On computation of attractors for invertible expanding linear operators in $\mathbb{Z}^k$

Attila Kovács*
Department of Computer Algebra
Eötvös Loránd University, Budapest, Hungary
attila@compalg.inf.elte.hu

Dedicated to Prof. Imre Kátai on the occasion of his 60th birthday

"In mathematics there exists only one truth." — *I. Kátai*

Abstract

It is well-known that if $M : \mathbb{R}^k \to \mathbb{R}^k$ is a linear operator with complex eigenvalues $\lambda_1, \ldots, \lambda_k$ satisfying $|\lambda_i| < 1$ $(i = 1, 2, \ldots, k)$ then $M^n \underline{v} \to \underline{0}$ for each $\underline{v} \in \mathbb{R}^k$. If $|\lambda_i| > 1$ $(i = 1, 2, \ldots, k)$ then $\|M^n \underline{v}\| \to \infty$ as $n \to \infty$ $(\underline{v} \neq \underline{0})$. In the latter case $\underline{0}$ is a repelling fixed point of M and iterates of all points except $\underline{0}$ recede from $\underline{0}$. In [9] Kátai gave a method providing attracting dynamics for the case of invertible expanding operators with integer components. This leads to the notion of number systems. In this paper we shall give an effective algorithm determining the attractors for a given invertible expanding linear operator of $\mathbb{R}^k$ mapping $\mathbb{Z}^k$ into $\mathbb{Z}^k$ and for a given finite set of appropriate integer vectors as digits.

1 Definitions and basic properties

Part of the definitions and notations can be found in the earlier works of I. Kátai. We shall summarize and extend them according to our purposes. Let M be an invertible linear operator of $\mathbb{R}^k$ mapping $\mathbb{Z}^k$ into $\mathbb{Z}^k$. Assume that M is expanding, i.e. the eigenvalues $\lambda_1, \ldots, \lambda_k$ of M satisfy

$$| \lambda_i | > 1 \ (i = 1, \ldots, k).$$

Let $\mathcal{L} = M\mathbb{Z}^k$. Then $\mathcal{L}$ is a subgroup (lattice) in $\mathbb{Z}^k$, the order of the factorgroup $\mathbb{Z}^k/M\mathbb{Z}^k$ is $t = | \det(M) |$. Let $t \geq 2$ and A_j $(j = 0, \ldots, t-1)$ denote the cosets of this group. If $\underline{z}, \underline{z}' \in \mathbb{Z}^k$ are in the same residue class then we will say that

*Supported by the FKFP–0144.

they are congruent modulo M and we will denote this by $\underline{z} \equiv \underline{z}' \pmod{M}$. For each j choose an arbitrary element $\underline{a}_j$ from A_j and let

$$\mathcal{A} := \{\underline{a}_0, \underline{a}_1, \ldots, \underline{a}_{t-1}\}.$$

For arbitrary subsets X, Y of $\mathbb{Z}^k$ let

$$\Phi_Y(X) := \{M^{-1}(\underline{z} - \underline{d}) \in \mathbb{Z}^k : \underline{z} \in X,\ \underline{d} \in Y,\ \underline{z} \equiv \underline{d} \pmod{M}\}. \tag{1}$$

If Y is a complete residue system modulo M then for the one-element subset $\{\underline{\mathrm{x}}\}$ the set $\Phi_Y(\{\underline{x}\})$ has only one element, say $\underline{y}$. In this case we will write $\Phi_Y(\underline{x}) = \underline{y}$. If it is clear which Y is used, we shall only write $\Phi(\underline{x}) = \underline{y}$. In the following let $Y := \mathcal{A}$ and let Φ^l denote the l-fold iterate of Φ, $\Phi^0(\underline{z}_0) = \underline{z}_0$.

Definition: The sequence of integer vectors $\Phi^j(\underline{z}_0) = \underline{z}_j$ $(j = 0, 1, 2, \ldots)$ is called the path of the dynamical system generated by Φ. It is also called the orbit of $\underline{z}_0$ generated by Φ.

Since the spectral radius $\rho(M^{-1}) < 1$ therefore there exists a norm on $\mathbb{R}^k$ such that for the corresponding operator norm

$$\|M^{-1}\| = \sup_{\|\underline{x}\| \le 1} \|M^{-1}\underline{x}\| \tag{2}$$

the inequality $\|M^{-1}\| < 1$ holds (see [6]). Throughout this article $\|\,.\,\|$ denotes this vector and the appropriate operator norm. Let furthermore

$$K := \max_{\underline{b} \in \mathcal{A}} \|\underline{b}\|, \quad r := \|M^{-1}\|, \quad L := \frac{Kr}{1-r}. \tag{3}$$

In virtue of (1) and (3) we get that

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

Hence we obtain the following

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

Since the inequality $\|\underline{x}\| \le L$ holds only for finitely many integer vectors $\underline{x}$ therefore the path $\underline{z}, \Phi(\underline{z}), \Phi^2(\underline{z}), \ldots$ is ultimately periodic for all $\underline{z} \in \mathbb{Z}^k$.

Definition: $\underline{p} \in \mathbb{Z}^k$ is called periodic if there exist a $j \in \mathbb{N}$ such that $\Phi^j(\underline{p}) = \underline{p}$. The smallest such j is the length of the period of $\underline{p}$ generated by Φ.

Let $\mathcal{P}$ denote the set of all periodic elements. Let $\underline{p} \in \mathcal{P}$ be of period length l. The set of the periodic elements $\{\Phi(\underline{p}), \ldots, \Phi^l(\underline{p})\}$ will be denoted by $\mathcal{C}(\underline{p})$.

Definition: Suppose that $\underline{p} \in \mathcal{P}$. Then the basin of attraction of $\underline{p}$ consists of all $\underline{z} \in \mathbb{Z}^k$ for which there exists a $j \in \mathbb{N}$ such that $\Phi^j(\underline{z}) = \underline{p}$ and is denoted by $\mathcal{B}(\underline{p})$. Let $X \subseteq \mathcal{P}$. In a similar way, $\mathcal{B}(X)$ denotes all the $\underline{z} \in \mathbb{Z}^k$ for which there exists a $j \in \mathbb{N}$ and $\underline{p}' \in X$ such that $\Phi^j(\underline{z}) = \underline{p}'$.

The following assertions are clearly true:

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

$$\underline{p} = \underline{a}_0 + M\underline{a}_1 + \ldots + M^{l-1}\underline{a}_{l-1} + M^l\underline{p}, \quad \underline{a}_j \in \mathcal{A}, \tag{4}$$

- if $\underline{p}_1, \underline{p}_2 \in \mathcal{P}$ then either $\mathcal{C}(\underline{p}_1) = \mathcal{C}(\underline{p}_2)$ or $\mathcal{C}(\underline{p}_1) \cap \mathcal{C}(\underline{p}_2) = \emptyset$,
- if $\underline{p}_1, \underline{p}_2 \in \mathcal{P}$, $\underline{p}_1 \neq \underline{p}_2$ and $\mathcal{C}(\underline{p}_1) = \mathcal{C}(\underline{p}_2)$ then their length of period are equal,
- $\mathcal{B}(\mathcal{P}) = \mathbb{Z}^k$,
- if $\underline{p}_1, \underline{p}_2 \in \mathcal{P}$ then $\mathcal{B}(\underline{p}_1) = \mathcal{B}(\underline{p}_2)$ if and only if $\mathcal{C}(\underline{p}_1) = \mathcal{C}(\underline{p}_2)$,
- if $\underline{p}_1, \underline{p}_2 \in \mathcal{P}$, $\mathcal{C}(\underline{p}_1) \neq \mathcal{C}(\underline{p}_2)$ then $\mathcal{B}(\underline{p}_1) \cap \mathcal{B}(\underline{p}_2) = \emptyset$.

Definition: Let $\mathcal{G}(\mathcal{P})$ be the directed graph defined on the set $\mathcal{P}$ by drawing an edge from $\underline{p} \in \mathcal{P}$ to $\Phi(\underline{p})$. 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 of $\mathbb{Z}^k$ generated by Φ.

2 Number systems in $\mathbb{Z}^k$

Definition: Let M be an invertible expanding linear operator of $\mathbb{R}^k$ mapping $\mathbb{Z}^k$ into itself and $\mathcal{A}$ be a complete residue system modulo M. The pair $(M, \mathcal{A})$, $\mathcal{A} = \{\underline{0}, \underline{a}_1, \ldots, \underline{a}_{t-1}\}$, is called a number system in $\mathbb{Z}^k$ if for each $\underline{z} \in \mathbb{Z}^k$ there exist an $m \in \mathbb{N}_0$ and $\underline{a}_j \in \mathcal{A}$ $(j = 0, 1, \ldots, m)$ such that

$$\underline{z} = \sum_{j=0}^{m} M^j \underline{a}_j.$$

The uniqueness of the expansion follows from the assumption that any two elements of $\mathcal{A}$ are incongruent modulo M.

Remark: The condition $\underline{0} \in \mathcal{A}$ suites the traditional number system concept well: every integer element has a finite and unique representation. On the other hand all the following theorems remain valid if we substitute the vector $\underline{0}$ with a congruent element modulo M.

The next two theorems are simple generalizations of the results achieved in algebraic number fields. The proofs can be found among others in [8, 1, 7, 10].

Theorem 1: With the notations above the pair $(M, \mathcal{A})$ is a number system if and only if for each $\underline{z} \in \mathbb{Z}^k$ there is an $n \in \mathbb{N}_0$ such that $\Phi^n(\underline{z}) = \underline{0}$.

Proof: The condition $\Phi(\underline{z}) = \underline{0}$ is equivalent with $\underline{z} \equiv \underline{a}_0$ for some $\underline{a}_0 \in \mathcal{A}$. By induction $\Phi^n(\underline{z}) = \underline{0}$ if and only if $\underline{z}$ can be written in the form

$$\underline{z} = \underline{a}_0 + M\underline{a}_1 + \ldots + M^{n-1}\underline{a}_{n-1}$$

with some $\underline{a}_0, \underline{a}_1, \ldots, \underline{a}_{n-1} \in \mathcal{A}$. □

Corollary: The pair $(M, \mathcal{A})$ is a number system in $\mathbb{Z}^k$ if and only if $\mathcal{P} = \{\underline{0}\}$.

Since $\mathcal{P}$ is a finite set, this theorem can be used to decide for a given system $(M, \mathcal{A})$ whether it is a number system or not. Before we continue in this way consider the set of "fractions" in the system $(M, \mathcal{A})$:

$$H := \left\{ \sum_{n=1}^{\infty} M^{-n}\underline{d}_n : \underline{d}_n \in \mathcal{A} \right\} \subseteq \mathbb{R}^k.$$

It is well-known that the set H is compact in the metric space $\mathbb{R}^k$.

Theorem 2: Let $I(-H) := \mathbb{Z}^k \cap (-H)$. Then for each $\underline{z} \in \mathbb{Z}^k$ there is an $m \in \mathbb{N}_0$ such that $\Phi^m(\underline{z}) \in I(-H)$.

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

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

where $N_\varepsilon(-H)$ denotes the open ε-neighbourhood of $-H$. Let we choose an arbitrary $\underline{z} \in \mathbb{Z}^k$. Then we get that

$$\underline{z}_m = \Phi^m(\underline{z}) = M^{-m}\underline{z} - (M^{-1}\underline{d}_1 + M^{-2}\underline{d}_2 + \ldots + M^{-m}\underline{d}_m)$$

for the corresponding sequence $\underline{d}_1, \underline{d}_2, \ldots, \underline{d}_m \in \mathcal{A}$. If m is large enough then the first term of the right hand side has norm less then ε. Hence $\underline{z}_m \in \mathbb{Z}^k \cap (-H)$. □

Corollary :

a) For each $\underline{z} \in \mathbb{Z}^k$ the orbit of $\underline{z}$ must "pass through" the set $I(-H)$.

b) If for each $\underline{z} \in I(-H)$ there is an $m \in \mathbb{N}_0$ such that $\Phi^m(\underline{z}) = \underline{0}$ then $(M, \mathcal{A})$ is a number system.

The corollary suggests that in order to determine the attractors of the system $(M, \mathcal{A})$ it would be enough to find the integer points in $-H$, or which is computational equivalent, in H. Then we have only to apply Φ for these vectors and watching the "cycles" to be formed.

The straightforward way to compute the set $I(H) := \mathbb{Z}^k \cap H$ could be the following. It is obvious that

$$H = \bigcup_{\underline{b} \in \mathcal{A}} (\underline{b} + M^{-1}H).$$

If we could find a set T_0, $H \subseteq T_0$, for which the integers of the set $M^{-1}T_0$ can be computed easily then we would be ready, because in this case $H \subseteq T_1 := \bigcup_{\underline{b}\in\mathcal{A}}(\underline{b} + M^{-1}T_0)$ and only the convex hull of the integer points of 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. Many authors investigated the structure and properties of the set H in special cases in the last decades, see [1, 2, 3, 4, 5, 9, 9, 7, 10]. Our next aim is to determine a set T, $H \subseteq T$, for which the set of integers belonging to T can be computed simply and it contains possibly a small number of integer vectors.

3 Determining the set T

We shall consider two approaches. One of them uses coverings of the set H while the other one will be given by effectively computing the operator norm defined in (2).

3.1 Using a covering of the set H

Let $\underline{x} = (x_1, x_2, \ldots, x_k)^T \in \mathbb{R}^k$ and $\|x\|_\infty = \max_{1\le i\le k} \mid x_i \mid$. Let us denote by $\|.\|_\infty$ the corresponding operator norm. If M is an invertible expanding linear operator of $\mathbb{R}^k$ mapping $\mathbb{Z}^k$ into $\mathbb{Z}^k$ then there exists a smallest $c_0 \in \mathbb{N}$ such that for every $c \ge c_0,\ c \in \mathbb{N}$ the inequality $\|M^{-c}\|_\infty < 1$ holds. Let $C \ge c_0,\ C \in \mathbb{N}$ be fixed. Then

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

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

$$\gamma := \frac{1}{1 - \|M^{-C}\|_\infty} \ge \|(I - M^{-C})^{-1}\|_\infty. \tag{5}$$

Here I denotes the k-dimensional identity matrix. Using the notations introduced in the first section let

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

and let

$$\xi_m^{(j)} := \max_{\underline{b}\in\mathcal{A}} |c_m^{(j)}(\underline{b})|,\ (m = 1, \ldots, k),$$

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

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

Obviously, $M^{-j}\underline{b} \in I_j$ for each $\underline{b} \in \mathcal{A}$. Let

$$\mathcal{W} := \left\{ \underline{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_k \end{bmatrix}, |y_m| \leq \sum_{j=1}^{C} \xi_m^{(j)}, 1 \leq m \leq k \right\}. \tag{6}$$

It is clear that

$$\sum_{j=1}^{C} M^{-j}\underline{b}_j \in \mathcal{W}$$

for an arbitrary sequence of vectors $\underline{b}_j \in \mathcal{A}$. Hence,

$$H \subseteq \mathcal{W} + M^{-C}\mathcal{W} + M^{-2C}\mathcal{W} + \dots. \tag{7}$$

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

$$\begin{bmatrix} t_1 \\ \vdots \\ t_k \end{bmatrix}, \ -\alpha_m \leq t_m \leq \alpha_m, \alpha_m = \lceil \gamma \sum_{j=1}^{C} \xi_m^{(j)} \rceil, \ 1 \leq m \leq k. \tag{8}$$

Then by (5), (6) and (7) we get that $H \subseteq T'$ and the integer vectors in T' can be computed efficiently.

Remarks:

1. To determine the vertices of T' for a fixed C our method requires
   - a matrix inverse computation,
   - $C - c_0$ times matrix-matrix multiplications and
   - $(C - c_0 + 1)tk^2$ multiplications over $\mathbb{R}$.

   The "good choice" for the constant C in (5) depends on k, t and on the matrix M. A simple method could be choosing the constant C in such a way that starting with $C \leftarrow c$ increase C while $\|M^{-C}\|_\infty$ is less than or equal to a fixed constant. (This can be done because of the continuity of the norm.) Another approach may require much more arithmetical operations: start with $C \leftarrow c$ and increment C until the volume of T' does not change.

2. Even if $M^{-n}\underline{v} \to \underline{0}$ $(n \to \infty)$ for any $\underline{v} \in \mathbb{R}^k$ 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 $k = 5$, $M = \text{tridiag}(0, -2, -2^{10})$. (In the following 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.

3. Because of we are interested in only the integers in T' in equation (8) the floor function can also be applied. Then the integers in T' still covers the integers in H.

3.2 Using the operator norm

Let $\underline{x} \in H$ be an arbitrary vector. Then

$$\|\underline{x}\| = \|\sum_{j=1}^{\infty} M^{-j} \underline{b}_j\| \tag{9}$$

for any well-defined vector norm in $\mathbb{R}^k$, where $\underline{b}_j \in \mathcal{A}$ $(j = 1, 2, \ldots)$. For the estimation of the right hand side we will consider a vector norm introduced in (2).

Let M be an invertible expanding 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

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

so the operator norm induced by the vector norm $\|S\underline{x}\|_\infty$. Let $J = TM^{-1}T^{-1} = \mathrm{diag}(\Lambda_j)$ be the Jordan 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 $D_j := \mathrm{diag}_{1\le i\le m}(\mu_j^{m-i})$ 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 on let us denote $\|\cdot\| := \|\cdot\|_*$ as we used it in the earlier sections. 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

$$\|(I - M^{-1})^{-1}\| \le \frac{1}{1 - \|M^{-1}\|}. \tag{10}$$

By using (3), (9) and (10) we get that

$$\|S\underline{x}\|_\infty = \|\underline{x}\| = \|\sum_{j=1}^{\infty} M^{-j}\underline{b}_j\| \le \frac{Kr}{1-r} = L. \tag{11}$$

Now we are looking for those $\underline{x} \in \mathbb{Z}^k$ for which (11) is satisfied. If $\|\underline{x}\|_\infty \le L/\|S\|_\infty$ then (11) is clearly true. Let $\underline{y} := S\underline{x}$. Then $S^{-1}\underline{y} = \underline{x}$, hence

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

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

$$\beta_i := \lceil L\|S^{-1}\|_\infty \rceil. \tag{12}$$

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 $\|M^{-1}\| \le \rho(M^{-1}) + \varepsilon$. This is a known result (see [6]).

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 μ_j can be chosen arbitrary according to their definition but in the computer implementations the floating point overflows (e.g. μ_j-s are too small) must be avoided. The best solution would be to optimize μ_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, in (12) the floor function can also be applied. Then the integers in T'' covers the integers in H.

4. The computation of the vector norm $\|.\|$ for an arbitrary vector $\underline{x} \in \mathbb{R}^k$ requires a matrix multiplication with S and during these operations a maximum searching.

Constituting the intersection of T' and T'' we proved the following theorem:

Theorem 3: Let the set of integer points I_T be defined as follows:

$$I_T := \left\{ \begin{bmatrix} t_1 \\ \vdots \\ t_k \end{bmatrix} \in \mathbb{Z}^k, \ -\kappa_m \le t_m \le \kappa_m, \text{ where} \right.$$

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

Then $I(H) \subseteq I_T, \ I(-H) \subseteq I_T$ and the set I_T can be rapidly computed.

4 Computation of the function Φ

For calculation of the function Φ one needs a fast procedure to determine for an arbitrary $\underline{z} \in \mathbb{Z}^k$ the corresponding congruent element $\underline{d} \in \mathcal{A}$ modulo M. Our first method is a straightforward generalization of the method used for the case of Gaussian integers in [10].

4.1 The adjoint method

Applying the notations already adopted let $\underline{z}$ be an arbitrary element of $\mathbb{Z}^k$ and let $\mathcal{A} = \{\underline{a}_0, \underline{a}_1, \dots, \underline{a}_{t-1}\}$ be a complete residue system modulo M. If

$$\underline{z} \equiv \underline{a}_j \text{ modulo } M$$

then

$$M^*\underline{z} \equiv M^*\underline{a}_j \text{ modulo } tI,$$

where M^* denotes the adjoint of M, I the identity matrix and $t = |\det(M)|$. Here the notion "adjoint of the operator M" denotes the integer matrix, for which the elements are the adjoints of the appropriate subdeterminants. Let

$$\bar{\mathcal{A}} := M^*\mathcal{A} \pmod{tI} = \{\underline{b}_0, \underline{b}_1, \dots, \underline{b}_{t-1}\}, \tag{13}$$

where

$$\underline{b}_j = M^*\underline{a}_j \pmod{tI} = \begin{bmatrix} b_1^{(j)} \\ \vdots \\ b_k^{(j)} \end{bmatrix} \in \mathbb{Z}^k,\ 0 \le b_i^{(j)} < t,\ (i = 1, \dots, k). \tag{14}$$

Due to the complete residue system property of $\mathcal{A}$ for every $\underline{z} \in \mathbb{Z}^k$ there exists a unique $\underline{b}_j \in \bar{\mathcal{A}}$ such that $\underline{b}_j = M^*\underline{z} \pmod{tI}$. Then from (13) and (14) it follows that $\underline{z} \equiv \underline{a}_j$ modulo M.

In order to obtain for an arbitrary $\underline{z} \in \mathbb{Z}^k$ the congruent element in $\mathcal{A}$ 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 positive. Suppose that there exists an $i \in \mathbb{N}$, $1 \le i \le k$ for which $b_i^{(j)} (j = 0, 1, \dots, t-1)$ in (14) are all different. Then the inner product of an arbitrary $\underline{z} \in \mathbb{Z}^k$ by the i-th row of M^* modulo t uniquely determines the index j for which $\underline{z} \equiv \underline{a}_j$ 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 section 6. But what can be made when such an i does not exists? Then one has to investigate further the set $\bar{\mathcal{A}}$ and to figure out a strategy to minimize the number of multiplications to obtain for an arbitrary $\underline{z} \in \mathbb{Z}^k$ the appropriate $\underline{b}_j \in \bar{\mathcal{A}}$ for which $\underline{b}_j = M^*\underline{z}$ modulo tI. Beside the optimization the strategy needs GCD 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 so called Smith normal form (see [7]).

4.2 Using Smith normal form

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

$$UMV = D$$

has diagonal form in the standard basis with positive integer elements $d_1, \ldots, d_k$ in the diagonal such that $d_i \mid d_{i+1}$ for $i = 1, 2, \ldots, k-1$ and

$$\prod_{i=1}^{k} d_i = |\det(M)|.$$

The Smith normal form can be obtained by doing elementary row and column operations of M. We remark that U and V have determinants ± 1 and they are also invertible having integer components.

Theorem 4: For an invertible M with the notations above let for $\underline{z}_1, \underline{z}_2 \in \mathbb{Z}^k$ the numbers $u_1, u_2, \ldots, u_k$ and $\hat{u}_1, \hat{u}_2, \ldots, \hat{u}_k$ denote the coordinates of $U\underline{z}_1$ and $U\underline{z}_2$ respectively. Then

$$\underline{z}_1 \equiv \underline{z}_2 \text{ modulo } M$$

if and only if

$$u_i \equiv \hat{u}_i \text{ modulo } d_i$$

for all $i = 1, 2, \ldots, k$.

Proof: $\underline{z}_1 \equiv \underline{z}_2$ modulo M if and only if

$$M^{-1}(\underline{z}_1 - \underline{z}_2) \in \mathbb{Z}^k.$$

This is equivalent with the condition

$$V^{-1}M^{-1}(\underline{z}_1 - \underline{z}_2) \in \mathbb{Z}^k.$$

But

$$V^{-1}M^{-1} = D^{-1}U,$$

hence the equations

$$u_i \equiv \hat{u}_i \text{ modulo } d_i$$

must be satisfied for all $i = 1, 2, \ldots, k$. □

From computational point of view to obtain the congruent element in $\mathcal{A}$ modulo M for the vector $\underline{z} \in \mathbb{Z}^k$ in a general case at the first sight there is no gain: in the first step one has to multiply $\underline{z}$ by the integer matrix U (mod D) instead of M^* (mod tI). But if there exists a positive integer s for which $d_i = 1$,

$i = 1, \ldots, s, s < k$ then $u_i \equiv 0 \pmod{d_i}$ for all $i = 1, \ldots, s$ and for all $\underline{z} \in \mathbb{Z}^k$, hence enough to perform only k integer multiplications modulo d_j, for each $j = s+1, \ldots, k$. Let

$$\hat{\mathcal{A}} := U\mathcal{A} \pmod{D} = \{\hat{\underline{b}}_0, \hat{\underline{b}}_1, \ldots, \hat{\underline{b}}_{t-1}\}, \tag{15}$$

where

$$\hat{\underline{b}}_j = U\underline{a}_j \pmod{D} = \begin{bmatrix} b_1^{(j)} \\ \vdots \\ b_k^{(j)} \end{bmatrix} \in \mathbb{Z}^k,\ 0 \le b_i^{(j)} < d_i,\ (i = 1, \ldots, k). \tag{16}$$

We get that for every $\underline{z} \in \mathbb{Z}^k$ there exists a unique $\hat{\underline{b}}_j \in \hat{\mathcal{A}}$ such that $\hat{\underline{b}}_j = U\underline{z} \pmod{D}$. From (15) and (16) we have that $\underline{z} \equiv \underline{a}_j$ modulo M.

4.3 Computer implementation

In computer implementations once the computation $M^*\underline{z}$ modulo tI or $U\underline{z}$ modulo D was performed for the vector $\underline{z} \in \mathbb{Z}^k$ the result must be looked up in the table $T(\bar{\mathcal{A}})$ or in $T(\hat{\mathcal{A}})$, respectively, obtaining the index j for which $\underline{a}_j \equiv \underline{z}$ modulo M, $\underline{a}_j \in \mathcal{A}$. 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.

Theorem 5: Using the notations above let us define the function h by

$$h(\underline{z}) = \sum_{i=s+1}^{k} (u_i \bmod d_i) \prod_{j=s+1}^{i-1} d_j.$$

Then h is an integer valued function with values $0, \ldots, t-1$, and $h(\underline{z}_1) = h(\underline{z}_2)$ if and only if $\underline{z}_1 \equiv \underline{z}_2$ modulo M.

The proof is easy, it is left to the reader. (See also [7].)

Remark: The set $\bar{\mathcal{A}}$ can be generated only from $\mathcal{A}$ but the set $\hat{\mathcal{A}}$ can be produced also directly from D. A complete residue system $\pmod{M}$ can be generated from $\hat{\mathcal{A}}$ $(\bar{\mathcal{A}})$ by multiplying the elements with U^{-1} (M), respectively.

We summarize our results for the computation of the function Φ:

- For a given vector $\underline{z} \in \mathbb{Z}^k$ computing $M^*\underline{z} \pmod{tI}$ needs k^2 integer multiplications over $\mathbb{Z}_t$, computing $U\underline{z} \pmod{D}$ requires k integer multiplications over $\mathbb{Z}_{d_j}$ for each $j = s+1, \ldots, k$, where s depends on the matrix M.
- Looking up the congruent element $\underline{a}_j$ in the table $T(\mathcal{A})$ either a searching has to be performed in $T(\bar{\mathcal{A}})$ or in $T(\hat{\mathcal{A}})$ to obtain the index j or a hashing has to be done.

- To perform the function Φ 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 $\mathbb{R}$.

5 Algorithm for computing the attractors

In the previous section we presented a fast procedure to compute the function Φ. But how many times should we apply it for an arbitrary vector $\underline{z}$ until the orbit of $\underline{z}$ reaches the attractor?

5.1 The length of the standard expansion

Let $\underline{z} \in \mathbb{Z}^k$ be an arbitrary vector. If $\underline{z}_0 := \underline{z} \notin \mathcal{P}$ then there is a unique $l \in \mathbb{N}$ and $\underline{a}_0, \underline{a}_1, \ldots, \underline{a}_{l-1} \in \mathcal{A}$ such that

$$\underline{z}_j = \underline{a}_j + M\underline{z}_{j+1} \ (j = 0, \ldots, l-1), \ \underline{z}_l \in \mathcal{P}$$

and none of $\underline{z}_0, \underline{z}_1, \ldots, \underline{z}_{l-1}$ do belong to $\mathcal{P}$. Then let the expansion of $\underline{z}$ be denoted by

$$(\underline{a}_0, \underline{a}_1, \ldots, \underline{a}_{l-1} \mid \underline{p}), \ (\underline{p} = \underline{z}_l). \tag{17}$$

If such an expansion is given then $\underline{z}$ can be computed by

$$\underline{z} = \underline{a}_0 + M\underline{a}_1 + \ldots + M^{l-1}\underline{a}_{l-1} + M^l\underline{p}. \tag{18}$$

If $\underline{z} \in \mathcal{P}$ then its expansion in $(M, \mathcal{A})$ will be denoted by $(* \mid \underline{z})$.

Definition: We shall say that (17) is the standard expansion of the vector $\underline{z}$ given by (18) and l is the length of the standard expansion.

For an arbitrary sequence of vectors $\underline{a}_0, \underline{a}_1, \ldots, \underline{a}_{l-1} \in \mathcal{A}$ and $\underline{p} \in \mathcal{P}$ the expression $(\underline{a}_0, \underline{a}_1, \ldots, \underline{a}_{l-1} \mid \underline{p})$ means the vector $\underline{z}$ given by $\underline{z} = \sum_{j=0}^{l-1} M^j \underline{a}_j + M^l \underline{p}$. This expansion is the standard expansion of the vector $\underline{z}$ if and only if $\Phi^{l-1}\underline{z} = \underline{a}_{l-1} + M\underline{p} \notin \mathcal{P}$. Observe that if $\underline{p} \in \mathcal{P}$ than all $\underline{z} \in \mathcal{B}(\underline{p}) \setminus \mathcal{C}(\underline{p})$ have a standard expansion $(\underline{a}_0, \underline{a}_1, \ldots, \underline{a}_l \mid \hat{\underline{p}})$ for some $\underline{a}_i \in \mathcal{A}$ $(i = 0, \ldots, l)$, $l \in \mathbb{N}$ and $\hat{\underline{p}} \in \mathcal{C}(\underline{p})$.

Back to the original question: what can be stated about the length of the standard expansion for an arbitrary $\underline{z} \in \mathbb{Z}^k$? Let $l(\underline{z})$ be the smallest integer $l \geq 0$ for which $\Phi^l(\underline{z}) \in \mathcal{P}$.

Theorem 6: There is a constant $c = c(M, \mathcal{A})$ for which

$$\left| l(\underline{z}) - \frac{\log \|\underline{z}\|}{\log \|M\|} \right| < c,$$

where $\underline{z} \in \mathbb{Z}^k \setminus \{\underline{0}\}$.

The proof is a straightforward generalization of the one-dimensional case given by Kátai in [9].

5.2 The algorithm

In this subsection an algorithm will be presented for determining the attractors denoted by $\mathcal{G}(\mathcal{P})$ for a given invertible expanding linear operator M of $\mathbb{R}^k$ mapping $\mathbb{Z}^k$ into $\mathbb{Z}^k$ and for a given set $\mathcal{A}$ of digits.

For arbitrary finite subsets X, Y of $\mathbb{Z}^k$ let

$$\Psi_Y(X) := \{M\underline{z} + \underline{d} \in \mathbb{Z}^k : \underline{z} \in X, \underline{d} \in Y\}.$$

In the following let $Y := \mathcal{A}$ and for simplicity we shall write $\Psi(X)$. Observe, that in certain sense Ψ — which acts as a left shift operator — is an inverse of Φ, which acts as a right shift operator. Let Ψ^l denote the l-fold iterate of Ψ, $\Psi^0(X) = X$. Then the following assertions are obviously true:

- $X \subset \Psi(X)$ if and only if $\mathcal{P} \subseteq X$,
- if for all $\underline{z} \in X$, $\underline{z} \notin \mathcal{P}$ then $X \cap \Psi^j(X) = \emptyset$ for all $j \in \mathbb{N}$,
- if $\underline{p} \in \mathcal{P}$, the length of the period of $\underline{p}$ is q, $X \cap \mathcal{C}(\underline{p}) = \{\underline{p}\}$ and $\underline{p} \in \Psi^j(X)$ for some $j \in \mathbb{N}$ then $q \mid j$.

Remark: The computation of Ψ requires k^2 integer multiplications and tk integer additions.

Algorithm for determining the attractors for a given system $(M, \mathcal{A})$:

Recall that the points of the attractor are in the set I_T, which can be computed efficiently.

1. Suppose that $\underline{p} \in \mathcal{P} \subseteq I_T$ is known. Then (iterating the function Φ) $\mathcal{C}(\underline{p})$ can be easily computed.
2. The operator Ψ can be used to generate the (finite) set $(\mathcal{B}(\underline{p}) \cap I_T) \setminus \mathcal{C}(\underline{p})$.
3. Omit these points from I_T.
4. Find a new periodic element and repeat the process until integer points in $I_T \setminus \cup\mathcal{C}(\underline{p}_i)$ remain.

Generating the orbit of an arbitrary $\underline{z} \in I_T$ one can easily get a new periodic element $\underline{p}$ and even the set $\mathcal{C}(\underline{p})$. I_T is finite so the algorithm terminates.

We have seen that the function Φ can be applied to generate the periodic elements. We will show that to determine the periodic elements with small periods there is a faster way to compute.

Theorem 7: Let $l \in \mathbb{N}$ be fixed. Suppose that the matrix $(I - M^l)$ is regular. Then $\underline{p} \in \mathcal{P}$ if and only if

$$(I - M^l)^{-1}(\underline{a}_0 + M\underline{a}_1 + \ldots + M^{l-1}\underline{a}_{l-1}) \in \mathbb{Z}^k,$$

where $\underline{a}_i \in \mathcal{A}$, $(i = 0, 1, \ldots, l-1)$.

Proof: It was stated in (4) that $\underline{p} \in \mathcal{P}$ if and only if there is an $l > 0$ such that

$$\underline{p} = \underline{a}_0 + M\underline{a}_1 + \ldots + M^{l-1}\underline{a}_{l-1} + M^l\underline{p}, \quad \underline{a}_j \in \mathcal{A}.$$

This is equivalent with

$$\underline{p} = (I - M^l)^{-1}(\underline{a}_0 + M\underline{a}_1 + \ldots + M^{l-1}\underline{a}_{l-1}) \in \mathbb{Z}^k, \tag{19}$$

which was to be proved. $\square$

For a given $l \in \mathbb{N}$ this theorem can be used to generate *all* the periodic elements with period length l_1, where $l_1 \mid l$ ($l_1 \in \mathbb{N}$). Putting the vectors $\underline{a}_i \in \mathcal{A}$ in all possible ways into the equation (19) one has only to check whether the right hand side is an integer vector or not. The method requires a matrix inversion computation, $k^2(l-1)(t+1)$ integer and $t^l k^2$ floating point multiplications.

Remarks:

1. From computational aspect this procedure is preferable only for small l and t.

2. The method requires exact floating point computations, e.g. use of computer algebra.

Recall that the attractor of a number system may contain only the loop $\underline{0} \to \underline{0}$. To answer the question whether a given system $(M, \mathcal{A})$ is a number system or not — applying Theorem 7 — first check the number of small periods (in most cases only the loops). Next, — if it is necessary — the operator Ψ can be iterated for the one element initial set $X = \{\underline{0}\}$. If all the points in I_T can be produced in this way then it is a number system otherwise it isn't. The termination follows from Lemma 1 and from the fact that I_T is finite.

6 Canonical number systems in $\mathbb{Z}^k$

Definition: A set of vectors $\mathcal{C}_{M,j} \subset \mathbb{Z}^k$ is called j-canonical with respect to the matrix M ($1 \leq j \leq k$) if all the elements have the form $\nu\underline{e}_j$, where $\underline{e}_j$ denotes the j-th unit vector, $\nu = 0, 1, \ldots, t-1$ and $t = |\det(M)|$. If the set $\mathcal{C}_{M,j}$ forms a complete residue system — CRS for brevity — modulo M then we will denote it by $\mathcal{A}_j$. If there exists a j for which $(M, \mathcal{A}_j)$ is a number system then it is called j-canonical number system.

Canonical number systems play a significant role in mathematics and computer science. In the following we shall analyze the existence of j-canonical complete residue systems.

Theorem 8: Let M be an invertible expanding linear operator of $\mathbb{R}^k$ mapping $\mathbb{Z}^k$ into itself and let $\underline{c} = [c_1, c_2, \ldots, c_k]^T \in \mathbb{Z}^k$ be the j-th column of the matrix M^* (adjoint of M). Let $\mathrm{GCD}(c_l, t) := \delta_l$ ($l = 1, \ldots, k$), 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

$$\bar{\mathcal{A}}_c = \left\{ \nu \underline{c} \bmod t = \begin{bmatrix} \nu c_1 \bmod t \\ \vdots \\ \nu c_k \bmod t \end{bmatrix}, \ \nu = 0, 1, \ldots, t-1 \right\}$$

has exactly t elements.
3. $\mathrm{LCM}(\tau_1, \ldots, \tau_k) = t$.

Proof:
(1) $\Leftrightarrow$ (2).
The proof immediately follows from the construction of $\bar{\mathcal{A}}$ (see section 4.1).
(1) $\Leftrightarrow$ (3).
Due to the CRS property of the set $\mathcal{A}_j$ all their elements are incongruent modulo M and the set $\mathcal{A}_j$ 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 $\underline{\eta} = [\eta_1, \eta_2, \ldots, \eta_k]^T \in \mathbb{Z}^k$. Hence it is enough to examine the solvability of the system of equations

$$\begin{aligned} hc_1 &= t\eta_1 \\ &\vdots \\ hc_k &= t\eta_k. \end{aligned} \tag{20}$$

Case 1. There exists a c_l $(1 \le l \le k)$ such that $\mathrm{GCD}(c_l, t) = 1$. In this case from the equation $hc_l = t\eta_l$ it follows that $t \mid h$. Hence the system of equations (20) has no integer solution.
Case 2. Suppose that $\mathrm{GCD}(c_l, t) = \delta_l > 1$ for all $l = 1, 2, \ldots, k$. Let $c_l^* = c_l/\delta_l$. Then $hc_l^* = \tau_l\eta_l$ $(l = 1, \ldots, k)$. Since $\mathrm{GCD}(c_l^*, \tau_l) = 1$, therefore $\tau_l \mid h$ for all $l = 1, \ldots, k$. It means that $\mathrm{LCM}(\tau_1, \tau_2, \ldots, \tau_k) \mid h$. Hence the system of equations (20) has no solution if and only if $\mathrm{LCM}(\tau_1, \tau_2, \ldots, \tau_k) \ge t$. On the other hand $\mathrm{LCM}(\tau_1, \ldots, \tau_k) \mid t$. Therefore $\mathrm{LCM}(\tau_1, \ldots, \tau_k) = t$. (If $\tau_l = t$ for some l then $\mathrm{GCD}(c_l, t) = 1$.) We have that there exists j-canonical CRS modulo M if and only if $\mathrm{LCM}(\tau_1, \ldots, \tau_k) = t$. □

Remarks:

1. If there exists a $c_i \in \mathbb{Z} \setminus 0$ in the j-th column of the matrix M^* for which $\mathrm{GCD}(c_i, t) = 1$ modulo t then there is a j-canonical complete residue system modulo M. Theorem 8 shows that the converse of this statement is not always true.

2. If t is prime then always exist j-canonical CRS for all $1 \le j \le k$.

Lemma 2: Using the notations above suppose that for a given M there exists a j-canonical CRS. Then there is an $i \in \mathbb{N},\ 1 \le i \le k$ for which $\mathrm{GCD}(c_i, t) = 1$ 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 the lemma then it is enough to perform only k multiplications modulo t to determine for an arbitrary $\underline{z} \in \mathbb{Z}^k$ the element $\underline{b} = (M^*\underline{z} \text{ modulo } tI) \in \bar{\mathcal{A}}$ (see section 4).

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

7 Examples

First we have to call the reader's attention to the connection between the invertible expanding linear operators and the ring of integers of a given algebraic number field.

Let Θ be an algebraic integer with minimal polynomial $f(x) = x^k + c_{k-1}x^{k-1} + \ldots + c_1 x + c_0$. Let us denote the conjugates of Θ over $\mathbb{Q}$ by $\Theta_1, \Theta_2, \ldots, \Theta_k$. Assume that $|\Theta_i| > 1$ $(i = 1, \ldots, k)$. Let $\mathbb{Z}[\Theta]$ be the set of elements of form $u_0 + u_1\Theta + \ldots + u_{k-1}\Theta^{k-1}$ $(u_j \in \mathbb{Z})$. Then $\mathbb{Z}[\Theta]$ is a ring. For the addition it is isomorphic with the additive group $\mathbb{Z}^k$. Let $\alpha \in \mathbb{Z}[\Theta]$. The map $\alpha \to \Theta\alpha$ can be formulated as a linear transformation, which has a simple form in the basis $\{1, \Theta, \Theta^2, \ldots, \Theta^k\}$, namely the matrix

$$\begin{bmatrix} 0 & \ldots & & & -c_0 \\ 1 & 0 & \ldots & & \vdots \\ 0 & \ddots & & & \\ \vdots & & & & \\ 0 & \ldots & & 1 & -c_{k-1} \end{bmatrix}$$

Therefore all the problems can be formulated in $\mathbb{Z}^k$ instead of making it in $\mathbb{Z}[\Theta]$.

Example 1. Let

$$M = \begin{bmatrix} 1 & -1 \\ -2 & -1 \end{bmatrix}, \ \mathcal{A} = \left\{ \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -2 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}.$$

Then $t = |\det(M)| = 3$, the eigenvalues of M are $\pm\sqrt{3}$, the Jordan transformation matrix $S = \begin{bmatrix} 1 & \frac{1-\sqrt{3}}{2} \\ 1 & \frac{1+\sqrt{3}}{2} \end{bmatrix}$, $r = \|M^{-1}\| = \frac{\sqrt{3}}{3}$, $K = 2$, $L = 1 + \sqrt{3}$, $I_T = \left\{ \begin{bmatrix} t_1 \\ t_2 \end{bmatrix} \in \mathbb{Z}^2 \middle| -2 \le t_1, t_2 \le 2 \right\}$, $\bar{\mathcal{A}} = \left\{ \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 2 \end{bmatrix} \right\}$ and $\hat{\mathcal{A}} = \left\{ \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \end{bmatrix} \right\}$. The attractors can be seen in Figure 1.

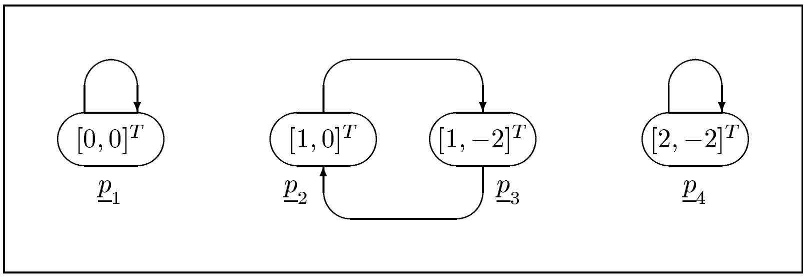

Figure 1

Figure 2 shows the basin of attractions in a region of $\mathbb{Z}^2$. $\mathcal{B}(\underline{p}_1) =$ gray, $\mathcal{B}(\underline{p}_2) =$ black, $\mathcal{B}(\underline{p}_4) =$ white.

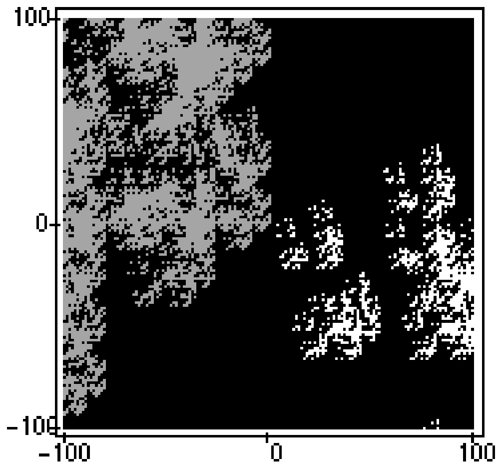

Figure 2

Example 2. Let us choose for the same matrix a different digit set. Let

$$M = \begin{bmatrix} 1 & -1 \\ -2 & -1 \end{bmatrix}, \ \mathcal{A} = \left\{ \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -2 \\ -1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \end{bmatrix} \right\}.$$

In this case the corresponding (different) values are $K = \frac{\sqrt{3}+5}{2}$, $L = \frac{3\sqrt{3}+4}{2}$, $I_T = \left\{ \begin{bmatrix} t_1 \\ t_2 \end{bmatrix} \in \mathbb{Z}^2 \middle| -2 \leq t_1 \leq 2, -3 \leq t_2 \leq 3 \right\}$. Figure 3 shows the attractors of the system $(M, \mathcal{A})$.

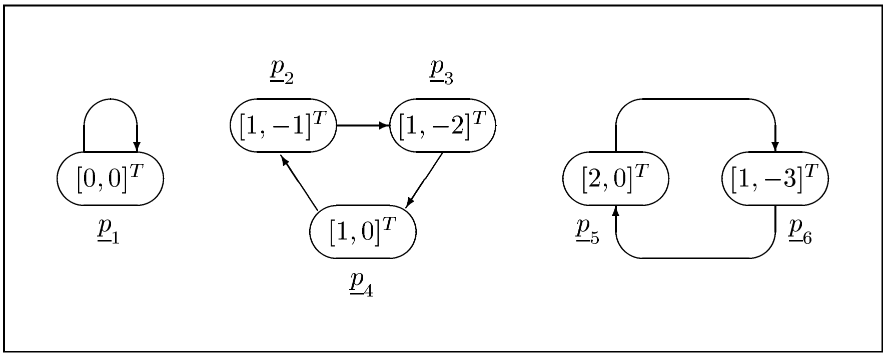

Figure 3

Since $M^* = -M$ therefore it is easy to check that there exist 1- and 2-canonical complete residue systems. It remains for the reader to show that 1- or 2-canonical number systems do not exist.

Example 3. Let us consider the ring of Gaussian integers $\mathbb{Z}[i] = \{a + bi \mid a, b \in \mathbb{Z}\}$. Let $\Theta = A + Bi \in \mathbb{Z}[i]$. Hence $M = \begin{bmatrix} A & -B \\ B & A \end{bmatrix}$, $t = A^2 + B^2$ and $M^* = \begin{bmatrix} A & B \\ -B & A \end{bmatrix}$. From Theorem 8 it follows that if $\mathrm{GCD}(A, B) = 1$ then always exist 1- and also 2-canonical CRS.

Let $\Theta = 2 + i \in \mathbb{Z}[i]$. Then $M = \begin{bmatrix} 2 & -1 \\ 1 & 2 \end{bmatrix}$ and $t = 5$. Let us consider the canonical digit set $\mathcal{A}_1$. In this case the eigenvalues of M are $2 \pm i$, the Jordan transformation matrix $S = \begin{bmatrix} 1 & -i \\ 1 & i \end{bmatrix}$, $r = \|M^{-1}\| = \frac{\sqrt{5}}{5}$, $K = 4$, $L = 1 + \sqrt{5}$, $I_T = \left\{ \begin{bmatrix} t_1 \\ t_2 \end{bmatrix} \in \mathbb{Z}^2 \middle| -2 \le t_1, t_2 \le 2 \right\}$. The attractors can be seen in Figure 4.

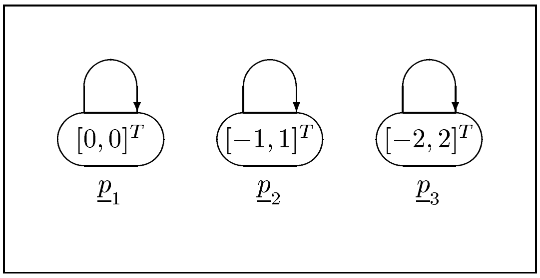

Figure 4

Figure 5 shows the basin of attractions in a region of $\mathbb{Z}^2$. $\mathcal{B}(\underline{p}_1) =$ black, $\mathcal{B}(\underline{p}_2) =$ white, $\mathcal{B}(\underline{p}_3) =$ gray.

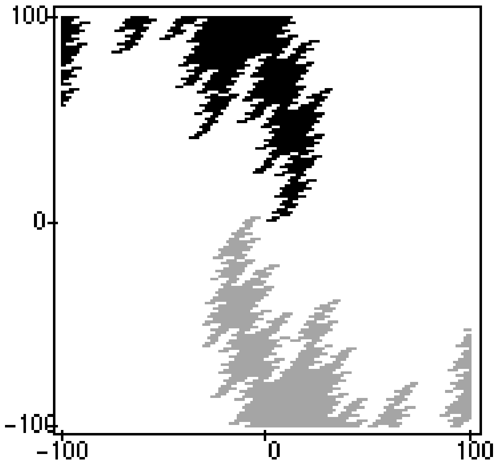

Figure 5

Example 4. Let our last example be the matrix $M = \begin{bmatrix} -1 & -1 \\ 1 & -1 \end{bmatrix}$ with the digit set $\mathcal{A} = \left\{ \begin{bmatrix} -3 \\ 3 \end{bmatrix}, \begin{bmatrix} 2 \\ -3 \end{bmatrix} \right\}$. Then the eigenvalues of M are $-1 \pm i$, the Jordan transformation matrix $S = \begin{bmatrix} 1 & -i \\ 1 & i \end{bmatrix}$, $r = \|M^{-1}\| = \frac{\sqrt{2}}{2}$, $K = 3\sqrt{2}$, $L = 3(2+\sqrt{2})$, $I_T = \left\{ \begin{bmatrix} t_1 \\ t_2 \end{bmatrix} \in \mathbb{Z}^2 \middle| -7 \le t_1 \le 7,\ -5 \le t_2 \le 5 \right\}$. In this case there are two attractors. One of them has period length 3 while the other one has period length 60 (see [10]).

8 Further questions and problems

1. It is known that if $\underline{p} \in \mathcal{P}$ then the maximum of the period length of $\underline{p}$ can be estimated with the number of integers in the k-dimensional ball centered at $\underline{0}$ with radius L. Is there a better estimation?

2. Is there a good upper estimation for the number of the different sets $\mathcal{C}(\underline{p})$?

3. Give all the invertible expanding linear operator M in $\mathbb{Z}^k$ for which there exists a complete residue system $\mathcal{A}$ modulo M such that $(M, \mathcal{A})$ is a number system.

4. How can we characterize the geometric/algebraic stucture of the sets $\mathcal{B}(\underline{p})$, $\underline{p} \in \mathcal{P}$?

5. What can be stated about the attractors in case of special invertible expanding operators, e.g. matrices generated by the ring of integers of a given algebraic field?

6. The problem of characterizing the j-canonical number systems seems to be very interesting.

A remark to the third problem: It is not true that for a given invertible expanding linear operator M there exists always a complete residue system $\mathcal{A}$ modulo M such that $(M, \mathcal{A})$ is a number system. Let a counter-example be the matrix $M = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$ for which does not exists any appropriate digit set $\mathcal{A} = \left\{ \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} a \\ b \end{bmatrix} \right\}$, $(a, b \in \mathbb{Z})$ such that $(M, \mathcal{A})$ would be a number system. It is easy to check that there is always a loop $\begin{bmatrix} -b \\ a \end{bmatrix} \to \begin{bmatrix} -b \\ a \end{bmatrix}$.

The author is grateful to professors Antal Járai and Imre Kátai for their remarks, comments and suggestions.

References

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

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

[3] K.H. Indlekofer, I. Kátai, P. Racskó. *Some remarks on generalized number systems*, Acta Sci. Math. 57 (1993), 543–553.

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

[5] K.H. Indlekofer, A. Járai, I. Kátai. *On topological properties of attractors of some iterated function systems*, Acta Sci. Math., Szeged 60 (1995), 411–427.

[6] E. Isaacson, H. Keller. *Analysis of Numerical Methods*, Wiley, New York, 1966.

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

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

[9] I. Kátai. *Generalized number systems and fractal geometry*, Janus Pannonius University, Pécs, Hungary, 1995.

[10] A. Kovács, N. Harnos. *Fractals and number systems*, Techn. Report, University of Paderborn, CeBIT, 1993.