Journal of Statistical Physics, Vol. 95, Nos. 1/2, 1999

Computing the Topological Entropy for Piecewise Monotonic Maps on the Interval

Thomas Steinberger[1]

Received April 13, 1998; final October 30, 1998

A new method for computing the topological entropy of a piecewise monotonic transformation on the interval is presented. It uses a transition matrix associated with the transformation. For this matrix we give a spectral theorem. This can be used for an estimation of the accuracy of the algorithm.

KEY WORDS: Topological entropy; piecewise monotonic transformation; least square procedure; algorithm; dynamical system; Markov diagram.

1. INTRODUCTION

Dynamical systems generated by iterated maps are used to describe the chaotic behavior of physical and other natural phenomena. The topological entropy of a map is one of the quantitative measures of the complexity of these dynamical systems. In this paper we consider maps on the interval.

A map $T: [0, 1] \to [0, 1]$ is called piecewise monotonic if there are $c_i \in [0, 1]$ for $0 \leqslant i \leqslant N$ with $0 = c_0 < \cdots < c_N = 1$ such that $T|_{(c_i, c_{i+1})}$ is continuous and strictly monotone for $0 \leqslant i < N$. Set $\mathscr{Z} = \{(c_i, c_{i+1}): 0 \leqslant i < N\}$ and $\mathscr{Z}_n = \bigvee_{i=0}^{n} T^{-i}\mathscr{Z} := \{\bigcap_{i=0}^{n} T^{-i}Z_i \neq \emptyset: Z_i \in \mathscr{Z}\}$. Then $\mathscr{Z}_n$ is the set of intervals on which monotonicity of T^{n+1} holds. This is proved in Lemma 1. For $n \geqslant 0$ denote the number of elements of $\mathscr{Z}_n$ by $c_n(T)$.

For a piecewise monotonic map T on the interval we define the topological entropy $h_{\text{top}}(T)$ by

$$h_{\text{top}}(T) := \lim_{n \to \infty} \frac{1}{n} \log(c_n(T))$$

[1] Institut für Mathematik, Universität Wien, A-1090 Wien, Austria.

0022-4715/99/0400-0287$16.00/0

Observe that $c_{n+m}(T) \leqslant c_n(T)\, c_m(T)$ holds, which implies the existence of the limit in the above definition (see ref. 16). In refs. 11 and 13 it is shown, that this definition is equivalent to the usual definition of topological entropy as it can be found in ref. 16.

The definition of topological entropy suggests an obvious method of computing $h_{\text{top}}(T)$: Obtain a suitable estimate of $c_n(T)$ for large n, then $\log(c_n(T))/n$ is an estimate of $h_{\text{top}}(T)$. The problem of this method is the slow convergence rate of the expression $\log(c_n(T))/n$ to $h_{\text{top}}(T)$. Even under the best circumstances, if $c_n(T) = C \exp(h_{\text{top}}(T)\, n)$, the convergence rate is $O(1/n)$, except of the special case when $C = 1$. In this paper we show that $c_n(T) = c_n e^{h_{\text{top}}(T)\, n} + r_n$ where the c_n are bounded and periodic and $|r_n| \leqslant K\alpha^n$ for a constant K and $\alpha < e^{h_{\text{top}}(T)}$. To prove this we use a transition matrix associated with T. In such a case a good method of computing $h_{\text{top}}(T)$ is to compute $c_{n_1}(T),..., c_{n_k}(T)$ for a sequence of values $n_1 < \cdots < n_k$ and to calculate any suitable line of best fit (e.g. the line of least squares) to the data pairs $(n_1, \log c_{n_1}(T)),..., (n_k, \log c_{n_k}(T))$. Then the slope of this best fit line is a good estimate of $h_{\text{top}}(T)$. Furthermore the associated transition matrix allows us to compute $c_n(T)$ in a computation time growing with n^2.

Algorithms for computing the topological entropy for the special case of unimodal maps are described in refs. 4 and 5. The approach of ref. 4 is extended in ref. 3 to maps with three monotone pieces.

Algorithms for computing the topological entropy of general piecewise monotonic transformations of the interval are presented in refs. 2, 1, and 7. All three papers use methods motivated by maps with the Markov property.

In ref. 14 two algorithms for computing entropy for higher dimensional dynamical systems are presented.

2. COMPUTING $c_n(T)$ USING THE MARKOV DIAGRAM

As described above the problem of computing $h_{\text{top}}(T)$ reduces to the problem of computing $c_n(T)$ for arbitrary $n \geqslant 0$ We present a method which uses the so called Markov diagram of T over $\mathscr{Z}$. This concept is due to Hofbauer and developed in refs. 9–12.

For an open subinterval D of a $Z_0 \in \mathscr{Z}$ the nonempty intervals among $T(D) \cap Z$ for $Z \in \mathscr{Z}$ are called successors of D. We write $D \to C$ if C is a successor of D. All successors of a D are again open subintervals of a $Z_1 \in \mathscr{Z}$ since T is piecewise monotonic. Hence building successors can be iterated.

Set $\mathscr{D}_0 = \mathscr{Z}$ and define $\mathscr{D}_i = \mathscr{D}_{i-1} \cup \{D: \exists C \in \mathscr{D}_{i-1} \text{ with } C \to D\}$. Then all the $\mathscr{D}_i$ are finite sets since $\mathscr{Z}$ is a finite set and each $D \in \mathscr{D}_{i-1}$ has not more than $card(\mathscr{Z})$ successors. Hence $\mathscr{D} = \bigcup_{i=0}^{\infty} \mathscr{D}_i$ is at most countable and we call the oriented graph $(\mathscr{D}, \rightarrow)$ with arrows $C \rightarrow D$, if D is a successor of C, the Markov diagram of $([0,1], T)$ over $\mathscr{Z}$.

A finite or infinite sequence $D_0 D_1 \ldots$ with $D_i \in \mathscr{D}$ is called path if $D_i \rightarrow D_{i+1}$ holds for $i \geqslant 0$.

The Markov diagram can be conceived as a matrix $\mathbf{M}$. Let $\mathbf{M}$ be the $\mathscr{D} \times \mathscr{D}$ matrix with $\mathbf{M}_{CD} = 1$ if $C \rightarrow D$ and $\mathbf{M}_{CD} = 0$ else. Since the number of successors of a $C \in D$ is bounded by $card(\mathscr{Z}) = N$ we get $\sum_{D \in \mathscr{D}} \mathbf{M}_{CD} \leqslant N$. Hence $\mathbf{u} \rightarrow \mathbf{uM}$ is an $l^1(\mathscr{D})$-operator and $\mathbf{v} \rightarrow \mathbf{Mv}$ is an $l^\infty(\mathscr{D})$-operator. Both operators have the same norm, denoted by $\|\mathbf{M}\|$, which equals $\sup_{C \in \mathscr{D}} |\sum_{D \in \mathscr{D}} \mathbf{M}_{CD}|$. Hence both operators have the same spectral radius denoted by $r(\mathbf{M})$ since $r(\mathbf{M}) = \lim_{n \to \infty} \sqrt[n]{\|\mathbf{M}^n\|}$. In Theorem 7 of ref. 11 it is proved that $h_{\text{top}}(T) = \log r(\mathbf{M})$ holds.

In order to investigate the matrix $\mathbf{M}$ we start with the following lemma.

Lemma 1. Suppose $Z_i \in \mathscr{Z}$ for $i \geqslant 0$ and let D be an open interval with $D \subseteq Z_0$. Set $D_0 = D$ and $D_{i+1} = T(D_i) \cap Z_{i+1}$ for $i \geqslant 0$. Furthermore set $A_i = D \cap T^{-1}(Z_1) \cap \cdots \cap T^{-i}(Z_i)$ for $i \geqslant 0$. Then

(1) $A_i = \bigcap_{j=0}^{i} T^{-j}(D_j)$ and A_i is an interval for $i \geqslant 0$

(2) $T^i(A_i) = D_i$ for $i \geqslant 0$

(3) T^{i+1} is monotone on A_i for $i \geqslant 0$.

Proof. The lemma is proved by induction. For $i = 0$ we have $A_0 = D_0 = D$ and $T^0(A_0) = A_0$. Suppose the assertions are proved for $i \geqslant 0$. Since T^{i+1} is monoton on A_i we get that $A_{i+1} = A_i \cap T^{-(i+1)}(Z_{i+1})$ is an interval. We use $f(A \cap f^{-1}(B)) = f(A) \cap B$. Since $D_j \subseteq Z_j$ for all j, we get $\bigcap_{j=0}^{i+1} T^{-j}(D_j) \subseteq A_{i+1}$. On the other hand

$$\begin{aligned}
\bigcap_{j=0}^{i+1} T^{-j}(D_j) &= \bigcap_{j=0}^{i} T^{-j}(D_j) \cap T^{-(i+1)}(T(D_i) \cap Z_{i+1}) \\
&= \bigcap_{j=0}^{i} T^{-j}(D_j) \cap T^{-(i+1)}(T(D_i \cap T^{-1}(Z_{i+1}))) \\
&\supseteq \bigcap_{j=0}^{i} T^{-j}(D_j) \cap T^{-i}(D_i) \cap T^{-(i+1)}(Z_{i+1}) \\
&= A_i \cap T^{-(i+1)}(Z_{i+1}) = A_{i+1}
\end{aligned}$$

which proves (1). To prove (2) observe

$$T^{i+1}(A_{i+1}) = T^{i+1}(A_i \cap T^{-(i+1)}(Z_{i+1})) = T^{i+1}(A_i) \cap Z_{i+1}$$
$$= T(D_i) \cap Z_{i+1} = D_{i+1}$$

The proof of (3) uses that T^{i+1}: $A_{i+1} \to T^{i+1}(A_{i+1}) = D_{i+1}$ is assumed to be monotonic, since $A_{i+1} \subseteq A_i$. Because of $D_{i+1} \subseteq Z_{i+1}$ we get that T: $D_{i+1} \to T(D_{i+1})$ is monotonic. Hence T^{i+2}: $A_{i+1} \to T(D_{i+1})$ is monotonic. ▮

The next result gives the theoretical background of the main tool for computing the topological entropy.

Theorem 1. Suppose T: $[0, 1] \to [0, 1]$ is piecewise monotonic. Then for $k \geqslant 0$

(1) there is a bijection between the set $\mathscr{Z}_k$ and the set of paths in $\mathscr{D}$ with length k starting in an element of $\mathscr{D}_0 = \mathscr{Z}$

(2) we get $c_k(T) = \|\mathbf{u}\mathbf{M}^k\|_1$ where $\mathbf{u} \in l^1(\mathscr{D})$ is given by $\mathbf{u}_D = 1$ if $D \in \mathscr{Z}$ and $\mathbf{u}_D = 0$ else.

Proof. To proof (1) we attach to every element $A = \bigcap_{i=0}^{k} T^{-i}(Z_i) \in \mathscr{Z}_k$ the path $D_0 \cdots D_k$ where $D_0 = Z_0$ and $D_{i+1} = T(D_i) \cap Z_{i+1}$ for $1 \leqslant i < k$ holds.

By Lemma 1 we get $D_i = T^i(A)$ and hence $D_i \neq \emptyset$.

This map is injective: Suppose $C_0 \cdots C_k$ is a path which is attached to $A = \bigcap_{i=0}^{k} T^{-i}(Z_i)$ as well as to $B = \bigcap_{i=0}^{k} T^{-i}(Y_i)$. Then $C_0 = Z_0 = Y_0$ holds. If $Z_i = Y_i$ is proved for $i < j \leqslant k$, then also $Z_j = Y_j$ holds, since $T(C_{j-1}) \cap Z_j = C_j = T(C_{j-1}) \cap Y_j$ and since the elements of $\mathscr{Z}$ are disjoint.

This map is surjective: Let $C_0 \cdots C_k$ be a path of length k starting in $\mathscr{Z}$. Because of Lemma 1 we get with $B := \bigcap_{i=0}^{k} T^{-i}(C_i)$ that $T^k(B) = C_k$ and hence $B \neq \emptyset$. Again by Lemma 1, $B \in \mathscr{Z}_k$ follows and the path $C_0 \cdots C_k$ is attached to B.

To prove assertion (2) observe that $\|\mathbf{u}\mathbf{M}^k\|_1$ equals the number of paths in $\mathscr{D}$ with length k starting in an element of $\mathscr{Z}$, since $\mathbf{u}_D = 1$ if $D \in \mathscr{Z}$ and $\mathbf{u}_D = 0$ else. With (1) the equation $c_k(T) = \|\mathbf{u}\mathbf{M}^k\|_1$ follows. ▮

3. THE ALGORITHM

The results, which we have presented until now, suggest the following algorithm for computing the topological entropy of a piecewise monotonic transformation T on $[0, 1]$. Let $(n_1,..., n_k)$ be any sequence of numbers for which $c_{n_j}(T)$ should be evaluated. To get an estimate of $h_{\text{top}}(T)$ one has to do the following:

(1) Build up the finite parts $\mathscr{D}_{n_j}$ of the Markov diagram for $1 \leqslant j \leqslant k$ successively. As we will see below, the orbits of length n_j of the endpoints of the intervals of monotonicity determine $\mathscr{D}_{n_j}$ completely.

(2) Compute $c_{n_j}(T)$ for $1 \leqslant j \leqslant k$ with use of Proposition 1 below. To compute $c_{n_j}(T)$ it suffices to know the finite part of the Markov diagram $\mathscr{D}_{n_j}$ since all paths of length n_j starting in $\mathscr{Z}$ are contained in $\mathscr{D}_{n_j}$.

To compute the number $c_m(T)$ for an $m \geqslant 0$ there is a fast inductive method. For $C \in \mathscr{D}_m$ and $0 \leqslant j \leqslant m$ let $N_C(j)$ be the number of paths of length j which start in an element in $\mathscr{Z}$ and end in C. Then $N_C(0) = 1$ for $C \in \mathscr{Z}$ and $N_C(0) = 0$ for $C \notin \mathscr{Z}$. Furthermore we get $N_C(j+1) = \sum_{D \to C} N_D(j)$ for $0 \leqslant j < m$ and $c_m(T) = \sum_{C \in \mathscr{D}_m} N_C(m)$. Observe that $N_C(m) = 0$ if C is not an element of $\mathscr{D}_m$.

(3) Calculate the least squares line (or any other line of best fit) through the data pairs $(n_1, \log c_{n_1}(T)),..., (n_k, \log c_{n_k}(T))$. The slope of this line is an estimate of $h_{\text{top}}(T)$.

A piecewise monotonic transformation $T: [0, 1] \to [0, 1]$ is called piecewise increasing if T is increasing on its intervals of monotonicity. This assumption simplifies the computation of the Markov diagram since a piecewise increasing transformation preserves the ordering. In ref. 10 it is described how to construct a piecewise increasing transformation $\hat{T}: [0, 1] \to [0, 1]$ for every piecewise monotonic transformation T such that $c_n(\hat{T}) = 2c_n(T)$ holds for every $n \geqslant 0$. Hence it suffices to describe the Markov diagram of piecewise increasing transformations.

Let $T: [0, 1] \to [0, 1]$ be piecewise increasing and let $\mathscr{Z} = \{Z_i = (x_i, y_i): 1 \leqslant i \leqslant N\}$ be the intervals of monotonicity, where $x_{i+1} = y_i$. Furthermore set $y_0 := x_1 = 0$ and $x_{N+1} := y_N = 1$.

If there is a $j > 0$ with $\lim_{y \searrow x_i} T^j(y) = x_k$ for $1 \leqslant k \leqslant N$ set $n^+(x_i) = \min\{j > 0: \lim_{y \searrow x_i} T^j(y) = x_k$ for a k with $1 \leqslant k \leqslant N\}$. Otherwise set $n^+(x_i) = \infty$.

Analogously we define: If there is a $j > 0$ with $\lim_{y \nearrow y_i} T^j(y) = y_k$ for $1 \leqslant k \leqslant N$ set $n^-(y_i) = \min\{j > 0: \lim_{y \nearrow y_i} T^j(y) = y_k$ for a k with $1 \leqslant k \leqslant N\}$. Otherwise set $n^-(y_i) = \infty$.

Set $d(i, 0) = i$ and for $0 < j < n^+(x_i)$ choose $d(i, j)$ with $1 \leqslant d(i, j) \leqslant N$, such that $T^j(x_i) \in (x_{d(i, j)}, y_{d(i, j)})$. For $j = n^+(x_i) < \infty$ choose $d(i, j)$ with $0 \leqslant d(i, j) < N$, such that $\lim_{y \searrow x_i} T^j(y) = x_{d(i, j)+1}$.

Analogous set $e(i, 0) = i$ and choose for $0 < j < n^-(y_i)$ the number $e(i, j)$ with $1 \leqslant e(i, j) \leqslant N$, such that $T^j(y_i) \in (x_{e(i, j)}, y_{e(i, j)})$. For $j = n^-(y_i) < \infty$ let $e(i, j)$ with $1 < e(i, j) \leqslant N+1$ be such that $\lim_{y \nearrow y_i} T^j(y) = y_{e(i, j)-1}$.

Set $R(i, 0) = S(i, 0) = 0$ and define inductively for $j \geqslant 1$, as long as all appearing numbers exist:

$$r(i, j) = \min\{k \geqslant 1: d(i, R(i, j-1)+k) \neq e(d(i, R(i, j-1)), k)\}$$

and if $r(i, j) < \infty$ set $R(i, j) = R(i, j-1) + r(i, j)$

$$s(i, j) = \min\{k \geqslant 1: d(e(i, S(i, j-1)), k) \neq e(i, S(i, j-1)+k)\}$$

and if $s(i, j) < \infty$ set $S(i, j) = S(i, j-1) + s(i, j)$.

Observe that $r(i, 1) = s(i, 1)$ holds. If $n^+(x_i) < \infty$, then $R(i, j) = n^+(x_i)$ for a $j \geqslant 1$. In this case $r(i, k)$ and $R(i, k)$ remain undefined for $k > j$. The same holds if $n^-(y_i) < \infty$.

Proposition 1. Set $A^i_0 = Z_i$ for $1 \leqslant i \leqslant N$ and for $j \geqslant 1$ set

$$A^i_{R(i, j)} = (T^{R(i, j)}(x_i), y_{d(i, R(i, j))})$$

and

$$B^i_{S(i, j)} = (x_{e(i, S(i, j))}, T^{S(i, j)}(y_i))$$

If $k \neq R(i, j)$ set $A^i_k = T(A^i_{k-1})$ and if $k \neq S(i, j)$ set $B^i_k = T(B^i_{k-1})$. Then the Markov diagram $(\mathscr{D}, \rightarrow)$ consists of the following elements and arrows:

(1) $A^i_{k-1} \rightarrow A^i_k$ for $1 \leqslant i \leqslant N$ and $1 \leqslant k < n^+(x_i)$

(2) $A^i_{R(i, 1)-1} \rightarrow B^i_{r(i, 1)}$ if $r(i, 1) < n^-(y_i)$

(3) $B^i_{k-1} \rightarrow B^i_k$ for $1 \leqslant i \leqslant N$ and $S(i, 1) < k < n^-(y_i)$

(4) $A^i_{R(i, j)-1} \rightarrow B^{d(i, R(i, j-1))}_{r(i, j)}$ for $j > 1$ and $r(i, j) < n^-(y_{d(i, R(i, j-1))})$

(5) $A^i_{R(i, j)-1} \rightarrow A^l_0$ for $d(i, R(i, j)) < l < e(d(i, R(i, j-1)), r(i, j))$

(6) $B^i_{S(i, j)-1} \rightarrow A^{e(i, S(i, j-1))}_{s(i, j)}$ for $j > 1$ and $s(i, j) < n^+(x_{e(i, S(i, j-1))})$

(7) $B^i_{S(i, j)-1} \rightarrow A^l_0$ for $j > 1$ and $d(e(i, S(i, j-1)), s(i, j)) < l < e(i, S(i, j))$.

Proof. The proof is given in Part II of ref. 9. ∎

The above proposition shows that the Markov diagram is determined by the orbits of the endpoints of the intervals of monotonicity. It implies that $\mathscr{D}_n$ contains at most $N(2n-1)$ different sets. Also the inductive method of computing $N_C(m)$ given above can be described more exactly with Proposition 1. Suppose $(N_C(m))_{C \in \mathscr{D}_m}$ is computed, then $(N_D(m+1))_{D \in \mathscr{D}_{m+1}}$ is built up in the following way:

Start with $N_D(m+1)=0$ for all $D \in \mathscr{D}_{m+1}$. Add $N_{A^i_{k-1}}(m)$ to $N_{A^i_k}(m+1)$ and $N_{B^i_{k-1}}(m)$ to $N_{B^i_k}(m+1)$. This corresponds to Proposition 1 (1) and (3). If $k \neq R(i, j)$ for all j then this is the only computation one has to do with $N_{A^i_{k-1}}(m)$ and if $k \neq S(i, j)$ for all j then this is the only computation one has to do with $N_{B^i_{k-1}}(m)$. If $k = R(i, j)$ for some j then add $N_{A^i_{k-1}}(m)$ to $N_D(m+1)$ for all D appearing in (2), (4) and (5), and if $k = S(i, j)$ for some j then use (6) and (7).

For an estimate of the computation time of $c_m(T)$ observe that computing $d(i, j)$ or $e(i, j)$ for some i, j needs at most N comparisons of numbers. Hence the computation for all the $d(i, j)$ and $e(i, j)$ needs at most $2N^2m$ comparisons of numbers. Similar, computing all the $r(i, j)$, $s(i, j)$, $R(i, j)$ and $S(i, j)$ needs at most $N(2m-1)$ comparisons of numbers and additions. The computation of $c_m(T)$ needs at most $N^2m(2m-1)$ additions and hence the total computation time for computing $c_m(T)$ is $O(m^2)$.

4. EXAMPLE

We use the well known logistic map $f(x)=ax(1-x)$, for a less but near 4, to illustrate the results of the previous part. First of all we restrict our attention to the interval $[0, f(c)]$ where $c=1/2$ is the critical point and $f(c)=a/4$ holds. This makes the transformation onto and clearly does not change entropy.

Next we use the method of ref. 10 to construct the associated piecewise increasing transformation. First we shrink the graph of f to the square $[0, f(c)/2]^2$. This means that we consider $f'(x)=f(2x)/2$. Then we mirror the decreasing part of f' on the line $y=f(c)/2$ such that it becomes increasing. Hence we get

$$f''(x)=\begin{cases} f'(x) & \text{if} \quad x \text{ is in the increasing part of } f' \\ f(c)-f'(x) & \text{if} \quad x \text{ is in the decreasing part of } f' \end{cases}$$

At last we define f''' on $[f(c)/2, f(c)]$ by $f'''(x)=f(c)-f''(f(c)-x)$ which means that we mirror the graph of f'' on the vertical line $x=f(c)/2$ and on the horizontal line $y=f(c)/2$. Then the associated piecewise increasing transformation T is given by $T(x)=f''(x)$ if $x \in [0, f(c)/2)$ and

$T(x) = f'''(x)$ if $x \in (f(c)/2, f(c)]$ and we let T undefined in the point $f(c)/2$. In other words we consider T: $[0, f(c)] \mapsto [0, f(c)]$ given by

$$T(x) = \begin{cases} \dfrac{f(2x)}{2} & \text{if } 0 \leqslant x \leqslant \dfrac{1}{4} \\[2ex] f(c) - \dfrac{f(2x)}{2} & \text{if } \dfrac{1}{4} \leqslant x < \dfrac{f(c)}{2} \\[2ex] \dfrac{f(2(f(c)-x))}{2} & \text{if } \dfrac{f(c)}{2} < x \leqslant f(c) - \dfrac{1}{4} \\[2ex] f(c) - \dfrac{f(2(f(c)-x))}{2} & \text{if } f(c) - \dfrac{1}{4} \leqslant x \leqslant f(c) \end{cases}$$

Then $1 - T(x) = T(1-x)$ holds and in ref. 10 it is shown that $h_{\text{top}}(T) = h_{\text{top}}(f)$ holds.

In Fig. 1 and for the following explicit computation of the finite part $\mathscr{D}_{10}$ of the Markov diagram we have chosen $a = 3.89$.

Next we explain how one computes the Markov diagram step by step. The sets which appear in $\mathscr{D}_{10}$ are drawn in the figure.

Compute the successors under T of C_0 and D_0, the two intervals of monotonicity. C_0 has two successors namely C_0 itselve ($C_0 = T(C_0) \cap C_0$)

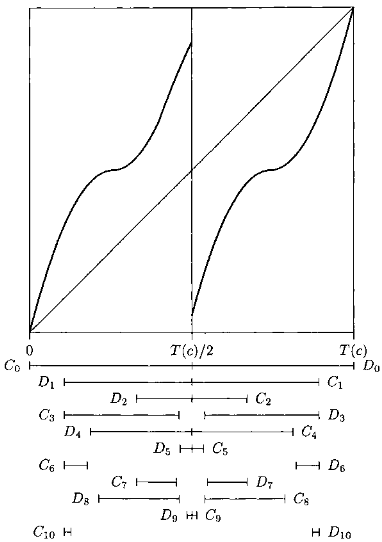
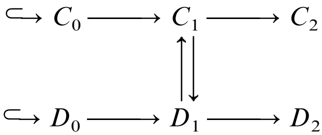


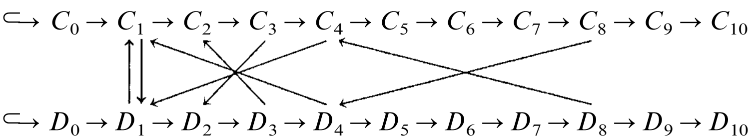
Figure 1.

and $C_1 = T(C_0) \cap D_0$. The same is true for D_0, which is a successor of itselve and the other successor is $D_1 = T(D_0) \cap C_0$.

Next compute the successors of C_1 and D_1. C_1 has two successors $T(C_1) \cap C_0 = D_1$ and $T(C_1) \cap D_0 = C_2$ and for D_1 we get the two successors $T(D_1) \cap D_0 = C_1$ and $T(D_1) \cap C_0 = D_2$. We get $\mathscr{D}_2$:



When computing the successors again, one sees that C_2 has only one successor $C_3 = T(C_2)$. But C_3 has 2 successors $T(C_3) \cap D_0 = C_4$ and $T(C_3) \cap C_0$ which equals D_2 since C_2 has its left endpoint in common with D_0. C_4 has the two successors $T(C_4) \cap D_0 = C_5$ and $T(C_4) \cap C_0$ which equals D_1 since C_4 has its left endpoint in common with D_0. Then we get that $T(C_5) = C_6$, $T(C_6) = C_7$ and $T(C_7) = C_8$ are the only successors. C_8 has the two successors $T(C_8) \cap D_0 = C_9$ and $T(C_8) \cap C_0$ which equals D_4 again since C_5 has its left endpoint in common with D_0. At last C_{10} is the only successor of C_9. Using the symmetry of the transformation, we get $\mathscr{D}_{10}$



The second method to compute the finite part $\mathscr{D}_{10}$ of the Markov diagram is that one we presented in the last paragraph. This method can be implemented on a computer directly. We now use the notation of paragraph 3.

We have $Z_1 = (x_1, y_1) = (0, f(c)/2)$ and $Z_2 = (x_2, y_2) = (f(c)/2, f(c))$. Clearly $\lim_{y \searrow x_1} T(y) = x_1$ and $\lim_{y \nearrow y_2} T(y) = y_2$ holds. This implies $n^+(x_1) = 1$ and $n^-(y_2) = 1$. The first step in computing $\mathscr{D}_{10}$ is to compute $d(i, j)$ and $e(i, j)$ for $i = 1, 2$ and $0 \leqslant j \leqslant 10$. With the choice $a = 3.89$ we get

$$d_1 = (d(1, 0), d(1, 1)) = (1, 0) \qquad \text{since} \quad n^+(x_1) = 1$$

$$e_1 = (e(1, 0), \dots, e(1, 10)) = (1, 2, 2, 1, 2, 2, 1, 1, 2, 2, 1)$$

$$d_2 = (d(2, 0), \dots, d(2, 10)) = (2, 1, 1, 2, 1, 1, 2, 2, 1, 1, 2)$$

$$e_2 = (e(2, 0), e(2, 1)) = (2, 3) \qquad \text{since} \quad n^-(y_2) = 1$$

This sequences include the whole information to compute $\mathscr{D}_{10}$. We have to set $R(1,0)=S(1,0)=0$ and $R(2,0)=S(2,0)=0$. Next we compute $r(1,1)=\min\{k\geqslant 1: d(1,k)\neq e(1,k)\}$ and $r(1,1)=1$ follows. This implies $R(1,1)=s(1,1)=S(1,1)=1$ and since $R(1,1)=n^+(x_1)$ holds, $r(1,j)$ and $R(1,j)$ for $j>1$ remain undefined. The next calculations are

$$\begin{aligned} s(1,2) &= \min\{k\geqslant 1: d(e(1,1),k)\neq e(1,k+1)\} \\ &= \min\{k\geqslant 1: d(2,k)\neq e(1,k+1)\} = 1 \end{aligned}$$

which implies $S(1,2)=2$

$$\begin{aligned} s(1,3) &= \min\{k\geqslant 1: d(e(1,2),k)\neq e(1,k+2)\} \\ &= \min\{k\geqslant 1: d(2,k)\neq e(1,k+2)\} = 2 \end{aligned}$$

and $S(1,3)=4$

$$\begin{aligned} s(1,4) &= \min\{k\geqslant 1: d(e(1,4),k)\neq e(1,k+4)\} \\ &= \min\{k\geqslant 1: d(2,k)\neq e(1,k+4)\} = 1 \end{aligned}$$

and $S(1,2)=5$

$$\begin{aligned} s(1,5) &= \min\{k\geqslant 1: d(e(1,5),k)\neq e(1,k+5)\} \\ &= \min\{k\geqslant 1: d(2,k)\neq e(1,k+5)\} = 4 \end{aligned}$$

which gives at last $S(1,2)=9$.

The symmetry of the transformation T implies $s(1,j)=r(2,j)$ and $r(1,j)=s(2,j)$ and we can summerize

$$r(1,1)=r(2,1)=s(1,1)=s(2,1)=1,\ R(2,1)=S(1,1)=1$$

$$r(2,2)=s(1,2)=1,\ R(2,2)=S(1,2)=2$$

$$r(2,3)=s(1,3)=2,\ R(2,3)=S(1,3)=4$$

$$r(2,4)=s(1,4)=1,\ R(2,4)=S(1,4)=5$$

$$r(2,5)=s(1,5)=4,\ R(2,5)=S(1,5)=9$$

and $r(1,j)$, $R(1,j)$, $s(2,j)$ and $S(2,j)$ stay undefined for $j>1$.

Now we can build up $\mathscr{D}_{10}$ according to Proposition 1:

(1) gives $A_0^2\to A_1^2\to\cdots\to A_{10}^2$ and no arrow departing from A_0^1 since $n^+(x_1)=1$.

(2) gives $A_0^1\to B_1^1$ and no arrow from A_0^2 since $n^-(y_2)=1=r(2,1)$.

(3) gives the arrows $B_1^1 \to \cdots \to B_{10}^1$.

(4) For $i=1$ we have nothing to do since the values $r(1, j)$ are undefined for $j>1$.

So consider $i=2$ and $j=2$. The arrow $A^2_{R(2,2)-1} = A^2_1 \to B^{d(2,R(2,1))}_{r(2,2)} = B^1_1$ is inserted since $r(2,2)=1<n^-(y_{d(2,R(2,1))})=n^-(y_1)>10$.

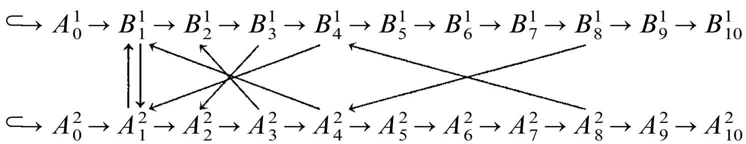
For $j=3$ the arrow $A^2_{R(2,3)-1}=A^2_3 \to B^{d(2,R(2,2))}_{r(2,3)}=B^1_2$ is inserted since $r(2,3)=2<n^-(y_{d(2,R(2,2))})=n^-(y_1)>10$.

For $j=4$ we get the arrow $A^2_{R(2,4)-1}=A^2_4 \to B^{d(2,R(2,3))}_{r(2,4)}=B^1_1$ since $r(2,4)=1<n^-(y_{d(2,R(2,3))})=n^-(y_1)>10$ and since $r(2,5)=4<n^-(y_{d(2,R(2,4))})=n^-(y_1)>10$ we insert the arrow $A^2_{R(2,5)-1}=A^2_8 \to B^{d(2,R(2,4))}_{r(2,5)}=B^1_4$. Since the transformation T is symmetric we get by condition (6) the symmetric arrows to that of condition (4). This also can be easily computed. We get the arrows $B^1_1 \to A^2_1$, $B^1_3 \to A^2_2$, $B^1_4 \to A^2_1$ and $B^1_8 \to A^2_4$.

(5) We start with $i=1$ and $j=1$ and get $A^1_0 \to A^l_0$ for $d(1,1)=0<l<e(d(1,0), r(1,1))=e(1,1)=2$. This is the arrow $A^1_0 \to A^1_0$. The values $r(1, j)$ do not exist for $j>1$.

So we consider $i=2$ and $j=1$ and obtain $A^2_0 \to A^l_0$ for $d(2,1)=1<l<e(d(2,0), r(2,1))=e(2,1)=3$. This gives $A^2_0 \to A^2_0$. The case $i=2$ and $j=1$ gives no arrow since there is no integer l with $d(2,2)=1<l<e(d(2,1), r(2,2))=e(1,1)=2$. In the same way one can see that $j=3, 4, 5$ give no arrow and using the symmetry, or computing (7) explicitly shows that we have found all arrows.

This gives the $\mathscr{D}_{10}$, which is the same diagram as we have computed above step by step:



5. SPECTRAL PROPERTIES OF THE MARKOV DIAGRAM

As described above, we conceive the Markov diagram of a piecewise monotonic transformation as a $\mathscr{D} \times \mathscr{D}$-matrix $\mathbf{M}$. We will prove a spectral theorem for this matrix (see Theorem 2 below).

A subset $\mathscr{C} \subseteq \mathscr{D}$ is called closed, if $D \in \mathscr{C}$ and $D \to C$ imply $C \in \mathscr{C}$. It is called irreducible, if, whenever $C \in \mathscr{C}$ and $D \in \mathscr{C}$, a path leads from D to C and no subset of $\mathscr{D}$ which contains $\mathscr{C}$ has this property.

Lemma 2. Suppose that T is topologically transitive, $h_{\mathrm{top}}(T)>0$ and $\mathscr{Z}$ is a generator. Then there is an irreducible closed subset $\mathscr{C}\subseteq\mathscr{D}$, such that all irreducible subsets of $\mathscr{D}\backslash\mathscr{C}$ consist only of a single closed path and for all $D\in\mathscr{D}\backslash\mathscr{C}$ there is a path starting in D which leads to $\mathscr{C}$.

Proof. The prove can be found in ref. 8. ▌

Lemma 3. Under the same assumptions as in Lemma 2 there is a strictly positive vector $\mathbf{v}\in l^{\infty}(\mathscr{D})$ with $\mathbf{M}\mathbf{v}=r(\mathbf{M})\,\mathbf{v}$.

Proof. Since $h_{\mathrm{top}}(T)>0$ is assumed, $r(\mathbf{M})>1$ follows. The existence of a nonnegative right eigenvector $\mathbf{v}\in l^{\infty}(\mathscr{D})$ for the eigenvalue $r(\mathbf{M})$ is stated in Corollary 1 to Theorem 9 of ref. 11.

Since $r(\mathbf{M})>1$ holds and because of Lemma 2 the vector $\mathbf{v}$ must have a strictly positive entry at one of the coordinates in the set $\mathscr{C}$ of Lemma 2. Observe that $\mathbf{v}_A>0$ implies $\mathbf{v}_B>0$ for $B\in\mathscr{D}$, whenever there is a path starting in B which meets A. Hence Lemma 2 implies that all entries of $\mathbf{v}$ are positive. ▌

Lemma 4. Suppose T is a piecewise monotonic transformation and $h_{\mathrm{top}}(T)>0$. Then there is $k\in\mathbb{N}$ and a finite dimensional operator $\mathbf{E}$ on $l^1(\mathscr{D})$ with $\|\mathbf{M}^k-\mathbf{E}\|<r(\mathbf{M})^k$.

Proof. Choose $k\in\mathbb{N}$ such that $r(\mathbf{M})^k>2$ holds. Define $\Pi_{\mathscr{D}_k}$: $l^1(\mathscr{D})\to l^1(\mathscr{D})$ by

$$(\Pi_{\mathscr{D}_k}(\mathbf{u}))_D=\begin{cases}\mathbf{u}_D & \text{for}\quad D\in\mathscr{D}_k\\ 0 & \text{for}\quad D\in\mathscr{D}\backslash\mathscr{D}_k\end{cases}$$

Then the image of $\Pi_{\mathscr{D}_k}\circ\mathbf{M}^k$ is contained in a finite dimensional subspace of $l^1(\mathscr{D})$. Hence the operator $\Pi_{\mathscr{D}_k}\circ\mathbf{M}^k$ is finite dimensional. For $\mathbf{u}\in l^1(\mathscr{D})$ we compute

$$\begin{aligned}\|\mathbf{u}(\mathbf{M}^k-\Pi_{\mathscr{D}_k}\circ\mathbf{M}^k)\|_1&=\sum_{D\in\mathscr{D}}\left|\sum_{C\in\mathscr{D}}\mathbf{u}_C(\mathbf{M}^k_{CD}-(\Pi_{\mathscr{D}_k}\circ\mathbf{M}^k)_{CD})\right|\\&\leqslant\sum_{D\in\mathscr{D}}\sum_{C\in\mathscr{D}}|\mathbf{u}_C|\,(\mathbf{M}^k_{CD}-(\Pi_{\mathscr{D}_k}\circ\mathbf{M}^k)_{CD})\\&=\sum_{D\in\mathscr{D}\backslash\mathscr{D}_k}\sum_{C\in\mathscr{D}}|\mathbf{u}_C|\,\mathbf{M}^k_{CD}=\sum_{C\in\mathscr{D}}|\mathbf{u}_C|\,N^k(C)\end{aligned}$$

where $N^k(C)$ denotes the number of paths of length k which start in C and end in $\mathscr{D}\backslash\mathscr{D}_k$.

We get $N^k(C) \leqslant 2$ for all $C \in \mathscr{D}$: In the same way as in the proof of Theorem 1 we have a bijection between the paths of length k starting in C and the elements $Z \in \mathscr{Z}_k$ with $Z \cap C \neq \emptyset$. If $A \in \mathscr{Z}_k$ corresponds to $C = C_0 \to \cdots \to C_k$ then $C_k = T^k(C \cap A)$ holds because of Lemma 1. There are at most 2 elements $Z \in \mathscr{Z}_k$ such that $C \cap Z \neq Z$ holds. All paths which correspond to the other $Z \in \mathscr{Z}_k$ have $T^k(Z)$ as their last element and end in $\mathscr{D}_k$ because of Lemma 1.

Hence $N^k(C) \leqslant 2$ and

$$\|\mathbf{u}(\mathbf{M}^k - \Pi_{\mathscr{D}_k} \circ \mathbf{M}^k)\|_1 \leqslant 2\ \|\mathbf{u}\|_1 < r(\mathbf{M})^k\ \|\mathbf{u}\|_1 \quad \blacksquare$$

Lemma 8.2 in ref. 6 and Lemma 4 imply the following representation of $\mathbf{M}^n$.

Proposition 2. Let T: $[0, 1] \to [0, 1]$ be piecewise monotonic and $h_{\text{top}}(T) > 0$. Then there is $k \in \mathbb{N}$ such that for all $n \in \mathbb{N}$

$$\mathbf{M}^n = \sum_{i=1}^{k} \lambda_i^n (\mathbf{P}_i + \mathbf{N}_i)^n + (\mathbf{PM})^n$$

where λ_i are isolated eigenvalues with $|\lambda_i| = r(\mathbf{M})$, $\mathbf{P}_i$ are projections onto the finite dimensional generalized eigenspace of λ_i and $\mathbf{N}_i$ are nilpotent for $1 \leqslant i \leqslant k$. For $i \neq j$ with $1 \leqslant i, j \leqslant k$ we get $\mathbf{P}_i\mathbf{P}_j = \mathbf{P}_j\mathbf{P}_i = 0$, $\mathbf{P}_i\mathbf{N}_i = \mathbf{N}_i\mathbf{P}_i = \mathbf{N}_i$ and $\mathbf{PP}_i = \mathbf{P}_i\mathbf{P} = 0$. Furthermore $\mathbf{P}^2 = \mathbf{P}$ and $r(\mathbf{PM}) < r(\mathbf{M})$.

Set $\hat{\mathbf{M}} = \mathbf{M}/r(\mathbf{M})$. For $\mathbf{r} \in l^1(\mathscr{D})$ and $\mathbf{s} \in l^\infty(\mathscr{D})$ set $|\mathbf{r}| = (|\mathbf{r}_D|)_{D \in \mathscr{D}}$ and $\langle \mathbf{r}, \mathbf{s} \rangle = \sum_{D \in \mathscr{D}} \mathbf{r}_D \mathbf{s}_D$.

Lemma 5. Suppose the assumptions of Lemma 2 are fulfilled and let $\mathbf{N}_i$ be the nilpotent operators of the above proposition. Then $\mathbf{N}_i \equiv 0$ for $1 \leqslant i \leqslant k$.

Proof. Because of Lemma 3 a strictly positive eigenvector $\mathbf{v} \in l^\infty(\mathscr{D})$ of $\hat{\mathbf{M}}$ for the eigenvalue 1 exists.

Suppose there is i with $\mathbf{N}_i \not\equiv 0$. Then $l > 1$ exists with $\mathbf{N}_i^l \equiv 0$ and $\mathbf{N}_i^{l-1} \not\equiv 0$. Choose $\mathbf{p} \in l^1(\mathscr{D})$ such that $\mathbf{pN}_i^{l-1} =: \mathbf{w} \neq 0$. Then $\mathbf{wN}_i = \mathbf{pN}_i^l = 0$ and $\mathbf{wP}_i = \mathbf{pN}_i^{l-1}\mathbf{P}_i = \mathbf{pN}_i^{l-1} = \mathbf{w}$. Set $\mathbf{u} := \mathbf{pN}_i^{l-2}$. We get $\mathbf{uN}_i = \mathbf{w} \neq 0$ and $\mathbf{uN}_i^n = 0$ for all $n \geqslant 2$. Furthermore $\mathbf{uP}_j = \mathbf{u}$ for $j = i$ and $\mathbf{uP}_j = 0$ for $j \neq i$ holds. At last $\mathbf{uN}_j = 0$ for $j \neq i$ and $\mathbf{uP} = 0$.

Since $|\mathbf{w}| \not\equiv 0$ and since $\mathbf{v}$ is strictly positive we get that $\langle |\mathbf{w}|, \mathbf{v} \rangle > 0$ holds. Hence we can choose $n \in \mathbb{N}$ such that $n > (2\langle |\mathbf{u}|, \mathbf{v} \rangle)/(\langle |\mathbf{w}|, \mathbf{v} \rangle)$. We compute

$$\mathbf{u}\hat{\mathbf{M}}^n = \frac{1}{r(\mathbf{M})^n}\left(\sum_{j=1}^{k} \lambda_j^n \mathbf{u}(\mathbf{P}_j+\mathbf{N}_j)^n + \mathbf{u}(\mathbf{PM})^n\right)$$
$$= \left(\frac{\lambda_i}{r(\mathbf{M})}\right)^n \left(\mathbf{uP}_i^n + \binom{n}{1}\mathbf{uP}_i^{n-1}\mathbf{N}_i + \binom{n}{2}\mathbf{uP}_i^{n-2}\mathbf{N}_i^2 + \cdots\right)$$
$$= \left(\frac{\lambda_i}{r(\mathbf{M})}\right)^n \mathbf{u} + \left(\frac{\lambda_i}{r(\mathbf{M})}\right)^n n\mathbf{w}$$

This implies $n\,|\mathbf{w}| \leqslant |\mathbf{u}| + |\mathbf{u}|\,\hat{\mathbf{M}}^n$, since $\hat{\mathbf{M}}^n$ is positive and $|\lambda_i| = r(\mathbf{M})$ holds. We get

$$n\langle |\mathbf{w}|, \mathbf{v}\rangle \leqslant \langle |\mathbf{u}|, \mathbf{v}\rangle + \langle |\mathbf{u}|\,\hat{\mathbf{M}}^n, \mathbf{v}\rangle$$
$$= \langle |\mathbf{u}|, \mathbf{v}\rangle + \langle |\mathbf{u}|, \hat{\mathbf{M}}^n\mathbf{v}\rangle = 2\langle |\mathbf{u}|, \mathbf{v}\rangle$$

This contradicts the choice of n. ∎

Lemma 6. Suppose again the assumptions of Lemma 2 are fulfilled. Then $\sup_{n\geqslant 1} \|\hat{\mathbf{M}}^n\| < \infty$.

Proof. Because of Proposition 2 and Lemma 5 we get

$$\hat{\mathbf{M}} = \sum_{i=1}^{k} \frac{\lambda_i}{r(\mathbf{M})}\mathbf{P}_i + \mathbf{R} \qquad \text{with} \quad \mathbf{R} = \mathbf{P}\hat{\mathbf{M}}$$

Furthermore $r(\mathbf{R}) = r((1/r(\mathbf{M}))\,\mathbf{PM}) = r(\mathbf{PM})/r(\mathbf{M}) < 1$ follows. Hence we get $M_0 := \sup_{n\geqslant 1} \|\mathbf{R}^n\| < \infty$. We compute for an arbitrary $\mathbf{u} \in l^1(\mathscr{D})$

$$\|\mathbf{u}\hat{\mathbf{M}}^n\|_1 \leqslant \sum_{i=1}^{k} \left|\frac{\lambda_i}{r(\mathbf{M})}\right|^n \|\mathbf{uP}_i\|_1 + \|\mathbf{uR}^n\|_1 \leqslant k\,\|\mathbf{u}\|_1 + M_0\,\|\mathbf{u}\|_1$$

and the result follows. ∎

We have checked all assumptions of Theorem 8.8 in ref. 6 and get the following result, where we denote the spectrum of an operator $\mathbf{M}$ by $\sigma(\mathbf{M})$.

Theorem 2. Let $T: [0,1] \to [0,1]$ be piecewise monotonic and topologically transitive. Suppose $\mathscr{Z}$ is a generator and $h_{\text{top}}(T) > 0$. Then $\sigma(\mathbf{M}) = \sigma \cup \{\lambda_1,..., \lambda_k\}$, where σ is a closed subset of $\{z \in \mathbb{C}: |z| < \alpha\}$ for an $\alpha < r(\mathbf{M})$. For $1 \leqslant j \leqslant k$ we get that λ_j is an eigenvalue of $\mathbf{M}$ and there is $\theta_j \in \mathbb{Q}$ with $\lambda_j = r(\mathbf{M})\, e^{2\pi i\theta_j}$. Furthermore projections $\mathbf{P}_1,..., \mathbf{P}_k$ and $\mathbf{P}$ exist with $\mathbf{P}_1 + \cdots + \mathbf{P}_k + \mathbf{P} = Id$, where $\mathbf{P}_j$ is the projection onto the eigenspace of λ_j. With $\mathbf{R} = \mathbf{PM}$ we get

$$\mathbf{M}^m = \sum_{j=1}^{k} r(\mathbf{M})^m \, e^{2\pi i\theta_j m}\mathbf{P}_j + \mathbf{R}^m \qquad \text{and} \qquad \sup_{m\in\mathbb{N}} \frac{1}{\alpha^m} \|\mathbf{R}^m\| < \infty$$

We use the above theorem to compute $c_m(T)$ for $m \geqslant 0$.

Theorem 3. Under the assumptions of Theorem 2 there are $p \in \mathbb{N}$ and numbers c_i with $0 < c_i < \infty$ for $1 \leqslant i \leqslant p$ such that for all $m \in \mathbb{N}$ and with $i = m \bmod(p)$

$$c_m(T) = c_i r(\mathbf{M})^m + r_m$$

holds, where $|r_m| \leqslant K\alpha^m$ for a constant K and the α of Theorem 2.

Proof. Set $\mathbf{w}_D = 1$ for all $D \in \mathscr{D}$ and $\mathbf{v}_D = 1$ for $D \in \mathscr{D}_0$ and $\mathbf{v}_D = 0$ for $D \notin \mathscr{D}_0$. Theorem 1 and Theorem 2 imply

$$c_m(T) = \|\mathbf{vM}^m\|_1 = \langle \mathbf{vM}^m, \mathbf{w} \rangle = r(\mathbf{M})^m \sum_{j=1}^{k} e^{2\pi i\theta_j m} \langle \mathbf{vP}_j, \mathbf{w} \rangle + \langle \mathbf{vR}^m, \mathbf{w} \rangle$$

For $m \in \mathbb{N}$ set $c_m = \sum_{j=1}^{k} e^{2\pi i\theta_j m} \langle \mathbf{vP}_j, \mathbf{w} \rangle$ and $r_m = \langle \mathbf{vR}^m, \mathbf{w} \rangle$. Since $\mathbf{w}_D = 1$ for all $D \in \mathscr{D}$ we get

$$|r_m| = |\langle \mathbf{vR}^m, \mathbf{w} \rangle| \leqslant \langle |\mathbf{vR}^m|, \mathbf{w} \rangle = \|\mathbf{vR}^m\|_1 \leqslant \|\mathbf{R}^m\| \, \|\mathbf{v}\|_1 \leqslant K\alpha^m$$

for a constant K and $\alpha < r(\mathbf{M})$ by Theorem 2.

Choose $p \in \mathbb{N}$ such that $\theta_j p \in \mathbb{N}$ holds for $1 \leqslant j \leqslant k$. Then $e^{2\pi i\theta_j p} = 1$ for $1 \leqslant j \leqslant k$ and $c_{p+l} = c_l$ for all $l \geqslant 1$ follow. The equation $c_m(T) = c_m r(\mathbf{M})^m + r_m$, valid for all $m \in \mathbb{N}$, implies

$$c_l = \frac{1}{r(\mathbf{M})^{l+jp}} \left(c_{l+jp}(T) - r_{l+jp}\right) \qquad \text{for all} \quad j \geqslant 0$$

With Lemma 4 in ref. 15 we get $c_m(T) \geqslant \|\mathbf{M}^m\| \geqslant r(\mathbf{M})^m$ for all $m \geqslant 0$. Since $|r_m| \leqslant K\alpha^m$ for all $m \geqslant 0$ we get $c_i > 0$ for $1 \leqslant i \leqslant p$ and the result follows. ∎

6. REGRESSION METHOD FOR COMPUTING ENTROPY

We obtain by Theorem 3 for $m \in \mathbb{N}$

$$\log c_m(T) = m \log r(\mathbf{M}) + \log c_m + R(m)$$

with $R(m) = \log(1 + (r_m/c_m r(\mathbf{M})^m)$. Observe that $\lim_{m \to \infty} |R(m)| = 0$ since $|r_m/c_m r(\mathbf{M})^m|$ tends to zero exponentially by Theorem 3. This suggests to compute the line of best fit to the datsa pairs $(n_1, \log c_{n_1}(T))$,..., $(n_k, \log c_{n_k}(T))$ where the n_i are large enough such that $R(n_i)$ is negligible. Furthermore statistical effects, when computing the line of best fit, decrease the influence of the $R(n_i)$ if the number of data points is large.

Furthermore Theorem 3 states, that the c_m are periodic. Suppose $c_{i+kp} = c_i$ for $1 \leqslant i \leqslant p$ and all $k \geqslant 0$. Then we get

$$\log c_{i+kp}(T) = (i + kp) \log r(\mathbf{M}) + \log c_i + R(i + kp)$$

and the best fit process gives a line with slope $\log r(\mathbf{M}) = h_{\text{top}}(T)$ and intercept $\log c_i$ for every i with $1 \leqslant i \leqslant p$.

If the period p is not too large it can be detected by a plot of the data pairs $(n_1, \log c_{n_1}(T))$,..., $(n_k, \log c_{n_k}(T))$. When a period is detected, the estimate of $h_{\text{top}}(T)$ can be improved by setting $n_j = n_0 + jp$.

If the transformation T is topologically mixing then the eigenvalue $\lambda = r(\mathbf{M})$ is the only eigenvalue of $\mathbf{M}$ with $|\lambda| = r(\mathbf{M})$. Then $c_m(T) = cr(\mathbf{M})^m + r_m$ for a constant $c > 0$ follows with r_m as in Theorem 3.

ACKNOWLEDGMENTS

The author would like to thank Franz Hofbauer for permanent encouragement and Peter Raith for helpfull comments. This work was supported by Projekt Nr. P11579-MAT of the Austrian Fonds zur Förderung der wissenschaftlichen Forschung.

REFERENCES

1. S. L. Baldwin and E. E. Slaminka, Calculating topological entropy, *J. Statist. Phys.* **85**:1017–1033 (1997).
2. N. J. Balmforth, E. A. Spiegel, and C. Tresser, Topological entropy of one-dimensional maps: Approximations and bounds, *Phys. Rev. Lett.* **72**:80–83 (1994).
3. L. Block and J. Keesling, Computing the topological entropy of maps of the interval with three monotone pieces, *J. Statist. Phys.* **66** (1992).
4. L. Block, J. Keesling, S. Li, and K. Peterson, An improved algorithm for computing topological entropy, *J. Statist. Phys.* **55**:929–939 (1989).
5. P. Collet, J. P. Crutchfield, and J. P. Eckmann, Computing the topological entropy of maps, *Comm. Math. Phys.* **88**:257–262 (1983).
6. N. Dunford and J. T. Schwartz, *Linear Operators Part I* (Interscience Publishers, New York, 1958).
7. P. Gora and A. Boyarsky, Computing the topological entropy of general one-dimensional maps, *Trans. Am. Math. Soc.* **323**:39–49 (1991).
8. F. Hofbauer, The box dimension of completely invariant subsets for expanding piecewise monotonic transformations, *Mh. Math.* **121**:199–211 (1995).

9. F. Hofbauer, On intrinsic ergodicity of piecewise monotonic transformations with positive entropy i, *Israel J. Math.* **34**:213–237 (1979).
10. F. Hofbauer, On intrinsic ergodicity of piecewise monotonic transformations with positive entropy ii, *Israel J. Math.* **38**:107–115 (1981).
11. F. Hofbauer, Piecewise invertible dynamical systems, *Probab. Theory Relat. Fields* **72**:359–386 (1986).
12. F. Hofbauer, The structure of piecewise monotonic transformations, *Ergod. Th. Dynam. Sys.* **1**:159–178 (1981).
13. M. Misiurewicz and W. Szlenk, Entropy of piecewise monotone mappings, *Sudia Math.* **67**:45–63 (1980).
14. S. Newhouse and T. Pignataro, On the estimation of topological entropy, *J. Statist. Phys.* **72**:1331–1351 (1993).
15. P. Raith, Continuity of the Hausdorff dimension for piecewise monotonic maps, *Israel J. Math.* **80**:97–133 (1992).
16. P. Walters, *An Introduction to Ergodic Theory* (Springer Verlag, Berlin/Heidelberg/New York, 1982).