Rigorous lower bounds for the topological entropy via a verified optimization technique

Balázs Bánhelyi
University of Szeged
Institute of Informatics
H-6701 Szeged, Hungary
banhelyi@inf.u-szeged.hu

Tibor Csendes
University of Szeged
Institute of Informatics
H-6701 Szeged, Hungary
csendes@inf.u-szeged.hu

Barnabás M. Garay
Budapest University of Technology
Institute of Mathematics
H-1521 Budapest, Hungary
garay@math.bme.hu

Abstract

Our automatic method developed for the detection of chaos is used for finding rigorous lower bounds for the topological entropy of the classical Hénon mapping. We do this within the abstract framework created by Galias and Zgliczynski in 2001, and focus on covering graphs involving different iterations. Our results are compared to those obtained by them.

1. Introduction

Starting from the landmark paper by Mischaikow and Mrozek [11] on the Lorenz equation, computer–assisted proofs for chaos have become an integral part of dynamical systems theory. They upgraded the importance of verified numerics to general mathematics considerably.

What the computer is actually used for is the rigorous checking of a finite number of inclusions of the form

$$\mathcal{T}_j(W_j) \subset U_j \ , \quad j = 1, 2, \ldots, M, \tag{1}$$

where the W_j's and U_j's are subsets of the phase space and the $\mathcal{T}_j$'s are functions associated with the dynamics. The collection of inclusions (1) is structured by the concept of covering relations [16] and forms a sufficient condition for chaos, more precisely, for the embeddability of a certain type of symbolic dynamics into the mapping or differential equation under investigation.

The topological entropy of the embedded symbolic dynamics can be easily determined on the basis of inclusions (1) and is a lower bound for the topological entropy of the mapping or of a Poincaré mapping of the differential equation considered. Computer–assisted proofs for chaos lead to positive lower bounds for topological entropy. Positive topological entropy is an important qualitative indicator for chaos. The larger the topological entropy the stronger chaos. For diffeomorphisms, positive topological entropy implies positivity of the maximal Lyapunov exponent.

The real task is not to check inclusions (1) but to find the W_j's, U_j's, and $\mathcal{T}_j$'s involved in a sufficient condition for chaos.

In a recent paper of ours [4], this task is formulated as a constraint satisfaction problem in optimization theory and thus the possibility of applying algorithms of global optimization in computer–assisted proofs for chaos opened. As for the first application, we reproved Zgliczynski's result [16] on embeddability of the full shift on two symbols—shortly: the existence of Σ_2–chaos—in the seventh iterate of the widely investigated Hénon mapping

$$\mathcal{H} : \mathbb{R}^2 \to \mathbb{R}^2 \ , \ (x, y) \to (1 + y - ax^2, bx)$$

with the classical parameters

$$a = 1.4 \ \text{ and } \ b = 0.3 \, .$$

Our proof in [4] differs strikingly from all computer–assisted proofs for chaos (we mean those based on embedding symbolic dynamics we are aware of): There is no trial and error interaction between computer and computer scientist, and the tedious task of adjusting the distinguished sets W_j, U_j in (1) by hand is left entirely to the computer. Nevertheless, a reasonable choice of the search domain in the optimization procedure requires a "good initial guess" for the W_j's and U_j's exploiting a priori numerical and theoretical results on the dynamics. This is the only reason why our method cannot be termed fully automatic.

The present paper is a continuation of our previous work [4]. We focus on computing lower bounds for the topological entropy of the classical Hénon mapping based on covering graphs involving different iterations.

Though they are valid in a much greater generality [10], [12], all abstract results in Section 2 are formulated only in two dimension. We follow the presentation in Galias and

Zgliczynski [9] very closely. Covering relation and topological entropy, the two most relevant notions of dynamical systems theory are defined and thoroughly discussed in Subsections 2.1 and 2.2, respectively. We assume that the reader is (at least, intuitively) familiar with Smale's classical horseshoe presented in Figure 4(a), a well–known example for chaos. The type of combinatorial and dynamical complexity of Smale's classical horseshoe is termed as Σ_2–chaos. For dynamical systems theory in general, we refer to [14].

Section 3 is devoted to describing the interval arithmetic based checking algorithm and the optimization model for the collection of inclusions (1).

In Section 4 applications to the classical Hénon mapping are presented. The compact set in (2) is chosen for the standard trapping region, a positively invariant trapezoid for $\mathcal{H}$. It is actually the continuous mapping $\varphi = \mathcal{H}|X$, the restriction of $\mathcal{H}$ to X to which the abstract results of Section 2 and the computational methods of Section 3 are applied. Our results are compared to those obtained by Galias and Zgliczynski [6, 8, 9, 16, 17]. The greater part of them confirms earlier intuition but a small part of them gives new geometric insight on the existence of Σ_2–chaos for small and large iterates of $\mathcal{H}$.

We cannot conceal that the combination of what are called "educated guesses" and computer experimentation with human overhead is still better. The 2002 lower bound of Galias [8] for the topological entropy of the classical Hénon mapping remains the best known to date. Nevertheless, we think that automatic methods for rigorous chaos detection – including those that lead to rigorous estimates for quantitative chaos indicators – have a huge potential for the future and, eventually, from the view–point of effectiveness, they supersede those requiring repeated human interference. In particular, with further improvements, we hope that the model case applications of our verified optimization method [4] in Section 4 to bound topological entropy (the only new results of the present paper) will be competitive to those governed by hands.

2. Some abstract results in advance

2.1 The covering relation

Throughout this paper, let $Q_1, Q_2, \ldots, Q_N$ denote pairwise disjoint, closed, solid quadrangles in $\mathbb{R}^2$ with pairs of opposite edges termed horizontal and vertical. Continuing the vertical edges by parallel half–lines, two closed, solid, vertical stripes—an upper stripe and a lower stripe—are attached to each quadrangle. The upper stripe and the quadrangle share the upper horizontal edge whereas the lower stripe and the quadrangle share the lower horizontal edge of the quadrangle. Together with the attached stripes, each quadrangle separates $\mathbb{R}^2$ to a left and a right part. They are open, topological half–planes.

For $i = 1, 2, \ldots, N$, the union of the two attached stripes is denoted by E_i. Left and right vertical edges of Q_i are denoted by e_i^L and e_i^R, respectively. The corresponding left and right topological half–planes are components of the set $\mathbb{R}^2 \setminus (Q_i \cup E_i)$. They are denoted by $\mathcal{O}_i^L$ and $\mathcal{O}_i^R$, respectively.

Let X be a compact set with the property

$$\bigcup_{i=1}^{N} Q_i \subset X \subset \mathbb{R}^2, \tag{2}$$

and consider a continuous mapping $\varphi : X \to X$.

Definition 1. *Following Galias and Zgliczynski* [9]*, we say that Q_i φ^k–covers Q_j and use the notation $Q_i \overset{k}{\Longrightarrow} Q_j$ ($k = k(i,j)$ is a positive integer) if*

(i) *the image of Q_i under φ^k is located between the horizontal edges of Q_j or, equivalently,*

$$\varphi^k(Q_i) \subset \mathbb{R}^2 \setminus E_j$$

and

(ii) *the images of the vertical edges of Q_i under φ^k have empty intersection with $Q_j \cup E_j$ and they are located on the opposite sides thereof. In other words, one of the following two alternatives holds true: Either*

$$\varphi^k(e_i^L) \subset \mathcal{O}_j^L \text{ and } \varphi^k(e_i^R) \subset \mathcal{O}_j^R$$

(see Figure 1) or

$$\varphi^k(e_i^L) \subset \mathcal{O}_j^R \text{ and } \varphi^k(e_i^R) \subset \mathcal{O}_j^L \ .$$

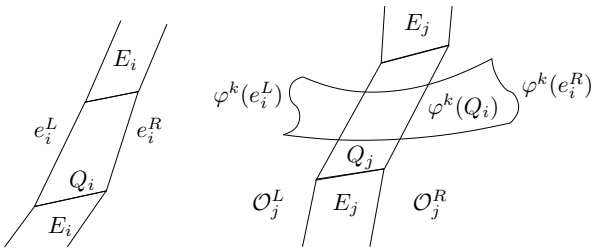


Figure 1. The covering properties

Definition 1 leads naturally to the concept of the covering graph. The covering graph is a directed graph with vertices $Q_1, Q_2, \ldots, Q_N$ and weighted edges. The pair (Q_i, Q_j) belongs to the edge set of the covering graph if $Q_i \overset{k}{\Longrightarrow} Q_j$ for some positive integer k and integer k is the weight on

this edge. Thus the covering relation $Q_i \stackrel{k}{\Longrightarrow} Q_j$ represents a weighted edge of the covering graph and vice versa. Multiple edges have different weights. Loop edges are also allowed. The *covering matrix* is an $N \times N$ matrix C with elements $\{c_{ij}\}_{i,j=1}^{N}$ where $c_{ij} = k > 0$ if $Q_i \stackrel{k}{\Longrightarrow} Q_j$ for a unique $k = k(i,j)$ and 0 otherwise. Thus the covering matrix is defined only in the absence of multiple edges.

Of course everything depends on the choice of the pairwise disjoint quadrangles $Q_1, Q_2, \ldots, Q_N$. The more careful this choice, the richer the structure of the covering graph can be, and the higher the estimate for $h(\varphi)$, the topological entropy of mapping φ.

The *expanded covering graph* is defined as follows. For $i = 1, 2, \ldots, N$, set

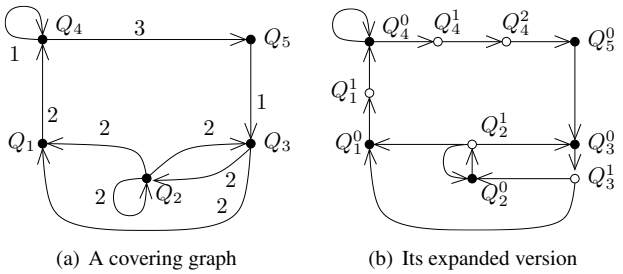
$$s(i) = \max\{k \mid Q_i \stackrel{k}{\Longrightarrow} Q_j \ , \ j = 1, 2, \ldots, N\} \, ,$$

the maximum weight on outgoing edges at Q_i. In case there are no outgoing edges from Q_i, we take $s(i) = 0$. If $0 \leq s(i) < 2$, vertex Q_i is renamed as Q_i^0. If $s(i) \geq 2$, then vertex Q_i is replaced by the string of $s(i)$ vertices and the connecting $s(i) - 1$ directed edges

$$Q_i^0 \to Q_i^1 \to \ \ldots \ \to Q_i^{s(i)-1} \, .$$

In both cases, vertex Q_i^0 of the expanded covering graph is identified with vertex Q_i of the covering graph. For $i \in \{1, 2, \ldots, N\}$ with $s(i) \geq 2$, the new, intermediate vertices $Q_i^1, Q_i^2, \ldots, Q_i^{s(i)-1}$ of the expanded covering graph are identified with the sets $\varphi(Q_i), \varphi^2(Q_i), \ldots, \varphi^{s(i)-1}(Q_i)$, respectively. Finally, edge $Q_i \stackrel{k}{\Longrightarrow} Q_j$ of the covering graph is replaced by the 'not–in–the–string' edge $Q_i^{k-1} \to Q_j^0$ of the expanded covering graph. For an example, see Figure 2.

The *expanded covering matrix* A is defined as the adjacency matrix of the expanded covering graph. In other words, A is the $N^E \times N^E$ 0–1 matrix with $a_{pq} = 1$ if there is an edge of the expanded covering graph starting from the p–th and arriving at the q–th vertex and 0 otherwise. Note that N^E is the sum of N (the number of the original vertices of the covering graph) and of $\sum_i \{(s(i) - 1) \mid s(i) \geq 2\}$ } (the number of the new, intermediate vertices).



(a) A covering graph (b) Its expanded version

Figure 2. Expanding covering graphs

The most important property of the concept of covering is summarized in the following fundamental result stating that paths and circles of the covering graph can be shadowed by true trajectories.

Theorem 1. [9] *Suppose we are given a finite chain of coverings*

$$Q_{i_0} \stackrel{k_0}{\Longrightarrow} Q_{i_1} \stackrel{k_1}{\Longrightarrow} \ \ldots \ \stackrel{k_M}{\Longrightarrow} Q_{i_{M+1}} . \tag{3}$$

For brevity, we write $\ell_0 = 0$ and, inductively, $\ell_{m+1} = \ell_m + k_m$, $m = 0, 1, \ldots, M$. Then there exists a finite sequence of points

$$x_{\ell_0} \in Q_{i_0}, x_{\ell_1} \in Q_{i_1}, \ldots, x_{\ell_{M+1}} \in Q_{i_{M+1}}$$

such that

$$\varphi^{\ell_m}(x_0) = x_{\ell_m}, \ m = 0, 1, \ldots, M+1.$$

If, in addition,

$$i_0 = i_{M+1},$$

then

$$\big(Q_{i_0} = Q_{i_{M+1}} \ and \ \big) \ x_{\ell_0} = x_{\ell_{M+1}}.$$

For $s = 0, 1, \ldots, k_m - 1$ and $m = 0, 1, \ldots, M$, define $x_{\ell_m+s} = \varphi^{\ell_m+s}(x_0)$ and observe that $x_{\ell_m+s} \in Q_{i_m}^s = \varphi^s(Q_{i_m})$. Note also that $x_{j+1} = \varphi(x_j)$ whenever $j = 0, 1, \ldots, \ell_{M+1} - 1$. Thus we are justified in saying that $\{x_j\}_{j=0}^{\ell_{M+1}}$ is a *φ–trajectory segment induced by the chain of coverings* (3). The length of this trajectory segment is $L = \ell_{M+1}$. The distance between two trajectory segments $\{x'_j\}_0^L$ and $\{x''_j\}_0^L$ of equal length is $\max_j |x'_j - x''_j|$.

2.2 Symbolic dynamics and topological entropy

For convenience, we recall here the definition of the topological entropy. As before, we assume that X is a compact set with property (2) and function $\varphi : X \to X$ is continuous.

Definition 2. *For $n = 1, 2, \ldots$ and $\varepsilon > 0$, a (necessarily finite) set $Y \subset X$ is called (n, ε)–separated if for every two different points $y, \tilde{y} \in Y$, there exists an integer $j \in [0, n)$ such that*

$$|\varphi^j(y) - \varphi^j(\tilde{y})| > \varepsilon \, .$$

With $\sharp Y$ denoting the cardinality of Y, set

$$s_n(\varepsilon) = \max \{ \, \sharp Y \mid Y \subset X \ \textit{is} \ (n, \varepsilon)\textit{-separated} \, \} \, .$$

Finally, the topological entropy of φ is defined as

$$h(\varphi) = \lim_{\varepsilon \to 0} \limsup_{n \to \infty} \frac{1}{n} \log s_n(\varepsilon) \, .$$

Roughly speaking, topological entropy measures the scaled number of metrically different trajectory segments of increasing length. ($s_n(\varepsilon)$ itself means the maximum number of discrete points that can be packed into X before there exists two of them that do not separate by ε after n iterates.)

Theorem 2. [9] *Using the terminology adopted in the last paragraph of the previous subsection, assume that*

(a) *for some $\varepsilon_0 > 0$, the distance between any two φ–trajectory segments of equal length which are induced by different chains of coverings is at least ε_0.*

Then

$$h(\varphi) \geq \log \lambda_1(A),$$

where $h(\varphi)$ is the topological entropy of φ and $\lambda_1(A)$ is the dominant eigenvalue of the expanded covering matrix.

Note that assumption (a) is considerably weaker than each of the standard assumptions

(b) all iterates involved in the covering graph are the same, i.e., all edges of the covering graph have the same weight, say k^*

or

(c) the original quadrangles $\{Q_i^0\}_{i=1}^N$ and the new, intermediate sets $\{Q_i^1, Q_i^2, \ldots, Q_i^{s(i)-1}\}_{\{i \mid s(i)\geq 2\}}$ altogether are pairwise disjoint.

The dominant eigenvalue $\lambda_1(A)$ satisfies inequality $\lambda_1(A) \geq |\lambda_j(A)|$ for all eigenvalues of A. The existence of such an eigenvalue is part of the Frobenius–Perron theorem on nonnegative matrices [14]. Note that $\lambda_1(A) > 1$ if and only if the covering graph has two different but intersecting directed circles (a loop edge is not excluded). In view of Theorem 2, inequality $\lambda_1(A) > 1$ is a sufficient condition for chaos.

It is worth mentioning here that assumption (b) implies

$$\lambda_1(C) = k^* \left(\lambda_1(A)\right)^{k^*}$$

for the dominant eigenvalue of the covering matrix. On the other hand, assumption (c) implies that φ is semiconjugate to a subshift of finite type with transition matrix A. None of these two latter properties is a consequence of assumption (a).

The crucial task is to guarantee that assumption (a) is satisfied. The problem is to control the position of the intermediate sets. It is clearly enough to check the existence of a positive integer T with the property as follows.

Property (P) *Given any two φ–trajectory segments $\{x'_j\}_0^L$ and $\{x''_j\}_0^L$ of equal length*

$$L = k'_0 + k'_1 + \cdots + k'_{M'} = k''_0 + k''_1 + \cdots + k''_{M''}$$

and which are induced by the different chains of coverings

$$Q_{i'_0} \overset{k'_0}{\Longrightarrow} Q_{i'_1} \overset{k'_1}{\Longrightarrow} \ \ldots \ \overset{k'_{M'}}{\Longrightarrow} Q_{i'_{M'+1}}$$

and

$$Q_{i''_0} \overset{k''_0}{\Longrightarrow} Q_{i''_1} \overset{k''_1}{\Longrightarrow} \ \ldots \ \overset{k''_{M''}}{\Longrightarrow} Q_{i''_{M''+1}},$$

there exist integers $n \in [0, L]$ and $t \in [0, T]$ such that

$$x'_n \in Q_{i'_p} \quad \textit{for some integer} \quad p \in [0, M'+1]\,, \tag{4}$$

$$x''_n \in \varphi^t(Q_{i''_r}) \quad \textit{for some integer} \quad r \in [0, M''+1] \tag{5}$$

with $n = \#\left\{\textit{iterates up to } Q_{i''_r}\right\} + t$ and

$$Q_{i'_p} \cap \varphi^t(Q_{i''_r}) = \emptyset, \tag{6}$$

or, alternatively,

$$x'_n \in \varphi^t(Q_{i'_p}) \quad \textit{for some integer} \quad p \in [0, M'+1]\,, \tag{7}$$

$$x''_n \in Q_{i''_r} \quad \textit{for some integer} \quad r \in [0, M''+1] \tag{8}$$

with $n = \#\left\{\textit{iterates up to } Q_{i'_p}\right\} + t$ and

$$\varphi^t(Q_{i'_p}) \cap Q_{i''_r} = \emptyset. \tag{9}$$

If property (P) is satisfied, ε_0 in assumption (a) can be chosen as

$$\varepsilon_0 = \min\{\ \text{distance}(Q_i, \varphi^t(Q_j))\ \},$$

where the minimum is taken for all $i, j = 1, 2, \ldots, N$ and $t = 0, 1, \ldots, T$ with $Q_i \cap \varphi^t(Q_j) = \emptyset$.

3. Computer procedures to analyze the covering property

In order to check subset relations of the form $\mathcal{T}(W) \subset U$ in a rigorous way, several algorithms were developed in the last decade. They form an integral part of what is called set–valued numerics and are surveyed in [5]. The key task is, however, to establish the subset relations themselves. We assume that $\mathcal{T} : \mathbb{R}^n \to \mathbb{R}^n$ is continuous, $W \subset \mathbb{R}^n$ is compact, $U \subset \mathbb{R}^n$ is open. The major assumption is that $\mathcal{T}$, W, and U depend on some vector $\lambda \in \Lambda$ of parameters where Λ is a compact subset of $\mathbb{R}^m$. The parameter vector λ_0 has to be specified in such a way that the resulting subset relation $\mathcal{T}(\lambda_0)(W(\lambda_0)) \subset U(\lambda_0)$ is fulfilled.

In what follows the task of finding successful subset relations is modeled as a constrained optimization problem. The checking algorithm presented first will then be built in a framework optimization algorithm and numerical results provided.

3.1 A checking algorithm

The checking algorithm is a branch–and–bound procedure using interval arithmetic based inclusion functions [1, 13]. The point–to–point transformation $\mathcal{T} : \mathbb{R}^n \to \mathbb{R}^n$ is replaced by its natural interval extension $T : \mathbb{I}^n \to \mathbb{I}^n$ where $\mathbb{I}^n$ stands for the set of all closed and axis-aligned rectangles in $\mathbb{R}^n$. Note that $\mathcal{T}(x) \in T(I)$ whenever $I \in \mathbb{I}^n$ with $x \in I$. For $I, J \in \mathbb{I}^n$, $I \subset J$ implies $T(I) \subset T(J)$. The width of the rectangle I for which the i-th component is $[\underline{x}_i, \overline{x}_i]$ is defined as $w(I) = \max\{|\overline{x}_i - \underline{x}_i| \,\big|\, i = 1, 2, \ldots, n\}$. For any bounded subset S in $\mathbb{R}^n$, note that $w(T(I_j)) \to 0$ holds for all interval sequences $\{I_j\}$ with $I_j \subset S$ for all $j = 1, 2, \ldots$ and $w(I_j) \to 0$. To enclose the rounding errors and to provide verified numerical results we use the outward rounding that gives computer representable result intervals containing all the points of the real operations.

ALGORITHM 1 *The Checking Routine*

Inputs:
- ε: the user set limit size of subintervals,
- W: the argument set,
- U: the aimed set for which $\mathcal{T}(W) \subset U$ is to be checked.

1. Calculate the initial interval $I \supset W$
2. Push the initial interval into the stack
3. **while** (the stack is nonempty)
4. Pop an interval I out of the stack
5. Calculate the transformed interval $J = T(I)$
6. **if** $I \cap W \neq \emptyset$, and the condition $J \subset U$ does not hold, **then**
7. **if** the width of interval I is less than ε **then**
8. push the subintervals into the output list ($\mathcal{J}$)
9. **else** bisect I along the widest side: $I = I_1 \cup I_2$
10. push the subintervals into the stack
11. **endif**
12. **endif**
13. **end while**
14. **print** that $\mathcal{T}(W) \subset U$ is proven and **stop**

For details as well as for a formal proof of the correctness of Algorithm 1, see our paper [4].

3.2 The accompanying optimization problem

Each relation $\mathcal{T}_j(W_j) \subset U_j$, $j = 1, 2, \ldots, M$ is analyzed separately. The j–th execution of Algorithm 1 may result in an interval $I_j = I_j(\lambda)$ such that $I_j \cap W_j \neq \emptyset$ but $T_j(I_j) \subset U_j$ does not hold true. This means that the j–th execution of Algorithm 1 ends at Step 8 — let $\mathcal{J} = \mathcal{J}(\lambda) = \{j_1, \ldots, j_\ell\} \subset \{1, 2, \ldots, M\}$ denote the set of such indices. Otherwise, for $j \notin \{j_1, \ldots, j_\ell\}$, the j–th execution of Algorithm 1 ends at Step 14.

Consider the optimization problem

$$\min_{\lambda \in \Lambda} g(\lambda)$$

where

$$g(\lambda) = p\left(\sum_{j \in \mathcal{J}(\lambda)} \max_{v \in T_j(I_j(\lambda))} \inf_{u \in U_j(\lambda)} |u - v|\right), \quad (10)$$

with $p(r) = r + 1$ if r is positive and $p(r) = 0$ otherwise. Here $I_j(\lambda)$ is the interval returned by the checking routine for $j \in \mathcal{J}(\lambda)$ (and the empty set for $j \notin \mathcal{J}(\lambda)$), $\Lambda \subset \mathbb{R}^m$ is the search set (the compact set of admissible parameter values), and $p : \mathbb{R} \to \mathbb{R}$ is the penalty function. Note that the j's summation term $\max_{v \in T_j(I_j(\lambda))} \inf_{u \in U_j(\lambda)} |u - v|$ of the argument of the penalty function in (10) is a nondifferentiable function of λ and stands for the Hausdorff distance of the transformed subinterval $T_j(I_j(\lambda))$ to the set $U_j(\lambda)$, a value proportional to the measure of how much condition $J \subset U$ (i.e. condition $T_j(I_j(\lambda)) \subset U_j(\lambda)$, $j \in \mathcal{J}(\lambda)$ in Step 6 of Algorithm 1) is violated.

The computation of this Hausdorff distance requires some geometry. It is an elementary task provided that $n = 2$ and that the boundary of each $U_j(\lambda)$ consists of a moderate number of finite or infinite straight line segments.

The penalty function p adds a fixed penalty term in case at least one of the constraints is not satisfied. Hence, if an optimization algorithm leads to a parameter vector λ_0 with $g(\lambda_0) = 0$, then – at the same time – the built-in checking routine provides a guaranteed reliability computational proof of the respective subset relations $\mathcal{T}_j(\lambda_0)(W_j(\lambda_0)) \subset U_j(\lambda_0)$, $j = 1, 2, \ldots, M$. Unfortunately, due to the high degree of nonlinearity of the problem, it is well possible that the output of the optimization algorithm is inconclusive, even if $\min_{\lambda \in \Lambda} g(\lambda) = 0$.

The emerging global optimization problem has been solved by a reliable clustering stochastic optimization method which goes back to [3]. This method is able to find all global optimizer points in search domains of moderate dimension and does not use the differentiability of the objective function. For a detailed discussion of this optimization model and of the relevant techniques of global optimization, see our paper [4].

In all the applications below, a typical parameter is a coordinate of a vertex of a quadrangle. The search domains for the coordinates of the individual vertices are suggested by the position of the periodic points of the Hénon map [7] as well as by basic facts on homoclinic saddles.

4. Applications

4.1 A local improvement

Galias and Zgliczynski [9] considered the following configuration of five quadrangles

$$Q_i = \mathrm{conv}\{V_{ul}^{Q_i}, V_{ur}^{Q_i}, V_{ll}^{Q_i}, V_{lr}^{Q_i}\}\ ,\quad i = 1,2,3,4,5,$$

the closed convex hulls of their respective upper left, upper right, lower left, and lower right vertices. The coordinates of these vertices are

$$\begin{array}{rclrcl}
V_{ul}^{Q_1} &=& (-0.95, 0.39)\,, & V_{ur}^{Q_1} &=& (-0.81, 0.38)\,,\\
V_{ll}^{Q_1} &=& (-0.95, 0.29)\,, & V_{lr}^{Q_1} &=& (-0.77, 0.28)\,,\\
V_{ul}^{Q_2} &=& (-0.80, 0.39)\,, & V_{ur}^{Q_2} &=& (-0.31, 0.34)\,,\\
V_{ll}^{Q_2} &=& (-0.75, 0.29)\,, & V_{lr}^{Q_2} &=& (-0.22, 0.24)\,,\\
V_{ul}^{Q_3} &=& (0.24, 0.30)\,, & V_{ur}^{Q_3} &=& (0.44, 0.25)\,,\\
V_{ll}^{Q_3} &=& (0.07, 0.20)\,, & V_{lr}^{Q_3} &=& (0.36, 0.15)\,,\\
V_{ul}^{Q_4} &=& (0.638, 0.28)\,, & V_{ur}^{Q_4} &=& (0.72, 0.28)\,,\\
V_{ll}^{Q_4} &=& (0.518, 0.07)\,, & V_{lr}^{Q_4} &=& (0.62, 0.07)\,,\\
V_{ul}^{Q_5} &=& (0.74, 0.23)\,, & V_{ur}^{Q_5} &=& (0.88, 0.21)\,,\\
V_{ll}^{Q_5} &=& (0.70, 0.12)\,, & V_{lr}^{Q_5} &=& (0.85, 0.10)\,.
\end{array}$$

All these data were found by human experimentation.

The covering graph and the expanded covering graph are those shown in Figure 2.

All the covering relations and properties

$$\mathcal{H}(Q_4) \cap Q_5 = \emptyset\,, \tag{11}$$

$$Q_5 \cap \mathcal{H}(Q_5) = \emptyset \tag{12}$$

were checked by interval computation. Properties (11) and (12) are crucial in proving that property (P) is fulfilled. In fact, since all the outgoing edges from Q_1, Q_2, and Q_3 have the same weight, it is enough to consider φ–trajectory segments $x'_0, x'_1, \ldots$ and $x''_0, x''_1, \ldots$ induced by the different chains of coverings $Q_4 \overset{1}{\Longrightarrow} Q_4 \Longrightarrow \ldots$ and $Q_4 \overset{3}{\Longrightarrow} Q_5 \Longrightarrow \ldots$, respectively. If

$$Q_4 \overset{1}{\Longrightarrow} Q_4 \overset{1}{\Longrightarrow} Q_4 \Longrightarrow \ldots\,, \tag{13}$$

then $x'_3 \in \mathcal{H}(Q_4)$ by $x'_2 \in Q_4$. But $x''_3 \in Q_5$ and recall (11). If

$$Q_4 \overset{1}{\Longrightarrow} Q_4 \overset{3}{\Longrightarrow} Q_5 \Longrightarrow \ldots\,, \tag{14}$$

then $x'_4 \in Q_5$. But $x''_4 \in \mathcal{H}(Q_5)$ by $x''_3 \in Q_5$ and recall (12). We see that, in both cases, relations (4)–(6) resp. (7)–(9) are satisfied. Thus Theorem 2 applies and yields that [9]

$$h(\mathcal{H}) > 0.338\,.$$

The distinction between subcases (13)–(14) is necessary because the *direct way* is blocked by $\mathcal{H}^2(Q_4) \cap Q_5 \neq \emptyset$.

Keeping Q_1, Q_2, Q_3, and Q_5 fixed, we look for a new quadrangle $\tilde{Q}_4$ so that $\tilde{Q}_4 \overset{2}{\Longrightarrow} Q_5$ and thus the modified expanded covering graph has only a single new vertex between $\tilde{Q}_4$ and Q_5. In other words, we try to replace element $c_{45} = 3$ of the original covering matrix

	Q_1	Q_2	Q_3	Q_4	Q_5
Q_1				2	
Q_2	2	2	2		
Q_3	2	2			
Q_4				1	**3**
Q_5			1		

.

by $\tilde{c}_{45} = 2$. If also property (cf. (11))

$$\mathcal{H}(\tilde{Q}_4) \subset \mathbb{R}^2 \setminus Q_5 \tag{15}$$

is fulfilled, then Theorem 2 applies in a *direct way* and yields that

$$h(\mathcal{H}) > 0.357\,,$$

the logarithm of the dominant eigenvalue of the modified expanded covering matrix $\tilde{A}$.

We are facing an optimization problem with constraints (15) and $Q_1 \overset{2}{\Longrightarrow} \tilde{Q}_4, \tilde{Q}_4 \overset{1}{\Longrightarrow} \tilde{Q}_4, \tilde{Q}_4 \overset{2}{\Longrightarrow} Q_5$. The tangent of the slope of the parallel half–lines determining union $\tilde{E}_4$ of the two attached stripes is taken for 10. This particular choice for the tangent was motivated by our earlier experience [2]. Thus we have eight parameters, the horizontal and the vertical coordinates of the four vertices of quadrangle $\tilde{Q}_4$. Fortunately, the optimization was successful and the vertices of the modified quadrangle $\tilde{Q}_4$ are

$$\begin{array}{rclrcl}
V_{ul}^{\tilde{Q}_4} &=& (0.638, 0.28)\,, & V_{ur}^{\tilde{Q}_4} &=& (0.75, 0.28)\,,\\
V_{ll}^{\tilde{Q}_4} &=& (0.518, 0.07)\,, & V_{lr}^{\tilde{Q}_4} &=& (0.65, 0.07)
\end{array}$$

– we rounded the nine–digit numbers supplied by the computer and checked the constraints again.

4.2 A global search

Now we are looking for the quadrangles Q_1, Q_2, Q_3, Q_4 with the covering matrix

	Q_1	Q_2	Q_3	Q_4
Q_1	1	2		
Q_2			2	2
Q_3			2	2
Q_4	2	2		

.

In order to apply Theorem 2, we add inclusion (cf. (15))

$$\mathcal{H}(Q_1) \subset \mathbb{R}^2 \setminus Q_2$$

as a constraint to those determined by the covering relations.

As before, the tangent of the slope of the parallel half–lines is taken for 10. The vertical coordinates of the 16 vertices of the four quadrangles are also fixed. They are chosen according to the position of the unstable manifold of the famous homoclinic saddle point $P = (0.631\ldots, 0.189\ldots)$ [14] of $\mathcal{H}$. This is an example of how a priori knowledge on the geometry helps to decrease the number of parameters in the relevant constraint satisfaction problem. More explanation will be given in Subsection 4.3 below.

We were lucky again. The solution

$$\begin{array}{rclrcl}
V_{ul}^{Q_1} &=& (0.52, 0.30)\,, & V_{ur}^{Q_1} &=& (0.78, 0.30)\,,\\
V_{ll}^{Q_1} &=& (0.38, 0.05)\,, & V_{lr}^{Q_1} &=& (0.65, 0.05)\,,\\
V_{ul}^{Q_2} &=& (0.24, 0.30)\,, & V_{ur}^{Q_4} &=& (0.51, 0.30)\,,\\
V_{ll}^{Q_2} &=& (0.07, 0.20)\,, & V_{lr}^{Q_4} &=& (0.33, 0.13)\,,\\
V_{ul}^{Q_3} &=& (-0.62, 0.34)\,, & V_{ur}^{Q_2} &=& (-0.38, 0.34)\,,\\
V_{ll}^{Q_3} &=& (-0.51, 0.24)\,, & V_{lr}^{Q_2} &=& (-0.20, 0.24)\,,\\
V_{ul}^{Q_4} &=& (-0.92, 0.38)\,, & V_{ur}^{Q_3} &=& (-0.67, 0.38)\,,\\
V_{ll}^{Q_4} &=& (-0.89, 0.28)\,, & V_{lr}^{Q_3} &=& (-0.63, 0.28)
\end{array}$$

given by the computer is illustrated in Figure 6.

The characteristic polynomial of the expanded covering matrix is

$$p(\lambda) = \lambda^8 - \lambda^7 - \lambda^6 + \lambda^5 - \lambda^4 + \lambda^3 - \lambda^2, \qquad (16)$$

and the accompanying entropy estimate is

$$h(\mathcal{H}) > 0.382. \qquad (17)$$

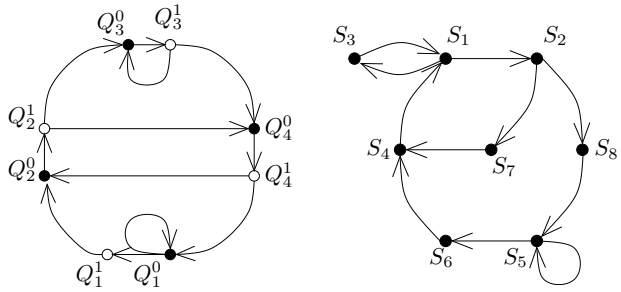
The expanded covering graph of the four–quadrangle configuration found by the computer is shown in Figure 3(a).

The (expanded) covering graph of the eight–quadrangle configuration found in Galias [8] by hand is shown in Figure 3(b). Note that assumption (b) is satisfied in Galias's example with $k^* = 1$. Hence the (expanded) covering matrix is actually the transition matrix of the embedded shift dynamics. Remarkably, the characteristic polynomial is (16) again. This means we could reproduce Galias's estimate (17) by using only four quadrangles, the half of the number of the quadrangles he needed.

4.3 On the underlying geometry

The first global search we completed successfully is illustrated in Figure 5. We looked for two quadrangles with the covering matrix

	Q_1	Q_2
Q_1	2	2
Q_2	2	2

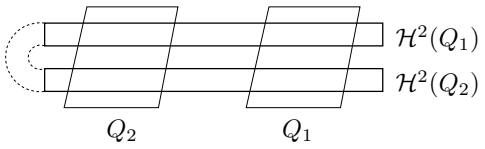


(a) Found by the computer (b) Found by Galias

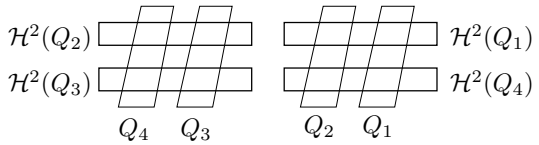
Figure 3. Expanded covering graphs with the same characteristic polynomial

and had 16 parameters again, this time the vertical and the horizontal coordinates of the eight vertices. (As before, the tangent of the slope of the parallel lines was taken to be 10.) The positive outcome of the optimization procedure means embeddability of the full shift on two symbols—shortly: the existence of Σ_2–chaos—in the second iterate of $\mathcal{H}$.

Analogies to the first steps of constructing the maximal invariant set in Smale's classical, piecewise affine horseshoe [14] and the strong resemblance of Figure 5 to the schematic picture of the piecewise linear horseshoe in Figure 4(a) suggests we might look for the four–quadrangle configuration shown schematically in Figure 4(b).



(a) Full shift on two symbols



(b) Towards full shift on four symbols

Figure 4. Piecewise affine horseshoes

But this would only repeat the entropy estimate

$$h(\mathcal{H}) \geq 2^{-1} \log 2 > 0.346$$

obtained on the basis of Figure 5. The possibility of improving the $2^{-1}\log 2$ lower bound is opened via replacing the $c_{11} = 2$ element of the covering matrix that corresponds to Figure 4(b) by $\tilde{c}_{11} = 1$, i.e., via considering the covering matrix investigated in the previous subsection. The $\tilde{c}_{11} = 1$ choice is promising because the homoclinic saddle point P,

a fixed point of $\mathcal{H}$, is contained in $Q_1 \cap \mathcal{H}^2(Q_1)$. (A periodic point of period two is contained in $Q_3 \cap \mathcal{H}^2(Q_3)$, and a periodic point of period 4 is contained in $Q_4 \cap \mathcal{H}^2(Q_2)$ and in $Q_2 \cap \mathcal{H}^2(Q_4)$ each.) This was the line of argumentation that motivated Subsection 4.2 above.

Remaining at two quadrangles, we firmly hope our optimization procedure leads to a proof of the following conjecture generalizing of what is presented in Figure 5 for $\ell = 2$.

Conjecture 1. *The full shift on two symbols embeds in the ℓ–s iterate of $\mathcal{H}$ if and only if $\ell = 2, 4$ or $\ell \geq 6$.*

Abstract theory implies embeddability only for ℓ sufficiently large. The proof of the Conjecture begins by showing that the *sufficiently large* number equals 12. This is also computer-assisted, accompanying the abstract argumentation in [14]. Cases $\ell = 2, 4, 6, 7, 8, 9, 10, 11$ had to be checked individually. The remaining cases $\ell = 1, 3, 5$ are excluded by Szymczak [15].

5. Conclusion

In the last paper of Galias [8] devoted to the subject, he presents the configuration of 29 polygons leading to the rigorous entropy estimate

$$h(\mathcal{H}) > 0.430\,.$$

This is quite near to the generally approved value of $h(\mathcal{H}) = 0.465\ldots$ conjectured on the basis of the number of periodic points of low periods. All the 29 polygons are narrow quadrangles (or quadrangles with some of the vertices "chopped off") situated along the unstable manifold of the homoclinic saddle. They were found by hand, based on earlier computer search for periodic points of low periods. If a global search is concentrated only on finding 29 segments of the unstable manifold, we need 58 parameters. This is too much for our optimization procedure which can hardly work with more than 20 parameters. So we are not able to reproduce Galias's best lower bound by our method at the moment. It remains open if a bootstrap application of our optimization procedure, keeping the number of parameters under say 10 at each step of the gradual improvements by the consecutive local searches, can reach a better entropy estimate.

Nevertheless, we think that in the near future automatic methods for rigorous chaos detection – including those that lead to rigorous estimates for quantitative chaos indicators – will be competitive to those requiring repeated human interference.

6. Acknowledgements

This work has been partially supported by the Hungarian National Science Foundation Grant OTKA No. T048377, T046822, and T049819. Barnabás Garay is also affiliated to Analogic and Neural Computing Laboratory, Computer and Automation Institute, H-1111 Budapest. The authors are grateful to the two anonymous referees for their suggestions that improved the presentation of the paper considerably.

References

[1] G. Alefeld and G. Mayer. Interval analysis: Theory and applications. *J. Comput. Appl. Math.*, 121:421–464, 2000.

[2] B. Bánhelyi, T. Csendes, and B. M. Garay. Optimization and the Miranda approach in detecting horseshoe-type chaos by computer. *International Journal of Bifurcation and Chaos*, 2007. Accepted for publication, available at `www.inf.u-szeged.hu/~csendes/publ.html`.

[3] T. Csendes. Nonlinear parameter estimation by global optimization - efficiency and reliability. *Acta Cybernetica*, 8:361–370, 1988.

[4] T. Csendes, B. M. Garay, and B. Bánhelyi. A verified optimization technique to locate chaotic regions of a Hénon system. *Journal of Global Optimization*, 35:145–160, 2006.

[5] M. Dellnitz and O. Junge. *Handbook of Dynamical Systems*, volume 3, chapter Set oriented numerical methods for dynamical systems. North-Holland, 2002.

[6] Z. Galias. Rigorous numerical studies of the existence of periodic orbits for the Hénon map. *J. Universal Comp. Sci*, 4:114–124, 1998.

[7] Z. Galias. Interval methods for rigorous investigations of periodic orbits. *Int. J. Bifurcation and Chaos*, 11:2427–2450, 2001.

[8] Z. Galias. Obtaining rigorous bounds for topological entropy for discrete time dynamical systems. In *Proc. Internat. Symposium on Nonlinear Theory and its Applications*, pages 619–622, 2002.

[9] Z. Galias and P. Zgliczynski. Abundance of homoclinic and heteroclinic connections and rigorous bounds for the topological entropy of the Hénon map. *Nonlinearity*, 14:903–932, 2001.

[10] M. Gidea and P. Zgliczynski. Covering relations for multidimensional dynamical systems. *J. Differ. Eq.*, 202:32–58, 2004.

[11] K. Mischaikow and M. Mrozek. Chaos in the Lorenz equations: a computer-assisted proof. *Bull. Amer. Math. Soc.*, 32:66–72, 1995.

[12] M. Pireddu and F. Zanolin. Fixed points for dissipative–repulsive systems and topological dynamics of mappings defined on n–dimensional cells. *Advanced Nonlinear Studies*, 5:411–440, 2005.

[13] H. Ratschek and J. Rokne. *New Computer Methods for Global Optimization*. Ellis Horwood, 1988.

[14] C. Robinson. *Dynamical Systems. Stability, Symbolic Dynamics, and Chaos*. RCR Press, 1999.

[15] A. Szymczak. A combinatorial procedure for finding isolated neighbourhoods and index pairs. *Proc. R. Soc. Edinb. A*, 127:1075–1088, 1997.

[16] P. Zgliczynski. Computer assisted proof of chaos in the Rössler equations and in the Hénon map. *Nonlinearity*, 10:243–252, 1997.

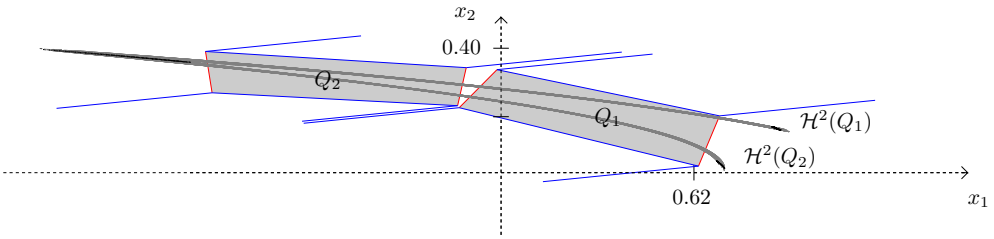


Figure 5. Two quadrangles found by computer

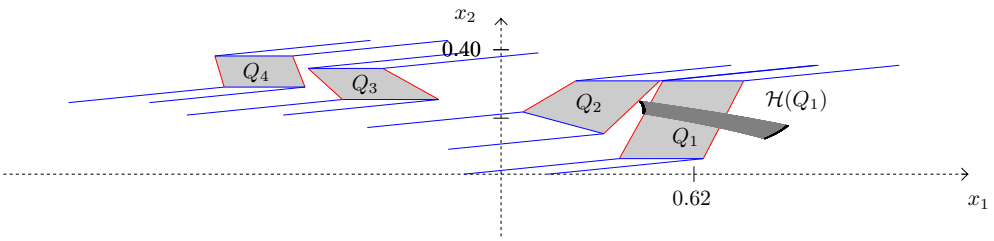


Figure 6. Four quadrangles found by computer

[17] P. Zgliczynski. Computer assisted proof of the horseshoe dynamics in the Hénon map. *Random Comput. Dynam.*, 5:1–17, 1997.