Attila Kovács

# Radix expansion in lattices
# PhD thesis

Eötvös Loránd University

Supervisor: Imre Kátai

Program: Informatics, Head: J. Demetrovics

Budapest, 2001

# Table of contents

# Preface

*"Number theory is a building of rare beauty and harmony."*
*— D. Hilbert*

The development of modern science and technology always strongly depended on the development of adequate methods for representing integers and doing integer arithmetic. The history of number representation is a fascinating story, since it parallels the development of civilization itself. See D. E. Knuth [60] for further details.

In this work number expansions in lattices are analyzed. Chapter one contains the concept of number systems, dynamic properties of expansions and estimates for the length of expansions. Chapter two deals with classification of expansions. An effective algorithm is presented. Chapter three contains methods for constructing number systems of several types. The connection between number expansion in lattices and number expansion in the ring of integers of a given algebraic number field is discussed, canonical, polygonal and simultaneous radix systems are analyzed. Generalized binary number systems are also treated. For general radix systems a sufficient condition is proved to be able to construct number systems. In chapter four the number, location and structural properties of periodic elements are described for radix systems of imaginary quadratic fields using canonical digit sets. Chapter five deals with the geometry of expansions. Some properties of the set of numbers with zero integer part are analyzed and the notion of self-affine lattice tilings are discussed. These tilings arise in image processing, computer vision and many other topics of mathematics and physics [107]. The boundary of the tiles often have non-integral Hausdorff dimension. Methods for estimating, or in some cases computing this dimension are presented, an example is also

given. In chapter six after a short summary some open problems and further directions are mentioned.

## Acknowledgments

*'What does your Master teach?'*
*asked a visitor.*
*'Nothing,' said the disciple.*
*'Then why does he give discourses?'*
*He only points the way — he teaches nothing.'*
*— Antony de Mello, One Minute Wisdom*

I would like to acknowledge the assistance of several people.

I wish to express my special thanks to my supervisor Prof. Imre Kátai. He answered all my questions and he has always disposed to explain the mysteries and beauty of number theory.

I would like to thank all my colleagues and students. I acknowledge with great appreciation Prof. Antal Járai for his useful remarks and suggestions.

Finally, my heartfelt thanks go also to my wife for her patient.

# Chapter 1

# Expansion in lattices

*"There are two kinds of generalizations. One is cheap and the other is valuable. It is easy to generalize by diluting a little idea with a big terminology. It is much more difficult to prepare a refined and condensed extract from several good ingredients."*
*— Gy. Pólya*

A *lattice* in $\mathbb{R}^k$ is the set of all integer combinations of $k$ linearly independent vectors. Let $\Lambda$ be a lattice, which can be viewed either geometrically as a set of points in a Euclidean space, or algebraically, a $\mathbb{Z}$-module or as a finitely generated free Abelian group. Let $M : \Lambda \to \Lambda$ be a group endomorphism and let $D$ be a finite subset of $\Lambda$ containing 0. Clearly, $M$ can be taken as an arbitrary square non-singular matrix. Moreover, if the basis of $M$ is chosen in $\Lambda$ then $M$ is an integer matrix.

## 1.1 Concept of number systems

The triple $(\Lambda, M, D)$ is called a *number system* (or having the unique representation property) if every element $n$ of $\Lambda$ has a unique finite representation of the form

$$n = a_0 + Ma_1 + M^2a_2 + \ldots + M^la_l = (a_la_{l-1}\ldots a_1a_0)_M, \tag{1.1}$$

where $a_i \in D$. The endomorphism $M$ is called the *base* or *radix*, $D$ is the *digit set*. The *length of expansion* of $n$ in (1.1) is $l + 1$.

One of the main problems concerning radix representation is to give conditions under which $(\Lambda, M, D)$ is a number system. For a lattice $\Lambda$, both $\Lambda$ and $M\Lambda$ are Abelian groups under addition. The order of the factor group $\Lambda/M\Lambda$ is $t = |\det(M)|$. Let $A_j$, $(j = 1 \dots t)$ denote the cosets of this group. If $z_1, z_2 \in A_j$, i.e. they are in the same residue class then we will say that they are congruent modulo $M$ and we will denote this by $z_1 \equiv z_2 \pmod{M}$.

The following result was known and used by I. Kátai and co-workers as well as by W. Gilbert in algebraic number fields (see section 3.2.1). Moreover, it can be found implicitly in A. Vince's paper [106]. Recall that a linear map is called expansive if all eigenvalues have modulus greater than one.

**Assertion 1.** *(Necessary conditions for the number system property)*
*If $(\Lambda, M, D)$ is a number system then*
*(a) $D$ must be a complete set of residues modulo $M$,*
*(b) $M$ must be expansive and*
*(c)* $\det(I - M) \neq \pm 1$.
PROOF: Concerning (a) if $z \in \Lambda$ is represented by $(a_m a_{m-1} \dots a_1 a_0)_M$ then $z \equiv a_0 \pmod{M}$. Hence the digit set $D$ must contain a complete residue system modulo $M$. Now suppose that two digits $c$ and $d$ are congruent modulo $M$. Then $c - d = Me$ for some $e \in \Lambda$. Represent $e$ by $(a_l a_{l-1} \dots a_1 a_0)_M$ so that

$$(c)_M = c = Me + d = (a_l a_{l-1} \dots a_1 a_0 d)_M.$$

Hence $c \in \Lambda$ has two different representations, which is a contradiction. Statement (b) was proved in [106]. Concerning (c) first observe that $(I - M^n)$ is nonsingular for any positive integer $n$. Otherwise 1 would be an eigenvalue of $M^n$, hence $M$ would have an eigenvalue of modulus one. Second, it is also clear that if $(\Lambda, M, D)$ is a number system then there is not any $\pi \in \Lambda$ and $l \in \mathbb{N}$ for which $\pi = a_0 + Ma_1 + \dots + M^{l-1}a_{l-1} + M^l\pi$, where $a_i \in D$. In other words $(I - M^l)^{-1}(a_0 + Ma_1 + \dots + M^{l-1}a_{l-1}) \in \Lambda$ can never be happen. But if $\det(I - M) = \pm 1$ then $(I - M)\Lambda = (I - M)^{-1}\Lambda = \Lambda$, which is a contradiction. □

COROLLARY. *Suppose that an arbitrary $z \in \Lambda$ has a finite expansion of form (1.1). Then the uniqueness of the representation follows from the assumption that any two elements of $D$ are incongruent modulo $M$.*

If for a given triple $(\Lambda, M, D)$ the conditions (a) and (b) in Assertion 1 hold then we say that it is a *radix system*. Assertion 1(c) explains why it is

impossible to find appropriate digit sets for the matrices $\left(\begin{smallmatrix}1 & 1\\ 1 & m\end{smallmatrix}\right)$, $\left(\begin{smallmatrix}0 & -m\\ 1 & m\end{smallmatrix}\right)$ or for the matrix $2I + S$, where $S$ is strictly upper (or lower) triangular.

**Assertion 2.** *(Sufficient condition for the number system property)*
*If for a given radix system $(\Lambda, M, D)$ (a) there is a basis for the lattice $\Lambda$ for which all the basis vectors have some finite representation and (b) all the elements of the set $D \pm D$ have some expansion of form $a_0 + M^j a_j$ $(a_0, a_j \in D,\ j \in \mathbb{N})$ then $(\Lambda, M, D)$ is a number system.*

PROOF: By the corollary of Assertion 1 it is enough to show that every lattice point $z$ has a finite representation. Let us denote the basis vectors — which have all finite representations — by $b_1, b_2, \ldots, b_k$. Then $z = \sum \alpha_i b_i$ for some $\alpha_i \in \mathbb{Z}$. The proof is by induction of the number of summands $n$. The case $n = 1$ is obvious. By induction, assume that the sum of first $n-1$ terms has the form $x = (a_l a_{l-1} \ldots a_0)$. It is clear that if we add a lattice point $d \in \pm D$ to $x$ then $x + d \in \Lambda$ and the length of expansion of $x + d$ is less then or equal to $l + s + 1$ where $s$ is the length of the longest expansion in $D \pm D$. In the same way, if we add an arbitrary basis vector $b_i$ to $x$ then — adding digit by digit — $x + b_i$ must have a bounded length of expansion, therefore a finite representation. □

The theorem is a simple generalization of A. Vince's theorem [106]. Unfortunately, in order to decide the number system property for a given triple $(\Lambda, M, D)$ this theorem can be applied in very few cases. Fortunately, as we will see in section 3.4, there is a sufficient condition for the base $M$, in which case the unique representation property holds for some digit set $D$. Moreover, the digit set can easily be constructed.

**Assertion 3.** *(Equivalence of number systems)*
*Let the matrices $M_1$ and $M_2$ are similar via the matrix $Q$. Then the number system property for $(\Lambda, M_1, D)$ and for $(Q\Lambda, M_2, QD)$ holds at exactly the same time.*

PROOF: $D$ is a full residue system modulo $M_1$ in $\Lambda$ iff $QD$ is a full residue system modulo $M_2$ in $Q\Lambda$. Moreover, $z = \sum_{i=0}^{l} M_1^i a_i$ iff $Qz = \sum_{i=0}^{l} Q M_1^i a_i = \sum_{i=0}^{l} M_2^i (Q a_i)$ $(a_i \in D)$. □

This equivalence is essentially a change of basis for the matrix $M_1$, therefore there exist similar matrices — bases — in several forms. Moreover, if we change the basis in $\Lambda$, a similar integer matrix $M_2 : \mathbb{Z}^k \to \mathbb{Z}^k$ is obtained. Hence the number system property can be examined without loss of generality on the cubic lattice $\mathbb{Z}^k$. This has a computational advantage, since $M_2$ and its characteristic polynomial have integer coefficients (see also [106]).

## 1.2 Dynamic of expansions

Further we analyze the expansions in the radix system $(\Lambda, M, D)$. The system $(\Lambda, M, D)$ can be used to represent all the lattice points in $\Lambda$ even if it is not a number system. Clearly, for each $\gamma \in \Lambda$ there exist a unique $a_j \in D$ such that $M \mid \gamma - a_j$. Let $\gamma_1 = M^{-1}(\gamma - a_j)$ and let us define the function $\Phi : \Lambda \to \Lambda$ by $\Phi(\gamma) = \gamma_1$. Let $\Phi^l$ denote the $l$-fold iterate of $\Phi$, $\Phi^0(\gamma) = \gamma_0$. The sequence of integer vectors $\Phi^j(z_0) = z_j$ $(j = 0, 1, 2, \dots)$ is called the *path of the dynamical system* generated by $\Phi$. It is also called the *orbit* of $z_0$ generated by $\Phi$.[1] Since the spectral radius $\rho(M^{-1}) < 1$ therefore there exists a norm on $\mathbb{R}^k$ such that for the corresponding operator norm

$$\|M^{-1}\| = \sup_{\|x\| \le 1} \|M^{-1}x\| \tag{1.2}$$

the inequality $\|M^{-1}\| < 1$ holds [43]. Throughout this work $\|\,.\,\|$ denotes this vector and the appropriate operator norm. Let furthermore

$$K := \max_{b \in D} \|b\|, \quad r := \|M^{-1}\|, \quad L := \frac{Kr}{1-r}. \tag{1.3}$$

In virtue of (1.3) and the definition of $\Phi$ we get that

$$\|\Phi(z)\| = \|M^{-1}z - M^{-1}b\| \le r\|z\| + Kr.$$

Hence we obtain the following

**Lemma 1.** *(a) if* $\|z\| \le L$ *then* $\|\Phi(z)\| \le r(L+K) = L$*, (b) if* $\|z\| > L$ *then* $\|\Phi(z)\| \le r\|z\| + L(1-r) < \|z\|(r+1-r) = \|z\|$.

Since the inequality $\|x\| \le L$ holds only for finitely many lattice points $x$ therefore the path $z, \Phi(z), \Phi^2(z), \dots$ is ultimately periodic for all $z \in \Lambda$. The vector $p \in \Lambda$ is called *periodic* if there exist a $j \in \mathbb{N}$ such that $\Phi^j(p) = p$. The smallest such $j$ is the *length of period of* $p$ generated by $\Phi$. Let $\mathcal{P}$ denote the set of all periodic elements. Let $p \in \mathcal{P}$ be of period length $l$. The set of periodic elements $\{\Phi(p), \dots, \Phi^l(p)\}$ is called the *cycle* generated by $p$ and is denoted by $\mathcal{C}(p)$. Suppose that $p \in \mathcal{P}$. Then the *domain of attraction* of $p$ or *basin of attraction* of $p$ consists of all $z \in \Lambda$ for which there exists a $j \in \mathbb{N}$ such

[1] Historical remark: the function $\Phi$ was introduced by D. W. Matula [86] for rational integers in order to examine number systems. Somewhat later, independently, I. Kátai and W. Gilbert used it for constructing number systems in algebraic extensions [54, 33].

that $\Phi^j(z) = p$ and is denoted by $\mathcal{B}(p)$. Let $X \subseteq \mathcal{P}$. In a similar way, $\mathcal{B}(X)$ denotes all the $z \in \Lambda$ for which there exists a $j \in \mathbb{N}$ and $q \in X$ such that $\Phi^j(z) = q$. The function $\Phi$ defines a *discrete dynamic* on $\Lambda$ in the following way: let $\mathcal{G}(\mathcal{P})$ be the directed graph defined on the set $\mathcal{P}$ by drawing an edge from $p \in \mathcal{P}$ to $\Phi(p)$. Then $\mathcal{G}(\mathcal{P})$ is a disjoint union of directed cycles, where loops are allowed. We shall also call $\mathcal{G}(\mathcal{P})$ the *attractor set* of $\Lambda$ generated by $\Phi$.

The graph $\mathcal{G}(\mathcal{P})$ has the following properties [64]:

- $\mathcal{P}$ is finite;
- if $p \in \mathcal{P}$ then $\Phi(p) \in \mathcal{P}$;
- if $p \in \mathcal{P}$ then $\|p\| \leq L$;
- $p \in \mathcal{P}$ if and only if there is an $l > 0$ such that

$$p = a_0 + Ma_1 + \ldots + M^{l-1}a_{l-1} + M^l p, \quad a_j \in D; \tag{1.4}$$

- if $p_1, p_2 \in \mathcal{P}$ then either $\mathcal{C}(p_1) = \mathcal{C}(p_2)$ or $\mathcal{C}(p_1) \cap \mathcal{C}(p_2) = \emptyset$;
- if $p_1, p_2 \in \mathcal{P},\ p_1 \neq p_2$ and $\mathcal{C}(p_1) = \mathcal{C}(p_2)$ then their length of period are equal;
- $\mathcal{B}(\mathcal{P}) = \Lambda$;
- if $p_1, p_2 \in \mathcal{P}$ then $\mathcal{B}(p_1) = \mathcal{B}(p_2)$ if and only if $\mathcal{C}(p_1) = \mathcal{C}(p_2)$;
- if $p_1, p_2 \in \mathcal{P},\ \mathcal{C}(p_1) \neq \mathcal{C}(p_2)$ then $\mathcal{B}(p_1) \cap \mathcal{B}(p_2) = \emptyset$.

For a given radix system $(\Lambda, M, D)$ the computation of the graph $\mathcal{G}(\mathcal{P})$ determines a classification of radix expansions. Two lattice points $x, y \in \Lambda$ are in the same class iff $\Phi^{l_1}(x) = \Phi^{l_2}(y)$ for some non-negative integers $l_1, l_2$, or in other words, iff there is a $p \in \mathcal{P}$ for which $x, y \in \mathcal{B}(p)$. In chapter 2 we show an effective way to perform the classification.

We end this section by giving a necessary and sufficient condition for the unique representation property.

**Assertion 4.** *(Necessary and sufficient condition for the number system property) The triple* $(\Lambda, M, D)$ *is a number system if and only if for each* $z \in \Lambda$ *there is an* $n \in \mathbb{N}_0$ *such that* $\Phi^n(z) = 0$.
PROOF: The condition $\Phi(z) = 0$ is equivalent with $z \equiv a_0$ for some $a_0 \in D$. By induction, $\Phi^n(z) = 0$ if and only if $z$ can be written in the form

$$z = a_0 + Ma_1 + \ldots + M^{n-1}a_{n-1}$$

with some $a_0, a_1, \ldots, a_{n-1} \in D$. □

Assertion 4 has a very important corollary.

**Lemma 2.** *The triple* $(\Lambda, M, D)$ *is a number system if and only if* $\mathcal{P} = \{0\}$, *in which case*

$$\bigcap_{i=1}^{\infty} M^i\Lambda = \{0\}.$$

## 1.3 Length of expansions

Let $z \in \Lambda$ be an arbitrary vector. If $z_0 := z \notin \mathcal{P}$ then there is a unique $l \in \mathbb{N}$ and $a_0, a_1, \ldots, a_{l-1} \in D$ such that

$$z_j = a_j + Mz_{j+1} \ (j = 0, \ldots, l-1), \ z_l \in \mathcal{P}$$

and none of $z_0, z_1, \ldots, z_{l-1}$ do belong to $\mathcal{P}$. Let the expansion of $z$ be denoted by

$$(a_0, a_1, \ldots, a_{l-1} \mid p), \ (p = z_l). \tag{1.5}$$

If such an expansion is given then $z$ can be computed by

$$z = a_0 + Ma_1 + \ldots + M^{l-1}a_{l-1} + M^l p. \tag{1.6}$$

If $z \in \mathcal{P}$ then its expansion in $(\Lambda, M, D)$ will be denoted by $(* \mid z)$. We shall say that (1.5) is the *standard expansion* of the vector $z$ given by (1.6) and $l$ is the length of the standard expansion. For an arbitrary sequence of vectors $a_0, a_1, \ldots, a_{l-1} \in D$ and $p \in \mathcal{P}$ the expression $(a_0, a_1, \ldots, a_{l-1} \mid p)$ means the vector $z$ given by $z = \sum_{j=0}^{l-1} M^j a_j + M^l p$. This expansion is the

standard expansion of the vector $z$ if and only if $\Phi^{l-1}(z) = a_{l-1} + Mp \notin \mathcal{P}$. Observe that if $p \in \mathcal{P}$ then all $z \in \mathcal{B}(p) \setminus \mathcal{C}(p)$ have a standard expansion $(a_0, a_1, \ldots, a_l \mid \hat{p})$ for some $a_i \in D$ $(i = 0, \ldots, l)$, $l \in \mathbb{N}$ and $\hat{p} \in \mathcal{C}(p)$.

Now we give an estimate for the length of expansions in the radix system $(\Lambda, M, D)$.

Let us denote in $\mathbb{R}^k$ a vector norm and the corresponding operator norm by $\|\cdot\|$ for which $r = \|M^{-1}\| < 1$, let $K = \max\{\|d\|, d \in D\}$ and $L = Kr/(1-r)$ as before. Let $z \in \Lambda \setminus \{0\}$ be fixed. Let us define the path of $z = z_0$ in $\Lambda$ by $z_j = a_j + Mz_{j+1}$ $(j = 0, \ldots, )$. Let $T = l(z)$ be the smallest non-negative integer for which $\|z_T\| \leq L$. The existence of such a $T$ follows from Lemma 1.

**Assertion 5.** *There is a constant $c$ for which*

$$l(z) \leq \frac{\log \|z\|}{\log (1/\|M^{-1}\|)} + c. \tag{1.7}$$

PROOF: It is enough to examine the case $\|z\| > L$, $z \in \Lambda$. Since $z_j = a_j + Mz_{j+1}$ therefore $z_{j+1} = M^{-1}z_j - M^{-1}a_j$, hence $\|z_{j+1}\| \leq r(\|z_j\| + K)$. Let $t = t(z_0)$ be the smallest non-negative integer for which $\|z_t\| \leq 2KL$. Since the ball $\|\omega\| \leq 2KL$ contains finitely many lattice points therefore the inequality

$$l(z) \leq t(z) + c_1 \tag{1.8}$$

holds for an appropriate constant $c_1$. On the other hand $2KL < \|z_{t-1}\| \leq r(\|z_{t-2}\| + K) \leq r^2(\|z_{t-3}\| + K) + rK \leq \ldots \leq r^{t-1}\|z_0\| + KL$. It means that $KL \leq r^{t-1}\|z_0\|$, hence

$$\log KL \leq (t-1)\log r + \log \|z_0\|,$$

from which we can deduce that

$$(t-1)\log 1/r \leq \log \|z_0\| - \log KL,$$

i.e.,

$$t \leq \frac{\log \|z_0\|}{\log (1/r)} + c_2$$

for an appropriate $c_2$. Using the inequality (1.8) the assertion follows immediately. □

Assertion 5 extends the results of E. H. Grossman [37], I. Kátai, I. Környei [54] and B. Kovács, A. Pethő [79].

# Chapter 2

# Classification of expansions

*"There are problems that one poses,*
*and there are problems that pose themselves."*
*— H. Poincaré*

In the previous chapter it was pointed out that the function $\Phi$ defines a classification of the system $(\Lambda, M, D)$. The aim of this chapter is to give an effective algorithm to construct all these classes. Via the construction of the attractor set we also have a fast method to decide whether the radix system $(\Lambda, M, D)$ has the unique representation property.

Consider the set of "fractions" in the system $(\Lambda, M, D)$:

$$H := \mathcal{F}(M, D) = \left\{ \sum_{n=1}^{\infty} M^{-n} a_n : a_n \in D \right\} \subseteq \mathbb{R}^k. \tag{2.1}$$

This set is called the *fundamental domain* or the *set of fractions* of the system $(\Lambda, M, D)$. In chapter 5 we shall show that the set $H$ is compact in the metric space $\mathbb{R}^k$. Let $E$ be an arbitrary compact set in $\mathbb{R}^k$ and let us denote the set of lattice points in $E$ by $I(E)$, i.e., $I(E) := E \cap \Lambda$.

**Lemma 3.** *For each $z \in \Lambda$ there is an $m_0 \in \mathbb{N}_0$ such that for each $m \geq m_0$ : $\Phi^m(z) \in I(-H)$.*

PROOF: Since $H$ is a compact subset of $\mathbb{R}^k$, there exists an $\varepsilon > 0$ such that there is no element of $\Lambda$ in the set

$$N_\varepsilon(-H) \setminus -H,$$

where $N_\varepsilon(-H)$ denotes the open $\varepsilon$-neighborhood of $-H$. Let us choose an arbitrary $z \in \Lambda$. Then we get that

$$z_m = \Phi^m(z) = M^{-m}z - (M^{-1}a_1 + M^{-2}a_2 + \ldots + M^{-m}a_m)$$

for the corresponding sequence $a_1, a_2, \ldots, a_m \in D$. If $m$ is large enough, say $m \geq m_0$, then the norm of the first term of the right hand side is less than $\varepsilon$. Hence, $z_m \in \Lambda \cap (-H)$ for all $m \geq m_0$. □

COROLLARY. *(a) For each $z \in \Lambda$ the orbit of $z$ must run into the set $I(-H)$ and can never leave it. (b) If for each $z \in I(-H)$ there is an $m \in \mathbb{N}_0$ such that $\Phi^m(z) = 0$ then $(\Lambda, M, D)$ is a number system.*

The corollary suggests that in order to determine the attractors of the system $(\Lambda, M, D)$ it would be enough to find the lattice points in $-H$, or, which is computational equivalent, in $H$. Then one has only to apply the function $\Phi$ for these vectors and watching the "cycles" to be formed.

The straightforward way to compute the set $I(H)$ could be the following. It is obvious (see section 5.1) that

$$H = \bigcup_{a \in D} M^{-1}(a + H).$$

If we could find a set $T_0$, $H \subseteq T_0$, for which the lattice points of the set $M^{-1}T_0$ can be computed easily then we would be ready, because in this case $H \subseteq T_1 := \bigcup_{a \in D} M^{-1}(a + T_0)$ and only the convex hull of the lattice points in $T_1$ has to be computed. Unfortunately, to find the "smallest possible" such set $T_0$ is not easy, since the shape of the set $H$ is in almost every case rather complicated.

Our next aim is to determine a set $T$, $H \subseteq T$, for which the set of lattice points belonging to $T$ can be computed simply and which contains possibly a small number of them. We consider two approaches. One of them uses covering of the set $H$ while the other one is given by effectively computing the operator norm defined in (1.2).

## 2.1 Covering construction

Let $x = (x_1, x_2, \ldots, x_k)^T \in \mathbb{R}^k$ and $\|x\|_\infty = \max_{1 \leq i \leq k} |x_i|$. Let us denote by $\|\cdot\|_\infty$ the corresponding operator norm. If $M$ is an invertible expansive linear operator of $\mathbb{R}^k$ mapping $\Lambda$ into $\Lambda$ then there exists a smallest $c_0 \in \mathbb{N}$

such that for every $c \geq c_0,\ c \in \mathbb{N}$ the inequality $\|M^{-c}\|_\infty < 1$ holds. Let $C \geq c_0,\ C \in \mathbb{N}$ be fixed. Then

$$\|M^{-C}\|_\infty < 1,$$

therefore $(I - M^{-C})^{-1}$ exists and

$$\gamma := \frac{1}{1 - \|M^{-C}\|_\infty} \geq \|(I - M^{-C})^{-1}\|_\infty. \tag{2.2}$$

Here $I$ denotes the $k$-dimensional identity matrix. Using the notations introduced in the previous chapter let

$$M^{-j}a = \begin{bmatrix} c_1^{(j)}(a) \\ \vdots \\ c_k^{(j)}(a) \end{bmatrix},$$

and let

$$\xi_m^{(j)} := \max_{a \in D} |c_m^{(j)}(a)|,\ (m = 1, \ldots, k),$$

where $1 \leq j \leq C$. Furthermore, define the sets $I_j$ $(1 \leq j \leq C)$ as follows:

$$I_j := \left\{ x = \begin{bmatrix} x_1 \\ \vdots \\ x_k \end{bmatrix}, |x_m| \leq \xi_m^{(j)}, 1 \leq m \leq k \right\}.$$

Obviously, $M^{-j}a \in I_j$ for each $a \in D$. Let

$$\mathcal{W} := \left\{ y = \begin{bmatrix} y_1 \\ \vdots \\ y_k \end{bmatrix}, |y_m| \leq \sum_{j=1}^{C} \xi_m^{(j)}, 1 \leq m \leq k \right\}. \tag{2.3}$$

It is clear that

$$\sum_{j=1}^{C} M^{-j}a_j \in \mathcal{W}$$

for an arbitrary sequence of vectors $a_j \in D$. Hence,

$$H \subseteq \mathcal{W} + M^{-C}\mathcal{W} + M^{-2C}\mathcal{W} + \ldots. \tag{2.4}$$

Let us define the points of the $k$-dimensional rectangle $T'$ by

$$\begin{bmatrix} t_1 \\ \vdots \\ t_k \end{bmatrix}, \ -\alpha_m \le t_m \le \alpha_m, \alpha_m = \lceil \gamma \sum_{j=1}^{C} \xi_m^{(j)} \rceil, \ 1 \le m \le k. \tag{2.5}$$

Then by (2.2), (2.3) and (2.4) we get that $H \subseteq T'$ and the lattice points in the $k$-dimensional rectangle $T'$ can be computed efficiently.

*Remarks. (1)* The "good choice" for the constant $C$ in (2.2) strongly depends on the matrix $M$. A simple method could be to start with $C \leftarrow c$ and increase $C$ while $\|M^{-C}\|_\infty$ is less than or equal to a fixed constant. Another approach may require much more arithmetical operations: start with $C \leftarrow c$ and increment $C$ until the volume of $T'$ changes less than a pre-defined constant $\delta > 0$.

*(2)* Even if $M^{-n}v \to 0$ $(n \to \infty)$ for any $v \in \mathbb{R}^k$ one should be careful with raising to powers the matrix $M^{-1}$. In computer implementations using traditional programming languages on certain cases arithmetical overflow can occur. Let an example be $k = 5$, $M = \text{tridiag}(0, -2, -2^{10})$ (diag() and tridiag() denote the diagonal and tridiagonal matrices, respectively). Then $M_{1,5}^{-4} = 150323855360 > 2^{32}$. In these cases (among others) computer algebra softwares can be used (about computer algebra see [69]).

*(3)* Suppose that $\Lambda = \mathbb{Z}^k$. This can be achieved by a simple basis transformation. Then, we are interested in the integers in $T'$. It means that in equation (2.5) the floor function can also be applied. Clearly, the integers in $T'$ still cover the integers in $H$.

## 2.2 Operator norm construction

Let $x \in H$ be an arbitrary vector of $\mathbb{R}^k$. Then

$$\|x\| = \|\sum_{j=1}^{\infty} M^{-j} a_j\| \tag{2.6}$$

for any well-defined vector norm in $\mathbb{R}^k$, where $a_j \in D$ $(j = 1, 2, \dots)$. Let $M$ be an invertible expansive linear operator of $\mathbb{R}^k$. We shall construct a vector norm — throughout this subsection denoted by $\|\cdot\|_*$ —, such that for the corresponding operator norm the inequality $\|M^{-1}\|_* < 1$ holds. This

operator norm can be given using a basis transformation with the aid of an appropriate regular matrix $S$ and the maximum norm in the form

$$\|M^{-1}\|_* := \|SM^{-1}S^{-1}\|_\infty.$$

This follows from the fact that

$$\|M^{-1}x\|_* = \|SM^{-1}x\|_\infty \leq \|SM^{-1}S^{-1}\|_\infty \|Sx\|_\infty,$$

so the operator norm induced by the vector norm $\|Sx\|_\infty$.

Let $J = TM^{-1}T^{-1} = \mathrm{diag}(\Lambda_j)$ be the Jordan canonical form of the matrix $M^{-1}$. Let us choose $S := T$. Hence,

$$\|M^{-1}\|_* := \|J\|_\infty = \max_j \|\Lambda_j\|_\infty.$$

If $J$ is simple (i.e. $J$ consists of $k$ Jordan blocks) then

$$\|J\|_\infty = \rho(M^{-1}) < 1.$$

Suppose now that the eigenvalues of the matrix $M$ are not all distinct. Let $\Lambda_j = \mathrm{tridiag}(0, \lambda_j, 1) \in \mathbb{C}^{m\times m}$ be a non-trivial Jordan block ($m < k$). In this case

$$\|\Lambda_j\|_\infty > 1,$$

therefore we use the similarity transformation $D_j := \mathrm{diag}_{1\leq i\leq m}(\mu_j^{m-i})$ to obtain $D_j\Lambda_j D_j^{-1} = \mathrm{tridiag}(0, \lambda_j, \mu_j)$, where $\mu_j > 0$ and it can be chosen in such a way that $\mu_j + |\lambda_j| < 1$. Hence

$$\|D_j\Lambda_j D_j^{-1}\|_\infty < 1.$$

Putting all together, in case of trivial Jordan blocks let $D_j := 1$, moreover, $S := \mathrm{diag}(D_j)T$. Then

$$\|M^{-1}\|_* = \|SM^{-1}S^{-1}\|_\infty = \|D_j\Lambda_j D_j^{-1}\|_\infty < 1.$$

Further, let us denote $\|\cdot\| := \|\cdot\|_*$ as we used it earlier. Then $(I - M^{-1})^{-1}$ exists, it has the geometric series expansion $(I - M^{-1})^{-1} = I + M^{-1} + M^{-2} + \ldots + M^{-n} + \ldots$, and

$$\|(I - M^{-1})^{-1}\| \leq \frac{1}{1 - \|M^{-1}\|}. \tag{2.7}$$

By using (1.3), (2.6) and (2.7) we get that

$$\|Sx\|_\infty = \|x\| = \|\sum_{j=1}^{\infty} M^{-j} a_j\| \leq \frac{Kr}{1-r} = L. \tag{2.8}$$

Now we are looking for those $x \in \Lambda$ for which (2.8) is satisfied. If $\|x\|_\infty \leq L/\|S\|_\infty$ then (2.8) is clearly true. Let $y := Sx$. Then $S^{-1}y = x$, hence

$$\|x\|_\infty \leq \|S^{-1}\|_\infty \|y\|_\infty = \|S^{-1}\|_\infty \|Sx\|_\infty \leq L\|S^{-1}\|_\infty.$$

Let $T''$ be the $k$-dimensional hypercube centered at 0 with vertex coordinates $\pm\beta_i \quad (i = 1, \ldots, k)$, where

$$\beta_i := \lceil L\|S^{-1}\|_\infty \rceil. \tag{2.9}$$

It follows from the construction that $H \subseteq T''$.

*Remarks. (1)* By virtue of the construction for a given $\varepsilon > 0$ there is an operator matrix norm for which $\|M^{-1}\| \leq \rho(M^{-1}) + \varepsilon$. This is a well-known result.

*(2)* To determine the vertices of $T''$ one needs

- a Jordan block computation of $M$ and
- a matrix inverse computation of $S$.

Clearly, the matrix $S$ is not unique. The constants $\mu_j$ can be chosen arbitrary according to their definition but in computer implementations the floating point overflows (e.g. $\mu_j$-s are too small) must be avoided. The best solution would be to optimize $\mu_j$-s obtaining the smallest value for $\|S^{-1}\|$ but it could have high computational time. Nevertheless, in some cases it is worth the trouble.

*(3)* By similar arguments as we did earlier, if $\Lambda = \mathbb{Z}^k$ then in (2.9) the floor function can also be applied. Obviously, the integers in $T''$ cover the integers in $H$.

Let $\Lambda = \mathbb{Z}^k$. This can be assumed without loss of generality. Forming the intersection of $T'$ and $T''$ we proved the following theorem:

**Theorem 1.** *Let the set of integer points $I(T)$ be defined as follows:*

$$I(T) := \left\{ \begin{bmatrix} t_1 \\ \vdots \\ t_k \end{bmatrix} \in \mathbb{Z}^k,\ -\kappa_m \le t_m \le \kappa_m,\ \textit{where} \right.$$

$$\left. \kappa_m = \min(\ \lfloor \gamma \sum_{j=1}^{C} \xi_m^{(j)} \rfloor,\ \lfloor L \| S^{-1} \|_\infty \rfloor\ ),\ 1 \le m \le k \right\}.$$

*Then $I(H) \subseteq I(T)$ and $I(-H) \subseteq I(T)$.*

Computer experiments show that in many cases the covering construction is preferable to the operator norm construction. Clearly, applying Theorem 1 one can construct a $k$-dimensional rectangle $T$. Unfortunately, the number of lattice points in $T$ can be much higher than the number of periodic elements. This construction can be a first step towards a better approach.

## 2.3 Applying an iterated function system

A finite set of contractions $\{f_i\}$ mapping from $\mathbb{R}^k$ to $\mathbb{R}^k$ is called an *iterated function system* (IFS). On the space $S$ of compact subsets of $\mathbb{R}^k$, with respect to the Hausdorff metric $\delta(A,B) = \inf\{r : A \subseteq N_r(B) \text{ and } B \subseteq N_r(A)\}$, where $N_r(A)$ is the open r-neighborhood of $A$, define $f : S \to S$ by $f(X) = \bigcup_{i=1}^{l} f_i(X)$, for any compact set $X$. Clearly, $f$ is a contraction on $S$ and hence, by Hutchinson's theorem [39], $f$ has a unique fixed point or *attractor* $T$ satisfying

$$T = \bigcup_{i=1}^{l} f_i(T)$$

and given by

$$T = \lim_{n \to \infty} f^{(n)}(X_0),$$

where $f^{(n)}$ denotes the $n$th iterate of $f$, $X_0$ is an arbitrary compact subset of $\mathbb{R}^k$, and the limit is with respect to the Hausdorff metric.

For each digit $d \in D$ we define the function $f_d : \mathbb{R}^k \to \mathbb{R}^k$ by $f_d(z) = M^{-1}(z+d)$. These are linear contraction maps. If $z \in H$ then $f_d(z) \in H$. Clearly, $f_d$ is a right-shift map and furthermore $H = \bigcup_{d\in D} f_d(H)$ so $H$ is the unique invariant set determined by Hutchinson's theorem applied to the functions $f_d$. The set $H$ is self-affine with respect to these functions.

It was already mentioned that we are interested in the lattice points in the set $-H$. Let $\pi \in -H$. Then

$$-\pi - (M^{-1}d_1 + \ldots + M^{-J}d_J) = M^{-(J+1)}d_{J+1} + M^{-(J+2)}d_{J+2} + \ldots , \tag{2.10}$$

for the appropriate sequence $d_i \in D$. Fortunately, for the right hand side of (2.10) a good estimate can be given. Let $\Lambda = \mathbb{Z}^k$. The following algorithm provides the set $W$, for which the integers in $W$ cover the integers in $H$.

Number Expansion Classification Algorithm in $\mathbb{Z}^k$ for a given expansive matrix $M$ and digit set $D$. Let $\hat{M} \in \mathbb{Z}^{k\times k}$ be similar to $M$ via the matrix $Q$ and let $Q$ be an optional argument of the algorithm. If it is not given then let $Q$ be the identity matrix. Let $\hat{D} = QD$. Further, $B$ and $C$ are constants depending on the given computer hardware (word size, memory capacity) and on the matrix $\hat{M}$. $B$ is an integer and $C < 1$ a real number.

1. $q := \min\{j \in \mathbb{N},\ \|\hat{M}^{-j}\|_\infty < 1\}$;
2. $s := \min\{j \in \mathbb{N},\ (r := \|\hat{M}^{-j}\|_\infty) < C\}$;
3. $f := (f_1, \ldots, f_k)^T \in \mathbb{R}^k$, $f_m = 1/(1-r)\sum_{l=1}^{s} \max_{b\in\hat{D}} |c_m^{(l)}(b)|$, $1 \le m \le k$, where $(c_1^{(l)}(b), \ldots, c_k^{(l)}(b))^T = \hat{M}^{-l}b$;
4. minvol:=infinity; Chose an appropriate $B$, $q \le B \le s$;
5. **for** $j$ **from** $q$ **to** $B$ **do** {
   **if** $(\|\hat{M}^{-j}\|_\infty < 1)$ {
   Compute the vector $v^{(j)} = (v_1^{(j)}, \ldots, v_k^{(j)})^T \in \mathbb{R}^k$,
   $v_m^{(j)} = \sum_{l=1}^{k} |\hat{M}^{-j}_{m,l} f_l|$, $1 \le m \le k$;
   **if** $((\omega := \prod_{l=1}^{k} v_l^{(j)}) <$ minvol) { minvol:= $\omega$; $J := j$;}}}
6. $U := \{-\sum_{i=1}^{J} \hat{M}^{-i}b, b \in \hat{D}\}$;
7. $S := \bigcup_{u\in U}(u+P)$, where $P$ denotes the $k$-dimensional rectangle $P = \{(p_1, \ldots, p_k)^T \in \mathbb{R}^k,\ |p_i| \le v_i^{(J)},\ 1 \le i \le k\}$;
8. $W := \{w = (w_1, \ldots, w_k)^T \in \mathbb{Z}^k,\ Qw \in S\}$;
9. Apply the function $\Phi$ determined by the system $(\mathbb{Z}^k, M, D)$ for the points of $W$ and the arising cycles mean the required classification.

The lines 1-3 provide the $k$-dimensional rectangle $\hat{G} = \{(g_1, \ldots, g_k)^T \in \mathbb{R}^k, |g_i| \leq f_i,\ 1 \leq i \leq k\}$. Let us analyze the second assignment in line 4. If we increase $B$, the time complexity of the algorithm grows exponentially in $t = |\det(M)|$. Unfortunately, in some cases $q$ can be rather big, which means that the convergence of $M^{-i}$ $(i \to \infty)$ is slow. In these cases this algorithm can be ineffective, even if keeping the running time moderate one choose $B$ close to $q$. The reason is that the set $\hat{G}$ can also be rather big. Let an example be the Frobenius matrix (companion matrix) of the irreducible polynomial $2 + 3x + 4x^2 + 4x^3 + 4x^4 + 3x^5 + 2x^6 + x^7$ with the canonical[1] (binary) digit set, $Q = I,\ C = 0.01$. Then $s = 188, q = 53$ and the number of integers in $\hat{G}$ is 15319297125. Using other kinds of matrices, during the computation of $s$ problems can arise with the matrix elements (see section 2.1, Remark 2). Line 5 tries to keep the index $J$ small. The lines 6-8 are the application of Hutchinson's theorem in (2.10). Concerning line 8 one can observe that the number of elements of the set $W$ depends also on $|\det(Q)|$. Concerning line 9, a fast algorithm for computing the function $\Phi$ is the subject of the next section. The termination of the algorithm is clear.

It must be emphasized that the running time of the algorithm depends strongly on the matrices $M$ and $Q$, i.e., on the basis of the lattice determined by the matrix $M$. In other words one has to choose the matrix $Q$ in a way that the convergence of $\hat{M}^{-i} = (QMQ^{-1})^{-i}$ $(i \to \infty)$ is fast, $|\det(Q)|$ is big and the volume of $\hat{G}$ is as small as possible. It seems to be rather hard. Sometimes the simple idea of choosing the matrix $Q$ in a way that $\hat{M} = M^T$ can help. Fortunately, for a large class of matrices the algorithm is quite effective even if we choose $Q$ for the identity matrix. The author implemented the CLASSIFICATION ALGORITHM in C language. In order to perform computations in the lattice effectively the elements of $\mathbb{Z}^k$ were transformed to $\mathbb{Z}$ using mixed radix representation. During the computation of elements of the set $S$ a hashing table was used.

## 2.4 Computation of the function $\Phi$

Let a radix system $(\mathbb{Z}^k, M, D)$ be given. For calculation of the function $\Phi$ one needs a fast procedure to determine for an arbitrary $z \in \mathbb{Z}^k$ the corresponding congruent element $d \in D$ modulo $M$. Our first method is a straightforward generalization of the method used for the case of Gaussian integers in [62].

[1] For the definition see section 3.1.

### 2.4.1 Adjoint method

Applying the notations already adopted let $z$ be an arbitrary element of $\mathbb{Z}^k$ and let $D = \{a_0, a_1, \ldots, a_{t-1}\}$ be a complete residue system modulo $M$. If $z \equiv a_j$ modulo $M$ then $M^*z \equiv M^*a_j$ modulo $\det(M)I$, where $M^*$ denotes the adjoint of $M$ and $I$ the identity matrix. Here by "adjoint of the operator $M$" we mean the integer matrix, for which the elements are the adjoints of the appropriate sub-determinants. Let $t = |\det(M)|$ as before. Let

$$D_1 := M^*D \pmod{tI} = \{b_0, b_1, \ldots, b_{t-1}\}, \tag{2.11}$$

where

$$b_j = M^*a_j \pmod{tI} = \begin{bmatrix} b_1^{(j)} \\ \vdots \\ b_k^{(j)} \end{bmatrix} \in \mathbb{Z}^k,\ 0 \le b_i^{(j)} < t,\ (i = 1, \ldots, k). \tag{2.12}$$

Due to the complete residue system property of $D$ for every $z \in \mathbb{Z}^k$ there exists a unique $b_j \in D_1$ such that $b_j = M^*z \pmod{tI}$. Then from (2.11) and (2.12) it follows that $z \equiv a_j$ modulo $M$.

In order to obtain for an arbitrary $z \in \mathbb{Z}^k$ the congruent element in $D$ modulo $M$ one has to perform a multiplication by the matrix $M^* \pmod{tI}$, which requires $k^2$ integer multiplication over $\mathbb{Z}_t = \mathbb{Z}/t\mathbb{Z}$. Can the number of operations be reduced? Fortunately, in many cases the answer is yes. Suppose that there exists an $i \in \mathbb{N},\ 1 \le i \le k$ for which $b_i^{(j)} (j = 0, 1, \ldots, t-1)$ in (2.12) are all different. Then the inner product of an arbitrary $z \in \mathbb{Z}^k$ by the $i$-th row of $M^*$ modulo $t$ uniquely determines the index $j$ for which $z \equiv a_j$ modulo $M$. This requires only $k$ integer multiplications over $\mathbb{Z}_t$. The question, in which cases such an $i$ exists will be answered in chapter 3. But what can be made when such an $i$ does not exists? Then one has to investigate further the set $D_1$ and to figure out a strategy to minimize the number of multiplications to obtain for an arbitrary $z \in \mathbb{Z}^k$ the appropriate $b_j \in D_1$ for which $b_j = M^*z$ modulo $tI$. Beside the optimization the strategy requires greatest common divisor computations, which suggests the existence of another (a simpler) approach. Indeed, essentially the same can be reached via another way, which is based on the Smith canonical form of $M$ (see [45]).

### 2.4.2 Smith normal form method

Let $M$ be an invertible linear operator mapping $\mathbb{Z}^k$ into $\mathbb{Z}^k$. Then there are linear transformations $U$ and $V$ mapping $\mathbb{Z}^k$ onto itself such that $UMV = G$ has diagonal form in the standard basis with positive integer elements $g_1, \ldots, g_k$ in the diagonal such that $g_i \mid g_{i+1}$ for $i = 1, 2, \ldots, k-1$ and $\prod_{i=1}^{k} g_i = |\det(M)|$. The Smith normal form can be obtained by doing elementary row and column operations of $M$. We remark that $U$ and $V$ have determinants $\pm 1$ and they are also invertible having integer components.

**Lemma 4.** *For an invertible $M$ with the notations above let for $z_1, z_2 \in \mathbb{Z}^k$ the numbers $u_1, u_2, \ldots, u_k$ and $\hat{u}_1, \hat{u}_2, \ldots, \hat{u}_k$ denote the coordinates of $Uz_1$ and $Uz_2$ respectively. Then $z_1 \equiv z_2$ modulo $M$ if and only if $u_i \equiv \hat{u}_i$ modulo $g_i$ for all $i = 1, 2, \ldots, k$.*

PROOF: $z_1 \equiv z_2$ modulo $M$ if and only if $M^{-1}(z_1 - z_2) \in \mathbb{Z}^k$. This is equivalent with the condition $V^{-1}M^{-1}(z_1 - z_2) \in \mathbb{Z}^k$. But $V^{-1}M^{-1} = G^{-1}U$, hence the equations $u_i \equiv \hat{u}_i$ modulo $g_i$ must be satisfied for all $i = 1, 2, \ldots, k$. □

From a computational point of view, at the first sight there is no gain. In the first step one has to multiply $z \in \mathbb{Z}^k$ by the integer matrix $U \pmod{G}$ instead of $M^* \pmod{tI}$. But if there exists a positive integer $s$ for which $g_i = 1$, $i = 1, \ldots, s, s < k$ then $u_i \equiv 0 \pmod{g_i}$ for all $i = 1, \ldots, s$ and for all $z \in \mathbb{Z}^k$, hence enough to perform only $k$ integer multiplications modulo $g_j$, for each $j = s+1, \ldots, k$. Let

$$D_2 := UD \pmod{G} = \{c_0, c_1, \ldots, c_{t-1}\}, \tag{2.13}$$

where

$$c_j = Ua_j \pmod{G} = \begin{bmatrix} c_1^{(j)} \\ \vdots \\ c_k^{(j)} \end{bmatrix} \in \mathbb{Z}^k, \ 0 \le c_i^{(j)} < g_i, \ (i = 1, \ldots, k). \tag{2.14}$$

We get that for every $z \in \mathbb{Z}^k$ there exists a unique $c_j \in D_2$ such that $c_j = Uz \pmod{G}$. From (2.13) and (2.14) we have that $z \equiv a_j$ modulo $M$.

### 2.4.3 Computer implementation

In computer implementations once the computation $M^*z$ modulo $tI$ or $Uz$ modulo $G$ was performed for the vector $z \in \mathbb{Z}^k$ the result must be looked up

in the table $T(D_1)$ or in $T(D_2)$, respectively, obtaining the index $j$ for which $a_j \equiv z$ modulo $M$, $a_j \in D$. This can be done using searching strategies or hashing. Let us see an example for such a hash function in the case of Smith normal form. The idea comes from the mixed radix representation.

**Lemma 5.** *Using the notations above let us define the function $h$ by*

$$h(z) = \sum_{i=s+1}^{k} (u_i \mod g_i) \prod_{j=s+1}^{i-1} g_j.$$

*Then $h$ is an integer valued function with values $0, \dots, t-1$, and $h(z_1) = h(z_2)$ if and only if $z_1 \equiv z_2$ modulo $M$.*

PROOF: It is easy to see that $h$ has the given range. If $z_1 \equiv z_2$ then $u_i \equiv \hat{u}_i \bmod g_i$ for all $i = 1, 2, \dots, k$, hence $h(z_1) = h(z_2)$. In the other direction, if $h(z_1) = h(z_2)$, then taking the remainder of both side with respect to $g_1$ we get that $u_1 \equiv \hat{u}_1 \pmod{g_1}$. Subtracting this common term and dividing with $g_1$ one can continue with $g_2$, etc. □

*Remark.* The set $D_1$ can be generated only from $D$ but the set $D_2$ can be produced also directly from $G$. A complete residue system (mod $M$) can be generated from $D_2$ ($D_1$) by multiplying the elements with $U^{-1}$ ($M$), respectively.

We summarize our results for the computation of the function $\Phi$:

- For a given vector $z \in \mathbb{Z}^k$ computing $M^* z \pmod{tI}$ needs $k^2$ integer multiplications over $\mathbb{Z}_t$, computing $Uz \pmod G$ requires $k$ integer multiplications over $\mathbb{Z}_{g_j}$ for each $j = s+1, \dots, k$, where $s$ depends on the matrix $M$.

- Looking up the congruent element $a_j$ in the table $T(D)$ either a searching has to be performed in $T(D_1)$ or in $T(D_2)$ to obtain the index $j$ or a hashing has to be done.

- To perform the function $\Phi$, after a vector subtraction a matrix multiplication must be applied either with $M^*$ over $\mathbb{Z}$ and then dividing by $t$ or with $M^{-1}$ over $\mathbb{R}$.

# Chapter 3

# Number system constructions

*"Number theory is an inexhaustible storehouse of interesting truth."*
*— C. F. Gauss*

This chapter contains number system constructions of several types. First a necessary and sufficient condition is given establishing canonical digit sets. Then, we deal with polynomial constructions including the complete list of generalized binary number systems up to degree 8. Polygonal and simultaneous constructions are also mentioned. We end this chapter by proving a sufficient condition for the general case.

## 3.1 Canonical digit sets

Let $\Lambda = \mathbb{Z}^k$ and let $M : \mathbb{Z}^k \to \mathbb{Z}^k$ be a matrix satisfying Assertion 1(b)-(c). Further, we examine special kinds of digit sets. A set of vectors $D_M^{(j)} \subset \mathbb{Z}^k$ is called *j-canonical* with respect to the matrix $M$ $(1 \leq j \leq k)$ if all the elements have the form $\nu e_j$, where $e_j$ denotes the $j$-th unit vector, $\nu = 0, \ldots, |\det(M)| - 1$. If the set $D_M^{(j)}$ forms a complete residue system modulo $M$ — CRS for brevity — then we call it a *j-canonical digit set* and denote it by $D^{(j)}$. If there exists a $j$ for which $(\mathbb{Z}^k, M, D^{(j)})$ is a number system then it is called *j-canonical number system*. Furthermore, 1-canonical digit sets are called simply canonical. In the following we analyze the existence of $j$-canonical complete residue systems.

**Theorem 2.** *Let $M$ be an invertible expansive linear operator of $\mathbb{R}^k$ mapping $\mathbb{Z}^k$ into itself and let $\underline{c} = [c_1, c_2, \dots, c_k]^T \in \mathbb{Z}^k$ be the $j$-th column of the matrix $M^*$ (adjoint of $M$). Let $\delta_l := \gcd(c_l, t)$ $(l = 1, \dots, k)$, where $t = |\det(M)|$. Let furthermore $\tau_l := t/\delta_l$. Then the following statements are equivalent:*

1 *There exists $j$-canonical CRS modulo $M$.*

2 *The set*

$$D^{(j)} = \left\{ \nu\underline{c} \bmod t = \begin{bmatrix} \nu c_1 \bmod t \\ \vdots \\ \nu c_k \bmod t \end{bmatrix}, \ \nu = 0, 1, \dots, t-1 \right\}$$

*has exactly $t$ elements.*

3 $\operatorname{lcm}(\tau_1, \dots, \tau_k) = t$.

*(Here* gcd *and* lcm *means the greatest common divisor and least common multiply of the integer elements, resp.)*

PROOF: (1) $\Leftrightarrow$ (2). The proof immediately follows from the construction of $D_1$ in (2.11). (1) $\Leftrightarrow$ (3). Due to the CRS property of the set $D^{(j)}$ all their elements are incongruent modulo $M$ and the set $D^{(j)}$ has $t$ elements. This means that the equation $h\underline{e}_j = M\underline{\eta}$ has no solution for any $h \in \mathbb{N}$, $0 < h < t$ and any $\underline{\eta} = [\eta_1, \eta_2, \dots, \eta_k]^T \in \mathbb{Z}^k$. Hence it is enough to examine the solvability of the system of equations

$$\begin{aligned} hc_1 &= t\eta_1, \\ &\vdots \\ hc_k &= t\eta_k. \end{aligned} \tag{3.1}$$

Case 1. There exists a $c_l$ $(1 \le l \le k)$ such that $\gcd(c_l, t) = 1$. In this case from the equation $hc_l = t\eta_l$ it follows that $t \mid h$. Therefore the system of equations (3.1) has no integer solution.

Case 2. Suppose that $\gcd(c_l, t) = \delta_l > 1$ for all $l = 1, 2, \dots, k$. Let $c_l^* = c_l/\delta_l$. Then $hc_l^* = \tau_l\eta_l$ $(l = 1, \dots, k)$. Since $\gcd(c_l^*, \tau_l) = 1$, therefore $\tau_l \mid h$ for all $l = 1, \dots, k$. It means that $\operatorname{lcm}(\tau_1, \tau_2, \dots, \tau_k) \mid h$. Hence the system of equations (3.1) has no solution if and only if $\operatorname{lcm}(\tau_1, \tau_2, \dots, \tau_k) \ge t$. On the other hand $\operatorname{lcm}(\tau_1, \dots, \tau_k) \mid t$. Therefore $\operatorname{lcm}(\tau_1, \dots, \tau_k) = t$. (If $\tau_l = t$ for some $l$ then $\gcd(c_l, t) = 1$.) We have that there exists $j$-canonical CRS modulo $M$ if and only if $\operatorname{lcm}(\tau_1, \dots, \tau_k) = t$. □

*Remarks. (1)* If there exists a $c_i \in \mathbb{Z} \setminus 0$ in the $j$-th column of the matrix $M^*$ for which $\gcd(c_i, t) = 1$ modulo $t$ then there is a $j$-canonical complete residue system modulo $M$. Theorem 2 shows that the converse of this statement is not always true.

*(2)* If $t$ is prime then always exists $j$-canonical CRS for all $1 \leq j \leq k$.

**Lemma 6.** *Using the notations above suppose that for a given $M$ there exists a $j$-canonical CRS. Then there is an $i \in \mathbb{N}$, $1 \leq i \leq k$ for which $\gcd(c_i, t) = 1$ modulo $t$ if and only if the set $\{\nu c_i$ modulo $t$, $\nu = 0, 1, \ldots, t-1\}$ forms a CRS modulo $t$.*

The proof is obvious.

COROLLARY. *If for a given $M$ there exist $j$-canonical CRS and $c_i$ according to Lemma 6 then it is enough to perform only $k$ multiplications modulo $t$ to determine for an arbitrary $z \in \mathbb{Z}^k$ the element $b = (M^* z$ modulo $tI) \in D_1$ (see section 2.4.1).*

The converse of this statement is not true. Let a counter-example be the matrix $M = \left(\begin{smallmatrix} 2 & 4 \\ 6 & 3 \end{smallmatrix}\right)$. Then $t = 18$ and $M^* = \left(\begin{smallmatrix} -3 & 4 \\ 6 & -2 \end{smallmatrix}\right)$. Using the Smith normal form for every $z \in \mathbb{Z}^k$ there is enough to perform $k = 2$ multiplications to obtain the appropriate $b \in D_1$ but there is no 1- or 2-canonical CRS and $\gcd(c_i, t) > 1$ modulo $t$ for all $c_i$.

## 3.2 Polynomial construction

Consider the polynomial

$$f(x) = c_k x^k + c_{k-1} x^{k-1} + \ldots + c_0 = (x - \theta_1) \ldots (x - \theta_k), \; c_k = 1 \quad (3.2)$$

over $\mathbb{Z}[x]$. Let us denote the quotient ring $\mathbb{Z}[x]/(f)$ by $\Lambda_f$. Let $\beta = x + (f)$ denote the image of $x$ in $\Lambda_f$. Then $\Lambda_f$ has the structure of a free Abelian group with basis $\{1, \beta, \beta^2, \ldots, \beta^{k-1}\}$. Hence, $\Lambda_f$ is a lattice, addition and multiplication of lattice points is just addition and multiplication in the ring $\mathbb{Z}[x]/(f)$. To be more precise consider the polynomial $f(x)$ in (3.2) and assume that $|\theta_i| > 1$ $(i = 1, \ldots, k)$. Observe that $\Lambda_f$ is the set of elements of form $u_0 + u_1\beta + \ldots + u_{k-1}\beta^{k-1}$ $(u_j \in \mathbb{Z})$. For the addition it is isomorphic with the additive group $\mathbb{Z}^k$. Clearly, $I_\beta = \{\beta\sigma : \sigma \in \Lambda_f\}$ is an ideal in $\Lambda_f$, the number of residue classes in the factor ring $\Lambda_f / I_\beta$ is $t = |\theta_1 \ldots \theta_k|$. Choosing an element from each residue class the digit set can be defined as $D_\beta = \{a_0 = 0, a_1, \ldots, a_{t-1}\} \subseteq \Lambda_f$. Let $\alpha \in \Lambda_f$. Then there exists a

unique $a \in D_\beta$ and a unique $\alpha_1 \in \Lambda_f$ for which $\alpha = a + \beta\alpha_1$. The function $\Phi : \Lambda_f \to \Lambda_f$ is defined as $\Phi(\alpha) = \alpha_1$. Observe that the map $\alpha \to \beta\alpha$ can be formulated as a linear transformation, which has a simple form in the basis $\{1, \beta, \beta^2, \ldots, \beta^{k-1}\}$, namely the Frobenius matrix

$$M_f = \begin{pmatrix} 0 & \ldots & & -c_0 \\ 1 & 0 & \ldots & \vdots \\ 0 & \ddots & & \\ \vdots & & & \\ 0 & \ldots & 1 & -c_{k-1} \end{pmatrix}. \tag{3.3}$$

Hence, all the problems regarding number expansions can be formulated in $\mathbb{Z}^k$ instead of making it in $\Lambda_f$. The digit set for $M_f$ must have $|c_0|$ elements. Clearly, $|c_0|$ must be greater than or equal to 2.

### 3.2.1 Radix representation of algebraic integers

In the special case, when $f(x)$ is irreducible over $\mathbb{Z}[x]$ then $\Lambda_f = \mathbb{Z}[x]/(f)$ is isomorphic with $\mathbb{Z}[\theta]$, where $\theta$ is any root of $f(x)$ in an appropriate extension field of the rationals. Hence, we may replace $\beta$ to $\theta$ in the previous reasoning. The next lemma provides a sufficient condition for $\mathbb{Z}[x]/(f)$ being isomorphic with $\mathbb{Z}[\theta]$.

**Lemma 7.** *Consider the polynomial $f(x)$ in (3.2) and assume that $|\theta_i| > 1$, $(1 \leq i \leq k)$. If $f(0) = c_0$ is prime then $f(x)$ is irreducible.*
PROOF: Suppose indirectly that $f(x) = u(x)v(x)$, $u, v \in \mathbb{Z}[x]$, $\deg(u) \geq 1, \deg(v) \geq 1$ and both $u$ and $v$ are monic. Since $c_0 = f(0) = u(0)v(0)$ is prime therefore either $u(0)$ is $\pm 1$ or $v(0)$ is $\pm 1$. Assume that $u(0)$ is $\pm 1$. Since the constant term of $u(x)$ is the product of some roots of $f$ in module, this is impossible. □

In the following we shortly summarize the results obtained by representing algebraic integers in some extension field of the rationals. Let $\theta$ be any rational integer greater than one. It is well-known that every non-negative integer $n$ has a unique representation of the form $n = a_0 + a_1\theta + \ldots + a_k\theta^k$, where the integers $a_j$ are selected from the set $\{0, 1, \ldots, \theta - 1\}$. The decimal $(\theta = 10)$ and binary $(\theta = 2)$ systems are the most familiar. Both positive and negative integers can be uniquely represented without a sign prefix in any

negative base $\theta < -1$ using the digits from $\{0, 1, \ldots, |\theta| - 1\}$. Conditions under which each rational integer has a unique radix representation have been investigated by D. W. Matula [86], A. M. Odlyzko [91] and by B. Kovács, A. Pethő [77].

A straightforward way to extend radix systems is choosing the radix to an algebraic integer. The first non-real base radix system was introduced by D. E. Knuth [60], who suggested that $\theta = 2i$ can be used as base for the complex numbers with the digit set $D = \{0, 1, 2, 3\}$, i.e. all complex number $\gamma$ has an expansion of form $\gamma = \sum_{i=-\infty}^{l} d_i \theta^i$ for some $l \in \mathbb{N}_0$ ($d_i \in D$). However, in order to represent all the Gaussian integers, it is necessary to use one negative radix place; for example $1+5i = 3(2i)^1+1(2i)^0+2(2i)^{-1}$. W. Penney in 1965 noticed [94], that every complex number can be represented in binary form using the base $-1+i$, moreover, all the Gaussian integers can be written in the form $\sum_{j=0}^{n} a_j(-1+i)^j$, where $a_j = 0$ or 1.

The systematic research of positional number systems in algebraic extensions was initiated by I. Kátai and J. Szabó [55]. They proved that if $\theta$ is a Gaussian integer of norm $N \geq 2$ and the digit set is $D = \{0, 1, \ldots, N-1\}$ then every Gaussian integer $\gamma$ can be uniquely represented as $\gamma = a_0 + a_1\theta + \ldots + a_m\theta^m$, $a_j \in D$, $a_m \neq 0$ if and only if $\theta = -n \pm i$ for some positive integer $n$.

If the digit set $D$ is restricted to be a set of non-negative numbers, we get a straightforward generalization of the traditional number systems in $\mathbb{Z}$. The set $D = \{0, 1, \ldots, N-1\}$ is called *canonical digit set*. If the radix system $(\mathbb{Z}[\theta], \theta, D)$ satisfies the unique representation property with some canonical digit set $D$ then it is called a *canonical number system*. In this case all those integers $\theta$ in quadratic number fields can be given, for which $(\mathbb{Z}[\theta], \theta, D)$ are number systems [27, 52, 53]: if $\theta$ is a quadratic integer with minimal polynomial $x^2 + Ex + F$ and $D = \{0, 1, \ldots, |F| - 1\}$ then $(\mathbb{Z}[\theta], \theta, D)$ is a number system if and only if $F \geq 2$ and $-1 \leq E \leq F$.

Using canonical digit sets S. Körmendi [80] determined all the integers $\theta \in \mathbb{Q}(\sqrt[3]{2})$ for which $(\mathbb{Z}[\theta], \theta, D)$ is a number system. B. Kovács [72] gave a necessary and sufficient condition for the *existence* of canonical number systems in $\mathbb{Z}[\theta]$, i.e., in the ring of integers $\mathbb{Q}[\theta]$ of a $k^{th}$ degree extension of $\mathbb{Q}$ ($k \geq 3$) there exists canonical number system iff there exists an $\alpha \in \mathbb{Q}[\theta]$ such that $\{1, \alpha, \ldots, \alpha^{k-1}\}$ is an integer basis in $\mathbb{Q}[\theta]$. B. Kovács and A. Pethő [78] characterized all those integral domains that have canonical number systems.

### 3.2.2 Cns-polynomials

The concept of canonical number systems was extended to arbitrary square-free polynomials $f(x) \in \mathbb{Z}[x]$ with leading coefficient one by A. Pethő [95] and to arbitrary monic polynomials $f(x) \in \mathbb{Z}[x]$ by S. Akiyama and A. Pethő [1]. Concerning (3.3) it is easy to see that $M_f^*[k,1] = (-1)^{k+1}$ therefore by Theorem 2 canonical digit set always exist. Here $M^*$ means the adjoint of $M$.

Let a canonical radix system $(\Lambda_f, M_f, D)$ be given. Computing the Smith normal form of $M_f$ by $UM_fV = G$ it is easy to see that

$$U = \begin{pmatrix} 0 & 1 & & 0 \\ \vdots & & \ddots & \\ 0 & 0 & & 1 \\ -\operatorname{sgn}(c_0) & 0 & \dots & 0 \end{pmatrix}$$

and $G = \operatorname{diag}(1, \dots, 1, |c_0|)$. Hence, by Lemma 4 the function $\Phi$ can be given as

$$\begin{aligned} &\Phi(\underline{x}) = \Phi([x_1, \dots, x_k]^T) = \\ &= [-\frac{c_1}{c_0}x^* + x_2, -\frac{c_2}{c_0}x^* + x_3, \dots, -\frac{c_{k-1}}{c_0}x^* + x_k, -\frac{x^*}{c_0}]^T \end{aligned} \tag{3.4}$$

where $x^* = x_1 - d, 0 \le d < |c_0|$ and $c_0 \mid x^*$. Using the notation $y = \lfloor x_1/c_0 \rfloor$ in (3.4) the function $\Phi$ can also be written as

$$\Phi(\underline{x}) = [-c_1 y + x_2, -c_2 y + x_3, \dots, -c_{k-1} y + x_k, -y]^T. \tag{3.5}$$

If the system $(\Lambda_f, M_f, D)$ is a canonical number system then we call the polynomial $f(x)$ as a *cns-polynomial*, or we say that the polynomial $f(x)$ has the *cns-property*. Recall that in this case for every $\underline{x} \in \mathbb{Z}^k$ there is a $j \in \mathbb{N}_0$ for which $\Phi^j(\underline{x}) = 0$.

### 3.2.3 Necessary conditions for the cns-property

In order to construct canonical number systems via cns-polynomials we give some necessary conditions. These conditions are quite obvious, many of them were used in different research papers by W. J. Gilbert, I. Kátai and A. Pethő. We prove them for the sake of completeness.

**Lemma 8.** *If $(\Lambda_f, M_f, D)$ is a canonical number system defined by the cns-polynomial (3.2) then*
*(a) $c_0 \geq 2$;*
*(b) if $-1 \leq r \in \mathbb{R}$ then $f(r) > 0$, if $-1 \leq z \in \mathbb{Z}$ then $f(z) \geq 1$;*
*(c) $f(1) \geq c_0$;*
*(d) if $k$ is even then $f(-c_0) \geq 1$, if $k$ is odd then $f(-c_0) \leq -1$;*
*(e) $\sum_{i=0}^{\lfloor k/2 \rfloor} c_{2i} \geq \lfloor (c_0+1)/2 \rfloor$.*
PROOF: (a) It is clear that each real root of $f(x)$ (if exists) must be less than $-1$. Hence, $c_0 = (-1)^k \theta_1 \dots \theta_k > 1$. Concerning (b) the previous idea can also be applied. (c) It is known that the only periodic element in the number system $(\Lambda_f, M_f, D)$ is the null vector. Now we analyze how can we avoid the loops $\Phi(\underline{x}) = \underline{x}$ different from $0 \to 0$. Suppose that there is a loop. Using (3.5) the following system of equations can be set up: $\{x_1 = x_2 - c_1 y, x_2 = x_3 - c_2 y, \dots, x_{k-1} = x_k - c_{k-1} y, x_k = -y\}$. From these equations it is easy to deduce that $x_k(1 + c_{k-1} + \dots + c_0) = d \in D$. If $x_k = 0$ then $\underline{x} = 0$ which is a known case. If $x_k \neq 0$ then applying (a) the number of loops is $\lfloor (c_0 - 1)/f(1) \rfloor$. Hence, if $c_0 \leq f(1)$ then there does not exist any loop. Concerning (d) if $\theta_i \in \mathbb{C} \setminus \mathbb{R}$ for all $0 \leq i \leq k$ then the assertion is obvious. On the other hand observe that there does not exist any real $\theta_i$ for which $\theta_i \leq -c_0$, otherwise there would be a $\theta_j$ for which $|\theta_j| < 1$. Hence $-c_0 < \theta_i < -1$ for all real roots of $f(x)$. It means that if $k$ is even then $f(-c_0) \geq 1$, if $k$ is odd then $f(-c_0) \leq -1$. (e) is immediately follows from (a) and (b) by $z = -1$. □

Let $c_0 \geq 2$ and $k$ be fixed. Since all roots of the polynomial $f(x)$ has moduli greater then one — we also say that the polynomial satisfies the root-condition —, therefore the number of cns-polynomials is finite. Next, we provide upper bounds for the absolute value of the coefficients $c_i, 1 \leq i \leq k-1$ in (3.2).

**Lemma 9.** *Let $f(x)$ be the cns-polynomial defined by (3.2) and let $2 \leq k \leq 9$. Then the coefficients of $f(x)$ can be bounded as*

$$|c_j| \leq s(1 - c_0) + c_0 \binom{k}{j} - 1,$$
$$|c_{k-j}| \leq s(c_0 - 1)(1 - \lfloor k/j \rfloor) + c_0 \binom{k}{j} - 1,$$
$$\text{where } s = \left\lfloor \binom{k}{j} / \lfloor k/j \rfloor \right\rfloor, \quad 1 \leq j \leq \lfloor k/2 \rfloor.$$

PROOF: We use the relationship between roots and coefficients of polynomials and the inequalities

$$\alpha+\beta<1+\alpha\beta \quad \text{and} \quad \frac{1}{\alpha}+\frac{1}{\beta}<1+\frac{1}{\alpha\beta} \tag{3.6}$$

where $\alpha,\beta>1$. For brevity let $z_i=\mid \theta_i \mid$. To have a better view into the formulas let us consider the special case $k=7, j=2$. Then $\sum_{1\le i_1<i_2\le 7} z_{i_1}z_{i_2} < z_1z_2\ z_4z_5\ z_6z_7+z_1z_3\ z_2z_5\ z_4z_7+z_1z_4\ z_2z_6\ z_3z_7+z_1z_5\ z_2z_4\ z_3z_6+z_1z_6\ z_2z_3\ z_5z_7+z_1z_7\ z_3z_4\ z_5z_6+z_2z_7\ z_3z_5\ z_4z_6+2\cdot 7<7c_0+14$. In the given range $2\le k\le 9$ such a sort is always possible. Hence,

$$\mid c_{k-j} \mid = \sum_{1\le i_1<\ldots<i_j\le k} z_{i_1}\ldots z_{i_j} < sc_0+s(\lfloor k/j\rfloor-1) \quad \text{and}$$
$$\mid c_j \mid = c_0 \sum_{1\le i_1<\ldots<i_j\le k} \frac{1}{z_{i_1}}\ldots\frac{1}{z_{i_j}} < c_0(\frac{s}{c_0}+s(\lfloor k/j\rfloor-1)),$$

from which the lemma follows. □

*Remarks. (1)* These estimates are good enough for searching canonical number systems algorithmically.

*(2)* By using these formulas we got the following estimates ($c_k=1$):
$k=2, |c_1|\le c_0$;
$k=3, |c_1|\le 2c_0, |c_2|\le c_0+1$;
$k=4, |c_1|\le 3c_0, |c_2|\le 3c_0+2, |c_3|\le c_0+2$;
$k=5, |c_1|\le 4c_0, |c_2|\le 5c_0+4, |c_3|\le 5c_0+4, |c_4|\le c_0+3$;
$k=6, |c_1|\le 5c_0, |c_2|\le 10c_0+4, |c_3|\le 10c_0+9, |c_4|\le 5c_0+9, |c_5|\le c_0+4$;
$k=7, |c_1|\le 6c_0, |c_2|\le 14c_0+6, |c_3|\le 18c_0+16, |c_4|\le 18c_0+16, |c_5|\le 7c_0+13, |c_6|\le c_0+5$;
$k=8, |c_1|\le 7c_0, |c_2|\le 21c_0+6, |c_3|\le 28c_0+27, |c_4|\le 35c_0+34, |c_5|\le 28c_0+27, |c_6|\le 7c_0+20, |c_7|\le c_0+6$;
$k=9, |c_1|\le 8c_0, |c_2|\le 27c_0+8, |c_3|\le 56c_0+27, |c_4|\le 63c_0+62, |c_5|\le 63c_0+62, |c_6|\le 28c_0+55, |c_7|\le 9c_0+26, |c_8|\le c_0+7$.

### 3.2.4 Some results

It was observed that a wide class of polynomials can serve for constructing canonical number systems. B. Kovács [72] proved that if $f(x)\in\mathbb{Z}[x]$ is irreducible, its zeroes have moduli greater than one and if $c_k\le c_{k-1}\le\ldots\le$

$c_0 \geq 2$ then $f(x)$ is a cns-polynomial. His proof can be applied for reducible polynomials as well. Moreover, if $c_0$ is "big enough" then S. Akiyama and A. Pethő gave a method determining the cns-property of arbitrary polynomials [1]. They also proved that if $c_2, \ldots, c_{k-1}, \sum_{i=1}^{k} c_i \geq 0$ and $c_0 > 2\sum_{i=1}^{k} |c_i|$ then $f(x)$ is a cns-polynomial and the last inequality can be replaced by $c_0 \geq 2\sum_{i=1}^{k} |c_i|$ when all $c_i \neq 0$.

Recently, H. Brunotte provided an algorithm [10], which attempt to prove the cns-property for a given irreducible monic polynomial $f(x) \in \mathbb{Z}[x]$ satisfying the root-condition. His algorithm works for arbitrary monic polynomials in $\mathbb{Z}[x]$ as well. His method differs essentially from the method of S. Akiyama and A. Pethő. Instead of using power basis he chose a different one. In H. Brunotte's basis the function $\overline{\Phi} : \mathbb{Z}^k \to \mathbb{Z}^k$ has the form

$$\overline{\Phi}([x_1, \ldots, x_k]^T) = [-\operatorname{sign}(c_0) \left\lfloor \frac{\sum_{j=1}^{k-1} c_j x_j + x_k}{|c_0|} \right\rfloor, x_1, \ldots, x_{k-1}]^T$$

His algorithm based on the following theorem. Suppose that the set $E \subseteq \mathbb{Z}^k$ has the recursive definition *(i)* $[0, \ldots, 0]^T, [-1, 0, \ldots, 0]^T, [0, \ldots, 0, -1]^T \in E$, *(ii)* for every $[x_1, \ldots, x_k]^T \in E$ and $d \in D = \{0, 1, \ldots, |c_0| - 1\}$ the element $\overline{\Phi}([x_1, \ldots, x_{k-1}, x_k + d]^T)$ belongs to $E$. If for every $e \in E$ there exists a $j_e \in \mathbb{N}_0$ such that $\overline{\Phi}^{j_e}(e) = 0$ then the polynomial $f(x)$ has the cns-property.

Let us see some examples. Let $k = 2$. Then by Lemma 8 and Lemma 9 we get that $-1 \leq c_1 \leq c_0$. It is easy to see that in these cases the roots of $f(x)$ are outside the complex unit disc. Using the previous algorithm of H. Brunotte it is also not hard to see that $E \subseteq \{[x_1, x_2]^T, x_1, x_2 \in \{-1, 0, 1\}\}$ and applying the function $\overline{\Phi}$ we have that the cns-property always holds. In fact, we got a kind of generalization of the result of I. Kátai, B. Kovács [52, 53] and of W. Gilbert [27].

If $k = 3$ then we are only able to write a set of inequalities between the coefficients of $f(x)$ (see also [1, 10]). Nevertheless, the following assertion holds.

**Assertion 6.** *The following polynomials are cns-polynomials in* $\mathbb{Z}[x]$*:*
*(i)* $x^k + c_1 x + c_0$ *for every* $k \geq 3$ *iff* $-1 \leq c_1 \leq c_0 - 2, c_0 \geq 2$*;*
*(ii)* $x^k + px^{k-1} + px^{k-2} + \ldots + px + p$ *for all* $2 \leq p \in \mathbb{N}$*;*
*(iii)* $x^k + x^{k-1} + x^{k-2} + \ldots + x + p$ *for all* $2 \leq p \in \mathbb{N}$*;*
*(iv)* $x^k + px^{k-1} + p^2x^{k-2} + \ldots + p^{k-1}x + p^k$ *for all* $2 \leq p \in \mathbb{N}$*.*
PROOF: The case (i) was proved in [10]. In order to check that the roots

of the polynomials (ii) and (iii) are outside the complex unit disc one can use the method of Lehmer-Schur [84]. The proof is easy, we leave it to the reader. It is also obvious that the moduli of the roots of polynomial (iv) are equal and greater than one. Since the coefficients of the polynomials (ii)-(iv) are positive and monotonically increasing, the theorem of B. Kovács can be applied. The proof is finished. □

*Remarks. (1)* We proved that there are infinitely many cns-polynomials (therefore canonical number systems) for each dimension $k$ even if the constant term of the polynomial is "small".

*(2)* The polynomials (iv) and (i) for $c_1 = 0$ show that for every $e > 1$ there is a base $M$ such that $(\Lambda, M, D)$ is a canonical number system and the moduli of each eigenvalues of $M$ are smaller than or equal to $e$. This shows that the second necessary condition in Assertion 1 for satisfying the unique representation property is sharp.

*(3)* Consider the Frobenius matrix $M$ of the polynomial (iv). Note that all eigenvalues of $M$ have the same moduli. The importance of these systems appears in chapter 5, in examining the Hausdorff dimension of the boundary of their fundamental domain.

### 3.2.5 Searching for cns-polynomials

Now we provide an algorithm for searching canonical number systems. To decide whether the polynomial $f(x)$ has a root inside the complex unit disc the method of Lehmer-Schur can be used. To analyze the possible roots in the unit circle we have the following well-known lemma.

**Lemma 10.** *Let* $Q(x) = q_0 + q_1 x + \ldots + q_k x^k \in \mathbb{Z}[x], Q(\gamma_i) = 0, |\gamma_i| \geq 1$. *Then* $|\gamma_i| > 1$ *if and only if* $\gcd(Q(x), x^k Q(1/x))$ *is a constant polynomial.*

ALGORITHM: CNS-SIEVE. Searching for all candidates of cns-polynomials in case of given inputs constant term $c_0$ and degree $k$ of the monic polynomial $f(x) \in \mathbb{Z}[x]$.

1. Let $S$ be the finite set of polynomials determined by Lemma 9;
2. **if** $S \neq \emptyset$ **then** $p$ :=get-a-new-candidate($S$); $S := S \setminus \{p\}$;
   **else goto** step 5;
3. **if** Lemma 8 (e), (b) with $z = -1$, (c) and (d) hold for the polynomial $p$
   **then goto** step 4; **else goto** step 2;
4. Apply Lehmer-Schur and Lemma 10 for the polynomial $p$;

**if** all roots of $p$ have moduli greater than one **then print**$(p)$;
**goto** step 2;
5. STOP;

The algorithm terminates since $S$ is a finite set. Observe that the CNS-SIEVE algorithm contains computationally easy-to-check methods. Moreover, if Lemma 8 fails for the polynomial $p$ then possibly more than one polynomials can be deleted from the set $S$, depending on which part of Lemma 8 does not hold. Clearly, the CNS-SIEVE algorithm can also be applied for $k > 9$ but in this case bounds for the coefficients of $f(x)$ must be determined.

### 3.2.6 Cns-polynomials with constant term $c_0 = 2$

Now we turn our attention to generalized binary number expansions, i.e. $c_0 = 2$. The case $k = 1$ is well-known, and the case $k = 2$ was analyzed in section 3.2.4. Let $k \geq 3$. Suppose that the polynomial $f(x)$ is obtained by the CNS-SIEVE ALGORITHM for some $k$. Then, a periodic element $0 \neq \pi \in \mathcal{P}$ would be a test proving that $f(x)$ is not a cns-polynomial. If one does not find such a $\pi$ by searching a small finite portion of the space systematically or randomly then one can use the CLASSIFICATION ALGORITHM or H. Brunotte's algorithm [10] to prove that $f(x)$ is really a cns-polynomial. If $f(x)$ is not a cns-polynomial then these algorithms serve also the test.

The author implemented the CNS-SIEVE ALGORITHM in C language. The following table shows the results up to degree 8.

| Degree $(k)$ | Output of CNS-SIEVE ALGORITHM (number of polynomials) | Number of cns-polynomials |
|---|---|---|
| 3 | 5 | 4 |
| 4 | 22 | 12 |
| 5 | 18 | 7 |
| 6 | 73 | 25 |
| 7 | 62 | 12 |
| 8 | 215 | 20 |

Table 1

Further, we enumerate the computed cns-polynomials.

$k = 3$, $2 - x + x^3, 2 + x^3, 2 + x + x^2 + x^3, 2 + 2x + 2x^2 + x^3$.

$k = 4$, $2 - x + x^4, 2 + x^4, 2 - x^2 + x^4, 2 + x^2 + x^4, 2 + 2x^2 + x^4, 2 + x + x^3 + x^4, 2 + x + x^2 + x^3 + x^4, 2 + 2x + x^2 + x^3 + x^4, 2 + x + 2x^2 + x^3 + x^4, 2 + 2x + 2x^2 + x^3 + x^4, 2 + 2x + 2x^2 + 2x^3 + x^4, 2 + 3x + 3x^2 + 2x^3 + x^4$.

$k = 5$, $2 - x + x^5, 2 + x^5, 2 - x + x^2 + x^5, 2 + x^2 + x^3 + x^5, 2 + x + x^4 + x^5, 2 + x + x^2 + x^3 + x^4 + x^5, 2 + 2x + 2x^2 + 2x^3 + 2x^4 + x^5$.

$k = 6$, $2 - x + x^6, 2 - x^2 + x^6, 2 - x^3 + x^6, 2 + x^6, 2 + x^3 + x^6, 2 + 2x^3 + x^6, 2 + x^2 - x^3 + x^4 + x^6, 2 + x^2 + x^4 + x^6, 2 + x^2 + x^3 + x^4 + x^6, 2 + 2x^2 + 2x^4 + x^6, 2 + x - x^2 - x^3 + x^5 + x^6, 2 + x - x^3 + x^5 + x^6, 2 + x + x^5 + x^6, 2 + x + x^2 + x^3 + x^4 + x^5 + x^6, 2 + 2x + x^2 + x^3 + x^4 + x^5 + x^6, 2 + 2x + 2x^2 + x^3 + x^4 + x^5 + x^6, 2 + x + x^2 + 2x^3 + x^4 + x^5 + x^6, 2 + 2x + 2x^2 + 2x^3 + x^4 + x^5 + x^6, 2 + x + 2x^2 + x^3 + 2x^4 + x^5 + x^6, 2 + 2x + 2x^2 + 2x^3 + 2x^4 + x^5 + x^6, 2 + 2x + 3x^2 + 2x^3 + 2x^4 + x^5 + x^6, 2 + 2x + 2x^2 + 2x^3 + 2x^4 + 2x^5 + x^6, 2 + 3x + 3x^2 + 3x^3 + 3x^4 + 2x^5 + x^6, 2 + 3x + 4x^2 + 4x^3 + 3x^4 + 2x^5 + x^6, 2 + x + x^2 + x^4 + x^5 + x^6$.

$k = 7$, $2 - x + x^7, 2 - 2x + 2x^2 - x^3 + x^5 - x^6 + x^7, 2 - x + x^2 + x^4 + x^7, 2 + x^3 + x^4 + x^7, 2 + x^2 + x^5 + x^7, 2 + x + x^6 + x^7, 2 + x + x^2 + x^3 + x^4 + x^5 + x^6 + x^7, 2 + 2x + 2x^2 + x^3 + x^4 + x^5 + x^6 + x^7, 2 + 2x + 2x^2 + 2x^3 + 2x^4 + x^5 + x^6 + x^7, 2 + 2x + 2x^2 + 2x^3 + 2x^4 + 2x^5 + 2x^6 + x^7, 2 + 3x + 4x^2 + 4x^3 + 4x^4 + 3x^5 + 2x^6 + x^7$.

$k = 8$, $2 - x + x^8, 2 - x^2 + x^8, 2 - x^4 + x^8, 2 + x^8, 2 + x^4 + x^8, 2 + 2x^4 + x^8, 2 + x^3 + x^5 + x^8, 2 + x^2 + x^6 + x^8, 2 + x^2 + x^4 + x^6 + x^8, 2 + 2x^2 + x^4 + x^6 + x^8, 2 + x^2 + 2x^4 + x^6 + x^8, 2 + 2x^2 + 2x^4 + x^6 + x^8, 2 + 2x^2 + x^3 + x^4 + x^5 + x^6 + x^8, 2 + 2x^2 + 2x^4 + 2x^6 + x^8, 2 + 3x^2 + 3x^4 + 2x^6 + x^8, 2 + x + x^7 + x^8, 2 + x + x^2 + x^4 + x^6 + x^7 + x^8, 2 + x + x^2 + x^3 + x^5 + x^6 + x^7 + x^8, 2 + x + x^2 + x^3 + x^4 + x^5 + x^6 + x^7 + x^8, 2 + 2x + x^2 + x^3 + x^4 + x^5 + x^6 + x^7 + x^8, 2 + 2x + 2x^2 + 2x^3 + x^4 + x^5 + x^6 + x^7 + x^8, 2 + 2x + 2x^2 + 2x^3 + 2x^4 + x^5 + x^6 + x^7 + x^8, 2 + 2x + 2x^2 + 2x^3 + 2x^4 + 2x^5 + 2x^6 + x^7 + x^8, 2 + x + x^2 + x^3 + 2x^4 + x^5 + x^6 + x^7 + x^8, 2 + x + 2x^2 + 2x^3 + x^4 + 2x^5 + x^6 + x^7 + x^8, 2 + x + 2x^2 + x^3 + 2x^4 + x^5 + 2x^6 + x^7 + x^8, 2 + x + 3x^2 + 2x^3 + 3x^4 + 2x^5 + 2x^6 + x^7 + x^8, 2 + 2x + 3x^2 + 3x^3 + 3x^4 + 2x^5 + 2x^6 + x^7 + x^8, 2 + 3x + 3x^2 + 3x^3 + 3x^4 + 3x^5 + 3x^6 + 2x^7 + x^8, 2 + 3x + 4x^2 + 5x^3 + 5x^4 + 4x^5 + 3x^6 + 2x^7 + x^8, 2 + 4x + 5x^2 + 5x^3 + 5x^4 + 4x^5 + 3x^6 + 2x^7 + x^8$.

The output of the CNS-Sieve Algorithm shows that the estimates in Lemma 8 and Lemma 9 may be complemented and improved. It is also clear that the time complexity of the algorithm is exponential in $k$. Moreover, in higher dimensions proving that a given polynomial obtained by the CNS-Sieve Algorithm is really a cns-polynomial is hard. The following conjecture would help, but the author was unable to prove this.

**Conjecture.** *Suppose that the lattice $\Lambda$ is generated with the power basis and the polynomial $f(x)$ is obtained by the* CNS-Sieve Algorithm. *If there does not exist any periodic element $\pi$ for which $\|\pi\|_\infty = 1$ then $f(x)$ is a cns-polynomial.*

Obviously, if such a $\pi$ exist then the polynomial is not a cns-polynomial. We used this idea to test the output of the CNS-SIEVE ALGORITHM.

*Remarks. (1)* The case $k = 3$ in Table 1 was known to A. Járai (unpublished).

*(2)* Suppose that the polynomial $f(x)$ is obtained by the CNS-SIEVE ALGORITHM and it is not a cns-polynomial. Then, the CLASSIFICATION ALGORITHM provides more than one periods. The following questions are quite interesting: how many such periods exist and what are the length of them? The general characterization seems to be hard. The following table shows some computational results.

| the polynomial $f(x)$ | $\pi \in \mathcal{P}$ $\|\pi\|_\infty = 1$ | the length of period of $\pi$ |
|---|---|---|
| $2 + x + x^2 + x^4$ | $[-1, 1, 0, 0]^T$ | 11 |
| $2 + x + 2x^2 + 2x^3 + x^4 + x^5$ | $[-1, -1, -1, 0, 0]^T$ | 21 |
| $2 + x + x^3 + x^4 + x^5 + x^6$ | $[-1, -1, -1, 0, 0, 0]^T$ | 33 |
| $2 + x + 2x^3 + 2x^4 + x^6 + x^7$ | $[-1, -1, 1, -1, 0, 1, 0]^T$ | 47 |
| $2 + 2x + x^2 + x^6 + 2x^7 + x^8$ | $[-1, -1, 0, 0, 0, 0, 0, 0]^T$ | 64 |

Table 2

*(3)* In order to decide the cns-property of a given polynomial the algorithm of H. Brunotte is preferable. The author is grateful to J. Sziliczi who programmed this algorithm in C++ in a very fine way. This shows among others that for the cns-polynomial $2 + x + 2x^2 + x^3 + 2x^4 + x^5 + 2x^6 + x^7 + x^8$ the algorithm uses 344 iteration steps, the number of integer vectors in the set $E$ is 143123, while for the cns-polynomial $2 + 3x + 3x^2 + 3x^3 + 3x^4 + 3x^5 + 3x^6 + 2x^7 + x^8$ the algorithm uses 253 iteration steps and number of integer vectors in the set $E$ is 241719.

### 3.2.7 Polygonal construction

Let $f(x) = x^k + x^{k-1} + \ldots + 1$ and let $\Lambda_f = \mathbb{Z}[x]/(f)$ be the corresponding $k$-dimensional lattice as earlier. Let $\omega = x + (f)$ denote the image of $x$ in $\Lambda_f$ and let $\theta = n - \omega, n \in \mathbb{Z}$. Note that $\omega^{k+1} = 1$. Clearly, the corresponding

matrix

$$M_{pol} = \begin{pmatrix} n & 0 & 0 & \dots & 0 & 1 \\ -1 & n & 0 & \dots & 0 & 1 \\ 0 & -1 & n & \dots & 0 & 1 \\ \vdots & \vdots & \vdots & \dots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & n & 1 \\ 0 & 0 & 0 & \dots & -1 & n+1 \end{pmatrix}$$

acts on the cubic lattice $\mathbb{Z}^k$ with respect to the basis $\{1, \omega, \dots, \omega^{k-1}\}$. The determinant of $M_{pol}$ is $(n^{k+1}-1)/(n-1)$. A. Vince [106] considered the interesting case $n = 2$. In this case $\det(M_{pol}) = 2^{k+1} - 1$. Let

$$D = \{\epsilon_0 + \epsilon_1\omega + \ldots + \epsilon_k\omega^k : \epsilon_i \in \{0,1\}, \text{ not all } \epsilon_i \text{ is } 1\}.$$

It can be seen that $D$ is a full residue system modulo $M_{pol}$. For $k = 1$ we have that $\Lambda = \mathbb{Z}, M_{pol} = (3), D = \{-1, 0, 1\}$, which is the *balanced ternary* representation of the integers. For $k = 2$ the matrix $M_{pol} = \left(\begin{smallmatrix} 2 & 1 \\ -1 & 3 \end{smallmatrix}\right)$ and

$$D = \{0, 1, \omega, \dots, \omega^5 : \omega \text{ is a primitive 6th root of unity}\}.$$

A. Vince called these systems as the *generalized balanced ternary* (GBT). In these systems, addition and multiplication can be carried out by simple and fast bit string routines, since each digit can be represented by the binary string $\epsilon_0\epsilon_1 \dots \epsilon_k$. Moreover, using the two-dimensional GBT, a planar database management system was developed (see [105, 106] and the references there). We call the reader's attention to an interesting fact regarding generalized balanced ternary, which was observed by A. Vince. The eigenvalues of $M_{pol}$ for the GBT are $\{2 - \omega : \omega \text{ is an } (n+1)\text{th root of unity}, \omega \neq 1\}$. Therefore the minimum modulus of an eigenvalue tends to 1 as $k \to \infty$. Since GBT systems are number systems for all $k$, we got again that Assertion 1(b) is sharp.

Radix systems, where the digit set has the form $\{0, 1, \zeta, \zeta^2, \dots, \zeta^{k-1}\}$, $\zeta = \exp(2\pi i/k)$ is the primitive $k$-th root of unity, are very important in computer science, since they enables fast addition and on-line multiplication. We refer the interested reader to [98, 100].

## 3.3 Simultaneous construction

The following radix system was introduced by K-H. Indlekofer, I. Kátai and P. Racskó [41]. Let $N_1, N_2, \dots, N_k$ be mutual co-prime integers, none

of them is $0, \pm 1$. Let $M_s = \mathrm{diag}(N_1, N_2, \ldots, N_k)$ and $D = \{\delta e\}$, where $e = [1, \ldots, 1]^T, \delta = 0, 1, \ldots, t-1, t = \mid N_1 \ldots N_k \mid$. Clearly, the set $D$ is a full residue system modulo $M_s$. The proper work of the function $\Phi$ is based on the Chinese remainder theorem. In dimension two let $2 \leq N_1 < N_2$. The above mentioned authors proved that the system $(\mathbb{Z}^k, M_s, D)$ is a number system if and only if $N_2 = N_1 + 1$.

## 3.4 General construction

A further question concerning radix expansions is the following: for a given $M$ satisfying criterion (b) and (c) in Assertion 1 is there any digit set $D$ for which $(\Lambda, M, D)$ is a number system? How many such digit sets exist and how to construct them? In imaginary quadratic fields due to G. Steidl [102] and I. Kátai [48] we know that to be able to construct number systems the conditions in Assertion 1 are also sufficient. Remarkable results are obtained by G. Farkas in real quadratic fields [20, 21, 23]. Moreover, if $M$ is similar to the Frobenius matrix of an irreducible monic polynomial over $\mathbb{Z}$ then some results are also available [50]. The above mentioned authors gave the constructions as well. For the general case, A. Vince proved [106] that if all the singular values of $M$ are greater than $3\sqrt{k}$ then the digit set $D$ can be constructed. In dimension 2 this value can be made sharper to 2. Now we prove the following.

**Assertion 7.** *(Sufficient condition for the number system property)*
*Suppose that the conditions for $M, D$ in Assertion 1 hold. Let us denote in $\mathbb{R}^k$ a vector norm and the corresponding operator norm by $\|\cdot\|$ for which $r = \|M^{-1}\| < 1$. Let $K = \max\{\|d\|, d \in D\}$ and $L = Kr/(1-r)$. Let furthermore $R$ be a positive real number for which $z \in \Lambda, \|z\| \leq R$ implies $z \in D$. If $r \leq R/(R+K)$ then $(\Lambda, M, D)$ is a number system.*

PROOF: It follows from Lemma 1 that if $\pi$ is a periodic element then $\|\pi\| \leq L$. Hence, if we could prove that $L \leq R$ then we would be ready, since in this case the only periodic element is the null vector. But if $r = \|M^{-1}\| \leq R/(R+K)$ then $Kr \leq R(1-r)$, by which $L = Kr/(1-r) \leq R$. □

The construction of the digit set is as follows: enumerate all integers in a 'big enough' ball around the origin, order them using the appropriate norm and select a full residue system keeping the norm of the elements as small as possible.

Assertion 7 has an important corollary. Recall that a basis transform-

ation does not change the number system property, i.e. if $M_1$ and $M_2$ are similar via the matrix $Q$ then the number system property of $(\Lambda, M_1, D)$ and $(Q\Lambda, M_2, QD)$ holds at exactly the same time. Let $U = [-\frac{1}{2}, \frac{1}{2})^k$ denote the $k$-dimensional half-open unit cube centered at the origin. Recall that the $k$-dimensional parallelotop $V = MU$ has volume $|\det(M)|$ and the appropriate integers in $V$ constitute a full residue system modulo $M$. Suppose that the norm in $\mathbb{R}^k$ is the Euclidean norm. Then, performing a basis transformation, the full residue system $V$ can be transformed to the half-open unit cube $U$, in which case $\frac{K}{R}$ is equal to $\sqrt{k}$. Hence, we proved the following:

**Assertion 8.** *For a given expansive $M$ suppose that $\|M^{-1}\|_2 \leq 1/(1+\sqrt{k})$. Then there exists a digit set $D$ for which $(\Lambda, M, D)$ is a number system.*

Our result is stronger than that one of A. Vince except in dimension 2. Applying Assertions 1 and 8 in dimension 1 shows that if $2 < \theta \in \mathbb{Z}$ then every rational integer has a unique base $\theta$ radix representation with $D = \{-\lfloor(|\theta| - 1)/2\rfloor, \ldots, \lfloor|\theta|/2\rfloor\}$, which is well-known. Consider the ring of Gaussian integers $\mathbb{Z}[i] = \{a + bi : a, b \in \mathbb{Z}\}$ and let $\theta = A + Bi \in \mathbb{Z}[i]$. In this case $M_\theta = \left(\begin{smallmatrix} A & -B \\ B & A \end{smallmatrix}\right)$ and $\|M_\theta^{-1}\|_2 = 1/\sqrt{A^2 + B^2}$, which is, apart from a few cases, always smaller than $1/(1+\sqrt{2})$. Keeping in mind Assertion 1, Assertion 7 and [55] these cases are easy to handle. We got the following: for any Gaussian integer $\theta$ of modulus larger than one, except 2 and $1 \pm i$, there exists a full residue system $D$ so that $(\mathbb{Z}^2, M_\theta, D)$ is a number system. Hence, as a special case of Assertion 8 we have the result of G. Steidl[1]. If we consider the Eisenstein integers $\mathbb{Z}[\omega] = \{a + b\omega : a, b \in \mathbb{Z}\}$, where $\omega$ is the complex cube root of unity, and we perform the above mentioned computations, we obtain the same conclusion. Nevertheless, it is not any surprise: I. Kátai solved the problem in all imaginary quadratic fields. If we consider the real quadratic fields — without going into the details — it is possible to reprove the result of G. Farkas [20]. The interesting is that the above mentioned authors gave the digit sets explicitly which is different from our construction. This suggests that the unique representation property depends mainly on the radix, and if any, than several different digit sets can be constructed.

[1] Historical remark: for the first proof of this result there is a research report by M. Davio, J.P. Deschamps and C. Gossart [14] dated back to 1978.

# Chapter 4

# Analyzing expansions in $Q[i\sqrt{F}]$

*"The imaginary number is a fine and*
*wonderful recourse of the divine spirit,*
*almost an amphibian between being and not being."*
*— G. W. Leibniz*

In this chapter we analyze the attractor set of special radix systems. Using the notations already adopted the following questions arise: (a) What can be stated about the attractor set of an arbitrary radix system $(\Lambda, M, D)$? (b) How the structure of the periodic elements looks like? (c) It is known that if $\pi \in \mathcal{P}$ then the maximum of the period length of $\pi$ can be estimated with the number of lattice points covered by the disk with radius $L$ centered at the origin. Is there a better estimation? (d) Is there a good upper estimation for the number of the different sets $\mathcal{C}(\pi)$? The purpose of this chapter is to answer these questions using bases as integers in imaginary quadratic fields and canonical digit sets. It must be noted that the results of this section for the case of Gaussian integers was proved in the author's paper [65] using a different technique. We remark that there are also some results in the real quadratic field $\mathbb{Q}(\sqrt{2})$ using a different kind of digit set [22, 23].

Let $F = 1$ or $F \geq 2$ be a square-free integer. Let $\mathbb{Q}(i\sqrt{F})$ be an imaginary quadratic extension of $\mathbb{Q}$, $I$ be the set of integers in $\mathbb{Q}(i\sqrt{F})$. It is known, that if $F \not\equiv 3 \pmod 4$ then $\{1, \delta\}$, while for $F \equiv 3 \pmod 4$ $\{1, \omega\}$ is an integer basis in $I$, where $\delta = i\sqrt{F}$, $\omega = (1 + i\sqrt{F})/2$. The lattice generated by the basis $\{1, \delta\}$ will be called the $\delta$-lattice and denoted by $\Lambda_\delta$, while the lattice generated by the basis $\{1, \omega\}$ is the $\omega$-lattice $\Lambda_\omega$.

Let $\alpha_1 = a + b\delta$ and $\alpha_2 = a + b\omega$, $a, b \in \mathbb{Z}$, $b \neq 0$, $E = (F+1)/4$. In these cases the corresponding linear operators in $\mathbb{Z}^2$ are $M_1 = \left(\begin{smallmatrix} a & -Fb \\ b & a \end{smallmatrix}\right)$ and $M_2 = \left(\begin{smallmatrix} a & -Eb \\ b & a+b \end{smallmatrix}\right)$. Clearly, $\det(M_1) = a^2 + Fb^2$ and $\det(M_2) = a^2 + ab + Eb^2$; the first column of the adjoint of the matrices $M_1$ and $M_2$ are $[a, -b]^T$ and $[a+b, -b]^T$, accordingly. Suppose that $\gcd(a, b) > 1$. It follows from Theorem 2 that in these cases the sets $\{0, 1, \ldots, a^2 + Fb^2 - 1\}$ and $\{0, 1, \ldots, a^2 + ab + Eb^2 - 1\}$ can not be complete residue systems modulo $M_1$ and $M_2$, accordingly. Hence the following lemma holds.

**Lemma 11.** *For a given* $\alpha \in \mathbb{Q}[i\sqrt{F}]$ ($\alpha = a + b\delta$ *or* $\alpha = a + b\omega$) *the set* $D = \{0, 1, \ldots, \mathrm{Norm}(\alpha) - 1\}$ *is a complete residue system if and only if* $\gcd(a, b) = 1$.

Throughout this chapter we shall always assume that $\gcd(a, b) = 1$. For the sake of brevity we use the notation $(x, y)$ for $\gcd(x, y)$.

## 4.1 Periodic elements of period length one

Consider the $\delta$-lattice and let $\alpha = a + b\delta,\ a, b \neq 0,\ (a, b) = 1$.

**Lemma 12.1.** *In the system* $(\Lambda_\delta, \alpha, D)$ *the periodic elements of period length one are* $\pi_j = \frac{1-a+b\delta}{(1-a,b)} j,\ j = 0, \ldots, k$, *where* $k = \left\lfloor (1-a, b)(1 + 2\frac{a-1}{(1-a)^2+b^2F}) \right\rfloor$.

Proof: It follows from (1.4) that $\pi \in \mathcal{P}$ is a periodic element of period length one if and only if $\pi = d + \alpha\pi$ for some $d \in D$. It means that $(1-\alpha)\pi = d \in D$, hence $\pi = \frac{d}{1-\alpha} = \frac{d(1-a)}{(1-a)^2+b^2F} + \delta\frac{db}{(1-a)^2+b^2F}$. Since $\pi \in I$ therefore $(1-a)^2 + b^2F \mid d(1-a, b)$. On the other hand $0 \leq d \leq a^2 + b^2F - 1$ by which the proof is completed. □

Consider now the $\omega$-lattice. Let $\alpha = a + b\omega,\ b \neq 0,\ (a, b) = 1,\ N = \mathrm{Norm}(\alpha) = a^2 + ab + b^2E,\ E = (F+1)/4$. If $E = 1$, $a = 0, b = \pm 1$ or $E = 1, a = -b = \pm 1$ then $|\alpha| = 1$, so in the following we always exclude these cases. Using the same idea as before the next lemma can be easily proved. We leave it to the reader.

**Lemma 12.2.** *In* $(\Lambda_\omega, \alpha, D)$ *the periodic elements of period length one are* $\pi_j = \frac{1-a-b+b\omega}{(1-a-b,b)} j,\ j = 0, \ldots, k$, *where* $k = \left\lfloor (1-a-b, b)(1 + \frac{2a+b-2}{(1-a)^2-(1-a)b+b^2E}) \right\rfloor$.

*Remarks. (1)* Let $b > 0$ be fixed. From these lemmas we can calculate the maximal number of loops. In the $\delta$-lattice this can be achieved by $b \mid a - 1$, $a \geq 1$, in which case it is $b + 1$ if $F \geq 2$ and $b + 2$ if $F = 1$. In the $\omega$-lattice we have two cases depending on the value of $E$. If $E \geq 2$ then the maximal

number of loops is $b+1$ by $b \mid a-1, 2a+b \geq 2$. If $E=1$ then this value is $b+2$, by $a=1$ or by $b \geq 1, a=b+1$.

*(2)* If $a$ is positive then the element $1-a+b\delta \in \mathcal{P}$ of period length one. In the $\omega$-lattice, if $2a+b \geq 2$ then the element $1-a-b+b\omega \in \mathcal{P}$ of period length one. Moreover, if $E=1, \#\mathcal{P}=b+2$ then $(1-a-b)(b+1)/b+(b+1)\omega \in \mathcal{P}$ of period length one.

## 4.2 Location of periodic elements

Before we continue our analysis, we have some useful observations.
(1) Let $\gamma \in I,\ \gamma \equiv 0 \bmod \alpha, \gamma_1 = \gamma + x,\ \gamma_2 = \gamma + y,\ x, y \in D$. Then

$$\Phi(\gamma_1) = \Phi(\gamma_2). \tag{4.1}$$

(2) Let $\alpha \in I, \pi \in \mathcal{P}$, that is, $\pi = a_0 + a_1\alpha + \ldots + a_{l-1}\alpha^{l-1} + \pi\alpha^l,\ a_j \in D$. Then

$$\overline{\pi} = a_0 + a_1\overline{\alpha} + \ldots + a_{l-1}\overline{\alpha}^{l-1} + \overline{\pi\alpha}^l, \quad a_j \in D. \tag{4.2}$$

It means that if $\pi \in \mathcal{P}$ in $(\Lambda, \alpha, D)$ then $\overline{\pi} \in \mathcal{P}$ in $(\Lambda, \overline{\alpha}, D)$. If $\alpha = a+b\delta$ then $\overline{\alpha} = a-b\delta$, if $\alpha = a+b\omega$ then $\overline{\alpha} = a+b-b\omega$, so it is enough to examine the cases $b \geq 1$.
(3) It follows from Lemma 3 that if $\pi \in \mathcal{P}$ then

$$-\pi = \frac{d_1}{\alpha} + \frac{d_2}{\alpha^2} + \frac{d_3}{\alpha^3} + \cdots$$

for some $d_i \in D$. It means that

$$\left| -\pi - \frac{d_1}{\alpha} \right| \leq \sum_{i=2}^{\infty} \frac{|d_i|}{|\alpha|^i} \leq \frac{(N-1)}{|\alpha|^2} \frac{1}{1-1/|\alpha|} = \frac{|\alpha|+1}{|\alpha|} = 1 + \frac{1}{|\alpha|}.$$

Hence,

$$\left| -\pi - \frac{d_1\overline{\alpha}}{N} \right| \leq 1 + \frac{1}{|\alpha|}. \tag{4.3}$$

**Lemma 13.** *Let* $\alpha \in I$ ($\alpha = a + b\delta$ *or* $\alpha = a + b\omega$), $(a, b) = 1$, $e \in \mathbb{Z}$. *If* $\alpha \mid e$ *then* $N = \alpha\overline{\alpha} \mid e$.
PROOF: If $a+b\delta = \alpha \mid e$ then $(a+b\delta)(c+d\delta) = ac-bdF+(ad+bc)\delta = e$ for some $c, d \in \mathbb{Z}$. Since $(a, b) = 1$ therefore $c = ap$ and $d = -bp$ for some $p \in \mathbb{Z}$. Hence $a^2p + b^2Fp = e$, which means that $a^2 + b^2F \mid e$. If $a + b\omega = \alpha \mid e$ then $(a + b\omega)(c + d\omega) = ac - bdE + (ad + bc + bd)\omega = e$ for some $c, d \in \mathbb{Z}$. Again, since $(a, b) = 1$ therefore $c = (a + b)p$ and $d = -bp$ for some $p \in \mathbb{Z}$. It means that $a(a + b)p + Eb^2p = e$, by which the proof is finished. □

### 4.2.1 Case $\alpha = a + ib\sqrt{F}$

Let $\pi = U + V\delta \in \mathcal{P}$ and let $\Phi(\pi) = U_1 + V_1\delta$. By the definition of $\Phi$ we have the following equations:

$$U = d + aU_1 - bFV_1 \tag{4.4}$$
$$V = bU_1 + aV_1, \tag{4.5}$$

for some $d \in D$. On the other hand using (4.3) we have that

$$\left| \left( -U - \frac{d_1 a}{N} \right) + \left( -V + \frac{d_1 b}{N} \right)\delta \right| \leq 1 + \frac{1}{|\alpha|}. \tag{4.6}$$

**Theorem 3.1.** Let $\alpha = a + b\delta$, $f = -a$, $b \geq 1$, $S_1 = \{U + V\delta,\ -a + 1 \leq U \leq 0,\ 0 \leq V \leq b\}$, $S_2 = \{U + V\delta,\ 0 \leq U \leq f,\ 0 \leq V \leq b - 1\}$. Let $\pi = U + V\delta \in \mathcal{P}$. If $a \geq 1$ and ($F \geq 2$ or $F = 1, a \neq b + 1$) then $\pi \in S_1$, if $a \geq 1, F = 1, a = b + 1$ then $\pi \in S_1 \cup \{-a + ai\}$, if $-a \geq 1$ then $\pi \in S_2$.
PROOF: Let $|\alpha| \geq 2$. Suppose that $|U| \geq |a| + 2$. Then using (4.6) we get that

$$|a| + 2 \leq |U| \leq \frac{3}{2} + \frac{|d_1 a|}{N} \leq \frac{3}{2} + \frac{N - 1}{N}|a|,$$

which is a contradiction. Therefore $|U| \leq |a| + 1$. Now suppose that $F \geq 4$ and $|V| \geq |b| + 1$. Then

$$b + 1 \leq |V| \leq \frac{3}{2|\delta|} + \frac{|d_1|b}{N} \leq \frac{3}{2\sqrt{F}} + \frac{N - 1}{N}b,$$

which is a contradiction again. Hence, if $F \geq 4$ then $|V| \leq b$. In the same way, if $F = 1$ or $F = 2$ then it is easy to see that $|V| \leq b + 1$. On the other

hand it follows from (4.6) that

$$\left| -U - \frac{d_1 a}{N} \right| \leq \frac{3}{2},$$

therefore if $a > 0$ then $U \leq 1$, if $a < 0$ then $U \geq -1$. It is obvious as well that

$$\left| -V + \frac{d_1 b}{N} \right| \leq \frac{3}{2|\delta|},$$

hence if $F > 2$ then $V \geq 0$, if $F = 1$ or $F = 2$ then $V \geq -1$.

Case $a \geq 1$. If $a = 1, b = 2, F = 1$ then it is easy to check that $\mathcal{G}(\mathcal{P}) = \{0 \to 0, i \to i, 2i \to 2i\}$, so the theorem holds. In the following we exclude this case. Let $F = 1$ or $F = 2$. Consider equation (4.5) and suppose that $V_1 = b + 1$. Then we have that $V = b(a + U_1) + a \leq b + 1$, therefore either $U_1 = 0, a = 1$ or $U_1 \leq -a$. In the first case, if $a = 1$ then $b > 1$ and by (4.4) we get that $U \leq -bF(b + 1) + N - 1 = -bF < -2 = -(a + 1)$ which is a contradiction. It means that $(b+1)\delta$ can not be periodic. On the other hand, if $U_1 = -a-1$ then $U \leq -a^2-a-bF(b+1)+N-1 = -a-bF-1 < -(a+1)$ which is a contradiction again. Let $U_1 = -a$ and $F = 2$. Then $V = a \leq b+1$, therefore if $a \geq 3$ then $U \leq -a^2 - bF(b + 1) + N - 1 = -bF - 1 \leq (-a + 1)F - 1 = -2a + 1 < -(a + 1)$ which is not possible. If $a = 2$ then $a = b + 1$, therefore $b = 1$ and it is easy to check that in this case $\mathcal{G}(\mathcal{P}) = \{-1 + \delta \to -1 + \delta, 0 \to 0\}$. If $a = 1, b \geq 2$ then $U \leq -2b - 1 \leq -5$ which is a contradiction again. Let $U_1 = -a$ and $F = 1$. Then $V = a \leq b+1$ and $U \leq -b - 1 \leq -a$ hence in both cases equality must be satisfied, i.e. $a = b + 1$. But now $U = U_1 = -a, V = V_1 = b + 1$ and this is the only periodic element with $V = b+1$. Hence if $U + V\delta \in \mathcal{P}$ then $V \leq b$ except the case $F = 1, a = b + 1$, in which case $U + V\delta = -a + ai$. Suppose now that $F = 1$ or $F = 2$ and $V_1 = -1$. Then by (4.5) we get that $-1 \leq V = bU_1 - a$, therefore $U_1 \geq 0$. If $U_1 = 1$ or $U_1 = 0, bF \geq 2$ then $U \geq aU_1 + bF \geq 2$ which is a contradiction. If $U_1 = 0, b = F = 1$ then by (4.5) we have that $-1 \leq V = -a$, hence $a = 1$ which is a contradiction again. Let $F \geq 1$ and suppose that $U_1 = 1$. Then by (4.5) we have that $V = b + aV_1 \leq b$. Hence $V_1 = 0$, but obviously 1 can not be periodic. Suppose that $U_1 = -a-1$. Then (4.5) shows that $V \leq -b$ which is impossible. If $U_1 = -a$ then we have that $V_1 \geq b$. It follows from Lemma 13, (4.1) and from the remark of Lemma 12.2 that if $x \in D$ then $-x - 1 + \alpha \in \mathcal{B}(1 - \overline{\alpha})$. Lastly, since $2a + 1 \leq a^2 + b^2F$, therefore $-a + b\delta$ can be periodic iff $F = 1, a = b + 1$.

Case $-a = f \geq 1$. Suppose that ($F = 1$ or $F = 2$) and $V_1 = b + 1$. Then $-1 \leq V = bU_1 - f(b+1) = b(U_1 - f) - f$, therefore $f - 1 \leq b(U_1 - f)$, so $U_1 = f = 1$ or $U_1 \geq f+1$. In the first case it follows from (4.4) that $U \leq -1 - bF(b+1) + N - 1 = -bF - 1 < -1$ which is a contradiction. In the second case, if $U_1 = f+1$ then $U \leq -f^2 - f - bF(b+1) + N - 1 = -f - bF - 1 < -1$ which is not possible as well. Hence if $U + V\delta \in \mathcal{P}$ then $V \leq b$. Now suppose that ($F = 1$ or $F = 2$) and $V_1 = -1$. Then by (4.5) we get that $V = bU_1 + f \leq b$, therefore $U_1 \leq 0$ and if $U_1 = 0$ then $f \leq b$, if $U_1 = -1$ then $f \leq 2b$. If $F = 2$ then using (4.4) we have that $f+1 \geq U \geq -fU_1 + bF \geq 2f$ and equality holds iff $f = b = 1$ which is a contradiction. If $F = 1$ then $f + 1 \geq U \geq -fU_1 + b$. Clearly, if $U_1 = -1$ then $b = f = 1$, which is impossible. If $U_1 = 0$ then $f + 1 \geq b \geq f$, and since $(f, b) = 1$ therefore $b = f + 1$ ($b = f = 1$ is not valid). Hence $U = f + 1, V = f$. But if $U_1 = f + 1, V_1 = f, b = f + 1$ then $V = b(f + 1) - f^2 = 2f + 1 \leq f + 1$, which is a contradiction again. It is known [55] that if $f = 2, b = 1, F = 1$ then $\mathcal{G}(\mathcal{P}) = \{0 \to 0\}$. Excluding this case it is also clear that $a + x + b\delta \in \mathcal{B}(0)$ ($x \in D$) and $2f + 2 \leq f^2 + b^2F$, therefore by (4.1) we have that $V_1 \leq b-1$. If $U_1 = -1$ then $0 \leq V = -b - fV_1$, therefore $V_1 < 0$, which is a contradiction. Suppose that $U_1 = f + 1$. Then using (4.5) we get that $V = b(f + 1) - fV_1 \leq b - 1$, therefore $V_1 > b$, which is a contradiction as well.

If $|\alpha| < 2$ then keeping in mind [52, 53] we have to check only the following cases. If $a = b = 1, F = 1$ or $F = 2$ then it is easy to see that $\mathcal{G}(\mathcal{P}) = \{\delta \to \delta, 0 \to 0\}$. The proof is complete. □

**Lemma 14.1.** *If $a \geq 1$ then $\#\mathcal{P} \leq b + 1$, if $-a \geq 1$ then $\#\mathcal{P} \leq b$.*

PROOF: We have seen that if $a \geq 1$ and $\pi = U + b\delta \in \mathcal{P}$ then $\pi = 1 - \overline{\alpha}$. It is obvious that $d \in \mathcal{B}(0)$ for each $d \in D$. Now we shall examine the expansion of $-1$. Clearly, $-1 = -1 + N - \alpha\overline{\alpha}$ and $-\overline{\alpha} = -2a + \alpha$. Since $a \geq 1$ therefore $-2a + \alpha = -2a + N - \alpha\overline{\alpha} + \alpha$. Moreover, $0 < N - 2a < N - 1$ and $1 - \overline{\alpha} \in \mathcal{P}$ therefore $-1 \in \mathcal{B}(1 - \overline{\alpha})$. Hence the only rational integer periodic element is 0. Considering Theorem 3.1 observe that there does not exist any $\rho \in S_1 \cup S_2$, ($\rho \neq 0$) for which $\rho \equiv 0$ ($\alpha$). In virtue of (4.1) it is easy to see that if $U + V\delta \in \mathcal{P}$ then there is not any $Z, (Z \neq U)$ for which $Z + V\delta \in \mathcal{P}$. The proof is finished. □

### 4.2.2 Case $\alpha = a + b\omega$

Let $\pi = U + V\omega \in \mathcal{P}$ and let $\Phi(\pi) = U_1 + V_1\omega$. By the definition of $\Phi$ we have the equations

$$U = d + aU_1 - bEV_1 \tag{4.7}$$
$$V = b(U_1 + V_1) + aV_1, \tag{4.8}$$

for some $d \in D$. On the other hand using (4.3) we have that

$$\left| \left( -U - \frac{d_1(a+b)}{N} \right) + \left( -V + \frac{d_1 b}{N} \right)\omega \right| \leq 1 + \frac{1}{|\alpha|}. \tag{4.9}$$

**Theorem 3.2.** *Let* $\alpha = a + b\omega,\ f = -a,\ b \geq 1,\ T_1 = \{U + V\omega,\ -a - b + 1 \leq U \leq 0,\ 0 \leq V \leq b - 1\},\ T_2 = \{U + V\omega, 0 \leq U \leq f - b,\ 0 \leq V \leq b - 1\}$. *Let* $\pi = U + V\omega \in \mathcal{P}$.
*If* $E = 1, a = 1$ *then* $\pi \in T_1 \cup \{1 - \overline{\alpha}, -b - 1 + (b+1)\omega\}$,
*if* $E = 1, a = b + 1$ *then* $\pi \in T_1 \cup \{1 - \overline{\alpha}, -a - b - 1 + (b+1)\omega\}$,
*if* $E \geq 2$ *or* $E = 1, a > 1$ *and* $a \neq b + 1$ *then* $\pi \in T_1 \cup \{1 - \overline{\alpha}\}$,
*if* $1 \leq f < b$ *and* $2a + b \geq 2$ *then* $\pi \in T_1 \cup \{1 - \overline{\alpha}\}$,
*if* $1 \leq f < b$ *and* $2a + b < 2$ *then* $\pi \in T_1$,
*if* $f > b$ *then* $\pi \in T_2$.
PROOF: Let $|\alpha| \geq 3$. Suppose that $|U| \geq |a + b| + 2$. Then using (4.9) we get that

$$|a + b| + 2 \leq |U| \leq \frac{4}{3} + \frac{|d_1(a+b)|}{N} \leq \frac{4}{3} + \frac{N-1}{N}|a + b|,$$

which is a contradiction. Therefore $|U| \leq |a + b| + 1$. Suppose that $E \geq 2$ and $|V| \geq |b| + 1$. Then

$$b + 1 \leq |V| \leq \frac{4}{3|\omega|} + \frac{|d_1|b}{N} \leq \frac{4}{3\sqrt{E}} + \frac{N-1}{N}b,$$

which is a contradiction again. Hence, if $E \geq 2$ then $|V| \leq b$. In the same way, if $E = 1$ then it is easy to see that $|V| \leq b + 1$. On the other hand it follows from (4.9) that

$$\left| -U - \frac{d_1(a+b)}{N} \right| \leq \frac{4}{3},$$

therefore if $a+b>0$ then $U\leq 1$, if $a+b<0$ then $U\geq -1$. It is obvious as well that

$$\left| -V+\frac{d_1 b}{N}\right| \leq \frac{4}{3|\omega|},$$

hence if $E\geq 2$ then $V\geq 0$, if $E=1$ then $V\geq -1$.

Case $a\geq 1$. Let $E=1$. Consider equation (4.8) and suppose that $V_1=b+1$. Then we have that $V=b(U_1+a+b+1)+a\leq b+1$. Hence either $a=1, U_1=-b-1$ or $U_1=-a-b-1, 1\leq a\leq b+1$. If $a=1, U_1=-b-1$ then by Lemma 12.2 we have that $-b-1+(b+1)\omega\in\mathcal{P}$ of period length one. If $U_1=-a-b-1$ then by (4.7),(4.8) we get that $1\leq V=a\leq b+1$ and $U=-a-b-1$. Now, suppose that $U_1=-a-b-1, V_1=a$. In virtue of (4.8) we have that $-1\leq V=-b^2-b+a^2$. It means that $b(b+1)\leq a^2+1$, therefore $a=b+1$. Hence, if $a=1$ then $-b-1+(b+1)\omega\in\mathcal{P}$, if $a\geq 2$ and $a=b+1$ then $-a-b-1+(b+1)\omega\in\mathcal{P}$ of period length one and does not exist any other periodic element $X+Y\omega$ with $Y=b+1$. Let $E=1$ and $V_1=-1$. Then by (4.8) we have that $-1\leq V=b(U_1-1)-a$, therefore $U_1=1, a=1$. Using (4.7) we get that $U\geq b+1\geq 2$ which is a contradiction. Let furthermore $E\geq 1$. Since $2a+b\geq 2$ always holds, therefore by the remark of Lemma 12.2 the element $1-a-b+b\omega\in\mathcal{P}$ of period length one. Clearly, $a^2+ab+b^2E>2a+b+1$ therefore there is not any other element $X+Y\omega\in\mathcal{P}$ with $Y=b$. Suppose that $U_1=1$ Then by (4.8) we have that $0\leq V=V_1(a+b)+b\leq b-1$ which is a contradiction. Suppose that $U_1=-a-b-c, (c=-1,0,1)$ and $0\leq V_1\leq b-1$. Then by (4.8) we get that $V=b(-a-b-c+V_1)+aV_1\leq b(-a-1-c)+ab-a=-a-b-bc<0$ which is a contradiction again.

Case $-a=f\geq 1$. Let $E=1$. Suppose that $V_1=b+1$. Then by (4.8) we have that $-1\leq V=b(U_1+b-f+1)-f\leq b+1$, therefore either $f=1, U_1=-b\leq -4, V=-1$ or $U_1\geq f-b$. In the first case using (4.7) we get that $U\leq b-b(b+1)+N-1=-b$. Suppose that $U_1=-b, V_1=-1, f=1$. It follows from (4.8) that $V=-b^2-b+1<-1$ which is a contradiction. In the second case, by (4.7) we get that $U\leq -fU_1-b(b+1)+N-1\leq -b-1<-b$ which is a contradiction as well. Hence if $U+V\omega\in\mathcal{P}$ then $V\leq b$. Suppose that $V_1=-1$. Then using (4.8) we have that $-1\leq V=b(U_1-1)+f\leq b$ therefore $U_1\leq 1$. Since $U\geq -fU_1+b$ therefore $U_1\geq 0$. If $U_1=0$ then $b-1\leq f\leq 2b$ and by (4.8) we get that $V=f-b$. Suppose that $U_1\geq b, V_1=f-b$. Then by (4.8) we have that $V=b(U_1+f-b)-f(f-b)=2fb-f^2+bc\leq b$ $(c\geq 0)$. It means that $c=0$ or 1. Moreover, in both cases the only solution

is $2b = f$, which contradicts either to $(f, b) = 1$ or to $|\alpha| \geq 3$. Suppose that $U_1 = 1, V_1 = -1$. It follows from (4.7), (4.8) that $U \geq b - f$ and $V = f \leq b$. This can happen iff $b = f + 1$. Now, suppose that $U_1 = 1, V_1 = b - 1$. Then using (4.8) we get that $V = b^2 - f(b-1) = b + f$, which is not possible. It means that if $U + V\omega \in \mathcal{P}$ then $0 \leq V \leq b$. Let furthermore $E \geq 1$. Suppose that $V_1 = 0$. Clearly, it is enough to consider the expansion of $-1$. Since $-1 = -1 + N - \alpha\overline{\alpha}$, $-\overline{\alpha} = \alpha - 2a - b$ therefore if $2a + b \leq 0$ then $-1 \in \mathcal{B}(0)$, if $2a + b > 0$ then $-\overline{\alpha} = \alpha - 2a - b + N - \alpha\overline{\alpha}$. Obviously $2a + b \leq N - 1$ and $1 - \overline{\alpha} = \alpha - 2a - b + 1$ therefore we can conclude that if $2a + b \geq 2$ then $-1 \in \mathcal{B}(1 - \overline{\alpha})$ else $-1 \in \mathcal{B}(0)$. Suppose that $V_1 = b$. It follows from (4.8) that $V = b(U_1 + b - f) \leq b$. Clearly, it is enough to consider the case $U_1 = f - b + 1$. The previous deduction shows that $U_1 + V_1\omega \in \mathcal{P}$ iff $2a + b \geq 2$. We can also notice that there is not other periodic element with $V_1 = b$.

Sub-case $f < b$. Suppose that $U_1 = 1$. Then by (4.8) we have that $V = V_1(b - f) + b \leq b$, therefore $V_1 = 0$ which is a known case. Suppose that $U_1 = f - b - c$ $(c = 0, 1)$. Now $0 \leq V = V_1(b - f) + b(f - b - c) \leq b$, therefore $V_1 = b, c = 0$ which is known as well. It means that if $f < b$ and $U + V\omega \in \mathcal{P}$ then $f - b + 1 \leq U \leq 0$.

Sub-case $b < f$. Suppose that $U_1 = -1$. Then using (4.8) we have that $0 \leq V = V_1(b - f) - b \leq b$ therefore $V_1 < 0$ which is a contradiction. Suppose that $U_1 = f - b + 1$. Then $0 \leq V = V_1(b - f) + b(f - b + 1) \leq b$, hence $V_1 = b$ which is a known case.

If $|\alpha| < 3$ then by Lemma 11 and by [52, 53] the following cases remain. If $a = 2, b = 1, E = 1$ then $-4 + 2\omega, -2 + \omega, 0 \in \mathcal{P}$ of period length one, if $a = 2, b = 1, E = 2$ then $-2 + \omega, 0 \in \mathcal{P}$ of period length one, if $a = 1, b = 1, E = 1$ then $-2 + 2\omega, -1 + \omega, 0 \in \mathcal{P}$ of period length one, if $a = 1, b = 1, E = 2, \ldots, 6$ then $-1 + \omega, 0 \in \mathcal{P}$ of period length one, if $a = 1, b = 2, E = 1$ then $-3 + 3\omega, -2 + 2\omega, -1 + \omega, 0 \in \mathcal{P}$ of period length one, if $a = -1, b = 2, E = 1, 2$ then $\omega, 0 \in \mathcal{P}$ of period length one, if $a = -1, b = 3, E = 1$ then $\mathcal{G}(\mathcal{P}) = \{\omega \to -1 + 2\omega \to \omega, 0 \to 0\}$, if $a = -2, b = 3, E = 1$ then $2\omega, \omega, 0 \in \mathcal{P}$ of period length one, if $a = -3, b = 2, E = 1$ then $1 + \omega, 0 \in \mathcal{P}$ of period length one. The proof is completed. □

**Lemma 14.2.** *If* $E = 1, a = 1$ *or if* $E = 1, a = b + 1$ *then* $\#\mathcal{P} \leq b + 2$, *if* $E \geq 2$ *or* $E = 1, a > 1, a \neq b + 1$ *or* $1 \leq f < b$ *and* $2a + b \geq 2$ *then* $\#\mathcal{P} \leq b + 1$, *else* $\#\mathcal{P} \leq b$.
PROOF: Since there does not exist any $\rho \in T_1$ (resp. $T_2$), $(\rho \neq 0)$ for which $\rho \equiv 0\ (\alpha)$ therefore by (4.1) and by Theorem 3.2 we have that if $U + V\omega \in \mathcal{P}$ then there is not any $Z, (Z \neq U)$ for which $Z + V\omega \in \mathcal{P}$. □

## 4.3 Structure of periodic elements

Let $b \geq 2$, $\beta = \delta$ or $\omega$ and $\mathcal{L}_\mu = \{P + Q\beta \in I,\ (b, Q) = \mu\}$. Obviously, $I = \bigcup_{\mu | b} \mathcal{L}_\mu$. Now, we shall examine the case $\mu < b$. In virtue of (4.5) and (4.8) it is easy to see that if $(V, b) = \mu$ then $(V_1, b) = \mu$. Hence the function $\Phi$ maps $\mathcal{L}_\mu$ to $\mathcal{L}_\mu$ for each $\mu \mid b$. Let $b_\mu = b/\mu$.

**Theorem 4.** *There is a finite decomposition of* $\mathcal{L}_\mu$ *into* $\mathcal{L}_\mu = \cup_{j=0}^{l_\mu - 1} \mathcal{L}_\mu^{(j)}$ *for which if* $\pi \in \mathcal{L}_\mu^{(j)}$ *then* $\Phi(\pi) \in \mathcal{L}_\mu^{(j)}$ *for every* $\pi \in \mathcal{P}$. *The length of period of* $\pi \in \mathcal{P}$ *is* $\varphi(b_\mu)/l_\mu$, *where* $\varphi$ *denotes the Euler totient function.*
PROOF: Let $X = V/\mu$, $X_1 = V_1/\mu$. Then from (4.5) we have that $X = b_\mu U_1 + aX_1$ and from (4.8) we get that $X = b_\mu(U_1 + V_1) + aX_1$. Clearly, in both cases $X \equiv aX_1 \pmod{b_\mu}$, $(X, b_\mu) = (X_1, b_\mu) = 1$. Let us denote by $\mathbb{Z}_{b_\mu}^*$ the set of reduced residue classes modulo $b_\mu$, i.e., $\mathbb{Z}_{b_\mu}^* = \{m \pmod{b_\mu},\ (m, b_\mu) = 1\}$. Let $T_\mu$ denotes the cyclic subgroup $< a >$ in $\mathbb{Z}_{b_\mu}^*$ and let $t_\mu = \mathrm{ord}(a)$. By Lagrange theorem, $\varphi(b_\mu) = l_\mu t_\mu$, hence the order of the factor group $\mathbb{Z}_{b_\mu}^*/T_\mu$ is $l_\mu$. So we have a decomposition $\mathbb{Z}_{b_\mu}^* = H_0 \cup H_1 \cup \ldots \cup H_{l_\mu - 1}$, where $H_0 = T_\mu$. Let $\mathcal{L}_\mu^{(j)} = \{\gamma = P + Q\beta,\ \gamma \in \mathcal{L}_\mu,\ Q/\mu \pmod{b_\mu} \in H_j\}$. Finally, we have the decomposition by $\mathcal{L}_\mu = \mathcal{L}_\mu^{(0)} \cup \mathcal{L}_\mu^{(1)} \cup \ldots \cup \mathcal{L}_\mu^{(l_\mu - 1)}$. The proof is completed. □

*Remark.* Consider the graph $\mathcal{G}(\mathcal{P})$. Theorem 4 states that for a fixed $a$ and $b$ $(b \geq 2)$ there are $\tau(b)$ different sets $\mathcal{L}_\mu$, in each there exist $l_\mu = \varphi(b_\mu)/\mathrm{ord}_{b_\mu} a$ cycles with period length $t_\mu = \mathrm{ord}_{b_\mu}$. If $b_\mu$ is prime then there is only one cycle in $\mathcal{L}_\mu$ with period length $b_\mu - 1$. If $a \equiv 1 \pmod{b_\mu}$ then there are only loops in $\mathcal{L}_\mu$ and the number of them is $\varphi(b_\mu)$.

## 4.4 Number of periodic elements

We have seen in the previous section that for each $\mu \mid b$ and for each $j = 0, 1, \ldots, l_\mu - 1$ there exist at least one period-cycle in $\mathcal{L}_\mu^{(j)}$. The length of a period in $\mathcal{L}_\mu^{(j)}$ is a multiple of $t_\mu$, so it is at least $t_\mu$. This means that $\#\mathcal{P} \geq \sum_{\mu|b} t_\mu l_\mu$. Since $t_\mu l_\mu = \varphi(b_\mu)$ therefore $\#\mathcal{P} \geq \sum_{\mu|b} \varphi(b/\mu) = b$. Keeping in mind the theorems and lemmas proved in this chapter we have the following result.

**Theorem 5.** *Let $b \geq 1$. Let $\alpha = a + b\delta$. If $a \geq 1$ and ($F \geq 2$ or $F = 1, a \neq b + 1$) then $\#\mathcal{P} = b + 1$, if $a \geq 1, F = 1, a = b + 1$ then $\#\mathcal{P} = b + 2$ and if $-a \geq 1$ then $\#\mathcal{P} = b$. Let $\alpha = a + b\omega$. If $E = 1, a = 1$ or if $E = 1, a = b + 1$ then $\#\mathcal{P} = b + 2$, if $E \geq 2$ or $E = 1, a > 1, a \neq b + 1$ or $1 \leq f < b$ and $2a + b \geq 2$ then $\#\mathcal{P} = b + 1$, else $\#\mathcal{P} = b$. If $b \leq -1$ then apply* (4.2).

## 4.5 Expansions in the Gaussian ring

In the following we analyse expansions in the ring of Gaussian integers. Let $\theta = a + bi$. Recall that the elements of the ring $\mathbb{Z}[\theta]$ have the form $\{m + n\theta \ : \ m, n \in \mathbb{Z}\}$. Then $\mathbb{Z}[\theta] = \{u + vbi \ : \ u, v \in \mathbb{Z}\} = \{\beta \in \mathbb{Z}[i] \ : \ b \mid \mathrm{Im}(\beta)\}$.

**Theorem 6.** *Let $\theta = a + bi,\ a = -f,\ f > 0,\ (f, b) = 1$. Then $\mathbb{Z}[\theta] = \mathcal{B}(0)$.*

PROOF: Let $N = \theta\overline{\theta}$. First let $\beta$ be an arbitrary Gaussian integer such that $\beta \in \mathcal{B}(0)$. It means that $\beta = b_0 + b_1\theta + \ldots + b_l\theta^l$ for some $l \in \mathbb{N}$ and $b_j \in D$. Clearly, $b \mid \mathrm{Im}(\theta^j)$. Since $D \subset \mathbb{N}_0$ therefore $b \mid \mathrm{Im}(\beta)$. By the previous remark we have that $\beta \in \mathbb{Z}[\theta]$, hence $\mathcal{B}(0) \subseteq \mathbb{Z}[\theta]$. Suppose now that $\beta \in \mathbb{Z}[\theta]$. Then

$$\beta = m + n\theta. \tag{4.10}$$

On the other hand the expansion of $-1$ is

$$-1 = N - 1 + \theta\overline{\theta} = N - 1 + 2f\theta + \theta^2. \tag{4.11}$$

Since $1, N-1, 2f \in D$ therefore $-1$ has the finite expansion (4.11). Equations (4.10) and (4.11) mean that $\beta$ has also an expansion of the form

$$\beta = u_0 + u_1\theta + u_2\theta^2 + u_3\theta^3, \tag{4.12}$$

where $u_j \geq 0 \quad (j = 0, 1, 2, 3)$. The idea of the following lemma is originated to I. Kátai and J. Szabó [55].

**Lemma 15.** *(Clearing Lemma) Suppose that* $\beta$ *has the following expansion:*

$$\beta = u_0 + u_1\theta + \ldots + u_m\theta^m, \tag{4.13}$$

*where* $u_j \geq 0 \quad (j = 0, \ldots, m)$. *Let* $T = \sum_{j=0}^{m} u_j$. *Then for every* $s \geq 0$ *there exists an expansion* $\beta = v_0 + v_1\theta + \ldots + v_s\theta^s + \ldots + v_l\theta^l$ *such that* $\sum_{j=0}^{l} v_j = T$ *and* $0 \leq v_j < N \quad (j = 0, 1, \ldots, s)$.

PROOF: First we shall examine the expansion of $N = a^2 + b^2$.

$$\begin{aligned} N &= (-2f - \theta)\theta = \big((N - 2f) - \theta\overline{\theta} - \theta\big)\theta = \\ &= (N - 2f)\theta + (-\overline{\theta} - 1)\theta^2 = (N - 2f)\theta + (2f - 1)\theta^2 + \theta^3. \end{aligned} \tag{4.14}$$

Observe that $N - 2f, 2f - 1, 1 \in D$ and the sum of the digits of the expansion in (4.14) is $N$. Clearly, if $u_0 < N$ in (4.13) then the lemma holds for $s = 0$. In the opposite case, if $N \leq u_0$, then let $u_0 = pN + q,\ p \geq 1,\ 0 \leq q < N$. Let us take $u_0 = q + p\big(0 + (N - 2f)\theta + (2f - 1)\theta^2 + \theta^3\big)$ into the equation (4.13). Then we have that $\beta = u_0' + u_1'\theta + \ldots + u_m'\theta^{m'}$, where $u_0' = q,\ u_1' = u_1 + p(N - 2f),\ u_2' = u_2 + p(2f - 1),\ u_3' = u_3 + p,\ u_j' = u_j\ (j \geq 4)$. Observe that $\sum u_j' = T$, so the sum of the digits does not change in the new expansion. Hence, the case $s = 0$ is satisfied. We can continue the process for $j = 1, 2, \ldots, s$. □

Now, we can apply the Clearing Lemma for the expansion of $\beta$ in (4.12). Let $T_0 = T = u_0 + u_1 + u_2 + u_3$. If $u_0 \geq N$ then by the lemma we have that $\beta = v_0 + \theta\beta_1,\ 0 \leq v_0 < N,\ \beta_1 = y_1 + y_2\theta + \ldots + y_m\theta^m,\ y_j \geq 0$. Let $T_1 = T_0 - v_0$. Clearly, $T_1 = y_1 + \ldots + y_m$. Applying the Clearing Lemma again and again we have a monotonically decreasing sequence $T_0, T_1, T_2, \ldots$, and expansions $\beta, \beta_1, \beta_2, \ldots$. If there exists an $h \in \mathbb{N}$ such that $T_h = 0$ then the expansion of $\beta$ is finite having the digits from the set $D$, so Theorem 6 is proved. If such an $h$ does not exist then there is a suitable large $h_0$ such that $T_{h_0} = T_{h_0+1} = \ldots = r > 0$. But in this case $\beta_{h_0} = \theta\beta_{h_0+1} = \theta^2\beta_{h_0+2} = \ldots = \theta^j\beta_{h_0+j} = \ldots$, therefore $\gamma := \beta - (v_0 + v_1\theta + \ldots + v_{h_0}\theta^{h_0}) = \theta^{h_0+t}\beta_{h_0+t}$ for every $t \in \mathbb{N}$. Observe that $\theta^{h_0+t} \mid \gamma \quad (t = 1, 2, \ldots)$ and this holds only if $\gamma = 0$. This means that $\beta$ has a finite expansion with digits from the set $D$. The proof of Theorem 6 is complete. □

# Chapter 5

# Geometry of expansions

*"The mathematical sciences particularly exhibit order, symmetry and limitation; and these are the greatest forms of the beautiful."*
*— Aristoteles*

In this chapter we investigate the set of fractions of radix systems and lattice tilings with these sets.

## 5.1 Set of fractions $H$

In section 2 for a given radix system $(\Lambda, M, D)$ the set of fractions was defined as $H = \mathcal{F}(M, D) = \{ \sum_{n=1}^{\infty} M^{-n} a_n \ : \ a_n \in D\} \subset \mathbb{R}^k$. Recall that $\mathcal{P}$ denotes the set of periodic elements. Let furthermore $\Gamma_l$ be the set of lattice points of form $a_0 + Ma_1 + \cdots + M^l a_l$, $(a_i \in D)$. Then $D = \Gamma_0 \subseteq \Gamma_1 \subseteq \cdots$. Let $\Gamma = \bigcup \Gamma_l$. Thus, $\Gamma$ is the set of those lattice points $z$ which have finite expansions in the radix system $(\Lambda, M, D)$. Let $\lambda$ be the Lebesgue measure on $\mathbb{R}^k$.

**Assertion 9.** *The fundamental domain $H$ has the following properties: (i) $H$ is compact. (ii) $H$ has interior points. More specifically $\bigcup_{p\in\mathcal{P}}(p+H)$ contains a neighborhood of the origin. (iii) $\mathbb{R}^k = \bigcup(H + \Lambda)$ (iv) For every $x \in \mathbb{R}^k$ there is a $z \in \Lambda$ and $h \in H$ such that $x = z + h$. (v) $\lambda(H) > 0$. (vi) $\lambda(H + z_1 \cap H + z_2) = 0$ for all $z_1 \neq z_2 \in \Gamma$.*
PROOF: Concerning (i) most of the proofs applies Cantor's diagonal principle.

Now a different method will be given. Let $F \subseteq H$ be infinite. For each digit $a \in D$ let

$$F(a) := \left\{ x \in F : x = \sum_{j=1}^{\infty} M^{-j} d_j, d_j \in D, d_1 = a \right\}.$$

Then we have $F = \bigcup_{a \in D} F(a)$. Since $F$ is infinite, at least one of the sets $F(a)$ is also infinite. Choose $a_1 \in D$ so that $F(a_1)$ is infinite. Then let

$$F(a_1, a) := \left\{ x \in F : x = \sum_{j=1}^{\infty} M^{-j} d_j, d_j \in D, d_1 = a_1, d_2 = a \right\}.$$

There is an $a_2 \in D$ so that $F(a_1, a_2)$ is infinite. We may continue this process to obtain a sequence $(a_j) \in D$ so that $F(a_1, a_2, \ldots, a_n)$ is infinite for all $n$. Then in the metric space $\mathbb{R}^k$ the vector $\sum_{j=1}^{\infty} M^{-j} a_j$ is an accumulation point of the set $F$. This shows that $H$ is compact. Concerning (ii) the proof can be found in [107, Theorem 1] or in [83, Theorem 1.1]. The equivalent assertions (iii) and (iv) are easy consequences of (ii). Since $H$ is compact, it is measurable, and $\lambda(H) = 0$ would imply that $\lambda(\mathbb{R}^k) = 0$. Therefore $\lambda(H) > 0$, which was stated in (v). Concerning (vi) suppose that $z_1, z_2 \in \Gamma$ and $z_1 \neq z_2$. Then

$$\lambda(H) \mid \det M \mid^l = \lambda(M^l H) \leq \bigcup_{z \in \Gamma_{l-1}} \lambda(H + z) = \mid \det M \mid^l \lambda(H).$$

If there is a couple $z_1, z_2 \in \Gamma_{l-1}, z_1 \neq z_2$ for which $\lambda(H + z_1 \cap H + z_2) > 0$, then the "less than or equal to" can be changed to "less than", which is impossible. □

Since $H$ is compact, it is possible to draw it. In section 2.3 we noted that $H$ is self-affine with respect to the linear contraction maps $f_a : \mathbb{R}^k \to \mathbb{R}^k, f_a(z) = M^{-1}(z + a), a \in D$ and $H$ is the unique attractor of the iterated function system $\{f_a : a \in D\}$. A computer can be used to generate rapidly the attractor of an iterated function system by repeatedly applying the maps of the system with equal probabilities and plotting the resulting points. The same can be achieved by plotting the points of the set $H_l = \{\sum_{i=1}^{l} M^{-i} a_i : a_i \in D\}$. However, this is not the best method because of the size limitations of the graphics device. Hereinafter we follow the method of B. Mandelbrot [85].

ESCAPE ALGORITHM FOR PLOTTING THE SET $H$.
Consider the radix system $(\Lambda, M, D)$. For each digit $a \in D$ define the function $g_a : \mathbb{R}^k \to \mathbb{R}^k$ by $g_a(z) = Mz - a$. Let $K(H)$ be a bounded subset of $\mathbb{R}^k$ that contains $H$ and easy to decide whether an arbitrary $y \in \mathbb{R}^k$ is in $K(H)$. Such a set — a $k$-dimensional rectangle — was constructed in chapter 2. Given any number $z \in \mathbb{R}^k$ construct the sequence of sets $S_0, S_1, S_2, \ldots$ as follows. Let the initial set $S_0$ be $\{z\}$, if $z \in K(H)$ or empty otherwise. Let

$$S_j = \{g_{a_i}(z) : z \in S_{j-1}, a_i \in D, g_{a_i}(z) \in K(H)\}.$$

Stop the algorithm, if the set $S_j$ becomes empty or the number of sets $j$ reaches some predetermined limit $l$. If any of the sets $S_j$ are empty, then $z$ does not lie in the set $H$. If $l$ is large and $S_l$ is non-empty, then $z$ either lies in $H$, or it is very close to it. A point $x$ in the set $S_l$ is of the form $x = g_{a_l} \circ g_{a_{l-1}} \circ \ldots g_{a_1}(z)$, where each $a_i \in D$ and $z$ is approximately $\sum_{i=1}^{l} M^{-i} a_i \in H$. If $S_j$ is the first empty set, then $j$ is a measure of time taken by $z$ to escape from $K(H)$ under iterations of maps $g_a$.

Applying the ESCAPE ALGORITHM for all the points of a compact region of $\mathbb{R}^k$ (according to the graphics device) one can color these points via their escape time $j$. Plenty of pictures of fundamental sets was generated by the author in the Gaussian ring. These pictures had much more success in the exhibition CeBIT'93 than the mathematics behind them [62]. Some of those pictures can be seen in the home page of my project leader[1]. A few fundamental sets can be found in appendix B.

## 5.2 Just touching coverings and the boundary of $H$

In the previous section we analyzed the fundamental domain $H$ and the translates of $H$ to the points of $\Gamma$. It is easily seen that the elements of $\Gamma$ are not necessary closed for the addition. Clearly, if $(\Lambda, M, D)$ is a number system then $\Gamma = \Lambda$, consequently $\lambda(H + z_1 \cap H + z_2) = 0$ holds for each $z_1, z_2 \in \Lambda, z_1 \neq z_2$. This suggest the following definition: the radix system $(\Lambda, M, D)$ is called a *just touching covering (JTC)* system, if $\lambda(H + z_1 \cap H + z_2) = 0$ holds for each $z_1, z_2 \in \Lambda, z_1 \neq z_2$.

[1] http://math.uni-paderborn.de/~k-heinz

Let $S$ denote the set of those elements $0 \neq z$ of $\Lambda$ for which $H \cap (H+z) \neq \emptyset$. I. Kátai and co-workers proved [40] that the covering $\bigcup(H+\Lambda) = \mathbb{R}^k$ is JTC if and only if $\Gamma - \Gamma = \Lambda$, or equivalently, a covering is JTC iff for each element $z \in S$ may be written as $z = \sum_{j=0}^{m} M^j b_j$ with $b_j \in \mathcal{B} := D - D$.

In order to examine the points of $\partial H$ (boundary of $H$), let us introduce the set $B(z) = (z + H) \cap H, z \in \Lambda$. Clearly, $S$ is a set of those $z \in \Lambda \setminus \{0\}$ for which $B(z)$ is nonempty. It is obvious that $B(z)$ is nonempty iff $z$ has an expansion of the form $z = \sum_{i=1}^{\infty} M^{-i} b_i$, where $b_i \in \mathcal{B}$. Hence $S \subset H - H$, therefore $\|z\| \leq 2L$ for all $z \in S$. In [40, 41] it was suggested to use the *transition graph* $\mathcal{G}(\mathcal{S})$.

Algorithm for constructing the Transition Graph $\mathcal{G}(\mathcal{S})$.
Let $K(H)$ be the $k$-dimensional rectangle centered at the origin determined in chapter 2 and let $U = 2K(H)$. Clearly, if $z \in S$ then $z \in U$. For all $z \in \Lambda \cap U, z \neq 0$ calculate $z_b = Mz - b$, where $b \in \mathcal{B}$. Let $m(b)$ be the number of possibilities to write $b \in \mathcal{B}$ in the form $b = a_i - a_j,\ (a_i, a_j \in D)$. If $z_b \in U$ then direct $m(b)$ edges with labels $a_i$ from $z$ to $z_b$. Delete $z$ if no edge leaves it and delete all edges that end in $z$. Continue this process until no appropriate $z$ remains. The resulting graph is $\mathcal{G}(\mathcal{S})$. The process terminates because the number of nodes is finite.

Observe that the graph $\mathcal{G}(\mathcal{S})$ has symmetry properties: if the graph contains an edge from $x$ to $y$ with label $a$, then there is an edge from $-z$ to $-y$ with $-a$. It is also not hard to see, that every node in the graph $\mathcal{G}(\mathcal{S})$ has incoming edge(s). The transition graph is a tool for computing $\partial H$ without computing the interior points.

Algorithm for computing the boundary of $H$.
Let start from an arbitrary node $y \in \mathcal{G}(\mathcal{S})$, and walk on the transition graph writing down the randomly chosen sequence of labels $a_1, a_2, \ldots$. Then, $z \in B(y)$ iff $z = \sum_{j=1}^{\infty} M^{-j} a_j$. From computational point of view it is enough to generate some finite steps (depending on the graphics device) of the walk. Repeat the process.

The transition graph $\mathcal{G}(\mathcal{S})$ can also be used to decide whether the radix system $(\Lambda, M, D)$ is a JTC system. The property holds iff for each node $z$ there is a path $\alpha$ in $\mathcal{G}(\mathcal{S})$ for which $i(\alpha) \in \mathcal{B},\ t(\alpha) = z$. Here $i(\alpha)$ and $t(\alpha)$ denote the initial and terminal node of the path $\alpha$.

## 5.3 Hausdorff dimension of $\partial H$

First, we recall the different notions of dimensions which are used in this chapter. The *Hausdorff dimension* of a Borel set $E$ is defined as follows: let $\{U_i\}_{i=1}^{\infty}$ be an $\varepsilon$-cover of $E$, i.e. $E \subseteq \bigcup_{i=1}^{\infty} U_i$ and $\operatorname{diam}(U_i) < \varepsilon$, where $\operatorname{diam}(U_i)$ denotes the diameter of $U_i$. Then the $s$-dimensional Hausdorff measure of $E$ is given by

$$\mathcal{H}^s(E) = \lim_{\varepsilon \to 0} \big( \inf \big\{ \sum_{i=1}^{\infty} \operatorname{diam}(U_i)^s : \{U_i\}_{i=1}^{\infty} \text{ is an } \varepsilon\text{-cover of } E \big\} \big).$$

The Hausdorff dimension of $E$ is now defined by

$$\dim_H(E) = \inf\{s : \mathcal{H}^s(E) = 0\} = \sup\{s : \mathcal{H}^s(E) = \infty\}.$$

There are several difficulties in evaluating the Hausdorff dimension in a concrete case. The *box-counting dimension* simplifies this problem by replacing the terms $\operatorname{diam}(U_i)^s$ by the terms $\delta^s$ in $\mathbb{R}^k$. A formal definition of the box dimension $\dim_B$ of any bounded subset $E$ of $\mathbb{R}^k$ proceed as follows. Let $N_\delta(E)$ be the smallest number of sets of diameter at most $\delta$ which cover $E$. Since $E$ is bounded we can always assume that the cover is finite. Then

$$\dim_B(E) = \lim_{\delta \to 0} \frac{\log N_\delta(E)}{\log 1/\delta},$$

provided that the limit exists. If it exists and is not an integer, then $E$ is said to have *fractal dimension*. It may take non-integral values, but yields the usual dimension for the most ordinary spaces. A *fractal set* is one whose Hausdorff dimension is strictly greater than its topological dimension. The term fractal was introduced by the mathematician Benoit Mandelbrot. Examples of fractal sets are the Cantor set and the boundary of Koch's snowflake. Unfortunately, it is not true that the Hausdorff dimension and the box dimension are always the same. But it is true that $\dim_H(E) \leq \dim_B(E)$. For further discussion of these and other kinds of dimensions we refer to [17, 18, 19].

Second, a brief survey will be given for the concept of graph self-similarity which was introduced by R. D. Mauldin and S. C. Williams [87, 17], and by Falconer [18]. A *directed multi-graph* consists of two (finite) sets $V$ and $E$, and two functions $i : E \to V$ and $t : E \to V$. The elements of $V$ are called *vertices*

or *nodes*; the elements of $E$ are called *edges* or *arrows*. For an edge $e$, we call $i(e)$ the *initial vertex* of $e$, and we call $t(e)$ the *terminal vertex* of $e$. We will often write $E_{uv}$ for the set of all edges $e$ with $i(e) = u$ and $t(e) = v$. A directed multi-graph is *strongly connected* iff, for each pair $u, v$ of vertices, there is a path from $u$ to $v$. A *path* in a directed multi-graph is a sequence of edges, taken in some order. A path will often be identified with a string made up of the labels of the edges. Let $\rho$ be a metric on $\mathbb{R}^k$. A mapping $f : \mathbb{R}^k \to \mathbb{R}^k$ is called a *contraction* if $\rho(f(x), f(y)) \leq c\rho(x, y)$ $(x, y \in \mathbb{R}^k)$ holds for some constant $c < 1$. We call the infimum of these constants $c$, for which the inequality holds, the *ratio* of the contraction $f$. A contraction, which maps any subset of $\mathbb{R}^k$ to a geometrically similar set is called a *contracting similarity*.

A directed multi-graph $(V, E, i, t)$ together with a function $r : E \to (0, \infty)$, will be called a *Mauldin-Williams graph*. Suppose that $(V, E, i, t, r)$ is a Mauldin-Williams graph. An iterated function system realizing the graph is made up of metric spaces $S_v$, one for each vertex $v$, and similarities $f_e$, one for each edge $e \in E$, such that $f_e : S_v \to S_u$ if $e \in E_{uv}$, and $f_e$ has ratio $r(e)$. An *invariant list* for such an iterated function system is a list of nonempty compact sets $K_v \subseteq S_v$, one for each node $v \in V$, such that

$$K_u = \bigcup_{v \in V, e \in E_{uv}} f_e[K_v]$$

for all $u \in V$. Each of the nonempty compact sets $K_v$ satisfying such equations will be said to have *graph self-similarity*. A Mauldin-Williams graph $(V, E, i, t, r)$ will be called *strictly contracting* if the conditions $r(e) < 1$ are satisfied, in which case there is a unique list $(K_v)_{v \in V}$ of nonempty compact sets $(K_v \subseteq S_v)$ satisfying the previous equation.

A non-negative square matrix $M$ is called *primitive* if $M^j > 0$ for some positive integer $j$. A square matrix is called *reducible* if there exist a permutation that puts into the form $M_\theta = \left(\begin{smallmatrix} M_{11} & M_{12} \\ 0 & M_{22} \end{smallmatrix}\right)$, where $M_{11}$ and $M_{22}$ are square matrices. Otherwise $M$ is called *irreducible*. An irreducible non-negative matrix $M$ always has a positive eigenvalue $\lambda$. The moduli of all the other eigenvalues do not exceed $\lambda$. Moreover, there is an eigenvector associated to $\lambda$ with all positive entries. Let a Mauldin-Williams graph be given. For all $t \geq 0, u, v \in V$ define

$$A_{uv}(t) = \sum_{e \in E_{uv}} r(e)^t$$

and the matrix $A(t)$ by $A(t)[u,v] = A_{uv}(t)$. Then, by the Perron-Frobenius theorem, the spectral radius of $A(t)$ takes the value 1 for a uniquely determined value of $t = t_0$. This $t_0$ is called the *graph dimension* of the Mauldin-Williams graph. Consider a strongly connected Mauldin-Williams graph $(V, E, i, t, r)$. When the invariant set list is found, each of the sets will be similar to a subset of each of the others. So they will all have the same Hausdorff dimension. In order to determine the graph dimension, first we need to find the proper sort of *Perron numbers*. If $s$ is a positive real number, then the $s$-dimensional Perron numbers for the graph are positive numbers $q_v$, one for each vertex $v \in V$, such that

$$q_u^s = \sum_{v \in V, e \in E_{uv}} r(e)^s \cdot q_v^s.$$

There is exactly one positive number $s$ such that $s$-dimensional Perron numbers exist. This unique number is equal to the graph dimension of the Mauldin-Williams graph.

If $(f_e)$ is a realization of $(V, E, i, t, r)$ in $\mathbb{R}^k$, then we say it satisfies the *graph open set condition* iff there exist nonempty open sets $U_v$, one for each $v \in V$, with $f_e[U_v] \subseteq U_u$ for all $u, v \in V$ and $e \in E_{uv}$; and $f_e[U_v] \cap f_{e'}[U_{v'}] = \emptyset$ for all $u, v, v' \in V, e \in E_{uv}, e' \in E_{uv'}$ with $e \neq e'$.

The graph dimension can be used to calculate an upper bound of the Hausdorff dimension of the sets of the invariant list. Let $(V, E, i, t, r)$ be a strongly connected contracting Mauldin-Williams graph describing the graph self-similarity of a list $(K_v)_{v \in V}$ of nonempty compact sets in $\mathbb{R}^k$. Let $s > 0$ be such that $s$-dimensional Perron numbers exist. Then $\dim K_v \leq s$ for all $v$. If, in addition, the realization satisfies the open set condition, then $\dim K_v = s$.

How can we compute the graph dimension if the graph is not strictly connected? Let $SC(V)$ be the set of all strictly connected components of $V$. Let $s$ be the graph dimension of $V$ and $s_W$ be the graph dimension of $W \in SC(V)$. Then $s = \max_{W \in SC(V)} s_W$. Let furthermore $K = \bigcup_{v \in V} K_v$. Then $\dim K \leq s$, and, if the open set condition is satisfied then equality holds [87].

Suppose that the all the eigenvalues of $M$ are distinct and greater than one in module. Let us examine the transition graph $\mathcal{G}(\mathcal{S})$ from similarity aspects. First, let us define the sets $B(z)$ for each node $z$. Clearly, the sets $B(z)$ are compact for all $z$. Suppose that the graph contains some edges from $x$ to $y$ with labels $d_i$. Let us define the maps $f_{d_i} : B(y) \to B(x)$, $f_{d_i}(z) = M^{-1}(z + d_i)$ for each label, where $d_i \in D$. We will prove that $f_{d_i}(B(y)) \subset$

$B(x)$. Indeed, if $z \in B(x)$, then $z$ can be written as $z = \sum_{j=1}^{\infty} M^{-j}a_j = x + \sum_{j=1}^{\infty} M^{-j}b_j$ $(a_j, b_j \in D)$. Thus, $x = \sum_{j=1}^{\infty} M^{-j}(a_j - b_j)$. Therefore $y = Mx - (a_1 - b_1) = \sum_{j=1}^{\infty} M^{-j}(a_{j+1} - b_{j+1})$ where $\delta = a_1 - b_1 \in \mathcal{B}$. It means that if $z_1 = \sum_{j=1}^{\infty} M^{-j}a_{j+1} = y + \sum_{j=1}^{\infty} M^{-j}b_{j+1} \in B(y)$ then $f_{a_1}(z_1) = z \in B(x)$. So we have that

$$B(x) = \bigcup_{y \in V, e \in E_{xy}} f_e[B(y)],$$

in other words the sets $B(z)$ form an invariant list of the iterated function system $\{f_e\}$. Since the mappings $f_{d_i}$ are contracting similarities, the graph $\mathcal{G}(S)$ is a strictly contracting Mauldin-Williams graph so its graph dimension can be determined by the previously described way.

Unfortunately, in most cases the open set condition does not hold, so the Hausdorff dimension is hard to determine. But under certain circumstances the Hausdorff dimension of $\partial H$ is equal to its box counting dimension and the open set condition satisfies. Recall that a finite directed graph is *primitive*, if it is strongly connected and the greatest common divisor of the length of its closed directed walks is one [9]. In this case the accompanying matrix of $\mathcal{G}(\mathcal{S})$ has a unique (positive real) eigenvalue of largest modulus. The following theorem was proved in [90]. Let $(\Lambda, M, D)$ a JTC radix system and assume that all eigenvalues of $M$ have the same modulus $\lambda$. Assume further that the associated transition graph is primitive and denote by $\mu_{\max}$ the unique eigenvalue of largest modulus of its accompanying matrix. Then the Hausdorff dimension of $\partial H$ is equal its box dimension and is given by $s = \frac{\log \mu_{\max}}{\log \lambda}$. Moreover, if the transition graph $\mathcal{G}(\mathcal{S})$ is not primitive but there is a primitive subgraph of $\mathcal{G}(\mathcal{S})$ which has the same maximal eigenvalue as $\mathcal{G}(\mathcal{S})$ then their graph dimension are equal and the Hausdorff dimension of $\partial H$ can be computed in the above described way.

If the moduli of the eigenvalues of $M$ are not all the same then using the graph $\mathcal{G}(\mathcal{S})$ the box dimension of $\partial H$ can be computed. In real quadratic fields using canonical digit sets it was calculated by J. M. Thuswaldner [104].

## 5.4 Just touching coverings in special cases

Let $D = \{a_1, a_2, \ldots, a_N\} \subset \mathbb{Z}$, where $a_i \equiv i \pmod N$ and let $\mathcal{B} = D - D$. The set of integers expressible in the form $\sum_{i=0}^{l} b_i N^i$, for some $l$ with $b_i \in \mathcal{B}$,

is denoted by $Z_{\mathcal{B}}$. Then $Z_{\mathcal{B}} = d\mathbb{Z}$ iff $(a_1 - a_N, a_2 - a_N, \dots, a_N - a_N) = d$. Assume that $a_N = 0$ Then $Z_{\mathcal{B}} = \mathbb{Z}$ iff $(a_1, a_2, \dots, a_N) = 1$. This theorem was conjected by I. Kátai and was proved by G. E. Michalek for $N = 3$ in [88] and for arbitrary $N$ in [89]. Consider now the Gaussian integers $\mathbb{Z}[i]$.

**Proposition 1.** *Let $\theta = a + bi \in \mathbb{Z}[i], N = \mathrm{Norm}(\theta) = a^2 + b^2 \geq 2, (a, b) = 1, D = \{0, 1, \dots, N-1\}$. The system $(\mathbb{Z}[i], \theta, D)$ is JTC system if and only if $b = \pm 1$. In these cases the Hausdorff dimensions of boundaries of the fundamental domains are $\frac{\log(\lambda_{\max})}{\log(1/(a^2+1))}$, where $\lambda_{\max}$ is the largest (positive) real root of the polynomial $(a^2+1)z^3 + (a^2 - 2a + 1)z^2 + (2a - 1)z - 1$.*

PROOF: Let $D \subset \mathbb{Z}$ be an arbitrary complete residue system modulo $\theta$. If $\alpha \in \mathbb{Z}[i]$ can be represented in the form $\alpha = \sum_{i=0}^{l} \theta^i d_i$ $(d_i \in \mathcal{B} = D - D)$ then $b \mid \mathrm{Im}(\alpha)$, since $b \mid \mathrm{Im}(\theta^l)$ $(l = 1, 2, \dots)$. Hence the JTC property implies that $b = \pm 1$, i.e. $\theta$ is of form $\theta = a \pm i$. Observe that if $(\mathbb{Z}[i], \theta, D)$ is JTC radix system then $(\mathbb{Z}[i], \overline{\theta}, D)$ is as well and the Hausdorff dimensions of boundaries of their fundamental domains are the same. Hence it is enough to examine the case $b = 1$. But due to [55] we know that if $a \geq 1$ then $(\mathbb{Z}[i], -a - i, D)$ is a number system, therefore JTC system. This implies that $(\mathbb{Z}[i], a + i, D)$ is also a JTC system with the same Hausdorff dimensions of their $\partial H$. W. Gilbert computed the box dimensions of the boundaries of fundamental domains of the number systems $(\mathbb{Z}[i], -a + i, D)$ $(a \in \mathbb{N})$ by successive approximations [31]. S. Ito computed the Hausdorff dimensions of $\partial H$ for all canonical number systems in imaginary quadratic fields using group endomorphism [44]. Moreover, if $\alpha$ is a non-real quadratic integer and $D$ is a canonical digit set modulo $\alpha$ then I. Környei determined the Hausdorff dimension of $\partial H$ [80] using the linear recursive method of K-H. Indlekofer, I. Kátai and P. Racskó [40]. So the proposition is essentially proved. But the aim of this section is to provide a proof using graph constructions, which is different from the above mentioned methods.

Let $a \geq 5$. In order to compute the transition graph we follow the method of section 2.1. The corresponding matrix belonging to $\theta = a + i$ is $M = \left(\begin{smallmatrix} a & -1 \\ 1 & a \end{smallmatrix}\right)$. Then

$$\|M^{-1}\|_\infty = \left\| \begin{pmatrix} a/(a^2+1) & 1/(a^2+1) \\ -1/(a^2+1) & a/(a^2+1) \end{pmatrix} \right\|_\infty = \frac{a+1}{a^2+1} < 1$$

where $\|\cdot\|_\infty$ is the matrix norm induced by the maximum norm of $\mathbb{R}^2$.

Therefore $(I - M^{-1})^{-1}$ exists and

$$\eta = \frac{1}{1 - \|M^{-1}\|_\infty} = \frac{a^2+1}{a(a-1)} \geq \|(I - M^{-1})^{-1}\|_\infty,$$

where $I$ is the two dimensional identity matrix. Let $\underline{v}$ be the first column vector of $M^{-1}$ and let $\underline{\mu} = \underline{v}d = [\mu_1(d), \mu_2(d)]^T$ $(d \in \mathbb{Z})$. In this case

$$\xi_1 = \max_{d \in D} |\mu_1(d)| = \frac{a^3}{a^2+1} < a \quad \text{and} \quad \xi_2 = \max_{d \in D} |\mu_2(d)| = \frac{a^2}{a^2+1} < 1.$$

This means that

$$\eta\xi_1 = \frac{a^2}{a-1} < a+2 \quad \text{and} \quad \eta\xi_2 = \frac{a^2}{a(a-1)} < 2.$$

Since we are interested in only the integers in $H$ therefore we can conclude that if $\gamma \in H \cap \mathbb{Z}[i]$ then $|\,\mathrm{Re}(\gamma)\,| \leq a+1$ and $|\,\mathrm{Im}(\gamma)\,| \leq 1$. Obviously, if $\gamma \in \mathcal{G}(\mathcal{S})$ then

$$|\,\mathrm{Re}(\gamma)\,| \leq 2(a+1) \quad \text{and} \tag{5.1}$$
$$|\,\mathrm{Im}(\gamma)\,| \leq 2. \tag{5.2}$$

Suppose that there is an edge in $\mathcal{G}(\mathcal{S})$ from $X+Yi$ to $A+Bi$. Then $A+Bi = (a+i)(X+Yi) - \delta$, where

$$\delta \in \mathcal{B} = \{-a^2, \ldots, a^2\}. \tag{5.3}$$

Hence, using (5.1),(5.2) we have the equations

$$A = aX - Y - \delta,\ |A\,| \leq 2(a+1) \tag{5.4}$$
$$B = aY + X,\ |B\,| \leq 2. \tag{5.5}$$

One can immediately observe that $|Y| \geq 3$ contradicts to (5.5), therefore $|Y| \leq 2$. Let $Y = 2$. Using equation (5.5) we have the cases $X = -2a - 2, \ldots, -2a+2$. Now, equations (5.3),(5.4) show that none of them are valid. The same can be stated about $Y = -2$. Hence $|Y| \leq 1$ and we can modify equation (5.5) to

$$B = aY + X,\ |B\,| \leq 1. \tag{5.6}$$

*Case* $Y = 0$. Equation (5.6) shows that in this case $|X| \leq 1$. Let $X = 1$. In virtue of (5.3),(5.4),(5.6) and by the symmetry property of the $\mathcal{G}(\mathcal{S})$ we have some candidates for the nodes of $\mathcal{G}(\mathcal{S})$:

$$1 \to \eta + i,\ -1 \to -\eta - i,\ \eta = -2(a+1), \dots, 2(a+1). \tag{5.7}$$

*Case* $Y = \pm 1$. In virtue of (5.6) we have the cases

$$Y = 1, X = -a-1, -a, -a+1,\quad Y = -1, X = a-1, a, a+1. \tag{5.8}$$

If $X = -a-1$ then by equations (5.4) and (5.6) we get new candidates for the nodes of $\mathcal{G}(\mathcal{S})$:

$$-a - 1 + i \to \eta - i,\ \eta = -2(a+1), \dots, -(a+1).$$

It follows from (5.7) and (5.8) that the only valid case could be $-a-1+i \to -a-1-i$, but it obviously can not happen. Using the symmetry of $\mathcal{G}(\mathcal{S})$ it is easy to see that $a+1-i \to a+1+i$ can not happen as well. If $X = \pm a$ then by using the result of the case $Y = 0$ we have that

$$-a + i \to -1,\ a - i \to 1. \tag{5.9}$$

Finally, if $X = -a+1$ then we have the candidates

$$-a + 1 + i \to \eta + i,\ \eta = -a, -a+1, \tag{5.10}$$

and if $X = a - 1$ then

$$a - 1 - i \to \eta - i,\ \eta = a, a-1. \tag{5.11}$$

Using equations (5.7),(5.9),(5.10) and (5.11) we can construct the graph $\tilde{\mathcal{G}}(\mathcal{S})$:

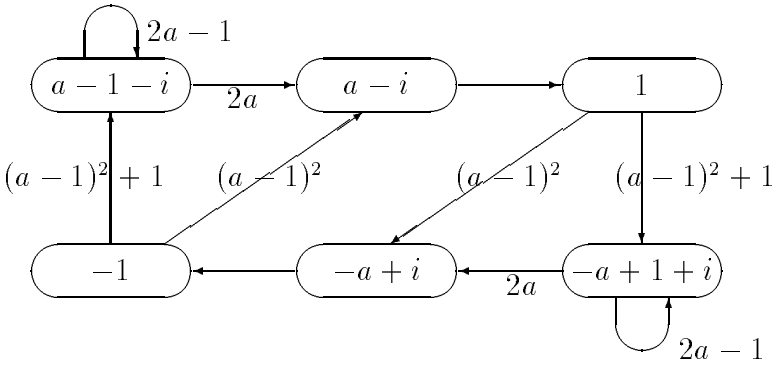


Figure 1

The difference from $\mathcal{G}(\mathcal{S})$ is that the labels of the graph $\tilde{\mathcal{G}}(\mathcal{S})$ show the $m(\delta)$ multiplicities of the edges of $\mathcal{G}(\mathcal{S})$, which can be easily get from equation (5.4). The same graph can be constructed also for the cases $a = 3, 4$. The JTC property clearly holds. In case of $a = 2$ the graph $\mathcal{G}(\mathcal{S})$ is a bit different but the JTC property still true. In this case the only strongly connected component of $\tilde{\mathcal{G}}(\mathcal{S})$ having more than one node is the graph above. The remaining nodes does not influence the graph dimension. If $a = 1$ then the graph $\tilde{\mathcal{G}}(\mathcal{S})$ is again the same by canceling two edges labeled above with $(a-1)^2$.

Since the eigenvalues of $M_\theta = \left(\begin{smallmatrix} a & -1 \\ 1 & -a \end{smallmatrix}\right)$ have the same moduli $(a^2+1)^{1/2}$, using the graph $\tilde{\mathcal{G}}(\mathcal{S})$ we can calculate the Hausdorff dimension of $\partial H$ for all $a \geq 1$. Solving the system of equations $x_1 = (2a-1)\lambda x_1 + 2a\lambda x_2, x_2 = \lambda x_3, x_3 = (a-1)^2\lambda x_5 + ((a-1)^2+1)\lambda x_6, x_4 = ((a-1)^2+1)\lambda x_1 + (a-1)^2\lambda x_2, x_5 = \lambda x_4, x_6 = (2a-1)\lambda x_6 + 2a\lambda x_5$ by substituting $x_1 = 1$ we have that $\lambda$ is the root of the polynomial $(a^2+1)z^3+(a^2-2a+1)z^2+(2a-1)z-1$. Let us denote by $\lambda_{\max}$ the largest (positive) real root of this polynomial. Then the Hausdorff dimension of $\partial H$ is $\frac{\log(\lambda_{\max})}{\log(1/(a^2+1))}$. □

*Remarks. (1)* Recall that a metric space $(X, d)$ is connected if it cannot be expressed as the union of two disjoint nonempty closed subsets. A subset $S \subset X$ is connected if the metric space $(S, d)$ is connected. $S$ is totally disconnected provided that the only nonempty connected subsets of $S$ are subsets consisting of single points. Let $S \subset X$ be a subset of a metric space $(X, d)$. Then $S$ is arcwise connected if, for each pair of points $x$ and $y$ in $S$, there is a continuous function $f : [0,1] \to S$, from the metric space ($[0,1]$, Euclidean) into the metric space $(S, d)$ such that $f(0) = x$ and $f(1) = y$. $S$ is arcwise disconnected if it is not arcwise connencted. There is a brand-new result of P. Talabér (personal communication) who presented a simple method of proving the connectedness of fundamental domains. As a special case, using canonical digit sets, if $1 \in \mathcal{G}(\mathcal{S})$ then $\mathcal{F}(M, D)$ is always connected (see also [38]). Moreover, our construction shows that in case of Gaussian integers using canonical digit sets the condition is also sufficient. Hence, we have that in the Gaussian ring using canonical digit sets *all* fundamental domains of JTC systems are connected. We note that in this case a stronger result — the arcwise connectedness — is known due to S. Akiyama and J. M. Thuswaldner [2].

*(2)* Let $D_1$ be the canonical digit set $\{0, 1, \ldots, N-1\}$ and $D_2$ be the symmetric digit set $\{\lfloor(-N+2)/2\rfloor, \ldots, \lfloor N/2\rfloor\}$. Since $\mathcal{B} = D_1 - D_1$ is equal to $D_2 - D_2$ therefore the JTC property of $(\mathbb{Z}[i], \theta, D_1)$ and $(\mathbb{Z}[i], \theta, D_2)$ holds at exactly the same time. Moreover, the Hausdorff dimensions of the boundaries of their fundamental domains are the same. In contrast to the number system property, for the Gaussian integer $\theta = a + bi$ the system $(\mathbb{Z}[i], \theta, D_2)$ is a number system iff $b = \pm 1$ and $a \neq 0, 1, 2, -2, 3$ (see [49]).

Consider now the radix system defined by the cns-polynomial (iv) in Assertion 6 with canonical digit set. Let $k$ be fixed. Suppose that the associated transition graph is primitive. Then, according to the results of [90] and of section 3.2.4 it is possible to determine the Hausdorff dimension of $\partial H$.

## 5.5 Tiles and tilings

A *tiling* is a collection $\mathcal{T}$ of nonempty compact subsets of $\mathbb{R}^k$, called *tiles*, such that (1) each tile is the closure of its interior, (2) $\bigcup_{T_i \in \mathcal{T}} T_i = \mathbb{R}^k$ and (3) the distinct tiles are non-overlapping. Non-overlapping means that the interiors are disjoint. A tiling is a *periodic tiling* if it is invariant under $k$ linearly independent translations, *non-periodic* otherwise. A *lattice tiling* is a tiling by translates of a single tile to the points of a lattice. Note that lattice tilings are periodic tilings. A *self-replicating tiling* is a tiling $\mathcal{T}$ by translates of a single tile such that there is a linear expansive map $A$ with the following property. For each tile $T \in \mathcal{T}$ the image of $A(T)$ is tiled by copies of tiles in $\mathcal{T}$. It must be noted that there are self-replicating tilings which are not lattice tilings. Let an example be the following in $\mathbb{R}$ (see [5, 82, 83]). Let $T_i = [i, i+1] \cup [i+2, i+3], i \in \mathbb{Z}$ and let $A(T) = 4T$. Clearly, $A$ is expansive and $A(T_j) = T_{4j} \cup T_{4j+1} \cup T_{4j+8} \cup T_{4j+9}$. It is a periodic tiling with period lattice $4\mathbb{Z}$ but it is not a lattice tiling.

A *self-affine tile* in $\mathbb{R}^k$ is a nonempty compact set $T$ of positive Lebesgue measure with $A(T) = \bigcup_{a \in D}(a + T)$, where $A$ is an expanding $k \times k$ real matrix with $|\det(A)| = t$ an integer, $D = \{a_1, \ldots, a_t\} \subseteq \mathbb{R}^k$ is a set of $t$ digits and the union is non-overlapping. We remark that for any expanding matrix $A$ and finite set $D$ in $\mathbb{R}^k$ the previous equation determines a unique compact set $T$, the set of numbers with zero integer part. However, uniqueness does not hold in the converse direction. In fact, any self-affine tile $T$ arises from infinitely many different pairs $(\tilde{A}, \tilde{D})$. Self-affine tiles arises in many topics, see [5, 107, 108] and the references there. A *self-similar tile* is a special kind

of self-affine tile, for which the matrix $A$ is a similarity, i.e., $A = \lambda Q$ where $\lambda > 1$ and $Q$ is an orthogonal matrix. Self-similar tiles are somewhat easier to analyze than general self-affine tiles. Self-similar tiles are sometimes called *rep-tiles*.

There is a nice connections between self-replicating tilings and self-affine tiles. R. Kenyon proved [58] that all the tiles in any self-replicating tiling are necessarily self-affine tiles $H = \mathcal{F}(M, D)$ for some digit set $D$. Conversely, every self-affine tile $H$ serves as a prototile for some self-replicating tiling [82]. The following result is the Tiling theorem of self-affine tiles [83]. If $H = \mathcal{F}(M, D)$ is a self-affine tile containing an open set then there exists a set $\mathcal{L} \subseteq \Gamma - \Gamma$ such that $\mathcal{L} + H$ tiles $\mathbb{R}^k$. Note, that no lattice is mentioned in the theorem. On the other hand, if $\mathcal{L} = \Lambda = \Gamma - \Gamma$ then $(\Lambda, M, D)$ has the JTC property.

Recall that the set $\Gamma$ is $M$-invariant, i.e., $M(\Gamma) \subseteq \Gamma$. In the same way, $\Gamma - \Gamma$ is $M$-invariant as well. Let $\mathbb{Z}(M, D)$ denote the smallest $M$-invariant lattice containing $\mathcal{B} = D - D$. A self-affine tile $H = \mathcal{F}(M, D)$ has a lattice tiling with the lattice $\mathbb{Z}(M, D)$ if and only if $\Gamma - \Gamma = \mathbb{Z}(M, D)$ [83].

It is easy to see the connection between JTC systems and self-affine lattice tilings. With the notations already adopted we have the following result.

**Assertion 9.** *If $(\Lambda, M, D)$ is a JTC system then (1) the fundamental domain $H$ is a self-affine tile with $0 \in \mathrm{int}(H)$, (2) the tiling is a lattice tiling and (3) $\Lambda$ is the smallest $M$-invariant lattice containing $D - D$.*

Summarizing the results of this chapter with respect to the tiling properties an algorithm was provided that determines for a given radix system $(\Lambda, M, D)$ whether or not it is a JTC system. Recall that in chapter 3 number system constructions, hence, constructions of self-affine lattice tilings were discussed. More details about existence, structure and tiling properties of general self-affine tiles can be found in the paper of J. C. Lagarias and Y. Wang [83]. We end this chapter with an interesting conjecture of A. Vince: if $(\Lambda, M, D)$ is a radix system then there is some lattice tiling using only translates of $H = \mathcal{F}(M, D)$.

# Chapter 6

# Summary and further directions

*"The art of asking the right questions in mathematics is more important than the art of solving them."*
*— G. Cantor*

In this chapter we summarize the results of this work, enumerate some open problems and provide further directions related to number expansions in lattices.

The results are as follows:
**A.** Concerning the examination of number expansions:

1 In case of a given endomorphism $M : \Lambda \to \Lambda$ and digit set $D \subset \Lambda$, $0 \in D$ a necessary and a sufficient condition were given for satisfying the unique representation property (Assertions 1 and 2).

2 It was stated that a basis transformation in $\Lambda$ does not change the number system property (Assertion 3).

3 Generating the digits of an expansion the function $\Phi$ was considered. It was observed that the path $z, \Phi(z), \Phi^2(z) \ldots$ is ultimately periodic for all $z \in \Lambda$. The set of periodic elements were denoted by $\mathcal{P}$. With the aid of the function $\Phi$ the attractor set $\mathcal{G}(\mathcal{P})$ of $\Lambda$ was defined. It was proved that the radix system $(\Lambda, M, D)$ is a number system if and only if $\mathcal{G}(\mathcal{P}) = \{0 \to 0\}$ (Assertion 4).

4 It was shown that for any radix system $(\Lambda, M, D)$ the lattice points are classified by the attractor set $\mathcal{G}(\mathcal{P})$, i.e. two lattice points $x, y \in \Lambda$ are in the same class if and only if $\Phi^{l_1}(x) = \Phi^{l_2}(y)$ for some non-negative integers $l_1, l_2$. In order to obtain the classification it was proved that all the periodic elements are inside a compact set $-H$ where $H$ is the set of fractions (or fundamental domain) in $\mathbb{R}^k$. Determining the lattice points inside the fundamental domain two approaches (a covering construction and an operator norm construction) were used (Theorem 1). Then, applying an iterated function system, an effective algorithm was presented in order to perform the classification (CLASSIFICATION ALGORITHM).

5 Methods were developed for the fast computation of the function $\Phi$ (section 2.4).

6 For the length of expansion of an arbitrary $z \in \Lambda$ an estimate was proved (Assertion 5).

**B.** Concerning number system constructions:

1 It was introduced the notion of $j$-canonical number systems and equivalent statements were proved for the existence of $j$-canonical complete residue systems (Theorem 2).

2 It was stated that number expansions in algebraic number fields are special cases of number expansions in $\mathbb{Z}^k$. In these cases, the linear transformation $M$ has a simple form in the appropriate power basis, namely the Frobenius matrix of a monic irreducible polynomial over $\mathbb{Z}[x]$. It was shown how to extend this concept to arbitrary monic polynomials over $\mathbb{Z}[x]$ obtaining canonical radix constructions. We called these polynomials as cns-polynomials (or having the cns-property). Necessary conditions for the cns-property were discussed (Lemmas 8 and 9). A large family of polynomials in $\mathbb{Z}[x]$ was proved to be cns-polynomials (Assertion 6). Indeed, it was shown that there are infinitely many cns-polynomials (therefore canonical number systems) in each dimension even if the constant term of the polynomial is "small".

3 Searching for all cns-polynomials in case of a given degree and constant term an algorithm was presented (CNS-SIEVE ALGORITHM).

4 There were given all cns-polynomials up to the degree 8 with constant term $c_0 = 2$.

5 In general, for a given radix $M$ a sufficient condition was given, in which case there is a digit set $D$ for which $(\Lambda, M, D)$ is a number system (Assertions 7 and 8). The digit set can be constructed. This theorem, which is sharper than the earlier results, shows that a wide class of matrices can serve as bases for some number systems.

**C.** Concerning canonical expansions in imaginary quadratic fields:

1 In case of imaginary quadratic fields using canonical digit sets the attractor set $\mathcal{G}(\mathcal{P})$ was completely described, i.e, the number, location and structure of periodic elements was fully determined (Theorems 3.1, 3.2, 4, and 5).

2 In the Gaussian ring for certain bases a special property was proved (Theorem 6).

**D.** Concerning the geometry of expansions:

1 An algorithm was presented for plotting the points of the fundamental domain $H$ (ESCAPE ALGORITHM). This set is the unique invariant (or attractor) set of an iterated function system determined by the radix system $(\Lambda, M, D)$.

2 It was analyzed the just touching covering property of radix systems and with the aid of the transition graph an algorithm was given to decide this property (TRANSITION GRAPH CONSTRUCTION ALGORITHM). It was also given an algorithm for computing the boundary of the set $H$ without computing the interior points.

3 Via the construction of the transition graph it was determined all just touching covering systems in the Gaussian ring using canonical digit sets, included the exact values of the Hausdorff dimension of the boundary of their fundamental domain (Proposition 1).

4 Finally, some remarks were made on just touching covering properties of radix systems.

The author's main results are: A1 (Assertion 2), A4 (CLASSIFICATION ALGORITHM), A5, A6, B1, B2 (Assertion 6), B3, B4, B5, C1, C2.

Now consider some open problems and further directions.

**1.** Let a radix system $(\Lambda, M, D)$ be given. The following questions arise naturally (see also page 37 and [64]). It is known that if $p \in \mathcal{P}$ then the maximum of the period length of $p$ can be estimated with the number of lattice points in the $k$-dimensional ball centered at 0 with radius $L$. Is there a better estimation? Is there a good upper estimation for the number of different sets $\mathcal{C}(p)$? Give all the bases $M$ mapping $\Lambda$ to $\Lambda$ for which there exist a complete residue system $D$ modulo $M$ such that $(\Lambda, M, D)$ is a number system. How can be characterized the geometric,–algebraic structure of the sets $\mathcal{B}(p)$, $p \in \mathcal{P}$ (e.g. symmetry)? What can be stated about the attractor set in case of special operators, e.g. matrices generated by the ring of integers of a given algebraic number field? The problem of characterizing the $j$-canonical number systems seems to be interesting. It is known that if $z \in \mathcal{B}(0)$ for all $\|z\| \leq L$ then the unique representation property holds. Instead of $L$ is there a better estimation? This is a critical problem for examining number systems algorithmically, since $L$ can be very large.

**2.** Let a number system $(\Lambda, M, D)$ be given. Desing and implement the basic operations (addition, subtraction, multiplication, division) in this system. For special digit sets — where the digits are the $k$-th root of unity — some important results are available [100]. What about the canonical digit sets? The real problem is the division. For the ring of Gaussian integers it was analyzed by W. Gilbert [32] and by I. Kátai [47] independently, using different methods. It seems that the method of I. Kátai can be generalized.

**3.** Topological questions are also very interesting. Let a radix system be given. Is the fundamental domain $H$ (arcwise) connected,–disconnected? When the projections of $H$ to lines are intervals? What about the geometric,–algebraic,–measure theoretic properties of a non-empty intersection of $H$ with a hyperplane of $\mathbb{R}^k$? The question of characterizing JTC systems in different domains using various digit sets seems to be very hard.

**4.** Let the standard expansion of $z \in \mathbb{Z}^k$ be $\sum_{i=0}^{j-1} M^i a_i + M^j \pi, a_i \in D, \pi \in \mathcal{P}$. I. Kátai introduced the set of $(M, D)$-additive and $(M, D)$-multiplicative functions by $\mathcal{E}_{(M,D)} = \{f : \mathbb{Z}^k \to \mathbb{R}, f(M^r \pi) = 0$ for every $\pi \in \mathcal{P}, r \in \mathbb{N}_0$ and for every $z \in \mathbb{Z}^k$ $f(z) = \sum_{i=0}^{j-1} f(M^i a_i)\}$ and by $\mathcal{M}_{(M,D)} = \{g : \mathbb{Z}^k \to \mathbb{C}, g(M^r \pi) = 0$ for every $\pi \in \mathcal{P}, r \in \mathbb{N}_0$ and for every $z \in \mathbb{Z}^k$ $g(z) = \prod_{i=0}^{j-1} g(M^i a_i)\}$. There are lots of interesting questions which can be stated, we refer the reader to [51].

**5.** Canonical number systems can be one of the links between number theory and theoretical computer science via automatic sequences. A sequence is called $(M, D)$-automatic if — roughly speaking — its $n$-th term can be generated by a finite state automaton from the digits of the radix expansion of $n$. This concept was studied by many authors, see [3] and the references there. The positional (or radix) systems are special cases of numeration systems generated by a strictly increasing sequence $G = \left(G_n\right)_{n\geq 0}$ of positive integers with $G_0 = 1$. Such a sequence is called *G-scale*. Using the greedy algorithm (see e.g. A. S. Fraenkel [24]) every natural number can be expanded in the form

$$n = \varepsilon_0(n)G_0 + \ldots + \varepsilon_l(n)G_l, \tag{6.1}$$

where the digits $\varepsilon_j(n) \in \mathbb{N}_0$ satisfy $0 \leq \varepsilon_j(n) < G_{j+1}/G_j$. The so-called *G-expansion* in (6.1) is unique provided that $\varepsilon_0(n)G_0 + \ldots + \varepsilon_j(n)G_j < G_{j+1}$ for all $j$ $(0 \leq j \leq l)$. In this way the natural numbers can be identified to a sequence of non-negative integers by $n \to m0^\infty = e_0e_1 \ldots e_l0^\infty$, $(e_l \neq 0)$. The set $\mathcal{L}(G)$ of words $m$ is called the source language of G. If $\mathcal{L}(G)$ is regular (i.e. recognizable by an automaton) then G must be a linear recurrent sequence with integer coefficients (see J. Shallit [101]). Another direction of the investigations is the sum-of-digit function of the G-expansions. It has been extensively studied because of its nice structural properties ([35, 36]).

**6.** In this work we considered only number expansions in lattices. Clearly, number expansions can be defined in many different ways. The most common is the following. The *$\beta$-expansion* of $x \in [0, 1]$ is a sequence of integers of $\{0, 1, \ldots, \lfloor\beta\rfloor\}$ with $d_n = \lfloor\beta f_\beta^{n-1}(x)\rfloor, n \geq 1$, where $f_\beta(x) = \beta x - \lfloor\beta x\rfloor = \beta x \mod 1$. These expansions were studied by many authors, see e.g. [8, 56, 92]. The concept was generalized to interval filling sequences and to univoque sequences by Z. Daróczy and I. Kátai [11, 12, 13]. Recently, there is a PhD thesis on univoque numbers [46]. There are many other kinds of number expansions (e.g. Balkema-Oppenheim expansions [61], etc.) which are rather different from our construction. Finally, a brand new theory opens in examining number expansions if one leaves the lattice for some non-Euclidean space.

# A Applications

*"A man who loves practice without theory is like the sailor*
*who boards ship without a rubber and compass*
*and never knows where he may cast."*
*— Leonardo da Vinci*

In this section we points out some possible applications, mainly referring to some papers.

Generalized number systems can be very interesting in computer algebra, since they enables us error-free computations. Recall that the problems regarding number expansions in algebraic number fields are special cases of problems in $\mathbb{Z}^k$. Computing efficiently in an algebraic number field one might choose an appropriate number system representation in order to perform fast calculations either sequentially or parallel. Obviously, one has to choose systems — if it is possible at all —, for which the basic operations can be made efficiently (see also section 6 Problem 2).

A. Pethő proposed a public key cryptosystem based on canonical number systems in $\mathbb{Z}^k$ [95]. His cryptosystem is related to the Merkle-Hellman knapsack scheme.

It is not yet clear in which cases and how generalized canonical number systems can be applied for data compression or in telecommunication in order to reduce the number of transmitted packets. Nevertheless, this research direction could be very interesting. (See also Example 1 in section B.)

A. Vince in his nice introductory exposition [107] enumerates many topics, where recently self-replicating tilings come under investigation. Without giving the exact references — which can be found in his paper — we mention a few of them.

- Wavelet bases construction;

- Multi-resolution analysis;
- Crystallographic;
- Finite state machines and Markov partitions in dynamical systems;
- Ergodic theory and statistical mechanics;
- Image processing and computer vision.

# B Examples

*"In Riemann, Hilbert or in Banach space*
*Let superscripts and subscripts go their ways.*
*Our asymptotes no longer out of phase,*
*We shall encounter, counting, face to face."*
*— Stanislaw Lem, Cyberiad*

This chapter contains some examples regarding number expansions in lattices.

**Example 1.** Let a 13 decimal digit number $n = 1003462401565$ be given. Let us denote the Frobenius matrix of the cns-polynomial $2 - x + x^4$ by $M_1$. Using the correspondences $0 = [0,0,0,0]^T$ and $1 = [1,0,0,0]^T$ we have that

$$(n)_{10} = (1110100110100011000001010001111000011101)_2 =$$

$$= ([29,0,0,0]^T)_{M_1}$$

and 29 is only a 2 decimal digit number.

In the same way, let $n = 20220584137211351918878806846970568755 37$ be a 40 decimal and 131 binary digit number. Again, if we consider the Frobenius matrix $M_2$ of the cns-polynomial $2 + 4x + 5x^2 + 5x^3 + 5x^4 + 4x^5 + 3x^6 + 2x^7 + x^8$ and the appropriate abbreviations as above, we get that the expansion of $n$ is

$$(n)_{10} = ([29,0,0,0,0,0,0,0]^T)_{M_2}$$

and 29 has only 5 binary digits, $(29)_{10} = (11101)_2$. It would be interesting to characterize all the rational integers which have shorter expansions in *some*

generalized binary number system than in the traditional (one-dimensional) binary case.

**Example 2.** Let $\Lambda = \mathbb{R}, M = (3), D = \{-2, 0, 2\}$.
Clearly, $\mathcal{G}(\mathcal{P}) = \{-1 \to -1, 0 \to 0, 1 \to 1\}, \mathcal{B}(1) = \{\text{positive odd numbers}\}$, $\mathcal{B}(-1) = \{\text{negative odd numbers}\}, \mathcal{B}(0) = \{\text{even numbers}\}$. The fundamental set $H$ is the interval $[-1, 1]$. The system $(\Lambda, M, D)$ is not a number system, not a just touching covering system, but $0 \in \text{int}(H)$ and it is a self-affine lattice tiling with the lattice $2\mathbb{Z}$.

**Example 3.** Let $\Lambda = \mathbb{Z}^2, M = \left(\begin{smallmatrix} 0 & -3 \\ 1 & 0 \end{smallmatrix}\right), D = \{\binom{0}{0}, \binom{1}{0}, \binom{-1}{1}\}$.
Now, $\mathcal{G}(\mathcal{P}) = \{\binom{-1}{0} \to \binom{-1}{0}, \binom{0}{0} \to \binom{0}{0}\}$. Hence it is not a number system. On the other hand, computations show that it is a JTC system and the graph $\mathcal{G}(\mathcal{S})$ has two strongly connected components. Let us denote the domain of attraction $\mathcal{B}(\binom{-1}{0})$ by black and $\mathcal{B}(\binom{0}{0})$ by white. Figure 2 shows the $400 \times 400$ region of $\mathbb{Z}^2$ centered at the origin.

**Example 4.** Let $\Lambda = \mathbb{Z}[i]$ be the ring of Gaussian integers.

(a) Let $M = \left(\begin{smallmatrix} 2 & -1 \\ 1 & 2 \end{smallmatrix}\right)$ and $D$ be the canonical digit set. Then the eigenvalues of $M$ are $2 \pm i$, $r = \|M^{-1}\| = \sqrt{5}/5$. The attractor set $\mathcal{G}(\mathcal{P})$ is $\{0 \to 0, -1 + i \to -1 + i, -2 + 2i \to -2 + 2i\}$. Let us denote the domain of attraction $\mathcal{B}(0)$ by black, $\mathcal{B}(-1 + i)$ by white and $\mathcal{B}(-2 + 2i)$ by gray. Figure 3 shows the $400 \times 400$ region of $\mathbb{Z}[i]$ centered at the origin. The fundamental domain $H$ in the region $\{(x, y), x \in [-0.5, 2.5], y \in [-2.5, 0.5]\}$ can be seen in Figure 4. The set $H$ is arcwise connected, its boundary has the Hausdorff dimension approximately 1.6087. The system is a JTC system.

(b) Let $M = \left(\begin{smallmatrix} 3 & 0 \\ 0 & 3 \end{smallmatrix}\right)$, $D_1 = \{0, \pm 1, \pm i, \pm 1 \pm i, \pm 1 \mp i\}$ and $D_2 = \{0, 1, 2, i, 2i, 1 + 2i, 2 + i, -1 + 2i, -2 + i\}$. The fundamental domain $\mathcal{F}(M, D_1)$ is just the unit square centered at the origin. The set $\mathcal{F}(M, D_2)$ in the region $\{(x, y), x \in [-1, 1], y \in [0, 1]\}$ can be seen in Figure 5. It is proved to be connected. The system $(\Lambda, M, D_2)$ is not a number system but it is a JTC system. The radix representations in these systems essentially separate a complex number into its real and imaginary parts.

(c) Let $M_1 = \left(\begin{smallmatrix} 2 & -1 \\ 1 & 2 \end{smallmatrix}\right)$, $M_2 = \left(\begin{smallmatrix} -2 & -1 \\ 1 & -2 \end{smallmatrix}\right)$ and let $D \subset \{a + bi, a, b \in \mathbb{Z}, -3 \leq a, b \leq 3\}$ be a full residue system that contains 0. Then $(\Lambda, M_1, D)$ is a number system in 127 different cases while $(\Lambda, M_2, D)$ is a number system in 2488 different cases. The boundary of the fundamental domain $H = \mathcal{F}(M_1, \{0, \pm 1, \pm i\})$ can be seen in Figure 6. Its Hausdorff dimension is approximately 1.3652. The set $H$ is the same as the set constructed by B. Mandelbrot from a generalized Koch curve [85].

**Example 5.** Let $\Lambda = \mathbb{Z}^2$, $M = \left(\begin{smallmatrix} 1 & -2 \\ 1 & 1 \end{smallmatrix}\right)$ and $D = \left\{\left(\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}\right), \left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right), \left(\begin{smallmatrix} -1 \\ 0 \end{smallmatrix}\right)\right\}$.
Then $(\Lambda, M, D)$ is a number system, its fundamental domain can be seen in Figure 7. About this polygonal radix system see [100] for further references.
**Example 6.** Let us consider the ring of Gaussian integers with base $\theta = A + Bi$ and a canonical digit set. Let $A = 5$ and $B = 12$. We shall use the notations of section 4.3.

If $\mu = 1$ then $\varphi(B_1) = 4$, $\mathrm{ord}_{B_1} A = 2$. Therefore there are two cycles with period lenght 2. The periodic elements are $i \to -2 + 5i \to i$ and $-2 + 7i \to -4 + 11i \to -2 + 7i$.

If $\mu = 2$ then $\varphi(B_2) = 2$, $\mathrm{ord}_{B_2} A = 2$. Therefore there is one cycle with period length 2, namely $2i \to -4 + 10i \to 2i$.

If $\mu = 3$ then $\varphi(B_3) = 2$, $\mathrm{ord}_{B_3} A = 1$. Therefore there are two cycles with period length 1, namely $-1 + 3i \to -1 + 3i$ and $-3 + 9i \to -3 + 9i$.

If $\mu = 4$ then $\varphi(B_4) = 2$, $\mathrm{ord}_{B_4} A = 2$. Therefore there is one cycle with period length 2, namely $-1 + 4i \to -3 + 8i \to -1 + 4i$.

If $\mu = 6$ then $\varphi(B_6) = 1$, $\mathrm{ord}_{B_6} A = 1$. Therefore there is one cycle with period length 1, namely $-2 + 6i \to -2 + 6i$.

If $\mu = 12$ then there are two cycles with period length 1, namely $0 \to 0$ and $-4 + 12i \to -4 + 12i$.

**Example 7.** Let $\Lambda = \mathbb{Z}[i]$ be the ring of Gaussian integers. Let $\theta = -3 + i$ and consider the canonical digit set $D = \{0, 1, \ldots, 9\}$. The system $(\Lambda, \theta, D)$ is a number system, its fundamental domain can be seen in Figure 8. It is arcwise connected, its boundary has the Hausdorff dimension approximately 1.5495. Observe that the system is a staightforward generalization of the traditional decimal number system.

**Example 8.** Let $\Lambda = \mathbb{Z}[i]$ be the ring of Gaussian integers again.

(a) Let $M = \left(\begin{smallmatrix} 2 & -1 \\ 1 & 2 \end{smallmatrix}\right)$ and $D = \left\{\left(\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}\right), \left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right), \left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right), \left(\begin{smallmatrix} 0 \\ -1 \end{smallmatrix}\right)\left(\begin{smallmatrix} -6 \\ -5 \end{smallmatrix}\right)\right\}$. Then $(\Lambda, M, D)$ is a number system, its fundamental domain $H$ can be seen in Figure 9. The set $H$ is disconnected, but clearly it is a lattice tiling.

(b) Let $M = \left(\begin{smallmatrix} 2 & -1 \\ 1 & 2 \end{smallmatrix}\right)$ and $D = \left\{\left(\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}\right), \left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right), \left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right), \left(\begin{smallmatrix} 0 \\ -1 \end{smallmatrix}\right)\left(\begin{smallmatrix} -2 \\ -3 \end{smallmatrix}\right)\right\}$. Then $(\Lambda, M, D)$ is a number system, its fundamental domain $H$ can be seen in Figure 10. The set $H$ is disconnected. It is a lattice tiling. The approximation of the fundamental domain by the ESCAPE ALGORITHM can be seen in Figure 11.

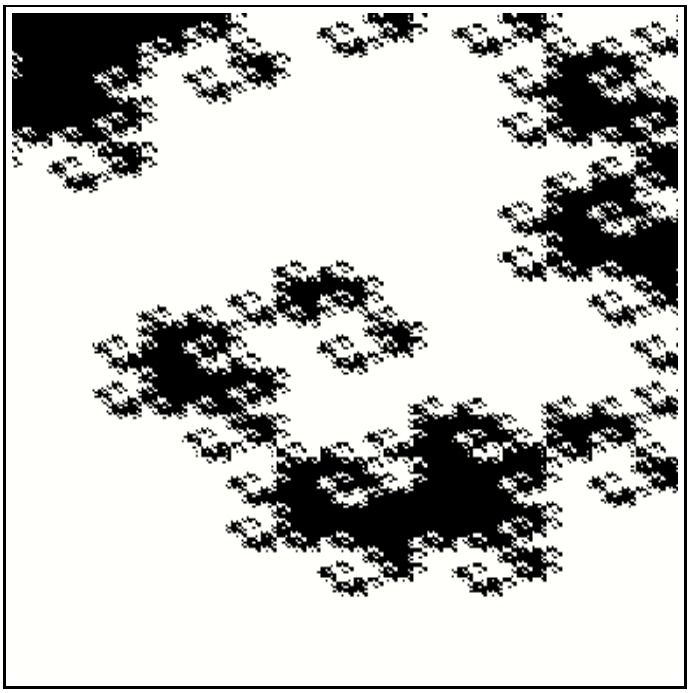

Figure 2

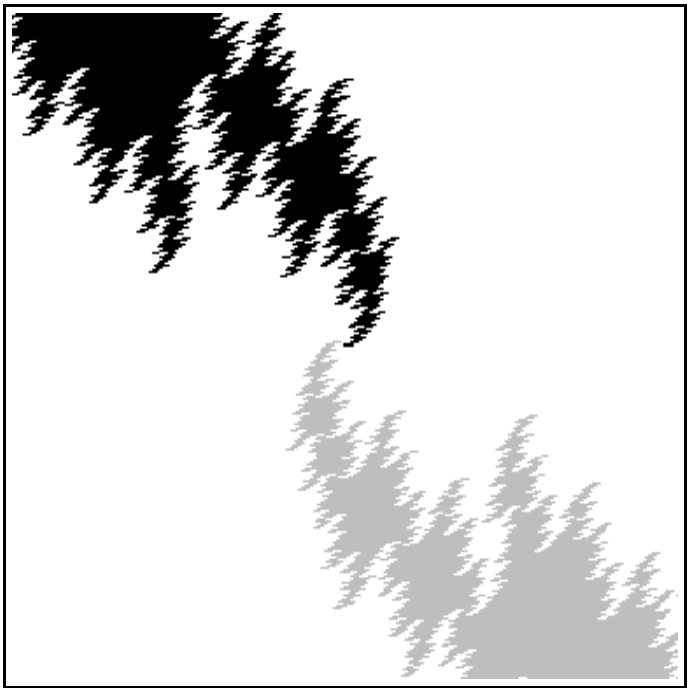

Figure 3

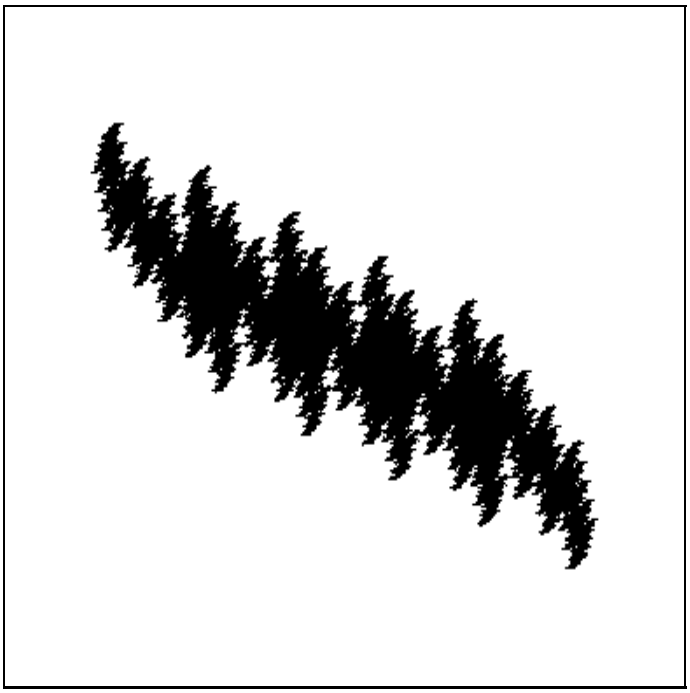

Figure 4

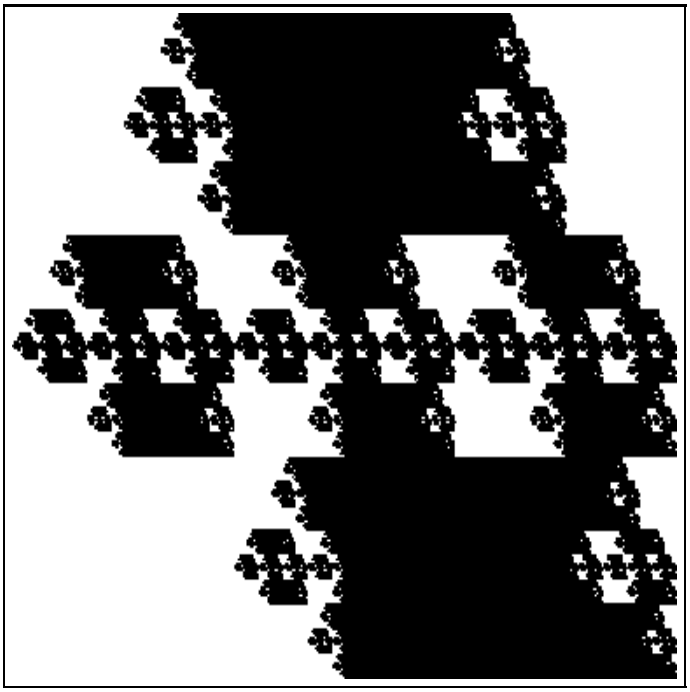

Figure 5

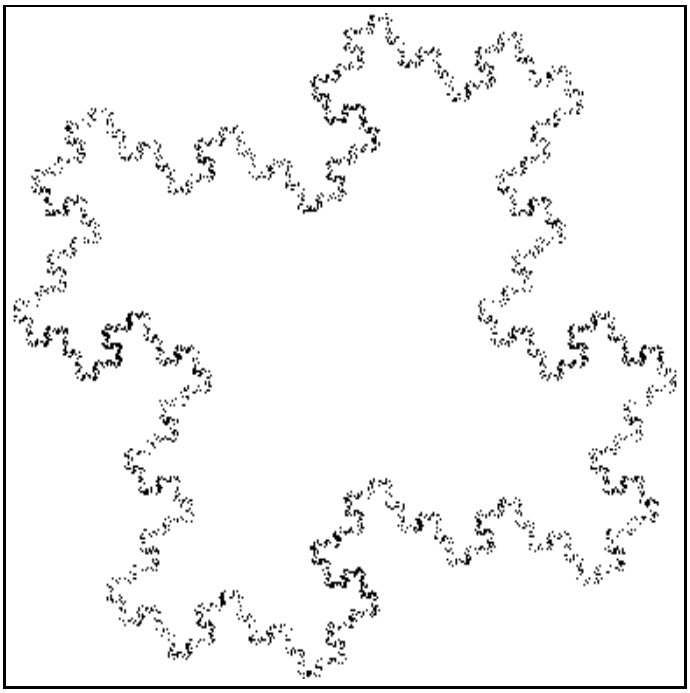

Figure 6

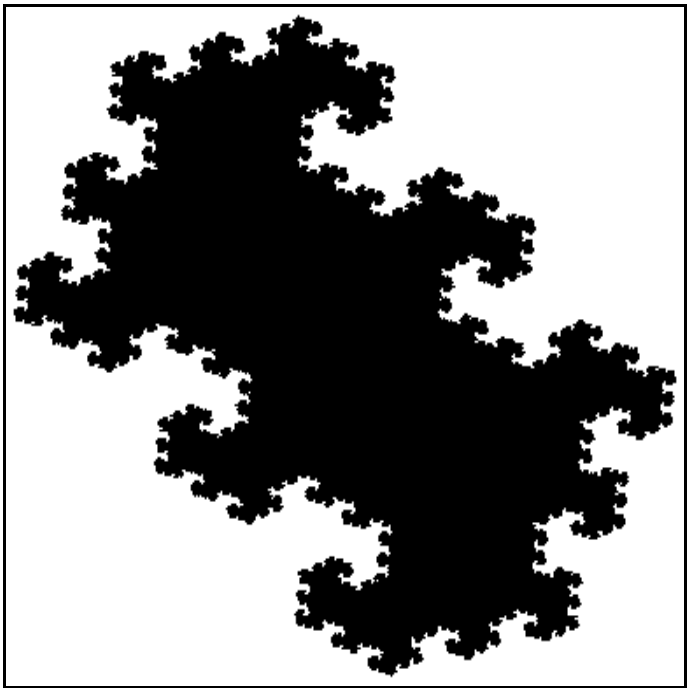

Figure 7

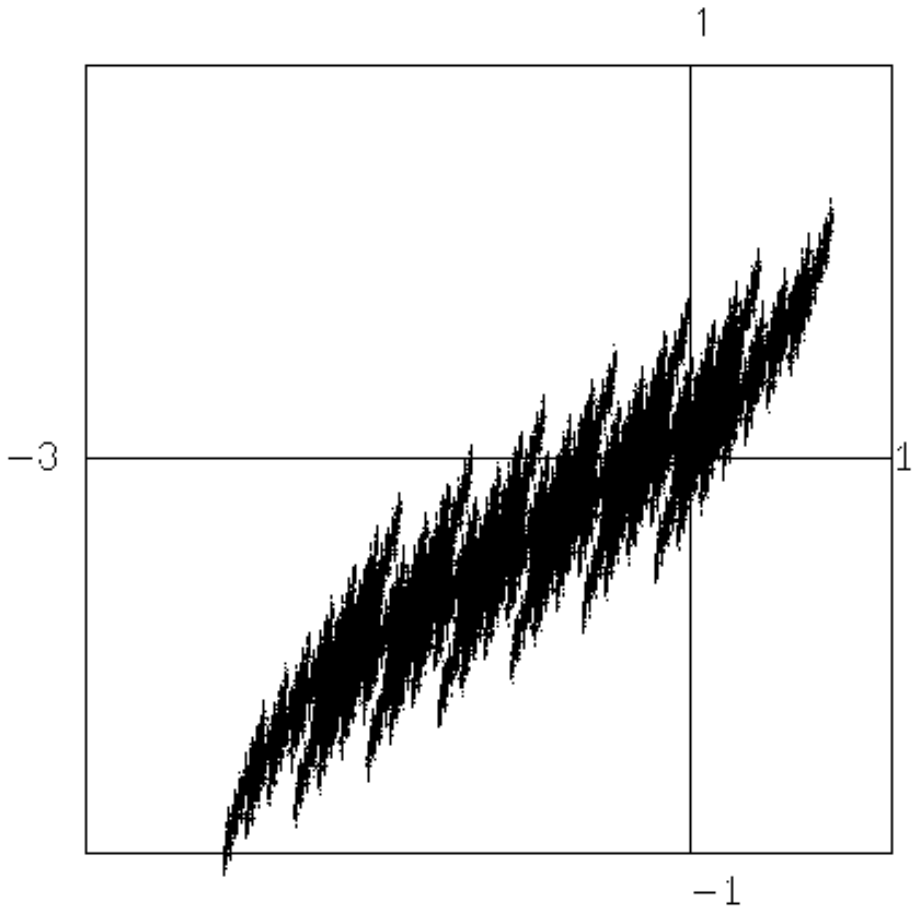


Figure 8

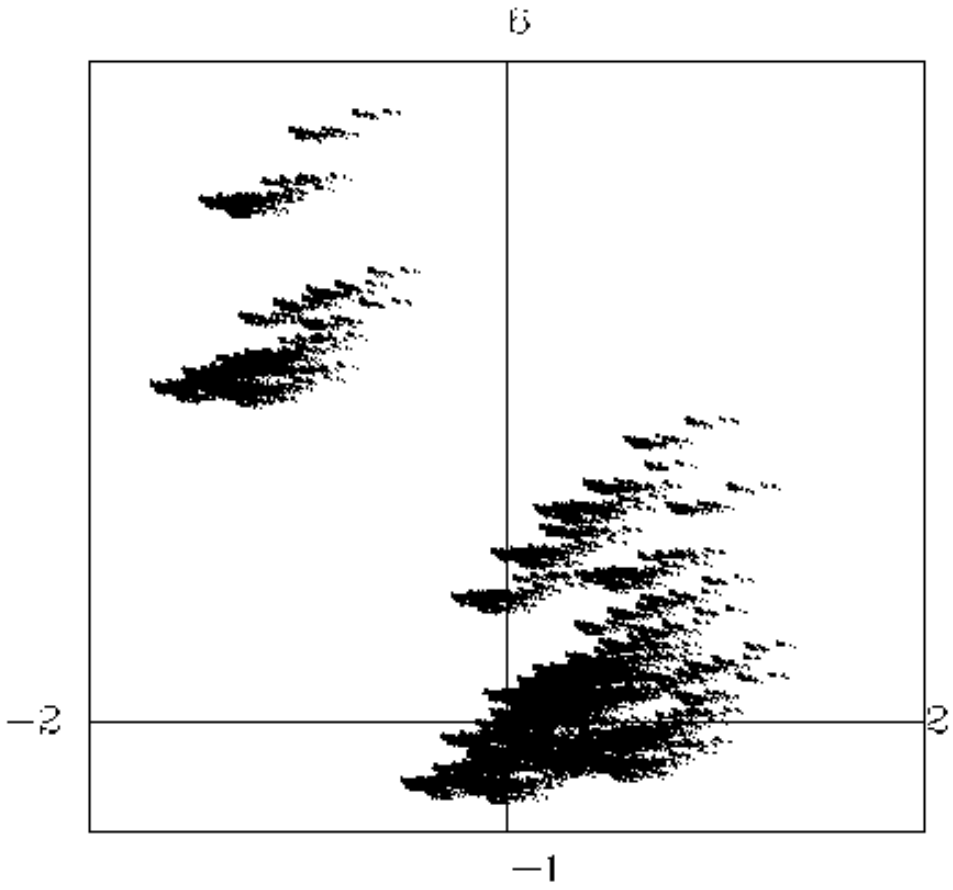


Figure 9

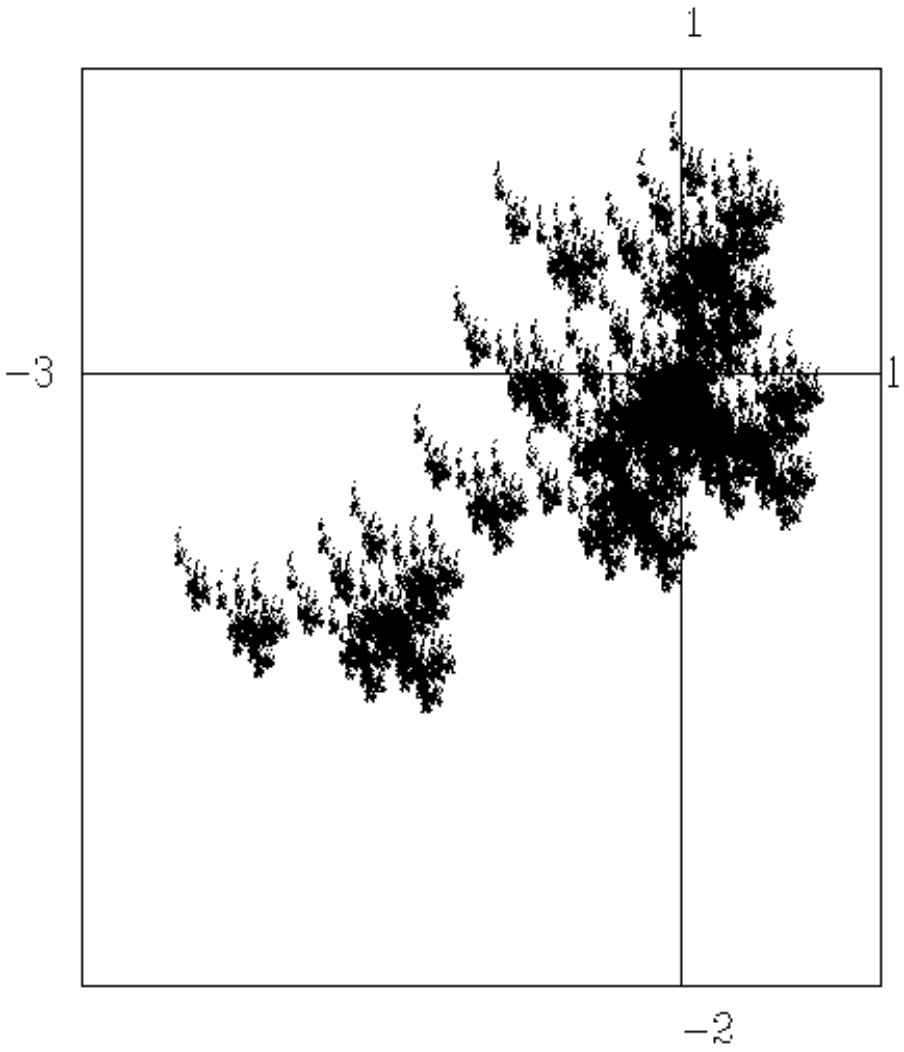


Figure 10

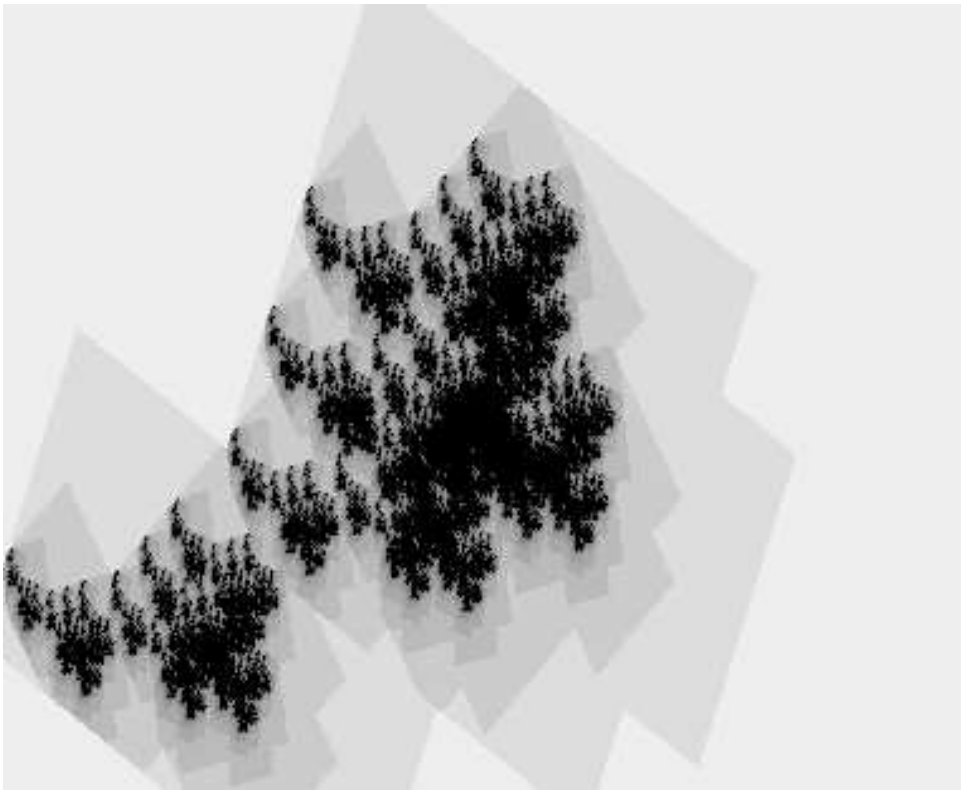

Figure 11

# C Bibliography

*"Knowledge is of two kinds. We know a subject ourselves, or we know where we can find information upon it."*
*— Samuel Johnson*


[1] Akiyama, S., Pethő, A., *On canonical number systems*, preprint, 1999, 1–12, submitted to Theor. Comp. Sci.

[2] Akiyama, S., Thuswaldner, J. M., *Topological properties of two dimensional number systems*, Journal de Théorie des Nombres de Bordeaux **12**, 2000, 69–79.

[3] Allouche, J.-P., Cateland, E., Gilbert, W. J., Peitgen, H.-O., Shallit, J., Skordev, G., *Automatic maps in exotic numeration systems*, Theory Comp. Syst. (Math. System Theory) **30**, 1997, 285–331.

[4] Bandt, C., *Self-similar sets 5. Integer matrices and fractal tilings of* $\mathbb{R}^n$, Proc. Am. Math. Soc. **112**, 1991, 549–562.

[5] Bandt, C., *Classification of self-affine lattice tilings*, J. London Math. Soc. **50**/3, 1994, 581–593.

[6] Barnsley, M., *Fractals everywhere*, Academic Press Inc. 1988.

[7] Benedek, A., Panzone, R., *On the Eisenstein set*, Revista de la Unión Matemática Argentina **41**/2, 1998, 177–185.

[8] Blanchard, F., $\beta-$*expansions and symbolic dynamics*, Theor. Comp. Sci. **65**, 1989, 131–141.

[9] Brualdi, R., Ryser, H., *Combinatorial Matrix Theory*, Encyclopedia of Mathematics and its Applications, Cambridge Univ. Press, 1991.

[10] Brunotte, H., *On trinomial bases of radix representations of algebraic integers*, preprint, 2000, 1–9, submitted to Acta Sci. Math.

[11] Daróczy, Z., Kátai, I., *Generalized number systems in the complex plane*, Acta Math. Hung. **51**/3-4, 1988, 409–416.

[12] Daróczy, Z., Kátai, I., *Univoque sequences*, Publ. Math. Debrecen **42**, 1993, 397–407.

[13] Daróczy, Z., Kátai, I., *On the structure of univoque numbers*, Publ. Math. Debrecen **46**, 1995, 385–408.

[14] Davio, M., Deschamps, J. P., Gossart, C., *Complex arithmetic*, Philips MBLE Research Lab. Report R369, Brussels, May 1978.

[15] Dekking, F. M., *Recurrent sets*, Adv. Math. **44**, 1982, 78–103.

[16] Dekking, F. M., *A self-similar tiling of Euclidean space by two shapes in two sizes*, J. Phys. A. **29**/9, 1996, 2123–2126.

[17] Edgar, G. A., *Measure, topology and fractal geometry*, Springer, 1990.

[18] Falconer, K.J., *Fractal geometry*, Wiley Inc., 1990.

[19] Falconer, K.J., *Techniques in Fractal Geometry*, Wiley Inc., 1997.

[20] Farkas, G., *Number systems in real quadratic fields*, Annales Univ. Sci. Bud., Sect. Comp. **18**, 1999, 47–59.

[21] Farkas, G., *Digital expansions in real algebraic quadratic fields*, Mathematica Pannonica **10**/2, 1999, 235–248.

[22] Farkas, G., *Location and number of periodic elements in* $\mathbb{Q}(\sqrt{2})$, Annales Univ. Sci. Bud., Sect. Comp. **20**, 2001, to appear

[23] Farkas, G., *Investigation of periodic elements in* $\mathbb{Q}(\sqrt{2})$, Proc. 5th International Conference on Applied Informatics, Eger, Hungary, submitted to Comp. Math. Appl.

[24] Fraenkel, A. S., *Systems of numeration*, Amer. Math. Monthly **92**, 1985, 105–114.

[25] Frougny, C., *Representation of numbers and finite automata*, Math. System Theory **25**, 1992, 37–60.

[26] Gilbert, W. J., *Arithmetic in Complex Bases*, Math. Mag. **57**/2, March 1984, 77–81.

[27] Gilbert, W. J., *Radix representation of quadratic fields*, J. Math. Anal. Appl. **83**, 1991, 264–274.

[28] Gilbert, W. J., *Fractal geometry derived from complex basis*, Math. Intell. **4**, 1982, 78–86.

[29] Gilbert, W. J., *Geometry of radix representation*, in: The Geometric Vein, Springer, 1981, 129–139.

[30] Gilbert, W. J., *Complex numbers with three radix expansions*, Can. J. Math. **34**, 1982, 1335–1348.

[31] Gilbert, W. J., *The fractal dimension of sets derived from complex bases*, Can. Math. Bull. **29**, 1986, 495–500.

[32] Gilbert, W. J., *The division algorithm in complex bases*, Can. Math. Bull. **39**/1, 1996, 47–54.

[33] Gilbert, W. J., *Gaussian integers as bases for exotic number systems*, The Mathematical Heritage of C. F. Gauss, (ed. by Rassias, G. M.) World Scientific Publ. Co., unpublished manuscript, 1994.

[34] Goffinet, D., *Number systems with a complex base: a fractal tool for teaching topology*, Amer. Math. Monthly **98**, 1991, 249–255.

[35] Grabner, P. J., Liardet, P., *Harmonic properties of the sum-of-digit function for complex bases*, preprint, 1998, 1–19.

[36] Grabner, P. J., Liardet, P., Tichy, R., *Odometers and systems of numeration*, Acta Arithm. **70**, 1995, 103–122.

[37] Grossman, E. H., *Number bases in quadratic fields*, Studia Sci. Math. Hung. **20**, 1985, 55–58.

[38] Hata, M., *On the structure of self-similar sets*, Japan J. Appl. Math. **2**, 1985, 381–414.

[39] Hutchinson, J. E., *Fractals and self-similarity*, Indiana Univ. Math. J. **30**, 1981, 713–747.

[40] Indlekofer, K.-H., Kátai, I., Racskó, P., *Some remarks on generalized number systems*, Acta Sci. Math. **37**, 1993, 543–553.

[41] Indlekofer, K.-H., Kátai, I., Racskó, P., *Number systems and fractal geometry*, Probability Theory and its Applications (Eds. Galambos, J. and Kátai, I.), Kluwer Ac. Publ., 1992, 319–334.

[42] Indlekofer, K.-H., Járai, A., Kátai, I., *On some properties of attractors generated by iterated function systems*, Acta Sci. Math. **60**, 1995, 411–427.

[43] Isaacson, E., Keller, H., *Analysis of numerical methods*, Wiley Inc., 1966.

[44] Ito, S., *On the fractal curves induced from the complex radix expansion*, Tokyo J. Math. **12**, 1989, 300–319.

[45] Járai, A., *Fractals and number systems on computers*, manuscript, 1994.

[46] Kallós, G., *Univoque számok*, PhD thesis, preprint, University of Eötvös Loránd, Budapest, 2001.

[47] Kátai, I., *On the division algorithm for* $\Theta = -1 + i$, $\mathcal{A} = \{0, 1\}$, manuscript, 1993.

[48] Kátai, I., *Number systems in imaginary quadratic fields*, Annales Univ. Sci. Bud., Sect. Comp. **14**, 1994, 91–103.

[49] Kátai, I., *Generalized number systems and fractal geometry*, monograph, Janus Pannonius Univ., Pécs, Hungary, 1995, 1–40.

[50] Kátai, I., *Construction of number systems in algebraic number fields*, Annales Univ. Sci. Bud., Sect. Comp. **18**, 1999, 103–107.

[51] Kátai, I., *On q-additive and q-multiplicative functions*, submitted, 2000, 1–16.

[52] Kátai, I., Kovács, B., *Kanonische Zahlensysteme bei reelen quadratischen algebraischen Zahlen*, Acta Sci. Math. **42**, 1980, 99–107.

[53] Kátai, I., Kovács, B., *Canonical number systems in imaginary quadratic fields*, Acta Math. Hung. **37**, 1981, 159–164.

[54] Kátai, I., Környei, I., *On number systems in algebraic number fields*, Publ. Math. Debrecen **41**, 1992, 289–294.

[55] Kátai, I., Szabó, J., *Canonical number systems for complex integers*, Acta Sci. Math. **37**, 1975, 255–260.

[56] Kempner, A.J., *Anormal system of numeration*, Amer. Math. Monthly **43**, 1936, 610–617.

[57] Keng, H.L., *Introduction to Number Theory*, Springer, 1982.

[58] Kenyon, R., *Self-replicating tilings*, in "Symbolic Dynamics and Its Applications" (Walters, P. Ed.), Contemporary Mathematics, Amer. Math. Soc. **135**, 1992, 239–264.

[59] Kimberling, C., *Numeration systems and fractal sequences*,, Acta Arithm. **73**, 1995. 103–117.

[60] Knuth, D. E., *The art of computer programming*, Vol. 2, Seminumerical Algorithms, 3rd updated and revised ed., Addison-Wesley, 1998.

[61] Kovács, A., *Simulation analysis of complex number systems and Balkema-Oppenheim expansions*, Diploma-thesis, University of Eötvös Loránd, Budapest, 1991.

[62] Kovács, A., Harnos, N., *Fractals and number systems*, Technical Report, Univ. of Paderborn, CeBIT, 1993, 1–62.

[63] Kovács, A., *Sets of complex numbers generated from a polynomial functional equation*, Annales Univ. Sci. Bud., Sect. Comp. **18**, 1999, 115–124.

[64] Kovács, A., *On computation of attractors for invertible expanding linear operators in* $\mathbb{Z}^k$, Proc. Numbers, Functions, Equations '98, Noszvaj, Hungary, Leaflets in Mathematics, Janus Pannonius Univ., (Pécs), 1998, 108–109, Publ. Math. Debrecen **56**/1-2, 2000, 97–120.

[65] Kovács, A., *On expansions of Gaussian integers with non-negative digits*, Math. Pannonica **10**/2, 1999, 177–191.

[66] Kovács, A., *Canonical expansions of integers in imaginary quadratic fields*, submitted to Acta Math. Hung.

[67] Kovács, A., *On number expansion in lattices*, Proc. 5*th* Intern. Conf. on Applied Informatics, Eger, Hungary, 2001, submitted to Comp. Math. Appl.

[68] Kovács, A., *Generalized binary number systems*, Annales Univ. Sci. Bud., Sect. Comp. **20**, 2001, to appear.

[69] Kovács, A., *Computer Algebra: Impact and Perspectives*, Nieuw Archief voor Wiskunde **17**/1, 1999, 29–55.

[70] Kovács, A., *Mathematical background for fractals*, University of Eötvös Loránd, manuscript, http://compalg.inf.elte.hu/∼attila, 1998, 1–13.

[71] Kovács, A., *Komputeralgebra a tudományokban és a gyakorlatban*, Alkalmazott Matematikai Lapok **18**, 1998, 181–202.

[72] Kovács, B., *Canonical number systems in algebraic number fields*, Acta Math. Hung. **37**/4, 1981, 405–407.

[73] Kovács, B., *Integral domains with canonical number systems*, Publ. Math. Debrecen **36**, 1989, 153–156.

[74] Kovács, B., *Representation of complex numbers in number systems*, Acta Math. Hung. **58**, 1991, 113–120.

[75] Kovács, B., *Number systems*, Computational Number Theory, Proc., Walter de Gruyter Publ. Comp., 1991, 21–25.

[76] Kovács, B., Környei, I., *On the periodicity of the radix representation*, Annales Univ. Sci. Bud., Sect. Comp. **13**, 1992, 129–133.

[77] Kovács, B., Pethő, A., *Canonical number systems in the ring of integers*, Publ. Math. Debrecen **30**, 1983, 39–44.

[78] Kovács, B., Pethő, A., *Number systems in integral domains, especially in orders of algebraic number fields*, Acta Sci. Math. **55**, 1991, 287–299.

[79] Kovács, B., Pethő, A., *On a representation of algebraic integers*, Studia Sci. Math. Hung. **27**, 1992, 169–172.

[80] Környei, I., *The Hausdorff dimension of the boundary of sets* $\{z \in \mathbb{C}, z = \sum_{k=1}^{\infty} a_k/\alpha^k, 0 \leq a_k < N(\alpha), deg(\alpha) = 2\}$, Annales Univ. Sci. Bud., Sect. Comp. **36**, 1993, 179–191.

[81] Körmendi, S., *Canonical number systems in* $\mathbb{Q}(\sqrt[3]{2})$, Acta Sci. Math. **50**, 1986, 351–357.

[82] Lagarias, J. C., Wang, Y., *Integral self-affine tiles in* $\mathbb{R}^n$ *I. Standard and nonstandard digit sets*, J. London Math. Soc. **54**/2, 1996, 161–179.

[83] Lagarias, J. C., Wang, Y., *Self-affine tiles in* $\mathbb{R}^n$, Adv. in Math. **121**, 1996, 21–49.

[84] Lehmer, D. H., *A machine method for solving polynomial equations*, J. Assoc. Comp. Mach. **2**, 1961, 151–162.

[85] Mandelbrot B. B., *The fractal geometry of nature*, Freeman, 1983.

[86] Matula, D. W., *Basic digit sets for radix representation*, J. ACM **29**, 1982, 1131–1143.

[87] Mauldin, R. D., Williams, S. C., *Hausdorff dimension in graph directed constructions*, Transaction of AMS **309**/2, 1988, 811–829.

[88] Michalek, G. E., *Base three just touching covering systems*, Publ. Math. Debrecen **51**/3-4, 1997, 241–263.

[89] Michalek, G. E., *Base N just touching covering systems*, preprint, 1999.

[90] Müller, W., Thuswaldner, J. M., Tichy, R. F., *Fractal properties of number systems*, Periodica Math. Hung., to appear.

[91] Odlyzko, A.M., *Non-negative digit sets in positional number systems*, Proc. London Math. Soc. **37**, 1978, 213–229.

[92] Parry, W., *On the* $\beta$*-expansions of real numbers*, Acta Math. Hung. **11**, 1960, 401–406.

[93] Peitgen, H.-O., Jürgens, H., Saupe, D., *Chaos and fractals*, Springer, 1992.

[94] Penney, W., *A "binary" system for complex numbers*, J. ACM **12**, 1965, 247–248.

[95] Pethő, A., *On a polynomial transformation and its application to the construction of a public key cryptosystem*, Computational Number Theory, Proc., Walter de Gruyter Publ. Comp., 1991, 31–44.

[96] Pethő, A., *On the periodic expansion of algebraic numbers*, Annales Univ. Sci. Bud., Sect. Comp. **18**, 1999, 167–174.

[97] Ralston, A., *A first course in numerical analysis*, McGraw-Hill Inc., 1965.

[98] Robert, A., *A good basis for the computing with complex numbers*, Elemente der Mathematik **49**, 1994, 111-117.

[99] Safer, T., *Radix representation of algebraic numbers and finite automata*, Proc. STACS'98, 1999, 356–365.

[100] Safer, T., *Polygonal radix representation of complex numbers*, Theor. Comp. Sci. **210**, 1999, 159–171.

[101] Shallit, J., *A generalization of automatic sequences*, Theor. Comp. Sci. **61**, 1988, 1–16.

[102] Steidl, G., *On symmetric representation of gaussian integers*, BIT **29**, 1989, 563–571.

[103] Thuswaldner, J., *Fractal dimension of sets induced by bases of imaginary quadratic fields*, Math. Slovaca **48**, 1998, 365–371.

[104] Thuswaldner, J., *Fractals and number systems in real quadratic number fields*, Acta Math. Hung. **90**/(3), 2001, 253–269.

[105] Vince, A., *Radix representation and rep-tiling*, Proc. 24*th* Southeastern Intern. Conf. on Combinatorics, Graph Theory and Computing, Boca Raton, FL, 1993, Congr. Numer. **98**, 1993, 199–212.

[106] Vince, A., *Replicating tesselations* SIAM J. Discrete Math. **6**, 1993, 501–521.

[107] Vince, A., *Rep-tiling euclidean space*, Aequationes Math. **50**, 1995, 191–213.

[108] Vince, A., *Self-replicating tiles and their boundary*, Discrete Comp. Geom. **21**/3, 1999, 463–476.

# Index