THE TEACHING OF MATHEMATICS
2024, Vol. XXVII, No. 1, pp. 15–26
DOI: 10.57016/TM-GWYW1150

TWO VIEWS ON ENTROPY IN DYNAMICAL SYSTEMS

Jelena Katić

Abstract. Banach's fixed point theorem is a part of standard curriculum of several university courses. It is also an example of a discrete dynamical system that is very regular – in the limit, the orbit of each point "ends" at a single fixed point. This is the starting point for this article. We begin by analyzing how small changes in the assumptions of this theorem affect the regularity of the system. We then discuss how the concept of regularity and chaos can be formalized. With this goal in mind, we talk about topological entropy.

We give definitions and some examples of topological and polynomial entropy in dynamical systems. We also explain two ways of looking at these dynamical invariants.

We also consider points that are in a sense the opposite to fixed points, namely *wandering points* and at the end we explain the role of wandering points in measuring the complexity of a dynamical system.

MathEduc Subject Classification: I95

AMS Subject Classification: 97I99

Key words and phrases: Topological entropy; dynamical system; fixed point; wandering point.

1. Introduction

A *topological dynamical system* consists of a pair (X, f), where (X, d) is a compact metric space and $f : X \to X$ is a continuous map. If f is additionally a homeomorphism, we say that the dynamical system is *reversible*.

We define f^0 to be the identity map id, and for $n \in \mathbb{N}$, $f^n := f \circ \ldots \circ f$ (n times). If f is reversible, we also define $f^{-n} := f^{-1} \circ \ldots \circ f^{-1}$. Since $f^{m+n} = f^m \circ f^n$, the set of maps $\{f^n\}$ forms a semigroup, for $n \in \mathbb{N}$, or – when f is reversible – a group, for $n \in \mathbb{Z}$.

The *positive orbit* of an element $x \in X$ is the set $\mathcal{O}_+(x) := \{f^n(x) \mid n \geq 0\}$. When f is reversible, we can define the *full orbit* $\mathcal{O}(x) := \{f^n(x) \mid n \in \mathbb{Z}\}$. One can also consider the finite orbits, of length n.

The simplest orbits are those of fixed points, they consist only of the point itself. The next simplest orbits are *periodic orbits*, i.e. the orbits of periodic points (a point x is *periodic* if there exists an integer $k > 0$ with $f^k(x) = x$).

We are interested in predicting the behavior of particular orbits, or a group of some orbits, or all orbits together, when $n \to \infty$.

Example 1. **(Banach's fixed point theorem)** Let $f : X \to X$ be a contraction, i.e. there exists $\lambda \in [0, 1)$ with $d(f(x), f(y)) \leq \lambda \cdot d(x, y)$ for all $x, y \in X$.

Then, according to the Banach fixed point theorem, there exists a unique fixed point p of f, and all orbits $f^n(x)$ converge to p, when $n \to \infty$. The Banach fixed point theorem holds in more general spaces, namely complete ones (in this paper, however, we are mainly concerned with compact spaces).

For the proof of this classical theorem see, e.g., [12].

The previous example illustrates a very simple asymptotic behaviour, since the sequence of iterates of any point converges to the same point.

If f is not a contraction but an isometry, we cannot expect asymptotic behaviour as in Example 1.

Example 2. Let $X = \mathbb{S}^1 = \{z \in \mathbb{C} \mid |z| = 1\}$ be the unit circle and $f = \rho_\theta$ the rotation

$$\rho_\theta(z) := e^{2\pi\theta i} z.$$

Then, if $\theta \in \mathbb{Q}$, every orbit of ρ_θ is periodic, with the same period, and if $\rho \notin \mathbb{Q}$, there are no periodic orbits. Moreover, if $\rho \notin \mathbb{Q}$, the orbit of every point is dense in $\mathbb{S}^1$.

We give an example of a slightly more complicated dynamical system.

Example 3. Let $f : [0, 1] \to [0, 1]$ be a non-decreasing continuous map. Then there exists a fixed point of f, possible not unique (this is a special case of the more general Brouwer fixed point theorem, see [10]). It is not hard to prove that for any $x \in [0, 1]$, the sequence $f^n(x)$ converges to some fixed point. It does not have to be a unique limit: take, for example, a strictly increasing map with the set of fixed points $p_1 < p_2 < \ldots < p_m$. Then we have:

- if $f(x) > x$ holds for all $x \in (p_{j-j}, p_j)$, then $f^n(x) \to p_{j-1}$ for all $x \in (p_{j-j}, p_j)$,
- if $f(x) < x$ holds for all $x \in (p_{j-j}, p_j)$, then $f^n(x) \to p_j$ for all $x \in (p_{j-j}, p_j)$,

see Figure 1.

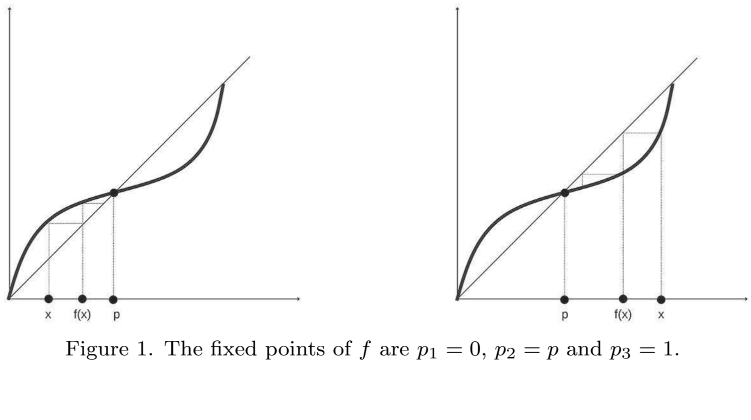

Figure 1. The fixed points of f are $p_1 = 0$, $p_2 = p$ and $p_3 = 1$.

The question that interests us is which dynamical system is more complicated than the other. We can discuss this question using Examples 1, 2 and 3. The first natural conclusion is that a contraction has the simplest orbital behavior, for two reasons: all orbits converge and they all converge to the same point. Moreover, it is obvious that an irrational rotation is more complex than a rational rotation. Furthermore, looking at the behavior of a single point orbit, it is safe to say that an irrational rotation has the most complicated orbits of all three examples. However, since it is an isometry, it is not possible to clearly determine whether an irrational rotation is more or less complex than a non-decreasing interval map when looking at all orbits. This motivates the introduction of a finer invariant that measures the complexity of a dynamical system, namely different types of entropy. These invariants must vanish for all isometries (see Example 6), so from this point of view, a non-decreasing interval map becomes more complex than any isometry.

What all these examples have in common, however, is that all orbits behave in the same way, either they all converge to some point, or they are all periodic, or they are all dense; this means that the dynamical system is not chaotic.

2. Topological entropy

The topological entropy is one of classical measures of the complexity of a dynamical system.

Let (X, d) and f be as above. Denote by $d_n^f(x, y)$ the *dynamic metric* (induced by f and d):

$$d_n^f(x, y) := \max_{0 \leq k \leq n-1} d(f^k(x), f^k(y)).$$

Two points x and y are close to each other (with respect to the metric d_n^f) if all iterates in the corresponding orbits of length n are close to each other.

EXAMPLE 4. If f is an isometry or a contraction, it holds $d(f^k(x), f^k(y)) \leq d(x, y)$, for all $k \geq 0$, so $d_n^f(x, y) = d(x, y)$. This means that the points do not move away from each other. This is not the case for any non-decreasing interval map. For example, if $f(x) = \sqrt{x}$, and $y \neq 0$, than $d(f^k(0), f^k(y)) \leq d(f^{k+1}(0), f^{k+1}(y))$, so $d_n^f(0, y) = d(f^{n-1}(0), f^{n-1}(y)) = \sqrt[2n-2]{y}$.

For $\varepsilon > 0$, we say that a finite set $E \subset X$ is (n, ε)*-separated* if for every $x, y \in E$ it holds $d_n^f(x, y) \geq \varepsilon$. Let $\operatorname{sep}(n, \varepsilon)$ denote the maximal cardinality of an (n, ε)-separated set E.

The number $\operatorname{sep}(n, \varepsilon)$ gives the maximal number of orbits of length n that can be distinguished, up to an error ε (meaning that the error is less than ε). This means that the orbit of any other point x of length n is less than or ε-close to the orbit of some point in E. The number $\operatorname{sep}(n, \varepsilon)$ increases with respect to n and decreases with respect to ε.

We are interested in the asymptotic behaviour of the number $\operatorname{sep}(n, \varepsilon)$ when $n \to \infty$ and $\varepsilon \to 0^+$.

DEFINITION 5. The *topological entropy* of the map f is defined by

$$h_{top}(f) := \lim_{\varepsilon\to 0} \limsup_{n\to\infty} \frac{\log \operatorname{sep}(n,\varepsilon)}{n} \in [0,\infty]. \tag{1}$$

EXAMPLE 6. If f is a Lipschitz map with a Lipschitz constant $\lambda \leq 1$, then $d_n^f = d$, so $\operatorname{sep}(n,\varepsilon) = \operatorname{sep}(1,\varepsilon)$ and $h_{top}(f) = 0$. This is the case in Example 1 and Example 2. If f is as in Example 3, it also holds $h_{top}(f) = 0$, but for some other reasons. However, another type of entropy will not vanish for such a map, see Example 16. This gives another, finer measure of the complexity of a dynamical system, which we will analyze further.

The topological entropy measures the *exponential* growth of the distinguished orbits, because if $\operatorname{sep}(n,\varepsilon) \sim \varphi(\varepsilon)e^{\alpha n}$, then $h_{top}(f) = \alpha$. It is also possible to measure a different growth, e.g. polynomial, logarithmic, etc. The reason for the choice of exponential growth in the basic definition of topological entropy lies in the definition of *the entropy of a random variable* in information theory (see [9]) as well as *metric entropy* in a measurable space (see [6]), both of which historically preceded the concept of topological entropy.

REMARK 7. There are other measures of complexity of a dynamical system. Let us mention one of them – chaos. As we have already said, a system is not chaotic if all iterates behave similarly. One way to describe chaos is that, informally, there are many periodic points, as well as points that are far from being periodic. More precisely, the system is *Devaney chaotic* if both:

- the set of periodic points, and
- the set of points x whose (positive) orbits are dense in X

are dense in X.[1]

In some situations, positive topological entropy implies Devaney chaos, and in some vice versa, Devaney chaos implies $h_{top} > 0$. In general, however, these two invariants are not equivalent.

In the rest of the paper we will show two ways of looking at topological entropy: a geometric one (as a measure of stretching) and a combinatorial one (as a measure of all possible codings of orbits).

2.1. Entropy as a measure of streching

We have already said that $h_{top}(f) = 0$ if f is a Lipschitz map with a Lipschitz constant $\lambda \leq 1$. Let us prove the following simple generalization of this fact.

Let $D(X)$ denote the *ball dimension* of a compact metric space, defined as

$$D(X) := \lim_{\varepsilon\to 0^+} \frac{\log \operatorname{cov}(\varepsilon)}{-\log \varepsilon},$$

[1] This is one of the definitions of chaos, due to Devaney. There are many more, e.g. Li-Yorke chaos, Block-Coppel chaos, etc. They are generally not equivalent, although in certain situations one type of chaos may imply another.

where $\operatorname{cov}(\varepsilon)$ is the minimal cardinality of the covering $\mathcal{U}_\varepsilon$ of X, consisting of balls of radii ε.

It is not difficult to see that the ball dimension of a closed cube in $\mathbb{R}^N$ or a compact manifold is equal to the standard dimension. Indeed, let $X = [0,1]^N \subset \mathbb{R}^N$ and

$$d\left((x_1, \ldots, x_N), (y_1, \ldots, y_N)\right) = \max\{|x_j - y_j| \mid j \in \{1, \ldots, N\}\}. \tag{2}$$

Then the balls of radius $1/n$ are the small cubes with the side lengths $1/n$ and

$$n^N \le \operatorname{cov}(1/n) \le (n+1)^N.$$

By taking the limit as $n \to \infty$ we obtain $D(X) = N$. It is easy to see that ball dimension coincides for equivalent metrics. The case when X is a compact manifold can be proven in a similar way.

Proposition 8. *If f is a Lipschitz map with a Lipschitz constant λ, then*

$$h_{top}(f) \le D(X) \cdot \max\{0, \log\lambda\}. \tag{3}$$

Here we want to notice how an asymptotic property of a dynamical system (the left-hand side in (3)) is controlled by the first step (Lipschitz constant) and a property of the space (the right-hand side in (3)).

Proof of Proposition 8. We have already discussed the case $\lambda \le 1$ in Example 6, so let us assume $\lambda > 1$. Since $d(f(x), f(y)) \le \lambda\, d(x,y)$, we have $d(f^k(x), f^k(y)) \le \lambda^k d(x,y)$ so

$$d_n^f(x,y) = \max_{0 \le k \le n-1} d(f^k(x), f^k(y)) \le \lambda^{n-1} d(x,y).$$

Therefore

$$\operatorname{sep}(n,\varepsilon) \le \operatorname{sep}\left(1, \frac{\varepsilon}{\lambda^{n-1}}\right). \tag{4}$$

It is easy to see that

$$\operatorname{sep}(1,\alpha) \le \operatorname{cov}\left(\frac{\alpha}{2}\right). \tag{5}$$

Indeed, let $\operatorname{sep}(1,\alpha) = m$ and $E = \{x_1, \ldots, x_m\}$ be $(1,\alpha)$-separated. If $\operatorname{cov}(\alpha) > m$, i.e. $\operatorname{cov}(\alpha) \ge m+1$, then there exist two points x_i and x_j from E that belong to the same open ball of radius $\alpha/2$. This implies $d(x_i, x_j) < \alpha$, which contradicts the fact that E is $(1,\alpha)$-separated.

Fix $\varepsilon > 0$. From (4) and (5) we get:

$$\begin{aligned}\frac{\log \operatorname{sep}(n,\varepsilon)}{n} &\le \frac{\log \operatorname{cov}\left(\frac{\varepsilon}{2\lambda^{n-1}}\right)}{n} = \frac{\log \operatorname{cov}\left(\frac{\varepsilon}{2\lambda^{n-1}}\right)}{-\log \frac{\varepsilon}{2\lambda^{n-1}}} \cdot \frac{-\log \frac{\varepsilon}{2\lambda^{n-1}}}{n} \\ &= \frac{\log \operatorname{cov}\left(\frac{\varepsilon}{2\lambda^{n-1}}\right)}{-\log \frac{\varepsilon}{2\lambda^{n-1}}} \cdot \frac{-\log\varepsilon + \log 2 + (n-1)\log\lambda}{n}.\end{aligned}$$

If we take the upper limit as $n \to \infty$ we obtain

$$\limsup_{n\to\infty} \frac{\log \operatorname{sep}(n,\varepsilon)}{n} \leq D(X)\cdot \log\lambda,$$

and we finish the proof by letting $\varepsilon \to 0$. ■

The previous proposition shows that the topological entropy is in a correlation with the maximal streching that the map f can make. In this sense, the following two examples are not surprising (but the rigorous proof is not so simple).

Example 9. One can define the topological entropy for a non-compact metric space X. If $K \subseteq X$ is compact, we denote by $\operatorname{sep}(n,\varepsilon;K)$ the maximal cardinality of an (n,ε)-separated set contained in K and

$$h_{top}(f;K) := \lim_{\varepsilon\to 0}\limsup_{n\to\infty} \frac{\log \operatorname{sep}(n,\varepsilon;K)}{n}.$$

Now we define

$$h_{top}(f) := \sup\{h_{top}(f;K) \mid K \subset X \text{ is compact}\}.$$

Let X be $\mathbb{R}^N$ and $f = L : \mathbb{R}^N \to \mathbb{R}^N$ be a linear map. Then

$$h_{top}(L) = \sum \log|\lambda_i|, \tag{6}$$

where the sum is taken over all eigenvalues λ of L with $|\lambda| > 1$. If all eigenvalues satisfy $|\lambda| \leq 1$, then $h_{top}(L) = 0$.

The proof of (6) is easy for $N = 1$ and for the case when the matrix of L is diagonalizable[2] and we leave it to the reader. In the general case, the proof is slightly more complicated and relies on the Jordan normal form of the linear map.

Example 10. Let $X = \mathbb{T}^2$ be the two-torus and

$$A := \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}.$$

Since A is a matrix with integer entries, the linear map $L_A : \mathbb{R}^2 \to \mathbb{R}^2$ defined by A induces the map T_A on $\mathbb{T}^2$. This map is known as *Arnold's cat map* and it can be proved that $h_{top}(T_A) = \log(3+\sqrt{5})/2$, as in the case of a linear map (note that the eigenvalues of A are $(3\pm\sqrt{5})/2$).

We would like to mention an important and deep result, due to Yomdin [11] and Newhouse [8], which has a similar flavor to the previous examples, in the sense that it concerns the relation between the entropy and a kind of dilatation of f (see also [2]).

Let M be a smooth Riemannian manifold (by smooth we mean C^∞-smooth). Denote by Σ the set of all smooth compact submanifolds of M. For a Riemannian

[2]For the latter the following property of entropy can be used: if $f \times g : X \times Y \to X \times Y$ is defined as $f \times g(x,y) := (f(x), g(y))$, then $h_{top}(f\times g) = h_{top}(f) + h_{top}(g)$.

metric g, and $\sigma \in \Sigma$ of dimension k, let $\mathrm{Vol}_g(\sigma)$ denote the k-dimensional volume of σ computed with respect to the measure on σ induced by g. Let $f : M \to M$ be a smooth diffeomorphism. Denote by

$$v(f, \sigma, n) = \mathrm{Vol}_g(f^n(\sigma)),$$

and

$$v(f, \sigma) := \limsup_{n\to\infty} \frac{\log v(f, \sigma, n)}{n}.$$

Finally, define the *volume growth* of f as

$$v(f) := \sup_{\sigma\in\Sigma} v(f, \sigma).$$

THEOREM 11. [Newhouse, Yomdin] *Let M and f be as above. Then $h_{top}(f) = v(f)$.* ■

3. Coding

In this section, we will try to look at the meaning of topological entropy from a different angle, which is purely combinatorial.

3.1. Topological entropy via open covers

The following equivalent definition of the topological entropy involves only open sets. From the fact that the two definitions coincide, it follows that the topological entropy depends only on the topology defined by the given metric, and not on the metric itself (hence the term topological).

Let $\mathcal{U}$ be an open cover of X, i.e. $X = \bigcup_{U\in\mathcal{U}} U$ and all $U \in \mathcal{U}$ are open. For a compact set X every open cover has a finite subcover. If we have m open covers $\mathcal{U}_1, \ldots, \mathcal{U}_m$, define their *join* by:

$$\bigvee_{j=1}^{m} \mathcal{U}_j := \left\{ \bigcap_{j=1}^{m} U_j \mid U_j \in \mathcal{U}_j \right\}.$$

If $f : X \to X$ is a continuous map, and $m \geq 0$ an integer, define

$$f^{-m}\mathcal{U} := \left\{ f^{-m}(U) \mid U \in \mathcal{U} \right\}.$$

Note that both $\bigvee_{j=1}^{m} \mathcal{U}_j$ and $f^{-m}\mathcal{U}$ are open covers of X. Denote by

$$\mathcal{U}_f^n := \bigvee_{j=0}^{n-1} f^{-j}(\mathcal{U}).$$

Denote by $N(\mathcal{U})$ the minimal cardinality of all open finite subcovers of $\mathcal{U}$. The *topological entropy of f relative to the open cover $\mathcal{U}$* is defined as

$$h_{top}(f, \mathcal{U}) := \lim_{n\to\infty} \frac{\log N(\mathcal{U}_f^n)}{n}. \tag{7}$$

One can prove that the sequence $a_n := \log N\left(\mathcal{U}_f^n\right)$ satisfies $a_{m+n} \leq a_n + a_m$, so the above limit exists (and is finite) due to Fekete's subadditive lemma.

The following theorem gives the relation between topological entropy defined by the metric, and the topological entropy defined via open covers.

THEOREM 12. *It holds:*

$$h_{top}(f) = \sup_{\mathcal{U}} h_{top}(f, \mathcal{U}), \tag{8}$$

where the above supremum is taken over all open covers $\mathcal{U}$ of X induced by the metric d. ∎

The proof of Theorem 12 is quite elementary, but very long and technical, and it can be found in most classical textbooks on dynamical systems (see [6, 7]).

REMARK 13. Let us discuss the definition of $N(\mathcal{U}_f^n)$, for a fixed finite open cover $\mathcal{U}$ that has no proper subcover (meaning that the union of any proper subset of $\mathcal{U}$ is not the whole space X). Let $\mathcal{U} = \{U_1, \dots, U_k\}$. We say that the set $\mathcal{U}$ is an *alphabet*, and that any finite sequence of U_j's is a *word*. We say that a word

$$(U_{i_0}, U_{i_1}, \dots, U_{i_{l-1}}) \in \mathcal{U}^l$$

is a *coding* for a finite sequence

$$(x_0, \dots, x_{l-1}) \in X^l,$$

if $x_j \in U_{i_j}$ for every $j \in \{0, \dots, l-1\}$. Note that the coding of a given sequence does not have to be unique.

Consider the set $\mathcal{U}_f^n$. An element $U \in \mathcal{U}_f^n$ is of the form

$$U = U_{i_0} \cap f^{-1}(U_{i_1}) \cap \dots \cap f^{-(n-1)}\left(U_{i_{n-1}}\right),$$

for $U_{i_j} \in \mathcal{U}$. If $x \in U$, then $f^j(x) \in U_{i_j}$ for $j \in \{0, \dots, n-1\}$, so the word $(U_{i_0}, U_{i_1}, \dots, U_{i_{n-1}})$ is a coding for the finite orbit $(x, f(x), \dots, f^{n-1})$. So $N(\mathcal{U}_f^n)$ is the number of all possible codings for all possible orbits (of length n) of all points in X. Therefore the term $h_{top}(f, \mathcal{U})$ measures the exponential growth of the number of words from the fixed alphabet $\mathcal{U}$, that are needed to describe all possible orbits of the system.

In certain situations, it turns out that the supremum in (8) does not have to be taken over all open covers of X. This can be particularly convenient for computations or estimates of the entropy.

Let $\mathcal{U}$ be a finite open cover, we define its *diameter* as

$$\mathrm{diam}(\mathcal{U}) := \max\{\mathrm{diam}(U) \mid U \in \mathcal{U}\}.$$

The following proposition provides a slight simplification for the calculations. The proof is in the same spirit as that of Theorem 12.

Proposition 14. *Let* $\mathcal{U}_n$ *be a sequence of open covers that satisfy* $\text{diam}(\mathcal{U}_n) \to 0$*, as* $n \to \infty$*. Then the limit* $\lim_{n\to\infty} h_{top}(f, \mathcal{U}_n)$ *exists and*

$$h_{top}(f) = \lim_{n\to\infty} h_{top}(f, \mathcal{U}_n). \quad \blacksquare$$

3.2. Polynomial entropy in wandering setting

Let us try to illustrate a possible reduction of the set of all finite open covers to a smaller set. We consider an invariant that measures polynomial instead of exponential orbit growth, i.e., polynomial entropy, in a special (wandering) environment.

A subset $A \subset X$ that satisfies the condition $f^n(A) \cap A = \emptyset$ for all $n \geq 1$, is called a *wandering set*. A point $x \in X$ is *wandering* if it has a wandering neighbourhood. The point that is not wandering is said to be *non-wandering*. More precisely, the point $x \in X$ is non-wandering if for every neighbourhood $U \ni x$ there exists $n \geq 1$ with $f^n(U) \cap U \neq \emptyset$.

Example 15. If $f : \mathbb{S}^1 \to \mathbb{S}^1$ is a rational rotation as in Example 2, then every subset of $\mathbb{S}^1$ is not wandering and every point $x \in \mathbb{S}^1$ is non-wandering. More generally, for any dynamical system every periodic (and in particular fixed) point is non-wandering.

In the case of a non-decreasing continuous interval map, every point that is not fixed is wandering. Namely, if $f(x_0) > x_0$ then for some $\delta > 0$ the interval $(f(x_0-\delta), f(x_0+\delta)) = f(x_0-\delta, x_0+\delta)$ is to the right of the interval $(x_0-\delta, x_0+\delta)$ (i.e. $f(x_0 - \delta) > x_0 + \delta$), and since f is non-decreasing, this is true for all $f^n(x_0 - \delta, x_0 + \delta)$.

If $f : X \to X$ is a contraction, then every point except the unique fixed point p is wandering. To see this, take $\varepsilon > 0$ and n_0 such that:

- $B(x_0; \varepsilon) \cap B(p; \varepsilon) = \emptyset$,
- for $n \geq n_0$ $f^n(x) \in B(p; \varepsilon)$.

All points $x_0, f(x_0), \dots, f^{n_0}(x_0)$ are different from each other, because otherwise f would have a periodic point, which contradicts the fact that f is a contraction.[3] Choose $\delta \leq \varepsilon$ such that the balls $B(f^j(x_0); \delta)$ are pairwise disjoint for $j \in \{0, 1, \dots, n_0\}$. The ball $B(x_0; \delta)$ is a wandering neighbourhood of x_0.

Note also that a point x can be such that the set $\{x\}$ is wandering but the point x itself is non-wandering. This is the case for every point $x \in \mathbb{S}^1$ and an irrational rotation f.

Let $Y \subset X$ be a set consisting only of wandering points. Denote

$$M(Y) := \sup_{x\in X} \sharp \{n \in \mathbb{N} \mid f^n(x) \in Y\},$$

[3]Indeed, suppose $f^m(y) = y$ for $m \geq 2$ and y is not fixed. Then it holds $d(f^m(y), f^{m+1}(y)) = d(y, f(y))$ which is impossible since $d(f^m(y), f^{m+1}(y)) \leq \lambda^m d(y, f(y)) < d(y, f(y))$.

where $\sharp A$ stands for the cardinality of the set A. If Y is compact, it can be covered by a finite number of open wandering sets. No orbit can intersect a wandering set twice, since

$$f^k(x), f^m(x) \in A,\ m > k \ \Rightarrow\ f^{m-k}(A) \cap A \ni f^m(x).$$

Therefore $M(Y) < \infty$.

Suppose that $Y_j \subset X$ are compact sets consisting only of wandering points, for $j \in \{1, \dots, k\}$. Denote by $Y_\infty := X \setminus \bigcup_j Y_j$. We think of $\{Y_1, \dots, Y_k, Y_\infty\}$ as of an alphabet. Denote by

$$\mathcal{A}_n := \mathcal{A}_n(f; Y_1, \dots, Y_n)$$

the set of all codings of all orbits $\{x, f(x), \dots, f^{n-1}(x)\}$ of length n, i.e. $\mathcal{A}_n$ is the set of all n-tuples $(w_0, w_1, \dots, w_{n-1})$, where $w_j \in \{Y_1, \dots, Y_k, Y_\infty\}$ and there exists $x \in X$ such that $f^j(x) \in w_j$, for all $j \in \{0, \dots, n-1\}$. (Of course, the family $\{Y_1, \dots, Y_k, Y_\infty\}$ is not an open cover of X, but the coding is still well defined.)

Since all Y_j's (except Y_∞) are compact and contain only wandering points, each Y_j can occur at most $M(Y_j)^n$ times in each word $(w_0, w_1, \dots, w_{n-1})$. Therefore

$$\sharp \mathcal{A}_n \le \prod_{j=1}^{k} M(Y_j)^n, \tag{9}$$

i.e. the growth of the number of words corresponding to an alphabet chosen as described is at most polynomial.

This is a motivation for the definition of the *polynomial entropy*. The definition differs from (5) only in the denominator in the limits. More precisely, we define:

$$h_{pol}(f) := \lim_{\varepsilon \to 0} \limsup_{n \to \infty} \frac{\log \operatorname{sep}(n, \varepsilon)}{\log n}.$$

Example 16. As we saw in Example 6, if f is a Lipschitz map with a Lipschitz constant $\lambda \le 1$, then $\operatorname{sep}(n, \varepsilon) = \operatorname{sep}(1, \varepsilon)$ and $h_{pol}(f)$ also vanishes. This is the case of Example 1 and Example 2.

Let us prove that this does not have to hold for Example 3. Let $f : [0,1] \to [0,1]$ be a continuous strictly increasing map different from the identity and $x_0 \in [0,1]$ such that $f(x_0) \ne x_0$. It is easy to see that there exists $\delta > 0$ such that $f(x_0 - \delta, x_0 + \delta) \cap (x_0 - \delta, x_0 + \delta) = \emptyset$. Since f is monotone, it holds $f^k(x_0 - \delta, x_0 + \delta) \cap (x_0 - \delta, x_0 + \delta) = \emptyset$ for all $k \ge 1$. We claim that the points

$$x_0, f^{-1}(x_0), \dots, f^{-(n-1)}(x_0)$$

are (n, δ)-separated. Indeed, let $l, m \in \{0, \dots, n-1\}$ and $l < m$. We have:

$$d(f^l(f^{-m}(x_0), f^m(f^{-m}(x_0)) = d(f^{-(m-l)}(x_0), x_0) > \delta,$$

since $f^{-(m-l)}(x_0) \notin (x_0 - \delta, x_0 + \delta)$. We conclude that $\mathrm{sep}(n, \delta) \geq n$ so

$$h_{pol}(f) = \lim_{\varepsilon\to 0} \limsup_{n\to\infty} \frac{\log \mathrm{sep}(n,\varepsilon)}{\log n} \geq \limsup_{n\to\infty} \frac{\log \mathrm{sep}(n,\delta)}{\log n} \geq 1.$$

There is an analogous definition of the polynomial entropy via open covers, one should put $\log n$ instead of n in the denominator in (7). It is obvious that

$$h_{top}(f) > 0 \;\Rightarrow\; h_{pol}(f) = \infty.$$

Considering Remark 13 and the polynomial growth (9), one can guess that there is a connection between the cardinality of the set of all codings by a family of wandering sets and the polynomial entropy. More precisely, we have the following proposition.

Proposition 17. [3] *Suppose that a reversible dynamical system $f : X \to X$ has a unique non-wandering point. Let $Y_j \subset X$ be compact sets consisting only of wandering points, for $j \in \{1, \dots, k\}$ and $Y_\infty := X \setminus \bigcup_j Y_j$. If we define the polynomial entropy of the family $\{Y_1, \dots, Y_k\}$ as*

$$h_{pol}(f; Y_1, \dots, Y_k) := \limsup_{n\to\infty} \frac{\log\, \sharp\mathcal{A}_n(f; Y_1, \dots, Y_k)}{\log n},$$

then:

$$h_{pol}(f) = \sup\{h_{pol}(f; Y_1, \dots, Y_k)\},$$

where the supremum is taken over all finite families $\{Y_1, \dots, Y_k\}$ of compact sets consisting only of wandering points. ■

We can also reduce the set of all possible codings to the set of all finite sets. Let us assume that the points $x_1, \dots, x_k$ are wandering and choose k decreasing sequences $U_{j,n}$, for $j \in \{1, \dots, k\}$ such that $\{U_{j,n}\}_{n\in\mathbb{N}}$ forms a basis of neighbourhoods of x_j. In [3], it is proved that the limit $\lim_{n\to\infty} h_{pol}(f; U_{1,n}, \dots, U_{k,n})$ exists and does not depend on the choice of $U_{j,n}$ but only on the points $x_1, \dots, x_k$. Define

$$h^{loc}_{pol}(f; x_1, \dots, x_k) := \lim_{n\to\infty} h_{pol}(f; U_{1,n}, \dots, U_{k,n}).$$

The following proposition has a similar flavour to Proposition 14 (since in both cases the sets that are crucial in coding have small diameters).

Proposition 18. [3] *Let $f : X \to X$ be a reversible system with a unique non-wandering point. Then*

$$h_{pol}(f) = \sup\left\{h^{loc}_{pol}(f; x_1, \dots, x_k)\right\},$$

where the supremum is taken over all finite sets $\{x_1, \dots, x_k\}$ of wandering points. ■

Propositions 17 and 18 can be generalized for the case when f is not reversible and the set of non-wandering points is not a singleton but finite (see [4]).

In this subsection we have seen that the polynomial entropy is well suited for the wandering setting. With the topological entropy the situation is exactly the opposite. Indeed, if we denote by $NW(f)$ the set of all non-wandering points, then $h_{top}(f) = h_{top}\left(f|_{NW(f)}\right)$. The estimate (9) shows that the growth of the cardinality of the set of codings is at most polynomial for the wandering part of a dynamical system. However, a rigorous proof of the formula $h_{top}(f) = h_{top}\left(f|_{NW(f)}\right)$ is more complicated. For example, if f is a contraction from Example 1 we have $NW(f) = \{p\}$, and since obviously $h_{pol}\left(f|_{\{p\}}\right) = 0$, we get another proof of the fact mentioned in Example 6.

This is an interesting difference between the topological and the polynomial entropy: although their definitions are similar, and they share many properties (which we have not listed here), the first only recognizes the non wandering set, while the second also recognizes the wandering part.

Acknowledgements. The author would like to thank Darko Milinković for several valuable suggestions and Maša Đorić for the drawing of the figure. The author also thanks the anonymous referee for many useful remarks.

REFERENCES

[1] M. Brin and G. Stuck, *Introduction to Dynamical Systems*, Cambridge: Cambridge University Press, 2002.

[2] M. Gromov, *Entropy, homology and semialgrebraic geometry*, Séminaire Bourbaki 28 (1985-1986), 225–240.

[3] L. Hauseux and F. Le Roux, *Entropie polynomiale des homémorphismes de Brouwer*, Annales Henri Lebesgue **2** (2019), 39–57.

[4] J. Katić, M. Perić, *On the polynomial entropy for Morse gradient systems*, Math. Slovaca **69**, 3 (2019), 611–624.

[5] A. Katok and B. Hasselblatt, *Introduction to the Modern Theory of Dynamical Systems* (Encyclopedia of Mathematics and its Applications), Cambridge: Cambridge University Press, 1995.

[6] A. N. Kolmogorov, *New Metric Invariant of Transitive Dynamical Systems and Endomorphisms of Lebesgue Spaces*, Doklady of Russian Academy of Sciences, **119**, 5 (1958), 861–864.

[7] W. Merry, *Sixty Lectures on Dynamical Systems*, preprint.

[8] S. Newhouse, *Entropy and volume*, Egrodic Theory and Dynamical Systems, **8*** (Conley Memorial Issue), 283–300, 1988.

[9] C. E. Shannon, *A Mathematical Theory of Communication*, The Bell System Technical Journal, 27, 379–423, 1948.

[10] E. Spanier, *Algebraic Topology*, Springer-Verlag, Berlin, 1966.

[11] Y. Yomdin, *Volume growth and entropy*, Israel J. Math., **57** (1987), 285–300.

[12] V. A. Zorich, *Mathematical Analysis I*, 2nd ed., Spinger, 2015.

Faculty of Mathematics, Studentski trg 16, 11000 Beograd, Serbia
ORCID: 0000-0001-8927-0506
E-mail: `jelena.katic@matf.bg.ac.rs`

Received: 29.01.2024, in revised form 07.03.2024
Accepted: 11.03.2024