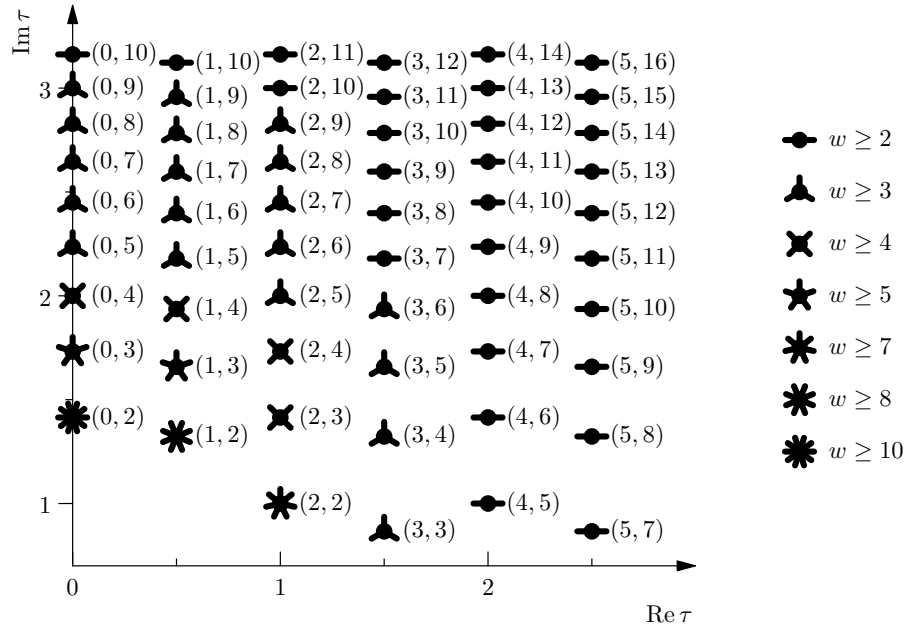


Figure 7.2. Optimality of a $(w-1)$-NAF-expansion with a digit set used for w-NAFs. Each symbol is labelled with $(|p|, q)$ and represents the minimal w for which there is an optimal $(w-1)$-NAF-expansion of each element of $\mathbb{Z}[\tau]$.

This means $0 < T^2(5, q, 3) < 1$ for all $q \geq 7$. As $p \geq 5$ implies $q \geq 7$, using monotonicity again, the case (ii) is done for the odd p.

The condition given in (v) is exactly

$$T^2(p, q, 2) < 1$$

for odd p, so the result follows immediately.

Since we have now analysed all the conditions, the proof is finished. □

Now we can prove the following optimality corollary, which is a consequence of Theorem 3.2.

Corollary 7.2. *Suppose that one of the conditions (i) to (v) of Lemma 7.1 holds. Then the width-w non-adjacent form expansion for each element of $\mathbb{Z}[\tau]$ is optimal.*

Proof. Lemma 7.1 implies that the digit set $\mathcal{D}$ is w-subadditive, therefore Theorem 3.2 can be used directly to get the desired result. □

Remark 7.3. We have the following weaker optimality result. Let p, q and w be integers with $|p| \geq p_0$, $q \geq q_0$ and $w \geq w_0$ for a $(p_0, q_o, w_0) \in Y$, where

$$Y = \{(0,10,2),(0,5,3),(0,4,4),(0,3,5),(0,2,10),\\ (1,2,8),(2,3,4),(2,2,7),(3,7,2),(3,3,3),(4,5,2)\}.$$

Then we can show that the minimal norm representatives digit set modulo τ^w coming from a τ with (p,q) is w-weak-subadditive, and therefore, by Remark 3.4, we obtain optimality of a $(w-1)$-NAF of each element of $\mathbb{Z}[\tau]$. The results are visualised graphically in Figure 7.2.

To show that the digit set is w-weak-subadditive we proceed in the same way as in the proof of Lemma 7.1. We have to show the condition

$$T'(p,q,w) < 1$$

where

$$T'(p,q,w) = 2\left(|\tau|^{-2} + 2\,|\tau|^{-w}\right)|V|$$

with $|\tau| = \sqrt{q}$. When p is even, we have

$$|V| = \frac{1}{2}\sqrt{1 + q - \frac{p^2}{4}},$$

and when p is odd, we have

$$|V| = \frac{1}{2}\left(q - \frac{p^2}{4}\right)^{-1/2}\left(q - \frac{p^2}{4} + \frac{1}{4}\right).$$

Using monotonicity arguments as in the proof of Lemma 7.1 yields the list Y of "critical points".

8. The *p*-is-3-*q*-is-3-Case

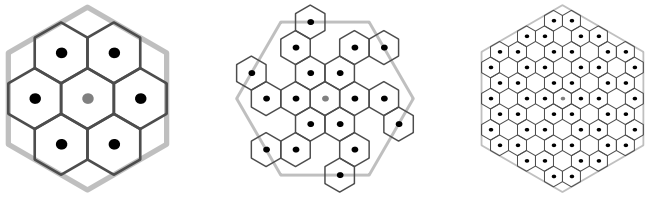
One important case can be proved by using the Optimality Theorem of Section 3, too, namely when τ comes from a Koblitz curve in characteristic 3. We specialise the setting of Section 7 to $p = 3\mu$ with $\mu \in \{-1, 1\}$ and $q = 3$. We continue looking at w-NAF-numeral systems with minimal norm representative digit set modulo τ^w with $w \geq 2$. Some examples of those digit sets are shown in Figure 8.1. We have the following optimality result.

Corollary 8.1. *With the setting above, the width-w non-adjacent form expansion for each element of* $\mathbb{Z}[\tau]$ *is optimal.*

Proof. Using the statement of Lemma 7.1 and Theorem 3.2 yields the optimality for all $w \geq 4$.

Let $w = 2$. Then our minimal norm representatives digit set is

$$\mathcal{D} = \{0\} \cup \bigcup_{0 \leq k < 6} \zeta^k \{1\},$$



(a) Digit set for $\tau = \frac{3}{2} + \frac{i}{2}\sqrt{3}$ and $w = 2$. (b) Digit set for $\tau = \frac{3}{2} + \frac{i}{2}\sqrt{3}$ and $w = 3$. (c) Digit set for $\tau = \frac{3}{2} + \frac{i}{2}\sqrt{3}$ and $w = 4$.

FIGURE 8.1. Minimal norm representatives digit sets modulo τ^w. For each digit η, the corresponding Voronoi cell V_η is drawn. The large scaled Voronoi cell is $\tau^w V$.

where ζ is a primitive sixth root of unity, see Avanzi, Heuberger and Prodinger [3]. Therefore we obtain $|\mathcal{D}| = 1$ and $\mathcal{D} = -\mathcal{D}$. For $k \in \{0, 1\}$ we get

$$\tau^k\mathcal{D} + \mathcal{D} + \mathcal{D} \subseteq \overline{\mathcal{B}}\Big(0, \sqrt{3}+2\Big) \subseteq \sqrt{3}^4\, \mathcal{B}\Big(0, \frac{1}{2}\Big) \subseteq \tau^{2w} \operatorname{int}(V)\,,$$

so the digit set $\mathcal{D}$ is w-subadditive by the same arguments as in the beginning of the proof of Lemma 7.1, and we can apply the Optimality Theorem to get the desired result.

Let $w = 3$. Then our minimal norm representatives digit set is

$$\mathcal{D} = \{0\} \cup \bigcup_{0 \le k < 6} \zeta^k \{1, 2, 4 - \mu\tau\}\,,$$

where ζ is again a primitive sixth root of unity, again [3]. Therefore we obtain $|\mathcal{D}| = |4 - \mu\tau| = \sqrt{7}$ and again $\mathcal{D} = -\mathcal{D}$. For $k \in \{0, 1, 2\}$, we get $|\tau|^k\, |\mathcal{D}| \le 3\sqrt{7}$, because $|\tau| = \sqrt{3}$. Therefore

$$\tau^k\mathcal{D} + \mathcal{D} + \mathcal{D} \subseteq \overline{\mathcal{B}}\Big(0, 5\sqrt{7}\Big) \subseteq \sqrt{3}^6\, \mathcal{B}\Big(0, \frac{1}{2}\Big) \subseteq \tau^{2w} \operatorname{int}(V)\,,$$

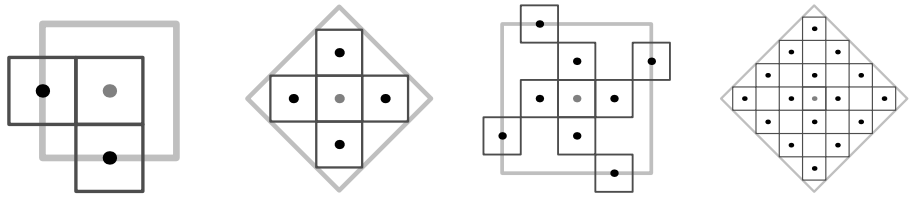
so we can use Theorem 3.2 again to get the optimality. □

9. The *p*-is-2-*q*-is-2-Case

In this section we look at another special base τ. We assume that $p \in \{-2, 2\}$ and $q = 2$. Again, we continue looking at w-NAF-numeral systems with minimal norm representative digit set modulo τ^w with $w \ge 2$. Some examples of those digit sets are shown in Figure 9.1.

For all possible τ of this section, the corresponding Voronoi cell can be written explicitly as

$$V = \mathsf{polygon}\Big(\Big\{\tfrac{1}{2}(1+i), \tfrac{1}{2}(-1+i), \tfrac{1}{2}(-1-i), \tfrac{1}{2}(1-i)\Big\}\Big)\,.$$



(a) Digit set for $\tau = 1 + i$ and $w = 2$.

(b) Digit set for $\tau = 1 + i$ and $w = 3$.

(c) Digit set for $\tau = 1 + i$ and $w = 4$.

(d) Digit set for $\tau = 1 + i$ and $w = 5$.

FIGURE 9.1. Minimal norm representatives digit sets modulo τ^w. For each digit η, the corresponding Voronoi cell V_η is drawn. The large scaled Voronoi cell is $\tau^w V$.

Remark that V is an axis-parallel square and that we have

$$\tau V = \mathsf{polygon}\Big(\Big\{ i^j \;\Big|\; j \in \{0,1,2,3\} \Big\}\Big).$$

In this section we will prove that the w-NAFs are optimal if and only if w is odd. The first part, optimality for odd w, is written down as the theorem below. The non-optimality part for even w can be found as Proposition 9.3.

Theorem 9.1. *Let w be an odd integer with $w \geq 3$, and let $z \in \mathbb{Z}[\tau]$. Then the width-w non-adjacent form expansion of z is optimal.*

Remark 9.2. Let w be an odd integer with $w \geq 3$. Let $z \in \tau^w V \cap \mathbb{Z}[\tau]$, then z can be represented as a w-NAF expansion with weight at most 1. To see this, consider the boundary of $\tau^w V$. Its vertices are $2^{(w-1)/2} i^m$ for $m \in \{0,1,2,3\}$. All elements of $\partial(\tau^w V) \cap \mathbb{Z}[\tau]$ can be written as $2^{(w-1)/2} i^m + k(1+i)i^n$ for some integers k, m and n. Further, all those elements are divisible by τ. Therefore each digit lies in the interior of $\tau^w V$, and for each $z \in \tau^w V \cap \mathbb{Z}[\tau]$ there is an integer $\ell \geq 0$ such that $\tau^{-\ell} z \in \mathcal{D}$, because $\tau^{-1} V \subseteq V$ and $|\tau| > 1$.

Proof of Theorem 9.1. We prove that the digit set $\mathcal{D}$ is w-subadditive. Optimality then follows using Theorem 3.2. Using the remark above, $\mathcal{D} = -\mathcal{D}$ and the ideas of Proposition 3.3, it is sufficient to show

$$\tau^{-w} \left(\tau^k \mathcal{D} + \mathcal{D} + \mathcal{D} \right) \cap \mathbb{Z}[\tau] \subseteq \tau^w V$$

for $k \in \{0, \ldots, w-1\}$.

Let $k = w - 1$. We show that

$$\left(\mathcal{D} + \tau^{-(w-1)} (\mathcal{D} + \mathcal{D}) \right) \cap \tau \mathbb{Z}[\tau] \subseteq \tau^{w+1} V. \tag{9.1}$$

So let $y = b + a$ be an element of the left hand side of (9.1) with $b \in \mathcal{D}$ and $a \in \tau^{-(w-1)} (\mathcal{D} + \mathcal{D})$. We can assume $y \neq 0$. Since $y \in \mathbb{Z}[\tau]$ and $\mathcal{D} \subseteq \mathbb{Z}[\tau]$,

we have $a \in \mathbb{Z}[\tau]$. Since $\mathcal{D} \subseteq \tau^w V$, we obtain

$$\tau^{-(w-1)} (\mathcal{D} + \mathcal{D}) \subseteq 2\tau V.$$

The case $b = 0$ is easy, because $2\tau V = \tau^3 V \subseteq \tau^w V$. So we can assume $b \neq 0$. This means $\tau \nmid b$. Since $\tau \mid y$, we have $\tau \nmid a$. The set $2\tau V \cap \mathbb{Z}[\tau]$ consists exactly of 0, i^m, $2i^m$ and τi^m for $m \in \{0, 1, 2, 3\}$. The only elements in that set not divisible by τ are the i^m. Therefore $a = i^m$ for some m. The digit b is in the interior of $\tau^w V$, thus $y = b + a$ is in $\tau^w V \subseteq \tau^{w+1} V$.

Now let $k \in \{0, \dots, w-2\}$. If $w \geq 5$, then

$$\tau^{-w} \left(\tau^k \mathcal{D} + \mathcal{D} + \mathcal{D}\right) \subseteq \tau^{w-2} V + 2V,$$

using $\mathcal{D} \subseteq \tau^w V$ and properties of the Voronoi cell V. Consider the two squares $\tau^{w-2} V$ and $\tau^w V = 2\tau^{w-2} V$. The distance between the boundaries of them is at least $\frac{1}{2} |\tau|^{w-2}$, which is at least $\sqrt{2}$. Since $2V$ is contained in a disc with radius $\sqrt{2}$, we obtain $\tau^{w-2} V + 2V \subseteq \tau^w V$.

We are left with the case $w = 3$ and $k \in \{0, 1\}$. There the digit set $\mathcal{D}$ consists of 0 and i^m for $m \in \{0, 1, 2, 3\}$. Therefore we have $\mathcal{D} \subseteq \tau V$ (instead of $\mathcal{D} \subseteq \tau^3 V$). By the same arguments as in the previous paragraph we get

$$\tau^{-3} \left(\tau^k \mathcal{D} + \mathcal{D} + \mathcal{D}\right) \subseteq \tfrac{1}{2} (\tau V + 2V) \subseteq \tau^3 V,$$

so the proof is complete. □

The next result is the non-optimality result for even w.

Proposition 9.3. *Let w be an even integer with $w \geq 2$. Then there is an element of $\mathbb{Z}[\tau]$ whose w-NAF-expansion is non-optimal.*

Again, some examples of the digit sets used are shown in Figure 9.1. The proof of the proposition is split up: Lemma 9.4 handles the general case for even $w \geq 4$ and Lemma 9.5 gives a counter-example (to optimality) for $w = 2$.

For the remaining section—it contains the proof of Proposition 9.3—we will assume $\tau = 1 + i$. All other cases are analogous.

Lemma 9.4. *Let the assumptions of Proposition 9.3 hold and suppose $w \geq 4$. Define $A := |\tau|^w \frac{1}{2}(1 - i)$ and $B := \frac{1}{\tau} A$ and set $s = -i^{1-w/2}$. Then*

(a) 1, i, -1 and $-i$ are digits,
(b) $A - 1$ is a digit,
(c) $-B - 1$ is a digit,
(d) $i\tau^{w-1} - s^{-1}$ is a digit, and
(e) we have

$$(A-1)\tau^{w-1} + (-s^{-1}) = s\tau^{2w} + (-B-1)\tau^w + (i\tau^{w-1} - s^{-1}).$$

Figure 9.2 shows the digits used in Lemma 9.4 for a special configuration.

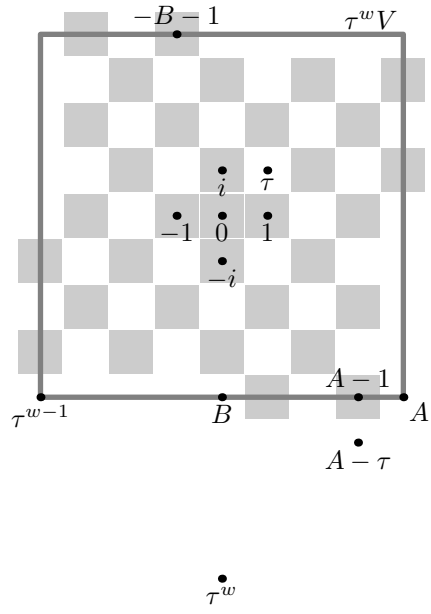


FIGURE 9.2. The w-is-even situation. The figure shows the configuration $p = 2$, $q = 2$, $w = 6$, $s = 1$. A polygon filled grey represents a digit, a dot represents a point of interest in Lemma 9.4.

Proof. (a) A direct calculation shows that the lattice elements 1, i, -1 and $-i$ are in the interior of

$$\tau^w V = \left[-2^{w/2-1}, 2^{w/2-1}\right] + \left[-2^{w/2-1}, 2^{w/2-1}\right] i$$

and are not divisible by τ. So all of them are digits.

(b) We can rewrite A as

$$A = 2^{w/2-1}(1-i) = -2^{w/2-1} i\tau,$$

therefore $\tau^2 \mid A$. We remark that A is a vertex (the lower-right vertex) of the scaled Voronoi cell $\tau^w V$ and that the edges of $\tau^w V$ are parallel to the real and imaginary axes. This means that $A-1$ is on the boundary, too, and its real part is larger than 0. By using the construction of the restricted Voronoi cell, cf. Definition 5.2, we know that $A-1$ is in $\tau^w \widetilde{V}$. Since it is clearly not divisible by τ, it is a digit.

(c) We have

$$B = \tfrac{1}{\tau} A = -2^{w/2-1} i.$$

Therefore $\tau \mid B$, and we know that B halves the edge at the bottom of the Voronoi cell $\tau^w V$. By construction of the scaled restricted Voronoi

cell $\tau^w\widetilde{V}$, cf. Definition 5.2, we obtain that $B+1$ is a digit, and therefore, by symmetry, $-B-1$ is a digit, too.

(d) Rewriting yields

$$i\tau^{w-1} - s^{-1} = s^{-1}(is\tau^{w-1} - 1),$$

and we obtain

$$s\tau^w = -i^{1-w/2}(1+i)^w = -2^{w/2}i,$$

since $(1+i)^2 = 2i$. Further we can check that the vertices of $\tau^w V$ are $i^k\tau^{w-1}$ for an appropriate $k \in \mathbb{Z}$.

Now consider $is\tau^{w-1}$. This is exactly the lower-right vertex A of $\tau^w V$. Therefore, we have

$$i\tau^{w-1} - s^{-1} = s^{-1}(A-1).$$

Using that $A-1$ is a digit and the rotational symmetry of the restricted Voronoi cell, $i\tau^{w-1} - s^{-1}$ is a digit.

(e) As before, we remark that $s\tau^w = -2^{w/2}i$. Therefore we obtain

$$B - 1 - s\tau^w = -B - 1.$$

Now, by rewriting, we get

$$\begin{aligned}(A-1)\tau^{w-1} + (-s^{-1}) &= (A-\tau)\tau^{w-1} + (i\tau^{w-1} - s^{-1})\\ &= (B-1)\tau^w + (i\tau^{w-1} - s^{-1})\\ &= s\tau^{2w} + (B - 1 - s\tau^w)\tau^w + (i\tau^{w-1} - s^{-1})\\ &= s\tau^{2w} + (-B-1)\tau^w + (i\tau^{w-1} - s^{-1}),\end{aligned}$$

which was to prove. □

Lemma 9.5. *Let the assumptions of Proposition 9.3 hold and suppose $w = 2$. Then*

(a) -1 and $-i$ are digits and

(b) we have

$$-\tau - 1 = -i\tau^6 - \tau^4 - i\tau^2 - i.$$

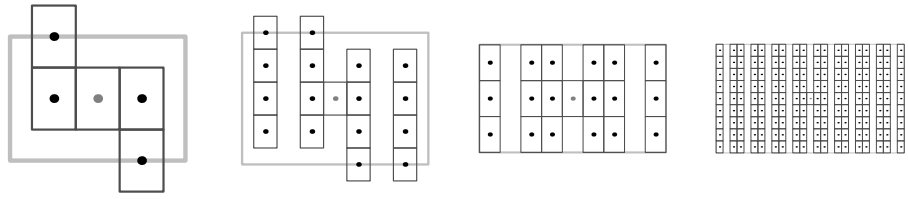
Proof. (a) The elements -1 and $-i$ are on the boundary of the Voronoi cell $\tau^2 V$, cf. Figure 9.1(a). More precisely, each is halving an edge of the Voronoi cell mentioned. The construction of the restricted Voronoi cell, together with the rotation and scaling of $\tau^2 = 2i$, implies that -1 and $-i$ are in $\tau^2\widetilde{V}$. Since none of them is divisible by τ, both are digits.

(b) The element i has the 2-NAF-representation

$$i = -i\tau^4 - \tau^2 - i.$$

Therefore we obtain

$$-\tau - 1 = (-1+i)\tau + (i\tau - 1) = i\tau^2 + (-i) = -i\tau^6 - \tau^4 - i\tau^2 - i$$



(a) Digit set for $\tau = i\sqrt{2}$ and $w = 3$. (b) Digit set for $\tau = i\sqrt{2}$ and $w = 5$. (c) Digit set for $\tau = i\sqrt{3}$ and $w = 3$. (d) Digit set for $\tau = i\sqrt{3}$ and $w = 5$.

FIGURE 10.1. Minimal norm representatives digit sets modulo τ^w. For each digit η, the corresponding Voronoi cell V_η is drawn. The large scaled Voronoi cell is $\tau^w V$.

as required. □

Finally, we are able to prove the non-optimality result.

Proof of Proposition 9.3. Let $w \geq 4$. Everything needed can be found in Lemma 9.4: We have the equation

$$(A-1)\tau^{w-1} + (-s^{-1}) = s\tau^{2w} + (-B-1)\tau^w + (i\tau^{w-1} - s^{-1}),$$

in which the left and the right hand side are both valid expansions (the coefficients are digits). The left hand side has weight 2 and is not a w-NAF, whereas the right hand side has weight 3 and is a w-NAF.

Similarly the case $w = 2$ is shown in Lemma 9.5: We have the equation

$$-\tau - 1 = -i\tau^6 - \tau^4 - i\tau^2 - i,$$

which again is a counter-example to the optimality of the 2-NAFs. □

10. The *p*-is-0-Case

This section contains another special base τ. We assume that $p = 0$ and that we have an integer $q \geq 2$. Again, we continue looking at w-NAF-numeral systems with minimal norm representative digit set modulo τ^w with $w \geq 2$. Some examples of the digit sets used are shown in Figure 10.1.

For all possible τ of this section, the corresponding Voronoi cell can be written explicitly as

$$V = \mathsf{polygon}\Big(\Big\{\tfrac{1}{2}(\tau+1), \tfrac{1}{2}(\tau-1), \tfrac{1}{2}(-\tau-1), \tfrac{1}{2}(-\tau+1)\Big\}\Big).$$

Remark that V is an axis-parallel rectangle.

In this section we prove the following non-optimality result.

Proposition 10.1. *Let w be an odd integer with $w \geq 3$ and the setting as above. Then there is an element of $\mathbb{Z}[\tau]$ whose w-NAF-expansion is non-optimal.*

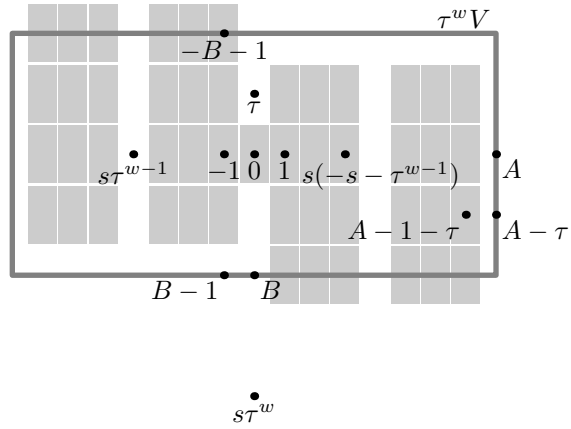


FIGURE 10.2. The q-is-even situation. The figure shows the configuration $p = 0$, $q = 4$, $\tau = 2i$, $w = 3$, $s = 1$. A polygon filled grey represents a digit, a dot represents a point of interest in Lemma 10.2.

For the remaining section, which contains the proof of the proposition above, we will assume $\tau = i\sqrt{q}$. The case $\tau = -i\sqrt{q}$ is analogous. Before we start with the proof of Proposition 10.1, we need the following two lemmata.

Lemma 10.2. *Let the assumptions of Proposition 10.1 hold, and suppose that q is even. Define $A := \frac{1}{2}|\tau|^{w+1}$ and $B := \frac{1}{\tau}A$, and set $s = (-1)^{\frac{1}{2}(w+1)}$. Then*

(a) 1 and -1 are digits,
(b) $A - 1 - \tau$ is a digit,
(c) $-B - 1$ is a digit,
(d) $-s - \tau^{w-1}$ is a digit, and
(e) we have

$$(A - 1 - \tau)\tau^{w-1} - s = s\tau^{2w} + (-B - 1)\tau^w + (-s - \tau^{w-1}).$$

Figure 10.2 shows the digits used in Lemma 10.2 for a special configuration.

Proof. (a) A direct calculation shows that -1 and 1 are in an open disc with radius $\frac{1}{2}|\tau|^w$, which itself is contained in $\tau^w V$. Both are not divisible by τ, so both are digits.

(b) Because w is odd, q is even and $\tau = i\sqrt{q}$, we can rewrite the point A as

$$A = \frac{1}{2}|\tau|^{w+1} = \frac{q}{2}q^{\frac{1}{2}(w-1)}$$

and see that A is a (positive) rational integer and that $\tau^{w-1} \mid A$. Furthermore, A halves an edge of $\tau^w V$. Therefore, $A - 1$ is inside $\tau^w V$.

If $q \geq 4$ or $w \geq 5$, the point $A - 1 - \tau$ is inside $\tau^w V$, too, since the vertical (parallel to the imaginary axis) side-length of $\tau^w V$ is $|\tau|^w$ and $|\tau| < \frac{1}{2}|\tau|^w$. Since $\tau^2 \mid A$, we obtain $\tau \nmid A - 1 - \tau$, so $A - 1 - \tau$ is a digit. If $q = 2$ and $w = 3$, we have $A - 1 - \tau = 1 - \tau$. Due to the definition of the restricted Voronoi cell $\widetilde{V}$, cf. Definition 5.2, we obtain that $1 - \tau$ is a digit.

(c) Previously we saw $\tau^{w-1} \mid A$. Using the definition of B and $w \geq 3$ yields $\tau \mid B$. It is easy to check that $B = \frac{1}{2}s\tau^w$. Furthermore, we see that B is on the boundary of the Voronoi cell $\tau^w V$. By a symmetry argument we get the same results for $-B$. By the construction of the restricted Voronoi cell $\widetilde{V}$, cf. Definition 5.2, we obtain that $-B - 1$ is in $\tau^w \widetilde{V}$ and since clearly $\tau \nmid (-B - 1)$, we get that $-B - 1$ is a digit.

(d) We first remark that $\tau^{w-1} \in \mathbb{Z}$ and that $\left|\tau^{w-1}\right| \leq A$. Even more, we get $0 < -s\tau^{w-1} \leq A$. Since A is on the boundary of $\tau^w V$, we obtain $-1 - s\tau^{w-1} \in \tau^w \operatorname{int}(V)$. By symmetry the result is true for $-s - \tau^{w-1}$ and clearly $\tau \nmid (-s - \tau^{w-1})$, so $-s - \tau^{w-1}$ is a digit.

(e) We get

$$\begin{aligned}(A - 1 - \tau)\tau^{w-1} + (-s) &= (A - \tau)\tau^{w-1} + (-s - \tau^{w-1}) \\ &= (B - 1)\tau^w + (-s - \tau^{w-1}) \\ &= s\tau^{2w} + (B - 1 - s\tau^w)\tau^w + (-s - \tau^{w-1}) \\ &= s\tau^{2w} + (-B - 1)\tau^w + (-s - \tau^{w-1}),\end{aligned}$$

which can easily be verified. We used $B = \frac{1}{\tau}A$. □

Lemma 10.3. *Let the assumptions of Proposition 10.1 hold, and suppose that q is odd. Define $A' := \frac{1}{2}|\tau|^{w+1}$, $B' := \frac{1}{\tau}A$, $A := A' - \frac{1}{2}$ and $B := B' + \frac{\tau}{2}$, and set $C = -A$, $t = (q+1)/2$ and $s = (-1)^{\frac{1}{2}(w+1)} \in \{-1, 1\}$. Then*

(a) 1 and -1 are digits,
(b) $A - \tau$ is a digit,
(c) sC is a digit,
(d) $-B - 1$ is a digit,
(e) $sC - t\tau^{w-1}$ is a digit, and
(f) we have

$$(A - \tau)\tau^{w-1} + (sC) = s\tau^{2w} + (-B - 1)\tau^w + (sC - t\tau^{w-1}).$$

Figure 10.3 shows the digits used in Lemma 10.3 for a special configuration.

Proof. (a) See the proof of Lemma 10.2.

(b) We can rewrite the point A as

$$A = \frac{1}{2}|\tau|^{w+1} - \frac{1}{2} = \frac{1}{2}\left(q^{\frac{1}{2}(w+1)} - 1\right).$$

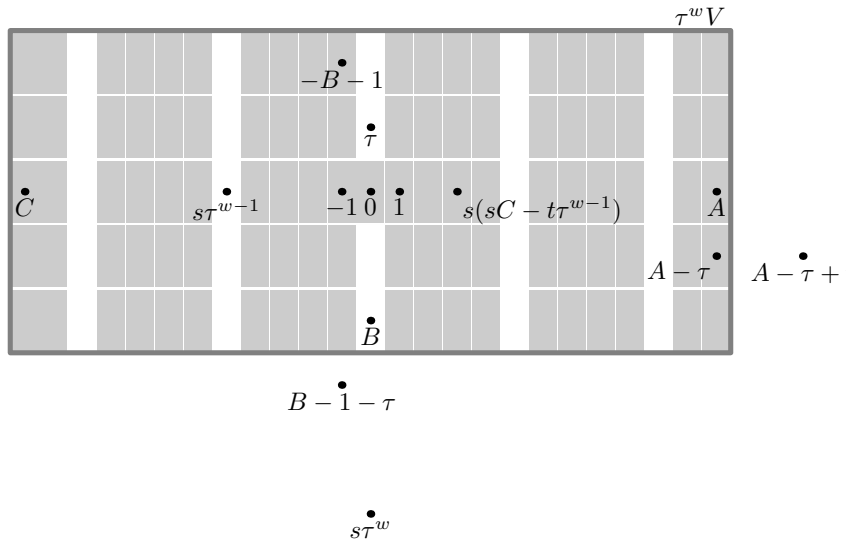


FIGURE 10.3. The q-is-odd situation. The figure shows the configuration $p=0$, $q=5$, $\tau=i\sqrt{5}$, $w=3$, $s=1$. A polygon filled grey represents a digit, a dot represents a points of interest in Lemma 10.3.

Since q is odd with $q \geq 3$ and w is odd with $w \geq 3$, we obtain $A \in \mathbb{Z}$ with $0 < A < \frac{1}{2}|\tau|^{w+1}$ and $q \nmid A$. Therefore $\tau \nmid A$ and A is in the interior of the Voronoi cell $\tau^w V$. The vertical (parallel to the imaginary axis) side-length of $\tau^w V$ is $|\tau|^w$ and $|\tau| < \frac{1}{2}|\tau|^w$, so $A-\tau$ is in the interior of $\tau^w V$, too. Since $\tau \nmid A-\tau$, the element $A-\tau$ is a digit.

(c) We got $\tau \nmid A$ and A is in the interior of the Voronoi cell $\tau^w V$. Therefore A is a digit, and—by symmetry—sC is a digit, too.

(d) We obtain

$$B = -\frac{1}{2} i\sqrt{q}\,|\tau|^{w-1} + i\frac{1}{2}\sqrt{q} = \frac{1}{2}\tau\left(-|\tau|^{w-1}+1\right),$$

which is inside $\tau^w V$. Therefore the same is true for $-B$. The horizontal (parallel to the real axis) side-length of $\tau^w V$ is larger than 2, therefore $-B-1$ is inside $\tau^w V$, too. Since $\tau \mid B$ we get $\tau \nmid (-B-1)$, so $-B-1$ is a digit.

(e) We obtain

$$\begin{aligned} 0 < s\left(sC - t\tau^{w-1}\right) &= \frac{1}{2}\left((q+1)|\tau|^{w-1} - |\tau|^{w+1} + 1\right) \\ &= \frac{1}{2}\left(|\tau|^{w-1}+1\right) < \frac{1}{2}|\tau|^{w+1}. \end{aligned}$$

This means that $sC - t\tau^{w-1}$ is in the interior of the Voronoi cell $\tau^w V$. Since $\tau \nmid (-A) = C$, the same is true for $sC - t\tau^{w-1}$, i.e. it is a digit.

(f) We get

$$\begin{aligned}(A-\tau)\tau^{w-1} + (sC) &= (A - \tau + t)\tau^{w-1} + (sC - t\tau^{w-1}) \\ &= (B - 1 - \tau)\tau^w + (sC - t\tau^{w-1}) \\ &= s\tau^{2w} + (B - 1 - \tau - s\tau^w)\tau^w + (sC - t\tau^{w-1}) \\ &= s\tau^{2w} + (-B - 1)\tau^w + (sC - t\tau^{w-1}),\end{aligned}$$

which can be checked easily. □

The two lemmata above now allow us to prove the non-optimality result of this section.

Proof of Proposition 10.1. Let q be even. In Lemma 10.2 we got

$$(A - 1 - \tau)\tau^{w-1} - s = s\tau^{2w} + (-B - 1)\tau^w + (-s - \tau^{w-1})$$

and that all the coefficients there were digits, i.e. we have valid expansions on the left and right hand side. The left hand side has weight 2 and is not a w-NAF, whereas the right hand side has weight 3 and is a w-NAF. Therefore a counter-example to the optimality was found.

The case q is odd works analogously. We got the counter-example

$$(A - \tau)\tau^{w-1} + (sC) = s\tau^{2w} + (-B - 1)\tau^w + (sC - t\tau^{w-1})$$

in Lemma 10.3. □

11. Computational Results

This section contains computational results on the optimality of w-NAFs for some special imaginary quadratic bases τ and integers w. We assume that we have a τ coming from integers p and q with $q > p^2/4$. Again, we continue looking at w-NAF-numeral systems with minimal norm representative digit set modulo τ^w with $w \geq 2$.

As mentioned in Section 3, the condition w-subadditivity-condition—and therefore optimality—can be verified by finding a w-NAF-expansion with weight at most 2 in $w\,(\#\mathcal{D} - 1)$ cases. The computational results can be found in Figure 11.1. The calculations were performed in Sage [27].

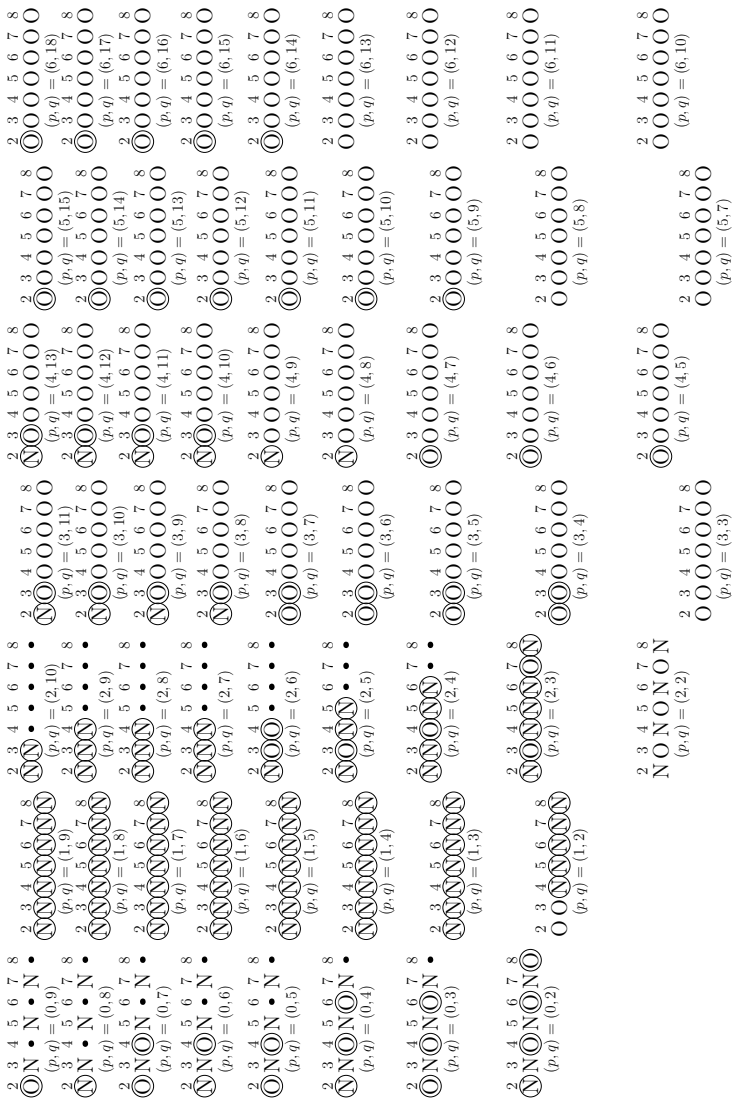


FIGURE 11.1. The optimality map including computational results. Below each block the parameters $|p|$ and q (fulfilling $\tau^2 - p\tau + q = 0$) are printed. A block is positioned according to τ in the complex plane. Above each block are the w. The symbol O means that the w-NAF-expansions are optimal, N means there are non-optimal w-NAF-expansions. If a result is surrounded by a circle, then it is a computational result. Otherwise, if there is no circle, then the result comes from a theorem given here or was already known. A dot means that there is no result available.

References

[1] Roberto Avanzi, Clemens Heuberger, and Helmut Prodinger, *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 (B. Preneel and S. Tavares, eds.), Lecture Notes in Comput. Sci., vol. 3897, Springer, Berlin, 2006, pp. 332–344.

[2] ———, *Scalar multiplication on Koblitz curves. Using the Frobenius endomorphism and its combination with point halving: Extensions and mathematical analysis.* Algorithmica **46** (2006), 249–270.

[3] ———, *Arithmetic of supersingular Koblitz curves in characteristic three.* Tech. Report 2010-8, Graz University of Technology, 2010, `http://www.math.tugraz.at/fosp/pdfs/tugraz_0166.pdf`, also available as Cryptology ePrint Archive, Report 2010/436, `http://eprint.iacr.org/`.

[4] Roberto Maria 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 (H. Handschuh and A. Hasan, eds.), Lecture Notes in Comput. Sci., vol. 3357, Springer-Verlag, Berlin, 2004, pp. 130–143.

[5] Ian F. Blake, V. Kumar Murty, and Guangwu Xu, *Efficient algorithms for Koblitz curves over fields of characteristic three.* J. Discrete Algorithms **3** (2005), no. 1, 113–124.

[6] ———, *A note on window τ-NAF algorithm.* Inform. Process. Lett. **95** (2005), 496–502.

[7] ———, *Nonadjacent radix-τ expansions of integers in Euclidean imaginary quadratic number fields.* Canad. J. Math. **60** (2008), no. 6, 1267–1282.

[8] László Germán and Attila Kovács, *On number system constructions.* Acta Math. Hungar. **115** (2007), no. 1-2, 155–167.

[9] Daniel M. Gordon, *A survey of fast exponentiation methods.* J. Algorithms **27** (1998), 129–146.

[10] Clemens Heuberger, *Redundant τ-adic expansions II: Non-optimality and chaotic behaviour.* Math. Comput. Sci. **3** (2010), 141–157.

[11] Clemens Heuberger and Daniel Krenn, *Analysis of width-w non-adjacent forms to imaginary quadratic bases.* J. Number Theory **133** (2013), 1752–1808.

[12] Clemens Heuberger and Helmut Prodinger, *Analysis of alternative digit sets for non-adjacent representations.* Monatsh. Math. **147** (2006), 219–248.

[13] Thomas W. Hungerford, *Algebra.* Graduate Texts in Mathematics, vol. 73, Springer, 1996.

[14] Jonathan Jedwab and Chris J. Mitchell, *Minimum weight modified signed-digit representations and fast exponentiation.* Electron. Lett. **25** (1989), 1171–1172.

[15] Donald E. Knuth, *Seminumerical algorithms.* third ed., The Art of Computer Programming, vol. 2, Addison-Wesley, 1998.

[16] Neal Koblitz, *CM-curves with good cryptographic properties.* Advances in cryptology—CRYPTO '91 (Santa Barbara, CA, 1991) (J. Feigenbaum, ed.), Lecture Notes in Comput. Sci., vol. 576, Springer, Berlin, 1992, pp. 279–287.

[17] ———, *An elliptic curve implementation of the finite field digital signature algorithm.* Advances in cryptology—CRYPTO '98 (Santa Barbara, CA, 1998), Lecture Notes in Comput. Sci., vol. 1462, Springer, Berlin, 1998, pp. 327–337.

[18] Markus Kröll, *Optimality of digital expansions to the base of the Frobenius endomorphism on Koblitz curves in characteristic three.* Tech. Report 2010-09, Graz University of Technology, 2010, available at `http://www.math.tugraz.at/fosp/pdfs/tugraz_0167.pdf`.

[19] M. Lothaire, *Algebraic combinatorics on words.* Encyclopedia of Mathematics and its Applications, vol. 90, Cambridge University Press, Cambridge, 2002.

[20] Willi Meier and Othmar Staffelbach, *Efficient multiplication on certain nonsupersingular elliptic curves.* Advances in cryptology—CRYPTO '92 (Santa Barbara, CA, 1992) (Ernest F. Brickell, ed.), Lecture Notes in Comput. Sci., vol. 740, Springer, Berlin, 1993, pp. 333–344.

[21] James A. Muir and Douglas 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.

[22] ———, *Minimality and other properties of the width-w nonadjacent form*. Math. Comp. **75** (2006), 369–384.

[23] Braden Phillips and Neil Burgess, *Minimal weight digit set conversions*. IEEE Trans. Comput. **53** (2004), 666–677.

[24] George W. Reitwiesner, *Binary arithmetic*. Advances in computers, vol. 1, Academic Press, New York, 1960, pp. 231–308.

[25] Jerome 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.

[26] ———, *Efficient arithmetic on Koblitz curves*. Des. Codes Cryptogr. **19** (2000), 195–249.

[27] William A. Stein et al., *Sage Mathematics Software (Version 4.7.1)*. The Sage Development Team, 2011, `http://www.sagemath.org`.

[28] Christiaan van de Woestijne, *The structure of Abelian groups supporting a number system (extended abstract)*. Actes des rencontres du CIRM **1** (2009), no. 1, 75–79.

Clemens Heuberger
Institute of Mathematics
Alpen-Adria-Universität Klagenfurt
Universitätsstraße 65–67, 9020 Klagenfurt am Wörthersee, Austria
E-mail: `clemens.heuberger@aau.at`

Daniel Krenn
Institute of Optimisation and Discrete Mathematics (Math B)
Graz University of Technology
Steyrergasse 30/II, 8010 Graz, Austria
E-mail: `math@danielkrenn.at`
E-mail: `krenn@math.tugraz.at`