ONE-DIMENSIONAL DYNAMICS

MICHAŁ MISIUREWICZ

1. Introduction

In these notes I present some basics of the dynamical systems in dimension 1. By a *dynamical system* we mean a continuous map $f : X \to X$ of a compact metric space X to itself, or, more precisely, the iterates $f^n = f \circ f \circ \cdots \circ f$ of this map. Additionally we define f^0 to be the identity on X. If f is a homeomorphism, then we also consider maps $f^{-n} = (f^{-1})^n$. While in several cases I present general results, most of the time X will be a real one-dimensional space: a compact interval or the circle.

As in other areas of mathematics, sometimes two systems are practically the same ("isomorphic"). Here this "isomorphism" is called *conjugacy* (or *topological conjugacy*, in order to distinguish from stronger notions). The map $f : X \to X$ is *conjugate* to $g : Y \to Y$ if there is a homeomorphism $h : X \to Y$ such that the diagram

$$\begin{array}{ccc} X & \xrightarrow{f} & X \\ {\scriptstyle h}\downarrow & & \downarrow{\scriptstyle h} \\ Y & \xrightarrow{g} & Y \end{array}$$

commutes. Clearly, in this case h also conjugates f^n with g^n for all n. If h is only a continuous surjection, then we say that it is a *semiconjugacy* and g is a *factor* of f.

While in general we want to study what happens with $f^n(x)$ as $n \to \infty$, there are some points for which studying this asymptotic behavior reduces to looking at finitely many iterates. For a map $f : X \to X$ a point $x \in X$ is called a *periodic point* if there is a positive integer n such that $f^n(x) = x$. The smallest such n is called the *period* of x. In particular, if $n = 1$, then x is called a *fixed point*.

As I mentioned at the beginning, these notes contain only some basic things. More information on one-dimensional dynamics can be found for instance in 3 books: theory of smooth maps in [3]; theory of continuous maps in [2]; combinatorial dynamics and entropy in [1]. References included in those books consist of several hundreds items. Papers with newer results can be found via MathSciNet.

2. Sharkovsky's Theorem

The story begins in 1964, when A. N. Sharkovsky published a beautiful theorem, known now as *Sharkovsky's Theorem*. It characterizes the possible sets of periods of periodic points of maps of a compact interval into itself.

Date: May 12, 2013.

Consider the *Sharkovsky ordering* of the set of natural numbers:

$$3 \prec 5 \prec 7 \prec 9 \prec \dots \prec 3 \cdot 2 \prec 5 \cdot 2 \prec 7 \cdot 2 \prec 9 \cdot 2 \prec \dots$$
$$\prec 3 \cdot 2^2 \prec 5 \cdot 2^2 \prec 7 \cdot 2^2 \prec 9 \cdot 2^2 \prec \dots \prec 2^3 \prec 2^2 \prec 2 \prec 1.$$

Now Sharkovsky's Theorem can be split into three parts. Let I be a compact interval.

Theorem 2.1. *Let $f : I \to I$ be a continuous map. If f has a periodic orbit of period n and if n appears before k in the Sharkovsky ordering, then f has a periodic orbit of period k.*

Theorem 2.2. *For every k there exists a continuous map $f : I \to I$ that has a periodic orbit of period n, but has no periodic orbits of period k for any k appearing before n in the Sharkovsky ordering.*

Theorem 2.3. *There exists a continuous map $f : I \to I$ that has a periodic orbit of period 2^n for every n and has no periodic orbits of any other periods.*

The basic notion used in the standard proof of Theorem 2.1 is f-covering. We say that an interval $K \subset I$ *f-covers* an interval $L \subset I$ if $L \subset f(K)$. Note that in such a case there is a subinterval $J \subset K$ such that $f(J) = L$. When we speak of f-covering, we always assume that the intervals we mention are compact.

The following simple lemma allows us to use this notion for finding periodic orbits.

Lemma 2.4. *Let $f : I \to I$ be a continuous map and let $J_1, J_2, \dots, J_n$ be subintervals of I such that J_i f-covers J_{i+1} for $i = 1, 2, \dots, n-1$ and J_n f-covers J_1. Then there exist points $x_i \in J_i$ such that $f(x_i) = x_{i+1}$ for $i = 1, 2, \dots, n-1$ and $f(x_n) = x_1$.*

Given a periodic orbit $P = \{x_1, x_2, \dots, x_n\}$ of f of period n, such that $x_1 < x_2 < \dots < x_n$, we call the intervals $[x_i, x_{i+1}]$, $i = 1, 2, \dots, n-1$, *P-basic intervals*. Then we construct a directed graph that we call the *P-graph*. Its vertices are the P-basic intervals, and we draw an arrow (a directed edge) from $[x_i, x_{i+1}]$ to $[x_j, x_{j+1}]$ if the interval with endpoints $f(x_i)$ and $f(x_{i+1})$ contains $[x_j, x_{j+1}]$ (note that then $[x_i, x_{i+1}]$ f-covers $[x_j, x_{j+1}]$). Looking at the loops (as in Lemma 2.4) in this graph we find periodic orbits of f of various periods. In general, the period of such orbit will be equal to the length of the loop, except when the loop is a repetition of a shorter loop (but in most cases, this situation is easy to rule out). A loop that is not a repetition of a shorter one will be called *simple*.

We will give here the proof of Theorem 2.1 in the case when n is an odd integer larger than 1. This is the main part of the proof. The rest of the proof is not difficult, but involves many details.

Let $f : I \to I$ be a continuous interval map, let n be an odd integer larger than 1, and let $P = \{x_1, x_2, \dots, x_n\}$ be a periodic orbit of f of period n. We say that P is a *Štefan orbit* if either

$$x_n < x_{n-2} < \dots < x_5 < x_3 < x_1 < x_2 < x_4 < \dots < x_{n-3} < x_{n-1}$$

or

$$x_{n-1} < x_{n-3} < \dots < x_4 < x_2 < x_1 < x_3 < x_5 < \dots < x_{n-2} < x_n$$

(see Figure 1).

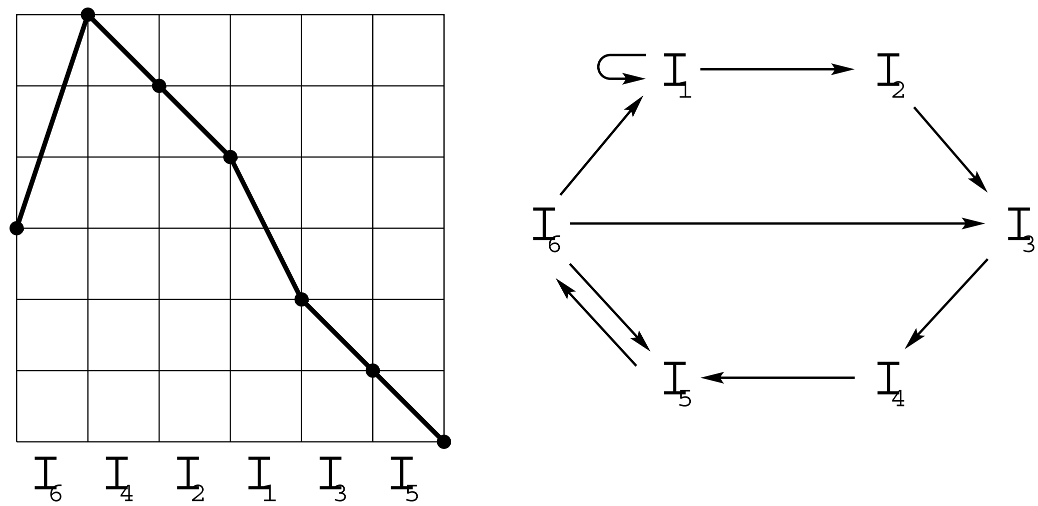

FIGURE 1. Štefan orbit P of period 7 and the P-graph.

Lemma 2.5. *Let P be a periodic orbit of a continuous map $f : I \to I$ of period n odd and larger than 1. Assume that f has no periodic orbit of odd period k such that $1 < k < n$. Then P is Štefan.*

Proof. Clearly, $f(\max P) < \max P$ and $f(\min P) > \min P$. Take the largest $z \in P$ for which $f(z) > z$ and denote by I_1 the P-basic interval such that $\inf I_1 = z$. We have $f(z) \geq \sup I_1$, $f(\sup I_1) \leq z$ and, hence $f(I_1) \supset I_1$.

Now, we define by induction sets J_i: $J_1 = I_1$ and J_{i+1} is the union of P-basic intervals f-covered by J_i. Clearly, J_1 is an interval. If J_i is an interval then $f(J_i)$ is also an interval and so is J_{i+1}. This proves that all J_i are intervals. Since $f(J_1) \supset J_1$, we have $J_2 \supset J_1$. If $J_{i+1} \supset J_i$ then $f(J_{i+1}) \supset f(J_i) \supset J_{i+1}$ and therefore $J_{i+2} \supset J_{i+1}$. This proves that $J_{i+1} \supset J_i$ for all i.

Thus we get a sequence of intervals $J_1 \subset J_2 \subset J_3 \dots$, and each of them is a union of P-basic intervals. If J_i is not the convex hull of P, then the set $P \cap J_i$ is a proper subset of P. This subset is not invariant and hence $f(J_i)$ contains some P-basic interval which is not contained in J_i. Therefore $J_{i+1} \neq J_i$. Thus we see that the number of P-basic intervals contained in J_i increases with i until J_i is the convex hull of P. Let ℓ be the smallest number such that J_ℓ is the convex hull of P.

Since n is odd, there are $x, y \in P$ such that x and y are on the same side of I_1 but $f(x)$ and $f(y)$ are on different sides of I_1. Therefore there exists a P-basic interval I' such that $I' \neq I_1$ and I' f-covers I_1. Let k be the smallest $i \in \{1, 2, \dots, \ell\}$ such that $I' \subset J_i$. Set $I_k = I'$. Since $I_k \subset f(J_{k-1})$, we see that J_{k-1} f-covers I_k. Then there exists a P-basic interval I_{k-1} such that $I_{k-1} \subset J_{k-1}$ and I_{k-1} f-covers I_k. If $I_{k-1} \neq I_1$, we can repeat the argument until for some $m \leq k$ we get the loop $(I_1, I_m, I_{m+1}, \dots, I_k)$. This shows that there is a loop from I_1 to itself of length not greater than k, which is not a repetition of the length 1 loop from I_1 to itself. Take the shortest such loop $\alpha = (I_1, K_2, K_3, \dots, K_r)$. Since it is the shortest one, it is simple. Then f has a periodic orbit Q associated to α (in the sense of Lemma 2.4). Since $1 < \mathrm{Card}(Q) = r \leq k \leq n - 1$, by the hypothesis r is even.

If $n = 3$ then we are done, because every periodic orbit of period 3 is Štefan. Thus we assume that $n > 3$. If $r \le n-3$ then the loop $(I_1, I_1, K_2, K_3, \dots, K_r)$ gives a cycle of f of odd period smaller than n; a contradiction. Since r is even, we get $r = n-1$.

We have $n-1 = r \le k \le \ell \le n-1$. Hence, $\ell = n-1$. Therefore, the cardinality of $J_j \cap P$ is $j+1$ for $j = 1, 2, \dots, n-1$. Since α is the shortest loop from I_1 to I_1 which is not a repetition of the length 1 loop from I_1 to itself, the interval K_j is not contained in J_{j-1} for $j = 2, 3, \dots, n-1$. Hence, K_j is contained in the closure of $J_j \setminus J_{j-1}$. Since the cardinality of $(J_j \setminus J_{j-1}) \cap P$ is 1, the closure of $J_j \setminus J_{j-1}$ is a P-basic interval, and hence it is equal to K_j.

Since the cardinality of $J_2 \cap P$ is 3, one of the endpoints of I_1 is mapped by f to the other endpoint. Therefore we may assume that $P = \{x_1, x_2, \dots, x_n\}$ with the temporal labeling (that is, $f(x_1) = x_2$, $f(x_2) = x_3$, etc.), and I_1 has endpoints x_1 and x_2. In fact, we may assume that $x_1 < x_2$. Then $x_3 = f(x_2) < x_2$ by our choice of I_1. Therefore $x_3 < x_1 < x_2$ and $K_2 = [x_3, x_1]$. If $x_4 = f(x_3) < x_3$ then $f(K_2) \supset I_1$. Then we get the loop (I_1, K_2) shorter than α; a contradiction. Therefore $x_4 > x_2$ and $K_3 = [x_2, x_4]$. In the same way we obtain $x_5 < x_3$ and $K_4 = [x_5, x_3]$ etc. This proves that P is a Štefan cycle. □

Now in order to finish the proof of Theorem 2.1 in the case when n is an odd integer larger than 1, it is enough to do two things. One is to check by the straightforward inspection that the lengths of simple loops in the P-graph of a Štefan periodic orbit of period n are exactly the numbers that follow n in the Sharkovsky ordering. The second thing is the observation that if f is monotone on every P-basic interval and the endpoints of I belong to P then every periodic orbit of f of period k corresponds to a loop of length k in the P-graph.

Remark 2.6. In the proof of Theorem 2.1 the convex hull of the periodic orbit of period k that we construct is contained in the convex hull of the initial periodic orbit of period n.

There are several ways of proving Theorems 2.2 and 2.3. Perhaps the simplest one is to use truncated tent maps. The *tent map* is the map of the interval $[0, 1]$ onto itself, given by the formula $T(x) = 1 - |2x - 1|$ (see Figure 2).

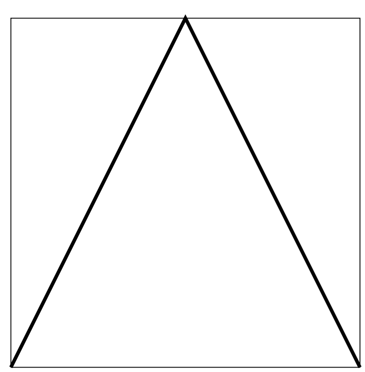

FIGURE 2. The tent map.

By looking at the P-graph of the tent map T, where $P = \{0, 1/2, 1\}$ (it is not a periodic orbit, but the notion of the P-graph still makes sense for it), one can see that T has periodic orbits of all periods. Moreover, on each P-basic interval T expands distances, so for every n it has only finitely many periodic orbits of period n. For a given n look at the rightmost (largest) point of each periodic orbit of period n, and choose the leftmost (smallest) of them. Call it a. Then the *truncated tent map* T_a, given by $T_a(x) = \min(T(x), a)$ (see Figure 3), has exactly one periodic orbit of period n (call it Q).

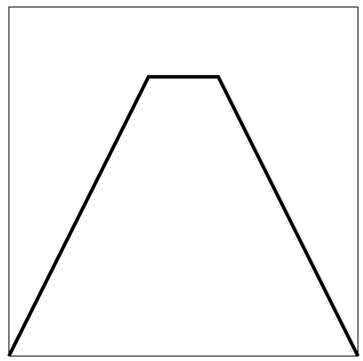

FIGURE 3. Truncated tent map.

The convex hull of Q is $[T(a), a]$, and this interval is invariant for T_a. Every periodic orbit of T_a, except $\{0\}$, eventually enters this interval and stays there, so it has to be contained in it. Thus, if k appears before n in the Sharkovsky ordering and T_a has a periodic orbit R of period k, then the convex hull of R is contained in the convex hull of Q. However, since k appears before n in the Sharkovsky ordering, by Remark 2.6, T_a has a periodic orbit of period n, whose convex hull is contained in the convex hull of R. Hence, this orbit must be different than Q, a contradiction. This proves Theorem 2.2.

To get an example necessary to prove Theorem 2.3, we choose as the level a at which we truncate the tent map the limit of the levels a_{2^n} corresponding to the periods 2^n. The property that no other period exists for T_a follows easily form the fact that in the Sharkovsky ordering there is no period followed by all powers of 2 and only by them.

The methods of the proof of Theorem 2.1 suggest that one can refine Sharkovsky's Theorem by considering not only the periods of periodic orbits, but also their permutations (or permutations modulo orientation; they are called *patterns*). Then there is *forcing relation* between patterns: a pattern A *forces* pattern B if every continuous interval map having a periodic orbit of pattern A has also a periodic orbit of pattern B. It turns out that the forcing relation is a partial order. Investigation of this relation and related objects constitutes an area of the dynamical systems theory called *combinatorial dynamics*.

3. Topological entropy for interval maps

Let X be a compact Hausdorff (often we want it to be metric) topological space, and let $f : X \to X$ be a continuous map. *Topological entropy* of f measures how complicated the dynamical system given by f is: how many very different orbits it has, how fast it "mixes" together various sets, etc.

A family $\mathcal{A}$ of subsets of X is called a *cover* if their union is all of X. For open covers $\mathcal{A}_1, \mathcal{A}_2, \dots, \mathcal{A}_n$ of X we denote

$$\bigvee_{i=1}^{n} \mathcal{A}_i = \mathcal{A}_1 \vee \mathcal{A}_2 \vee \cdots \vee \mathcal{A}_n = \left\{ \bigcap_{i=1}^{n} A_i : A_i \in \mathcal{A}_i,\ \bigcap_{i=1}^{n} A_i \neq \emptyset \right\}.$$

Note that $\bigvee_{i=1}^{n} \mathcal{A}_i$ is also an open cover. For an open cover $\mathcal{A}$ we denote $f^{-n}(\mathcal{A}) = \{f^{-n}(A) : A \in \mathcal{A}\}$ and $\mathcal{A}^n = \bigvee_{i=0}^{n-1} f^{-i}(\mathcal{A})$. For each i, $f^{-i}(\mathcal{A})$ is an open cover, so $\mathcal{A}^n$ is also an open cover. Next, we denote by $\mathcal{N}(\mathcal{A})$ the minimal possible cardinality of a subcover of $\mathcal{A}$ (i.e., a subset of $\mathcal{A}$ which is also a cover of X).

The following simple inequalities hold:

$$\mathcal{N}(\mathcal{A} \vee \mathcal{B}) \leq \mathcal{N}(\mathcal{A})\mathcal{N}(\mathcal{B}), \tag{3.1}$$

$$\mathcal{N}(f^{-n}\mathcal{A}) \leq \mathcal{N}(\mathcal{A}). \tag{3.2}$$

We have $\mathcal{A}^{k+n} = \mathcal{A}^k \vee f^{-k}(\mathcal{A}^n)$, and hence we obtain the next useful inequality

$$\mathcal{N}(\mathcal{A}^{k+n}) \leq \mathcal{N}(\mathcal{A}^k)\mathcal{N}(\mathcal{A}^n). \tag{3.3}$$

By (3.3), the sequence $(\log \mathcal{N}(\mathcal{A}^n))_{n=1}^{\infty}$ is subadditive, and therefore the limit

$$h(f, \mathcal{A}) = \lim_{n\to\infty} \frac{1}{n} \log \mathcal{N}(\mathcal{A}^n) \tag{3.4}$$

exists and is equal to the infimum of $(1/n) \log \mathcal{N}(\mathcal{A}^n)$. Clearly, we have $h(f, \mathcal{A}) \geq 0$. The number $h(f, \mathcal{A})$ is called *the (topological) entropy of f on the cover $\mathcal{A}$*. Now we take

$$h(f) = \sup h(f, \mathcal{A}), \tag{3.5}$$

where supremum is taken over all open covers $\mathcal{A}$ of X, and $h(f)$ is the *topological entropy* of f.

We say that a cover $\mathcal{A}$ is *finer* than a cover $\mathcal{B}$, and write $\mathcal{A} \geq \mathcal{B}$, when each element of $\mathcal{A}$ is contained in some element of $\mathcal{B}$. In such a case we have $h(f, \mathcal{A}) \geq h(f, \mathcal{B})$. Since X is compact, from every open cover $\mathcal{A}$ of X we can choose a finite subcover $\mathcal{B}$. This $\mathcal{B}$ is finer than $\mathcal{A}$. Therefore it is enough to take in (3.5) the supremum only over all open finite covers.

Some basic properties of topological entropy are:

(a) $h(f^n) = n \cdot h(f)$,
(b) if $Y \subset X$ is an invariant closed set then $h(f|_Y) \leq h(f)$,
(c) if g is a factor of f, then $h(g) \leq h(f)$,
(d) if g is conjugate to f, then $h(g) = h(f)$.

Note also that $h(f^n, \mathcal{A}^n) = n \cdot h(f, \mathcal{A})$.

Now we consider the case when the space X is a compact interval I. We denote the set of all continuous maps from I into itself by $\mathcal{I}$.

The first nice property of the interval which we can use is that for every open cover there exists a finer open cover consisting of intervals. Therefore we may take in (3.5) the supremum only over all open finite covers of I by intervals. We can ask, do they have to be open? Formally, we can repeat the whole procedure from the definition of topological entropy using finite covers made of arbitrary intervals (we may consider also degenerate intervals, i.e., sets consisting of one point, as intervals). The question is whether we get also $h(f)$ after taking the supremum.

The answer to this question is positive for an important class of interval maps, namely piecewise monotone ones. A *partition* is a cover whose elements are pairwise disjoint. A map $f \in \mathcal{I}$ is called *piecewise monotone* if there exists a finite partition of I into intervals such that on each element of this partition f is monotone. We shall denote the set of all piecewise monotone maps $f \in \mathcal{I}$ by $\mathcal{M}$.

Lemma 3.1. *If $f \in \mathcal{I}$ and $\mathcal{A}$ is a finite cover of I by intervals then there exists an open cover $\mathcal{B}$ of I such that $h(f, \mathcal{A}) \le h(f, \mathcal{B}) + \log 3$.*

Proof. For every $A \in \mathcal{A}$ we take the partition $\mathcal{C}_A$ of I into A and the components of $I \setminus A$. Then $\mathcal{C} = \bigvee_{A \in \mathcal{A}} \mathcal{C}_A$ is a finite partition of I into intervals, finer than $\mathcal{A}$.

Since $\mathcal{C} \ge \mathcal{A}$, we have $h(f, \mathcal{C}) \ge h(f, \mathcal{A})$. For each element C of $\mathcal{C}$ there exists an interval $B(C)$, open in I, containing C and intersecting at most 3 elements of $\mathcal{C}$. Set $\mathcal{B} = \{B(C) : C \in \mathcal{C}\}$. If $\mathcal{D}$ is a subcover of $\mathcal{B}^n$, then each element of $\mathcal{D}$ intersects at most 3^n elements of $\mathcal{C}^n$. Since $\mathcal{D}$ is a cover, we obtain $\mathcal{N}(\mathcal{C}^n) \le 3^n \mathcal{N}(\mathcal{B}^n)$. In the limit, $h(f, \mathcal{C}) \le \log 3 + h(f, \mathcal{B})$. Consequently, $h(f, \mathcal{A}) \le h(f, \mathcal{B}) + \log 3$. □

Proposition 3.2. *If $f \in \mathcal{M}$ then* (3.5) *holds also if the supremum is taken over all finite covers of I by intervals (not necessarily open).*

Proof. Denote by h the supremum of $h(f, \mathcal{A})$ over all finite covers $\mathcal{A}$ of I by intervals. Let $\mathcal{A}$ be such a cover. Since f is piecewise monotone, for each positive integer n there exists a finite cover $\mathcal{C}$ of I by intervals which is finer than $\mathcal{A}^n$. From Lemma 3.1 we obtain $h(f^n, \mathcal{A}^n) \le h(f^n, \mathcal{C}) \le h(f^n) + \log 3$. Moreover, $h(f^n, \mathcal{A}^n) = n \cdot h(f, \mathcal{A})$ and $h(f^n) = n \cdot h(f)$. Therefore $h(f, \mathcal{A}) \le h(f) + (1/n) \log 3$. Since n is arbitrary, we obtain $h(f, \mathcal{A}) \le h(f)$. Since $\mathcal{A}$ is arbitrary, we obtain $h \le h(f)$.

On the other hand, for any open cover $\mathcal{B}$ we can find a finite cover $\mathcal{A}$ consisting of intervals and finer than $\mathcal{B}$. Then $h(f, \mathcal{B}) \le h(f, \mathcal{A}) \le h$. Since $\mathcal{B}$ is arbitrary, we obtain $h(f) \le h$. This completes the proof. □

If $f \in \mathcal{M}$, we shall call a cover $\mathcal{A}$ of I *f-mono* if it is finite, consists of intervals and f is monotone on each element of $\mathcal{A}$. Notice that then $\mathcal{A}^n$ is also f^n-mono.

Proposition 3.3. *If $f \in \mathcal{M}$ and $\mathcal{A}$ is an f-mono cover of I then $h(f, \mathcal{A}) = h(f)$.*

Proof. Let $\widetilde{\mathcal{B}}$ be a finite cover of I by intervals. Set $\mathcal{B} = \widetilde{\mathcal{B}} \vee \mathcal{A}$. Let $\mathcal{C}$ be a subcover of $\mathcal{A}^n$. Take $A \in \mathcal{C}$. The map $f^k|_A$ is monotone for $k = 1, 2, \dots, n-1$. Therefore for any $B \in \mathcal{B}$ the set $A \cap f^{-k}(B) = (f^k|_A)^{-1}(B)$ is an interval (unless it is empty). Each interval has at most 2 endpoints (notice that a degenerate interval has only 1 endpoint). Let x be an endpoint of an element D of $\mathcal{B}^n$. Then there exists $A \in \mathcal{C}$ such that $A \cap D \ne \emptyset$ and x is an endpoint of $A \cap D$. Since $D = \bigcap_{k=0}^{n-1} f^{-k}(B_k)$ for some $B_0, \dots, B_{n-1} \in \mathcal{B}$ and each of the sets $f^{-k}(B_k)$ is a union of finite number of intervals, x is an endpoint of some component of $f^{-k}(B_k)$, for some $k \in \{0, 1, \dots, n-1\}$. Hence, x is an endpoint of the interval $A \cap f^{-k}(B_k)$ for this k. In each $A \in \mathcal{C}$ there are at most $2n \operatorname{Card} \mathcal{B}$ such endpoints. The number of possible intervals with endpoints in a given set is not larger than 4 times the square of the cardinality of this set (we multiply by 4 because intervals with given endpoints may contain them or not). Therefore $\operatorname{Card} \mathcal{B}^n|_A \leq 4(2n \operatorname{Card} \mathcal{B})^2$. Hence, $\mathcal{N}(\mathcal{B}^n) \leq 4(2n \operatorname{Card} \mathcal{B})^2 \operatorname{Card} \mathcal{C}$. Since $\mathcal{C}$ was arbitrary, we obtain $\mathcal{N}(\mathcal{B}^n) \leq 4(2n \operatorname{Card} \mathcal{B})^2 \mathcal{N}(\mathcal{A}^n)$. In the limit we get $h(f, \widetilde{\mathcal{B}}) \leq h(f, \mathcal{B}) \leq h(f, \mathcal{A})$. By Proposition 3.2, since $\widetilde{\mathcal{B}}$ was arbitrary, we obtain $h(f) \leq h(f, \mathcal{A})$, and consequently $h(f) = h(f, \mathcal{A})$. □

Let c_n be the minimum of cardinalities of all f^n-mono covers of I. In other words, c_n is the number of *laps* of f^n (maximal intervals of monotonicity of f^n).

Theorem 3.4. *If $f \in \mathcal{M}$ then*

$$\lim_{n \to \infty} \frac{1}{n} \log c_n = h(f)$$

and $(1/n) \log c_n \geq h(f)$ for any n.

Proof. Let $\mathcal{A}_n$ be an f^n-mono cover of minimal cardinality (i.e., of cardinality c_n), $n = 1, 2, \dots$. For $m, k > 0$, the cover $f^{-k}(\mathcal{A}_m) \vee \mathcal{A}_k$ is an f^{m+k}-mono cover. Therefore

$$c_{m+k} \leq \operatorname{Card}(f^{-k}(\mathcal{A}_m) \vee \mathcal{A}_k) \leq c_m c_k.$$

Thus, the sequence $(\log c_n)_{n=1}^{\infty}$ is subadditive, and therefore the limit $\lim_{n \to \infty} (1/n) \log c_n$ exists and is equal to the infimum of $(1/n) \log c_n$ over all n. For each n we have from Proposition 3.3,

$$h(f) = \frac{1}{n} h(f^n) = \frac{1}{n} h(f^n, \mathcal{A}_n) \leq \frac{1}{n} \log \operatorname{Card} \mathcal{A}_n = \frac{1}{n} \log c_n.$$

On the other hand, for each n and k, the cover $\mathcal{B} = \bigvee_{i=0}^{k-1} f^{-in}(\mathcal{A}_n)$ is an f^{nk}-mono cover and therefore

$$\lim_{k \to \infty} \frac{1}{k} \log c_k = \lim_{k \to \infty} \frac{1}{nk} \log c_{nk} \leq \lim_{k \to \infty} \frac{1}{nk} \log \mathcal{N}(\mathcal{B}) = \frac{1}{n} h(f^n, \mathcal{A}_n) = \frac{1}{n} h(f^n) = h(f).$$

This completes the proof. □

Corollary 3.5. *If $f \in \mathcal{I}$ is a homeomorphism then $h(f) = 0$.*

Example 3.6. An example of the application of Theorem 3.4 is the computation of the entropy of a map f_n which looks as on Figure 4. The graph of f_n goes from the "bottom" to the "top" or vice versa n times. Then it is easy to see that the graph of f_n^k does the same n^k times. Therefore the number of laps of f_n^k is n^k and we get $h(f_n) = \log n$.

We note that we can define a map $f \in \mathcal{I}$ which is conjugate to f_{2n-1} on $[2^{-n}, 2^{-n+1}]$ for each $n \in \mathbb{N}$ and such that additionally $f(0) = 0$ (see Figure 5). Then we have $h(f) = +\infty$.

If $f \in \mathcal{I}$ and $s \geq 2$, then an *s-horseshoe* for f is an interval $J \subset I$ and a partition $\mathcal{D}$ of J into s subintervals such that the closure of each element of $\mathcal{D}$ f-covers J. For instance, the map from Figure 4 has a 6-horseshoe (as well as an s-horseshoe for all $s < 6$), and the map from Figure 5 has s-horseshoes for each s. We want to establish connections between the entropy of f and the existence of horseshoes for f and its iterates.

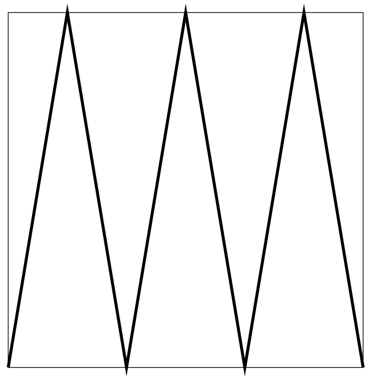

FIGURE 4. The graph of the map f_6.

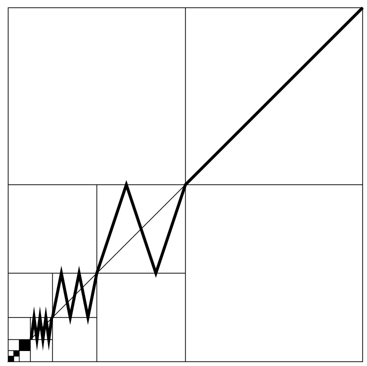

FIGURE 5. Interval map with infinite entropy.

Lemma 3.7. *If* $(J, \mathcal{D})$ *is an* s*-horseshoe for* f *then there exists a partition* $\mathcal{D}_n$ *of* J *such that* $(J, \mathcal{D}_n)$ *is an* s^n*-horseshoe for* f^n*.*

Proof. We use induction. For $n = 1$ there is nothing to prove.

Suppose that the induction hypothesis holds for some n. Let $D \in \mathcal{D}$ and $C \in \mathcal{D}_n$. Since $f(D)$ contains the interior of J, which contains the interior of C, there exists a subinterval $B(D, C)$ of D such that $f(B(D, C))$ is equal to the interior of C. The closure of the interval $B(D, C)$ f^{n+1}-covers J. The family

$$\{B(D, C) : D \in \mathcal{D},\ C \in \mathcal{D}_n\}$$

consists of pairwise disjoint intervals and therefore there exists a partition

$$\mathcal{D}_{n+1} = \{A(D, C) : D \in \mathcal{D},\ C \in \mathcal{D}_n\}$$

of J into intervals such that $B(D, C) \subset A(D, C)$ for each D, C. Then $(J, \mathcal{D}_{n+1})$ is an s^{n+1}-horseshoe for f^{n+1}. □

Proposition 3.8. *If* f *has an* s*-horseshoe then* $h(f) \geq \log s$*.*

Proof. Fix $n > 0$. By Lemma 3.7, f^n has an s^n-horseshoe. Call it $(J, \mathcal{D})$. Remove the first and the last element of $\mathcal{D}$ and call the remaining interval K. Since K is contained in the interior of J, each element D of $\mathcal{D}$ contains in its interior a closed interval which f^n-covers K. For each D we choose such an interval and call it $A(D)$. Then there exists an open cover $\mathcal{B}$ of I such that for each $D \in \mathcal{D}|_K$, the interval $A(D)$ intersects only one element $B(D)$ of $\mathcal{B}$ (then it has to be contained in it) and if $D, D' \in \mathcal{D}|_K$ with $D \neq D'$ then $B(D) \neq B(D')$. For $D_0, D_1, \dots, D_{k-1} \in \mathcal{D}|_K$ the set $\bigcap_{i=0}^{k-1} f^{-in}(A(D_i))$ is non-empty and intersects only one element of $\bigvee_{i=0}^{k-1} f^{-in}(\mathcal{B})$, namely $\bigcap_{i=0}^{k-1} f^{-in}(B(D_i))$. Therefore the sets $\bigcap_{i=0}^{k-1} f^{-in}(A(D_i))$ are different for different sequences $(D_0, D_1, \dots, D_{k-1})$, and thus

$$\mathcal{N}(\mathcal{B}_{f^n}^k) \geq (\operatorname{Card} \mathcal{D} - 2)^k.$$

Hence we obtain

$$h(f) = \frac{1}{n} h(f^n) \geq \frac{1}{n} h(f^n, \mathcal{B}) \geq \frac{1}{n} \log(\operatorname{Card} \mathcal{D} - 2) = \frac{1}{n} \log(s^n - 2).$$

Since n is arbitrary, we obtain $h(f) \geq \log s$. □

Remark 3.9. In the above proof, for a fixed n, we can replace f by any map $g \in \mathcal{I}$ which is sufficiently close to f in the topology of uniform convergence, and we get then

$$h(g) \geq \frac{1}{n} \log(s^n - 2).$$

Therefore, if f has an s-horseshoe then for every $\varepsilon > 0$ there exists a neighborhood U of f in the topology of uniform convergence such that for every $g \in U$ we have $h(g) \geq \log s - \varepsilon$.

The following theorem is basic in the investigation of the topological entropy for interval maps. It says that the entropy comes from horseshoes. Its proof is quite long and in order to keep these notes relatively short, we omit it.

Theorem 3.10. *Assume that $f \in \mathcal{I}$ has positive entropy. Then there exist sequences $(k_n)_{n=1}^\infty$ and $(s_n)_{n=1}^\infty$ of positive integers such that $\lim_{n\to\infty} k_n = \infty$, for each n the map f^{k_n} has an s_n-horseshoe and*

$$\lim_{n\to\infty} \frac{1}{k_n} \log s_n = h(f).$$

This theorem has several important consequences. For instance, from Remark 3.9 and Theorem 3.10 we get immediately the following corollary.

Corollary 3.11. *Topological entropy as a function from $\mathcal{I}$ (with the topology of uniform convergence) to $[0, \infty]$ is lower semicontinuous, that is, for every $f \in \mathcal{I}$ and $\varepsilon > 0$ there exists a neighborhood U of f in $\mathcal{I}$ such that $h(g) \geq h(f) - \varepsilon$ for every $g \in U$.*

Another straightforward corollary of Theorem 3.10 is the following result.

Corollary 3.12. *If $f \in \mathcal{I}$ then*

$$\limsup_{n\to\infty} \frac{1}{n} \log \operatorname{Card}\{x \in I : f^n(x) = x\} \geq h(f).$$

Let us recall that if $\varphi : I \to \mathbb{R}$ is a function then its *variation* is the supremum of the numbers

$$\sum_{i=0}^{s-1} |\varphi(x_i) - \varphi(x_{i+1})|$$

over all finite sequences $x_0 < x_1 < \cdots < x_s$ of elements of I. It is denoted $\operatorname{Var} \varphi$ or $\operatorname{Var}_I \varphi$.

Lemma 3.13. *If $f \in \mathcal{I}$ is a surjection then* $\operatorname{Var} f^{n+1} \geq \operatorname{Var} f^n$ *for all n.*

Proof. If f is a surjection then there exists a monotone (not necessarily continuous) map $g : I \to I$ such that $f \circ g$ is the identity on I. Let $x_0 < x_1 < \cdots < x_s$ be a sequence of points of I. Then we have either $g(x_0) < g(x_1) < \cdots < g(x_s)$ or $g(x_0) > g(x_1) > \cdots > g(x_s)$ and hence

$$\operatorname{Var} f^{n+1} \geq \sum_{i=0}^{s-1} \left| f^{n+1}(g(x_i)) - f^{n+1}(g(x_{i+1})) \right| = \sum_{i=0}^{s-1} \left| f^n(x_i) - f^n(x_{i+1}) \right|.$$

Taking supremum over all choices of $x_0 < x_1 < \cdots < x_s$, we get $\operatorname{Var} f^{n+1} \geq \operatorname{Var} f^n$. □

Let us denote the length of an interval J by $|J|$.

Theorem 3.14. *For any map $f \in \mathcal{M}$ we have*

$$\lim_{n\to\infty} \max\left(0, \frac{1}{n} \log \operatorname{Var} f^n\right) = h(f).$$

Proof. Clearly, $\operatorname{Var} f^n \leq |I| \cdot c_n$. Therefore, by Theorem 3.4,

$$\limsup_{n\to\infty} \frac{1}{n} \log \operatorname{Var} f^n \leq \lim_{n\to\infty} \frac{1}{n} \log c_n = h(f).$$

Let us prove the reverse inequality. If $h(f) = 0$ then there is nothing to prove. Assume that $h(f) > 0$. The set $J = \bigcap_{i=0}^{\infty} f^i(I)$ is a closed interval. It is easy to see that $h(f|_J) = h(f)$. Therefore the interval J is proper. Clearly, $\operatorname{Var}_I f^n \geq \operatorname{Var}_J f^n$. Therefore in order to prove that $\liminf_{n\to\infty}(1/n) \log \operatorname{Var} f^n \geq h(f)$ it is enough to show that $\liminf_{n\to\infty}(1/n) \log \operatorname{Var}_J f^n \geq h(f|_J)$. Thus, we may continue the proof under the assumption that $J = I$. This means that $f(I) = I$, that is, f is a surjection.

By Theorem 3.10 there are s_n-horseshoes $(J_n, \mathcal{D}_n)$ for f^{k_n} and

$$\lim_{n\to\infty} \frac{1}{k_n} \log s_n = h(f).$$

By Lemma 3.7, there are s_n^r-horseshoes $(J_n, \mathcal{D}_{n,r})$ for f^{rk_n}. Then we have $\operatorname{Var} f^{rk_n} \geq |J_n| \cdot s_n^r$. Thus,

$$\liminf_{r\to\infty} \frac{1}{rk_n} \log \operatorname{Var} f^{rk_n} \geq \frac{1}{k_n} \log s_n.$$

Let $i = rk_n + j$ with $0 \leq j < k_n$. Then by Lemma 3.13

$$\frac{1}{i} \log \operatorname{Var} f^i \geq \frac{1}{i} \log \operatorname{Var} f^{rk_n} \geq \left(1 - \frac{k_n}{i}\right) \frac{1}{rk_n} \log \operatorname{Var} f^{rk_n}.$$

Consequently, we get

$$\liminf_{i\to\infty} \frac{1}{i} \log \operatorname{Var} f^i \geq h(f).$$

This completes the proof. □

We shall say that f *has constant slope* s if on each of its laps it is affine with the slope coefficient of absolute value s.

Corollary 3.15. *Assume that $f \in \mathcal{M}$ has constant slope s. Then $h(f) = \max(0, \log s)$.*

Proof. If f has constant slope s then f^n has constant slope s^n. Then $\operatorname{Var} f^n = |I| \cdot s^n$. Therefore, by Theorem 3.14, $h(f) = \max(0, \log s)$. □

Notice that the proof of the inequality

$$\liminf_{n\to\infty} \frac{1}{n} \log \operatorname{Var} f^n \geq h(f)$$

did not use the assumption that f is piecewise monotone. However, the reverse inequality, as well as Corollary 3.15, uses this assumption.

Let us mention connections with combinatorial dynamics. If A is a pattern, we can define the entropy of A as the minimal topological entropy of a continuous interval map with a periodic orbit of this pattern. Clearly, if A forces B, then the entropy of A is larger than or equal to the entropy of B.

It is easy to compute the entropy of a pattern A. If the period of A is 1, then of course the entropy is 0. Otherwise, take a periodic orbit P of this pattern and the P-graph G. Then consider the 0-1 *transition matrix* M of G (the i, j-th entry of M is 1 if there is an arrow from the i-th vertex to the j-th one in G, and 0 otherwise). Then the entropy of A is equal to the logarithm of the spectral radius (the largest eigenvalue) of M.

It turns out that the minimal entropy of the pattern of odd (and larger than 1) period pattern is $\log \lambda_n$, where λ_n is the largest zero of the polynomial $x^n - 2x^{n-2} - 1$. This is of course the entropy of the Štefan pattern. Additionally we set $\lambda_1 = 1$. If the period is $2^k \cdot n$, where n is odd, the minimal entropy is $(1/2^k) \log \lambda_n$.

4. Invariant measures for interval maps

Let us recall several facts connected with invariant measures. Let X be a nonempty compact metric space and $f : X \to X$ a continuous map. A probability measure μ on X is called *invariant* if for every Borel set A we have $\mu(f^{-1}(A) = \mu(A)$. We will denote the space of all such measures by $\mathfrak{M}(X, f)$. We can look at them from another point of view. Let $C(X)$ be the space of all continuous functions on X with the supremum norm. Then the formula $\mu(\varphi) = \int \varphi \, d\mu$ gives a one-to-one correspondence between the space $\mathfrak{M}(X)$ of all probability measures on X and the cone of all continuous positive functionals on $C(X)$ of norm 1. Consequently, we may endow $\mathfrak{M}(X)$ with the weak-* topology: $\mu_n \to \mu$ if and only if for every $\varphi \in C(X)$ we have $\mu_n(\varphi) \to \mu(\varphi)$. By the Alaoglu theorem, $\mathfrak{M}(X)$ is compact in this topology. The set $\mathfrak{M}(X, f)$ is a closed subset of $\mathfrak{M}(X)$, so it is also compact.

The map $f : X \to X$ induces a map $f^* : C(X) \to C(X)$ by $f^*(\varphi) = \varphi \circ f$. Then, f^* induces $f_* : \mathfrak{M}(X) \to \mathfrak{M}(X)$ by $(f_*(\mu))(\varphi) = \mu(f^*(\varphi))$. It is easy to see that μ is f-invariant if and only if $f_*(\mu) = \mu$.

This gives us a natural way of looking for invariant measures. It is well illustrated by the proof of the following Krylov-Bogolubov Theorem.

Theorem 4.1. *The set $\mathfrak{M}(X, f)$ is nonempty.*

Proof. We start with an arbitrary measure $\mu \in \mathfrak{M}(X)$ and look at the sequence

$$\left(\frac{1}{n}\sum_{k=0}^{n-1} f_*^k(\mu)\right)_{n=1}^{\infty}.$$

It has a subsequence convergent to some measure $\nu \in \mathfrak{M}(X)$. For every $\varphi \in C(X)$ we have:

$$\left|\left(f_*\left(\frac{1}{n}\sum_{k=0}^{n-1} f_*^k(\mu)\right)\right)(\varphi) - \left(\frac{1}{n}\sum_{k=0}^{n-1} f_*^k(\mu)\right)(\varphi)\right| = \left|\frac{1}{n}(f_*^n(\mu))(\varphi) - \frac{1}{n}\mu(\varphi)\right| \le \frac{2}{n}\|\varphi\|.$$

Since $(2/n)\|\varphi\| \to 0$ as $n \to \infty$, we get $(f_*(\nu))(\varphi) = \nu(\varphi)$. This holds for all $\varphi \in C(X)$, and thus, ν is invariant. □

Since we will be interested in measures absolutely continuous with respect to a given one, we will need the following simple lemma.

Lemma 4.2. *Let $\nu \in \mathfrak{M}(X)$. Let measures μ_n ($n = 1, 2, \dots$) be absolutely continuous with respect to ν with the densities ρ_n, such that $\rho_n \le \rho$ for all n and for some ν-integrable function ρ. Suppose that $\mu_n \to \mu$ in the weak-* topology. Then μ is absolutely continuous with respect to ν.*

Proof. It is enough to show that if a set A has measure ν zero then also $\mu(A) = 0$. Suppose that $\nu(A) = 0$ but $\mu(A) > 0$. Then we can find: a compact set $B \subset A$ such that $\mu(B) > 0$, an open set $C \supset B$ such that $\int_C \rho\, d\nu < \mu(B)$, and a continuous function φ such that $0 \le \varphi \le 1$, $\varphi = 0$ outside C, and $\varphi = 1$ on B. We have

$$\mu_n(\varphi) = \int_X \varphi\rho_n\, d\mu \le \int_X \varphi\rho\, d\mu \le \int_C \rho\, d\nu$$

for all n, and consequently

$$\mu(B) \le \mu(\varphi) \le \int_C \rho\, d\nu < \mu(B),$$

a contradiction. □

Let us go back to the case when X is an interval. If f is smooth then the image of a measure which is absolutely continuous with respect to the Lebesgue measure, is also absolutely continuous. It is easy to find a density of the image of a measure with a given density ρ. It is equal to $P_f(\rho)$, where

$$(P_f(\rho))(x) = \sum_{y \in f^{-1}(x)} \frac{\rho(y)}{|f'(y)|}.$$

The operator P_f is called the *Perron-Frobenius operator* for f.

Now, in view of Lemma 4.2, using the same construction as in the proof of Theorem 4.1, we get the following result (we use the acronym *acim* for an absolutely continuous [with respect to the Lebesgue measure] invariant [probability] measure).

Proposition 4.3. *If all functions $P_f^n(1)$ are bounded by the same integrable function, then f has an acim.*

It turns out that the formalism and methods described above work also if a map is piecewise continuous (with finitely many pieces). In particular, we can define a *piecewise expanding map* as a map $f : [0,1] \to [0,1]$ (in fact, it can be any compact interval, but we want to make things simple) such that there exist points $a_0 = 0 < a_1 < a_2 < \cdots < a_m = 1$ such that on every $[a_{i-1}, a_i]$ the map f is differentiable and there is a constant $\alpha > 1$ such that $|f'| \geq \alpha$. At the points a_i by $f(a_i)$ and $f'(a_i)$ we mean the corresponding one-sided limits; we are not very much interested in what the actual values at these points are. Similarly we think of higher derivatives (if they exist). In particular, if we speak about class C^2, we mean that f'' exists and is continuous on closed intervals $[a_{i-1}, a_i]$ (and again, at the endpoints we take one-sided limits).

The following theorem is known as *Lasota-Yorke Theorem.*

Theorem 4.4. *Let f be a C^2 piecewise expanding interval map. Then it has an acim.*

We will prove here its weaker form (the general proof uses additionally bounded variation).

Theorem 4.5. *Let $f : [0,1] \to [0,1]$ be a C^2 piecewise expanding interval map, which maps every interval of continuity/monotonicity into $[0,1]$. Then it has an acim.*

Proof. Let K be an interval on which f^n is continuous and monotone. We will estimate the *distortion* of f^n on K, that is $\sup\{|(f^n)'(x)| : x \in K\}/\inf\{|(f^n)'(x)| : x \in K\}$. We want this estimate to be independent on n and K.

Set $\gamma = \sup\{|f''(x)/f'(x)| : x \in [0,1]\}$. Then the function $\log|f'|$ is Lipschitz continuous with constant γ on each interval on which f is continuous and monotone. Therefore, if f^n is continuous and monotone on an interval K, we have for $x, y \in K$

$$\begin{aligned}\log \frac{|(f^n)'(x)|}{|f^n)'(y)|} &= \sum_{k=0}^{n-1} (\log|f'(f^k(x)) - \log|f'(f^k(y))|) \\ &\leq \gamma \sum_{k=0}^{n-1} |f^k(x) - f^k(y)| \leq \sum_{k=0}^{n-1} |f^k(K)|.\end{aligned}$$

Since f is expanding with constant $\alpha > 1$, we have $|f^k(K)| \leq \alpha^{k-n}$, and therefore

$$\sum_{k=0}^{n-1} |f^k(K)| \leq \sum_{i=0}^{\infty} \alpha^{-i} = \frac{\alpha}{\alpha - 1}.$$

In such way we get

$$\frac{\sup\{|(f^n)'(x)| : x \in K\}}{\inf\{|(f^n)'(x)| : x \in K\}} \leq \delta,$$

where

$$\delta = e^{\alpha/(\alpha-1)}.$$

We can rewrite our inequality as

$$\sup\left\{\frac{1}{|(f^n)'(x)|} : x \in K\right\} \leq \delta \inf\left\{\frac{1}{|(f^n)'(x)|} : x \in K\right\}.$$

Each maximal interval K on which f^n is continuous and monotone is mapped by f^n onto all of $[0,1]$. Therefore, summing the above inequalities over all such K, we see that

$$\sup\{(P_f^n(1))(x) : x \in [0,1]\} \le \delta \inf\{(P_f^n(1))(x) : x \in [0,1]\}.$$

However, we have (by integration by substitution) $\int P_f^n(1) = \int 1 = 1$, so we get $\sup\{(P_f^n(1))(x) : x \in [0,1]\} \le \delta$ for all n. By Proposition 4.3, f has an acim. □

Now we want to investigate smooth interval maps. Unless they are globally monotone (and then they are not interesting from our point of view), they have critical points (where the derivative is 0). In particular, there is no chance of having expansion everywhere. However, there are smooth maps that behave similarly, except at critical points.

Assume that our map f is of class C^3. Its *Schwarzian derivative* is defined at all regular (non-critical) points by

$$Sf(x) = \frac{f'''(x)}{f'(x)} - \frac{3}{2}\left(\frac{f''(x)}{f'(x)}\right)^2.$$

Elementary computations show that

$$S(f \circ g) = ((Sf) \circ g) \cdot (g')^2 + Sg.$$

Therefore, if $Sf < 0$ and $Sg < 0$ then $S(f \circ g) < 0$. In particular, by induction we get that if $Sf < 0$ then $S(f^n) < 0$ for all n. Of course, when we write $Sf < 0$, we mean that this holds at regular points of f.

There are very important consequences of negative Schwarzian derivative for the dynamics of the map. A periodic orbit P of f is called *attracting* if $|(f^n)'(x)| < 1$ for $x \in P$, where n is the period of P. In fact, by the chain rule, $(f^n)'(x)$ does not depend on the choice of $x \in P$; it is equal to the product of the derivatives of f at all points of P.

As before, for simplicity we will assume that the domain of f is the interval $[0,1]$, although of course we can take any compact interval.

Before we state the first theorem, we prove a useful lemma.

Lemma 4.6. *If $Sg < 0$ then $|g'|$ has no positive local minima.*

Proof. Suppose that g' has a non-zero local extremum at x. Then $g''(x) = 0$, and consequently $g'''(x)/g'(x) = Sg(x) < 0$. Hence, if $g'(x) > 0$ then $(g')''(x) = g'''(x) < 0$ and g' has a local maximum at x; if $g'(x) < 0$ then $(g')''(x) > 0$ and g' has a local minimum at x. In both cases, $|g'|$ has a local maximum at x. □

Theorem 4.7. *If $Sf < 0$ then for every periodic attracting orbit there is a critical point of f or an endpoint of $[0,1]$, which is attracted by this orbit.*

Proof. Let P be an attracting periodic orbit of f of period n. If P contains a critical point c_1 and c_2 of f or an endpoint of $[0,1]$ then we are done. If not, take a point $x \in P$ and look at the nearest critical points of f^n (to the left and right; one of them may be an endpoint of $[0,1]$ instead). By Lemma 4.6, $|(f^n)'|$ does not have any non-zero local minima, so between c_1 and c_2 it has only one local extremum. Therefore, one of the points c_1, c_2 (call it c) has the property that $|(f^n)'|$ is monotone on the interval J joining x and c. Since $|(f^n)'(x)| < 1$ and $|(f^n)'(c) = 0|$, we have $|(f^n)'| < 1$

on the whole interval J. Thus, under the iterates of f^n this interval is contracted towards x. By the chain rule $f^k(c)$ is a critical point of f for some $k < n$ (again, it may be instead an endpoint of $[0,1]$), and this completes the proof. □

Thus, if we know that the endpoints of $[0,1]$ are not attracted to an attracting periodic orbit (for instance, $f(0) = f(1) = 0$ and $f'(0) > 1$), then the number of attracting periodic orbits is not larger than the number of critical points. In particular, if the map f is unimodal (has one critical point), this means that there is at most one attracting periodic orbit. This applies to maps like $x \mapsto ax(1-x)$ or $x \mapsto a\sin(\pi x)$, where it is easy to check that the Schwarzian derivative is negative.

There are many theorems about the existence of acim for smooth maps. Let us concentrate on the class $\mathcal{C}$ of unimodal maps for which $f(0) = f(1) = 0$, $f'(0) > 1$, and $f'' < 0$. The strongest result is *Jakobson's Theorem*. We will give one of its simple statements; it is possible to make it slightly stronger.

Theorem 4.8. *Let $f \in \mathcal{C}$ be such that $f([0,1]) = [0,1]$. Then there is a set $A \subset (0,1]$ of positive Lebesgue measure, and for which 1 is a Lebesgue density point, such that for every $a \in A$ the map af has an acim.*

There exist several proofs of this theorem. All of them are very complicated, and most of them contain many minor errors and gaps.

While Jakobson's Theorem deals with a family of maps, one can also ask whether there is an acim for concrete maps. One of the answers was given by the author of these notes.

Theorem 4.9. *Let $f \in \mathcal{C}$ be of class C^3, have negative Schwarzian derivative, and assume that there is $\varepsilon > 0$ such that $|f^n(c) - c| \geq \varepsilon$ for all $n > 0$, where c is the critical point of f. Then f has an acim.*

The assumption that there is $\varepsilon > 0$ such that $|f^n(c) - c| \geq \varepsilon$ for all $n > 0$ is not the strongest possible. One can for instance replace it by the assumption that the *Collet-Eckmann condition*

$$\liminf_{n\to\infty} \frac{1}{n} \log |(f^n)'(f(c))| > 0$$

is satisfied, which turn out to be weaker, but still implies the existence of an acim. There are also even weaker conditions that can be used here.

As with other theorems, we will present the proof of a special case of Theorem 4.9.

Theorem 4.10. *Let $f \in \mathcal{C}$ be of class C^3, have negative Schwarzian derivative, and assume that $f([0,1]) = [0,1]$. Then f has an acim.*

Proof. We start by estimating $P_f^n(1)$ by a function which is integrable far from the images of the critical point c under the iterates of f (i.e., in this case, far from 0 and 1). We will show that for all $n \geq 0$ and $x \in (0,1)$

$$(P_f^n(1))(x) \leq \frac{1}{\min(x, 1-x)}. \tag{4.1}$$

Let $a < b$ be two consecutive critical points of f^n (one of them may be instead an endpoint of $[0,1]$). Then either $f(a) = 0$ and $f(b) = 1$, or vice versa. There is exactly one $y \in (a,b)$ such that $f^n(y) = x$. By Lemma 4.6, $|(f^n)'|$ is bounded by $|(f^n)'(y)|$ on one of the intervals $[a,y]$, $[y,b]$. Therefore,

$$\min(x, 1-x) = \min\left(\int_a^y |(f^n)'|, \int_y^b |(f^n)'|\right) \le |(f^n)'(y)|(b-a).$$

Hence,

$$\frac{1}{|(f^n)'(y)|} \le \frac{b-a}{\min(x,1-x)}.$$

Summing these inequalities over all intervals of monotonicity of f^n, we get

$$\sum_{y \in f^{-n}(x)} \frac{1}{|(f^n)'(y)|} \le \frac{1}{\min(x,1-x)}.$$

However, the left-hand side is equal to $(P_f^n(1))(x)$, and (4.1) follows.

Now we want to check that all functions $P_f^n(1)$ are bounded by the same integrable function. Unfortunately, the function $1/\min(x,1-x)$ is not integrable in neighborhoods of 0 and 1. Therefore, we have to get additional bounds there.

Since f is of class C^2 (it is even C^3, but we do not need it here) and $f''(c) \ne 0$, for given $\varepsilon > 0$ there exist constants $\alpha, \omega > 0$ such that if $|x-c| < \varepsilon$ then

$$\alpha|x-c| \le |f'(x)| \le \omega|x-c|.$$

Consequently,

$$\frac{1}{2}\alpha(x-c)^2 \le 1 - f(x) \le \frac{1}{2}\omega(x-c)^2.$$

Hence, if $n \ge 1$, $x \ne y$, $f(x) = f(y) = z$, $|x-c| < \varepsilon$, and $|y-c| < \varepsilon$, then

$$\begin{aligned}(P_f^n(1))(z) = (P_f(P_f^{n-1}(1)))(z) &= \frac{(P_f^{n-1}(1))(x)}{|f'(x)|} + \frac{(P_f^{n-1}(1))(y)}{|f'(y)|} \\ &\le \frac{1}{\min(c,1-c)-\varepsilon}\left(\frac{1}{|f'(x)|} + \frac{1}{|f'(y)|}\right).\end{aligned}$$

However,

$$|f'(x)| \ge \alpha|x-c| \ge \alpha(2\omega^{-1}(1-z))^{1/2},$$

and the same is true with y instead of x. Consequently,

$$(P_f^n(1))(z) \le \eta(1-z)^{-1/2},$$

where

$$\eta = \frac{(2\omega)^{1/2}}{\alpha(\min(c,1-c)-\varepsilon)}.$$

This holds for all z such that $1 - z < \delta$, where $\delta = \alpha\varepsilon^2/2$. If ε is sufficiently small then our estimate holds also for $n = 0$.

Now look at a neighborhood of 0. If $\zeta < c$ then (since 0 is repelling) there are constants $\gamma \ge \beta > 1$, such that $\beta \le f'(x) \le \gamma$ for all $x \in [0,\zeta]$. There are also constants $\gamma_1 \ge \beta_1 > 0$, such that $\beta_1 \le |f'(x)| \le \gamma_1$ for all $x \in [1-\delta, 1]$. If ζ is sufficiently small then we may make β and γ close to each other (but not too close to 1). In particular, we may assume that $\gamma < \beta^2$.

If $f^n(y) = x$ then we look at the last k such that $f^k(y) \in (c, 1]$. According to this, we rewrite the formula for $(P_f^n(1))(x)$:

$$(P_f^n(1))(x) = \sum_{k=0}^{n} \frac{(P_f^k(1))(x_k)}{|(f^{n-k})'(x_k)|},$$

where for $k = 0$ we have $x_0, f(x_0), \dots, f^n(x_0) \in [0, \zeta]$, for $k > 0$ we have $x_k \in (c, 1]$ and $f(x_k), \dots, f^{n-k}(x_k) \in [0, \zeta]$, and $f^{n-k}(x_k) = x$ for all k. If ζ is sufficiently small then $x_k \in (1 - \delta, 1]$ for $k = 1, 2, \dots, n$. Hence,

$$(P_f^n(1))(x) \le \frac{1}{|(f^n)'(x_0)|} + \sum_{k=1}^{n} \frac{\eta(1 - x_k)^{-1/2}}{|(f^{n-k})'(x_k)|}.$$

We have $x \le \gamma_1 \gamma^{n-k-1}(1 - x_k)$ for $k \ge 1$. Moreover, $|(f^n)'(x_0)| \ge \beta^n \ge 1$ and $|(f^{n-k})'(x_k)| \ge \beta_1 \beta^{n-k-1}$ for $k \ge 1$. We get:

$$(P_f^n(1))(x) \le 1 + \sum_{k=1}^{n} \frac{\eta(x/(\gamma_1 \gamma^{n-k-1}))^{-1/2}}{\beta_1 \beta^{n-k-1}} < 1 + \frac{\eta \gamma_1^{1/2}}{\beta_1} \cdot \sum_{j=0}^{\infty} \left(\frac{\gamma^{1/2}}{\beta} \right)^j \cdot x^{-1/2}.$$

Since we have assumed $\gamma^{1/2}/\beta < 1$, we get $(P_f^n(1))(x) \le \vartheta x^{-1/2}$ for some constant $\vartheta > 0$ independent of n.

Since on neighborhoods of 0 and 1 the functions $\vartheta x^{-1/2}$ and $\eta(1 - x)^{-1/2}$ respectively are integrable, and outside these neighborhoods the function $1/\min(x, 1 - x)$ is integrable, the proof is complete. □

5. Rotation numbers

Let $\mathbb{S}$ be a circle. We shall consider it to be the unit circle $\{z : |z| = 1\}$ in the complex plane. The natural projection $e : \mathbb{R} \to \mathbb{S}$ is given by $e(X) = \exp(2\pi i X)$. Notice that $e(X) = e(Y)$ if and only if $X - Y \in \mathbb{Z}$.

If $f : \mathbb{S} \to \mathbb{S}$ is a continuous map then there exists a *lifting* of f, that is a continuous map $F : \mathbb{R} \to \mathbb{R}$ such that the diagram

$$\begin{array}{ccc} \mathbb{R} & \xrightarrow{F} & \mathbb{R} \\ {\scriptstyle e}\downarrow & & \downarrow{\scriptstyle e} \\ \mathbb{S} & \xrightarrow{f} & \mathbb{S} \end{array}$$

commutes. For every $k \in \mathbb{Z}$ the map $F + k$ (that is the map given by $(F + k)(X) = F(X) + k$ for $X \in \mathbb{R}$) is also a lifting of f. On the other hand, if G is a lifting of f then $G = F + k$ for some $k \in \mathbb{Z}$.

If $X \in \mathbb{R}$ then, since $e(X) = e(X + 1)$,

$$e(F(X)) = f(e(X)) = f(e(X + 1)) = e(F(X + 1)).$$

Therefore $d = F(X + 1) - F(X)$ is an integer. By continuity, d does not depend on the choice of X. It is called the *degree* of f. It depends really on f, not on F. If we take $G = F + k$ then

$$G(X + 1) - G(X) = (F(X + 1) + k) - (F(X) + k) = F(X + 1) - F(X) = d.$$

If x moves around the circle once then $f(x)$ moves around the circle d times (negative d means the opposite direction).

We shall concentrate our attention on the orientation preserving homeomorphisms. We shall denote the space of all of them by $\mathcal{L}$. They, and their iterates, have degree 1. The central role in their study is played by the *rotation numbers*.

Theorem 5.1. *Let $f \in \mathcal{L}$ and let F be its lifting. Then the limit*

$$\rho(F) = \lim_{n\to\infty} \frac{F^n(X) - X}{n}$$

exists for every $X \in \mathbb{R}$ and is independent of X.

Proof. If $X, Y \in \mathbb{R}$ then there exists an integer k such that $X + k \le Y \le X + k + 1$. We have

$$\begin{aligned} F^n(X) - X - 1 = F^n(X+k) - (X+k+1) &\le F^n(Y) - Y \\ &\le F^n(X+k+1) - (X+k) = F^n(X) - X + 1, \end{aligned}$$

so

$$\big|(F^n(X) - X) - (F^n(Y) - Y)\big| \le 1. \tag{5.1}$$

Since

$$\begin{aligned} F^{mn}(X) - X = (F^n(X) - X) + (F^n(F^n(X)) - (F^n(X))) + \dots \\ + (F^n(F^{(m-1)n}(X)) - F^{(m-1)n}(X)), \end{aligned}$$

by (5.1) we get $|m(F^n(X) - X) - (F^{mn}(X) - X)| \le m$. Therefore

$$\left|\frac{F^n(X) - X}{n} - \frac{F^{mn}(X) - X}{mn}\right| \le \frac{1}{n}.$$

By switching the roles of n and m we get

$$\left|\frac{F^m(X) - X}{m} - \frac{F^{mn}(X) - X}{mn}\right| \le \frac{1}{m}.$$

Therefore

$$\left|\frac{F^n(X) - X}{n} - \frac{F^m(X) - X}{m}\right| \le \frac{1}{n} + \frac{1}{m}.$$

This proves that the sequence $((F^n(X) - X)/n)_{n=1}^{\infty}$ is a Cauchy sequence and therefore it converges. By (5.1), the limit does not depend on X. □

We write $\rho(F)$ rather than $\rho(f)$, because the rotation number depends on the choice of a lifting. If we take another lifting $F + k$ then $(F + k)^n(X) = F^n(X) + nk$, so $\rho(F + k) = \rho(F) + k$. Therefore $e(\rho(F)) \in \mathbb{S}$ is already independent of the choice of a lifting. However, it is sometimes simpler to think about rotation numbers as real numbers, so we shall use rather $\rho(F) \in \mathbb{R}$, referring to it also as a rotation number of f.

It is easy to prove that the rotation number depends continuously on the map (in the topology of uniform convergence).

Set $R_\alpha(z) = z \cdot \exp(2\pi i\alpha)$. We call R_α the *rotation by* α (although strictly speaking this is the rotation by the angle $2\pi\alpha$). The map T_α given by $T_\alpha(X) = X + \alpha$ (the translation by α) is a lifting of R_α. We have $T_\alpha^n(X) - X = n\alpha$, so $\rho(T_\alpha) = \alpha$.

If α is rational, R_α is called a *rational rotation*; if α is irrational, R_α is called an *irrational rotation*. Note that for a rational rotation all points are periodic.

There is a basic difference between the dynamical properties of maps from $\mathcal{L}$ with rational and irrational rotation numbers.

Theorem 5.2. *Let $f \in \mathcal{L}$ and let F be its lifting. If $\rho(F)$ is irrational then f has no periodic points. If $\rho(F) = p/q$, where p and q are integers, $q > 0$, and p and q are relatively prime, then f has a periodic point and all periodic points of f have period q.*

Proof. Suppose that x is a periodic point of f of period q. Take $X \in \mathbb{R}$ with $e(X) = x$. Then $e(F^q(X)) = f^q(x) = x$, so $F^q(X) = X + p$ for some $p \in \mathbb{Z}$. We have $F^{nq}(X) = X + np$, so $\rho(F) = \lim_{n\to\infty}(F^{nq}(X) - X)/(nq) = p/q$. This shows that if $\rho(F)$ is irrational then f has no periodic points.

Assume now that $\rho(F) = p/q$, where p and q are integers, $q > 0$, and p and q are relatively prime. Look at $G = F^q - p$. It is a lifting of f^q. If G has no fixed points then there is $\varepsilon > 0$ such that either $G(X) - X \geq \varepsilon$ for every $X \in [0, 1]$ or $G(X) - X \leq -\varepsilon$ for all $X \in [0, 1]$. Since for $k \in \mathbb{Z}$ we have $G(X+k) = G(X)+k$, we get either $G(X) - X \geq \varepsilon$ for every $X \in \mathbb{R}$ or $G(X) - X \leq -\varepsilon$ for all $X \in \mathbb{R}$. Since $G^n(Y) - Y = \sum_{i=0}^{n-1}(G(G^i(Y)) - G^i(Y))$, we get either $\rho(G) \geq \varepsilon$ or $\rho(G) \leq -\varepsilon$. On the other hand, $G^n(Y)-Y = F^{nq}(Y)-np$, so $\rho(G) = q\rho(F)-p = 0$, a contradiction. This proves that G has a fixed point. If Z is this fixed point and $e(Z) = z$ then $f^q(z) = e(F^q(Z)) = e(Z + p) = z$, so f has a periodic point.

Let x be a periodic point of f of period m. Take $X \in \mathbb{R}$ with $e(X) = x$. We have $e(G^m(X)) = e(F^{mq}(X) - mp) = e(F^{mq}(X)) = f^{mq}(x) = x$, so $G^m(X) = X + l$ for some $l \in \mathbb{Z}$. If Z is a fixed point of G then for every $k \in \mathbb{Z}$ the point $Z + k$ is also a fixed point of G. Therefore, if X is not a fixed point of G then there are fixed points Z_1 and Z_2 of G such that $Z_1 < X < Z_2$, $Z_2 - Z_1 \leq 1$, and there are no fixed points of G in (Z_1, Z_2). Hence either $F(Y) > Y$ for all $Y \in (Z_1, Z_2)$ or $F(Y) < Y$ for all $Y \in (Z_1, Z_2)$. Therefore either $\lim_{n\to\infty} G^n(X) = Z_2$ or $\lim_{n\to\infty} G^n(X) = Z_1$. In both cases we get a contradiction with $G^m(X) = X + l$. This proves that X is a fixed point of G. Therefore $f^q(x) = e(F^q(X)) = e(X + p) = x$. This proves that m divides q. We have $e(F^m(X)) = e(X)$, so $F^m(X) = X + r$ for some $r \in \mathbb{Z}$. If $q/m = s$ then $F^q(X) = X + rs$, so $p = rs$. Thus, s is a common divisor of p and q. We assumed that those two numbers were relatively prime and hence $s = 1$. Therefore $q = m$. This proves that all periodic points of f have period q. □

In fact, from the above proof we get even more information about the dynamics of a map from $\mathcal{L}$ with a rational rotation number. Namely, if this rotation number is p/q with p and q as above, then for every $x \in \mathbb{S}$ the sequence $(f^{nq}(x))_{n=0}^{\infty}$ converges to a periodic point of f.

Now we start studying closer maps from $\mathcal{L}$ with irrational rotation numbers.

Theorem 5.3. *For an irrational rotation all orbits are dense.*

Proof. Let α be an irrational number. Suppose that R_α has an orbit which is not dense. This means that there exists $X \in \mathbb{R}$ and an interval $I \subset \mathbb{R}$ such that for every $n, k \in \mathbb{Z}$ with $n \geq 0$ the point $X + n\alpha - k$ does not belong to I.

There is a positive integer m such that the length of I is larger than $1/m$. Divide the circle into m arcs of equal length. Among the points $x, R_\alpha(x), R_\alpha^2(x), \dots, R_\alpha^m(x)$ (where $x = e(X)$) there are two which fall into the same arc. This means that there are integers i, j with $0 \le i < j \le m$ such that for some $r \in \mathbb{Z}$ the distance between $X + i\alpha$ and $X + j\alpha + r$ is smaller than $1/m$. Hence $|s\alpha + r| < 1/m$, where $s = j - i > 0$. Since $|s\alpha + k|$ is smaller than the length of I, among the points $X, X + s\alpha + r, X + 2s\alpha + 2r, X + 3s\alpha + 3r, \ldots$ there must be one which belongs to $I + p$ for some $p \in \mathbb{Z}$ (we define $I + p$ as $\{Y + p : Y \in I\}$). This means that $X + n\alpha$ belongs to $I + k$ for some $n, k \in \mathbb{Z}$ with $n \ge 0$, and this is what we wanted to prove. □

Theorem 5.4. *Let $f \in \mathcal{L}$ have an irrational rotation number α. Then it is semiconjugate to the rotation by α via a degree 1 map with monotone lifting.*

Proof. Let F be the lifting of f with $\rho(F) = \alpha$. We are looking for a map H such that the diagram

$$\begin{array}{ccc} \mathbb{R} & \xrightarrow{F} & \mathbb{R} \\ {\scriptstyle H}\downarrow & & \downarrow{\scriptstyle H} \\ \mathbb{R} & \xrightarrow{T_\alpha} & \mathbb{R} \end{array}$$

commutes and H is a monotone lifting of some degree 1 circle map h. Then h will be the desired semiconjugacy.

We start by choosing some $X \in \mathbb{R}$ and setting $H(X) = 0$. Then for $k, n \in \mathbb{Z}$ with $n \ge 0$ we have to set $H(F^n(X) + k) = n\alpha + k$. This we can do, since by Theorem 5.2 f has no periodic points, and consequently for different pairs (k, n) we get different points $F^n(X) + k$. We claim that the function H defined up to now (that is, on the set $A = \{F^n(X) + k : k, n \in \mathbb{Z},\ n \ge 0\}$) is increasing. To prove it assume that $F^n(X)+k < F^m(X)+l$. If $n = m$ then $k < l$ and $n\alpha+k < m\alpha+l$. Assume that $n < m$ (the proof in the case $m < n$ is similar). We have then $F^{m-n}(F^n(X))-F^n(X) > k-l$. Set $G = F^{m-n} - (k - l)$. Since f has no periodic points, G has no fixed points. Since $G(Z) > Z$ for $Z = F^n(X)$, we have the same inequality for all $Z \in \mathbb{R}$. As in the proof of Theorem 5.2 this leads to the conclusion that there exists $\varepsilon > 0$ such that $G(Z) \ge Z + \varepsilon$ for all $Z \in \mathbb{R}$, and consequently $\rho(G) \ge \varepsilon > 0$. Since $\rho(G) = (m - n)\rho(F) - (k - l)$, we get $\alpha = \rho(F) > (k - l)/(m - n)$. Therefore $k - l < (m - n)\alpha$, that is $n\alpha + k < m\alpha + l$. This proves the claim.

Now for any $Y \in \mathbb{R}$ we set $H(Y) = \sup\{H(Z) : Z \in A \text{ and } Z \le Y\}$. Since by Theorem 5.3 the set $H(A)$ is dense in $\mathbb{R}$, and $H|_A$ is increasing, we have also $H(Y) = \inf\{H(Z) : Z \in A \text{ and } Z \ge Y\}$. The map H defined in such a way is clearly nondecreasing. Since $H(A)$ is dense, H has no jump discontinuities, and therefore it is continuous. For $Z \in A$ we have $H(Z + 1) = H(Z) + 1$, and thus the same holds for any $Z \in \mathbb{R}$. Therefore H is a lifting of a degree 1 circle map. For $Z \in A$ we have $H(F(Z)) = H(Z) + \alpha = T_\alpha(H(Z))$, so the same holds for all $Z \in \mathbb{R}$. Hence, H is the semiconjugacy between F and T_α. This completes the proof. □

With a semiconjugacy h as above, the inverse images of points under h are either points or closed arcs. The inverse images of different points are disjoint, so there can be at most countable number of points with arcs as inverse images. Thus any map from $\mathcal{L}$ with irrational rotation number α can be obtained (up to a conjugacy) from the rotation by α by blowing up a point to an arc at most countable number of points. Notice that by continuity, if we blow up some point, we have to blow up all points of its full (positive and negative) trajectory. This describes almost fully the dynamics of those maps from topological point of view.

We continue our study of orientation preserving circle homeomorphisms. Let $f \in \mathcal{L}$ have an irrational rotation number. Let h be its semiconjugacy with a rotation, as in Theorem 5.4. When $y \in \mathbb{S}$ and $h^{-1}(y)$ is an arc, we will call the interior of this arc a *wandering arc*.

Notice that if A is a wandering arc (or any arc contained in a wandering arc) then $h(A)$ consists of one point y, and since the rotation number α of f is irrational, $R_\alpha^n(y) \neq y$ for any $n \neq 0$. Therefore $f^n(A) \cap A = \emptyset$ for any $n \neq 0$. Therefore we get $f^i(A) \cap f^j(A) = \emptyset$ if $i, j \in \mathbb{Z},\ i \neq j$.

On the other hand, if A is an open set not contained in the union of wandering arcs, then $h(A)$ contains an arc B. By Theorem 5.3, if $y \in B$ then there exists $n > 0$ such that $R_\alpha^n(y) \in B$, so $R_\alpha(B) \cap B \neq \emptyset$. Therefore there exists $n > 0$ such that $f^n(A) \cap A \neq \emptyset$.

Clearly, if there are no wandering arcs then h is a conjugacy. Thus, if we want to prove that some $f \in \mathcal{L}$ with an irrational rotation number α is conjugate to a rotation by α then it is enough to show that it has no wandering arcs.

Let α be an irrational number. Fix a point $x \in \mathbb{S}$ and look at the (forward) R_α-trajectory of x. By Theorem 5.3 it is dense in the circle, so from time to time it approaches x closer than ever before. Thus we get a sequence of *closest approaches*. It is closely related to the number-theoretical properties of α, in particular to the continued fraction representation of α. On the other hand, it is related to the existence of smooth conjugacies of diffeomorphisms with rotation number α with R_α. However, all that we will need here is the existence of this sequence. If n is one of those closest approaches then $R_\alpha^n(x)$ is closer to x than any of the points $R_\alpha^i(x),\ 0 < i < n$. This holds for every x, since R_α commutes with every rotation.

When f is smooth then we can consider its derivative $f'(x)$ at a point $x \in \mathbb{S}$ defined as the derivative $F'(X)$ of its lifting F at a point $X \in e^{-1}(x)$ (of course it does not depend on the particular choice of X).

Theorem 5.5. *Let $f : \mathbb{S} \to \mathbb{S}$ be an orientation preserving C^1 diffeomorphism with an irrational rotation number α. Assume that f' has bounded variation. Then f is conjugate to the rotation by α.*

Proof. Take a point $y \in \mathbb{S}$ and n which is one of the closest approaches. Let h be a semiconjugacy between f and R_α, as in Theorem 5.4. Take any point $z \in \mathbb{S}$ and set $z_i = f^i(z),\ y_i = h(z_i) = R_\alpha^i(h(z)),\ i \in \mathbb{Z}$. We claim that the points $y_0, y_1, \dots, y_{n-1}$ and $y_{-n}, y_{-n+1}, \dots, y_{-1}$ alternate on the circle, that is for any $k \in \{0, 1, \dots, n-1\}$ in any of the open arcs with endpoints y_k, y_{k-n} there are no points y_i with $-n \le i \le n-1$. This follows immediately from the property that y_0 is closer to y_n than any of the points y_i, $0 < i < n$, and that the distance from y_j to $y_{j\pm i}$ is the same as from y_0 from y_i. Since the semiconjugacy h preserves the cyclic order of the points on the circle, the points $z_0, z_1, \dots, z_{n-1}$ and $z_{-n}, z_{-n+1}, \dots, z_{-1}$ also alternate.

By the assumption, f' has bounded variation. Since f is a diffeomorphism, f' is bounded away from 0. Therefore $\log f'$ also has bounded variation. If this variation is bounded by V then we get

$$\big|(\log f'(z_0) - \log f'(z_{-n})) + (\log f'(z_1) - \log f'(z_{-n+1}) \\ + \cdots + (\log f'(z_{n-1}) - \log f'(z_{-1}))\big| \le V. \tag{5.2}$$

The expression at the left-hand side of (5.2) is equal to

$$\begin{aligned}\big|(\log f'(z_0) + \log f'(z_1) + \cdots + \log f'(z_{n-1}))& \\ - (\log f'(z_{-n}) + \log f'(z_{-n+1}) + \cdots + \log f'(z_{-1}))\big|& \\ = \big|\log(f^n)'(z_0) - \log(f^n)'(z_{-n})\big|&.\end{aligned}$$

Furthermore, we have $(f^n)'(z_{-n}) = 1/(f^{-n})'(z_0)$, so the expression at the left-hand side of (5.2) is equal to $\big|\log((f^n)'(z_0) \cdot (f^{-n})'(z_0))\big|$. Thus we get from (5.2)

$$\exp(-V) \le (f^n)'(z_0) \cdot (f^{-n})'(z_0) \le \exp(V), \tag{5.3}$$

where $z_0 = z$ is an arbitrary point of the circle.

Let A be an arc and set $A_i = f^i(A)$ for $n \in \mathbb{Z}$. We have $|A_i| = \int_A (f^i)'(x)\,dx$, so

$$|A_n| + |A_{-n}| = \int_A ((f^n)'(x) + (f^{-n})(x))\,dx.$$

Since the arithmetic mean of $(f^n)'(x)$ and $(f^{-n})(x)$ is greater than or equal to their geometric mean, we get from (5.3)

$$|A_n| + |A_{-n}| \ge 2\int_A ((f^n)'(x) \cdot (f^{-n})(x))^{1/2}\,dx \ge 2\exp(-V/2)|A|. \tag{5.4}$$

If A is a wandering arc then we know that $f^i(A) \cap f^j(A) = \emptyset$ for $i, j \in \mathbb{Z},\ i \ne j$. Therefore $\lim_{k\to\infty}(|A_k| + |A_{-k}|) = 0$, so (5.4) can hold only for finitely many n's. However, it holds for every n from the sequence of closest approaches, and there are infinitely many of those. Hence we get a contradiction. This proves that there is no wandering arc. Consequently, f is conjugate to the rotation by α. □

Let us make a couple of comments about the Denjoy Theorem. The assumption on the bounded variation of f' may seem difficult to verify. However, it is not. Namely, if f is of class C^2 then $|f''|$ is bounded by some constant M, and the same constant bounds the derivative of f'. Therefore the variation of f' is bounded by M times the length of the circle (which is 1 in our model). Thus, any orientation preserving C^2 diffeomorphism of the circle with irrational rotation number is conjugate to an irrational rotation. The second comment is that the density of orbits is an invariant of conjugacies. Thus, if a circle homeomorphism is conjugate to an irrational rotation then all its orbits are dense.

It turns out that the assumption on the bounded variation of f' is essential. Namely, for every irrational α there exists a C^1 orientation preserving circle diffeomorphism with rotation number α and not conjugate to R_α.

One can ask why in the definition of rotation numbers we take the expression $(F^n(X) - X)/n$ instead of $F^n(X)/n$, which gives the same result in the limit, but is simpler. The reason is that we can consider the *displacement function* $\varphi : \mathbb{S} \to \mathbb{R}$

by $\varphi(x) = F(X) - X$, where $e(X) = x$ (clearly, it does not depend on the choice of $X \in e^{-1}(x)$), and then we have

$$\frac{F^n(X) - X}{n} = \sum_{k=0}^{n-1} \varphi(f^k(x)).$$

Thus, $(F^n(X) - X)/n$ is the n-th *ergodic average* of the displacement function. By the *Ergodic Theorem*, ergodic averages of integrable functions converge almost everywhere for every invariant probability measure. Here they converge everywhere. If the rotation number is irrational, the reason for that is that the system is *uniquely ergodic*, that is, there is only one invariant probability measure. This implies that the ergodic averages of every continuous function converge uniformly to a constant.

Theorem 5.6. *Let $f : X \to X$ be a continuous map of a nonempty compact metric space X into itself. Assume that f is uniquely ergodic. Let μ be the unique element of $\mathfrak{M}(X, f)$. Then for every continuous function $\varphi : X \to \mathbb{R}$ its ergodic averages converge uniformly to $\int \varphi \, d\mu$.*

Proof. Suppose that there is a continuous function $\varphi : X \to \mathbb{R}$ such that its ergodic averages do not converge uniformly to $\int \varphi \, d\mu$. This means that there exists $\varepsilon > 0$ such that for every i there exists $n_i \geq i$ and $x_i \in X$ such that

$$\left| \frac{1}{n_i} \sum_{k=0}^{n_i-1} \varphi(f^k(x_i)) - \int \varphi \, d\mu \right| \geq \varepsilon.$$

Let δ_{x_i} be the Dirac delta at x_i, that is, the probability measure concentrated at the point x_i. Set

$$\nu_i = \frac{1}{n_i} \sum_{k=0}^{n_i-1} f_*^k(\delta_{x_i}).$$

Some subsequence of the sequence (ν_i) is convergent to a measure $\nu \in \mathfrak{M}(X)$. In the same way as in the proof of Theorem 4.1 we see that ν is invariant, and thus, $\nu = \mu$.

We have $f_*^k(\delta_{x_i}) = \delta_{f^k(x_i)}$, so from the definition of ν_i we get

$$\int \varphi \, d\nu_i = \frac{1}{n_i} \sum_{k=0}^{n_i-1} \varphi(f^k(x_i)).$$

Therefore by our assumption $\left| \int \varphi \, d\nu_i - \int \varphi \, d\mu \right| \geq \varepsilon$ for every i, which contradicts the fact that a subsequence of the sequence (ν_i) converges to μ. This completes the proof. □

Rotation theory for circle homeomorphisms can be generalized to other cases, where $f : X \to X$ is a continuous map of a compact metric space to itself and $\varphi : X \to \mathbb{R}$ (or more general, $\varphi : X \to \mathbb{R}^n$) a continuous function (an *observable*). In that case, the rotation number is replaced by the *rotation set*, that is, the set of all possible limits of ergodic averages of φ. Still, the most interesting results come from the cases where φ is the displacement function; for continuous degree 1 circle maps, homeomorphisms of the torus isotopic to the identity, and homeomorphisms of the annulus (we can think about the annulus as the product of a circle and an interval; the displacement is measured only in the direction of the circle).

References

[1] Ll. Alsedà, J. Llibre and M. Misiurewicz, *Combinatorial dynamics and entropy in dimension one*, World Scientific (Advanced Series in Nonlinear Dynamics, vol. 5), Singapore 1993 (Second Edition: 2000).

[2] L. S. Block and W. A. Coppel, *Dynamics in one dimension*, Lecture Notes in Math. **1513**, Springer, Berlin, 1992.

[3] W.-de Melo and S. van Strien, *One–dimensional dynamics*, Springer, Berlin, 1993.

Department of Mathematical Sciences, IUPUI, 402 N. Blackford Street, Indianapolis, IN 46202

E-mail address: mmisiure@math.iupui.edu