Topological Pressure and Topological Entropy
- Seminar Talk -

Stilianos Louca

July 14, 2011

Contents

What this is

This is a small collection of material prepared by me for a seminar talk on *Dynamical Systems and Fractals*, given at the Friedrich Schiller University of Jena, Germany in May 2011. Most of the material and proofs were taken from [1, 2] and [3], although some definitions presented here may differ from what is found in one or the other.

1 Preliminaries

1.0.1 Definition: Iterated dynamical system

Let X be some set and $T : X \to X$ some mapping. Then the tuple (X, T) shall be called an **iterated dynamical system** and T the **generator of its dynamic**. It is called **continuous** if X is a topological space and T is continuous. The tuple $(X, \mathfrak{S}, \mu, T)$ is called a **measure-preserving, iterated dynamical system**[1], if $(X, \mathfrak{S}, \mu)$ is a measure space and $T : X \to X$ measurable and measure preserving, that is $\mu \circ T^{-1} = \mu$.

1.0.2 Definition: Orbit

Let (X, T) be an iterated dynamical system. For some point $x \in X$, we call the set $\bigcup_{n\in\mathbb{N}_0} \{T^n x\}$ **orbit** of x under T.

1.0.3 Definition: Conjugated dynamical systems

Two iterated dynamical systems (X, T) and $(\widetilde{X}, \widetilde{T})$ shall be called **conjugated** if there exists a bijection $\varphi : X \to \widetilde{X}$ such that $\varphi \circ T = \widetilde{T} \circ \varphi$. If the systems are continuous, we call them **continuously conjugated** if the mapping φ can be chosen to be continuous in both directions. If $(X, \mathfrak{S}, \mu, T)$, $(\widetilde{X}, \widetilde{\mathfrak{S}}, \widetilde{\mu}, \widetilde{T})$ are measure preserving systems and φ measurable in both directions such that $\mu \circ \varphi^{-1} = \widetilde{\mu}$, then the systems are called **measurably conjugated**.

Note that conjugacy is an equivalence relation between iterated dynamical systems. Similarly, continuous (measurable) conjugacy is an equivalence relation between continuous (measure-preserving), iterated dynamical systems.

2 Dynamical systems on probability spaces

2.1 Information and entropy of partitions

2.1.1 Definition: Partition and generated σ-algebras

A **measurable partition** $\mathcal{A}$ of a measurable space $(X, \mathfrak{S})$ is a family of measurable, disjoint subsets of X, covering the whole space X. For measurable partitions $\mathcal{A}_1, .., \mathcal{A}_n$ of $(X, \mathfrak{S})$ we call the measurable partition $\bigvee_{k=1}^n \mathcal{A}_k := \{\bigcap_{k=1}^n A_k : A_k \in \mathcal{A}_k\}$ the **refinement** of all $\mathcal{A}_1, .., \mathcal{A}_n$. We call a measurable partition $\mathcal{B}$ **finer** than $\mathcal{A}$ ($\mathcal{A}$ **coarser** than $\mathcal{B}$) and write $\mathcal{A} \preccurlyeq \mathcal{B}$ if $\mathcal{A} \vee \mathcal{B} = \mathcal{B}$. Note that $\preccurlyeq$ is a partial order on the system of measurable partitions of $(X, \mathfrak{S})$ and that $\mathcal{A} \preccurlyeq \mathcal{A} \vee \mathcal{B}$ for all measurable partitions $\mathcal{A}, \mathcal{B}$.

We say that an increasing sequence of σ-algebras $(\mathfrak{B}_n)_{n\in\mathbb{N}}$ converges to the σ-algebra $\mathfrak{B}$ and write $\mathfrak{B}_n \uparrow \mathfrak{B}$ if $\mathfrak{B} = \sigma\left[\bigcup_{n\in\mathbb{N}} \mathfrak{B}_n\right]$. For a countable number of measurable partitions $(\mathcal{A}_n)_{n\in\mathbb{N}}$ we write $\bigvee_{n\in\mathbb{N}} \mathcal{A}_n := \sigma\left(\bigcup_{n\in\mathbb{N}} \bigvee_{k=1}^n \mathcal{A}_k\right)$. Thus for an increasing sequence $\mathcal{A}_1 \preccurlyeq \mathcal{A}_2 \preccurlyeq ..$ of measurable partitions $\sigma(\mathcal{A}_n) \uparrow \bigvee_{k\in\mathbb{N}} \mathcal{A}_k$.
We shall note $\mathcal{Z}(\mathfrak{S})$ the system of all countable, measurable partitions of $(X, \mathfrak{S})$.

2.1.2 Definition: Information and entropy of partitions

Let $(X, \mathfrak{S}, \mu)$ be a probability space, $\mathcal{A} \in \mathcal{Z}(\mathfrak{S})$ a countable, measurable partition of $(X, \mathfrak{S})$ and $\mathfrak{B} \subseteq \mathfrak{S}$. Then the function

$$I_\mu(\mathcal{A}|\mathfrak{B}) := -\sum_{A\in\mathcal{A}} 1_A \cdot \ln \mu(A|\sigma(\mathfrak{B})) \tag{2.1}$$

is called **information of $\mathcal{A}$ conditional upon $\mathfrak{B}$**. Here, $\mu(A|\sigma(\mathfrak{B}))$ is the probability of A conditional upon the σ-algebra $\sigma(\mathfrak{B})$ generated by $\mathfrak{B}$. We call the number

$$H_\mu(\mathcal{A}|\mathfrak{B}) := \mathbb{E}_\mu I_\mu(\mathcal{A}|\mathfrak{B}) \tag{2.2}$$

entropy of $\mathcal{A}$ conditional upon $\mathfrak{B}$. For $\mathfrak{B} = \{\emptyset, X\}$ we call $I_\mu(\mathcal{A}) := I_\mu(\mathcal{A}|\mathfrak{B})$ simply **information** and $H_\mu(\mathcal{A}) := H_\mu(\mathcal{A}|\mathfrak{B})$ **entropy** of $\mathcal{A}$. We shall note $\mathcal{Z}_1(\mathfrak{S}, \mu)$ the system of countable, measurable partitions of $(X, \mathfrak{S})$ which posses finite entropy with respect to μ.

[1]In literature[1] often simply called a dynamical system.

2.1.3 Proposition: Representation of entropy of partitions

Let $(X, \mathfrak{S}, \mu)$ be a probability space, $\mathcal{A}, \mathcal{B} \in \mathcal{Z}(\mathfrak{S})$ countable, measurable partitions of $(X, \mathfrak{S})$. Then:

1. $I_\mu(\mathcal{A}) = -\sum_{A\in\mathcal{A}} 1_A \cdot \ln \mu(A)$ μ-almost everywhere.
2. $H_\mu(\mathcal{A}) = -\sum_{A\in\mathcal{A}} \mu(A) \cdot \ln \mu(A)$.
3. For any $A \in \mathfrak{S}$ and μ-almost all $x \in X$ one has $\mu(A|\sigma(\mathcal{B}))(x) = \mu(A|B)$ with $x \in B \in \mathcal{B}$.
4. $H_\mu(\mathcal{A}|\mathcal{B}) = -\displaystyle\sum_{B\in\mathcal{B}} \mu(B) \sum_{A\in\mathcal{A}} \mu(A|B) \ln \mu(A|B)$.

Proof: Follows from the definitions of information and entropy.

2.1.4 Properties of information and entropy of partitions

Let $(X, \mathfrak{S}, \mu)$ be a probability space, $\mathcal{A}, \mathcal{B}, \mathcal{D} \in \mathcal{Z}(\mathfrak{S})$ countable, measurable partitions of $(X, \mathfrak{S})$ and $\mathfrak{B} \subseteq \mathfrak{S}$. Then:

1. $I_\mu(\mathcal{A}|\mathfrak{B}) \geq 0$ μ-almost everywhere.
2. $I_\mu(\mathcal{A}|\mathfrak{B}) = 0$ μ-almost everywhere if and only if $A \in \sigma(\mathfrak{B})$ for all $A \in \mathfrak{S}$ with $\mu(A) > 0$.
3. If $\mathcal{A} \preccurlyeq \mathcal{B}$ then $I_\mu(\mathcal{A}|\mathfrak{B}) \leq I_\mu(\mathcal{B}|\mathfrak{B})$ μ-almost everywhere.
4. If $\mathcal{B} \preccurlyeq \mathcal{D}$ then $H_\mu(\mathcal{A}|\mathcal{D}) \leq H_\mu(\mathcal{A}|\mathcal{B})$. In particular $H_\mu(\mathcal{A}|\mathcal{D}) \leq H_\mu(\mathcal{A})$.
5. $H_\mu(\mathcal{A}) \leq \ln |\mathcal{A}|$.
6. (Triangle inequality) $H_\mu(\mathcal{A}|\mathcal{D}) \leq H_\mu(\mathcal{A}|\mathcal{B}) + H_\mu(\mathcal{B}|\mathcal{D})$.
7. $I_\mu(\mathcal{A} \vee \mathcal{B}|\mathcal{D}) = I_\mu(\mathcal{A}|\mathcal{D}) + I_\mu(\mathcal{B}|\mathcal{A} \vee \mathcal{D})$ μ-almost everywhere.
8. $H_\mu(\mathcal{A} \vee \mathcal{B}|\mathcal{D}) = H_\mu(\mathcal{A}|\mathcal{D}) + H_\mu(\mathcal{B}|\mathcal{A} \vee \mathcal{D}) \leq H_\mu(\mathcal{A}|\mathcal{D}) + H_\mu(\mathcal{B}|\mathcal{D})$.
9. Consequently, $H_\mu(\mathcal{A}) - H_\mu(\mathcal{B}) = H_\mu(\mathcal{A}|\mathcal{B}) - H_\mu(\mathcal{B}|\mathcal{A})$ if both sides are well-defined.
10. If $\sigma(\mathcal{A})$ and $\sigma(\mathcal{B})$ are independent, then $I_\mu(\mathcal{A}|\mathcal{B}) = I_\mu(\mathcal{A})$ μ-almost everywhere and $H_\mu(\mathcal{A}|\mathcal{B}) = H_\mu(\mathcal{A})$.
11. If $\sigma(\mathcal{A})$ and $\sigma(\mathcal{B})$ are independent, then $I_\mu(\mathcal{A} \vee \mathcal{B}) = I_\mu(\mathcal{A}) + I_\mu(\mathcal{B})$ μ-almost everywhere.

Proof: For (5) use the concavity of $\ln(\cdot)$ and Jensen's inequality A.2.7. For the rest, see [1], Proposition 20 & Korollar 15 and [2], Theorem 1.3.1.

2.1.5 Lemma: Continuity of information and entropy

Let $(X, \mathfrak{S}, \mu)$ be a probability space, $\mathcal{A}, \mathcal{A}_n \in \mathcal{Z}_1(\mathfrak{S}, \mu)$ be countable, measurable partitions with finite entropy and $\mathfrak{B}, \mathfrak{B}_n \subseteq \mathfrak{S}$ be σ-algebras. Suppose that $\mathfrak{B}_n \uparrow \mathfrak{B}$ (that is, $\sigma(\bigcup_n \mathfrak{B}_n) = \mathfrak{B}$) and $\sigma(\mathcal{A}_n) \uparrow \sigma(\mathcal{A})$. Then:

1. $I_\mu(\mathcal{A}_n|\mathfrak{B}) \uparrow I_\mu(\mathcal{A}|\mathfrak{B})$ μ-almost everywhere.
2. $H_\mu(\mathcal{A}_n|\mathfrak{B}) \uparrow H_\mu(\mathcal{A}|\mathfrak{B})$.
3. $I_\mu(\mathcal{A}|\mathfrak{B}_n) \overset{n\to\infty}{\longrightarrow} I_\mu(\mathcal{A}|\mathfrak{B})$ μ-almost everywhere.
4. $H_\mu(\mathcal{A}|\mathfrak{B}_n) \downarrow H_\mu(\mathcal{A}|\mathfrak{B})$.

Proof: See [1], Satz 103.

2.2 Kolmogorov-Sinai entropy of dynamical systems

We have up to now developed the notions of entropy for mere partitions of probability spaces without reference to any underlying dynamic. We shall now use these preconsiderations to introduce the Kolmogorov-Sinai entropy for measure-preserving, iterated dynamical systems $(X, \mathfrak{S}, \mu, T)$ on probability spaces $(X, \mathfrak{S}, \mu)$.

2.2.1 Definition: Entropy of a dynamical system

Let $(X, \mathfrak{S}, \mu)$ be a probability space, $(X, \mathfrak{S}, \mu, T)$ a measure-preserving, iterated dynamical system and $\mathcal{A} \in \mathcal{Z}(\mathfrak{S})$ a countable, measurable partition of $(X, \mathfrak{S})$. Then

$$h_\mu(T, \mathcal{A}) := H_\mu\left[\mathcal{A} | \bigvee_{k\in\mathbb{N}} T^{-k}(\mathcal{A})\right] \tag{2.3}$$

is called **mean entropy** of $\mathcal{A}$ within the dynamical system. It corresponds in a sense to the entropy of the partition conditional upon all its pre-images. The supremum

$$h_\mu(T) := \sup\left\{h_\mu(T, \mathcal{A}) : \mathcal{A} \in \mathcal{Z}_1(\mathfrak{S}, \mu)\right\} \tag{2.4}$$

is called the (**measure-theoretic** or **Kolmogorov-Sinai**) **entropy** of the dynamical system.

2.2.2 Proposition: Invariance of the Kolmorov-Sinai entropy

Two measurably conjugated, measure-preserving, iterated dynamical systems have same Kolmogorov-Sinai entropy.

Proof: See [1], Satz 109.

2.2.3 Theorem: Representation of mean entropies [Shannon-McMillan-Breiman]

Let $(X, \mathfrak{S}, \mu)$ be a probability space, $(X, \mathfrak{S}, \mu, T)$ a measure-preserving, iterated dynamical system and $\mathcal{A} \in \mathcal{Z}_1(\mathfrak{S}, \mu)$ a countable, measurable partition of $(X, \mathfrak{S})$ with finite entropy. Then

$$h_\mu(T, \mathcal{A}) = \lim_{n\to\infty} \frac{1}{n} H_\mu\left[\bigvee_{k=0}^{n-1} T^{-k}\mathcal{A}\right], \tag{2.5}$$

with the sequence on the right hand side monotonically decreasing.

Proof: See [1], Satz 104.

2.2.4 Lemma: Properties of the mean entropy

Let $(X, \mathfrak{S}, \mu)$ be a probability space, $(X, \mathfrak{S}, \mu, T)$ a measure-preserving, iterated dynamical system and $\mathcal{A}, \mathcal{B} \in \mathcal{Z}_1(\mathfrak{S}, \mu)$ countable, measurable partitions of $(X, \mathfrak{S})$ with finite entropy. Then:

1. $h_\mu(T, \mathcal{A}) \leq H_\mu(\mathcal{A})$.
2. $h_\mu(T, \mathcal{A} \vee \mathcal{B}) \leq h_\mu(T, \mathcal{A}) + h_\mu(T, \mathcal{B})$.
3. If $\mathcal{A} \preccurlyeq \mathcal{B}$ then $h_\mu(T, \mathcal{A}) \leq h_\mu(T, \mathcal{B})$.
4. $h_\mu(T, \mathcal{A}) \leq h_\mu(T, \mathcal{B}) + H_\mu(\mathcal{A}|\mathcal{B})$.

Proof: See [2], theorem 1.4.4 and [4].

2.2.5 Lemma: Persistency of entropy approximation

Let $(X, \mathfrak{S}, \mu)$ be a probability space and $(X, \mathfrak{S}, \mu, T)$ a measure-preserving, iterated dynamical system. Let $\mathcal{A}, \mathcal{B} \in \mathcal{Z}_1(\mathfrak{S}, \mu)$ be countable, measurable partitions of the space such that $H_\mu(\mathcal{A}|\mathcal{B}) \leq \varepsilon$. Then

$$H_\mu\left[\bigvee_{k=0}^{n-1} T^{-k}\mathcal{A}\right] \leq H_\mu\left[\bigvee_{k=0}^{n-1} T^{-k}\mathcal{B}\right] + n\varepsilon \tag{2.6}$$

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

Proof: The following proof was taken from [2]. Note $\mathcal{A}^{(n)} := \bigvee_{k=0}^{n-1} T^{-k}\mathcal{A}$ and $\mathcal{B}^{(n)} := \bigvee_{k=0}^{n-1} T^{-k}\mathcal{B}$. Then

$$\begin{aligned} H_\mu(\mathcal{A}^{(n)}) &\overset{2.1.4(6)}{\leq} H_\mu(\mathcal{B}^{(n)}) + H_\mu(\mathcal{A}^{(n)}|\mathcal{B}^{(n)}) \overset{2.1.4(8)}{\leq} H_\mu(\mathcal{B}^{(n)}) + \sum_{k=0}^{n-1} H_\mu(T^{-k}\mathcal{A}|\mathcal{B}^{(n)}) \\ &\overset{2.1.4(4)}{\leq} H_\mu(\mathcal{B}^{(n)}) + \sum_{k=0}^{n-1} \underbrace{H_\mu(T^{-k}\mathcal{A}|T^{-k}\mathcal{B})}_{H_\mu(\mathcal{A}|\mathcal{B})\leq\varepsilon} \leq H_\mu(\mathcal{B}^{(n)}) + n\varepsilon \end{aligned} \tag{2.7}$$

as claimed.

□

2.2.6 Proposition: Characterization of entropy by finite partitions

Let $(X, \mathfrak{S}, \mu)$ be a probability space and $(X, \mathfrak{S}, \mu, T)$ a measure-preserving, iterated dynamical system. Then its Kolmogorov-Sinai entropy may be taken as the supremum

$$h_\mu(T) = \sup\left\{h_\mu(T, \mathcal{A}) : \mathcal{A} \in \mathcal{Z}(\mathfrak{S}),\ |\mathcal{A}| < \infty\right\} \tag{2.8}$$

over all finite, measurable partitions of $(X, \mathfrak{S})$.

Proof: Clearly, it suffices to show inequality "$\leq$" in (2.9). For that, it suffices to show that for any $\varepsilon > 0$ and any countable, measurable partition $\mathcal{A} \in \mathcal{Z}_1(\mathfrak{S}, \mu)$ with finite entropy, there exists a finite, measurable partition $\mathcal{B}$ such that $h_\mu(T, \mathcal{A}) \leq h_\mu(T, \mathcal{B}) + \varepsilon$. By lemma 2.2.5 and representation (2.5), it suffices to show that $H_\mu(\mathcal{A}|\mathcal{B}) \leq \varepsilon$. Now let $\mathcal{A} = \{A_1, A_2, ..\} \in \mathcal{Z}_1(\mathfrak{S}, \mu)$ be given. As $\mathcal{A}$ has finite entropy, we can fix an $m \in \mathbb{N}$ such that

$$-\sum_{n=m+1}^{\infty} \mu(A_n) \ln \mu(A_n) \leq \varepsilon. \tag{2.9}$$

Set $B_n := A_n$ for $n \in \{1, .., m\}$, $B_0 := \bigcup_{n=m+1}^{\infty} A_n$ and $\mathcal{B} := \{B_0, .., B_m\}$. Then $\mathcal{B}$ is a finite, measurable partition and satisfies

$$\begin{aligned} H_\mu(\mathcal{A}|\mathcal{B}) &\overset{2.1.3(4)}{=} -\sum_{n=1}^{m}\sum_{k=1}^{\infty} \underbrace{\mu(A_k \cap B_n)}_{0 \text{ if } n\neq k} \underbrace{\ln \mu(A_k|B_n)}_{0 \text{ if } n=k} - \sum_{k=1}^{\infty} \underbrace{\mu(A_k \cap B_0)}_{0 \text{ if } k\leq m} \ln(A_k|B_0) \\ &= -\sum_{k=m+1}^{\infty} \mu(A_k \cap B_0) \ln \mu(A_k|B_0) = -\sum_{k=m+1}^{\infty} \mu(A_k) \ln \frac{\mu(A_k)}{\mu(B_0)} \\ &\leq -\sum_{k=m+1}^{\infty} \mu(A_k) \ln \mu(A_k) \leq \varepsilon \end{aligned} \tag{2.10}$$

as intended.

□

2.2.7 Theorem: Representation of entropy [Kolmogorov-Sinai]

Let $(X, \mathfrak{S}, \mu)$ be a probability space and $(X, \mathfrak{S}, \mu, T)$ a measure-preserving, iterated dynamical system. Let $(\mathcal{A}_n)_{n\in\mathbb{N}} \subseteq \mathcal{Z}_1(\mathfrak{S}, \mu)$ be a fining sequence of countable, measurable partitions of $(X, \mathfrak{S})$ with finite entropy, such that $\sigma\left[\bigcup_{k\in\mathbb{N}_0, n\in\mathbb{N}} T^{-k}\mathcal{A}_n\right] = \mathfrak{S}$. Then

$$h_\mu(T) = \lim_{n\to\infty} h_\mu(T, \mathcal{A}_n). \tag{2.11}$$

In particular, if $\mathcal{A} \in \mathcal{Z}_1(\mathfrak{S}, \mu)$ is such that $\bigvee_{k\in\mathbb{N}_0} T^{-k}\mathcal{A} = \mathfrak{S}$ (we say that $\mathcal{A}$ is a **generator** for $(X, \mathfrak{S}, T)$), then

$$h_\mu(T) = h_\mu(T, \mathcal{A}). \tag{2.12}$$

Proof: See [1], Satz 106.

2.2.8 Lemma: Entropy of higher order dynamics

Let $(X, \mathfrak{S}, \mu)$ be a probability space and $(X, \mathfrak{S}, \mu, T)$ a measure-preserving, iterated dynamical system. Then

$$h_\mu(T^n) = n \cdot h_\mu(T) \tag{2.13}$$

for any $n \in \mathbb{N}_0$.

Proof: See [1], Satz 105.

2.2.9 Theorem: Entropy as an affine mapping

Let $(X, \mathfrak{S})$ be a measurable space, $T : X \to X$ measurable and $\mathfrak{M}(\mathfrak{S}, T)$ the convex set of all probability measures on $(X, \mathfrak{S})$ preserved by T. Let $\mathcal{A} \in \mathcal{Z}(\mathfrak{S})$ be a countable, measurable partition of the space. Then:

1. The mapping $\mathfrak{M}(\mathfrak{S}, T) \to [0, \infty], \ \mu \mapsto h_\mu(T, \mathcal{A})$ is affine, that is

$$h_{\lambda\mu+(1-\lambda)\nu}(T, \mathcal{A}) = \lambda \cdot h_\mu(T, \mathcal{A}) + (1 - \lambda) \cdot h_\nu(T, \mathcal{A}) \tag{2.14}$$

   for all $\mu, \nu \in \mathfrak{M}(\mathfrak{S}, T)$ and $\lambda \in [0, 1]$.

2. The mapping $\mathfrak{M}(\mathfrak{S}, T) \to [0, \infty], \ \mu \mapsto h_\mu(T)$ is affine.

Proof: See [5], proposition 10.13.

2.2.10 Example: Entropy of finite dynamical systems

Let $(X, \mathfrak{S}, \mu)$ be a probability space and $(X, \mathfrak{S}, \mu, T)$ a measure-preserving, iterated dynamical system. If the σ-algebra $\mathfrak{S}$ is finite, the system's Kolmogorov-Sinai entropy vanishes.

Proof: The system of atoms $\mathcal{A}$ of $\mathfrak{S}$ is the finest measurable partition of $(X, \mathfrak{S})$. It generates already by its self the σ-algebra $\mathfrak{S}$, thus by Kolmogorov-Sinai 2.2.7 $h_\mu(T) = h_\mu(T, \mathcal{A})$. Now, for any $n \in \mathbb{N}$ one has $\mathcal{A} \preccurlyeq \bigvee_{k=0}^{n-1} T^{-k}\mathcal{A}$, so that actually $\mathcal{A} = \bigvee_{k=0}^{n-1} T^{-k}\mathcal{A}$. By Shannon-McMillan-Breiman this implies $h_\mu(T, \mathcal{A}) = \lim\limits_{n\to\infty} \frac{1}{n} H_\mu(\mathcal{A}) = 0$.

□

3 Dynamical systems on metric spaces

3.1 Topological pressure

In the following we shall restrict ourselves to iterated dynamical systems on compact metric spaces.

3.1.1 Definition: Bowen-Metric

Let (X, d) be a metric space and (X, T) an iterated dynamical system. For $n \in \mathbb{N}_0$ we note

$$d_n(x, y) := \sup_{0 \le k \le n-1} d(T^k x, T^k y) \tag{3.1}$$

for any $x, y \in X$. The so defined metric $d_n : X \times X \to \mathbb{R}_+$ is called the n-th **Bowen-metric** of the dynamical system. It can be interpreted as the *maximum distance* between the two orbits $\{T^k\}_{k=0}^{n-1}$ and $\{T^k y\}_{k=0}^{n-1}$.

Remarks:

(i) The Bowen-metrics d_n are growing in n, that is $d_n \le d_{n+1}$ for any $n \in \mathbb{N}_0$.

(ii) Each Bowen-metric is equivalent to the intrinsic metric d, provided that T is continuous.

3.1.2 Definition: (d,ε)-separated sets

Let (X,d) be a metric space and $\varepsilon > 0$. A subset $N \subseteq X$ is called **(d,ε)-separated** if

$$\inf\{d(x,y) : x \neq y \in N\} \geq \varepsilon. \tag{3.2}$$

It is called **maximally (d,ε)-separated** if it is maximal by inclusion with that property, that is, $N \cup \{x\}$ is no longer (d,ε)-separated for any other point $x \in X \setminus N$.

3.1.3 Proposition: Extension to maximally separated sets

Let (X,d) be a metric space, $\varepsilon > 0$ and $N \subseteq X$ some (d,ε)-separated subset. Then there exists a maximally (d,ε)-separated subset $\widetilde{N} \subseteq X$ containing N.

Proof: Let $\mathcal{N} \subseteq 2^X$ be the system of (d,ε)-separated subsets containing N, equipped with the partial order of inclusion. Then every totally ordered subset $\mathcal{M} \subseteq \mathcal{N}$ has an upper bound in $\mathcal{N}$, namely $\bigcup_{M \in \mathcal{M}} M$. By Zorn's lemma, $\mathcal{N}$ has a maximal element $\widetilde{N} \in \mathcal{N}$. Obviously $\widetilde{N}$ is maximally (d,ε)-separated and includes N.

□

3.1.4 Definition: Topological pressure

Let (X,d) be a compact metric space and (X,T) an iterated dynamical system with n-th Bowen metric d_n. For $\varepsilon > 0$, $n \in \mathbb{N}$ and any continuous function $f \in \mathcal{C}(f,\mathbb{R})$ we shall call

$$S(d,\varepsilon,f,T,n) := \sup\left\{\sum_{x\in N} \exp\left[\sum_{k=0}^{n-1} f(T^k x)\right] : N \text{ is } (d,\varepsilon)\text{-separating set}\right\} \tag{3.3}$$

f-weighted ε-capacities of the system. The *infinitesimal asymptotic growth rate*

$$P_{\text{top}}(T,f) := \lim_{\varepsilon\to 0^+} \limsup_{n\to\infty} \frac{1}{n} \ln S(d_n,\varepsilon,f,T,n) \tag{3.4}$$

is called the **topological pressure** of the system for the **potential** f. For the special case $f = 0$ we call $P_{\text{top}}(T,0) =: h_{\text{top}}(T)$ **topological entropy**[3] of the system. The induced mapping $P_{\text{top}}(T,\cdot) : \mathcal{C}(X) \to \mathbb{R}$ is called **topological pressure function** of the system.

Remarks:

(i) For the special case $f = 0$ we shall call

$$S(d,\varepsilon) := \sup\{|N| : N \subseteq X \text{ is } (d,\varepsilon) \text{ -separated}\} \tag{3.5}$$

ε-capacity of the space (X,d). Note that this definition is not linked to T or any potential.

(ii) As (X,d) is totally bounded, the ε-capacity $S(d,\varepsilon)$ is finite and equal to the cardinality of some maximally (d,ε)-separated set. It corresponds to the maximum number of open ε-balls one can choose in X such that neither one includes the center of the other.

(iii) For two metrics $d \leq \widetilde{d}$ on X one has $S(d,\varepsilon,f,T,n) \leq S(\widetilde{d},\varepsilon,f,T,n)$.

(iv) For $0 < \varepsilon \leq \widetilde{\varepsilon}$ one has $S(d,\widetilde{\varepsilon},f,T,n) \leq S(d,\varepsilon,f,T,n)$.

(v) For any $f \in \mathcal{C}(X)$ one has $S(d,\varepsilon) \cdot e^{n\inf(f)} \leq S(d,\varepsilon,f,T,n) \leq S(d,\varepsilon) \cdot e^{n\cdot\sup(f)}$.

(vi) By remark (iv), the limit (3.1.4) exists.

3.1.5 Theorem: Topological pressure as a topological invariant

Let (X,d) be a compact metric space and (X,T) an iterated dynamical system. Then its topological pressure function $P_{\text{top}}(T,\cdot) : \mathcal{C}(T) \to \mathbb{R}$ is independent of the metric generating the topology of the space.

Proof: We shall provide a proof inspired by [1]. Let $\widetilde{d}$ be a topologically equivalent metric to d. Let $P_{\text{top}}(d,T,f)$ and $P_{\text{top}}(\widetilde{d},T,f)$ be the topological pressures of some function $f \in \mathcal{C}(X)$ defined with respect to the metric d and $\widetilde{d}$ respectively. It suffices to show that $P_{\text{top}}(d,T,f) \leq P_{\text{top}}(\widetilde{d},T,f)$. As the space is compact, by A.2.4 the metrics $d, \widetilde{d}$ are uniformly equivalent, which means that for any given $\varepsilon > 0$ there exists an $\widetilde{\varepsilon}_\varepsilon > 0$ so that $d(x,y) \geq \varepsilon$ implies $\widetilde{d}(x,y) \geq \widetilde{\varepsilon}_\varepsilon$. This relation is inherited directly by the Bowen-metrics d_n and $\widetilde{d}_n$, with the same $\widetilde{\varepsilon}_\varepsilon$. Thus every (d_n,ε)-separated set is also a $(\widetilde{d}_n,\widetilde{\varepsilon}_\varepsilon)$-separated one, meaning that $S(d_n,\varepsilon,f,T,n) \leq S(\widetilde{d}_n,\widetilde{\varepsilon}_\varepsilon,f,T,n)$ for every $\varepsilon > 0$. As this holds for all $n \in \mathbb{N}$, one finds that

$$\limsup_{n\to\infty} \frac{1}{n} \ln S(d_n,\varepsilon,f,T,n) \leq \limsup_{n\to\infty} \frac{1}{n} \ln S(\widetilde{d}_n,\widetilde{\varepsilon}_\varepsilon,f,T,n). \tag{3.6}$$

Now, since $\widetilde{\varepsilon}_\varepsilon$ can be chosen so that $\widetilde{\varepsilon}_\varepsilon \overset{\varepsilon\to 0}{\longrightarrow} 0$, (3.6) implies

$$\begin{aligned} P_{\text{top}}(d,T,f) &= \lim_{\varepsilon\to 0^+} \limsup_{n\to\infty} \frac{1}{n} \ln S(d_n,\varepsilon,f,T,n) \\ &\leq \lim_{\widetilde{\varepsilon}\to 0^+} \limsup_{n\to\infty} \frac{1}{n} \ln S(\widetilde{d}_n,\widetilde{\varepsilon},f,T,n) = P_{\text{top}}(\widetilde{d},T,f), \end{aligned} \tag{3.7}$$

as claimed.

□

3.1.6 Lemma: Estimation of capacities

Let (X,d) be some metric space. If $0 < 2\varepsilon \leq \widetilde{\varepsilon}$ and $N \subseteq X$ is maximally (d,ε)-separated, then $S(d,\widetilde{\varepsilon}) \leq |N|$.

Proof: We show that $|M| \leq |N|$ for every $(d,\widetilde{\varepsilon})$-separated $M \subseteq X$ by defining an injection $g : M \to N$. For any point $x \in M$ there exists at least one point $y \in N$ such that $d(x,y) < \varepsilon$, because otherwise N would not be maximally (d,ε)-separated. We set $g(x) = y$. Now for different points $x, \widetilde{x} \in M$ one necessarily has $g(x) \neq g(\widetilde{x})$, because otherwise $d(x,\widetilde{x}) < 2\varepsilon \leq \widetilde{\varepsilon}$, a contradiction to the fact that M is $(d,\widetilde{\varepsilon})$-separated. Thus $g : M \to N$ is injective.

□

3.1.7 Lemma: Estimation of weighted capacities

Let (X,d) be a compact metric space and (X,T) an iterated dynamical system with Bowen-metrics $(d_n)_n$. Let $f \in \mathcal{C}(X)$ be some function with continuity module ω_f (see def. A.2.5). Then for $0 < 2\varepsilon \leq \widetilde{\varepsilon}$ and any maximally (d_n,ε)-separated subset $N_{n,\varepsilon} \subseteq X$ the inequality

$$S(d_n,\widetilde{\varepsilon},f,T,n) \leq e^{n\omega_f(\varepsilon)} \cdot \sum_{y\in N_{n,\varepsilon}} \exp\left[\sum_{k=0}^{n-1} f(T^k y)\right] \tag{3.8}$$

holds. Statement 3.1.6 thus becomes a special case for $f = 0$.

Proof: Note $\Sigma_T^n f := \sum_{k=0}^{n-1} f \circ T^k$. Let $M_{n,\widetilde{\varepsilon}} \subseteq X$ be some arbitrary $(d_n,\widetilde{\varepsilon})$-separated subset. Then just as in 3.1.6, one finds that there exists an injection $g : M_{n,\widetilde{\varepsilon}} \to N_{n,\varepsilon}$ such that $d_n(x,g(x)) < \varepsilon$ for all $x \in M_{n,\widetilde{\varepsilon}}$, meaning that $d(T^k x, T^k g(x)) < \varepsilon$ for all $x \in M_{n,\widetilde{\varepsilon}}$ and $k \in \{0,..,n-1\}$. Thus $\left|f(T^k x) - f(T^k g(x))\right| \leq \omega_f(\varepsilon)$

and consequently $|(\Sigma_T^n f)(x) - (\Sigma_T^n f)(g(x))| \leq n \cdot \omega_f(\varepsilon)$. Thus

$$\begin{aligned}
\sum_{x\in M_{n,\widetilde{\varepsilon}}} e^{\Sigma_T^n f(x)} &= \sum_{g(x)\in g(M_{n,\widetilde{\varepsilon}})} e^{(\Sigma_T^n f)(g(x))+(\Sigma_T^n f)(x)-(\Sigma_T^n f)(g(x))} \\
&\leq \sum_{y\in g(M_{n,\widetilde{\varepsilon}})} e^{(\Sigma_T^n f)(y)} \cdot e^{n\omega_f(\varepsilon)} \\
&\leq e^{n\omega_f(\varepsilon)} \cdot \sum_{y\in g(M_{n,\widetilde{\varepsilon}})} e^{(\Sigma_T^n f)(y)} + e^{n\omega_f(\varepsilon)} \cdot \sum_{y\in N_{n,\varepsilon}\setminus g(M_{n,\widetilde{\varepsilon}})} e^{\Sigma_T^n f(x)} \\
&= e^{n\omega_f(\varepsilon)} \cdot \sum_{y\in N_{n,\varepsilon}} e^{\Sigma_T^n f(y)},
\end{aligned} \tag{3.9}$$

as claimed.

□

3.1.8 Theorem: Characterization of the topological pressure

Let (X,d) be a compact metric space and (X,T) an iterated dynamical system with n-th Bowen metric d_n. For $n\in\mathbb{N}$ and $\varepsilon>0$ let $N_{n,\varepsilon}\subseteq X$ be arbitrary, maximally (d_n,ε)-separated sets. Then the topological pressure of any function $f\in\mathcal{C}(X)$ takes the form

$$P_{\text{top}}(T,f) = \lim_{\varepsilon\to 0^+} \limsup_{n\to\infty} \frac{1}{n} \ln \sum_{x\in N_{n,\varepsilon}} \exp\left[\sum_{k=0}^{n-1} f(T^k x)\right]. \tag{3.10}$$

Proof: Note $\Sigma_T^n f := \sum_{k=0}^{n-1} f\circ T^k$. Clearly inequality "$\geq$" holds in (3.10) as $S(d_n,\varepsilon,f,T,n) \geq \sum_{x\in N_{n,\varepsilon}} e^{\Sigma_T^n f(x)}$ for all $\varepsilon>0$ and $n\in\mathbb{N}$. We shall show that for any $\varepsilon>0$ the inequality

$$\limsup_{n\to\infty} \frac{1}{n} \ln S(d_n, 2\varepsilon, f, T, n) \leq \omega_f(\varepsilon) + \limsup_{n\to\infty} \frac{1}{n} \ln \sum_{x\in N_{n,\varepsilon}} e^{\Sigma_T^n f(x)} \tag{3.11}$$

holds, with ω_f as the continuity module of f. By taking the limit $\varepsilon\to 0^+$ this would readily imply inequality "$\leq$" in (3.10), as f is uniformly continuous and thus $\omega_f(\varepsilon) \xrightarrow{\varepsilon\to 0^+} 0$. But (3.11) follows directly from lemma 3.1.7 applied to $\widetilde{\varepsilon} := 2\varepsilon$.

□

3.1.9 Lemma: Properties of the topological pressure

Let (X,d) be a compact metric space and (X,T) an iterated dynamical system. Then the topological pressure function $P(T,\cdot):\mathcal{C}(X)\to\mathbb{R}$ satisfies:

1. Positiveness, that is $0\leq P_{\text{top}}(T,f)$ for any $0\leq f\in\mathcal{C}(X)$.
2. Lipschitz continuity with Lipschitz constant 1.
3. Convexity.
4. Sub-additivity.
5. $\inf(f) + h_{\text{top}}(T) \leq P_{\text{top}}(T,f) \leq \sup(f) + h_{\text{top}}(T)$ for any $f\in\mathcal{C}(X)$.
6. $P_{\text{top}}(T,f)$ is either always finite or always infinite, corresponding to the cases $h_{\text{top}}(T)<\infty$ and $h_{\text{top}}(T)=\infty$.

Proof:

1. As $\ln \exp[\Sigma_T^n f(x)] \geq 0$ for any positive $0 \leq f \in \mathcal{C}(X)$, the topological pressure is clearly positive.
2. Choose for every $n \in \mathbb{N}$, $\varepsilon > 0$ some maximally (d_n, ε)-separated $N_{n,\varepsilon} \subseteq X$, then by theorem 3.1.8

$$|P_{\mathrm{top}}(T,f) - P_{\mathrm{top}}(T,g)| \leq \lim_{\varepsilon \to 0^+} \limsup_{n\to\infty} \frac{1}{n} \left| \ln \sum_{x \in N_{n,\varepsilon}} e^{\Sigma_T^n f(x)} - \ln \sum_{y \in N_{n,\varepsilon}} e^{\Sigma_T^n g(y)} \right| \tag{3.12}$$

for any $f, g \in \mathcal{C}(X)$. Furthermore

$$\begin{aligned} \left| \ln \sum_{x \in N_{n,\varepsilon}} e^{\Sigma_T^n f(x)} - \ln \sum_{y \in N_{n,\varepsilon}} e^{\Sigma_T^n g(y)} \right| &= \left| \ln \sum_{x \in N_{n,\varepsilon}} e^{\Sigma_T^n g(x)} e^{\Sigma_T^n (f-g)(x)} - \ln \sum_{y \in N_{n,\varepsilon}} e^{\Sigma_T^n g(y)} \right| \\ &\leq \ln e^{\|\Sigma_T^n (f-g)\|_\infty} \leq n \cdot \|f-g\|_\infty , \end{aligned} \tag{3.13}$$

so that (3.12) implies $|P_{\mathrm{top}}(T,f) - P_{\mathrm{top}}(T,g)| \leq \|f-g\|_\infty$.
3. Let $\lambda \in [0,1]$, then by Hölder's inequality (see A.2.6)

$$\begin{aligned} \frac{1}{n} \ln \sum_{x \in N_{n,\varepsilon}} e^{\Sigma_T^n (\lambda f + (1-\lambda) g)(x)} &= \frac{1}{n} \ln \sum_{x \in N_{n,\varepsilon}} \left[e^{\Sigma_T^n f(x)} \right]^\lambda \cdot \left[e^{\Sigma_T^n g(x)} \right]^{(1-\lambda)} \\ &\overset{\text{Hölder}}{\leq} \frac{1}{n} \ln \left[\sum_{x \in N_{n,\varepsilon}} e^{\Sigma_T^n f(x)} \right]^\lambda \left[\sum_{y \in N_{n,\varepsilon}} e^{\Sigma_T^n g(y)} \right]^{(1-\lambda)} \\ &= \frac{\lambda}{n} \cdot \ln \sum_{x \in N_{n,\varepsilon}} e^{\Sigma_T^n f(x)} + \frac{(1-\lambda)}{n} \cdot \ln \sum_{y \in N_{n,\varepsilon}} e^{\Sigma_T^n g(y)} . \end{aligned} \tag{3.14}$$

By theorem 3.1.8 thus

$$P_{\mathrm{top}}(T, \lambda f + (1-\lambda) g) \leq \lambda P_{\mathrm{top}}(T,f) + (1-\lambda) P_{\mathrm{top}}(T,g). \tag{3.15}$$

4. For $f, g \in \mathcal{C}(X)$ one has

$$\begin{aligned} \frac{1}{n} \ln \sum_{x \in N_{n,\varepsilon}} e^{\Sigma_T^n (f+g)(x)} &\leq \frac{1}{n} \ln \sum_{x \in N_{n,\varepsilon}} e^{\Sigma_T^n f(x)} \sum_{y \in N_{n,\varepsilon}} e^{\Sigma_T^n g(y)} \\ &= \frac{1}{n} \ln \sum_{x \in N_{n,\varepsilon}} e^{\Sigma_T^n f(x)} + \frac{1}{n} \ln \sum_{y \in N_{n,\varepsilon}} e^{\Sigma_T^n g(y)} , \end{aligned} \tag{3.16}$$

which implies the sub-additivity of $P_{\mathrm{top}}(T, \cdot)$.
5. Follows readily from remark 3.1.4(v).
6. Follows from (5).

□

3.1.10 Proposition: Topological pressure of higher order dynamics

Let (X,d) be a compact metric space and (X,T) an iterated dynamical system. Then

$$P_{\mathrm{top}}(T^m, \Sigma_T^m f) = m \cdot P_{\mathrm{top}}(T,f) \tag{3.17}$$

for any $m \in \mathbb{N}_0$ and $f \in \mathcal{C}(X)$, with $\Sigma_T^m f := \sum_{k=0}^{m-1} f \circ T^k$.

Proof: Let d_m be the m-th Bowen-metric of (X,T) with respect to the metric d. By remark 3.1.1(ii) the metrics d and $\widetilde{d} := d_m$ are equivalent. By theorem 3.1.5 it suffices to show the equality of $P_{\text{top}}(T^m, \Sigma_T^m f)$ and $P_{\text{top}}(T,f)$ taken with respect to the metrics $\widetilde{d}$ and d respectively. Let $\widetilde{d}_n$ be the n-th Bowen-metric of the system (X,T^m) with respect to the metric $\widetilde{d}$, then $\widetilde{d}_n = d_{nm}$. Similarly

$$\Sigma_{T^m}^n \Sigma_T^m f = \sum_{k=0}^{n-1} (\Sigma_T^m f)\circ T^{mk} = \sum_{k=0}^{n-1}\sum_{l=0}^{m-1} f\circ T^{l+mk} = \Sigma_T^{nm} f, \tag{3.18}$$

so that $S(\widetilde{d}_n, \varepsilon, \Sigma_T^m f, T^m, n) = S(d_{nm}, \varepsilon, f, T, nm)$. Consequently

$$\begin{aligned} P_{\text{top}}(T^m, \Sigma_T^m f) &= \lim_{\varepsilon\to 0^+} \limsup_{n\to\infty} \frac{1}{n} \ln S(\widetilde{d}_n, \varepsilon, \Sigma_T^m f, T^m, n) \\ &= m\cdot \lim_{\varepsilon\to 0^+} \limsup_{n\to\infty} \frac{1}{nm} \ln S(d_{nm}, \varepsilon, f, T, nm) \\ &= m\cdot P_{\text{top}}(T,f). \end{aligned} \tag{3.19}$$

□

3.2 Topological entropy

In the following we shall consider an iterated dynamical system (X,T) on a compact metric space (X,d) with d_n as its n-th Bowen-metric, defined in 3.1.1. We have introduced in 3.1.4 the topological entropy of such a system as topological pressure $P_{\text{top}}(T,0)$ of the zero-potential, taking the form of the *infinitesimal asymptotic growth* rate of capacities

$$h_{\text{top}}(T) := \lim_{\varepsilon\to 0^+} \limsup_{n\to\infty} \frac{1}{n} \ln S(d_n, \varepsilon) \tag{3.20}$$

of the space. As such, it is by theorem 3.1.5 a topological invariant of the system (X,T) and equal for continuously conjugated, iterated dynamical systems.

If we consider two points $x,y\in X$ as *distinguishable* if and only if $d(x,y)\geq\varepsilon$ for some given $\varepsilon>0$, then $d_n(x,y)\geq\varepsilon$ if and only if the orbits of x,y can at some point of time in $\{0,..,n-1\}$ be distinguished.

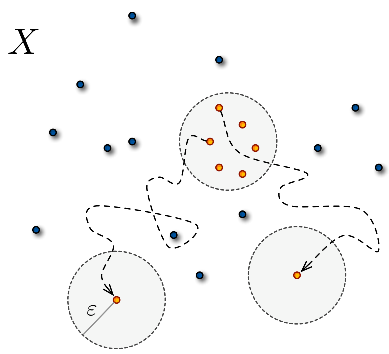

Figure 3.1: On the interpretation of the Bowen-metric and the capacity $S(d_n,\varepsilon)$. Initially undistinguishable points become along their orbits distinguishable if their Bowen-distance is big enough.

$S(d_n,\varepsilon)$ gives thus the maximum number of orbits, all pairwise distinguishable within some time between 0 and $n-1$. The topological entropy can therefore be interpreted as asymptotic growth rate of the maximum number of distinguishable orbits with time in the limit $\varepsilon\to 0^+$. It is a measure for the *dispersal* of *nearby* orbits and thus for the complexity of the system, uniquely defined by the topological nature of the underlying dynamics. Theorem 3.5.5 and corollary 3.5.6 provide with some further interpretations for an important class of iterated dynamical systems, so called *expansive* ones. Example 3.5.10 gives the topological entropy for affine transformations on the torus.

3.2.1 Theorem: Topological entropy as a topological invariant

Let (X, d) be a compact metric space and (X, T) an iterated dynamical system. Then its topological entropy is independent of the metric generating the topology of the space. In particular, continuously conjugated, iterated dynamical systems have equal topological entropy.

Proof: The theorem is simply a special case of the topological invariance of topological pressures, as stated in 3.1.5. We shall nonetheless give a *light* version of the same proof to clarify the underlying ideas. Let $\widetilde{d}$ be a topologically equivalent metric to d. Let $h_{\text{top}}(d, T)$ and $h_{\text{top}}(\widetilde{d}, T)$ be the topological entropies of the system defined with respect to the metric d and $\widetilde{d}$ respectively. It suffices to show that $h_{\text{top}}(d, T) \leq h_{\text{top}}(\widetilde{d}, T)$. As the space is compact, by A.2.4 the metrics $d, \widetilde{d}$ are uniformly equivalent, which means that for any given $\varepsilon > 0$ there exists an $\widetilde{\varepsilon}_\varepsilon > 0$ so that $d(x, y) \geq \varepsilon$ implies $\widetilde{d}(x, y) \geq \widetilde{\varepsilon}_\varepsilon$. This relation is inherited directly by the Bowen-metrics d_n and $\widetilde{d}_n$, with the same $\widetilde{\varepsilon}_\varepsilon$. Thus every (d_n, ε)-separated set is also a $(\widetilde{d}_n, \widetilde{\varepsilon}_\varepsilon)$-separated one, meaning that $S(d_n, \varepsilon) \leq S(\widetilde{d}_n, \widetilde{\varepsilon}_\varepsilon)$. As this holds for all $n \in \mathbb{N}$, one finds that

$$\limsup_{n\to\infty} \frac{1}{n} \ln S(d_n, \varepsilon) \leq \limsup_{n\to\infty} \frac{1}{n} \ln S(\widetilde{d}_n, \widetilde{\varepsilon}_\varepsilon). \tag{3.21}$$

Now, since $\widetilde{\varepsilon}_\varepsilon$ can be chosen so that $\widetilde{\varepsilon}_\varepsilon \xrightarrow{\varepsilon\to 0} 0$, (3.21) implies

$$h_{\text{top}}(d, T) = \lim_{\varepsilon\to 0^+} \limsup_{n\to\infty} \frac{1}{n} \ln S(d_n, \varepsilon) \leq \lim_{\widetilde{\varepsilon}\to 0^+} \limsup_{n\to\infty} \frac{1}{n} \ln S(\widetilde{d}_n, \widetilde{\varepsilon}) = h_{\text{top}}(\widetilde{d}, T), \tag{3.22}$$

as claimed.

□

3.2.2 Theorem: Characterization of topological entropy

Let (X, d) be a compact metric space and (X, T) an iterated dynamical system. For $n \in \mathbb{N}$ and $\varepsilon > 0$ let $N_{n,\varepsilon} \subseteq X$ be arbitrary, maximally (d_n, ε)-separated sets. Then

$$h_{\text{top}}(T) = \lim_{\varepsilon\to 0^+} \limsup_{n\to\infty} \frac{1}{n} \ln |N_{n,\varepsilon}| . \tag{3.23}$$

Proof: Clearly inequality "$\geq$" holds in (3.23) as $S(d_n, \varepsilon) \geq |N_{n,\varepsilon}|$ for every $n \in \mathbb{N}, \varepsilon > 0$. By lemma 3.1.6, for each $\varepsilon > 0$ and $n \in \mathbb{N}$ one has $S(d_n, \varepsilon) \leq \left|N_{n,\varepsilon/2}\right|$. Thus

$$\begin{aligned} h_{\text{top}}(T) = \lim_{\varepsilon\to 0^+} \limsup_{n\to\infty} \frac{1}{n} \ln S(d_n, \varepsilon) &\leq \lim_{\varepsilon\to 0^+} \limsup_{n\to\infty} \frac{1}{n} \ln \left|N_{n,\varepsilon/2}\right| \\ &= \lim_{\varepsilon\to 0^+} \limsup_{n\to\infty} \frac{1}{n} \ln |N_{n,\varepsilon}| \end{aligned} \tag{3.24}$$

which was to be shown.

□

3.2.3 Proposition: Topological entropy of isometries

Let (X, d) be a compact metric space and (X, T) an iterated dynamical system, such that $T : X \to X$ is an isometry of the space. Then its topological entropy is zero.

Proof: As T is an isometry, all Bowen-metrics equal the intrinsic space metric. Thus, the capacities $S(d_n, \varepsilon)$ stay constant over n, so that $h_{\text{top}}(T) = 0$.

□

3.2.4 Proposition: Topological entropy of higher order dynamics

Let (X, d) be a compact metric space and (X, T) an iterated dynamical system. Then $h_{\text{top}}(T^m) = m \cdot h_{\text{top}}(T)$ for any $m \in \mathbb{N}_0$.

Proof: The proof is merely a special case of 3.1.10. We present it nonetheless to clarify some concepts. Let d_m be the m-th Bowen-metric of (X,T) with respect to the metric d. By remark 3.1.1(ii) the metrics d and $\widetilde{d} := d_m$ are equivalent. By theorem 3.2.1 it suffices to show the equality of $h_{\text{top}}(T^m)$ and $h_{\text{top}}(T)$ taken with respect to the metrics $\widetilde{d}$ and d respectively. Let $\widetilde{d}_n$ be the n-th Bowen-metric of the system (X,T^m) with respect to the metric $\widetilde{d}$. Then $\widetilde{d}_n = d_{nm}$, thus $S(\widetilde{d}_n,\varepsilon) = S(d_{nm},\varepsilon)$. Consequently

$$h_{\text{top}}(T^m) = \lim_{\varepsilon\to 0^+}\limsup_{n\to\infty}\frac{1}{n}\ln S(\widetilde{d}_n,\varepsilon) = m\cdot\lim_{\varepsilon\to 0^+}\limsup_{n\to\infty}\frac{1}{nm}\ln S(d_{nm},\varepsilon) = m\cdot h_{\text{top}}(T). \tag{3.25}$$

□

3.2.5 Proposition: Topological entropy of periodic dynamics

Let (X,d) be a compact metric space and (X,T) an iterated dynamical system such that T is periodic, that is, $T^m = \text{Id}$ for some $m\in\mathbb{N}$. Then the topological entropy $h_{\text{top}}(T)$ is zero.

Proof: Using 3.2.4 we obtain $h_{\text{top}}(T) = h_{\text{top}}(T^{m+1}) = (m+1)\cdot h_{\text{top}}(T)$, implying $h_{\text{top}}(T)=0$.

□

3.3 Connecting topological pressure to the Kolmogorov-Sinai entropy

In sections 2.2.1 and 3.1.4 two different kinds of *entropies* were introduced for iterated dynamical systems. The Kolmogorov-Sinai entropy requires solely a probability space and a measure-preserving map on it. The topological entropy (or pressure for that matter) only required a compact metric space with a continuous generator of the dynamic and was defined through the infinitesimal, asymptotic growth rate with time of the number of distinguishable orbits. In this section we shall present the so called **Variational Principle**, which connects the two concepts for probability measures defined on the Borel-σ-algebra of the metric space. It was proven for the zero-potential around 1970 by [6, 7, 8] and [9].

If $\mathscr{B}$ is the Borel-σ-algebra of the underlying metric space (X,d), let $\mathfrak{M}(\mathscr{B})$ denote the convex set of all Borel probability measures on $(X,\mathscr{B})$. Recall that as X is compact, $\mathfrak{M}(\mathscr{B})$ is weakly* sequentially compact. Let $\mathfrak{M}(\mathscr{B},T)\subseteq\mathfrak{M}(\mathscr{B})$ denote the set of all T-invariant Borel probability measures on $(X,\mathscr{B})$ for some $T: X\to X$. Then if T is continuous, $\mathfrak{M}(\mathscr{B},T)$ is sequentially closed in the weak* topology and thus its self sequentially compact.

3.3.1 Lemma: Approximation of measure-theoretic entropies

Let (X,d) be a compact metric space with Borel-σ-algebra $\mathscr{B}$ and $(X,\mathscr{B},\mu)$ a probability space. Let $\mathcal{A}\in\mathcal{Z}_1(\mathscr{B},\mu)$ be a countable (finite), measurable partition of the space with finite entropy. Let $\varepsilon>0$ be given. Then there exists a countable (finite), measurable partition $\mathcal{B}=\{B_0,B_1,B_2,..\}\in\mathcal{Z}_1(\mathscr{B},\mu)$ such that:

- All $B_1,B_2,..$ are compacts.
- $|H_\mu(\mathcal{A})-H_\mu(\mathcal{B})|\le\varepsilon$.
- $H_\mu(\mathcal{B}|\mathcal{A})\le\varepsilon$ and $H_\mu(\mathcal{A}|\mathcal{B})\le\varepsilon$.

Proof: We shall restrict our selfs to the countable case, the finite one is proven in the same way. Let $\mathcal{A}=\{A_1,A_2,..\}$. As X is compact, the measure μ is regular. Also recall that $x\ln\frac{1}{x}\overset{x\to 0^+}{\longrightarrow}0$. Thus for every $n\in\mathbb{N}$ there exists some compact $B_n\subseteq A_n$ such that

$$\mu(A_n\setminus B_n)\cdot\ln\frac{1}{\mu(A_n\setminus B_n)}\le\frac{\varepsilon}{2}\cdot 2^{-n} \tag{3.26}$$

and

$$\mu(A_n\setminus B_n)\cdot\ln\frac{1}{\mu(B_n)}\le\frac{\varepsilon}{6}\cdot 2^{-n} \tag{3.27}$$

and

$$\mu(A_n)\cdot\ln\frac{\mu(A_n)}{\mu(B_n)}\le\frac{\varepsilon}{6}\cdot 2^{-n}. \tag{3.28}$$

Let $B_0 := X \setminus \bigcup_{n\in\mathbb{N}} B_n$, then $\mu(B_0) = \sum_n \mu(A_n \setminus B_n)$. We can thus even assume that $\mu(B_0)\ln\frac{1}{\mu(B_0)} \leq \varepsilon/6$. Now

$$H_\mu(\mathcal{A}|\mathcal{B}) = -\mu(B_0)\sum_{m\in\mathbb{N}}\mu(A_m|B_0)\ln\mu(A_m|B_0) - \sum_{n\in\mathbb{N}}\mu(B_n)\underbrace{\sum_{m\in\mathbb{N}}\underbrace{\mu(A_m|B_n)}_{\delta_{nm}}\ln\mu(A_m|B_n)}_{0}$$

$$= -\sum_{m\in\mathbb{N}}\underbrace{\mu(A_m\cap B_0)}_{\mu(A_m\setminus B_m)}\ln\underbrace{\mu(A_m|B_0)}_{\frac{\mu(A_m\setminus B_m)}{\mu(B_0)}} = \sum_{m\in\mathbb{N}}\mu(A_m\setminus B_m)\ln\frac{\mu(B_0)}{\mu(A_m\setminus B_m)} \tag{3.29}$$

$$\leq \sum_{m\in\mathbb{N}}\mu(A_m\setminus B_m)\ln\frac{1}{\mu(A_m\setminus B_m)} \overset{(3.26)}{\leq} \frac{\varepsilon}{2}\cdot\sum_{m\in\mathbb{N}}2^{-n} = \frac{\varepsilon}{2}.$$

On the other hand

$$|H_\mu(\mathcal{A}) - H_\mu(\mathcal{B})| \leq |\mu(B_0)\ln\mu(B_0)| + \sum_{n=1}^{\infty}\Big|\mu(A_n)\left[\ln\mu(A_n) - \ln\mu(B_n)\right]\Big| + \sum_{n=1}^{\infty}\Big|\left[\mu(A_n) - \mu(B_n)\right]\cdot\ln\mu(B_n)\Big|$$

$$= \underbrace{\mu(B_0)\ln\frac{1}{\mu(B_0)}}_{\leq\frac{\varepsilon}{6}} + \underbrace{\sum_{n=1}^{\infty}\mu(A_n)\ln\frac{\mu(A_n)}{\mu(B_n)}}_{\leq\frac{\varepsilon}{6}\text{ by }(3.28)} + \underbrace{\sum_{n=1}^{\infty}\mu(A_n\setminus B_n)\cdot\ln\frac{1}{\mu(B_n)}}_{\leq\frac{\varepsilon}{6}\text{ by }(3.27)}$$

$$\leq \frac{\varepsilon}{2}. \tag{3.30}$$

Finally, by 2.1.4(9) $H_\mu(\mathcal{B}|\mathcal{A}) \leq \varepsilon$.

□

3.3.2 Lemma: Existence of fine partitions

Let (X,d) be a compact metric space with Borel-σ-algebra $\mathscr{B}$ and $(X,\mathscr{B},\mu)$ a probability space. Then for every $\varepsilon > 0$ there exists a finite, measurable partition $\mathcal{A}$ of $(X,\mathscr{B})$ such that $\operatorname{diam}(A) \leq \varepsilon$ and $\mu(\partial A) = 0$ for every $A \in \mathcal{A}$.

Proof: We present the proof given in [2]. Let $\varepsilon > 0$ be given, then by total boundedness of X there exist points $x_1, ..., x_n \in X$ such that $X = \bigcup_{i=1}^n B(x_i, \varepsilon/4)$. For every fixed $i \in \{1,..,n\}$ the sets $\{x : d(x,x_i) = r\}$, $\varepsilon/4 < r < \varepsilon/2$ are closed and disjoint, so that only countably many of them can have positive measure. We can thus choose an $\varepsilon/4 < r_0 < \varepsilon/2$ such that

$$\mu\left(\{x \in X : d(x,x_i) = r_0\}\right) = 0 \tag{3.31}$$

for every $i \in \{1,..,n\}$. Set $A_1 := \{x \in X : d(x,x_i) \leq r_0\}$ and inductively

$$A_i := \{x \in X : d(x,x_i) \leq r_0\} \setminus \bigcup_{k=1}^{i-1} A_i \tag{3.32}$$

for $i \in \{2,..,n\}$. Then $\mathcal{A} := \{A_1,..,A_n\}$ is a partition of X satisfying $\operatorname{diam}(A_i) \leq \varepsilon$ for all $i \in \{1,..,n\}$. Furthermore, $\mu(\partial A_i) = 0$ for every $i \in \{1,..,n\}$ as generally $\partial(A \setminus B) \subseteq \partial A \cup \partial B$ for any subsets $A, B \subseteq X$.

□

3.3.3 Theorem: The Variational Principle

Let (X,d) be a compact metric space and (X,T) a continuous, iterated dynamical system. Let $\mathscr{B}$ be the Borel-σ-algebra on X and $\mathfrak{M}(\mathscr{B},T)$ the system of all T-invariant probability measures on $(X,\mathscr{B})$. Then the pressure of any continuous $f \in \mathcal{C}(X)$ is given by

$$\boxed{P_{\text{top}}(T,f) = \sup\Big\{h_\mu(T) + \mathbb{E}_\mu f : \mu \in \mathfrak{M}(\mathscr{B},T)\Big\}.} \tag{3.33}$$

In particular, is the topological entropy given by the supremum

$$h_{\text{top}}(T) = \sup\Big\{h_\mu(T) : \mu \in \mathfrak{M}(\mathscr{B}, T)\Big\}. \tag{3.34}$$

Proof: We shall combine the proofs found in [1] and [2]. We shall in the first step show that $P_{\text{top}}(T, f) \geq h_\mu(T) + \mathbb{E}_\mu f$ for any $\mu \in \mathfrak{M}(\mathscr{B}, T)$. Let $\mathcal{A} \in \mathcal{Z}_1(\mathscr{B}, \mu)$ be a finite, measurable partition. Let $\varepsilon > 0$ be given and choose $\mathcal{B} = \{B_0, .., B_s\} \in \mathcal{Z}_1(\mathscr{B}, \mu)$ as described in lemma 3.3.1, that is with $B_1, .., B_s$ as compacts and $H_\mu(\mathcal{A}|\mathcal{B}) \leq 1$. Note $\mathcal{A}^{(n)} := \bigvee_{k=0}^{n-1} T^{-k}\mathcal{A}$ and $\mathcal{B}^{(n)} := \bigvee_{k=0}^{n-1} T^{-k}\mathcal{B}$ for every $n \in \mathbb{N}$. Then by lemma 2.2.5 $H_\mu(\mathcal{A}^{(n)}) \leq H_\mu(\mathcal{B}^{(n)}) + n$. The T-invariance of μ implies

$$\begin{aligned}
\frac{1}{n}H_\mu(\mathcal{A}^{(n)}) + \int f\ d\mu &\leq 1 + \frac{1}{n}H_\mu(\mathcal{B}^{(n)}) + \int f\ d\mu \\
&= 1 - \frac{1}{n}\sum_{C\in\mathcal{B}^{(n)}} \mu(C)\ln\mu(C) + \frac{1}{n}\sum_{k=0}^{n-1}\int f\circ T^k\ d\mu \\
&= 1 + \frac{1}{n}\sum_{C\in\mathcal{B}^{(n)}} \int_C \Sigma_T^n f\ d\mu - \mu(C)\ln\mu(C) \\
&= 1 + \frac{1}{n}\sum_{C\in\mathcal{B}^{(n)}} \mu(C)\ln\left[\frac{1}{\mu(C)}\exp\left[\frac{1}{\mu(C)}\int_C \Sigma_T^n f\ d\mu\right]\right],
\end{aligned} \tag{3.35}$$

which by concavity of $\ln(\cdot)$ and Jensen's inequality A.2.7 implies

$$\begin{aligned}
\frac{1}{n}H_\mu(\mathcal{A}^{(n)}) + \int f\ d\mu &\overset{\text{Jensen}}{\leq} 1 + \frac{1}{n}\ln\sum_{C\in\mathcal{B}^{(n)}}\exp\left[\frac{1}{\mu(C)}\int_C \Sigma_T^n f\ d\mu\right] \\
&\leq 1 + \frac{1}{n}\ln\sum_{C\in\mathcal{B}^{(n)}}\exp\left[\sup_{x\in C}\Sigma_T^n f(x)\right].
\end{aligned} \tag{3.36}$$

As X is compact, the closure $\overline{C}$ of every $C \in \mathcal{B}^{(n)}$ is compact as well. By continuity of f, we can thus choose some $x_{C,n} \in \overline{C}$ such that $\Sigma_T^n f(x_{C,n}) = \sup_{x\in C}\Sigma_T^n f(x)$. Now choose $\delta > 0$ such that[2]

$$\delta < \frac{1}{2}\inf\{d(B_k, B_l) : 1 \leq k \neq l \leq s\} \tag{3.37}$$

and $d(x, y) \leq \delta$ implies $|f(x) - f(y)| \leq 1$. For each $n \in \mathbb{N}$, let $N_{n,\delta} \subseteq X$ be some maximally (d_n, δ)-separated set, with d_n as n-th Bowen metric of the system. Then for each $n \in \mathbb{N}$ and $C \in \mathcal{B}^{(n)}$ there exists some $y_{C,n} \in N_{n,\delta}$ with $d_n(x_{C,n}, y_{C,n}) < \delta$ and thus by (3.36)

$$\frac{1}{n}H_\mu(\mathcal{A}^{(n)}) + \int f\ d\mu \leq 1 + \frac{1}{n}\ln\sum_{C\in\mathcal{B}^{(n)}}\exp\left[\Sigma_T^n f(y_{C,n}) + n\right]. \tag{3.38}$$

Taking the limit $\limsup_{n\to\infty}$ in (3.38) and applying the Shannon-McMillan-Breiman theorem 2.2.3, one obtains

$$h_\mu(T, \mathcal{A}) + \int f\ d\mu \leq \limsup_{n\to\infty}\frac{1}{n}\ln\sum_{C\in\mathcal{B}^{(n)}} e^{\Sigma_T^n f(y_{C,n})} + 2. \tag{3.39}$$

Now let $M(n, \delta, \mathcal{B})$ be an upper bound for the cardinality of pre-images of the mapping $\mathcal{B}^{(n)} \to N_{n,\delta}$, $C \mapsto y_{C,n}$. Then

$$\sum_{C\in\mathcal{B}^{(n)}} e^{\Sigma_T^n f(y_{C,n})} \leq \sum_{y\in N_{n,\delta}} e^{\Sigma_T^n f(y)} \cdot M(n, \delta, \mathcal{B}) \tag{3.40}$$

and (3.39) takes the form

$$\begin{aligned}
h_\mu(T, \mathcal{A}) + \int f\ d\mu &\leq 2 + \limsup_{n\to\infty}\frac{1}{n}\ln\underbrace{\sum_{y\in N_{n,\delta}} e^{\Sigma_T^n f(y)}}_{\leq S(d_n,\delta,f,T,n)} + \limsup_{n\to\infty}\frac{1}{n}\ln M(n, \delta, \mathcal{B}) \\
&\leq 2 + P_{\text{top}}(T, f) + \limsup_{n\to\infty}\frac{1}{n}\ln M(n, \delta, \mathcal{B}).
\end{aligned} \tag{3.41}$$

[2]Recall that all $B_1, B_2, ..$ are compact and pairwise disjoint.

Claim: $M(n,\delta,\mathcal{B}) \leq 2^n$.

Proof: It suffices to show that

$$\left|\left\{C \in \mathcal{B}^{(n)} : \overline{C} \cap B^o_{d_n}(x,\delta) \neq \emptyset\right\}\right| \leq 2^n. \tag{3.42}$$

for any $x \in X$. Now by choice of δ, one has by (3.37) for each $y \in X$

$$\left|\left\{C \in \mathcal{B} : \overline{C} \cap B^o_d(y,\delta) \neq \emptyset\right\}\right| \leq 2, \tag{3.43}$$

as any ball $B^o_d(y,\delta)$ can intersect at most one $\overline{B_k}$, $1 \leq k \leq s$ (and perhaps $\overline{B_0}$). Suppose that $\overline{C} \cap B^o_{d_n}(x,\delta) \neq \emptyset$ for some $C \in \mathcal{B}^{(n)}$. Now C is of the form $C = \bigcap_{k=0}^{n-1} T^{-k} C_k$ with $C_0, .., C_{n-1} \in \mathcal{B}$. Furthermore

$$B^o_{d_n}(x,\delta) = \bigcap_{k=0}^{n-1} T^{-k}(B^o_d(T^k x,\delta)) \tag{3.44}$$

so that

$$\emptyset \neq \overline{\bigcap_{k=0}^{n-1} T^{-k}(C_k)} \bigcap \bigcap_{k=0}^{n-1} T^{-k}(B^o_d(T^k x,\delta)) \subseteq \bigcap_{k=0}^{n-1} T^{-k}\left[\overline{C_k} \cap B^o_d(T^k x,\delta)\right], \tag{3.45}$$

where we have used the fact that T is continuous. This implies that $\overline{C}_k \cap B^o_d(T^k x,\delta) \neq \emptyset$ for every $k \in \{0,..,n-1\}$. By (3.43) there exist for each k at most two $C_k \in \mathcal{B}$ satisfying this, so that there exist overall at most 2^n possible $C \in \mathcal{B}^{(n)}$ whose closure intersects $B^o_{d_n}(x,\delta)$. Therefore $M(n,\delta,\mathcal{B}) \leq 2^n$ as claimed.

Consequently, (3.41) takes the form

$$h_\mu(T,\mathcal{A}) + \int f \, d\mu \leq 2 + P_{\text{top}}(T,f) + \ln 2. \tag{3.46}$$

Recall that (3.46) holds for all finite, measurable partitions $\mathcal{A}$, so that by 2.2.6 it follows

$$h_\mu(T) + \int f \, d\mu \leq 2 + P_{\text{top}}(T,f) + \ln 2. \tag{3.47}$$

Now replacing T with T^m and f with $\Sigma_T^m f$, (3.46) gives

$$h_\mu(T^m) + \int \Sigma_T^m f \, d\mu \leq 2 + P_{\text{top}}(T^m, \Sigma_T^m f) + \ln 2, \tag{3.48}$$

which by T-invariance of μ, lemma 3.1.10 and lemma 2.2.8 implies

$$m \cdot h_\mu(T) + m \cdot \int f \, d\mu \leq 2 + m \cdot P_{\text{top}}(T,f) + \ln 2. \tag{3.49}$$

Dividing by m and taking the limit $m \to \infty$ in (3.48) yields

$$h_\mu(T) + \int f \, d\mu \leq P_{\text{top}}(T,f) \tag{3.50}$$

for any $\mu \in \mathfrak{M}(\mathscr{B},T)$. We shall now show inequality "$\leq$" in (3.33). We shall show that for any $\eta > 0$ one can find a $\nu \in \mathfrak{M}(\mathscr{B},T)$ and $\mathcal{A} \in \mathcal{Z}_1(\mathscr{B},\nu)$ such that $P_{\text{top}}(T,f) - \eta \leq h_\nu(T,\mathcal{A}) + \mathbb{E}_\nu f$.

Let ν be some probability measure on $(X, \mathscr{B})$ (not necessarily T-invariant) and $n, M \in \mathbb{N}$. Then the probability measure $\nu^{(nM)} := \frac{1}{nM}\sum_{k=0}^{nM-1} \nu \circ T^{-k}$ satisfies

$$\begin{aligned}
H_{\nu^{(nM)}}\left(\mathcal{A}^{(M)}\right) &= -\sum_{A\in\mathcal{A}^{(M)}} \frac{1}{nM}\sum_{k=0}^{nM-1} \nu(T^{-k}(A)) \ln \frac{1}{nM}\sum_{k=0}^{nM-1} \nu(T^{-k}(A)) \\
&\overset{(\clubsuit)}{\geq} -\sum_{A\in\mathcal{A}^{(M)}} \frac{1}{nM}\sum_{k=0}^{nM-1} \nu(T^{-k}(A)) \ln \nu(T^{-k}(A)) \\
&= \frac{1}{nM}\sum_{k=0}^{nM-1} H_\nu(T^{-k}\mathcal{A}^{(M)}) = \frac{1}{nM}\sum_{j=0}^{M-1}\sum_{k=0}^{n-1} H_\nu(T^{-j-kM}\mathcal{A}^{(M)}) \\
&\overset{2.1.4(8)}{\geq} \frac{1}{nM}\sum_{j=0}^{M-1} H_\nu\left[\bigvee_{k=0}^{n-1}\bigvee_{l=j+kM}^{j+kM+M-1} T^{-l}\mathcal{A}\right] \\
&= \frac{1}{nM}\sum_{j=0}^{M-1} H_\nu\left[\bigvee_{l=j}^{j+nM-1} T^{-l}\mathcal{A}\right] \overset{2.1.4(3)}{\geq} \frac{1}{nM}\sum_{j=0}^{M-1} H_\nu\left[\bigvee_{l=j}^{nM-1} T^{-l}\mathcal{A}\right] \\
&\overset{(\spadesuit)}{\geq} \frac{1}{nM}\sum_{j=0}^{M-1} \underbrace{H_\nu\left[\bigvee_{l=0}^{nM-1} T^{-l}\mathcal{A}\right]}_{H_\nu(\mathcal{A}^{(nM)})} - \frac{1}{nM}\sum_{j=0}^{M-1} \underbrace{H_\nu\left[\bigvee_{l=0}^{j-1} T^{-l}\mathcal{A}\right]}_{\substack{\leq H_\nu(\mathcal{A}^{(M)}) \\ \text{by } 2.1.4(3)}} \\
&\geq \frac{1}{n}H_\nu\left(\mathcal{A}^{(nM)}\right) - \frac{1}{nM}\sum_{j=0}^{M-1} \underbrace{H_\nu\left(\mathcal{A}^{(M)}\right)}_{\substack{\leq \ln\left|\mathcal{A}^{(M)}\right| \\ \text{by } 2.1.4(5)}} \geq \frac{1}{n}H_\nu(\mathcal{A}^{(nM)}) - \frac{1}{n}\ln \underbrace{\left|\mathcal{A}^{(M)}\right|}_{\leq |\mathcal{A}|^M} \\
&\geq \frac{1}{n}H_\nu(\mathcal{A}^{(nM)}) - \frac{M}{n}\ln|\mathcal{A}|
\end{aligned} \tag{3.51}$$

for any finite, measurable partition $\mathcal{A}$. In step $(\clubsuit)$ we used the concavity of $x \mapsto -x\ln x$ and Jensen's inequality A.2.7 applied to the integral operator $\frac{1}{nM}\sum_{k=0}^{nM-1}(\cdot)$ and integrable function $h(k) := \nu(T^{-k}(A))$. In step $(\spadesuit)$ we used the fact that $H_\nu(\mathcal{B}) \geq H_\nu(\mathcal{B}\vee\mathcal{D}) - H_\nu(\mathcal{D})$ for any $\mathcal{B}, \mathcal{D} \in \mathcal{Z}_1(\mathfrak{S}, \nu)$.
Let $\eta > 0$ be given. By theorem 3.1.8 we can find an $\varepsilon > 0$ and an infinite subset $K \subseteq \mathbb{N}$ together with maximally (d_k, ε)-separated subsets $N_{k,\varepsilon} \subseteq X$ such that

$$\frac{1}{k}\ln\sum_{x\in N_{k,\varepsilon}} e^{\Sigma_T^k f(x)} \geq P_{\text{top}}(T, f) - \eta \tag{3.52}$$

for all $k \in K$. For every $k \in K$ define the probability measure μ_k on $(X, \mathscr{B})$ by

$$\mu_k := \frac{\sum_{x\in N_{k,\varepsilon}} e^{\Sigma_T^k f(x)}\cdot\delta_x}{\sum_{y\in N_{k,\varepsilon}} e^{\Sigma_T^k f(y)}}, \tag{3.53}$$

with δ_x as Dirac-measure at point x. Consider the *time-average* probability measures

$$\widetilde{\mu}_k := \frac{1}{k}\sum_{j=0}^{k-1} \mu_k \circ T^{-j}. \tag{3.54}$$

As X is compact, the set of probability measures $\mathfrak{M}(\mathscr{B})$ is weak* sequentially compact. We can thus suppose the sequence $(\widetilde{\mu}_k)_{k\in K}$ to have a weak* limit probability measure $\nu \in \mathfrak{M}(\mathscr{B})$. From definition (3.54) it is easily verifiable that the sequences $(\widetilde{\mu}_k)_k$ and $(\widetilde{\mu}_k \circ T^{-1})_k$ have same weak* limits, thus by continuity of T the limit ν is T-invariant.
By 3.3.2 we can choose a finite partition $\mathcal{A}$ of $(X, \mathscr{B})$ such that $\operatorname{diam}(A) < \varepsilon$ and $\nu(\partial A) = 0$ for all $A \in \mathcal{A}$. Now fix $n \in \mathbb{N}$ and let $\mathcal{A}^{(n)} := \bigvee_{k=0}^{n-1} T^{-k}\mathcal{A}$. For each $x \in N_{n,\varepsilon}$ let $A_x \in \mathcal{A}^{(n)}$ be the unique atom containing x.
Claim: Each atom $A \in \mathcal{A}^{(n)}$ can contain at most one point of $N_{n,\varepsilon}$.

Proof: Supposing that $x \neq y \in A \cap N_{n,\varepsilon}$, then $d_n(x,y) \geq \varepsilon$, that is $d(T^k x, T^k y) \geq \varepsilon$ for some $k \in \{0,..,n-1\}$. By choice of $\mathcal{A}$ this means that $T^k x$ and $T^k y$ cannot lie in the same atom of $\mathcal{A}$. Now A is of the form $A = \bigcap_{l=0}^{n-1} T^{-l} A_l$ with $A_0,..,A_{n-1} \in \mathcal{A}$, so that $T^k x, T^k y \in T^k(A) \subseteq T^k T^{-k}(A_k) \subseteq A_k \in \mathcal{A}$, a contradiction!

Now fix $M \in \mathbb{N}$. Let $K = \{k_1, k_2, ..\}$ with $k_1 \leq k_2 \leq ..$ and choose for each $l \in \mathbb{N}$ natural numbers $n_l \in \mathbb{N}$ and $0 \leq r_l < M$ such that $k_l = n_l \cdot M + r_l$. Define

$$\mu_{k_l}^{(n_l M)} := \frac{1}{n_l M} \sum_{j=0}^{n_l M - 1} \mu_{k_l} \circ T^{-j}. \tag{3.55}$$

Claim: $\mu_{k_l}^{(n_l M)} \underset{\text{weak*}}{\overset{l\to\infty}{\longrightarrow}} \nu$.

Proof: Notice that $\mu_{k_l}^{(n_l M)} = \frac{k_l}{n_l M} \cdot \widetilde{\mu}_{k_l} - \frac{1}{n_l M} \sum_{j=n_l M}^{k_l - 1} \mu_{k_l} \circ T^{-j}$. As $\frac{k_l}{n_l M} \to 1$ and $\widetilde{\mu}_{k_l} \to \nu$, we need to show that $\frac{1}{n_l M} \sum_{j=n_l M}^{k_l-1} \mu_{k_l} \circ T^{-j} \underset{\text{weak*}}{\overset{l\to\infty}{\longrightarrow}} 0$. Well, for any $g \in \mathcal{C}(X)$

$$\left| \frac{1}{n_l M} \sum_{j=n_l M}^{k_l - 1} \int g \; d(\mu_{k_l} \circ T^{-j}) \right| \leq \frac{1}{n_l M} \sum_{j=n_l M}^{k_l-1} \underbrace{\int \left| g \circ T^j \right| \; d\mu_{k_l}}_{\leq \|g\|_\infty} \leq \|g\|_\infty \cdot \frac{r_l}{n_l M} \overset{l\to\infty}{\longrightarrow} 0 \tag{3.56}$$

which was to be shown.

Since the border of every $A \in \mathcal{A}^{(M)}$ has zero ν-measure it follows $\mu_{k_l}^{(n_l M)}(A) \overset{l\to\infty}{\longrightarrow} \nu(A)$. Since $\mathcal{A}^{(M)}$ is finite, we conclude that $H_{\mu_{k_l}^{(n_l M)}}(\mathcal{A}^{(M)}) \overset{l\to\infty}{\longrightarrow} H_\nu(\mathcal{A}^{(M)})$. Applying (3.51) to $\mu_{k_l}^{(n_l M)}$ we find

$$\begin{aligned}
H_\nu(\mathcal{A}^{(M)}) + M \cdot \int f \; d\nu &= \lim_{l\to\infty} \left[H_{\mu_{k_l}^{(n_l M)}}(\mathcal{A}^{(M)}) + M \cdot \int f \; d\mu_{k_l}^{(n_l M)} \right] \\
&\overset{(3.51)}{\geq} \limsup_{l\to\infty} \left[\frac{1}{n_l} H_{\mu_{k_l}}(\mathcal{A}^{(n_l M)}) - \frac{M}{n_l} \ln|\mathcal{A}| + \frac{M}{n_l M} \cdot \int \Sigma_T^{n_l M} f \; d\mu_{k_l} \right] \\
&\overset{(\clubsuit)}{\geq} \limsup_{l\to\infty} \left[\frac{1}{n_l} H_{\mu_{k_l}}(\mathcal{A}^{(k_l)}) - \frac{2M}{n_l} \ln|\mathcal{A}| + \frac{1}{n_l} \int \Sigma_T^{n_l M} f \; d\mu_{k_l} \right] \\
&\overset{(\spadesuit)}{=} \limsup_{l\to\infty} \left[\frac{1}{n_l} H_{\mu_{k_l}}(\mathcal{A}^{(k_l)}) - \frac{2M}{n_l} \ln|\mathcal{A}| + \frac{1}{n_l} \int \Sigma_T^{k_l} f \; d\mu_{k_l} \right] \\
&= \limsup_{l\to\infty} \frac{1}{n_l} \sum_{A \in \mathcal{A}^{(k_l)}} \underbrace{\mu_{k_l}(A)}_{\substack{0 \text{ if} \\ A \cap N_{k_l,\varepsilon} = \emptyset}} \ln \frac{1}{\mu_{k_l}(A)} \exp\left[\frac{1}{\mu_{k_l}(A)} \int_A \Sigma_T^{k_l} f \; d\mu_{k_l} \right] \\
&\overset{(3.53)}{=} \limsup_{l\to\infty} \frac{1}{n_l} \sum_{\substack{A \in \mathcal{A}^{(k_l)} \\ A \cap N_{k_l,\varepsilon} = \{x_A\} \neq \emptyset}} \mu_{k_l}(A) \ln \frac{\sum_{y \in N_{k_l,\varepsilon}} e^{\Sigma_T^{k_l} f(y)}}{\exp\left[\Sigma_T^{k_l} f(x_A)\right]} \exp\left[\Sigma_T^{k_l} f(x_A)\right] \\
&= \limsup_{l\to\infty} \frac{M}{k_l} \underbrace{\left(1 + \frac{r_l}{n_l M}\right)}_{\overset{l\to\infty}{\longrightarrow} 1} \underbrace{\sum_{A \in \mathcal{A}^{(k_l)}} \mu_{k_l}(A)}_{1} \ln \sum_{y \in N_{k_l,\varepsilon}} e^{\Sigma_T^{k_l} f(y)} \\
&\overset{(3.52)}{\geq} M \cdot (P_{\text{top}}(T,f) - \eta).
\end{aligned} \tag{3.57}$$

Here x_A denotes the unique point in $A \cap N_{k_l,\varepsilon}$, if existing. In step ($\clubsuit$) we used the fact that

$$\begin{aligned}
H_{\mu_{k_l}}(\mathcal{A}^{(n_l M)}) &\geq H_{\mu_{k_l}}(\mathcal{A}^{(k_l)}) - H_{\mu_{k_l}} \left[\bigvee_{j=n_l M}^{k_l-1} T^{-j} \mathcal{A} \right] \\
&\geq H_{\mu_{k_l}}(\mathcal{A}^{(k_l)}) - \ln \underbrace{\left| \bigvee_{j=n_l M}^{k_l-1} T^{-j} \mathcal{A} \right|}_{\leq |\mathcal{A}|^{k_l - n_l M} \leq |\mathcal{A}|^M} \geq H_{\mu_{k_l}}(\mathcal{A}^{(k_l)}) - M \ln |\mathcal{A}|.
\end{aligned} \tag{3.58}$$

In step (♠) we used the fact that

$$\left|\frac{1}{n_l}\int \Sigma_T^{n_l M} f\ d\mu_{k_l} - \frac{1}{n_l}\int \Sigma_T^{k_l} f\ d\mu_{k_l}\right| \leq \frac{1}{n_l}\underbrace{\int \sum_{j=n_l M}^{k_l-1} \left|f\circ T^j\right|\ d\mu_{k_l}}_{\leq r_l\cdot\|f\|_\infty} \overset{l\to\infty}{\longrightarrow} 0. \tag{3.59}$$

Dividing (3.57) by M yields

$$\frac{1}{M}H_\nu(\mathcal{A}^{(M)}) + \int f\ d\nu \geq P_{\text{top}}(T,f) - \eta. \tag{3.60}$$

The limit $M\to\infty$ yields by the Shannon-McMillan-Breiman theorem 2.2.3 the inequality

$$h_\nu(T,\mathcal{A}) + \mathbb{E}_\nu f \geq P_{\text{top}}(T,f) - \eta \tag{3.61}$$

as needed.

□

3.3.4 Corollary: Variational principle for ergodic measures

Let (X,d) be a compact metric space and (X,T) a continuous, iterated dynamical system. Let $\mathfrak{M}_{\text{er}}(\mathscr{B},T)$ be the system of all T-invariant, ergodic Borel-probability measures on X. Then the pressure of any continuous $f\in\mathcal{C}(X)$ is given by

$$P_{\text{top}}(T,f) = \sup\left\{h_\mu(T) + \mathbb{E}_\mu f : \mu\in\mathfrak{M}_{\text{er}}(\mathscr{B},T)\right\}. \tag{3.62}$$

Proof: See [2], corollary 2.4.3.

3.4 Equilibrium measures

The Variational Principle proven in 3.3.3 gives a direct connection between topological pressure and the Kolmogorov-Sinai entropy of an iterated dynamical system on a compact metric space. Strongly related to this principle are so called equilibrium measures, maximizing certain functionals of invariant probability measures. We shall outline here only some basic results stemming from considerations of the pressure function.

3.4.1 Definition: Equilibrium measures

Let (X,d) be a compact metric space with Borel-σ-algebra $\mathscr{B}$ and (X,T) a continuous, iterated dynamical system. A T-invariant probability measure $\mu\in\mathfrak{M}(\mathscr{B},T)$ is called an **equilibrium measure** for the potential $f\in\mathcal{C}(X)$ if it satisfies

$$P_{\text{top}}(T,f) = h_\mu(T) + \mathbb{E}_\mu f. \tag{3.63}$$

For the special case $f=0$ it is also called **maximum entropy measure**. Note that such measures need not always exist[3]! The study of the functional $P_{\text{top}}(T,\cdot):\mathcal{C}(X)\to\mathbb{R}$ gives insight into questions of existence and unicity of equilibrium measures[2].

3.4.2 Proposition: The set of equilibrium measures

Let (X,d) be a compact metric space with Borel-σ-algebra $\mathscr{B}$ and (X,T) a continuous, iterated dynamical system. Let $\mathfrak{M}_{\text{eq}}(\mathscr{B},T,f)\subseteq\mathfrak{M}(\mathscr{B},T)$ denote the set of all equilibrium measures for the potential $f\in\mathcal{C}(X)$. Then:

1. $\mathfrak{M}_{\text{eq}}(\mathscr{B},T,f)$ is a convex set.
2. If $\mathfrak{M}_{\text{eq}}(\mathscr{B},T,f)\neq\emptyset$ then $\mathfrak{M}_{\text{eq}}(\mathscr{B},T,f)$ contains ergodic measures.

[3] See [2], example 2.5.2.

Proof:

1. Follows from the fact that $\mu \mapsto h_\mu(T)$ is an affine mapping on $\mathfrak{M}(\mathscr{B}, T)$ (see 2.2.9).
2. See [2], proposition 2.5.1.

3.4.3 Definition: Tangent functional

Let E be a $\mathbb{K}$-linear space ($\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$) and $\Phi : E \to \mathbb{K}$ convex. A linear functional $a : E \to \mathbb{K}$ is called a **tangent functional** to Φ at $x \in E$ if

$$\Phi(x+y) - \Phi(x) \geq a(y) \tag{3.64}$$

for all $y \in E$.

3.4.4 Proposition over tangent functionals

Let E be a $\mathbb{K}$-Banach space and $\Phi : E \to \mathbb{K}$ convex, Gâteaux-derivable at $x \in E$ with Gâteaux-derivative $d_x\Phi$. Then $d_x\Phi$ is linear if and only if Φ has a tangent functional at x, in which case that functional is exactly $d_x\Phi$.

Proof:

Direction "⇐": Let $a : E \to \mathbb{K}$ be a tangent functional to Φ at x. Then by definition 3.4.3, for arbitrary $y \in E$ one has

$$a(y) = \lim_{\lambda \to 0^+} \frac{a(\lambda y)}{\lambda} \leq \lim_{\lambda \to 0^+} \frac{\Phi(x + \lambda y) - \Phi(x)}{\lambda} = d_x\Phi y, \tag{3.65}$$

showing that $a \leq d_x\Phi$. By $\mathbb{K}$-homogeneity of both sides this implies $a = d_x\Phi$.

Direction "⇒": Let $d_x\Phi : E \to \mathbb{K}$ be linear. Then for every $\lambda \in [0,1]$ and $y \in E$ one has by convexity

$$\underbrace{\Phi(\lambda x + (1-\lambda)(x+y))}_{\Phi(x+(1-\lambda)y)} \leq \underbrace{\lambda\Phi(x) + (1-\lambda)\Phi(x+y)}_{(1-\lambda)\Phi(x+y)-(1-\lambda)\Phi(x)+\Phi(x)} \tag{3.66}$$

and thus

$$\frac{\Phi(x+(1-\lambda)y) - \Phi(x)}{(1-\lambda)} \leq \Phi(x+y) - \Phi(x). \tag{3.67}$$

Taking the limit $\lambda \to 1^-$ in (3.67) one finds

$$d_x\Phi(y) \leq \Phi(x+y) - \Phi(x), \tag{3.68}$$

that is $d_x\Phi$ is indeed a tangent functional to Φ at x.

□

3.4.5 Proposition: Equilibrium measures as tangents

Let (X, d) be a compact metric space and (X, T) a continuous, iterated dynamical system. If a T-invariant Borel-probability measure $\mu \in \mathfrak{M}(\mathscr{B}, T)$ is an equilibrium measure for the potential $f \in \mathcal{C}(X)$, then the linear functional $\mathcal{C}(X, \mathbb{R}) \to \mathbb{R},\ g \mapsto \mathbb{E}_\mu g$ is tangent to the pressure function $P_{\text{top}}(T, \cdot)$ at f. Consequently, if $P_{\text{top}}(T, \cdot)$ is Gâteaux derivable at f, there exists by 3.4.4 at most one equilibrium measure for f.

Proof: By definition of equilibrium measures one has $h_\mu(T) + \mathbb{E}_\mu f = P_{\text{top}}(T, f)$. Furthermore, for $h \in \mathcal{C}(X)$ one has by the Variational Principle $h_\mu(T) + \mathbb{E}_\mu(f + h) \leq P_{\text{top}}(T, f + h)$. Subtracting the inequality from the previous equality yields $\mathbb{E}_\mu h \leq P_{\text{top}}(T, f + h) - P_{\text{top}}(T, f)$, which by definition 3.4.3 was to be shown.

□

3.4.6 Definition: Upper semicontinuity

A function $f : X \to \mathbb{R} \cup \{\pm\infty\}$ on a metric space (X, d) is called **upper semicontinuous** if for any convergent sequence $x_n \overset{n\to\infty}{\longrightarrow} x \in X$ one has $\limsup_{n\to\infty} f(x_n) \leq f(x)$.

3.4.7 Proposition: Existence of equilibrium measures

Let (X,d) be a compact metric space with Borel-σ-algebra $\mathscr{B}$ and (X,T) a continuous, iterated dynamical system. If the function $\mathfrak{M}(\mathscr{B},T)\to\mathbb{R},\ \mu\mapsto h_\mu(T)$ is upper semicontinuous in the weak* topology, then each potential $f\in\mathcal{C}(X)$ has an equilibrium measure.

Proof: By the Variational Principle 3.3.3 there exists a sequence $(\mu_n)_{n\in\mathbb{N}}\subseteq\mathfrak{M}(\mathscr{B},T)$ of T-invariant Borel probability measures such that $h_{\mu_n}(T)+\int f\,d\mu_n\uparrow P_{\text{top}}(T,f)$. As $\mathfrak{M}(\mathscr{B},T)$ is weakly* sequentially compact, we can suppose the sequence (μ_n) to have a weak* limit μ in $\mathfrak{M}(\mathscr{B},T)$, thus by upper semicontinuity of $h_{(\cdot)}(T)$

$$P_{\text{top}}(T,f)=\limsup_{n\to\infty}\left[h_{\mu_n}(T)+\int f\,d\mu_n\right]\le h_\mu(T)+\int f\,d\mu \tag{3.69}$$

as claimed. The other inequality direction follows from the Variational Principle 3.3.3.

□

3.5 Expansive dynamical systems

In the following we consider so called expansive, iterated dynamical systems on compact metric spaces and present expressions for their topological entropy.

3.5.1 Definition: Expansive dynamical system

Let (X,d) be a metric space. We call an iterated dynamical system (X,T) **expansive**[1], if there exists a constant $\vartheta>0$ such that $d(T^nx,T^ny)<\vartheta\ \forall n\in\mathbb{N}_0$ implies $x=y$ for all $x,y\in X$. We call such a ϑ an **expansivity constant** for (X,T).

Remarks:

(i) If $\vartheta>0$ is an expansivity constant of (X,T) and $0<\widetilde{\vartheta}<\vartheta$, then $\widetilde{\vartheta}$ is one as well.

(ii) If X is a compact metric space, then by lemma A.2.4 expansivity of iterated dynamical systems on X does not depend on the metric generating the topology. As a consequence, continuously conjugated, compact, iterated dynamical systems are either both or none of them expansive.

(iii) If ϑ is an expansivity constant of (X,T), then each finite subset $N\subseteq X$ becomes (d_n,ϑ)-separated for n large enough. In particular $S(d_n,\vartheta)\overset{n\to\infty}{\longrightarrow}\infty$ provided $|X|=\infty$. When considering any $x,y\in X$ to be distinguishable iff $d(x,y)\ge\vartheta$, then any finite number of orbits are distinguishable after some long enough time.

3.5.2 Lemma: Uniform expansivity constants

Let (X,d) be a compact metric space and (X,T) an expansive, continuous, iterated dynamical system with expansivity constant ϑ. Then there exists an expansivity constant $0<\widetilde{\vartheta}\le\vartheta$ such that

$$\sup_{x\in X}\left\{\operatorname{diam}_d B_{d_n}(x,\widetilde{\vartheta})\right\}\overset{n\to\infty}{\longrightarrow}0, \tag{3.70}$$

whereas the balls $B_{d_n}(x,\widetilde{\vartheta})$ are defined with respect to the n-th Bowen-metric d_n of the system but their diameter $\operatorname{diam}_d B_{d_n}(x,\widetilde{\vartheta})$ with respect to the intrinsic metric d. We shall call such a $\widetilde{\vartheta}$ a **uniform expansivity constant**.

Proof: We present the proof given by [1]. Let $B_d^o(z_i,\vartheta/4),\ i\in I$ be a finite open covering of X by open balls and $0<\widetilde{\vartheta}\le\vartheta$ a Lebesgue-number (see A.2.1) for that covering. Suppose (3.70) to be false, then there exists an $\varepsilon>0$, numbers $(n_k)_{k\in\mathbb{N}}\subseteq\mathbb{N}$ and points $(x_k)_k\subseteq X$ such that $n_k\to\infty$ and $\operatorname{diam}_d B_{d_{n_k}}(x_k,\widetilde{\vartheta})>2\varepsilon$ for all $k\in\mathbb{N}$. We thus find points $(y_k)_k\subseteq X$ such that $d_{n_k}(x_k,y_k)\le\widetilde{\vartheta}$ but $d(x_k,y_k)\ge\varepsilon$. As X is compact, we can suppose that $x_k\overset{k\to\infty}{\longrightarrow}x\in X$ and $y_k\overset{k\to\infty}{\longrightarrow}y\in X$. Note that $x\ne y$.
Now $d_{n_k}(x_k,y_k)\le\widetilde{\vartheta}$ means in particular that $d(T^lx_k,T^ly_k)\le\widetilde{\vartheta}$ for all $l\in\{0,..,n_k\}$. Fix $l\in\mathbb{N}$, then by choice of $\widetilde{\vartheta}$ one finds that $T^lx_k,T^ly_k\in B_d^o(z_i,\vartheta/4)$ for some $i\in I$ and an infinite number of $k\in\mathbb{N}$. By continuity of T this implies that $T^lx\in B_d(z_i,\vartheta/4)$ and $T^ly\in B_d(z_i,\vartheta/4)$, thus $d(T^lx,T^ly)\le\vartheta/2$. This holds for every $l\in\mathbb{N}$. But ϑ was an expansivity constant for (X,T), which is a contradiction!

□

3.5.3 Definition: Expanding dynamical system

Let (X, d) be a metric space. We call an iterated dynamical system (X, T) **expanding**[1], if there exist constants $\Lambda > 1$ and $\vartheta > 0$ such that

$$d(Tx, Ty) \geq \Lambda \cdot d(x, y) \tag{3.71}$$

for all x, y with $d(x, y) \leq \vartheta$.

Remarks:

(i) Every expanding system is expansive with the above-mentioned constant ϑ as expansivity constant.

(ii) Wether an iterated dynamical system is expanding or not depends on the actual underlying metric!

3.5.4 Lemma: Expansive dynamical systems as expanding ones

Let (X, d) be a compact metric space and (X, T) an expansive, continuous, iterated dynamical system. Then there exists on X a metric $\widetilde{d}$ equivalent to d, under which the system becomes expanding.

Proof: Can be found in [10] or [1], Satz 46.

3.5.5 Theorem: Topological entropy of expansive systems

Let (X, d) be a compact metric space and (X, T) a continuous, expansive, iterated dynamical system. Let ϑ be a uniform expansivity constant for the system as postulated in 3.5.2. Let $N_n \subseteq X$ be arbitrary, maximally (d_n, ϑ)-separated subsets. Then the system's topological entropy $h_{\text{top}}(T)$ is given by the asymptotic growth rate

$$\limsup_{n\to\infty} \frac{1}{n} \ln |N_n| \tag{3.72}$$

of the cardinalities $|N_n|$. In particular

$$h_{\text{top}}(T) = \limsup_{n\to\infty} \frac{1}{n} \ln S(d_n, \vartheta). \tag{3.73}$$

Proof: As X is compact, by remark 3.1.4(ii) one can suppose that

$$h_{\text{top}}(T) = \lim_{\varepsilon\to 0^+} \limsup_{n\to\infty} \frac{1}{n} \ln |M_{n,\varepsilon}| \tag{3.74}$$

for some maximally (d_n, ε)-separated sets $M_{n,\varepsilon} \subseteq X$. Note that $S(d_n, \varepsilon) \geq |N_n|$ for all $0 < \varepsilon < \vartheta$ and $n \in \mathbb{N}$, so that $h_{\text{top}}(T)$ is certainly greater or equal to (3.72). It thus suffices to show that for any $0 < \varepsilon < \vartheta$ one has

$$\limsup_{n\to\infty} \frac{1}{n} \ln |M_{n,\varepsilon}| \leq \limsup_{n\to\infty} \frac{1}{n} \ln |N_n|. \tag{3.75}$$

As ϑ satisfies (3.5.2), one can find an $m_\varepsilon \in \mathbb{N}$ such that $d_m(x, y) < \vartheta$ implies $d(x, y) < \varepsilon/4$ for all $m \geq m_\varepsilon$ and $x, y \in X$. In particular

$$\forall x, y \in X,\ n \in \mathbb{N} : d_n(x, y) \geq \varepsilon/4 \ \Rightarrow\ d_{n+m_\varepsilon}(x, y) \geq \vartheta. \tag{3.76}$$

Now fix $n \in \mathbb{N}$ and $0 < \varepsilon < \vartheta$. For each $x \in M_{n,\varepsilon}$ there exists an $y \in N_{n+m_\varepsilon}$ such that $d_n(x, y) < \varepsilon/4$, because otherwise $d_n(x, y) \geq \varepsilon/4\ \forall y \in N_{n+m_\varepsilon}$ would by (3.76) imply $d_{n+m_\varepsilon}(x, y) \geq \vartheta\ \forall y \in N_{n+m_\varepsilon}$, a contradiction to the maximality of N_{n+m_ε} as $(d_{n+m_\varepsilon}, \vartheta)$ separating set. We pose $f(x) := y$. The so defined function $f : M_{n,\varepsilon} \to N_{n+m_\varepsilon}$ is injective because $f(x) = f(\widetilde{x})$ for two $x \neq \widetilde{x} \in M_{n,\varepsilon}$ would imply

$$d_n(x, \widetilde{x}) \leq d_n(x, f(x)) + d_n(\widetilde{x}, f(\widetilde{x})) < \frac{\varepsilon}{4} + \frac{\varepsilon}{4} = \frac{\varepsilon}{2}, \tag{3.77}$$

a contradiction as $M_{n,\varepsilon}$ is (d_n, ε)-separated. Thus $|M_{n,\varepsilon}| \leq |N_{n+m_\varepsilon}|$ and

$$\begin{aligned}\limsup_{n\to\infty} \frac{1}{n} \ln |M_{n,\varepsilon}| &\leq \limsup_{n\to\infty} \frac{1}{n} \ln |N_{n+m_\varepsilon}| \\ &= \limsup_{n\to\infty} \frac{1}{n+m_\varepsilon} \ln |N_{n+m_\varepsilon}| \\ &= \limsup_{n\to\infty} \frac{1}{n} \ln |N_n|\end{aligned} \tag{3.78}$$

as claimed.

□

3.5.6 Corollary: Estimation of topological entropy of expansive systems

Let (X, d) be a compact metric space and (X, T) a continuous, expansive, iterated dynamical system. For $n \in \mathbb{N}$ let $P_n \subseteq X$ be the set of fixed points of T^n. Then

$$h_{\text{top}}(T) \geq \limsup_{n\to\infty} \frac{1}{n} \ln |P_n| . \tag{3.79}$$

Proof: Let ϑ be a uniform expansivity constant for the system. Then by 3.5.5 it suffices to show that $S(d_n, \vartheta) \geq |P_n|$, or that P_n is (d_n, ϑ)-separated for that matter. But this is clearly the case, as any pair $x \neq y \in P_n$ with $d_n(x, y) < \vartheta$ would satisfy $d(T^k x, T^k y) < \vartheta$ for all times $k \in \mathbb{N}_0$, a contradiction to the expansivity of the system.

□

3.5.7 Lemma: Generators for expansive systems

Let (X, d) be a compact metric space with Borel-σ-algebra $\mathscr{B}$ and (X, T) an expansive, iterated dynamical system with uniform expansivity constant $\vartheta > 0$. Then each countable, measurable partition $\mathcal{A}$ of $(X, \mathscr{B})$ with $\text{diam}(A) \leq \vartheta$ for each $A \in \mathcal{A}$ is a generator for $(X, \mathscr{B}, T)$, that is $\bigvee_{n\in\mathbb{N}_0} T^{-n}(\mathcal{A}) = \mathscr{B}$.

Proof: Noting $\mathcal{A}^{(n)} := \bigvee_{k=0}^{n-1} T^{-k}(\mathcal{A})$, we need to show that $\sigma\left[\bigcup_{n\in\mathbb{N}_0} \mathcal{A}^{(n)}\right] = \mathscr{B}$. As each $\mathcal{A}^{(n)}$ is countable, it suffices to show that the countable system $\mathcal{B} := \bigcup_{n\in\mathbb{N}_0} \mathcal{A}^{(n)}$ "generates" the topology of the space, in the sense that each open set is union of elements in $\mathcal{B}$. It suffices to show that for each $x \in X$ and $\varepsilon > 0$ there exists an $A \in \mathcal{B}$ such that $x \in A \subseteq B_d^o(x, \varepsilon)$.

Now by definition of uniform expansivity constants there exists an $n_\varepsilon \in \mathbb{N}$ such that $d_{n_\varepsilon}(x, y) \leq \vartheta$ implies $d(x, y) < \varepsilon$ for any $x, y \in X$, or inversely, $d(x, y) \geq \varepsilon$ implies $d(T^k x, T^k y) > \vartheta$ for some $k \in \{0, .., n_\varepsilon - 1\}$. Choose $A \in \mathcal{A}^{(n_\varepsilon)}$ such that $x \in A$. Then A is of the form $A = \bigcap_{k=0}^{n_\varepsilon-1} T^{-k}(A_k)$ with $A_k \in \mathcal{A}$. Thus, any $y \in A$ satisfies $d(x, y) < \varepsilon$, since otherwise $d(T^k x, T^k y) > \vartheta$ for some $k \in \{0, .., n_\varepsilon - 1\}$, a contradiction to the fact that $T^k x, T^k y \in A_k$ and $\text{diam}(A_k) \leq \vartheta$. Therefore $x \in A \subseteq B_d^o(x, \varepsilon)$ as intended.

□

3.5.8 Theorem: Kolmogorov-Sinai entropy of expansive systems

Let (X, d) be a compact metric space with Borel-σ-algebra $\mathscr{B}$ and (X, T) an expansive, iterated dynamical system with uniform expansivity constant $\vartheta > 0$. Let $\mu \in \mathfrak{M}(\mathscr{B}, T)$ be a T-invariant probability measure and $\mathcal{A} \in \mathscr{Z}_1(\mathscr{B}, \mu)$ a countable, measurable partition with finite entropy, such that $\text{diam}(A) \leq \vartheta$ for each $A \in \mathcal{A}$. Then the Kolmogorov-Sinai entropy $h_\mu(T)$ of the system $(X, \mathscr{B}, \mu, T)$ is given by the average entropy $h_\mu(T, \mathcal{A})$ of the partition.

Proof: By lemma 3.5.7 the partition $\mathcal{A}$ is a generator for $\mathscr{B}$. By the Kolmogorov-Sinai theorem 2.2.7 this implies $h_\mu(T) = h_\mu(T, \mathcal{A})$.

□

3.5.9 Theorem: Existence of equilibrium measures for expansive systems

Let (X, d) be a compact metric space with Borel-σ-algebra $\mathscr{B}$ and (X, T) a continuous, expansive, iterated dynamical system. Then $\mathfrak{M}_{\text{eq}}(\mathscr{B}, T, f) \neq \emptyset \ \forall f \in \mathcal{C}(X)$, that is, each continuous potential $f \in \mathcal{C}(X)$ has an equilibrium measure.

Proof: The following proof was taken from [2]. By proposition 3.4.7 it suffices to show that the mapping $\mathfrak{M}(\mathscr{B}, T) \to \mathbb{R}$, $\mu \mapsto h_\mu(T)$ is upper semicontinuous. Let $\vartheta > 0$ be a uniform expansivity constant for the system and $\mu \in \mathfrak{M}(\mathscr{B}, T)$ be given. Then by 3.3.2 there exists a finite, measurable partition $\mathcal{A}$ of $(X, \mathscr{B})$ such that $\text{diam}(A) \leq \vartheta$ and $\mu(\partial A) = 0$ for all $A \in \mathcal{A}$. Note $\mathcal{A}^{(m)} := \bigvee_{k=0}^{m-1} T^{-k}(\mathcal{A})$, then each $A \in \mathcal{A}^{(m)}$ satisfies $\mu(\partial A) = 0$ since $\partial(B \cap C) \subseteq \partial B \cap \partial C$ and $\partial T^{-1}(B) \subseteq T^{-1}(\partial B)$ for any $B, C \subseteq X$. As $\mathcal{A}^{(m)}$ is finite, this implies that the function

$$\mathcal{H}_m : \mathfrak{M}(\mathscr{B}, T) \to \mathbb{R} \quad , \quad \mathcal{H}_m(\nu) := \frac{1}{m} H_\nu(T, \mathcal{A}^{(m)}) \tag{3.80}$$

is weak* continuous at μ for each $m \in \mathbb{N}$. Furthermore, the sequence of functions $(\mathcal{H}_m)_m$ is by 2.2.3 pointwise decreasing with $\mathcal{H}(\nu) := \inf_m \mathcal{H}_m(\nu) = h_\nu(T, \mathcal{A})$, which by lemma A.2.8 implies that $\mathcal{H}(\nu)$ is upper semicontinuous at μ. Finally, note that by construction of $\mathcal{A}$ and theorem 3.5.8 one has $h_\nu(T) = h_\nu(T, \mathcal{A}) = \mathcal{H}(\nu)$ for every $\nu \in \mathfrak{M}(\mathscr{B}, T)$, which concludes the proof.

□

3.5.10 Example: Affine flows on the torus

Let $\mathbb{T}^m$ be the m-dimensional torus and $T : \mathbb{T}^m \to \mathbb{T}^m$ an affine transformation defined by $T(x) := (\alpha + Ax)\text{mod}\mathbb{Z}^m$ for some $\alpha \in \mathbb{R}^m$ and $A \in \text{GL}(\mathbb{Z}^m)$. Let $\lambda_1, .., \lambda_m \in \mathbb{C}$ be the eigenvalues of A. Then the iterated dynamical system $(\mathbb{T}^m, T)$ satisfies:

1. It is expansive, provided that $|\lambda_1|, .., |\lambda_m| > 1$.
2. It has topological entropy

$$h_{\text{top}}(T) = \sum_{|\lambda_i|>1} \ln |\lambda_i| \, . \tag{3.81}$$

Proof:

1. By 3.5.3(i) it suffices to show that $(\mathbb{T}^m, T)$ is expanding in some convenient metric. For that, it suffices to show that for some large enough $n \in \mathbb{N}$ the matrix A^n is a strictly expansive map on $\mathbb{R}^m$, that is $\|A^n x\| \geq \Lambda \cdot \|x\|$ for every $x \in \mathbb{R}^m$ and some convenient $\Lambda > 1$. Equivalently, it suffices to show that A^{-n} is a strictly contractive map on $\mathbb{R}^m$, that is $\|A^{-n}\| < 1$, for n large enough. Indeed, as all eigenvalues $\lambda_1^{-1}, .., \lambda_n^{-1} \in \mathbb{C}$ of A^{-1} are by absolute value strictly smaller than 1, one has $A^{-n} \overset{n\to\infty}{\longrightarrow} 0$ in the operator norm.
2. See [1], Satz 56.

□

A Appendix

A.1 Classical thermodynamics

A.1.1 The free-energy formalism

For a classical thermodynamical system modeled by a countable phase space X with probability measure μ on $(X, 2^X)$, one defines its **thermodynamical entropy** as a functional[4]

$$S[\mu] := -\hbar \sum_{x \in X} \mu(x) \ln \mu(x) \tag{A.1}$$

[4] Supposing the Boltzmann constant k_B to be normalized to 1.

solely of the underlying measure. The **free energy** $F_\mu(T)$ of the system, a function of the system's temperature T, is defined as

$$F_\mu(T) = \mathbb{E}_\mu E - TS_\mu, \tag{A.2}$$

with E as energy function on X. Now classical thermodynamics teaches us that $F_\mu(T)$ is minimized in thermodynamic equilibrium in case of constant temperature T, fixed particle number and fixed volume. This corresponds to a variational problem in the measure μ, which translates into minimizing

$$\frac{1}{\hbar T} \sum_{x \in X} E(x)\mu(x) + \sum_{x \in X} \mu(x) \ln \mu(x) \tag{A.3}$$

under the constraint $\sum_{x \in X} \mu(x) = 1$. Using the method of Lagrange multipliers this leads to the equilibrium measure

$$\boxed{\mu(x) = \frac{e^{-E(x)/\hbar T}}{\sum\limits_{y \in X} e^{-E(y)/\hbar T}},} \tag{A.4}$$

known as **Boltzmann distribution**. Note the formal similarity to the definition of equilibrium measures in iterated dynamical systems for a potential $-E/(\hbar T)$, maximizing the expression $h_\mu + \mathbb{E}_\mu(-E/\hbar T)$, or equivalently, minimizing $\mathbb{E}_\mu(E/\hbar T) - h_\mu$.

A.1.2 Maximizing entropy under energy constraints

For a given system with constant particle number and volume, its temperature corresponds to a certain expected energy $\mathbb{E}_\mu E$. The Boltzman distribution (A.4) can also be derived as the probability measure μ on the system $(X, 2^X)$ for which $S[\mu]$ is maximized under the condition of a given expected energy $\mathbb{E}_\mu E \overset{!}{=} \mathcal{E}$. Using the method of Lagrange multipliers, this variational problem leads to the sole possibility

$$\mu(x) = \frac{e^{-\beta E(x)}}{\sum\limits_{y \in X} e^{-\beta E(y)}}, \tag{A.5}$$

with the constant β such that

$$\mathcal{E} \cdot \sum_{x \in X} e^{-\beta E(x)} \overset{!}{=} \sum_{x \in X} E(x) \cdot e^{-\beta E(x)}, \tag{A.6}$$

heavily depending on the energy function E. Comparing (A.5) with (A.4), one interprets β as the inverse temperature $1/\hbar T$ of the system.

A.2 Auxiliary statements

A.2.1 On the existence of Lebesgue-numbers

Let (X, d) be a compact metric space and $(U_i)_{i \in I}$ an open covering of X. Then there exists an $\varepsilon > 0$, called **Lebesgue number** for the covering, such that

$$\forall x \in X : \exists i \in I : B(x, \varepsilon) \subseteq U_i. \tag{A.7}$$

A.2.2 Definition: Topological equivalence of metrics

Two metrics $d, \widetilde{d}$ on some set X are called **topologically equivalent** if they induce the same open sets.

A.2.3 Definition: Uniform equivalence of metrics

Two metrics $d, \widetilde{d}$ on some set X are called **uniformly equivalent** if the identity mappings $\mathrm{Id} : (X, d) \to (X, \widetilde{d})$ and $\mathrm{Id} : (X, \widetilde{d}) \to (X, d)$ are both uniformly continuous. This is equivalent to demanding that for every $\varepsilon > 0$ there exists a $\delta > 0$ such that $d(x, y) \leq \delta$ implies $\widetilde{d}(x, y) \leq \varepsilon$ and $\widetilde{d}(x, y) \leq \delta$ implies $d(x, y) \leq \varepsilon$ for all $x, y \in X$. Note that uniformly equivalent metrics are topologically equivalent.

A.2.4 Lemma: Uniform equivalence of metrics on compact spaces

Let $d, \widetilde{d}$ be two topologically equivalent metrics on some set X such that the metric space (X, d) is compact. Then d and $\widetilde{d}$ are uniformly equivalent metrics.

Proof: As d and $\widetilde{d}$ are equivalent, the identity $\mathrm{Id} : (X, d) \to (X, \widetilde{d})$ is continuous in both directions. As (X, d) and $(X, \widetilde{d})$ are compact, it is uniformly continuous in both directions.

□

A.2.5 Definition: Continuity module

Let (X_1, d_1) and (X_2, d_2) be two metric spaces and $f : X_1 \to X_2$ some function. Then the mapping $\omega_f : [0, \infty) \to \mathbb{R} \cup \{\infty\}$ defined by

$$\omega_f(\delta) := \sup \{d_2(f(x), f(y)) : x, y \in X_1,\ d_1(x, y) \leq \delta\} \tag{A.8}$$

is called **continuity module** of f.

Remarks:

(i) If f is bounded then ω_f is real and bounded.

(ii) If f is uniformly continuous then ω_f is continuous in 0.

A.2.6 Hölder's inequality

Let $(\Omega, \mathfrak{S}, \mu)$ be a measure space and $p, q \in [1, \infty]$ such that $\frac{1}{p} + \frac{1}{q} = 1$. Then for any measurable $f, g : \Omega \to \mathbb{C}$ one has

$$\|f \cdot g\|_1 \leq \|f\|_p \cdot \|g\|_q \,. \tag{A.9}$$

Equivalently, for $0 \leq \alpha, \beta \leq 1$ satisfying $\alpha + \beta = 1$ one has

$$\left\|f^\alpha \cdot g^\beta\right\|_1 \leq \|f\|_1^\alpha \cdot \|g\|_1^\beta \,. \tag{A.10}$$

A.2.7 Jensen's inequality

Let $(\Omega, \mathfrak{S}, \mu)$ be a probability space and $\Phi : [0, \infty) \to \mathbb{R}$ concave. Then for any non-negative function $f \in L^1(\Omega, \mathfrak{S}, \mu)$ one has

$$\int \Phi(f)\ d\mu \leq \Phi\left[\int f\ d\mu\right]. \tag{A.11}$$

A.2.8 Lemma on upper semicontinuity

Let (X, d) be a metric space and $(f_m)_{m \in \mathbb{N}}$ a sequence of functions $f_m : X \to \mathbb{R} \cup \{\pm\infty\}$, all upper semicontinuous in some point $x \in X$. Then their pointwise infimum $f = \inf_{m \in \mathbb{N}} f_m$ is also upper semi-continuous in x.

Proof: Suppose the contrary, that is $\limsup_{m \to \infty} f(x_n) > f(x)$ for some convergent sequence $x_n \overset{n \to \infty}{\longrightarrow} x \in X$. Then we can w.l.o.g. suppose $f(x_n) \geq f(x) + \varepsilon\ \forall n \in \mathbb{N}$ for some $\varepsilon > 0$. Choose $m_0 \in \mathbb{N}$ such that $f_{m_0}(x) \leq f(x) + \frac{\varepsilon}{2}$, then $\inf_m f(x_n) \geq f_{m_0}(x) + \frac{\varepsilon}{2}$ for all $n \in \mathbb{N}$. In particular $f_{m_0}(x_n) \geq f_{m_0}(x) + \frac{\varepsilon}{2}$, so that $\limsup_{n \to \infty} f_{m_0}(x_n) > f_{m_0}(x)$, a contradiction to the upper semicontinuity of f_{m_0} at x_0.

□

B Symbols and abbreviations

$\mathbb{K}$: Standing for $\mathbb{R}$ or $\mathbb{C}$.

$B(x,r)$**:** Closed ball of radius r centered at point $x \in X$ of some metric space X.

$B_d(x,r)$**:** Closed ball of radius r centered at point $x \in X$ of some metric space (X,d).

$B^o(x,r)$**:** Open ball of radius r centered at point $x \in X$ of some metric space X.

$B^o_d(x,r)$**:** Open ball of radius r centered at point $x \in X$ of some metric space (X,d).

$\operatorname{diam}_d A$**:** Diameter of some subset $A \subseteq X$ of a metric space (X,d).

$\bigvee_{k=1}^n \mathcal{A}_k$**:** Refinement of measurable partitions $\mathcal{A}_1,..,\mathcal{A}_n$. See 2.1.1.

$\bigvee_{k=1}^\infty \mathcal{A}_k$**:** Limit σ-algebra of all $\sigma\left[\bigvee_{k=1}^n \mathcal{A}_k\right]$, $n \in \mathbb{N}$. See 2.1.1.

$\mathcal{Z}(\mathfrak{S})$**:** System of all countable, measurable partitions of the measurable space $(X,\mathfrak{S})$. See 2.1.1.

$\mathcal{Z}_1(\mathfrak{S},\mu)$**:** System of all countable, measurable partitions with finite entropy. See 2.1.2.

$\Sigma_T^n f$**:** Defined as $\Sigma_T^n f := \sum_{k=0}^{n-1} f \circ T^k$ for any function $f : A \to \mathbb{C}$, $T : A \to A$ and $n \in \mathbb{N}$.

ω_f**:** Continuity module of function f. See A.2.5.

$\mathfrak{M}(\mathfrak{S})$**:** Set of all probability measures on the measurable space $(X,\mathfrak{S})$.

$\mathfrak{M}(\mathfrak{S},T)$**:** T-invariant probability measures on the measurable space $(X,\mathfrak{S})$, with $T : X \to X$ measurable.

$\mathfrak{M}_{\mathrm{er}}(\mathfrak{S},T)$**:** T-invariant, ergodic probability measures on the measurable space $(X,\mathfrak{S})$, with $T : X \to X$ measurable.

$\mathfrak{M}_{\mathrm{eq}}(\mathscr{B},T,f)$**:** T-invariant equilibrium measures for the function $f \in \mathcal{C}(X)$ on the metric space (X,d), with $T : X \to T$ continuous. See 3.4.2.

References

[1] Denker, M. *Einführung in die Analysis dynamischer Systeme*. Springer-Lehrbuch (Springer, 2004). URL http://books.google.com/books?id=ignIM55Xj0sC.

[2] Przytycki, F. & Urbański, M. *Conformal Fractals: Ergodic Theory Methods*. London Mathematical Society Lecture Note Series (Cambridge University Press, 2010). URL http://www.math.unt.edu/~urbanski/pubook/pu090517.pdf.

[3] Sarig, O. Lecture notes on ergodic theory (2009).

[4] Walters, P. *An introduction to ergodic theory*. Graduate texts in mathematics (Springer, 2000). URL http://books.google.fr/books?id=eCoufOp7ONMC.

[5] Denker, M., Grillenberger, C., Sigmund, K. & en ligne), S. S. *Ergodic theory on compact spaces*, vol. 527 (Springer-Verlag, 1976).

[6] Dinaburg, E. On the relations among various entropy characteristics of dynamical systems. *Izvestiya: Mathematics* **5**, 337–378 (1971).

[7] Dinaburg, E. The relation between topological entropy and metric entropy. In *Soviet Math. Dokl*, vol. 11, 13–16 (1970).

[8] Goodman, T. Relating topological entropy and measure entropy. *Bulletin of the London Mathematical Society* **3**, 176 (1971).

[9] Goodwyn, L. The product theorem for topological entropy. *Transactions of the American Mathematical Society* **158**, 445–452 (1971).

[10] Reddy, W. L. Expanding maps on compact metric spaces. *Topology and its Applications* **13**, 327 – 334 (1982). URL http://www.sciencedirect.com/science/article/B6V1K-45F51BS-1N/2/7811eb3522c496aec784c03c2e09d58d.