Radboud University Nijmegen

Faculty of Science

Topological entropy on compact spaces

Bachelor Thesis in Mathematics

Supervisor:
dr. A.Y. Burtscher
dr. J. Hedicke

Author:
Tamar Hinkema

Second reader:
dr. A. Javan Peykar

July 31, 2026

Contents

1 Introduction

Topological entropy is essentially a measure for the complexity of a topological dynamical system. It assigns a number in $\mathbb{R}_{\geq 0} \cup \{\infty\}$ to each continuous map $f : X \to X$ working on a compact topological space X. That number represents the exponential growth rate of the number of different orbit segments of length $n \in \mathbb{N}$. Two orbit segments are considered to be "different" if they are not too close to each other. In a metric space, we can say that two orbit segments have to be at least $\varepsilon > 0$ apart to be considered as separate orbit segments.

The first introduction of the concept of topological entropy was made by Roy Adler, Alan Konheim, and Harry McAndrew [1] in 1965. They introduced it as an invariant for continuous maps. In Section 3.3 of this thesis we will see that topological entropy is indeed invariant under topological conjugacy. That is, if there exists a homeomorphism $X \to Y$ such that $H \circ f = g \circ H$ for continuous functions $f : X \to X$ and $g : Y \to Y$, then the entropy for f and g is equal.

In 1970, Efim Isaakovich Dinaburg [7] gave an alternative definition of topological entropy using separating and spanning sets. This definition specifically requires a compact metric space, whereas the original definition works for any compact topological space. Then, in January 1971, Rufus Bowen [4] gave a similar definition, but he extended it so that it could be applied to any metric space that is not necessarily compact. He did so using compact subsets of the metric space X. Both the definition from Adler, Konheim, and McAndrew and the one from Bowen and Dinaburg are described only for single-valued functions. The notion of topological entropy can however be altered in a way that makes it applicable to set-valued functions as well. This altered version preserves most of the properties that hold for the single-value definitions, such as invariance under topological conjugacy. The definition of topological entropy for set-valued functions and its properties can be found in the paper of James Kelly and Tim Tennant [8].

Topological entropy has applications in fields like computer science and even biology. It can be used to define the complexity of formal languages. According to [11] it is suggested that so-called simple formal languages are languages that have zero entropy and it has been shown that regular languages have zero entropy. Furthermore, in [9] the notion of topological entropy was used to conclude that the entropy of introns is higher than that of exons in the human DNA. That is, the parts of the DNA that do not directly code for proteins and that are removed during RNA splicing (the introns) are more random than the parts of DNA that do provide coding for proteins (the exons). We will discuss these applications in more detail in Chapter 6.

In this thesis, we will mostly look at topological entropy as a way to be able to say that some dynamical systems are more "complex" than other dynamical systems. Consider for example the unit circle and two operations we can apply to it. One thing we can do, is rotate the circle with an angle $2\pi\alpha$ for some $\alpha \in \mathbb{R}$. After rotation, we get the same circle back. Intuitively, we can say that by rotating the circle, our system that consists of the unit circle and the rotation map does not change in complexity. Another operation we can do on the unit circle is to expand it by a factor $n \in \mathbb{N}$, and then fold it $n-1$ times in such a way that we obtain an n-cover of the unit circle. That would look like n times the unit circle placed on top of itself, which is decisively more complex than just the unit circle itself. Our goal is to find a measure that tells us exactly this. Topological entropy is such a measure.

The outline of this thesis is as follows. We start in Chapter 2 by introducing the notion of a discrete dynamical system and some of its properties. We will then look at two examples of discrete dynamical systems, using the unit circle. They are the two examples that were already shortly listed above. Chapter 3 is all about topological entropy. We first introduce the definition given by Roy Adler, Alan Konheim, and Harry McAndrew. We will also see the more recent definition of topological entropy that was independently given by Rufus Bowen and by Efim Isaakovich Dinaburg. One of the most important results mentioned in this thesis is the fact that these definitions are equivalent on compact metric spaces. This result is stated and proved at the end of Section 3.2. We will also discuss some properties of topological entropy and calculate the entropy in three examples. Lastly, we are interested in when the entropy is equal to zero. We will see several cases in which the entropy is zero in Chapter 5, but the cases discussed in this thesis are certainly not the only ones that are known. We will end the thesis by giving a short outlook on some applications of topological entropy.

Data management

No data were generated or used for the preparation of this thesis.

Use of generative AI

No generative AI was used for the preparation of this thesis.

2 Discrete dynamical systems

2.1 Definition

In this thesis we will consider discrete-time dynamical systems, which we will refer to as discrete dynamical systems or simply as dynamical systems.

Definition 2.1.1. A *discrete dynamical system* consists of a set $X \neq \emptyset$ and a function $f : X \to X$ working on that set.

We assume that, given a structure on X, the map f preserves the structure of X. For example, if X is a smooth manifold, then f has to differentiable. A structure-preserving map on a topological space X is a continuous map. Namely, by definition inverse images of open sets in X under a continuous function $f : X \to X$ are open in X. Later in Chapter 3 we will also be looking at metric spaces with a continuous function working on the space, but we will consider them still as topological dynamical systems, and not metric dynamical systems, since a continuous function does not necessarily preserve the structure of a metric space.

In a discrete dynamical system, we can consider certain "time-steps". For $n \in \mathbb{N}$, the n-th time-step of the dynamical system (X, f) is the system after composing f with itself n times. This is denoted by $f^n = f \circ \cdots \circ f$. We can also look at previous time-steps by considering $f^{-n} = f^{-1} \circ \cdots \circ f^{-1}$ if f is invertible. We consider f^0 to be the identity map Id. Throughout this thesis, we assume that $0 \notin \mathbb{N}$, and we will denote the set $\mathbb{N} \cup \{0\}$ as $\mathbb{N}_0$. Note that by definition

$$f^n \circ f^m = f^{n+m}$$

for all $n, m \in \mathbb{N}_0$ in general, and for all $n, m \in \mathbb{Z}$ if f is invertible. For each point in X we can look at where the point is in each time-step. What we then look at is called the orbit of the point.

Definition 2.1.2. Let $X \neq \emptyset$ be a set and let $f : X \to X$. The set

$$\{f^n(x) : n \in \mathbb{N}_0\}$$

is called the *positive semiorbit* of x, for some fixed $x \in X$. If f is invertible, we call the set

$$\{f^{-n}(x) : x \in \mathbb{N}_0\}$$

the *negative semiorbit* of x. The *orbit* of x is then given by the set

$$\{f^n(x) : n \in \mathbb{Z}\}.$$

Another important notion within the field of dynamical systems is that of periodic points.

Definition 2.1.3. Let $X \neq \emptyset$ and $f : X \to X$. If $f^n(x) = x$ for some $n \in \mathbb{N}$, we say that x is *periodic* with *period* n. If $f^n(x) = x$ and there exists no natural number $m < n$ such that $f^m(x) = x$, we say that n is the *minimal period* of x.

Any periodic point has a finite orbit. In fact, if $x \in X$ with minimal period $n \in \mathbb{N}$, then

$$f^k(x) = f^{k \mod n}(x)$$

and the orbit of x is given by $\{f^k(x) : k = 0, \dots, n-1\}$, assuming that f is invertible.

Remark 2.1.4. We do not consider continuous-time dynamical systems, because we only define the notion of topological entropy for discrete-time topological dynamical systems. It can however be extended to the continuous-time case.

2.2 Example: circle maps

In this section we will look at two examples of topological dynamical systems, namely rotations and expansions of the unit circle. Later, we will be able to compute the topological entropy of the maps given here. More information on the basic concept of dynamical systems can be found in [5].

The first example of a dynamical system we will be looking at is the rotation of a circle. We define the unit circle as $S^1 = [0,1]/\{0,1\}$. This notation of the unit circle can be thought of as taking the interval $[0,1]$ and identifying the point 1 with 0 to turn the interval into a circle, as is shown in Figure 1.

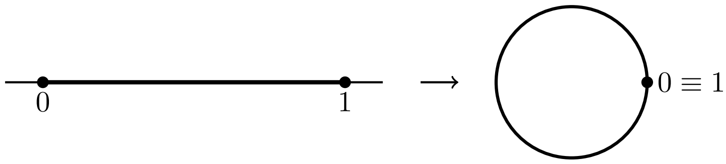

Figure 1: Representation of the unit circle

We can define the distance

$$d(x,y) := \min\{|x-y|, 1-|x-y|\} \tag{2.1}$$

on the circle. This distance is induced by the natural distance. The natural distance on the interval $[0,1]$ is given by

$$d_{nat}(a,b) := |a-b|$$

for $a, b \in [0,1]$. Let $x, y \in S^1$. Then we can walk from point x to point y either clockwise or counter-clockwise over the circle. The distance between x and y that is induced by the natural distance on $[0,1]$ is then either $|x-y|$ when we walk clockwise, or $1-|x-y|$ if we walk counter-clockwise. Note here that the length of the circle is one. Since the distance between two points is always the length of the shortest path between two points, it should be the shortest of the two distances described above. We find that

$$d(x,y) = \min\{|x-y|, 1-|x-y|\}$$

is a distance on S^1 induced by the natural distance on $[0,1]$.

Any point $x \in [0,1]$ relates to a point on S^1 by the relationship $x \sim x+1$. We can define the function $R_\alpha : S^1 \to S^1$, given by

$$R_\alpha(x) = x + \alpha \mod 1,$$

as being the rotation of the unit circle by the angle $2\pi\alpha$ for $\alpha \in \mathbb{R}$. Figure 2 is a representation of such a rotation of the circle.

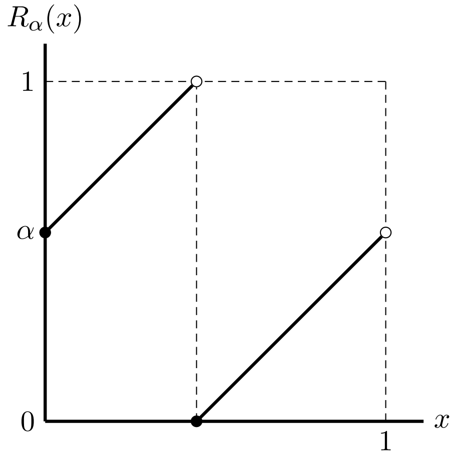

Figure 2: Rotation of the unit circle by degree $2\pi\alpha$

Proposition 2.2.1. *The function R_α is an isometry on the unit circle for any $\alpha \in \mathbb{R}$.*

Proof. Let $\alpha \in \mathbb{R}$ and let $x, y \in S^1$ be points on the unit circle. To prove that R_α is an isometry we need to show that it preserves the distance between the arbitrarily chosen points x and y. Using the notion of distance defined in 2.1 we get

$$\begin{aligned} d(R_\alpha(x), R_\alpha(y)) &= \min\{|R_\alpha(x) - R_\alpha(y)|, 1 - |R_\alpha(x) - R_\alpha(y)|\} \\ &= \min\{|x + \alpha - y - \alpha \mod 1|, 1 - |x + \alpha - y - \alpha \mod 1|\} \\ &= \min\{|x - y \mod 1|, 1 - |x - y \mod 1|\} \end{aligned}$$

This is exactly equal to $d(x, y)$. Namely, since $x, y \in S^1$, it holds that $x - y \in (-1, 1)$. If $x - y \in [0, 1)$, then

$$|x - y \mod 1| = |x - y|$$

which implies that $d(R_\alpha(x), R_\alpha(y)) = d(x, y)$. In the case that $x - y \in (-1, 0)$ it holds that

$$|x - y \mod 1| = 1 - |x - y|$$

which also implies $d(R_\alpha(x), R_\alpha(y)) = d(x, y)$. We can conclude that R_α is an isometry. $\square$

Proposition 2.2.2. *[3, Proposition 2.1] Let $\alpha \in \mathbb{R}$ and $R_\alpha : S^1 \to S^1$ be the rotation map. If $\alpha = \frac{p}{q} \in \mathbb{Q}$ with p and q coprime, then every point $x \in S^1$ is a periodic point of R_α with period q. Moreover, if $\alpha \in \mathbb{R} \setminus \mathbb{Q}$, then no point of S^1 is a periodic point of R_α.*

Proof. For any $n \in \mathbb{Z}$ and $x \in S^1$ it holds that

$$R_\alpha^n(x) = x + n\alpha \mod 1.$$

This follows directly from the way R_α is constructed. First assume that $\alpha = \frac{p}{q} \in \mathbb{Q}$ with p and q coprime, and let $x \in S^1$. Then x is periodic with period q if and only if

$$x + q\alpha \mod 1 = x.$$

This is the case if and only if $q\alpha \in \mathbb{Z}$. We have $q\alpha = p$, hence x is periodic with period q. Since $x \in S^1$ was chosen arbitrarily, we can conclude that every point of S^1 is periodic and has period q. Now assume that $\alpha \in \mathbb{R} \setminus \mathbb{Q}$ and assume that $x \in S^1$ is periodic. Then

$$x + n\alpha \mod 1 = x$$

for some $n \in \mathbb{Z}$ and it follows that $\alpha = \frac{p}{n}$ for some $p \in \mathbb{Z}$. However, that would mean that $\alpha \in \mathbb{Q}$ which is a contradiction with our assumption. We can conclude that x is not periodic and since $x \in S^1$ was chosen arbitrarily, no point of S^1 is periodic. $\square$

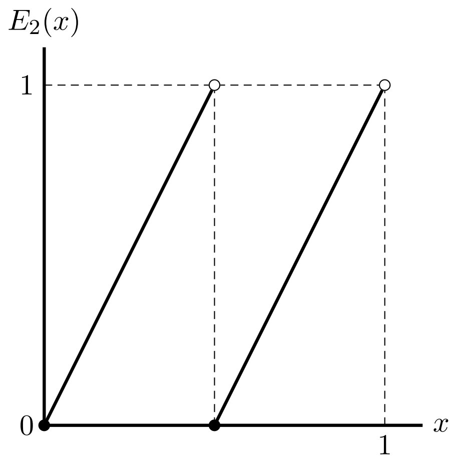

Figure 3: 2-fold of the unit circle

Aside from rotations, we can also look at a different type of operation on the unit circle. Namely, we can look at *expanding maps* $E_n : S^1 \to S^1$ of the circle for n a natural number. These maps are given by

$$E_n(x) = nx \mod 1$$

and it is shown in Figure 3 for $n = 2$. Let us take for example the expanding map E_2. The operation this map describes can be thought of as first expanding the circle by a factor two, that is to make it twice as large. Then we can fold the larger circle over in the middle, to make a sort of eight-figure consisting of two connected circles. We can flip one of the circles over to lay it on top of the other and we get a double cover of the original circle. We call this a *2-fold* of the circle, and it is shown in Figure 4.

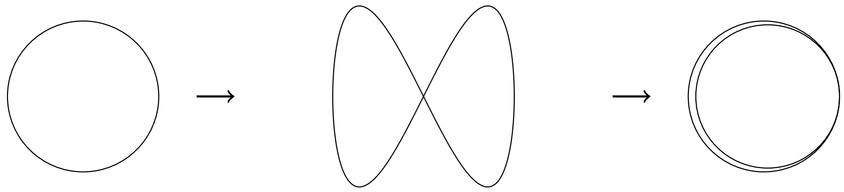

Figure 4: Representation of a 2-fold

For a higher number $n \in \mathbb{N}$ we can go through the same process, but instead of expanding the circle by a factor 2, we expand it by a factor n. Then we fold it over $n-1$ times to make n smaller connected circles and we can lay all those circles on top of each other, resulting in an n-fold cover of the original circle. As is clear from this explanation, the map E_n is more complex for bigger n. With the notion of topological entropy we discuss in Chapter 3, we can make this concrete and compare different n-folds.

3 Basic notions and properties

Topological entropy was first defined in 1965 by Adler, Konheim and McAndrew [1]. They defined topological entropy on compact topological spaces using open covers. Later, Bowen [4] and Dinaburg [7] both independently gave an equivalent definition for compact metric spaces, and a similar definition can even be used for metric spaces that are not necessarily compact. We will not consider the definition for non-compact spaces, as this goes beyond the scope of this thesis, but this concept is explained in detail in [13].

In the first section of this chapter we will define topological entropy as it was originally done by Adler, Konheim and McAndrew, and show that this definition is well-defined. In the second section of the chapter we will see the alternative definition by Bowen and Dinaburg and we will compare the two definitions. The most important result in this section is that the two definitions are equivalent on compact metric spaces. In the last section we will see some important properties of topological entropy.

3.1 Topological entropy on compact topological spaces

We define topological entropy of a continuous map f based on the entropy of an open cover $\mathcal{A}$ of X and of f with respect to the open cover $\mathcal{A}$. We need X to be compact for this because in a compact space, every open cover has a finite subcover. Later, we will also see that we need the map f to be continuous so that the preimage of f on an open cover is again an open cover. Unless specified otherwise, all definitions, lemmas and theorems in this section are based on the original article by Adler, Konheim, and McAndrew [1].

Definition 3.1.1. Let X be a compact topological space and $\mathcal{A}$ an open cover of X. We define the *entropy of* $\mathcal{A}$ as

$$H(\mathcal{A}) := \log N(\mathcal{A})$$

where $N(\mathcal{A})$ denotes the number of sets in a subcover of $\mathcal{A}$ of minimal cardinality.

A subcover is of minimal cardinality if there is no other subcover of $\mathcal{A}$ that has fewer elements. Such a subcover is said to be *minimal*. Since X is a compact space, every open cover of X has a finite subcover. Therefore $N(\mathcal{A}) < \infty$ for every open cover $\mathcal{A}$. Note that $N(\mathcal{A})$ is always a positive integer and thus $H(\mathcal{A}) \geq 0$ for every open cover $\mathcal{A}$.

Definition 3.1.2. Let X be a compact topological space. For open covers $\mathcal{A}$ and $\mathcal{B}$ of X we define their *join* as

$$\mathcal{A} \vee \mathcal{B} := \{A \cap B : A \in \mathcal{A}, B \in \mathcal{B}\}.$$

We can easily see that for any two open covers $\mathcal{A}$ and $\mathcal{B}$ of some topological space X, their join is also an open cover of X. Namely, for any $x \in X$ there exist $A \in \mathcal{A}$ and $B \in \mathcal{B}$ such that $x \in A$ and $x \in B$. Then $x \in A \cap B \in \mathcal{A} \vee \mathcal{B}$. It follows that $X \subseteq \mathcal{A} \vee \mathcal{B}$. Furthermore, the intersection of two open sets is again an open set, hence $\mathcal{A} \vee \mathcal{B}$ is an open cover of X. We will need this to prove the following lemma.

Lemma 3.1.3. *Let X be a compact topological space and let $\mathcal{A}$ and $\mathcal{B}$ be open covers of X. Then*

$$H(\mathcal{A}) \leq H(\mathcal{A} \vee \mathcal{B}).$$

Proof. Let $A \cap B \in \mathcal{A} \vee \mathcal{B}$. Note that $A \cap B \subseteq A$, hence every set in $\mathcal{A} \vee \mathcal{B}$ is a subset of some set in $\mathcal{A}$. Define $N := N(\mathcal{A} \vee \mathcal{B})$ and let $\{C_i\}_{i=1}^N$ be a minimal subcover of $\mathcal{A} \vee \mathcal{B}$. For every $i = 1, \ldots, N$ there exists $A_i \in \mathcal{A}$ such that $C_i \subseteq A_i$ and it holds trivially that the family $\{A_i\}_{i=1}^N$ is a subcover of $\mathcal{A}$. Namely

$$X = \bigcup_{i=1}^{N} C_i$$

and $\{C_i\}_{i=1}^N \subseteq \{A_i\}_{i=1}^N$. Then

$$N(\mathcal{A}) \leq N = N(\mathcal{A} \vee \mathcal{B}).$$

Since the log function is increasing, the result

$$H(\mathcal{A}) = \log N(\mathcal{A}) \leq \log N(\mathcal{A} \vee \mathcal{B}) = H(\mathcal{A} \vee \mathcal{B})$$

follows. □

The following lemmas are needed for the proof of the well-definedness of the topological entropy in Theorem 3.1.10.

Lemma 3.1.4. *Let X be a compact topological space and let $\mathcal{A}$ and $\mathcal{B}$ be open covers of X. Then $N(\mathcal{A} \vee \mathcal{B}) \leq N(\mathcal{A}) \cdot N(\mathcal{B})$ and $H(\mathcal{A} \vee \mathcal{B}) \leq H(\mathcal{A}) + H(\mathcal{B})$.*

Proof. Define $N := N(\mathcal{A})$ and $M := N(\mathcal{B})$. Let $\{A_i\}_{i=1}^N$ be a minimal subcover of $\mathcal{A}$ and let $\{B_j\}_{j=1}^M$ be a minimal subcover of $\mathcal{B}$. Then the family

$$\{A_i \cap B_j : i = 1, \ldots, N, j = 1, \ldots, M\}$$

is a subcover of $\mathcal{A} \vee \mathcal{B}$ with $N \cdot M$ elements. It follows that

$$N(\mathcal{A} \vee \mathcal{B}) \leq N \cdot M = N(\mathcal{A}) \cdot N(\mathcal{B})$$

and since the log function is increasing we can conclude that

$$H(\mathcal{A} \vee \mathcal{B}) \leq H(\mathcal{A}) + H(\mathcal{B}),$$

which was to be shown. □

Lemma 3.1.5. *Let X be a compact topological space, $f : X \to X$ be a continuous function and let $\mathcal{A}$ be an open cover of X. Then*

$$f^{-1}(\mathcal{A}) = \{f^{-1}(A) : A \in \mathcal{A}\}$$

is also an open cover of X. Moreover

$$f^{-1}(\mathcal{A} \vee \mathcal{B}) = f^{-1}(\mathcal{A}) \vee f^{-1}(\mathcal{B})$$

if $\mathcal{B}$ is an open cover of X.

Proof. First we need to show that $f^{-1}(\mathcal{A})$ is an open cover of X. Since f is continuous, it follows by definition that $f^{-1}(A)$ is open for any open set $A \in \mathcal{A}$. Furthermore, we have

$$\bigcup_{A \in \mathcal{A}} A = X$$

and

$$f^{-1}(\mathcal{A}) = \bigcup_{A \in \mathcal{A}} f^{-1}(A) = f^{-1}\left(\bigcup_{A \in \mathcal{A}} A\right) = f^{-1}(X) = X.$$

From this we can conclude that $f^{-1}(\mathcal{A})$ is an open cover of X. The second statement can be proved easily as follows: let $A \in \mathcal{A}, B \in \mathcal{B}$ be open sets. It holds that $f(x) \in A \cap B$ for $x \in X$ if and only if $f(x) \in A$ and $f(x) \in B$. Then $x \in f^{-1}(A \cap B)$ if and only if $x \in f^{-1}(A) \cap f^{-1}(B)$. Since this holds for any set $A \cap B \in \mathcal{A} \vee \mathcal{B}$, we can conclude. $\square$

As stated at the beginning of this section, we need to define the entropy of a continuous map $f : X \to X$ with respect to a fixed open cover, before we can define the topological entropy of f. By looking at each open cover separately, we can eventually say something about the behaviour of the dynamical system in the limit.

Definition 3.1.6. Let X be a compact topological space, $\mathcal{A}$ an open cover of X and $f : X \to X$ a continuous function. We define the *entropy of f with respect to the cover $\mathcal{A}$* as

$$h(f, \mathcal{A}) := \lim_{n \to \infty} \frac{H(\mathcal{A} \vee f^{-1}(\mathcal{A}) \vee \cdots \vee f^{-n+1}(\mathcal{A}))}{n}.$$

In this definition, we need $\mathcal{A} \vee f^{-1}(\mathcal{A}) \vee \cdots \vee f^{-n+1}(\mathcal{A})$ to be an open cover of X. Only then the notation $H(\mathcal{A} \vee f^{-1}(\mathcal{A}) \vee \cdots \vee f^{-n+1}(\mathcal{A}))$ makes sense. We have already seen that the join of two open covers is again an open cover, and from that it follows that any finite join of open covers is an open cover of X. We just need $f^{-k}(\mathcal{A})$ to be an open cover of X for every $k = 1, \ldots, n-1$. We saw in the previous lemma that $f^{-1}(\mathcal{A})$ is an open cover of X and the proof given in the lemma can be done exactly the same for any $k \in \mathbb{N}$, hence it follows that $f^{-k}(\mathcal{A})$ is an open cover of X for any $k = 1, \ldots, n-1$. We will show later that the entropy $h(f, \mathcal{A})$ is well-defined. For that we still need to show that the limit exists and is finite, and we need the following two lemmas to do so.

Lemma 3.1.7. *Let X be a compact topological space, $f : X \to X$ be a continuous function and let $\mathcal{A}$ be an open cover of X. Then*

$$H(f^{-1}(\mathcal{A})) \leq H(\mathcal{A}).$$

Proof. Define $N := N(\mathcal{A})$ and let $\{A_i\}_{i=1}^N$ be a minimal subcover of $\mathcal{A}$. Since $\mathcal{A}$ is an open cover of X and f is continuous, it holds that $f^{-1}(\mathcal{A})$ is an open cover of X. For the same reason $f^{-1}(\{A_i\}_{i=1}^N) = \{f^{-1}(A_i)\}_{i=1}^N$ is an open cover of X. It also holds that $\{f^{-1}(A_i)\}_{i=1}^N \subseteq f^{-1}(\mathcal{A})$, hence $\{f^{-1}(A_i)\}_{i=1}^N$ is a subcover of $f^{-1}(\mathcal{A})$. This cover does not have to be minimal, hence it follows that

$$N(f^{-1}(\mathcal{A})) \leq N(\mathcal{A}).$$

We can conclude that

$$H(f^{-1}(\mathcal{A})) \leq H(\mathcal{A})$$

by definition. $\square$

Lemma 3.1.8 (Fekete's Lemma)**.** *Let $(a_n)_{n\in\mathbb{N}}$ be a subadditive sequence of real numbers, i.e. $a_{n+m} \leq a_n + a_m$ for all $n, m \in \mathbb{N}$. Then*

$$\lim_{n\to\infty} \frac{a_n}{n} = \inf_{n\in\mathbb{N}} \frac{a_n}{n},$$

hence the sequence $\frac{a_n}{n}$ converges to its infimum.

We will not proof Fekete's Lemma here, but the statement including a proof can be found in [12, Lemma 1.2.1]. Now we have everything we need to define topological entropy on a compact topological space, and to show that this notion is well-defined.

Definition 3.1.9. Let X be a compact topological space. The *topological entropy of a continuous map* $f : X \to X$ is defined as

$$h(f) := \sup_{\mathcal{A}} h(f, \mathcal{A})$$

where $\mathcal{A}$ is an open cover of X.

Theorem 3.1.10 (Well-definedness of topological entropy)**.** *Topological entropy of Definition 3.1.9 is well-defined.*

Proof. For any $n \in \mathbb{N}$ and $\mathcal{A}$ an open cover of X, define

$$H_n(\mathcal{A}) := H(\mathcal{A} \vee \dots f^{-1}(\mathcal{A}) \vee \dots \vee f^{-n+1}(\mathcal{A})).$$

We need to show that the limit

$$h(f, \mathcal{A}) = \lim_{n\to\infty} \frac{H_n(\mathcal{A})}{n}$$

exists and is finite. Then we can take the supremum over all $h(f, \mathcal{A})$ because each $h(f, \mathcal{A})$ is a real number. The supremum will either be a real number as well, or it will be infinite. Let $\mathcal{A}$ be an open cover of X and let $n, m \in \mathbb{N}$. Then

$$\begin{aligned}
H_{m+n}(\mathcal{A}) &= H(\mathcal{A} \vee \dots \vee f^{-m-n+1}(\mathcal{A})) \\
&= H(\mathcal{A} \vee \dots \vee f^{-m+1}(\mathcal{A}) \vee f^{-m}(\mathcal{A} \vee \dots \vee f^{-n+1}(\mathcal{A}))) \\
&\leq H(\mathcal{A} \vee \dots \vee f^{-m+1}(\mathcal{A})) + H(f^{-m}(\mathcal{A} \vee \dots \vee f^{-n+1}(\mathcal{A}))) \\
&\leq H(\mathcal{A} \vee \dots \vee f^{-m+1}(\mathcal{A})) + H(\mathcal{A} \vee \dots \vee f^{-n+1}(\mathcal{A})) = H_m(\mathcal{A}) + H_n(\mathcal{A}).
\end{aligned}$$

The first equation follows from Lemma 3.1.5, the second equation follows from Lemma 3.1.4, and the last inequality from Lemma 3.1.7. Now that we know that $H_{m+n}(\mathcal{A}) \leq H_m(\mathcal{A}) + H_n(\mathcal{A})$, it follows directly from Fekete's Lemma that

$$lim_{n\to\infty} \frac{H_n(\mathcal{A})}{n} = \inf_{n\in\mathbb{N}} \frac{H_n(\mathcal{A})}{n}.$$

Note that this cannot be $-\infty$, since we have seen that $H_n(\mathcal{A}) \geq 0$ for all $n \in \mathbb{N}$. Therefore, the limit exists and is finite, and we can conclude that topological entropy is well-defined on compact topological spaces. $\square$

3.2 Topological entropy on metric spaces

In this section we consider topological entropy on a compact metric space (X, d). The first part of this section is based on [3]. We can define the entropy using a sequence of metrics $(d_n)_{n\in\mathbb{N}}$ that are as follows: let $f : X \to X$ be a continuous function and define

$$d_n(x, y) := \max \left\{ d\left(f^k(x), f^k(y)\right) : k = 0, \ldots, n-1 \right\}$$

for each $n \in \mathbb{N}$ and $x, y \in X$. We can interpret this as taking the largest distance between the orbits of x and y in the first n time-steps.

Proposition 3.2.1. *Let (X, d) be a compact metric space, $f : X \to X$ a continuous function and $n \in \mathbb{N}$. Then d_n is a distance on X.*

Proof. To be able to conclude that d_n is a distance on X, we need to show that it is definite, symmetric and that the triangle inequality holds. We start by proving definiteness. Let $x, y \in X$. We know that d is a distance on X, hence

$$d\left(f^k(x), f^k(y)\right) \geq 0$$

for every $k = 0, \ldots, n-1$. Therefore $d_n(x, y) \geq 0$. Assume that $x = y$. Then $f^k(x) = f^k(y)$ and

$$d\left(f^k(x), f^k(y)\right) = 0$$

for all $k = 0, \ldots, n-1$. It follows that $d_n(x, y) = 0$. Now assume that $d_n(x, y) = 0$. Then

$$\max \left\{ d\left(f^k(x), f^k(y)\right) \right\} = 0$$

and since $d(x, y) \geq 0$ for all $x, y \in X$ it follows that

$$d\left(f^k(x), f^k(y)\right) = 0,$$

which implies that $f^k(x) = f^k(y)$ for all $k = 0, \ldots, n-1$. We can conclude that

$$x = f^0(x) = f^0(y) = y$$

which concludes definiteness. Now we will prove symmetry. Let $x, y \in X$. We can again use that d is a distance on X, hence symmetry holds for d. It then holds that

$$\begin{aligned} d_n(x, y) &= \max \left\{ d\left(f^k(x), f^k(y)\right) : k = 0, \ldots, n-1 \right\} \\ &= \max \left\{ d\left(f^k(y), f^k(x)\right) : k = 0, \ldots, n-1 \right\} = d_n(y, x). \end{aligned}$$

It follows that d_n is symmetric. Lastly, we need to show that the triangle inequality holds. Let $x, y, z \in X$. We have

$$\begin{aligned}
d_n(x,y) &= \max\left\{d\left(f^k(x), f^k(y)\right) : k = 0, \ldots, n-1\right\} \\
&\leq \max\left\{d\left(f^k(x), f^k(z)\right) + d\left(f^k(z), f^k(y)\right) : k = 0, \ldots, n-1\right\} \\
&\leq \max\left\{d\left(f^k(x), f^k(z)\right) : k = 0, \ldots, n-1\right\} + \max\left\{d\left(f^k(z), f^k(y)\right) : k = 0, \ldots, n-1\right\} \\
&= d_n(x,z) + d_n(z,y).
\end{aligned}$$

The first inequality follows from the triangle inequality for the distance d. We can conclude that the triangle inequality holds for d_n and therefore d_n is a distance on X for every $n \in \mathbb{N}$. $\square$

Topological entropy informs us about the complexity of the orbits of points in X. In a metric space, we always have a distance, so we can look at the points of X with a certain resolution. For $\varepsilon > 0$, we do not distinguish between points $x, y \in X$ with $d(x,y) < \varepsilon$. Note that we cannot look at each point of X and it's orbit separately because the number of orbits will then almost always be infinite. We can make the concept of a resolution more precise by looking at (n,ε)-separated sets.

Definition 3.2.2. Let (X,d) be a compact metric space and $f : X \to X$ a continuous function. Let $n \in \mathbb{N}$ a natural number and fix $\varepsilon > 0$. A subset $E \subseteq X$ is said to be *(n,ε)-separated with respect to f* if $d_n(x,y) > \varepsilon$ for any $x, y \in E$ such that $x \neq y$. We denote by $N(n,\varepsilon)$ the largest cardinality of an (n,ε)-separated subset of X with respect to f, for a fixed $n \in \mathbb{N}$ and $\varepsilon > 0$.

We can for example look at an (n,ε)-separated subset of X for $n = 1$ and some $\varepsilon > 0$. Then $d_n(x,y) = d(x,y)$ for every $x, y \in X$. In Figure 5 such an example is given. Here we can see that the points x_1, x_2 and x_3 form a $(1,\varepsilon)$-separated set. Namely

$$d_1(x_i, x_j) = d(x_i, x_j) > \varepsilon$$

for $i, j = 1, 2, 3$ and $i \neq j$. The set $\{x_1, x_2, x_3, x_4\}$ is not a $(1,\varepsilon)$-separated set however. Even though $d_1(x_4, x_2) > \varepsilon$ and $d_1(x_4, x_3) > \varepsilon$ we can see that $d_1(x_4, x_1) < \varepsilon$. Therefore the requirement for a $(1,\varepsilon)$-separated set is not met.

The notion $N(n,\varepsilon)$ can be thought of as taking the maximum number of points $x_1, \ldots, x_m$ in X such that

$$d_n(x_i, x_j) > \varepsilon$$

for all $i, j = 1, \ldots, m$ with $i \neq j$. These points are then the points of X that we can distinguish from one another, and we look at the orbits of these points. For consistency, and to make the notation look more clean, we will define

$$H(n,\varepsilon) := \log N(n,\varepsilon)$$

and use this notation in Definition 3.2.4 of topological entropy.

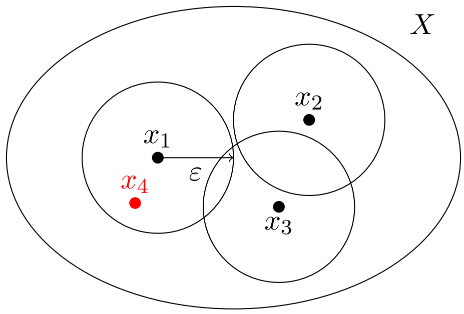

Figure 5: $\{x_1, x_2, x_3\} \subseteq X$ is a $(1, \varepsilon)$-separated set

Lemma 3.2.3. *For any natural number n and $\varepsilon > 0$, the cardinality $N(n, \varepsilon)$ is finite.*

Proof. We take a cover of X, consisting of open balls with radius $\frac{\varepsilon}{2}$. Since X is compact, there exists a finite subcover $B_1, \ldots, B_m$. Let E be an (n, ε)-separated set. Then any two points of E can never be in the same open ball B_i, $i \in \{1, \ldots, m\}$. It follows that $N(n, \varepsilon) \leq m$ and therefore it is finite. $\square$

Definition 3.2.4. Let (X, d) be a compact metric space and let $f : X \to X$ be a continuous function. The *topological entropy of f* is defined by

$$\tilde{h}(f) := \lim_{\varepsilon \to 0} \limsup_{n \to \infty} \frac{H(n, \varepsilon)}{n}.$$

Remark 3.2.5. To avoid confusion, I have chosen to define the topological entropy in terms of (n, ε)-separated sets. It can however also be defined using (n, ε)-spanning sets or coverings of the set X by sets with a d_n-diameter that is less than ε. All three versions of the definition can be found in [5].

The proof that this notion is well-defined can be found in section 2.5 of [5]. We will not prove it separately here but we show that on compact metric spaces the definition of Adler, Konheim and McAndrew and the definition of Bowen and Dinaburg are equivalent, and therefore $\tilde{h}$ is well-defined by Theorem 3.1.10. We will be proving the equivalence below in several steps, which are based on the proof in [13], specifically in Chapter 7. First, we need to introduce (n, ε)-spanning sets.

Definition 3.2.6. Let (X, d) be a compact metric space and $f : X \to X$ a continuous function. Let $n \in \mathbb{N}$ and fix $\varepsilon > 0$. A subset $F \subseteq X$ is said to be *(n, ε)-spanning with respect to f* if for any $x \in X$, there exists a $y \in F$ with $d_n(x, y) \leq \varepsilon$. We denote the minimal cardinality of any (n, ε)-spanning set over X with respect to f by $M(n, \varepsilon)$.

Let us consider a $(1, \varepsilon)$-spanning set in X for some $\varepsilon > 0$. Note that by definition a set $F \subseteq X$ is (n, ε)-spanning if and only if the set of ε-balls around each point of F form an open cover of X. This can also be seen in Figure 6. In this figure, the set $\{x_1, \ldots, x_6\}$ form a $(1, \varepsilon)$-spanning subset of X. Namely, the ε-balls around the points $x_1, \ldots, x_6$

cover X and therefore it holds that for any $x \in X$, there exists an $i \in \{1, \dots, 6\}$ such that

$$d_n(x, x_i) = d_1(x, x_i) \leq \varepsilon.$$

Of course, this notion still holds when we consider the set $\{x_1, \dots, x_6, x_7\}$, hence $\{x_1, \dots, x_7\}$ is also a $(1, \varepsilon)$-spanning set. A trivial (n, ε)-spanning set of any compact metric space (X, d) is actually X itself.

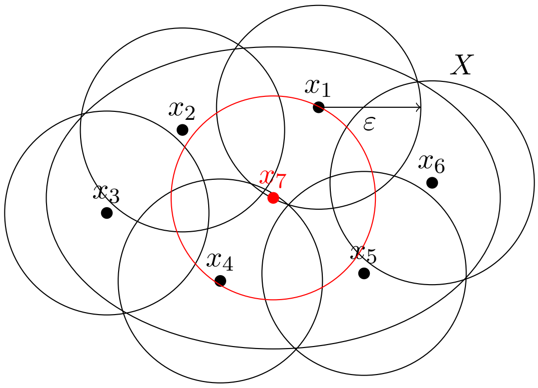

Figure 6: A $(1, \varepsilon)$-spanning set in X

Since X is compact and therefore bounded, we can find a finite amount of points $x_1, \dots, x_m \in X$ such that the set X is covered by balls of radius ε around each point $x_1, \dots, x_m$. Thus for any $n \in \mathbb{N}$ and $\varepsilon > 0$ it holds that $M(n, \varepsilon)$ is a finite number.

Note that any (n, ε)-separated set of X with respect to f of maximum cardinality is also an (n, ε)-spanning set. Namely, let $E = \{x_1, \dots, x_{N(n,\varepsilon)}\} \subseteq X$ be such a set of cardinality $N(n, \varepsilon)$. Then

$$d_n(x_i, x_j) > \varepsilon$$

for all $i, j = 1, \dots, N(n, \varepsilon)$ with $i \neq j$. Take $x \in X$ and assume that

$$d_n(x, x_i) > \varepsilon$$

for all $i = 1, \dots, N(n, \varepsilon)$. This cannot be because then $E \cup \{x\}$ would also be an (n, ε)-separated set and that would imply that E is not of maximal cardinality. It follows that

$$d_n(x, x_i) \leq \varepsilon$$

for all $i = 1, \dots, N(n, \varepsilon)$ and we can conclude that E is an (n, ε)-spanning set.

Proposition 3.2.7. *Let (X, d) be a compact metric space and let $f : X \to X$ be a continuous function. Fix $n \in \mathbb{N}$ and $\varepsilon > 0$. Then the following hold:*

(i) $M(n, \varepsilon) \leq N(n, \varepsilon)$

(ii) For the open cover $\mathcal{A}_{2\varepsilon}$ of X consisting of all open balls of radius 2ε, it holds that

$$N\left(\bigvee_{i=0}^{n-1} f^{-i}\left(\mathcal{A}_{2\varepsilon}\right)\right) \leq M(n,\varepsilon).$$

(iii) For $\mathcal{A}_{\frac{\varepsilon}{2}}$ an open cover of X consisting of balls with radius $\frac{\varepsilon}{2}$, it holds

$$N(n,\varepsilon) \leq N\left(\bigvee_{i=0}^{n-1} f^{-i}\left(\mathcal{A}_{\frac{\varepsilon}{2}}\right)\right).$$

Proof. (i) This follows directly from the fact that any (n,ε)-separated set of maximum cardinality $N(n,\varepsilon)$ is also an (n,ε)-spanning set. Since $M(n,\varepsilon)$ is the minimal cardinality of an (n,ε)-spanning set, we have $M(n,\varepsilon) \leq N(n,\varepsilon)$.

(ii) Note that

$$\bigvee_{i=0}^{n-1} f^{-i}(\mathcal{A}_{2\varepsilon}) = \bigcup_{x\in X} \bigcap_{i=0}^{n-1} f^{-i}\left(B\left(f^{-i}(x), 2\varepsilon\right)\right)$$

Let F be an (n,ε)-spanning set of cardinality $M(n,\varepsilon)$. Since for every $x \in X$, there exists a $y \in F$ such that

$$\max\left\{d\left(f^k(x), f^k(y)\right) : k = 0,\ldots,n-1\right\} \leq \varepsilon,$$

it holds that

$$X = \bigcup_{y\in F} \bigcap_{i=0}^{n-1} f^{-i}\left(B\left(f^{-i}(y), \varepsilon\right)\right).$$

Each ball $B\left(f^{-i}(y), \varepsilon\right)$ is a subset of a member of $\mathcal{A}_{2\varepsilon}$, because the balls in $\mathcal{A}_{2\varepsilon}$ have radius 2ε. It follows that

$$X = \bigcup_{y\in F} \bigcap_{i=0}^{n-1} f^{-i}\left(B\left(f^{-i}(y), 2\varepsilon\right)\right) \subseteq \bigvee_{i=0}^{n-1} f^{-i}(\mathcal{A}_{2\varepsilon})$$

which implies that

$$N\left(\bigvee_{i=0}^{n-1} f^{-i}\left(\mathcal{A}_{2\varepsilon}\right)\right) \leq M(n,\varepsilon).$$

(iii) Let $E \subseteq X$ be an (n,ε)-separated set of cardinality $N(n,\varepsilon)$. For $x, y \in E$ with $x \neq y$ it holds that $d_n(x,y) > \varepsilon$, hence

$$\max\left\{d\left(f^k(x), f^k(y)\right) : k = 0,\ldots,n-1\right\} > \varepsilon.$$

For $x \in E$, the set

$$\bigcap_{i=0}^{n-1} f^{-i}\left(B\left(f^{-i}(x), \frac{\varepsilon}{2}\right)\right)$$

contains no other point of E because there exists an $i \in \{0, \dots, n-1\}$ such that $d\left(f^{-i}(x), f^{-i}(y)\right) > \varepsilon$. Then it follows that no member the cover

$$\bigvee_{i=0}^{n-1} f^{-i}(\mathcal{A}_{\frac{\varepsilon}{2}})$$

can contain more than one point of E. A minimal subcover of this cover therefore contains at least $N(n,\varepsilon)$ sets and we can conclude that

$$N(n,\varepsilon) \leq N\left(\bigvee_{i=0}^{n-1} f^{-i}\left(\mathcal{A}_{\frac{\varepsilon}{2}}\right)\right).$$

□

Corollary 3.2.8. *Let (X,d) be a compact metric space and $f : X \to X$ a continuous function, and $\varepsilon > 0$ fixed. Let $\mathcal{A}_{2\varepsilon}$ be a cover of X consisting of all open balls with radius 2ε and let $\mathcal{A}_{\frac{\varepsilon}{2}}$ be the cover of X consisting of open balls with radius $\frac{\varepsilon}{2}$. Then*

$$h(f, \mathcal{A}_{2\varepsilon}) \leq \limsup_{n\to\infty} \frac{H(n,\varepsilon)}{n} \leq h(f, \mathcal{A}_{\frac{\varepsilon}{2}}).$$

Proof. From 3.2.7 we can conclude that

$$N\left(\bigvee_{i=0}^{n-1} f^{-i}(\mathcal{A}_{2\varepsilon})\right) \leq N(n,\varepsilon) \leq N\left(\bigvee_{i=0}^{n-1} f^{-i}(\mathcal{A}_{\frac{\varepsilon}{2}})\right).$$

The log function is increasing, hence

$$\log N\left(\bigvee_{i=0}^{n-1} f^{-i}(\mathcal{A}_{2\varepsilon})\right) \leq \log N(n,\varepsilon) \leq \log N\left(\bigvee_{i=0}^{n-1} f^{-i}(\mathcal{A}_{\frac{\varepsilon}{2}})\right)$$

which also implies that

$$\frac{H\left(\bigvee_{i=0}^{n-1} f^{-i}(\mathcal{A}_{2\varepsilon})\right)}{n} \leq \frac{H(n,\varepsilon)}{n} \leq \frac{H\left(\bigvee_{i=0}^{n-1} f^{-i}(\mathcal{A}_{\frac{\varepsilon}{2}})\right)}{n}.$$

This holds for arbitrary $n \in \mathbb{N}$, hence

$$\limsup_{n\to\infty} \frac{H\left(\bigvee_{i=0}^{n-1} f^{-i}(\mathcal{A}_{2\varepsilon})\right)}{n} \leq \limsup_{n\to\infty} \frac{H(n,\varepsilon)}{n} \leq \limsup_{n\to\infty} \frac{H\left(\bigvee_{i=0}^{n-1} f^{-i}(\mathcal{A}_{\frac{\varepsilon}{2}})\right)}{n}.$$

We know for the left and right-hand side of the equation that the limit exists, so we can substitute the lim sup by the limit, and the result follows. □

To be able to prove that $h(f, \mathcal{A}_n)$ goes to $h(f)$ in the limit when the diameter of the cover $\mathcal{A}_n$ becomes arbitrarily small, we need to understand the concept of Lebesgue numbers. A Lebesgue number can be attributed to any open cover of a compact metric space (X, d). It is a positive number, usually denoted by $\delta > 0$, such that every subset $A \subseteq X$ with $\text{diam}(A) < \delta$ is contained in a member of the open cover. Naturally, a Lebesgue number is not unique to the cover, but every open cover of a compact metric space does admit a Lebesgue number. This result is known as the Lebesgue Number Lemma, which I will not state formally or prove here, but it is stated and proved in [10].

Lemma 3.2.9. *Let (X, d) be a compact metric space and $f : X \to X$ a continuous function. Let $(\mathcal{A}_n)_{n\in\mathbb{N}}$ be a sequence of open covers of X such that diam$(\mathcal{A}_n) \to 0$ as $n \to \infty$. That is, the maximum diameter of all sets in $\mathcal{A}_n$ becomes arbitrarily small as n approaches infinity. Then*

$$\lim_{n\to\infty} h(f, \mathcal{A}_n) = h(f)$$

if $h(f) < \infty$, and

$$\lim_{n\to\infty} h(f, \mathcal{A}_n) = \infty$$

if $h(f) = \infty$.

Proof. First, we assume that $h(f) < \infty$. Fix $\varepsilon > 0$ and let $\mathcal{A}$ be an open cover of X such that

$$h(f, \mathcal{A}) > h(f) - \varepsilon.$$

Let $\delta_1 > 0$ be a Lebesgue number for $\mathcal{A}$. Since $\text{diam}(\mathcal{A}_n) \to 0$ as $n \to \infty$ we can choose an $N \in \mathbb{N}$ such that $\text{diam}(\mathcal{A}_n) < \delta_1$ for every $n \geq N$. Then for $n \geq N$ every subset of X that is a member of $\mathcal{A}_n$ is contained in a member of $\mathcal{A}$. By definition it follows that

$$h(f) \geq h(f, \mathcal{A}_n) \geq h(f, \mathcal{A}) > h(f) - \varepsilon$$

for any $n \geq N$. Taking the limit $n \to \infty$ then gives us

$$h(f) \geq \lim_{n\to\infty} h(f, \mathcal{A}_n) > h(f) - \varepsilon.$$

Because $\varepsilon > 0$ was chosen arbitrarily, we can conclude that

$$h(f) = \lim_{n\to\infty} h(f, \mathcal{A}_n).$$

Now assume that $h(f) = \infty$. Let $a > 0$ be a real number and let $\mathcal{B}$ be an open cover of X such that

$$h(f, \mathcal{B}) > a.$$

Let $\delta_2 > 0$ be a Lebesgue number for $\mathcal{B}$ and choose $M \in \mathbb{N}$ such that $\text{diam}(\mathcal{A}_n) < \delta_2$ for all $n \geq M$. Then, by the same argument as before, we have

$$\infty = h(f) \geq h(f, \mathcal{A}_n) \geq h(f, \mathcal{B}) > a$$

and it follows that

$$\lim_{n\to\infty} h(f, \mathcal{A}_n) > a.$$

Since a can be arbitrarily large it holds that

$$\lim_{n\to\infty} h(f, \mathcal{A}_n) = \infty$$

which concludes our proof. $\square$

Theorem 3.2.10. *Let (X, d) be a compact metric space and $f : X \to X$ a continuous function. Then $h(f) = \tilde{h}(f)$, hence the Definition 3.1.9 and Definition 3.2.4 are equivalent.*

Proof. Let $\varepsilon = \frac{1}{n}$ for some $n \in \mathbb{N}$ and let $\mathcal{A}_{2\varepsilon}$ and $\mathcal{A}_{\frac{\varepsilon}{2}}$ be as in Corollary 3.2.8. Then we know that

$$h(f, \mathcal{A}_{2\varepsilon}) \leq \limsup_{n\to\infty} \frac{H(n,\varepsilon)}{n} \leq h(f, \mathcal{A}_{\frac{\varepsilon}{2}}).$$

By letting $n \to \infty$ we get $\varepsilon \to 0$ and we can take this limit to obtain

$$\lim_{\varepsilon\to 0} h(f, \mathcal{A}_{2\varepsilon}) \leq \lim_{\varepsilon\to 0} \limsup_{n\to\infty} \frac{H(n,\varepsilon)}{n} \leq \lim_{\varepsilon\to 0} h(f, \mathcal{A}_{\frac{\varepsilon}{2}}).$$

Since the sequences of open covers $\mathcal{A}_{2\varepsilon}$ and $\mathcal{A}_{\frac{\varepsilon}{2}}$ both have diameters going to zero as $n \to \infty$, it follows by Lemma 3.2.9 that

$$h(f) \leq \tilde{h}(f) \leq h(f).$$

We can conclude directly that $h(f) = \tilde{h}(f)$. $\square$

In a lot of cases involving compact metric spaces, it is easier to explicitly calculate the topological entropy of a continuous map $f : X \to X$ by using the distance d_n, than it is to use open covers of the space. The above result is therefore very useful when calculating topological entropy.

3.3 Properties of the topological entropy

In this section we will look at some important and interesting properties of topological entropy. The first property we will look at states that topological entropy is invariant, meaning that two topologically conjugate maps have the same topological entropy. It is a useful property when calculating the topological entropy of a map f on a set X is relatively difficult, but there exists a topologically conjugate map $g : Y \to Y$ whose topological entropy is easier to calculate. In that case we can calculate the entropy for the easier map to find the entropy of the original map. We will also look at two properties that are useful for calculation of the topological entropy.

Definition 3.3.1. Let X and Y be topological spaces and let $f : X \to X$ and $g : Y \to Y$. Then f and g are *topologically conjugate* if there exists a homeomorphism $H : X \to Y$ such that

$$H \circ f = g \circ H.$$

The map H is called a *topological conjugacy.*

Theorem 3.3.2. *[3, Theorem 3.3] Let (X, d_1) and (Y, d_2) be two metric spaces and let $f : X \to X$ and $g : Y \to Y$ be continuous functions. If f and g are topologically conjugate, then $h(f) = h(g)$.*

Proof. Assume f and g are topologically conjugate with $H : X \to Y$ a topological conjugacy between f and g. Since H is a homeomorphism on a compact set, it is continuous on a compact set and therefore H is uniformly continuous. Fix $\varepsilon > 0$. Then there exists $\delta > 0$ such that

$$d_2(H(x), H(y)) < \varepsilon$$

for all $x, y \in X$ with $d_1(x, y) < \delta$. Since H is a topological conjugacy it holds by definition that

$$H \circ f = g \circ H$$

and it follows that

$$\begin{aligned} H \circ f \circ f &= g \circ H \circ f \\ H \circ f^2 &= g \circ g \circ H \\ H \circ f^2 &= g^2 \circ H. \end{aligned}$$

Then we can easily prove by induction that

$$H \circ f^k = g^k \circ H$$

for any $k \in \mathbb{N}$. Assume that $p_1, \ldots, p_k \in X$ are such that

$$d_n(H(p_i), H(p_j)) = \max\{d_2(g^k(H(p_i)), g^k(H(p_j))) : k = 0, \ldots, n-1\} \geq \varepsilon$$

for $i \neq j$ and fixed $n \in \mathbb{N}$. Then

$$d_n(p_i, p_j) = \max\{d_1(f^k(p_i), f^k(p_j)) : k = 0, \ldots, n-1\} \geq \delta$$

for $i \neq j$ because H is uniformly continuous. It follows that

$$N_f(n, \delta) \geq N_g(n, \varepsilon)$$

where $N_f(n, \delta)$ and $N_g(n, \varepsilon)$ denote the largest cardinality of an (n, δ)-separated subset of X with respect to f and an (n, ε)-separated subset of Y with respect to g, respectively. We can now conclude that

$$\limsup_{n\to\infty} \frac{\log N_f(n, \delta)}{n} \geq \limsup_{n\to\infty} \frac{\log N_g(n, \varepsilon)}{n}$$

because the log function is increasing. If we send $\varepsilon \to 0$, then $\delta \to 0$ and we find $h(f) \geq h(g)$. To prove the opposite inequality we rewrite

$$\begin{gathered} H \circ f = g \circ H \\ H^{-1} \circ H \circ f \circ H^{-1} = H^{-1} \circ g \circ H \circ H^{-1} \\ f \circ H^{-1} = H^{-1} \circ g \end{gathered}$$

and we can repeat the process above, substituting H for H^{-1}. We then find $h(g) \geq h(f)$ and we can conclude that $h(f) = h(g)$. □

The other implication of this theorem does not always hold true. Namely, when $h(f) = h(g)$, it does not mean that f and g are also topologically conjugate. This fact is demonstrated in Example 4.0.2 and Example 4.0.3. In the first example, we will see that the entropy of the expanding map $E_2 : S^1 \to S^1$ is $\log 2$ and in the second example we find that the entropy of the shifting map on sequences with binary elements is also $\log 2$. However, the circle and the space X from the second example are not homeomorphic. This is because the circle is a connected set and X is not (in fact, X is totally disconnected), and therefore there does not exist a homeomorphism $S^1 \to X$ and the expanding map and the shifting map cannot be topologically conjugate.

Theorem 3.3.3. *[1, Theorem 2] Let (X, d) be a compact metric space and let $f : X \to X$ be a continuous function. Then*

$$h(f^m) = m \cdot h(f)$$

for any $m \in \mathbb{N}$.

Proof. We know by Theorem 3.2.10 that h is equivalent with h. Therefore it suffices to prove that

$$h(f^m) = m \cdot h(f)$$

for any $m \in \mathbb{N}$. Fix $m \in \mathbb{N}$ and let $\mathcal{A}$ be an open cover of X. Since $\mathcal{A} \vee f^{-1}(\mathcal{A}) \vee \cdots \vee f^{-n+1}(\mathcal{A})$ is also an open cover of X, it holds by definition that

$$h(f^m) \geq h\left(f^m, \bigvee_{i=0}^{m-1} f^{-i}(\mathcal{A})\right).$$

By applying Lemma 3.1.5 we obtain

$$\begin{aligned} h(f^m) &\geq \lim_{n\to\infty} \frac{H\left(\bigvee_{i=0}^{m-1} f^{-i}(\mathcal{A} \vee f^{-1}\left(\bigvee_{i=0}^{m-1} f^{-i}(\mathcal{A}\right) \vee \cdots \vee f^{-n+1}\left(\bigvee_{i=0}^{m-1} f^{-i}(\mathcal{A}\right)\right)}{n} \\ &= \lim_{n\to\infty} \frac{H(\mathcal{A} \vee f^{-1}\left(\mathcal{A}\right) \vee \cdots \vee f^{-n-m+2}(\mathcal{A}))}{n} \\ &= m \lim_{n\to\infty} \frac{H(\mathcal{A} \vee f^{-1}\left(\mathcal{A}\right) \vee \cdots \vee f^{-n-m+2}(\mathcal{A}))}{mn} = m \cdot h(f, \mathcal{A}). \end{aligned}$$

Since $\mathcal{A}$ is arbitrary, we can conclude that $h(f^m) \geq m \cdot h(f)$. For $\mathcal{A}$ an open cover of X consider the open covers

$$\mathcal{A} \vee (f^m)^{-1}(\mathcal{A}) \vee \cdots \vee (f^m)^{-n+1}(\mathcal{A}) \qquad \text{and} \qquad \mathcal{A} \vee f^{-1}(\mathcal{A}) \vee \cdots \vee f^{-mn+1}(\mathcal{A}).$$

Note that every set on the left-hand side is also a set on the right-hand side, hence the covers are of the form $\mathcal{C}$ and $\mathcal{C} \vee \mathcal{D}$ respectively, where $\mathcal{C}$ and $\mathcal{D}$ are open covers of X. By using Lemma 3.1.3 we obtain

$$\begin{aligned} H\left(\mathcal{A} \vee (f^m)^{-1}(\mathcal{A}) \vee \cdots \vee (f^m)^{-n+1}(\mathcal{A})\right) &\leq H\left(\mathcal{A} \vee f^{-1}(\mathcal{A}) \vee \cdots \vee f^{-mn+1}(\mathcal{A})\right) \\ \lim_{n\to\infty} \frac{H\left(\mathcal{A} \vee (f^m)^{-1}(\mathcal{A}) \vee \cdots \vee (f^m)^{-n+1}(\mathcal{A})\right)}{mn} &\leq \lim_{n\to\infty} \frac{H\left(\mathcal{A} \vee f^{-1}(\mathcal{A}) \vee \cdots \vee f^{-mn+1}(\mathcal{A})\right)}{mn} \\ \frac{h(f^m, \mathcal{A})}{m} &\leq h(f, \mathcal{A}). \end{aligned}$$

It follows that $h(f^m) \leq m \cdot h(f)$ and we can conclude that $h(f^m) = m \cdot h(f)$. ☐

Theorem 3.3.4. *[5, Proposition 2.5.5][2, Theorem 4.6] Let (X, d) be a compact metric space and let $f : X \to X$ be a continuous function. If f is invertible, then $h(f^{-1}) = h(f)$.*

Proof. For $x, y \in X$ we have

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

Note that the left-hand side is $d_n(x, y)$ with respect to f, which we will denote by $d_{n,f}(x, y)$. Similarly, the right-hand side is $d_n(f^{n-1}(x), f^{n-1}(y))$ with respect to f^{-1}, which we will denote by $d_{n,f^{-1}}(f^{n-1}(x), f^{n-1}(y))$. Let $x_1, \ldots, x_k \in X$ be such that

$$d_{n,f}(x_i, x_j) > \varepsilon$$

for all $i \neq j$. Then the points $x_1, \ldots, x_k$ form an (n, ε)-separated set with respect to f and

$$d_{n,f^{-1}}(f^{n-1}(x_i), f^{n-1}(x_j)) > \varepsilon$$

meaning that the points $f^{n-1}(x_1), \ldots, f^{n-1}(x_k)$ also form an (n, ε)-separated set with respect to f^{-1}. It follows that

$$N_f(n, \varepsilon) \leq N_{f^{-1}}(n, \varepsilon).$$

Now let $f^{n-1}(y_1), \ldots, f^{n-1}(y_k)$ be points in X such that

$$d_{n,f^{-1}}(f^{n-1}(x_i), f^{n-1}(x_j)) > \varepsilon$$

for all $i \neq j$. Then

$$d_{n,f}(x_i, x_j) > \varepsilon$$

and it follows that

$$N_{f^{-1}}(n, \varepsilon) \leq N_f(n, \varepsilon).$$

We can conclude equality and therefore $h(f^{-1}) = h(f)$. ☐

By combining the results of Theorem 3.3.3 and of Theorem 3.3.4, we find the following result.

Corollary 3.3.5. *For a continuous map $f : X \to X$ on a compact metric space (X, d) it holds that*

$$h(f^m) = |m| \cdot h(f)$$

for any $m \in \mathbb{Z}$.

Remark 3.3.6. All results mentioned in this chapter also hold for compact topological spaces, not only for compact metric spaces.

4 Examples

In Chapter 2 we have already seen two examples of circle maps, the rotation of the circle and expanding maps. For both these types of maps it is possible to calculate the topological entropy using Definition 3.2.4. First, we look at the rotation map.

Example 4.0.1 (Rotation map)**.** First, fix $\alpha \in \mathbb{R}$ and $\varepsilon > 0$. Let $x, y \in S^1$ be two points on the unit circle. Since R_α is an isometry by Proposition 2.2.1, it holds that

$$d_n(x,y) = \max\left\{d\left(R_\alpha^k(x), R_\alpha^k(y)\right) : k = 0, \ldots, n-1\right\} = d(x,y).$$

Namely

$$d\left(R_\alpha^k(x), R_\alpha^k(y)\right) = d(x,y)$$

for every $k = 0, \ldots, n-1$, hence the maximum over all these distances is also equal to $d(x,y)$. It follows that $N(n,\varepsilon) = N(1,\varepsilon)$ for every $n \in \mathbb{N}$, and therefore $H(n,\varepsilon)$ is a constant. We then know that

$$\limsup_{n\to\infty} \frac{H(n,\varepsilon)}{n} = 0$$

and since $\varepsilon > 0$ was chosen arbitrarily, we can conclude that

$$h(R_\alpha) = \lim_{\varepsilon\to 0} \limsup_{n\to\infty} \frac{H(n,\varepsilon)}{n} = 0.$$

The other map on S^1 that we saw was the expanding map E_n. For this map we can also calculate the topological entropy. This calculation can be found in [3, Example 3.12] for $n = 2$ and in [2, Example 4.13] for n any natural number, but we will also work out both cases below.

Example 4.0.2 (Expanding map)**.** Consider the expanding map $E_2 : S^1 \to S^1$ and a sequence $(\varepsilon_k)_{k\in\mathbb{N}}$ with $\varepsilon_k = 2^{-(k+1)}$. Then $\varepsilon_k \to 0$ as $k \to \infty$, hence we can write

$$h(E_2) = \lim_{k\to\infty} \limsup_{n\to\infty} \frac{H(n,\varepsilon_k)}{n}.$$

To be able to calculate this limit, we want to show that $N(n,\varepsilon_k) = 2^{n+k}$ for $n \in \mathbb{N}$. Fix $n, k \in \mathbb{N}$. Take $x, y \in S^1$ such that $d(x,y) < 2^{-n}$. Remember that we use the distance on the circle that is induced by the natural distance on $[0,1]$. Then

$$\begin{aligned} d_n(x,y) &= \max\{d\left(E_2^k(x), E_2^k(y)\right) : k = 0, \ldots, n-1\} \\ &= d\left(E_2^{n-1}(x), E_2^{n-1}(y)\right) = 2^{n-1} d(x,y). \end{aligned} \tag{4.1}$$

Of course, the maximum distance between $E_2^k(x)$ and $E_2^k(y)$ is $d\left(E_2^{n-1}(x), E_2^{n-1}(y)\right)$ when k ranges from 0 to $n-1$, because we keep expanding the circle by a factor 2 for each increasing k and thus the points x and y are mapped increasingly further away from each other. Consider the points $p_i := \frac{i}{2^{n+k}} \mod 1 \in S^1$ for $i = 0, 1, \ldots, 2^{n+k} - 1$. The distance between points p_i and p_{i+1} is

$$d(p_i, p_{i+1}) = 2^{-(n+k)} < 2^{-n}$$

hence we can use 4.1 to obtain

$$d_n(p_i, p_{i+1}) = 2^{-(k+1)}$$

for all $i = 0, 1, \ldots, 2^{n+k} - 2$. Because there is no point $p_j = \frac{j}{2^{n+k}} \mod 1$ such that $p_i < p_j < p_{i+1}$ for any $i = 0, 1, \ldots, 2^{n+k} - 2$, it holds that

$$d_n(p_i, p_j) \geq 2^{-(k+1)}$$

for all $i \neq j$. Note that the points p_i form an (n, ε_k)-separated set, containing 2^{n+k} points. Therefore $N(n, \varepsilon_k) \geq 2^{n+k}$. Now consider a subset $A \subseteq S^1$ with cardinality at least $2^{n+k} + 1$. It is easy to see that there must be two points $x \neq y \in A$ such that

$$d(x, y) < 2^{-(n+k)} \leq 2^{-(k+1)}.$$

Namely, it is impossible to divide the interval $[0, 1]$ into $2^{n+k} + 1$ non-overlapping segments of length $2^{-(n+k)}$. It follows that A is not (n, ε_k)-separated and therefore $N(n, \varepsilon_k) \leq 2^{n+k}$. We can conclude that $N(n, \varepsilon_k) = 2^{n+k}$ and use this to calculate

$$\begin{aligned} h(E_2) &= \lim_{k\to\infty} \limsup_{n\to\infty} \frac{1}{n} \log N(n, \varepsilon_k) \\ &= \lim_{k\to\infty} \limsup_{n\to\infty} \frac{1}{n} \log 2^{n+k} \\ &= \lim_{k\to\infty} \limsup_{n\to\infty} \frac{n+k}{n} \log 2 = \log 2. \end{aligned}$$

For any natural number $m \in \mathbb{N}$, the calculation of the topological entropy can be done in exactly the same way as shown above when we substitute 2 for m. We then find that

$$h(E_m) = \log m.$$

Example 4.0.3 (Shift map)**.** We consider the space

$$X := \{(a_n)_{n\in\mathbb{Z}} : a_n \in \{0, 1\}\}$$

and the map $f : X \to X$ that shifts the sequence by sending $(a_n)_{n\in\mathbb{Z}} \mapsto (a_{n+1})_{n\in\mathbb{Z}}$. We can also write the the set X as

$$X = \prod_{i\in\mathbb{Z}} X_i$$

where $X_i = \{0, 1\}$ for every $i \in \mathbb{Z}$. Since each X_i is compact, X is compact in the product topology. This follows by *Tychonoff's Theorem* [10, Theorem 37.3], which states that any arbitrary product of compact topological spaces is compact with respect to the product topology. We define a distance on X as

$$d(a,b) = \begin{cases} 2^{-s} & \text{if } a \neq b \\ 0 & \text{if } a = b \end{cases}$$

where s is the smallest number in $\mathbb{N}_0$ such that $a_n = b_n$ or $a_{-n} = b_{-n}$. We can now calculate the entropy of $f : X \to X$, based on the calculation in [3, Proposition 7.2] for the shift map on one-sided sequences.

Let $n, p \in \mathbb{N}$ and let $a = (a_n)_{n\in\mathbb{Z}}, b = (b_n)_{n\in\mathbb{Z}} \in X$. Consider the sequence $(\varepsilon_k)_{k\in\mathbb{N}} \subseteq \mathbb{R}$ with $\varepsilon_k = 2^{-k}$. Since $\varepsilon_k \to 0$ as $k \to \infty$, we can write

$$h(f) = \lim_{k\to\infty} \limsup_{n\to\infty} \frac{H(n,\varepsilon_k)}{n}.$$

For any $k \in \mathbb{N}$ we have

$$d(f^k(a), f^k(b)) \geq 2^{-p} = \varepsilon_p$$

if and only if $s \in \{k, \dots, p+k-1\}$. Since

$$d_n(a,b) = \max\{d(f^k(a), f^k(b)) : k = 0, \dots, n-1\}$$

it follows that $d_n(a,b) \geq 2^{-p}$ if and only if $s \in \{0, \dots, p+n-2\}$. We can choose $p+n-1$ elements out of $\{0,1\}$ in order in 2^{p+n-1} ways. Therefore, the number of sequences that differ in the first $p+n-1$ elements is 2^{p+n-1} and

$$N(n,\varepsilon_p) \leq 2^{p+n-1}.$$

A sequence $x = (x_n)_{n\in\mathbb{Z}} \in X$ is periodic if $f^m(x) = x$ for some $m \in \mathbb{N}$. That is, x is periodic if and only if the elements $x_0, \dots, x_m$ are infinitely repeated throughout the whole sequence. The number sequences in X that are periodic for $m \in \mathbb{N}$ is 2^m, because these sequences are defined precisely by their first m elements, starting from zero. It follows that the number of sequences in X that are periodic for $p+n-1$ is 2^{p+n-1}. Assume that $a \neq b \in X$ are such sequences. Then $s \in \{0, \dots, n-1\}$ and it follows that

$$d_n(a,b) \geq 2^{p+n-1}.$$

We can conclude that

$$N(n,\varepsilon_p) \geq 2^{p+n-1}$$

and thus

$$N(n,\varepsilon) = 2^{p+n-1}.$$

We can use this to calculate

$$\begin{aligned} h(f) &= \lim_{p\to\infty} \limsup_{n\to\infty} \frac{\log N(n,\varepsilon_p)}{n} \\ &= \lim_{p\to\infty} \limsup_{n\to\infty} \frac{\log(2^{p+n-1})}{n} \\ &= \lim_{p\to\infty} \limsup_{n\to\infty} \frac{p+n-1}{n} \log 2 = \log 2. \end{aligned}$$

The above calculation can be done for any natural number $N \in \mathbb{N}$. Then we define

$$X := \{(a_n)_{n\in\mathbb{Z}} : a_n \in \{0, \ldots, N-1\}\}$$

and we find that the topological entropy of the shift map is

$$h(f) = \log N.$$

The set X_i such that $a_n \in X_i$ for every $n \in \mathbb{Z}$ does not have to be finite or even countable. When X_i is a compact Hausdorff space that has an infinite number of points, such as the set S^1, then the entropy of the shift map on $X = \prod_{i\in\mathbb{Z}} X_i$ is infinite. We can see this when we consider the number of periodic sequences in X. Any constant sequence, that is $(a_n)_{n\in\mathbb{Z}} \in X$ with $a_n = c$ for every $n \in \mathbb{Z}$ and some $c \in S^1$, is periodic with period 1. Because S^1 has infinitely many points, there are infinitely many periodic sequences with period 1. Therefore

$$N(n, \varepsilon) = \infty$$

for some $n \in \mathbb{N}$ and $\varepsilon > 0$. Therefore the entropy of the shift map is

$$h(f) = \lim_{\varepsilon\to 0} \limsup_{n\to\infty} \frac{\log N(n, \varepsilon)}{n} = \infty.$$

We conclude that the entropy of the shift map completely depends on the choice of X_i. It is congruent with $|X_i|$ the number of elements in X_i.

5 When is the topological entropy equal to zero?

In Example 4.0.1 we saw that the rotation map on the unit circle had entropy $h(R_\alpha) = h(R_\alpha) = 0$. This example can be expanded to show that for every isometry on a compact metric space the topological entropy is equal to zero.

Theorem 5.0.1. *Let (X, d) be a compact metric space and let $f : X \to X$ be an isometry. Then the topological entropy*

$$h(f) = 0.$$

Proof. Since $f : X \to X$ is an isometry on (X, d), it preserves the distance, meaning that

$$d(x, y) = d(f(x), f(y))$$

for every $x, y \in X$. It follows immediately that

$$d(x, y) = d(f^k(x), f^k(y))$$

for every $k \in \mathbb{N}$ which implies that

$$d(x, y) = d_n(x, y)$$

for every $n \in \mathbb{N}$. Then by the same reasoning we followed for the rotation map R_α we can conclude that $h(f) = 0$. ☐

Of course, this makes sense intuitively as well. Because the distance between any two points stays the same, no matter where we are in the time, the complexity of the system does not actually change at all.

On the unit circle, the entropy is zero for an even bigger class of functions, namely that of homeomorphisms on the unit circle. Note that every isometry on S^1 is also a homeomorphism because it is continuous by preserving distances, its inverse exists because an isometry on S^1 is bijective, and the inverse is also continuous because it preserves distances.

Theorem 5.0.2. *[13, Theorem 7.14] Let $f : S^1 \to S^1$ be a homeomorphism on the unit circle. Then*

$$h(f) = 0.$$

Proof. We will only give an idea of the proof. The whole proof can be read in [13]. This proof is done using (n, ε)-spanning sets and the alternative definition

$$h(f) = \lim_{\varepsilon \to 0} \limsup_{n \to \infty} \frac{\log M(n, \varepsilon)}{n}$$

of topological entropy. We can choose $\varepsilon > 0$ such that

$$d\left(f^{-1}(x), f^{-1}(y)\right) \leq \frac{1}{4}$$

for every $x, y \in S^1$ such that $d(x,y) \leq \varepsilon$. We then have $M(1,\varepsilon) \leq \left\lceil \frac{1}{\varepsilon} \right\rceil$, where $\lceil \cdot \rceil$ denotes the ceiling function, because we can cover the unit circle with $\left\lceil \frac{1}{\varepsilon} \right\rceil$ balls with radius ε. It can be shown by induction that $M(n,\varepsilon) \leq n \left\lceil \frac{1}{\varepsilon} \right\rceil$. It follows that

$$h(f) = 0$$

by the definition used in this proof. □

Corollary 5.0.3. *[13, Corollary 7.14.1] Let $f : [0,1] \to [0,1]$ be a homeomorphism. Then*

$$h(f) = 0.$$

We can relate the unit circle to the interval $[0,1]$ in the way described in Chapter 2, hence it makes sense that since any homeomorphism on the unit circle has zero entropy, any homeomorphism on the interval $[0,1]$ also has zero entropy.

We have also seen in Theorem 3.3.3 that $h(f^m) = m \cdot h(f)$ for $f : X \to X$ a continuous function on a compact metric space, and any $m \in \mathbb{N}$. The following statement follows immediately.

Corollary 5.0.4. *Let (X,d) be a compact metric space and $f : X \to X$ a continuous function. If there exists $m \in \mathbb{N}$ such that $f^m = Id$, then $h(f) = 0$.*

Recall that if $f^m(x) = x$ for $m \in \mathbb{N}$, then x is a periodic point of f with period m. Since $f^m = Id$, every point $x \in X$ is periodic with period m. We can see now that if every $x \in X$ is periodic under f, then the entropy of f is zero. Note that we do not need every point to have the same minimal period. We only need the set

$$Q = \{q \in \mathbb{N} : q \text{ is the period of } \ x \in X\}$$

is a finite set. Namely, then

$$f^{\prod_{q \in Q} q}(x) = x$$

for every $x \in X$.

Finally, we want to look at a case where the topological entropy is always positive. It relies on a topological dynamical system having the *specification property.* This property is somewhat technical, so I will refer you to Chapter 21 in the book by Manfred Denker, Christian Grillenberger and Karl Sigmund [6] for the precise definition and the proof of the result we state below.

Theorem 5.0.5. *Let (X,d) be a metric space and $f : X \to X$ a continuous function. Assume that the cardinality of X is at least two. If the dynamical system (X,f) satisfies the specification property, then*

$$h(f) > 0.$$

The shift map we defined in Example 4.0.3 satisfies the specification property, hence by this theorem the topological entropy of the shift map is positive. That corresponds with the results we have seen in the example.

Of course, the results mentioned in this chapter are not the only cases of which we know something about the topological entropy. One can do more research into when the entropy is zero and when it is positive or infinite.

6 Outlook on applications

In the introduction of this thesis it was briefly mentioned that topological entropy has applications in several fields. In this chapter, we will look at some of these applications.

Application 6.0.1 (Formal languages)**.** The first application we discuss is in the field of computer science. In this field, topological entropy can be used to calculate the complexity of a formal language. A formal language needs an *alphabet*, which is a non-empty set. The language is a subset of all possible finite string, called *words*, we can make using elements of the alphabet. One can for example look at a *Dyck language*. The alphabet of a Dyck language consists of matching parentheses such as (and) or { and }. Any word in the Dyck language that uses the set $\{(,)\}$ as its alphabet has an equal number of opening and closing brackets, and before any closing bracket, there must be more opening than closing brackets. The strings () and (())() are thus words in this language, but)(and (())) are not.

When classifying formal languages in terms of their complexity, the *Chomsky hierarchy* is often used. The Chomsky hierarchy places formal languages in levels of growing complexity. It starts with the *regular* languages, which are accepted by finite automata and are considered to be the most simple. Friedrich Martin Schneider and Daniel Borchmann [11] state in their paper that using the Chomsky hierarchy leads to some contra-intuitive classifications of formal languages. Take the Dyck language with only one pair of parentheses we saw above. Intuitively, we can see that this language is relatively "simple". However, because it is not a regular language, it is placed in a class of more complicated languages by the Chomsky hierarchy. They propose a different way to classify the complexity of formal languages using just one computation model that works for all languages, namely *topological automata*. A topological automaton is a topological dynamical system, hence we can use the notion of topological entropy on it. It turns out that when we do this, all regular languages have entropy zero, as expected, and the Dyck language with one pair of parentheses also has zero entropy and can thus be considered to be simple.

Application 6.0.2 (DNA sequences)**.** Another application of topological entropy can be found in the field of biology. Specifically, topological entropy can be used to analyse the randomness of human DNA. In 2011, David Koslicki [9] defined an approximation to the notion of topological entropy we have seen in this paper that can be used for finite sequences.

The goal of the study is to compare the randomness of introns and exons in the human DNA. Introns are parts of the DNA that do not directly code for proteins and that are removed during RNA splicing. Exons do code for proteins and we would expect that exons should be less random than introns. In earlier research using topological entropy to measure the randomness of introns and exons, it was concluded that the entropy of exons is higher than the entropy of introns, meaning that exons are actually more random than introns. According to Koslicki this conclusion in incorrect, due to finite-sample effects and high-dimensionality problems with the definition of topological entropy used. With his adapted definition, Koslicki avoids these problems and the results are as expected. The entropy of introns is higher than that of exons, but both entropies are lower than that of a completely random sequence.

Of course, topological entropy also has mathematical applications. It is closely related to measure theoretic entropy and ergodic theory, as can be seen in most of the references. One important mathematical property of topological entropy is that it is the supremum of the metric entropies over all invariant probability measures, such as the well-known Shannon entropy. We will not go into the mathematical applications of topological entropy any further here, but for the aforementioned result I will refer you to [2, Theorem 4.7].

7 References

[1] Roy L Adler, Alan G Konheim, and M Harry McAndrew. Topological entropy. *Transactions of the American Mathematical Society*, 114(2):309–319, 1965.

[2] Luis Barreira. *Ergodic theory, hyperbolic dynamics and dimension theory*. Springer, 2012.

[3] Luis Barreira and Claudia Valls. *Dynamical systems: An introduction*. Springer Science & Business Media, 2012.

[4] Rufus Bowen. Entropy for group endomorphisms and homogeneous spaces. *Transactions of the American Mathematical Society*, 153:401–414, 1971.

[5] Michael Brin and Garrett Stuck. *Introduction to dynamical systems*. Cambridge university press, 2002.

[6] Manfred Denker, Christian Grillenberger, and Karl Sigmund. *Ergodic theory on compact spaces*. Springer, 2006.

[7] Efim Isaakovich Dinaburg. A correlation between topological entropy and metric entropy. In *Doklady Akademii Nauk*, volume 190, pages 19–22. Russian Academy of Sciences, 1970.

[8] James Kelly and Tim Tennant. Topological entropy on set-valued functions. *arXiv preprint arXiv:1509.08413*, 2015.

[9] David Koslicki. Topological entropy of dna sequences. *Bioinformatics*, 27(8):1061–1067, 2011.

[10] James Munkres. Topology james munkres second edition, 2014.

[11] Friedrich Martin Schneider and Daniel Borchmann. Topological entropy of formal languages. In *Semigroup Forum*, volume 94, pages 556–581. Springer, 2017.

[12] J Michael Steele. *Probability theory and combinatorial optimization*. SIAM, 1997.

[13] Peter Walters. *An introduction to ergodic theory*, volume 79. Springer Science & Business Media, 2000.