Daniel KRENN

Analysis of Digital Expansions to Imaginary Quadratic Bases

MASTER'S THESIS

written to obtain the academic degree of a Master of Science (MSc)

Master's degree Mathematical Computer Science

Graz University of Technology
Graz University of Technology

Supervisor:
Ao.Univ.-Prof. Dipl.-Ing. Dr.techn. Clemens HEUBERGER

Institute of Optimisation and Discrete Mathematics (Math B)

Graz, September 2010

Statutory Declaration

I declare that I have authored this thesis independently, that I have not used other than the declared sources/resources, and that I have explicitly marked all material which has been quoted either literally or by content from the used sources.

....
date (signature)

Abstract

Elliptic curves over finite fields can be used in public-key cryptography. There, the scalar multiplication in the group of rational points on the curve is the essential operation performed, and clearly, the aim is to make it as efficient as possible. Beside the double-and-add methods, Frobenius-and-add algorithms are attractive, since the Frobenius endomorphism can be evaluated very fast in finite fields. Due to the correspondence between the Frobenius endomorphism and an algebraic integer τ, we may consider τ-adic expansions for elements of $\mathbb{Z}[\tau]$.

Let w be an integer with $w \geq 2$, and let the digit set consist of zero and all minimal norm representatives modulo τ^w not divisible by τ. We consider width-w τ-adic non-adjacent forms (w-NAFs for short). This means that in an expansion with base τ every block of w consecutive digits contains at most one non-zero digit. This thesis deals with analysing the occurrences of a fixed non-zero digit in such expansions. The major result counts these occurrences in all w-NAFs in a region (e.g. disc) asymptotically. The theorem is proved for imaginary quadratic algebraic integers τ. Beside the main term, a second order term, which is periodically oscillating, is given. Further the necessary tools and prerequisites were developed, and it is shown that every element of $\mathbb{Z}[\tau]$ admits a unique w-NAF. Moreover, some properties of the fundamental domain of a w-NAF number systems can be found.

keywords: τ-adic expansions, non-adjacent forms, redundant digit sets, elliptic curve cryptography, Koblitz curves, Frobenius endomorphism, scalar multiplication, Hamming weight, sum of digits, fractals, fundamental domain

contact:
Daniel Krenn
e-mail: `mail@danielkrenn.at` or `daniel.krenn@tugraz.at`
web: `www.danielkrenn.at`

Kurzfassung

In asymmetrischen Kryptosystemen (Public-Key-Verfahren) können elliptische Kurven über endlichen Körpern verwendet werden. Dabei ist die Skalarmultiplikation in der Punktgruppe der Kurve von besonderem Interesse. Klarerweise ist ein Ziel, diese möglichst effizient auszuführen. Neben den Double-and-Add Verfahren können auch Frobenius-and-add Algorithmen eingesetzt werden, da der Frobenius Endomorphismus in endlichen Körpern sehr schnell auszuführen ist. Wegen des Zusammenhangs des Frobenius Endomorphismus mit einer ganzalgebraischen Zahl τ ist die Betrachtung von τ-adischen Entwicklungen von Elementen in $\mathbb{Z}[\tau]$ interessant.

Sei w eine natürliche Zahl mit $w \geq 2$, und sei eine Ziffernmenge, bestehend aus Null und Repräsentanten modulo τ^w mit minimaler Norm und teilerfremd zu τ, gegeben. Wir betrachten die Width-w τ-adic non-adjacent Form (kurz w-NAF). Dabei ist in dieser τ-adischen Entwicklung in jedem Block von w aufeinanderfolgenden Ziffern maximal eine nicht-Null. In dieser Arbeit wird das Vorkommen einer fixen Ziffer ungleich Null in solchen Entwicklungen analysiert. Im Hauptresultat wird eben dieses Vorkommen einer Ziffer in allen w-NAFs in einem Gebiet (z.B. Kreisscheibe) gezählt. Der entsprechende Satz wurde dabei für imaginär-quadratische ganzalgebraische τ bewiesen. Das Ergebnis besteht neben dem Hauptterm auch aus einem periodisch oszillierenden Term zweiter Ordnung. Weiters wurden die dafür nötigen Hilfsmittel entwickelt und es wurde gezeigt, dass jedes Element von $\mathbb{Z}[\tau]$ eine eindeutige w-NAF-Entwicklung besitzt. Außerdem wurde der Fundamentalbereich des w-NAF Zahlensystems untersucht.

Schlüsselworter: τ-adische Entwicklung, Non-adjacent Form, redundante Ziffernmengen, elliptische Kurven Kryptografie, Koblitz-Kurve, Frobenius Endomorphismus, Skalarmultiplikation, Hamming-Gewicht, Summe von Ziffern, Fraktale, Fundamentalbereich

Kontakt:
Daniel Krenn
E-mail: `mail@danielkrenn.at` or `daniel.krenn@tugraz.at`
Web: `www.danielkrenn.at`

Contents

List of Figures

List of Tables

Chapter 1

Introduction

In cryptography one major area of study is *public-key cryptography*. The main idea is to use two different keys: one for encrypting messages and one for decrypting them. This is in contrast to *symmetric-key cryptography*, where the same key is used for encrypting and decrypting a message. One approach in public-key cryptography is elliptic curve cryptography. There the algebraic structure of an elliptic curve over a finite field is used. The essential operation performed is building multiples of a rational point on the elliptic curve. Clearly one goal is to make this *scalar multiplication* as efficient as possible.

The first part of this thesis, Chapter 2, contains some background on elliptic curve cryptography. All information there is already known and of course the references given. The chapter, in particular, describes different methods performing the mentioned scalar multiplication. More detailed, Section 2.1 starts with some facts about elliptic curves, in Section 2.2 the idea of elliptic curve cryptography is described, and Section 2.3 explains the used notations. The remaining sections of Chapter 2 provide information about two special methods used for the scalar multiplication operation. Those are described in the following paragraphs.

The basic idea in scalar multiplication behind those methods is to represent the scalar in an appropriate number system and to use a *Horner scheme* for evaluation. Using base 2 in such a number system leads to *double-and-add methods*, see Section 2.4. There, if we take 0 and 1 as digits, then we get the standard binary system. Clearly the resulting Horner scheme is more efficient to evaluate, when there are a lot of zeros inside a representation. So one may ask, if there are representations with fewer non-zero digits. Since the binary representation is unique, we cannot have such a representation with digits 0 and 1, but by extending the digit set and still using base 2, our representations become non-unique.

The first idea is to use the digit -1 additionally, because inversion of a point on the elliptic curve can be performed fast, see Section 2.4.1 for details. To avoid non-uniqueness in the resulting number system, the concept of non-adjacent forms can be used. There in every representation each two consecutive digits contain at most one non-zero digit; we call such numbers width-2 non-adjacent forms, or 2-NAFs for short. This idea can be generalised in several ways. One is to use another digit instead of -1, see Section 2.4.3. A different generalisation is to use a larger digit set and a generalisation of the non-adjacency condition, namely the width-w non-adjacent form, or w-NAF for short. There, in every representation each w consecutive digits contain at most one non-zero digit, see Section 2.4.2.

A different approach is to use more properties of the elliptic curves and the finite field over which the curves are defined. In particular, we want to use the Frobenius endomorphism, which

is very cheap to perform, especially when normal bases are used. Now consider the elliptic curve

$$\mathcal{E}_3 \colon Y^2 = X^3 - X - \mu \qquad \text{with } \mu \in \{-1, 1\},$$

called *Koblitz curve*, defined over $\mathbb{F}_3$. We are interested in the group $\mathcal{E}_3(\mathbb{F}_{3^m})$ of rational points over a field extension $\mathbb{F}_{3^m}$ of $\mathbb{F}_3$ for an $m \in \mathbb{N}$. The Frobenius endomorphism

$$\varphi \colon \mathcal{E}_3(\mathbb{F}_{3^m}) \longrightarrow \mathcal{E}_3(\mathbb{F}_{3^m}), \quad (x, y) \longmapsto \left(x^3, y^3\right)$$

satisfies the relation $\varphi^2 - 3\mu\varphi + 3 = 0$. So φ may be identified with the imaginary quadratic number $\tau = \frac{3}{2}\mu + \frac{1}{2}\sqrt{-3}$, which is a solution of the mentioned relation. Thus we have an isomorphism between $\mathbb{Z}[\tau]$ and the endomorphism ring of $\mathcal{E}_3(\mathbb{F}_{3^m})$.

Let $z \in \mathbb{Z}[\tau]$ and $P \in \mathcal{E}_3(\mathbb{F}_{3^m})$. If we write the element z as $\sum_{j=0}^{\ell-1} z_j \tau^j$ for some digits z_j belonging to a digit set $\mathcal{D}$, then we can compute the action zP as $\sum_{j=0}^{\ell-1} z_j \, \varphi^j(P)$ via a Horner scheme. The resulting Frobenius-and-add method, see Section 2.5, is much faster than the classic double-and-add scalar multiplication.

Another example is the elliptic curve

$$\mathcal{E}_2 \colon Y^2 + XY = X^3 + aX^2 + 1 \qquad \text{with } a \in \{0, 1\}$$

defined over $\mathbb{F}_2$, which is also called *Koblitz curve*. There we get the relation $\varphi^2 - \mu\varphi + 2 = 0$ with $\mu = (-1)^{1-a}$ for the Frobenius endomorphism φ, and thus $\tau = \frac{1}{2}\mu + \frac{1}{2}\sqrt{-7}$. More details of those two mentioned examples can be found in Section 2.5.1 and Section 2.5.3, respectively.

So, in general, let $\tau \in \mathbb{C}$ be an algebraic integer. We are interested in a τ-adic expansion for an element of $\mathbb{Z}[\tau]$ such that the mentioned computation of the action is as efficient as possible.

The fewer non-zero digits there are in an expansion, the faster the main loop of the Horner scheme can be calculated. But usually fewer non-zero coefficients means larger digit sets and thus a higher pre-computation effort. So for optimal performance, a balance between digit set size and number of non-zeros has to be found.

Again we will use the concept of width-w non-adjacent forms. This will allow us to get expansions with a low number of non-zero entries. As digit set we use zero and a minimal norm representative from each residue class modulo τ^w in $\mathbb{Z}[\tau]$ not divisible by τ. It is commonly known that such expansions, if they exist, are unique, whereas the existence is not known in general. For the τ corresponding to curves $\mathcal{E}_2$ and $\mathcal{E}_3$ and an integer $w \geq 2$ this was shown. This, as well as other properties in conjunction with those τ, can be found in Sections 2.5.1 to 2.5.4. The existence results for some other τ are listed in Section 2.5.5.

The next chapter, Chapter 3, contains new results. One is an existence result for all imaginary quadratic algebraic integers τ and all $w \geq 2$. As pointed out in the previous paragraph, this was only known for some special cases. As digit set the mentioned minimal norm representatives were used and we got that every element of $\mathbb{Z}[\tau]$ admits a unique w-NAF, see Section 3.6. Additionally a simple algorithm for calculating those expansions is given. Further we get that every element of $\mathbb{C}$ has a w-NAF-expansion of the form $\xi_{\ell-1} \ldots \xi_1 \xi_0 . \xi_{-1} \xi_{-2} \ldots$, where the right hand side of the τ-point is allowed to be of infinite length. In Section 3.7 we consider numbers of the form $0 . \xi_{-1} \xi_{-2} \ldots$. The set of all values of such numbers is called the fundamental domain $\mathcal{F}$. It will be shown that $\mathcal{F}$ is compact and its boundary has Hausdorff dimension smaller than 2. Further a characterisation of the boundary is given and also a tiling property with scaled versions of $\mathcal{F}$ for the complex plane. Additionally we can calculate the Lebesgue measure of the fundamental domain.

The main part of Chapter 3 deals with analysing the occurrences of a fixed non-zero digit η. It starts in Section 3.1 with analysing the occurrence of a digit in the case of the curve $\mathcal{E}_3$ and $w = 2$. This analysing of the rational integers results in a formula containing a main term and

a second order term in form of a periodic nowhere differentiable function. The proof is very similar to the proof in the case of a balanced ternary number system, because there is a known connection between the two. In Section 3.4 we define a random variable $X_{n,w,\eta}$ for the number of occurrences of η in all w-NAFs of a fixed length n. It is assumed that all those w-NAFs are equally likely. For an arbitrary algebraic integer τ an explicit expressions for the expectation and the variance of $X_{n,w,\eta}$ is given. Asymptotically we get $\mathbb{E}(X_{n,w,\eta}) \sim e_w n$ and $\mathbb{V}(X_{n,w,\eta}) \sim v_w n$ for constants e_w and v_w depending on w and the norm of τ. The proof uses a regular expression describing the w-NAFs. This will then be translated into a generating function. Further in this section it is shown that $X_{n,w,\eta}$ satisfies a central limit theorem.

A more general question is, what the number of occurrences $Z_{\tau,w,\eta}$ of the non-zero digit η is, when we look at all w-NAFs with absolute value smaller than a given N. For imaginary quadratic τ and a region $U \subseteq \mathbb{C}$ (e.g. the unit disc for the absolute value), the answer is in Section 3.10. We prove that $Z_{\tau,w,\eta} \sim e_w N^2 \lambda(U) \log_{|\tau|} N$. This is not surprising, since intuitively there are about $N^2 \lambda(U)$ w-NAFs in the region NU, and each of them can be represented as a w-NAF with length $\log_{|\tau|} N$. We even get a more precise result. If the region is "nice", there is a periodic oscillation of order N^2 in the formula. The structure of the result — main term, oscillation term, smaller error term — is not uncommon in the context of digits counting. The proof follows the ideas of Delange. In Section 3.8 and Section 3.9 the necessary tools for the proof will be developed and the characteristic sets will be analysed.

At last a short overview of the not mentioned sections until now. In Section 3.2 Voronoi cells and their basic properties are discussed. Section 3.3 deals with the digit sets, as well as the formal definition of the non-adjacent forms and some basic results. In Section 3.5 we give bounds for the value of a w-NAF.

The last chapter is Chapter 4. It contains some concluding remarks, as well as some open problems.

Chapter 2

Background and Known Results

Our main interest lies in the scalar multiplication of points on an elliptic curve. This is used in public key cryptography. This chapter deals with some background information concerning this type of cipher, as well as already known results. In particular, digital expansion used for efficient scalar multiplication are considered.

The first section, Section 2.1, contains some facts about elliptic curves. This includes a zeta function and its connection to the Frobenius endomorphism. Section 2.2 deals with the principle of elliptic curve cryptography. This is explained on the example Diffie-Hellman key exchange. There are also some notes on the secureness of such systems. Since we will often use a special digital expansion, namely the non-adjacent form, Section 2.3 will define this expansion. Further all notations used will be explained there.

The last two sections in this chapter deal with two methods for the scalar multiplication. In Section 2.4 double-and-add methods will be described. For the digital expansion there, we always use base 2. The subsections in there are split according to the digits used. In Section 2.5 the Frobenius-and-add method is described. There are different subsections for expansions coming from Koblitz curves in characteristic 2 and Koblitz curves in characteristic 3, and there is one subsection for the other cases.

2.1 Some Facts about Elliptic Curves

This section will deal with the definition of an elliptic curve and some basic facts about it. See for example Silverman [62] or Koblitz [41] for details.

Definition 2.1.1. An *elliptic curve* is a pair $(\mathcal{E}, \mathbf{0})$, where $\mathcal{E}$ a smooth algebraic curve of genus 1 and $\mathbf{0}$ a point (identity) on the curve. The elliptic curve is defined over a field K, if $\mathcal{E}$ is defined over K and $\mathbf{0} \in \mathcal{E}(K)$.

We will usually write just $\mathcal{E}$ for the elliptic curve. The following characterisation of elliptic curves can be proved, cf. Silverman [62, III Proposition 3.1].

Proposition 2.1.2. *Let $\mathcal{E}$ be an elliptic curve defined over K. There exists a $\Phi\colon \mathcal{E} \longrightarrow \mathbb{P}^2$, such that Φ is an isomorphism of $\mathcal{E}$ defined over K onto a curve given by a* Weierstrass equation

$$C\colon Y^2Z + a_1XYZ + a_3YZ^2 = X^3 + a_2X^2Z + a_4XZ^2 + a_6Z^3$$

with coefficients $a_1, \ldots, a_6 \in K$ and such that $\mathbf{0} \longmapsto [0,1,0]$. Any two such Weierstrass equations for $\mathcal{E}$ are related by a linear change of variables.

Conversely, every smooth cubic curve C given by a Weierstrass equation is an elliptic curve defined over K with $\mathbf{0} = [0, 1, 0]$.

We will also write $\mathcal{E}$ for the Weierstrass equation. The set $\mathcal{E}(K)$ then consists of the point at infinity $[0, 1, 0]$ of all points of K^2 (now written in affine coordinates) that fulfil this Weierstrass equation. If the characteristic of the field K is neither 2 nor 3, then the Weierstrass form can be simplified to

$$C\colon Y^2 = X^3 + aX + b.$$

The points $\mathcal{E}(K)$ of an elliptic curve $\mathcal{E}$ over a field K form an Abelian group. The addition in this group has a nice geometric interpretation if, for example, $K = \mathbb{R}$. The group law in general is described in e.g. Silverman [62, III Section 2] and in Koblitz [41, VI Section 1]. The group operation will be notated by $+$, the inverse of a point P by $-P$, and we will use $nP = P + \cdots + P$ for $n \in \mathbb{N}_0$ and $nP = (-P) + \cdots + (-P)$, when n is a negative integer.

Now we look at elliptic curves over finite fields. So we set $K = \mathbb{F}_q$, where q is a power of a prime. First we want to know, how many points there are on the curve. Clearly, an upper bound is $2q+1$. This follows from the fact that each value of $x \in \mathbb{F}_q$ yields at most two values of $y \in \mathbb{F}_q$ on

$$\mathcal{E}\colon Y^2 + a_1XY + a_3Y = X^3 + a_2X^2 + a_4X + a_6.$$

But a better bound can be proved. It is stated in the following theorem, cf. for example Silverman [62, V Theorem 1.1] or Koblitz [41, VI Section 1].

Theorem 2.1.3. *Let $\mathcal{E}$ be an elliptic curve over $\mathbb{F}_q$, then*

$$|\#\mathcal{E}(\mathbb{F}_q) - q - 1| \leq 2\sqrt{q}.$$

Next we want to define the zeta function of the elliptic curve.

Definition 2.1.4. The *zeta function of $\mathcal{E}$ over $\mathbb{F}_q$* is the formal power series

$$Z_{\mathbb{F}_q}(T) = \exp\left(\sum_{n\in\mathbb{N}} \#\mathcal{E}(\mathbb{F}_{q^n})\frac{T^n}{n}\right).$$

Further we set

$$\zeta_{\mathbb{F}_q}(s) := Z_{\mathbb{F}_q}\left(q^{-s}\right).$$

The function $\zeta_{\mathbb{F}_q}$ is also called *zeta function*.

Directly related to the zeta function are the *Weil conjectures*[1], for example cf. Silverman [62, V Theorem 2.2] or Koblitz [41, VI Section 1]. The following theorem are the Weil conjectures for elliptic curves. A proof and further details can be found, again for example, in Silverman [62, V Section 2].

Theorem 2.1.5 (Weil Conjectures for Elliptic Curves)**.** *We get the following statements:*

(a) (Rationality) We get

$$Z_{\mathbb{F}_q}(T) \in \mathbb{Q}(T).$$

(b) (Functional Equation) The zeta function fulfils the functional equation

$$Z_{\mathbb{F}_q}\left(\frac{1}{qT}\right) = Z_{\mathbb{F}_q}(T).$$

[1] In Koblitz [41, VI Section 1] the Weil conjecture is also called *Delinge's theorem*.

(c) (Riemann Hypothesis) There is a factorisation

$$Z_{\mathbb{F}_q}(T) = \frac{(1-\alpha T)(1-\beta T)}{(1-T)(1-qT)}$$

with $|\alpha| = |\beta| = \sqrt{q}$.

The numerator of the zeta function $Z_{\mathbb{F}_q}(T)$ is called *L-polynomial*. We write

$$L_{\mathbb{F}_q}(T) = (1-\alpha T)(1-\beta T).$$

From the functional equation for $Z_{\mathbb{F}_q}$, we get the functional equation for $\zeta_{\mathbb{F}_q}$, namely

$$\zeta_{\mathbb{F}_q}(1-s) = \zeta_{\mathbb{F}_q}(s)\,.$$

The structure of this functional equation is one reason, why the last statement in Theorem 2.1.5 on the facing page is called Riemann hypothesis. The other is that $\zeta_{\mathbb{F}_q}(s) = 0$ implies $|q^s| = \sqrt{q}$. This means that $\mathrm{Re}(s) = \frac{1}{2}$.

Further, see Silverman [62, V Sections 2–4], we get

$$(1-\alpha T)(1-\beta T) = 1 - aT + qT^2$$

with $a \in \mathbb{Z}$. The integer a fulfils the relation

$$\#\mathcal{E}(\mathbb{F}_q) = 1 - a + q,$$

and we get

$$\#\mathcal{E}(\mathbb{F}_{q^n}) = 1 - \alpha^n - \beta^n + q^n$$

for field extensions. Moreover, the characteristic polynomial of the qth power Frobenius endomorphism is $(T-\alpha)(T-\beta)$, i.e., the reciprocal polynomial of the L-polynomial, cf. Silverman [62, V Sections 2–4]. Thus we obtain

$$\varphi^2 - a\varphi + q = 0$$

for the Frobenius endomorphism φ.

2.2 Elliptic Curve Cryptography

In this section a description of cryptography with elliptic curves on the example of Diffie-Hellman key exchange, cf. Diffie and Hellman [18], will be given.

Consider the following situation. We have two communication partners Alice and Bob, who want to exchange data securely by a symmetric key cipher. Therefore they need a shared secret key. The problem is that there is only a insecure communications channel. The *Diffie-Hellman key exchange* algorithm gives a solution for this problem. This is an asymmetric algorithm, i.e., there each communication partner has a public key that is commonly known and a private key only known by itself.

The "classical" Diffie-Hellman key exchange works in the following way. Both communication partners agree for a prime $p \in \mathbb{N}$ and a primitive root g modulo p. These two parameters are public. Alice chooses a secret integer a and sends $A = g^a$ modulo p to Bob. Bob does the same, i.e., chooses an secret integer b and sends $B = g^b$ modulo p to Alice. Now Bob calculates A^b modulo p and Alice B^a modulo p. Both get the same number g^{ab} modulo p, so they can use this as shared secret key.

The secureness of the "classical" Diffie-Hellman key exchange depends on the secureness of the *discrete logarithm problem*, for example cf. Koblitz [41, IV Section 3]. In our case, this means finding the secret a of Alice by the knowledge of g and $A = g^a$ or the secret b of Bob. The discrete logarithm problem is believed to be "hard" to solve, i.e., up to now, there is no efficient classical algorithm known to compute the discrete logarithm. Clearly this is only true, if the parameters are chosen appropriately, e.g. a very large prime p.

Now we want to use elliptic curves for cryptographic algorithms. This was first mentioned independently in Miller [46] and Koblitz [38]. In the following we want to describe the *elliptic curve Diffie-Hellman key exchange*. This was also the example used in Miller [46]. Koblitz [38] used other examples. Let $\mathcal{E}$ be an elliptic curve over a finite field $\mathbb{F}$ and let $\mathcal{E}(\mathbb{F})$ be its group of rational points. This curve is known public, as well as a point $G \in \mathcal{E}(\mathbb{F})$. Instead of taking publicly known powers as in the "classical" Diffie-Hellman algorithm, the scalar multiplication in the point group is used. This means, that for an $n \in \mathbb{N}$ multiples $nG = G + \cdots + G$ are calculated. The algorithm is then as follows. Alice chooses an integer a and sends her public key, the group point $A = aG$ to Bob. Bob chooses an integer b and sends $B = bG$ to Alice. Both compute $abG = bA = aB$ and therefore get a common shared secret key.

The secureness of the elliptic curve Diffie-Hellman key exchange algorithm depends on the secureness of the *elliptic curve discrete logarithm problem*, i.e., finding n, when G and nG are known. Again, using appropriate parameters, this is believed to be "hard".

The group of points of an elliptic curve over $\mathbb{F}_q$ can be embedded into the multiplicative group of $\mathbb{F}_{q^\kappa}$ for an appropriate K. This uses Weil pairing, cf. Menezes, Okamoto and Vanstone [45]. Therefore the elliptic curve discrete logarithm problem can be reduced to the discrete logarithm problem. This Menezes-Okamoto-Vanstone attack is useful, if K is small. If the elliptic curve is supersingular, then K can be chosen out of $\{1, 2, 3, 4, 6\}$, cf. Menezes, Okamoto and Vanstone [45]. If the curve is non-supersingular, K is usually much larger, cf. Balasubramanian and Koblitz [10].

2.3 Notations concerning Non-adjacent Forms

Let τ be an algebraic integer, and let $\mathcal{D}$ be a finite subset of $\mathbb{Z}[\tau]$ containing 0. This set $\mathcal{D}$ is the used *digit set*. For simplicity we set $\mathcal{D}^\bullet := \mathcal{D} \setminus \{0\}$, and with D^* we denote all finite words over the alphabet $\mathcal{D}$. For a digit η we set $\bar{\eta} := -\eta$. Let $w \in \mathbb{N}$. Usually we assume $w \geq 2$. We have the following definition for non-adjacent forms.

Definition 2.3.1 (Width-w τ-adic Non-Adjacent Forms)**.** Let $\boldsymbol{\eta} = (\eta_j)_{j\in\mathbb{Z}} \in \mathcal{D}^{\mathbb{Z}}$. The sequence $\boldsymbol{\eta}$ is called a *width-w τ-adic non-adjacent form*, or *w-NAF* for short, if each factor $\eta_{j+w-1} \dots \eta_j$, i.e., each block of length w, contains at most one non-zero digit.

Let $J = \{j \in \mathbb{Z} \,|\, \eta_j \neq 0\}$. We call $\sup(\{0\} \cup (J + 1))$ the *left-length of the w-NAF $\boldsymbol{\eta}$* and $-\inf(\{0\} \cup J)$ the *right-length of the w-NAF $\boldsymbol{\eta}$*.

Let λ and ρ be elements of $\mathbb{N}_0 \cup \{\mathsf{fin}, \infty\}$, where fin means finite. We denote the *set of all w-NAFs of left-length at most λ and right-length at most ρ* by $\mathbf{NAF}_w^{\lambda,\rho}$. If $\rho = 0$, then we will simply write $\mathbf{NAF}_w^{\lambda}$. The elements of the set $\mathbf{NAF}_w^{\mathsf{fin}}$ will be called *integer w-NAFs*.

For $\boldsymbol{\eta} \in \mathbf{NAF}_w^{\mathsf{fin},\infty}$ we call

$$\mathsf{value}(\boldsymbol{\eta}) := \sum_{j\in\mathbb{Z}} \eta_j \tau^j$$

the *value of the w-NAF $\boldsymbol{\eta}$*.

The following notations and conventions are used. A block of zero digits is denoted by $\mathbf{0}$. For a digit η and $k \in \mathbb{N}_0$ we will use

$$\eta^k := \underbrace{\eta \dots \eta}_{k},$$

with the convention $\eta^0 := \varepsilon$, where ε denotes the empty word. A w-NAF $\boldsymbol{\eta} = (\eta_j)_{j\in\mathbb{Z}}$ will be written as $\boldsymbol{\eta}_I.\boldsymbol{\eta}_F$, where $\boldsymbol{\eta}_I$ contains the η_j with $j \geq 0$ and $\boldsymbol{\eta}_F$ contains the η_j with $j < 0$. $\boldsymbol{\eta}_I$ is called *integer part*, $\boldsymbol{\eta}_F$ *fractional part*, and the dot is called τ*-point*. Left-leading zeros in $\boldsymbol{\eta}_I$ will can be skipped, except η_0, and right-leading zeros in $\boldsymbol{\eta}_F$ can be skipped as well. If $\boldsymbol{\eta}_F$ is a sequence containing only zeros, the τ-point and this sequence is not drawn.

Further, for a w-NAF $\boldsymbol{\eta}$ (a bold, usually small Greek letter) we will always use η_j (the same letter, but indexed and not bold) for the elements of the sequence.

The term "non-adjacent form" goes back to Reitwiesner [58]. There NAF meant 2-NAF with base 2 used. Solinas [63] and [64] used the term τ*-NAF* for a τ-adic 2-NAF. In this thesis, the τ or τ-adic is usually skipped, since it is clear in the context what τ is.

The next definition in this section deals with a special type of digit set.

Definition 2.3.2 (Width-w Non-Adjacent Digit Set)**.** A digit set $\mathcal{D}$ is called a *width-w non-adjacent digit set*, or w-NADS for short, if every element $z \in \mathbb{Z}[\tau]$ admits a unique w-NAF $\boldsymbol{\eta} \in \mathbf{NAF}_w^{\mathsf{fin}}$, i.e., $\mathsf{value}(\boldsymbol{\eta}) = z$. When this is the case, the function

$$\mathsf{value}|_{\mathbf{NAF}_w^{\mathsf{fin}}} \colon \mathbf{NAF}_w^{\mathsf{fin}} \longrightarrow \mathbb{Z}[\tau]$$

is bijective, and we will denote its inverse function by NAF_w. If $\mathcal{D}$ is not a w-NADS it is called a w-non-NADS.

For a subset $S \subseteq \mathbb{Z}[\tau]$ we call the digit set $\mathcal{D}$ a w-NADS for S, if every element of S admits a unique w-NAF.

Sometimes there will be written NADS and non-NADS instead of w-NADS and w-non-NADS, respectively, if w is clear from context.

Definition 2.3.3. For a w-NAF $\boldsymbol{\eta} \in \mathbf{NAF}_w^{0,\infty}$ we define $\mathsf{weight}(\boldsymbol{\eta})$ as the number of non-zero digits, i.e.,

$$\mathsf{weight}(\boldsymbol{\eta}) := \#(\{j \in \mathbb{Z} \,|\, \eta_j \neq 0\})\,.$$

Sometimes it is useful not to see a w-NAF as infinite sequence, but as a finite string. Note that two strings differing only on leading zeros denote the same w-NAF. We will also call a string a w-NAF, if the corresponding infinite sequence is a w-NAF. We denote the length of a string $\boldsymbol{\beta}$ by $|\boldsymbol{\beta}|$. For two strings $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ we write $\boldsymbol{\alpha} \parallel \boldsymbol{\beta}$ for their concatenation.

2.4 Expansions for Double-and-Add Scalar Multiplication Methods

One possibility to perform the scalar multiplication nP, $n \in \mathbb{N}_0$, P a point of the elliptic curve, is to write n in its standard binary expansion and to use a *Horner scheme* to evaluate nP, see Knuth [37]. Since base 2 is used, this is known as *double-and-add* method. One goal is to make this operation as efficient as possible. This main idea is using other (larger) digit sets than $\{0, 1\}$. The resulting expansions of n can be evaluated faster, since normally more digits means more zeros in the expansion. Most of the material in this section is related to the non-adjacent form defined in Section 2.3, or some variants and generalisations of it.

Section 2.4.1 will start with the digit set $\{-1, 0, 1\}$. The following two sections will generalise in different ways. In Section 2.4.2 the digit set will be expanded by some odd numbers to get more zeros in the expansion, whereas in Section 2.4.3 we will use a digit set containing 0, 1 and one additional digit.

Of course there are other numeral systems — not mentioned in detail in this thesis — which can be used in the scalar multiplication method. In Phillips and Burgess [55] representations

with base q, $q \in \mathbb{N}$, $q \geq 2$ and digit set $\{a \in \mathbb{Z} \,|\, \ell \leq a \leq u\}$ for some $\ell \leq 0$ and some $u \geq 1$ are considered. They give a construction that produces minimal weight representations from right to left. In Heuberger and Muir [29] an algorithm to calculate such expansions (with base $q = 2$) from the binary representation from left to right is given.

Another problem is the computation of linear combinations $nP + mQ$ of points P and Q. One can use *joint expansions* to calculate at once instead of nP and mQ separately. Further information can be found in Solinas [65], Grabner, Heuberger and Prodinger [24], Proos [57], Heuberger and Muir [28], Heuberger, Katti, Prodinger and Ruan [27].

2.4.1 2-NAFs with Digits 0, $+1$ and -1

The first idea of extending the standard binary digit set $\{0, 1\}$ is to add the digit -1. This seems to be efficient, since $-P$ can be calculated fast on the curve. Thus, no pre-computation is needed. This was first mentioned by Morain and Olivos [49]. In this section, some facts on non-adjacent representations with base 2 and this digit set $\mathcal{D} = \{0, 1, -1\}$ will be presented. Such representations go back to Reitwiesner [58].

Note that Section 2.4.2 and Section 2.4.3 are generalisations of the 2-NAFs here, so the results there might be interesting for this section, too.

2.4.1.1 Existence and Uniqueness

First we want to know, when such 2-NAF expansions exists and when they are unique. Luckily this is true for all integers. This results can be found in Reitwiesner [58]. We have the following theorem, which (including a proof) can also be found in Shallit [61, Theorems 1.2 and 1.3].

Theorem 2.4.1. *Every integer has exactly one 2-NAF-representation with digits* $\{0, +1, -1\}$.

Note that in this work the NAFs are defined via sequences, so leading zeros in the corresponding word of digits have not be mentioned extra.

2.4.1.2 Length and Density

Another result, which goes back to Reitwiesner [58], is that the length of the 2-NAF of n is at most one digit longer than the binary representation of n. This is also valid for the general case of a w-NAF, see Section 2.4.2.2.

The average density of non-zero coefficients in all 2-NAFs of length ℓ is

$$\frac{2^\ell(3\ell - 4) - (-1)^\ell(6\ell - 4)}{9(\ell - 1)(2^\ell - (-1)^\ell)}.$$

This fact was stated by Solinas [64, Section 3.1]. Therefore, we get asymptotically $\frac{1}{3}$, cf. Morain and Olivos [49], too. A more precisely result can be found in Thuswaldner [66].

Further for an integer n we get $2^\ell < 3n < 2^{\ell+1}$, where ℓ denotes the length of the 2-NAF of n, again cf. Solinas [64, Section 3.1] or Morain and Olivos [49].

2.4.1.3 Calculating the 2-NAFs

The 2-NAF of the integer n can of course be calculated directly from n by an algorithm. See, for example, Reitwiesner [58] or Solinas [63, Algorithm 2] or [64, Routine 4]. Jedwab and Mitchell [34] gave an algorithm that can compute the 2-NAF out of any redundant expansion with digits 0, 1, -1.

Further, there is also an explicit expression for the digits. In Prodinger [56, Section 2] the following formula can be found to calculate the digits of the 2-NAF-expansion and therefore the 2-NAF itself. For an integer n we get

$$n = \sum_{k\geq 0} \left(\left\lfloor \frac{n}{2^{k+2}} + \frac{5}{6} \right\rfloor - \left\lfloor \frac{n}{2^{k+2}} + \frac{4}{6} \right\rfloor - \left\lfloor \frac{n}{2^{k+2}} + \frac{2}{6} \right\rfloor + \left\lfloor \frac{n}{2^{k+2}} + \frac{1}{6} \right\rfloor \right) 2^k.$$

Note that

$$\left\lfloor \frac{n}{2^{k+2}} + \frac{5}{6} \right\rfloor - \left\lfloor \frac{n}{2^{k+2}} + \frac{4}{6} \right\rfloor - \left\lfloor \frac{n}{2^{k+2}} + \frac{2}{6} \right\rfloor + \left\lfloor \frac{n}{2^{k+2}} + \frac{1}{6} \right\rfloor = \begin{cases} 1, & \text{if } \left\{ \frac{n}{2^{k+2}} \right\} \in \left(\frac{1}{6}, \frac{2}{6} \right), \\ -1, & \text{if } \left\{ \frac{n}{2^{k+2}} \right\} \in \left(\frac{4}{6}, \frac{5}{6} \right), \\ 0, & \text{else,} \end{cases}$$

so these intervals may be seen as characteristic sets of the digits, cf. also Sections 2.4.3.9 and 3.9. A generalisation of the explicit formula can be found in Heuberger and Prodinger [30].

2.4.1.4 Number of Representations

Now we change the setting a little bit and omit the non-adjacency condition. Then we may ask, how many representations an integer has using base 2 and the digit set $\mathcal{D} = \{0, 1, -1\}$. The following result is given in Shallit [61, Theorem 1.1].

Theorem 2.4.2. *Every non-zero integer has an infinite number of signed-digit expansions (with digits* $\{0, +1, -1\}$*.*

Of course they usually do not fulfil the NAF-condition. Another question is, what the number of representations of a given length is. We have the following theorem, cf. for example Shallit [61, Theorem 1.5].

Theorem 2.4.3. *There are* $t_\ell = \frac{1}{3}\left(2^{\ell+2} - (-1)^\ell\right)$ *distinct 2-NAF representations of length* ℓ*.*

2.4.1.5 Optimality

We know from the previous section that there are a lot of representations of an integer, namely infinitely many. So a natural question is, is there a "good" representation (in some sense). Since our primary goal is to optimise the scalar multiplication on elliptic curves, we want expansions with a lot of zeros inside. The 2-NAF fulfils this, and indeed it can be shown that the 2-NAF representation is optimal in the sense that it minimises the number of non-zero digits. This is an result of Reitwiesner [58]. It can also be found in Jedwab and Mitchell [34] or Gordon [22]. Of course this minimum is not unique, since for example the integer 3 has expansions 11 and $10\bar{1}$.

2.4.1.6 Number of Optimal Representations

We continue with the topics of the previous two sections. From Section 2.4.1.5 we know that there is one optimal representation, namely the 2-NAF mentioned by Reitwiesner [58]. Usually there are more representations, see the example at the end of the previous section. So the question is, how many optimal (minimal) representation of an integer are there?

Let $f(n)$ denote the number of minimal expansions of the integer n. In Grabner and Heuberger [23, Theorem 1] we can find the following upper bound for $f(n)$.

Theorem 2.4.4. *For all integers* n*, the number of optimal expansions can be bounded by*

$$f(n) \leq F_{\lfloor \log_4 |n| \rfloor + 3},$$

where F_j denotes the Fibonacci sequence $F_0 = 0$, $F_1 = 1$, $F_{j+2} = F_{j+1} + F_j$. This bound is sharp for infinitely many values of n. Less precisely, we have

$$f(n) = \mathcal{O}\big(n^{\log_4 \varphi}\big) \quad \textit{with} \quad \varphi = \tfrac{1+\sqrt{5}}{2}.$$

The proof uses the automaton mentioned in Heuberger and Prodinger [31, Remark 20]. This automaton accepts an expansion if and only if it is optimal.

The next is the study of the summatory function $\sum_{0 \le n < N} f(n)$. It describes the average behaviour of $f(n)$. The following theorem can be found in Grabner and Heuberger [23, Theorems 2 and 3].

Theorem 2.4.5. *The counting function $f(n)$ of the representations of n with minimal weights satisfies*

$$\sum_{0 \le n < N} f(n) = N^{\log_2 \alpha}\, \Psi(\log_2 N) + \mathcal{O}\big(N^{\log_2 \alpha - \theta}\big)\,,$$

where Ψ denites a continuous periodic function of period 1, $\alpha = 2.17009\ldots$, and $\theta = 0.2168\ldots$. Furthermore, Ψ is Hölder continuous with exponent $\beta = 0.770632\ldots$. The function Ψ is differentiable almost everywhere and singular in the sense that it is not the integral of its derivative.

Further, the function Ψ admits an absolutely and uniformly convergent Fourier series. Its coefficients can be calculated.

2.4.2 Windowing Methods and w-NAFs

Now we will use the *width-w window method* to generalise the the mentioned 2-NAFs. For such methods in general cf. Gordon [22, Section 3]. The w-NAFs were described independently by Cohen, Miyaji and Ono [47], Blake, Seroussi and Smart [13] and Solinas [63, 64].

In this section, w is an integer with $w \ge 2$. The digit set $\mathcal{D}$ consists of zero and all odd numbers with absolute value less than 2^{w-1}.

2.4.2.1 Existence and Uniqueness

Here we will use the statements made in Solinas [63] and Solinas [64, Section 3.2]. Additionally the existence and uniqueness can be found in Muir and Stinson [53, Sections 2.1 and 2.2].

Theorem 2.4.6. *Let $w \in \mathbb{N}$ with $w \ge 2$. Then every positive integer has a unique w-NAF $\boldsymbol{\eta}$, where the digits are in $\mathcal{D}$.*

Clearly this theorem is true for non-positive integers as well. The idea behind is the following. We take a window of width w and let it slide over the binary expansion of an integer from right to left. If the value in this block is even, we get a zero, else an odd digit modulo 2^w. Based on this, Solinas [64, Routine 9] gave a simple algorithm to compute the w-NAF-expansion. Additionally they gave an algorithm to calculate multiples of an elliptic curve point using this w-NAFs. Because we are using more digits here, more pre-computation is needed, but in comparison the "main loop" is more efficiently.

The problem mentioned also in Solinas [64, Section 3.2] is that the w-NAF is computed from right to left, whereas the elliptic curve point scalar multiplication is done from left to right.

2.4.2.2 Length and Density

We can find the following proposition in Muir and Stinson [53, Section 2.3].

Proposition 2.4.7. *For any integer n the length of the w-NAF of n is at most one digit longer than the binary representation of $|n|$.*

It seems that this fact was stated first by Möller [48], but without a proof. Of course for the 2-NAF this is known a long time, e.g. cf. Reitwiesner [58].

Considering all w-NAFs of length n, the average density of non-zero digits is approximately (asymptotically) $1/(w+1)$, cf. Muir and Stinson [53, Section 6] and Cohen [16].

2.4.2.3 Optimality

Muir and Stinson [53, Theorem 3.3] proved the following theorem. There the w-NAFs are denoted as strings.

Theorem 2.4.8. *If $\boldsymbol{\alpha}$ is a w-NAF then for any $\boldsymbol{\beta} \in \mathcal{D}^*$ with $\mathsf{value}(\boldsymbol{\alpha}) = \mathsf{value}(\boldsymbol{\beta})$, we have $\mathsf{weight}(\boldsymbol{\alpha}) \leq \mathsf{weight}(\boldsymbol{\beta})$.*

This means that the w-NAF-expansion is optimal, i.e., minimises the Hamming weight among all expansions with digits $\mathcal{D}$. For the 2-NAF this was already mentioned by Reitwiesner [58]. This optimality result for w-NAFs was independently shown in Avanzi [3, Theorem 2.3], too.

Further, in Muir and Stinson [53, Theorem 4.1], we have the following generalisation of the result of Section 2.4.2.2. The generalisation concerns all minimal weight representations.

Theorem 2.4.9. *If $\boldsymbol{\alpha}$ is optimal, i.e., for any $\boldsymbol{\beta} \in \mathcal{D}^*$ with $n = \mathsf{value}(\boldsymbol{\alpha}) = \mathsf{value}(\boldsymbol{\beta})$ we have $\mathsf{weight}(\boldsymbol{\alpha}) \leq \mathsf{weight}(\boldsymbol{\beta})$, then we get that the length of $\boldsymbol{\alpha}$ is at most $\lfloor \log_2 |n| \rfloor + 2$.*

Remark that the standard binary expansion of a positive integer n has length $\lfloor \log_2 n \rfloor + 1$.

In general the w-NAF-representation need not be the unique minimal Hamming weight expansion of all expansions. Consider the following example found in Muir and Stinson [53, Section 5]. The integer 5 has the 3-NAF $100\bar{3}$. But we also have

$$5 = \mathsf{value}(101) = \mathsf{value}(13) = \mathsf{value}(3\bar{1})\,.$$

All those expansions have Hamming weight 2 and therefore, since the 3-NAF is optimal, i.e., minimal, are optimal. Muir and Stinson [53, Section 5] gave a method to compare different expansions and by means of this, the w-NAF is the uniquely defined minimum.

Consider a w-NAF $\boldsymbol{\alpha}$ as string. We build another string $\boldsymbol{\alpha}' \in \{0,1\}^*$ by

$$\alpha'_j = \begin{cases} 0 & \text{if } \alpha_j = 0, \\ 1 & \text{otherwise.} \end{cases}$$

Now for $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$, which are assumed to be representations of n, the order $\boldsymbol{\alpha} \preceq \boldsymbol{\beta}$ is defined by the lexicographic right-to-left ordering of $\boldsymbol{\alpha}'$ and $\boldsymbol{\beta}'$. We get the following result, cf. Muir and Stinson [53, Theorem 5.1].

Theorem 2.4.10. *Of all the representations of n with digits in $\mathcal{D}$, the w-NAF of n is the unique smallest representation under the order $\preceq$.*

2.4.2.4 Other Optimal Expansions than w-NAFs

As remarked in the previous section, the w-NAF is not the only optimal (minimal) representation. Muir and Stinson [52] gave another minimal expansion that uses the same digit set $\mathcal{D}$ as the w-NAF. They generalised the ideas of Joye and Yen [35]. The advantage of this expansion is that it can be calculated from left to right. Therefore it can be used efficiently in the double-and-add algorithm, because the digits do not have to be stored in memory.

2.4.3 2-NAFs with Digits 0, 1 and x

In the previous section we considered 2-NAFs with digit set $\mathcal{D} = \{0, +1, -1\}$. Now we change the digit set a little bit and generalise. The digit -1 is replaced by an arbitrary integer $x \in \mathbb{Z}$. Such digit sets were mentioned by Muir and Stinson [50]. This article was extended in Muir and Stinson [51]. In the following we will usually refer to the latter.

Throughout this section our base $\tau = 2$ is fixed, we have an integer $x \in \mathbb{Z}$, the digit set is $\mathcal{D} = \{0, 1, x\}$, and we will consider only 2-NAFs.

2.4.3.1 Notations

For an integer $n \in \mathbb{Z}$ we define

$$R_{\mathcal{D}}(n) := \begin{cases} \alpha & \text{when there exists a 2-NAF } \alpha \text{ with } \mathsf{value}(\alpha) = n, \\ \perp & \text{otherwise.} \end{cases}$$

The symbol $\perp$ is just some symbol not in $\mathcal{D}$. If there are more 2-NAFs α fulfilling $\mathsf{value}(\alpha) = n$, then one α is chosen. We will get a uniqueness result later in this section. Further we define

$$\mathsf{NAF}(\mathcal{D}) := \{n \in \mathbb{Z} \mid R_{\mathcal{D}}(n) \neq \perp\},$$

i.e., the set of integers which have a 2-NAF-representation.

2.4.3.2 Necessary Condition for a NADS

In Muir and Stinson [51, Theorem 3.1] we can find the following necessary condition that a set $\{0, 1, x\}$ is a NADS.

Theorem 2.4.11. *If there exists an $n \in \mathsf{NAF}(\mathcal{D})$ with $n \equiv 3 \pmod 4$, then $x \equiv 3 \pmod 4$.*

This necessary condition gives us now candidates x for NADS or formulated in another way rules out x not fulfilling the condition. This will be often used in Section 2.4.3.4 and Section 2.4.3.5.

2.4.3.3 Uniqueness

Of course, we want that our expansion is unique. Muir and Stinson [51, Theorem 3.3] proved the following result.

Theorem 2.4.12. *If $x \equiv 3 \pmod 4$, then any integer has at most one 2-NAF-representation with digit set $\{0, 1, x\}$.*

2.4.3.4 Non-adjacent Digit Sets for Positive x

Solinas [65] remarked that the digit set $\{0, 1, 3\}$ is a NADS for $\mathbb{N}_0$. But more can be proved, cf. Muir and Stinson [51, Theorem 3.2].

Theorem 2.4.13. *Let $x \geq 0$. The only NADS for $\mathbb{N}_0$ of the form $\{0, 1, x\}$ is $\{0, 1, 3\}$.*

The proof that $\{0, 1, 3\}$ is a NADS contains an algorithm to get a 2-NAF from the binary representation of a number. There the binary string is scanned from right to left and each 11 is replaced by 03. Clearly this does not change the value. The example

$$237 = \mathsf{value}(11101101) = \mathsf{value}(10300301)$$

mentioned in Muir and Stinson [51, Section 3.1] illustrates this.

For the other part of the proof, i.e., that there is no NADS for $x > 3$, it is shown that 3 has no representation as 2-NAF. The fact that for $x \in \{0, 1, 2\}$ the resulting digit set is not a NADS is easy to see.

2.4.3.5 Non-adjacent Digit Sets for Negative x

Now fix an $x < 0$ with $x \equiv 3 \pmod 4$. Muir and Stinson [51, Lemmata 4.1–4.4] showed the following lemma.

Lemma 2.4.14. *For $n \in \mathbb{N}_0$ we get the following statements:*

(a) If $n \equiv 0 \pmod 4$, then

$$n \in \mathsf{NAF}(\mathcal{D}) \iff n/4 \in \mathsf{NAF}(\mathcal{D})$$

Further, if $n \equiv 0 \pmod 4$, then $R_{\mathcal{D}}(n) = R_{\mathcal{D}}(n/4) \parallel 00$.

(b) If $n \equiv 1 \pmod 4$, then

$$n \in \mathsf{NAF}(\mathcal{D}) \iff (n-1)/4 \in \mathsf{NAF}(\mathcal{D})$$

Further, if $n \equiv 1 \pmod 4$, then $R_{\mathcal{D}}(n) = R_{\mathcal{D}}((n-1)/4) \parallel 01$.

(c) If $n \equiv 2 \pmod 4$, then

$$n \in \mathsf{NAF}(\mathcal{D}) \iff n/2 \in \mathsf{NAF}(\mathcal{D})$$

Further, if $n \equiv 2 \pmod 4$, then $R_{\mathcal{D}}(n) = R_{\mathcal{D}}(n/2) \parallel 0$.

(d) If $n \equiv 3 \pmod 4$, then

$$n \in \mathsf{NAF}(\mathcal{D}) \iff (n-x)/4 \in \mathsf{NAF}(\mathcal{D})$$

Further, if $n \equiv 3 \pmod 4$, then $R_{\mathcal{D}}(n) = R_{\mathcal{D}}((n-x)/4) \parallel 0x$.

Now consider the following two examples. We take $\mathcal{D} = \{0, 1, -9\}$ and in the first example $n = 7$. We have

$$R_{\mathcal{D}}(7) = R_{\mathcal{D}}(4) \parallel 0\bar{9} = R_{\mathcal{D}}(1) \parallel 00 \parallel 0\bar{9} = 1 \parallel 00 \parallel 0\bar{9} = 1000\bar{9}.$$

The process is stopped, since we got to evaluate $R_{\mathcal{D}}(0)$, which is clearly the empty word. For the second example we choose $n = 3$ and get

$$R_{\mathcal{D}}(3) = R_{\mathcal{D}}(3) \parallel 0\bar{9} = R_{\mathcal{D}}(3) \parallel 0\bar{9} \parallel 0\bar{9} = R_{\mathcal{D}}(3) \parallel 0\bar{9} \parallel 0\bar{9} \parallel 0\bar{9} = \ldots,$$

so this process does not stop.

Therefore, by means of the previous lemma, we get a simple recursive procedure to evaluate $R_{\mathcal{D}}(n)$. Using

$$f_{\mathcal{D}}(n) := \begin{cases} n/4 & \text{if } n \equiv 0 \pmod 4 \\ (n-1)/4 & \text{if } n \equiv 1 \pmod 4 \\ n/2 & \text{if } n \equiv 2 \pmod 4 \\ (n-x)/4 & \text{if } n \equiv 3 \pmod 4 \end{cases} \quad \text{and} \quad g_{\mathcal{D}}(n) := \begin{cases} 00 & \text{if } n \equiv 0 \pmod 4 \\ 01 & \text{if } n \equiv 1 \pmod 4 \\ 0n/2 & \text{if } n \equiv 2 \pmod 4 \\ 0x & \text{if } n \equiv 3 \pmod 4 \end{cases}$$

the procedure does in every step $\boldsymbol{\eta} \leftarrow g_{\mathcal{D}}(n) \parallel \boldsymbol{\eta}$ and $n \leftarrow f_{\mathcal{D}}(n)$ and runs until $n = 0$. Muir and Stinson [51, Section 4] showed that this procedure terminates if and only if $n \in \mathsf{NAF}(\mathcal{D})$.

The procedure cannot evaluate all n, since there is a problem when $R_{\mathcal{D}}(n) = \perp$. If this is the case, then the procedure fails to terminate. Let $n \in \mathbb{N}_0$. It can be shown that there is an i such that $0 \leq f_{\mathcal{D}}^j(n) < -x/3$ for all $j \geq i$. So if there is a problem in terminating, then the procedure "cycles" on the elements less than $-x/3$. Adding a simple cycle-detection to the procedure leads to an algorithm, that returns $\perp$, if a cycle is detected. This algorithm has running time $\mathcal{O}(\log n + |x|)$. Further we get the following theorem, cf. Muir and Stinson [51, Theorem 4.8].

Theorem 2.4.15. *If every element in the set*

$$\{n \in \mathbb{N}_0 \,|\, n \leq \lfloor -x/3 \rfloor\}$$

has a 2-NAF-expansion with digits $\{0, 1, x\}$*, then* $\{0, 1, x\}$ *is a NADS for* $\mathbb{N}_0$*.*

Therefore the question if $\{0, 1, x\}$ is a NADS can be answered by the computational method above. We need $\lfloor -x/3 \rfloor$ calls of our algorithm. Muir and Stinson [51, Corollary 4.9] improved this result, so that there are only $\lfloor -x/12 \rfloor$ calls needed.

Corollary 2.4.16. *If every element in the set*

$$\{n \in \mathbb{N}_0 \,|\, n \leq \lfloor -x/3 \rfloor,\ n \equiv 3 \pmod 4\}$$

has a 2-NAF-expansion with digits $\{0, 1, x\}$*, then* $\{0, 1, x\}$ *is a NADS for* $\mathbb{N}_0$*.*

They used this algorithm to compute all NADS with x greater than -10^6. The list starts with

$$3, -1, -5, -13, -17, -25, -29, -37, -53, -61, -65, -113, \ldots.$$

The mentioned algorithm was further improved by Avoine, Monnerat and Peyrin [9, Theorems 7 and 8]. They got the following two theorems.

Theorem 2.4.17. *Let* $3 \nmid x$*. If every element in the set*

$$\{n \in \mathbb{N}_0 \,|\, n \leq \lfloor -x/6 \rfloor,\ n \equiv 3 \pmod 4\}$$

has a 2-NAF-expansion with digits $\{0, 1, x\}$*, then* $\{0, 1, x\}$ *is a NADS for* $\mathbb{N}_0$*.*

Theorem 2.4.18. *Let* $3 \nmid x$ *and* $7 \nmid x$*. If every element in the set*

$$\{n \in \mathbb{N}_0 \,|\, n \leq \lfloor -x/12 \rfloor,\ n \equiv 3 \pmod 4\} \cup \{n \in \mathbb{N}_0 \,|\, \lfloor -x/6 \rfloor \leq n \leq \lfloor -x/6 \rfloor,\ n \equiv 3 \pmod 4\}$$

has a 2-NAF-expansion with digits $\{0, 1, x\}$*, then* $\{0, 1, x\}$ *is a NADS for* $\mathbb{N}_0$*.*

Further Avoine, Monnerat and Peyrin [9, Conjecture 1] gave the following conjecture.

Conjecture 2.4.19. Let $3 \nmid x$ and $7 \nmid x$. If every element in the set

$$\{n \in \mathbb{N}_0 \,|\, n \leq \lfloor -x/12 \rfloor,\ n \equiv 3 \pmod 4\}$$

has a 2-NAF-expansion with digits $\{0, 1, x\}$, then $\{0, 1, x\}$ is a NADS for $\mathbb{N}_0$.

Another result is given in Heuberger and Prodinger [31, Proposition 4].

Proposition 2.4.20. *Define the directed graph* $G := (V, A)$ *by* $V = \{0, \ldots, \lfloor |x| / 3 \rfloor\}$ *and*

$$A := \left\{(m, n) \in V^2 \,\middle|\, n \in \{2m, 4m + 1, 4m + x\}\right\}.$$

Then $\{0, 1, x\}$ *is a NADS for* $\mathbb{N}_0$ *if and only if every* $n \in V$ *is reachable from* 0*.*

This result is useful working with automata and their underlying directed graph. Further they remarked the following.

Proposition 2.4.21. *Let* $\mathcal{D} = \{0, 1, x\}$ *with* $x < 0$ *be a NADS for* $\mathbb{N}_0$*. Then* $\mathcal{D}$ *is a NADS for* $\mathbb{Z}$*.*

2.4.3.6 Infinite Families of Non-NADS

Again we fix an $x < 0$ with $x \equiv 3 \pmod 4$. Muir and Stinson [51, Corollary 6.2] gave the following statement.

Theorem 2.4.22. *If $(2^s - 1) \mid x$ for any $s \geq 2$, then $\{0, 1, x\}$ is not a NADS.*

This means that we get *non-allowable factors* of x. These are $3, 7, 31, \ldots$. But beside that numbers there are other non-allowable factors. This list starts with $73, 85, 89, 337, 451, 1103, \ldots$ and depends on the following result of Muir and Stinson [51, Corollary 6.2].

Theorem 2.4.23. *Suppose x_0 is an integer. If there is a $\boldsymbol{\beta} \in \{00, 0, 0x_0\}^*$ such that $\mathsf{value}(\boldsymbol{\beta}) \neq 0$ and $2^{|\boldsymbol{\beta}|} - 1 \mid \mathsf{value}(\boldsymbol{\beta})$, then x_0 is a non-allowable factor.*

Other results are the following three.

Theorem 2.4.24. *If $(3 - x)/4 = 11 \cdot 2^i$, where $i \geq 0$, then $\{0, 1, x\}$ is not a NADS.*

Theorem 2.4.25. *If $(3 - x)/4 = 7 \cdot 2^i$, where $i \geq 0$, then $\{0, 1, x\}$ is an NADS if and only if $i \in \{0, 1\}$.*

For the third, we define

$$m_i := \left\lfloor \frac{2^{i+1} - 1}{3} \right\rfloor$$

for $i \geq 0$. The following theorem holds.

Theorem 2.4.26. *Let x be an integer such that $4m_i - 1 < -x < 3 \cdot 2^i$ for some $i \geq 0$. If there exists $n \in \{1, 2, \ldots, \lfloor -x/3 \rfloor\}$ with $n \equiv 3 \pmod 4$ then $\{0, 1, x\}$ is not a NADS.*

All those results were proved in Muir and Stinson [51, Corollaries 6.3 and 6.4 and Theorem 6.5]. Other results were developed in Avoine, Monnerat and Peyrin [9, Theorems 9 and 10]. They gave generators for infinite families of non-NADS. Their results are given in the following two theorems.

Theorem 2.4.27. *If $x = -60k + 15$, $x = -60k + 11$ or $x = -28k + 7$ with $k \in \mathbb{N}$, then $\{0, 1, x\}$ is not a NADS.*

Theorem 2.4.28. *Let $t \geq 2$ and $k > 0$ be two integers and $x = -(4k - 1)(2^{2t-1} - 1)$, then $\{0, 1, x\}$ is not a NADS.*

Further the term worst non-NADS is defined in Avoine, Monnerat and Peyrin [9, Section 3.3] as follows.

Definition 2.4.29. The set $\mathcal{D} = \{0, 1, x\}$ is a *worst non-NADS* if for all $n \leq -x/3$ with $n \equiv 3 \pmod 4$, $n \notin \mathsf{NAF}(\mathcal{D})$.

They also characterised all worst non-NADS by means of the following theorem.

Theorem 2.4.30. *The set $\{0, 1, x\}$ is a worst non-NADS if and only if there exists an $i \geq 2$ such that $4m_i - 1 < -x < 3 \cdot 2^i$.*

Using those results, a much more effective algorithm can be created to determine whether a set $\{0, 1, x\}$ is a NADS or not, cf. Avoine, Monnerat and Peyrin [9, Section 4]. According to them this algorithm for calculating the NADS for x greater than -10^7 is about a factor 3 faster than the best known algorithm in Muir and Stinson [50, 51].

2.4.3.7 Infinite Families of NADS

Again we fix an $x < 0$ with $x \equiv 3 \pmod 4$. Let $w(n)$ denote the Hamming weight of n in its binary representation, i.e., the number of ones.

The following two theorems presents families of x, where the $\{0, 1, x\}$ is a NADS. Those results including proofs can be found in Muir and Stinson [51, Theorems 7.1 and 7.2] and [50, Theorem 15].

Theorem 2.4.31. *If $w((3-x)/4) = 1$, then $\{0, 1, x\}$ is a NADS.*

For negative x, the condition $w((3-x)/4) = 1$ is equivalent to $(3-x)/4 = 2^t$ for an $t \geq 0$. Thus we get that $\{0, 1, x\}$ is a NADS, when x is

$$-1, -5, -13, -29, -61, \ldots.$$

Theorem 2.4.32. *If $w((3-x)/4) = 2$ and $2^s - 1$ does not divide x for any $s \in \mathbb{N}$ with $s \geq 2$, then $\{0, 1, x\}$ is a NADS.*

2.4.3.8 Calculating the NAF from Right to Left

Let $\mathcal{D} = \{0, 1, x\}$ with $x \equiv 3 \pmod 4$. Since we want to use the 2-NAF expansion for performing scalar multiplication, we have to calculate it. Some results are stated in this section. Note that the calculation here is from right to left, but for the Horner scheme in the double-and-add scalar multiplication the expansion is needed from left to right. Therefore the digits must be stored somewhere.

Define $\eta_0 \colon \mathbb{Z} \longrightarrow \mathcal{D}$ and $r \colon \mathbb{Z} \longrightarrow \mathbb{Z}$ by

$$\eta_0(n) := \begin{cases} 0 & \text{if } n \equiv 0 \pmod 2, \\ 1 & \text{if } n \equiv 1 \pmod 4, \\ x & \text{if } n \equiv 3 \pmod 4, \end{cases} \qquad r(n) := \frac{n - \eta_0(n)}{2}.$$

Heuberger and Prodinger [31, Section 3] defined a transducer that calculates the NAF for an given input out of its binary representation. This transducer $\mathcal{T}_0$ is defined as follows. The input alphabet is $\{0, 1\}$, the output alphabet $\mathcal{D}$. The set of states $\mathcal{Q}_0$ consists of the initial state $\mathcal{I}$ and $\{0, \ldots, 2 + |x|\}$ representing carries. The terminal state is 0. The set of transitions is defined by

$$\mathcal{E}_0 = \left\{ \mathcal{I} \xrightarrow{0|\varepsilon} 0, \mathcal{I} \xrightarrow{1|\varepsilon} 1 \right\} \cup \left\{ m \xrightarrow{d|\eta_0(2d+m)} r(2d+m) \,\middle|\, 0 \leq m \leq 2 + |x|, d \in \{0, 1\} \right\}.$$

There ε denotes the empty word. We define the transducer $\mathcal{T}$ by deleting non-accessible states from $\mathcal{T}_0$. The states are denoted by $\mathcal{Q}$ and the transitions by $\mathcal{E}$. Further in Heuberger and Prodinger [31, Theorem 7] the following theorem is given.

Theorem 2.4.33. *Let $\mathcal{T}$ be the transducer constructed above. Then the following holds:*

(a) $\#\mathcal{Q} \leq |x| + 4$

(b) A integer n with binary expansion $\mathbf{d} = d_J \ldots d_0$ has a 2-NAF with digits $\mathcal{D}$ if and only if there is a successful path with input label $d_{J+\#\mathcal{Q}-2} \ldots d_0$ in $\mathcal{T}$. In this case, the output label of this successful path is the 2-NAF with digits $\mathcal{D}$ of n.

(c) The set $\mathcal{D}$ is a NADS if and only if the only cycle in $\mathcal{T}$ with input label $0 \ldots 0$ is $0 \xrightarrow{0|0} 0$.

2.4.3.9 Frequency of Digits

Let $\mathcal{D} = \{0, 1, x\}$ be a NADS. For a non-negative integer n and the corresponding 2-NAF $\boldsymbol{\eta}(n)$ we define the number of occurrences of the digit $d \in \mathcal{D}$ by

$$f_d(n) := \sum_{j \geq 0} [\eta_j(n) = d] .$$

Set $\mathbf{f}(n) := (f_1(n), f_x(n))^t$. Further let X_N be a random variable, uniformly distributed on $\{0, \ldots, 2^N - 1\}$, and define $\mathbf{F}_N := (F_{1,N}(n), F_{x,N}(n))^t := \mathbf{f}(X_N)$. Heuberger and Prodinger [31, Theorem 9] gave the following distribution result for the random vector $\mathbf{F}_N$.

Theorem 2.4.34. *Let $\mathcal{D} = \{0, 1, x\}$ be a NADS and $\mathbf{F}_N = (F_{1,N}, F_{x,N})^t$ the number of occurrences of the digits 1 and x in the NAF of a randomly chosen integer in the set $\{0, \ldots, 2^N - 1\}$. Then we have*

$$\mathbb{E}(\mathbf{F}_N) = \frac{1}{6} N \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \mathbf{e} + \mathcal{O}\left(\frac{1}{2^N}\right),$$

$$\mathbb{V}(F_{d,N}) = \frac{11}{108} N + v_d + \mathcal{O}\left(\frac{N}{2^N}\right), \quad d \in \{1, x\},$$

$$\operatorname{Cov}(F_{1,N}, F_{x,N}) = -\frac{7}{108} N + w + \mathcal{O}\left(\frac{N}{2^N}\right)$$

for some constants $\mathbf{e} = (e_1, e_x)^t$, v_1, v_x, and w depending on x, which can be computed explicitly. Furthermore, the central limit theorem

$$\mathbb{P}\left(\frac{\mathbf{F}_N - \frac{1}{6} N \begin{pmatrix} 1 \\ 1 \end{pmatrix}}{\sqrt{N}} \leq \mathbf{z} \right) = \frac{9}{\sqrt{2}\pi} \iint\limits_{\mathbf{y} \leq \mathbf{z}} \exp\left(-\frac{1}{2} \mathbf{y}^t \cdot \begin{pmatrix} \frac{33}{2} & \frac{21}{2} \\ \frac{21}{2} & \frac{33}{2} \end{pmatrix} \cdot \mathbf{y} \right) \mathrm{d}\mathbf{y} + \mathcal{O}\left(\frac{1}{\sqrt{N}}\right)$$

holds uniformly with respect to $\mathbf{z}$, $\mathbf{z} \in \mathbb{R}^2$. Here $\mathbf{y} \leq \mathbf{z}$ means $y_j \leq z_j$ for $j \in \{1, 2\}$.

The proof makes use of probability generating functions and their properties. In Heuberger and Prodinger [31, Sections 5–7] there is also another approach for counting digits. For $N \in \mathbb{N}$ define

$$H_d(N) := \sum_{n=0}^{N-1} f_d(n) = \sum_{n=0}^{N-1} \sum_{j \geq 0} [\eta_j(n) = d] .$$

They proved the following theorem.

Theorem 2.4.35. *Let $\mathcal{D} = \{0, 1, x\}$ be a NADS, $d \in \{1, x\}$ and $N \in \mathbb{N}$. Then the number of occurrences of the digit d in the 2-NAFs with digits $\mathcal{D}$ of the integers $0, \ldots, N - 1$ equals*

$$H_d(N) = \frac{1}{6} N \log_2 N + N \, \psi_d(\log_2 N) + \mathcal{O}(N^\alpha),$$

where ψ_d is a 1-periodic continuous function and $\alpha < 2$ computable.

The proof uses a geometric approach based on the ideas of Delange [17]. In the following a short summary and some of the needed results were given, cf. Heuberger and Prodinger [31, Sections 5–7].

For $d \in \mathcal{D}$ there exists disjoint open subsets W_d of the unit interval $[0, 1]$ such that

$$\eta_\ell(n) = d \iff \left\{ n / 2^{\ell+2} \right\} \in W_d \tag{2.4.1}$$

for $\ell \geq 0$. The sum of the Lebesgue measures of the W_d equals 1. We may call the W_d *characteristic sets*. Even more, it can be proved that $\lambda(W_1) = \lambda(W_x) = \frac{1}{6}$ and therefore $\lambda(W_0) = \frac{2}{3}$.

In the case $x = -1$ those characteristic sets are just a finite union of intervals, cf. Prodinger [56] and Heuberger and Prodinger [30]. In general those sets have a fractal structure.

Of special interest is the Hausdorff dimension of the boundary of $W_0 \cup W_1 \cup W_x$, because this is exactly the exponent α of the error term in Theorem 2.4.35 on the previous page. It turns out that

$$\dim_H \partial(W_0 \cup W_1 \cup W_x) = \log_2 \rho(M_2)\,,$$

where $\rho(M_2)$ is the spectral radius of the adjacency matrix M_2 of an auxiliary automaton $\mathcal{A}_2$. The construction of $\mathcal{A}_2$ is skipped here, it can be found in Heuberger and Prodinger [31, Section 6]. There also bounds for $\rho(M_2)$ are given, as well as a table of the values of $\dim_H \partial(W_0 \cup W_1 \cup W_x)$ for some x.

With those results, Theorem 2.4.35 on the preceding page can be proved. The idea is to use the equivalence (2.4.1). The characteristic set W_d will be replaced by an appropriate approximation and the sum rewritten as integral. This integral is split up to get the main term, the periodic oscillating term and the error term. This is similar to the approach in Grabner, Heuberger and Prodinger [24, Section 4].

2.4.3.10 Non-Optimality

Now we want to discuss the optimality of the 2-NAFs with digits $\mathcal{D} = \{0, 1, x\}$. Optimality means that this 2-NAF-expansion for an integer n has minimal Hamming weight, i.e., number of non-zero digits, amongst all $\mathcal{D}$-expansions for n.

For the case $x = -1$ Reitwiesner [58] showed optimality, see also Section 2.4.1. For $x = 3$ optimality can be shown. For $x \leq -5$ the expansions are non-optimal. These results are stated in the following theorem of Heuberger and Prodinger [31, Theorem 18].

Theorem 2.4.36. *Let $\mathcal{D} = \{0, 1, x\}$ be a NADS. Then the Hamming weight of the 2-NAF of n with digits $\mathcal{D}$ is minimal among all $\mathcal{D}$-expansions of n (for all n) if and only if $x = -1$ or $x = 3$.*

For the proof that the expansion is non-optimal for $x \leq -5$, an counterexample is given in Heuberger and Prodinger [31, Section 8]. If $|x| + 3 = 2^g$ for some $g \geq 4$, then the integer $n = 2^{g+1} + 7$ is considered, otherwise the integer 3. In both cases an expansion with digits $\mathcal{D}$ and smaller Hamming weight than the 2-NAF-expansion is given.

The proof that the 2-NAF-expansion is optimal for $x = 3$ is of algorithmic nature. There a transducer is used.

2.5 Expansions for Frobenius-and-Add Scalar Multiplication Methods

In the previous section, we used the double-and-add method to perform the scalar multiplication on an elliptic curve. But there is also another way to do that. Let $\mathcal{E}$ be an elliptic curve defined over a field $\mathbb{F}_q$. We look at the group $\mathcal{E}(\mathbb{F}_{q^m})$ of rational points over a field extension $\mathbb{F}_{q^m}$ of $\mathbb{F}_q$ for an $m \in \mathbb{N}$. Consider the qth-power Frobenius endomorphism

$$\varphi\colon\ \mathcal{E}(\mathbb{F}_{q^m}) \longrightarrow \mathcal{E}(\mathbb{F}_{q^m})\,, \quad (x, y) \longmapsto (x^q, y^q)\,.$$

This map satisfies $f(\varphi) = 0$ for a quadratic monic polynomial $f \in \mathbb{Z}[T]$, cf. Koblitz [41, VI Section 1] or Silverman [62, V], or see Section 2.1. This means that for every point P on the

elliptic curve, we have $f(\varphi)(P) = \mathbf{0}$, where $\mathbf{0}$ denotes the neutral element of the point group. Thus we may identify φ with a solution $\tau \in \mathbb{C}$ satisfying $f(\tau) = 0$ and therefore we have an isomorphism between $\mathbb{Z}[\tau]$ and the endomorphism ring of $\mathcal{E}(\mathbb{F}_{q^m})$.

One example is the elliptic curve

$$\mathcal{E}_2 \colon Y^2 + XY = X^3 + aX^2 + 1 \qquad \text{with } a \in \{0, 1\}$$

defined over $\mathbb{F}_2$, cf. Koblitz [39]. The Frobenius map on this Koblitz curve[2] satisfies the relation $\varphi^2 - \mu\varphi + 2 = 0$ with $\mu = (-1)^{1-a}$. Thus the imaginary quadratic number $\tau = \frac{1}{2}\mu + \frac{1}{2}\sqrt{-7}$ can be used. Another example is the Koblitz curve

$$\mathcal{E}_3 \colon Y^2 = X^3 - X - \mu \qquad \text{with } \mu \in \{-1, 1\}$$

defined over $\mathbb{F}_3$. This curve was studied in Koblitz [40]. There we have $\varphi^2 - 3\mu\varphi + 3 = 0$ and therefore get $\tau = \frac{3}{2}\mu + \frac{1}{2}\sqrt{-3}$.

Now let $z \in \mathbb{Z}[\tau]$ and $P \in \mathcal{E}(\mathbb{F}_{q^m})$. If we write the element z as

$$z = \sum_{j=0}^{\ell-1} z_j \tau^j$$

for some digits z_j belonging to a digit set $\mathcal{D}$, then we can compute the action zP — or scalar multiplication if $z \in \mathbb{Z}$ — as

$$zP = \sum_{j=0}^{\ell-1} z_j \, \varphi^j(P) \,.$$

The evaluation can be done via a Horner scheme as used by the double-and-add method, except that instead of the doubling operation the Frobenius operation is used. The resulting method is therefore called *Frobenius-and-add method*, cf. Koblitz [39] and Solinas [63, 64].

Since the Frobenius operation on a point can be done much faster than the doubling of a point — especially when using normal bases — the Frobenius-and-add method is much faster than the classic double-and-add scalar multiplication. To get more details on normal bases, see for example Ash, Blake and Vanstone [1]. Beside other results, they showed that for m prime, a normal basis for $\mathbb{F}_{2^m}$ exists and is easy to construct.

So our interest is to make this Frobenius-and-add calculation as efficient as possible. In the following two sections the τ of the two examples $\mathcal{E}_2$ and $\mathcal{E}_3$ from above will be used to analyse τ-adic expansions of elements of $\mathbb{Z}[\tau]$.

2.5.1 Koblitz Curves in Characteristic Two and 2-NAFs

Let $\tau = \frac{1}{2}\mu + \frac{1}{2}\sqrt{-7}$ with $\mu \in \{-1, 1\}$. Such a τ comes from an elliptic curve

$$\mathcal{E}_2 \colon Y^2 + XY = X^3 + aX^2 + 1 \qquad \text{with } a \in \{0, 1\}$$

defined over $\mathbb{F}_2$, cf. Koblitz [39, Sections 2 and 6] and $\mu = (-1)^{1-a}$. We denote the group of rational points on the curve over $\mathbb{F}_{3^m}$ by $\mathcal{E}_2(\mathbb{F}_{2^m})$. Using τ-adic expansions for the scalar multiplication on that Koblitz curve was already mentioned by Koblitz [39, Section 6] and Meier and Staffelbach [43]. But none of them used the non-adjacency property. However, Solinas [64, Section 4.2] defines a τ-adic 2-NAF. As a digit set, $\mathcal{D} = \{0, +1, -1\}$ is used.

[2] In Koblitz [39] those curves were called *anomalous binary curves*. Later, for example see Solinas [64], such curves were called *Koblitz curves*. Sometimes the term *subfield curve* is used, too.

2.5.1.1 Existence and Uniqueness

First we want to know, whether the mentioned expansions exist, and whether they are unique or not. In Solinas [64, Theorem 1] we find the following.

Theorem 2.5.1. *Every element of the ring $\mathbb{Z}[\tau]$ has a unique τ-adic 2-NAF with digit set $\{0, +1, -1\}$.*

Additionally an algorithm is given to compute this τ-adic 2-NAF. Next, we want to know, how many expansions of a given length are there. We get the following, which is clearly equal to the 2-NAF with base 2 and digits $\{0, +1, -1\}$, cf. Section 2.4.1.4.

Theorem 2.5.2. *There are $t_\ell = \frac{1}{3}\left(2^{\ell+2} - (-1)^\ell\right)$ distinct 2-NAF representations of length ℓ.*

In conjunction with Koblitz curves, this was stated by Solinas [64, Section 3].

2.5.1.2 Length and Density

The average density is asymptotically $\frac{1}{3}$. This and other results on length and density are equal to the ones in Section 2.4.1.2, because the same digit set is used and the statements are independent from the used base.

Next we want to look at the expansion length of an element of $\mathbb{Z}[\tau]$. Bounds for that would be interesting. Solinas [64, Theorem 2] stated the following Theorem.

Theorem 2.5.3. *Let $d \in \mathbb{N}_0$, let $\ell > 2d$, and let α be a length-ℓ element of $\mathbb{Z}[\tau]$. Then*

$$\left(\sqrt{N_{\min}(d)} - \frac{\sqrt{N_{\max}(d)}}{2^{d/2} - 1}\right)^2 2^{\ell-d} < \mathcal{N}(\alpha) < \frac{\sqrt{N_{\max}(d)}}{(2^{d/2} - 1)^2} 2^\ell.$$

There, $N_{\min}(d)$ and $N_{\max}(d)$ is the minimum and the maximum, respectively, of the norm of all length d elements.

For $d = 15$ this means that

$$\log_2 \mathcal{N}(\alpha) - 0.5462682713 < \ell < \log_2 \mathcal{N}(\alpha) + 3.51559412,$$

when $\ell > 30$. Further in Solinas [64, Section 4.3] there are elements of $\mathbb{Z}[\tau]$ given, which are "close" to the mentioned bounds.

Another result in Solinas [64, Section 4.3] is that the Hamming weight of the τ-adic 2-NAF of the integer n is asymptotically $\frac{2}{3}\log_2 n$. Compared to the Hamming weight of the ordinary 2-NAF with base 2, cf. Section 2.4.1, this is a factor 2 larger. So the concept of τ-adic NAFs eliminated the doubling on the elliptic curve, but doubled the number of elliptic additions. This problem will be solved in the next section using reduced NAFs.

2.5.1.3 Reduced NAFs

In the previous section it was mentioned that the τ-adic NAF has twice the length of the ordinary NAF with base 2. To "fix" this disadvantage, we will introduce reduced NAFs, cf. Solinas [64, Section 6.3]. To do this, we need the following definition found in Solinas [64, Section 6.1].

Definition 2.5.4. Let $\mathcal{G}$ be a set of points on a Koblitz curve. Let $\boldsymbol{\gamma}$ and $\boldsymbol{\rho}$ be w-NAFs. Then $\boldsymbol{\gamma}$ and $\boldsymbol{\rho}$ are equivalent with respect to $\mathcal{G}$, if $\mathsf{value}(\boldsymbol{\gamma})\, P = \mathsf{value}(\boldsymbol{\rho})\, P$ for all $P \in \mathcal{G}$.

One way to get equivalent NAFs is mentioned in Meier and Staffelbach [43]. There the following result can be found.

Proposition 2.5.5. *If γ and ρ are elements of $\mathbb{Z}[\tau]$ with*

$$\gamma \equiv \rho \pmod{\tau^m - 1},$$

then

$$\gamma P = \rho P$$

for all $P \in \mathcal{E}_2(\mathbb{F}_{2^m})$. Thus $\mathsf{NAF}_w(\gamma)$ and $\mathsf{NAF}_w(\rho)$ are equivalent with respect to $\mathcal{E}_2(\mathbb{F}_{2^m})$.

We want to use this equivalence-concept on the Koblitz curve. So consider the group $\mathcal{E}_2(\mathbb{F}_{2^m})$ of rational points of the curve $\mathcal{E}_2$. It should be chosen in a way that the computation of the discrete logarithms of its elements is difficult. For example, the order of $\mathcal{E}_2(\mathbb{F}_{2^m})$ should be divisible by a large prime, cf. Menezes, Oorschot and Vanstone [44]. If m would be not a prime, then there would be a divisor 2^d with $d \mid m$.

An integer n is called *very nearly prime*, if it is of the form $N = f \cdot r$ with $f \in \{2, 4\}$ and $r > 2$. In Solinas [64, Section 4.1] a list of m (up to 512) is given, where the order of $\mathcal{E}_2(\mathbb{F}_{2^m})$ is very nearly prime. If this is the case, the subgroup of order r is called the *main subgroup*. Cryptographic operations are commonly performed on the main subgroup, cf. [33] and [54].

So let the order of $\mathcal{E}_2(\mathbb{F}_{2^m})$ be $f \cdot r$ and set $\delta := (\tau^m - 1)/(\tau - 1)$. This element has norm r, cf. Solinas [64, Section 4.1]. We get the following theorem.

Theorem 2.5.6. *Let P be a point in the main subgroup in a Koblitz curve of very nearly prime order. Let γ and ρ be elements of $\mathbb{Z}[\tau]$ with*

$$\gamma \equiv \rho \pmod{\delta}.$$

Then

$$\gamma P = \rho P.$$

Thus $\mathsf{NAF}_w(\gamma)$ and $\mathsf{NAF}_w(\rho)$ are equivalent with respect to the main subgroup.

This result can be found in Solinas [64, Theorem 3]. Now we are ready to define the reduced NAF, cf. Solinas [64, Section 6.3]

Definition 2.5.7. The *reduced 2-NAF* of a positive integer n with $n < r/2$ is defined as $\mathsf{NAF}_2(\rho)$, where $\rho := n \bmod \delta$.

According to the previous theorem, the reduced 2-NAF of n and the 2-NAF of n are equivalent with respect to the main subgroup. Thus we can use the reduced 2-NAF for curve operations (on the main subgroup). The advantage of the reduced 2-NAF can be found in Solinas [64, Theorem 4].

Theorem 2.5.8. *The average Hamming weight among reduced 2-NAFs is asymptotically $m/3$.*

The algorithms to get the reduced 2-NAFs and to perform the elliptic scalar multiplication are described in Solinas [64, Sections 7.1 and 7.2]. The concept of reduced 2-NAFs can be generalised to reduced w-NAFs. This is analogously to the 2-NAF case. See Solinas [64, Section 7.3] for details.

2.5.1.4 Optimality

The binary non-adjacent form minimises the Hamming weight, cf. Reitwiesner [58]. The same is true for our τ-adic 2-NAFs. In Avanzi, Heuberger and Prodinger [5, Theorem 1] and [4, Section 3] the following theorem is given. This result can also be found in Gordon [22, Theorem 3].

Theorem 2.5.9. *Let $z \in \mathbb{Z}[\tau]$. Then the Hamming weight of the τ-adic 2-NAF of z is minimal amongst all τ-expansions of z.*

Avanzi, Heuberger and Prodinger showed this result with two different proofs. One version is a "direct" proof, the other one an "automatic" proof. The latter one uses the transducers presented in Avanzi, Heuberger and Prodinger [5, Figures 1 and 2]. Those transducers, one for the case $\mu = -1$ and one for $\mu = 1$, compute the τ-adic NAFs of an integer from any other τ-expansion from right to left.

Additionally they gave two automata, see Avanzi, Heuberger and Prodinger [5, Figures 3 and 4], which accept minimal Hamming weight expansions. More precisely, they showed that those automata accept a τ-expansion if and only if it has minimal Hamming weight amongst all τ-expansion. This is formulated as Theorem 2 in Avanzi, Heuberger and Prodinger [5].

2.5.1.5 Point Halving

Consider the generic elliptic curve

$$\mathcal{E}\colon Y^2 + XY = X^3 + aX^2 + b$$

defined over $\mathbb{F}_{2^m}$, with $a, b \in \mathbb{F}_{2^m}$. Let $\mathcal{G} \leq \mathcal{E}(\mathbb{F}_{2^m})$ be a subgroup of large prime order, cf. "main subgroup" in Solinas [64] or Section 2.5.1.3. For a given point $P \in \mathcal{G}$ we want to find $R \in \mathcal{G}$ such that $2R = P$. This R is unique on $\mathcal{G}$, since point halving is an automorphism of $\mathcal{G}$. Such an R can be calculated by solving an equation system over $\mathbb{F}_{2^m}$. According to Knudsen [36] or Schroeppel [60, 59], this can be done efficiently. Have also a look at Fong, Hankerson, López and Menezes [21, Section 4] for details on point halving. Using this point halving the double-and-add method can be replaced by a halve-and-add algorithm.

There is a connection of using point halving in the scalar multiplication and τ-adic NAF representations, cf. Avanzi, Heuberger and Prodinger [5] and [7, Section 2.4]. The halving corresponds to use the digits $\pm\overline{\tau}$ additionally to the digit set $\{0, \pm 1\}$.

2.5.2 Koblitz Curves in Characteristic Two and Width-w NAFs

As in the previous section, we let $\tau = \frac{1}{2}\mu + \frac{1}{2}\sqrt{-7}$ with $\mu \in \{-1, 1\}$. As mentioned, such a τ comes from an elliptic curve

$$\mathcal{E}_2\colon Y^2 + XY = X^3 + aX^2 + 1 \qquad \text{with } a \in \{0, 1\}$$

defined over $\mathbb{F}_2$, cf. Koblitz [39, Sections 2 and 6] and $\mu = (-1)^{1-a}$. Again, we can generalise the 2-NAF results. We will take a general w and of course we need other (larger) digit sets. Clearly more digits mean that the resulting scalar multiplication method is more efficient, but there also more precomputation is needed.

2.5.2.1 Existence and Uniqueness

Let $w \geq 2$ and consider width-w non-adjacent forms. This was first mentioned by Solinas in [64, Section 7.3]. The digit set used consists of 0 and for every odd r with $1 \leq r \leq 2^w - 1$ we take one representative α of minimal norm with $\alpha \equiv r \pmod{\tau^w}$. Such a digit set is uniquely determined, cf. Avanzi, Heuberger and Prodinger [7, Theorem 2].

Solinas gave an algorithm to compute the w-NAF for an element of $\mathbb{Z}[\tau]$, cf. Solinas [64, Algorithm 4]. It was also noticed that it is sufficient to take the reduced w-NAFs, cf. Section 2.5.1.3.

The Solinas algorithm was changed (generalised) a little bit in Blake, Kumar Murty and Xu [14, Section 3]. For each $u \in \{1, 3, \ldots, 2^w - 1\}$ we take an $\alpha \in \mathbb{Z}[\tau]$ with

$$\alpha \equiv u \pmod{\tau^w}$$

and use these, and of course 0, as digit set. Note that α must not be a representative of minimal norm. They gave an algorithm to calculate the w-NAF and proved that this algorithm terminates under a suitable condition. This is formulated in the following theorem, cf. Blake, Kumar Murty and Xu [14, Theorem 2].

Theorem 2.5.10. *If $1 \in \mathcal{D}$ and representatives with norm smaller than 2^w are used, then the algorithm terminates, i.e., every element of $\mathbb{Z}[\tau]$ has a τ-adic w-NAF expansion with digits in $\mathcal{D}$.*

A simple corollary to this theorem is that the Solinas algorithm terminates, too, cf. Blake, Kumar Murty and Xu [14, Corollary 3]. Further, in Blake, Kumar Murty and Xu [12, Theorems 3.1 and 3.8] a uniqueness result was shown. Thus, the digit set used in the previous theorem is a w-NADS.

2.5.2.2 Width-w Non-adjacent Digit Sets

In the previous section, we had that a digit set of minimal norm representatives is a w-NADS. Now we want to consider also other digit sets. The question, whether a digit set is a w-NADS or not was studied in Avanzi, Heuberger and Prodinger [6] and [7, Section 2.1]. There they gave the following algorithmic characterisation.

Theorem 2.5.11. *Let $\mathcal{D}$ be a finite subset of $\mathbb{Z}[\tau]$ containing 0 and $w \geq 1$ be an integer. Let*

$$M := \left\lfloor \frac{\max\left\{ |d|^2 \,\middle|\, d \in \mathcal{D} \right\}}{(2^{w/2} - 1)^2} \right\rfloor.$$

Consider the directed graph $G = (V, A)$ defined by its set of vertices

$$V := \{0\} \cup \left\{ z \in \mathbb{Z}[\tau] \,\middle|\, |z|^2 \leq M, \tau \nmid z \right\}$$

and set of arcs

$$A := \left\{ (y, z) \in V^2 \,\middle|\, \text{There exists } d \in \mathcal{D}^\bullet \text{ and an integer } v \geq w \text{ s.t. } z = \tau^v y + d \right\}.$$

Then every element of $\mathbb{Z}[\tau]$ has a w-NAF representation with digits in $\mathcal{D}$ if and only if the following conditions are both satisfied:

1. *The set $\mathcal{D}$ contains a reduced residue system modulo τ^w.*
2. *In $G = (V, A)$, each vertex $z \in V$ is reachable from 0.*

If the previous equivalent conditions are fulfilled and $\mathcal{D}^\bullet$ is a reduced residue system modulo τ^w, then $\mathcal{D}$ is a w-NADS.

This theorem can be used with some special digit sets. First let

$$\mathcal{D} = \{0\} \cup \left\{ \pm 1, \pm 3, \ldots, \pm(2^{w-1} - 1) \right\},$$

where the second part of this union is a reduced residue system modulo τ^w. Avanzi, Heuberger and Prodinger [7, Example 2.10] showed that this digit set is a w-NADS for

$$w \in \{2, 3, 4, 5, 7, 8, 9, 10\}.$$

In the case $w = 6$, this is not true. The element $1 - \mu\tau$ has no 6-NAF.

The next considered digit set, see Avanzi, Heuberger and Prodinger [7, Section 2.3], is the *set of short τ-NAF representations for τ^w*. There a τ-NAF is meant to be a 2-NAF with digits 0, -1 and 1. Let $w \geq 1$ and $\mathcal{D}$ be a subset of

$$\{0\} \cup \{\mathsf{value}(\boldsymbol{\eta}) \mid \boldsymbol{\eta} \text{ is a } \tau\text{-NAF of length at most } w \text{ with } \eta_0 \neq 0\}$$

consisting of 0 and a reduced residue system modulo τ^w. Such a digit set is in almost all cases a w-NADS. This can be found in Avanzi, Heuberger and Prodinger [7, Theorem 3]. They gave also a list of exceptional cases consisting of four entries. All those cases fulfil $w = 3$. Additionally they gave bounds for the length of an expansion using short τ-NAFs. This length is approximately $2\log_2|z|$ for an $z \in \mathbb{Z}[\tau]$. Here approximately means that the difference to the true length can be bounded by a constant (depending on w).

An example for a set of short τ-NAF representations is the digit set

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

In this case $\mathcal{D}$ is a w-NADS for all $w \geq 2$, see also Avanzi, Heuberger and Prodinger [7, Theorem 3].

The digit set defined below can be used in conjunction with point halving, see Section 2.5.1.5 and Avanzi, Heuberger and Prodinger [7, Section 2.4], [4], and [5]. It was mentioned that a halving in the scalar multiplication corresponds to the extension of the digit set $\{0, -1, 1\}$ by $\pm\overline{\tau}$. For $w \geq 2$ the digit set is now defined by

$$\mathcal{D} = \{0\} \cup \left\{\pm\overline{\tau}^k \,\middle|\, 0 \leq k < 2^{w-2}\right\}.$$

It can be shown that $\mathcal{D}^\bullet$ is indeed a reduced residue system modulo τ^w, cf. Avanzi, Heuberger and Prodinger [7, Theorem 5]. Further they showed in that $\mathcal{D}$ is a w-NADS if $w \in \{2, 3, 4, 5, 6\}$ and that it is not a w-NADS if $w \in \{7, 8, 9, 10, 11, 12\}$. This can be found in [7, Theorem 6].

A comparison of the mentioned digit sets can be found in [7, Section 2.5]. In conjunction with applications to Koblitz curves those digit sets are discussed in Avanzi, Heuberger and Prodinger [7, Section 3]. There a detailed analysis of the number of needed operations is given, too.

2.5.2.3 Non-Optimality

First we will use minimal norm representatives for the digit set $\mathcal{D}$, cf. Section 2.5.2.1. Again, an expansion is optimal, if it minimises the Hamming weight among all expansion with the same digit set, but without the NAF-condition. The case $w = 2$ was handled in Section 2.5.1.4; optimality can be shown. For $w = 3$ optimality can be shown, too, see Avanzi, Heuberger and Prodinger [4] and [5]. For the cases $w \in \{4, 5, 6\}$ non-optimality was shown by Heuberger [26, Section 3 and Table 1]. Even more, a chaotic behaviour was mentioned which is stated in the following theorem, cf. Heuberger [26, Theorem 1].

Theorem 2.5.12. *For every positive integer ℓ, there exists elements $z_\ell, z'_\ell \in \mathbb{Z}[\tau]$ with the following properties:*

1. *The numbers z_ℓ and z'_ℓ are congruent modulo τ^ℓ.*

2. *For all optimal expansions* $\boldsymbol{\eta}$ *and* $\boldsymbol{\eta}'$ *of* z_ℓ *and* z'_ℓ*, respectively, the least significant digits* η_0 *and* η'_0 *differ.*

A consequence is that an optimal expansion cannot be computed by a deterministic transducer automaton or an online algorithm from right to left.

Those non-optimality and chaotic behaviour results stay true, if other digit sets like the ones in the previous section are used. If short τ-NAF representations are used as digit set, the statements are true for $w \in \{4, 5, 6\}$. When using powers of $\overline{\tau}$, then it can be shown for $w \in \{4, 5\}$. See Heuberger [26, Section 3 and Theorem 1] for details.

In contrast to the previous result is the following theorem.

Theorem 2.5.13. *Let* $w \geq 1$ *and* $\mathcal{D}$ *be a* w*-NADS. Then there is an algorithm to compute an optimal expansion of* $y \in \mathbb{Z}[\tau]$ *with digits in* $\mathcal{D}$ *in* $\mathcal{O}(\log |y|)$ *time, where the implicit constant depends on* $\mathcal{D}$*.*

This can be found in Heuberger [26, Theorem 2]. However, the $\mathcal{O}$-constant in the theorem may be quite huge, so practical miracles cannot be expected.

2.5.3 Koblitz Curves in Characteristic Three

Consider the Koblitz curve

$$\mathcal{E}_3 \colon Y^2 = X^3 - X - \mu \qquad \text{with } \mu \in \{-1, 1\}$$

defined over $\mathbb{F}_3$. This curve was studied in Koblitz [40]. We denote the group of rational points on the curve over $\mathbb{F}_{3^m}$ by $\mathcal{E}_3(\mathbb{F}_{3^m})$. There we have $\varphi^2 - 3\mu\varphi + 3 = 0$ for the Frobenius endomorphism φ and therefore get

$$\tau = \frac{3}{2}\mu + \frac{1}{2}\sqrt{-3}.$$

This τ will be used for our τ-adic expansions.

Let $\zeta \in \mathbb{Z}[\tau]$ be a sixth root of unity. We fix

$$\zeta = \frac{1}{2} - \frac{1}{2}\mu\sqrt{-3},$$

cf. Koblitz [40, Section 2] and Avanzi, Heuberger and Prodinger [8, Section 2]. Then clearly $\mathbb{Z}[\tau] = \mathbb{Z}[\zeta]$. Let our digit set be

$$\mathcal{D} := \{0\} \cup \left\{ \zeta^k \,\middle|\, 0 \leq k < 6 \right\}.$$

We will use $w = 2$ with this digit set, i.e., we will consider 2-NAFs.

2.5.3.1 Existence and Uniqueness

We get the following existence and uniqueness result, see Koblitz [40, Theorem 1].

Theorem 2.5.14. *Every element of* $\mathbb{Z}[\tau]$ *reduced modulo* $\tau^m - 1$ *has a unique* τ*-adic* 2*-NAF expansion with digits* $\mathcal{D}$*, in which at most* $(m+1)/2$ *digits are non-zero. Asymptotically on the average* 60% *of the digits are zero.*

As in Meier and Staffelbach [43] and Solinas [64, Section 6] it is sufficient to look at elements of $\mathbb{Z}[\tau]$ modulo $\tau^m - 1$, i.e., the remainders by dividing by $\tau^m - 1$. The reason is that we get the same point after scalar multiplication, because $(\tau^m - 1)P = \varphi^m P - P = \mathbf{0}$.

The average density of non-zero coefficients in a 2-NAF of length ℓ is $\frac{2}{5}\ell$, cf. Koblitz [40]. To calculate the 2-NAF for rational integers we can use the following connection between the τ-adic expansion of an n and its balanced ternary expansion. The theorem can be found in Avanzi, Heuberger and Prodinger [8, Theorem 1].

Theorem 2.5.15. *Let n be a rational integer given by its balanced ternary expansion $n = \sum_{j=0}^{\ell-1} x_j 3^j$ for $x_j \in \{0, 1, -1\}$. Then the 2-NAF of n is given by $\eta_{2\ell-2} \dots \eta_0$, where*

$$\eta_j = \begin{cases} 0 & \text{if } j \text{ is odd,} \\ x_{j/2}\zeta^{(j/2) \bmod 6} & \text{if } j \text{ is even.} \end{cases}$$

2.5.3.2 Elliptic Curve Algorithm vs. Non Elliptic Curve Algorithm

As the elliptic curve discrete logarithm can be reduced to the discrete logarithm, cf. Menezes, Okamoto and Vanstone [45], a comparison of an algorithm using elliptic curves and one "classical" is possible. The curve $\mathcal{E}_3$ is supersingular with $K = 6$, cf. end of Section 2.2.

In Koblitz [40, Section 5] the comparison is done for the *digital signature algorithm* DSA. There the elliptic curve $\mathcal{E}$ is taken and the number of operations in DSA and ECDSA is compared. Koblitz uses the Menezes-Okamoto-Vanstone embedding from $\mathcal{E}_3(\mathbb{F}_{3^m})$ to $\mathbb{F}_{3^{6m}}^{\times}$. Because of this embedding, both algorithms have the same security — using the field extension $\mathbb{F}_{3^m}$ for ECDSA and $\mathbb{F}_{3^{6m}}$ for DSA — and can be compared. The result is that ECDSA is approximately 12 times faster than DSA. This value does not depend on m.

2.5.4 Koblitz Curves in Characteristic Three and Width-w NAFs

Again, as in the previous section, consider the Koblitz curve

$$\mathcal{E}_3 \colon Y^2 = X^3 - X - \mu \qquad \text{with } \mu \in \{-1, 1\}$$

defined over $\mathbb{F}_3$, cf. Koblitz [40] with the corresponding

$$\tau = \frac{3}{2}\mu + \frac{1}{2}\sqrt{-3}.$$

Let $\zeta \in \mathbb{Z}[\tau]$ again be a sixth root of unity with

$$\zeta = \frac{1}{2} - \frac{1}{2}\mu\sqrt{-3}.$$

The idea is the same that Solinas [64] used for Koblitz curves in characteristic 2. Let $w \in \mathbb{N}$ with $w \geq 2$. As in the characteristic 2 case, the main loop in the Frobenius-and-add scalar multiplication is more efficient the larger w is, but also more pre-computation is needed. Blake, Kumar Murty and Xu [15, Section 3] proposed to use the digit set $\mathcal{D}$ constructed out of

$$\widetilde{\mathcal{D}} = \left\{ x + y\tau \;\middle|\; 0 \leq x \leq 3^{\lceil w/2 \rceil} - 1,\, 0 \leq y \leq 3^{\lfloor w/2 \rfloor} - 1 \text{ and } 3 \nmid x \right\}$$

by taking an element with least norm of the congruence class of $x + y\tau \in \mathcal{D}$ modulo τ^w. By means of this $\mathcal{D}$ they gave an algorithm, [15, Algorithm 3.1], which calculates the width-w τ-adic non-adjacent form of an element of $\mathbb{Z}[\tau]$. Its termination was shown in Blake, Kumar Murty and Xu [15, Theorem 3.1]. Further they studied the performance and gave the results in tables in [15, Section 4]. Have also a look at the later work of Blake, Kumar Murty and Xu, namely [12].

The average density of non-zero coefficients in a w-NAF of length ℓ is given by

$$\frac{2}{2w+1}n,$$

cf. Blake, Kumar Murty and Xu [15, Section 4.3].

The digit set of minimal norm representatives mentioned above can also be written down explicitly. This is a result of Avanzi, Heuberger and Prodinger [8, Theorem 2]. We have the following.

Theorem 2.5.16. *Let $w \geq 2$ and set*

$$\mathcal{D}_w = \left\{ a + b\mu\tau \,\middle|\, a \in \mathbb{Z}, b \in \mathbb{Z}, 3 \nmid a, 1 \leq a \leq 3^{w/2} - 2 \text{ and } -\frac{a}{3} < b < 3^{w/2-1} - \frac{2a}{3} \right\}$$

if w is even and

$$\mathcal{D}_w = \left\{ a + b\mu\tau \,\middle|\, a \in \mathbb{Z}, b \in \mathbb{Z}, 3 \nmid a, -3^{\lfloor \frac{w}{2} \rfloor} + 2 \leq b \leq 0, 1 - 2b \leq a \leq 3^{\lfloor \frac{w}{2} \rfloor} - b - 1 \right\}$$
$$\cup \left\{ (3^{\lfloor \frac{w}{2} \rfloor} - b) + b\mu\tau \,\middle|\, b \in \mathbb{Z}, 3 \nmid b, -\frac{3^{\lfloor \frac{w}{2} \rfloor} - 1}{2} \leq b \leq 0 \right\}$$

if w is odd. Set

$$\mathcal{D} := \{0\} \cup \bigcup_{0 \leq k < 6} \zeta^k \mathcal{D}_w \ .$$

Then $\mathcal{D}$ consists of 0 and exactly one representative of minimum norm of every residue class modulo τ^w. In particular, $\mathcal{D}$ is a w-NADS.

2.5.5 Other Bases

A more general result is given in Blake, Kumar Murty and Xu [12]. They analysed Euclidean imaginary quadratic number fields. In particular, they gave existence and uniqueness results for τ-adic w-NAFs concerning the fields $\mathbb{Q}(\sqrt{-1})$, $\mathbb{Q}(\sqrt{-2})$, $\mathbb{Q}(\sqrt{-3})$, $\mathbb{Q}(\sqrt{-7})$ and $\mathbb{Q}(\sqrt{-11})$.

The first result is general. Let F be an Euclidean imaginary quadratic number field and O_F its ring of integers. Let $\tau \in O_F$ with $|\tau| > 1$, $w \in \mathbb{N}$, and suppose that $|\tau^w|^2 \geq 12$. The digit set is defined as follows. Let

$$R = \{k \in O_F \mid \tau \nmid k\} .$$

From each congruence class C of R modulo τ^w coprime to τ we fix a digit c in the following way. If there is a unit in C, we choose this unit as c. Otherwise an element with $|c| < |\tau|$ is fixed. Our digit set $\mathcal{D}$ consists then of 0 and all such c. Blake, Kumar Murty and Xu [12, Theorem 3.1] proved the following.

Theorem 2.5.17. *Every element $k \in O_F$ has a unique τ-adic w-NAF expansion with digits in $\mathcal{D}$.*

In the following, we will give the results for the fields mentioned above, except the two that were already handled in the Sections 2.5.1 to 2.5.4. For all those cases algorithms to calculate the w-NAF expansions can be found in Blake, Kumar Murty and Xu [12, Section 4].

2.5.5.1 Integers in $\mathbb{Q}(\sqrt{-1})$

We use $\tau = 1 + \sqrt{-1}$. Then we have $\mathbb{Z}[\tau] = \mathbb{Z}[\sqrt{-1}]$. In Blake, Kumar Murty and Xu [12, Section 3] the following is stated. Let

$$R = \left\{ x + y\tau \,\middle|\, 0 \leq x \leq 2^{\lceil w/2 \rceil} - 1,\ 0 \leq y \leq 2^{\lfloor w/2 \rfloor} - 1 \text{ and } 2 \nmid x \right\}$$

and let the digit set $\mathcal{D}$ consist of 0, the units 1, -1, $\sqrt{-1}$, and $-\sqrt{-1}$, and $\widetilde{z} \in \mathbb{R}$ with $1 < |\widetilde{z}| < |\tau^w|$, such that $\widetilde{z}$ is in the same residue class as $z \in R$ modulo τ^w. We get the following theorem, cf. Blake, Kumar Murty and Xu [12, Theorem 3.2].

Theorem 2.5.18. *If $w > 2$ then every element of $\mathbb{Z}[\tau]$ has a unique τ-adic w-NAF expansion with digits in $\mathcal{D}$.*

This theorem is not true for $w \leq 2$. There, we get the following results.

Theorem 2.5.19. *Every element of $\mathbb{Z}[\tau]$ has a τ-adic 2-NAF expansion with digits in*

$$\mathcal{D} = \left\{0, 1, -1, \sqrt{-1}, -\sqrt{-1}\right\}.$$

Theorem 2.5.20. *Every element of $\mathbb{Z}[\tau]$ has a τ-adic expansion (1-NAF) with digits in*

$$\mathcal{D} = \{0, 1, -1\}.$$

Those results can be found in Blake, Kumar Murty and Xu [12, Theorem 3.3]. Note that the expansions in the previous two theorems must not be unique. Counterexamples can be found in Blake, Kumar Murty and Xu [12], too.

2.5.5.2 Integers in $\mathbb{Q}\left(\sqrt{-2}\right)$

We use $\tau = 1 + \sqrt{-2}$. In Blake, Kumar Murty and Xu [12, Section 3] the following is stated. Let

$$R = \left\{x + y\tau \,\middle|\, 0 \leq x \leq 2^{\lceil w/2 \rceil} - 1,\, 0 \leq y \leq 2^{\lfloor w/2 \rfloor} - 1 \text{ and } 2 \nmid x\right\}$$

and let the digit set $\mathcal{D}$ consist of 0, the units 1 and -1, and $\widetilde{z} \in \mathbb{R}$ with $1 < |\widetilde{z}| < |\tau^w|$, such that $\widetilde{z}$ is in the same residue class as $z \in R$ modulo τ^w. We get the following theorem, cf. Blake, Kumar Murty and Xu [12, Theorem 3.4].

Theorem 2.5.21. *If $w > 2$ then every element of $\mathbb{Z}[\tau]$ has a unique τ-adic w-NAF expansion with digits in $\mathcal{D}$.*

Again, as in the $\mathbb{Q}\left(\sqrt{-1}\right)$ case, this theorem is not true for $w \leq 2$. There, we get the following results.

Theorem 2.5.22. *Every element of $\mathbb{Z}[\tau]$ has a τ-adic 2-NAF expansion with digits in*

$$\mathcal{D} = \{0, 1, -1, 1 + \tau\}.$$

Theorem 2.5.23. *Every element of $\mathbb{Z}[\tau]$ has a τ-adic expansion (1-NAF) with digits in*

$$\mathcal{D} = \{0, 1, -1\}.$$

Those results can be found in Blake, Kumar Murty and Xu [12, Theorem 3.5]. Note that there, again, is no uniqueness result given.

2.5.5.3 Integers in $\mathbb{Q}\left(\sqrt{-11}\right)$

We use $\tau = \frac{1}{2} + \frac{1}{2}\sqrt{-11}$. In Blake, Kumar Murty and Xu [12, Section 3] the following is stated. Let

$$\mathcal{D} = \{0\} \cup \{-1, 1\} \cup \{c_i \,|\, 1 < i < 3^w - 1 \text{ with } 3 \nmid i,\, c_i \equiv i \pmod{\tau^w} \text{ and } |c_i| < |\tau^w|\}$$

We get the following theorem, cf. Blake, Kumar Murty and Xu [12, Theorem 3.10].

Theorem 2.5.24. *If $w \in \mathbb{N}$ then every element of $\mathbb{Z}[\tau]$ has a unique τ-adic w-NAF expansion with digits in $\mathcal{D}$.*

Chapter 3

New Results

This chapter contains new results, which were developed during the work on the master's thesis.

In Section 3.1 an analysis of the digits in the case of Koblitz curves in Characteristic Three can be found. There w-NAF expansions of the rational integers are considered.

The Sections 3.2 to 3.10 contain the analysis of the occurrence of a digit in a general case and the necessary prerequisites for the proof. Those sections are a joint work with my supervisor Clemens Heuberger.

An overview of the requirements on τ and digit set $\mathcal{D}$ for the different sections, definitions, theorems, etc. can be found in Table 3.0.1 on the next page.

3.1 Analysis of 2-NAFs in Conjunction with Koblitz Curves in Characteristic Three

Let q and r be integers satisfying $q \geq 2$ and $0 \leq r \leq q-2$. Let the $\langle q, r\rangle$ number system be the positional number system with base q and digits $-r, 1-r, \ldots, q-1-r$. Flajolet and Ramshaw [20, Theorem P] gave the following theorem.

Theorem 3.1.1. *Let d be a non-zero digit in the $\langle q, r\rangle$ number system. Let $n \in \mathbb{N}$, let $\rho(n)$ denote the number of times that the digit d is used when n is expressed in the $\langle q, r\rangle$ number system, and let $F(d, n)$ denote the appropriately truncated summation of ρ, in particular,*

$$F(d,n) = \left(1 - \frac{r}{q-1}\right)\rho(0) + \rho(1) + \rho(2) + \cdots + \rho(n-1) + \left(\frac{r}{q-1}\right)\rho(n)$$

Then, there exists a continuous, nowhere differentiable function $P : \mathbb{R} \longrightarrow \mathbb{R}$, periodic with period 1, such that

$$F(d,n) = \frac{n \log_q n}{q} + n\,P\big(\log_q n\big) \quad \text{for } n \geq 1.$$

The Fourier series $P(x) = \sum_{\mu\in\mathbb{Z}} p_\mu e^{2\pi i \mu x}$ of P converges absolutely. Finally, if we set

$$m := d \bmod q$$

and define ξ and η by the formulas

$$\xi = \frac{m}{q} - \frac{r}{q(q-1)}$$

	short description	τ	digit set $\mathcal{D}$
Section 3.2	Voronoi cells	i-q	——
Lemma 3.3.3 on page 44	complete residue system	alg	——
Definition 3.3.5 on page 45	minimal norm representatives digit set	i-q	——
Definition 3.3.7 on page 46	width-w non-adjacent forms	gen	fin
Proposition 3.3.8 on page 46	continuity of value	gen	fin
Definition 3.3.11 on page 47	width-w non-adjacent digit set	gen	fin
Theorem 3.4.1 on page 48	full block length distribution theorem	alg	RRS
Section 3.5	bounds for the value	i-q	MNR
Theorem 3.6.1 on page 60	existence theorem for lattice points	i-q	MNR
Theorem 3.6.5 on page 62	existence theorem for $\mathbb{C}$	i-q	MNR
Definition 3.7.1 on page 62	fundamental domain $\mathcal{F}$	gen	fin
Proposition 3.7.2 on page 62	compactness of the fundamental domain	gen	fin
Corollary 3.7.4 on page 63	tiling property	i-q	MNR
Remark 3.7.5 on page 63	iterated function system	gen	fin
Proposition 3.7.7 on page 64	characterisation of the boundary	i-q	MNR
Proposition 3.7.8 on page 65	upper bound for the dimension of $\partial\mathcal{F}$	i-q	MNR
Section 3.8	cell rounding operations	i-q	——
Section 3.9	characteristic sets	i-q	MNR
Theorem 3.10.1 on page 78	counting the occurences of a digit	i-q	MNR

Abbreviations for τ (general: $\tau \in \mathbb{C}$ with $|\tau| > 1$)

gen	$\tau \in \mathbb{C}$
alg	τ algebraic integer
i-q	τ imaginary quadratic algebraic integer

Abbreviations for digit sets (general: $\mathcal{D} \subseteq \mathbb{Z}[\tau]$, $0 \in \mathcal{D}$)

fin	finite digit set
RRS	reduced residue system digit set
MNR	minimal norm representatives digit set

Table 3.0.1: Overview of requirements.

and

$$\eta = \frac{m+1}{q} - \frac{r}{q(q-1)},$$

the coefficients p_μ *are given by*

$$p_0 = \log_q \Gamma(\xi) - \log_q \Gamma(\eta) - \frac{1}{q \ln q} - \frac{1}{2q}$$

and

$$p_\mu = \frac{\zeta(\chi_\mu, \xi) - \zeta(\chi_\mu, \eta)}{(\ln q)\, \chi_\mu\, (1 + \chi_\mu)} \qquad \text{for } \chi_\mu = \frac{2\pi i \mu}{\ln q} \text{ and } \mu \neq 0.$$

Flajolet and Ramshaw [20, Section 3] used this theorem to calculate the occurrence of digits in the balanced ternary case, i.e., in the number system $\langle 3, 1\rangle$.

We modify this theorem a little bit. We want to count digits in given positions in the $\langle q, r\rangle$ number system. More precisely, we are interested in the occurrence of a digit in all numbers up to n at positions with index in a residue class $b + M\mathbb{Z}$. The result is the following theorem.

Theorem 3.1.2. *Let d be a non-zero digit in the $\langle q, r\rangle$ number system. Let $b \in \mathbb{Z}$ and $M \in \mathbb{N}$. Let $n \in \mathbb{N}$, let $\rho(n)$ denote the number of times that the digit d is used at positions in $b + M\mathbb{Z}$*

when n is expressed in the $\langle q, r\rangle$ number system, and let $F_{b+M\mathbb{Z}}(d,n)$ denote the appropriately truncated summation of ρ, in particular,

$$F_{b+M\mathbb{Z}}(d,n) = \left(1 - \frac{r}{q-1}\right)\rho(0) + \rho(1) + \rho(2) + \cdots + \rho(n-1) + \left(\frac{r}{q-1}\right)\rho(n)\,.$$

Then, there exists a piecewise continuous, piecewise nowhere differentiable function $P\colon \mathbb{R} \longrightarrow \mathbb{R}$, periodic with period M and a piecewise constant function $Q\colon \mathbb{R} \longrightarrow \mathbb{R}$, periodic with period M, such that

$$F_{b+M\mathbb{Z}}(d,n) = \frac{n\log_q n}{qM} + n\,P\big(\log_q n\big) + n\,Q\big(\log_q n\big) \quad \text{for } n \geq 1.$$

We have $P(x) = P_c(x)$ where the parameters c and x fulfil the relation

$$c = (\lfloor x \rfloor + 1 - b) \bmod M.$$

If we set

$$m := d \bmod q$$

and define ξ and η by the formulas

$$\xi = \frac{m}{q} - \frac{r}{q(q-1)}$$

and

$$\eta = \frac{m+1}{q} - \frac{r}{q(q-1)},$$

the coefficients $p_{c,\mu} = c_\mu + d_\mu$ of the Fourier series

$$P_c(x) = \sum_{\mu\in\mathbb{Z}} p_{c,\mu} e^{2\pi i\mu x}$$

of P_c are given by

$$c_\mu = \begin{cases} \frac{1}{q} + \frac{1}{2qM} & \text{for } \mu = 0, \\ \frac{1}{2\pi i\mu qM} & \text{for } \mu \neq 0 \end{cases}$$

and

$$d_\mu = \frac{1}{\ln q} \sum_{0\leq k} H_k(1 + \chi_\mu)$$

with $\chi_\mu = 2\pi i\mu/\ln q$ and

$$H_k(z) = \frac{1}{qzQ_k^z}\left(\frac{1}{q^z} - 1\right) - \frac{1}{qz}\left(V_z(qQ_k) - V_z(Q_k)\right) + \frac{1}{z}\sum_{j=Q_k}^{qQ_k-1}\left(V_z(j+\eta) - V_z(j+\xi)\right).$$

with $Q_k = q^{c+kM-2}$ and $V_1(v) = \ln v$ and $V_z(v) = \frac{1}{1-z}\frac{1}{v^{z-1}}$ for $z \neq 1$.

The function Q is given by

$$Q(x) = -\frac{1}{q}\left\{\frac{1 + \lfloor x \rfloor}{M}\right\}.$$

The coefficients q_μ of its Fourier series

$$Q(x) = \sum_{\mu\in\mathbb{Z}} q_\mu e^{2\pi i\mu x/M}$$

are given by

$$q_0 = -\frac{M-1}{2qM}.$$

and

$$q_\mu = -\frac{ie^{2\pi i\mu/M}}{2\pi q\mu}.$$

for $\mu \neq 0$.

If we set $M = 1$, then we get the statement of Theorem 3.1.1 on page 31, at least when the expression for the Fourier coefficients is simplified (which is possible in this case).

Remark 3.1.3. The Fourier coefficients of P can be calculated from the Fourier coefficients of the M different P_c, since

$$P(x) = \sum_{c=0}^{M-1} P_c(x)\, R_c(x)\,.$$

There the function R_c is the characteristic function of the set $\bigcup_{y\in c-1+b+M\mathbb{Z}} [y, y+1)$. The function R_c is clearly M periodic and its Fourier coefficients $r_{c,\mu}$ are easy to calculate. The Fourier coefficients p_μ of the Fourier series

$$P(x) = \sum_{\mu\in\mathbb{Z}} p_\mu e^{2\pi i\mu x/M}$$

then follow by "rescaling" the coefficients $p_{c,\mu}$ to period M, a convolution of them with the $r_{c,\mu}$, and summing up over all c.

The proof of Theorem 3.1.2 on page 32 follows the proof of Theorem 3.1.1 on page 31 which can be found in the appendix of Flajolet and Ramshaw [20]. The difference is, that here in our sums not all summands are taken. Thus the Fourier coefficients can not be written in such a nice form as in Theorem 3.1.1 on page 31.

The idea of the proof comes from the balanced ternary case, cf. Flajolet and Ramshaw [20, Sections 2 and 3]. In Table 3.1.1 on the next page the column $k = 0$ consists of the repetition of the block $01\bar{1}$. Unfortunately, the kth column does not consists of the repetition of the block $0^{3^k}1^{3^k}\bar{1}^{3^k}$. This problem can be fixed using blocks of the form $0^{3^k/2}1^{3^k}\bar{1}^{3^k}0^{3^k/2}$. This means that our first column $k = 0$ consists of blocks $0^{1/2}1\bar{1}0^{1/2}$. The summation function $F_{b+M\mathbb{Z}}$ will take this in account.

Proof of Theorem 3.1.2. Let d be a fixed non-zero digit. At the beginning of this proof we want to consider each position k separately. Let n be written in the $\langle q, r\rangle$ system, and let $\rho_k(n)$ be 1 if the digit d appears in the kth position of n. Further define

$$F_k(d,n) := \left(1 - \frac{r}{q-1}\right)\rho_k(0) + \rho_k(1) + \rho_k(2) + \cdots + \rho_k(n-1) + \left(\frac{r}{q-1}\right)\rho_k(n)\,.$$

Each position will contain d with probability $1/q$, so we set

$$t_k(n) := F_k(d,n) - \frac{n}{q}.$$

Those functions $t_k(n)$ can be expressed as a scaled version of a function $t\colon \mathbb{R} \longrightarrow \mathbb{R}$, which is defined by

$$t(x) = \begin{cases} -\frac{x}{q} & \text{if } 0 \le x \le \xi, \\ \frac{q-1}{q}x - \xi & \text{if } \xi \le x \le \eta, \\ \frac{1}{q} - \frac{x}{q} & \text{if } \eta \le x \le 1 \end{cases}$$

n	$k=3$	$k=2$	$k=1$	$k=0$
0				0
1				1
2			1	$\bar{1}$
3			1	0
4			1	1
5		1	$\bar{1}$	$\bar{1}$
6		1	$\bar{1}$	0
7		1	$\bar{1}$	1
8		1	0	$\bar{1}$
9		1	0	0
10		1	0	1
11		1	1	$\bar{1}$
12		1	1	0
13		1	1	1
14	1	$\bar{1}$	$\bar{1}$	$\bar{1}$
15	1	$\bar{1}$	$\bar{1}$	0

Table 3.1.1: Balanced Ternary System

and $t(x+1) = t(x)$. We get

$$t_k(n) = q^{k+1}\, t\big(n/q^{k+1}\big)\,,$$

cf. Flajolet and Ramshaw [20, Proof of Theorem P] Each n can be expressed by $\ell+1$ digits, where $\ell = \lfloor \log_q n \rfloor + 1$.

Now consider our fixed residue class $b + M\mathbb{Z}$. We have

$$F_{b+M\mathbb{Z}}(d,n) = \sum_{0\le k\le \ell} [k \equiv b \pmod M]\, F_k(d,n) = \sum_{0\le k\le \ell} [k \equiv b \pmod M] \left(\frac{n}{q} + t_k(n)\right).$$

Pulling out the summand n/q, inserting $t_k(n)$ and changing the summation index from k to $\ell - k$ yields

$$F_{b+M\mathbb{Z}}(d,n) = \frac{n}{q}\left(\left\lfloor \frac{\ell}{M} \right\rfloor + 1\right) + \sum_{0\le k\le \ell} [\ell - k \equiv b \pmod M]\, q^{\ell-k+1}\, t\left(\frac{n}{q^{\ell-k+1}}\right).$$

Since $t\big(n/q^{\ell-k+1}\big)$ is zero for $k > \ell$, we can rewrite this as

$$F_{b+M\mathbb{Z}}(d,n) = \frac{n}{q}\left(\left\lfloor \frac{\ell}{M} \right\rfloor + 1\right) + q^{\ell+1}\, h\left(\frac{n}{q^{\ell+1}}\right)$$

where

$$h(x) = \sum_{0\le k} [k \equiv \ell - b \pmod M] \frac{t\big(q^k x\big)}{q^k} = \sum_{0\le k} \frac{t\big(q^{c+kM} x\big)}{q^{c+kM}}$$

with $c = (\ell - b) \bmod M$. Since t is bounded and $q > 1$, the sum in $h(x)$ converges absolutely and uniformly by the Weierstrass M-test. Therefore h is continuous because t is continuous.

To see that h is nowhere differentiable, assume by contradiction that h is differentiable at y. Consider its difference quotient over the intervals I_j, $j \in \mathbb{N}$ with

$$I_j = [a_j, b_j] = \left[\frac{p_j}{q^j} - \frac{r}{q^j(q-1)}, \frac{p_j+1}{q^j} - \frac{r}{q^j(q-1)}\right].$$

There the p_j are chosen such that $y \in I_j$. The intervals I_j are nested, since

$$\frac{p_j}{q^j} - \frac{r}{q^j(q-1)} = \frac{p_j q - r}{q^{j+1}} - \frac{r}{q^{j+1}(q-1)}.$$

The difference quotient

$$\frac{h(b_j) - h(a_j)}{b_j - a_j} = \sum_{0 \le k} \frac{t\big(q^{c+kM} b_j\big) - t\big(q^{c+kM} a_j\big)}{q^{c+kM-j}}$$

would converge to $h'(y)$ for $j \to \infty$. If $c + kM \ge j$, then the kth term is zero, because of the periodicity of t. If $c+kM < j$, then we will get either $-1/q$ or $1-1/q$, since our scaled t is linear in the interval I_j. But this means that the limit $j \to \infty$ does not exist, and so h is nowhere differentiable.

By inserting ℓ and using $z = \lfloor z \rfloor + \{z\}$ we obtain

$$\begin{aligned} F_{b+M\mathbb{Z}}(d,n) &= \frac{n}{q}\left(\frac{\lfloor \log_q n \rfloor + 1}{M} - \left\{\frac{\lfloor \log_q n \rfloor + 1}{M}\right\} + 1\right) \\ &\quad + q^{\log_q n - \{\log_q n\} + 2}\, h\Big(q^{\log_q n - \lfloor \log_q n \rfloor - 2}\Big) \\ &= \frac{n}{q}\frac{\log_q n}{M} - \frac{n}{q}\left\{\frac{1 + \lfloor \log_q n \rfloor}{M}\right\} + \frac{n}{q} + \frac{n}{q}\frac{1 - \{\log_q n\}}{M} \\ &\quad + nq^{2-\{\log_q n\}}\, h\Big(q^{\{\log_q n\}-2}\Big), \end{aligned}$$

and therefore

$$F_{b+M\mathbb{Z}}(d,n) = \frac{n \log_q n}{qM} + n\, P_c\big(\log_q n\big) + n\, Q\big(\log_q n\big)$$

with

$$P_c(x) := \frac{1}{q} + \frac{1 - \{x\}}{qM} + q^{2-\{x\}}\, h\Big(q^{\{x\}-2}\Big)$$

and

$$Q(x) := -\frac{1}{q}\left\{\frac{1 + \lfloor x \rfloor}{M}\right\}.$$

Clearly P_c is 1-periodic and piecewise continuous (possible discontinuities at integer values), and Q is M-periodic and piecewise constant with discontinuities at $x \in \mathbb{Z}$. From the construction of P out of the M different P_c, the properties of P follow directly, especially that P is M-periodic. We want to compute the Fourier coefficients of P_c and Q. So let

$$P_c(x) = \sum_{\mu \in \mathbb{Z}} p_\mu e^{2\pi i \mu x}$$

with

$$p_\mu = \int_{u=0}^{1} P_c(u)\, e^{-2\pi i \mu u}\, \mathrm{d}u = \underbrace{\int_{u=0}^{1} \left(\frac{1}{q} + \frac{1-u}{qM}\right) e^{-2\pi i \mu u}\, \mathrm{d}u}_{=:c_\mu} + \underbrace{\int_{u=0}^{1} q^{2-u}\, h\big(q^{u-2}\big)\, e^{-2\pi i \mu u}\, \mathrm{d}u}_{=:d_\mu}.$$

It is easy to see that

$$c_\mu = \begin{cases} \frac{1}{q} + \frac{1}{2qM} & \text{for } \mu = 0, \\ \frac{1}{2\pi i \mu q M} & \text{for } \mu \neq 0. \end{cases}$$

Now consider the coefficient d_μ. Inserting h and changing the order of integration and summation yields

$$d_\mu = \sum_{0 \le k} \int_{u=0}^{1} q^{2-u-c-kM}\, t\big(q^{c+kM+u-2}\big)\, e^{-2\pi i \mu u}\, \mathrm{d}u.$$

Now we substitute $v = q^{c+kM+u-2}$, $\mathrm{d}v = v \ln q\, \mathrm{d}u$, $u = \log_q v - c - kM + 2$ and set $Q_k = q^{c+kM-2}$ to obtain

$$d_\mu = \frac{1}{\ln q} \sum_{0 \le k} \int_{v=Q_k}^{qQ_k} \frac{t(v)}{v^2} e^{-2\pi i \mu \left(\log_q v - c - kM + 2\right)}\, \mathrm{d}v.$$

By rewriting

$$e^{-2\pi i \mu \left(\log_q v - c - kM + 2\right)} = e^{-2\pi i \mu \log_q v} = e^{-2\pi i \mu \ln v / \ln q} = v^{-2\pi i \mu / \ln q} = v^{-\chi_\mu}$$

with $\chi_\mu = 2\pi i \mu / \ln q$ we get

$$d_\mu = \frac{1}{\ln q} \sum_{0 \le k} \int_{v=Q_k}^{qQ_k} \frac{t(v)}{v^{2+\chi_\mu}}\, \mathrm{d}v.$$

Now we define

$$H_k(z) := \int_{v=Q_k}^{qQ_k} \frac{t(v)}{v^{1+z}}\, \mathrm{d}v$$

and therefore get

$$d_\mu = \frac{1}{\ln q} \sum_{0 \le k} H_k(1 + \chi_\mu)$$

The function $t(v)$ can be written

$$t(v) = \int_{x=0}^{v} \left(\lfloor x - \xi \rfloor - \lfloor x - \eta \rfloor - \frac{1}{q} \right) \mathrm{d}x,$$

so we can use integration by parts at H_k to obtain

$$H_k(z) = -\left. \frac{t(v)}{z v^z} \right|_{v=Q_k}^{qQ_k} + \frac{1}{z} \int_{v=Q_k}^{qQ_k} \frac{1}{v^z} \left(\lfloor x - \xi \rfloor - \lfloor x - \eta \rfloor - \frac{1}{q} \right) \mathrm{d}v.$$

The difference $\lfloor x - \xi \rfloor - \lfloor x - \eta \rfloor$ is 1-periodic and piecewise constant. More precisely in the interval $[0, 1)$ the function is 1 on $[\xi, \eta)$ and 0 elsewhere. We set

$$V_z(v) = \int \frac{1}{v^z}\, \mathrm{d}v.$$

Thus we get $V_1(v) = \ln v$ and $V_z(v) = \frac{1}{1-z} \frac{1}{v^{z-1}}$ for $z \neq 1$. Using that t is $-\frac{1}{q}$ at integer values and the results above yields

$$H_k(z) = \frac{1}{qzQ_k^z} \left(\frac{1}{q^z} - 1 \right) - \frac{1}{qz} \left(V_z(qQ_k) - V_z(Q_k) \right) + \frac{1}{z} \sum_{j=Q_k}^{qQ_k - 1} \left(V_z(j + \eta) - V_z(j + \xi) \right).$$

Next we want to compute the Fourier coefficients of Q. Let

$$Q(x) = \sum_{\mu\in\mathbb{Z}} q_\mu e^{2\pi i\mu x/M}$$

with

$$q_\mu = \frac{1}{M}\int_{u=0}^{M} Q(u)\, e^{-2\pi i\mu u/M}\,\mathrm{d}u = \frac{1}{M}\int_{u=-1}^{M-1} Q(u)\, e^{-2\pi i\mu u/M}\,\mathrm{d}u$$

Since Q is piecewise constant, we obtain

$$q_\mu = \frac{1}{qM^2}\int_{u=-1}^{M-1} (1+\lfloor u\rfloor)\, e^{-2\pi i\mu u/M}\,\mathrm{d}u = \frac{1}{qM^2}\sum_{j=-1}^{M-2}(1+j)\int_{u=j}^{j+1} e^{-2\pi i\mu u/M}\,\mathrm{d}u,$$

and using computer algebra software we get for $\mu \neq 0$

$$q_\mu = \frac{ie^{2\pi i\mu/M}}{2\pi q\mu}.$$

Clearly

$$q_0 = \frac{M-1}{2qM}.$$

Therefore all Fourier coefficients are calculated and the proof is finished. □

We will now use Theorem 3.1.2 on page 32 to count the non-zero digits in the numbers up to $n \in \mathbb{N}$, when they are written with base

$$\tau = \frac{3}{2}\mu + \frac{1}{2}\sqrt{-3}$$

and digit set

$$\mathcal{D} = \{0\} \cup \left\{\zeta^k \,\middle|\, 0 \le k < 6\right\}$$

where $\zeta \in \mathbb{Z}[\tau]$ is a sixth root of unity, cf. Koblitz [40].

Avanzi, Heuberger and Prodinger [8] proved the following connection between the τ-adic 2-NAF of an n and its balanced ternary expansion.

Theorem 3.1.4. *Let n be a rational integer given by its balanced ternary expansion $n = \sum_{j=0}^{\ell-1} x_j 3^j$ for $x_j \in \{0,1,-1\}$. Then the 2-NAF of n is given by $\eta_{2\ell-2}\dots\eta_0$, where*

$$\eta_j = \begin{cases} 0 & \text{if } j \text{ is odd,} \\ x_{j/2}\zeta^{(j/2) \bmod 6} & \text{if } j \text{ is even.} \end{cases}$$

Combining this result with Theorem 3.1.2 on page 32 leads directly to the following corollary.

Corollary 3.1.5. *Let $\eta \in \mathcal{D}$ be a non-zero digit. Let $n \in \mathbb{N}$, let $\rho(n)$ denote the number of times that the digit η is used when n is expressed as τ-adic 2-NAF, and let $F(\eta, n)$ denote the appropriately truncated summation of ρ, in particular,*

$$F(\eta,n) = \tfrac{1}{2}\,\rho(0) + \rho(1) + \rho(2) + \cdots + \rho(n-1) + \tfrac{1}{2}\,\rho(n)\,.$$

Then, there exists piecewise continuous, piecewise nowhere differentiable function $P_+, P_- \colon \mathbb{R} \longrightarrow \mathbb{R}$, periodic with period 6 and a piecewise constant function $Q\colon \mathbb{R} \longrightarrow \mathbb{R}$, periodic with period 6, such that

$$F(\eta,n) = \tfrac{1}{9}n\log_3 n + n\,P_+(\log_3 n) + n\,P_-(\log_3 n) + 2n\,Q(\log_3 n) \quad \text{for } n \ge 1.$$

The Fourier coefficients of P_+, P_- can be calculated, Q is given explicit.

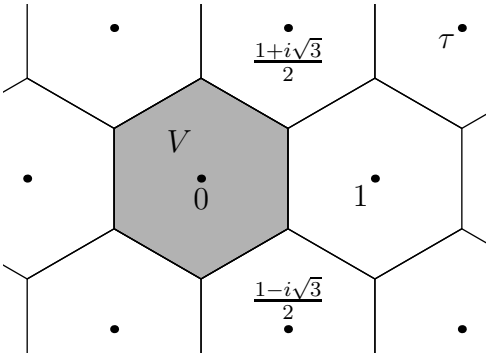


Figure 3.2.1: Voronoi cell V for 0 corresponding to the set $\mathbb{Z}[\tau]$ with $\tau = \frac{3}{2} + \frac{1}{2}\sqrt{-3}$.

Proof. Consider the τ-adic expansion of n. Theorem 3.1.4 on the facing page tells us that the digit η can occur either at an index in $b' + 6\mathbb{Z}$ coming from a 1 in the balanced ternary expansion of n or at an index in $b' + 3 + 6\mathbb{Z}$ coming from a -1. The latter one comes from the fact that $-\zeta^{b'} = \zeta^{b'+3}$, since ζ is a sixth root of unity.

We now use Theorem 3.1.2 on page 32 with parameters $q = 3$ and $r = 1$ for balanced ternary number system and $M = 6$, since according to Theorem 3.1.4 on the facing page every sixth digit of the balanced ternary expansion corresponds to the same power of ζ in its τ-adic expansion. For P_+ we use the P of Theorem 3.1.2 on page 32 with parameters $d = 1$ and the $b = b'$ from the previous paragraph. For P_- we use P with parameters $d = -1$ and $b = b' + 3$. The result follows by adding the two counting functions $F_{b'+6\mathbb{Z}}$ and $F_{b'+3+6\mathbb{Z}}$. □

3.2 Voronoi Cells

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

We will use the digit set of minimal norm representatives. In order to describe this digit set, we will rewrite the minimality condition in terms of the Voronoi cell for the lattice $\mathbb{Z}[\tau]$, cf. Gordon [22].

Definition 3.2.1 (Voronoi Cell)**.** We set

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

V is 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

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

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 3.2.1. Whenever the word "cells" is used in this paper, these Voronoi cells or scaled Voronoi cells will be meant.

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 a restricted version of V. This is very similar to the construction used in Avanzi, Heuberger and Prodinger [8].

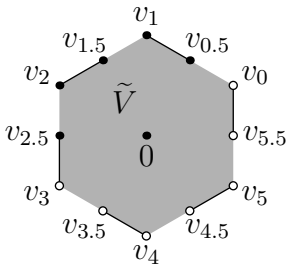


Figure 3.2.2: Restricted Voronoi cell $\widetilde{V}$ for 0 corresponding to the set $\mathbb{Z}[\tau]$ with $\tau = \frac{3}{2} + \frac{1}{2}\sqrt{-3}$.

Definition 3.2.2 (Restricted Voronoi Cell)**.** Let V_u be a Voronoi cell as above and u its centre. Let $v_0, \ldots, v_{m-1}$ 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 $k \in \left\{0, \ldots, \left\lfloor \frac{m}{2} \right\rfloor - 1\right\}$, and
- the points v_k for $k \in \left\{1, \ldots, \left\lfloor \frac{m}{3} \right\rfloor\right\}$.

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

In Figure 3.2.2 the restricted Voronoi cell for 0 is shown. The second condition is used, because it benefits symmetries. The third condition is just to make the midpoints unique. Obviously, other rules could have been used to define the restricted Voronoi cell.

As a generalisation of the usual fractional part of elements in $\mathbb{R}$ with respect to the integers, we define the fractional part of an element of $\mathbb{C}$ corresponding to the restricted Voronoi cell $\widetilde{V}$ and thus corresponding to the lattice $\mathbb{Z}[\tau]$.

Definition 3.2.3 (Fractional Part in $\mathbb{Z}[\tau]$)**.** Let $z \in \mathbb{C}$, $z = u + v$ with $u \in \mathbb{Z}[\tau]$ and $v \in \widetilde{V}$. Then we define the *fractional part corresponding to the lattice* $\mathbb{Z}[\tau]$ by $\{z\}_{\mathbb{Z}[\tau]} := v$.

This definition is valid, because of the construction of the restricted Voronoi cell. The fractional part of a point $z \in \mathbb{C}$ simply means, to search for the nearest lattice point u of $\mathbb{Z}[\tau]$ and returning the difference $z - u$.

Throughout this paper we will use the following notation for discs in the complex plane.

Definition 3.2.4 (Opened and Closed Discs)**.** Let $z \in \mathbb{C}$ and $r \geq 0$. The *open disc* $\mathcal{B}(z, r)$ *with centre* z *and radius* r is denoted by

$$\mathcal{B}(z, r) := \{y \in \mathbb{C} \,|\, |z - y| < r\}$$

and the *closed disc* $\overline{\mathcal{B}}(z, r)$ *with centre* z *and radius* r by

$$\overline{\mathcal{B}}(z, r) := \{y \in \mathbb{C} \,|\, |z - y| \leq r\}\,.$$

The disc $\mathcal{B}(0, 1)$ is called *unit disc*.

We will need suitable bounds for the digits in our digit set. These require precise knowledge on the Voronoi cells, such as the position of the vertices and bounds for the size of V. Such information is derived in the following proposition.

Proposition 3.2.5 (Properties of Voronoi Cells)**.** *We get the following properties:*

(a) The vertices of V are given by

$$\begin{aligned}
v_0 &= 1/2 + \frac{i}{2\operatorname{Im}(\tau)}\left(\operatorname{Im}(\tau)^2 + \{\operatorname{Re}(\tau)\}^2 - \{\operatorname{Re}(\tau)\}\right),\\
v_1 &= \{\operatorname{Re}(\tau)\} - \frac{1}{2} + \frac{i}{2\operatorname{Im}(\tau)}\left(\operatorname{Im}(\tau)^2 - \{\operatorname{Re}(\tau)\}^2 + \{\operatorname{Re}(\tau)\}\right),\\
v_2 &= -1/2 + \frac{i}{2\operatorname{Im}(\tau)}\left(\operatorname{Im}(\tau)^2 + \{\operatorname{Re}(\tau)\}^2 - \{\operatorname{Re}(\tau)\}\right) = v_0 - 1,\\
v_3 &= -v_0,\\
v_4 &= -v_1
\end{aligned}$$

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 the bounds

$$\overline{\mathcal{B}}\left(0, \tfrac{1}{2}\right) \subseteq V \subseteq \overline{\mathcal{B}}(0, |\tau|\, c_V)$$

with $c_V = \sqrt{\frac{7}{12}}$.

(d) The Lebesgue measure of V in the complex plane is

$$\lambda(V) = |\operatorname{Im}(\tau)|\,.$$

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

In the proof we will use some properties of Voronoi cells, which can, for example, be found in Aurenhammer [2].

Proof. (a) Since V is point-symmetric with respect to 0, we get $v_0 = -v_3$, $v_1 = -v_4$ and $v_2 = -v_5$. Thus we suppose without loss of generality $\operatorname{Im}(\tau) > 0$. Strict greater holds, because τ is imaginary quadratic. Even more, we get $\operatorname{Im}(\tau) \geq \frac{\sqrt{3}}{2}$, since

$$\tau = \frac{p}{2} \pm \frac{i}{2}\sqrt{4q - p^2}$$

is solution of $\tau^2 - p\tau + q = 0$ for $p, q \in \mathbb{Z}$ and either $4q - p^2 \equiv 0 \pmod 4$ or $4q - p^2 \equiv -1 \pmod 4$.

All elements of the lattice $\mathbb{Z}[\tau]$ can be written as $a + b\tau$, since τ is quadratic. We have to consider the neighbours of 0 in the lattice. The Voronoi cell is the area enclosed by the line segment bisectors of the lines from each neighbour to zero, see Figure 3.2.3.

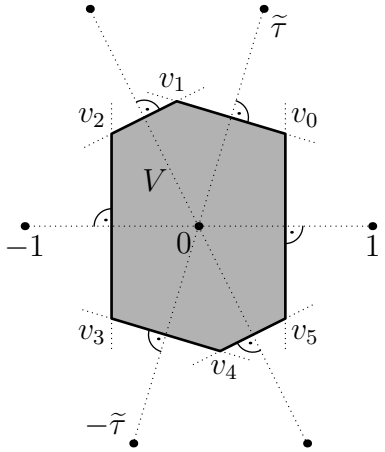


Figure 3.2.3: Construction of the Voronoi cell V for 0. The picture shows a general situation. Since τ is an imaginary quadratic algebraic integer, we will have $\mathrm{Re}(\widetilde{\tau}) \in \left\{0, \frac{1}{2}\right\}$.

Clearly $\mathrm{Re}(v_0) = \frac{1}{2}$ and $\mathrm{Re}(v_2) = -\frac{1}{2}$, since -1 and 1 are neighbours. Set $\widetilde{\tau} = \{\mathrm{Re}(\tau)\} + i\,\mathrm{Im}(\tau)$. Consider the line from 0 to $\widetilde{\tau}$ with midpoint $\frac{1}{2}\widetilde{\tau}$. We get

$$v_0 = \frac{1}{2}\widetilde{\tau} - x_A i\widetilde{\tau}$$

and

$$v_1 = \frac{1}{2}\widetilde{\tau} + x_B i\widetilde{\tau}$$

for some $x_A \in \mathbb{R}_{\geq 0}$ and $x_B \in \mathbb{R}_{\geq 0}$. Analogously, for the line from 0 to $\widetilde{\tau} - 1$, we have

$$v_1 = \frac{1}{2}\left(\widetilde{\tau} - 1\right) - x_C i\left(\widetilde{\tau} - 1\right)$$

and

$$v_2 = \frac{1}{2}\left(\widetilde{\tau} - 1\right) + x_D i\left(\widetilde{\tau} - 1\right).$$

for some $x_C \in \mathbb{R}_{\geq 0}$ and $x_D \in \mathbb{R}_{\geq 0}$. Solving this system of linear equations leads to the desired result. An easy calculation shows that $|v_0| = |v_1| = |v_2|$.

Until now, we have constructed the Voronoi cell of the points

$$P := \{0, 1, -1, \widetilde{\tau}, \widetilde{\tau} - 1, -\widetilde{\tau}, -(\widetilde{\tau} - 1)\}.$$

We want to rule out all other points, i.e., make sure, that none of the other points changes the already constructed cell. So let $z = x + iy \in \mathbb{Z}[\tau]$ and consider $\frac{z}{2}$. Because of symmetry reasons, we can assume $x \geq 0$ and $y \geq 0$. Clearly all points $z \in \mathbb{Z}$ with $z \geq 2$ do not change the Voronoi cell, since $\frac{z}{2} > \frac{1}{2}$ and the corresponding line segment bisector is vertical. So we can assume $y > 0$.

Now we will proceed in the following way. A point z can be ruled out, if the absolute value of $\frac{z}{2}$ is larger than

$$R = |v_0| = |v_1| = |v_2|.$$

Let $\tau = a + ib$. If $\{a\} = 0$, then $R^2 = \frac{1}{4}\left(1 + b^2\right)$. We claim that

$$R^2 < \frac{x^2 + y^2}{4} \Longleftrightarrow 1 + b^2 < x^2 + y^2.$$

Since $y > 0$, we have $y \geq b$. If $y = b$, then points with $x > 1$ need not be taken into account. But the remaining points are already in P (at least using symmetry and $\widetilde{\tau} + 1$ instead of $\widetilde{\tau} - 1$). If $y \geq 2b$, then all points except the ones with $x = 0$ can be ruled out, since $1 - b^2 \leq 1 - \frac{3}{4} = \frac{1}{4} < x$. But the points z with $x = 0$ can be ruled out, too, because there is already the point ib in P.

So let $\{a\} = \frac{1}{2}$. Then $R^2 = \frac{1}{4}\left(\frac{1}{2} + b^2 + \frac{1}{16b^2}\right)$ and we claim that

$$R^2 < \frac{x^2 + y^2}{4} \Longleftrightarrow \frac{1}{2} + b^2 + \frac{1}{16b^2} < x^2 + y^2.$$

If $y = b$, then $x > \sqrt{\frac{7}{12}}$ suffices to rule out a point z, since $b \geq \frac{\sqrt{3}}{2}$. But the only point z with $x \leq \sqrt{\frac{7}{12}}$ is $\frac{1}{2} + ib$, which is already in P. If $y \geq 2b$, then $\frac{1}{2} - b^2 + \frac{1}{16b^2} \leq \frac{1}{2} - \frac{3}{4} + \frac{1}{12} < 0$, so all points can be ruled out.

(b) Follows directly from the fact that all vertices have the same absolute value.

(c) From

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

we obtain

$$\frac{|v_0|}{|\widetilde{\tau}|} \leq \frac{|1 + i\operatorname{Im}(\widetilde{\tau})|}{2\,|\widetilde{\tau}|} \leq \frac{\operatorname{Im}(\widetilde{\tau})}{2\,|\widetilde{\tau}|} \sqrt{\frac{1}{\operatorname{Im}(\widetilde{\tau})^2} + 1} \leq \sqrt{\frac{7}{12}} =: c_V$$

since $\frac{\sqrt{3}}{2} \leq \operatorname{Im}(\widetilde{\tau}) \leq |\widetilde{\tau}|$. Therefore $V \subseteq \overline{\mathcal{B}}(0, |\tau|\, c_V)$.

Since $0 \leq \operatorname{Re}(\widetilde{\tau}) \leq 1$, we see that $\operatorname{Im}(v_1) \geq \operatorname{Im}(v_0) = \operatorname{Im}(v_2)$. By construction, the line from 0 to $\widetilde{\tau}$ intersects the line from v_0 to v_1 at $\frac{1}{2}\widetilde{\tau}$, so $\frac{1}{2}\,|\widetilde{\tau}|$ is an upper bound for the largest circle inside V. Analogously we get $\frac{1}{2}\,|\widetilde{\tau} - 1|$ as a bound, and from the line from 0 to 1 we get $\frac{1}{2}$. Since $\widetilde{\tau}$ and $\widetilde{\tau} - 1$ are lattice points and not zero, their norms are at least 1, so $\overline{\mathcal{B}}\left(0, \frac{1}{2}\right)$ is inside V.

(d) The area of V can be calculated easily, because $\operatorname{Im}(v_0) = \operatorname{Im}(v_2)$. Thus, splitting up the region in a rectangle and a triangle and using symmetry, the result follows.

(e) Let $x \in \tau^{-1}V$. Thus $x = \tau^{-1}z$ for an appropriate $z \in V$. For every $y \in \mathbb{Z}[\tau]$ we obtain

$$|x| = \left|\tau^{-1}\right| |z| \leq \left|\tau^{-1}\right| |z - y| = \left|x - \tau^{-1}y\right|.$$

For an arbitrary $u \in \mathbb{Z}[\tau]$ we can choose $y = \tau u$, and therefore $|x| \leq |x - u|$, i.e., $x \in V$. □

3.3 Digit Sets and Non-Adjacent Forms

In this section $\tau \in \mathbb{C}$ will be an algebraic integer with $|\tau| > 1$, and let $w \in \mathbb{N}$ with $w \geq 2$. Further let $\mathcal{N}\colon \mathbb{Z}[\tau] \longrightarrow \mathbb{Z}$ denote the norm function. We want to build a numeral system for the elements of $\mathbb{Z}[\tau]$ with base τ. Thus we need a digit set $\mathcal{D}$, which will be a finite subset of $\mathbb{Z}[\tau]$ containing 0.

Definition 3.3.1 (Reduced Residue Digit Set)**.** Let $\mathcal{D} \subseteq \mathbb{Z}[\tau]$. The set $\mathcal{D}$ is called a *reduced residue digit set modulo* τ^w, if it is consists of 0 and exactly one representative for each residue class of $\mathbb{Z}[\tau]$ modulo τ^w that is not divisible by τ.

From now on suppose $\mathcal{D}$ is a reduced residue digit set modulo τ^w. The following two auxiliary results are well-known; we include a proof for the sake of completeness.

Lemma 3.3.2. *Let c be a rational integer. Then τ divides c in $\mathbb{Z}[\tau]$ if and only if $\mathcal{N}(\tau)$ divides c in $\mathbb{Z}$.*

Proof. From the minimal polynomial, it is clear that τ divides $\mathcal{N}(\tau)$ in $\mathbb{Z}[\tau]$, so $\mathcal{N}(\tau) \mid c$ implies $\tau \mid c$.

For the converse direction, assume that $\tau \cdot \left(\sum_{j=0}^{d-1} x_j \tau^j\right) = c$ for some rational integers x_j. There d is the degree of τ. Write the minimal polynomial of τ as

$$\tau^d + \sum_{j=0}^{d-1} a_j \tau^j = 0.$$

Thus we obtain

$$c = -x_{d-1}a_0 + \sum_{j=1}^{d-1} (x_{j-1} - x_{d-1}a_j)\tau^j.$$

Comparing coefficients in τ^j yields $c = -x_{d-1}a_0$, which implies that $a_0 = (-1)^d \mathcal{N}(\tau)$ divides c in $\mathbb{Z}$, as required. □

Next, we determine the cardinality of $\mathcal{D}$ by giving an explicit system of representatives of the residue classes.

Lemma 3.3.3. *A complete residue system modulo τ^w is given by*

$$\sum_{j=0}^{w-1} a_j \tau^j \text{ with } a_j \in \{0, \ldots, \mathcal{N}(\tau) - 1\} \text{ for } 0 \leq j < w. \tag{3.3.1}$$

In particular, there are $\mathcal{N}(\tau)^w$ residue classes modulo τ^w in $\mathbb{Z}[\tau]$.

A representative $\sum_{j=0}^{w-1} a_j \tau^j$ with $a_j \in \{0, \ldots, \mathcal{N}(\tau) - 1\}$ is divisible by τ if and only if $a_0 = 0$. In particular, the cardinality of $\mathcal{D}$ equals $\mathcal{N}(\tau)^{w-1}(\mathcal{N}(\tau) - 1) + 1$.

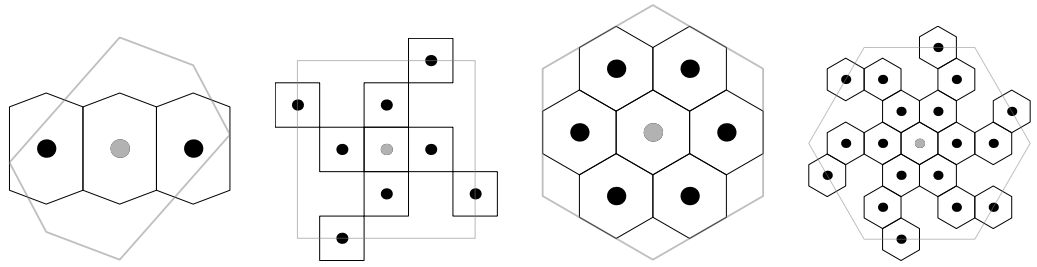
Proof. Every element z of $\mathbb{Z}[\tau]$ can be written as

$$z = x\tau^w + \sum_{j=0}^{w-1} a_j \tau^j$$

for some $a_j \in \{0, \ldots, \mathcal{N}(\tau) - 1\}$ and an appropriate $x \in \mathbb{Z}[\tau]$: Take the expansion of z with respect to the $\mathbb{Z}$-basis τ^j, $0 \leq j < d$ and subtract appropriate multiples of the minimal polynomial of τ in order to enforce $0 \leq a_j < \mathcal{N}(\tau)$ for $0 \leq j \leq w - 1$. This shows that (3.3.1) indeed covers all residue classes modulo τ^w.

Assume that $\sum_{j=0}^{w-1} a_j \tau^j \equiv \sum_{j=0}^{w-1} b_j \tau^j \pmod{\tau^w}$ for some $a_j, b_j \in \{0, \ldots, \mathcal{N}(\tau) - 1\}$, but $a_j \neq b_j$ for some j. We choose $0 \leq j_0 \leq w - 1$ minimal such that $a_{j_0} \neq b_{j_0}$. We obtain

$$\sum_{j=j_0}^{w-1} a_j \tau^{j-j_0} \equiv \sum_{j=j_0}^{w-1} b_j \tau^{j-j_0} \pmod{\tau^{w-j_0}},$$



(a) Digit set for $\tau = \frac{1}{2} + \frac{1}{2}\sqrt{-7}$ and $w = 2$.

(b) Digit set for $\tau = 1 + \sqrt{-1}$ and $w = 4$.

(c) Digit set for $\tau = \frac{3}{2} + \frac{1}{2}\sqrt{-3}$ and $w = 2$.

(d) Digit set for $\tau = \frac{3}{2} + \frac{1}{2}\sqrt{-3}$ and $w = 3$.

Figure 3.3.1: Minimal norm representatives digit sets modulo τ^w for different τ and w. For each digit η, the corresponding Voronoi cell V_η is drawn. The large scaled Voronoi cell is $\tau^w V$.

which implies that $a_{j_0} \equiv b_{j_0} \pmod{\tau}$. By Lemma 3.3.2 on the facing page, this implies that $a_{j_0} = b_{j_0}$, contradiction. Thus (3.3.1) is indeed a complete system of residues modulo τ^w.

From Lemma 3.3.2 on the preceding page we also see that exactly the $\mathcal{N}(\tau)^{w-1}$ residue classes $\sum_{j=1}^{w-1} a_j \tau^j$ are divisible by τ. We conclude that $\#\mathcal{D} = \mathcal{N}(\tau)^{w-1}(\mathcal{N}(\tau) - 1) + 1$. □

Since our digit set $\mathcal{D}$ is constructed of residue classes, we want a uniqueness in choosing the representative. We have the following definition, where the restricted Voronoi $\widetilde{V}$ for the point 0 from Definition 3.2.2 on page 40 is used.

Definition 3.3.4 (Representatives of Minimal Norm)**.** Let τ be an algebraic integer, imaginary quadratic, and let $\eta \in \mathbb{Z}[\tau]$ be not divisible by τ. Then η is called a *representative of minimal norm of its residue class*, if $\eta \in \tau^w \widetilde{V}$.

With this definition we can define the following digit set, cf. Solinas [63, 64] or Blake, Kumar Murty and Xu [15].

Definition 3.3.5 (Minimal Norm Representatives Digit Set)**.** Let τ be an algebraic integer, imaginary quadratic, and let $\mathcal{D}$ be a reduced residue digit set modulo τ^w consisting of representatives of minimum norm of its residue classes. Then we will call such a digit set *minimal norm representatives digit set modulo* τ^w.

From now on we will suppose that our digit set $\mathcal{D}$ is a minimal norm representatives digit set modulo τ^w. Some examples are shown in Figure 3.3.1.

The following remark summarises some basic properties of minimal norm representatives and the defined digit sets.

Remark 3.3.6. Let τ be an algebraic integer, imaginary quadratic. We have the following equivalence. The condition

$$|\eta| \leq |\xi| \text{ for all } \xi \in \mathbb{Z}[\tau] \text{ with } \eta \equiv \xi \pmod{\tau^w}$$

is fulfilled, if and only if $\eta \in \tau^w V$. The advantage of using the restricted Voronoi cell in Definition 3.3.4 is that also points on the boundary are handled uniquely.

Further we get for all $\eta \in \mathcal{D}$ that $|\eta| \leq |\tau|^{w+1} c_V$. On the other side, if an element of $\mathbb{Z}[\tau]$, which is not divisible by τ, has norm less than $\frac{1}{2}\tau^w$, cf. Proposition 3.2.5 on page 41, it is a digit. See also Lemma 3.3.3 on the facing page.

Since $\mathcal{D} \subseteq \mathbb{Z}[\tau]$, all non-zero elements have norm at least 1.

We can assume that $0 \leq \arg(\tau) \leq \frac{\pi}{2}$. Using any other τ lead to the same digit sets, except some mirroring at the real axis, imaginary axis, or at the origin. By adapting the definition of the boundary of the restricted Voronoi cell, Definition 3.2.2 on page 40, these mirroring effects can be handled.

Now we are ready to define the numbers built with our digit set $\mathcal{D}$.

Definition 3.3.7 (Width-w τ-adic Non-Adjacent Forms)**.** Let $\boldsymbol{\eta} = (\eta_j)_{j\in\mathbb{Z}} \in \mathcal{D}^{\mathbb{Z}}$. The sequence $\boldsymbol{\eta}$ is called a *width-w τ-adic non-adjacent form*, or *w-NAF* for short, if each factor $\eta_{j+w-1}\dots\eta_j$, i.e., each block of length w, contains at most one non-zero digit.

Let $J = \{j \in \mathbb{Z} \,|\, \eta_j \neq 0\}$. We call $\sup(\{0\} \cup (J+1))$ the *left-length of the w-NAF $\boldsymbol{\eta}$* and $-\inf(\{0\} \cup J)$ the *right-length of the w-NAF $\boldsymbol{\eta}$*.

Let λ and ρ be elements of $\mathbb{N}_0 \cup \{\mathsf{fin}, \infty\}$, where fin means finite. We denote the *set of all w-NAFs of left-length at most λ and right-length at most ρ* by $\mathbf{NAF}_w^{\lambda,\rho}$. If $\rho = 0$, then we will simply write $\mathbf{NAF}_w^{\lambda}$. The elements of the set $\mathbf{NAF}_w^{\mathsf{fin}}$ will be called *integer w-NAFs*.

For $\boldsymbol{\eta} \in \mathbf{NAF}_w^{\mathsf{fin},\infty}$ we call

$$\mathsf{value}(\boldsymbol{\eta}) := \sum_{j\in\mathbb{Z}} \eta_j \tau^j$$

the *value of the w-NAF $\boldsymbol{\eta}$*.

The following notations and conventions are used. A block of zero digits is denoted by $\mathbf{0}$. For a digit η and $k \in \mathbb{N}_0$ we will use

$$\eta^k := \underbrace{\eta\dots\eta}_{k},$$

with the convention $\eta^0 := \varepsilon$, where ε denotes the empty word. A w-NAF $\boldsymbol{\eta} = (\eta_j)_{j\in\mathbb{Z}}$ will be written as $\boldsymbol{\eta_I}.\boldsymbol{\eta_F}$, where $\boldsymbol{\eta_I}$ contains the η_j with $j \geq 0$ and $\boldsymbol{\eta_F}$ contains the η_j with $j < 0$. $\boldsymbol{\eta_I}$ is called *integer part*, $\boldsymbol{\eta_F}$ *fractional part*, and the dot is called *τ-point*. Left-leading zeros in $\boldsymbol{\eta_I}$ can be skipped, except η_0, and right-leading zeros in $\boldsymbol{\eta_F}$ can be skipped as well. If $\boldsymbol{\eta_F}$ is a sequence containing only zeros, the τ-point and this sequence is not drawn.

Further, for a w-NAF $\boldsymbol{\eta}$ (a bold, usually small Greek letter) we will always use η_j (the same letter, but indexed and not bold) for the elements of the sequence.

To see where the values, respectively the fractional values of our w-NAFs lie in the complex plane, have a look at Figure 3.9.1 on page 73. There some examples are drawn.

The set $\mathbf{NAF}_w^{\mathsf{fin},\infty}$ can be equipped with a metric. It is defined in the following way. Let $\boldsymbol{\eta} \in \mathbf{NAF}_w^{\mathsf{fin},\infty}$ and $\boldsymbol{\xi} \in \mathbf{NAF}_w^{\mathsf{fin},\infty}$, then

$$\mathrm{d}_{\mathsf{NAF}}(\boldsymbol{\eta},\boldsymbol{\xi}) := \begin{cases} |\tau|^{\max\{j\in\mathbb{Z}\,|\,\eta_j\neq\xi_j\}} & \text{if } \boldsymbol{\eta} \neq \boldsymbol{\xi}, \\ 0 & \text{if } \boldsymbol{\eta} = \boldsymbol{\xi}. \end{cases}$$

So the largest index, where the two w-NAFs differ, decides their distance. See for example Edgar [19] for details on such metrics.

We get the following continuity result.

Proposition 3.3.8. *The value function* value *is Lipschitz continuous on* $\mathbf{NAF}_w^{\mathsf{fin},\infty}$.

Proof. Let $c_{\mathcal{D}}$ be a bound for the absolute value of the digits in the digit set $\mathcal{D}$. Let $\boldsymbol{\eta} \in \mathbf{NAF}_w^{\mathsf{fin},\infty}$ and $\boldsymbol{\xi} \in \mathbf{NAF}_w^{\mathsf{fin},\infty}$, $\boldsymbol{\eta} \neq \boldsymbol{\xi}$, with $\mathrm{d}_{\mathsf{NAF}}(\boldsymbol{\eta},\boldsymbol{\xi}) = |\tau|^J$. Since $\boldsymbol{\eta}$ and $\boldsymbol{\xi}$ are equal on all digits with index larger than J we obtain

$$|\mathsf{value}(\boldsymbol{\eta}) - \mathsf{value}(\boldsymbol{\xi})| \leq \sum_{j\leq J} |\eta_j - \xi_j|\,|\tau|^j \leq 2c_{\mathcal{D}} \frac{|\tau|^J}{1-|\tau|^{-1}} = \frac{2c_{\mathcal{D}}}{1-|\tau|^{-1}}\,\mathrm{d}_{\mathsf{NAF}}(\boldsymbol{\eta},\boldsymbol{\xi})\,.$$

Thus Lipschitz continuity is proved. □

Furthermore, we get a compactness result on the metric space $\mathbf{NAF}_w^{\ell.\infty} \subseteq \mathbf{NAF}_w^{\mathsf{fin}.\infty}$ in the proposition below. The metric space $\mathbf{NAF}_w^{\mathsf{fin}.\infty}$ is not compact, because if we fix a non-zero digit η, then the sequence $\left(\eta 0^j\right)_{j\in\mathbb{N}_0}$ has no convergent subsequence, but all $\eta 0^j$ are in the set $\mathbf{NAF}_w^{\mathsf{fin}.\infty}$.

Proposition 3.3.9. *For every $\ell \geq 0$ the metric space $\left(\mathbf{NAF}_w^{\ell.\infty}, \mathrm{d}_{\mathsf{NAF}}\right)$ is compact.*

Proof. Let $(\boldsymbol{\xi}_{0,j})_{j\in\mathbb{N}_0}$ be a sequence with $\boldsymbol{\xi}_{0,j} \in \mathbf{NAF}_w^{\ell.\infty}$. We can assume $\boldsymbol{\xi}_{0,j} \in \mathbf{NAF}_w^{0.\infty}$, therefore each word $\boldsymbol{\xi}_{0,j}$ has digits zero for non-negative index. Now consider the digit with index -1. There is a subsequence $(\boldsymbol{\xi}_{1,j})_{j\in\mathbb{N}_0}$ of $(\boldsymbol{\xi}_{0,j})_{j\in\mathbb{N}_0}$, such that digit -1 is a fixed digit η_{-1}. Next there is a subsequence $(\boldsymbol{\xi}_{2,j})_{j\in\mathbb{N}_0}$ of $(\boldsymbol{\xi}_{1,j})_{j\in\mathbb{N}_0}$, such that digit -2 is a fixed digit η_{-2}. This process can be repeated for each $k \geq 1$ to get sequences $(\boldsymbol{\xi}_{k,j})_{j\in\mathbb{N}_0}$ and digits η_{-k}.

The sequence $(\boldsymbol{\vartheta}_j)_{j\in\mathbb{N}_0}$ with $\boldsymbol{\vartheta}_j := \boldsymbol{\xi}_{j,j}$ converges to $\boldsymbol{\eta}$, since for $\varepsilon > 0$ there is an $J \in \mathbb{N}_0$ such that for all $j \geq J$

$$\mathrm{d}_{\mathsf{NAF}}(\boldsymbol{\eta}, \boldsymbol{\vartheta}_j) \leq |\tau|^{-(J+1)} < \varepsilon.$$

It is clear that $\boldsymbol{\eta}$ is indeed an element of $\mathbf{NAF}_w^{0.\infty}$, as its first k digits coincide with $\boldsymbol{\xi}_{k,k} \in \mathbf{NAF}_w^{0.\infty}$ for all k. So we have found a converging subsequence of $(\boldsymbol{\xi}_{0,j})_{j\in\mathbb{N}_0}$, which proves the compactness. □

Remark 3.3.10. The compactness of $\left(\mathbf{NAF}_w^{\ell.\infty}, \mathrm{d}_{\mathsf{NAF}}\right)$ can also be deduced from general theory. As a consequence of Tychonoff's Theorem the set $\mathcal{D}^{\mathbb{N}}$ is a compact space, the product topology (of the discrete topology on $\mathcal{D}$) coincides with the topology induced by the obvious generalisation of the metric $\mathrm{d}_{\mathsf{NAF}}$. The subset $\mathbf{NAF}_w^{0.\infty} \subseteq \mathcal{D}^{\mathbb{N}}$ is closed and therefore compact, too.

We want to express all integers in $\mathbb{Z}[\tau]$ by finite w-NAFs. Thus we restrict ourselves to suitable digit sets, cf. Muir and Stinson [50].

Definition 3.3.11 (Width-w Non-Adjacent Digit Set)**.** A digit set $\mathcal{D}$ is called a *width-w non-adjacent digit set*, or w-NADS for short, when every element $z \in \mathbb{Z}[\tau]$ admits a unique w-NAF $\boldsymbol{\eta} \in \mathbf{NAF}_w^{\mathsf{fin}}$, i.e., $\mathsf{value}(\boldsymbol{\eta}) = z$. When this is the case, the function

$$\mathsf{value}|_{\mathbf{NAF}_w^{\mathsf{fin}}} \colon \mathbf{NAF}_w^{\mathsf{fin}} \longrightarrow \mathbb{Z}[\tau]$$

is bijective, and we will denote its inverse function by NAF_w.

Later, namely in Section 3.6, we will see that the digit set of minimal norm representatives is a w-NADS if τ is imaginary quadratic.

3.4 Full Block Length Analysis of Non-Adjacent Forms

Let $\tau \in \mathbb{C}$ be an algebraic integer, $w \in \mathbb{N}$ with $w \geq 2$, and $\mathcal{D}$ be a reduced residue digit set, cf. Definition 3.3.1 on page 44. Let $\mathcal{N} \colon \mathbb{Z}[\tau] \longrightarrow \mathbb{Z}$ denote the norm function.

Further, in this section all w-NAFs will be out of the set $\mathbf{NAF}_w^{\mathsf{fin}}$, and with *length* the left-length is meant.

This general setting allows us to analyse digit frequencies under the *full block length modell,*, i.e., we assume that all w-NAFs of given length are equally likely. We will prove the following theorem.

Theorem 3.4.1 (Full Block Length Distribution Theorem)**.** *We denote the number of w-NAFs of length $n \in \mathbb{N}_0$ by $C_{n,w}$, i.e., $C_{n,w} = \#(\mathbf{NAF}_w^n)$, and we get*

$$C_{n,w} = \frac{1}{(\mathcal{N}(\tau)-1)w+1}\,\mathcal{N}(\tau)^{n+w} + \mathcal{O}((\rho\mathcal{N}(\tau))^n)\,,$$

where $\rho = (1+\frac{1}{\mathcal{N}(\tau)w^3})^{-1} < 1$.

Further let $0 \neq \eta \in \mathcal{D}$ be a fixed digit and define the random variable $X_{n,w,\eta}$ to be the number of occurrences of the digit η in a random w-NAF of length n, where every w-NAF of length n is assumed to be equally likely.

Then the following explicit expressions hold for the expectation and the variance of $X_{n,w,\eta}$:

$$\mathbb{E}(X_{n,w,\eta}) = e_w n + \frac{(\mathcal{N}(\tau)-1)(w-1)w}{\mathcal{N}(\tau)^{w-1}\left((\mathcal{N}(\tau)-1)w+1\right)^2} + \mathcal{O}(n\rho^n) \tag{3.4.1}$$

$$\begin{aligned}\mathbb{V}(X_{n,w,\eta}) &= v_w n \\ &\quad + \frac{(w-1)w\big(-(w-1)^2-\mathcal{N}(\tau)^2w^2+(\mathcal{N}(\tau)-1)\mathcal{N}(\tau)^{w-1}((\mathcal{N}(\tau)-1)w+1)^2+2\mathcal{N}(\tau)\big(w^2-w+1\big)\big)}{\mathcal{N}(\tau)^{2w-2}\left((\mathcal{N}(\tau)-1)w+1\right)^4} \\ &\quad + \mathcal{O}\big(n^2\rho^n\big)\,,\end{aligned} \tag{3.4.2}$$

where

$$e_w = \frac{1}{\mathcal{N}(\tau)^{w-1}\left((\mathcal{N}(\tau)-1)w+1\right)},$$

and

$$v_w = \frac{\mathcal{N}(\tau)^{w-1}\left((\mathcal{N}(\tau)-1)w+1\right)^2 - \big((\mathcal{N}(\tau)-1)w^2+2w-1\big)}{\mathcal{N}(\tau)^{2w-2}\left((\mathcal{N}(\tau)-1)w+1\right)^3}.$$

Furthermore, $X_{n,w,\eta}$ satisfies the central limit theorem

$$\mathbb{P}\left(\frac{X_{n,w,\eta}-e_w n}{\sqrt{v_w n}} \leq x\right) = \Phi(x) + \mathcal{O}\left(\frac{1}{\sqrt{n}}\right),$$

uniformly with respect to $x \in \mathbb{R}$, where $\Phi(x) = (2\pi)^{-1/2}\int_{t\leq x} e^{-t^2/2}\,dt$ is the standard normal distribution.

For the proof we need estimates for the zeros of a polynomial which will be needed for estimating the non-dominant roots of our generating function.

Lemma 3.4.2. *Let $t \geq 2$ and*

$$f(z) = 1 - \frac{1}{t}z - \left(1-\frac{1}{t}\right)z^w.$$

Then $f(z)$ has exactly one root with $|z| \leq 1+\frac{1}{tw^3}$, namely $z = 1$.

Proof. It is easily checked that $f(1) = 0$. Assume that $z \neq 1$ is another root of f. As the coefficients of f are reals, it is sufficient to consider z with $\mathrm{Im}(z) \geq 0$. If $|z| < 1$, then

$$1 = \left|\frac{1}{t}z + \left(1-\frac{1}{t}\right)z^w\right| \leq \frac{1}{t}|z| + \left(1-\frac{1}{t}\right)|z|^w < \frac{1}{t} + \left(1-\frac{1}{t}\right) = 1,$$

which is a contradiction. Therefore, we have $|z| \geq 1$. We write $z = re^{i\psi}$ for appropriate $r \geq 1$ and $0 \leq \psi \leq \pi$. For $r > 0$, $f(r)$ is strictly decreasing, so we can assume that $\psi > 0$.

For $\psi < \pi/w$, we have $\sin(w\psi) > 0$ and $\sin(\psi) > 0$, which implies that $\operatorname{Im}\big(f(re^{i\psi})\big) = -\frac{1}{t}\sin\psi - \big(1-\frac{1}{t}\big)\sin(\psi w) < 0$, a contradiction. We conclude that $\psi \geq \pi/w$.

Next, we see that $f(re^{i\psi}) = 0$ implies that

$$1 - \frac{2r}{t}\cos\psi + \frac{r^2}{t^2} = \left|1 - \frac{1}{t}re^{i\psi}\right|^2 = \left(1-\frac{1}{t}\right)^2 r^{2w}.$$

We have $\cos\psi \leq \cos(\pi/w)$, which implies that

$$\left(1-\frac{1}{t}\right)^2 r^{2w} \geq 1 - \frac{2r}{t}\cos\frac{\pi}{w} + \frac{r^2}{t^2} = \left(1-\frac{r}{t}\right)^2 + \left(1-\cos\frac{\pi}{w}\right)\frac{2r}{t}. \tag{3.4.3}$$

For $w \geq 4$ and $r < \sqrt{2}$, the right hand side of (3.4.3) is decreasing and the left hand side is increasing. Thus, for $r \leq 1 + 1/(tw^3)$, (3.4.3) yields

$$\left(1-\frac{1}{t}\right)^2\left(1+\frac{1}{tw^3}\right)^{2w} \geq \left(1-\frac{1}{t}\cdot\left(1+\frac{1}{tw^3}\right)\right)^2 + \left(1-\cos\frac{\pi}{w}\right)\frac{2}{t}\left(1+\frac{1}{tw^3}\right).$$

Using the estimates $(1+\frac{1}{tw^3})^{2w} \leq 1 + \frac{2}{tw^2} + \frac{2}{t^2w^4}$ and $\cos\left(\frac{\pi}{w}\right) \leq 1 - \frac{\pi^2}{2w^2} + \frac{\pi^4}{24w^4}$, we obtain

$$\left(\frac{2-\pi^2}{t} - \frac{4}{t^2} + \frac{2}{t^3}\right)\frac{1}{w^2} + \left(\frac{2}{t^2} - \frac{2}{t^3}\right)\frac{1}{w^3} + \left(\frac{\pi^4}{12t} + \frac{2}{t^2} - \frac{4}{t^3} + \frac{2}{t^4}\right)\frac{1}{w^4} - \frac{\pi^2}{t^2w^5} - \frac{1}{t^4w^6} + \frac{\pi^4}{12t^2w^7} \geq 0,$$

which is a contradiction for $w \geq 4$ and $t \geq 2$.

For $w = 3$, we easily check that $|z| = \sqrt{\frac{t}{t-1}} \geq 1 + 1/(27t)$; similarly, for $w = 2$, we have $|z| = t/(t-1) > 1 + 1/(4t)$. □

Proof of Theorem 3.4.1. For simplicity we set $\mathcal{D}^\bullet := \mathcal{D} \setminus \{0\}$. A w-NAF can be described by the regular expression

$$\left(\varepsilon + \sum_{d\in\mathcal{D}^\bullet}\sum_{k=0}^{w-2} 0^k d\right)\left(0 + \sum_{d\in\mathcal{D}^\bullet} 0^{w-1}d\right)^*$$

Let a_{mn} be the number of w-NAFs of length n containing exactly m occurrences of the digit η. We consider the generating function $G(Y,Z) = \sum_{m,n} a_{mn}Y^mZ^n$. From the regular expression we see that

$$G(Y,Z) = \frac{1 + (Y + (\#\mathcal{D}^\bullet - 1))\frac{Z^w - Z}{Z-1}}{1 - Z - YZ^w - (\#\mathcal{D}^\bullet - 1)Z^w}.$$

We start with determining the number of w-NAFs of length n. This amounts to extracting the coefficient of Z^n of

$$G(1,Z) = \frac{1 + (\#\mathcal{D}^\bullet - 1)Z - \#\mathcal{D}^\bullet \cdot Z^w}{(1-Z)(1 - Z - \#\mathcal{D}^\bullet \cdot Z^w)}.$$

This requires finding the dominant root of the denominator. Setting $z = \mathcal{N}(\tau)\,Z$ in the second factor yields

$$1 - \frac{1}{\mathcal{N}(\tau)}z - \left(1 - \frac{1}{\mathcal{N}(\tau)}\right)z^w.$$

From Lemma 3.4.2 on page 48, we see that the dominant root of the denominator of $G(1,Z)$ is $Z = 1/\mathcal{N}(\tau)$, and that all other roots satisfy $|Z| \geq 1/\mathcal{N}(\tau) + 1/(\mathcal{N}(\tau)^2 w^3)$. Extracting the coefficient of Z^n of $G(1,Z)$ then yields the number $C_{n,w}$ of w-NAFs of length n as

$$\frac{1}{(\mathcal{N}(\tau)-1)w+1}\,\mathcal{N}(\tau)^{n+w} + \mathcal{O}((\rho\mathcal{N}(\tau))^n)\,, \tag{3.4.4}$$

where $\rho = (1 + \frac{1}{\mathcal{N}(\tau)w^3})^{-1}$.

The number of occurrences of the digit μ amongst all w-NAFs of length n is

$$\begin{aligned}[Z^n]\,\frac{\partial G(Y,Z)}{\partial Y}\bigg|_{Y=1} &= \frac{Z}{\left(1 - Z - \left(1 - \frac{1}{\mathcal{N}(\tau)}\right)(\mathcal{N}(\tau)\,Z)^w\right)^2} \\ &= \frac{1}{((\mathcal{N}(\tau)-1)w+1)^2} n\,\mathcal{N}(\tau)^{n+1} + \frac{(\mathcal{N}(\tau)-1)(w-1)w}{((\mathcal{N}(\tau)-1)w+1)^3}\,\mathcal{N}(\tau)^{n+1} + \mathcal{O}((\rho\mathcal{N}(\tau))^n)\,.\end{aligned}$$

Dividing this by (3.4.4) yields (3.4.1).

In order to compute the second moment, we compute

$$\begin{aligned}[Z^n]\,\frac{\partial^2 G(Y,Z)}{\partial Y^2}\bigg|_{Y=1} &= \frac{2Z^{w+1}}{\left(1 - Z - \left(1 - \frac{1}{\mathcal{N}(\tau)}\right)(\mathcal{N}(\tau)\,Z)^w\right)^3} \\ = \frac{1}{((\mathcal{N}(\tau)-1)w+1)^3} n^2\,\mathcal{N}(\tau)^{n-w+2} &+ \frac{\left((\mathcal{N}(\tau)-1)w^2 - 2w\,\mathcal{N}(\tau) + 1\right)}{((\mathcal{N}(\tau)-1)w+1)^4} n\,\mathcal{N}(\tau)^{n-w+2} \\ &- \frac{(w-1)w\left(w\,\mathcal{N}(\tau)^2 - 2\,\mathcal{N}(\tau) - w + 1\right)}{((\mathcal{N}(\tau)-1)w+1)^5}\,\mathcal{N}(\tau)^{n-w+2} + \mathcal{O}((\rho\mathcal{N}(\tau))^n)\,,\end{aligned}$$

which after division by (3.4.4) yields

$$\begin{aligned}\mathbb{E}(X_{n,w,\eta}(X_{n,w,\eta}-1)) = \frac{1}{\mathcal{N}(\tau)^{2w-2}\,((\mathcal{N}(\tau)-1)w+1)^2} n^2 &+ \frac{\left((\mathcal{N}(\tau)-1)w^2 - 2w\,\mathcal{N}(\tau) + 1\right)}{\mathcal{N}(\tau)^{2w-2}\,((\mathcal{N}(\tau)-1)w+1)^3} n \\ &- \frac{(w-1)w\left(w\,\mathcal{N}(\tau)^2 - 2\,\mathcal{N}(\tau) - w + 1\right)}{\mathcal{N}(\tau)^{n-w+2}\,((\mathcal{N}(\tau)-1)w+1)^4} + \mathcal{O}\big(n^2\rho^n\big)\,.\end{aligned}$$

Adding $\mathbb{E}(X_{n,w,\eta}) - \mathbb{E}(X_{n,w,\eta})^2$ yields the variance given in (3.4.2).

The asymptotic normality follows from Hwang's Quasi-Power-Theorem [32]. □

3.5 Bounds for the Value of Non-Adjacent Forms

Let $\tau \in \mathbb{C}$ be an algebraic integer, imaginary quadratic with minimal polynomial $X^2 - pX + q$ with $p, q \in \mathbb{Z}$ such that $4q - p^2 > 0$. Suppose that $|\tau| > 1$. Let $w \in \mathbb{N}$ with $w \geq 2$. Further let $\mathcal{D}$ be a minimal norm representatives digit set modulo τ^w as in Definition 3.3.5 on page 45.

In this section the *fractional value* of a w-NAF means the value of a w-NAF of the form $0.\boldsymbol{\eta}$. The term *most significant digit* is used for the digit η_{-1}.

So let us have a closer look at the fractional value of a w-NAF. We want to find upper bounds and if we fix a digit, e.g. the most significant one, a lower bound. We need two different approaches to prove those results. The first one is analytic. The results there are valid for all combinations

of τ and w except finitely many. These exceptional cases will be called "problematic values". To handle those, we will use an other idea. We will show an equivalence, which directly leads to a simple procedure to check, whether a condition is fulfilled. If this is the case, the procedure terminates and returns the result. This idea is similar to a proof in Matula [42].

The following proposition deals with three upper bounds, one for the absolute value and two give us regions containing the fractional value.

Proposition 3.5.1 (Upper Bounds for the Fractional Value)**.** *Let $\boldsymbol{\eta} \in \mathbf{NAF}_w^{0.\infty}$, and let*

$$f_U = \frac{|\tau|^w c_V}{1 - |\tau|^{-w}}.$$

Then the following statements are true:

(a) We get

$$|\mathsf{value}(\boldsymbol{\eta})| \leq f_U.$$

(b) Further we have

$$\mathsf{value}(\boldsymbol{\eta}) \in \bigcup_{z \in \tau^{w-1}V} \overline{\mathcal{B}}\Big(z, |\tau|^{-w} f_U\Big).$$

(c) The following two statements are equivalent:

(1) There is an $\ell \in \mathbb{N}_0$, such that for all $\boldsymbol{\xi} \in \mathbf{NAF}_w^{0.\ell}$ the condition

$$\overline{\mathcal{B}}\Big(\mathsf{value}(\boldsymbol{\xi}), |\tau|^{-\ell} f_U\Big) \subseteq \tau^{2w-1} \operatorname{int}(V)$$

is fulfilled.

(2) There exists an $\varepsilon > 0$, such that for all $\boldsymbol{\vartheta} \in \mathbf{NAF}_w^{0.\infty}$ the condition

$$\mathcal{B}(\mathsf{value}(\boldsymbol{\vartheta}), \varepsilon) \subseteq \tau^{2w-1} V$$

holds.

(d) We get

$$\mathsf{value}(\boldsymbol{\eta}) \in \tau^{2w-1} V.$$

(e) For $\ell \in \mathbb{N}_0$ we have

$$\mathsf{value}(0.\eta_{-1} \dots \eta_{-\ell}) + \tau^{-\ell} V \subseteq \tau^{2w-1} V.$$

Proof. (a) We have

$$|\mathsf{value}(\boldsymbol{\eta})| = \left| \sum_{j=1}^{\infty} \eta_{-j} \tau^{-j} \right| \leq \sum_{j=1}^{\infty} |\eta_{-j}| \, |\tau|^{-j}.$$

We consider w-NAFs, which have $\eta_{-j} \neq 0$ for $-j \equiv 1 \pmod{w}$. For all other w-NAFs the upper bound is smaller. To see this, assume that there are more than $w-1$ adjacent zeros in a w-NAF or the first digits are zero. Then we could build a larger upper bound by shifting digits to the left, i.e., multiplying parts of the sum by $|\tau|$, since $|\tau| > 1$.

We get

$$\begin{aligned} |\mathsf{value}(\boldsymbol{\eta})| &\leq \sum_{j=1}^{\infty} |\eta_{-j}| \, |\tau|^{-j} = \sum_{j=1}^{\infty} [-j \equiv 1 \pmod{w}] \, |\eta_{-j}| \, |\tau|^{-j} \\ &\leq |\tau|^{-1} \sum_{k=0}^{\infty} \left|\eta_{-(wk+1)}\right| |\tau|^{-wk}, \end{aligned}$$

in which we changed the summation index according to $wk + 1 = j$ and the Iversonian notation $[expr] = 1$ if $expr$ is true and $[expr] = 0$ otherwise, cf. Graham, Knuth and Patashnik [25], has been used. Using $\left|\eta_{-(wk+1)}\right| \leq |\tau|^{w+1} c_V$, see Remark 3.3.6 on page 45, yields

$$|\mathsf{value}(\boldsymbol{\eta})| \leq |\tau|^{-1} |\tau|^{w+1} c_V \frac{1}{1 - |\tau|^{-w}} = \underbrace{\frac{|\tau|^w c_V}{1 - |\tau|^{-w}}}_{=:f_U}.$$

(b) There is nothing to show if the w-NAF $\boldsymbol{\eta}$ is zero, and it is sufficient to prove it for $\eta_{-1} \neq 0$. Otherwise, let $k \in \mathbb{N}$ be minimal, such that $\eta_{-k} \neq 0$. Then

$$\tau^{-(k-1)}\tau^{k-1}\,\mathsf{value}(\boldsymbol{\eta}) \in \tau^{-(k-1)} \bigcup_{z \in \tau^{w-1}V} \overline{\mathcal{B}}\Big(z, |\tau|^{-w} f_U\Big) \subseteq \bigcup_{z \in \tau^{w-1}V} \overline{\mathcal{B}}\Big(z, |\tau|^{-w} f_U\Big),$$

since $|\tau| > 1$ and $\tau^{-1}V \subseteq V$, see Proposition 3.2.5 on page 41.

Since $\eta_{-1} \in \tau^w V$, see Remark 3.3.6 on page 45, we obtain $\eta_{-1}\tau^{-1} \in \tau^{w-1}V$. Thus, using (a), yields

$$\left|\tau^w \left(\mathsf{value}(\boldsymbol{\eta}) - \eta_{-1}\tau^{-1}\right)\right| \leq f_U,$$

i.e.,

$$\mathsf{value}(\boldsymbol{\eta}) \in \overline{\mathcal{B}}\Big(\eta_{-1}\tau^{-1}, |\tau|^{-w} f_U\Big),$$

which proves the statement.

(c) *(1) $\Longrightarrow$ (2).* Suppose there exists such an $\ell \in \mathbb{N}$. Then there exists an $\varepsilon > 0$ such that

$$\overline{\mathcal{B}}\Big(\mathsf{value}(\boldsymbol{\xi}), |\tau|^{-\ell} f_U + \varepsilon\Big) \subseteq \tau^{2w-1}V$$

for all $\boldsymbol{\xi} \in \mathbf{NAF}_w^{0.\ell}$, since there are only finitely many $\boldsymbol{\xi} \in \mathbf{NAF}_w^{0.\ell}$. Let $\boldsymbol{\vartheta} \in \mathbf{NAF}_w^{0.\infty}$. Then there is a $\boldsymbol{\xi} \in \mathbf{NAF}_w^{0.\ell}$ such that the digits from index -1 to $-\ell$ of $\boldsymbol{\xi}$ and $\boldsymbol{\vartheta}$ coincide. By using (a) we obtain

$$|\mathsf{value}(\boldsymbol{\vartheta}) - \mathsf{value}(\boldsymbol{\xi})| \leq |\tau|^{-\ell} f_U,$$

and thus

$$\overline{\mathcal{B}}(\mathsf{value}(\boldsymbol{\vartheta}), \varepsilon) \subseteq \overline{\mathcal{B}}\Big(\mathsf{value}(\boldsymbol{\xi}), |\tau|^{-\ell} f_U + \varepsilon\Big) \subseteq \tau^{2w-1}V.$$

(2) $\Longrightarrow$ (1). Now suppose there is such an $\varepsilon > 0$. Since there is an $\ell \in \mathbb{N}$ such that $|\tau|^{-\ell} f_U < \varepsilon$, the statement follows.

(d) We know from Proposition 3.2.5 on page 41 that $\overline{\mathcal{B}}\Big(0, \frac{1}{2}|\tau|^{2w-1}\Big) \subseteq \tau^{2w-1}V$. Therefore, if the upper bound found in (a) fulfils

$$f_U = \frac{|\tau|^w c_V}{1 - |\tau|^{-w}} \leq \frac{1}{2}|\tau|^{2w-1},$$

the statement follows.

The previous inequality is equivalent to

$$\nu := \frac{1}{2} - \frac{|\tau|\, c_V}{|\tau|^w - 1} \geq 0.$$

The condition is violated for $w = 2$ and $|\tau|$ equal to $\sqrt{2}$, $\sqrt{3}$ or $\sqrt{4}$, and for $w = 3$ and $|\tau| = \sqrt{2}$, see Table 3.5.1 on the following page. Since ν is monotonic increasing for $|\tau|$ and for w, there are no other "problematic cases".

For those cases we will use (c). For each of the "problematic cases" an ℓ satisfying the condition (1) of equivalences in (c) was found, see Table 3.5.2 on the next page for the results. Thus the statement is proved.

(e) Analogously to the proof of (a), except that we use ℓ for the upper bound of the sum, we obtain for $v \in V$

$$\begin{aligned}\left|\mathsf{value}(0.\eta_{-1}\dots\eta_{-\ell}) + \tau^{-\ell}v\right| &\le \left|\mathsf{value}(0.\eta_{-1}\dots\eta_{-\ell})\right| + \left|\tau^{-\ell}\right| |\tau|\, c_V \\ &\le \frac{|\tau|^w c_V}{1-|\tau|^{-w}}\left(1-|\tau|^{-w\left\lfloor\frac{\ell-1+w}{w}\right\rfloor}\right) + |\tau|^{-\ell+1} c_V \\ &\le \frac{|\tau|^w c_V}{1-|\tau|^{-w}}\left(1-|\tau|^{-\ell+1-w} + |\tau|^{-\ell+1-w}\left(1-|\tau|^{-w}\right)\right) \\ &= \frac{|\tau|^w c_V}{1-|\tau|^{-w}}\left(1-|\tau|^{-\ell+1-2w}\right).\end{aligned}$$

Since $1-|\tau|^{-\ell+1-2w} < 1$ we get

$$\left|\mathsf{value}(0.\eta_{-1}\dots\eta_{-\ell}) + \tau^{-\ell}V\right| \le \frac{|\tau|^w c_V}{1-|\tau|^{-w}} = f_U$$

for all $\ell \in \mathbb{N}_0$.

Let $z \in \mathbb{C}$. Have again a look at the proof of (d). If $\nu > 0$ there, we get that $|z| \le f_U$ implies $z \in \tau^{2w-1}V$.

Combining these two results yields the inclusion for $\nu > 0$, i.e., the "problematic cases" are left. Again, each of these cases has to be considered separately.

For each of the problematic cases, we find a $k \in \mathbb{N}_0$ such that

$$\overline{\mathcal{B}}\Big(\mathsf{value}(\boldsymbol{\xi}), 2\,|\tau|^{-k} f_U\Big) \subseteq \tau^{2w-1}V$$

holds for all $\boldsymbol{\xi} \in \mathbf{NAF}_w^{0.k}$. These k are listed in Table 3.5.3 on page 55.

For $\ell > k$ and $v \in V$, we obtain

$$\begin{aligned}\left|\mathsf{value}\big(0.0\dots0\eta_{-(k+1)}\dots\eta_{-\ell}\big) + \tau^{-\ell}v\right| &\le |\tau|^{-k} f_U + |\tau|^{-\ell}|\tau|\, c_V \\ &\le |\tau|^{-k} f_U\left(1+\frac{c_V}{f_U}\right) \\ &= |\tau|^{-k} f_U\left(1+\frac{1-|\tau|^{-w}}{|\tau|^w}\right) \\ &\le 2\,|\tau|^{-k} f_U\end{aligned}$$

using (a), Proposition 3.2.5 on page 41, and $|\tau|^w > 1$. Thus the desired inclusion follows for $\ell > k$.

For the finitely many $\ell \le k$ we additionally check all possibilities, i.e., whether for all combinations of $\boldsymbol{\vartheta} \in \mathbf{NAF}_w^{0.\ell}$ and vertices of the boundary of $\mathsf{value}(\boldsymbol{\vartheta}) + \tau^{-\ell}V$ the corresponding value is inside $\tau^{2w-1}V$. Convexity of V is used here. All combinations were valid, see last column of Table 3.5.3 on page 55, thus the inclusion proved. □

	$w = 2$	$w = 3$	$w = 4$
$\lvert\tau\rvert = \sqrt{2}$	−0.58012	−0.09074	0.13996
$\lvert\tau\rvert = \sqrt{3}$	−0.16144	0.18474	0.33464
$\lvert\tau\rvert = \sqrt{4}$	−0.00918	0.28178	0.39816
$\lvert\tau\rvert = \sqrt{5}$	0.07304	0.33224	0.42884

Table 3.5.1: Values (given five decimal places) of $\nu = \frac{1}{2} - \frac{\lvert\tau\rvert c_V}{\lvert\tau\rvert^{w-1}}$ for different $\lvert\tau\rvert$ and w. A negative sign means that this value is a "problematic value".

$q = \lvert\tau\rvert^2$	p	$\mathrm{Re}(\tau)$	$\mathrm{Im}(\tau)$	w	ℓ found?	ℓ	$\lvert\tau\rvert^{-\ell} f_U$	ε
2	−2	−1	1	2	*true*	8	0.1909	0.03003
2	−1	−0.5	1.323	2	*true*	4	0.7638	0.02068
2	0	0	1.414	2	*true*	6	0.3819	0.1484
2	1	0.5	1.323	2	*true*	4	0.7638	0.02068
2	2	1	1	2	*true*	7	0.27	0.08352
2	−2	−1	1	3	*true*	2	1.671	0.4505
2	−1	−0.5	1.323	3	*true*	2	1.671	0.4726
2	0	0	1.414	3	*true*	2	1.671	0.4505
2	1	0.5	1.323	3	*true*	2	1.671	0.4726
2	2	1	1	3	*true*	2	1.671	0.4505
3	−3	−1.5	0.866	2	*true*	1	1.984	0.03641
3	−2	−1	1.414	2	*true*	2	1.146	0.5543
3	−1	−0.5	1.658	2	*true*	2	1.146	0.4581
3	0	0	1.732	2	*true*	1	1.984	0.03641
3	1	0.5	1.658	2	*true*	2	1.146	0.4581
3	2	1	1.414	2	*true*	2	1.146	0.5543
3	3	1.5	0.866	2	*true*	1	1.984	0.03641
4	−3	−1.5	1.323	2	*true*	1	2.037	0.9164
4	−2	−1	1.732	2	*true*	1	2.037	0.4633
4	−1	−0.5	1.936	2	*true*	1	2.037	0.3227
4	0	0	2	2	*true*	1	2.037	0.9633
4	1	0.5	1.936	2	*true*	1	2.037	0.3227
4	2	1	1.732	2	*true*	1	2.037	0.4633
4	3	1.5	1.323	2	*true*	1	2.037	0.9164

Table 3.5.2: Upper bound inclusion $\mathsf{value}(\boldsymbol{\eta}) \in \tau^{2w-1}V$ checked for "problematic values" of $\lvert\tau\rvert$ and w, cf. (d) of Proposition 3.5.1 on page 51. The dependence of p, q and τ is given by $\tau^2 - p\tau + q = 0$. We have $p^2 < 4q$, since τ is assumed to be imaginary quadratic.

$q = \lvert\tau\rvert^2$	p	$\mathrm{Re}(\tau)$	$\mathrm{Im}(\tau)$	w	k found?	k	$2\lvert\tau\rvert^{-k} f_U$	ε	valid for $\ell \leq k$?
2	−2	−1	1	2	*true*	10	0.1909	0.03003	*true*
2	−1	−0.5	1.323	2	*true*	7	0.5401	0.138	*true*
2	0	0	1.414	2	*true*	8	0.3819	0.1484	*true*
2	1	0.5	1.323	2	*true*	7	0.5401	0.138	*true*
2	2	1	1	2	*true*	9	0.27	0.03933	*true*
2	−2	−1	1	3	*true*	4	1.671	0.2737	*true*
2	−1	−0.5	1.323	3	*true*	4	1.671	0.2682	*true*
2	0	0	1.414	3	*true*	4	1.671	0.0969	*true*
2	1	0.5	1.323	3	*true*	4	1.671	0.2682	*true*
2	2	1	1	3	*true*	4	1.671	0.2737	*true*
3	−3	−1.5	0.866	2	*true*	3	1.323	0.5054	*true*
3	−2	−1	1.414	2	*true*	3	1.323	0.04922	*true*
3	−1	−0.5	1.658	2	*true*	4	0.7638	0.4729	*true*
3	0	0	1.732	2	*true*	3	1.323	0.5054	*true*
3	1	0.5	1.658	2	*true*	4	0.7638	0.4729	*true*
3	2	1	1.414	2	*true*	3	1.323	0.04922	*true*
3	3	1.5	0.866	2	*true*	3	1.323	0.5054	*true*
4	−3	−1.5	1.323	2	*true*	2	2.037	0.9164	*true*
4	−2	−1	1.732	2	*true*	2	2.037	0.4633	*true*
4	−1	−0.5	1.936	2	*true*	2	2.037	0.3227	*true*
4	0	0	2	2	*true*	2	2.037	0.9633	*true*
4	1	0.5	1.936	2	*true*	2	2.037	0.3227	*true*
4	2	1	1.732	2	*true*	2	2.037	0.4633	*true*
4	3	1.5	1.323	2	*true*	2	2.037	0.9164	*true*

Table 3.5.3: Upper bound inclusion $\mathsf{value}(\eta_1 \dots \eta_\ell) + \tau^{-\ell} V \subseteq \tau^{2w-1} V$ checked for "problematic values" of $|\tau|$ and w, cf. (e) of Proposition 3.5.1 on page 51. The dependence of p, q and τ is given by $\tau^2 - p\tau + q = 0$. We have $p^2 < 4q$, since τ is assumed to be imaginary quadratic.

Next we want to find a lower bound for the fractional value of a w-NAF. Clearly the w-NAF 0 has fractional value 0, so we are interested in cases, where we have a non-zero digit somewhere.

Proposition 3.5.2 (Lower Bound for the Fractional Value)**.** *The following is true:*

(a) The following two statements are equivalent:

(1) There is an $\ell \in \mathbb{N}_0$, such that for all $\boldsymbol{\xi} \in \mathbf{NAF}_w^{0.\ell}$ with non-zero most significant digit the condition

$$|\mathsf{value}(\boldsymbol{\xi})| > |\tau|^{-\ell} f_U$$

is fulfilled.

(2) There exists a $\widetilde{\nu} > 0$, such that for all $\boldsymbol{\vartheta} \in \mathbf{NAF}_w^{0.\infty}$ with non-zero most significant digit the condition

$$|\mathsf{value}(\boldsymbol{\vartheta})| \geq |\tau|^{-1} \widetilde{\nu}.$$

holds.

(b) Let $\boldsymbol{\eta} \in \mathbf{NAF}_w^{0.\infty}$ with non-zero most significant digit. Then

$$|\mathsf{value}(\boldsymbol{\eta})| \geq |\tau|^{-1} f_L$$

with $f_L = \nu$ if $\nu > 0$, where

$$\nu = \frac{1}{2} - \frac{|\tau|\, c_V}{|\tau|^w - 1}.$$

If $\nu \leq 0$, see Table 3.5.1 on page 54, then we set $f_L = \widetilde{\nu}$ from Table 3.5.4 on page 58.

Proof of Proposition 3.5.2. (a) We have to prove both directions.

(1) $\Longrightarrow$ (2). Suppose there exists such an $\ell \in \mathbb{N}_0$. We set

$$\widetilde{\nu} = \min\left\{ |\tau| \left(|\mathsf{value}(\boldsymbol{\xi})| - |\tau|^{-\ell} f_U \right) \,\middle|\, \boldsymbol{\xi} \in \mathbf{NAF}_w^{0.\ell} \text{ with } \xi_{-1} \neq 0 \right\}.$$

Then clearly $\widetilde{\nu} > 0$. Using (a) of Proposition 3.5.1 on page 51 with digits shifted ℓ to the right, i.e., multiplication by $\tau^{-\ell}$, the desired result follows by using the triangle inequality.

(2) $\Longrightarrow$ (1). Now suppose there exists such a lower bound $\widetilde{\nu} > 0$. Then there is an $\ell \in \mathbb{N}_0$ such that $|\tau|^{-\ell} f_U < |\tau|^{-1} \widetilde{\nu}$. Since

$$|\mathsf{value}(0.\eta_{-1} \dots \eta_{-\ell})| \geq |\tau|^{-1} \widetilde{\nu} > |\tau|^{-\ell} f_U$$

for all w-NAFs $0.\eta_{-1} \dots \eta_{-\ell}$, the statement follows.

(b) Set

$$M := \tau^w \, \mathsf{value}(\boldsymbol{\eta}) = \eta_{-1}\tau^{w-1} + \sum_{i=2}^{\ell} \eta_{-i}\tau^{w-i}$$

Since $\eta_{-1} \neq 0$, we can rewrite this to get

$$M = \eta_{-1}\tau^{w-1} + \sum_{i=w+1}^{\ell} \eta_{-i}\tau^{w-i} = \eta_{-1}\tau^{w-1} + \sum_{k=1}^{\ell-w} \eta_{-(w+k)}\tau^{-k}.$$

Now consider the Voronoi cell $V_{\eta_{-1}}$ for η_{-1} and $V_0 = V$ for 0. Since $\eta_{-1} \neq 0$, these two are disjoint, except parts of the boundary, if they are adjacent.

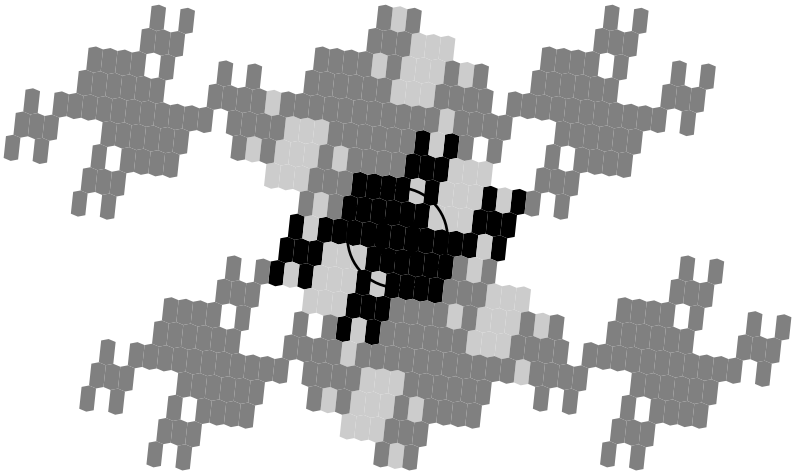


Figure 3.5.1: Lower bound for $\tau = \frac{1}{2} + \frac{1}{2}\sqrt{-11}$ and $w = 2$. The procedure stopped at $\ell = 6$. The large circle has radius $|\tau|^{-\ell} f_U$, the small circle is our lower bound with radius $f_L = \widetilde{\nu}$. The dot inside represents zero. The grey region has most significant digit zero, the black ones non-zero.

We know from (b) of Proposition 3.5.1 on page 51, that

$$M - \eta_{-1}\tau^{w-1} = \sum_{k=1}^{\ell-w} \eta_{-(w+k)}\tau^{-k} \in \bigcup_{z \in \tau^{w-1}V} \overline{\mathcal{B}}\Big(z, |\tau|^{-w} f_U\Big),$$

so

$$M \in \bigcup_{z \in \tau^{w-1}V_{\eta_{-1}}} \overline{\mathcal{B}}\Big(z, |\tau|^{-w} f_U\Big).$$

This means that M is in $\tau^{w-1}V_{\eta_{-1}}$ or in a $|\tau|^{-w} f_U$-strip around this cell.

Now we are looking at $\tau^{w-1}V_0$ and using Proposition 3.2.5 on page 41, from which we know that $\overline{\mathcal{B}}\Big(0, \frac{1}{2}|\tau|^{w-1}\Big)$ is inside such a Voronoi cell. Thus, we get

$$|M| \geq \frac{1}{2}|\tau|^{w-1} - |\tau|^{-w} f_U = \frac{1}{2}|\tau|^{w-1} - \frac{c_V}{1 - |\tau|^{-w}} = |\tau|^{w-1}\nu$$

for our lower bound of M and therefore, by multiplying with τ^{-w} one for $\mathsf{value}(\boldsymbol{\eta})$.

Looking in Table 3.5.1 on page 54, we see that there are some values where ν is not positive. As in Proposition 3.5.1 on page 51, this is the case, if $w = 2$ and $|\tau|$ is $\sqrt{2}$, $\sqrt{3}$ or $\sqrt{4}$, and if $w = 3$ and $|\tau| = \sqrt{2}$. Since ν is monotonic increasing with $|\tau|$ and monotonic increasing with w, there are no other non-positive values of ν than the above mentioned.

For those finite many problem cases, we use (a) to find a $\widetilde{\nu}$. The results are listed in Table 3.5.4 on the following page and an example is drawn in Figure 3.5.1. □

Combining the previous two Propositions leads to the following corollary, which gives an upper and a lower bound for the absolute value of a w-NAF by looking at the largest non-zero index.

Corollary 3.5.3 (Bounds for the Value)**.** *Let $\boldsymbol{\eta} \in \mathbf{NAF}_w^{\mathsf{fin},\infty}$, then we get*

$$\mathrm{d}_{\mathsf{NAF}}(\boldsymbol{\eta}, \mathbf{0})\, f_L \leq |\mathsf{value}(\boldsymbol{\eta})| \leq \mathrm{d}_{\mathsf{NAF}}(\boldsymbol{\eta}, \mathbf{0})\, f_U\, |\tau|\,.$$

$q = \lvert\tau\rvert^2$	p	$\mathrm{Re}(\tau)$	$\mathrm{Im}(\tau)$	w	ℓ	$\lvert\tau\rvert^{-\ell} f_U$	$\widetilde{\nu}$	$\log_{\lvert\tau\rvert}(f_U/\widetilde{\nu})$
2	−2	−1	1	2	9	0.135	0.004739	18.66
2	−1	−0.5	1.323	2	7	0.27	0.105	9.726
2	0	0	1.414	2	8	0.1909	0.07422	10.73
2	1	0.5	1.323	2	7	0.27	0.105	9.726
2	2	1	1	2	9	0.135	0.04176	12.39
2	−2	−1	1	3	6	0.4177	0.1126	9.782
2	−1	−0.5	1.323	3	6	0.4177	0.04999	12.13
2	0	0	1.414	3	6	0.4177	0.0153	15.54
2	1	0.5	1.323	3	6	0.4177	0.04999	12.13
2	2	1	1	3	6	0.4177	0.1126	9.782
3	−3	−1.5	0.866	2	4	0.3819	0.003019	12.81
3	−2	−1	1.414	2	5	0.2205	0.04402	7.933
3	−1	−0.5	1.658	2	5	0.2205	0.08717	6.689
3	0	0	1.732	2	4	0.3819	0.003019	12.81
3	1	0.5	1.658	2	5	0.2205	0.08717	6.689
3	2	1	1.414	2	5	0.2205	0.04402	7.933
3	3	1.5	0.866	2	4	0.3819	0.003019	12.81
4	−3	−1.5	1.323	2	4	0.2546	0.07613	5.742
4	−2	−1	1.732	2	5	0.1273	0.03807	6.742
4	−1	−0.5	1.936	2	4	0.2546	0.0516	6.303
4	0	0	2	2	5	0.1273	0.07035	5.856
4	1	0.5	1.936	2	4	0.2546	0.0516	6.303
4	2	1	1.732	2	5	0.1273	0.0467	6.447
4	3	1.5	1.323	2	4	0.2546	0.07613	5.742

Table 3.5.4: Lower bounds for "problematic values" of $|\tau|$ and w, cf. (b) of Proposition 3.5.2 on page 56. The dependence of p, q and τ is given by $\tau^2 - p\tau + q = 0$. We have $p^2 < 4q$, since τ is assumed to be imaginary quadratic.

Proof. Follows directly from Proposition 3.5.1 on page 51 and Proposition 3.5.2 on page 56. □

Last in this section, we want to find out, if there are special w-NAFs, for which we know for sure that all their expansions start with a certain finite w-NAF. We will show the following lemma.

Lemma 3.5.4. *Let*

$$k \geq k_0 = \max\{19, 2w+5\},$$

let $\boldsymbol{\eta} \in \mathbf{NAF}_w^{0,\infty}$ *start with the word* 0^k*, i.e.,* $\eta_{-1} = 0, \ldots, \eta_{-k} = 0$*, and set* $z = \mathsf{value}(\boldsymbol{\eta})$*. Then we get for all* $\boldsymbol{\xi} \in \mathbf{NAF}_w^{\mathsf{fin},\infty}$ *that* $z = \mathsf{value}(\boldsymbol{\xi})$ *implies* $\boldsymbol{\xi} \in \mathbf{NAF}_w^{0,\infty}$*.*

Proof. Let $\boldsymbol{\xi}_I.\boldsymbol{\xi}_F \in \mathbf{NAF}_w^{\mathsf{fin},\infty}$. Then $|\mathsf{value}(\boldsymbol{\xi}_I.\boldsymbol{\xi}_F)| < f_L$ implies $\boldsymbol{\xi}_I = \mathbf{0}$, cf. Proposition 3.5.2 on page 56. For our $\boldsymbol{\eta}$ we obtain $z = |\mathsf{value}(\boldsymbol{\eta})| \leq |\tau|^{-k} f_U$, cf. Proposition 3.5.1 on page 51. So we have to show that

$$|\tau|^{-k} f_U < f_L,$$

which is equivalent to

$$k > \log_{|\tau|} \frac{f_U}{f_L}.$$

For the "non-problematic cases", cf. Propositions 3.5.1 on page 51 and 3.5.2 on page 56, we obtain

$$k > 2w - 1 + \log_{|\tau|} A$$

with

$$A := \frac{1}{\nu} \frac{|\tau|\, c_V}{|\tau|^w - 1} = \left(\frac{|\tau|^w - 1}{2\,|\tau|\, c_V} - 1 \right)^{-1} > 0,$$

where we just inserted the formulas for f_U, f_L and ν, and used $\nu > 0$.

Consider the partial derivation of $\log_{|\tau|} A$ with respect to $|\tau|$. We get

$$\frac{\partial \log_{|\tau|} A}{\partial\, |\tau|} = \underbrace{\frac{1}{\log_e |\tau|}}_{>0} \underbrace{\nu}_{>0} \underbrace{\frac{|\tau|^w - 1}{|\tau|\, c_V}}_{>0} \underbrace{\frac{\partial A}{\partial\, |\tau|}}_{<0} < 0,$$

where we used $|\tau| > 1$, $w \geq 2$, and the fact that the quotient of polynomials $\frac{|\tau|^w - 1}{2|\tau| c_V}$ is monotonic increasing with $|\tau|$. Further we see that A is monotonic decreasing with w, therefore $\log_{|\tau|} A$, too.

For $|\tau| = \sqrt{5}$ and $w = 2$ we get $\log_{|\tau|} A = 5.84522$, for $|\tau| = \sqrt{3}$ and $w = 3$ we get $\log_{|\tau|} A = 1.70649$, and for $|\tau| = \sqrt{2}$ and $w = 4$ we get $\log_{|\tau|} A = 2.57248$. Using the monotonicity from above yields $k \geq 2w + 5$ for the "non-problematic cases".

For our "problematic cases", the value of $\log_{|\tau|} \frac{f_U}{f_L}$ is calculated in Table 3.5.4 on the facing page. Therefore we obtain $k \geq 19$. □

3.6 Numeral Systems with Non-Adjacent Forms

Let $\tau \in \mathbb{C}$ be an algebraic integer, imaginary quadratic. Suppose that $|\tau| > 1$. Let $w \in \mathbb{N}$ with $w \geq 2$. Further let $\mathcal{D}$ be a minimal norm representatives digit set modulo τ^w as in Definition 3.3.5 on page 45.

We are now able to show that in this setting, the digit set of minimal norm representatives is indeed a width-w non-adjacent digit set. This is then extended to infinite fractional expansions of elements in $\mathbb{C}$.

Theorem 3.6.1 (Existence and Uniqueness Theorem concerning Lattice Points)**.** *For each lattice point $z \in \mathbb{Z}[\tau]$ there is a unique element $\boldsymbol{\xi} \in \mathbf{NAF}_w^{\mathsf{fin}}$, such that $z = \mathsf{value}(\boldsymbol{\xi})$. Thus $\mathcal{D}$ is a width-w non-adjacent digit set. The w-NAF $\boldsymbol{\xi}$ can be calculated using Algorithm 3.6.1, i.e., this algorithm terminates and is correct.*

The uniqueness result is well known. The existence result is only known for special τ and w. For example in Koblitz [40] the case $\tau = \pm\frac{3}{2} + \frac{1}{2}\sqrt{-3}$ and $w = 2$ was shown. There the digit set $\mathcal{D}$ consists of 0 and powers of primitive sixth roots of unity. Blake, Kumar Murty and Xu [15] generalised that for $w \geq 2$. Another example is given in Solinas [64]. There $\tau = \pm\frac{1}{2} + \frac{1}{2}\sqrt{-7}$ and $w = 2$ is used, and the digit set $\mathcal{D}$ consists of 0 and ± 1. This result was generalised by Blake, Kumar Murty and Xu [14] for $w \geq 2$. The cases $\tau = 1 + \sqrt{-1}$, $\tau = \sqrt{-2}$ and $\tau = \frac{1}{2} + \frac{1}{2}\sqrt{-11}$ were studied in Blake, Kumar Murty and Xu [12].

Algorithm 3.6.1 Algorithm to calculate a w-NAF $\boldsymbol{\xi} \in \mathbf{NAF}_w^{\mathsf{fin}}$ for an element $z \in \mathbb{Z}[\tau]$.

```
 1: ℓ ← 0
 2: y ← z
 3: while y ≠ 0 do
 4:    if τ | y then
 5:       ξ_ℓ ← 0
 6:    else
 7:       Let ξ_ℓ ∈ D such that ξ_ℓ ≡ y (mod τ^w)
 8:    end if
 9:    y ← (y − ξ_ℓ)/τ
10:    ℓ ← ℓ + 1
11: end while
12: ξ ← ξ_{ℓ−1} ξ_{ℓ−2} ... ξ_0
13: return ξ
```

The proof follows a similar idea as in Section 3.5 on page 50 and in Matula [42]. There are again two parts, one analytic part for all but finitely many cases, and the other, which proves the remaining by the help of a simple procedure.

Proof. First we show that the algorithm terminates. Let $y \in \mathbb{Z}[\tau]$ and consider Algorithm 3.6.1 in cycle ℓ. If $\tau \mid y$, then in the next step the norm $|y|^2 \in \mathbb{N}_0$ becomes smaller since $|\tau|^2 > 1$.

Let $\tau \nmid y$. If $|y| < \frac{1}{2}|\tau|^w$, then $y \in \mathcal{D}$, cf. Proposition 3.2.5 on page 41 and Remark 3.3.6 on page 45. Thus the algorithm terminates in the next cycle. If

$$|y| > \frac{|\tau|\, c_V}{1 - |\tau|^{-w}} = \frac{|\tau|^{w+1} c_V}{|\tau|^w - 1},$$

we obtain

$$|y|\,|\tau|^w > |y| + |\tau|^{w+1} c_V \geq |y| + |\xi_\ell| \geq |y - \xi_\ell|\,,$$

which is equivalent to

$$\left|\frac{y - \xi_\ell}{\tau^w}\right| < |y|\,.$$

So if the condition

$$\frac{|\tau|^{w+1} c_V}{|\tau|^w - 1} < \frac{1}{2}|\tau|^w \iff \nu = \frac{1}{2} - \frac{|\tau|\, c_V}{|\tau|^w - 1} > 0,$$

with the same ν as in Proposition 3.5.2 on page 56, is fulfilled, the norm $|y|^2 \in \mathbb{N}_0$ is descending and therefore the algorithm terminating.

Now we consider the case, when $\nu \leq 0$. According to Table 3.5.1 on page 54 there are the same finitely many combinations of τ and w to check as in Proposition 3.5.1 on page 51 and Proposition 3.5.2 on page 56. For each of them, there is only a finite number of elements $y \in \mathbb{Z}[\tau]$ with

$$|y| \leq \frac{|\tau|^{w+1} c_V}{|\tau|^w - 1},$$

so altogether only finitely many $y \in \mathbb{Z}[\tau]$ left to check, whether they admit a w-NAF or not. The results can be found in the table in Appendix A. Every element that was to check, has a w-NAF.

To show the correctness, again let $y \in \mathbb{Z}[\tau]$ and consider Algorithm 3.6.1 on the preceding page in cycle ℓ. If τ divides y, then we append a digit $\xi_\ell = 0$. Otherwise y is congruent to a non-zero element ξ_ℓ of $\mathcal{D}$ modulo τ^w, since the digit set $\mathcal{D}$ was constructed in that way, cf. Definitions 3.3.1 and 3.3.5. The digit ξ_ℓ is appended. Because τ^w divides $y - \xi_\ell$, the next $w - 1$ digits will be zero. Therefore a correct w-NAF is produced.

For the uniqueness let $\boldsymbol{\xi} \in \mathbf{NAF}_w^{\mathsf{fin}}$ be an expansions for the element $z \in \mathbb{Z}[\tau]$. If $\tau \mid z$, then

$$0 \equiv z = \mathsf{value}(\boldsymbol{\xi}) \equiv \xi_0 \pmod{\tau},$$

so $\tau \mid \xi_0 \in \mathcal{D}$. Therefore $\xi_0 = 0$. If $\tau \nmid z$, then $\tau \nmid \xi_0$ and so $\xi_0 \neq 0$. This implies $\xi_1 = 0, \ldots, \xi_{w-1} = 0$. This means ξ_0 lies in the same residue class modulo τ^w as exactly one non-zero digit of $\mathcal{D}$ (per construction of the digit set, cf. Definitions 3.3.1 and 3.3.5), hence they are equal. Induction finishes the proof of the uniqueness. □

So we have that all elements of our lattice $\mathbb{Z}[\tau]$ have a unique expansion. Now we want to get a step further and look at all elements of $\mathbb{C}$. We will need the following three lemmata, to prove that every element of $\mathbb{C}$ has a w-NAF-expansion.

Lemma 3.6.2. *The function* $\mathsf{value}|_{\mathbf{NAF}_w^{\mathsf{fin.fin}}}$ *is injective.*

Proof. Let $\boldsymbol{\eta}$ and $\boldsymbol{\xi}$ be elements of $\mathbf{NAF}_w^{\mathsf{fin.fin}}$ with $\mathsf{value}(\boldsymbol{\eta}) = \mathsf{value}(\boldsymbol{\xi})$. This implies that $\tau^J \mathsf{value}(\boldsymbol{\eta}) = \tau^J \mathsf{value}(\boldsymbol{\xi}) \in \mathbb{Z}[\tau]$ for some $J \in \mathbb{Z}$. By uniqueness of the integer w-NAFs, see Theorem 3.6.1 on the facing page, we conclude that $\boldsymbol{\eta} = \boldsymbol{\xi}$. □

Lemma 3.6.3. *We have* $\mathsf{value}\Big(\mathbf{NAF}_w^{\mathsf{fin.fin}}\Big) = \mathbb{Z}[1/\tau]$.

Proof. Let $\boldsymbol{\eta} \in \mathbf{NAF}_w^{\mathsf{fin.fin}}$ and $\eta_j = 0$ for all $|j| > J$ for some $J \geq 1$. Then $\tau^J \mathsf{value}(\boldsymbol{\eta}) \in \mathbb{Z}[\tau]$, which implies that there are some $a, b \in \mathbb{Z}$ such that

$$\mathsf{value}(\boldsymbol{\eta}) = a\tau^{-(J-1)} + b\tau^{-J} \in \mathbb{Z}[1/\tau].$$

Conversely, if

$$z = \sum_{j=0}^{J} \eta_j \tau^{-j} \in \mathbb{Z}[1/\tau],$$

we have $\tau^J z \in \mathbb{Z}[\tau]$. Since every element of $\mathbb{Z}[\tau]$ admits an integer w-NAF, see Theorem 3.6.1 on the preceding page, there is an $\boldsymbol{\xi} \in \mathbf{NAF}_w^{\mathsf{fin.fin}}$ with $\mathsf{value}(\boldsymbol{\xi}) = z$. □

Lemma 3.6.4. $\mathbb{Z}[1/\tau]$ *is dense in* $\mathbb{C}$.

Proof. Let $z \in \mathbb{C}$ and $K \geq 0$. Then $\tau^K z = u + v\tau$ for some reals u and v. We have

$$\left| z - \frac{\lfloor u \rfloor + \lfloor v \rfloor \tau}{\tau^K} \right| < \frac{1 + |\tau|}{|\tau|^K},$$

which proves the lemma. □

Now we can prove the following theorem.

Theorem 3.6.5 (Existence Theorem concerning $\mathbb{C}$)**.** *Let $z \in \mathbb{C}$. Then there is an $\boldsymbol{\eta} \in \mathbf{NAF}_w^{\mathsf{fin}.\infty}$ such that $z = \mathsf{value}(\boldsymbol{\eta})$, i.e., each complex number has a w-NAF-expansion.*

Proof. By Lemma 3.6.4 on the preceding page, there is a sequence $z_n \in \mathbb{Z}[1/\tau]$ converging to z. By Lemma 3.6.3 on the previous page, there is a sequence $\boldsymbol{\eta}_n \in \mathbf{NAF}_w^{\mathsf{fin}.\mathsf{fin}}$ with $\mathsf{value}(\boldsymbol{\eta}_n) = z_n$ for all n. By Corollary 3.5.3 on page 57 the sequence $\mathrm{d}_{\mathsf{NAF}}(\boldsymbol{\eta}_n, 0)$ is bounded from above, so there is an ℓ such that $\boldsymbol{\eta}_n \in \mathbf{NAF}_w^{\ell.\mathsf{fin}} \subseteq \mathbf{NAF}_w^{\ell.\infty}$. By Proposition 3.3.9 on page 47, we conclude that there is a convergent subsequence $\boldsymbol{\eta}'_n$ of $\boldsymbol{\eta}_n$. Set $\boldsymbol{\eta} := \lim_{n\to\infty} \boldsymbol{\eta}'_n$. By continuity of value, see Proposition 3.3.8 on page 46, we conclude that $\mathsf{value}(\boldsymbol{\eta}) = z$. □

3.7 The Fundamental Domain

Let $\tau \in \mathbb{C}$ be an algebraic integer, imaginary quadratic. Suppose that $|\tau| > 1$. Let $w \in \mathbb{N}$ with $w \geq 2$. Further let $\mathcal{D}$ be a minimal norm representatives digit set modulo τ^w as in Definition 3.3.5 on page 45.

We now derive properties of the *Fundamental Domain*, i.e., the set of complex numbers representable by w-NAFs which vanish left of the τ-point. The boundary of the fundamental domain is shown to correspond to complex numbers which admit more than one w-NAF differing left of the τ-point. Finally, an upper bound for the Hausdorff dimension of the boundary is derived.

Definition 3.7.1 (Fundamental Domain)**.** The set

$$\mathcal{F} := \mathsf{value}\big(\mathbf{NAF}_w^{0.\infty}\big) = \left\{ \mathsf{value}(\boldsymbol{\xi}) \,\middle|\, \boldsymbol{\xi} \in \mathbf{NAF}_w^{0.\infty} \right\}.$$

is called *fundamental domain.*

The pictures in Figure 3.9.1 on page 73 can also be reinterpreted as fundamental domains for the τ and w given there. The definition of the fundamental domain for a general $\tau \in \mathbb{C}$ and a general finite digit set containing zero is meaningful, too. The same is true for following proposition, which is also valid for general $\tau \in \mathbb{C}$ and a general finite digit set including zero.

Proposition 3.7.2. *The fundamental domain $\mathcal{F}$ is compact.*

Proof. The set $\mathbf{NAF}_w^{0.\infty}$ is compact, cf. Proposition 3.3.9 on page 47. The compactness of the fundamental domain $\mathcal{F}$ follows, since $\mathcal{F}$ is the image of $\mathbf{NAF}_w^{0.\infty}$ under the continuous function value, cf. Proposition 3.3.8 on page 46. □

We can also compute the Lebesgue measure of the fundamental domain. This result can be found in Remark 3.9.3 on page 77. We will need the results of Sections 3.8 and 3.9 for calculating $\lambda(\mathcal{F})$.

Next we want to get more properties of the fundamental domain. We will need the following proposition, which will be extended in Proposition 3.7.7 on page 64.

Proposition 3.7.3. *Let $z \in \mathcal{F}$. If there exists a w-NAF $\boldsymbol{\xi}_I.\boldsymbol{\xi}_F \in \mathbf{NAF}_w^{\mathsf{fin}.\infty}$ with $\boldsymbol{\xi}_I \neq \mathbf{0}$ and such that $z = \mathsf{value}(\boldsymbol{\xi}_I.\boldsymbol{\xi}_F)$, then $z \in \partial\mathcal{F}$.*

Proof. Assume that $z \in \operatorname{int}\mathcal{F}$. Then there is an $\varepsilon_z > 0$ such that $\mathcal{B}(z,\varepsilon_z) \subseteq \operatorname{int}\mathcal{F}$. Let ε_z be small enough such that there exists a $y \in \mathcal{B}(z,\varepsilon_z) \cap \mathbb{Z}[1/\tau]$ and a $\boldsymbol{\vartheta} = \boldsymbol{\vartheta}_I.\boldsymbol{\vartheta}_F \in \mathbf{NAF}_w^{\mathsf{fin.fin}}$ with $y = \mathsf{value}(\boldsymbol{\vartheta})$ and such that $\boldsymbol{\vartheta}_I \neq \mathbf{0}$. Let k be the right-length of $\boldsymbol{\vartheta}$.

Choose $0 < \varepsilon_y < \tau^{-k-w} f_L$ such that $\mathcal{B}(y,\varepsilon_y) \subseteq \operatorname{int}\mathcal{F}$. Since $y \in \mathcal{F}$ there is an $\boldsymbol{\eta} \in \mathbf{NAF}_w^{0.\infty}$ with $y = \mathsf{value}(\boldsymbol{\eta})$. Therefore, there is a $y' \in \mathcal{B}(y,\varepsilon_y) \cap \mathbb{Z}[1/\tau]$ with $y' = \mathsf{value}(\boldsymbol{\eta}')$ for some $\boldsymbol{\eta}' = \boldsymbol{\eta}'_I.\boldsymbol{\eta}'_F \in \mathbf{NAF}_w^{\mathsf{fin.fin}}$ with $\boldsymbol{\eta}'_I = \mathbf{0}$ (by "cutting" the infinite right side of $\boldsymbol{\eta}$).

As $y' - y \in \mathbb{Z}[1/\tau]$, there is a $\boldsymbol{\xi} \in \mathbf{NAF}_w^{\mathsf{fin.fin}}$ with $\mathsf{value}(\boldsymbol{\xi}) = y' - y$. By Corollary 3.5.3 we obtain $\mathrm{d}_{\mathsf{NAF}}(\boldsymbol{\xi},\mathbf{0}) < \varepsilon_y / f_L < \tau^{-k-w}$. Thus $\xi_\ell = 0$ for all $\ell \geq -k-(w-1)$.

Now $y' = y + (y'-y)$ and we get a $\boldsymbol{\vartheta}' = \boldsymbol{\vartheta}'_I.\boldsymbol{\vartheta}'_F \in \mathbf{NAF}_w^{\mathsf{fin.fin}}$ with $\mathsf{value}(\boldsymbol{\vartheta}') = y'$ by digit-wise addition of $\boldsymbol{\vartheta}$ and $\boldsymbol{\xi}$. Note that at each index at most one summand (digit) is non-zero and that the w-NAF-condition is fulfilled. We have $\boldsymbol{\vartheta}'_I \neq \mathbf{0}$, since $\boldsymbol{\vartheta}_I \neq \mathbf{0}$.

So we got two different w-NAFs in $\mathbf{NAF}_w^{\mathsf{fin.fin}}$ for one element $y' \in \mathbb{Z}[1/\tau]$, which is impossible due to uniqueness, see Lemma 3.6.2 on page 61. Thus we have a contradiction. □

The complex plane has a tiling property with respect to the fundamental domain. This fact is stated in the following corollary to Theorem 3.6.1 on page 60 and Theorem 3.6.5 on the preceding page.

Corollary 3.7.4 (Tiling Property)**.** *The complex plane can be tiled with scaled versions of the fundamental domain $\mathcal{F}$. Only finitely many different size are needed. More precisely: Let $K \in \mathbb{Z}$, then*

$$\mathbb{C} = \bigcup_{\substack{k \in \{K, K+1, \ldots, K+w-1\} \\ \boldsymbol{\xi} \in \mathbf{NAF}_w^{\mathsf{fin}} \\ k \neq K+w-1 \text{ implies } \boldsymbol{\xi}_0 \neq 0}} \left(\tau^k \mathsf{value}(\boldsymbol{\xi}) + \tau^{k-w+1}\mathcal{F}\right),$$

and the intersection of two different $\tau^k \mathsf{value}(\boldsymbol{\xi}) + \tau^{k-w+1}\mathcal{F}$ and $\tau^{k'-w+1} \mathsf{value}(\boldsymbol{\xi}') + \tau^{k'}\mathcal{F}$ in this union is a subset of the intersection of their boundaries.

Later, after Proposition 3.7.8 on page 65, we will know that the intersection of the two different sets of the tiling in the previous corollary has Lebesgue measure 0.

Proof of Corollary 3.7.4. Let $z \in \mathbb{C}$. Then, according to Theorem 3.6.5 on the preceding page, there is a $\boldsymbol{\eta} \in \mathbf{NAF}_w^{\mathsf{fin}.\infty}$ with $z = \mathsf{value}(\boldsymbol{\eta})$. We look at the block $\eta_{K+w-1} \ldots \eta_{K+1}\eta_K$. If this block is $\mathbf{0}$, then set $k = K+w-1$, otherwise there is at most one non-zero digit in it, which we call η_k. So the digits $\eta_{k-1}, \ldots, \eta_{k-w+1}$ are always zero. We set $\boldsymbol{\xi} = \ldots \eta_{k+1}\eta_k.0\ldots$, and we obtain

$$z - \tau^k \mathsf{value}(\boldsymbol{\xi}) \in \tau^{k-w+1}\mathcal{F}.$$

Now set $F = \tau^k \mathsf{value}(\boldsymbol{\xi}) + \tau^{k-w+1}\mathcal{F}$ and $F' = \tau^{k'} \mathsf{value}(\boldsymbol{\xi}') + \tau^{k'-w+1}\mathcal{F}$ in a way that both are in the union of the tiling with $(k,\boldsymbol{\xi}) \neq (k',\boldsymbol{\xi}')$ and consider their intersection I. Since every point in there has two different representations per construction, we conclude that $I \subseteq \partial F$ and $I \subseteq \partial F'$ by Proposition 3.7.3 on the facing page. □

Remark 3.7.5 (Iterated Function System). Let $\tau \in \mathbb{C}$ and $\mathcal{D}$ be a general finite digit set containing zero. We have two possibilities building the elements $\boldsymbol{\xi} \in \mathbf{NAF}_w^{0.\infty}$ from left to right. We can either append 0, what corresponds to a division through τ, so we define $f_0(z) = \frac{z}{\tau}$. Or we can append a non-zero digit $\vartheta \in \mathcal{D}^\bullet$ and then add $w-1$ zeros. In this case, we define $f_\vartheta(z) = \frac{\vartheta}{\tau} + \frac{z}{\tau^w}$. Thus we get the *iterated function system* $(f_\vartheta)_{\vartheta \in \mathcal{D}}$, cf. Edgar [19] or Barnsley [11]. All f_ϑ are *similarities*, and the iterated function system realizes the *ratio list* $(r_\vartheta)_{\vartheta \in \mathcal{D}}$ with $r_0 = |\tau|^{-1}$ and

for $\vartheta \in \mathcal{D}^\bullet$ with $r_\vartheta = |\tau|^{-w}$. So our set can be rewritten as

$$\mathcal{F} = \bigcup_{\vartheta\in\mathcal{D}} f_\vartheta(\mathcal{F}) = \frac{1}{\tau}\mathcal{F} \cup \bigcup_{\vartheta\in\mathcal{D}^\bullet} \left(\frac{\vartheta}{\tau} + \frac{1}{\tau^w}\mathcal{F}\right).$$

Furthermore, if we have an imaginary quadratic algebraic integer τ and a minimal norm representatives digit set, the iterated function system $(f_\vartheta)_{\vartheta\in\mathcal{D}}$ fulfils *Moran's open set condition*[1], cf. Edgar [19] or Barnsley [11]. The *Moran open set* used is $\operatorname{int}\mathcal{F}$. This set satisfies

$$f_\vartheta(\operatorname{int}\mathcal{F}) \cap f_{\vartheta'}(\operatorname{int}\mathcal{F}) = \emptyset$$

for $\vartheta \neq \vartheta' \in \mathcal{D}$ and

$$\operatorname{int}\mathcal{F} \supseteq f_\vartheta(\operatorname{int}\mathcal{F})$$

for all $\vartheta \in \mathcal{D}$. We remark that the first condition follows directly from the tiling property in Corollary 3.7.4 on the previous page with $K = -1$. The second condition follows from the fact that f_ϑ is an open mapping.

Now we want to have a look at a special case.

Remark 3.7.6 (Koch snowflake). Let $\tau = \frac{3}{2} + \frac{1}{2}\sqrt{-3}$ and $w = 2$. Then our digit set consists of 0 and powers of primitive sixth roots of unity, i.e., $\mathcal{D} = \{0\} \cup \left\{\zeta^k \,\middle|\, k \in \mathbb{N}_0 \text{ with } 0 \leq k < 6\right\}$ with $\zeta = e^{i\pi/3}$, cf. Koblitz [40].

We get

$$\mathcal{F} = \frac{1}{\tau}\mathcal{F} \cup \bigcup_{0\leq k<6} \left(\frac{\zeta^k}{\tau} + \frac{1}{\tau^2}\mathcal{F}\right).$$

Since the digit set is invariant with respect to multiplication by ζ, i.e., rotation by $\frac{\pi}{3}$, the same is true for $\mathcal{F}$. Using this and $\tau = \sqrt{3}e^{i\pi/6}$ yields

$$\mathcal{F} = \frac{e^{i\pi/2}}{\sqrt{3}}\mathcal{F} \cup \bigcup_{0\leq k<6} \left(\frac{e^{ik\pi/3+i\pi/2}}{\sqrt{3}} + \frac{1}{3}\mathcal{F}\right).$$

This is an iterated function system of the *Koch snowflake*[2], it is drawn in Figure 3.9.1c on page 73.

Next we want to have a look at the Hausdorff dimension of the boundary of $\mathcal{F}$. We will need the following characterisation of the boundary, which is an extension to Proposition 3.7.3 on page 62.

Proposition 3.7.7 (Characterisation of the Boundary)**.** *Let $z \in \mathcal{F}$. Then $z \in \partial\mathcal{F}$ if and only if there exists a w-NAF $\boldsymbol{\xi}_I.\boldsymbol{\xi}_F \in \mathbf{NAF}_w^{\mathsf{fin}.\infty}$ with $\boldsymbol{\xi}_I \neq \mathbf{0}$, such that $z = \mathsf{value}(\boldsymbol{\xi}_I.\boldsymbol{\xi}_F)$.*

Proof. Let $z \in \partial\mathcal{F}$. For every $\varepsilon > 0$, there is a $y \in \mathcal{B}(z,\varepsilon)$, such that $y \notin \mathcal{F}$. Thus we have a sequence $(y_j)_{j\geq 1}$ converging to z, where the y_j are not in $\mathcal{F}$. Therefore each y_j has a w-NAF-representation $\boldsymbol{\eta}_j \in \mathbf{NAF}_w^{\mathsf{fin}.\infty}$ with non-zero integer part. Now we will use the tiling property stated in Corollary 3.7.4 on the preceding page. The fundamental domain $\mathcal{F}$ can be surrounded by only finitely many scaled versions of $\mathcal{F}$. So there is a subsequence $(\boldsymbol{\vartheta}_j)_{j\in\mathbb{N}_0}$ of $(\boldsymbol{\eta}_j)_{j\in\mathbb{N}_0}$ with fixed integer part $\boldsymbol{\xi}_I \neq \mathbf{0}$. Due to compactness of $\mathcal{F}$, cf. Proposition 3.7.2 on page 62, we find a $\boldsymbol{\xi} = \boldsymbol{\xi}_I.\boldsymbol{\xi}_F$ with value z as limit of a converging subsequence of $(\boldsymbol{\vartheta}_j)_{j\in\mathbb{N}_0}$.

The other direction is just Proposition 3.7.3 on page 62, thus the proof is finished. □

[1] "Moran's open set condition" is sometimes just called "open set condition"

[2] The fact that the Koch snowflake has the mentioned iterated function system seems to be commonly known, although we were not able to find a reference, where this statement is proved. Any hints are welcome.

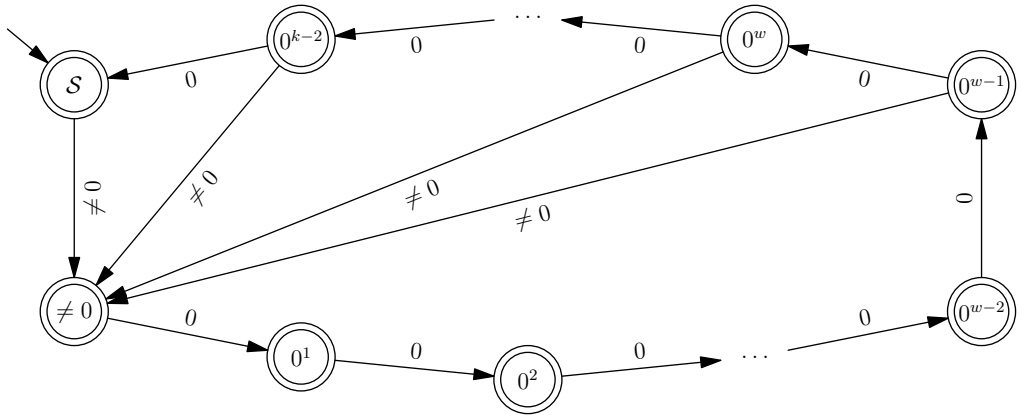


Figure 3.7.1: Automaton $\mathcal{A}$ recognising $\bigcup_{j\in\mathbb{N}} \widetilde{U}_j$ from right to left, see proof of Proposition 3.7.8. The state $\mathcal{S}$ is the starting state, all states are valid end states. An edges marked with $\neq 0$ means one edge for each non-zero digit in the digit set $\mathcal{D}$. The state $\neq 0$ means that there was an non-zero digit read, a state 0^ℓ means that ℓ zeros have been read.

The following proposition deals with the Hausdorff dimension of the boundary of $\mathcal{F}$.

Proposition 3.7.8. *For the Hausdorff dimension of the boundary of the fundamental domain we get* $\dim_H \partial\mathcal{F} < 2$.

The idea of this proof is similar to a proof in Heuberger and Prodinger [31].

Proof. Set $k := k_0 + w - 1$ with k_0 from Lemma 3.5.4 on page 59. For $j \in \mathbb{N}$ define

$$U_j := \left\{\boldsymbol{\xi} \in \mathbf{NAF}_w^{0,j} \,\middle|\, \xi_{-\ell}\xi_{-(\ell+1)}\dots\xi_{-(\ell+k-1)} \neq 0^k \text{ for all } \ell \text{ with } 1 \le \ell \le j-k+1\right\}.$$

The elements of U_j — more precisely the digits from index -1 to $-j$ — can be described by the regular expression

$$\left(\varepsilon + \sum_{d\in\mathcal{D}^\bullet}\sum_{\ell=0}^{w-2} 0^\ell d\right)\left(\sum_{d\in\mathcal{D}^\bullet}\sum_{\ell=w-1}^{k-1} 0^\ell d\right)^*\left(\sum_{\ell=0}^{k-1} 0^\ell\right).$$

This can be translated to the generating function

$$G(Z) = \sum_{j\in\mathbb{N}} \#U_j Z^j = \left(1 + \#\mathcal{D}^\bullet \sum_{\ell=0}^{w-2} Z^{\ell+1}\right) \frac{1}{1 - \#\mathcal{D}^\bullet \sum_{\ell=w-1}^{k-1} Z^{\ell+1}} \left(\sum_{\ell=0}^{k-1} Z^\ell\right)$$

used for counting the number of elements in U_j. Rewriting yields

$$G(Z) = \frac{1-Z^k}{1-Z}\,\frac{1 + (\#\mathcal{D}^\bullet - 1)Z - \#\mathcal{D}^\bullet Z^w}{1 - Z - \#\mathcal{D}^\bullet Z^w + \#\mathcal{D}^\bullet Z^{k+1}},$$

and we set

$$q(Z) := 1 - Z - \#\mathcal{D}^\bullet Z^w + \#\mathcal{D}^\bullet Z^{k+1}.$$

Now we define

$$\widetilde{U}_j := \{\boldsymbol{\xi} \in U_j \mid \xi_{-j} \neq 0\}$$

and consider $\widetilde{U} := \bigcup_{j\in\mathbb{N}} \widetilde{U}_j$. The w-NAFs in this set — more precisely the finite strings from index -1 to the index of the largest non-zero digit — will be recognised by the automaton $\mathcal{A}$ which reads its input from right to left, see Figure 3.7.1 on the preceding page. It is easy to see that the underlying directed graph $G_{\mathcal{A}}$ of the automaton $\mathcal{A}$ is strongly connected, therefore its adjacency matrix $M_{\mathcal{A}}$ is irreducible. Since there are cycles of length w and $w+1$ in the graph and $\gcd(w, w+1) = 1$, the adjacency matrix is primitive. Thus, using the Perron-Frobenius theorem we obtain

$$\begin{aligned}\#\widetilde{U}_j &= \#(\text{walks in } G_{\mathcal{A}} \text{ of length } j \text{ from starting state } \mathcal{S} \text{ to some other state})\\ &= \begin{pmatrix}1 & 0 & \dots & 0\end{pmatrix} M_{\mathcal{A}}^j \begin{pmatrix}1\\ \vdots\\ 1\end{pmatrix} = \widetilde{c}\left(\sigma\,|\tau|^2\right)^j \left(1+\mathcal{O}\big(s^j\big)\right)\end{aligned}$$

for a $\widetilde{c} > 0$, a $\sigma > 0$, and an s with $0 \le s < 1$. Since the number of w-NAFs of length j is $\mathcal{O}\Big(|\tau|^{2j}\Big)$, see Theorem 3.4.1 on page 48, we get $\sigma \le 1$.

We clearly have

$$U_j = \biguplus_{\ell=j-k+1}^{j} \widetilde{U}_\ell,$$

so we get

$$\#U_j = \left[Z^j\right] G(Z) = c\left(\sigma\,|\tau|^2\right)^j \left(1+\mathcal{O}\big(s^j\big)\right)$$

for some constant $c > 0$.

To rule out $\sigma = 1$, we insert the "zero" $|\tau|^{-2}$ in $q(Z)$. We obtain

$$\begin{aligned}q\Big(|\tau|^{-2}\Big) &= 1 - |\tau|^{-2} - \#\mathcal{D}^\bullet\,|\tau|^{-2w} + \#\mathcal{D}^\bullet\,|\tau|^{-2(k+1)}\\ &= 1 - |\tau|^{-2} - |\tau|^{2(w-1)}\left(|\tau|^2-1\right)|\tau|^{-2w} + |\tau|^{2(w-1)}\left(|\tau|^2-1\right)|\tau|^{-2(k+1)}\\ &= \left(|\tau|^2-1\right)|\tau|^{2(w-k-2)} > 0,\end{aligned}$$

where we used the cardinality of $\mathcal{D}^\bullet$ from Lemma 3.3.3 on page 44 and $|\tau| > 1$. Therefore we get $\sigma < 1$.

Define

$$U := \left\{\mathsf{value}(\boldsymbol{\xi}) \,\middle|\, \boldsymbol{\xi} \in \mathbf{NAF}_w^{0,\infty} \text{ with } \xi_{-\ell}\xi_{-(\ell+1)}\dots\xi_{-(\ell+k-1)} \neq 0^k \text{ for all } \ell \ge 1\right\}.$$

We want to cover U with squares. Let S be the closed paraxial square with centre 0 and width 2. Using Proposition 3.5.1 on page 51 yields

$$U \subseteq \bigcup_{z\in\mathsf{value}(U_j)} \left(z + f_U\,|\tau|^{-j}\,S\right)$$

for all $j \in \mathbb{N}$, i.e., U can be covered with $\#U_j$ boxes of size $2f_U\,|\tau|^{-j}$. Thus we get for the upper box dimension, cf. Edgar [19],

$$\overline{\dim}_B U \le \lim_{j\to\infty} \frac{\log \#U_j}{-\log(2f_U\,|\tau|^{-j})}.$$

Inserting the cardinality $\#U_j$ from above, using the logarithm to base $|\tau|$ and $0 \le s < 1$ yields

$$\overline{\dim}_B U \le \lim_{j\to\infty} \frac{\log_{|\tau|} c + j \log_{|\tau|}(\sigma\,|\tau|^2) + \log_{|\tau|}(1+\mathcal{O}(s^j))}{j+\mathcal{O}(1)} = 2 + \log_{|\tau|}\sigma.$$

Since $\sigma < 1$, we get $\overline{\dim}_B U < 2$.

Now we will show that $\partial\mathcal{F} \subseteq U$. Clearly $U \subseteq \mathcal{F}$, so the previous inclusion is equivalent to $\mathcal{F} \setminus U \subseteq \operatorname{int}(\mathcal{F})$. So let $z \in \mathcal{F} \setminus U$. Then there is a $\boldsymbol{\xi} \in \mathbf{NAF}_w^{0.\infty}$ such that $z = \mathsf{value}(\boldsymbol{\xi})$ and $\boldsymbol{\xi}$ has a block of at least k zeros somewhere on the right hand side of the τ-point. Let ℓ denote the starting index of this block, i.e.,

$$\boldsymbol{\xi} = 0.\underbrace{\xi_{-1}\dots\xi_{-(\ell-1)}}_{=:\boldsymbol{\xi}_A} 0^k \xi_{-(\ell+k)}\xi_{-(\ell+k+1)}\cdots.$$

Let $\boldsymbol{\vartheta} = \boldsymbol{\vartheta}_I.\boldsymbol{\vartheta}_A\vartheta_{-\ell}\vartheta_{-(\ell+1)}\ldots \in \mathbf{NAF}_w^{\mathsf{fin}.\infty}$ with $\mathsf{value}(\boldsymbol{\vartheta}) = z$. We have

$$z = \mathsf{value}(0.\boldsymbol{\xi}_A) + \tau^{-\ell-w} z_\xi = \mathsf{value}(\boldsymbol{\vartheta}_I.\boldsymbol{\vartheta}_A) + \tau^{-\ell-w} z_\vartheta$$

for appropriate z_ξ and z_ϑ. By Lemma 3.5.4 on page 59, all expansions of z_ξ are in $\mathbf{NAF}_w^{0.\infty}$. Thus all expansions of

$$\mathsf{value}(\boldsymbol{\vartheta}_I\boldsymbol{\vartheta}_A) + \tau^{-(w-1)} z_\vartheta - \mathsf{value}(\boldsymbol{\xi}_A) = \tau^{\ell-1} z - \mathsf{value}(\boldsymbol{\xi}_A) = \tau^{-(w-1)} z_\xi$$

start with 0.0^{w-1}, since our choice of k is k_0+w-1. As the unique NAF of $\mathsf{value}(\boldsymbol{\vartheta}_I\boldsymbol{\vartheta}_A) - \mathsf{value}(\boldsymbol{\xi}_A)$ concatenated with any NAF of $\tau^{-(w-1)} z_\vartheta$ gives rise to such an expansion, we conclude that $\mathsf{value}(\boldsymbol{\vartheta}_I\boldsymbol{\vartheta}_A) - \mathsf{value}(\boldsymbol{\xi}_A) = 0$ and therefore $\boldsymbol{\vartheta}_I = \mathbf{0}$ and $\boldsymbol{\vartheta}_A = \boldsymbol{\xi}_A$. So we conclude that all representations of z as a w-NAF have to be of the form $0.\boldsymbol{\xi}_A 0^{w-1}\boldsymbol{\eta}$ for some w-NAF $\boldsymbol{\eta}$. Thus, by using Proposition 3.7.7 on page 64, we get $z \notin \partial\mathcal{F}$ and therefore $z \in \operatorname{int}(\mathcal{F})$.

Until now we have proved

$$\overline{\dim}_B \partial\mathcal{F} \le \overline{\dim}_B U < 2.$$

Because the Hausdorff dimension of a set is at most its upper box dimension, cf. Edgar [19] again, the desired result follows. □

3.8 Cell Rounding Operations

Let $\tau \in \mathbb{C}$ be an algebraic integer, imaginary quadratic. In this section define operators working on subsets (regions) of the complex plane. These will use the lattice $\mathbb{Z}[\tau]$ and the Voronoi cells defined in Section 3.2. They will be a very useful concept to prove Theorem 3.10.1 on page 78.

Definition 3.8.1 (Cell Rounding Operations)**.** Let $B \subseteq \mathbb{C}$ and $j \in \mathbb{R}$. We define the *cell packing of B* ("floor B")

$$\lfloor B \rfloor_{\bigcirc} := \bigcup_{\substack{z\in\mathbb{Z}[\tau]\\ V_z\subseteq B}} V_z \qquad \text{and} \qquad \lfloor B \rfloor_{\bigcirc,j} := \frac{1}{\tau^j}\left\lfloor \tau^j B \right\rfloor_{\bigcirc},$$

the *cell covering of B* ("ceil B")

$$\lceil B \rceil_{\bigcirc} := \overline{\lfloor B^C \rfloor_{\bigcirc}^C} \qquad \text{and} \qquad \lceil B \rceil_{\bigcirc,j} := \frac{1}{\tau^j}\left\lceil \tau^j B \right\rceil_{\bigcirc},$$

the *fractional cells of B*

$$\{B\}_{\bigcirc} := B \setminus \lfloor B \rfloor_{\bigcirc} \qquad \text{and} \qquad \{B\}_{\bigcirc,j} := \frac{1}{\tau^j} \{\tau^j B\}_{\bigcirc},$$

the *cell covering of the boundary of B*

$$\partial(B)_{\bigcirc} := \overline{\lceil B \rceil_{\bigcirc} \setminus \lfloor B \rfloor_{\bigcirc}} \qquad \text{and} \qquad \partial(B)_{\bigcirc,j} := \frac{1}{\tau^j} \partial\big(\tau^j B\big)_{\bigcirc},$$

the *cell covering of the lattice points inside B*

$$\lfloor B \rceil_{\bigcirc} := \bigcup_{z \in B \cap \mathbb{Z}[\tau]} V_z \qquad \text{and} \qquad \lfloor B \rceil_{\bigcirc,j} := \frac{1}{\tau^j} \lfloor \tau^j B \rceil_{\bigcirc}$$

and the *number of lattice points inside B* as

$$\#(B)_{\bigcirc} := \#(B \cap \mathbb{Z}[\tau]) \qquad \text{and} \qquad \#(B)_{\bigcirc,j} := \#\big(\tau^j B\big)_{\bigcirc}$$

To get a slight feeling what those operators do, have a look at Figure 3.8.1 on the next page. There brief examples are given. For the cell covering of a set B an alternative, perhaps more intuitive description can be given by

$$\lceil B \rceil_{\bigcirc} := \bigcup_{\substack{z \in \mathbb{Z}[\tau] \\ V_z \cap B \neq \emptyset}} V_z.$$

The following proposition deals with some basic properties that will be helpful, when working with those operators.

Proposition 3.8.2 (Basic Properties of Cell Rounding Operations)**.** *Let $B \subseteq \mathbb{C}$ and $j \in \mathbb{R}$.*

(a) We have the inclusions

$$\lfloor B \rfloor_{\bigcirc,j} \subseteq B \subseteq \overline{B} \subseteq \lceil B \rceil_{\bigcirc,j} \tag{3.8.1a}$$

and

$$\lfloor B \rfloor_{\bigcirc,j} \subseteq \lfloor B \rceil_{\bigcirc,j} \subseteq \lceil B \rceil_{\bigcirc,j}. \tag{3.8.1b}$$

For $B' \subseteq \mathbb{C}$ with $B \subseteq B'$ we get $\lfloor B \rfloor_{\bigcirc,j} \subseteq \lfloor B' \rfloor_{\bigcirc,j}$, $\lfloor B \rceil_{\bigcirc,j} \subseteq \lfloor B' \rceil_{\bigcirc,j}$ and $\lceil B \rceil_{\bigcirc,j} \subseteq \lceil B' \rceil_{\bigcirc,j}$, i.e., monotonicity with respect to inclusion

(b) The inclusion

$$\{B\}_{\bigcirc,j} \subseteq \partial(B)_{\bigcirc,j} \tag{3.8.2}$$

holds.

(c) $\partial B \subseteq \partial(B)_{\bigcirc,j}$ and for each cell V' in $\partial(B)_{\bigcirc,j}$ we have $V' \cap \partial B \neq \emptyset$, so $\partial(B)_{\bigcirc,j}$ is the smallest union of cells that contains ∂B.

(d) For $B' \subseteq \mathbb{C}$ with B' disjoint from B, we get

$$\#(B \cup B')_{\bigcirc,j} = \#(B)_{\bigcirc,j} + \#(B')_{\bigcirc,j}, \tag{3.8.3}$$

and therefore the number of lattice points operation is monotonic with respect to inclusion, i.e., for $B'' \subseteq \mathbb{C}$ with $B'' \subseteq B$ we have $\#(B'')_{\bigcirc,j} \leq \#(B)_{\bigcirc,j}$. Further we get

$$\#(B)_{\bigcirc,j} = \#\Big(\lfloor B \rceil_{\bigcirc,j}\Big)_{\bigcirc,j} = |\tau|^{2j} \, \frac{\lambda\Big(\lfloor B \rceil_{\bigcirc,j}\Big)}{\lambda(V)} \tag{3.8.4}$$

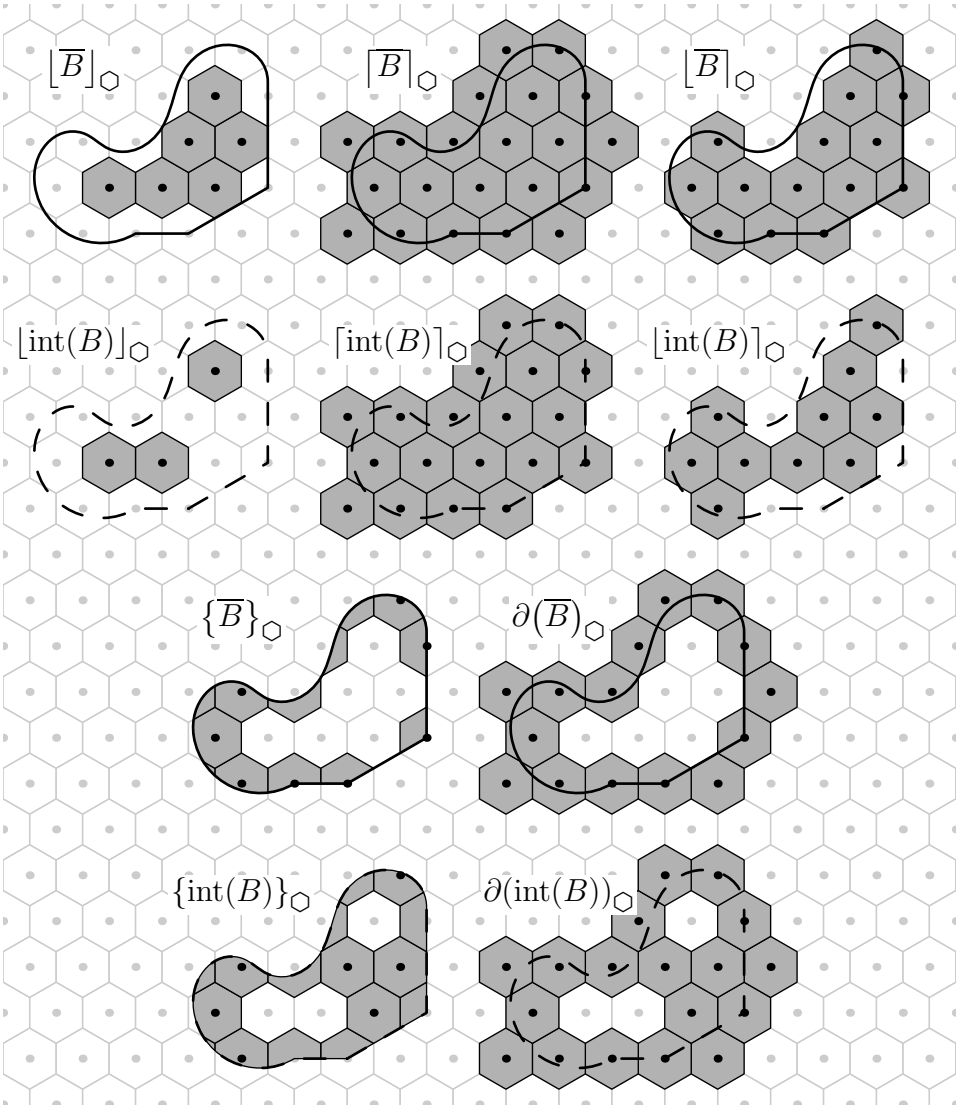


Figure 3.8.1: Examples of the cell rounding operators of Definition 3.8.1 on page 67. As lattice $\mathbb{Z}[\tau]$ with $\tau = \frac{3}{2} + \frac{1}{2}\sqrt{-3}$ was used here.

Proof. (a) $\lfloor B \rfloor_{\bigcirc,j} \subseteq B$ follows directly from the definition. Since $\lfloor B^C \rfloor_{\bigcirc,j} \subseteq B^C$, we get

$$\lceil B \rceil_{\bigcirc,j} = \overline{\lfloor B^C \rfloor_{\bigcirc,j}^C} \supseteq \overline{(B^C)^C} = \overline{B}.$$

The inclusion $\lfloor B \rfloor_{\bigcirc,j} \subseteq \lfloor B \rceil_{\bigcirc,j}$ follows directly from the definitions and $\lfloor B \rceil_{\bigcirc,j} \subseteq \lceil B \rceil_{\bigcirc,j}$ again by considering the complement, because $\overline{\lfloor B^C \rceil_{\bigcirc,j}^C} = \lfloor B \rceil_{\bigcirc,j}$. Similarly, the monotonicity can be shown.

(b) We have

$$\{B\}_{\bigcirc,j} = B \setminus \lfloor B \rfloor_{\bigcirc,j} \subseteq \lceil B \rceil_{\bigcirc,j} \setminus \lfloor B \rfloor_{\bigcirc,j} = \partial(B)_{\bigcirc,j}\,.$$

(c) We assume $j = 0$. Using (a) yields $\partial B \subseteq \overline{B} \subseteq \lceil B \rceil_{\bigcirc}$. Let $x \in \partial B$. If $x \notin B$, then $\lfloor B \rfloor_{\bigcirc} \subseteq B$ implies that $x \notin \lfloor B \rfloor_{\bigcirc}$. So we get

$$x \in \lceil B \rceil_{\bigcirc} \setminus \lfloor B \rfloor_{\bigcirc} \subseteq \overline{\lceil B \rceil_{\bigcirc} \setminus \lfloor B \rfloor_{\bigcirc}} = \partial(B)_{\bigcirc}\,.$$

Now suppose $x \in B$. Consider all Voronoi cells V_i, $i \in I$, for a suitable finite index set I, such that $x \in V_i$. We get $x \in \operatorname{int}\bigl(\bigcup_{i\in I} V_i\bigr)$. If all of the V_i are a subset of $\lfloor B \rfloor_{\bigcirc}$, then $x \in \operatorname{int}(B)$, which is a contradiction to $x \in \partial B$. So there is at least one cell V' that is not a subset of $\lfloor B \rfloor_{\bigcirc}$. Since

$$\lceil B \rceil_{\bigcirc}^C = \overline{\lfloor B^C \rfloor_{\bigcirc}^C}^C = \operatorname{int}\left(\bigcup_{\substack{z\in\mathbb{Z}[\tau] \\ V_z \subseteq B^C}} V_z \right)$$

and $x \notin B^C$, V' is not in this union of cells. So V' is in the complement, i.e., $V' \subseteq \lceil B \rceil_{\bigcirc}$. And therefore the statement follows.

Now we want to show that there is a subset of the boundary in each V-cell V' of $\partial(B)_{\bigcirc}$. Assume $V' \cap \partial B = \emptyset$. If $V' \cap B = \emptyset$, then $V' \subseteq B^C$, so V' is not a subset of $\lceil B \rceil_{\bigcirc}$, contradiction. If $V' \cap B \neq \emptyset$, then $V' \subseteq B$, since V' does not contain any boundary. But then, $V' \subseteq \lfloor B \rfloor_{\bigcirc}$, again a contradiction.

(d) Since the operator just counts the number of lattice points, the first statement follows.

In the other statement, the first equality follows, because $z \in B \cap \mathbb{Z}[\tau] \iff V_z \subseteq \lfloor B \rceil_{\bigcirc}$ holds. Since $\lfloor B \rceil_{\bigcirc,j}$ consists of cells each with area $\lambda\bigl(\tau^{-j}V\bigr)$, the second equality is just, after multiplying by $\lambda\bigl(\tau^{-j}V\bigr)$, the equality of the areas. □

We will need some more properties concerning cardinality. We want to know the number of points inside a region after using one of the operators. Especially we are interested in the asymptotic behaviour, i.e., if our region becomes scaled very large. The following proposition provides information about that.

Proposition 3.8.3. *Let $U \subseteq \mathbb{C}$ bounded, measurable, and such that*

$$\#\bigl(\partial(NU)_{\bigcirc}\bigr)_{\bigcirc} = \mathcal{O}\Bigl(|N|^{\delta}\Bigr) \tag{3.8.5}$$

for $N \in \mathbb{C}$.

(a) We get

$$\#\big(\lfloor NU\rfloor_{\bigcirc}\big)_{\bigcirc} = |N|^2\,\frac{\lambda(U)}{\lambda(V)} + \mathcal{O}\Big(|N|^{\delta}\Big)\,,$$

$$\#\big(\lceil NU\rceil_{\bigcirc}\big)_{\bigcirc} = |N|^2\,\frac{\lambda(U)}{\lambda(V)} + \mathcal{O}\Big(|N|^{\delta}\Big)$$

and

$$\#(NU)_{\bigcirc} = \#\big(\lfloor NU\rceil_{\bigcirc}\big)_{\bigcirc} = |N|^2\,\frac{\lambda(U)}{\lambda(V)} + \mathcal{O}\Big(|N|^{\delta}\Big)\,.$$

(b) We get

$$\#((N+1)U \setminus NU)_{\bigcirc} = \mathcal{O}\Big(|N|^{\delta}\Big)\,.$$

Proof. (a) Considering the areas yields

$$\#\big(\lfloor NU\rfloor_{\bigcirc}\big)_{\bigcirc}\,\lambda(V) \le \lambda(NU) = |N|^2\,\lambda(U) \le \#\big(\lceil NU\rceil_{\bigcirc}\big)_{\bigcirc}\,\lambda(V)\,,$$

since $\lfloor NU\rfloor_{\bigcirc} \subseteq NU \subseteq \lceil NU\rceil_{\bigcirc}$, see Proposition 3.8.2 on page 68. If we use $\lceil NU\rceil_{\bigcirc} = \lfloor NU\rfloor_{\bigcirc} \cup \partial(NU)_{\bigcirc}$, we obtain

$$0 \le |N|^2\,\frac{\lambda(U)}{\lambda(V)} - \#\big(\lfloor NU\rfloor_{\bigcirc}\big)_{\bigcirc} \le \#\big(\partial(NU)_{\bigcirc}\big)_{\bigcirc}$$

Because $\#\big(\partial(NU)_{\bigcirc}\big)_{\bigcirc} = \mathcal{O}\Big(|N|^{\delta}\Big)$ we get

$$\left||N|^2\,\frac{\lambda(U)}{\lambda(V)} - \#\big(\lfloor NU\rfloor_{\bigcirc}\big)_{\bigcirc}\right| = \mathcal{O}\Big(|N|^{\delta}\Big)\,,$$

and thus the result follows.

Combining the previous result and Proposition 3.8.2 on page 68 proves the other two statements.

(b) Let $d \in \mathbb{R}$ such that $U \subseteq \mathcal{B}(0,d)$. Let $y \in (N+1)U \setminus NU$. Obviously, this is equivalent to $y/(N+1) \in U$ and $y/N \notin U$, so there is a $z \in \partial U$ on the line from y/N to $y/(N+1)$. We get

$$\left|z - \frac{y}{N}\right| \le |y| \cdot \left|\frac{1}{N} - \frac{1}{N+1}\right| = \frac{|y|}{|N+1|} \cdot \frac{1}{|N|} \le \frac{d}{|N|},$$

and therefore

$$(N+1)U \setminus NU \subseteq \bigcup_{z\in\partial(NU)} \mathcal{B}(z,d)\,.$$

Since the boundary $\partial(NU)$ can be covered by $\mathcal{O}\Big(|N|^{\delta}\Big)$ cells, cf. (c) of Proposition 3.8.2 on page 68 and the discs in $\bigcup_{z\in\partial(NU)} \mathcal{B}(z,d)$ have a fixed size, the result follows. □

If the geometry of U is simple, e.g. U is a disc or U is a polygon, then we can check the covering condition (3.8.5) of Proposition 3.8.3 on the preceding page by means of the following proposition.

Proposition 3.8.4. *Let $U \subseteq \mathbb{C}$ such that the boundary of U consists of finitely many rectifiable curves. Then we get*

$$\#\big(\partial(NU)_{\bigcirc}\big)_{\bigcirc} = \mathcal{O}(|N|)$$

for $N \in \mathbb{C}$.

Proof. Without loss of generality, we may assume that the boundary of U is a rectifiable curve $\gamma\colon [0,L] \longrightarrow \mathbb{C}$, which is parametrised by arc length. For any $t \in [1/(2\,|N|), L - 1/(2\,|N|)]$, we have

$$\gamma\left(\left[t - \frac{1}{2\,|N|}, t + \frac{1}{2\,|N|}\right]\right) \subseteq \mathcal{B}\left(\gamma(t), \frac{1}{2\,|N|}\right),$$

as the straight line from $\gamma(t)$ to $\gamma(t')$ is never longer than the arc-length of $\gamma([t,t'])$. Thus ∂U can be covered by $\mathcal{O}(L\,|N|)$ discs of radius $1/(2\,|N|)$ and consequently, $\partial(NU)$ can be covered by $\mathcal{O}(L\,|N|)$ discs of radius $\frac{1}{2}$. As $\mathbb{Z}[\tau]$ is a lattice, each disc with radius $\frac{1}{2}$ is contained in at most 4 Voronoi-cells, cf. Proposition 3.2.5 on page 41. Therefore, $\mathcal{O}(N)$ cells suffice to cover $\partial(NU)$. □

3.9 The Characteristic Sets W_η

Let $\tau \in \mathbb{C}$ be an algebraic integer, imaginary quadratic. Suppose that $|\tau| > 1$. Let $w \in \mathbb{N}$ with $w \geq 2$. Further let $\mathcal{D}$ be a minimal norm representatives digit set modulo τ^w as in Definition 3.3.5 on page 45. We denote the norm function by $\mathcal{N}\colon \mathbb{Z}[\tau] \longrightarrow \mathbb{Z}$, and we simply have $\mathcal{N}(\tau) = |\tau|^2$. Again for simplicity we set $\mathcal{D}^\bullet := \mathcal{D} \setminus \{0\}$.

In this section we define characteristic sets for a digit at a specified position in the w-NAF expansion and prove some basic properties of them. Those will be used in the proof of Theorem 3.10.1 on page 78.

Definition 3.9.1 (Characteristic Sets)**.** Let $\eta \in \mathcal{D}^\bullet$. For $j \in \mathbb{N}_0$ define

$$\mathcal{W}_{\eta,j} := \left\{\mathsf{value}(\boldsymbol{\xi}) \;\middle|\; \boldsymbol{\xi} \in \mathbf{NAF}_w^{0.j+w} \text{ with } \xi_{-w} = \eta\right\}.$$

We call $\lfloor \mathcal{W}_{\eta,j} \rceil_{\bigcirc, j+w}$ the *jth approximation of the characteristic set for* η, and we define

$$W_{\eta,j} := \left\{ \lfloor \mathcal{W}_{\eta,j} \rceil_{\bigcirc, j+w} \right\}_{\mathbb{Z}[\tau]}.$$

Further we define the *characteristic set for* η

$$\mathcal{W}_\eta := \left\{\mathsf{value}(\boldsymbol{\xi}) \;\middle|\; \boldsymbol{\xi} \in \mathbf{NAF}_w^{0.\infty} \text{ with } \xi_{-w} = \eta\right\}.$$

and

$$W_\eta := \{\mathcal{W}_\eta\}_{\mathbb{Z}[\tau]}.$$

For $j \in \mathbb{N}_0$ we set

$$\beta_{\eta,j} := \lambda\Big(\lfloor \mathcal{W}_{\eta,j} \rceil_{\bigcirc, j+w}\Big) - \lambda(\mathcal{W}_\eta).$$

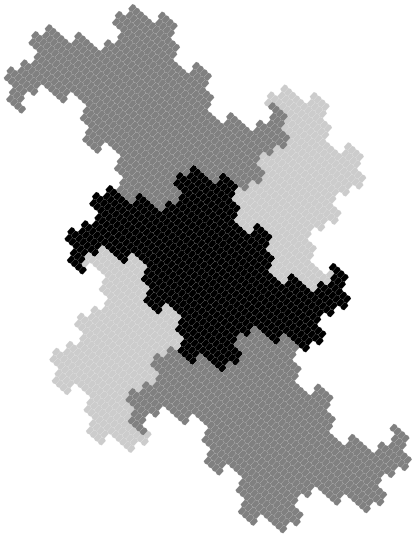
Note that sometimes the set W_η will also be called *characteristic set for* η, and analogously for the set $W_{\eta,j}$. In Figure 3.9.1 on the next page some of these characteristic sets — more precisely some approximations of the characteristic sets — are shown. The following proposition will deal with some properties of those defined sets,

Proposition 3.9.2 (Properties of the Characteristic Sets)**.** *Let $\eta \in \mathcal{D}^\bullet$.*

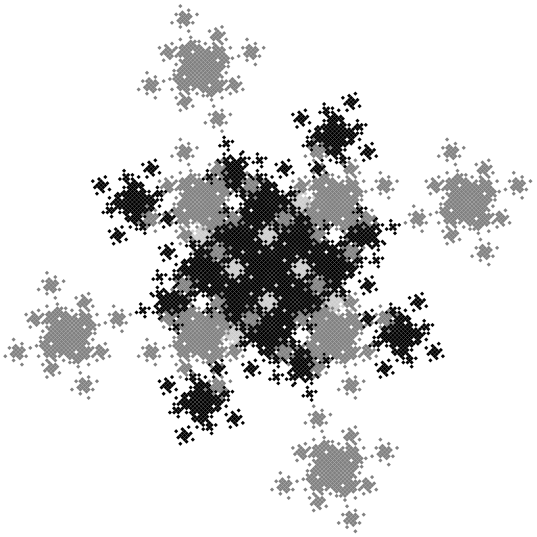
(a) We have

$$\mathcal{W}_\eta = \eta\tau^{-w} + \tau^{-2w+1}\mathcal{F}.$$

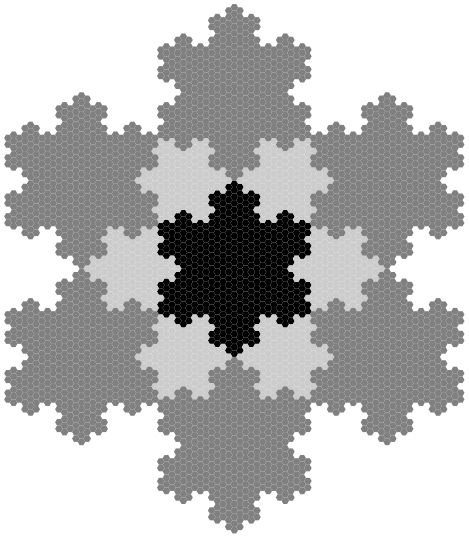
(b) The set $\mathcal{W}_\eta$ is compact.



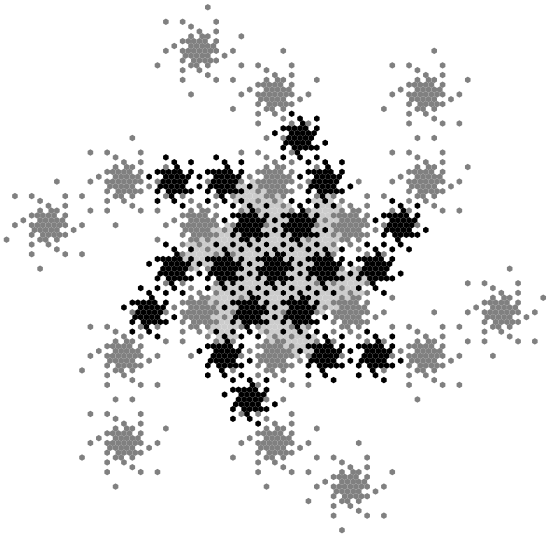
(a) $\mathcal{W}_{\eta,j}$ for $\tau = \frac{1}{2} + \frac{1}{2}\sqrt{-7}$, $w = 2$ and $j = 11$



(b) $\mathcal{W}_{\eta,j}$ for $\tau = 1 + \sqrt{-1}$, $w = 4$ and $j = 11$



(c) $\mathcal{W}_{\eta,j}$ for $\tau = \frac{3}{2} + \frac{1}{2}\sqrt{-3}$, $w = 2$ and $j = 7$



(d) $\mathcal{W}_{\eta,j}$ for $\tau = \frac{3}{2} + \frac{1}{2}\sqrt{-3}$, $w = 3$ and $j = 6$

Figure 3.9.1: Characteristic sets $\mathcal{W}_\eta$. Each figure can either be seen as approximation $\mathcal{W}_{\eta,j}$ for $\mathcal{W}_\eta$, or as values of w-NAFs of length j, where a scales Voronoi cell is drawn for each point. Different colours correspond to the digits 1 and w from the left in the w-NAF. They are "marked" whether they are zero or non-zero.

(c) We get

$$\mathcal{W}_\eta = \overline{\bigcup_{j\in\mathbb{N}_0} \mathcal{W}_{\eta,j}} = \overline{\lim_{j\to\infty} \mathcal{W}_{\eta,j}}.$$

(d) The set $\lfloor \mathcal{W}_{\eta,j} \rceil_{\bigcirc,j+w}$ is indeed an approximation of $\mathcal{W}_\eta$, i.e., we have

$$\mathcal{W}_\eta = \overline{\liminf_{j\in\mathbb{N}_0} \lfloor \mathcal{W}_{\eta,j} \rceil_{\bigcirc,j+w}} = \overline{\limsup_{j\in\mathbb{N}_0} \lfloor \mathcal{W}_{\eta,j} \rceil_{\bigcirc,j+w}}.$$

(e) We have $\operatorname{int} \mathcal{W}_\eta \subseteq \liminf_{j\in\mathbb{N}_0} \lfloor \mathcal{W}_{\eta,j} \rceil_{\bigcirc,j+w}$.

(f) We get $\mathcal{W}_\eta - \eta\tau^{-w} \subseteq V$, and for $j \in \mathbb{N}_0$ we obtain $\lfloor \mathcal{W}_{\eta,j} \rceil_{\bigcirc,j+w} - \eta\tau^{-w} \subseteq V$.

(g) For the Lebesgue measure of the characteristic set we obtain $\lambda(\mathcal{W}_\eta) = \lambda(W_\eta)$ and for its approximation $\lambda\Big(\lfloor \mathcal{W}_{\eta,j} \rceil_{\bigcirc,j+w}\Big) = \lambda(W_{\eta,j})$.

(h) Let $j \in \mathbb{N}_0$. If $j < w-1$, then the area of $\lfloor \mathcal{W}_{\eta,j} \rceil_{\bigcirc,j+w}$ is

$$\lambda\Big(\lfloor \mathcal{W}_{\eta,j} \rceil_{\bigcirc,j+w}\Big) = |\tau|^{-2(j+w)}\, \lambda(V)\,.$$

If $j \geq w-1$, then the area of $\lfloor \mathcal{W}_{\eta,j} \rceil_{\bigcirc,j+w}$ is

$$\lambda\Big(\lfloor \mathcal{W}_{\eta,j} \rceil_{\bigcirc,j+w}\Big) = \lambda(V)\, e_w + \mathcal{O}\big(\rho^j\big)$$

with e_w and ρ from Theorem 3.4.1 on page 48.

(i) The area of W_η is

$$\lambda(W_\eta) = \lambda(V)\, e_w,$$

again with e_w from Theorem 3.4.1 on page 48.

(j) Let $j \in \mathbb{N}_0$. We get

$$\beta_{\eta,j} = \int_{x\in V} \big(\mathbb{1}_{W_{\eta,j}} - \mathbb{1}_{W_\eta}\big)(x)\ \mathrm{d}x.$$

If $j < w-1$, then its value is

$$\beta_{\eta,j} = \Big(|\tau|^{-2(j+w)} - e_w\Big)\, \lambda(V)\,.$$

If $j \geq w-1$, then we get

$$\beta_{\eta,j} = \mathcal{O}\big(\rho^j\big)\,.$$

Again e_w and ρ can be found in Theorem 3.4.1 on page 48.

Proof. (a) Is clear, since we have the digit η at index $-w$ and an arbitrary w-NAF starting with index $-2w$. Note that the elements in $\mathbf{NAF}_w^{0.\infty}$ start with index -1.

(b) Follows directly from (a), because $\mathcal{F}$ is compact according to Proposition 3.7.2 on page 62.

(c) Clearly we have $\mathcal{W}_{\eta,j} \subseteq \mathcal{W}_\eta$. Thus $\bigcup_{j\in\mathbb{N}_0} \mathcal{W}_{\eta,j} \subseteq \mathcal{W}_\eta$, and because $\mathcal{W}_\eta$ is closed, the inclusion $\overline{\bigcup_{j\in\mathbb{N}_0} \mathcal{W}_{\eta,j}} \subseteq \mathcal{W}_\eta$ follows. Now let $z \in \mathcal{W}_\eta$, and let $\boldsymbol{\xi} \in \mathbf{NAF}_w^{0,\infty}$, such that $\mathsf{value}(\boldsymbol{\xi}) = z$. Then there is a sequence of w-NAFs $(\boldsymbol{\xi}_\ell)_{\ell\in\mathbb{N}_0}$ with finite right-lengths that converges to $\boldsymbol{\xi}$ and clearly

$$\mathsf{value}(\boldsymbol{\xi}_\ell) \in \bigcup_{k\in\mathbb{N}_0} \mathcal{W}_{\eta,k}.$$

Since evaluating the value is a continuous function, see Proposition 3.3.8 on page 46, we get

$$z = \mathsf{value}(\boldsymbol{\xi}) = \mathsf{value}\left(\lim_{\ell\to\infty} \boldsymbol{\xi}_\ell\right) = \lim_{\ell\to\infty} \mathsf{value}(\boldsymbol{\xi}_\ell) \in \overline{\bigcup_{k\in\mathbb{N}_0} \mathcal{W}_{\eta,j}}.$$

The equality $\bigcup_{j\in\mathbb{N}_0} \mathcal{W}_{\eta,j} = \lim_{j\to\infty} \mathcal{W}_{\eta,j}$ is obvious, since $\mathcal{W}_{\eta,j}$ is monotonic increasing.

(d) First we show that we have

$$\limsup_{j\to\infty} \lfloor \mathcal{W}_{\eta,j} \rceil_{\bigcirc,j+w} \subseteq \mathcal{W}_\eta.$$

Let

$$z \in \limsup_{j\to\infty} \lfloor \mathcal{W}_{\eta,j} \rceil_{\bigcirc,j+w} = \bigcap_{j\in\mathbb{N}_0} \bigcup_{k\geq j} \lfloor \mathcal{W}_{\eta,k} \rceil_{\bigcirc,k+w}.$$

Then there is a $j_0 \geq 0$ such that $z \in \lfloor \mathcal{W}_{\eta,j_0} \rceil_{\bigcirc,j_0+w}$. Further, for $j_{\ell-1}$ there is a $j_\ell \geq j_{\ell-1}$, such that $z \in \lfloor \mathcal{W}_{\eta,j_\ell} \rceil_{\bigcirc,j_\ell+w}$. For each $\ell \in \mathbb{N}_0$ there is a $z_\ell \in \mathcal{W}_{\eta,j_\ell} \subseteq \mathcal{W}_\eta$ with

$$|z - z_\ell| \leq c_V \, |\tau| \, |\tau|^{-j_\ell - w},$$

since $\lfloor \mathcal{W}_{\eta,j_\ell} \rceil_{\bigcirc,j_\ell+w}$ consists of cells $|\tau|^{-j_\ell-w} V$ with centres out of $\mathcal{W}_{\eta,j_\ell}$. Refer to Proposition 3.2.5 on page 41 for the constant $c_V \, |\tau|$. Thus we get $z = \lim_{\ell\to\infty} z_\ell \in \mathcal{W}_\eta$, since $|\tau|^{-j_\ell-w}$ tends to 0 for large ℓ and $\mathcal{W}_\eta$ is closed.

Using the closeness property of $\mathcal{W}_\eta$ again yields

$$\overline{\limsup_{j\to\infty} \lfloor \mathcal{W}_{\eta,j} \rceil_{\bigcirc,j+w}} \subseteq \mathcal{W}_\eta.$$

Now we are ready to show the stated equalities. We obtain

$$\mathcal{W}_\eta = \overline{\lim_{j\to\infty} \mathcal{W}_{\eta,j}} = \overline{\liminf_{j\to\infty} \mathcal{W}_{\eta,j}} \subseteq \overline{\liminf_{j\to\infty} \lfloor \mathcal{W}_{\eta,j} \rceil_{\bigcirc,j+w}} \subseteq \overline{\limsup_{j\to\infty} \lfloor \mathcal{W}_{\eta,j} \rceil_{\bigcirc,j+w}} \subseteq \mathcal{W}_\eta,$$

so equality holds everywhere.

(e) Let $z \in \operatorname{int} \mathcal{W}_\eta$. Then there exists an $\varepsilon > 0$ such that $\mathcal{B}(z,\varepsilon) \subseteq \operatorname{int} \mathcal{W}_\eta$. For each $k \in \mathbb{N}_0$ there is a $y \in \tau^{-k-w}\mathbb{Z}[\tau]$ with the property that z is in the corresponding Voronoi cell, i.e., $z \in y + \tau^{-k-w} V$. For this y, there is also an $\boldsymbol{\xi} \in \mathbf{NAF}_w^{\mathsf{fin},k+w}$ such that $y = \mathsf{value}(\boldsymbol{\xi})$.

Clearly, if k is large enough, say $k \geq j$, we obtain $y + \tau^{-k-w} V \subseteq \mathcal{B}(z,\varepsilon)$. From Proposition 3.7.7 on page 64 (combined with (a)) we know that all w-NAFs corresponding to the values in $\operatorname{int} \mathcal{W}_\eta$ must have η at digit $-w$ and integer part $\mathbf{0}$. But this means that $\mathsf{value}(\boldsymbol{\xi}) \in \mathcal{W}_{\eta,k}$ and therefore $z \in \lfloor \mathcal{W}_{\eta,k} \rceil_{\bigcirc,k+w}$. So we conclude

$$z \in \bigcup_{j\in\mathbb{N}_0} \bigcap_{k\geq j} \lfloor \mathcal{W}_{\eta,k} \rceil_{\bigcirc,k+w} = \liminf_{j\to\infty} \lfloor \mathcal{W}_{\eta,j} \rceil_{\bigcirc,j+w}.$$

(f) Each w-NAF $\boldsymbol{\xi} \in \mathbf{NAF}_w^{0.\infty}$ corresponding to a value in $\mathcal{W}_\eta - \eta\tau^{-w}$ starts with $2w-1$ zeros from the left. Therefore

$$\mathsf{value}(\boldsymbol{\xi}) = \tau^{1-2w}\,\mathsf{value}(\boldsymbol{\vartheta})$$

for an appropriate w-NAF $\boldsymbol{\vartheta} \in \mathbf{NAF}_w^{0.\infty}$. Thus, using $\mathsf{value}(\boldsymbol{\vartheta}) \in \tau^{2w-1}V$ from Proposition 3.5.1 on page 51, the desired inclusion follows.

The set $\lfloor \mathcal{W}_{\eta,j} \rceil_{\bigcirc,j+w} - \eta\tau^{-w} \subseteq V$ consists of cells of type $\tau^{-j-w}V$, where their centres are the fractional value of an element $\boldsymbol{\xi} \in \mathbf{NAF}_w^{0.j+w}$. Again the first $2w-1$ digits are zero, so

$$\mathsf{value}(\boldsymbol{\xi}) = \tau^{1-2w}\,\mathsf{value}(\boldsymbol{\vartheta})$$

for an appropriate w-NAF $\boldsymbol{\vartheta} \in \mathbf{NAF}_w^{0.j-w+1}$. Suppose $j \geq w-1$. Using $\mathsf{value}(\boldsymbol{\vartheta}) + \tau^{-(j-w+1)}V \subseteq \tau^{2w-1}V$ again from Proposition 3.5.1 on page 51, the statement follows. If $j < w-1$, then $\mathsf{value}(\boldsymbol{\vartheta}) = 0$ and it remains to show that $\tau^{-j-w}V \subseteq V$. But this is clearly true, since $\tau^{-1}V \subseteq V$ according to Proposition 3.2.5 on page 41.

(g) As a shifted version of the sets $\mathcal{W}_\eta$ and $\lfloor \mathcal{W}_{\eta,j} \rceil_{\bigcirc,j+w}$ is contained in V by (f), so the equality of the Lebesgue measures follows directly.

(h) The set $\lfloor \mathcal{W}_{\eta,j} \rceil_{\bigcirc,j+w}$ consists of of cells of type $\tau^{-j-w}V$, where their centres are the value of an element $\boldsymbol{\xi} \in \mathbf{NAF}_w^{0.j+w}$. The intersection of two different cells is contained in the boundary of the cells, so a set of Lebesgue measure zero.

Suppose $j \geq w-1$. Since the digit $\xi_{-w} = \eta$ is fixed, the first $2w-1$ digits from the left are fixed, too. The remaining word $\xi_{-2w} \ldots \xi_{-(j+w)}$ can be an arbitrary w-NAF of length $j-w+1$, so there are $C_{j-w+1,w}$ choices, see Theorem 3.4.1 on page 48.

Thus we obtain

$$\lambda\Big(\lfloor \mathcal{W}_{\eta,j} \rceil_{\bigcirc,j+w}\Big) = C_{j-w+1,w}\,\lambda\Big(\tau^{-(w+j)}V\Big) = C_{j-w+1,w}\,|\tau|^{-2(w+j)}\,\lambda(V)\,.$$

Inserting the results of Theorem 3.4.1 on page 48 yields

$$\begin{aligned}\lambda\Big(\lfloor \mathcal{W}_{\eta,j} \rceil_{\bigcirc,j+w}\Big) &= \left(\frac{|\tau|^{2(j-w+1+w)}}{\Big(|\tau|^2-1\Big)\,w+1} + \mathcal{O}\bigg(\Big(\rho\,|\tau|^2\Big)^{j-w+1}\bigg)\right)|\tau|^{-2(w+j)}\,\lambda(V)\\ &= \lambda(V)\,\underbrace{\frac{1}{|\tau|^{2(w-1)}\,\Big(\Big(|\tau|^2-1\Big)\,w+1\Big)}}_{=e_w} + \mathcal{O}\big(\rho^j\big)\,.\end{aligned}$$

If $j < w-1$, then $\lfloor \mathcal{W}_{\eta,j} \rceil_{\bigcirc,j+w}$ consists of only one cell of size $\tau^{-j-w}V$, so the stated result follows directly.

(i) Using (d), (e) and the continuity of the Lebesgue measure yields

$$\begin{aligned}\lambda\bigg(\liminf_{j\in\mathbb{N}_0} \lfloor \mathcal{W}_{\eta,j} \rceil_{\bigcirc,j+w}\bigg) &\leq \liminf_{j\in\mathbb{N}_0} \lambda\Big(\lfloor \mathcal{W}_{\eta,j} \rceil_{\bigcirc,j+w}\Big) \leq \limsup_{j\in\mathbb{N}_0} \lambda\Big(\lfloor \mathcal{W}_{\eta,j} \rceil_{\bigcirc,j+w}\Big)\\ &\leq \lambda\bigg(\limsup_{j\in\mathbb{N}_0} \lfloor \mathcal{W}_{\eta,j} \rceil_{\bigcirc,j+w}\bigg) \leq \lambda\Big(\overline{\limsup_{j\in\mathbb{N}_0} \lfloor \mathcal{W}_{\eta,j} \rceil_{\bigcirc,j+w}}\Big)\\ &= \lambda(\mathcal{W}_\eta) \leq \lambda(\operatorname{int}\mathcal{W}_\eta) + \lambda(\partial\mathcal{W}_\eta)\\ &\leq \lambda\bigg(\liminf_{j\in\mathbb{N}_0} \lfloor \mathcal{W}_{\eta,j} \rceil_{\bigcirc,j+w}\bigg) + \lambda(\partial\mathcal{W}_\eta)\,.\end{aligned}$$

Since $\lambda(\partial\mathcal{W}_\eta) = 0$, combine (a) and Proposition 3.7.8 on page 65 to see this, we have equality everywhere, so

$$\lambda(\mathcal{W}_\eta) = \lim_{j\in\mathbb{N}_0} \lambda\Big(\lfloor \mathcal{W}_{\eta,j} \rceil_{\bigcirc,j+w}\Big)\,.$$

Thus the desired result follows from (h), because $\rho < 1$.

(j) Using (f) and (g) yields the first statement. The other result follows directly by using (h) and (i). □

Using the results of the previous proposition, we can finally determine the Lebesgue measure of the fundamental domain $\mathcal{F}$ defined in Section 3.7.

Remark 3.9.3 (Lebesgue Measure of the Fundamental Domain). We get

$$\lambda(\mathcal{F}) = |\tau|^{2w-1}\, e_w\, \lambda(V) = \frac{|\tau|\,|\mathrm{Im}(\tau)|}{(|\tau|^2-1)w+1},$$

using (a) and (i) from Proposition 3.9.2 on page 72, e_w from Theorem 3.4.1 on page 48, and $\lambda(V) = |\mathrm{Im}(\tau)|$ from Proposition 3.2.5 on page 41.

The next lemma makes the connection between the w-NAFs of elements of the lattice $\mathbb{Z}[\tau]$ and the characteristic sets $W_{\eta,j}$.

Lemma 3.9.4. *Let $\eta \in \mathcal{D}^\bullet$, $j \geq 0$. Let $n \in \mathbb{Z}[\tau]$ and let $\mathbf{n} \in \mathbf{NAF}_w^{\mathsf{fin}}$ be its w-NAF. Then the following statements are equivalent:*

(1) The jth digit of $\mathbf{n}$ equals η.

(2) The condition $\left\{\tau^{-(j+w)}n\right\}_{\mathbb{Z}[\tau]} \in W_{\eta,j}$ holds.

(3) The inclusion $\left\{\tau^{-(j+w)}V_n\right\}_{\mathbb{Z}[\tau]} \subseteq W_{\eta,j}$ holds.

Proof. Define $\mathbf{m}$ by

$$m_k := \begin{cases} n_k & \text{if } k < j+w, \\ 0 & \text{if } k \geq j+w \end{cases}$$

and $m := \mathsf{value}(\mathbf{m})$. Then, by definition, $m \equiv n \pmod{\tau^{j+w}}$,

$$\left\{\tau^{-(j+w)}n\right\}_{\mathbb{Z}[\tau]} = \left\{\tau^{-(j+w)}m\right\}_{\mathbb{Z}[\tau]}$$

and $\tau^{-(j+w)}m \in \mathcal{F}$. As the jth digit of $\mathbf{n}$ only depends on the $j+w$ least significant digits of $\mathbf{n}$, it is sufficient to show the equivalence of the assertions when $\mathbf{n}$ and n are replaced by $\mathbf{m}$ and m, respectively.

By definition, $m_j = \eta$ is equivalent to $\tau^{-(j+w)}m \in \mathcal{W}_{\eta,j}$.

(1) $\Longrightarrow$ (3). Assume now that $\tau^{-(j+w)}m \in \mathcal{W}_{\eta,j}$. Then $m \in \tau^{j+w}\mathcal{W}_{\eta,j}$ and

$$\tau^{-(j+w)}V_m \subseteq \lfloor \mathcal{W}_{\eta,j} \rceil_{\bigcirc,j+w}\,.$$

This implies $\left\{\tau^{-(j+w)}V_m\right\}_{\mathbb{Z}[\tau]} \subseteq W_j$.

(3) $\Longrightarrow$ (2). This implication holds trivially.

(2) $\Longrightarrow$ (1). So now assume that $\left\{\tau^{-(j+w)}m\right\}_{\mathbb{Z}[\tau]} \in W_{\eta,j}$. Thus there is an m' such that

$$\tau^{-(j+w)}m' \in \lfloor \mathcal{W}_{\eta,j} \rceil_{\bigcirc,j+w}$$

and

$$\tau^{-(j+w)}m - \tau^{-(j+w)}m' \in \mathbb{Z}[\tau].$$

This immediately implies $m' \in \mathbb{Z}[\tau]$ and $m \equiv m' \pmod{\tau^{j+w}}$. We also conclude that $m' \in \tau^{j+w} \lfloor \mathcal{W}_{\eta,j} \rceil_{\bigcirc,j+w}$. As $m' \in \mathbb{Z}[\tau]$, this is equivalent to $m' \in \tau^{j+w}\mathcal{W}_{\eta,j}$ and therefore $\tau^{-(j+w)}m' \in \mathcal{W}_{\eta,j}$. By definition of $\mathcal{W}_{\eta,j}$, there is a $0.\mathbf{m}' \in \mathbf{NAF}_w^{0.j+w}$ such that $\tau^{-(j+w)}m' = \mathsf{value}(0.\mathbf{m}')$, i.e., $m' = \mathsf{value}(\mathbf{m}')$, and $m'_j = \eta$. From $m' \equiv m \pmod{\tau^{j+w}}$ we conclude that $m_j = \eta$, too. (In fact, one can now easily show that we have $\mathbf{m}' = \mathbf{m}$, but this is not really needed.) □

3.10 Counting the Occurrences of a non-zero Digit in a Region

Let $\tau \in \mathbb{C}$ be an algebraic integer, imaginary quadratic. Suppose that $|\tau| > 1$. Let $w \in \mathbb{N}$ with $w \geq 2$. Further let $\mathcal{D}$ be a minimal norm representatives digit set modulo τ^w as in Definition 3.3.5 on page 45.

We denote the norm function by $\mathcal{N}\colon \mathbb{Z}[\tau] \longrightarrow \mathbb{Z}$, and we simply have $\mathcal{N}(\tau) = |\tau|^2$. We write $\tau = |\tau|\, e^{i\theta}$ for $\theta \in (-\pi, \pi]$. Further Iverson's notation $[expr] = 1$ if *expr* is true and $[expr] = 0$ otherwise, cf. Graham, Knuth and Patashnik [25], will be used.

In this section we will prove our main result on the asymptotic number of occurrences of a digit in a given region.

Theorem 3.10.1 (Counting Theorem)**.** *Let $0 \neq \eta \in \mathcal{D}$ and $N \in \mathbb{R}$ with $N \geq 0$. Further let $U \subseteq \mathbb{C}$ be measurable with respect to the Lebesgue measure, $U \subseteq \mathcal{B}(0,d)$ with d finite, i.e., U bounded, and set δ such that $\#\big(\partial(NU)_{\bigcirc}\big)_{\bigcirc} = \mathcal{O}\big(N^{\delta}\big)$. Assume $1 \leq \delta < 2$. We denote the number of occurrences of the digit η in all width-w non-adjacent forms with value in the region NU by*

$$Z_{\tau,w,\eta}(N) = \sum_{z \in NU \cap \mathbb{Z}[\tau]} \sum_{j \in \mathbb{N}_0} [j\text{th digit of } z \text{ in its } w\text{-NAF-expansion equals } \eta]\,.$$

Then we get

$$Z_{\tau,w,\eta}(N) = e_w N^2\, \lambda(U) \log_{|\tau|} N + N^2\, \psi_\eta\Big(\log_{|\tau|} N\Big) + \mathcal{O}\Big(N^{\alpha} \log_{|\tau|} N\Big) + \mathcal{O}\Big(N^{\delta} \log_{|\tau|} N\Big)\,,$$

in which the following expressions are used. The Lebesgue measure is denoted by λ. We have the constant of the expectation

$$e_w = \frac{1}{|\tau|^{2(w-1)} \left(\left(|\tau|^2 - 1\right) w + 1\right)},$$

cf. Theorem 3.4.1 on page 48. Then there is the function

$$\psi_\eta(x) = \psi_{\eta,\mathcal{M}}(x) + \psi_{\eta,\mathcal{P}}(x) + \psi_{\eta,\mathcal{Q}}(x)\,,$$

where

$$\psi_{\eta,\mathcal{M}}(x) = \lambda(U)\,(c + 1 - \{x\})\, e_w,$$

$$\psi_{\eta,\mathcal{P}}(x) = \frac{|\tau|^{2(c-\{x\})}}{\lambda(V)} \sum_{j=0}^{\infty} \int_{y \in \left\{|\tau|^{\{x\}-c}\, \widehat{\theta}(\lfloor x \rfloor) U\right\}_{\bigcirc, j-w}} \left(\mathbb{1}_W\Big(\left\{y\tau^{j-w}\right\}_{\mathbb{Z}[\tau]}\Big) - \lambda(W)\right) \mathrm{d}y,$$

with the rotation $\widehat{\theta}(x) = e^{-i\theta x - i\theta c}$*, and*

$$\psi_{\eta,\mathcal{Q}}(x) = \psi_{\eta,\mathcal{Q}} = \frac{\lambda(U)}{\lambda(V)} \sum_{j=0}^{\infty} \frac{\beta_j}{\lambda(V)}.$$

We have $\alpha = 2 + \log_{|\tau|} \rho < 2$*, with* $\rho = \left(1 + \frac{1}{|\tau|^2 w^3}\right)^{-1} < 1$*, and*

$$c = \left\lfloor \log_{|\tau|} d - \log_{|\tau|} f_L \right\rfloor + 1$$

with the constant f_L *of Proposition 3.5.2 on page 56.*

Further, if there is a $p \in \mathbb{N}$*, such that* $e^{2i\theta p} U = U$*, then* ψ_η *is* p*-periodic and continuous.*

Remark 3.10.2. Consider the main term of our result. When N tends to infinity, we get the asymptotic formula

$$Z_{\tau,w,\eta} \sim e_w N^2 \, \lambda(U) \log_{|\tau|} N.$$

This result is not surprising, since intuitively there are about $N^2 \, \lambda(U)$ w-NAFs in the region NU, and each of them can be represented as a w-NAF with length $\log_{|\tau|} N$. Therefore, using the expectation of Theorem 3.4.1 on page 48, we get an explanation for this term.

Remark 3.10.3. Using a disc as region U, e.g. $U = \mathcal{B}(0,1)$, yields that ψ_η is 1-periodic and continuous for all valid τ. The reason is that the condition $e^{i\theta p} U = U$ is then clearly fulfilled for every p, especially for $p = 1$.

The parameter δ is 1 for simple geometries like a disc or a polygon. See Proposition 3.8.4 on page 71 for details.

Remark 3.10.4. If $\delta = 2$ in the theorem, then the statement stays true, but degenerates to

$$Z_{\tau,w,\eta}(N) = \mathcal{O}\left(N^2 \log_{|\tau|} N\right).$$

This is a trivial result of Remark 3.10.2.

The proof of Theorem 3.10.1 on the facing page follows the ideas used by Delange [17]. By Remark 3.10.4 we restrict ourselves to the case $\delta < 2$.

We will use the following abbreviations. We set $Z(N) := Z_{\tau,w,\eta}(N)$, and we set $W := W_\eta$ and $W_j := W_{\eta,j}$ for our fixed η of Theorem 3.10.1 on the facing page. Further we set $\beta_j := \beta_{\eta,j}$, cf. Proposition 3.9.2 on page 72. By log we will denote the logarithm to the base $|\tau|$, i.e., $\log x = \log_{|\tau|} x$. These abbreviations will be used throughout the remaining section.

Proof of Theorem 3.10.1. We know from Theorem 3.6.1 on page 60 that every element of $\mathbb{Z}[\tau]$ is represented by a unique element of $\mathbf{NAF}_w^{\mathsf{fin}}$. To count the occurrences of the digit η in NU, we sum up 1 over all lattice points $n \in NU \cap \mathbb{Z}[\tau]$ and for each n over all digits in the corresponding w-NAF equal to η. Thus we get

$$Z(N) = \sum_{n \in NU \cap \mathbb{Z}[\tau]} \sum_{j \in \mathbb{N}_0} \left[\varepsilon_j(\mathsf{NAF}_w(n)) = \eta\right],$$

where ε_j denotes the extraction of the jth digit, i.e., for a w-NAF $\boldsymbol{\xi}$ we define $\varepsilon_j(\boldsymbol{\xi}) := \xi_j$. The inner sum over $j \in \mathbb{N}_0$ is finite, we will choose a large enough upper bound J later in Lemma 3.10.5 on page 82.

Using

$$\left[\varepsilon_j(\mathsf{NAF}_w(n)) = \eta\right] = \mathbb{1}_{W_j}\left(\left\{\frac{n}{\tau^{j+w}}\right\}_{\mathbb{Z}[\tau]}\right)$$

from Lemma 3.9.4 on page 77 yields

$$Z(N) = \sum_{j=0}^{J} \sum_{n \in NU \cap \mathbb{Z}[\tau]} \mathbb{1}_{W_j}\left(\left\{\frac{n}{\tau^{j+w}}\right\}_{\mathbb{Z}[\tau]}\right),$$

where additionally the order of summation was changed. This enables us to rewrite the sum over n as integral

$$\begin{aligned} Z(N) &= \sum_{j=0}^{J} \sum_{n \in NU \cap \mathbb{Z}[\tau]} \frac{1}{\lambda(V_n)} \int_{x \in V_n} \mathbb{1}_{W_j}\left(\left\{\frac{x}{\tau^{j+w}}\right\}_{\mathbb{Z}[\tau]}\right) \mathrm{d}x \\ &= \frac{1}{\lambda(V)} \sum_{j=0}^{J} \int_{x \in \lfloor NU \rceil_{\bigcirc}} \mathbb{1}_{W_j}\left(\left\{\frac{x}{\tau^{j+w}}\right\}_{\mathbb{Z}[\tau]}\right) \mathrm{d}x. \end{aligned}$$

We split up the integrals into the ones over NU and others over the remaining region and get

$$Z(N) = \frac{1}{\lambda(V)} \sum_{j=0}^{J} \int_{x \in NU} \mathbb{1}_{W_j}\left(\left\{\frac{x}{\tau^{j+w}}\right\}_{\mathbb{Z}[\tau]}\right) \mathrm{d}x + \mathcal{F}_\eta(N),$$

in which $\mathcal{F}_\eta(N)$ contains all integrals (with appropriate signs) over regions $\lfloor NU \rceil_{\bigcirc} \setminus NU$ and $NU \setminus \lfloor NU \rceil_{\bigcirc}$.

Substituting $x = \tau^J y,\ \mathrm{d}x = |\tau|^{2J}\ \mathrm{d}y$ we obtain

$$Z(N) = \frac{|\tau|^{2J}}{\lambda(V)} \sum_{j=0}^{J} \int_{y \in \tau^{-J} NU} \mathbb{1}_{W_j}\left(\left\{y\tau^{J-j-w}\right\}_{\mathbb{Z}[\tau]}\right) \mathrm{d}y + \mathcal{F}_\eta(N).$$

Reversing the order of summation yields

$$Z(N) = \frac{|\tau|^{2J}}{\lambda(V)} \sum_{j=0}^{J} \int_{y \in \tau^{-J} NU} \mathbb{1}_{W_{J-j}}\left(\left\{y\tau^{j-w}\right\}_{\mathbb{Z}[\tau]}\right) \mathrm{d}y + \mathcal{F}_\eta(N).$$

We rewrite this as

$$\begin{aligned} Z(N) = {} & \frac{|\tau|^{2J}}{\lambda(V)} (J+1)\, \lambda(W) \int_{y \in \tau^{-J} NU} \mathrm{d}y \\ & + \frac{|\tau|^{2J}}{\lambda(V)} \sum_{j=0}^{J} \int_{y \in \tau^{-J} NU} \left(\mathbb{1}_W\left(\left\{y\tau^{j-w}\right\}_{\mathbb{Z}[\tau]}\right) - \lambda(W)\right) \mathrm{d}y \\ & + \frac{|\tau|^{2J}}{\lambda(V)} \sum_{j=0}^{J} \int_{y \in \tau^{-J} NU} \left(\mathbb{1}_{W_{J-j}}\left(\left\{y\tau^{j-w}\right\}_{\mathbb{Z}[\tau]}\right) - \mathbb{1}_W\left(\left\{y\tau^{j-w}\right\}_{\mathbb{Z}[\tau]}\right)\right) \mathrm{d}y \\ & + \mathcal{F}_\eta(N). \end{aligned}$$

With $\tau^{-J} NU = \left\lfloor \tau^{-J} NU \right\rfloor_{\bigcirc, j-w} \cup \left\{\tau^{-J} NU\right\}_{\bigcirc, j-w}$, see Figure 3.10.1 on the next page, for each integral region we get

$$Z(N) = \mathcal{M}_\eta(N) + \mathcal{Z}_\eta(N) + \mathcal{P}_\eta(N) + \mathcal{Q}_\eta(N) + \mathcal{S}_\eta(N) + \mathcal{F}_\eta(N),$$

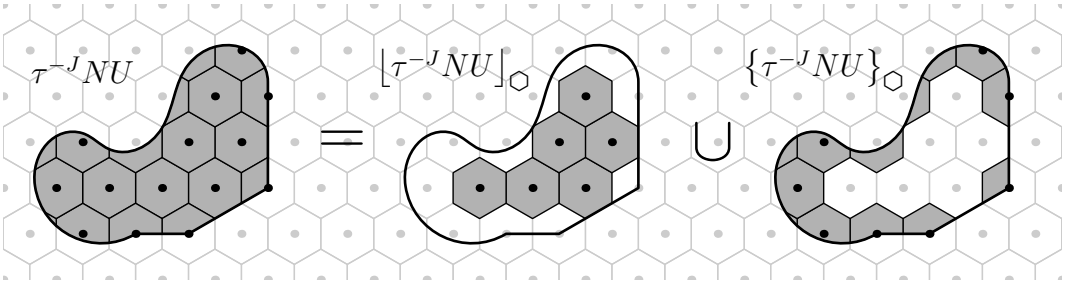


Figure 3.10.1: Splitting up the region of integration $\tau^{-J}NU$.

in which $\mathcal{M}_\eta$ is *"The Main Part"*, see Lemma 3.10.8 on the following page,

$$\mathcal{M}_\eta(N) = \frac{|\tau|^{2J}}{\lambda(V)}(J+1)\,\lambda(W)\int_{y\in\tau^{-J}NU}\mathrm{d}y, \tag{3.10.1a}$$

$\mathcal{Z}_\eta$ is *"The Zero Part"*, see Lemma 3.10.9 on page 83,

$$\mathcal{Z}_\eta(N) = \frac{|\tau|^{2J}}{\lambda(V)}\sum_{j=0}^{J}\int_{y\in\lfloor\tau^{-J}NU\rfloor_{\hexagon,j-w}}\Big(\mathbb{1}_W\Big(\big\{y\tau^{j-w}\big\}_{\mathbb{Z}[\tau]}\Big)-\lambda(W)\Big)\,\mathrm{d}y, \tag{3.10.1b}$$

$\mathcal{P}_\eta$ is *"The Periodic Part"*, see Lemma 3.10.10 on page 83,

$$\mathcal{P}_\eta(N) = \frac{|\tau|^{2J}}{\lambda(V)}\sum_{j=0}^{J}\int_{y\in\{\tau^{-J}NU\}_{\hexagon,j-w}}\Big(\mathbb{1}_W\Big(\big\{y\tau^{j-w}\big\}_{\mathbb{Z}[\tau]}\Big)-\lambda(W)\Big)\,\mathrm{d}y, \tag{3.10.1c}$$

$\mathcal{Q}_\eta$ is *"The Other Part"*, see Lemma 3.10.11 on page 85,

$$\mathcal{Q}_\eta(N) = \frac{|\tau|^{2J}}{\lambda(V)}\sum_{j=0}^{J}\int_{y\in\lfloor\tau^{-J}NU\rfloor_{\hexagon,j-w}}\big(\mathbb{1}_{W_{J-j}}-\mathbb{1}_W\big)\Big(\big\{y\tau^{j-w}\big\}_{\mathbb{Z}[\tau]}\Big)\,\mathrm{d}y, \tag{3.10.1d}$$

$\mathcal{S}_\eta$ is *"The Small Part"*, see Lemma 3.10.12 on page 86,

$$\mathcal{S}_\eta(N) = \frac{|\tau|^{2J}}{\lambda(V)}\sum_{j=0}^{J}\int_{y\in\{\tau^{-J}NU\}_{\hexagon,j-w}}\big(\mathbb{1}_{W_{J-j}}-\mathbb{1}_W\big)\Big(\big\{y\tau^{j-w}\big\}_{\mathbb{Z}[\tau]}\Big)\,\mathrm{d}y \tag{3.10.1e}$$

and $\mathcal{F}_\eta$ is *"The Fractional Cells Part"*, see Lemma 3.10.13 on page 88,

$$\begin{aligned}\mathcal{F}_\eta(N) = {}& \frac{1}{\lambda(V)}\sum_{j=0}^{J}\int_{x\in\lfloor NU\rceil_{\hexagon}\setminus NU}\mathbb{1}_{W_j}\Big(\Big\{\frac{x}{\tau^{j+w}}\Big\}_{\mathbb{Z}[\tau]}\Big)\,\mathrm{d}x\\ & -\frac{1}{\lambda(V)}\sum_{j=0}^{J}\int_{x\in NU\setminus\lfloor NU\rceil_{\hexagon}}\mathbb{1}_{W_j}\Big(\Big\{\frac{x}{\tau^{j+w}}\Big\}_{\mathbb{Z}[\tau]}\Big)\,\mathrm{d}x.\end{aligned} \tag{3.10.1f}$$

To complete the proof we have to deal with the choice of J, see Lemma 3.10.5 on the following page, as well as with each of the parts in (3.10.1), see Lemmata 3.10.8 to 3.10.13 on pages 82–88. The continuity of ψ_η is checked in Lemma 3.10.14 on page 88. □

Lemma 3.10.5 (Choosing J)**.** *Let $N \in \mathbb{R}_{\geq 0}$. Then every w-NAF of $\mathbf{NAF}_w^{\mathsf{fin}}$ with value in NU has at most $J+1$ digits, where*

$$J = \lfloor \log N \rfloor + c$$

with

$$c = \lfloor \log d - \log f_L \rfloor + 1$$

with f_L of Proposition 3.5.2 on page 56.

Proof. Let $z \in NU$, $z \neq 0$, with its corresponding w-NAF $\boldsymbol{\xi} \in \mathbf{NAF}_w^{\mathsf{fin}}$, and let $j \in \mathbb{N}_0$ be the largest index, such that the digit ξ_j is non-zero. By using Corollary 3.5.3 on page 57, we conclude that

$$|\tau|^j f_L \leq |z| < Nd.$$

This means

$$j < \log N + \log d - \log f_L,$$

and thus we have

$$j \leq \lfloor \log N + \log d - \log f_L \rfloor \leq \lfloor \log N \rfloor + \lfloor \log d - \log f_L \rfloor + 1.$$

Defining the right hand side of this inequality as J finishes the proof. □

Remark 3.10.6. For the parameter used in the region of integration in the proof of Theorem 3.10.1 on page 78 we get

$$\tau^{-J} N = |\tau|^{\{\log N\} - c}\, \widehat{\theta}(\log N)\,,$$

with the rotation $\widehat{\theta}(x) = e^{-i\theta\lfloor x \rfloor - i\theta c}$. In particular we get $\left|\tau^{-J} N\right| = \mathcal{O}(1)$.

Proof. With $\tau = |\tau|\, e^{i\theta}$ and the J of Lemma 3.10.5 we obtain

$$\tau^{-J} N = \tau^{-\lfloor \log N \rfloor - c}\, |\tau|^{\log N} = |\tau|^{-c - \lfloor \log N \rfloor + \log N}\, e^{-i\theta(\lfloor \log N \rfloor + c)} = |\tau|^{\{\log N\} - c}\, \widehat{\theta}(\log N)\,. \quad \square$$

Remark 3.10.7. Let $\gamma \in \mathbb{R}$ with $\gamma \geq 1$, then

$$\gamma^J = N^{\log \gamma} \gamma^{c - \{\log N\}} = \mathcal{O}\big(N^{\log \gamma}\big)\,.$$

In particular $|\tau|^{2J} = \mathcal{O}\big(N^2\big)$ and $|\tau|^J = \mathcal{O}(N)$.

Proof. We insert J from Lemma 3.10.5 and obtain

$$\begin{aligned} \gamma^J &= \gamma^{\lfloor \log N \rfloor + c} = \gamma^{\log N} \gamma^{c - \{\log N\}} = |\tau|^{\log N \log \gamma}\, \gamma^{c - \{\log N\}} \\ &= N^{\log \gamma} \gamma^{c - \{\log N\}} = \mathcal{O}\big(N^{\log \gamma}\big)\,. \end{aligned}$$

□

Lemma 3.10.8 (The Main Part)**.** *For* (3.10.1a) *in the proof of Theorem 3.10.1 on page 78 we get*

$$\mathcal{M}_\eta(N) = e_w N^2\, \lambda(U) \log N + N^2\, \psi_{\eta,\mathcal{M}}(\log N)$$

with a 1-periodic function $\psi_{\eta,\mathcal{M}}$,

$$\psi_{\eta,\mathcal{M}}(x) = \lambda(U)\, (c + 1 - \{x\})\, e_w$$

and

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

Proof. We have

$$\mathcal{M}_\eta(N) = \frac{|\tau|^{2J}}{\lambda(V)}(J+1)\,\lambda(W)\int_{y\in\tau^{-J}NU}\mathrm{d}y.$$

As $\lambda\big(\tau^{-J}NU\big) = |\tau|^{-2J}\,N^2\,\lambda(U)$ we obtain

$$\mathcal{M}_\eta(N) = \frac{\lambda(W)}{\lambda(V)}(J+1)N^2\,\lambda(U)\,.$$

By taking $\lambda(W) = \lambda(V)\,e_w$ from (i) of Proposition 3.9.2 on page 72 and J from Lemma 3.10.5 on the preceding page we get

$$\mathcal{M}_\eta(N) = N^2\,\lambda(U)\,e_w\,(\lfloor\log N\rfloor + c + 1)\,.$$

Finally, the desired result follows by using $x = \lfloor x\rfloor + \{x\}$. □

Lemma 3.10.9 (The Zero Part)**.** *For* (3.10.1b) *in the proof of Theorem 3.10.1 on page 78 we get*

$$\mathcal{Z}_\eta(N) = 0.$$

Proof. Consider the integral

$$I_j := \int_{y\in\lfloor\tau^{-J}NU\rfloor_{\bigcirc,j-w}}\Big(\mathbb{1}_W\Big(\big\{y\tau^{j-w}\big\}_{\mathbb{Z}[\tau]}\Big) - \lambda(W)\Big)\,\mathrm{d}y.$$

We can rewrite the region of integration as

$$\big\lfloor\tau^{-J}NU\big\rfloor_{\bigcirc,j-w} = \frac{1}{\tau^{j-w}}\,\big\lfloor\tau^{j-w}\tau^{-J}NU\big\rfloor_{\bigcirc} = \frac{1}{\tau^{j-w}}\bigcup_{z\in T_{j-w}}V_z$$

for some appropriate $T_{j-w}\subseteq\mathbb{Z}[\tau]$. Substituting $x = \tau^{j-w}y,\ \mathrm{d}x = |\tau|^{2(j-w)}\,\mathrm{d}y$ yields

$$I_j = \frac{1}{|\tau|^{2(j-w)}}\int_{x\in\bigcup_{z\in T_{j-w}}V_z}\Big(\mathbb{1}_W\Big(\{x\}_{\mathbb{Z}[\tau]}\Big) - \lambda(W)\Big)\,\mathrm{d}x.$$

We split up the integral and eliminate the fractional part $\{x\}_{\mathbb{Z}[\tau]}$ by translation to get

$$I_j = \frac{1}{|\tau|^{2(j-w)}}\sum_{z\in T_{j-w}}\underbrace{\int_{x\in V}\big(\mathbb{1}_W(x) - \lambda(W)\big)\,\mathrm{d}x}_{=0}\,.$$

Thus, for all $j\in\mathbb{N}_0$ we obtain $I_j = 0$, and therefore $\mathcal{Z}_\eta(N) = 0$. □

Lemma 3.10.10 (The Periodic Part)**.** *For* (3.10.1c) *in the proof of Theorem 3.10.1 on page 78 we get*

$$\mathcal{P}_\eta(N) = N^2\,\psi_{\eta,\mathcal{P}}(\log N) + \mathcal{O}\big(N^\delta\big)$$

with a function $\psi_{\eta,\mathcal{P}}$*,*

$$\psi_{\eta,\mathcal{P}}(x) = \frac{|\tau|^{2(c-\{x\})}}{\lambda(V)}\sum_{j=0}^{\infty}\int_{y\in\left\{|\tau|^{\{x\}-c}\,\widehat{\theta}(\lfloor x\rfloor)U\right\}_{\bigcirc,j-w}}\Big(\mathbb{1}_W\Big(\big\{y\tau^{j-w}\big\}_{\mathbb{Z}[\tau]}\Big) - \lambda(W)\Big)\,\mathrm{d}y,$$

with the rotation $\widehat{\theta}(x) = e^{-i\theta x - i\theta c}$*.*

If there is a $p\in\mathbb{N}$*, such that* $e^{i\theta p}U = U$*, then* $\psi_{\eta,\mathcal{P}}$ *is p-periodic.*

Proof. Consider

$$I_j := \int_{y \in \{\tau^{-J} NU\}_{\bigcirc, j-w}} \left(\mathbb{1}_W \left(\{ y \tau^{j-w} \}_{\mathbb{Z}[\tau]} \right) - \lambda(W) \right) \mathrm{d}y.$$

The region of integration satisfies

$$\{\tau^{-J} NU\}_{\bigcirc, j-w} \subseteq \partial(\tau^{-J} NU)_{\bigcirc, j-w} = \frac{1}{\tau^{j-w}} \bigcup_{z \in T_{j-w}} V_z \tag{3.10.2}$$

for some appropriate $T_{j-w} \subseteq \mathbb{Z}[\tau]$.

We use the triangle inequality and substitute $x = \tau^{j-w} y,\ \mathrm{d}x = |\tau|^{2(j-w)}\ \mathrm{d}y$ in the integral to get

$$|I_j| \leq \frac{1}{|\tau|^{2(j-w)}} \int_{x \in \bigcup_{z \in T_{j-w}} V_z} \underbrace{\left| \mathbb{1}_W \left(\{x\}_{\mathbb{Z}[\tau]} \right) - \lambda(W) \right|}_{\leq 1 + \lambda(W)} \mathrm{d}x.$$

After splitting up the integral and using translation to eliminate the fractional part, we get

$$|I_j| \leq \frac{1 + \lambda(W)}{|\tau|^{2(j-w)}} \sum_{z \in T_{j-w}} \int_{x \in V} \mathrm{d}x = \frac{1 + \lambda(W)}{|\tau|^{2(j-w)}} \lambda(V) \, \#(T_{j-w}) \, .$$

Using $\#\big(\partial(NU)_{\bigcirc}\big)_{\bigcirc} = \mathcal{O}\big(N^\delta\big)$ as assumed and Equation (3.10.2) we gain

$$\#(T_{j-w}) = \mathcal{O}\Big(|\tau|^{(j-w)\delta} \left| \tau^{-J} N \right|^\delta \Big) = \mathcal{O}\Big(|\tau|^{(j-w)\delta} \Big) \, ,$$

because $\left| \tau^{-J} N \right| = \mathcal{O}(1)$, see Remark 3.10.6 on page 82, and thus

$$|I_j| = \mathcal{O}\Big(|\tau|^{\delta(j-w) - 2(j-w)} \Big) = \mathcal{O}\Big(|\tau|^{(\delta-2)j} \Big) \, .$$

Now we want to make the summation in $\mathcal{P}_\eta$ independent from J, so we consider

$$I := \frac{|\tau|^{2J}}{\lambda(V)} \sum_{j=J+1}^{\infty} I_j$$

Again we use triangle inequality and we calculate the sum to obtain

$$|I| = \mathcal{O}\Big(|\tau|^{2J} \Big) \sum_{j=J+1}^{\infty} \mathcal{O}\Big(|\tau|^{(\delta-2)j} \Big) = \mathcal{O}\Big(|\tau|^{2J} \, |\tau|^{(\delta-2)J} \Big) = \mathcal{O}\Big(|\tau|^{\delta J} \Big) \, .$$

Note that $\mathcal{O}\Big(|\tau|^J \Big) = \mathcal{O}(N)$, see Remark 3.10.7 on page 82, so we obtain $|I| = \mathcal{O}\big(N^\delta\big)$.

Let us look at the growth of

$$\mathcal{P}_\eta(N) = \frac{|\tau|^{2J}}{\lambda(V)} \sum_{j=0}^{J} I_j.$$

We get

$$|\mathcal{P}_\eta(N)| = \mathcal{O}\Big(|\tau|^{2J} \Big) \sum_{j=0}^{J} \mathcal{O}\Big(|\tau|^{(\delta-2)j} \Big) = \mathcal{O}\Big(|\tau|^{2J} \Big) = \mathcal{O}\big(N^2\big) \, ,$$

using $\delta < 2$, and, to get the last equality, Remark 3.10.7 on page 82.

Finally, inserting the result of Remark 3.10.6 on page 82 for the region of integration, rewriting $|\tau|^{2J}$ according to Remark 3.10.7 on page 82 and extending the sum to infinity, as above described, yields

$$\begin{aligned}\mathcal{P}_\eta(N) &= \frac{|\tau|^{2J}}{\lambda(V)} \sum_{j=0}^{J} \int_{y\in\{\tau^{-J}NU\}_{\bigcirc,j-w}} \Big(\mathbb{1}_W\Big(\big\{y\tau^{j-w}\big\}_{\mathbb{Z}[\tau]}\Big) - \lambda(W)\Big)\,\mathrm{d}y \\ &= N^2 \underbrace{\frac{|\tau|^{2(c-\{\log N\})}}{\lambda(V)} \sum_{j=0}^{\infty} \int_{y\in\left\{|\tau|^{\{\log N\}-c}\,\widehat{\theta}(\lfloor\log N\rfloor)U\right\}_{\bigcirc,j-w}} \Big(\mathbb{1}_W\Big(\big\{y\tau^{j-w}\big\}_{\mathbb{Z}[\tau]}\Big) - \lambda(W)\Big)\,\mathrm{d}y}_{=:\psi_{\eta,\mathcal{P}}(\log N)} \\ &\quad + \mathcal{O}\big(N^\delta\big)\,,\end{aligned}$$

with the rotation $\widehat{\theta}(x) = e^{-i\theta x - i\theta c}$.

Now let

$$e^{i\theta p}U = U \Longleftrightarrow e^{-i\theta p}U = e^{-i\theta 0}U.$$

Clearly the region of integration in $\psi_{\eta,\mathcal{P}}(x)$ is p-periodic, since x occurs as $\{x\}$ and $\lfloor x\rfloor$. All other occurrences of x are of the form $\{x\}$, i.e., 1-periodic, so period p is obtained. □

Lemma 3.10.11 (The Other Part)**.** *For* (3.10.1d) *in the proof of Theorem 3.10.1 on page 78 we get*

$$\mathcal{Q}_\eta(N) = N^2\psi_{\eta,\mathcal{Q}} + \mathcal{O}(N^\alpha \log N) + \mathcal{O}\big(N^\delta\big)\,,$$

with

$$\psi_{\eta,\mathcal{Q}} = \frac{\lambda(U)}{\lambda(V)} \sum_{j=0}^{\infty} \frac{\beta_j}{\lambda(V)}$$

and $\alpha = 2 + \log\rho < 2$, where $\rho < 1$ can be found in Theorem 3.4.1 on page 48.

Proof. Consider

$$I_{j,\ell} := \int_{y\in\lfloor\tau^{-J}NU\rfloor_{\bigcirc,j-w}} \big(\mathbb{1}_{W_{\eta,\ell}} - \mathbb{1}_W\big)\Big(\big\{y\tau^{j-w}\big\}_{\mathbb{Z}[\tau]}\Big)\,\mathrm{d}y.$$

We can rewrite the region of integration and get

$$\big\lfloor\tau^{-J}NU\big\rfloor_{\bigcirc,j-w} = \frac{1}{\tau^{j-w}}\,\big\lfloor\tau^{j-w}\tau^{-J}NU\big\rfloor_{\bigcirc} = \frac{1}{\tau^{j-w}} \bigcup_{z\in T_{j-w}} V_z$$

for some appropriate $T_{j-w} \subseteq \mathbb{Z}[\tau]$, as in the proof of Lemma 3.10.9 on page 83. Substituting $x = \tau^{j-w}y,\ \mathrm{d}x = |\tau|^{2(j-w)}\,\mathrm{d}y$ yields

$$I_{j,\ell} = \frac{1}{|\tau|^{2(j-w)}} \int_{x\in\bigcup_{z\in T_{j-w}} V_z} \big(\mathbb{1}_{W_{\eta,\ell}} - \mathbb{1}_W\big)\Big(\{x\}_{\mathbb{Z}[\tau]}\Big)\,\mathrm{d}x$$

and further

$$I_{j,\ell} = \frac{1}{|\tau|^{2(j-w)}} \sum_{z\in T_{j-w}} \underbrace{\int_{x\in V} \big(\mathbb{1}_{W_{\eta,\ell}} - \mathbb{1}_W\big)(x)\,\mathrm{d}x}_{=\beta_\ell} = \frac{1}{|\tau|^{2(j-w)}}\,\#(T_{j-w})\,\beta_\ell,$$

by splitting up the integral, using translation to eliminate the fractional part and taking β_ℓ according to (j) of Proposition 3.9.2 on page 72. From Proposition 3.8.3 on page 70 we obtain

$$\frac{\#(T_{j-w})}{|\tau|^{2(j-w)}} = \frac{\left|\tau^{j-w}\tau^{-J}N\right|^2}{|\tau|^{2(j-w)}}\frac{\lambda(U)}{\lambda(V)} + \mathcal{O}\left(\frac{\left|\tau^{j-w}\tau^{-J}N\right|^\delta}{|\tau|^{2(j-w)}}\right) = \left|\tau^{-J}N\right|^2\frac{\lambda(U)}{\lambda(V)} + \mathcal{O}\Big(|\tau|^{(\delta-2)j}\Big),$$

because $\left|\tau^{-J}N\right| = \mathcal{O}(1)$, see Remark 3.10.6 on page 82.

Now let us have a look at

$$\mathcal{Q}_\eta(N) = \frac{|\tau|^{2J}}{\lambda(V)}\sum_{j=0}^{J} I_{j,J-j}.$$

Inserting the result above and using $\beta_\ell = \mathcal{O}\big(\rho^\ell\big)$, see (j) of Proposition 3.9.2 on page 72, yields

$$\mathcal{Q}_\eta(N) = |\tau|^{2J}\left|\tau^{-J}N\right|^2\frac{\lambda(U)}{\lambda(V)}\sum_{j=0}^{J}\frac{\beta_{J-j}}{\lambda(V)} + |\tau|^{2J}\sum_{j=0}^{J}\mathcal{O}\Big(|\tau|^{(\delta-2)j}\Big)\,\mathcal{O}\big(\rho^{J-j}\big)$$

We notice that $|\tau|^{2J}\left|\tau^{-J}N\right|^2 = N^2$.

Therefore, after reversing the order of the first summation, we obtain

$$\mathcal{Q}_\eta(N) = N^2\frac{\lambda(U)}{\lambda(V)}\sum_{j=0}^{J}\frac{\beta_j}{\lambda(V)} + |\tau|^{2J}\,\rho^J\sum_{j=0}^{J}\mathcal{O}\left(\Big(\rho\,|\tau|^{2-\delta}\Big)^{-j}\right).$$

If $\rho\,|\tau|^{2-\delta} \geq 1$, then the second sum is $J\,\mathcal{O}(1)$, otherwise the sum is $\mathcal{O}\Big(\rho^{-J}\,|\tau|^{(\delta-2)J}\Big)$. So we obtain

$$\mathcal{Q}_\eta(N) = N^2\frac{\lambda(U)}{\lambda(V)}\sum_{j=0}^{J}\frac{\beta_j}{\lambda(V)} + \mathcal{O}\Big(|\tau|^{2J}\,\rho^J J\Big) + \mathcal{O}\Big(|\tau|^{\delta J}\Big).$$

Using $J = \Theta(\log N)$, see Lemma 3.10.5 on page 82, Remark 3.10.7 on page 82, and defining $\alpha = 2 + \log\rho$ yields

$$\mathcal{Q}_\eta(N) = N^2\frac{\lambda(U)}{\lambda(V)}\sum_{j=0}^{J}\frac{\beta_j}{\lambda(V)} + \underbrace{\mathcal{O}\big(N^{2+\log\rho}\log N\big)}_{=\mathcal{O}(N^\alpha\log N)} + \mathcal{O}\big(N^\delta\big).$$

Now consider the first sum. Since $\beta_j = \mathcal{O}\big(\rho^j\big)$, see (j) of Proposition 3.9.2 on page 72, we obtain

$$N^2\sum_{j=J+1}^{\infty}\beta_j = N^2\,\mathcal{O}\big(\rho^J\big) = \mathcal{O}(N^\alpha).$$

Thus the lemma is proved, because we can extend the sum to infinity. □

Lemma 3.10.12 (The Small Part)**.** *For* (3.10.1e) *in the proof of Theorem 3.10.1 on page 78 we get*

$$\mathcal{S}_\eta(N) = \mathcal{O}(N^\alpha\log N) + \mathcal{O}\big(N^\delta\big)$$

with $\alpha = 2 + \log\rho < 2$ *and* $\rho < 1$ *from Theorem 3.4.1 on page 48.*

Proof. Consider

$$I_{j,\ell} := \int_{y \in \{\tau^{-J} NU\}_{\bigcirc, j-w}} (\mathbb{1}_{W_\ell} - \mathbb{1}_W)\Big(\{y\tau^{j-w}\}_{\mathbb{Z}[\tau]}\Big)\, \mathrm{d}y.$$

Again, as in the proof of Lemma 3.10.10 on page 83, the region of integration satisfies

$$\{\tau^{-J} NU\}_{\bigcirc, j-w} \subseteq \partial(\tau^{-J} NU)_{\bigcirc, j-w} = \frac{1}{\tau^{j-w}} \bigcup_{z \in T_{j-w}} V_z, \tag{3.10.3}$$

for some appropriate $T_{j-w} \subseteq \mathbb{Z}[\tau]$.

We substitute $x = \tau^{j-w} y,\ \mathrm{d}x = |\tau|^{2(j-w)}\ \mathrm{d}y$ in the integral to get

$$|I_{j,\ell}| = \frac{1}{|\tau|^{2(j-w)}} \left| \int_{x \in \bigcup_{z \in T_{j-w}} V_z} (\mathbb{1}_{W_\ell} - \mathbb{1}_W)\Big(\{x\}_{\mathbb{Z}[\tau]}\Big)\, \mathrm{d}x \right|.$$

Again, after splitting up the integral, using translation to eliminate the fractional part and the triangle inequality, we get

$$|I_{j,\ell}| \leq \frac{1}{|\tau|^{2(j-w)}} \sum_{z \in T_{j-w}} \underbrace{\left| \int_{x \in V} (\mathbb{1}_{W_\ell} - \mathbb{1}_W)(x)\, \mathrm{d}x \right|}_{=|\beta_\ell|} = \frac{1}{|\tau|^{2(j-w)}} \#(T_{j-w})\, |\beta_\ell|\,,$$

in which $|\beta_\ell| = \mathcal{O}\big(\rho^\ell\big)$ is known from (j) of Proposition 3.9.2 on page 72. Using $\#\big(\partial(NU)_{\bigcirc}\big)_{\bigcirc} = \mathcal{O}\big(N^\delta\big)$, Remark 3.10.6 on page 82, and Equation (3.10.3) we get

$$\#(T_{j-w}) = \mathcal{O}\Big(|\tau|^{(j-w)\delta} \left|\tau^{-J} N\right|^\delta\Big) = \mathcal{O}\Big(|\tau|^{\delta(j-w)}\Big)\,,$$

because $\left|\tau^{-J} N\right| = \mathcal{O}(1)$. Thus

$$|I_{j,\ell}| = \mathcal{O}\Big(\rho^\ell\, |\tau|^{(\delta-2)(j-w)}\Big) = \mathcal{O}\Big(\rho^\ell\, |\tau|^{(\delta-2)j}\Big)$$

follows by assembling all together.

Now we are ready to analyse

$$\mathcal{S}_\eta(N) = \frac{|\tau|^{2J}}{\lambda(V)} \sum_{j=0}^{J} I_{j,J-j}.$$

Inserting the result above yields

$$|\mathcal{S}_\eta(N)| = \frac{|\tau|^{2J}}{\lambda(V)} \sum_{j=0}^{J} \mathcal{O}\Big(\rho^{J-j}\, |\tau|^{(\delta-2)j}\Big) = \frac{\rho^J\, |\tau|^{2J}}{\lambda(V)} \sum_{j=0}^{J} \mathcal{O}\bigg(\Big(\rho\, |\tau|^{2-\delta}\Big)^{-j}\bigg)$$

and thus, by the same argument as in the proof of Lemma 3.10.11 on page 85,

$$|\mathcal{S}_\eta(N)| = \rho^J\, |\tau|^{2J}\, \mathcal{O}\Big(J + \rho^{-J}\, |\tau|^{(\delta-2)J}\Big) = \mathcal{O}\Big(\rho^J\, |\tau|^{2J}\, J\Big) + \mathcal{O}\Big(|\tau|^{\delta J}\Big)\,,$$

Finally, using Lemma 3.10.5 on page 82 and Remark 3.10.7 on page 82, we obtain

$$|\mathcal{S}_\eta(N)| = \mathcal{O}(N^\alpha \log N) + \mathcal{O}\big(N^\delta\big)$$

with $\alpha = 2 + \log \rho$. Since $\rho < 1$, we have $\alpha < 2$. □

Lemma 3.10.13 (The Fractional Cells Part)**.** *For* (3.10.1f) *in the proof of Theorem 3.10.1 on page 78 we get*

$$\mathcal{F}_\eta(N) = \mathcal{O}\big(N^\delta \log N\big)$$

Proof. For the regions of integration in $\mathcal{F}_\eta$ we obtain

$$NU \setminus \lfloor NU \rceil_{\bigcirc} \subseteq \lceil NU \rceil_{\bigcirc} \setminus \lfloor NU \rfloor_{\bigcirc} = \partial(NU)_{\bigcirc} = \bigcup_{z \in T} V_z$$

and

$$\lfloor NU \rceil_{\bigcirc} \setminus NU \subseteq \lceil NU \rceil_{\bigcirc} \setminus \lfloor NU \rfloor_{\bigcirc} = \partial(NU)_{\bigcirc} = \bigcup_{z \in T} V_z$$

for some appropriate $T \subseteq \mathbb{Z}[\tau]$ using Proposition 3.8.2 on page 68. Thus we get

$$|\mathcal{F}_\eta(N)| \le \frac{2}{\lambda(V)} \sum_{j=0}^{J} \int_{x \in \bigcup_{z \in T} V_z} \mathbb{1}_{W_j}\left(\left\{ \frac{x}{\tau^{j+w}} \right\}_{\mathbb{Z}[\tau]} \right) \mathrm{d}x \le \frac{2}{\lambda(V)} \sum_{j=0}^{J} \sum_{z \in T} \int_{x \in V_z} \mathrm{d}x,$$

in which the indicator function was replaced by 1. Dealing with the sums and the integral, which is $\mathcal{O}(1)$, we obtain

$$|\mathcal{F}_\eta(N)| = (J+1) \# T \, \mathcal{O}(1)\,.$$

Since $J = \mathcal{O}(\log N)$, see Lemma 3.10.5 on page 82, and $\# T = \mathcal{O}\big(N^\delta\big)$, the desired result follows. □

Lemma 3.10.14. *If the ψ_η from Theorem 3.10.1 on page 78 is p-periodic, then ψ_η is also continuous.*

Proof. There are two possible parts of ψ_η, where an discontinuity could occur. The first is $\{x\}$ for an $x \in \mathbb{Z}$, the second is building $\{\dots\}_{\bigcirc, j-w}$ in the region of integration in $\psi_{\eta,\mathcal{P}}$.

The latter is no problem, i.e., no discontinuity, since

$$\int_{y \in \left\{ |\tau|^{\{x\}-c}\, \widehat{\theta}(\lfloor x \rfloor) U \right\}_{\bigcirc, j-w}} \left(\mathbb{1}_W\Big(\big\{ y\tau^{j-w} \big\}_{\mathbb{Z}[\tau]} \Big) - \lambda(W) \right) \mathrm{d}y$$
$$= \int_{y \in |\tau|^{\{x\}-c}\, \widehat{\theta}(\lfloor x \rfloor) U} \left(\mathbb{1}_W\Big(\big\{ y\tau^{j-w} \big\}_{\mathbb{Z}[\tau]} \Big) - \lambda(W) \right) \mathrm{d}y,$$

because the integral of the region $\Big\lfloor |\tau|^{\{x\}-c}\, \widehat{\theta}(\lfloor x \rfloor)\, U \Big\rfloor_{\bigcirc, j-w}$ is zero, see proof of Lemma 3.10.9 on page 83.

Now we deal with the continuity for $x \in \mathbb{Z}$. Let $m \in x + p\mathbb{Z}$, let $M = |\tau|^m$, and consider

$$Z_\eta(M) - Z_\eta(M-1)\,.$$

For an appropriate $a \in \mathbb{R}$ we get

$$Z_\eta(M) = aM^2 \log M + M^2\, \psi_\eta(\log M) + \mathcal{O}(M^\alpha \log M) + \mathcal{O}\big(M^\delta \log M\big)\,,$$

and thus

$$Z_\eta(M) = aM^2 m + M^2 \underbrace{\psi_\eta(m)}_{=\psi_\eta(x)} + \mathcal{O}(M^\alpha m) + \mathcal{O}\big(M^\delta m\big)\,.$$

Further we obtain

$$Z_\eta(M-1) = a\,(M-1)^2 \log(M-1) + (M-1)^2\,\psi_\eta(\log(M-1)) \\ + \mathcal{O}((M-1)^\alpha \log(M-1)) + \mathcal{O}\Big((M-1)^\delta \log(M-1)\Big)\,,$$

and thus, using the abbreviation $L = \log\bigl(1 - M^{-1}\bigr)$ and $\delta \geq 1$,

$$Z_\eta(M-1) = aM^2m + M^2 \underbrace{\psi_\eta(m+L)}_{=\psi_\eta(x+L)} + \mathcal{O}(M^\alpha m) + \mathcal{O}\bigl(M^\delta m\bigr)\,.$$

Therefore we obtain

$$\frac{Z_\eta(M) - Z_\eta(M-1)}{M^2} = \psi_\eta(x) - \psi_\eta(x+L) + \mathcal{O}\bigl(M^{\alpha-2}m\bigr) + \mathcal{O}\bigl(M^{\delta-2}m\bigr)\,.$$

Since $\#(MU \setminus (M-1)\,U)_{\bigcirc}$ is clearly an upper bound for the number of w-NAFs with values in $MU \setminus (M-1)\,U$ and each of these w-NAFs has less than $\lfloor \log M \rfloor + c$ digits, see Lemma 3.10.5 on page 82, we obtain

$$Z_\eta(M) - Z_\eta(M-1) \leq \#(MU \setminus (M-1)\,U)_{\bigcirc}\,(m+c)\,.$$

Using (b) of Proposition 3.8.3 on page 70 yields then

$$Z_\eta(M) - Z_\eta(M-1) = \mathcal{O}\bigl(M^\delta m\bigr)\,.$$

Therefore we get

$$\psi_\eta(x) - \psi_\eta(x+L) = \mathcal{O}\bigl(M^{\delta-2}m\bigr) + \mathcal{O}\bigl(M^{\alpha-2}m\bigr) + \mathcal{O}\bigl(M^{\delta-2}m\bigr)\,.$$

Taking the limit $m \to \infty$ in steps of p, thus L tends to 0, and using $\alpha < 2$ and $\delta < 2$ yields

$$\psi_\eta(x) - \lim_{\varepsilon \to 0^-} \psi_\eta(x+\varepsilon) = 0,$$

i.e., ψ_η is continuous for $x \in \mathbb{Z}$. □

Chapter 4

Concluding Remarks and Some Open Problems

In this last chapter some concluding remarks on the results of Chapter 3 can be found. Further, there will also be mentioned some open problems in conjunction with those topics.

The analysis in Section 3.1 concerning Koblitz curves in characteristic 3 and 2-NAFs was similar to the analysis of the balanced ternary number system. The analysis is only for the rational integers and not for every element of the lattice $\mathbb{Z}[\tau]$.

In Section 3.2 we defined the Voronoi cell and the restricted Voronoi cell. In the latter definition, there was some freedom on choosing the boundary. In this work, one configuration was fixed, but it might be interesting what changes when taking other configurations.

Section 3.3 contained the definition of the digit sets used. They consisted of minimal norm representatives. But this is not the only meaningful choice. The question is: Are the presented results and theorems true for other digit sets? Or perhaps it can be shown that they are true in a more general setting.

In Section 3.3 there was also the definition of the τ-adic width-w non-adjacent form. One could think about using other concepts. One very interesting question concerns the optimality those representations. In the case of Koblitz curves in characteristic 2, the expansions are optimal, but this is not true in general in the characteristic 3 case, see Section 2.5 for details. Can there be proved a general statement which cases are optimal or non-optimal?

From Section 3.6 we know that every element of $\mathbb{Z}[\tau]$ admits a unique w-NAF for all imaginary quadratic algebraic integers τ and all $w \geq 2$. But is this true for a general algebraic integer τ? As mentioned above, the digit set was fixed with minimal norm representatives. What can be said, when other digit sets were used? Does every element of $\mathbb{Z}[\tau]$ still have a w-NAF representation?

Section 3.7 contained results on the fundamental domain. There it was shown that the dimension of the boundary is smaller than 2. But what is the exact value of it? Can this be calculated in general or is it only possible to calculate it for specific τ and w by using some computer algebra system?

One main result was the analysis of the w-NAFs in a certain region, e.g. in a disc (what means all w-NAFs with absolute value smaller than a given N). It can be found in Section 3.10. Here again this result was for imaginary quadratic algebraic integers τ. It was essential in the proof to have "imaginary quadratic", since in this case the Voronoi cell exists as described in Section 3.2. But is it possible to prove a similar result for general algebraic integer τ? Another question is, whether the exponent in the error term — it was called α — has a connection to the dimension of the boundary of the fundamental domain.

The last remark concerns w. It was always assumed that the integer w is at least 2. But what is in the case $w = 1$? Clearly this would mean that we do not have a non-adjacency condition and therefore a non-redundant number system. A lot of results are known for this case, but can anything be said in general (for all $w \in \mathbb{N}$) about that? What is the "correct" digit set to choose?

Appendix A

Existence of "Small" w-NAFs

Table A.0.1: w-NAF-expansions $\boldsymbol{\xi}$ of elements $z \in \mathbb{Z}[\tau]$ with "small" norm for "problematic values" of $|\tau|$ and w. "Small" norm means $|\cdot| \leq (1 - |\tau|^{-w})^{-1} |\tau| \, c_V$. The mapping $z \mapsto \boldsymbol{\xi}$ implies $\mathsf{value}(\boldsymbol{\xi}) = z$.

settings: $q = \|\tau\|^2 = 2$, $p = -2$, $\tau = -1 + 1i$, $w = 2$, $\|\cdot\|^2 \leq 4.667$ **digit set $\mathcal{D}$:** $0 = 0 + 0\tau$, $A = -1 + 0\tau$, $B = -1 - 1\tau$ **mapping** $z \mapsto \boldsymbol{\xi}$: $-2 - 2\tau \mapsto A0B0A00$, $-2 - 1\tau \mapsto A0B0$, $-1 - 1\tau \mapsto B$, $0 - 1\tau \mapsto A0$, $-2 + 0\tau \mapsto B00$, $-1 + 0\tau \mapsto A$, $0 + 0\tau \mapsto 0$, $1 + 0\tau \mapsto A0B0A$, $2 + 0\tau \mapsto A0B00$, $0 + 1\tau \mapsto A0B0A0$, $1 + 1\tau \mapsto A0B$, $2 + 1\tau \mapsto B0$, $2 + 2\tau \mapsto A00$
settings: $q = \|\tau\|^2 = 2$, $p = -1$, $\tau = -0.5 + 1.323i$, $w = 2$, $\|\cdot\|^2 \leq 4.667$ **digit set $\mathcal{D}$:** $0 = 0 + 0\tau$, $A = 1 + 0\tau$, $B = -1 + 0\tau$ **mapping** $z \mapsto \boldsymbol{\xi}$: $-2 - 1\tau \mapsto A00$, $-1 - 1\tau \mapsto A0A$, $0 - 1\tau \mapsto B0$, $1 - 1\tau \mapsto A00B$, $-2+0\tau \mapsto B0B0$, $-1+0\tau \mapsto B$, $0+0\tau \mapsto 0$, $1+0\tau \mapsto A$, $2+0\tau \mapsto A0A0$, $-1+1\tau \mapsto B00A$, $0 + 1\tau \mapsto A0$, $1 + 1\tau \mapsto B0B$, $2 + 1\tau \mapsto B00$
settings: $q = \|\tau\|^2 = 2$, $p = 0$, $\tau = 0 + 1.414i$, $w = 2$, $\|\cdot\|^2 \leq 4.667$ **digit set $\mathcal{D}$:** $0 = 0 + 0\tau$, $A = 1 + 0\tau$, $B = 1 + 1\tau$ **mapping** $z \mapsto \boldsymbol{\xi}$: $-1 - 1\tau \mapsto B0B$, $0 - 1\tau \mapsto A0A0$, $1 - 1\tau \mapsto A00B$, $-2 + 0\tau \mapsto A00$, $-1+0\tau \mapsto A0A$, $0+0\tau \mapsto 0$, $1+0\tau \mapsto A$, $2+0\tau \mapsto A0A00$, $-1+1\tau \mapsto A0B$, $0+1\tau \mapsto A0$, $1+1\tau \mapsto B$
settings: $q = \|\tau\|^2 = 2$, $p = 1$, $\tau = 0.5 + 1.323i$, $w = 2$, $\|\cdot\|^2 \leq 4.667$ **digit set $\mathcal{D}$:** $0 = 0 + 0\tau$, $A = -1 + 0\tau$, $B = 1 + 0\tau$ **mapping** $z \mapsto \boldsymbol{\xi}$: $-1 - 1\tau \mapsto B00B$, $0 - 1\tau \mapsto A0$, $1 - 1\tau \mapsto A0A$, $2 - 1\tau \mapsto A00$, $-2+0\tau \mapsto B0B0$, $-1+0\tau \mapsto A$, $0+0\tau \mapsto 0$, $1+0\tau \mapsto B$, $2+0\tau \mapsto A0A0$, $-2+1\tau \mapsto B00$, $-1+1\tau \mapsto B0B$, $0+1\tau \mapsto B0$, $1+1\tau \mapsto A00A$
settings: $q = \|\tau\|^2 = 2$, $p = 2$, $\tau = 1 + 1i$, $w = 2$, $\|\cdot\|^2 \leq 4.667$ **digit set $\mathcal{D}$:** $0 = 0 + 0\tau$, $A = -1 + 0\tau$, $B = 1 - 1\tau$ **mapping** $z \mapsto \boldsymbol{\xi}$: $2 - 2\tau \mapsto A00$, $0 - 1\tau \mapsto A0$, $1 - 1\tau \mapsto B$, $2 - 1\tau \mapsto B0$, $-2 + 0\tau \mapsto B0A0B00$, $-1 + 0\tau \mapsto A$, $0 + 0\tau \mapsto 0$, $1 + 0\tau \mapsto B0A$, $2 + 0\tau \mapsto B00$, $-2 + 1\tau \mapsto B0A0B0$, $-1 + 1\tau \mapsto B0A0B$, $0 + 1\tau \mapsto B0A0$, $-2 + 2\tau \mapsto B0A00$

---

settings: $q = |\tau|^2 = 2$, $p = -2$, $\tau = -1 + 1i$, $w = 3$, $|\cdot|^2 \leq 2.792$
digit set $\mathcal{D}$: $0 = 0 + 0\tau$, $A = 1 + 0\tau$, $B = 1 + 1\tau$, $C = -1 + 0\tau$, $D = -1 - 1\tau$
mapping $z \mapsto \boldsymbol{\xi}$: $-2 - 1\tau \mapsto B0$, $-1 - 1\tau \mapsto D$, $0 - 1\tau \mapsto C0$, $-1 + 0\tau \mapsto C$, $0 + 0\tau \mapsto 0$, $1 + 0\tau \mapsto A$, $0 + 1\tau \mapsto A0$, $1 + 1\tau \mapsto B$, $2 + 1\tau \mapsto D0$

---

settings: $q = |\tau|^2 = 2$, $p = -1$, $\tau = -0.5 + 1.323i$, $w = 3$, $|\cdot|^2 \leq 2.792$
digit set $\mathcal{D}$: $0 = 0 + 0\tau$, $A = 1 + 0\tau$, $B = 1 + 1\tau$, $C = -1 + 0\tau$, $D = -1 - 1\tau$
mapping $z \mapsto \boldsymbol{\xi}$: $-1 - 1\tau \mapsto D$, $0 - 1\tau \mapsto C0$, $-1 + 0\tau \mapsto C$, $0 + 0\tau \mapsto 0$, $1 + 0\tau \mapsto A$, $0 + 1\tau \mapsto A0$, $1 + 1\tau \mapsto B$

---

settings: $q = |\tau|^2 = 2$, $p = 0$, $\tau = 0 + 1.414i$, $w = 3$, $|\cdot|^2 \leq 2.792$
digit set $\mathcal{D}$: $0 = 0 + 0\tau$, $A = 1 + 0\tau$, $B = 1 + 1\tau$, $C = -1 + 0\tau$, $D = -1 - 1\tau$
mapping $z \mapsto \boldsymbol{\xi}$: $0 - 1\tau \mapsto C0$, $-1 + 0\tau \mapsto C$, $0 + 0\tau \mapsto 0$, $1 + 0\tau \mapsto A$, $0 + 1\tau \mapsto A0$

---

settings: $q = |\tau|^2 = 2$, $p = 1$, $\tau = 0.5 + 1.323i$, $w = 3$, $|\cdot|^2 \leq 2.792$
digit set $\mathcal{D}$: $0 = 0 + 0\tau$, $A = 1 + 0\tau$, $B = -1 + 1\tau$, $C = -1 + 0\tau$, $D = 1 - 1\tau$
mapping $z \mapsto \boldsymbol{\xi}$: $0 - 1\tau \mapsto C0$, $1 - 1\tau \mapsto D$, $-1 + 0\tau \mapsto C$, $0 + 0\tau \mapsto 0$, $1 + 0\tau \mapsto A$, $-1 + 1\tau \mapsto B$, $0 + 1\tau \mapsto A0$

---

settings: $q = |\tau|^2 = 2$, $p = 2$, $\tau = 1 + 1i$, $w = 3$, $|\cdot|^2 \leq 2.792$
digit set $\mathcal{D}$: $0 = 0 + 0\tau$, $A = 1 + 0\tau$, $B = -1 + 1\tau$, $C = -1 + 0\tau$, $D = 1 - 1\tau$
mapping $z \mapsto \boldsymbol{\xi}$: $0 - 1\tau \mapsto C0$, $1 - 1\tau \mapsto D$, $2 - 1\tau \mapsto D0$, $-1 + 0\tau \mapsto C$, $0 + 0\tau \mapsto 0$, $1 + 0\tau \mapsto A$, $-2 + 1\tau \mapsto B0$, $-1 + 1\tau \mapsto B$, $0 + 1\tau \mapsto A0$

---

settings: $q = |\tau|^2 = 3$, $p = -3$, $\tau = -1.5 + 0.866i$, $w = 2$, $|\cdot|^2 \leq 3.938$
digit set $\mathcal{D}$: $0 = 0 + 0\tau$, $A = 1 + 0\tau$, $B = 2 + 1\tau$, $C = 1 + 1\tau$, $D = -1 + 0\tau$, $E = -2 - 1\tau$, $F = -1 - 1\tau$
mapping $z \mapsto \boldsymbol{\xi}$: $-3 - 2\tau \mapsto C0$, $-3 - 1\tau \mapsto B0$, $-2 - 1\tau \mapsto E$, $-1 - 1\tau \mapsto F$, $0 - 1\tau \mapsto D0$, $-1 + 0\tau \mapsto D$, $0 + 0\tau \mapsto 0$, $1 + 0\tau \mapsto A$, $0 + 1\tau \mapsto A0$, $1 + 1\tau \mapsto C$, $2 + 1\tau \mapsto B$, $3 + 1\tau \mapsto E0$, $3 + 2\tau \mapsto F0$

---

settings: $q = |\tau|^2 = 3$, $p = -2$, $\tau = -1 + 1.414i$, $w = 2$, $|\cdot|^2 \leq 3.938$
digit set $\mathcal{D}$: $0 = 0 + 0\tau$, $A = 1 + 0\tau$, $B = 1 + 1\tau$, $C = -1 + 0\tau$, $D = -2 + 0\tau$, $E = -1 - 1\tau$, $F = 2 + 0\tau$
mapping $z \mapsto \boldsymbol{\xi}$: $-2 - 1\tau \mapsto A0B$, $-1 - 1\tau \mapsto E$, $0 - 1\tau \mapsto C0$, $-1 + 0\tau \mapsto C$, $0 + 0\tau \mapsto 0$, $1 + 0\tau \mapsto A$, $0 + 1\tau \mapsto A0$, $1 + 1\tau \mapsto B$, $2 + 1\tau \mapsto C0E$

---

settings: $q = |\tau|^2 = 3$, $p = -1$, $\tau = -0.5 + 1.658i$, $w = 2$, $|\cdot|^2 \leq 3.938$
digit set $\mathcal{D}$: $0 = 0 + 0\tau$, $A = 1 + 0\tau$, $B = 1 + 1\tau$, $C = -1 + 1\tau$, $D = -1 + 0\tau$, $E = -1 - 1\tau$, $F = 1 - 1\tau$
mapping $z \mapsto \boldsymbol{\xi}$: $-1 - 1\tau \mapsto E$, $0 - 1\tau \mapsto D0$, $-1 + 0\tau \mapsto D$, $0 + 0\tau \mapsto 0$, $1 + 0\tau \mapsto A$, $0 + 1\tau \mapsto A0$, $1 + 1\tau \mapsto B$

---

settings: $q = |\tau|^2 = 3$, $p = 0$, $\tau = 0 + 1.732i$, $w = 2$, $|\cdot|^2 \leq 3.938$
digit set $\mathcal{D}$: $0 = 0 + 0\tau$, $A = 1 + 0\tau$, $B = 1 + 1\tau$, $C = -1 + 1\tau$, $D = -1 + 0\tau$, $E = -1 - 1\tau$, $F = 1 - 1\tau$
mapping $z \mapsto \boldsymbol{\xi}$: $0 - 1\tau \mapsto D0$, $-1 + 0\tau \mapsto D$, $0 + 0\tau \mapsto 0$, $1 + 0\tau \mapsto A$, $0 + 1\tau \mapsto A0$

---

settings: $q = |\tau|^2 = 3$, $p = 1$, $\tau = 0.5 + 1.658i$, $w = 2$, $|\cdot|^2 \leq 3.938$
digit set $\mathcal{D}$: $0 = 0 + 0\tau$, $A = 1 + 0\tau$, $B = 1 + 1\tau$, $C = -1 + 1\tau$, $D = -1 + 0\tau$, $E = -1 - 1\tau$, $F = 1 - 1\tau$

---

mapping $z \mapsto \boldsymbol{\xi}$: $0-1\tau \mapsto D0$, $1-1\tau \mapsto F$, $-1+0\tau \mapsto D$, $0+0\tau \mapsto 0$, $1+0\tau \mapsto A$, $-1+1\tau \mapsto C$, $0+1\tau \mapsto A0$

---

settings: $q = |\tau|^2 = 3$, $p = 2$, $\tau = 1 + 1.414i$, $w = 2$, $|\cdot|^2 \leq 3.938$
digit set $\boldsymbol{\mathcal{D}}$: $0 = 0+0\tau$, $A = 2+0\tau$, $B = -1+1\tau$, $C = -1+0\tau$, $D = -2+0\tau$, $E = 1-1\tau$, $F = 1+0\tau$
mapping $z \mapsto \boldsymbol{\xi}$: $0-1\tau \mapsto C0$, $1-1\tau \mapsto E$, $2-1\tau \mapsto C0B$, $-1+0\tau \mapsto C$, $0+0\tau \mapsto 0$, $1+0\tau \mapsto F$, $-2+1\tau \mapsto F0E$, $-1+1\tau \mapsto B$, $0+1\tau \mapsto F0$

---

settings: $q = |\tau|^2 = 3$, $p = 3$, $\tau = 1.5 + 0.866i$, $w = 2$, $|\cdot|^2 \leq 3.938$
digit set $\boldsymbol{\mathcal{D}}$: $0 = 0+0\tau$, $A = -1+1\tau$, $B = -2+1\tau$, $C = -1+0\tau$, $D = 1-1\tau$, $E = 2-1\tau$, $F = 1+0\tau$
mapping $z \mapsto \boldsymbol{\xi}$: $3-2\tau \mapsto D0$, $0-1\tau \mapsto C0$, $1-1\tau \mapsto D$, $2-1\tau \mapsto E$, $3-1\tau \mapsto E0$, $-1+0\tau \mapsto C$, $0+0\tau \mapsto 0$, $1+0\tau \mapsto F$, $-3+1\tau \mapsto B0$, $-2+1\tau \mapsto B$, $-1+1\tau \mapsto A$, $0+1\tau \mapsto F0$, $-3+2\tau \mapsto A0$

---

settings: $q = |\tau|^2 = 4$, $p = -3$, $\tau = -1.5 + 1.323i$, $w = 2$, $|\cdot|^2 \leq 4.148$
digit set $\boldsymbol{\mathcal{D}}$: $0 = 0+0\tau$, $A = 1+0\tau$, $B = 2+0\tau$, $C = 3+1\tau$, $D = 2+1\tau$, $E = 1+1\tau$, $F = -1+1\tau$, $G = -1+0\tau$, $H = -2+0\tau$, $I = -3-1\tau$, $J = -2-1\tau$, $K = -1-1\tau$, $L = 1-1\tau$
mapping $z \mapsto \boldsymbol{\xi}$: $-3-1\tau \mapsto I$, $-2-1\tau \mapsto J$, $-1-1\tau \mapsto K$, $0-1\tau \mapsto G0$, $-2+0\tau \mapsto H$, $-1+0\tau \mapsto G$, $0+0\tau \mapsto 0$, $1+0\tau \mapsto A$, $2+0\tau \mapsto B$, $0+1\tau \mapsto A0$, $1+1\tau \mapsto E$, $2+1\tau \mapsto D$, $3+1\tau \mapsto C$

---

settings: $q = |\tau|^2 = 4$, $p = -2$, $\tau = -1 + 1.732i$, $w = 2$, $|\cdot|^2 \leq 4.148$
digit set $\boldsymbol{\mathcal{D}}$: $0 = 0+0\tau$, $A = 1+0\tau$, $B = 2+0\tau$, $C = 3+0\tau$, $D = 1+1\tau$, $E = -1+1\tau$, $F = -2+1\tau$, $G = -1+0\tau$, $H = -2+0\tau$, $I = -3+0\tau$, $J = -2-1\tau$, $K = -1-1\tau$, $L = 1-1\tau$
mapping $z \mapsto \boldsymbol{\xi}$: $-2-1\tau \mapsto J$, $-1-1\tau \mapsto K$, $0-1\tau \mapsto G0$, $-2+0\tau \mapsto H$, $-1+0\tau \mapsto G$, $0+0\tau \mapsto 0$, $1+0\tau \mapsto A$, $2+0\tau \mapsto B$, $0+1\tau \mapsto A0$, $1+1\tau \mapsto D$, $2+1\tau \mapsto G0J$

---

settings: $q = |\tau|^2 = 4$, $p = -1$, $\tau = -0.5 + 1.936i$, $w = 2$, $|\cdot|^2 \leq 4.148$
digit set $\boldsymbol{\mathcal{D}}$: $0 = 0+0\tau$, $A = 1+0\tau$, $B = 2+0\tau$, $C = 1+1\tau$, $D = 1+2\tau$, $E = -1+1\tau$, $F = -2+1\tau$, $G = -1+0\tau$, $H = -2+0\tau$, $I = -1-1\tau$, $J = -1-2\tau$, $K = 1-1\tau$, $L = 2-1\tau$
mapping $z \mapsto \boldsymbol{\xi}$: $-1-1\tau \mapsto I$, $0-1\tau \mapsto G0$, $-2+0\tau \mapsto H$, $-1+0\tau \mapsto G$, $0+0\tau \mapsto 0$, $1+0\tau \mapsto A$, $2+0\tau \mapsto B$, $0+1\tau \mapsto A0$, $1+1\tau \mapsto C$

---

settings: $q = |\tau|^2 = 4$, $p = 0$, $\tau = 0 + 2i$, $w = 2$, $|\cdot|^2 \leq 4.148$
digit set $\boldsymbol{\mathcal{D}}$: $0 = 0+0\tau$, $A = 1+0\tau$, $B = 2+0\tau$, $C = 1+1\tau$, $D = 2+2\tau$, $E = 1+2\tau$, $F = -1+1\tau$, $G = -2+1\tau$, $H = -1+0\tau$, $I = -1-1\tau$, $J = -1-2\tau$, $K = 1-1\tau$, $L = 2-1\tau$
mapping $z \mapsto \boldsymbol{\xi}$: $0-1\tau \mapsto H0$, $-2+0\tau \mapsto A0B$, $-1+0\tau \mapsto H$, $0+0\tau \mapsto 0$, $1+0\tau \mapsto A$, $2+0\tau \mapsto B$, $0+1\tau \mapsto A0$

---

settings: $q = |\tau|^2 = 4$, $p = 1$, $\tau = 0.5 + 1.936i$, $w = 2$, $|\cdot|^2 \leq 4.148$
digit set $\boldsymbol{\mathcal{D}}$: $0 = 0+0\tau$, $A = 1+0\tau$, $B = 2+1\tau$, $C = 1+1\tau$, $D = -1+2\tau$, $E = -1+1\tau$, $F = -1+0\tau$, $G = -2+0\tau$, $H = -2-1\tau$, $I = -1-1\tau$, $J = 1-2\tau$, $K = 1-1\tau$, $L = 2+0\tau$
mapping $z \mapsto \boldsymbol{\xi}$: $0-1\tau \mapsto F0$, $1-1\tau \mapsto K$, $-2+0\tau \mapsto G$, $-1+0\tau \mapsto F$, $0+0\tau \mapsto 0$, $1+0\tau \mapsto A$, $2+0\tau \mapsto L$, $-1+1\tau \mapsto E$, $0+1\tau \mapsto A0$

---

settings: $q = |\tau|^2 = 4$, $p = 2$, $\tau = 1 + 1.732i$, $w = 2$, $|\cdot|^2 \leq 4.148$
digit set $\boldsymbol{\mathcal{D}}$: $0 = 0+0\tau$, $A = 2+0\tau$, $B = 1+0\tau$, $C = 1+1\tau$, $D = -1+1\tau$, $E = -1+0\tau$, $F = -2+0\tau$, $G = -3+0\tau$, $H = -2-1\tau$, $I = -1-1\tau$, $J = 1-1\tau$, $K = 2-1\tau$, $L = 3+0\tau$

mapping $z \mapsto \boldsymbol{\xi}$: $0-1\tau \mapsto E0$, $1-1\tau \mapsto J$, $2-1\tau \mapsto K$, $-2+0\tau \mapsto F$, $-1+0\tau \mapsto E$, $0+0\tau \mapsto 0$, $1+0\tau \mapsto B$, $2+0\tau \mapsto A$, $-2+1\tau \mapsto B0K$, $-1+1\tau \mapsto D$, $0+1\tau \mapsto B0$
settings: $q = \lvert\tau\rvert^2 = 4$, $p = 3$, $\tau = 1.5 + 1.323i$, $w = 2$, $\lvert\cdot\rvert^2 \leq 4.148$ **digit set $\boldsymbol{\mathcal{D}}$:** $0 = 0+0\tau$, $A = 1+0\tau$, $B = 2+0\tau$, $C = 1+1\tau$, $D = -1+1\tau$, $E = -2+1\tau$, $F = -3+1\tau$, $G = -1+0\tau$, $H = -2+0\tau$, $I = -1-1\tau$, $J = 1-1\tau$, $K = 2-1\tau$, $L = 3-1\tau$ **mapping** $z \mapsto \boldsymbol{\xi}$: $0-1\tau \mapsto G0$, $1-1\tau \mapsto J$, $2-1\tau \mapsto K$, $3-1\tau \mapsto L$, $-2+0\tau \mapsto H$, $-1+0\tau \mapsto G$, $0+0\tau \mapsto 0$, $1+0\tau \mapsto A$, $2+0\tau \mapsto B$, $-3+1\tau \mapsto F$, $-2+1\tau \mapsto E$, $-1+1\tau \mapsto D$, $0+1\tau \mapsto A0$

Bibliography

[1] D. W. Ash, I. F. Blake, and S. A. Vanstone, *Low complexity normal bases*, Discrete Appl. Math. **25** (1989), no. 3, 191–210.
(Cited on page 21.)

[2] F. Aurenhammer, *Voronoi diagrams — a survey of a fundamental geometric data structure*, ACM Comput. Surv. **23** (1991), no. 3, 345–405.
(Cited on page 41.)

[3] 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.
(Cited on page 13.)

[4] R. M. Avanzi, C. Heuberger, and H. 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.
(Cited on pages 23 and 26.)

[5] ———, *Scalar multiplication on Koblitz curves. Using the Frobenius endomorphism and its combination with point halving: Extensions and mathematical analysis*, Algorithmica **46** (2006), 249–270.
(Cited on pages 23, 24, and 26.)

[6] ———, *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.
(Cited on page 25.)

[7] ———, *Redundant τ-adic expansions I: Non-adjacent digit sets and their applications to scalar multiplication*, Des. Codes Cryptogr. (2010), DOI 10.1007/s10623-010-9396-6, earlier version available at Report 2008-7, TU Graz, `http://www.math.tugraz.at/fosp/pdfs/tugraz_100.pdf` and as Cryptology ePrint Archive, Report 2008/148, `http://eprint.iacr.org/`.
(Cited on pages 24, 25, and 26.)

[8] R. Avanzi, C. Heuberger, and H. Prodinger, *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/`.
(Cited on pages 27, 28, 38, and 39.)

[9] G. Avoine, J. Monnerat, and T. Peyrin, *Advances in alternative non-adjacent form representations*, Progress in cryptology—INDOCRYPT 2004, Lecture Notes in Comput. Sci., vol. 3348, Springer, Berlin, 2004, pp. 260–274.
(Cited on pages 16 and 17.)

[10] R. Balasubramanian and N. Koblitz, *The improbability than an elliptic curve has subexponential discrete log problem under the menezes–okamoto–vanstone algorithm*, J. Cryptology **11** (1998), 141–145.
(Cited on page 8.)

[11] M. Barnsley, *Fractals everywhere*, Academic Press, Inc, 1988.
(Cited on pages 63 and 64.)

[12] I. Blake, K. Murty, and G. Xu, *Nonadjacent radix-τ expansions of integers in Euclidean imaginary quadratic number fields*, Canad. J. Math. **60** (2008), no. 6, 1267–1282.
(Cited on pages 25, 28, 29, 30, and 60.)

[13] I. Blake, G. Seroussi, and N. Smart, *Elliptic curves in cryptography*, London Mathematical Society Lecture Note Series, vol. 265, Cambridge University Press, 1999.
(Cited on page 12.)

[14] I. F. Blake, V. K. Murty, and G. Xu, *A note on window τ-NAF algorithm*, Inform. Process. Lett. **95** (2005), 496–502.
(Cited on pages 25 and 60.)

[15] I. F. Blake, V. Kumar Murty, and G. Xu, *Efficient algorithms for Koblitz curves over fields of characteristic three*, J. Discrete Algorithms **3** (2005), no. 1, 113–124.
(Cited on pages 28, 45, and 60.)

[16] H. Cohen, *Analysis of the sliding window powering algorithm*, J. Cryptology **18** (2005), no. 1, 63–76.
(Cited on page 13.)

[17] H. Delange, *Sur la fonction sommatoire de la fonction "somme des chiffres"*, Enseignement Math. (2) **21** (1975), 31–47.
(Cited on pages 19 and 79.)

[18] W. Diffie and M. E. Hellman, *New directions in cryptography*, IEEE Transactions on Information Theory **22** (1976), no. 6, 644–654.
(Cited on page 7.)

[19] G. A. Edgar, *Measure, topology, and fractal geometry*, second ed., Undergraduate Texts in Mathematics, Springer-Verlag, New York, 2008.
(Cited on pages 46, 63, 64, 66, and 67.)

[20] P. Flajolet and L. Ramshaw, *A note on Gray code and odd-even merge*, SIAM J. Comput. **9** (1980), 142–158.
(Cited on pages 31, 32, 34, and 35.)

[21] K. Fong, D. Hankerson, J. López, and A. Menezes, *Field inversion and point halving revisited*, IEEE Transactions on Computers **53** (2004), 1047–1059.
(Cited on page 24.)

[22] D. M. Gordon, *A survey of fast exponentiation methods*, J. Algorithms **27** (1998), 129–146.
(Cited on pages 11, 12, 23, and 39.)

[23] P. J. Grabner and C. Heuberger, *On the number of optimal base 2 representations of integers*, Des. Codes Cryptogr. **40** (2006), no. 1, 25–39.
(Cited on pages 11 and 12.)

[24] 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.
(Cited on pages 10 and 20.)

[25] R. L. Graham, D. E. Knuth, and O. Patashnik, *Concrete mathematics. A foundation for computer science*, second ed., Addison-Wesley, 1994.
(Cited on pages 52 and 78.)

[26] C. Heuberger, *Redundant τ-adic expansions II: Non-optimality and chaotic behaviour*, Math. Comput. Sci. **3** (2010), 141–157.
(Cited on pages 26 and 27.)

[27] C. Heuberger, R. Katti, H. Prodinger, and X. Ruan, *The alternating greedy expansion and applications to left-to-right algorithms in cryptography*, Theoret. Comput. Sci. **341** (2005), 55–72.
(Cited on page 10.)

[28] C. Heuberger and J. A. Muir, *Minimal weight and colexicographically minimal integer representations*, J. Math. Cryptol. **1** (2007), 297–328.
(Cited on page 10.)

[29] ———, *Unbalanced digit sets and the closest choice strategy for minimal weight integer representations*, Des. Codes Cryptogr. **52** (2009), 185–208.
(Cited on page 10.)

[30] C. Heuberger and H. Prodinger, *On minimal expansions in redundant number systems: Algorithms and quantitative analysis*, Computing **66** (2001), 377–393.
(Cited on pages 11 and 20.)

[31] ———, *Analysis of alternative digit sets for nonadjacent representations*, Monatsh. Math. **147** (2006), 219–248.
(Cited on pages 12, 16, 18, 19, 20, and 65.)

[32] H.-K. Hwang, *On convergence rates in the central limit theorems for combinatorial structures*, European J. Combin. **19** (1998), 329–343.
(Cited on page 50.)

[33] IEEE Std 1363-2000, *IEEE standard specifications for public-key cryptography*, IEEE Computer Society, August 29 2000.
(Cited on page 23.)

[34] J. Jedwab and C. J. Mitchell, *Minimum weight modified signed-digit representations and fast exponentiation*, Electron. Lett. **25** (1989), 1171–1172.
(Cited on pages 10 and 11.)

[35] M. Joye and S.-M. Yen, *Optimal left-to-right binary signed digit recoding*, IEEE Transactions on Computers **49** (2000), no. 7, 740–748.
(Cited on page 13.)

[36] E. W. Knudsen, *Elliptic Scalar Multiplication Using Point Halving*, Advances in Cryptology – Asiacrypt'99, Lecture Notes in Comput. Sci., vol. 1716, Springer-Verlag, Berlin, 1999, pp. 135–149.
(Cited on page 24.)

[37] D. E. Knuth, *Seminumerical algorithms*, third ed., The Art of Computer Programming, vol. 2, Addison-Wesley, 1998.
(Cited on page 9.)

[38] N. Koblitz, *Elliptic curve cryptosystems*, Math. Comp. **48** (1987), 203–209.
(Cited on page 8.)

[39] 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.
(Cited on pages 21 and 24.)

[40] ———, *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.
(Cited on pages 21, 27, 28, 38, 60, and 64.)

[41] N. Koblitz, *A course in number theory and cryptography*, 2nd ed., Graduate Texts in Mathematics, Springer, 1994.
(Cited on pages 5, 6, 8, and 20.)

[42] D. W. Matula, *Basic digit sets for radix representation*, J. Assoc. Comput. Mach. **29** (1982), no. 4, 1131–1143.
(Cited on pages 51 and 60.)

[43] W. Meier and O. Staffelbach, *Efficient multiplication on certain nonsupersingular elliptic curves*, Advances in cryptology—CRYPTO '92 (Santa Barbara, CA, 1992), Lecture Notes in Comput. Sci., vol. 740, Springer, Berlin, 1993, pp. 333–344.
(Cited on pages 21, 22, and 27.)

[44] A. J. Menezes, P. C. van Oorschot, and S. A. Vanstone, *Handbook of applied cryptography*, CRC Press, Boca Raton, FL, 1997, With a foreword by Ronald L. Rivest.
(Cited on page 23.)

[45] A. Menezes, S. Vanstone, and T. Okamoto, *Reducing elliptic curve logarithms to logarithms in a finite field*, STOC '91: Proceedings of the twenty-third annual ACM symposium on Theory of computing (New York, NY, USA), ACM, 1991, pp. 80–89.
(Cited on pages 8 and 28.)

[46] V. S. Miller, *Use of elliptic curves in cryptography*, Advances in cryptology – CRYPTO '85, Lecture Notes in Comput. Sci., vol. 218, Springer-Verlag, Berlin, 1986, pp. 417–426.
(Cited on page 8.)

[47] A. Miyaji, T. Ono, and H. Cohen, *Efficient elliptic curve exponentiation*, Information and communications security. 1st international conference, ICICS '97, Beijing, China, November 11–14, 1997. Proceedings (Y. e. a. Han, ed.), LNCS, vol. 1334, Springer-Verlag, 1997, pp. 282–290.
(Cited on page 12.)

[48] B. Möller, *Improved techniques for fast exponentiation*, Information Security and Cryptology — ICISC 2002 (P. J. Lee and C. H. Lim, eds.), LNCS, vol. 2587, Springer-Verlag, 2003, pp. 298–312.
(Cited on page 13.)

[49] 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.
(Cited on page 10.)

[50] J. A. Muir and D. R. Stinson, *Alternative digit sets for nonadjacent representations*, Selected areas in cryptography, Lecture Notes in Comput. Sci., vol. 3006, Springer, Berlin, 2004, pp. 306–319.
(Cited on pages 14, 17, 18, and 47.)

[51] ———, *Alternative digit sets for nonadjacent representations*, SIAM J. Discrete Math. **19** (2005), 165–191.
(Cited on pages 14, 15, 16, 17, and 18.)

[52] ———, *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.
(Cited on page 13.)

[53] ———, *Minimality and other properties of the width-w nonadjacent form*, Math. Comp. **75** (2006), 369–384.
(Cited on pages 12 and 13.)

[54] National Institute of Standards and Technology, *Digital signature standard*, FIPS Publication, vol. 186-2, February 2000.
(Cited on page 23.)

[55] B. Phillips and N. Burgess, *Minimal weight digit set conversions*, IEEE Trans. Comput. **53** (2004), 666–677.
(Cited on page 9.)

[56] H. Prodinger, *On binary representations of integers with digits* $-1, 0, 1$, Integers **0** (2000), A08, available at `http://www.integers-ejcnt.org/vol0.html`.
(Cited on pages 11 and 20.)

[57] J. Proos, *Joint sparse forms and generating zero columns when combing*, Tech. Report CORR 2003-23, Centre for Applied Cryptographic Research, University of Waterloo, 2003, available at `http://www.cacr.math.uwaterloo.ca/techreports/2003/corr2003-23.ps`.
(Cited on page 10.)

[58] G. W. Reitwiesner, *Binary arithmetic*, Advances in computers, vol. 1, Academic Press, New York, 1960, pp. 231–308.
(Cited on pages 9, 10, 11, 13, 20, and 23.)

[59] R. Schroeppel, *Elliptic curve point ambiguity resolution apparatus and method*, International Application Number PCT/US00/31014, filed 9 November 2000.
(Cited on page 24.)

[60] R. Schroeppel, *Point halving wins big*, Talk at the ECC 2001 Workshop, October 29–31, 2001, University of Waterloo, Ontario, Canada.
(Cited on page 24.)

[61] J. Shallit, *A primer on balanced binary representations*, Draft, 1992.
(Cited on pages 10 and 11.)

[62] J. H. Silverman, *The arithmetic of elliptic curves*, Graduate Texts in Mathematics, vol. 106, Springer, New York, 1992.
(Cited on pages 5, 6, 7, and 20.)

[63] 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.
(Cited on pages 9, 10, 12, 21, and 45.)

[64] ———, *Efficient arithmetic on Koblitz curves*, Des. Codes Cryptogr. **19** (2000), 195–249.
(Cited on pages 9, 10, 12, 21, 22, 23, 24, 27, 28, 45, and 60.)

[65] ———, *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`.
(Cited on pages 10 and 14.)

[66] J. M. Thuswaldner, *Summatory functions of digital sums occurring in cryptography*, Period. Math. Hungar. **38** (1999), 111–130.
(Cited on page 10.)