

Recent Development in Rigorous Computational Methods in Dynamical Systems

Zin ARAI*, Hiroshi KOKUBU† and Paweł PILARCZYK‡

**Hokkaido University, Creative Research Institution “Sousei”
N21W10 Kita-ku, Sapporo 001–0021, Japan and
PRESTO, Japan Science and Technology Agency, Japan
URL: <http://www.cris.hokudai.ac.jp/arai/>*

†*Kyoto University, Department of Mathematics
Kyoto 606–8502, Japan
URL: <http://www.math.kyoto-u.ac.jp/~kokubu/>
E-mail: kokubu@math.kyoto-u.ac.jp*

‡*University of Minho, Centre of Mathematics
Campus de Gualtar, 4710–057 Braga, Portugal
URL: <http://www.pawelpilarczyk.com/>*

Received May 19, 2008

Revised November 26, 2008

We highlight selected results of recent development in the area of rigorous computations which use interval arithmetic to analyse dynamical systems. We describe general ideas and selected details of different ways of approach and we provide specific sample applications to illustrate the effectiveness of these methods. The emphasis is put on a topological approach, which combined with rigorous calculations provides a broad range of new methods that yield mathematically reliable results.

Key words: dynamical system, rigorous computations, algorithm, Conley index, hyperbolic dynamics

1. Introduction

Considerable progress has been recently achieved in various applications of rigorous numerical methods using interval arithmetic to the study of dynamical systems. In particular, the combination of the interval arithmetic and some topological and graph algorithm techniques appears to be very successful for revealing global aspects of dynamics in both ODEs and iterates of maps in a mathematically rigorous manner. This article illustrates selected methods of this type and their applications done by the authors and their collaborators.

In the subsequent sections, we discuss the following topics: In Section 2, we show how a certain invariant set consisting of a periodic orbit and heteroclinic orbits of a concrete three-dimensional system of ODEs can be captured by a topological method using so-called h-sets and their covering relation. In Section 3, we describe a constructive method which uses the Conley index theory for proving existence of periodic orbits in ODE systems. Sections 4, 5 and 6 are devoted to the study

*Z. Arai was partially supported by Grant-in-Aid for Scientific Research (No. 17740054).

†H. Kokubu was partially supported by Grant-in-Aid for Scientific Research (No. 17340045).

‡P. Pilarczyk was partially supported by the JSPS Postdoctoral Fellowship (No. P06039) at the Department of Mathematics, Kyoto University, and by Grant-in-Aid for Scientific Research (No. 1806039).

of iterates of maps. In Section 4, we outline a method for giving a combinatorial outer-approximation of the dynamics of iterates of maps, from which one can easily obtain rigorous bounds on various types of dynamical information such as Morse decomposition, the number of attractors, the Conley index of isolated invariant sets, and bifurcations. A similar combinatorial representation can also be used to prove the hyperbolicity of the dynamics, if the method is applied to the induced maps on the tangent bundle, as illustrated in Section 5 in the case of the Hénon map family. A variant of the combinatorial representation which also takes into account the derivative of the map is described in Section 6 where it is applied to one-dimensional quadratic maps to prove uniform expansion outside some neighborhood of the critical point.

Throughout the paper, it is emphasized that the use of topological approach is crucial, and, in particular, an algebraic topological method known as the homological Conley index theory provides a powerful tool, especially that this index can be efficiently computed with the publically available CHomP software package [8] which is based on the theory introduced in [20, 29, 36] (see also references therein).

Needless to say, the interval arithmetic is used as a basic tool for constructing a rigorous outer enclosure of a map on subsets of the phase space, but for ODEs only advanced integrators (like those included in the CAPD software package [7]) based on interval arithmetic yield satisfactory results.

Finally, the software implementation of most algorithms described here is freely available for download with the source code distributed under the terms of the GNU General Public License; see Sections 3, 4, 6 and references therein for the details.

Since the research on dynamical systems involves a wide variety of techniques from various branches of mathematics, it would be impossible to gather all of the recent achievements in the applications of rigorous numerics to this area in a short survey paper. Therefore, we purposely omit some subjects and focus on the selected ones. Moreover, we would like to point out that in addition to the discussed results, there are also other important works in similar spirit which we do not mention here, for example, [4, 11, 18]. See also a survey article of K. Mischaikow [28].

Acknowledgement. The results presented here have been achieved within the framework of or inspired by our long-term collaboration with Konstantin Mischaikow (Rutgers University, U.S.A.), Marian Mrozek (Jagiellonian University, Poland) and several other scientists. We are grateful to all of them for joint work and fruitful discussions.

2. Cocoon bifurcation for the Michelson system

We begin the overview by introducing a rigorous computational method for a system of ODEs which involves both elementary application of interval arithmetic, and advanced rigorous integration of trajectories as well. Combined with topological arguments, this method can be used to prove the existence of an infinite cascade of bifurcations of certain type.

The Michelson system [25]

$$\begin{aligned}\dot{x} &= y, \\ \dot{y} &= z, \\ \dot{z} &= c^2 - y - \frac{1}{2}x^2\end{aligned}\tag{1}$$

is known as the travelling wave equation of the Kuramoto–Sivashinsky PDE, and it also arises as the limit family of an unfolding of a codimension three nilpotent vector field singularity (see [12]).

The Michelson system (1) satisfies the following as its basic properties:

- The time-reversibility with respect to the involution $R(x, y, z) = (-x, y, -z)$.
- Two equilibrium points for $c > 0$ at $x_{\pm} = (\pm\sqrt{2}c, 0, 0)$, both of which are of saddle-focus type with $\dim(W^u(x_+)) = \dim(W^s(x_-)) = 2$.

McCord [26] proved that there exists a unique transverse heteroclinic orbit connecting x_+ and x_- in (1) for sufficiently large c , and that the equilibrium points together with the heteroclinic orbit form the maximal bounded invariant set of (1).

When the parameter c decreases, it is numerically observed in [25, 24] that (1) exhibits an infinite sequence of heteroclinic bifurcations due to the tangency of $W^u(x_+)$ and $W^s(x_-)$, each of which creates a pair of new transverse heteroclinic orbits. Also numerically, the sequence of values of c for which these bifurcations occur converges to $\bar{c} \approx 1.2662$, at which there appears a saddle-node bifurcation that creates a periodic orbit γ_* . These sequences of bifurcations were numerically studied by Lau [24] who called them the “cocoon” bifurcation.

Later, Dumortier, Ibáñez and the second author of this paper [13] studied the cocoon bifurcation, and proved that its occurrence can be explained by the presence of a special invariant set called the *cusplike transverse heteroclinic chain*, which is defined below. However, the question of whether the Michelson system indeed has the cusplike transverse heteroclinic chain remained open until very recently [23]. In this section, we show how such a question can be answered by the use of rigorous numerical computation.

Throughout this section, let X_λ be a one-parameter family of vector fields on \mathbb{R}^3 having the same basic properties as (1); namely, the time-reversibility with respect to the involution R and the presence of the saddle-focus equilibrium points x_{\pm} . (See (H1) and (H2) in [13] for the detail.)

DEFINITION 2.1 (see [13, Definition 1.3]). *We say that the family X_λ exhibits a cocooning cascade of heteroclinic tangencies centered at λ_* , if there exists a closed solid torus T with $x_{\pm} \notin T$ and a monotone infinite sequence of parameters λ_n converging to λ_* , for which the corresponding vector field X_{λ_n} has a tangency of $W^u(x_+)$ and $W^s(x_-)$ such that the heteroclinic tangency orbit intersects with T and has its length within T tending to infinity as $n \rightarrow \infty$.*

DEFINITION 2.2 (see [13, Definition 1.4]). *The family X_λ has a cusp-transverse heteroclinic chain (CTHC for short) at $\lambda = \lambda_0$, if the following three conditions hold:*

- (C1) X_{λ_0} has a saddle-node periodic orbit γ_* which is symmetric under the involution R . Here the saddle-node periodic orbit is meant by a periodic orbit whose Poincaré map has the unity as its eigenvalue. Under the presence of the reversibility, this implies that the other eigenvalue of the linearized Poincaré map is also the unity.
- (C2) The saddle-node periodic orbit γ_* is generic and generically unfolded in X_λ under the reversibility with respect to R . Here the genericity means that some of the derivatives of the Poincaré map for the saddle-node periodic orbit are non-zero, see [13] for more details.
- (C3) $W^u(\gamma_*)$ and $W^s(x_-)$, as well as $W^s(\gamma_*)$ and $W^u(x_+)$, intersect transversely, where $W^u(\gamma_*)$ and $W^s(\gamma_*)$ stand for the stable and unstable sets of the non-hyperbolic periodic orbit γ_* .

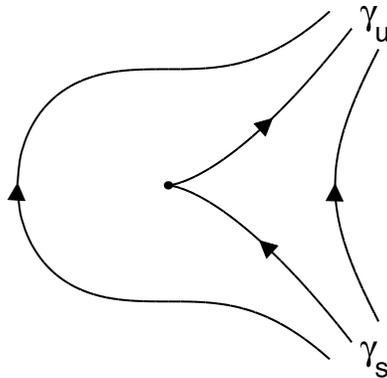


Fig. 1. The stable and unstable set of the saddle-node periodic orbit on the Poincaré section.

THEOREM 2.3 (see [13, Theorem 1.5]). *Let X_λ be a smooth family of reversible vector fields on \mathbb{R}^3 as above. Suppose that at $\lambda = \lambda_0$ the corresponding vector field X_{λ_0} has a cusp-transverse heteroclinic chain. Then the family exhibits a cocooning cascade of heteroclinic tangencies centered at λ_0 .*

In order to prove the existence of a CTHC in (1), one has to verify the conditions (C1)–(C3). All but (C3) are relatively easy to check by computation with the interval arithmetic, but rigorous enclosure of the stable and unstable sets is not enough to verify the transversality in (C3). Nevertheless, we can prove the existence of a CTHC and obtain the following theorem:

THEOREM 2.4 (see [23]). *The Michelson system (1) exhibits a cocooning cascade of heteroclinic tangencies centered at a*

$$c_\infty \in [1.2662323370670545, 1.2662323370713253].$$

The main idea of the proof of this theorem is to replace the condition (C3) by its topological version:

(C3t) $W^u(\gamma_*)$ and $W^s(x_-)$, as well as $W^s(\gamma_*)$ and $W^u(x_+)$, intersect. Moreover, these intersections are *topologically transverse*.

Notice that the proof of Theorem 2.3 in [13, Theorem 1.5] uses topological arguments, like the intermediate value theorem, and hence it remains valid even if (C3) is replaced by (C3t). Therefore, once we verify the weaker condition (C3t), we can conclude the existence of a cocooning cascade of heteroclinic tangencies centered at c_∞ .

For the verification of (C1, C2, C3t), we use rigorous numerical methods developed by the CAPD group [7, 41, 42]. The main tool for the verification is a method involving so-called h-sets and their covering relations. An h-set is a box-like structure which carries topological information of dynamics similar to the hyperbolicity. The covering relation from one h-set to another guarantees that the hyperbolic-like behavior can be carried over between these h-sets. See [40] and references therein for more details about the h-sets and covering relations. Fig. 2 shows some of the h-sets used in the proof. Notice that the figure is drawn on a cross section of the saddle-node periodic orbit which corresponds to the point v_∞ .

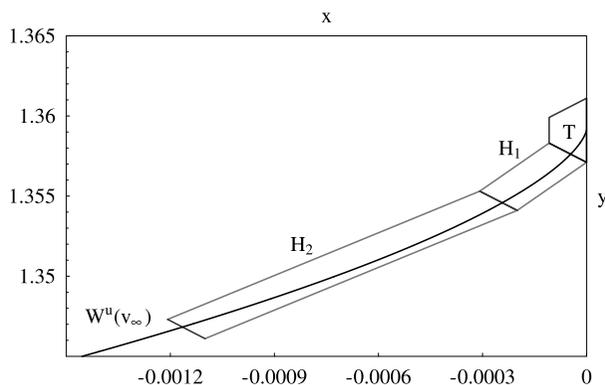


Fig. 2. The h-sets T, H_1, H_2 . The curve indicates the unstable set $W^u(v_\infty)$ of the saddle-node periodic point v_∞ of the Poincaré map which forms a horizontal disk in H_2 .

For the verification of the existence of a CTHC, one constructs h-sets that contain pieces of the stable and unstable manifolds of the equilibrium points and the saddle-node periodic orbit in such a way that these h-sets satisfy good covering relations that lead to a nice transverse intersection. Since the saddle-node periodic orbit is not hyperbolic, it is not straightforward how to carry out this strategy. For this purpose, we use two additional ideas in [23]: h-sets with cones and a Lyapunov

function near the saddle-node periodic orbit. The former carries the information on the derivative of the dynamics in the form of cones, and thus controls how a piece of stable or unstable manifolds can be mapped nicely to the next h-set. The latter can give some monotonicity property of a piece of the stable and unstable sets of the non-hyperbolic saddle-node periodic orbit, which helps to prove that these pieces are properly contained in the constructed h-sets. Combining this information, we can prove that certain part of the stable manifold of the hyperbolic equilibrium point x_- intersects topologically transversely, in an h-set, with the unstable set of the saddle-node periodic point, as it is shown in Fig. 3. See [23] for the detailed proof and related information.

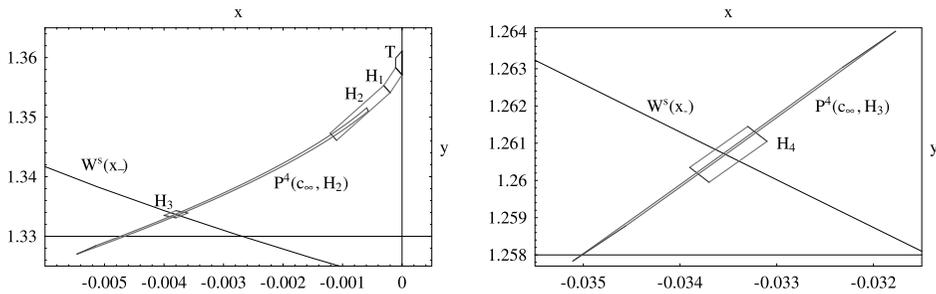


Fig. 3. Numerical illustration of topologically transverse intersections of intermediate h-sets H_2 , H_3 , H_4 and $W^s(x_-)$ under the Poincaré map $P^4(c_\infty, \cdot)$ along the saddle-node periodic orbit, which is one of the essential steps for the proof of Theorem 2.4.

Finally, we remark that the above methods are general and can be applied to other systems which exhibit a similar bifurcation.

3. Periodic trajectories in ODEs

Advanced numerical methods for rigorous integration of ODEs combined with topological tools (in particular, the Conley index) can be used to develop constructive methods for proving the existence of certain types of invariant sets for flows. In this section we describe an approach introduced in [33, 34, 35] for proving the existence of periodic trajectories.

3.1. Introduction

Assume we have an ordinary differential equation

$$\dot{x} = f(x), \tag{2}$$

where f is a C^1 -class vector field in \mathbb{R}^n . Suppose that numerical simulations indicate the existence of a periodic trajectory which seems to be a hyperbolic orbit, and we would like to prove the existence of a periodic trajectory near the observed orbit. A rough approximation of this trajectory is a starting point for an algorithm which constructs a possibly small neighborhood of this set satisfying certain assumptions

used further to prove that it contains a periodic trajectory indeed. As it will be pointed out in the sequel, limited amount of input is necessary from the user, and all the remaining computations are performed automatically. The theoretical part involves the Conley index theory [9] and will be explained below.

3.2. Preliminaries

If $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$ is a C^1 -class vector field which is bounded by a linear function for large arguments then the equation (2) induces a flow φ on \mathbb{R}^n . We can make this assumption without loss of generality, because we are only interested in the dynamics on some bounded region in \mathbb{R}^n .

A set $S \subset \mathbb{R}^n$ is called *invariant* with respect to the (continuous or discrete) dynamical system φ if $\varphi(S, T) = S$, where $T = \mathbb{R}$ or \mathbb{Z} , respectively. The *invariant part* of a set $N \subset \mathbb{R}^n$, denoted as $\text{inv}(N, \varphi)$, is defined as its largest invariant subset with respect to φ . If φ is a discrete dynamical system (with $T = \mathbb{Z}$) and $f = \varphi(\cdot, 1)$ is its generator then $\text{inv}(N, \varphi)$ is also denoted as $\text{inv}(N, f)$. The set N is called an *isolating neighborhood* if N is compact and $\text{inv}(N, \varphi) \subset \text{int} N$.

A compact subset Ξ of an $(n - 1)$ -dimensional hyperplane P is called a *local section* for the flow φ on \mathbb{R}^n induced by the equation (2) if the vector field f is transverse to P on Ξ . A local section Ξ is called a *Poincaré section for φ in an isolating neighborhood N* if $\Xi \cap N$ is closed and for every $x \in N$ there exists $t > 0$ such that $\varphi(x, t) \in \Xi$. If such a section exists for N then we say that N *admits a Poincaré section*.

3.3. The Conley index

The definition of the Conley index [9] for flows (valid also for semiflows in which only forward trajectories are well-defined) is based on the notion of an index pair (see Fig. 4).

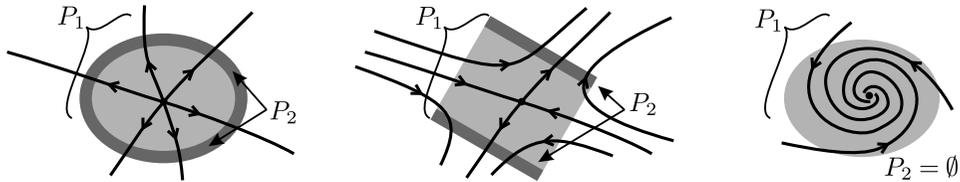


Fig. 4. Sample index pairs for some fixed points in \mathbb{R}^2 . P_2 is indicated in dark gray, $P_1 \setminus P_2$ in light gray.

DEFINITION 3.1. Let φ be a flow on \mathbb{R}^n . The pair of compact sets (P_1, P_2) such that $P_2 \subset P_1$ is called an *index pair* if the following properties hold true:

- (1) if $x \in P_1$ and $\varphi(x, s) \notin P_1$ for some $s > 0$ then there exists $t \geq 0$ such that $\varphi(x, [0, t]) \subset P_1$ and $\varphi(x, t) \in P_2$,
- (2) if $x \in P_2$ and $\varphi(x, [0, t]) \subset P_1$ then also $\varphi(x, [0, t]) \subset P_2$,
- (3) $\text{inv}(\text{cl}(P_1 \setminus P_2), \varphi) \subset \text{int}(P_1 \setminus P_2)$.

The *Conley index* of the index pair (P_1, P_2) is the homotopy type of the pointed quotient space $(P_1/P_2, [P_2])$. This index is independent of the choice of the index pair for the isolated invariant set $S := \text{inv}(\text{cl}(P_1 \setminus P_2), \varphi)$, and thus one can speak of the Conley index of an isolated invariant set (or of an isolating neighborhood). Among many useful properties of the index, the most basic one is that if the index is non-trivial then $S \neq \emptyset$.

By applying the cohomology functor to the Conley index, one obtains the *cohomological Conley index* which is the relative cohomology $H^*(P_1, P_2)$ of the index pair. If both P_1 and P_2 are compact polyhedra then without loss of generality one can compute the simpler homology of (P_1, P_2) instead, and this can be done by efficient algorithms and software [7, 8, 20] if P_1 and P_2 are sets of special type (cubical sets, see Section 3.5).

In the case of discrete dynamical systems, there are several generalizations of the definition of the Conley index. Below we introduce a definition based on [30] with the index pairs as defined in [38] (this definition is also valid for semidynamical systems, see Section 4):

DEFINITION 3.2. *Let φ be a discrete dynamical system on \mathbb{R}^n , and $f := \varphi(\cdot, 1)$ its generator. The pair of compact sets (P_1, P_2) such that $P_2 \subset P_1$ is called an index pair if the following properties hold true:*

- (1) $f(P_1 \setminus P_2) \subset P_1$,
- (2) $f(P_2) \cap P_1 \subset P_2$,
- (3) $\text{inv}(\text{cl}(P_1 \setminus P_2), f) \subset \text{int}(P_1 \setminus P_2)$.

The properties (1) and (2) ensure that the map

$$f_P: (P_1, P_2) \rightarrow (P_1 \cup f(P_2), P_2 \cup f(P_2))$$

is well defined and that the *excision property* $(P_1 \cup f(P_2)) \setminus P_1 = (P_2 \cup f(P_2)) \setminus P_2$ holds true. As a consequence, the inclusion map

$$i_P: (P_1, P_2) \rightarrow (P_1 \cup f(P_2), P_2 \cup f(P_2))$$

is an excision for the Alexander–Spanier cohomology, which implies that i_P^* is an isomorphism, and therefore, the *index map*

$$I_P^* := f_P^* \circ (i_P^*)^{-1}: H^*(P_1, P_2) \rightarrow H^*(P_1, P_2)$$

is well defined. The *cohomological Conley index* of (P_1, P_2) is defined as the *Leray reduction* (see [30]) of $(H^*(P_1, P_2), I_P^*)$.

As in the case of flows, the remark on computing the homological (instead of cohomological) Conley index is also valid here. Note that the algorithms in [8, 29, 36] are capable of computing the homomorphism induced in homology by a continuous map between cubical sets.

3.4. The time- t discretization of a flow

Given $t > 0$, the *time- t discretization* of the flow φ on \mathbb{R}^n is the discrete dynamical system defined as

$$\varphi_t: \mathbb{R}^n \times \mathbb{Z} \ni (x, \tau) \rightarrow \varphi(x, \tau t) \in \mathbb{R}^n. \tag{3}$$

Efficient rigorous numerical methods exist for integrating vector fields in \mathbb{R}^n which can be used to compute an outer approximation of $\varphi_t(Q)$ for a rectangular set $Q \subset \mathbb{R}^n$; see [7] for a sample effective implementation of such methods.

The following two results, proved in a more general setting in [32] as Theorem 1 and Corollary, state the correspondence between isolated invariant sets with respect to the flow φ and its time- t discretization φ_t , and relate their Conley indices:

THEOREM 3.3. *For a flow φ in \mathbb{R}^n and $S \subset \mathbb{R}^n$ the following three conditions are equivalent:*

- (1) S is an isolated invariant set with respect to φ ,
- (2) S is an isolated invariant set with respect to φ_t for all $t > 0$,
- (3) S is an isolated invariant set with respect to φ_t for some $t > 0$.

THEOREM 3.4. *The cohomological Conley index of an isolated invariant set with respect to a flow φ coincides with the corresponding index with respect to the discrete dynamical system φ_t for any $t > 0$.*

Moreover, the following lemma proved in [33] indicates an important feature of isolating neighborhoods for the time- t discretization of a flow:

LEMMA 3.5. *Let N be an isolating neighborhood with respect to a discretization φ_t of the flow φ . Then N is an isolating neighborhood with respect to the flow φ and $\text{inv}(N, \varphi_t) = \text{inv}(N, \varphi)$.*

These results imply that the problem of finding an isolating neighborhood for a flow and computing its Conley index can be replaced by the same task for its time- t discretization. In what follows, it will be shown that the latter can be completed algorithmically in a constructive way.

3.5. Cubical sets and combinatorial maps

A class of algorithmically representable subsets of \mathbb{R}^n can be defined with the use of a rectangular grid in \mathbb{R}^n . For fixed positive numbers d_1, \dots, d_n , the set

$$\mathcal{H} := \{[d_1 k_1, d_1(k_1 + 1)] \times \dots \times [d_n k_n, d_n(k_n + 1)]: k_i \in \mathbb{Z}\} \tag{4}$$

is formally called a *grid* in \mathbb{R}^n , and elements of \mathcal{H} are n -dimensional hypercubes, further called *cubes* for short.

If $\mathcal{A} \subset \mathcal{H}$ is finite then the union of the elements of \mathcal{A} is denoted by $|\mathcal{A}|$ and is called a *cubical set*. We say that $\mathcal{A} \subset \mathcal{H}$ represents $|\mathcal{A}| \subset \mathbb{R}^n$. A *minimal cover* of an arbitrary set $A \subset \mathbb{R}^n$ is defined as $\{Q \in \mathcal{H}: Q \cap A \neq \emptyset\}$; if A is bounded then its minimal cover represents a cubical set.

If $\mathcal{X}, \mathcal{Y} \subset \mathcal{H}$ are finite then any multivalued map $\mathcal{F}: \mathcal{X} \multimap \mathcal{Y}$ (which maps each element of \mathcal{X} to a subset of \mathcal{Y}) is called a *combinatorial map*. Such a map is called a *representation* of a continuous map $f: |\mathcal{X}| \rightarrow |\mathcal{Y}|$ if for all $Q \in \mathcal{X}$

$$f(Q) \subset \text{int}|\mathcal{F}(Q)|. \quad (5)$$

The notion of an index pair can be carried over to the combinatorial setting.

DEFINITION 3.6. *Let \mathcal{F} be a combinatorial map. The pair $(\mathcal{P}_1, \mathcal{P}_2)$ of finite subsets of \mathcal{H} such that $\mathcal{P}_2 \subset \mathcal{P}_1$ is called a *combinatorial index pair* if the following properties hold true:*

- (1) $\mathcal{F}(\mathcal{P}_1 \setminus \mathcal{P}_2) \subset \mathcal{P}_1$,
- (2) $\mathcal{F}(\mathcal{P}_2) \cap \mathcal{P}_1 \subset \mathcal{P}_2$.

PROPOSITION 3.7 (see [35]). *If \mathcal{F} is a representation of $f := \varphi(\cdot, 1)$ for some discrete dynamical system φ , and $(\mathcal{P}_1, \mathcal{P}_2)$ is a combinatorial index pair for \mathcal{F} then $(|\mathcal{P}_1|, |\mathcal{P}_2|)$ is an index pair for φ .*

If \mathcal{F} satisfies some additional assumptions then \mathcal{F} can be used to compute the homomorphism induced in homology by the map f on the index pair (see [29, 36]), and thus to compute the homological Conley index with respect to f . In [33, 35], an algorithm is developed for the construction of a combinatorial index pair $(\mathcal{P}_1, \mathcal{P}_2)$ such that $\mathcal{P}_1 \setminus \mathcal{P}_2$ contains an a priori given set $\mathcal{Q} \subset \mathcal{H}$.

3.6. Main results

The main theorem upon which the result of this approach is based is a special case of Corollary 1.4 in [27]:

THEOREM 3.8. *Assume N is an isolating neighborhood for the flow φ which admits a Poincaré section. If N has the cohomological Conley index of some hyperbolic periodic orbit then $\text{inv}(N, \varphi)$ contains a periodic orbit.*

Given an approximate rough covering $\mathcal{Q} \subset \mathcal{H}$ of a hypothetic periodic trajectory observed in numerical simulations, one can use the algorithm for the construction of a combinatorial index pair containing \mathcal{Q} mentioned in Section 3.5 for a representation \mathcal{F} of the map φ_t computed with the use of the software [7] for some a priori chosen $t > 0$. If the algorithm succeeds then it constructs a combinatorial index pair $(\mathcal{P}_1, \mathcal{P}_2)$ such that $N := |\mathcal{P}_1 \setminus \mathcal{P}_2|$ is an isolating neighborhood with respect to φ_t , and the homological Conley index of N with respect to φ_t can be automatically computed with algorithms available in [8, 29, 36]. By Lemma 3.5, N is an isolating neighborhood also for the flow φ , and $S := \text{inv}(N, \varphi_t) = \text{inv}(N, \varphi)$. Moreover, the computed Conley index coincides with the Conley index of S with respect to the flow φ . If this index is the same as the index of some hyperbolic periodic orbit then it only suffices to verify the admittance of a Poincaré section to conclude the existence of a periodic trajectory with respect to φ contained in N . The latter stage of computations needs some additional input from the user and is described in [33, 34, 35].

The effectiveness of this approach for proving existence of periodic trajectories in ODEs is illustrated in [33, 34, 35] by several examples, including the well-known Vanderpol equations, Rössler equations, and Lorenz equations (see Fig. 5). For the parameter values for which either attracting or unstable periodic trajectories are numerically observed, an index pair is automatically constructed as described above, the homological Conley index of this index pair is computed, and a Poincaré section is defined and verified.

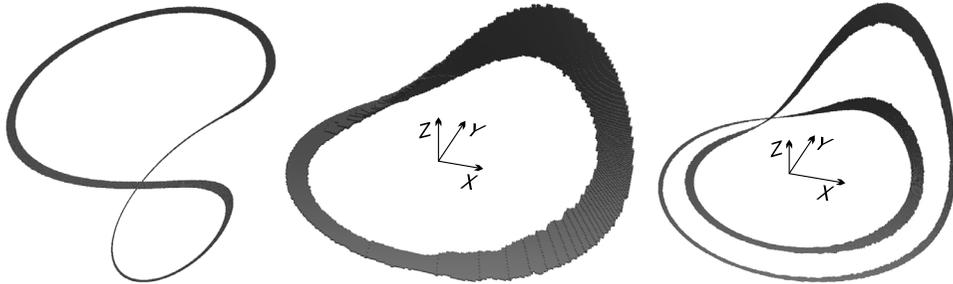


Fig. 5. Isolating neighborhoods constructed for a few stable periodic trajectories: One in the Lorenz equations (left), and two in the Rössler equations for different parameter values (middle and right) where a period doubling bifurcation is observed.

3.7. Remarks and comments

Among the benefits of the method described in this section it is worth to mention that it is stable with respect to small perturbations of the vector field. Even more, the vector field itself does not need to be known exactly, a relatively tight rigorous estimate is also sufficient. As a result of using this method, a bound for the location of the periodic trajectory whose existence is proved is obtained explicitly in terms of its isolating neighborhood in \mathbb{R}^n , unlike in methods based on the analysis of a Poincaré section alone.

Unfortunately, the stability type of the actual orbit whose existence is proven with this approach cannot be determined, but this is a general feature of topological methods. Another drawback of this method is its relatively high cost in terms of computation time and memory usage, especially if applied to unstable periodic orbits. Also the choice of various parameters of computations, like the grid sizes $d_i > 0$ and the time $t > 0$, may be difficult and in general must be done by the method of trial and error.

4. Combinatorial analysis of dynamics

The idea of subdividing a rectangular area in \mathbb{R}^n into a finite number of blocks using a rectangular grid and representing a continuous map in terms of a multi-valued map on grid elements (see Section 3.5) gives means to represent such a map as a graph, and then to apply fast graph algorithms for the automatic analysis of dynamics. In particular, the dynamics can be easily decomposed into recurrent

structures and the gradient-like part which allows further analysis of asymptotic dynamics. The automatization of this process allows one to scan large sets of parameters of a given family of dynamical systems to determine changes in dynamics automatically and to search for “interesting” regions of parameters worth further attention. In particular, [3] is a large project which benefits from these features, and we are going to describe it in this section.

4.1. Introduction

Assume we have a family of continuous maps

$$f: \mathbb{R}^n \times \mathbb{R}^d \ni (x, \lambda) \mapsto f_\lambda(x) \in \mathbb{R}^n \quad (6)$$

and we are interested in the analysis of global dynamics generated by these maps as the parameter λ is varied over \mathbb{R}^d . Since we do not assume that the maps f_λ are homeomorphisms, we must work in a more general setting of a semidynamical system in which only forward iterations are well-defined, and backward trajectories may be either ambiguous or undefined.

With typical applications in mind, we restrict our attention to bounded rectangular regions $R \subset \mathbb{R}^n$ and $A \subset \mathbb{R}^d$, and we introduce finite rectangular grids \mathcal{H} and \mathcal{L} in them (see Section 3.5). Then for each $L \in \mathcal{L}$, a common combinatorial representation \mathcal{F}_L valid for f_λ on R for all $\lambda \in L$ can be computed with the direct use of interval arithmetic. In what follows, we describe how the dynamics represented by \mathcal{F}_L is analysed, and how the relations between results of this analysis for \mathcal{F}_L and \mathcal{F}_K for adjacent parameter boxes $L, K \in \mathcal{L}$ are automatically determined to provide a global overview of the dynamics of (6) for the entire parameter region A .

4.2. Graphs

Since graphs are going to be used to represent various structures throughout this section, let us recall the definition of a graph. The pair $G = (V, E)$ is called a (*directed*) *graph* if V is a finite set and $E \subset V \times V$. Elements of V are called *vertices*, and elements of E are called *edges*. We say that the edge (v, w) goes *from* v *to* w . In illustrations, edges are indicated as arrows pointing from an object that represents v to the one for w . A *path* in the graph G is a finite sequence of vertices v_1, \dots, v_n such that $(v_i, v_{i+1}) \in E$ for all $i = 1, \dots, n-1$. One can also speak of an *undirected graph* if $(v, w) \in E$ implies $(w, v) \in E$ for all $v, w \in V$.

4.3. Conley–Morse decompositions

In a generic gradient flow on a 2-dimensional surface, the dynamics is very simple to describe: There are several isolated critical points (stable and unstable fixed points), and the remaining trajectories join one critical point with another. It turns out that every dynamical system on a compact invariant subset of \mathbb{R}^n can be described in a similar way, but instead of critical points one must use (possibly larger) isolated invariant sets:

DEFINITION 4.1 (see [9]). *Let $\varphi: X \times T^+ \rightarrow X$ be a semidynamical system. A Morse decomposition of X is a finite collection of disjoint isolated invariant*

sets S_1, \dots, S_n (called Morse sets) with strict partial ordering \prec on the index set $\{1, \dots, n\}$ such that for every $x \in X$ and every complete orbit γ through x there exist indices $i \prec j$ such that $\gamma(t) \rightarrow S_i$ and $\gamma(-t) \rightarrow S_j$ as $t \rightarrow \infty$.

A Morse decomposition can be represented in terms of a directed graph $G = (V, E)$ where $V = \{S_1, \dots, S_n\}$ and $(S_i, S_j) \in E$ iff $j \prec i$. This graph is called a *Morse graph*. If each Morse set is assigned its Conley index then such a structure is called a *Conley–Morse decomposition*, and the corresponding graph is called a *Conley–Morse graph*. In order to represent the computed Morse decomposition in a compact way, it is convenient to plot a Morse graph whose edges are determined by the *transitive reduction* of the relation \prec which is a minimal relation \prec' whose transitive closure retrieves \prec . This representation is used in Fig. 6.

4.4. Graph representation of dynamics

For each $L \in \mathcal{L}$, a combinatorial representation $\mathcal{F}_L: \mathcal{H} \rightarrow \mathcal{H}$ of f_λ in (6) can be represented by means of a directed graph $G = (V, E)$ where $V = \mathcal{H}$ and $(v, w) \in E$ iff $w \in \mathcal{F}(v)$. The analysis of G provides information on the asymptotic dynamics of f_λ represented by \mathcal{F}_L . For instance, each *combinatorial invariant set* defined as a set $\mathcal{S} \subset \mathcal{H}$ for which $\mathcal{S} \subset \mathcal{F}(\mathcal{S}) \cap \mathcal{F}^{-1}(\mathcal{S})$ represents an isolating neighborhood $|\mathcal{S}|$ with respect to f_λ . Moreover, a *combinatorial attractor* defined as a set $\mathcal{A} \subset \mathcal{H}$ such that $\mathcal{F}(\mathcal{A}) \subset \mathcal{A}$ represents an isolating neighborhood $|\mathcal{A}|$ whose invariant part is stable in the sense of Conley [9]: Every positive semitrajectory starting in some open neighborhood of A (actually, in $\text{int}|\mathcal{A}|$) approaches A ($\text{dist}(f_\lambda^n, A) \rightarrow 0$ as $n \rightarrow \infty$).

4.5. Combinatorial Morse decompositions

Extensive analysis of the dynamics via the graph G can be obtained by computing the *strongly connected components* of G , that is, maximal sets of vertices $C \subset V$ such that for each $v, w \in C$ there exists a path from v to w with vertices in C and also a path in the opposite direction (from w to v) through C . In [21] it is shown that all the strongly connected components of G form isolating neighborhoods for the union of all the chain recurrent sets of the dynamical system, and thus can serve as a *combinatorial Morse decomposition* $\{\mathcal{M}_i: i = 1, \dots, k\}$, for some $k > 0$, which represents a family of isolating neighborhoods $|\mathcal{M}_i|$. The sets \mathcal{M}_i are called *combinatorial Morse sets*. A partial order \prec between the computed combinatorial Morse sets can be determined by the analysis of paths in G connecting those sets: $i \prec j$ if $i \neq j$ and there exists a path in G from any vertex in \mathcal{M}_j to any vertex in \mathcal{M}_i . With the use of the graph G , a combinatorial Morse decomposition can be computed by algorithms introduced in [3, 5] based on the standard DFS (“depth-first search”) strategy. The complexity of these algorithms is linear (i.e., they are very fast).

The combinatorial Morse sets \mathcal{M}_i can be used directly for the computation of the Conley index using the approach introduced in [29] and refined in [36], where each \mathcal{M}_i gives rise to the pair $(\mathcal{M}_i \cup \mathcal{F}(\mathcal{M}_i), \mathcal{F}(\mathcal{M}_i) \setminus \mathcal{M}_i)$ satisfying the conditions for a combinatorial index pair (Definition 3.6) with respect to \mathcal{F}_L . In this way, a *combinatorial Conley–Morse decomposition* can be effectively computed.

Note that due to overestimates introduced in the map \mathcal{F}_L , the invariant part of some isolating neighborhoods $|\mathcal{M}_i|$ may be empty; such sets are called *spurious Morse sets*. An attempt in the algorithms is being made to find and eliminate these sets by taking a refinement of the grid \mathcal{H} and verifying whether the invariant part of $|\mathcal{M}_i|$ computed with this refinement can be proved to be empty indeed.

4.6. Continuation and bifurcations

Let $(\mathcal{M}_1^1, \dots, \mathcal{M}_{k_1}^1, \prec^1)$ be a combinatorial Morse decomposition for \mathcal{F}_{L_1} and let $(\mathcal{M}_1^2, \dots, \mathcal{M}_{k_2}^2, \prec^2)$ be a combinatorial Morse decomposition for \mathcal{F}_{L_2} , where $L_1, L_2 \subset \Lambda$ are such that $L_1 \cap L_2 \neq \emptyset$. We say that these decompositions are *equivalent* if $k := k_1 = k_2$ and there exists a bijection b of the set $\{1, \dots, k\}$ onto itself such that $\mathcal{M}_i^1 \cap \mathcal{M}_j^2 \neq \emptyset$ iff $j = b(i)$, and for no i, j one has $i \prec^1 j$ and $b(j) \prec^2 b(i)$. In this situation we talk about *continuation* of combinatorial Morse decompositions.

This definition provides a very weak notion of continuation; it is, in fact, the continuation of isolating neighborhoods in the sense of Conley (see [9]). In particular, the lack of continuation suggests a substantial change in dynamics, but does not necessarily imply one, because the perception of the dynamics is limited to the finite resolution of the grid. Nevertheless, one can use the Conley index and trace sequences of changes in combinatorial Morse sets over series of adjacent parameter boxes in order to detect behaviors resembling certain types of bifurcations, like period doubling or a saddle-node bifurcation, taking place at the level of isolating neighborhoods.

4.7. Applications

The practicality of the approach based on the analysis of combinatorial Morse decompositions is illustrated in [3] with the application of this method to the 2-dimensional overcompensatory Leslie population model and $f: \mathbb{R}^2 \times \mathbb{R}^4 \rightarrow \mathbb{R}^2$ in (6) given by

$$f(x_1, x_2; \theta_1, \theta_2, \kappa, p) = (\theta_1 x_1 + \theta_2 x_2 e^{-\kappa(x_1 + x_2)}, p x_1). \quad (7)$$

This model and its biological relevance is discussed in [39] where detailed numerical studies were conducted which indicate that this system exhibits a wide variety of dynamical behaviors.

In [3], a bounded rectangular region $R \subset \mathbb{R}^2$ is fixed which is proven in [39] to contain the dynamics of interest, and bounded ranges of parameters θ_1, θ_2 and p are considered. The parameter κ is a rescaling factor, so it is arbitrarily set to 0.1, like in [39]. The numerical analysis of the model done in [39] fixes $p = 0.7$ and assumes $\theta_1 = \theta_2$ varying in the interval $[10, 35]$ so that f defines a one-parameter family of discrete semidynamical systems. In contrast to this, the method discussed here allows to analyze the dynamics over full ranges of all parameters at the same time, without the limitations of the ad hoc assumptions.

A sample result of computations with this method is illustrated in Fig. 6. The parameter space $(\theta_1, \theta_2) \in [10, 50]^2$ was subdivided into 64×64 equal squares,

and the other parameters were fixed as $p = 0.7$, $\kappa = 0.1$. The region R was subdivided into 4096×4096 rectangles of the same size, and a combinatorial Conley–Morse decomposition was computed for each parameter box using this grid. In the picture, adjacent boxes in the parameter space (θ_1, θ_2) with equivalent Morse decompositions are plotted in the same shade of gray. The transitive reduction of the Morse graph is illustrated for all the regions; a square indicates an attractor, a filled circle corresponds to a Morse set with a nontrivial Conley index, and a hollow circle indicates a Morse set with the trivial Conley index. In addition to the information shown, the actual computations also provide the constructed isolating neighborhoods $|\mathcal{M}_i|$, rigorous estimates for the location of connecting orbits, upper and lower bounds for the attraction basins of attractors, and the Conley indices of all the constructed Morse sets.

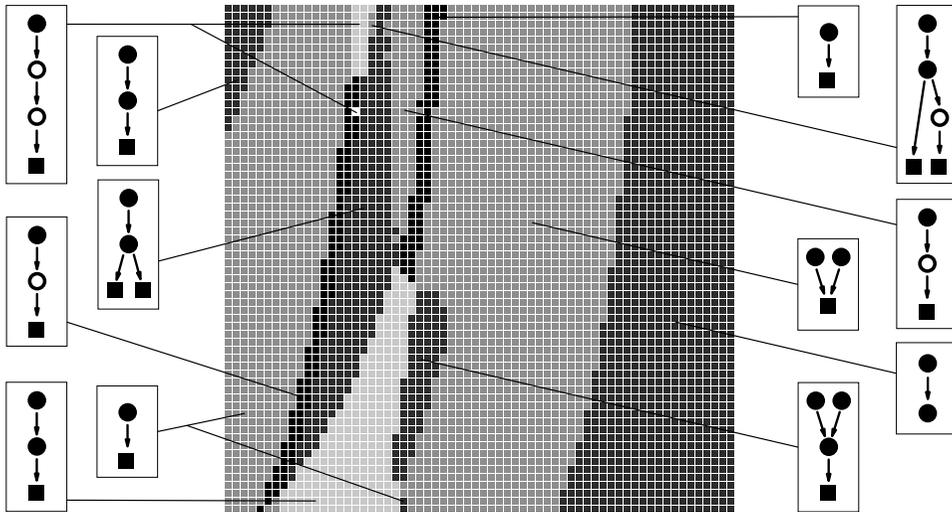


Fig. 6. Sample result of computations for the Leslie population model discussed in Section 4.7.

4.8. Remarks and comments

We would like to point out the fact that this method, based on outer approximations of families of maps, provides results which are stable under small perturbations, and thus cannot detect features that are not characterized by open conditions. In particular, the detection of bifurcations is limited, as it was argued in Section 4.6.

However, this wide and robust approach provides a very effective method for “scanning” comprehensive parameter ranges in search for changes in global dynamics. An undisputable advantage is also that the family of maps (6) does not even have to be defined by an explicit formula, a relatively tight rigorous approximation suffices in this approach. All this decides on the high potential of this approach and its wide applicability.

5. Uniform hyperbolicity

Applying the combinatorial techniques developed in the previous section to the analysis of the dynamics at the tangent bundle allows one to develop a rigorous computational method for proving uniform hyperbolicity of discrete dynamical systems. We briefly introduce this method in this section, and we refer the reader to [1] for details.

5.1. Introduction

Consider the problem of determining the set of parameter values for which the real Hénon map

$$H_{a,b}: \mathbb{R}^2 \rightarrow \mathbb{R}^2: (x, y) \mapsto (a - x^2 + by, x) \quad (a, b \in \mathbb{R})$$

is uniformly hyperbolic. If a dynamical system is uniformly hyperbolic, generally speaking, we can apply the so-called hyperbolic theory of dynamical systems and obtain many results on the behavior of the system. However, proving hyperbolicity is a difficult problem even for such simple polynomial maps.

The first mathematical result about the hyperbolicity of the Hénon map was obtained by Devaney and Nitecki [17]. They showed that for any fixed b , if a is sufficiently large then the non-wandering set of $H_{a,b}$ is uniformly hyperbolic and conjugate to the full horseshoe map. Besides this uniformly hyperbolic full horseshoe region, Davis, MacKay and Sannami [14] observed that there exist some parameter regions in which the non-wandering set of the Hénon map is uniformly hyperbolic. Their observation is based on numerical computations of the stable and unstable manifolds, and although the mechanism of hyperbolicity is clear by their observation, no mathematical proof of the uniform hyperbolicity has been obtained so far.

Applying our method to the real Hénon map, we obtain a computer assisted proof of the uniform hyperbolicity of Hénon map on many parameter regions including those discussed by Davis et al. (see Theorem 1.1 and 1.2 of [1]).

THEOREM 5.1 ([1, Theorem 1.1]). *There exists a set $P \subset \mathbb{R}^2$, which is the union of 8943 closed rectangles, such that if $(a, b) \in P$ then $\mathcal{R}(H_{a,b})$ is uniformly hyperbolic. The set P is illustrated in Fig. 7 (shaded regions). The complete list of these rectangles can be found at <http://www.cris.hokudai.ac.jp/arai/>.*

Here we denote by $\mathcal{R}(H_{a,b})$ the chain recurrent set of $H_{a,b}$. The uniform hyperbolicity of the chain recurrent set implies the \mathcal{R} -stability, and therefore, the theorem implies that no bifurcation of the chain recurrent set takes place on P . Hence numerical invariants such as the topological entropy, the number of periodic points, etc., are constant on each component of P .

Notice that the theorem claims nothing about parameter values outside P ; they may be hyperbolic or non-hyperbolic. For verifying the hyperbolicity of hyperbolic parameter values which is not contained in P , we need a longer computation.

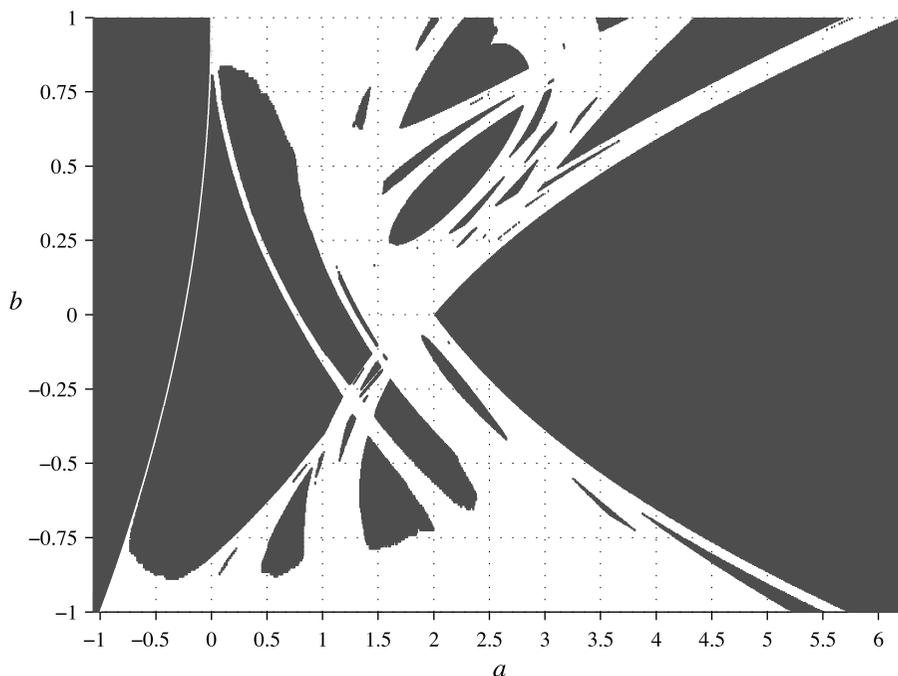


Fig. 7. Uniformly hyperbolic plateaus for the Hénon map.

We remark that Hruska [19] also constructed a rigorous numerical method for proving hyperbolicity of the complex Hénon map. The main difference between our method and Hruska’s method is that our method does not prove hyperbolicity directly. Instead, we prove quasi-hyperbolicity, which is equivalent to uniform hyperbolicity under the assumption of chain recurrence. This rephrasing enables us to avoid the computationally expensive procedure of constructing a metric adapted to the hyperbolic splitting.

We emphasize that our method can also be applied to higher dimensional dynamical systems. For example, a conjecture on the topology of the horseshoe locus of the complex Hénon map has been verified with this method [2].

5.2. Hyperbolicity and quasi-hyperbolicity

First we recall the definition of uniform hyperbolicity. Let f be a diffeomorphism on a manifold M and Λ a compact invariant set of f .

DEFINITION 5.2. *We say that f is uniformly hyperbolic on Λ if $TM|_{\Lambda}$ splits into a direct sum $TM|_{\Lambda} = E^s \oplus E^u$ of two Tf -invariant subbundles, and there exist constants $c > 0$ and $0 < \lambda < 1$ as well as a metric $\|\cdot\|$ on M , such that $\|Tf^n|_{E^s}\| < c\lambda^n$ and $\|Tf^{-n}|_{E^u}\| < c\lambda^n$ hold for all $n \geq 0$.*

Unfortunately, this conventional definition is not computation-friendly. According to this definition, we must control constants c , λ and at the same time, we also need to construct a metric adapted to the hyperbolic splitting.

To avoid this difficulty, we introduce a weaker notion of hyperbolicity.

DEFINITION 5.3. *We say that f is quasi-hyperbolic on Λ if there is no bounded orbit of $Tf: TM|_{\Lambda} \rightarrow TM|_{\Lambda}$ other than the orbit consisting of zero vectors.*

It is easy to see that uniform hyperbolicity implies quasi-hyperbolicity. The converse is false in general; however, if $f|_{\Lambda}$ is chain recurrent then they are equivalent.

THEOREM 5.4 ([6, 37]). *Assume that $f|_{\Lambda}$ is chain recurrent, that is, $\mathcal{R}(f|_{\Lambda}) = \Lambda$. Then f is uniformly hyperbolic on Λ if and only if f is quasi-hyperbolic on it.*

Note that the linearity of Tf in fibers implies that if $\{v_i\}_{i \in \mathbb{Z}}$ is a non-zero bounded orbit of Tf , then so is its multiplication $\{\mu \cdot v_i\}_{i \in \mathbb{Z}}$ where μ is a non-zero constant. Hence if f is not quasi-hyperbolic, any compact neighborhood N of the zero-section of $TM|_{\Lambda}$ contains a non-zero bounded orbit. Therefore, quasi-hyperbolicity is equivalent to saying that the zero section of $TM|_{\Lambda}$ is an isolated invariant set with respect to $Tf: TM|_{\Lambda} \rightarrow TM|_{\Lambda}$. Consequently, we have the following more computation-friendly condition for quasi-hyperbolicity.

PROPOSITION 5.5. *Assume that $N \subset TM|_{\Lambda}$ is an isolating neighborhood with respect to $Tf: TM|_{\Lambda} \rightarrow TM|_{\Lambda}$ and N contains the image of the zero-section of $TM|_{\Lambda}$. Then f is quasi-hyperbolic on Λ .*

5.3. Algorithm

For simplicity, we assume that $M = \mathbb{R}^n$ and consider a family of diffeomorphisms $f_{\lambda}: \mathbb{R}^n \rightarrow \mathbb{R}^n$ that depends on a d -dimensional parameter $\lambda \in \mathbb{R}^d$. We consider regions $K \subset M = \mathbb{R}^n$, $N \subset TM = \mathbb{R}^{2n}$ and $\Lambda \subset \mathbb{R}^d$, and then put rectangular grids \mathcal{H} , $\tilde{\mathcal{H}}$ and \mathcal{L} on K , N and Λ , respectively.

For a parameter set $L \in \mathcal{L}$, we choose common combinatorial representations \mathcal{F}_L and \mathcal{TF}_L of f_{λ} and Tf_{λ} , which are valid for all $\lambda \in L$. We denote their graph representations by $G(K, \mathcal{F}_L)$ and $G(N, \mathcal{TF}_L)$. For a directed graph G , the combinatorial invariant set and the strongly connected components of G are denoted by $\text{inv } G$ and $\text{scc } G$.

PROPOSITION 5.6. *For any $\lambda \in L$, $\text{inv}(K, f_{\lambda}) \subset |\text{inv } G(K, \mathcal{F}_L)|$. Furthermore, if $\mathcal{R}(f_{\lambda}) \subset \text{int } K$ holds for all $\lambda \in L$, then we have $\mathcal{R}(f_{\lambda}) \subset |\text{scc } G(K, \mathcal{F}_L)|$ for all $\lambda \in L$. The same statements hold for Tf_{λ} and $G(N, \mathcal{TF}_L)$.*

Now we describe our algorithm to prove quasi-hyperbolicity. The algorithm involves the subdivision algorithm [16], that is, if it fails to prove quasi-hyperbolicity, then it subdivides cubes in K and L to have a better approximation of the invariant set, and repeats the whole step until it succeeds with the proof.

ALGORITHM 5.7 (quasi-hyperbolicity verification for a fixed parameter set).

- (1) Find K such that $\mathcal{R}(f_{\lambda}) \subset \text{int } K$ holds for all $\lambda \in L$, and let $N := K \times [-1, 1]^n$.
- (2) Replace K with $|\text{scc } G(K, \mathcal{F}_L)|$.

- (3) Replace N with $N \cap (K \times [-1, 1]^n)$.
- (4) If $|\text{inv } G(N, \mathcal{TF}_L)| \subset K \times \text{int}[-1, 1]^n$ then stop. Otherwise, replace N with $|\text{inv } G(N, \mathcal{TF}_L)|$ and refine the grids \mathcal{H} and $\tilde{\mathcal{H}}$ by bisecting all grid elements. Then goto step (2).

THEOREM 5.8. *If Algorithm 5.7 stops, then f_λ is quasi-hyperbolic on $\mathcal{R}(f_\lambda)$ for every $\lambda \in L$.*

This theorem implies that if L contains a non-quasi-hyperbolic parameter value, then Algorithm 5.7 never stops. Therefore, if we want to apply the method for a large family of diffeomorphisms, the algorithm should involve an automatic selection of parameter values. We can also use the subdivision algorithm to realize such a procedure. That is, we start with a large parameter set Λ with a trivial grid $\mathcal{L} = \{\Lambda\}$, and then inductively decompose Λ into smaller cubes. On each subdivision level, we remove the cubes in \mathcal{L} on which we have proved quasi-hyperbolicity.

ALGORITHM 5.9 (with adaptive selection of parameters).

- (1) Let $\mathcal{L} = \{\Lambda\}$ and find K such that $\mathcal{R}(f_\lambda) \subset \text{int } K$ holds for all $\lambda \in L$. Let $N := K \times [-1, 1]^n$. Save the data of K and N make a link from Λ to these data.
- (2) If $\mathcal{L} = \emptyset$ then stop. Else, choose a parameter cube $L \in \mathcal{L}$ according to a certain “selection rule” and load the data of K and N linked from L .
- (3) Apply step (2), and (3) of Algorithm 5.7 once.
- (4) If $|\text{inv } G(N, \mathcal{TF}_L)| \subset K \times \text{int}[-1, 1]^n$ then remove L from \mathcal{L} and goto step (2). Otherwise, bisect L into L_0 and L_1 . Remove L from \mathcal{L} , and add L_0 and L_1 to \mathcal{L} . Subdivide the cubes in K and N and save these data. Make links from L_0 and L_1 to these data, then goto step (2).

It follows from Theorem 5.8 that if the cube L_i is removed in the procedure of Algorithm 5.9, then L_i consists of quasi-hyperbolic parameter values.

We did not specify the “selection rule” that appears in step (2) of Algorithm 5.9. Various rules can be applied and the effectiveness of a rule depends on a specific case [1].

5.4. Application to the Hénon map

Now we apply the method developed in Section 5.2 and Section 5.3 to the chain recurrent set of the Hénon map.

To complete the step (1) of Algorithm 5.9, we need to know a priori the size of $\mathcal{R}(H_{a,b})$. Further, to apply Theorem 5.4, we also need to check that the dynamics restricted to $\mathcal{R}(H_{a,b})$ is chain recurrent. In the case of the Hénon map, this step is straightforward. Let

$$R(a, b) := \frac{1}{2} (1 + |b| + \sqrt{(1 + |b|)^2 + 4a}),$$

$$S(a, b) := \{(x, y) \in \mathbb{R}^2 : |x| \leq R(a, b), |y| \leq R(a, b)\}.$$

Then we can prove the following [17].

LEMMA 5.10. *The chain recurrent set $\mathcal{R}(H_{a,b})$ is contained in $S(a,b)$. And $H_{a,b}$ restricted to $\mathcal{R}(H_{a,b})$ is chain recurrent.*

Next, we fix the parameter region to be investigated. Note that we only need to consider the case $b \in [-1, 1]$, because the inverse of the Hénon map $H_{a,b}$ is conjugate to the Hénon map $H_{a/b^2, 1/b}$, whose Jacobian is $1/b$, and the hyperbolicity of a diffeomorphism is equivalent to that of its inverse. Further, we can restrict our computation to the case $(a, b) \in [-1, 12] \times [-1, 1]$, for otherwise it follows from the proof of [17] that $\mathcal{R}(H_{a,b})$ is hyperbolic or empty.

Therefore, we start with the parameter set $(a, b) \in [-1, 12] \times [-1, 1]$. Let $K := [-8, 8] \times [-8, 8]$ and $L = K \times [-1, 1]^2$. Then Lemma 5.10 implies $\mathcal{R}(H_{a,b}) \subset \text{int } K$ holds for these parameter values. By running Algorithm 5.9 with this initial data, we can prove Theorem 5.1.

6. Rigorous estimates of uniform expansion

Another family of fast graph algorithms can be applied to obtain rigorous bounds for the uniform expansion exponent in C^1 maps of the interval outside some critical neighborhood. The computed quantity is a rigorous lower bound for the Lyapunov exponent which gives important information about the map on the selected interval. An open neighborhood of critical points must be excluded from these calculations, or otherwise the result is not computable.

The rough idea of this approach is to subdivide the complement of the critical neighborhood into a finite number of small intervals which are then taken as vertices in a graph. Edges connecting each pair of vertices related by the map are assigned weights corresponding to a lower estimate of the derivative exponent for points mapped from one interval to the other. Then algorithms which compute minimum path weights or minimum average path weights can be used to obtain a lower bound for the expansion exponent along trajectories. Details are explained in this section, based on [15].

6.1. Introduction

Let $f: I \rightarrow I$ be a continuous map of the interval $I \subset \mathbb{R}$, and let $D \subset I$ be an open set such that f is differentiable on $I \setminus D$ and $f'(x) \neq 0$ for $x \in I \setminus D$. The set D is called a *critical set*. Fix some $\delta > 0$ and let Δ be an open δ -neighborhood of D ; in particular, if $D = \{0\}$ then $\Delta = (-\delta, \delta)$. In what follows, a constructive computer-assisted method for proving the following theorem is described:

THEOREM 6.1. *There exist constants $C, \lambda > 0$ such that for every x and for every n such that $f^i(x) \notin \Delta$ for all $i = 0, \dots, n-1$ we have*

$$|Df^n(x)| \geq Ce^{\lambda n}. \quad (8)$$

For the purpose of interval arithmetic computations, a subdivision of $I \setminus \Delta$ into intervals is introduced. A finite collection of intervals $\mathcal{I} = \{I_j : j = 1, \dots, K\}$ is called an *admissible cover* of $I \setminus \Delta$ if $\text{int}(I_i \cap I_j) = \emptyset$ for $i \neq j$ and $I \setminus \Delta \subseteq \bigcup_{j=1}^K I_j$.

6.2. Graph algorithms and the analysis of f

A directed graph $G = (V, E)$ (see Section 4.2) equipped with the *weight function* $w: E \rightarrow \mathbb{R}$ is called a *weighted graph*. Such a graph is called a *representation* of f on $I \setminus \Delta$ if the following conditions are satisfied:

- $V = \mathcal{I} \cup \{\Delta\}$, where \mathcal{I} is an admissible cover of $I \setminus \Delta$,
- the set of edges E contains all the pairs $(I_1, I_2) \in \mathcal{I} \times V$ such that $f(I_1) \cap I_2 \neq \emptyset$,
- for all $I_1, I_2 \in V$, if $f(I_1) \cap I_2 \neq \emptyset$ then

$$w(I_1, I_2) \leq \min\{\log|Df(x)|: x \in I_1 \cap f^{-1}(I_2)\}. \tag{9}$$

Representing f in terms of a graph G satisfying these conditions provides a straightforward relationship between the weight of a path in G and the derivative along points whose orbit is described by this path. Namely, given a point $x \in I \setminus \Delta$ and a path (I_0, \dots, I_n) in G such that $f^j(x) \in I_j$, we have

$$\log|Df^n(x)| = \sum_{j=0}^{n-1} \log|Df(f^j(x))| \geq \sum_{j=0}^{n-1} w(I_j, I_{j+1}). \tag{10}$$

This observation implies the following result proved in [15]:

PROPOSITION 6.2. *If $\lambda \in \mathbb{R}$ is not greater than the mean weight of any cycle in G , then there exists $C > 0$ such that*

$$|Df^n(x)| \geq Ce^{\lambda n} \tag{11}$$

for any $n > 0$ and $x \in I$ such that $f^i(x) \notin \Delta$ for all $i = 0, \dots, n - 1$.

Both λ and C can be computed effectively using known algorithms developed for the analysis of general weighted graphs. Karp’s algorithm [22] can be used to determine λ as a lower bound for the minimum mean cycle weight. Then Floyd–Warshall algorithm [10, Section 25.2] or Johnson’s algorithm [10, Section 25.3] applied to the graph G with the modified weight function $w'(e) := w(e) - \lambda$ can be used to determine $\log C$ as the minimum path weight. If the computed value λ is positive then Theorem 6.1 is proven. Otherwise, either (8) does not hold true for any positive λ , C , or the expansion is so small that it cannot be detected due to overestimates introduced both by estimating the logarithm of the derivatives on entire intervals of the admissible cover \mathcal{I} of $I \setminus \Delta$, and by rounding errors that appear in the computation of the mean cycle weights in the graph with interval arithmetic.

6.3. Applications

As an example application of the algorithms introduced in this section, the well-known family of unimodal maps $f_a: I \rightarrow I$ given by

$$f_a(x) = x^2 - a, \tag{12}$$

where $I = [-2, 2]$ and $a \in [1.7, 2]$, is considered in [15]. First of all, the parameter value $a = 2$ is fixed and several computations and comparisons are carried out. Then other parameters of computations (δ , K , and \mathcal{T}) are fixed and the computations are carried out for several thousand equally spaced parameter values in the parameter range under consideration.

Since the graph G and the weight function w are both computed using interval arithmetic for the evaluation of f and f' , no substantial change in the algorithms is necessary to obtain results valid for all f_a for a in some (small) interval $[a_1, a_2]$. Fig. 8 shows the result of these calculations conducted with the software referred to in [15] to obtain explicit lower bounds for the expansion constant $\lambda(a)$ in (11) for all $a \in [1.5, 2]$, and thus proving Theorem 6.1 for those values of a for which the computed value of $\lambda(a)$ is positive.

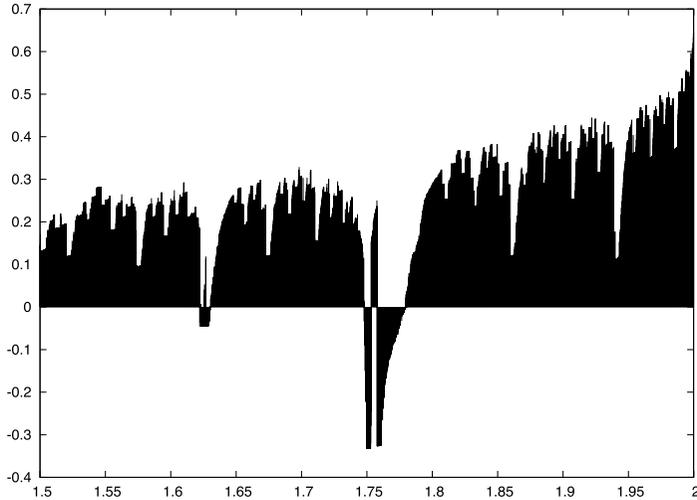


Fig. 8. Rigorous lower bound of the expansion exponent λ computed for 4096 adjacent intervals which fill the parameter interval $[1.5, 2]$ for the quadratic map (12).

6.4. Additional estimates

In [15], additional algorithms are also provided to prove another estimate similar to Theorem 6.1:

THEOREM 6.3. *There exists a constant $\lambda_0 > 0$ such that for every x and for every n such that $f^i(x) \notin \Delta$ for all $i = 0, \dots, n - 1$, and additionally $x \in f(\Delta)$ or $f^n(x) \in \Delta$, we have*

$$|Df^n(x)| \geq e^{\lambda_0 n}. \tag{13}$$

This estimate is in some sense stronger than the one provided by Theorem 6.1, because the positive constant C in (8) can be very small, and $C = 1$ in (13). However, this estimate only applies to sections of orbits that either leave or enter the

critical neighborhood. Obtaining a rigorous estimate for λ_0 includes the computation of a lower bound for the mean average weight of all simple paths, that is, paths that do not contain any cycle, with the additional condition that they begin at one of given vertices of the graph (the vertices corresponding to intervals intersecting $f(\Delta)$), or which end at a given vertex (the vertex corresponding to Δ). An algorithm based on Karp's algorithm is developed in [15] specifically for that purpose.

6.5. Remarks and comments

We would like to emphasize how little needs to be known on the function f in order to apply this algorithmic method for computing a lower bound for the expansion exponent λ and the constant C in (8). In addition to knowing the interval I and the critical neighborhood Δ , one only needs to be able to find a covering of $f(I_i)$ by intervals in \mathcal{I} for each $I_i \in \mathcal{I}$, where \mathcal{I} is an a priori fixed admissible cover of $I \setminus \Delta$, and also one needs to be able to provide a rigorous lower bound for the logarithm of the derivative of f on each interval $I_i \in \mathcal{I}$ for equation (9), preferably taking into consideration also the preimages of intervals in \mathcal{I} and the preimage of Δ . Note that this does not even require the knowledge of the exact formula for f , just some reasonable estimates for the above-mentioned quantities are sufficient. This is a very important feature of this method which decides on its wide applicability.

It turns out that the choice of the admissible cover \mathcal{I} of $I \setminus \Delta$ may have substantial impact on the accuracy of the results of computations. Some examples in [15] suggest that smaller intervals in the area in which $|f'|$ is large help provide a much better estimate for λ while keeping the total number K of intervals in \mathcal{I} constant, and thus not increasing the cost of computations significantly.

7. Concluding remarks

The results described in this paper illustrate new trends in the numerical approach to dynamical systems. Numerical simulations contributed to groundbreaking discoveries in the 60–70s (the detection of chaotic dynamics by Lorenz, for example) and their importance is still undeniable; however, the increasing power of contemporary computers prompted the development of more advanced methods which not only reveal various phenomena but also provide mathematically confident results.

In the applications described in this paper, one can see two opposite trends of using interval arithmetic for the analysis of dynamical systems. On the one hand, one can see the tendency of limiting the usage of rigorous numerics only to generate input for complicated combinatorial algorithms which provide the results through extensive computations involving fast graph algorithms and other abstract methods, like computational algebraic topology. On the other hand, the development of advanced numerical methods which use interval arithmetic not only to provide rigorous bounds on images of functions but also their derivatives of several orders, opens new possibilities of much more thorough analysis of dynamics.

It will be very interesting to see the directions of future development of applications of interval arithmetic in dynamical systems, combining rigorous numerics with variational methods, statistical methods, and other abstract constructs to provide automatized tools for instant and thorough investigation of dynamical systems.

References

- [1] Z. Arai, On hyperbolic plateaus of the Hénon map. *Experimental Math.*, **16** (2007), 181–188.
- [2] Z. Arai, On loops in the hyperbolic locus of the complex Hénon map and their monodromies. Preprint.
- [3] Z. Arai, W. Kalies, H. Kokubu, K. Mischaikow, H. Oka and P. Pilarczyk, A database schema for the analysis of global dynamics of multi-parameter systems. To appear in *SIAM J. Appl. Dynam. Sys.*
- [4] Z. Arai and K. Mischaikow, Rigorous computations of homoclinic tangencies. *SIAM J. Appl. Dynam. Sys.*, **5** (2006), 280–292.
- [5] H. Ban and W. Kalies, A computational approach to Conley’s decomposition theorem. *J. Comp. Nonlinear Dynam.*, **1** (2006), 312–319.
- [6] R.C. Churchill, J. Franke and J. Selgrade, A geometric criterion for hyperbolicity of flows. *Proc. Amer. Math. Soc.*, **62** (1977), 137–143.
- [7] Computer Assisted Proofs in Dynamics. <http://capd.wsb-nlu.edu.pl/>.
- [8] Computational Homology Project. <http://chomp.rutgers.edu/>.
- [9] C. Conley, *Isolated Invariant Sets and the Morse Index*. CBMS Regional Conference Series in Math., No. 38, Amer. Math. Soc., Providence, RI, 1978.
- [10] T. Cormen, C. Leiserson, R. Rivest and C. Stein, *Introduction to Algorithms*, 2nd edition. MIT Press, Cambridge, MA, 2001.
- [11] S. Day, O. Junge and K. Mischaikow, A rigorous numerical method for the global analysis of infinite dimensional discrete dynamical systems. *SIAM J. Appl. Dynam. Sys.*, **3** (2004), 117–160.
- [12] F. Dumortier, S. Ibáñez and H. Kokubu, New aspects in the unfolding of the nilpotent singularity of codimension three. *Dynam. Syst.*, **16** (2001), 63–95.
- [13] F. Dumortier, S. Ibáñez and H. Kokubu, Cocoon bifurcation in three dimensional reversible vector fields. *Nonlinearity*, **19** (2006), 305–328.
- [14] M.J. Davis, R.S. MacKay and A. Sannami, Markov shifts in the Hénon family. *Physica D*, **52** (1991), 171–178.
- [15] S. Day, H. Kokubu, S. Luzzatto, K. Mischaikow, H. Oka and P. Pilarczyk, Quantitative hyperbolicity estimates in one-dimensional dynamics. *Nonlinearity*, **21** (2008), 1967–1987.
- [16] M. Dellnitz and O. Junge, Set oriented numerical methods for dynamical systems. *Handbook of Dynamical Systems II*, North-Holland, 2002, 221–264.
- [17] R. Devaney and Z. Nitecki, Shift automorphisms in the Hénon mapping. *Commun. Math. Phys.*, **67** (1979), 137–146.
- [18] Y. Hiraoka, T. Ogawa and K. Mischaikow, Conley-index based numerical verification method for global bifurcations of the stationary solutions to the Swift–Hohenberg equation. *Trans. JJIAM*, **13** (2003), 191–211, in Japanese.
- [19] S.L. Hruska, A numerical method for proving hyperbolicity of complex Hénon mappings. *Found. Comp. Math.*, **6** (2006), 427–455.
- [20] T. Kaczynski, K. Mischaikow and M. Mrozek, *Computational Homology*. Applied Mathematical Sciences, Vol. 157, Springer-Verlag, New York, 2004.
- [21] W. Kalies, K. Mischaikow and R.C.A.M. VanderVorst, An algorithmic approach to chain recurrence. *Found. Comp. Math.*, **5** (2005), 409–449.
- [22] R. Karp, A characterization of the minimum cycle mean in a digraph. *Discrete Math.*, **23** (1978), 309–311.
- [23] H. Kokubu, D. Wilczak and P. Zgliczyński, Rigorous verification of cocoon bifurcations in the Michelson system. *Nonlinearity*, **20** (2007), 2147–2174.
- [24] Y.-T. Lau, The “cocoon” bifurcations in three-dimensional systems with two fixed points. *Int. J. Bif. Chaos*, **2** (1992), 543–558.

- [25] D. Michelson, Steady solutions of the Kuramoto–Sivashinsky equation. *Physica D*, **19** (1986), 89–111.
- [26] C.K. McCord, Uniqueness of connecting orbits in the equation $Y^{(3)} = Y^2 - 1$. *J. Math. Anal. Appl.*, **114** (1986), 584–592.
- [27] C. McCord, K. Mischaikow and M. Mrozek, Zeta functions, periodic trajectories and the Conley index. *J. Diff. Eq.*, **121** (1995), 258–292.
- [28] K. Mischaikow, Topological techniques for efficient rigorous computations in dynamics. *Acta Numerica*, **11** (2002), 435–477.
- [29] K. Mischaikow, M. Mrozek and P. Pilarczyk, Graph approach to the computation of the homology of continuous maps. *Found. Comp. Math.*, **5** (2005), 199–229.
- [30] M. Mrozek, Leray functor and cohomological Conley index for discrete dynamical systems. *Trans. Amer. Math. Soc.*, **318** (1990), 149–178.
- [31] M. Mrozek, Open index pairs, the fixed point index and rationality of zeta functions. *Ergod. Th. & Dynam. Sys.*, **10** (1990), 555–564.
- [32] M. Mrozek, The Conley index on compact ANR’s is of finite type. *Results in Mathematics*, **18** (1990), 306–313.
- [33] P. Pilarczyk, Computer assisted method for proving existence of periodic orbits. *Top. Methods Nonlinear Anal.*, **13** (1999), 365–377.
- [34] P. Pilarczyk, Topological-numerical approach to the existence of periodic trajectories in ODEs. *Discrete and Continuous Dynamical Systems 2003, A Supplement Volume: Dynamical Systems and Differential Equations*, 701–708.
- [35] P. Pilarczyk, The Conley index and rigorous numerics for hyperbolic periodic trajectories in ODEs. In preparation.
- [36] P. Pilarczyk and K. Stolot, Excision-preserving cubical approach to the algorithmic computation of the discrete Conley index. *Topology Appl.*, **155** (2008), 1149–1162.
- [37] R.J. Sacker and G.R. Sell, Existence of dichotomies and invariant splitting for linear differential systems I. *J. Diff. Eq.*, **27** (1974), 429–458.
- [38] A. Szymczak, The Conley index for discrete semidynamical systems. *Topology Appl.*, **66** (1995), 215–240.
- [39] I. Ugarčević and H. Weiss, Chaotic dynamics of a nonlinear density dependent population model. *Nonlinearity*, **17** (2004), 1689–1711.
- [40] D. Wilczak and P. Zgliczyński, Topological method for symmetric periodic orbits for maps with a reversing symmetry. *Discrete Cont. Dyn. Sys. A*, **17** (2007), 629–652.
- [41] P. Zgliczyński, C^1 -Lohner algorithm. *Found. Comp. Math.*, **2** (2002), 429–465.
- [42] D. Wilczak and P. Zgliczyński, C^r -Lohner algorithm. Preprint, available at <http://arxiv.org/abs/0704.0720>.

