SIAM J. APPLIED DYNAMICAL SYSTEMS
Vol. 10, No. 4, pp. 1480–1501

Tinkerbell Is Chaotic*

Alexandre Goldsztejn[†], Wayne Hayes[‡], and Pieter Collins[§]

Abstract. Shadowing is a method of backward error analysis that plays a important role in hyperbolic dynamics. In this paper, the shadowing by containment framework is revisited, including a new shadowing theorem. This new theorem has several advantages with respect to existing shadowing theorems: It does not require injectivity or differentiability, and its hypothesis can be easily verified using interval arithmetic. As an application of this new theorem, shadowing by containment is shown to be applicable to infinite length orbits and is used to provide a computer assisted proof of the presence of chaos in the well-known noninjective Tinkerbell map.

Key words. shadowing by containment, interval analysis, chaos computer assisted proof

AMS subject classifications. 37D45, 65G20, 65G30

DOI. 10.1137/100819011

1. Introduction.

The chaos theory emerged from the observation that very simple dynamical systems can have complicated behaviors. A famous example is the Tinkerbell map [2], whose iteration gives rise to the beautiful strange attractor depicted in Figure 1 (simply computed by iterating the map with initial condition $(-0.5, -0.5)$ using double precision computations). However, the observation itself of this strange attractor gives rise to a well-known seeming paradoxical difficulty: Indeed, we interpret Figure 1 as the strange attractor of the Tinkerbell map, while the presence of a strange attractor implies that the Tinkerbell map is chaotic. This in turn implies that the computations shown on Figure 1 suffer from exponential magnification of rounding errors and hence should not be considered as a trustworthy representation of the strange attractor.

In this situation, any attempt toward a forward error analysis is foreseen to fail due to the exponential growth of the forward error. However, the backward error analysis may be able to extract some information from a numerically computed trajectory, and indeed the above mentioned difficulty can be explained in the shadowing backward error analysis framework. A pseudo-orbit is a finite or infinite sequence of vectors whose local error is bounded. It is called an ϵ-pseudo-orbit when the bound is known to be ϵ. An exact trajectory $(\mathbf{y}_k)_{k\in E}$, where $E \subseteq \mathbb{Z}$ is a set of contiguous integers, is a δ-shadow of the pseudo-orbit $(\mathbf{x}_k)_{k\in E}$ if $\max_{k\in E} ||\mathbf{x}_k - \mathbf{y}_k||$ is less than δ. Shadowing theorems provide sufficient conditions for pseudo-orbits to possess shadows.

Shadowing first appeared in the 1970's with Anosov's work as a remarkable property of hyperbolic dynamical systems (cf. [19]): In the neighborhood of a hyperbolic set, for every

*Received by the editors December 21, 2010; accepted for publication (in revised form) by E. Sander June 7, 2011; published electronically December 8, 2011.
http://www.siam.org/journals/siads/10-4/81901.html
[†]CNRS – LINA, 44322 Nantes, France (Alexandre.Goldsztejn@univ-nantes.fr).
[‡]University of California Irvine, Irvine, CA 92697 (Wayne@ics.uci.edu).
[§]Maastricht University, 6200 MD, Maastricht, The Netherlands (Pieter.Collins@cwi.nl).

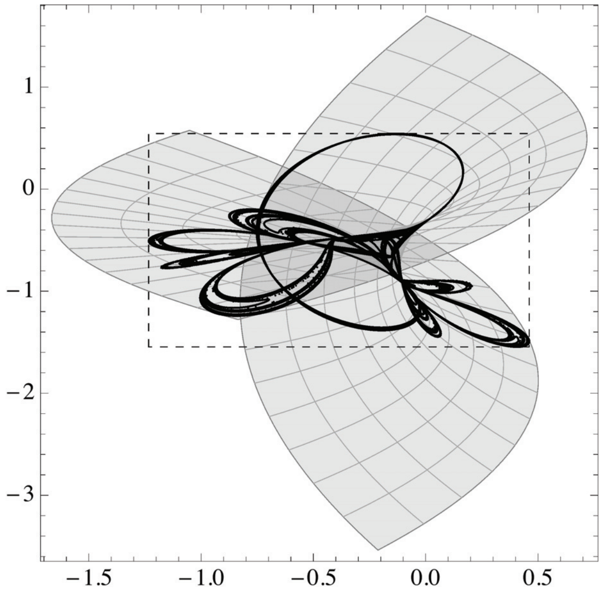

Figure 1. *The Tinkerbell map strange attractor: One million steps computed with standard double precision starting from* $(-0.5, -0.5)$*. The gray area is the image of the dashed rectangle. It shows that the Tinkerbell map is not injective.*

$\delta > 0$ there exists $\epsilon > 0$ such that every pseudo-orbit $\mathbf{x}_k$ that satisfies $||\mathbf{x}_{k+1} - \mathbf{f}(\mathbf{x}_k)|| \leq \epsilon$ has a δ-shadow. Later, some shadowing theorems allowing a numerical verification of the existence of shadows appeared (cf. [35, 41] and [16] for a survey of shadowing methods for numerical solutions of ordinary differential equations). These numerical shadowing theorems verify some property in the neighborhood of a given pseudo-orbit and therefore do not require global hyperbolicity.

Numerical shadowing allows us to quantify the backward error of numerical simulations of chaotic systems. For example, the method presented hereafter can be used to prove that the pseudo-orbit depicted in Figure 1 has a δ-shadow for $\delta = 10^{-6}$. Therefore, this pseudo-orbit is actually a good approximation of an exact trajectory for an initial condition close to $(-0.5, -0.5)$ and thus is representative of the system's dynamics. Such backward error analysis through shadowing is also used, e.g., in [14] in the context of galaxy simulations, in [37, 36] in the context of fluid mechanics, and in [35] for the rigorous numerical proof of the embedding of symbolic dynamics. Among different numerical shadowing methods, we will focus on the so-called containment method (see [13, 15, 41]). It roughly consists of proving the existence of a true trajectory by checking inclusions and exclusions of different sets (usually parallelotopes) and their images by the map.

The contribution of the present paper is threefold: First, the containment property is reformulated using the newly introduced *local containment property*, which is easier to formalize since it assumes that unstable and stable directions are aligned with axes. The general situations are then handled using some affine change of variables. Second, a new containment theorem is proved for finite length orbits, which is more general than the shadowing by containment theorem proposed in [41] since it does not require differentiability or injectivity. Its proof is conceptually very simple since it relies directly on the Poincaré–Miranda theorem.[1] Note that the containment theorem proposed in [21] also does not require differentiability or injectivity: However, its hypotheses are quite difficult to verify. Third, this new containment theorem is extended to bi-infinite orbits, thus reaching the scope of shadowing theorems by refinement like [35].

As an application, we show that this containment theorem for bi-infinite orbits can be used to prove that a dynamical system is chaotic in several senses (positive topological entropy, Li–Yorke chaos, and Devaney chaos). Different computed assisted proofs of the presence of chaos were proposed in, e.g., [34, 9, 42, 10, 26]. All these approaches are similar in the sense that they rely on some topological fixed point theorems whose hypotheses are translated to sufficient conditions for the presence of chaos. However, the usage of different topological fixed point theorems leads to different statements. The sufficient conditions we obtain here are remarkable by their simplicity, which allows an easy numerical verification by a computer.

2. Interval analysis.

Interval analysis is a branch of numerical analysis that was born in the 1960's. It consists of computing with intervals of reals instead of reals, providing a framework for handling uncertainties and verified computations (see [27, 1, 28, 17, 20] for a survey). Among other applications, interval analysis was used to provide rigorous numerical proofs of mathematical statements (resolution of the 14th Smale problem [38], properties of manifolds [6, 7], and properties of chaotic dynamical systems [3, 18, 4, 39, 33]).

2.1. Intervals, interval vectors, and interval matrices.

An interval is a closed connected subset of $\mathbb{R}$. Intervals are denoted by bracketed symbols, e.g., $[x] \subseteq \mathbb{R}$. When no confusion is possible, lower and upper bounds of an interval $[x]$ are denoted by $\underline{x} \in \mathbb{R}$ and $\overline{x} \in \mathbb{R}$, with $\underline{x} \leq \overline{x}$, i.e., $[x] = [\underline{x}, \overline{x}] = \{x \in \mathbb{R} : \underline{x} \leq x \leq \overline{x}\}$. Furthermore, a real number x will be identified with the degenerate interval $[x, x]$. The magnitude and mignitude of an interval, respectively, are defined by $|[x]| := \max\{|x| : x \in [x]\} = \max\{|\underline{x}|, |\overline{x}|\}$ and $\langle [x] \rangle := \min\{|x| : x \in [x]\}$; thus $\langle [x] \rangle = \min\{|\underline{x}|, |\overline{x}|\}$ if $0 \notin [x]$, while $\langle [x] \rangle = 0$ otherwise.

There are two equivalent ways of defining interval vectors. On the one hand, being given two vectors $\underline{\mathbf{x}} \leq \overline{\mathbf{x}} \in \mathbb{R}^n$ (where the inequality is defined componentwise), an interval of vectors is obtained by considering

$$[\mathbf{x}] := \{\mathbf{x} \in \mathbb{R}^n : \underline{\mathbf{x}} \leq \mathbf{x} \leq \overline{\mathbf{x}}\}. \tag{2.1}$$

On the other hand, being given intervals $[x_i] \in \mathbb{IR}$ for $i \in \{1, \ldots, n\}$, a vector of intervals is obtained by considering

$$[\mathbf{x}] := \{\mathbf{x} \in \mathbb{R}^n : \forall i \in \{1, \ldots n\}, x_i \in [x_i]\}. \tag{2.2}$$

These two definitions are obviously equivalent following the notational convention $\underline{\mathbf{x}} = (\underline{x}_i)$, $\overline{\mathbf{x}} = (\overline{x}_i)$, and $[x_i] = [\underline{x}_i, \overline{x}_i]$ and will be used interchangeably. In the whole paper, we consider the infinity norm and distance, so the closed ball $B(\mathbf{0}, 1)$ is actually the interval vector $([-1, 1], \ldots, [-1, 1])^T$.

Interval matrices are defined similarly to interval vectors as either intervals of matrices or matrices of intervals.

[1] The Poincaré–Miranda theorem is a generalization of the intermediate value theorem to arbitrary dimensions which is equivalent to the more famous Brouwer fixed point theorem [31, 25].

2.2. Interval arithmetic. Operations $\circ \in \{+, \times, -, \div\}$ are extended to intervals in the following way:

$$[x] \circ [y] := \{x \circ y : x \in [x], y \in [y]\}. \tag{2.3}$$

The division is defined for intervals $[\underline{y}, \overline{y}]$ that do not contain zero. Unary elementary functions $f(x)$ (like exp, ln, sin, etc.) are also be extended to intervals similarly:

$$f([x]) = \{f(x) : x \in [x]\}. \tag{2.4}$$

All these elementary interval extensions form the interval arithmetic (IA). As real numbers are identified to degenerated intervals, the IA actually generalizes the real arithmetic, and mixed operations like $1 + [1, 2] = [2, 3]$ are interpreted using (2.3).

Rounded computations. As real numbers are approximately represented by floating point numbers [11], the IA cannot match the definitions (2.3) and (2.4) exactly. In order to preserve the inclusion property, the IA has to be implemented using an outward rounding. For example, $[1, 3]/[10, 10] = [0.1, 0.3]$, while both 0.1 and 0.3 cannot be exactly represented with standard floating point numbers. Therefore, the computed result will be $[0.1^-, 0.3^+]$, where 0.1^- (respectively, 0.3^+) is a floating point number smaller than 0.1 (respectively, greater than 0.3) (we actually expect the greatest floating point number to be smaller than 0.1 and the smallest floating point number to be greater than 0.3, which is often achieved by IA implementations). Among other implementations of IA, we can cite the C/C++ libraries PROFIL/BIAS [22] and Gaol [12], the MATLAB toolbox INTLAB [32], and *Mathematica* [40]. The developments presented in the rest of the paper use the ideal real IA. The algorithms are finally implemented using outwardly rounded floating point interval arithmetic.

2.3. Interval extensions. An interval function $[\mathbf{f}] : \mathbb{IR}^n \longrightarrow \mathbb{IR}^m$ is an interval extension of the real function $\mathbf{f} : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ if for all $[\mathbf{x}] \in \mathbb{IR}^n$ we have $[\mathbf{f}]([\mathbf{x}]) \supseteq \{\mathbf{f}(\mathbf{x}) : \mathbf{x} \in [\mathbf{x}]\}$. Thus interval extensions allow computing enclosures of the real function's range over boxes. So-called natural interval extensions of a function are obtained by evaluating an expression of this function for interval arguments using the IA. In particular when every variable has one unique occurrence in the function's expression the natural interval extension is optimal; i.e., it computes the exact function range.

Example 1. Consider the function $\mathbf{f} : \mathbb{R}^2 \to \mathbb{R}^2$ defined by $\mathbf{f}(\mathbf{x}) = \big((x_1 + x_2)^2, (x_1 - 1)^2 \cos x_2\big)$. The image of the box $[\mathbf{x}] = ([-1, 1], [-1, 1])$ is depicted in Figure 2. Since the expression of each component of $\mathbf{f}$ contains one unique occurrence of each variable, their interval evaluation $\big(([x_1] + [x_2])^2, ([x_1] - 1)^2 \cos[x_2]\big) = ([0, 4], [0, 4])$ is the optimal enclosure of the image (it is depicted in Figure 2). The image enclosure can be pessimistic if other expressions of the function are used; e.g., $\big(x_1^2 + 2x_1x_2 + x_2^2, (x_1^2 - 2x_1 + 1) \cos x_2\big)$ evaluated for the same box gives rise to $([-2, 4], [-1, 4])$ (also depicted in Figure 2).

In the context of the developments proposed hereafter, interval extensions will be evaluated on small boxes (whose width is about 10^{-6}). For such small intervals, the order of convergence of interval extension provides meaningful information about its pessimism: The natural interval extension has a linear order of convergence, which means that in the worst case its pessimism is proportional to the width of its interval arguments (see [28] for a details). This is in general too pessimistic for our usage.

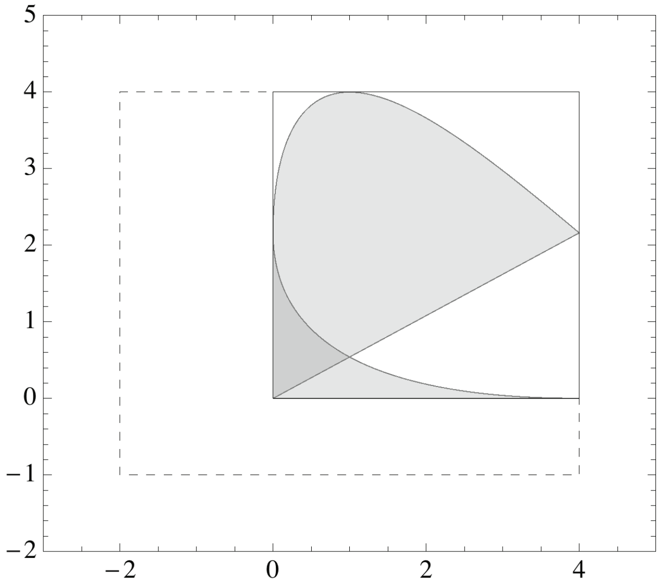

Figure 2. *In gray, the image of the box* $([-1,1]),[-1,1])$ *through the function of Example* 1 *(the darker area shows the self-overlapping of the image). Also displayed are the interval hull of this set (full line box) and the pessimistic enclosure (dashed line box) computed in Example* 1.

On the other hand, the centered forms described below enjoy a quadratic order of convergence, which means that in the worst case its pessimism is proportional to the square of the width of its interval arguments. This is a drastic and necessary improvement with respect to the linear order of convergence of the natural interval extensions when dealing with small boxes. A centered form has the following form:

$$[\mathbf{f}]([\mathbf{x}]) = \mathbf{f}(\tilde{\mathbf{x}}) + [A]([\mathbf{x}] - \tilde{\mathbf{x}}). \tag{2.5}$$

There are different ways of choosing $[A]$. When $\mathbf{f}$ is differentiable, $[A]$ can be chosen as some interval extension of the Jacobian of $\mathbf{f}$. Lipschitz interval matrices allow generalizing this property to nondifferentiable maps: The introduction of Lipschitz interval matrices allows a rigorous linearization of a map, i.e., a linearization with controlled truncation error. This gives rise to an interval extension called the centered interval extensions. An interval matrix $[L] \in \mathbb{IR}^{n\times n}$ is a Lipschitz interval matrix (LIM) for $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^n$ and $E \subseteq \mathbb{R}^n$ if and only if for all $\mathbf{x}, \mathbf{y} \in E$ there exists $L \in [L]$ such that $\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{y}) = L(\mathbf{x} - \mathbf{y})$. A LIM can be computed similarly to the interval evaluation of the Jacobian matrix: In case the function is differentiable, then an interval evaluation of its Jacobian expression gives rise to a LIM. In case it contains some absolute value, the following rule can be used for the Jacobian differentiation to give rise to a LIM: $\mathrm{abs}'([x]) = -1$ if $\underline{x} \leq 0$, $\mathrm{abs}'([x]) = 1$ if $\overline{x} \geq 0$, and $\mathrm{abs}'([x]) = [-1, 1]$ otherwise (see [28] for details).

2.4. The Poincaré–Miranda theorem. Interval extensions allow computing enclosures of the range of a function. One important application of this framework is the rigorous numerical proof of existence of solutions to systems of equations. This is done by using interval extensions to verify the hypothesis of some existence theorems. Among different existence theorems that can be used in conjunction with interval analysis, we will use the Poincaré–Miranda theorem [31] which can be formulated as follows.

Theorem 2.1 (Poincaré–Miranda theorem). *Let* $\mathbf{f} = B(\mathbf{0}, 1) \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}^n$ *be a continuous function. Associate each equation with a variable through a permutation* π *of* $\{1, \ldots, n\}$ *and assume that for all* $i \in \{1, \ldots, n\}$

- $\mathbf{x} \in B(\mathbf{0}, 1)$ *and* $x_{\pi(i)} = -1$ *implies* $f_i(\mathbf{x}) \leq 0$ *and*
- $\mathbf{x} \in B(\mathbf{0}, 1)$ *and* $x_{\pi(i)} = 1$ *implies* $f_i(\mathbf{x}) \geq 0$.

Then the system $\mathbf{f}(\mathbf{x}) = 0$ *has a solution inside* $B(\mathbf{0}, 1)$.

The Poincaré–Miranda theorem is actually equivalent to the more famous Brouwer fixed point theorem (see [25, 24]). Its hypotheses are obviously very well suited for a numerical verification by interval analysis.

3. Shadowing by containment.

3.1. Local containment theorems. In order to apply the results of this section in a more general context in section 3.3, we consider nonautonomous dynamical systems $(\mathbf{f}_i)_{i\in\mathbb{Z}}$. An orbit $(\mathbf{x}_i)_{i\in\mathbb{Z}}$ of such a system satisfies $\mathbf{x}_{i+1} = \mathbf{f}_i(\mathbf{x}_i)$. In the rest of the section, $(\mathbf{f}_i)_{i\in\mathbb{Z}}$ is a nonautonomous dynamical system with $\mathbf{f}_i : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ continuous.

The following definition of the local inductive containment property (LICP) focuses on the following situations: The pseudo-orbit is the bi-infinite identically null sequence $(\ldots, \mathbf{0}, \mathbf{0}, \mathbf{0}, \ldots)$; stable and unstable directions are close enough to canonical basis vectors; the maximum distance between the pseudotrajectory and its shadow is 1. Some affine change of variables will allow dealing with more general situations in section 3.3.

Definition 3.1. *Let* $\mathbf{f} = (f_1, \ldots, f_n)^T : \mathbb{R}^n \to \mathbb{R}^n$ *be a continuous map. Let* $\mathcal{S}$ *and* $\mathcal{U}$ *form a partition of* $\{1, \ldots, n\}$ *(*$\mathcal{S}$ *refers to the stable directions and* $\mathcal{U}$ *to the unstable directions). Then, we say that* $\mathbf{f}$ *satisfies the* $(\mathcal{S},\mathcal{U})$*-LICP if both of the following hold:*

(LICP1) *For all* $j \in \mathcal{S}$ *and for all* $\mathbf{x} \in B(\mathbf{0}, 1)$, $-1 \leq f_j(\mathbf{x}) \leq 1$.

(LICP2) *For all* $j \in \mathcal{U}$ *and for all* $\mathbf{x} \in B(\mathbf{0}, 1)$, *either*

1. $x_j = -1$ *implies* $f_j(\mathbf{x}) \leq -1$ *and*
2. $x_j = 1$ *implies* $f_j(\mathbf{x}) \geq 1$,

or

1'. $x_j = -1$ *implies* $f_j(\mathbf{x}) \geq 1$ *and*

2'. $x_j = 1$ *implies* $f_j(\mathbf{x}) \leq -1$.

Finally, the index set $\mathcal{U}$ *is split into* $\mathcal{U}^+$ *and* $\mathcal{U}^-$, *satisfying, respectively,* (LICP2.k) *and* (LICP2.k').

Note that the LICP does not require injectivity or differentiability. See Example 2 and Figure 3 for a function that satisfies Definition 3.1. Then, the following containment theorem for finite length shadows is proposed. Its proof relies on a simple application of the Poincaré–Miranda theorem to the system of equations formed with $\mathbf{x}_{k+1} = \mathbf{f}_k(\mathbf{x}_k)$.

Theorem 3.2. *Let* $\mathcal{S}$ *and* $\mathcal{U}$ *be a partition of* $\{1, \ldots, n\}$. *Consider* $a, b \in \mathbb{Z}$ *such that* $a < b$, *and suppose that* $\mathbf{f}_i$, *for* $i \in \{a, a+1, \ldots, b-1\}$, *satisfies the* $(\mathcal{S},\mathcal{U})$*-LICP. Then there exists a finite orbit* $(\mathbf{x}_i)_{a\leq i\leq b}$ *bounded in* $B(\mathbf{0}, 1)$ *such that* $\mathbf{x}_{i+1} = \mathbf{f}_i(\mathbf{x}_i)$ *for all* $i \in \{a, \ldots, b-1\}$.

Proof. Let $\mathbf{x}_i = (x_{i,1}, \ldots, x_{i,n})^T$ for $i \in \{a, \ldots, b\}$, and let $\mathbf{f}_i = (f_{i,1}, \ldots, f_{i,n})^T$ for $i \in \{a, \ldots, b-1\}$. We fix arbitrarily $x_{a,j} \in [-1, 1]$ for $j \in \mathcal{S}$ and $x_{b,j} \in [-1, 1]$ for $j \in \mathcal{U}$ and prove the existence of an orbit for this choice. So we consider the following system of $n(b-a)$ equations:

$$
\begin{aligned}
g_{i,j}(\mathbf{x}_a,\dots,\mathbf{x}_b) &:= -f_{i,j}(\mathbf{x}_i)+x_{i+1,j}=0 \quad \text{if } j\in\mathcal{S}\cup\mathcal{U}^-,\\
g_{i,j}(\mathbf{x}_a,\dots,\mathbf{x}_b) &:= \ \ f_{i,j}(\mathbf{x}_i)-x_{i+1,j}\ =0 \quad \text{if } j\in\mathcal{U}^+
\end{aligned} \tag{3.1}
$$

for $(i,j)\in\{a,\dots,b-1\}\times\{1,\dots,n\}=:\mathcal{E}$. Since $x_{a,j}$ are fixed for $j\in\mathcal{S}$ and $x_{b,j}$ are fixed for $j\in\mathcal{U}$, the $n(b-a)$ variables are $x_{i,j}$ for $(i,j)\in\{a+1,\dots,b\}\times\mathcal{S}\cup\{a,\dots,b-1\}\times\mathcal{U}=:\mathcal{V}$. A solution to this system of equations obviously corresponds to an exact orbit. In order to apply the Poincaré–Miranda theorem we associate equations and variables through a permutation $\pi:\mathcal{E}\to\mathcal{V}$ defined as follows: $\pi(i,j)=(i+1,j)$ if $j\in\mathcal{S}$ and $\pi(i,j)=(i,j)$ if $j\in\mathcal{U}$. It remains just to verify that the hypotheses of the Poincaré–Miranda theorem do hold; that is, we need to prove that for all $(i,j)\in\mathcal{V}$, for all $x_{i,j}\in[0,1]$

$$x_{\pi(i,j)}=-1 \implies g_{i,j}(\mathbf{x}_a,\dots,\mathbf{x}_b)\le 0, \tag{3.2}$$

$$x_{\pi(i,j)}=1 \implies g_{i,j}(\mathbf{x}_a,\dots,\mathbf{x}_b)\ge 0. \tag{3.3}$$

So consider arbitrary $x_{i,j}\in[0,1]$ for $(i,j)\in\mathcal{V}$, so $\mathbf{x}_a,\dots,\mathbf{x}_b\in B(\mathbf{0},1)$ since $x_{i,j}$ are fixed in $[0,1]$ for $(i,j)\notin\mathcal{V}$, and let us prove that (3.2) and (3.3) hold: We consider the equation (i,j) which is associated with the variable $\pi(i,j)$.

Let us suppose that $x_{\pi(i,j)}=-1$. If $j\in\mathcal{S}$ (so $x_{i+1,j}=-1$), then by (LICP1) we have $-1\le f_{i,j}(\mathbf{x}_i)\le 1$. Thus $x_{i+1,j}\le f_{i,j}(\mathbf{x}_i)$, which means $g_{i,j}(\mathbf{x}_a,\dots,\mathbf{x}_b)\le 0$. Now if $j\in\mathcal{U}^+$ (so $x_{i,j}=-1$), then by (LICP2.1) we have $f_{i,j}(\mathbf{x}_i)\le -1$. Thus $f_{i,j}(\mathbf{x}_i)\le x_{i,j}$, which means $g_{i,j}(\mathbf{x}_a,\dots,\mathbf{x}_b)\le 0$. Finally, if $j\in\mathcal{U}^-$ (so $x_{i,j}=-1$), then by (LICP2.1') we have $f_{i,j}(\mathbf{x}_i)\ge 1$. Thus $f_{i,j}(\mathbf{x}_i)\ge x_{i,j}$, which means $g_{i,j}(\mathbf{x}_a,\dots,\mathbf{x}_b)\le 0$. So (3.2) is satisfied in all cases.

Similarly, let us suppose that $x_{\pi(i,j)}=1$. If $j\in\mathcal{S}$ (so $x_{i+1,j}=1$), then by (LICP1) we have $-1\le f_{i,j}(\mathbf{x}_i)\le 1$. Thus $f_{i,j}(\mathbf{x}_i)\le x_{\pi(i,j)}$, which reads as $g_{i,j}(\mathbf{x}_a,\dots,\mathbf{x}_b)\ge 0$. Now if $j\in\mathcal{U}^+$ (so $x_{i,j}=1$), then by (LICP2.2) we have $f_{i,j}(\mathbf{x}_i)\ge 1$. Thus $f_{i,j}(\mathbf{x}_i)\ge x_{\pi(i,j)}$, which means $g_{i,j}(\mathbf{x}_a,\dots,\mathbf{x}_b)\ge 0$. Finally, if $j\in\mathcal{U}^-$ (so $x_{i,j}=1$), then by (LICP2.2') we have $f_{i,j}(\mathbf{x}_i)\le -1$. Thus $f_{i,j}(\mathbf{x}_i)\le x_{i,j}$, which means $g_{i,j}(\mathbf{x}_a,\dots,\mathbf{x}_b)\ge 0$. So (3.3) is satisfied in all cases. ■

Theorem 3.2 is now extended to bi-infinite sequences.

Theorem 3.3. *Let $\mathcal{S}$ and $\mathcal{U}$ be a partition of $\{1,\dots,n\}$. Consider a sequence $(\mathbf{f}_i)_{i\in\mathbb{Z}}$ of continuous maps, and suppose that $\mathbf{f}_i$ satisfies the $(\mathcal{S},\mathcal{U})$-LICP for all $i\in\mathbb{Z}$. Then there exists a bi-infinite orbit $(\mathbf{x}_i)_{i\in\mathbb{Z}}$ bounded in $B(\mathbf{0},1)$.*

Proof. For an arbitrary $m\in\mathbb{N}$, we can apply Theorem 3.2 with $a=-m$ and $b=m$, hence proving for all $m\in\mathbb{N}$ the existence of finite length orbits

$$\left(\mathbf{x}^m_{-m},\mathbf{x}^m_{-m+1},\dots,\mathbf{x}^m_{m-1},\mathbf{x}^m_{m}\right) \tag{3.4}$$

such that $\mathbf{x}^m_{k+1}=\mathbf{f}_k(\mathbf{x}^m_k)$. Note that, from another point of view, with each $k\in\mathbb{Z}$ can be associated the sequence

$$\mathcal{S}_k:=(\mathbf{x}_k^{|k|+i})_{i\in\mathbb{N}}, \tag{3.5}$$

and that the subsequences of $\mathcal{S}_k$ are $(\mathbf{x}_k^{|k|+\rho_i})_{i\in\mathbb{N}}$, where $(\rho_i)_{i\in\mathbb{N}}$ is a strictly increasing sequence of nonnegative integers. Let us denote the accumulation points of $\mathcal{S}_k$ by $\mathcal{A}_k$, which is nonempty since the sequence is bounded.

First, we show that for all $k \in \mathbb{Z}$

$$\mathbf{x}_k^* \in \mathcal{A}_k \implies \mathbf{f}_k(\mathbf{x}_k^*) \in \mathcal{A}_{k+1}. \tag{3.6}$$

Since $\mathbf{x}_k^* \in \mathcal{A}_k$, it is the limit of a subsequence $(\mathbf{x}_k^{|k|+\rho_i})_{k\in\mathbb{N}}$ of $\mathcal{S}_k$, and $\mathbf{f}_k$ being continuous we obtain

$$\mathbf{f}_k(\mathbf{x}_k^*) = \lim_{i\to\infty} \mathbf{f}_k\big(\mathbf{x}_k^{|k|+\rho_i}\big). \tag{3.7}$$

Now for $|k|+\rho_i \geq |k+1|$ we have $\mathbf{f}_k(\mathbf{x}_k^{|k|+\rho_i}) = \mathbf{x}_{k+1}^{|k|+\rho_i}$, so $\mathbf{f}_k(\mathbf{x}_k^*)$ is the limit of a subsequence of $\mathcal{S}_{k+1}$. Therefore $\mathbf{f}_k(\mathbf{x}_k^*) \in \mathcal{A}_{k+1}$, which proves (3.6).

Second, we show that for all $k \in \mathbb{Z}$

$$\mathbf{x}_k^* \in \mathcal{A}_k \implies \exists \mathbf{x}_{k-1}^* \in \mathcal{A}_{k-1}, \mathbf{f}_{k-1}(\mathbf{x}_{k-1}^*) = \mathbf{x}_k^*. \tag{3.8}$$

Again $\mathbf{x}_k^*$ is the limit of a subsequence $(\mathbf{x}_k^{|k|+\rho_i})_{k\in\mathbb{N}}$ of $\mathcal{S}_k$. We consider now $(\mathbf{x}_{k-1}^{|k|+\rho_{i+1}})_{k\in\mathbb{N}}$, which is a subsequence of $\mathcal{S}_{k-1}$ since $|k|+\rho_{i+1} \geq |k-1|$. Since this sequence is bounded, it has a subsequence $(\mathbf{x}_{k-1}^{|k|+\rho'_i})_{k\in\mathbb{N}}$ which converges to some $\mathbf{x}_{k-1}^*$. Since this subsequence is also a subsequence of S_{k-1}, we have $\mathbf{x}_{k-1}^* \in \mathcal{A}_{k-1}$. Using the continuity of $\mathbf{f}_{k-1}$ we prove

$$\mathbf{f}_k(\mathbf{x}_{k-1}^*) = \lim_{i\to\infty} \mathbf{f}_k\big(\mathbf{x}_{k-1}^{|k|+\rho'_i}\big) = \lim_{i\to\infty} \mathbf{x}_k^{|k|+\rho'_i} = \mathbf{x}_k^*, \tag{3.9}$$

which proves (3.8).

Finally, since A_0 is nonempty, we consider $\mathbf{x}_0^* \in \mathcal{A}_0$, while (3.6) and (3.8) prove that it can be extended forward and backward to a bi-infinite orbit bounded inside $B(\mathbf{0}, 1)$. ■

Theorems 3.2 and 3.3 present several advantages with respect to usual shadowing theorems like the shadowing lemma [19], Stoffer and Palmer's shadowing lemma [35], and Young, Hayes, and Jackson's containment theorem [41]: First, they do not need $\mathbf{f}$ to be differentiable or injective, while the three former do require both. Second, the LICP requires only $2\,\mathrm{card}(\mathcal{S})+1$ simple interval evaluations to be verified (see Example 2). However, the price paid is the uniqueness of the shadow: The shadowing theorems of [19] and [35] do prove the uniqueness of the shadow in addition to its existence. It can be seen trivially that uniqueness does not hold in general for Theorem 3.3, as the identity matrix does satisfy the LICP while admitting every constant bi-infinite sequence as orbits. However, uniqueness is necessary for embedding purposes in section 4. Section 3.2 provides a uniqueness sufficient condition that will be used in the section 4.

The following example illustrates the power of Theorem 3.3: Three interval evaluations performed with rounded floating point computations allow proving the existence of a bi-infinite length orbit for a nonlinear dynamical system which lacks injectivity and differentiability.

Example 2. Consider the continuous function $\mathbf{f}(\mathbf{x}) = M\,\mathbf{g}(\mathbf{x})$, where

$$M = \begin{pmatrix} 0.5 & 0.1 \\ -0.1 & 2 \end{pmatrix} \quad \text{and} \quad \mathbf{g}(\mathbf{x}) = \begin{pmatrix} |\,x_1 - 0.5\,| \\ x_2 - 0.1x_1^2 \end{pmatrix}. \tag{3.10}$$

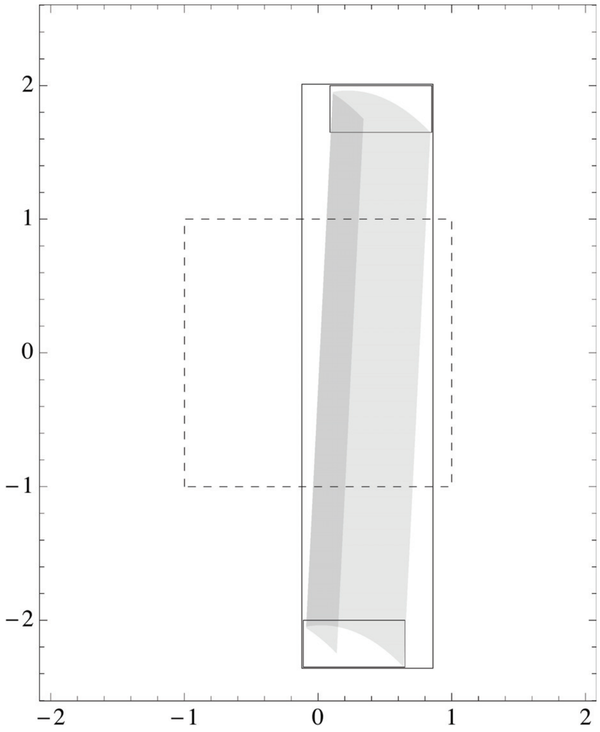

Figure 3. *Representation of the image of* $([-1,1],[-1,1])$ *by the function* $\mathbf{f}$ *defined by* (3.10) *(the darker area shows the self-overlapping of the image). The boundary of* $B(\mathbf{0},1)$ *is represented by a dashed rectangle. The image of* $B(\mathbf{0},1)$ *and of its upper and lower sides are enclosed by rectangles, which obviously show that* $\mathbf{f}$ *satisfies the* $(\{1\},\{2\})$*-LICP.*

An approximation of the image of $B(\mathbf{0},1) = ([-1,1],[-1,1])^T$ by $\mathbf{f}$ is depicted in Figure 3. It clearly shows that $\mathbf{f}$ is neither differentiable nor injective inside $B(\mathbf{0},1)$. Let $[\mathbf{f}]$ be the natural interval extension of $\mathbf{f}$. Then,

$$[\mathbf{f}](\ [-1,1]\ ,\ [-1,1]\ ) = ([-0.11, 0.85], [-2.35, 2]), \tag{3.11}$$

$$[\mathbf{f}](\ [-1,1]\ ,\ -1\ ) = ([-0.11, 0.65], [-2.35, -2.]), \tag{3.12}$$

$$[\mathbf{f}](\ [-1,1]\ ,\ 1\ ) = ([0.09, 0.85], [1.65, 2.]). \tag{3.13}$$

These three rectangles are also represented in Figure 3. These computations prove that $\mathbf{f}$ satisfies the $(\{1\},\{2\})$-LICP. As a consequence, one can apply Theorem 3.3 with $\mathbf{f}_i = \mathbf{f}$ for all $i \in \mathbb{Z}$, which proves the existence of a bi-infinite orbit $(\mathbf{x}_i)_{i\in\mathbb{Z}}$ bounded in $B(\mathbf{0},1)$.

3.2. A sufficient condition for the uniqueness of the shadows. Theorem 3.3 applies in situations where the maps are not hyperbolic. In order to obtain uniqueness of the shadow, an additional hypothesis on the Lipschitz matrices of the maps is required that somehow forces the maps to be hyperbolic. To this end, we introduce the following definition.

Definition 3.4. *For some $\epsilon > 0$ and $\mathcal{S}$ and $\mathcal{U}$ forming a partition of $\{1, \dots, n\}$, a matrix $A \in \mathbb{R}^{n\times n}$ is called ϵ-canonically*[2] *hyperbolic with respect to $\mathcal{S}$ and $\mathcal{U}$ (for short, ϵ-($\mathcal{S}$,$\mathcal{U}$) canonically hyperbolic) if for all $i \in \{1, \dots n\}$*

$$i \in \mathcal{S} \implies \sum_j |a_{ij}| \le 1 - \epsilon, \tag{3.14}$$

$$i \in \mathcal{U} \implies |a_{ii}| - \sum_{j\neq i} |a_{ij}| \ge 1 + \epsilon. \tag{3.15}$$

Furthermore, an interval matrix $[A] \in \mathbb{IR}^{n\times n}$ is called ϵ-($\mathcal{S}$,$\mathcal{U}$) canonically hyperbolic if for all $i \in \{1, \dots n\}$

$$i \in \mathcal{S} \implies \sum_j |[a_{ij}]| \le 1 - \epsilon, \tag{3.16}$$

$$i \in \mathcal{U} \implies \langle [a_{ij}] \rangle - \sum_{j\neq i} |[a_{ij}]| \ge 1 + \epsilon. \tag{3.17}$$

It is obviously seen that an ϵ-($\mathcal{S}$,$\mathcal{U}$) canonically hyperbolic interval matrix contains only ϵ-($\mathcal{S}$,$\mathcal{U}$) canonically hyperbolic real matrices.[3] As a typical example, the interval matrix

$$\begin{pmatrix} [-0.3, 0.3] & [-0.3, 0.3] & [-0.3, 0.3] \\ [-0.1, 0.1] & [1.3, 1.4] & [-0.1, 0.1] \\ [-0.3, 0.3] & [-0.3, 0.3] & [-0.3, 0.3] \end{pmatrix} \tag{3.18}$$

is ϵ-($\mathcal{S}$,$\mathcal{U}$) canonically hyperbolic for $\epsilon = 0.1$, $\mathcal{S} = \{1, 3\}$, and $\mathcal{U} = \{2\}$.

The following proposition is the main property of ϵ-($\mathcal{S}$,$\mathcal{U}$) canonically hyperbolic matrices that will be used later.

Proposition 3.5. *Let $A_k \in \mathbb{R}^{n\times n}$ for $k \in \{-K, \dots, K-1\}$ be ϵ-($\mathcal{S}$,$\mathcal{U}$) canonically hyperbolic matrices for $\epsilon > 0$. Consider $\mathbf{u}_k \in B(\mathbf{0}, 1) \subseteq \mathbb{R}^n$ for $k \in \{-K, \dots, K\}$ such that $\mathbf{u}_{k+1} = A_k \mathbf{u}_k$. Then $||\mathbf{u}_0|| \le (1 + \epsilon)^{-K}$.*

Proof. Define $m_k = \arg\max_i |u_{k\,i}|$, so $||\mathbf{u}_k|| = |u_{k\,m_k}| \le 1$. First, we claim that

$$i \in \mathcal{S} \implies |u_{k+1\,i}| \le (1 - \epsilon)||\mathbf{u}_k||. \tag{3.19}$$

Indeed, $|u_{k+1\,i}| = |(A_k\,\mathbf{u}_k)_i| \le \sum_j |a_{kij}||u_{kj}| \le ||\mathbf{u}_k||(\sum_j |a_{kij}|)$ and the claim comes by (3.14). As a consequence, we have

$$m_{k+1} \in \mathcal{S} \implies ||\mathbf{u}_{k+1}|| \le (1 - \epsilon)||\mathbf{u}_k||. \tag{3.20}$$

Now, two cases arise. First, $m_k \in \mathcal{S}$ for all $k \in \{-K, \dots, 0\}$. Using (3.20) inductively we show that $||\mathbf{u}_0|| \le ||\mathbf{u}_{-K}||(1 - \epsilon)^K \le (1 - \epsilon)^K$. Finally, note that $(1 - \epsilon)^K \le (1 + \epsilon)^{-K}$ since $(1 + \epsilon)(1 - \epsilon) = 1 - \epsilon^2 \le 1$.

[2]Here *canonically* emphasizes the fact that this class of matrices has its stable and unstable manifolds close to the canonical basis vectors.

[3]This is since by definition of magnitude and mignitude $a_{ij} \in [a_{ij}] \Rightarrow |a_{ij}| \le |[a_{ij}]|$ and $a_{ij} \in [a_{ij}] \Rightarrow |a_{ij}| \ge \langle [a_{ij}] \rangle$.

Second, there exists $k^* \in \{-K,\ldots,0\}$ such that $m_k \in \mathcal{U}$. We first claim that $k \in \mathcal{U}$ implies both $||\mathbf{u}_{k+1}|| \geq (1+\epsilon)||\mathbf{u}_k||$ and $m_{k+1} \in \mathcal{U}$. With m_k denoted by m for clarity,

$$||\mathbf{u}_{k+1}|| = |(A_k\,\mathbf{u}_k)_{m_{k+1}}| \geq |(A_k\,\mathbf{u}_k)_m| = |\sum_j a_{kmj}u_{kj}| \tag{3.21}$$

$$\geq |a_{kmm}||u_{k\,m}| - \sum_{j\neq m} |a_{kmj}||u_{kj}| \tag{3.22}$$

$$= |a_{kmm}|\,||\mathbf{u}_k|| - \sum_{j\neq m} |a_{kmj}||u_{kj}| \tag{3.23}$$

$$\geq |a_{kmm}|\,||\mathbf{u}_k|| - \sum_{j\neq m} |a_{kmj}|\,||\mathbf{u}_k|| \tag{3.24}$$

$$= \left(|a_{kmm}| - \sum_{j\neq m} |a_{kmj}|\right)||\mathbf{u}_k||, \tag{3.25}$$

and finally (3.15) shows that $||\mathbf{u}_{k+1}|| \geq (1+\epsilon)||\mathbf{u}_k||$. For the second part of the claim, suppose that $m_{k+1} \in \mathcal{S}$, so by (3.20) we have $||\mathbf{u}_{k+1}|| \leq (1-\epsilon)||\mathbf{u}_k||$, but this contradicts the previously shown inequality, and therefore $m_{k+1} \in \mathcal{U}$. Finally, applying the claim inductively from k^* to 0 shows that $m_0 \in \mathcal{U}$ and then from 0 to K shows that $||\mathbf{u}_K|| \geq (1+\epsilon)^K||\mathbf{u}_0||$. Since $||\mathbf{u}_K|| \leq 1$ we have proved that $||\mathbf{u}_0|| \leq (1+\epsilon)^{-K}$. ∎

Now, we have the following consequence of Proposition 3.5.

Theorem 3.6. *Let $\mathcal{S}$ and $\mathcal{U}$ be a partition of $\{1,\ldots,n\}$. Consider two orbits $(\mathbf{x}_k)_{k\in E}$ and $(\mathbf{y}_k)_{k\in E}$ of $(\mathbf{f}_k)_{k\in E\setminus\{m\}}$, where $E = \{k \in \mathbb{Z} : |k| \leq m\}$, that are bounded inside $B(\mathbf{0},1)$. Suppose that $[L_k]$ are LIMs for $\mathbf{f}_k$ and $B(\mathbf{0},1)$ that are ϵ-$(\mathcal{S},\mathcal{U})$ canonically hyperbolic for $\epsilon > 0$. Then $||\mathbf{x}_0 - \mathbf{y}_0|| \leq (1+\epsilon)^{-m}$, and furthermore if $m = \infty$, then $(\mathbf{x}_k)_{k\in\mathbb{Z}} = (\mathbf{y}_k)_{k\in\mathbb{Z}}$.*

Proof. For an arbitrary $k \in \{-m,\ldots,m-1\}$,

$$\mathbf{x}_{k+1} - \mathbf{y}_{k+1} = \mathbf{f}_k(\mathbf{x}_k) - \mathbf{f}_k(\mathbf{y}_k) = L_k\,(\mathbf{x}_k - \mathbf{y}_k) \tag{3.26}$$

for some $L_k \in [L_k]$. Thus the sequence $(\mathbf{h}_k)_{k\in\mathbb{Z}}$ defined by $\mathbf{h}_k = \frac{1}{2}(\mathbf{x}_k - \mathbf{y}_k)$ satisfies $\mathbf{h}_{k+1} = L_k\,\mathbf{h}_k$ and $||\mathbf{h}_k|| \leq 1$. As the interval matrices $[L_k]$ are ϵ-$(\mathcal{S},\mathcal{U})$ canonically hyperbolic for $\epsilon > 0$, so are the real matrices L_k, and we can apply Proposition 3.5, thus proving that $||\mathbf{h}_0|| \leq (1+\epsilon)^{-m}$. Now if $m = \infty$, we define instead $\mathbf{h}_k = \frac{1}{2}(\mathbf{x}_{k+c} - \mathbf{y}_{k+c})$ for an arbitrary $c \in \mathbb{Z}$, and prove similarly that $||\mathbf{h}_0|| \leq (1+\epsilon)^{-K}$ for an arbitrary $K \in \mathbb{N}$. Thus actually $\mathbf{h}_0 = \mathbf{0}$, which proves $(\mathbf{x}_k)_{k\in\mathbb{Z}} = (\mathbf{y}_k)_{k\in\mathbb{Z}}$. ∎

Example 3. Continuing Example 2, we obtain the following LIM computing the natural extension of the derivative of $\mathbf{f}$ (the absolute value is differentiated as explained in section 2.3):

$$[A] := \begin{pmatrix} [-0.52, 0.52] & 0.1 \\ [-0.5, 0.5] & 2 \end{pmatrix}. \tag{3.27}$$

The interval matrix [A] is 0.3-$(\mathcal{S},\mathcal{U})$ canonically hyperbolic. Therefore Theorem 3.6 (with $[A_k] = [A]$ and $\mathbf{f}_k = \mathbf{f}$) proves the uniqueness of the shadow whose existence was proved in Example 2. Note that since the dynamical system is autonomous (i.e., $\mathbf{f}_k = \mathbf{f}$ for all $k \in \mathbb{Z}$) the uniqueness of the shadow implies that this shadow is constant. Indeed, if it were not constant, then $(\mathbf{y}_k)_{k\in\mathbb{Z}}$ defined by $\mathbf{y}_k = \mathbf{f}(\mathbf{x}_k)$ would be a shadow different from $(\mathbf{x}_k)_{k\in\mathbb{Z}}$, which is impossible. Hence, we have finally proved that $\mathbf{f}$ has a unique fixed point in $B(\mathbf{0}, 1)$. And indeed, it can be verified formally that $\mathbf{f}$ actually has two fixed points which are approximately $(49.515, 495.248)$ and $(0.169, 0.0388)$.

3.3. Application to general dynamical systems.

The LICP is intended to be used for the pseudotrajectories $(\dots, \mathbf{0}, \mathbf{0}, \mathbf{0}, \dots)$ and unstable and stable directions close to axes. Let us now detail how it can be used together with an affine change of variables in the case where the pseudotrajectory is $(\tilde{\mathbf{x}}_i)_{i\in\mathbb{Z}}$, $\tilde{\mathbf{x}}_i \in \mathbb{R}^n$, and $(A_i)_{i\in\mathbb{Z}}$, $A_i \in \mathbb{R}^{n\times n}$, is a sequence of matrices whose columns are vectors approximating the unstable and stable directions at $\tilde{\mathbf{x}}_i$. Theorems 3.3 and 3.6 are now extended to such general orbits. Without loss of generality, we consider an autonomous dynamical system $\mathbf{f} : \mathbb{R}^n \longrightarrow \mathbb{R}^n$.

The local inductive containment property is generalized to arbitrary pseudo-orbits with local stable and unstable directions nonparallel to axes by using an affine change of variables. This is done rigorously using parallelotopes centered on the pseudo-orbit vectors. A parallelotope is the image of $B(\mathbf{0}, 1)$ by an affine transformation. The parallelotope $\{A\mathbf{u} + \tilde{\mathbf{x}} : \mathbf{u} \in B(\mathbf{0}, 1)\}$ with $A \in \mathbb{R}^{n\times n}$ and $\tilde{\mathbf{x}} \in \mathbb{R}^n$ is denoted by $\langle A, \tilde{\mathbf{x}}\rangle$. A parallelotope is said to be nonsingular if its characteristic is nonsingular. Then the LICP is generalized as follows.

Definition 3.7. *Let $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^n$ be a continuous map. Let $\mathcal{S}$ and $\mathcal{U}$ form a partition of $\{1, \dots, n\}$. Then we say that $\mathbf{f}$ satisfies the $(\mathcal{S},\mathcal{U})$-ICP from the nonsingular parallelotope $\langle A, \tilde{\mathbf{x}}\rangle$ to the nonsingular parallelotope $\langle B, \tilde{\mathbf{y}}\rangle$ if the function $\tilde{\mathbf{f}} : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ defined by*

$$\tilde{\mathbf{f}}(\mathbf{u}) := B^{-1}\big(\mathbf{f}(A\mathbf{u} + \tilde{\mathbf{x}}) - \tilde{\mathbf{y}}\big) \tag{3.28}$$

satisfies the $(\mathcal{S},\mathcal{U})$-LICP. Furthermore, we say that $\mathbf{f}$ satisfies the $(\mathcal{S},\mathcal{U})$-ICP for a sequence of parallelotopes $(\langle A_k, \tilde{\mathbf{x}}_k\rangle)_{k\in\mathbb{Z}}$ if it satisfies the $(\mathcal{S},\mathcal{U})$-ICP from $\langle A_k, \tilde{\mathbf{x}}_k\rangle$ to $\langle A_{k+1}, \tilde{\mathbf{x}}_{k+1}\rangle$ for all $k \in \mathbb{Z}$.

This definition can be interpreted directly in terms of parallelotopes, roughly speaking requiring that the image of the upper-side of the parallelotope $\langle A, \tilde{\mathbf{x}}\rangle$ be above the upper side of the parallelotope $\langle A, \tilde{\mathbf{x}}\rangle$, and so on. Such a formulation of Definition 3.7 was proposed in [41]. Then Theorem 3.3 is generalized to Corollary 3.8, which allows handling general orbits with stable and unstable directions that are not aligned axes. The hypotheses of Corollary 3.8 are similar to those of Theorem 1 of [41] but expressed in a simpler way using the LICP. Corollary 3.8 also generalizes Theorem 1 of [41] to continuous maps that are not diffeomorphisms and to bi-infinite pseudo-orbits.[4] We say that a sequence $(\mathbf{x}_i)_{i\in\mathbb{Z}}$ hits a sequence of sets $(E_i)_{i\in\mathbb{Z}}$ if and only if $\mathbf{x}_i \in E_i$ holds for all $i \in \mathbb{Z}$.

Corollary 3.8. *Let $\mathcal{S}$ and $\mathcal{U}$ be a partition of $\{1, \dots, n\}$. Suppose that $\mathbf{f}$ satisfies the $(\mathcal{S},\mathcal{U})$-ICP for a sequence of nonsingular parallelotopes $(\langle A_k, \tilde{\mathbf{x}}_k\rangle)_{k\in\mathbb{Z}}$. Then there exists a bi-infinite orbit $(\mathbf{x}_k)_{k\in\mathbb{Z}}$, i.e., $\mathbf{x}_{k+1} = \mathbf{f}(\mathbf{x}_k)$, that hits $(\langle A_k, \tilde{\mathbf{x}}_k\rangle)_{k\in\mathbb{Z}}$.*

Proof. Define for $k \in \mathbb{Z}$

$$\tilde{\mathbf{f}}_k(\mathbf{u}) := A_{k+1}^{-1}\big(\mathbf{f}(A_k\mathbf{u} + \mathbf{x}_k) - \mathbf{x}_{k+1}\big). \tag{3.29}$$

[4] For simplicity, Corollary 3.8 is stated for bi-infinite orbits only but obviously holds for infinite and finite orbits.

By definition of the $(\mathcal{S},\mathcal{U})$-ICP every $\tilde{\mathbf{f}}_k$ satisfies the $(\mathcal{S},\mathcal{U})$-LICP. Thus Theorem 3.3 proves the existence of $(\mathbf{u}_k)_{k\in\mathbb{Z}}$ such that $\mathbf{u}_{k+1} = \tilde{\mathbf{f}}_k(\mathbf{u}_k)$ and $\mathbf{u}_k \in B(\mathbf{0}, 1)$. Finally, it is easy to check that $\mathbf{x}_k := A_k\mathbf{u}_k + \tilde{\mathbf{x}}_k$ is the wanted orbit. ■

To rigorously verify that the LICP holds for (3.28), we need just to be able to evaluate an interval enclosure of $\tilde{\mathbf{f}}$ for some box $[\mathbf{u}] \subseteq B(\mathbf{0}, 1)$ (this box being either $B(\mathbf{0}, 1)$ when hypothesis (LICP1) is considered or a side of $B(\mathbf{0}, 1)$ when hypothesis (LICP2) is considered). Evaluating (3.28) using the naive interval evaluation

$$B^{-1}\big([\mathbf{f}](A[\mathbf{u}] + \tilde{\mathbf{x}}) - \tilde{\mathbf{y}}\big), \tag{3.30}$$

where $[\mathbf{f}]$ is, for example, the natural interval extension of $\mathbf{f}$, generally gives rise to very pessimistic enclosures that are useless. As usually done when such a change of basis is involved, we use here the centered interval extension of $\tilde{\mathbf{f}}$ with expansion point $\mathrm{mid}B(\mathbf{0}, 1) = \mathbf{0}$: For a box $[\mathbf{u}] \subseteq B(\mathbf{0}, 1)$ we have

$$\tilde{\mathbf{f}}([\mathbf{u}]) \subseteq [\tilde{\mathbf{f}}](\mathbf{0}) + [\tilde{L}]\,[\mathbf{u}], \tag{3.31}$$

where $[\tilde{\mathbf{f}}]$ is an interval extension of $\tilde{\mathbf{f}}$ and $[\tilde{L}]$ is a LIM for $\tilde{\mathbf{f}}$ and $\langle A_k, \mathbf{x}_k\rangle$. This gives rise to the following enclosure:

$$\tilde{\mathbf{f}}([\mathbf{u}]) \;\subseteq\; A_{k+1}^{-1}\big([\mathbf{f}](\tilde{\mathbf{x}}_k) - \tilde{\mathbf{x}}_{k+1}\big) + \big(A_{k+1}^{-1}[L]A_k\big)[\mathbf{u}], \tag{3.32}$$

where $[\mathbf{f}]$ is an interval extension of $\mathbf{f}$ (the natural extension can be used here) and $[L]$ is a LIM for $\mathbf{f}$ and $\square\langle A_k, \tilde{\mathbf{x}}_k\rangle = \tilde{\mathbf{x}}_k + \tilde{A}_k[\mathbf{u}]$. Note that when evaluating the product $\big(A_{k+1}^{-1}[L]A_k\big)[\mathbf{u}]$, it is important to evaluate the matrix/matrix products first and then the matrix/vector product. Indeed, the alternative evaluation $(A_{k+1}^{-1}([L](A_k[\mathbf{u}])))$ is much more pessimistic in general. Finally, in order to evaluate (3.32) rigorously, we need to compute a rigorous enclosure of the inverse of a real matrix. For low dimensional matrices, this can be done by evaluating the formal expression of the matrix inverse using IA. More generally, the method proposed in [30] can be used.

We end this section with the following corollary of Theorem 3.6, which generalizes this theorem to general orbits with stable and unstable directions that are not aligned axes.

Corollary 3.9. *Let $\mathcal{S}$ and $\mathcal{U}$ be a partition of $\{1, \ldots, n\}$. Consider two orbits $(\mathbf{x}_k)_{k\in E}$ and $(\mathbf{y}_k)_{k\in E}$, where $E = \{k \in \mathbb{Z} : |k| \leq m\}$, that hit the sequence of nonsingular parallelotopes $(\langle A_k, \tilde{\mathbf{x}}_k\rangle)_{k\in E}$. Suppose that $[L_k]$ are LIMs for $\mathbf{f}$ and $\langle A_k, \tilde{\mathbf{x}}_k\rangle$ such that $A_{k+1}^{-1}[L_k]A_k$ is ϵ-$(\mathcal{S},\mathcal{U})$ canonically hyperbolic for $\epsilon > 0$. Then*

$$||\mathbf{x}_0 - \mathbf{y}_0|| \leq ||A_0||(1 + \epsilon)^{-m}, \tag{3.33}$$

and thus if $m = \infty$, we have $(\mathbf{x}_k)_{k\in\mathbb{Z}} = (\mathbf{y}_k)_{k\in\mathbb{Z}}$.

Proof. Define $\mathbf{u}_k = A_k^{-1}(\mathbf{x}_k - \tilde{\mathbf{x}}_k)$ and $\mathbf{v}_k = A_k^{-1}(\mathbf{y}_k - \tilde{\mathbf{x}}_k)$ so that $(\mathbf{u}_k)_{k\in E}$ and $(\mathbf{v}_k)_{k\in E}$ are orbits of $\tilde{\mathbf{f}}_k(\mathbf{u}) := A_{k+1}^{-1}\big(\mathbf{f}(A_k\mathbf{u} + \tilde{\mathbf{x}}_k) - \tilde{\mathbf{x}}_{k+1}\big)$ bounded in $B(\mathbf{0}, 1)$. Note that $A_{k+1}^{-1}[L_k]A_k$ is a

LIM for $\mathbf{f}_k$ and $B(\mathbf{0},1)$, so we can apply Theorem 3.6, which proves that $||\mathbf{u}_0-\mathbf{v}_0|| \leq (1+\epsilon)^{-m}$, thus implying $||\mathbf{x}_0-\mathbf{y}_0|| \leq ||A_0||(1+\epsilon)^{-m}$. Theorem 3.6 also proves that if $m=\infty$, then $(\mathbf{u}_k)_{k\in\mathbb{Z}} = (\mathbf{v}_k)_{k\in\mathbb{Z}}$, which is equivalent to $(\mathbf{x}_k)_{k\in\mathbb{Z}} = (\mathbf{y}_k)_{k\in\mathbb{Z}}$. ■

4. Chaotic dynamical systems. In this section, we consider a continuous function $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^n$ and two length p sequences of parallelotopes $(\langle A_k, \tilde{\mathbf{x}}_k\rangle)_{0\leq k\leq p-1}$ and $(\langle B_k, \tilde{\mathbf{y}}_k\rangle)_{0\leq k\leq p-1}$. Also we consider $\mathcal{S}$ and $\mathcal{U}$ forming a partition of $\{1,\dots,n\}$. The theorems proposed below are based on the following hypotheses.

Hypothesis 1. *$\tilde{\mathbf{x}}_0 = \tilde{\mathbf{y}}_0$, $A_0 = B_0$.*

Hypothesis 2. *$\langle A_{k^*}, \tilde{\mathbf{x}}_{k^*}\rangle \cap \langle B_{k^*}, \tilde{\mathbf{y}}_{k^*}\rangle = \emptyset$ for some $k^* \in \{1,\dots,p-1\}$.*

We denote $\min\{d(\mathbf{x},\mathbf{y}) : \mathbf{x} \in \langle A_{k^*}, \tilde{\mathbf{x}}_{k^*}\rangle, \mathbf{y} \in \langle B_{k^*}, \tilde{\mathbf{y}}_{k^*}\rangle\}$ by ϵ^*, so $\epsilon^* > 0$.

Hypothesis 3. *$\mathbf{f}$ satisfies the $\{\mathcal{S},\mathcal{U}\}$-ICP*

1. *from $\langle A_k, \tilde{\mathbf{x}}_k\rangle$ to $\langle A_{k+1}, \tilde{\mathbf{x}}_{k+1}\rangle$ for $k \in \{0,\dots,p-2\}$,*
2. *from $\langle A_{p-1}, \tilde{\mathbf{x}}_{p-1}\rangle$ to $\langle A_0, \tilde{\mathbf{x}}_0\rangle$,*
3. *from $\langle B_k, \tilde{\mathbf{y}}_k\rangle$ to $\langle B_{k+1}, \tilde{\mathbf{y}}_{k+1}\rangle$ for $k \in \{0,\dots,p-2\}$,*
4. *from $\langle B_{p-1}, \tilde{\mathbf{y}}_{p-1}\rangle$ to $\langle B_0, \tilde{\mathbf{y}}_0\rangle$.*

The sequences $(\tilde{\mathbf{x}}_k)_{k\in\{0,\dots,p-1\}}$ and $(\tilde{\mathbf{y}}_k)_{k\in\{0,\dots,p-1\}}$ start with a common element, then diverge enough so that their enclosing parallelotopes are disjoint, and finally converge back to the initial common element. We call them branching periodic pseudo-orbits. Such pairs of pseudo-orbits were used, for example, in [35] to prove that some dynamical systems are chaotic. A simple but typical situation satisfying these hypotheses is shown in Figure 4. The depicted relationships between domains and images are similar to the Smale horseshoe but more general: The first step stretches the initial left-hand-side box, and its second step takes two different parts of the stretched image back to the initial box. In fact, the Smale horseshoe does satisfy the above hypotheses, as illustrated in Figure 4.

A fourth hypothesis is considered: For $k \in \{1,\dots,p-1\}$, let $[L_k^A]$ (respectively, $[L_k^B]$) be a LIM for $\mathbf{f}$ and $\langle A_k, \tilde{\mathbf{x}}_k\rangle$ (respectively, $\langle B_k, \tilde{\mathbf{y}}_k\rangle$). Then we introduce the following hypothesis.

Hypothesis 4. *For some $\epsilon > 0$ and a partition $\mathcal{S},\mathcal{U}$, the interval matrices $A_{k+1}^{-1}[L_k^A]A_k$ and $B_{k+1}^{-1}[L_k^B]B_k$ for $k \in \{0,\dots,p-2\}$, and $A_0^{-1}[L_{p-1}^A]A_{p-1}$ and $B_0^{-1}[L_{p-1}^B]B_{p-1}$, are ϵ-$\{\mathcal{S},\mathcal{U}\}$ canonically hyperbolic for some $\epsilon > 0$.*

The two theorems presented in the rest of the section use the above hypotheses to prove that a system is chaotic in various senses. The first theorem shows that under these hypotheses, the map $\mathbf{f}^p = \mathbf{f}\circ\cdots\circ\mathbf{f}$ is chaotic in the sense of Li and Yorke and in the sense of Devaney (see the next subsection for details). The second theorem provides a more accurate description of the chaotic behavior of the system: Under the additional hypothesis that $\mathbf{f}$ is injective, it allows us to show that the map $\mathbf{f}^p$ admits a subsystem that is topologically conjugate to the Bernoulli shift.

4.1. The full shift on two symbols and chaos. Let $\Sigma = \{0,1\}^{\mathbb{Z}}$ be the set of bi-infinite sequences of 0 and 1. This set is made a compact metric space using the distance $d(\alpha,\beta) := \max\{2^{-|k|} : k \in \mathbb{Z}, \alpha_k \neq \beta_k\}$. Note that two sequences $\alpha,\beta \in \Sigma$ satisfy $\alpha_k = \beta_k$ for all k such that $|k| \leq n$ if and only if $d(\alpha,\beta) \leq 2^{-n}$. Finally, the topological shift map $\sigma : \Sigma \to \Sigma$ is defined by $\sigma(\alpha)_k = \alpha_{k+1}$. The shift map is chaotic following most definitions of chaos.

For a dynamical system $f : X \longrightarrow X$, where X is compact, the topological entropy is defined as follows: First, a set $E \subseteq X$ is called (n,ϵ)-separated if for all $\mathbf{x},\mathbf{y} \in E$, $\mathbf{x} \neq \mathbf{y}$, there

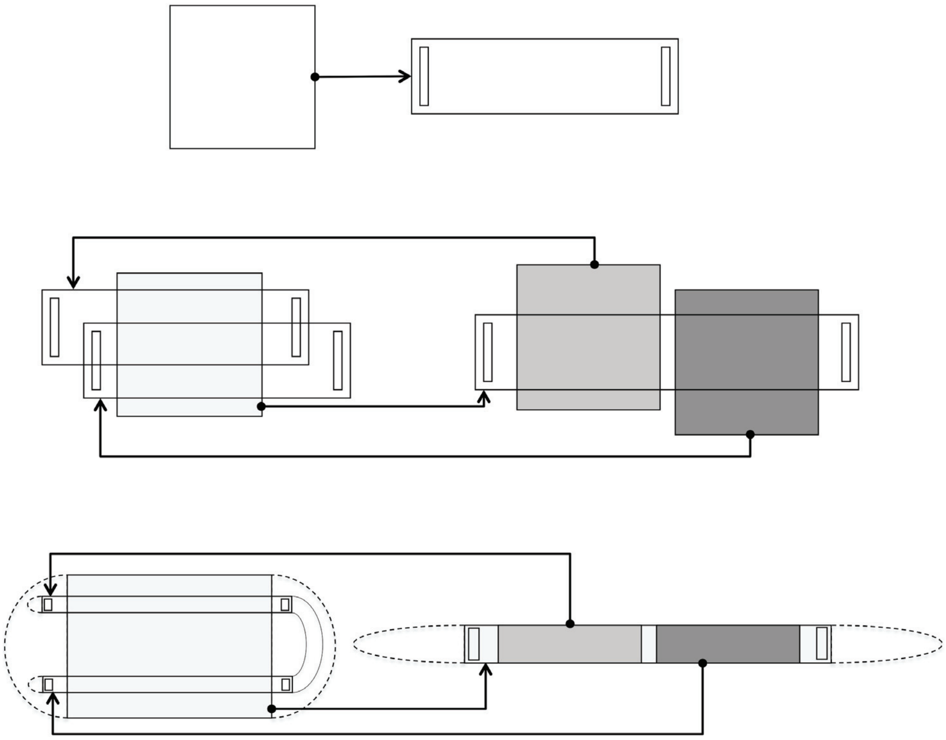

Figure 4. *The upper graphic shows a diagrammatic representation of the ICP: The image of the square on the left is enclosed by the large rectangle on the right, while the images of the left and right sides of the square are enclosed by the smaller interior rectangles on the right. The central graphic shows a typical situation that satisfies Hypotheses 1, 2, and 3 with $p = 2$: From lighter to darker gray, $\langle A_0, \tilde{\mathbf{x}}_0\rangle = \langle B_0, \tilde{\mathbf{y}}_0\rangle$, $\langle A_1, \tilde{\mathbf{x}}_1\rangle$, and $\langle B_1, \tilde{\mathbf{y}}_1\rangle$ (stable and unstable directions are parallel to the axis for simplicity). The lower graphic shows that the Smale horseshoe map actually satisfies the constraints required by the central graphic.*

exists $k \leq n$ such that $d(f^k(\mathbf{x}), f^k(\mathbf{y})) \geq \epsilon$. The maximum cardinality of an (n, ϵ)-separated is denoted by $N(n, \epsilon)$ (and is finite since $E \subseteq X$ is bounded). Then, the topological entropy $h_{\text{top}}(f)$ is defined by

$$h_{\text{top}}(f) = \lim_{\epsilon \to 0} \left(\limsup_{n \to \infty} \frac{1}{n} \log_2 N_f(n, \epsilon) \right). \tag{4.1}$$

The shift map has a strictly positive topological entropy, which implies Li–Yorke chaos (see [5]). The shift map is also Devaney chaotic, as it is sensitive (there exists $\epsilon > 0$ such that any $x \in X$ is a limit point for points $y \in X$ satisfying the condition $d(f^n(x), f^n(y)) \geq \epsilon$ for some positive n) and topologically mixing (for any pair of open sets $U, V \subseteq X$, there exists $k \geq 1$ such that $T^n(V) \cap U \neq \emptyset$ for all $n \geq k$) and the set $\text{per}(\sigma)$ of its periodic points (i.e., $\text{per}(f) := \{x \in X : \exists p > 0, f^p(x) = x\}$) is dense inside Σ. Details can be found in [19, 23] and the references therein.

4.2. Chaotic noninjective dynamical systems. Since the map $\mathbf{f}$ is not injective, we will not be able to make the usual use of topological factor or conjugacy with the shift map. Instead, we will use the following proposition.

Proposition 4.1. *Let F and G be metric spaces, and consider three maps $h : G \to F$, $g : G \to G$, and $f : F \to F$ satisfying h continuous and surjective, and $f \circ h = h \circ g$ (that is, f is a topological factor of g).*[5] *Then the following implications hold:*

- $\mathrm{per}(g)$ *dense in G implies* $\mathrm{per}(f)$ *dense in F.*
- *g topologically mixing implies f topologically mixing.*

Proof. As h is continuous, we have for all $x \in G$

$$\forall \epsilon > 0, \exists \delta > 0, h(B(x, \delta)) \subseteq B(h(x), \epsilon). \tag{4.2}$$

Let us prove the first implication. Consider an arbitrary $y \in F$. As h is surjective, there exists $x \in G$ such that $y = h(x)$. By assumption, $\mathrm{per}(g)$ is dense in G, so

$$\forall \delta > 0, \mathrm{per}(g) \cap B(x, \delta) \neq \emptyset. \tag{4.3}$$

As obviously $h(\mathrm{per}(g)) \subseteq \mathrm{per}(f)$, (4.3) implies for all $\delta > 0, \mathrm{per}(f) \cap h(B(x, \delta)) \neq \emptyset$. Finally, considering (4.2) we obtain for all $\epsilon > 0, \mathrm{per}(f) \cap B(h(x), \epsilon) \neq \emptyset$.

Let us prove the second implication. Suppose that the function g is mixing; that is, for all $x, x' \in G$ and all $\delta, \delta' > 0$ we have

$$\exists N > 0, \forall n \geq n,\ g^n(B(x, \delta)) \cap g^n(B(x', \delta')) \neq \emptyset. \tag{4.4}$$

Now, $g^n(B(x, \delta)) \cap g^n(B(x', \delta')) \neq \emptyset$ implies $h(g^n(B(x, \delta))) \cap h(g^n(B(x', \delta'))) \neq \emptyset$, which implies $f^n(h(B(x, \delta))) \cap f^n(h(B(x', \delta'))) \neq \emptyset$ since $f \circ h = h \circ g$. So we have proved that for all $x, x' \in G$ and all $\delta, \delta' > 0$ we have

$$\exists N > 0, \forall n \geq N,\ f^n(h(B(x, \delta))) \cap f^n(h(B(x', \delta'))) \neq \emptyset. \tag{4.5}$$

Together with (4.2) this proves that for all $x, x' \in G$ and all $\epsilon, \epsilon' > 0$ we have

$$\exists N > 0, \forall n \geq N,\ f^n(B(h(x), \epsilon)) \cap f^n(B(h(x'), \epsilon')) \neq \emptyset. \tag{4.6}$$

As h is surjective, the statement actually holds for all $y, y' \in F$ with $y = h(x)$ and $y' = h(x')$. ■

We can now prove the following theorem.

Theorem 4.2. *Assuming Hypotheses 1, 2, and 3, there exists a bounded set Λ such that $\mathbf{f}^p(\Lambda) \subseteq \Lambda$ and $h_{\mathrm{top}}(\mathbf{f}^p|_\Lambda)$ is strictly positive. If, furthermore, Hypothesis 4 holds, then $\mathbf{f}^p|_\Lambda$ is sensitive and mixing and its periodic points are dense in Λ.*

Proof. For every $\alpha \in \Sigma$, consider the bi-infinite sequence of parallelotopes $(\langle C_k^\alpha, \tilde{\mathbf{z}}_k^\alpha \rangle)_{k \in \mathbb{Z}}$ defined by

- $C_k^\alpha = A_{k \,\mathrm{mod} p}$ and $\tilde{\mathbf{z}}_k^\alpha = \tilde{\mathbf{x}}_{k \,\mathrm{mod} p}$ if $\alpha_{k \div p} = 0$,
- $C_k^\alpha = B_{k \,\mathrm{mod} p}$ and $\tilde{\mathbf{z}}_k^\alpha = \tilde{\mathbf{y}}_{k \,\mathrm{mod} p}$ if $\alpha_{k \div p} = 1$.

This corresponds to building a bi-infinite pseudo-orbit $(\tilde{\mathbf{z}}_k)_{k \in \mathbb{Z}}$ by switching between the pseudo-orbit $(\tilde{\mathbf{x}}_k)_{k \in \{1,\dots,p\}}$ and the pseudo-orbit $(\tilde{\mathbf{y}}_k)_{k \in \{1,\dots,p\}}$ depending on α_k. Then Hypotheses 1 and 3 obviously imply that for every $k \in \mathbb{Z}$ the function $\mathbf{f}$ satisfies the $\{\mathcal{S}, \mathcal{U}\}$-ICP

[5] Here, topological properties are inferred from g to its factor f, while the other direction is usually used (e.g., a well-known property is that if the factor f has a positive topological entropy, then so does g).

from $\langle C_k^\alpha, \tilde{\mathbf{z}}_k^\alpha\rangle$ to $\langle C_{k+1}^\alpha, \tilde{\mathbf{z}}_{k+1}^\alpha\rangle$, so we can apply Corollary 3.8 to prove that there exists a bi-infinite orbit $(\mathbf{z}_k)_{k\in\mathbb{Z}}$ that hits $(\langle C_k^\alpha, \tilde{\mathbf{z}}_k^\alpha\rangle)_{k\in\mathbb{Z}}$. We denote this orbit by $(\mathbf{z}_k^\alpha)_{k\in\mathbb{N}}$ and define $h : \Sigma\times\mathbb{Z}\to\mathbb{R}^n$ by $h(\alpha,k) = \mathbf{z}_k^\alpha$. Finally, we define

$$\Lambda := h(\Sigma, k^* + p\mathbb{Z}), \tag{4.7}$$

where k^* is defined in Hypothesis 2. The elements of Λ are thus $\mathbf{z}_{k^*+ip}^\alpha$ for some $\alpha\in\Sigma$ and $i\in\mathbb{Z}$, and we have $\mathbf{z}_{k^*+ip}^\alpha \in \langle A_{k^*}, \tilde{\mathbf{x}}_{k^*}\rangle$ if $\alpha_i = 0$ or $\mathbf{z}_{k^*+ip}^\alpha \in \langle B_{k^*}, \tilde{\mathbf{y}}_{k^*}\rangle$ if $\alpha_i = 1$. Therefore Λ is bounded since it is a subset of $\langle A_{k^*}, \tilde{\mathbf{x}}_{k^*}\rangle \cup \langle B_{k^*}, \tilde{\mathbf{y}}_{k^*}\rangle$. Finally note that $\mathbf{f}^p(\mathbf{z}_{k^*+ip}^\alpha) = \mathbf{z}_{k^*+(i+1)p}^\alpha \in \Lambda$, so $\mathbf{f}^p(\Lambda)\subseteq\Lambda$.

Let us prove that $h_{\text{top}}(\mathbf{f}^p|_\Lambda) \geq 1$. Let Σ_q be the set of q-periodic elements of Σ; thus $\operatorname{card}\Sigma_q = 2^q$. Then define $E_q = h(\Sigma_q, k^*)\subseteq\Lambda$. We are going to prove that E_q is (q,ϵ^*)-separated and that $h(\cdot,k^*)$ is injective inside Σ_q (so $\operatorname{card} E_q = 2^q$), which will imply $h_{\text{top}}(\mathbf{f}^p|_\Lambda)\geq 1$ by definition of the topological entropy. Consider $\mathbf{z}_{k^*}^\alpha, \mathbf{z}_{k^*}^\beta \in E_q$ with $\alpha\neq\beta$. As α and β are q-periodic and different, there exists $i\in\{0,1,\dots,q-1\}$ such that $\alpha_i\neq\beta_i$. Define $k^+ := k^*+ip$, and note that by Hypothesis 2 the minimal distance between $\langle C_{k^+}^\alpha, \tilde{\mathbf{z}}_{k^+}^\alpha\rangle$ and $\langle C_{k^+}^\beta, \tilde{\mathbf{z}}_{k^+}^\beta\rangle$ is ϵ^*. Finally, since $(\mathbf{f}^p)^i(\mathbf{z}_{k^*}^\alpha) = \mathbf{z}_{k^+}^\alpha \in \langle C_{k^+}^\alpha, \tilde{\mathbf{z}}_{k^+}^\alpha\rangle$ and $(\mathbf{f}^p)^i(\mathbf{z}_{k^*}^\beta) = \mathbf{z}_{k^+}^\beta \in \langle C_{k^+}^\beta, \tilde{\mathbf{z}}_{k^+}^\beta\rangle$, we have proved that $d((\mathbf{f}^p)^i(\mathbf{z}_{k^*}^\alpha), (\mathbf{f}^p)^i(\mathbf{z}_{k^*}^\beta))\geq\epsilon^*$ for some $0\leq i<q$ and therefore that the set E_q is (q,ϵ^*)-separated. This also proves that $\mathbf{z}_{k^*}^\alpha\neq\mathbf{z}_{k^*}^\beta$ (since their images through $(\mathbf{f}^p)^i$ are different) and therefore that $h(\cdot,k^*)$ is injective inside E_q.

Now suppose that Hypothesis 4 holds. Thus by Corollary 3.9, $(\mathbf{z}_k^\alpha)_{k\in\mathbb{N}}$ is the unique orbit that hits $(\langle C_k^\alpha, \tilde{\mathbf{z}}_k^\alpha\rangle)_{k\in\mathbb{Z}}$. The sequence $(\mathbf{f}^p(\mathbf{z}_k^\alpha))_{k\in\mathbb{N}}$ is equal to $(\mathbf{z}_{k+p}^\alpha)_{k\in\mathbb{N}}$ and hits $(\langle C_{k+p}^\alpha, \tilde{\mathbf{z}}_{k+p}^\alpha\rangle)_{k\in\mathbb{Z}}$, the latter being equal to $(\langle C^{\sigma(\alpha)_k}, \tilde{\mathbf{z}}^{\sigma(\alpha)_k}\rangle)_{k\in\mathbb{Z}}$. As there is only one orbit that hits this latter sequence of parallelepipeds, this proves that $(\mathbf{f}^p(\mathbf{z}_k^\alpha))_{k\in\mathbb{N}}$ is equal to $(\mathbf{z}_k^{\sigma(\alpha)})_{k\in\mathbb{N}}$. As a consequence, $(\mathbf{f}|_\Lambda)^p\circ h = h\circ\sigma$, where $h:\Sigma\to\langle A_{k^*}, \tilde{\mathbf{x}}_{k^*}\rangle\cup\langle B_{k^*}, \tilde{\mathbf{y}}_{k^*}\rangle$ is defined by $h(\alpha) := h(\alpha,k^*)$. This implies in particular that

$$(\mathbf{f}|_\Lambda)^{kp}\circ h = h\circ\sigma^k \tag{4.8}$$

and $h(\alpha, k^*+pk) = h(\sigma^k(\alpha), k^*)$, which proves that $\Lambda = h(\Sigma)$.

We proceed by proving that h is continuous inside Λ. Let us consider a sequence $(\alpha_k)_{k\in\mathbb{N}}$, $\alpha_k\in\Sigma$, that converges to $\beta\in\Sigma$. We just have to prove that $(h(\alpha_k))_{k\in\mathbb{N}}$ converges to $h(\beta)$. So for all $i\in\mathbb{N}$, there exists $K\in\mathbb{N}$, where $k\geq K$ implies $\alpha_{k,j}=\beta_j$ for all $j\in\{-i,\dots,i\}$. Therefore $(\mathbf{z}_j^{\alpha_k})_{j\in P}$, $P=\{-pi,\dots,pi\}$, and $(\mathbf{z}_j^\beta)_{j\in P}$ are orbits of $\mathbf{f}$ that both hit $(\langle C_k^\alpha, \tilde{\mathbf{z}}_k^\alpha\rangle)_{k\in P}$. In this situation, by Hypothesis 4 we can use Corollary 3.9, which proves that $||\mathbf{z}_0^{\alpha_k} - \mathbf{z}_0^\beta|| \leq ||A_0||(1+\epsilon)^{-pi}$, and thus $(\mathbf{z}_0^{\alpha_k})_{k\in\mathbb{N}}$ converges to $\mathbf{z}_0^\beta$. Finally, since $\mathbf{f}$ is continuous, $(\mathbf{z}_{k^*}^{\alpha_k})_{k\in\mathbb{N}}$ converges to $\mathbf{z}_{k^*}^\beta$, which actually means $(h(\alpha_k))_{k\in\mathbb{N}}$ converges to $h(\beta)$.

Since h is continuous and surjective, $\mathbf{f}^p|_\Lambda$ is a topological factor of σ. So we can apply Proposition 4.1 and prove that $\mathbf{f}^p|_\Lambda$ is both mixing and that its periodic points are dense in Λ.

It remains to prove that $\mathbf{f}^p|_\Lambda$ is sensitive. Consider an arbitrary $\mathbf{x}\in\Lambda$, so there exists $\beta\in\Sigma$ such that $\mathbf{x}=h(\beta)$. Define the sequence $(\alpha_k)_{k\in\mathbb{N}}$, $\alpha_k\in\Sigma$, by $\alpha_{ki}=\beta_i$ for $i\neq k$ and $\alpha_{kk} = 1-\beta_k$ (so $\sigma^k(\beta)_0\neq\sigma^k(\alpha_k)_0$). This sequence obviously converges to β. Furthermore, since h is continuous, $\mathbf{y}_k := h(\alpha_k)$ converges to $\mathbf{x}$. Now using (4.8) we prove that $\mathbf{f}^{pk}(\mathbf{x}) =$

$\mathbf{f}^{pk}(h(\beta)) = h(\sigma^k(\beta))$ and $\mathbf{f}^{pk}(\mathbf{y}^k) = \mathbf{f}^{pk}(h(\alpha_k)) = h(\sigma^k(\alpha_k))$. Finally, since $\sigma^k(\beta)_0 \neq \sigma^k(\alpha_k)_0$, both $\mathbf{f}^{pk}(\mathbf{x})$ and $\mathbf{f}^{pk}(\mathbf{y}^k)$ must lie in different parallelotopes $\langle A_{k^*}, \tilde{\mathbf{x}}_{k^*}\rangle$ and $\langle B_{k^*}, \tilde{\mathbf{y}}_{k^*}\rangle$, which proves by Hypothesis 2 that $d(\mathbf{f}^{pk}(\mathbf{x}), \mathbf{f}^{pk}(\mathbf{y}^k)) \geq \epsilon^*$, hence finally showing that $\mathbf{f}^p|_\Lambda$ is sensitive. ■

4.3. Chaotic injective dynamical systems. Theorem 4.2 can be sharpened under the additional assumption that the map is injective. We obtain the following theorem, which allows showing that an injective map is topologically conjugate with the full shift on two symbols, thus obtaining the same conclusion as [35] but under hypotheses based on the ICP.

Theorem 4.3. *Assume that* $\mathbf{f}$ *is injective in addition to Hypotheses* 1, 2, 3, *and* 4. *Then there exists a set* Λ *such that* $\mathbf{f}^p(\Lambda) \subseteq \Lambda$ *and the restriction* $\mathbf{f}^p|_\Lambda$ *of* $\mathbf{f}^p$ *to* Λ *is topologically conjugate to* σ*; i.e., there exists a homeomorphism* $h : \Sigma \to \Lambda$ *such that* $\mathbf{f}^p|_\Lambda \circ h = \sigma \circ h$.

Proof. From the proof of Theorem 4.2, we know that $\mathbf{f}^p|_\Lambda$ is a topological factor of σ. In order to prove that both maps are actually topologically conjugate, it remains just to prove that $h : \Sigma \longrightarrow \Lambda = h(\Sigma)$ is a homeomorphism. In fact, we prove just that h is injective, since a bijective map between compact metric spaces is a homeomorphism (see [8]). Consider $\alpha, \beta \in \Sigma$ such that $h(\alpha) = h(\beta)$, that is, $\mathbf{z}^\alpha_{k^*} = \mathbf{z}^\beta_{k^*}$. Since $\mathbf{f}$ is injective, this implies $(\mathbf{z}^\alpha_k)_{k\in\mathbb{Z}} = (\mathbf{z}^\beta_k)_{k\in\mathbb{Z}}$. Finally, by Hypothesis 2 this implies $\alpha = \beta$, and thus h is injective. ■

5. Numerical applications.

5.1. Computing branching periodic pseudo-orbits and their stable and unstable directions. In order to use Theorems 4.2 and 4.3 we need to find two sequences of parallelotopes $(\langle A_k, \tilde{\mathbf{x}}_k\rangle)_{0\leq k\leq p-1}$ and $(\langle B_k, \tilde{\mathbf{y}}_k\rangle)_{0\leq k\leq p-1}$ that satisfy Hypotheses 1, 2, 3, and 4. So that interval extensions are sharp and taking into account the simplicity of the system's expressions (there are only a few multiple occurrences of variables, which makes interval evaluation very efficient), we expect that the size δ of the parallelotopes can be around 10^{-3}, while we will need the distance between the branching periodic pseudo-orbits to be negligible with respect to this size, hence around 10^{-5}.

The first step is to find such branching periodic pseudo-orbits. The task is easily carried out by some random search since we expect the number of periodic points to grow exponentially with respect to the period. Let us denote these two pseudo-orbits by $(\tilde{\mathbf{x}}_k)_{k\in\{0,\dots,p-1\}}$ and $(\tilde{\mathbf{y}}_k)_{k\in\{0,\dots,p-1\}}$. Once these periodic pseudo-orbits are found, we need to compute approximations of their stable and unstable directions. In the two dimensional case, this is easily done by iterating forward the derivative in the tangent space for finding normed vectors approximating the unstable directions and backward the inverse of the derivative for finding normed vectors approximating the stable directions (for higher dimensions, a QR decomposition is usually additionally used to prevent all vectors from converging to the strongest unstable direction). The computed directions form the columns of the matrices $(A_k)_{k\in\{0,\dots,p-1\}}$ and $(B_k)_{k\in\{0,\dots,p-1\}}$ which are scaled so that the norm of each column is δ. Note that we expect that A_0 and B_0 will represent approximately the same unstable and stable directions. Finally, we change $\tilde{\mathbf{x}}_0$ and $\tilde{\mathbf{y}}_0$ to $0.5(\tilde{\mathbf{x}}_0 + \tilde{\mathbf{y}}_0)$, and A_0 and B_0 to $0.5(A_0 + B_0)$, so that $\langle A_0, \tilde{\mathbf{x}}_0\rangle = \langle B_0, \tilde{\mathbf{y}}_0\rangle$ and Hypothesis 1 is satisfied.

We have so far constructed two sequences of parallelotopes

$$(\langle A_k, \tilde{\mathbf{x}}_k\rangle)_{0\leq k\leq p-1} \quad \text{and} \quad (\langle B_k, \tilde{\mathbf{y}}_k\rangle)_{0\leq k\leq p-1}. \tag{5.1}$$

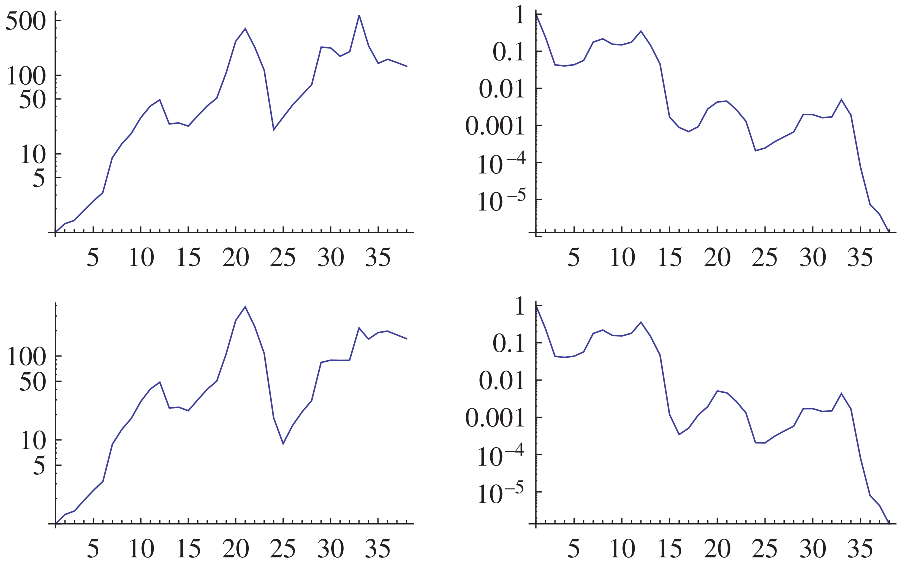

Figure 5. *Cumulative expansion (left) and contraction (right) for the two branching periodic pseudo-orbits (up and down) for the Tinkerbell map.*

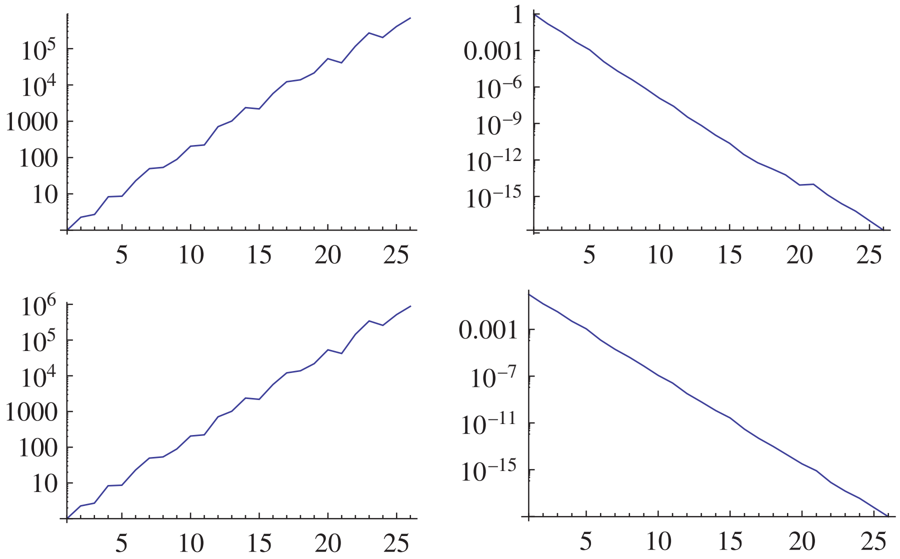

Figure 6. *Cumulative expansion (left) and contraction (right) for the two branching periodic pseudo-orbits (up and down) for the Hénon map.*

If the map were uniformly expanding and contracting, then Theorems 4.2 and 4.3 could be applied directly to them. However, Figures 5 and 6 show that the maps under study are not uniformly expanding and contracting. Therefore, we need to tune the norm of the columns of A_k and B_k so as to compensate for the lack of expansion and contraction, which can easily be done automatically.

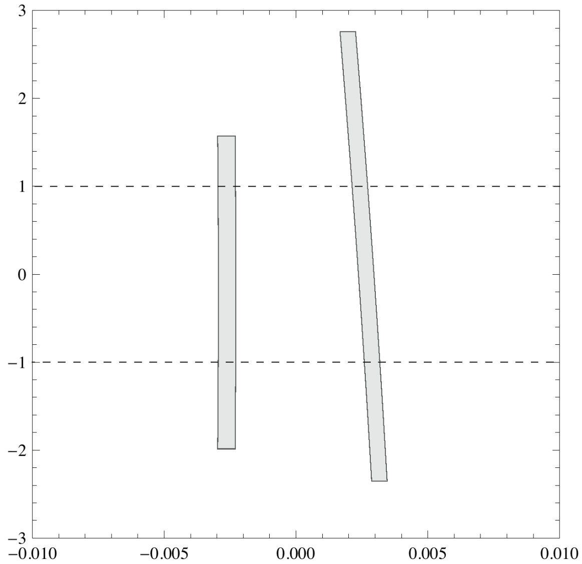

Figure 7. *Verification of the LICP for the last step of the Tinkerbell couple of branching periodic pseudo-orbits.*

5.2. The Tinkerbell map is chaotic.

The Tinkerbell map [29] is defined by

$$\mathbf{f}(\mathbf{x}) = \begin{pmatrix} x_1^2 + ax_1 - x_2^2 + bx_2 \\ cx_1 + 2x_1x_2 + dx_2 \end{pmatrix}. \tag{5.2}$$

We use the standard parameter values $a = 0.9$, $b = -0.6013$, $c = 2.$, and $d = 0.5$, which give rise to the strange attractor depicted in Figure 1. The following two vectors can be used to build branching periodic pseudo-orbits of period 37 and their unstable and stable directions:

$$\begin{pmatrix} -0.16148450580019486 \\ -0.43326600041803776 \end{pmatrix} ; \begin{pmatrix} -0.16144419627383078 \\ -0.433256643512554 \end{pmatrix}. \tag{5.3}$$

The parallelotopes resulting from the process described in the previous subsection can be found in 81901_01.zip [local/web 22.4KB] together with the C++ code to verify the hypotheses of Theorem 4.2. The following first and last parallelotopes can be extracted from this archive:

$$\langle A_0, \tilde{\mathbf{x}}_0 \rangle = \langle B_0, \tilde{\mathbf{y}}_0 \rangle \approx \left\langle \begin{pmatrix} 0.000352 & 0.000093 \\ -0.000935 & 0.000034 \end{pmatrix}, \begin{pmatrix} -0.161464 \\ -0.433261 \end{pmatrix} \right\rangle, \tag{5.4}$$

$$\langle A_{36}, \tilde{\mathbf{x}}_{36} \rangle \approx \left\langle \begin{pmatrix} 0.00000071 & 0.000033 \\ -0.00000073 & 0.000195 \end{pmatrix}, \begin{pmatrix} -0.146851 \\ -0.676514 \end{pmatrix} \right\rangle, \tag{5.5}$$

$$\langle B_{36}, \tilde{\mathbf{y}}_{36} \rangle \approx \left\langle \begin{pmatrix} 0.00000063 & 0.000049 \\ -0.00000064 & 0.000279 \end{pmatrix}, \begin{pmatrix} -0.146855 \\ -0.676457 \end{pmatrix} \right\rangle. \tag{5.6}$$

Figure 7 shows the images of $\langle A_{36}, \tilde{\mathbf{x}}_{36} \rangle$ and $\langle B_{36}, \tilde{\mathbf{y}}_{36} \rangle$ inside the basis of $\langle A_0, \tilde{\mathbf{x}}_0 \rangle$. We can clearly see that the LICP is satisfied. More generally, the execution of the provided code shows that the hypotheses of Theorem 4.2 are all satisfied and therefore that the Tinkerbell map is chaotic in the sense of Theorem 4.2.

5.3. The Hénon map is topologically conjugate to the full shift on two symbols. For the Hénon map, we use the same branching periodic pseudo-orbits as in [35]:

$$\begin{pmatrix} 0.7545200204672611 \\ 0.1664237457349109 \end{pmatrix} \; ; \; \begin{pmatrix} 0.7545213108637951 \\ 0.16640114136214962 \end{pmatrix}. \tag{5.7}$$

Parallelotopes obtained are available in the same archive as previously; the execution of the provided C++ code proves that the hypotheses of Theorem 4.3 are satisfied and therefore that the Hénon map is topologically conjugate to the full shift on two symbols. We therefore reproduce the result obtained in [35] by using a containment shadowing argument.

REFERENCES

[1] G. Alefeld and J. Herzberger, *Introduction to Interval Computations*, Comput. Sci. Appl. Math., Academic Press, New York, 1983.

[2] K.T. Alligood, T.D. Sauer, and J.A. Yorke, *Chaos: An Introduction to Dynamical Systems*, Springer-Verlag, Berlin, 1996.

[3] Z. Arai, *On hyperbolic plateaus of the hénon map*, Exp. Math., 16 (2007), pp. 181–188.

[4] B. Banhelyi, T. Csendes, B.M. Garay, and L. Hatvani, *A computer-assisted proof of Σ_3-chaos in the forced damped pendulum equation*, SIAM J. Appl. Dyn. Syst., 7 (2008), pp. 843–867.

[5] F. Blanchard, E. Glasner, S. Kolyada, and A. Maass, *On Li-Yorke pairs*, J. Reine Angew Math., 547 (2002), pp. 51–68.

[6] N. Delanoue, L. Jaulin, and B. Cottenceau, *Using interval arithmetic to prove that a set is path-connected*, Theoret. Comput. Sci., 351 (2006), pp. 119–128.

[7] N. Delanoue, L. Jaulin, and B. Cotteneau, *Guaranteeing the homotopy type of a set defined by non-linear inequalities*, Reliab. Comput., 13 (2007), pp. 381–398.

[8] J. Dieudonné, *Foundations of Modern Analysis*, Vol. 1, Academic Press, New York, 1969.

[9] Z. Galias and P. Zgliczynski, *Abundance of homoclinic and heteroclinic orbits and rigorous bounds for the topological entropy for the Hénon map*, Nonlinearity, 14 (2001), pp. 909–932.

[10] M. Gidea and P. Zgliczynski, *Covering relations for multidimensional dynamical systems*–II, J. Differential Equations, 202 (2004), pp. 59–80.

[11] D. Goldberg, *What every computer scientist should know about floating-point arithmetic*, Comput. Surveys, 23 (1991), pp. 5–48.

[12] F. Goualard, *GAOL* 3.1.1: *Not Just Another Interval Arithmetic Library*, 4.0 ed., Laboratoire d'Informatique de Nantes-Atlantique, Nantes, France, 2006.

[13] C. Grebogi, S.M. Hammel, J.A. Yorke, and T. Sauer, *Shadowing of physical trajectories in chaotic dynamics: Containment and refinement*, Phys. Rev. Lett., 65 (1990), pp. 1527–1530.

[14] W. Hayes, *Computer simulations, exact trajectories, and the gravitational n-body problem*, Amer. J. Phys., 72 (2004), pp. 1251–1257.

[15] W.B. Hayes and K.R. Jackson, *Rigorous shadowing of numerical solutions of ordinary differential equations by containment*, SIAM J. Numer. Anal., 41 (2003), pp. 1948–973.

[16] W. Hayes and K.R. Jackson, *A survey of shadowing methods for numerical solutions of ordinary differential equations*, Appl. Numer. Math., 53 (2005), pp. 299–321.

[17] L. Jaulin, M. Kieffer, O. Didrit, and E. Walter, *Applied Interval Analysis with Examples in Parameter and State Estimation, Robust Control and Robotics*, Springer-Verlag, London, 2001.

[18] T. Kapela and C. Simó, *Computer assisted proofs for nonsymmetric planar choreographies and for stability of the eight*, Nonlinearity, 20 (2007), pp. 1241–1255.

[19] A. Katok and B. Hasselblatt, *Introduction to the Modern Theory of Dynamical Systems*, Cambridge University Press, Cambridge, UK, 1995.

[20] R.B. Kearfott, *Interval computations: Introduction, uses, and resources*, Euromath Bull., 2 (1996), pp. 95–112.

[21] J. Kennedy and J.A. Yorke, *Shadowing in higher dimensions*, in Differential Equations, Chaos and Variational Problems, Vol. 75, Birkhäuser Verlag, Basel, 2008, pp. 241–246.
[22] O. Knueppel, *PROFIL/BIAS - A fast interval library*, Computing, 53 (1994), pp. 277–287.
[23] S.F. Kolyada, *Li-Yorke sensitivity and other concepts of chaos*, Ukrainian Math. J., 56 (2004), pp. 1242–1257.
[24] W. Kulpa, *The Poincaré-Miranda theorem*, Amer. Math. Monthly, 104 (1997), pp. 545–550.
[25] C. Miranda, *Un' osservazione su un teorema di Brouwer*, Boll. Un. Mat. Ital. (2), 3 (1940), pp. 5–7.
[26] J.D. Mireles James, *Adaptive set-oriented computation of topological horseshoe factors in area and volume preserving maps*, SIAM J. Appl. Dyn. Syst., 9 (2010), pp. 1164–1200.
[27] R. Moore, *Interval Analysis*, Prentice–Hall, Englewood Cliffs, NJ, 1966.
[28] A. Neumaier, *Interval Methods for Systems of Equations*, Cambridge University Press, Cambridge, UK, 1990.
[29] H.E. Nusse and J.A. Yorke, *Dynamics: Numerical Explorations*, Springer, New York, 1994.
[30] K. Ozaki, T. Ogita, S. Miyajima, S. Oishi, and S. M. Rump, *A method of obtaining verified solutions for linear systems suited for Java*, J. Comput. Appl. Math., 199 (2006), pp. 337–344.
[31] H. Poincaré, *Sur certaines solutions particulières du problème des trois corps*, Comptes Rendus de l'Académie des Sciences, 97 (1883), pp. 251–252.
[32] S.M. Rump, *INTLAB - INTerval LABoratory*, in Developments in Reliable Computing, Tibor Csendes, ed., Kluwer Academic Publishers, Dordrecht, The Netherlands, 1999, pp. 77–104.
[33] L. Sella and P. Collins, *Computation of symbolic dynamics for one-dimensional maps*, J. Comput. Appl. Math., 234 (2010), pp. 418–436.
[34] R. Srzednicki, *A generalization of the Lefschetz fixed point theorem and detection of chaos*, Proc. Amer. Math. Soc., 128 (2000), pp. 1231–1239.
[35] D. Stoffer and K.J. Palmer, *Rigorous verification of chaotic behaviour of maps using validated shadowing*, Nonlinearity, 12 (1999), pp. 1683–1698.
[36] E. Svensson, *Computational Characterization of Flows with Some Hyperbolicity*, Technical report 2006:20, Chalmers University of Technology and Göteborg University, Göteborg, Sweden, 2006.
[37] E. Svensson, *Computational Characterization of Mixing in Flows*, Ph.D. thesis, Chalmers University of Technology and Göteborg University, Göteborg, Sweden, 2006.
[38] W. Tucker, *A rigorous ODE solver and Smale's* 14*th problem*, Found. Comput. Math., 2 (2002), pp. 53–117.
[39] D. Wilczak and P. Zgliczyński, *Computer assisted proof of the existence of homoclinic tangency for the Hénon map and for the forced damped pendulum*, SIAM J. Appl. Dyn. Syst., 8 (2009), pp. 1632–1663.
[40] Wolfram Research Inc., *Mathematica*, Version 7.0, Champaign, IL, 2008.
[41] C. Young, W. Hayes, and K.R. Jackson, *Rigorous high-dimensional shadowing using containment: The general case*, Discrete Contin. Dyn. Syst., 14 (2006), pp. 329–342.
[42] P. Zgliczyński and M. Gidea, *Covering relations for multidimensional dynamical systems*, J. Differential Equations, 202 (2004), pp. 32–58.