Dynamics and classical entropies of $\mathbb{Z}$-actions

Matthieu Joseph and Christopher-Lloyd Simon

L'entropie est un concept anthropique.
É.Ghys

This is a six hours course given by Mikael de la Salle at Ens de Lyon as an introduction to the winter school *Entropies & soficity* held January 2018 in Lyon, France. The goal of the course was to study both measure-preserving and topological dynamical systems and to define different entropies of $\mathbb{Z}$-actions. In no case Mikael de la Salle can be held responsible for the errors that are left in this paper.

1 Dynamical systems

1.1 Measure-preserving dynamical system

Definition. *A measure-preserving dynamical system is an action of a group G on a probability space $(X, \mathcal{A}, \mu)$, such that for all $g \in G$, the map defined on X by $x \mapsto g \cdot x$ is measurable and measure-preserving: for all $A \in \mathcal{A}$,*

$$\mu(g^{-1}A) = \mu(A),$$

where $g^{-1}A = \{x \in X, g \cdot x \in A\}$.

Since G is a group, the maps $x \mapsto g \cdot x$ are all invertible. This definition of measure-preserving dynamical systems excludes very interesting examples, such as the angle doubling map on the circle.

The groups under consideration will be countable. We write $G \curvearrowright (X, \mathcal{A}, \mu)$ for the dynamical system given by the action of G on $(X, \mathcal{A}, \mu)$. If $G = \mathbb{Z}$ then the group action is given by the iteration of an invertible, bimeasurable and measure-preserving map $T : X \to X$, which we denote by $\mathbb{Z} \curvearrowright^T (X, \mathcal{A}, \mu)$.

The standard hypothesis. In this paper, we always assume that $(X, \mathcal{A}, \mu)$ is a *standard* probability space, this means that $(X, \mathcal{A}, \mu)$ is isomorphic to a compact metric space endowed with its completed Borel sigma-algebra with respect to a Borel probability measure.

Let us recall how to construct the completion of a sigma-algebra. Let $(X, \mathcal{A}_0, \mu)$ be a measure space. A subset B of X is *negligible* if there exists $A \in \mathcal{A}_0$ such that $B \subset A$ and $\mu(A) = 0$. The completed sigma-algebra $\mathcal{A}$ is the smallest sigma-algebra containing $\mathcal{A}_0$ and the negligible sets. It can be shown that

$$\mathcal{A} = \{A \cup B \mid A \in \mathcal{A}_0, B \text{ negligible set of } X\}.$$

The measure μ can be extended to a measure on $\mathcal{A}$ in the following way : for all $A \in \mathcal{A}_0$ and for all negligible set B,

$$\mu(A \cup B) = \mu(A).$$

Another way to proceed is to consider the semi-metric space $(2^X, d_0)$ where $d_0(A, B) = \mu_0(A \Delta B)$. This is indeed symetric and satisfies the triangle inequality. Then $\mathcal{A}$ is taken as the closure of $\mathcal{A}_0$ in 2^X for that semi-metric: $A_\infty \in \mathcal{A}$ if there is a sequence $A_n \in \mathcal{A}_0$ such that $d_0(A_n, A_\infty) \to 0$. The measure μ_0 extends by continuity to μ defined on $\mathcal{A}$.

Example. Every Borel measure on a *polish* space, that is a complete separable metric space, is standard. In fact, standard probability space are completely classified. Such a space is either isomorphic the compact interval $[0, 1]$ endowed with its Lebesgue measure, or to a countable set with some counting measure, or else to a disjoint union of the two previous.

Remark. The standard hypothesis prevents two undesirable phenomenae. The first appears when there are redundant information in the underlying set from the measure theoretic point of view. For instance with $X = [-1, 1]$ endowed with the symetric Borel algebra: $\{B \in \mathcal{B}(X) | B = -B\}$. The second happens when X is so big that $L^1(X, \mathcal{A}, \mu)$ is not separable. Actually, if $(Y, \mathcal{B}, \mu)$ is a space such that $L^1(Y, \mathcal{B}, \nu)$ is separable, then there exists a standard space $(X, \mathcal{A}, \mu)$ such that $L^1(Y, \mathcal{B}, \nu) = L^1(X, \mathcal{A}, \mu)$.

Shifts. The archetypal example of a measure-preserving dynamical system is the *Bernoulli shift*. Let $(X, \mathcal{A}, \mu)$ be a probability space. The Bernoulli shift on X is the map

$$S : \begin{array}{ccc} X^{\mathbb{Z}} & \longrightarrow & X^{\mathbb{Z}} \\ (a_n)_{n \in \mathbb{Z}} & \longmapsto & (a_{n-1})_{n \in \mathbb{Z}} \end{array} \tag{1}$$

The $\mathbb{Z}$-action defined by $n \cdot x = S^n(x)$ gives rise to a measure-preserving dynamical system on $(X^{\mathbb{Z}}, \mathcal{A}^{\otimes \mathbb{Z}}, \mu^{\otimes \mathbb{Z}})$.

More generally, if G is a countable discrete group, denote by X^G the topological product space, which is a compact metrizable space since G is countable. One can show that there exists a unique probability measure $\mu^{\otimes G}$ on X^G that behaves as a product measure. The action of G on $(X^G, \mathcal{A}^{\otimes G}, \mu^{\otimes G})$ given by $g \cdot (a_h)_{h \in G} = (a_{g^{-1}h})_{h \in G}$ is also a measure-preserving dynamical system.

Isomorphism between measure-preserving dynamical systems. Two measure-preserving dynamical systems $G \curvearrowright (X, \mathcal{A}, \mu)$ and $G \curvearrowright (Y, \mathcal{B}, \nu)$ are called *isomorphic*, or *conjugate*, if there exists a bimeasurable and measure-preserving bijection $\varphi : X' \to Y'$ between subsets $X' \subset X$ and $Y' \subset Y$ of full measure, such that $\varphi_* \mu = \nu$ and which is equivariant under the G-actions: for all $x \in X'$ and all $g \in G$,

$$\varphi(g \cdot x) = g \cdot \varphi(x).$$

The natural problem is to classify dynamical systems up to isomorphism. For instance, under which conditions are the Bernouilli shifts over two probability spaces $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ isomorphic? One way to answer this question is to find numerical invariants under conjugacy, such as the measure-theoretic entropy.

Factors. The dynamical system $G \curvearrowright (Y, \mathcal{B}, \nu)$ is a *factor* of $G \curvearrowright (X, \mathcal{A}, \mu)$ if there exists a measurable and measure-preserving map $q : X \to Y$ which is equivariant with respect to the G-actions: for all $x \in X$ and all $g \in G$,

$$q(g \cdot x) = g \cdot q(x).$$

One thinks of $G \curvearrowright (Y, \mathcal{B}, \nu)$ as being simpler than $G \curvearrowright (X, \mathcal{A}, \mu)$.

Exercise. $G \curvearrowright (Y, \mathcal{B}, \nu)$ is a factor of $G \curvearrowright (X, \mathcal{A}, \mu)$ if and only if $G \curvearrowright (Y, \mathcal{B}, \nu)$ is isomorphic to $G \curvearrowright (X, \mathcal{A}', \mu)$ where $\mathcal{A}'$ is a G invariant sub-sigma-algebra of $\mathcal{A}$.

Preliminary notions for the measure-theoretic entropy.

Definition. *A μ-partition (or partition) of the probability space $(X, \mathcal{A}, \mu)$ is a finite or countable collection $\mathcal{P} = \{P_n \in \mathcal{A}\}$ such that*

- $\mu(P_n \cap P_m) = 0$ *if* $m \neq n$,
- $\mu(P_n) > 0$,
- $\mu((\bigcup P_n)^c) = 0$.

A μ-partition $\mathcal{P}$ is a *refinement* of $\mathcal{Q}$, and we write $\mathcal{Q} \preccurlyeq \mathcal{P}$ if each element of $\mathcal{P}$ is contained in an element of $\mathcal{Q}$ (up to a negligible set). The *join* $\mathcal{P} \vee \mathcal{Q}$ of two μ-partitions $\mathcal{P}$ and $\mathcal{Q}$ is the μ-partition

$$\{P_n \cap Q_m \mid m, n \in \mathbb{N}, \mu(P_n \cap Q_m) > 0\}.$$

Thus, the μ-partition $\mathcal{P} \vee \mathcal{Q}$ is a refinement of $\mathcal{P}$ and $\mathcal{Q}$.

Remark. The $\vee$ notation refers to a supremum for the lattice structure on the μ-partitions of X induced by the partial order $\preccurlyeq$.

If $G \curvearrowright (X, \mathcal{A}, \mu)$ is a measure-preserving dynamical system and $\mathcal{P}$ a μ-partition, then for all $g \in G$, the g-pullback $g^{-1}\mathcal{P}$ of $\mathcal{P}$ given by $\{g^{-1}P_n \mid n \in \mathbb{N}\}$ is also a μ-partition.

We can provide a coding for a point $x \in X$ by looking at its trajectory in the parts of a partition, under the action of G:

$$\begin{array}{rccc} q: & X & \longrightarrow & \mathcal{P}^G \\ & x & \longmapsto & (\{P \in \mathcal{P} \mid g^{-1} \cdot x \in P\})_{g \in G} \end{array}$$

This map is equivariant under the G-actions: for all $g \in G$ and all $x \in X$, $q(g \cdot x) = g \cdot q(x)$.

Definition. *The μ-partition $\mathcal{P}$ of $(X, \mathcal{A}, \mu)$ is called generating for the G-action if the completion with respect to μ of the smallest σ-algebra containing all the $g^{-1}\mathcal{P}$ for $g \in G$ is equal to $\mathcal{A}$.*

Proposition. *If $\mathcal{P}$ is generating, then $q\colon (X, \mathcal{A}, \mu) \to (\mathcal{P}^G, q_*\mathcal{A}, q_*\mu)$ is an isomorphism.*

Remark. Beware, the action of G on $(\mathcal{P}^G, q_*\mathcal{A}, q_*\mu)$ is not a Bernoulli shift in the sense that the measure $q_*\mu$ is not the product measure over some fixed alphabet space.

1.2 Topological dynamical system

Definition. *A topological dynamical system is an action of a group G on a topological space X such that for all $g \in G$, the map defined on X by $x \mapsto g \cdot x$ is an homeomorphism.*

As in the measure preserving case, G is a countable group, and we write $G \curvearrowright X$. When $G = \mathbb{Z}$, the group action is given by the iteration of an homeomorphism $T : X \to X$, and we write $\mathbb{Z} \curvearrowright^T X$.

The compact metric hypothesis. In this paper, we always assume the topological space X to be a compact metrizable space.

Shifts. Let X be a finite set, and μ a probability measure on $(X, 2^X)$. The shift S defined in (1) is an homeomorphism if X is endowed with the discrete topology and $X^{\mathbb{Z}}$ the product topology. Thus, the action of $\mathbb{Z}$ on $X^{\mathbb{Z}}$ corresponding to the iteration of S defines a topological dynamical system.

More generally, the action of a countable discrete group G on X^G (endowed with the product topology) given by $g \cdot (a_h)_{h \in G} = (a_{g^{-1}h})_{h \in G}$ defines similarly a topological dynamical system.

Isomorphism between topological dynamical systems. Two topological dynamical systems $G \curvearrowright X$ and $G \curvearrowright Y$ are *isomorphic*, or *conjugate*, if there exists a G-equivariant homeomorphism $\varphi : X \to Y$.

2 Entropies of $\mathbb{Z}$-actions

In this section we focus on $\mathbb{Z}$-actions, which amount to the iteration of an invertible map on the underlying space. By convention, we extend by continuity the map $t \mapsto t\log(1/t)$ with $0\log(1/0) = 0$.

2.1 Entropy of a measure-preserving dynamical system

2.1.1 Shannon's entropy of partitions

Definition. *The entropy of a partition $\mathcal{P}$ of $(X, \mathcal{A}, \mu)$ is the quantity $H(\mathcal{P}, \mu) \in [0, \infty]$ defined by:*

$$H(\mathcal{P}, \mu) = \sum_{A \in \mathcal{P}} \mu(A) \log(1/\mu(A)).$$

If the measure μ is clear in the context, the entropy of the partition $\mathcal{P}$ is denoted as $H(\mathcal{P})$.

The entropy is to be interpreted as the average amount of binary digits of information one has to provide to tell in which part a random point $x \in X$ belongs to. Indeed, using dichotomy, one can specify any element of $A \in \mathcal{P}$ in roughly $\log(1/\mu(A))$ bits of information.

Remark. Note that $\mu \mapsto H(\mathcal{P}, \mu)$ is concave as a sum of $t \mapsto t\log(1/t)$:

$$(1-\lambda)H(\mathcal{P}, \mu) + \lambda H(\mathcal{P}, \nu) \leq H(\mathcal{P}, (1-\lambda)\mu + \lambda\nu).$$

Therefore, the entropy is maximal for a uniform measure.

Definition. *The conditional entropy of a partition $\mathcal{P}$ with respect to $\mathcal{Q}$ is:*

$$H((\mathcal{P}, \mu)|\mathcal{Q}) = \sum_{\substack{B \in \mathcal{Q} \\ \mu(B) > 0}} \mu(B) H(\mathcal{P}, \mu_B),$$

where $\mu_B = \mu(\cdot \cap B)/\mu(B)$.

As in the non-conditional case, the conditional entropy is the average amount of additional information one has to give in order to specify the $\mathcal{P}$-coding of a point if we already know its $\mathcal{Q}$-coding.

Lemma. *Let $\mathcal{P}$, $\mathcal{Q}$, $\mathcal{Q}'$ be partitions of $(X, \mathcal{A}, \mu)$ such that $\mathcal{Q}'$ refines $\mathcal{Q}$ and $H(\mathcal{Q}')$ is finite. Then*

- $H(\mathcal{P}|\mathcal{Q}) = H(\mathcal{P} \vee \mathcal{Q}) - H(\mathcal{Q})$,
- $H(\mathcal{P}|\mathcal{Q}') \leq H(\mathcal{P}|\mathcal{Q}) \leq H(\mathcal{P})$,
- $H(\mathcal{Q}) \leq H(\mathcal{Q}')$.

Proof. The first identity is immediate after developing $\log(\mu_B(A)) = \log(\mu(A \cap B)) - \log(B)$ and separating the sum. The second and the third follow by summing by parts $\sum_{B' \in Q} = \sum_{B \in Q} \sum_{B' \subset B}$ and using convexity of H. $\square$

Remark. If $\mathcal{P}$ and $\mathcal{Q}$ are two partitions, then

$$H(\mathcal{P} \vee \mathcal{Q}) \leq H(\mathcal{P}) + H(\mathcal{Q}),$$

with equality if and only if $\mathcal{P}$ and $\mathcal{Q}$ are "independant" in the following sense : for all $A \in \mathcal{P}, B \in \mathcal{Q}$, $\mu(A \cap B) = \mu(A)\mu(B)$.

Define the *mutual information* I and the *decorelation* d between two partitions:

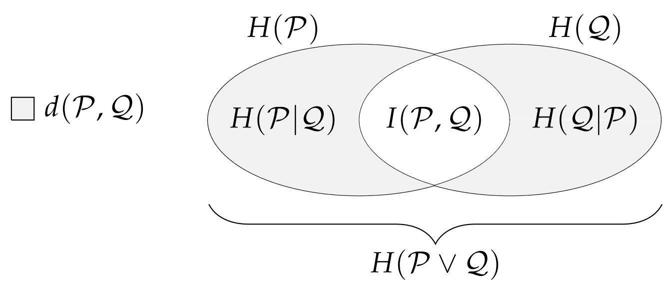
$$\begin{aligned} I(\mathcal{P}, \mathcal{Q}) &= H(\mathcal{P}) + H(\mathcal{Q}) - H(\mathcal{P} \vee \mathcal{Q}), \\ d(\mathcal{P}, \mathcal{Q}) &= H(\mathcal{P}|\mathcal{Q}) + H(\mathcal{Q}|\mathcal{P}). \end{aligned}$$

Those are non negative quantities, that are related by the formula

$$d(\mathcal{P}, \mathcal{Q}) = H(\mathcal{P}) + H(\mathcal{Q}) - 2I(\mathcal{P}, \mathcal{Q}).$$

The decorelation d satisfies $d(\mathcal{P}, \mathcal{Q}) = 0 \Leftrightarrow \mathcal{P} = \mathcal{Q}$ up to negligible sets. This is because $H(\mathcal{P}|\mathcal{Q}) = 0$ implies that $\mathcal{Q}$ refines $\mathcal{P}$. Moreover, one can show that if $\mathcal{P}, \mathcal{Q}$ and $\mathcal{R}$ are partitions, then $H(\mathcal{P}|\mathcal{Q}) \leq H(\mathcal{P}|\mathcal{R}) + H(\mathcal{R}|\mathcal{Q})$, which implies that d satisfies the triangle inequality. Thus d is a distance on the set of partitions up to negligible sets, known as the *Rokhlin distance*.

We can sum up all this in the following diagram.



2.1.2 Kolmogorov's entropy of a dynamical system

If $\mathbb{Z} \curvearrowright^T (X, \mathcal{A}, \mu)$ is a measure-preserving dynamical system and $\mathcal{P}$ a μ-partition of $(X, \mathcal{A}, \mu)$, for all integer n, we set

$$\mathcal{P}^n = \bigvee_{i=0}^{n-1} T^{-i}\mathcal{P}.$$

Definition. *Let $\mathcal{P}$ be a μ-partition of the measure-preserving dynamical system $\mathbb{Z} \curvearrowright^T (X, \mathcal{A}, \mu)$ with finite entropy. Then the following limit is well defined:*

$$h(\mathcal{P}, \mu, T) = \lim_{n\to\infty} \frac{1}{n} H(\mathcal{P}^n, \mu)$$

and it increases by taking refinements of the μ-partition $\mathcal{P}$. The number

$$h(\mu, T) = \sup\{h(\mathcal{P}, \mu, T) \mid H(\mathcal{P}) < \infty\}$$

is the entropy of the dynamical system $\mathbb{Z} \curvearrowright^T (X, \mathcal{A}, \mu)$. Entropy is invariant under isomorphism of measure-preserving dynamical system.

Proof. The sequence $(H(\mathcal{P}^n))_{n\in\mathbb{N}}$ is subadditive :

$$\begin{aligned} H(\mathcal{P}^{m+n}) &= H(\mathcal{P}^n \vee T^{-n}\mathcal{P}^m) \\ &= H(\mathcal{P}^n) + H(T^{-n}\mathcal{P}^m|\mathcal{P}^n) \\ &\leq H(\mathcal{P}^n) + H(\mathcal{P}^m). \end{aligned}$$

Hence, the limit is well defined. The previous lemma implies it is monotonous with respect to the μ-partition. □

The next result is one of the main tools to compute entropies of dynamical systems and to avoid working with a supremum over μ-partitions.

Theorem (Kolmogorov-Sinaï). *If a μ-partition $\mathcal{P}$ with finite entropy is generating, then*

$$h(\mu, T) = h(\mathcal{P}, \mu, T).$$

Corollary 1. *If (A, μ) is a finite probability space then,*

$$h(A^{\mathbb{Z}}, \mu^{\otimes\mathbb{Z}}, S) = H(A, \mu) = \sum_{x\in A} \mu(x)\log(1/\mu(x))$$

In particular, if the underlying base spaces have different entropies then the Bernoulli shifts over those alphabets are not isomorphic.

Corollary 2. *If $\mathbb{Z} \curvearrowright^T (X, \mathcal{A}, \mu)$ has a generating μ-partition with a finite number m of parts, then $h(\mu, T) \leq \log m$*

One can prove that the converse of Corollary 2 is true: if $h(\mu, T) \leq \log m$ then there exists a generating μ-partition with less than m parts.

Theorem (Entropy decreases through factors). *If $q\colon (X, \mathcal{A}, \mu) \to (Y, \mathcal{B}, \nu)$ is a factor map between two dynamical systems $\mathbb{Z} \curvearrowright^T (X, \mathcal{A}, \mu)$ and $\mathbb{Z} \curvearrowright^S (Y, \mathcal{B}, \nu)$, then $h(\nu, S) \leq h(\mu, T)$.*

2.1.3 Proof of the Kolmogorov-Sinaï theorem

We must show that for any other μ-partition $\mathcal{Q}$ with finite entropy, we have

$$h(\mathcal{Q}, \mu, T) \leq h(\mathcal{P}, \mu, T)$$

for $\mathcal{A} = \overline{\sigma}\left(\bigcup_{n\in\mathbb{N}} \bigvee_{|j|<N} T^{-j}\mathcal{P}\right)$.

Particular Case. There is an integer N such that $\bigvee_{-N\leq j\leq N} T^{-j}\mathcal{P}$ refines $\mathcal{Q}$. Then, leaving out μ and T from notations, we use monotony with respect to μ-partitions, the definiton of entropy h in terms of H and that of H. The last equality is where we make use of the amenability of the group $\mathbb{Z}$:

$$
\begin{aligned}
h(\mathcal{Q},T) &\leq h\Big(\bigvee_{|j|<N} T^{-j}\mathcal{P}\Big) \\
&= \lim_{n\to\infty} \frac{1}{n} H\Big(\bigvee_{i=-N}^{N+n-1} T^{-i}\mathcal{P}\Big) \\
&= \lim_{n\to\infty} \frac{1}{n} H(T^N\mathcal{P}^{2N+n}) \\
&= \lim_{n\to\infty} \frac{n+2N}{n}\frac{1}{n+2N} H(\mathcal{P}^{2N+n}) \\
&= h(\mathcal{P},T).
\end{aligned}
$$

General Case. We reduce it to the particular case with the following lemmas.

Lemma. *For all $\varepsilon > 0$, there exists a μ-partition $\mathcal{Q}'$ with finite entropy such that $d(\mathcal{Q},\mathcal{Q}') < \varepsilon$.*

Recall that $\mathcal{Q}^n = \bigvee_{i=0}^{n-1} T^{-i}\mathcal{Q}$.

Lemma. *$H(\mathcal{Q}^n|\mathcal{Q}'^n) \leq nH(\mathcal{Q}|\mathcal{Q}')$ and the same holds exchanging $\mathcal{Q}$ and $\mathcal{Q}'$.*

Proof. Using the definition of $\mathcal{Q}^n$, the subadditivity lemma and monotony:

$$
\begin{aligned}
H(\mathcal{Q}^n|\mathcal{Q}'^n) &= H\Big(\bigvee_{i=0}^{n-1} T^{-i}\mathcal{Q}|\mathcal{Q}'^n\Big) \\
&\leq \sum_0^{n-1} H(T^{-i}\mathcal{Q}|\mathcal{Q}'^n) \\
&\leq \sum_0^{n-1} H(T^{-i}\mathcal{Q}|T^{-i}\mathcal{Q}') \\
&\leq nH(\mathcal{Q}|\mathcal{Q}').
\end{aligned}
$$

□

2.2 Entropy of a topological dynamical system

Metric approach. Let $\mathbb{Z} \curvearrowright^T X$ be a topological dynamical system, where X is a compact metrizable space. A distance d on X inducing the topology can be refined under iterations of T by setting for all integer n:

$$d_n(x,y) = \max_{i\in\{0,\dots,n-1\}} d(T^ix, T^iy).$$

The function d_n is a distance inducing the topology of X whose open balls are:

$$B_{d_n}(x,\varepsilon) = \bigcap_{i=0}^{n-1} T^{-i}(B_d(T^ix,\varepsilon)).$$

For $\varepsilon > 0$, define the *covering number* and the *packing number* as

$$r_n(\varepsilon) = \min\{\#F \mid F \subset X, \bigcup_{x \in F} B_{d_n}(x, \varepsilon) = X\},$$
$$s_n(\varepsilon) = \max\{\#F \mid F \subset X, \forall x, y \in F, x \neq y \Rightarrow d_n(x, y) \geq \varepsilon\}.$$

These two numbers are related by the inequalities

$$s_n(2\varepsilon) \leq r_n(\varepsilon) \leq s_n(\varepsilon).$$

Remark. The covering and packing numbers give an estimate of the number of different orbits under the mapping T one can get with a spatial resolution to the order ε. The more T mixes points around in X, the greater the number of orbits. As before, the entropy shall measure the average exponential growth of this number, in the limit where the resolution gets thinner.

Definition. *The metric entropy of* $\mathbb{Z} \curvearrowright^T (X, d)$ *is defined by the nondecreasing* ε*-limits:*

$$h_d(T) = \lim_{\varepsilon \to 0} \limsup_{n \to \infty} \frac{1}{n} \log r_n(\varepsilon) = \lim_{\varepsilon \to 0} \limsup_{n \to \infty} \frac{1}{n} \log s_n(\varepsilon).$$

Remark. This definition depends sharply on the distance d, although we shall see that the actual invariant h does not. In fact, we define below an invariant h_{top} using topological concepts only, and show that it is equal to the metric entropy for any compatible metric structure. Therefore, the metric entropy actually only depends on the topology arising from the metric.

Topological approach. An *open cover* of X is a cover consisting of open sets. An open cover β is a *refinement* of an open cover α if each element of β is contained in an element of α. As for partitions, the *join* $\alpha \vee \beta$ of two open cover α and β is the cover $\{A \cap B \mid A \in \alpha, B \in \beta\}$. For a topological dynamical system $\mathbb{Z} \curvearrowright^T X$ and an open cover α of X, we set

$$\alpha^n = \bigvee_{i=0}^{n-1} T^{-i}\alpha.$$

A *subcover* of a cover α is a subfamily $\beta \subset \alpha$ which is also a cover. In particular it is a refinement of α. For an open cover α, set

$$N(\alpha) = \min\{\#\beta \mid \beta \text{ open subcover of } \alpha\}.$$

Definition. *Let* α *be an open cover of a topological dynamical system* $\mathbb{Z} \curvearrowright^T X$*. Then the following limit is well defined and increases by taking refinements of* α*:*

$$h_{top}(\alpha, T) = \lim_{n \to \infty} \frac{1}{n} \log N(\alpha^n).$$

Define the topological entropy of the system $\mathbb{Z} \curvearrowright^T X$*:*

$$h_{top}(T) = \sup\{h_{top}(\alpha, T) \mid \alpha \textit{ open cover}\}.$$

Entropy is invariant under isomorphism of topological dynamical systems.

Proof. The submultiplicativity $N(\alpha \vee \beta) \leq N(\alpha)N(\beta)$ ensures that the limit exists. If β is a refinement of α, then β^n is a refinement of α^n, and N increases by taking refinements. Therefore $h_{top}(\alpha, T)$ is monotonous with respect to the open cover. □

Theorem. *For any topological dynamical system $\mathbb{Z} \curvearrowright^T (X, d)$, $h_d(T) = h_{top}(T)$.*

Remark. If α is an open cover of X, there exists by compactness of X a Lebesgue number $\delta > 0$ such that every subset of X with diameter less than δ is contained in an element of α. The *diameter* of an open cover β is $\operatorname{diam}_d \beta = \sup\{\operatorname{diam}_d B \mid B \in \beta\}$.

Thus, if β is an open cover with $\operatorname{diam}_d \beta \leq \delta$, then β refines α. Therefore, if α_n is a sequence of open covers with $\operatorname{diam} \alpha_n \to 0$, then

$$h_{top}(T) = \lim_{n\to\infty} h_{top}(\alpha_n, T).$$

Proof that topological entropy and metric entropy coincide. Each of the following lemmas yields one inequality between h_d and h_{top}.

Lemma. *Let δ be a Lebesgue number of an open cover α. Then $N(\alpha^n) \leq r_n(\delta/2)$.*

Proof. Let F be a subset of X such that $\#F = r_n(\delta/2)$ and $B_{d_n}(F, \delta/2) = X$, where

$$B_{d_n}(F, \delta/2) = \bigcup_{x\in F} B_{d_n}(x, \delta/2) = \bigcup_{x\in F} \bigcap_{i=0}^{n-1} T^{-i}(B_d(T^i x, \delta/2)).$$

By definition of δ, for all $i \in \{0, \ldots, n-1\}$ and for all $x \in F$, the ball $B_d(T^i x, \delta/2)$ is contained in an element of α. Thus $B_{d_n}(x, \delta/2)$ is contained in an element α_x^n of α^n. Choosing those elements $(\alpha_x^n)_{x\in F}$ gives a subcover which implies the inequality in the lemma. □

Lemma. *Let α be an open cover of diameter less than ε. Then $s_n(\varepsilon) \leq N(\alpha^n)$.*

Proof. Let F be an ε-net of X for d_n, that is an ε-packing ($\forall x, y \in F$, $x \neq y \Rightarrow d_n(x, y) > \varepsilon$) with maximal cardinality $s_n(\varepsilon)$. By definition of d_n, any element of α^n has a d_n-diameter smaller than ε. So any element of a subcover β of α contains at most one point of F: $\#F \leq \#\beta$. □

Expansive maps.

Definition. *A map $T : X \to X$ is δ-expansive if for all $x, y \in X$, there exists an integer $n \in \mathbb{Z}$ such that*

$$d(T^n x, T^n y) > \delta.$$

Expansiveness of T does not depend on the distance d. In fact, if d' is another distance then by compactness of X, d and d' are uniformly equivalent: the identity map from (X, d) to (X, d') is uniformly continuous. This implies that T is δ'-expansive with respect to d'. For an expansive map T, there exists an open cover α that is *metrically generating* in the following sense :

$$\operatorname{diam}\left(\bigvee_{i=-n}^{n} T^{-i}\alpha\right) \xrightarrow[n\to\infty]{} 0.$$

The following theorem is a practical tool to compute entropy of topological dynamical systems and to avoid working with a supremum over open covers or with thinner and thinner resolution of ε-covers.

Theorem. *Let $T : X \to X$ be a δ-expansive map for d, and α a metrically generating open cover.*

- $h_d(T) = h_{top}(\alpha, T)$.
- $\forall \varepsilon < \frac{\delta}{4}$, $h(T) = \lim_{n\to\infty} \frac{1}{n} \log r_n(\varepsilon) = \lim_{n\to\infty} \frac{1}{n} \log s_n(\varepsilon)$.

Proof. For the first assumption, by definition of a metrically generating open cover and by the previous remark concerning covers whose diameter shrinks to 0:

$$h(T) = \lim_{n\to\infty} h_{top}\left(\bigvee_{i=-n}^{n} T^{-i}\alpha, T \right).$$

Then, the last equality is where we make use of the amenability of $\mathbb{Z}$:

$$\begin{aligned} h_{top}\left(\bigvee_{i=-n}^{n} T^{-i}\alpha, T \right) &= \lim_{k\to\infty} \frac{1}{k} \log N\left(\bigvee_{j=0}^{k-1} T^{-j} \Big(\bigvee_{i=-n}^{n} T^{-i}\alpha \Big) \right) \\ &= \lim_{k\to\infty} \frac{1}{k} \log N(\alpha^{2n+k-1}) \\ &= \lim_{k\to\infty} \frac{2n+k-1}{k} \frac{1}{2n+k-1} \log N(\alpha^{2n+k-1}) \\ &= h_{top}(\alpha, T). \end{aligned}$$

□

2.3 The variational principle

The main result of this section relates entropies of topological and measure-preserving dynamical systems.

Let X be a compact metrizable space, and $T : X \to X$ a homeomorphism. Denote by $\mathcal{M}(X, T)$ the set of Borel probability measures on X invariant under T. It is a convex compact subset of the dual to the topological vector space $(\mathcal{C}(X), \|\cdot\|_\infty)$ endowed with the weak-star topology.

Theorem (Markov-Kakutani)**.** *The set $\mathcal{M}(X, T)$ is nonempty.*

Proof. Let μ be a Borel probability measure on X. The sequence of means

$$\mu_n = \frac{1}{n} \sum_{k=0}^{n-1} T_*^k \mu$$

has a converging subsequence $\mu_{n_k} \overset{*}{\rightharpoonup} \mu_\infty$. Then, for all $f \in \mathcal{C}(X)$,

$$\begin{aligned} \left| \int_X f \circ T d\mu_\infty - \int_X f d\mu_\infty \right| &= \lim_{k\to\infty} \left| \int_X f \circ T d\mu_{n_k} - \int_X f d\mu_{n_k} \right| \\ &= \lim_{k\to\infty} \frac{1}{n_k} \left| \int_X (f \circ T^{n_k} - f) d\mu \right| \le \lim_{k\to\infty} \frac{2\|f\|_\infty}{n_k}. \end{aligned}$$

Therefore $\mu_\infty \in \mathcal{M}(X, T)$. □

Remark. The map $\mu \in \mathcal{M}(X,T) \mapsto h(\mu,T)$ is affine. By definition, an invariant measure is ergodic if it is an extreme point of $\mathcal{M}(X,T)$. The Krein-Milmann theorem implies that the compact convex set $\mathcal{M}(X,T)$ is the convex hull of its extremal points. So entropy takes its extremal values on ergodic invariant measures.

Theorem (Variational principle)**.** *Let $\mathbb{Z} \curvearrowright^T X$ be a topological dynamical system. Then*

$$h_{top}(T) = \sup_{\mu \in \mathcal{M}(X,T)} h(\mu,T).$$

Sketch of the proof. We first show that for all $\mu \in \mathcal{M}(X,T)$, $h(\mu,T) \leq h_{top}(T)$

Particular case. Fix $\mu \in \mathcal{M}(X,T)$. Let $\mathcal{P}$ be a μ-partition of X which is also an open cover. Then for all integer n, $H(\mu,\mathcal{P}^n) \leq \log \#\mathcal{P}^n = \log N(\mathcal{P}^n)$. Therefore,

$$h(\mathcal{P},\mu,T) \leq \lim_{n\to\infty} \frac{1}{n} \log N(\mathcal{P}^n) = h_{top}(\mathcal{P},T) \leq h_{top}(T).$$

General case. One can enlarge a little bit the pieces of a partition $\mathcal{P}$ to get open sets.

For the other inequality, given $\varepsilon > 0$ one constructs a measure $\mu \in \mathcal{M}(X,T)$ such that $\limsup_{n\to+\infty} \frac{1}{n} \log s_n(\varepsilon) \leq h(\mu,T)$. To do so, let F_n be an ε-net of X for d_n. Set

$$\nu_n = \frac{1}{s_n(\varepsilon)} \sum_{x\in F_n} \delta_x,$$

and

$$\mu_n = \frac{1}{n} \sum_{i=0}^{n-1} T^i_* \nu_n.$$

By compactness, (μ_n) has a converging subsequence $\mu_{n_k} \overset{*}{\rightharpoonup} \mu_\infty$. One can then show that $\mu_\infty \in \mathcal{M}(X,T)$ and that $\limsup_{n\to+\infty} \frac{1}{n} \log s_n(\varepsilon) \leq h(\mu,T)$. $\square$

3 Entropies of amenable groups

The theory of entropy for $\mathbb{Z}$-action has been generalized to actions of amenable groups by Ornstein and Weiss [3]. This section is a very brief overview of the subject. For more about entropies of amenable action, see [2]. In the following, the groups under consideration will be countable.

Definition. *A group G is amenable if there exists a Følner sequence : a sequence (F_n) of finite nonempty subsets of G such that for all $g \in G$,*

$$\frac{|gF_n \cap F_n|}{|F_n|} \underset{n\to\infty}{\longrightarrow} 1.$$

Measure-preserving entropy.

Definition. *Let $G \curvearrowright (X, \mathcal{A}, \mu)$ be a measure-preserving dynamical system, with G an amenable group. If $\mathcal{P}$ is a finite μ-partition of X, then the following limit*

$$h(\mathcal{P}, \mu, G) = \lim_{n\to\infty} \frac{1}{|F_n|} \log H(\mathcal{P}^{F_n}, \mu),$$

where $\mathcal{P}^{F_n} = \bigvee_{g \in F_n} g^{-1}\mathcal{P}$, exists and does not depend on the Følner sequence (F_n). Define the measure entropy of $G \curvearrowright (X, \mathcal{A}, \mu)$ as

$$h(\mu, G) = \sup\{h(\mathcal{P}, \mu, G) \mid \mathcal{P} \text{ finite } \mu\text{-partition}\}.$$

Topological entropy. As for $\mathbb{Z}$-actions, there are two equivalent ways of defining the topological entropy that coincide: one with a metric approach, another with a topological approach. We only give here the second one.

Definition. *Let $G \curvearrowright X$ be a topological dynamical system, with G an amenable group and X a compact metrizable space. If α is a finite open cover of X, the following limit*

$$h_{top}(\alpha, G) = \lim_{n\to\infty} \frac{1}{|F_n|} \log N(\alpha^{F_n}),$$

where $\alpha^{F_n} = \bigvee_{g \in F_n} g^{-1}\alpha$, exists and does not depend on the Følner sequence (F_n). Define the topological entropy of $G \curvearrowright X$ as

$$h_{top}(G) = \sup\{h_{top}(\alpha, G) \mid \alpha \text{ finite open cover}\}.$$

The variational principle. Let $G \curvearrowright X$ be a topological dynamical system, with G an amenable group and X a compact metrizable space. Denote by $\mathcal{M}(X, G)$ the set of Borel probability measures on X invariant under G, that is for all $g \in G$, and for all Borel set A, $\mu(g^{-1}A) = \mu(A)$. It is a convex compact subset of the dual to the topological vector space $(\mathcal{C}(X), \|\cdot\|_\infty)$ endowed with the weak-star topology. As for $\mathbb{Z}$-actions:

Theorem (Markov-Kakutani)**.** *The set $\mathcal{M}(X,T)$ is nonempty.*

Proof. Let μ be a Borel probability measure on X, and (F_n) a Følner sequence. The sequence of means

$$\mu_n = \frac{1}{|F_n|} \sum_{g \in F_n} g_* \mu$$

has a converging subsequence $\mu_{n_k} \xrightarrow{*} \mu_\infty$. So, for all $f \in \mathcal{C}(X)$, and for all $g \in G$:

$$\begin{aligned}
\left| \int_X f(g \cdot x) d\mu_\infty(x) - \int_X f(x) d\mu_\infty(x) \right| &= \lim_{k \to \infty} \left| \int_X f(g \cdot x) d\mu_{n_k}(x) - \int_X f(x) d\mu_{n_k}(x) \right| \\
&= \lim_{k \to \infty} \frac{1}{|F_{n_k}|} \left| \sum_{g' \in F_{n_k}} \int_X (f(g \cdot x) - f(x)) d(g'_* \mu)(x) \right| \\
&= \lim_{k \to \infty} \frac{1}{|F_{n_k}|} \left| \sum_{g' \in F_{n_k}} \int_X (f((gg') \cdot x) - f(g' \cdot x)) d\mu(x) \right| \\
&= \lim_{k \to \infty} \frac{1}{|F_{n_k}|} \left| \sum_{g' \in gF_{n_k} \Delta F_{n_k}} \int_X (f((gg') \cdot x) - f(g' \cdot x)) d\mu(x) \right| \\
&\leq 2\|f\|_\infty \times \lim_{k \to \infty} \frac{|gF_{n_k} \Delta F_{n_k}|}{|F_{n_k}|}.
\end{aligned}$$

Therefore $\mu_\infty \in \mathcal{M}(X,G)$. □

As for $\mathbb{Z}$-actions, the variational principle relates topological and measure entropies.

Theorem. *For a topological dynamical system $G \curvearrowright X$ with G amenable and X compact metrizable,*

$$h_{top}(G) = \sup_{\mu \in \mathcal{M}(X,G)} h(\mu, G).$$

4 Nonamenable groups : the example of Ornstein & Weiss

Problem. Can we define a convenient notion of entropy for measure preserving or topological dynamical systems under the action of any countable group?

A notion of entropy must be a scalar quantity defined for a dynamical system which is invariant under conjugation and whose size describes the rate at which the group mixes the elements in the space. By convenient, we mean satisfying the remarkable properties we have encountered up to now in the case of $\mathbb{Z}$, and which hold true in the more general context of amenable groups:

- The entropy of a Bernoulli shift is that of its base.
- Entropy decreases through factor maps.

The next example shows that it is not possible to do so for any countable group, and one has to be ready to let down some of those properties. It is the last which will be relaxed.

Example of Ornstein & Weiss. The free group Γ on two generators a, b acts on its Cayley graph with generating set $S = \{a^{\pm 1}, b^{\pm 1}\}$. This is an infinite simplicial complex of dimension 1 whose verticies are $\{g \mid g \in \Gamma\}$ and edges are $\{(g, gs) \mid g \in \Gamma, s \in S\}$.

Consider the Bernoulli shift $\Gamma \curvearrowright X = B^\Gamma$ over the base $B = \{0, 1\}$ with uniform distribution. This is the measure preserving action on the vector space of 0-cochains $\omega \colon \Gamma \to \mathbb{Z}/2\mathbb{Z}$ given by: $g \cdot (\omega_v)_{v\in\Gamma} = (\omega_{g^{-1}v})_{v\in\Gamma}$. If a correct notion of entropy could be defined for the group Γ, then we would expect the entropy of that shift to equal that of the base: $\log 2$.

The coboundary operator from 0-cochains to 1-cochains is surjective and equivariant with respect to the action of Γ:

$$\begin{array}{rccc} \partial : & B^\Gamma & \longrightarrow & B^\Gamma \times B^\Gamma = (B \times B)^\Gamma \\ & (\omega_g)_g & \longmapsto & (\omega_{ga} - \omega_g, \omega_{gb} - \omega_g)_g \end{array}$$

Since it is measure preserving, it defines a factor map from X to the shift over the base $B \times B$ which should have entropy $\log 4$. This wrecks any hope to define a notion of entropy satisfying the decreasing property through factor maps, even for such natural maps.

References

[1] T. Downarowicz, *Entropy in dynamical systems*, Cambridge University Press, 2011.

[2] D. Kerr and H. Li, *Ergodic theory: Independence and dichotomies*, Springer Monographs in Mathematics, 2016.

[3] D. Ornstein and B. Weiss, *Entropy and isomorphism theorems for actions of amenable groups*, J. Anal. Math. **48** (1987), 1–141.