Local Entropy Theory

Shahid Beheshti University

Mohammad Nourbakhsh Marvast

m.nourbakhsh2000@gmail.com

Supervisor: Dr. Bijan Ahmadi Kakavandi

August 2022

Contents

Chapter 1

Introduction to the Project

As one of the most interesting concepts of Dynamical Systems, entropy theory appear to be useful in characterising the dynamical systems and divide it into two categories: zero-entropy systems and non-zero-entropy ones. In this note, the guideline is the survey [8], "Local entropy theory". In the survey, Glasner and Ye, made an overview of this topics which has been in development since the early 1990s.

It had been contributed to dozens of phenomenons, definitions, adoptions, and in a word, history until we have today definition of entropy. It has take a long way from the definitions in physics through the Shannon entropy in information theory, to the measure-theoretical entropy and topological entropy. In the concept of measure-theoretical and topological entropy, Kolmogorov, Sinai, Bowen, Adler, Furstenberg, Glasner, Pinsker, Walters and so many other great mathematicians have gained lots of elaborations and achievements.

What's about the name? What is the meaning of 'entropy'?

As the explanation in [15]: The word entropy was invented in 1865 by the German physicist and mathematician Rudolf Clausius, one of the founding pioneers of Thermodynamics. In the theory of systems in thermodynamical equilibrium, the entropy quantifies the degree of "disorder" in the system.

The origin of the word came back to the Greek word 'entropia' meaning "a turning toward" (en: in; trope: transformation). From this word we have the German word Clausius used 'Entropie' meaning 'measure of the disorder of a system'[1].

What's in a name? In the case of Shannon's measure the naming was not accidental. In 1961 one of us (Tribus) asked Shannon what he had thought about when he had finally confirmed his famous measure. Shannon replied: "My greatest concern was what to call it. I thought of calling it 'information,' but the word was overly used, so I decided to call it 'uncertainty'. When I discussed it with John Von Neumann, he had a better idea. Von Neumann told me, 'You should call it entropy, for two reasons. In the first place your uncertainty function has been used in statistical mechanics under that name, so it already has a name. In the second place, and more important, no one knows what entropy really is, so in a debate you will always have the advantage' ".

Kolomogrov and Sinai developed the concept of entropy for a system in Ergodic Theory (nowadays we call it 'Kolmogorov-Sinai entropy').
It was in 1965 that Adler [1] introduced the concept of Topological Entropy. They assigned a number to open covers to measure it's 'size'.

In this note, we attempted to study the concept of entropy with both measure-theoretical aspects and topologicals along with their prerequisites. Unfortunately, we could not have covered the whole topics. Though, they were lots of papers and books we could relish reading them. The very books we used are mentioned in the notes; However, the epitomes of them are [15], [17], [4], and [7].

We try our best to give the proofs intricately; Although, in some parts we give the sketch of proofs and in some parts we just mention a reference to the proofs.

[1] www.etymonline.com/word/entropy

Chapter 2

Dynamical System

Dynamical system theory is the study of the qualitative properties of group actions on spaces with certain structures. In this survey we are mainly interested in actions by homeomorphisms on compact metric spaces with an additional structure of a Borel probability measure invariant under the action. We mostly consider Z-actions.

2.1 Topological Dynamical Systems (TDS)

By a *topological dynamical system (TDS)* we mean a pair (X, T), where X is a compact metric space and $T\colon X \to X$ is a self-homeomorphism. In some theorems we also emphasize on the properties of T, e.g. being surjective. In references the main definition of a TDS is a pair (X, T), where X is a compact metric space and $T\colon X \to X$ is a Continues map. Also in [6], Furstenberg use the word *"flow"* for a TDS; and use the word *"bilateral"* for a TDS with a homeomorphism $T\colon X \to X$. However, we consider a TDS along with a self-homeomorphism.

Definition 2.1.1. *A TDS is a pair (X, T), where X is a compact metric space and $T\colon X \to X$ is a homeomorphism.*

A TDS (X, T) is called trivial, if $\#X \leq 1$.

We will see that in many parts, the map T just needs to be surjective.

Definition 2.1.2. *Consider a TDS (X, T), and a point $x \in X$. The orbit of x is the set:*

$$\mathrm{orb}_T(x) := \{x, Tx, T^2x, \dots\}$$

Definition 2.1.3. *Considering a TDS (X, T) and a point $x \in X$, the orbit of x is called periodic if there exists some $n \in \mathbb{N}$ such that*

$$T^n x = x$$

Consequently, the orbit of x will be finite.

Example 2.1.1. *Consider $S^1 = \mathbb{R}/\mathbb{Z}$ with the rotation map* $\begin{cases} R_\alpha \colon S^1 \to S^1 \\ R_\alpha([x]) = [x + \alpha] \end{cases}$

1. *R_α is well-defined and one-to-one*

   $[x] = [y] \leftrightarrow \exists m \in \mathbb{Z} s.t. x - y = m \leftrightarrow$

   $x - \alpha - (y - \alpha) = m \leftrightarrow R_\alpha[x] = R_\alpha[y]$

2. *R_α is onto.*

   For each $[x] \in S^1$ consider the class $[x - \alpha]. \in S^1$

3. *R_α is continues.*

   Due to the $\epsilon - \delta$ definition of continuity:

   Since $d([x], [y]) = x - y \pmod 1$ for arbitrary $\epsilon > 0$,take $\delta < \epsilon$ and notice that $d([x], [y]) = d([x + \alpha], [y + \alpha])$.

4. *R_α^{-1} is continues.*

   Since $d([x], [y]) = d([x + \alpha], [y + \alpha])$, R_α is open map; So R_α^{-1} is continues.

5. *S^1 is a compact metric space.*

   The map $\begin{cases} i \colon [0, 1] \to \mathbb{R}/\mathbb{Z} \\ i(x) = [x] \end{cases}$ *is continues,onto and $[0, 1]$ is compact.*

Definition 2.1.4. *For a TDS (X, T) and open sets $U, V \subseteq X$, Let*
$N(U, V) := \{n \in \mathbb{N} \mid T^n U \cap V \neq \emptyset\}$

- *A TDS (X, T) is transitive if for every open sets $U, V \subseteq X$:*
  $N(U, V) \neq \emptyset$.
- *A TDS (X, T) is weakly mixing if for every open sets $U_1, V_1, U_2, V_2 \subseteq X$,*
  $N(U_1, V_1) \cap N(U_2, V_2) \neq \emptyset$.
- *A TDS (X, T) is mixing or strongly mixing if for every open sets $U, V \subseteq X$, there exists $n_0 \in \mathbb{N}$ such that:*
  $\{n \in \mathbb{N} \mid n > n_0\} \subseteq N(U, V)$.

Proposition 2.1.1. *If a TDS (X, T) be strongly mixing, then it's weakly mixing; and if (X, T) is weakly mixing, then it's transitive.*

Proof. The proof is straightforward. □

Proposition 2.1.2. *A TDS (X, T) is transitive iff $\bigcup_{n \in \mathbb{N}} \Delta_n \subseteq X \times X$ is dense; where $\Delta_n := \{(x, T^n x) \colon x \in X\}$.*

Proof. If (X, T) is transitive, for each open neighbourhood $U \times V \subseteq X \times X$ of $(x, y) \in X \times X$, since U is a neighbourhood of x and V is a neighbourhood of y, there is $n \in \mathbb{N}$ such that $T^n U \cap V \neq \emptyset$. Thus $(U \times T^n U) \cap (U \times V) \neq \emptyset$ so that there is $(z, T^n z) \in (U \times T^n U) \cap (U \times V)$. Therefore

$$(U \times V) \cap \bigcup_{n \in \mathbb{N}} \Delta_n \neq \emptyset$$

In other words, $\bigcup_{n \in \mathbb{N}} \Delta_n$ is dense subset of $X \times X$.
Conversely, if $\bigcup_{n \in \mathbb{N}} \Delta_n$ is dense subset of $X \times X$, for every open sets $U, V \subseteq X$, there is $n \in \mathbb{N}$ such that $(U \times V) \cap \Delta_n \neq \emptyset$ so that there is $(z, T^n z) \in U \times V$. Thus

$$T^n z \in T^n U \cap V \neq \emptyset$$

□

Proposition 2.1.3. *A TDS (X, T) is weakly mixing iff $(X \times X, T \times T)$ is transitive.*

Proof. The proof is straightforward. □

Example 2.1.2. *Consider the unit circle S^1 with rotation map* $\begin{cases} R_\alpha \colon S^1 \to S^1 \\ R_\alpha([x]) = [x + \alpha] \end{cases}$

1. *If $\alpha = 0$ the system is not transitive, considering two open set with no intersection.*
2. *If $\alpha = \frac{p}{q}$ for $p, q \in \mathbb{N}$, $p < q$, and $(p, q) = 1$ R_α is not transitive,*
   Define $\delta := \min\{n\frac{p}{q} - \lfloor n\frac{p}{q} \rfloor \mid 0 < n < q\}$. Consider the $\epsilon > 0$ such that $\epsilon < \delta$.
   consider the open sets $U := \{[x] \mid 0 < x < \epsilon\}$ and $V := \{[x] \mid \epsilon < x < \delta\}$.
   For these two, we have $N(U, V) = \emptyset$, since
   $T^n U = \{[x] \mid n\frac{p}{q} < x < n\frac{p}{q} + \epsilon\} = \{[x] \mid n\frac{p}{q} - \lfloor n\frac{p}{q} \rfloor < x < n\frac{p}{q} - \lfloor n\frac{p}{q} \rfloor + \epsilon\}$.
3. *If $\alpha \notin \mathbb{Q}$ R_α is transitive, since $\forall x \in S^1 \colon \overline{orb_{R_\alpha}(x)} = S^1$.*
   $\forall U, V \overset{\text{open}}{\subseteq} S^1 \colon \forall x \in U \colon \overline{orb_{R_\alpha}(x)} \cap V \neq \emptyset$
   $\Rightarrow \forall U, V \overset{\text{open}}{\subseteq} S^1 \colon \forall x \in U \colon \exists n \in \mathbb{N}$ *s.t.* $R_\alpha^n(x) \in V$
   $\Rightarrow \forall U, V \overset{\text{open}}{\subseteq} S^1 \colon N(U, V) \neq \emptyset$.

Definition 2.1.5. *Considering a TDS (X, T), a subset $A \subseteq X$ is called T-invariant, if*

$$TA \subseteq A$$

Definition 2.1.6. *Let (X, T) be a TDS. $x \in X$ is called a* ***recurrent point****, if there is an increasing sequence $\{n_i\}$ in $\mathbb{N}$ with $\lim_{i \to \infty} T^{n_i} x = x$.*
In other words, there exists a subsequence of $\mathrm{orb}_T(x)$ converging to x.

Example 2.1.3.
- *Consider any topological space X, every point of the TDS (X, id_X) is recurrent point.*

- *Every point of S^1, considering the rotation map of rational radius, is recurrent point.*
- *Each point of a TDS (X, T) with periodic orbit is recurrent point*

Definition 2.1.7. *Consider a TDS (X, T) and a closed subset $Y \subseteq X$. If $(Y, T\restriction_Y)$ be a TDS, we call $(Y, T\restriction_Y)$ a subsystem of (X, T).*

Definition 2.1.8. *A TDS (X, T) is minimal if it has no non-trivial subsystem.*

2.2 Measure Theoretical Dynamical Systems (MDS)

In this section we are going to study the fundamental definitions and theorems of measure theoretical dynamical systems. In this section and generally in the whole notes, by saying measure, we mean **probability measure**, i.e. the measure with values in $[0, 1]$.

Definition 2.2.1. *Consider the measure space $(X, \mathscr{A}, \mu)$. We say a measurable function $T\colon X \to X$ is μ-**invariant**, when*

$$\forall A \in \mathscr{A}\colon \mu(T^{-1}A) = \mu(A)$$

*or in general case, considering also the measure space $(Y, \mathscr{B}, \nu)$, we say a measurable function $\pi\colon Y \to X$ is **measure-preserving** when*

$$\forall A \in \mathscr{A}\colon \nu(\pi^{-1}A) = \mu(A)$$

As we'll see the proof in corollary (4.0.2), T is μ-invariant iff for any integrable function $\phi\colon X \to \mathbb{R}$,

$$\int_X \phi \circ T \,\mathrm{d}\mu = \int_X \phi \,\mathrm{d}\mu$$

Definition 2.2.2 (MDS)**.** *If $(X, \mathscr{A}, \mu)$ be a measure space and $T\colon X \to X$ be a μ-invariant transformation, we call $(X, \mathscr{A}, \mu, T)$ measure theoretic dynamical system and we write $(X, \mathscr{A}, \mu, T)$ is a **MDS**.*

Definition 2.2.3 (Orbit)**.** *Let $(X, \mathscr{A}, \mu, T)$ be a MDS.For every $x \in X$, the set*

$$\mathrm{orb}_T(x) := \{x, T(x), T^2(x), \dots\}$$

is called the orbit of x.

Definition 2.2.4. *Consider a MDS $(X, \mathscr{A}, \mu, T)$ we say the system is ergodic or simply T is ergodic if*

$$\forall A \in \mathscr{A} : T^{-1}A = A \implies \mu(A) = 0 \text{ or } \mu(A) = 1$$

2.3 Factor map and Homomorphism

In this section we are going to introduce the maps between dynamical systems -both TDSs and MDSs- that preserves the dynamical properties of the *'larger'* space.

Definition 2.3.1 (TDS Homomorphism, Factor map, Extension)**.** *Consider two TDS (X, T) and (Y, S). The continues surjective map $\pi : Y \to X$ is called a factor map or a homomorphism when*

$$\forall y \in Y : \pi \circ S(y) = T \circ \pi(y)$$

i.e. the diagram below commutes:

$$\begin{tikzcd} Y \arrow[r, "S"] \arrow[d, "\pi"'] \arrow[dr, dashed, "\pi\circ S = T\circ\pi"] & Y \arrow[d, "\pi"] \\ X \arrow[r, "T"'] & X \end{tikzcd}$$

(X, T) is called a ***factor*** *of (Y, S) and (Y, S) is called an* ***extension*** *of (X, T).*

Definition 2.3.2 (MDS Homomorphism, Factor map, Extension)**.** *Consider two MDS $(Y, \mathscr{B}, \nu, S)$ and $(X, \mathscr{A}, \mu, T)$ with a surjective measurable function $\pi : (Y, \mathscr{B}, \nu, S) \to (X, \mathscr{A}, \mu, T)$. π is called a factor map or homomorphism if*

- *π is measurable; i.e.*

$$\forall A \in \mathscr{A} : \pi^{-1}A \in \mathscr{B} \quad \textit{i.e. } \pi^{-1}\mathscr{A} \subseteq \mathscr{B}$$

- *π is measure invariant; i.e.*

$$\forall A \in \mathscr{A} : \nu(\pi^{-1}A) = \mu(A)$$

- *$\pi \circ S = T \circ \pi$; i.e the diagram below commutes:*

$$\begin{tikzcd}
(Y,\mathscr{B},\nu) \arrow[r, "S"] \arrow[d, "\pi"'] \arrow[dr, dashed, bend left, "\pi\circ S"] \arrow[dr, dashed, bend right, "T\circ\pi"'] \arrow[dr, phantom, "="] & (Y,\mathscr{B},\nu) \arrow[d, "\pi"] \\
(X,\mathscr{A},\mu) \arrow[r, "T"] & (X,\mathscr{A},\mu)
\end{tikzcd}$$

$(X,\mathscr{A},\mu)$ is called the ***factor*** *of $(Y,\mathscr{B},\nu)$ and $(Y,\mathscr{B},\nu)$ is denoted as an* ***extension*** *of $(X,\mathscr{A},\mu)$.*

Definition 2.3.3. *If X and Y is two dynamical system -either two TDS or two MDS- $\pi\colon Y \to X$ is isomorphism if*

- *π is bijective and continues.*
- *$\pi\colon Y \to X$ is a factor map.*
- *$\pi^{-1}\colon X \to Y$ is a factor map.*

Proposition 2.3.1. *Consider two dynamical system -either TDS or MDS- (X,T) and (Y,S) if $\pi\colon Y \to X$ is a factor map, then for each $y \in Y$,*

$$\pi\Big(\mathrm{orb}_S(y)\Big) = \mathrm{orb}_T\Big(\pi(y)\Big)$$

Proof.

$$\begin{aligned}
\mathrm{orb}_T\Big(\pi(y)\Big) &= \{\pi(y), T\circ\pi(y), \dots\} \\
&= \{\pi(y), \pi\circ S(y), \dots\} \\
&= \pi\Big(\{y, S(y), \dots\}\Big) \\
&= \pi\Big(\mathrm{orb}_S(y)\Big)
\end{aligned}$$

□

Proposition 2.3.2. *If there is a factor map $\pi\colon (Y,S)\to(X,T)$ between two TDSs, then considering the recurrent point $y\in Y$, $\pi(y)$ is a recurrent point of X.*

Proof. Since y is a recurrent point of Y, there is a sequence $\{y_i\}_{i\in\mathbb{N}}$ in $\mathrm{orb}_S(y)$ such that $\lim_{i\to\infty} y_i = y$.

Since π is continues, $\lim_{i\to\infty}\pi(y_i) = \pi(y)$. So $\{\pi(y_i)\}_{i\in\mathbb{N}}$ converges to $\pi(y)$ in the set $\pi\Big(\mathrm{orb}_S(y)\Big)$

By previous proposition, we have $\pi\Big(\mathrm{orb}_S(y)\Big) = \mathrm{orb}_T\Big(\pi(y)\Big)$.

Thus, the sequence $\{\pi(y_i)\}_{i\in\mathbb{N}}$ in $\mathrm{orb}_T\Big(\pi(y)\Big)$ converges to $\pi(y)$. Therefor $\pi(y)$ is a recurrent point of X. □

Corollary 2.3.1. *If there is a factor map $\pi\colon (Y,S)\to(X,T)$ between two TDSs, and the set of recurrent points of Y is dense, then the set of recurrent points of X is also dense.*

Proof. Since π is continues, it's project a dense set to a dense set and by the previous proposition, the result is obtained. □

Proposition 2.3.3. *If $\pi_1\colon Y_1\to X_1$ and $\pi_2\colon Y_2\to X_2$ be two factor map between dynamical systems -either TDS or MDS-*

then $\begin{cases} \pi_1 \times \pi_2 \colon Y_1 \times Y_2 \to X_1 \times X_2 \\ \pi_1 \times \pi_2(y_1, y_2) = (\pi_1(y_1), \pi_2(y_2)) \end{cases}$ *is a factor map between dynamical systems* $Y_1 \times Y_2$ *and* $X_1 \times X_2$.

Proof. The proof is straightforward. □

Remark 2.3.1. *If $h \colon Y \to X$ is a homeomorphism between topological spaces, then for every continues function $T \colon X \to X$ there exists $S \colon Y \to Y$ such that h is a factor map.*

Proof. $\forall T \colon X \to X$, consider the continues function

$$\begin{cases} S \colon Y \to Y \\ S(y) = h^{-1} \circ T \circ h(y) \end{cases}$$

we have

$$h \circ S = h \circ h^{-1} \circ T \circ h = T \circ h$$

□

Proposition 2.3.4. *Let $\pi \colon (Y, S) \to (X, T)$ be a factor map. If (Y, S) is transitive, (X, T) is also transitive.*

Proof. (Y, S) is transitive, so for every nonempty $U', V' \overset{\text{open}}{\subseteq} Y$, there exists $n \in \mathbb{N}$ such that $S^n(U') \cap V' \neq \emptyset$.
For every non empty $U, V \overset{\text{open}}{\subseteq} X$, Since (Y, S) is transitive, there is $n \in \mathbb{N}$ such that $S^n(\pi^{-1}U) \cap \pi^{-1}V \neq \emptyset$. Since

$$\begin{aligned} \pi^{-1}(T^n U \cap V) &= \pi^{-1}(T^n U) \cap \pi^{-1} V \\ &= S^n(\pi^{-1}U) \cap \pi^{-1}V \neq \emptyset \end{aligned}$$

we have that $T^n U \cap V \neq \emptyset$. Thus, (X, T) is transitive. □

Proposition 2.3.5. *Let* $\pi\colon (Y,S) \to (X,T)$ *be a factor map. If* (Y,S) *is weakly mixing,* (X,T) *is also weakly mixing.*

Proof. By proposition (2.1.3) We saw that (Y,S) is weakly mixing iff $(Y \times Y, S \times S)$ be transitive. Also by proposition (2.3.3) we saw that $(Y \times Y, S \times S)$ is an extension of $(X \times X, T \times T)$. By proposition (2.3.4) we obtain that $(X \times X, T \times T)$ is also transitive. Therefore (X,T) is weakly mixing. □

Proposition 2.3.6. *Let* $\pi\colon (Y,S) \to (X,T)$ *be a factor map. If* (Y,S) *is strongly mixing,* (X,T) *is also strongly mixing.*

Proof. The proof is straightforward. □

Definition 2.3.4 (Cantor set)**.** *The cantor dynamical system is the set*

$$C := \{(s_1, s_2, \dots) \mid \forall i \in \mathbb{N}\colon s_i \in \{0,2\}\}$$

with discrete topology and the dynamic map

$$\begin{cases} \sigma\colon C \to C \\ \sigma((s_1, s_2, \dots)) = (s_2, s_3, \dots) \end{cases}$$

One could check that C *is uncountable and compact, and the shift map* σ *is homeomorphism.*

Lemma 2.3.1. *Every compact metric space* X *is a continuous image of the cantor set* C*.*

Proof. Here there is a sketch of the proof:

1. (X,T) is embedded in $[0,1]^{\mathbb{N}}$; i.e. there exists $A \overset{\text{compact}}{\subseteq} [0,1]^{\mathbb{N}}$ such that X is homeomorphic to A.

2. $[0, 1]$ is a continues image of C by the continues function

$$\begin{cases} f\colon C \to [0,1] \\ f((s_1, s_2, \dots)) = \sum_{i=1}^{\infty} \frac{s_i}{2^{i+1}} \end{cases}$$

3. C is homeomorphic to $C^{\mathbb{N}}$ Since $\mathbb{N}$ can be written as a countable union of countable subsets $U_i := \{2^i n \colon n \text{ is an odd natural number}\}$.

   Thus,

$$C = \{0,2\}^{\mathbb{N}} = \{0,2\}^{\bigcup_{i\in\mathbb{N}} U_i} = \prod_{i=1}^{\infty} \{0,2\}^{U_i} \cong \prod_{i=1}^{\infty} \{0,2\}^{\mathbb{N}} = C^{\mathbb{N}}$$

4. The set $[0, 1]^{\mathbb{N}}$ is continues image of C, since $[0, 1]^{\mathbb{N}}$ is continues image of $C^{\mathbb{N}} \cong C$.
5. If K is a closed subset of the Cantor set C, then K is the continuous image of the Cantor set.

All in all,

$$\begin{tikzcd}
X & \overset{\text{compact}}{A} \subseteq [0,1]^{\mathbb{N}} \arrow[l, "\text{(by 1) } \exists h"'] & {[0,1]^{\mathbb{N}}} \\
C \arrow[u, "h \circ f_1 \circ f_2"] \arrow[r, "\text{(by 5) } \exists f_2"] & f_2^{-1}(A) \arrow[u, "f_1"'] & C \arrow[u, "\text{(by 4) } \exists f_1"]
\end{tikzcd}$$

Thus, the result is obtained.

□

Theorem 2.3.1. *For every TDS (X, T), there exists a dynamical system (Y, S), where (Y, S) is Cantor dynamical system and (X, T) is a factor of (Y, S) (i.e. there exists a factor map $\pi\colon Y \to X$).*

Proof. The proof is in 3 steps

1. As X is compact metric space by lemma (2.3.1), there exists a cantor set C and a continues surjective map $\pi'\colon C \to X$. Consider the set

$$Y := \left\{ y \in C^{\mathbb{Z}} \colon \forall n \in \mathbb{Z} \colon \pi'(y_{n+1}) = T \circ \pi'(y_n) \right\}$$

Let S be a shift map, i.e. $\begin{cases} S : Y \to Y \\ S(\dots, y_0, y_1, y_2, \dots) \\ \qquad\qquad = (\dots, y_1, y_2, y_3, \dots) \end{cases}$

and also consider $\begin{cases} \pi : Y \to X \\ \pi(\dots, y_{-1}, y_0, y_1, \dots) = \pi'(y_0) \end{cases}$

2. It's easy to check that S is well-defined. And $Y \subseteq C^{\mathbb{Z}}$ is closed in product topology since

$$Y = \bigcap_{n \in \mathbb{Z}} \left\{ (\dots, y_0, y_1, y_2, \dots) \in C^{\mathbb{Z}} \colon \pi'(y_{n+1}) = T \circ \pi'(y_n) \right\}$$

and for each $n \in \mathbb{Z}$

$$\begin{aligned} &\left\{ (\dots, y_0, y_1, y_2, \dots) \in C^{\mathbb{Z}} \colon \pi'(y_{n+1}) = T \circ \pi'(y_n) \right\} \\ &\qquad = \left(\prod_{i=-\infty}^{n} C \right) \times \left((\pi')^{-1}(T \circ \pi'(y_n)) \right) \times \left(\prod_{i=n+2}^{\infty} C \right) \end{aligned}$$

is closed and the intersection of countable closed sets is closed.

Therefore (Y, S) is a TDS.

3. $\pi \circ S = T \circ \pi$ since

$$\begin{aligned}
&\forall y = (\dots, y_{-1}, y_0, y_1, \dots) \in Y: \\
&\qquad \pi \circ S(y) = \pi((\dots, y_0, y_1, y_2, \dots)) = \pi'(y_1) \\
&\qquad T \circ \pi(y) = T(\pi'(y_0)) = \pi'(y_1) \\
&\implies \pi \circ S(y) = T \circ \pi(y)
\end{aligned}$$

□

Corollary 2.3.2. *For every TDS (X, T), there exists a zero-dimensional -by definition* (6.3.1)- *TDS (Y, S) such that (X, T) is a factor of (Y, S).*

Proof. As every Cantor set is zero-dimensional, by theorem (2.3.1) the result is obtained.

□

Definition 2.3.5 (Natural Extension)**.** *Let (X, T) be a TDS, T* ***be a surjective map****, and $d\colon X \to \mathbb{R}$ be our distance function on X. Define*

$$\tilde{X} := \{(x_1, x_2, \dots)\colon T(x_{i+1}) = x_i, x_i \in X, i \in \mathbb{N}\}$$

and consider the metric $\begin{cases} d_T : \tilde{X} \to \mathbb{R} \\ d_T\Big((x_1, x_2, \dots), (y_1, y_2, \dots)\Big) = \sum_{i=1}^{\infty} \frac{d(x_i, y_i)}{2^i} \end{cases}$ *and let* $\begin{cases} \sigma_T : \tilde{X} \to \tilde{X} \\ \sigma_T\Big((x_1, x_2, \dots)\Big) = (\end{cases}$
be shift homeomorphism.
It's easy to check that $\begin{cases} \pi : (\tilde{X}, \sigma_T) \to (X, T) \\ \pi\Big((x_1, x_2, \dots)\Big) = x_1 \end{cases}$ *is an* ***open map and homomorphism*** *between dynamical system and we call $(\tilde{X}, \sigma_T)$ the* ***natural extension*** *of dynamical sys-*

tem (X, T)*:*

$$\begin{tikzcd}
(x_1, x_2, \dots) \arrow[d, squiggly, "\pi_1"] \arrow[rrrr, squiggly, bend left, "\sigma_T"] & \in & \tilde{X} \arrow[r, "\sigma_T"] \arrow[d, "\pi_1"] & \tilde{X} \arrow[d, "\pi_1"] & \ni & (T(x_1), x_1, \dots) \arrow[d, squiggly, "\pi_1"] \\
x_1 \arrow[rrrrr, squiggly, bend right, "T"] & \in & X \arrow[r, "T"] & X & \ni & T(x_1)
\end{tikzcd}$$

Remark 2.3.2. *Considering* T *as a surjective map is necessary. e.g. consider any non-trivial compact topological space* X *- i.e.* $\#X \geq 2$ *-* $x_0 \in X$ *and the continues function*

$$\begin{cases} T\colon X \to X \\ T(x) = x_0 \end{cases}$$

(X, T) *as a TDS is not a factor of invertible system:*

If there's an invertible extension of (X, T)*, called* (Y, S) *with a factor map* $\pi\colon Y \to X$*, we should have*

$$\forall y \in Y : \pi \circ S(y) = T \circ \pi(y) = x_0$$

but since π *and* S *are surjective, we have* $X = \{x_0\}$*. this contradiction shows that not every TDSs* (X, T) *has a surjective extension. Although, surjective TDSs have an invertible extension called natural extension.*

2.4 Disjointness

The notion of disjointness of two TDSs and two MDSs was introduced by Furstenberg in [6].

Definition 2.4.1 (joint)**.** *Consider two TDS* $(X, T), (Y, S)$*.* $J \subseteq X \times Y$ *is called a joint of* X *and* Y *if*

- J *is non-empty closed subset of* $X \times Y$*.*

- $\pi_1(J) = X,$ *where* $\begin{cases} \pi_1 \colon X \times Y \to X \\ \pi_1((x, y)) = x \end{cases}$
- $\pi_2(J) = Y,$ *where* $\begin{cases} \pi_2 \colon X \times Y \to Y \\ \pi_2((x, y)) = y \end{cases}$

Example 2.4.1. $(X \times Y, T \times S))$ *is trivial joint of* (X, T) *and* (Y, S).

Definition 2.4.2 (TDS Disjointness)**.** *If every joint of* (X, T) *and* (Y, S) *is equal to* $(X \times Y, T \times S)$*, we say* X *and* Y *are disjoint and we denote* $(X, T) \perp (Y, S)$.

We shall notice that this is not the first definition Furstenberg introduced. We'll see the first definition in the proposition below:

Proposition 2.4.1. *Let* (X, T) *and* (Y, S) *be two TDSs.*
Then $(X, T) \perp (Y, S)$ *iff*
if there exists a common extension (Z, R) *of* (X, T) *and* (Y, S) *with the factor maps* $\alpha \colon Z \to X$ *and* $\beta \colon Z \to Y$*, then there exists a factor map* $\gamma \colon Z \to X \times Y$ *such that* $\alpha = \pi_1 \gamma$ *and* $\beta = \pi_2 \gamma$.

Proof. If $(X, T) \perp (Y, S)$ and there exists factor maps $\alpha \colon Z \to X$ and $\beta \colon Z \to Y$, define

$$\begin{cases} \gamma \colon Z \to X \times Y \\ \gamma(z) = (\alpha(z), \beta(z)) \end{cases}$$

is a homomorphism between dynamical systems.

1. γ is continues and surjective sinve α and β are.

2. $\gamma \circ R = (T \times S) \circ \gamma$ since α and β are factor maps. so the diagram below commutes:

$$
\begin{tikzcd}
& X \arrow[rr, dashed, "T"] \arrow[dd, dashed, "\alpha"] & & X \arrow[dd, dashed, "\alpha"] \\
Y \arrow[rr, dotted, "S"] \arrow[dr, dotted, "\beta"] & & Y \arrow[dr, dotted, "\beta"] & \\
& Z \arrow[rr, "R"] \arrow[d, "\gamma:=(\alpha,\beta)"] \arrow[drr, bend left] \arrow[drr, bend right] \arrow[drr, phantom, "="] & & Z \arrow[d, "\gamma:=(\alpha,\beta)"] \\
& X \times Y \arrow[rr, "T\times S"] & & X \times Y
\end{tikzcd}
$$

conversely, If there exists factor maps that the diagrams above commute, consider a closed set $Z \subseteq X \times Y$ that project onto X and Y - Z be a joint of X and Y; And since Z is closed, let $R = T \times S\restriction_Z$ so that (Z, R) be a TDS (a subsystem of $(X \times Y, T \times S)$). define

$$
\begin{cases} \alpha \colon Z \to X \\ \alpha(z) = \pi_1(z) \end{cases} \qquad \begin{cases} \beta \colon Z \to Y \\ \beta(z) = \pi_2(z) \end{cases}
$$

It's not difficult to see that α and β are homomorphism of dynamical systems.

By hypothesis, exists a factor map $\begin{cases} \gamma \colon Z \to X \times Y \\ \gamma(z) = (\alpha(z), \beta(z)) = (\pi_1(z), \pi_2(z)) = z \end{cases}$. Since $Z \subseteq X \times Y$ and γ is surjective, $Z = X \times Y$ $\square$

Definition 2.4.3 (MDS Disjointness). *Let $(X, \mathscr{A}, \mu, T)$ and $(Y, \mathscr{B}, \nu, S)$ be two MDS. $(X, \mathscr{A}, \mu, T) \perp (Y, \mathscr{B}, \nu, S)$ or simply $X \perp Y$ if every measure in $M_{T\times S}(X \times Y)$ with the σ-algebra $\mathscr{A} \times \mathscr{B}$ is equal to $\mu \times \nu$ i.e.*

$$
\forall \eta \in M_{T\times S}(X \times Y) \colon \eta = \mu \times \nu
$$

This definition, again is not the first introduced definition.

Proposition 2.4.2. *Let $(X, \mathscr{A}, \mu, T)$ and $(Y, \mathscr{B}, \nu, S)$ be two MDSs.*
Then $X \perp Y$ iff

if there exists a common extension $(Z, \mathscr{C}, \eta, R)$ *of* X *and* Y *with the factor maps* $\alpha \colon Z \to X$ *and* $\beta \colon Z \to Y$*, then there exists a factor map* $\gamma \colon Z \to X \times Y$ *such that* $\alpha = \pi_1 \gamma$ *and* $\beta = \pi_2 \gamma$.

Proof. If $X \perp Y$ and $\alpha \colon Z \to X$ and $\beta \colon Z \to Y$ are homomorphisms, then define

$$\begin{cases} \gamma \colon Z \to X \times Y \\ \gamma(z) = (\alpha(z), \beta(z)) \end{cases}$$

- γ is measurable since α and β are.
- γ is η-invariant since α and β are:

$$\forall A \times B \in \mathscr{A} \times \mathscr{B} \colon \eta(\gamma^{-1}(A \times B)) = \eta(\alpha^{-1}A \cap \beta^{-1}B)$$

  now we wish to show that $\eta(\alpha^{-1}A \cap \beta^{-1}B) = \mu(A)\nu(B)$.
  Assume that $\eta(\alpha^{-1}A \cap \beta^{-1}B) \neq \mu(A)\nu(B)$. So define a new measure on $X \times Y$ by

$$\eta'(A \times B) = \eta(\alpha^{-1}A \cap \beta^{-1}B), \forall A \in \mathscr{A}, B \in \mathscr{B}$$

  It's not difficult to see that η' is $T \times S$-invariant since α and β are factor maps. Thus by assumption $(X \perp Y)$ we have $\eta'(A \times B) = \eta(\alpha^{-1}A \cap \beta^{-1}B) = \mu(A)\nu(B)$. this contradiction shows that our desire is fulfilled.
- $\forall z \in Z$:

$$\begin{aligned} (T \times S) \circ \gamma(z) &= (T \times S)(\alpha(z), \beta(z)) = (T\alpha(z), S\beta(z)) \\ &= (\alpha R(z), \beta R(z)) = \gamma \circ R(z) \end{aligned}$$

  Thus γ is a factor map.
  Conversely, Let consider the space $(X \times Y, \mathscr{A} \times \mathscr{B}, \eta, R)$ when $\eta \in M_{T \times S}(X \times Y)$.
  Define

$$\begin{cases} \alpha\colon X \times Y \to X \\ \alpha((x,y)) = x \end{cases} \quad \begin{cases} \beta\colon X \times Y \to Y \\ \beta((x,y)) = y \end{cases}$$

α and β are homomorphisms, so by hypothesis, there exist a homomorphism

$$\gamma\colon (X \times Y, \mathscr{A} \times \mathscr{B}, \eta, R) \to (X \times Y, \mathscr{A} \times \mathscr{B}, \mu \times \nu, T \times S)$$

such that $\pi_1(\gamma) = \alpha$ and $\pi_2(\gamma) = \beta$. Since γ is measure-preserving, we have

$$\forall A \times B \in \mathscr{A} \times \mathscr{B}\colon \eta(\gamma^{-1}(A \times B)) = \mu \times \nu(A \times B)$$

but

$$\eta(\gamma^{-1}(A \times B)) = \eta(\alpha^{-1}A \cap \beta^{-1}B) = \eta((A \times Y) \cap (X \times B)) = \eta(A \times B)$$

All in all,

$$\eta = \mu \times \nu$$

□

Theorem 2.4.1. *If (X,T) and (Y,S) is two TDSs, then*

$$X \perp Y \implies X \textit{ is minimal or } Y \textit{ is minimal}$$

Proof. First remember the definition (2.1.8). Then assume neither X nor Y be minimal. So let the minimal subsystem of X be $A \subsetneqq X$ and the minimal subsystem of Y be $B \subsetneqq Y$. Consider $Z := (A \times Y) \cup (X \times B)$.

- $Z \subsetneqq X \times Y$
- Z is closed since A and B are closed.
- Z is project onto X and Y.

Thus Z is a non-trivial joint of X and Y so that $X \not\perp Y$. All in all, X is minimal or Y is minimal. $\square$

Proposition 2.4.3. *For two TDSs* (X, T) *and* (Y, S) *we have*

1. *If* $X \perp Y$*, any factor of* X *is disjoint from* Y*.*
2. *If* $X \perp Y$*, any factor of* X *is disjoint from any factor of* Y*.*
3. *Recalling the definition* (2.3.5)*,* $X \perp Y$ *iff* $\tilde{X} \perp \tilde{Y}$
4. *If* $X \perp Y$*, then* $\tilde{X} \perp Y$

Proof. 1. Let us consider (Z, R) as a factor of (X, T) and the commutative diagram below:

$$\begin{tikzcd} X \arrow[r, "T"] \arrow[d, "\pi"'] \arrow[dr, dashed, bend left, "\pi\circ T"] \arrow[dr, dashed, bend right, "R\circ\pi"'] & X \arrow[d, "\pi"] \\ Z \arrow[r, "R"] & Z \end{tikzcd}$$

Let $J \subseteq Z \times Y$ be a closed subset projecting onto Z and Y.

Consider the continues surjective map $X \times Y \xrightarrow{\pi \times id_Y} Z \times Y$.

$(\pi \times id_y)^{-1} J \subseteq X \times Y$ is

- closed.
- project onto X since $\pi_1\Big((\pi \times id_y)^{-1} J\Big) = \pi^{-1}\Big(\pi_1 J\Big) = \pi^{-1} Z = X$
- project onto Y.

Thus, since $X \perp Y$ we have $(\pi \times id_y)^{-1} J = X \times Y$ so that $J = Z \times Y$.

2. Considering (Z, R) as a factor of (X, T), by 1 we have $(Z, R) \perp (X, T)$. Now that $(Z, R) \perp (X, T)$, considering (W, K) as a factor of (Y, S), by using 1, we have $(Z, R) \perp (W, K)$.
3. Consider the commutative diagram below with the notations of definition (2.3.5)

and J be a joint of $\tilde{X}$ and $\tilde{Y}$:

$$\begin{tikzcd}
J \subseteq & \tilde{X}\times\tilde{Y} \arrow[r, "\sigma_T\times\sigma_S"] \arrow[d, "\pi_1\times\pi_1"] \arrow[dr, dashed] \arrow[dr, phantom, "="] & \tilde{X}\times\tilde{Y} \arrow[d, "\pi_1\times\pi_1"] \\
& X\times Y \arrow[r, "T\times S"] & X\times Y
\end{tikzcd}$$

$\pi_1 \times \pi_1$ is an open map so consider the set $\pi_1 \times \pi_1(J) \subseteq X \times Y$:

- it's closed.
- projects onto X Since $\pi_1(\tilde{X}) = X$.
- projects onto Y.

Thus $\pi_1 \times \pi_1(J) = X \times Y$ therefore $J = \tilde{X} \times \tilde{Y}$.

Conversely, by 2, if $\tilde{X} \perp \tilde{Y}$ we have $X \perp Y$.

4. By 3, if $X \perp Y$ then $\tilde{X} \perp \tilde{Y}$. By 1, if $\tilde{X} \perp \tilde{Y}$ then $\tilde{X} \perp Y$.

$\square$

Theorem 2.4.2. *Consider two TDS (X,T) and (Y,S) that $X \perp Y$. By the theorem above, let (Y,S) be minimal. Assuming that Y is non-trivial system, then the set of recurrent points of X is dense.*

Proof. The proof has two steps:

1. First assume that T is homeomorphism.

   Let $\Omega(T)$ be the closure of the set of all recurrent points of (X,T). Assume that $\Omega(T) \neq X$ so that there exists $x \in X \setminus \Omega(T)$.

   There exists $U \overset{\text{open}}{\subseteq} X$ such that

   $$x \in U \text{ and } \forall i \neq j \colon T^i U \cap T^j U = \emptyset$$

   otherwise, if for any neighborhood of x, there is $i, j \in \mathbb{N}$ such that $T^i(U) \cap T^j(U) \neq \emptyset$, then we have a point $x' \in T^i(U) \cap T^j(U)$ so $T^i(x) = T^j(x)$ or equivalently,

$T^{i-j}(x) = x$ so that $\mathrm{orb}_T(x')$ is periodic and indeed $x' \in \Omega \cap U$ thus $x \in \Omega(T)$.

Y is non-trivial i.e. it consists of at least two points. Let $y \in Y$ be arbitrary and define:

$$J := \overline{\left(\bigcup_{n\in\mathbb{Z}} T^n U \times \{S^n(y)\}\right)} \cup \left(\left(X \setminus \bigcup_{n\in\mathbb{Z}} T^n U\right) \times Y\right)$$

$J \subseteq X \times Y$ is a joint of X and Y since

- $J \subseteq X \times Y$ is closed.
- $\pi_1(J) = X$ since $\left(\bigcup_{n\in\mathbb{Z}} T^n U\right) \cup \left(X \setminus \bigcup_{n\in\mathbb{Z}} T^n U\right) = X$.
- $\pi_2(J) = Y$.

By the hypothesis that $X \perp Y$ we have $J = X \times Y$. considering the set $U \times (Y \setminus \{y\})$ we have:

$$\begin{aligned}
&U \times (Y \setminus \{y\}) \subseteq X \times Y = J \\
&U \times (Y \setminus \{y\}) \subseteq \overline{\left(\bigcup_{n\in\mathbb{Z}} T^n U \times \{S^n(y)\}\right)} \qquad \text{Since } U \cap \left(X \setminus \bigcup_{n\in\mathbb{Z}} T^n U\right) = \emptyset \\
&\exists n \in \mathbb{N} \setminus \{0\} \text{ s.t. } \Big(U \times (Y \setminus \{y\})\Big) \cap \Big(T^n U \times \{S^n(y)\}\Big) \neq \emptyset \qquad \text{Notice that } \Big(U \times (Y \setminus \{y\})\Big) \cap \overline{\Big(U \times \{y\}\Big)} = \emptyset \\
&\implies U \cap T^n U \neq \emptyset
\end{aligned}$$

by this contradiction we have $\Omega(T) = X$.

2. In general case, by proposition (2.4.3), part (4) since $X \perp Y$ we have $\tilde{X} \perp Y$. Thus by step 1, $\Omega(\sigma_T)$ is dense in $\tilde{X}$. By corollary (2.3.1) $\Omega(T)$ is dense in (X, T).

$\square$

Chapter 3

Combinatorial and Topological Entropy

3.1 Combinatorial entropy

If we have a set X, we'll denote C_X the set of all finite covers of X. If $\mathcal{U} \in C_X$, $\mathcal{N}(\mathcal{U})$ is denoted as the minimum cardinality of subcovers of $\mathcal{U}$: $\mathcal{N}(\mathcal{U}) := min\{\#\mathcal{V} \mid \mathcal{V} \in C_X,\ \mathcal{V} \subseteq \mathcal{U}\}$. Now consider a map $T \colon X \to X$. For a given integer $M \leq N$ and $\mathcal{U} \in C_X$, let $\mathcal{U}_M^N := \bigvee_{n=M}^{N} T^{-n}(\mathcal{U})$.

Definition 3.1.1 (Combinatiorial Entropy)**.** *The combinatorial entropy of $\mathcal{U}$ with respect to $T \colon X \to X$ is*

$$h_c(\mathcal{U}, T) := \lim_{n\to\infty} \frac{1}{n} \log(\mathcal{N}(\mathcal{U}_0^{n-1})) \tag{3.1}$$

Example 3.1.1. *Let X be a set and $\mathcal{P} \in C_X$ be a partition of X, consider the map $id \colon X \to X$ we have $\mathcal{N}(\mathcal{P}) = \#\mathcal{P}$ and $\mathcal{P}_0^n = \mathcal{P}$ for each $n \in \mathbb{Z}$.*
So due to definition 3.1, we have: $h_c(\mathcal{P}, id) = \lim_{n\to\infty} \frac{1}{n} \log(\#\mathcal{P}) = 0$

Proposition 3.1.1. *The limit define in 3.1 always exists.*

Proof. Let $a_m := \mathcal{N}(\mathcal{U}_0^{n-1}))$ we proof that $a_{m+n} \leq a_m a_n$. First notice that

$$\begin{aligned}
\mathcal{U}_0^{m+n-1} &= \bigvee_{i=0}^{m+n-1} T^{-i}\mathcal{U} \\
&= (\bigvee_{i=0}^{m-1} T^{-i}\mathcal{U}) \bigvee (\bigvee_{i=m}^{m+n-1} T^{-i}\mathcal{U}) \\
&= (\bigvee_{i=0}^{m-1} T^{-i}\mathcal{U}) \bigvee T^{-m}(\bigvee_{i=0}^{n-1} T^{-i}\mathcal{U}) \\
&= \mathcal{U}_0^{m-1} \bigvee T^{-m}\mathcal{U}_0^{n-1}
\end{aligned}$$

So if U_m is a subcover of $\mathcal{U}_0^{m-1}$ and U_n is a subcover of $\mathcal{U}_0^{n-1}$, both with minimal cardinality, $U_m \bigvee T^{-m}U_n$ is a subcover of $\mathcal{U}_0^{m+n-1}$.
Furthermore $\#(U_m \bigvee T^{-m}U_n) = a_m a_n$. Thus $a_{m+n} \leq a_m a_n$ □

3.2 Topological entropy

Now we move to topological entropy. If X be a topological space, let us consider C_X° the set of all open covers of X.

Definition 3.2.1. *If $\mathcal{U} \in C_X^\circ$ and T be a map on X, consider*

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

then we define an entropy of an open cover $\mathcal{U}$ as:

$$h_{\text{top}}(\mathcal{U}, T) := h_c(\mathcal{U}, T) = \lim_{n\to\infty} \frac{1}{n} H(\mathcal{U}_0^{n-1}) \tag{3.2}$$

Definition 3.2.2. *We define the topological entropy of T, as:*

$$h_{\text{top}}(T) := \sup\{h_{\text{top}}(\mathcal{U}, T) \mid \mathcal{U} \in C_X^\circ\} \tag{3.3}$$

Remark 3.2.1. *In some parts we writes $h(\mathcal{U},T)$ as for $h_{\text{top}}(\mathcal{U},T)$ and some books writes $h_{\text{top}}(T,\mathcal{U})$. Thus, They are all notations for one phenomenon.*

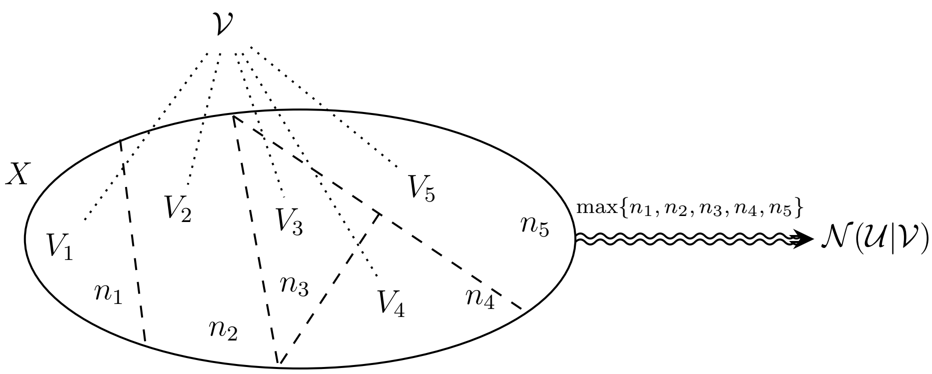
Conditional Topological Entropy was introduced in 1976 by Michal Misiurewicz in [12].

Now let us study conditional topological entropy: For a TDS (X,T), covers $\mathcal{U},\mathcal{V} \in C_X$, and any subset $\emptyset \neq Y \subseteq X$, define

$$\mathcal{N}(\mathcal{U} \mid Y) := \min\left\{ \#\mathcal{U}' \mid (\mathcal{U})' \subseteq \mathcal{U} \text{ and } Y \subseteq \bigcup_{U' \in \mathcal{U}'} U' \right\} \tag{3.4}$$

and

$$\mathcal{N}(\mathcal{U} \mid \mathcal{V}) := \max\{\mathcal{N}(\mathcal{U} \mid V) \mid V \in \mathcal{V}\} \tag{3.5}$$



$n_i := \mathcal{N}(\mathcal{U}|V_i)$ =At least how many elements of $\mathcal{U}$ are going to cover V_i

Example 3.2.1.
- $\mathcal{N}(\mathcal{U} \mid X) = \mathcal{N}(\mathcal{U})$
- $\mathcal{N}(\mathcal{U} \mid \{X,\emptyset\}) = \mathcal{N}(\mathcal{U})$

Theorem 3.2.1. *Considering a TDS (X,T), covers $\mathcal{U},\mathcal{V},\mathcal{U}',\mathcal{V}' \in C_X$*

1. $\mathcal{U} \succeq \mathcal{U}' \implies \mathcal{N}(\mathcal{U} \mid \mathcal{V}) \geq \mathcal{N}(\mathcal{U}' \mid \mathcal{V})$
2. $\mathcal{V} \succeq \mathcal{V}' \implies \mathcal{N}(\mathcal{U} \mid \mathcal{V}) \leq \mathcal{N}(\mathcal{U} \mid \mathcal{V}')$
3. $\mathcal{U}' \succeq \mathcal{U}, \mathcal{V} \succeq \mathcal{V}' \implies \mathcal{N}(\mathcal{U} \mid \mathcal{V}) \leq \mathcal{N}(\mathcal{U}' \mid \mathcal{V}')$

4. *$\mathcal{N}(T^{-1}\mathcal{U} \mid T^{-1}\mathcal{V}) \leq \mathcal{N}(\mathcal{U} \mid \mathcal{V})$.*
5. *$\mathcal{N}(T^{-1}\mathcal{U} \mid T^{-1}\mathcal{V}) = \mathcal{N}(\mathcal{U} \mid \mathcal{V})$, if T is surjective.*
6. *$\mathcal{N}(\mathcal{U} \vee \mathcal{U}' \mid \mathcal{V} \vee \mathcal{V}') \leq \mathcal{N}(\mathcal{U} \mid \mathcal{V})\mathcal{N}(\mathcal{U}' \mid \mathcal{V}')$*
7. *$\mathcal{N}(\mathcal{U} \vee \mathcal{U}'|\mathcal{V}) \leq \mathcal{N}(\mathcal{U}|\mathcal{V})\mathcal{N}(\mathcal{U}'|\mathcal{U} \vee \mathcal{V})$*
8. *$\mathcal{N}(\mathcal{U}|\mathcal{V} \vee \mathcal{V}') \leq \mathcal{N}(\mathcal{U}|\mathcal{V})\mathcal{N}(\mathcal{U}|\mathcal{V}')$*
9. *$\mathcal{N}(\mathcal{U}|\mathcal{V}) \leq \mathcal{N}(\mathcal{U}|\mathcal{V}')\mathcal{N}(\mathcal{V}'|\mathcal{V})$*
10. *$\mathcal{N}(\mathcal{U}|\mathcal{V}) = \mathcal{N}(\pi^{-1}\mathcal{U}|\pi^{-1}\mathcal{V})$, where $\pi\colon (Y,S) \to (X,T)$ is a factor map.*

Proof. 1. For any $\alpha = \{A_1, A_2, \ldots, A_l\} \subseteq \mathcal{U}$ that $V \subseteq \bigcup_{i=1}^{l} A_i$, let $A'_i \in \mathcal{U}'$ be a member that $A_i \subseteq A'_i$. Thus,

$$V \subseteq \bigcup_{i=1}^{l} A_i \subseteq \bigcup_{i=1}^{l} A'_i$$

Let $\alpha' := \{A'_1, A'_2, \ldots, A'_l\}$. Thus, for any $\alpha \subseteq \mathcal{U}$ there is $\alpha' \subseteq \mathcal{U}'$ with $\#\alpha' \leq \#\alpha$ Therefore:

$$\min\left\{\#\alpha' \,\middle|\, \alpha' \subseteq \mathcal{U}' \text{ and } V \subseteq \bigcup_{A' \in \alpha'} A'\right\} \\ \leq \min\left\{\#\alpha \,\middle|\, \alpha \subseteq \mathcal{U} \text{ and } V \subseteq \bigcup_{A \in \alpha} A\right\}$$

In other words, for each $V \in \mathcal{V}$ we have:

$$\mathcal{N}(\mathcal{U}|V) \geq \mathcal{N}(\mathcal{U}'|V)$$

So,

$$\mathcal{N}(\mathcal{U}|\mathcal{V}) \geq \mathcal{N}(\mathcal{U}'|\mathcal{V})$$

2. Consider any $V \in \mathcal{V}$. let $V' \in \mathcal{V}'$ be such that $V \subseteq V'$. since $V \subseteq V'$, for any $\alpha \subseteq \mathcal{U}$ that $V' \subseteq \bigcup_{A\in\alpha} A$, we have:

$$V \subseteq V' \subseteq \bigcup_{A \in \alpha} A$$

So that,

$$\left\{\alpha \subseteq \mathcal{U} \colon V' \subseteq \bigcup_{A\in\alpha} A\right\} \subseteq \left\{\alpha \subseteq \mathcal{U} \colon V \subseteq \bigcup_{A\in\alpha} A\right\}$$

Thus,

$$\min\left\{\alpha \subseteq \mathcal{U} \colon V \subseteq \bigcup_{A\in\alpha} A\right\} \leq \min\left\{\alpha \subseteq \mathcal{U} \colon V' \subseteq \bigcup_{A\in\alpha} A\right\}$$

Therefore,

$$\mathcal{N}(\mathcal{U}|V) \leq \mathcal{N}(\mathcal{U}|V')$$

Thus,

$$\mathcal{N}(\mathcal{U}|\mathcal{V}) \leq \mathcal{N}(\mathcal{U}|\mathcal{V}')$$

since for every $V \in \mathcal{V}$ there is $V' \in \mathcal{V}'$ that $\mathcal{N}(\mathcal{U}|V) \leq \mathcal{N}(\mathcal{U}|V')$.

3. Its immediate result of 1,2.

4. For any $\alpha \in \mathcal{U}$ by the facts that $A \subseteq B \implies T^{-1}A \subseteq T^{-1}B$ and $\#T^{-1}\alpha \leq \#\alpha$, for any $V \in \mathcal{V}$,

$$\min\{\#T^{-1}\alpha|T^{-1}\alpha \subseteq T^{-1}\mathcal{U}, T^{-1}V \subseteq \bigcup_{A\in\alpha} T^{-1}A\} \leq \min\{\#\alpha|\alpha \subseteq \mathcal{U}, V \subseteq \bigcup_{A\in\alpha} A\}$$

Thus

$$\mathcal{N}(T^{-1}\mathcal{U}|T^{-1}V) \leq \mathcal{N}(\mathcal{U}|V)$$

Since for any $V \in \mathcal{V}$ the inequality is true, we have:

$$\mathcal{N}(T^{-1}\mathcal{U}|T^{-1}\mathcal{V}) \leq \mathcal{N}(\mathcal{U}|\mathcal{V})$$

5. if T is surjective, $T^{-1}A \subseteq T^{-1}B \iff A \subseteq B$ for any $A \subseteq X$. Besides for every

$\alpha \in C_X^\circ$ we have $\#\alpha = \#T^{-1}\alpha$. Thus, for any $V \in \mathcal{V}$:

$$\left\{\#T^{-1}\alpha \middle| T^{-1}\alpha \subseteq T^{-1}\mathcal{U},\ T^{-1}V \subseteq \bigcup_{A\in\alpha} T^{-1}A\right\}$$
$$= \left\{\#\alpha \middle| \alpha \subseteq \mathcal{U},\ V \subseteq \bigcup_{A\in\alpha} A\right\}$$
$$\implies \mathcal{N}(\mathcal{U}|V) = \mathcal{N}(T^{-1}\mathcal{U}|T^{-1}V)$$

Since we have the equality above for every $V \in \mathcal{V}$:

$$\mathcal{N}(\mathcal{U}|\mathcal{V}) = \mathcal{N}(T^{-1}\mathcal{U}|T^{-1}\mathcal{V})$$

6. Since for any $\alpha \in \mathcal{U} \vee \mathcal{U}'$ there is $\beta \in \mathcal{U}$ and $\beta' \in \mathcal{U}'$ such that $\alpha \subseteq \beta \vee \beta'$. we have the inequality:

$$\#\alpha \leq \#\big(\beta \vee \beta'\big) \leq \#\beta\#\beta'$$

Thus, for any $V \cap V' \in \mathcal{V} \vee \mathcal{V}'$:

$$\min\left\{\#\alpha \middle| \alpha \subseteq \mathcal{U} \vee \mathcal{U}', V \cap V' \subseteq \bigcup_{A\in\alpha} A\right\}$$
$$\leq \min\left\{\#\beta\#\beta' \middle| \beta \subseteq \mathcal{U}, \beta' \subseteq \mathcal{U}', V \subseteq \bigcup_{B\in\beta} B, V' \subseteq \bigcup_{B'\in\beta'} B'\right\}$$
$$= \min\left\{\#\beta \middle| \beta \subseteq \mathcal{U}, V \subseteq \bigcup_{B\in\beta} B\right\} \min\left\{\#\beta' \middle| \beta' \subseteq \mathcal{U}', V' \subseteq \bigcup_{B'\in\beta'} B'\right\}$$
$$= \mathcal{N}(\mathcal{U}|V)\mathcal{N}(\mathcal{U}'|V')$$

Since for any $V \in \mathcal{V}$ and $V' \in \mathcal{V}'$ the inequality above is true:

$$\mathcal{N}(\mathcal{U} \vee \mathcal{U}'|\mathcal{V} \vee \mathcal{V}') \leq \mathcal{N}(\mathcal{U}|\mathcal{V})\mathcal{N}(\mathcal{U}'|\mathcal{V}')$$

7. Fix $V \in \mathcal{V}$. There exists $U_1, U_2, \ldots, U_l \in \mathcal{U}$ such that

$$V \subseteq \bigcup_{i=1}^{l} U_i$$

where $l := \mathcal{N}(\mathcal{U}|V)$.

For each $U_i, 1 \leq i \leq l$, consider the set $U_i \cap V$. For this set there exist some $U'_{i1}, U'_{i2}, \ldots, U'_{ir} \in \mathcal{U}'$ such that

$$U_i \cap V \subseteq \bigcup_{j=1}^{r} U'_{ij}$$

where $r := \mathcal{N}(\mathcal{U}'|U_i \cap V)$.

Now define the set:

$$\alpha := \{U_i \cap U'_{ij} \colon 1 \leq i \leq l, 1 \leq j \leq r\}$$

- $\alpha \subseteq \mathcal{U} \vee \mathcal{U}'$
- $\#\alpha \leq l.r = \mathcal{N}(\mathcal{U}|V)\mathcal{N}(\mathcal{U}'|U_i \cap V)$
- $V \subseteq \bigcup_{i=1}^{l} \bigcup_{j=1}^{r} U_i \cap U_{ij} = \bigcup_{A \in \alpha} A$

Thus, for any $V \in \mathcal{V}$:

$$\mathcal{N}(\mathcal{U} \vee \mathcal{U}'|V) \leq \mathcal{N}(\mathcal{U}|V)\mathcal{N}(\mathcal{U}'|U_i \cap V)$$

Therefore,

$$\mathcal{N}(\mathcal{U} \vee \mathcal{U}'|\mathcal{V}) \leq \mathcal{N}(\mathcal{U}|\mathcal{V})\mathcal{N}(\mathcal{U}'|U_i \cap \mathcal{V})$$

8. Take $\mathcal{U}' := \mathcal{U}$ in part (6).
9. By (1), (2), and (7):

$$\mathcal{N}(\mathcal{U}|\mathcal{V}) \leq \mathcal{N}(\mathcal{V}' \vee \mathcal{U}|\mathcal{V}) \leq \mathcal{N}(\mathcal{V}'|\mathcal{V})\mathcal{N}(\mathcal{U}|\mathcal{U} \vee \mathcal{V}') \leq \mathcal{N}(\mathcal{U}|\mathcal{V}')\mathcal{N}(\mathcal{V}'|\mathcal{V})$$

10. The proof is the same as part (4), for any $V \in \mathcal{V}$:

$$\begin{aligned}
\mathcal{N}(\pi^{-1}\mathcal{U}|\pi^{-1}V) &= \min\{\#\pi^{-1}\alpha|\pi^{-1}\alpha \subseteq \pi^{-1}\mathcal{U}, \pi^{-1}V \subseteq \bigcup_{A\in\alpha} \pi^{-1}A\} \\
&= \min\{\#\alpha|\alpha \subseteq \mathcal{U}, V \subseteq \bigcup_{A\in\alpha} A\} \\
&= \mathcal{N}(\mathcal{U}|V)
\end{aligned}$$

$$\implies \mathcal{N}(\pi^{-1}\mathcal{U}|\pi^{-1}\mathcal{V}) = \mathcal{N}(\mathcal{U}|\mathcal{V})$$

$\square$

Remark 3.2.2. *In part* (5) *of theorem* (3.2.1)*, the assumption that* T *is surjective is crucial. Consider the TDS* $[-1,1] \subseteq \mathbb{R}$ *and the map* $\begin{cases} T\colon [-1,1] \to [-1,1] \\ T(x) = 0 \end{cases}$ *Let*

$$\mathcal{U} := \{[-1,0), (-\frac{1}{2}, 1), (\frac{1}{2}, 1]\}$$

and

$$\mathcal{V} := \{[-1,1), (-\frac{3}{4}, 1]\}$$

be two open cover of $[-1,1]$*. Let us compute* $\mathcal{N}(\mathcal{U}|\mathcal{V})$ *and* $\mathcal{N}(T^{-1}\mathcal{U}|T^{-1}\mathcal{V})$*:*

First notice that $T^{-1}\mathcal{U} = \{[-1,1]\}$ *and* $T^{-1}\mathcal{V} = \{[-1,1]\}$

$$\begin{aligned}
\mathcal{N}(\mathcal{U}|[-1,1)) &= 2 \\
\mathcal{N}(\mathcal{U}|(-\frac{3}{4}, 1]) &= 3 \\
\mathcal{N}(\mathcal{U}|\mathcal{V}) &= \max\{2,3\} = 3 \\
\mathcal{N}(T^{-1}\mathcal{U}|T^{-1}\mathcal{V}) &= 1
\end{aligned}$$

Thus,

$$\mathcal{N}(\mathcal{U}|\mathcal{V}) \neq \mathcal{N}(T^{-1}\mathcal{U}|T^{-1}\mathcal{V})$$

Definition 3.2.3. *If $\mathcal{U}, \mathcal{V} \in C_X^{\circ}$ and $T : X \to X$ be continues surjective function. consider*

$$H(\mathcal{U}|\mathcal{V}) := \log(\mathcal{N}(\mathcal{U}|\mathcal{V}))$$

then we define an entropy of an open cover $\mathcal{U}$ relative to the cover $\mathcal{V}$ as:

$$h_{\text{top}}(\mathcal{U}, T|\mathcal{V}) := \lim_{n\to\infty} \frac{1}{n} H(\mathcal{U}_0^{n-1}|\mathcal{V}_0^{n-1}) \tag{3.6}$$

In this section we denote $h(.)$ as for $h_{\text{top}}(.)$.

Remark 3.2.3. *The limit* (3.6) *exists and is equal to* $\inf_{n\geq 1} \frac{1}{n} H(\mathcal{U}_0^{n-1}|\mathcal{V}_0^{n-1})$, *since by part* (6) *of theorem* (3.2.1),

$$\begin{aligned}
&H\big(\mathcal{U}_0^{m+n-1}|\mathcal{V}_0^{m+n-1}\big)\\
&\qquad = H\left(\bigvee_{i=0}^{m-1} T^{-i}\mathcal{U} \vee \bigvee_{i=m}^{m+n-1} T^{-i}\mathcal{U} \,\middle|\, \bigvee_{i=0}^{m-1} T^{-i}\mathcal{V} \vee \bigvee_{i=m}^{m+n-1} T^{-i}\mathcal{V}'\right)\\
&\qquad = H\left(\mathcal{U}_0^{m-1} \vee T^{-m}\mathcal{U}_0^{n-1} \,\middle|\, \mathcal{V}_0^{m-1} \vee T^{-m}\mathcal{V}_0^{n-1}\right)\\
&\qquad \leq H(\mathcal{U}_0^{m-1}|\mathcal{V}_0^{m-1}) + H(T^{-m}\mathcal{U}_0^{n-1}|T^{-m}\mathcal{V}_0^{n-1})\\
&\qquad \leq H(\mathcal{U}_0^{m-1}|\mathcal{V}_0^{m-1}) + H(\mathcal{U}_0^{n-1}|\mathcal{V}_0^{n-1}) \qquad \textit{By part (4)}
\end{aligned}$$

So H is subadditive and consequently, the limit exists and equals to $\inf_{n\geq 1} \frac{1}{n} H(\mathcal{U}_0^{n-1}|\mathcal{V}_0^{n-1})$.

Theorem 3.2.2. *Consider the TDS (X, T) and $\mathcal{U}, \mathcal{U}', \mathcal{V}, \mathcal{V}' \in C_X^{\circ}$:*

1. $H(\mathcal{U}|\mathcal{V}) \leq H(\mathcal{U})$
2. $\mathcal{U} \succeq \mathcal{U}' \implies h(\mathcal{U}, T|\mathcal{V}) \geq h(\mathcal{U}', T|\mathcal{V})$
3. $\mathcal{V} \succeq \mathcal{V}' \implies h(\mathcal{U}, T|\mathcal{V}) \leq h(\mathcal{U}, T|\mathcal{V}')$
4. $\mathcal{U} \succeq \mathcal{U}'$ *and* $\mathcal{V}' \succeq \mathcal{V} \implies h(\mathcal{U}, T|\mathcal{V}) \geq h(\mathcal{U}', T|\mathcal{V}')$
5. $h(T^{-1}\mathcal{U}, T|T^{-1}\mathcal{V}) \leq h(\mathcal{U}, T|\mathcal{V})$
6. $h(T^{-1}\mathcal{U}, T|T^{-1}\mathcal{V}) = h(\mathcal{U}, T|\mathcal{V})$, *where T is surjective.*
7. $h(\mathcal{U} \vee \mathcal{U}', T|\mathcal{V} \vee \mathcal{V}') \leq h(\mathcal{U}, T|\mathcal{V}) + h(\mathcal{U}', T|\mathcal{V}')$

8. $h(\mathcal{U} \vee \mathcal{U}', T|\mathcal{V}) \leq h(\mathcal{U}, T|\mathcal{V}) + h(\mathcal{U}', T|\mathcal{U} \vee \mathcal{V})$
9. $h(\mathcal{U}, T|\mathcal{V} \vee \mathcal{V}') \leq h(\mathcal{U}, T|\mathcal{V}) + h(\mathcal{U}, T|\mathcal{V}')$
10. $h(\mathcal{U}, T|\mathcal{V}) \leq h(\mathcal{U}, T|\mathcal{V}') + h(\mathcal{V}', T|\mathcal{V})$
11. $h(\pi^{-1}\mathcal{U}, S|\pi^{-1}\mathcal{V}) = h(\mathcal{U}, T|\mathcal{V})$, *where* $\pi\colon (Y, S) \to (X, T)$ *is a factor map.*

Proof. All the parts are results of theorem(3.2.1).

1. By (2), let $\mathcal{V}' := \{X, \emptyset\}$, we have:

$$H(\mathcal{U}|\mathcal{V}) = \log \mathcal{N}(\mathcal{U}|\mathcal{V}) \leq \log \mathcal{N}(\mathcal{U}|\mathcal{V}') = \log \mathcal{N}(\mathcal{U}) = H(\mathcal{U})$$

2. It's a direct result of (1).
3. It's a direct result of (2).
4. It's a direct result of (3).
5. It's a direct result of (4).
6. It's a direct result of (5).
7. It's a direct result of (6).
8. It's a direct result of (7).
9. It's a direct result of (8).
10. It's a direct result of (9).
11. It's a direct result of (10).

□

Definition 3.2.4. *The conditional entropy pf the system* (X, T) *relative to the cover* $\mathcal{V}$ *is defined as*

$$h(T|\mathcal{V}) := \sup_{\mathcal{U} \in C_X^{\circ}} h(T, \mathcal{U}|\mathcal{V})$$

Note that $h(T||\{X\}) = h(T)$ and $h(T, \mathcal{U}|\{X\}) = h(T, \mathcal{U})$ so we have:

Theorem 3.2.3. *For some TDSs* (X, T) *and* (Y, S)*, and the open covers* $\mathcal{U}$ *and* $\mathcal{U}'$ *we have:*

1. *$H(\mathcal{U}) \leq \log \#\mathcal{U}$*
2. *$\mathcal{U} \succeq \mathcal{U}' \implies h(\mathcal{U}, T) \succeq h(\mathcal{U}', T)$*
3. *$h(T^{-1}\mathcal{U}, T) \leq h(\mathcal{U}, T)$*
4. *$h(T^{-1}\mathcal{U}, T) = h(\mathcal{U}, T)$ if T is surjective.*
5. *$h(\mathcal{U} \vee \mathcal{U}', T) \leq h(\mathcal{U}, T) + h(\mathcal{U}', T)$*
6. *$h(\mathcal{U} \vee \mathcal{U}', T) \leq h(\mathcal{U}, T) + h(\mathcal{U}', T|\mathcal{U})$*
7. *$h(\mathcal{U}, T) \leq h(\mathcal{U}, T|\mathcal{V}) + h(\mathcal{V}, T)$*
8. *$h(\pi^{-1}\mathcal{U}, S) = h(\mathcal{U}, T)$, where $\pi\colon Y \to X$ is a factor map.*
9. *$h(T|T^{-1}\mathcal{U}) = h(T|\mathcal{U})$, if T is a homeomorphism.*
10. *$h(T|\mathcal{U}) \leq h(S|\pi^{-1}\mathcal{U})$, when $\pi\colon (Y, S) \to (X, T)$ is a factor map.*
11. *$h(T) \leq h(S)$, when $\pi\colon (Y, S) \to (X, T)$ is a factor map.*

Proof. Almost all of the parts are consequences of theorem (3.2.2).

1. $H(\mathcal{U}) = \log \mathcal{N}(\mathcal{U}) \leq \log \#\mathcal{U}$
2. Let $\mathcal{V} = \{X\}$ in (2).
3. Let $\mathcal{V} = \{X\}$ in (5).
4. Let $\mathcal{V} = \{X\}$ in (6).
5. Let $\mathcal{V} = \mathcal{V}' = \{X\}$ in (7).
6. Let $\mathcal{V} = \{X\}$ in (8).
7. Let $\mathcal{V} := \{X\}$, $\mathcal{V}' := \mathcal{V}$ in (10).
8. Let $\mathcal{V} = \{X\}$ in (11).
9. For any $\mathcal{V} \in C_X^\circ$, by (6) we have $h(T, \mathcal{V}|\mathcal{U}) = h(T, T^{-1}\mathcal{V}|T^{-1}\mathcal{U})$, Thus,

$$\{h(T, \mathcal{V}|\mathcal{U}) : \mathcal{V} \in C_X^\circ\} \subseteq \{h(T, \mathcal{V}'|T^{-1}\mathcal{U}) : \mathcal{V}' \in C_X^\circ\}$$

   Since for any $\mathcal{V} \in C_X^\circ$, $T^{-1}\mathcal{V} \in C_X^\circ$. Thus, $h(T|\mathcal{U}) \leq h(T|T^{-1}\mathcal{U})$.

   Conversely, by the fact that T is homeomorphism for any $\mathcal{V} \in C_X^\circ$ we have $T\mathcal{V} \in C_X^\circ$ and so by (6):

$$\{h(T, \mathcal{V}|T^{-1}\mathcal{U}) : \mathcal{V} \in C_X^\circ\} \subseteq \{h(T, \mathcal{V}'|\mathcal{U}) : \mathcal{V}' \in C_X^\circ\}$$

since $T\mathcal{V} \in C_X^\circ$. Thus $h(T|T^{-1}\mathcal{U}) \leq h(T|\mathcal{U})$.

10. By (11), for any $\mathcal{V} \in C_X^\circ$, $h(T, \mathcal{V}|\mathcal{U}) = h(S, \pi^{-1}\mathcal{V}|\pi^{-1}\mathcal{U})$, we have:

$$\{h(T, \mathcal{V}|\mathcal{U}) : \mathcal{V} \in C_X^\circ\} \subseteq \{h(T, \mathcal{V}'|\pi^{-1}\mathcal{U}) : \mathcal{V}' \in C_Y^\circ\}$$

Thus,

$$h(T|\mathcal{U}) \leq h(S|\pi^{-1}\mathcal{U})$$

11. Let $\mathcal{U} = \{X\}$ in (10).

□

Remark 3.2.4. *By* (11)*, we found out that each factor of a system has less entropy than the main system. It's natural to thought of a maximal zero-entropy factor of a system. It is called the* ***Pinsker factor*** *of system*[1]*.*

In theorem (3.2.3), part (8) and part (8), we saw a proof of the below propositions, although there's another proofs base on natural extension introduced in (2.3.5). These kind of proofs helps us in proving other famous and fantastic theorems in TDSs.

Proposition 3.2.1. *Let $\pi\colon (Y, S) \to (X, T)$ be a factor map between two TDSs and $\mathcal{U} \in C_X^\circ$. Then, π preserve the entropy of the cover $\mathcal{U}$ i.e.*

$$h_{\text{top}}(T, \mathcal{U}) = h_{\text{top}}(S, \pi^{-1}\mathcal{U})$$

Proof. Let $\mathcal{U} = \{U_1, U_2, \ldots, U_n\}$.

- For any cover $\mathcal{V} = \{V_1, \ldots, V_n\} \in C_X^\circ$,

$$\mathcal{N}(\mathcal{V}) = \mathcal{N}(\pi^{-1}\mathcal{V})$$

[1]For more details see [2]

since, any subcover α of $\mathcal{V}$, give us the cover $\pi^{-1}\alpha$ of $\pi^{-1}\mathcal{V}$. and any subcover $\beta = \{\pi^{-1}V_{n_1}, \ldots, V_{n_m}\}$ $(n_i \leq n, m \leq n)$ of $\pi^{-1}\mathcal{V}$, gives us the subcover $\{V_{n_1}, \ldots, V_{n_m}\}$ of $\mathcal{V}$.

- Consider $\mathcal{N}((\pi^{-1}\mathcal{U})_0^{n-1})$ we have:

$$\begin{aligned}
\mathcal{N}((\pi^{-1}\mathcal{U})_0^{n-1}) &= \mathcal{N}(\bigvee_{i=0}^{n-1} S^{-i}\pi^{-1}\mathcal{U}) \\
&= \mathcal{N}(\bigvee_{i=0}^{n-1} (\pi S^{i})^{-1}\mathcal{U}) \\
&= \mathcal{N}(\bigvee_{i=0}^{n-1} (T^{i}\pi)^{-1}\mathcal{U}) \\
&= \mathcal{N}(\bigvee_{i=0}^{n-1} \pi^{-1}T^{-i}\mathcal{U}) \\
&= \mathcal{N}(\pi^{-1}\bigvee_{i=0}^{n-1} T^{-i}\mathcal{U}) \\
&= \mathcal{N}(\bigvee_{i=0}^{n-1} T^{-i}\mathcal{U})
\end{aligned}$$

Thus,

$$h_{\text{top}}(T,\mathcal{U}) = h_{\text{top}}(S, \pi^{-1}\mathcal{U})$$

□

Proposition 3.2.2. *For a TDS (X, T) and an open cover $\mathcal{U} \in C_X^{\circ}$, if T is surjective, we have*

$$h_{\text{top}}(T,\mathcal{U}) = h_{\text{top}}(T, T^{-1}\mathcal{U})$$

Proof. We prove the proposition in 2 steps:

1. If T is homeomorphism,

$$\mathcal{N}(\bigvee_{i=0}^{n-1} T^{-i}T^{-1}\mathcal{U}) = \mathcal{N}(T^{-1}\bigvee_{i=0}^{n-1} T^{-i}\mathcal{U}) = \mathcal{N}(\bigvee_{i=0}^{n-1} T^{-i}\mathcal{U})$$

Thus $h_{\text{top}}(T, T^{-1}\mathcal{U}) = h_{\text{top}}(T, \mathcal{U})$

2. In general case, by definition (2.3.5), let $(\tilde{X}, \sigma_T)$ be a natural extension of (X, T) by factor map $\pi\colon \tilde{X} \to X$.

   by proposition (3.2.1) and step 1, we have:

$$\begin{aligned} h_{\text{top}}(T, T^{-1}\mathcal{U}) &= h_{\text{top}}(\sigma_T, \pi^{-1}T^{-1}\mathcal{U}) = h_{\text{top}}(\sigma_T, \sigma_T^{-1}\pi^{-1}\mathcal{U}) \\ &= h_{\text{top}}(\sigma_T, \pi^{-1}\mathcal{U}) = h_{\text{top}}(T, \mathcal{U}) \qquad \text{by step 1} \end{aligned}$$

□

3.3 U.P.E.

Now that we introduced topological entropy, we define another phenomenon of a TDS: **Uniformly Positive Entropy** property.

At the beginning of the 1990s, Blanchard began a search for a topological analogue of K-systems. First he tried to define topological K-systems by means of global notions; then he realized that a local viewpoint may become very useful.

Definition 3.3.1. *We say that an open cover is nontrivial or standard if none of its members is dense in X.*

A TDS (X, T) has U.P.E. if every nontrivial open cover, consisting of two member, has positive entropy; that is

$$\forall \mathcal{U} = \{U, V\} \in C_X^\circ \colon \left(\overline{U} \neq X\right) \wedge \left(\overline{V} \neq X\right) \implies h_{\text{top}}(\mathcal{U}, T) > 0$$

Lemma 3.3.1 (Petersen)**.** [*13*] *A TDS system (X, T) is weakly mixing iff*

$$\forall U, V \overset{\text{open}}{\subseteq} X \ s.t.\ \exists n \in \mathbb{Z} \colon U \cap T^n V \neq \emptyset \wedge U \cap T^n U \neq \emptyset$$

Proof. $(\longleftarrow)$ for every $U_1, V_1, U_2.V_2 \overset{\text{open}}{\subseteq} X$ we show that $N(U_1, V_1) \cap N(U_2, V_2) \neq \emptyset$.

By assumption we have

$\exists n_1 \in \mathbb{Z}\ s.t.\ E := V_2 \cap T^{n_1} V_1 \neq \emptyset$

$\exists n_2 \in \mathbb{Z}\ s.t.\ F := T^{n_1} U_1 \cap T^{n_2} E \neq \emptyset$

$\exists n_3 \in \mathbb{Z}\ s.t.\ \begin{cases} U_2 \cap T^{n_3} F \neq \emptyset \\ F \cap T^{n_3} F \neq \emptyset \end{cases}$

Now let $n := n_2 + n_3$.

$$\begin{aligned}
T^{n_1}(U_1 \cap T^n V_1) &= T^{n_1} U_1 \cap T^{n_1+n_2+n_3} V_1 \\
&\supseteq T^{n_1} U_1 \cap T^{n_2+n_3} V_2 \cap T^{n_1+n_2+n_3} V_1 \\
&= T^{n_1} U_1 \cap T^{n_2+n_3}(V_2 \cap T^{n_1} V_1) \\
&= T^{n_1} U_1 \cap T^n E \\
&\supseteq T^{n_1} U_1 \cap T^{n_1+n_3} U_1 \cap T^{n_2+n_3} E \\
&= T^{n_1} U_1 \cap T^{n_3} F \supseteq F \cap T^{n_3} F \neq \emptyset
\end{aligned}$$

So $n \in N(U_1, V_1)$.

$$\begin{aligned}
T^{n_1+n_2+n_3} V_1 \cap U_2 \cap T^n V_2 &= U_2 \cap T^n (T^{n_1} V_1 \cap V_2) \\
&= U_2 \cap T^n E \\
&\supseteq U_2 \cap T^{n_2+n_3} E \cap T^{n_1+n_3} U_1 \\
&= U_2 \cap T^{n_3}(T^{n_2} E \cap T^{n_1} U_1) \\
&= U_2 \cap T^{n_3} F \neq \emptyset.
\end{aligned}$$

So $n \in N(U_2, V_2)$. Thus $N(U_1, V_1) \cap N(U_2, V_2) \neq \emptyset$.

($\longrightarrow$) If the system is weakly mixing,
then for every $U, V \overset{\text{open}}{\subseteq} X$, $N(U,U) \cap N(U,V) \neq \emptyset$ so there exists $n \in \mathbb{Z}$ such that $U \cap T^n U \neq \emptyset$ and $U \cap T^n V \neq \emptyset$ □

Theorem 3.3.1. *U.P.E. implies weakly mixing.*

Proof. By using lemma 3.3.1, we will show that if a system is not weakly mixing then there is a zero-entropy nontrivial cover consisting of two member, in particular the system is not U.P.E.
If (X,T) is not weakly mixing, then

$$\exists U, V \overset{\text{open}}{\subseteq} X \ s.t. \ \forall n \in \mathbb{Z}\colon U \cap T^n U = \emptyset \vee U \cap T^n V = \emptyset$$

Assume that $\mathcal{R} := \{R_U, R_V\}$ be a nontrivial open cover of X in a way that $U^c \subseteq R_V$ and $V^c \subseteq R_U$ (consider the complement of a close ball with a radius less than the diameter of U, that's R_V!)
For every $n \in \mathbb{N}$ and $i \in \{0, 1, ..., n\}$ we have two states:

- $U \cap T^i U \neq \emptyset$. So that, $U \cap T^i V = \emptyset$. Thus

$$U \subseteq (T^i V)^c = T^i(V^c) \subseteq T^i R_U$$

- $U \cap T^i U = \emptyset$. So that

$$U \subseteq (T^i U)^c = T^i(U^c) \subseteq T^i R_V$$

So for every $i \in \{0, 1, ...n-1\}$, in each case $U \subseteq T^i R_V$ or $U \subseteq T^i R_U$.
Consequently, for every $n \in \mathbb{N}$,

$$U \subseteq W_0 \cap TW_1 \cap ... \cap T^n W_n$$

for $W_i \in \{R_U, R_V\}$.
Furthermore, for every $x \in X$, if $x \in U$, x is in the intersection above and if $x \in U^c$, since

$U^c \subseteq R_v$, $x \in R_v$.

Now let

$$\mathcal{R}' := \{R_v \cap TR_v \cap ... \cap T^{i-1}R_v \cap T^iW_0 \cap T^{i+1}W_1 \cap ... \cap T^nW_{n-i} \mid i = 0, 1, ..., n-1\}$$

Now we proof that $\mathcal{R}'$ is a subcover of $\mathcal{R}_0^{n-1}$.

- $\mathcal{R}' \subseteq \mathcal{R}_0^{n-1}\{R_0 \cap T(R_1) \cap ... \cap T^{n-1}(R_{n-1}) \mid \forall 0 \leq i \leq n-1\colon R_i \in \mathcal{R}\}$

- $\mathcal{R}'$ is a cover of X.
  $\forall x \in X\colon x \in U \implies x \in W_0 \cap TW_1 \cap ... \cap T^{n-1}W_{n-1} \in \mathcal{R}'$.

  $x \in U^c$ choose j such that

  $(x \in R_V) \wedge ... \wedge (T^{j-1}(x) \in R_V) \wedge (T^j(x) \in U)$ so

  $x \in R_V \cap TR_V \cap ... \cap T^{j-1}R_V \cap T^jW_0 \cap ... \cap T^nW_n - 1 \in \mathcal{R}'$.

Thus, $\forall n \in \mathbb{N}\colon H(\mathcal{R}_0^{n-1}) \leq H(\mathcal{R}') = \log(n)$

$$\implies h(\mathcal{R}, T) \leq \lim_{n\to\infty} \frac{1}{n} H(\mathcal{R}') = \lim_{n\to\infty} \frac{\log(n)}{n} = 0$$

$\square$

3.4 C.P.E.

After introducing U.P.E. we put a step further and define **Completely Positive Entropy**.

Definition 3.4.1. *A TDS (X, T) has C.P.E. if every nontrivial factor of X have positive entropy.*

Chapter 4

Weak* Topology

In this section we are going to learn about the topology on the space of measures. We mostly consider the theorems will be required. Since our guideline in this section is the book *'Foundations of ergodic theory'*, the proofs are omitted. The proofs of most theorems are straightforward and can simply find in the book.

Definition 4.0.1. *A functional is a map $\phi\colon X \to \mathbb{R}$ or $\mathbb{C}$ from a Hilbert space X to a scalar field $\mathbb{R}$ or $\mathbb{C}$.*
A linear functional $\phi\colon X \to \mathbb{R}$ or $\mathbb{C}$ is a functional with the property

$$\phi(ax_1 + x_2) = a\phi(x_1) + \phi(x_2)$$

Definition 4.0.2. *We said a linear functional is bounded when*

$$\|\phi\| := \sup\{\frac{|\phi(v)|}{\|v\|} : v \neq 0, v \in X\} < \infty$$

Definition 4.0.3 (Dual Space)**.** *The dual space of a Hilbert space X,denoted by X^* is the vector field consisted by all bounded linear functionals.*

Definition 4.0.4 (Weak* Topology). *Weak* topology is the smallest topology on* X^* *such that all of linear functionals* $\begin{cases} h\colon X \to X^* \\ h(x) = x^* \end{cases}$ *where* $\begin{cases} x^*\colon X \to \mathbb{R} \text{ or } \mathbb{C} \\ x^*(v) = v \cdot x \end{cases}$ *are continues.*

Note that for each measure μ on X, $\mu \in X^*$.

Remark 4.0.1. *Weak* topology is a topology on probability measures on* X *such that two measures are close if the integral of every bounded continues function with respect to the measures are close.*

Proposition 4.0.1. *Let* $M(X)$ *be the collection of all probability measures on* X *and* $\mu \in M(X)$. *The base of weak* topology on* X^* *is the set containing*

$$V(\mu, \phi, \epsilon) := \left\{ \nu \in M(X)\colon |\int_X \phi_i \,\mathrm{d}\nu - \int_X \phi_i \,\mathrm{d}\mu| < \epsilon \right\}$$

where ϕ *is a set of bounded continues functions* $\phi_i\colon X \to \mathbb{R}$ *and* $\epsilon > 0$.

Corollary 4.0.1. $A \subseteq M(X)$ *is open iff*

$$\forall \mu \in A\colon \exists \epsilon > 0, \exists \phi = \{\phi_1, \phi_2, \dots, \phi_n\} \text{ s.t. } V(\mu, \phi, \epsilon) \subseteq A$$

Proposition 4.0.2. *A sequence of probability measures* $\{\mu_n\}_{n\in\mathbb{N}}$ *converges to* $\mu \in M(X)$ *iff for every bounded continues function* $\phi\colon X \to \mathbb{R}$,

$$\lim_{n\to\infty} \int_X \phi \,\mathrm{d}\mu_n = \int_X \phi \,\mathrm{d}\mu$$

Theorem 4.0.1. *Weak* topology is metrizable.*

Proof. Let consider $B^\delta := \{x \in X\colon d(x, B) < \delta\}$ where $B \subseteq X$.

Define

$$\forall \mu, \nu \in M(X) \colon D(\nu, \mu) := \inf\{\delta > 0 \colon \mu(B) < \nu(B^{\delta}) + \delta \\ \text{and } \nu(B) < \mu(B^{\delta}) + \delta \\ \text{and } B \text{ be a Borel measurable set}\}$$

D is a distance on $M(X)$. □

Theorem 4.0.2. *Weak* topology is compact.*

Definition 4.0.5. *Let consider a compact metric space X. Every continues function $T \colon X \to X$ induces a function* $\begin{cases} T_* \colon M(X) \to M(X) \\ T_*(\mu)(A) = \mu(T^{-1}A), \forall A \overset{\text{measurable}}{\subseteq} X \end{cases}$

Proposition 4.0.3. *If $\phi \colon X \to \mathbb{R}$ be a continues function and $\mu \in M(X)$ then*

$$\int_X \phi \ \mathrm{d}T_*\mu = \int_X \phi \circ T \ \mathrm{d}\mu$$

Proof. First consider the result for characteristic functions -simple functions- $\begin{cases} \chi_A \colon X \to \mathbb{R} \\ \chi_A(x) = 1 \quad x \in A \\ \chi_A(x) = 0 \quad x \notin A \end{cases}$ □

Proposition 4.0.4. *Consider a measure $\mu \in M(X)$ and a continues function $T \colon X \to X$. μ is T-invariant iff*

$$T_*(\mu) = \mu$$

Corollary 4.0.2. *If (X, T) is a TDS and $\phi \colon X \to \mathbb{R}$ is a continues function then $\mu \in M_T(X)$ iff*

$$\int_X \phi \circ T \ \mathrm{d}\mu = \int_X \phi \ \mathrm{d}T_*\mu = \int_X \phi \ \mathrm{d}\mu$$

Proposition 4.0.5. *If $T\colon X \to X$ is continues then $T_*\colon M(X) \to M(X)$ is continues with respect to weak* topology.*

Theorem 4.0.3. *If (X, T) is a TDS, and $\{\nu_n\}_{n\in\mathbb{N}}$ be a sequence in $M(X)$, then each accumulation point of a sequence*

$$\left\{\frac{1}{n}\sum_{j=0}^{n-1} T_*^j \nu_n\right\}$$

is a T-invariant probability measure.

Corollary 4.0.3. *For each continues function $T\colon X \to X$, there exists at least one T-invariant probability measure.*

Lemma 4.0.1. *If $\pi\colon (Y, S) \to (X, T)$ is a homomorphism between two TDS, $\mu_Y \in M_S(Y)$, then $\mu_X = \pi_*\mu_Y \in M_T(X)$.*

Proof. For any $A \overset{\text{measurable}}{\subseteq} X$, Since π is continues we have $\mu_X \in M(X)$, and since

$$\begin{aligned}
\pi_*\mu_Y(T^{-1}A) &= \mu_Y(\pi^{-1} \circ T^{-1}A) = \mu_Y((T \circ \pi)^{-1}A) \\
&= \mu_Y((\pi \circ S)^{-1}A) = \mu_Y(S^{-1} \circ \pi^{-1}A) \\
&= \mu_Y(\pi^{-1}A) \\
&= \pi_*\mu_Y(A)
\end{aligned}$$

we have $\mu_X \in M_T(X)$. □

Lemma 4.0.2. *If $\pi\colon (Y, S) \to (X, T)$ is a homomorphism between two TDS, $\mu_Y \in M_S(Y)$, and $\mu_X = \pi_*\mu_Y \in M_T(X)$ then for any continues function $\phi\colon X \to \mathbb{R}$,*

$$\int_Y \phi \circ \pi \, \mathrm{d}\mu_Y = \int_X \phi \, \mathrm{d}\mu_X$$

Proof. First consider the result for characteristic functions -simple functions. □

Proposition 4.0.6. *If $\pi\colon (Y,S) \to (X,T)$ is a homomorphism between two TDS, $\mu_Y \in M_S(Y)$, and $\mu_X = \pi_*\mu_Y \in M_T(X)$, then for any continues function $\phi\colon X \to \mathbb{R}$,*

$$\int_Y \phi \circ \pi \ \mathrm{d}\mu_Y = \int_X \phi \ \mathrm{d}\mu_X$$

Proof.

$$\begin{aligned}\int_Y \phi \circ \pi \ \mathrm{d}\mu_Y &= \int_Y \phi \circ \pi \ \mathrm{d}S_*\mu_Y \\ &= \int_Y \phi \circ \pi \circ S \ \mathrm{d}\mu_Y \\ &= \int_Y \phi \circ T \circ \pi \ \mathrm{d}\mu_Y \\ &= \int_X \phi \circ T \ \mathrm{d}\pi_*\mu_Y \\ &= \int_X \phi \circ T \ \mathrm{d}\mu_X \\ &= \int_X \phi \ \mathrm{d}\mu_X\end{aligned}$$

□

For the sake of variational principles let us consider the lemma blew.

Lemma 4.0.3. *If (X,T) is a TDS, $\mu \in M(X)$, $\alpha \in \mathcal{P}_X$, and $\phi\colon X \to \mathbb{R}$ is a continues function, then*

$$\int_X \phi \ \mathrm{d}\mu \leq \sum_{A\in\alpha} \mu(A) \sup_{x\in A} \phi$$

Proof.

$$\int_X \phi \ \mathrm{d}\mu = \sum_{A\in\alpha} \int_A \phi \ \mathrm{d}\mu \leq \sum_{A\in\alpha} \int_A \sup_{x\in A} \phi \ \mathrm{d}\mu = \sum_{A\in\alpha} \mu(A) \sup_{x\in A} \phi$$

□

Chapter 5

Measure Theoretical Entropy

5.1 Measure Entropy of Partitions

Definition 5.1.1. *A partition of the measure space $(X, \mathscr{B}, \mu)$ is a disjoint collection of elements of $\mathscr{B}$ whose union is X.*
Denote the collection of all finite partitions of X by $\mathcal{P}_X$.

Proposition 5.1.1. *There's a one-to-one correspondence between finite sub σ-algebras and finite partitions.*

Proof. Let $(X, \mathscr{B}, \mu)$ be a measure space and $\alpha = \{A_1, A_2, \ldots, A_n\} \in \mathcal{P}_X$. $\mathscr{A} := \{\bigcup_{j=1}^n B_j \mid B_j = \emptyset \text{ or } B_j = A_j\}$ is a finite sub σ-algebra.
conversely, if $\mathscr{A} = \{B_1, B_2, \ldots, B_n\}$ is a sub σ-algebra, $\alpha := \{\bigcap_{j=1}^n A_j \mid A_j = B_j \text{ or } A_j = X \setminus B_j\}$ forms a finite partition of X. $\square$

Remark 5.1.1. *By the proposition above, As we can see in The book of Peter Walters [17], one can replace the definitions of entropy of a partition by entropy of finite sub σ-algebras. By using the notation of sub σ-algebra we mean the partition corresponding to that σ-algebra.*

Definition 5.1.2. *Let $(X, \mathscr{A}, \mu)$ be a measure space and $\alpha, \beta \in \mathcal{P}_X$ we said α is refinement of β or β is coarser than α and we write $\alpha \succeq \beta$ if*

$$\forall A \in \alpha \colon \exists B \in \beta \text{ s.t. } A \subseteq B$$

In the books and references there are two approaches, one with the definition above, and the other with the same definition but up to measure zero.
Though, both notations reaches to the same results.

Proposition 5.1.2. *Let $(X, \mathscr{A}, \mu)$ be a measure space and $\alpha, \beta \in \mathcal{P}_X$ that $\alpha \succeq \beta$ then*

$$\forall B \in \beta \colon \exists A_1, A_2, \dots, A_n \in \alpha \text{ s.t. } B = \cup_{i=1}^{n} A_i$$

Proof. for any $B \in \beta$,

$$B = \bigcup_{\substack{A \subseteq B \\ A \in \alpha}} A$$

Assume that there exist $B \in \beta$ such that $B \neq \bigcup_{\substack{A \subseteq B \\ A \in \alpha}} A$. that means there exists $A' \in \alpha$ such that $A \cap B \neq \emptyset$ and $A \nsubseteq B$. By assumption for $A' \in \alpha$ exists $B' \in \beta$ such that $A' \subseteq B'$. But Since $A' \cap B \neq \emptyset$ we have $B \cap B' \neq \emptyset$. Since β is a partition, we reach to the contradiction. Thus $B = \bigcup_{\substack{A \subseteq B \\ A \in \alpha}} A$ □

Definition 5.1.3. *Let $(X, \mathscr{A}, \mu)$ be a measure space, $\alpha = \{A_1, A_2, \dots, A_n\} \in \mathcal{P}_X$, and $\beta = \{B_1, B_2, \dots, B_m\} \in \mathcal{P}_X$ then we can reach to the natural refinement of both α and β by gathering the collection of the intersection of one's elements with the other. Denote $\alpha \vee \beta$ for this refinement, i.e.*

$$\alpha \vee \beta := \{A \cap B \mid A \in \alpha, B \in \beta\}$$

In sense of sub σ-algebras $\mathscr{A} \vee \mathscr{B}$ denote the smallest σ-algebra containing both $\mathscr{A}$ and $\mathscr{B}$.

Definition 5.1.4. *Let* $(X, \mathscr{A}, \mu)$ *be a measure space,* $\alpha = \{A_1, A_2, \ldots, A_n\} \in \mathcal{P}_X$*, and* $T\colon X \to X$ *be a measurable map. for* $n \in \mathbb{N}$*, define* $T^{-n}\alpha := \{T^{-n}A_1, T^{-n}A_2, \ldots T^{-n}A_n\}$*, Then* $T^{-n}\alpha \in \mathcal{P}$*.*

Remark 5.1.2. *Let* $(X, \mathscr{A}, \mu)$ *be a measure space,* $\alpha = \{A_1, A_2, \ldots, A_n\} \in \mathcal{P}_X$*, and* $T\colon X \to X$ *be a measurable map.*

$$T^{-n}(\alpha) \vee T^{-n}(\beta) = T^{-n}\Big(\alpha \vee \beta\Big)$$

$$\alpha \succeq \beta \implies T^{-n}\alpha \succeq T^{-n}\beta$$

Definition 5.1.5. *Let* $(X, \mathscr{A}, \mu)$ *be a measure space and* $\alpha \in \mathcal{P}_X$*. We call the* $H_\mu(\alpha)$ *the information entropy of a partition* α *or if it was clear from the context, the entropy of* α*, i.e.*

$$H_\mu(\alpha) = \sum_{A\in\alpha} -\mu(A)\log\mu(A) \tag{5.1}$$

If we consider the concave function $\begin{cases} \phi\colon [0,1] \to \mathbb{R} \\ \phi(x) = -x\log x \end{cases}$*, the entropy of a partition becomes*

$$H_\mu(\alpha) = \sum_{A\in\alpha} \phi(\mu(A)) \tag{5.2}$$

we call it 'information' as it comes from information theory and it is actually the mean of its information function [1]. It's worth mentioning that by concavity of the function ϕ, we mean:

$$\forall t_1, \ldots, t_k \in [0,1] \text{ that } \sum_{i=1}^{k} t_i = 1, \forall x_1 \ldots x_k \in \mathcal{R}^+\colon \sum_{i=1}^{k}\Big(t_i\phi(x_i)\Big) \le \phi\left(\sum_{i=1}^{k} t_i x_i\right)^2$$

and the equality holds iff $x_1 = x_2 = \cdots = x_k$.

[1]for more details see e.g. [15] or for even more details, see the chapter one of the book [5]

[2]For more detail see the book 'Convex Optimization' [3]

Proposition 5.1.3. *Every partition $\alpha \in \mathcal{P}_X$ has finite entropy; More precisely,*

$$\forall \alpha \in \mathcal{P}_X \colon H_\mu(\alpha) \leq \log \#\alpha$$

The equality holds iff $\forall A \in \alpha \colon \mu(A) = \frac{1}{\#\alpha}$

Proof. Consider $\alpha = \{A_1, A_2, \ldots, A_n\} \in \mathcal{P}_X$ using the concavity inequality explained above,

$$\frac{1}{n} H_\mu(\alpha) = \frac{1}{n} \sum_{i=1}^{n} \phi(\mu(A)) \leq \phi\left(\sum_{i=1}^{n} \frac{1}{n} \mu(A_i)\right) = \phi(\frac{1}{n}) = \frac{1}{n} \log n$$

According to the concavity inequality the equality holds iff

$$\mu(A_1) = \mu(A_2) = \cdots = \mu(A_n) = \frac{1}{n}$$

Since $\sum_{i=1}^{n} \mu(A_i) = 1$. $\square$

The next phenomenon is the *Conditional entropy* of a partition.

Definition 5.1.6. *The conditional entropy of a partition α with respect to another partition β is*

$$H_\mu(\alpha \mid \beta) = \sum_{B \in \beta} \mu(B) \sum_{A \in \alpha} \phi(\mu(A \mid B)) \tag{5.3}$$

where $\mu(A \mid B) = \frac{\mu(A \cap B)}{\mu(B)}$

Remark 5.1.3.

$$\begin{aligned} H_\mu(\alpha \mid \beta) &= \sum_{B \in \beta} \mu(B) \sum_{A \in \alpha} \phi(\mu(A \mid B)) \\ &= \sum_{B \in \beta} \mu(B) \sum_{A \in \alpha} -\frac{\mu(A \cap B)}{\mu(B)} \log\left(\frac{\mu(A \cap B)}{\mu(B)}\right) \\ &= \sum_{B \in \beta} \sum_{A \in \alpha} -\mu(A \cap B) \log\left(\frac{\mu(A \cap B)}{\mu(B)}\right) \end{aligned}$$

Example 5.1.1. *For any measure space $(X, \mathscr{A}, \mu)$ and $\alpha \in \mathcal{P}_X$,*

$$H_\mu(\alpha \mid \{X, \emptyset\}) = \sum_{A \in \alpha} -\mu(A) \log \mu(A) = H_\mu(\alpha)$$

Theorem 5.1.1. *Consider the measure space $(X, \mathscr{A}, \mu)$ and $\alpha, \beta, \xi \in \mathcal{P}_X$*

1. $H_\mu(\alpha \vee \beta \mid \xi) = H_\mu(\alpha \mid \xi) + H_\mu(\beta \mid \alpha \vee \xi)$
2. $H_\mu(\alpha \vee \beta) = H_\mu(\alpha) + H_\mu(\beta \mid \alpha)$
3. $\alpha \succeq \beta \implies H_\mu(\xi \mid \alpha) \leq H_\mu(\xi \mid \beta)$
4. $H_\mu(\xi \mid \alpha) \leq H_\mu(\xi)$
5. $\alpha \succeq \beta \implies H_\mu(\alpha \mid \xi) \geq H_\mu(\beta \mid \xi)$
6. $\alpha \succeq \beta \implies H_\mu(\alpha) \geq H_\mu(\beta)$
7. $H_\mu(\alpha \vee \beta \mid \xi) \leq H_\mu(\alpha \mid \xi) + H_\mu(\beta \mid \xi)$
8. $H_\mu(\alpha \vee \beta) \leq H_\mu(\alpha) + H_\mu(\beta)$
9. $\alpha \succeq \beta \iff H_\mu(\beta \mid \alpha) = 0$
10. *If $T \colon X \to X$ is a measure-preserving map,*

$$H_\mu(T^{-1}\alpha \mid T^{-1}\beta) = H_\mu(\alpha \mid \beta)$$

11. *If $T \colon X \to X$ is a measure-preserving map,*

$$H_\mu(T^{-1}\alpha) = H_\mu(\alpha)$$

12. *If $f \colon Y \to X$ is a measurable map, $\alpha, \beta \in \mathcal{P}_X$*

$$H_\mu(f^{-1}\alpha \mid f^{-1}\beta) = H_{f_*\mu}(\alpha \mid \beta)$$

13. *If* $f\colon Y \to X$ *is a measurable map,* $\alpha \in \mathcal{P}_X$

$$H_\mu(f^{-1}\alpha) = H_{f_*\mu}(\alpha)$$

14. *If* $T\colon X \to X$ *is bijective and measure-preserving transformation,*

$$H_\mu(T\alpha \mid T\beta) = H_\mu(\alpha \mid \beta)$$

15. *If* $T\colon X \to X$ *is bijective and measure-preserving transformation,*

$$H_\mu(T\alpha) = H_\mu(\alpha)$$

Proof. 1. By definition,

$$
\begin{aligned}
H_\mu(\alpha \vee \beta \mid \xi) &= \sum_{A,B,C} -\mu(A\cap B\cap C)\log\frac{\mu(A\cap B\cap C)}{\mu(C)}\\
&= \sum_{A,B,C} -\mu(A\cap B\cap C)\log\left(\frac{\mu(A\cap B\cap C)}{\mu(A\cap C)}\frac{\mu(A\cap C)}{\mu(C)}\right)\\
&= \sum_{A,B,C} -\mu(A\cap B\cap C)\log\frac{\mu(A\cap B\cap C)}{\mu(A\cap C)}\\
&+ \sum_{A,B,C} -\mu(A\cap B\cap C)\log\frac{\mu(A\cap C)}{\mu(C)}\\
&= \sum_{\substack{D\in\alpha\vee\xi\\ B\in\beta}} -\mu(D\cap B)\log\frac{\mu(D\cap B)}{\mu(D)}\\
&+ \sum_{\substack{A\in\alpha\\ C\in\xi}}\sum_{B\in\beta} -\mu(A\cap B\cap C)\log\frac{\mu(A\cap C)}{\mu(C)}\\
&= H_\mu(\beta \mid \alpha\vee\xi) + \sum_{\substack{A\in\alpha\\ C\in\xi}} -\mu(A\cap C)\log\frac{\mu(A\cap C)}{\mu(C)}\\
&= H_\mu(\beta \mid \alpha\vee\xi) + H_\mu(\alpha \mid \xi)
\end{aligned}
$$

2. From (1), let $\xi = \{X, \emptyset\}$ the result will be obtained.
3. Since for any $A \in \alpha$ and $C \in \xi$

$$
\frac{\mu(A\cap C)}{\mu(A)} = \sum_{\substack{B\subseteq A\\ B\in\beta}} \frac{\mu(B)}{\mu(A)}\frac{\mu(B\cap C)}{\mu(B)}
$$

and also notice that $\sum_{B\subseteq A}\frac{\mu(B)}{\mu(A)} = 1$ -with respect to the concavity property of the function $\phi(x) = -x \log x$ we have

$$\begin{aligned}
H_\mu(\xi \mid \alpha) &= \sum_{A\in\alpha} \mu(A) \sum_{C\in\xi} \phi\left(\frac{\mu(A\cap C)}{\mu(A)}\right) \\
&= \sum_{A\in\alpha} \mu(A) \sum_{C\in\xi} \phi\left(\sum_{\substack{B\subseteq A\\ B\in\beta}} \frac{\mu(B)}{\mu(A)} \frac{\mu(B\cap C)}{\mu(B)}\right) \\
&\le \sum_{A\in\alpha} \mu(A) \sum_{C\in\xi} \sum_{\substack{B\subseteq A\\ B\in\beta}} \frac{\mu(B)}{\mu(A)} \phi\left(\frac{\mu(B\cap C)}{\mu(B)}\right) \\
&= \sum_{\substack{C\in\xi\\ B\in\beta}} \mu(B)\phi\Big(\mu(C \mid B)\Big) \\
&= H_\mu(\xi \mid \beta)
\end{aligned}$$

4. The result will be gained From (3) by letting $\beta = \{X, \emptyset\}$.
5. If $\alpha \succeq \beta$, Notice that $B \subseteq A \implies \mu(A) \ge \mu(B)$, so

$$\begin{aligned}
H_\mu(\alpha \mid \xi) &= \sum_{\substack{C\in\xi\\ A\in\alpha}} \mu(A\cap C) \log \frac{\mu(A\cap C)}{\mu(C)} \\
&= \sum_{\substack{C\in\xi\\ A\in\alpha}} \sum_{\substack{B\subseteq A\\ B\in\beta}} \mu(B\cap C) \log \frac{\mu(A\cap C)}{\mu(C)} \\
&\ge \sum_{\substack{C\in\xi\\ A\in\alpha}} \sum_{\substack{B\subseteq A\\ B\in\beta}} \mu(B\cap C) \log \frac{\mu(B\cap C)}{\mu(C)} \\
&= \sum_{C\in\xi} \sum_{B\in\beta} \mu(B\cap C) \log \frac{\mu(B\cap C)}{\mu(C)} \\
&= H_\mu(\beta \mid \alpha)
\end{aligned}$$

6. From (5), let $\xi = \{X, \emptyset\}$.
7.

$$\begin{aligned}
H_\mu(\alpha \vee \beta \mid \xi) &= H_\mu(\alpha \mid \xi) + H_\mu(\beta \mid \alpha \vee \xi) && \text{From (1)} \\
&\le H_\mu(\alpha \mid \xi) + H_\mu(\beta \mid \xi) && \text{From (3), Since } \alpha \vee \xi \succeq \xi
\end{aligned}$$

8. From (7), let $\xi = \{X, \emptyset\}$.
9. Since for each $A \in \alpha$ and $B \in \beta$ we have $-\mu(A \cap B) \log \frac{\mu(A\cap B)}{\mu(A)} \geq 0$,

$$H_\mu(\beta \mid \alpha) = \sum_{\substack{A\in\alpha \\ B\in\beta}} -\mu(A \cap B) \log \frac{\mu(A \cap B)}{\mu(A)} = 0$$

iff $\mu(A \cap B) = 0$ or $\mu(A \cap B) = \mu(A)$ for every $A \in \alpha$ and $B \in \beta$. If $\alpha \succeq \beta$ by definition,

$$\forall A \in \alpha, B \in \beta \colon \mu(A \cap B) \neq 0 \implies \mu(A \cap B) = \mu(A)$$

If $\mu(A \cap B) = 0$ or $\mu(A \cap B) = \mu(A)$ for every $A \in \alpha$ and $B \in \beta$, by definition, $\alpha \succeq \beta$.
10. Since $\forall A \in \alpha, B \in \beta \colon \mu(T^{-1}A) = \mu(A), \mu(T^{-1}B) = \mu(B), \mu(T^{-1}A \cap T^{-1}B) = \mu(T^{-1}(A \cap B)) = \mu(A \cap B)$ the result will be obtained.
11. From (10), let $\beta = \{X, \emptyset\}$.
12.

$$\begin{aligned}
H_\mu(f^{-1}\alpha \mid f^{-1}\beta) &= \sum_{B\in\beta} \sum_{A\in\alpha} \mu(f^{-1}A \cap f^{-1}B) \log \frac{f^{-1}A \cap f^{-1}B}{f^{-1}B} \\
&= \sum_{B\in\beta} \sum_{A\in\alpha} \mu(f^{-1}(A \cap B)) \log \frac{\mu(f^{-1}(A \cap B))}{\mu(f^{-1}B)} \\
&= \sum_{B\in\beta} \sum_{A\in\alpha} f_*\mu(A \cap B) \log \frac{f_*\mu(A \cap B)}{f_*\mu(B)} \\
&= H_{f_*\mu}(\alpha \mid \beta)
\end{aligned}$$

13. From (12), let $\beta = \{X, \emptyset\}$.

14.

$$\begin{aligned}
H_\mu(T\alpha \mid T\beta) &= \sum_{B\in\beta}\sum_{A\in\alpha} \mu(TA\cap TB)\log\frac{\mu(TA\cap TB)}{\mu(TB)} \\
&= \sum_{B\in\beta}\sum_{A\in\alpha} \mu(T(A\cap B))\log\frac{\mu(T(A\cap B))}{\mu(TB)} \\
&= \sum_{B\in\beta}\sum_{A\in\alpha} \mu(A\cap B)\log\frac{\mu(A\cap B)}{\mu(B)} \\
&= H_\mu(\alpha \mid \beta)
\end{aligned}$$

15. From (14), let $\beta = \{X, \emptyset\}$.

$\square$

We saw that if α and β are partitions of X. $\alpha \vee \beta$ is partition of X finer than both α and β. Also we saw that $T^{-1}\alpha$ is a partition of X. So by considering these facts for each $n \in \mathbb{N}$, we represent a way of creating a partition α_0^{n-1} finer than α. We'll see that this construction help us define the entropy of a dynamical system. Let us introduce a notation:

If $(X, \mathscr{A}, \mu, T)$ is a MDS and $\alpha \in \mathcal{P}_X$, we set

$$\alpha_m^n := \bigvee_{i=m}^{n} T^{-i}\alpha \tag{5.4}$$

Since for each $i \in \mathbb{N}$, $T^{-i}\alpha$ is a partition of X and the join of partitions is a partition we obtain that for each $m \le n \in \mathbb{N}$, α_m^n is a partition of X. Now as we expect, for each $n \in \mathbb{N}$, α_0^{n-1} is a partition finer than α.

Also for each $x \in X$ we denote $\alpha(x)$ an element of α containing x. Therefore, $\alpha_0^{n-1}(x)$ specify the element of α_0^{n-1} containing x.

Remark 5.1.4. *If $(X, \mathscr{A}, \mu, T)$ is a MDS and $\alpha \in \mathcal{P}_X$,*

$$\alpha = \alpha_0^0 \preceq \alpha_0^1 \preceq \cdots \preceq \alpha_0^{n-1} \preceq \ldots$$

Thus, by Theorem(5.1.1) *part*(6)*, we have:*

$$H_\mu(\alpha) \leq H_\mu(\alpha_0^1) \leq \cdots \leq H_\mu(\alpha_0^{n-1}) \leq \ldots$$

Proposition 5.1.4. *For a MDS $(X, \mathscr{A}, \mu, T)$ and $\alpha \in \mathcal{P}_X$ the limit*

$$\lim_{n\to\infty} \frac{1}{n} H_\mu(\alpha_0^{n-1})$$

exists.

Proof. It is sufficient to show that the sequence $\{a_n\}_{n\in\mathbb{N}}$ that $a_n := H_\mu(\alpha_0^{n-1}), \forall n \in \mathbb{N}$ is subadditive i.e. $a_{n+m} \leq a_n + a_m$

Due to the the Theorem (5.1.1) part (8), part (11), and also the definition of join, we have:

$$\begin{aligned}
a_{m+n} &= H_\mu(\alpha_0^{m+n-1}) \\
&= H_\mu\left(\bigvee_{i=0}^{m+n-1} T^{-i}\alpha\right) \\
&= H_\mu\left(\bigvee_{i=0}^{n-1} T^{-i}\alpha \vee \bigvee_{i=n}^{m+n-1} T^{-i}\alpha\right) \\
&= H_\mu\left(\bigvee_{i=0}^{n-1} T^{-i}\alpha \vee T^{-n}\left(\bigvee_{i=0}^{m-1} T^{-i}\alpha\right)\right) \\
&\leq H_\mu\left(\bigvee_{i=0}^{n-1} T^{-i}\alpha\right) + H_\mu\left(T^{-n}\left(\bigvee_{i=0}^{m-1} T^{-i}\alpha\right)\right) && \text{from (8)} \\
&= H_\mu\left(\bigvee_{i=0}^{n-1} T^{-i}\alpha\right) + H_\mu\left(\bigvee_{i=0}^{m-1} T^{-i}\alpha\right) && \text{from (11)} \\
&= a_n + a_m
\end{aligned}$$

□

Definition 5.1.7. *Consider MDS $(X, \mathscr{A}, \mu, T)$ along with $\alpha \in \mathcal{P}_X$. The entropy of T with respect to the partition α is*

$$h_\mu(T, \alpha) := \lim_{n\to\infty} \frac{1}{n} H_\mu(\alpha_0^{n-1}) = \inf_{n\in\mathbb{N}} \frac{1}{n} H_\mu(\alpha_0^{n-1}) \tag{5.5}$$

Example 5.1.2. *Let $(X, \mathscr{A}, \mu, T)$ be a MDS.*

$$\begin{aligned}
h_\mu(T, \{X, \emptyset\}) &= \lim_{n\to\infty} \frac{1}{n} H_\mu(\{X, \emptyset\}_0^{n-1}) \\
&= \lim_{n\to\infty} \frac{1}{n} H_\mu(\{X, \emptyset\}) \\
&= 0
\end{aligned}$$

Example 5.1.3. *Let $(X, \mathscr{A}, \mu, T := id_X)$ be a MDS. for any $\alpha \in \mathcal{P}_X$*

$$\begin{aligned} h_\mu(id_X, \alpha) &= \lim_{n\to\infty} \frac{1}{n} H_\mu(\alpha_0^{n-1}) \\ &= \lim_{n\to\infty} \frac{1}{n} H_\mu(\alpha) \\ &= 0 \end{aligned}$$

Remark 5.1.5. *By definition of join, it is not difficult to see that*

$$\alpha_1 \succeq \beta_1 \text{ and } \alpha_2 \succeq \beta_2 \implies \alpha_1 \vee \alpha_2 \succeq \beta_1 \vee \beta_2$$

So if $\alpha \succeq \beta$ then $\alpha_0^{n-1} \succeq \beta_0^{n-1}$. Thus $H_\mu(\alpha_0^{n-1}) \geq H_\mu(\beta_0^{n-1})$

Definition 5.1.8. *The entropy of the system $(X, \mathscr{A}, \mu, T)$ or simply, the entropy of the system T is defined by*

$$h_\mu(T) := \sup_{\alpha \in \mathcal{P}_X} h_\mu(T, \alpha) \tag{5.6}$$

As Walters [17] write:

we think of an application of T as a passage of one day of time then $h_\mu(T)$ is the maximum average information per day obtainable by performing the same experiment daily.

Lemma 5.1.1. *If s_n is convergent sequence of real numbers, with $\lim_{n\to\infty} s_n = s$ then $\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} s_k = s$*

Proof. Since $\lim_{n\to\infty} s_n = s$, choose $m \in \mathbb{N}$ large enough that for any $k > m : |s_k - s| < \frac{\epsilon}{2}$. As m is specified, choose $N \in \mathbb{N}$ in a way that $\forall n > N$: $\frac{1}{n} \sum_{k=1}^{m} |s_k - s| < \frac{\epsilon}{2}$.

Now for any $n > N$ we have

$$\begin{aligned}
\left|\left(\frac{1}{n}\sum_{k=1}^{n} s_k\right) - s\right| &= \left|\frac{s_1 + s_2 + \cdots + s_n}{n} - s\right| \\
&= \left|\frac{s_1 - s + s_2 - s + \cdots + s_n - s}{n}\right| \\
&= \frac{1}{n}\left|\sum_{k=1}^{n}(s_k - s)\right| \\
&\leq \frac{1}{n}\sum_{k=1}^{n}|s_k - s| \\
&= \frac{1}{n}\sum_{k=1}^{m}|s_k - s| + \frac{1}{n}\sum_{k=m+1}^{n}|s_k - s| \\
&< \frac{\epsilon}{2} + \frac{n-m}{n}\frac{\epsilon}{2} \\
&< \epsilon
\end{aligned}$$

□

Theorem 5.1.2. *Let $(X, \mathscr{A}, \mu, T)$ be a MDS and $\alpha, \beta \in \mathcal{P}_X$, then*

1. $h_\mu(T, \alpha) \leq H_\mu(\alpha)$
2. $h_\mu(T, \alpha \vee \beta) \leq h_\mu(T, \alpha) + h_\mu(T, \beta)$
3. $\alpha \succeq \beta \implies h_\mu(T, \alpha) \geq h_\mu(T, \beta)$
4. $h_\mu(T, \alpha) \leq h_\mu(T, \beta) + H_\mu(\alpha \mid \beta)$
5. $h_\mu(T, T^{-1}\alpha) = h_\mu(T, \alpha)$
6. $\forall m \in \mathbb{N}\colon h_\mu\left(T, \alpha_0^{m-1}\right) = h_\mu(T, \alpha)$
7. *If T is invertible,*

$$\forall m \in \mathbb{N}\colon h_\mu\left(T, \alpha_{-m}^{m}\right) = h_\mu(T, \alpha)$$

8. *If α be a partition with finite entropy, then*

$$h_\mu(T, \alpha) = \lim_{n\to\infty} H_\mu(\alpha \mid \alpha_1^n) = H_\mu(\alpha \mid \alpha_0^\infty)$$

Proof. 1.

$$\begin{aligned}
h_\mu(T,\alpha) &= \lim_{n\to\infty} \frac{1}{n} H_\mu(\alpha_0^{n-1}) \\
&= \lim_{n\to\infty} \frac{1}{n} H_\mu(\bigvee_{i=0}^{n-1} T^{-i}\alpha) \\
&\leq \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} H_\mu(T^{-i}\alpha) \\
&= \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} H_\mu(\alpha) \\
&= H_\mu(\alpha)
\end{aligned}$$

2.

$$\begin{aligned}
h_\mu(T,\alpha\vee\beta) &= \lim_{n\to\infty} \frac{1}{n} H_\mu((\alpha\vee\beta)_0^{n-1}) \\
&= \lim_{n\to\infty} \frac{1}{n} H_\mu(\alpha_0^{n-1}\vee\beta_0^{n-1}) \\
&\leq \lim_{n\to\infty} \frac{1}{n} H_\mu(\alpha_0^{n-1}) + \frac{1}{n} H_\mu(\beta_0^{n-1}) \\
&= h_\mu(T,\alpha) + h_\mu(T,\beta)
\end{aligned}$$

3. If $\alpha \succeq \beta$, according to the remark (5.1.5),

$$\begin{aligned}
h_\mu(T,\alpha) &= \lim_{n\to\infty} \frac{1}{n} H_\mu(\alpha_0^{n-1}) \\
&\geq \lim_{n\to\infty} \frac{1}{n} H_\mu(\beta_0^{n-1}) \\
&= h_\mu(T,\beta)
\end{aligned}$$

4. Due to the theorem (5.1.1),

$$
\begin{aligned}
h_\mu(T,\alpha) &= \lim_{n\to\infty} \frac{1}{n} H_\mu(\alpha_0^{n-1}) \\
&\leq \lim_{n\to\infty} \frac{1}{n} H_\mu(\alpha_0^{n-1} \vee \beta_0^{n-1}) && \text{from (6)} \\
&= \lim_{n\to\infty} \frac{1}{n} \Bigg(H_\mu(\beta_0^{n-1}) + H_\mu(\alpha_0^{n-1} \mid \beta_0^{n-1}) \Bigg) && \text{from (2)} \\
&\leq \lim_{n\to\infty} \frac{1}{n} \Bigg(H_\mu(\beta_0^{n-1}) + \sum_{i=0}^{n-1} H_\mu(T^{-i}\alpha \mid \beta_0^{n-1}) \Bigg) && \text{from (8)} \\
&\leq \lim_{n\to\infty} \frac{1}{n} \Bigg(H_\mu(\beta_0^{n-1}) + \sum_{i=0}^{n-1} H_\mu(T^{-i}\alpha \mid T^{-i}\beta) \Bigg) && \text{from (3)} \\
&= \lim_{n\to\infty} \frac{1}{n} \Bigg(H_\mu(\beta_0^{n-1}) + \sum_{i=0}^{n-1} H_\mu(\alpha \mid \beta) \Bigg) && \text{from (10)} \\
&= h_\mu(T,\beta) + H_\mu(\alpha \mid \beta)
\end{aligned}
$$

5. Since $(T^{-1}\alpha)_0^{n-1} = T^{-1}\alpha_0^{n-1}$ and by noticing to the theorem (5.1.1), part (11) the result will be obtained.

6.

$$
\begin{aligned}
h_\mu\Bigg(T, \alpha_0^{m-1}\Bigg) &= \lim_{n\to\infty} \frac{1}{n} H_\mu\Bigg(\bigvee_{i=0}^{n-1} T^{-i}\alpha_0^{m-1} \Bigg) \\
&= \lim_{n\to\infty} \frac{1}{n} H_\mu\Bigg(\bigvee_{i=0}^{n-1} T^{-i}\Bigg(\bigvee_{j=0}^{m-1} T^{-j}\alpha \Bigg)\Bigg) \\
&= \lim_{n\to\infty} \frac{1}{n} H_\mu\Bigg(\bigvee_{i=0}^{n-1} \bigvee_{j=0}^{m-1} T^{-(i+j)}\alpha \Bigg) \\
&= \lim_{n\to\infty} \frac{1}{n} H_\mu\Bigg(\bigvee_{k=0}^{n+m-2} T^{-k}\alpha \Bigg) \\
&= \lim_{n\to\infty} \frac{1}{n} H_\mu\Bigg(\alpha_0^{m+n-2} \Bigg) \\
&= h_\mu(T,\alpha) && \text{Since if } n\to\infty, \text{ we got } m+n-2\to\infty
\end{aligned}
$$

7.

$$
\begin{aligned}
h_\mu\left(T, \alpha_{-m}^{m}\right) &= \lim_{n\to\infty} \frac{1}{n} H_\mu\left(\bigvee_{i=0}^{n-1} T^{-i}\alpha_{-m}^{m}\right) \\
&= \lim_{n\to\infty} \frac{1}{n} H_\mu\left(\bigvee_{i=0}^{n-1} T^{-i}\left(\bigvee_{j=-m}^{m} T^{-j}\alpha\right)\right) \\
&= \lim_{n\to\infty} \frac{1}{n} H_\mu\left(\bigvee_{i=0}^{n-1} \bigvee_{j=-m}^{m} T^{-(i+j)}\alpha\right) \\
&= \lim_{n\to\infty} \frac{1}{n} H_\mu\left(T^{m}\left(\bigvee_{i=0}^{n-1} \bigvee_{j=0}^{2m} T^{-(i+j)}\alpha\right)\right) && \text{Since } T \text{ is invertible} \\
&= \lim_{n\to\infty} \frac{1}{n} H_\mu\left(\bigvee_{i=0}^{n-1} \bigvee_{j=0}^{2m} T^{-(i+j)}\alpha\right) && \text{From (15)} \\
&= \lim_{n\to\infty} \frac{1}{n} H_\mu\left(\bigvee_{k=0}^{2m+n-1} T^{-k}\alpha\right) \\
&= \lim_{n\to\infty} \frac{1}{n} H_\mu\left(\alpha_{0}^{2m+n-1}\right) \\
&= h_\mu(T, \alpha)
\end{aligned}
$$

8. Using theorem (5.1.1), we have

$$\begin{aligned}
h_\mu(T,\alpha) &= \lim_{n\to\infty} \frac{1}{n} H_\mu(\alpha_0^{n-1}) \\
&= \lim_{n\to\infty} \frac{1}{n} H_\mu\left(\alpha \vee \alpha_1^{n-1}\right) \\
&= \lim_{n\to\infty} \frac{1}{n}\left(H_\mu(\alpha_1^{n-1}) + H_\mu(\alpha \mid \alpha_1^{n-1})\right) && \text{From (1)} \\
&= \lim_{n\to\infty} \frac{1}{n}\left(H_\mu(\alpha_0^{n-2}) + H_\mu(\alpha \mid \alpha_1^{n-1})\right) && \text{Since } T \text{ is } \mu\text{-invariant} \\
&= \lim_{n\to\infty} \frac{1}{n}\left(H_\mu(\alpha) + \sum_{k=1}^{n-1} H_\mu(\alpha \mid \alpha_1^{k})\right) && \text{Since by recurrence } H_\mu(\alpha_0^{n-2}) = H_\mu(\alpha_0^{n-3}) + H_\mu(\alpha \mid \alpha_1^{n-2}) \\
&= \lim_{n\to\infty} \frac{1}{n}\left(\sum_{k=1}^{n-1} H_\mu(\alpha \mid \alpha_1^{k})\right)
\end{aligned}$$

Since $\{H_\mu(\alpha \mid \alpha_1^k)\}_{n\in\mathbb{N}}$ is decreasing sequence of non-negative real numbers, from lemma (5.1.1):

$$\begin{aligned}
&= \lim_{n\to\infty} H_\mu(\alpha \mid \alpha_1^{n}) \\
&= H_\mu(\alpha \mid \alpha_1^{\infty})
\end{aligned}$$

□

Theorem 5.1.3. *Let $(X, \mathscr{A}, \mu, T)$ be a MDS,*

1. *$\forall k \in \mathbb{N}$: $h_\mu(T^k) = k h_\mu(T)$*
2. *If the system is invertible, $\forall k \in \mathbb{Z}$: $h_\mu(T^k) = |k| h_\mu(T)$*

Proof. 1. for each $k \in \mathbb{N}$ by theorem (5.1.2) part (6)

$$\begin{aligned}
h_\mu(T^k) &= \sup_{\alpha \in \mathcal{P}_X} h_\mu(T^k, \alpha) \\
&= \sup_{\alpha \in \mathcal{P}_X} h_\mu(T^k, \alpha_0^{k-1}) \\
&= \sup\left\{ \lim_{n\to\infty} \frac{1}{n} H_\mu\left(\bigvee_{i=0}^{n-1} T^{-ki} \alpha_0^{k-1} \right) \mid \alpha \in \mathcal{P}_X \right\} \\
&= \sup\left\{ \lim_{n\to\infty} \frac{1}{n} H_\mu\left(\bigvee_{i=0}^{n-1} T^{-ki} \left(\bigvee_{j=0}^{k-1} T^{-j} \alpha \right) \right) \mid \alpha \in \mathcal{P}_X \right) \right\} \\
&= \sup\left\{ \lim_{n\to\infty} \frac{1}{n} H_\mu\left(\bigvee_{i=0}^{n-1} \bigvee_{j=0}^{k-1} T^{-(ki+j)} \alpha \right) \mid \alpha \in \mathcal{P}_X \right\} \\
&= \sup\left\{ \lim_{n\to\infty} \frac{1}{n} H_\mu\left(\bigvee_{m=0}^{kn-1} T^{-m} \alpha \right) \mid \alpha \in \mathcal{P}_X \right\} \\
&= \sup\left\{ \lim_{n\to\infty} \frac{1}{n} H_\mu\big(\alpha_0^{kn-1}\big) \mid \alpha \in \mathcal{P}_X \right\} \\
&= k \sup\left\{ \lim_{n\to\infty} \frac{1}{kn} H_\mu\big(\alpha_0^{kn-1}\big) \mid \alpha \in \mathcal{P}_X \right\} \\
&= k \sup\left\{ \lim_{N\to\infty} \frac{1}{N} H_\mu\big(\alpha_0^{N-1}\big) \mid \alpha \in \mathcal{P}_X \right\} \qquad \text{Let } N := kn \\
&= k \sup_{\alpha \in \mathcal{P}_X} h_\mu(T, \alpha) \\
&= k h_\mu(T)
\end{aligned}$$

2. it is sufficient to show that $h_\mu(T^{-1}) = h_\mu(T)$

$$\begin{aligned}
h_\mu(T^{-1}) &= \sup_{\alpha \in \mathcal{P}_X} h_\mu(T^{-1}, \alpha) \\
&= \sup \left\{ \lim_{n\to\infty} \frac{1}{n} H_\mu \left(\bigvee_{i=0}^{n-1} T^i \alpha \right) \mid \alpha \in \mathcal{P}_X \right\} \\
&= \sup \left\{ \lim_{n\to\infty} \frac{1}{n} H_\mu \left(T^{n-1} \left(\bigvee_{i=0}^{n-1} T^{-i} \alpha \right) \right) \mid \alpha \in \mathcal{P}_X \right\} \\
&= \sup \left\{ \lim_{n\to\infty} \frac{1}{n} H_\mu \left(\bigvee_{i=0}^{n-1} T^{-i} \alpha \right) \mid \alpha \in \mathcal{P}_X \right\} && \text{From (15)} \\
&= \sup_{\alpha \in \mathcal{P}_X} h_\mu(T, \alpha) \\
&= h_\mu(T)
\end{aligned}$$

Now, $h_\mu(T^k) = h_\mu(T^{|k|}) = |k| h_\mu(T)$

□

Now that we define the entropy and see it's fundamental properties there are some others theorem that can help us in computation of entropy.

Theorem 5.1.4 (Kolmogorov-Sinai [3]). *Let $(X, \mathscr{A}, \mu, T)$ be a MDS and $\alpha_1 \prec \alpha_2 \prec \cdots \prec \alpha_n \prec \ldots$ be an increasing sequence of partitions with finite entropy that $\bigcup_{i=1}^{\infty} \alpha_i$ generates $\mathscr{A}$ up to measure zero. then*

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

[3]For the proof and more details see [17] or [15]

5.2 Measure Entropy of Open Covers

Consider a measure dynamical space $(X, \mathscr{A}, \mu, T)$ For a given $\mathcal{U} \in C_X$ Romagnoli [14] introduced the concept of measure entropies relative to $\mathcal{U}$, i.e

$$H_\mu(\mathcal{U}) = \inf\{H_\mu(\alpha) \colon \alpha \in P_X, \alpha \succeq \mathcal{U}\} \tag{5.7}$$

$$h_\mu^-(T,\mathcal{U}) = \lim_{n\to\infty} \frac{1}{n} H_\mu(\mathcal{U}_0^{n-1}) = \inf_{n\in\mathbb{N}} \frac{1}{n} H_\mu(\mathcal{U}_0^{n-1}) \tag{5.8}$$

$$h_\mu^+(T,\mathcal{U}) = \inf\{h_\mu(T,\alpha) \colon \alpha \in P_X, \alpha \succeq \mathcal{U}\} \tag{5.9}$$

Remark 5.2.1.

$$\begin{aligned} h_\mu^+(T,\mathcal{U}) &= \inf\{h_\mu(T,\alpha) \colon \alpha \in P_X, \alpha \succeq \mathcal{U}\} \\ &= \inf_{\alpha\succeq\mathcal{U}} \inf_{n\in\mathbb{N}} \frac{1}{n} H_\mu(\alpha_0^{n-1}) = \inf_{n\in\mathbb{N}} \inf_{\alpha\succeq\mathcal{U}} \frac{1}{n} H_\mu(\alpha_0^{n-1}) \\ &\geq \inf_{n\in\mathbb{N}} \frac{1}{n} \inf_{\beta\succeq\mathcal{U}_0^{n-1}} H_\mu(\beta) = \inf_{n\in\mathbb{N}} \frac{1}{n} H_\mu(\mathcal{U}_0^{n-1}) \\ &= h_\mu^-(T,\mathcal{U}) \end{aligned}$$

and the equality is acquired if the system is invertible.

Proposition 5.2.1. *Let (X,T) be a TDS, $\mu \in M_T(X)$, $\mathcal{U} \in C_X$, and $\mathcal{U}_0^{m-1} = \bigvee_{j=0}^{m-1} T^{-i}\mathcal{U}$ then for every $m \in \mathbb{N}$*

$$\frac{1}{m} h_\mu^-(T^m,\mathcal{U}_0^{m-1}) = h_\mu^-(T,\mathcal{U})$$

Proof.

$$\begin{aligned}
\frac{1}{m}h_\mu^-(T^m,\mathcal{U}_0^{m-1}) &= \lim_{n\to\infty}\frac{1}{mn}H_\mu\left((\mathcal{U}_0^{m-1})_0^{n-1}\right)\\
&= \lim_{n\to\infty}\frac{1}{mn}H_\mu\left(\bigvee_{i=0}^{n-1}T^{-im}\bigvee_{j=0}^{m-1}T^{-m}\mathcal{U}\right)\\
&= \lim_{n\to\infty}\frac{1}{mn}H_\mu\left(\bigvee_{i=0}^{n-1}\bigvee_{j=0}^{m-1}T^{-(im+j)}\mathcal{U}\right)\\
&= \lim_{n\to\infty}\frac{1}{mn}H_\mu\left(\bigvee_{k=0}^{mn-1}T^{-k}\mathcal{U}\right)\\
&= \lim_{n\to\infty}\frac{1}{mn}H_\mu\left(\mathcal{U}_0^{mn-1}\right)\\
&= h_\mu^-(T,\mathcal{U})
\end{aligned}$$

□

Proposition 5.2.2. *Considering (X,T) as a TDS, $\mu \in M_T(X)$, and $\mathcal{U} \in C_X$, we have:*

$$h_\mu^-(T,\mathcal{U}) = \lim_{m\to\infty}\frac{1}{m}h_\mu^+(T^m,\mathcal{U}_0^{m-1})$$

Proof. As we saw above, for any $m \in \mathbb{N}$ we have:

$$h_\mu^-(T,\mathcal{U}) = \frac{1}{m}h_\mu^-(T^m,\mathcal{U}_0^{m-1}) \le \frac{1}{m}h_\mu^+(T^m,\mathcal{U}_0^{m-1})$$

Conversely,

$$\begin{aligned}
\frac{1}{m}h_\mu^+(T^m,\mathcal{U}_0^{m-1}) &= \frac{1}{m}\inf_{\alpha\succeq\mathcal{U}_0^{m-1}} h_\mu(T^m,\alpha) \\
&= \frac{1}{m}\inf_{\alpha\succeq\mathcal{U}_0^{m-1}}\left(\inf_{n\in\mathbb{N}}\frac{1}{n}H_\mu(\alpha_0^{n-1})\right) \\
&= \frac{1}{m}\inf_{\alpha\succeq\mathcal{U}_0^{m-1}}\left(\inf_{n\in\mathbb{N}}\frac{1}{n}H_\mu(\bigvee_{i=0}^{n-1}T^{-im}\alpha)\right) \\
&\le \frac{1}{m}\inf_{\alpha\succeq\mathcal{U}_0^{m-1}}\left(\inf_{n\in\mathbb{N}}\frac{1}{n}\sum_{i=0}^{n-1}H_\mu(T^{-im}\alpha)\right) \\
&= \frac{1}{m}\inf_{\alpha\succeq\mathcal{U}_0^{m-1}}\left(\inf_{n\in\mathbb{N}}\frac{1}{n}\sum_{i=0}^{n-1}H_\mu(\alpha)\right) \\
&= \frac{1}{m}\inf_{\alpha\succeq\mathcal{U}_0^{m-1}}\left(\inf_{n\in\mathbb{N}}H_\mu(\alpha)\right) \\
&= \frac{1}{m}\inf_{\alpha\succeq\mathcal{U}_0^{m-1}} H_\mu(\alpha) \\
&= \frac{1}{m}H_\mu(\mathcal{U}_0^{m-1})
\end{aligned}$$

Thus,

$$h_\mu^-(T,\mathcal{U}) \le \frac{1}{m}h_\mu^+(T^m,\mathcal{U}_0^{m-1}) \le \frac{1}{m}H_\mu(\mathcal{U}_0^{m-1})$$

Therefore by taking $\lim_{m\to\infty}$ we reach to the result. □

Theorem 5.2.1. *If (X,T) is a TDS, $\mu\in M_T(X)$, and $\mathcal{U},\mathcal{V}\in C_X$ then*

1. $H_\mu(\mathcal{U}) \le \log\mathcal{N}(\mathcal{U})$
2. $\mathcal{U}\succeq\mathcal{V} \implies H_\mu(\mathcal{U}) \ge H_\mu(\mathcal{V})$
3. $H_\mu(\mathcal{U}\vee\mathcal{V}) \le H_\mu(\mathcal{U}) + H_\mu(\mathcal{V})$
4. $H_\mu(T^{-1}\mathcal{U}) = H_\mu(\mathcal{U})$

Proof. 1.

$$H_\mu(\mathcal{U}) = \inf_{\alpha\succeq\mathcal{U}} H_\mu(\alpha) = \inf_{\alpha\succeq\mathcal{U}}\sum_{A\in\alpha}\mu(A)\log\mu(A) \le \inf_{\alpha\succeq\mathcal{U}}(\log\#\alpha) = \log\mathcal{N}(\mathcal{U})$$

2.

$$H_\mu(\mathcal{U}) = \inf_{\alpha \succeq \mathcal{U}} H_\mu(\alpha) \geq \inf_{\alpha \succeq \mathcal{V}} H_\mu(\alpha) = H_\mu(\mathcal{V})$$

3.

$$\begin{aligned}
H_\mu(\mathcal{U} \vee \mathcal{V}) &= \inf_{\gamma \succeq \mathcal{U} \vee \mathcal{V}} H_\mu(\gamma) \\
&\leq \inf_{\substack{\alpha \succeq \mathcal{U} \\ \beta \succeq \mathcal{V}}} H_\mu(\alpha \vee \beta) \\
&= \inf_{\alpha \succeq \mathcal{U}} \inf_{\beta \succeq \mathcal{V}} H_\mu(\alpha \vee \beta) \\
&\leq \inf_{\alpha \succeq \mathcal{U}} \inf_{\beta \succeq \mathcal{V}} \Big(H_\mu(\alpha) + H_\mu(\beta) \Big) \\
&= \inf_{\alpha \succeq \mathcal{U}} H_\mu(\alpha) + \inf_{\beta \succeq \mathcal{V}} H_\mu(\beta) \\
&= H_\mu(\mathcal{U}) + H_\mu(\mathcal{V})
\end{aligned}$$

4.

$$H_\mu(T^{-1}\mathcal{U}) = \inf_{\alpha \succeq T^{-1}\mathcal{U}} H_\mu(\alpha) = \inf_{\beta \succeq \mathcal{U}} H_\mu(T^{-1}\beta) = \inf_{\beta \succeq \mathcal{U}} H_\mu(\beta) = H_\mu(\mathcal{U})$$

□

Let us introduce another notation that will help us in proving the local variational principle. For a TDS (X, T), consider $\mathcal{U} \in C_X$ and let $\alpha \in \mathcal{P}_X$ be a borel partition generated by $\mathcal{U}$ i.e. the partition created by the intersection of elements of $\mathcal{U}$. define

$$\mathcal{P}^*(\mathcal{U}) := \Big\{ \beta \in \mathcal{P}_X \mid \beta \succeq \mathcal{U} \text{ and each atom of } \beta \text{ is union of some atoms of } \alpha \Big\}$$

Remark 5.2.2. *Since each element of $\beta \in \mathcal{P}^*(\mathcal{U})$ is the union of some elements of $\alpha := \{A_1, A_2, \ldots, A_n\}$, there is a one-to-one correspondence between each $\beta \in \mathcal{P}^*(\mathcal{U})$ and each partition of the set $\{A_1, A_2, \ldots, A_n\}$. So $\#\mathcal{P}^*(\mathcal{U}) = \mathcal{B}_n$ where $\mathcal{B}_n$ is the n-th bell* *number i.e. the number of partitions of a n-element set; As α is finite, this number is also finite.*

Lemma 5.2.1. *Considering the function* $\begin{cases} \phi\colon [0,1] \to \mathbb{R} \\ \phi(x) = -x\log x \end{cases}$, *for* $\delta, x, y \in [0,1]$ *that* $0 < \delta \leq x \leq y$, *we have* $\phi(x-\delta) + \phi(y+\delta) \leq \phi(x) + \phi(y)$

Proof. Due to mean value theorem we have:

$$\exists t \in [x-\delta, x] \text{ s.t. } \frac{\phi(x) - \phi(x-\delta)}{\delta} = \phi'(t),$$
$$\exists s \in [y, y+\delta] \text{ s.t. } \frac{\phi(y+\delta) - \phi(y)}{\delta} = \phi'(s)$$

notice that $\phi'' \leq 0$ so ϕ' is decreasing function so since $t \leq s$ we have $\phi'(s) \leq \phi'(t)$. Thus

$$\begin{aligned} \frac{\phi(y+\delta) - \phi(y)}{\delta} &= \phi'(s) \\ &\leq \phi'(t) \\ &= \frac{\phi(x) - \phi(x-\delta)}{\delta} \end{aligned}$$

therefore

$$\frac{\phi(y+\delta) - \phi(y)}{\delta} \leq \frac{\phi(x) - \phi(x-\delta)}{\delta}$$
$$\implies \phi(x-\delta) + \phi(y+\delta) \leq \phi(x) + \phi(y)$$

□

Proposition 5.2.3. *Consider a TDS (X,T), and a Borel probability measure $\mu \in M_T(X)$. For any $\mathcal{U} \in C_X$*

$$H_\mu(\mathcal{U}) = \min_{\beta \in \mathcal{P}^*(\mathcal{U})} H_\mu(\beta)$$

Proof. It is sufficient to show that

$$\forall \beta \in \mathcal{P}_X : \mathcal{U} \preceq \beta \implies \exists \beta^* \in \mathcal{P}^*(\mathcal{U}) \text{ s.t. } H_\mu(\beta) \geq H_\mu(\beta^*)$$

Let $\mathcal{U} = \{U_1, U_2, \ldots, U_n\}$ and also consider any $\beta = \{B_1, B_2, \ldots, B_b\} \in \mathcal{P}_X$ that $\beta \succeq \mathcal{U}$. Choose $i_1 \in \{1, 2, \ldots, b\}$ such that

$$\mu(B_{i_1}) = \max\{\mu(B_i) : 1 \leq i \leq b\}$$

by the assumption that $\beta \succeq \mathcal{U}$, choose $j_1 \in \{1, 2, \ldots, n\}$ such that $B_{i_1} \subseteq U_{j_1}$. Define

$$\beta(j_1) := \{U_{j_1}\} \cup \{B_i \setminus U_{j_1} : 1 \leq i \leq b\}$$

Notice that $\beta(j_1) \in C_X$ and $\beta(j_1) \succeq \mathcal{U}$.
Let $I_1 := \{i \in \{1, 2, \ldots, b\} : B_i \setminus U_{j_1} \neq \emptyset\}$ and $b_1 := \# I_1$ and for each $i \in \{1, 2, \ldots, b_1\}$ let

$$r_i := \begin{cases} \min I_1 & i = 1 \\ \min(I_1 \setminus \bigcup_{j=1}^{i-1} r_j) & i \neq 1 \end{cases}$$

so that $I_1 = \{r_1, r_1, \ldots, r_{b_1}\}$.
Note that now, $\beta(j_1) = \{U_{j_1}, B_1 \setminus U_{j_1}, \ldots, B_{r_b} \setminus U_{j_1}\}$. For making all simple, let $B_i^{(1)} := B_{r_i} \setminus U_{j_1}$ for each $i \in \{1, 2, \ldots, b_1\}$. Thus,

$$\beta(j_1) = \{U_{j_1}, B_1^{(1)}, B_2^{(1)}, \ldots, B_{b_1}^{(1)}\}$$

If $X \setminus U_{j_1} = \emptyset$ then $\beta(j_1) = \{U_{j_1}\} = \{X\}$ and $\beta(j_1) \in \mathcal{P}^*(\mathcal{U})$. Let $\beta^* = \beta(j_1)$ then since $H_\mu(\beta^*) = 0$, we trivially have $H_\mu(\beta) \geq H_\mu(\beta^*) = 0$.
If $X \setminus U_{j_1} \neq \emptyset$ then we choose $i_2 \in \{1, 2, \ldots b_1\}$ such that

$$\phi(\mu(B_{i_2}^1)) = \max\{\phi(\mu(B_i^{(1)})) : 1 \leq i \leq b_1\}$$

Since $\beta(j_1) \succeq \mathcal{U}$ and for each $i \in \{1, 2, \ldots, b_1\}$, $B_i^{(1)} \cap U_{j_1} = \emptyset$, choose $j_2 \in \{1, 2, \ldots, n\} \setminus j_1$ such that $B_{i_2}^{(1)} \subseteq U_{j_2} \setminus U_{j_1}$. Define

$$\beta(j_1, j_2) := \{U_{j_1}, U_{j_2} \setminus U_{j_1}\} \cup \{B_i^{(1)} \setminus U_{j_2} \colon i \in \{1, 2, \ldots, b_1\}\}$$

Notice that $\beta(j_1, j_2) \in C_X$ and $\beta(j_1, j_2) \succeq \mathcal{U}$. Repeat all the process above, to find $b_2 \in \{1, 2, \ldots, b_1\}$ and $B_i^{(2)} := B_i^{(1)} \setminus U_{j_2}$ for $i \in \{1, 2, \ldots, b_2\}$; So that

$$\beta(j_1, j_2) = \{U_{j_1}, U_{j_2} \setminus U_{j_1}, B_1^{(2)}, \ldots, B_{b_2}^{(2)}\}$$

Now we have

$$\begin{aligned}
H_\mu(\beta(j_1, j_2)) &= \sum_{V \in \beta(j_1, j_2)} \phi(\mu(V)) \\
&= \phi(\mu(U_{j_1})) + \phi(\mu(U_{j_2} \setminus U_{j_1})) + \phi(\mu(B_1^{(2)})) + \cdots + \phi(\mu(B_{b_2}^{(2)})) \\
&\leq \phi(\mu(U_{j_1})) + \phi(\mu(B_1^{(1)})) + \cdots + \phi(\mu(B_{b_1}^{(1)})) \\
&= \sum_{V \in \beta(j_1)} \\
&= H_\mu(\beta(j_1, j_2))
\end{aligned}$$

if $(X \setminus U_{j_1}) \setminus (U_{j_2} \setminus U_{j_1}) = \emptyset$ then $\beta(j_1, j_2) = \{U_{j_1}, U_{j_2} \setminus U_{j_1}\} \in \mathcal{P}^*(\mathcal{U})$ and the result obtained. But if not, we use the structure above to obtain

$$\{j_1, j_2, \ldots, j_r\} \subseteq \{1, 2, \ldots n\}$$

for a $r \leq n$, such that

$$\bigcup_{i=1}^{r-1} U_{j_i} \neq X \text{ and } \bigcup_{i=1}^{r} U_{j_i} = X$$

and

$$\beta^* := \beta(j_1, \ldots, j_r) = \left\{U_{j_1}, U_{j_2} \setminus U_{j_1}, \ldots, U_{j_r} \setminus \bigcup_{i=1}^{r-1} U_{j_i}\right\} \in \mathcal{P}^*(\mathcal{U})$$

By induction, using lemma (5.2.1) we reach to the desired inequality

$$H_\mu(\beta^*) \leq H_\mu(\beta)$$

□

Theorem 5.2.2. *Consider a TDS (X, T), $\mu \in M_T(X)$, and $\mathcal{U}, \mathcal{V} \in C_X$.*

1. $h^-_\mu(T,\mathcal{U}) = \frac{1}{m} h^-_\mu(T^m, \mathcal{U}_0^{m-1}). \forall m \in \mathbb{N}$
2. $h^+_\mu(T,\mathcal{U}) \geq \frac{1}{m} h^+_\mu(T^m, \mathcal{U}_0^{m-1}), \forall m \in \mathbb{N}$
3. $h^-_\mu(T,\mathcal{U}) = \lim_{m\to\infty} \frac{1}{m} h^+_\mu(T^m, \mathcal{U}_0^{m-1})$
4. $h^-_\mu(T,\mathcal{U}) \leq h^+_\mu(T,\mathcal{U})$
5. $h^+_\mu(T,\mathcal{U} \vee \mathcal{V}) \leq h^+_\mu(T,\mathcal{U}) + h^+_\mu(T,\mathcal{V})$
6. $h^-_\mu(T,\mathcal{U} \vee \mathcal{V}) \leq h^-_\mu(T,\mathcal{U}) + h^-_\mu(T,\mathcal{V})$
7. $\mathcal{U} \succeq \mathcal{V} \implies h^-_\mu(T,\mathcal{U}) \geq h^-_\mu(T,\mathcal{V})$

Theorem 5.2.3. *Let $\pi\colon (X,T) \to (Y,S)$ be a factor map between two TDS and $\mathcal{U} \in C_Y^\circ$. Then for any $\mu \in M_T(X)$*

$$h^-_\mu(T, \pi^{-1}\mathcal{U}) = h^-_{\pi_*\mu}(S, \mathcal{U})$$

Proposition 5.2.4. *For a TDS (X, T), $\mu \in M_T(X)$*

$$h_\mu(T) = \sup\{h_\mu(T,\mathcal{U}) \mid \mathcal{U} \in C_X^\circ\}$$

Chapter 6

Pressure and Variational Principle

One of the prime theorem, maybe the most important of them, connecting a TDS to MDS is **Variational Principle**.

For a TDS (X, T) let $M_T(X)$ be the set of all Borel invariant probability measures. In 1969, Goodwyn [10] showed that $\forall \mu \in M_T(X) : h_\mu(T) \leq h_{\text{top}}(T)$. In 1971, Goodman [9] showed that $\sup\{h_\mu(T) \mid \mu \in M_T(X)\} \geq h_{\text{top}}(T)$.

In this section our guideline is "Foundations of ergodic theory" [15]. Here, it arises as a special case of a more general statement, the variational principle for the pressure, which is due to Walters [16].

6.1 Pressure

The pressure $P(T, \phi)$ is weighted version of topological entropy $h_{\text{top}}(T)$; Where the weights are determined with the continues function $\phi \colon X \to \mathbb{R}$. We call such a function a potential. In special case we'll see that:

$$P(T, 0) = h_{\text{top}}(T)$$

Where $0 \colon X \to \mathbb{R}$ is a zero function.

For every $n \geq 1$, let

$$\begin{cases} \phi_n \colon X \to \mathbb{R} \\ \phi_n(x) = \sum_{i=0}^{n-1} \phi \circ T^i(x) \end{cases}$$

Now for every $\alpha \in C_X^\circ$, let

$$P_n(T, \phi, \alpha) := \inf\{\sum_{U \in \gamma} \sup_{x \in U} \exp \phi_n(x) \mid \gamma \in C_X^\circ,\ \gamma \succeq \alpha_0^n\} \tag{6.1}$$

Remark 6.1.1. *if $\mathcal{U} \preceq \mathcal{V}$ denotes that $\mathcal{V}$ is a refinement of $\mathcal{U}$,*

$$\begin{aligned} \forall \beta,\ \eta \in C_X^\circ \colon \beta \preceq \eta \implies & \{\sum_{U \in \gamma} \sup_{x \in U} e^{\phi_n(x)} \mid \gamma \overset{\text{cover}}{\subseteq} \beta\} \subseteq \{\sum_{V \in \delta} \sup_{y \in V} e^{\phi_n(y)} \mid \delta \overset{\text{cover}}{\subseteq} \eta\} \\ \implies & \inf\{\sum_{V \in \delta} \sup_{y \in V} e^{\phi_n(y)} \mid \delta \overset{\text{cover}}{\subseteq} \eta\} \leq \inf\{\sum_{U \in \gamma} \sup_{x \in U} e^{\phi_n(x)} \mid \gamma \overset{\text{cover}}{\subseteq} \beta\} \\ \implies & P_n(T, \phi, \eta) \leq P_n(T, \phi, \beta) \end{aligned}$$

As we saw in proposition (5.2.3) here we the similar:

Proposition 6.1.1. *For each $n \in \mathbb{N}$ we have*

$$\inf\{\sum_{U \in \gamma} \sup_{x \in U} e^{\phi_n(x)} \mid\ \gamma \succeq \alpha\} = \min\{\sum_{U \in \gamma} \sup_{x \in U} e^{\phi_n(x)} \mid \gamma \in \mathcal{P}^*(\alpha)\}$$

Corollary 6.1.1. *For each $n \in \mathbb{N}$ we have*

$$P_n(T, \phi, \alpha) = \min\{\sum_{U \in \gamma} \sup_{x \in U} e^{\phi_n(x)} \mid \gamma \in \mathcal{P}^*(\alpha_0^{n-1})\}$$

Proposition 6.1.2. $\log P_n(T, \phi, \alpha)$ *is subadditiv; i.e.*

$$\log P_{n+m}(T, \phi, \alpha) \leq \log P_n(T, \phi, \alpha) + \log P_m(T, \phi, \alpha)$$

Proof. For simplicity, let P_n stands for $P_n(T,\phi,\alpha)$ and $\mathcal{U} \overset{\text{cover}}{\subseteq} \mathcal{V}$ denotes that $\mathcal{U}$ is a subcover of $\mathcal{V}$.

Since

$$\phi_{n+m} = \sum_{i=0}^{n+m-1} \phi \circ T^i = \phi_n + \sum_{i=n}^{m+n-1} \phi \circ T^i = \phi_n + (\sum_{i=0}^{m-1} \phi \circ T^i) \circ T^n = \phi_n + \phi_m \circ T^n,$$

we have

$$\begin{aligned}
P_{n+m} &= \inf\{\sum_{U\in\gamma} \sup_{x\in U} e^{\phi_{n+m}(x)} \mid \gamma \overset{\text{cover}}{\subseteq} \alpha_0^{n+m}\} \\
&= \inf\{\sum_{U\in\gamma} \sup_{x\in U} e^{\phi_n(x)+\phi_m\circ T^n(x)} \mid \gamma \overset{\text{cover}}{\subseteq} \alpha_0^{n+m}\} \\
&= \inf\{\sum_{U\in\gamma} \sup_{x\in U} e^{\phi_n(x)} e^{\phi_m\circ T^n(x)} \mid \gamma \overset{\text{cover}}{\subseteq} \alpha_0^{n+m}\} \\
&\leq \inf\{\sum_{U\in\gamma} \sup_{x\in U} e^{\phi_n(x)} \sum_{U\in\gamma} \sup_{x\in U} e^{\phi_m\circ T^n(x)} \mid \gamma \overset{\text{cover}}{\subseteq} \alpha_0^{n+m}\} \\
&= \inf\{\sum_{U\in\gamma} \sup_{x\in U} e^{\phi_n(x)} \mid \gamma \overset{\text{cover}}{\subseteq} \alpha_0^{n+m}\} \inf\{\sum_{U\in\gamma} \sup_{x\in U} e^{\phi_m\circ T^n(x)} \mid \gamma \overset{\text{cover}}{\subseteq} \alpha_0^{n+m}\} \\
&\leq \inf\{\sum_{U\in\gamma} \sup_{x\in U} e^{\phi_n(x)} \mid \gamma \overset{\text{cover}}{\subseteq} \alpha_0^{n+m}\} \inf\{\sum_{U\in\gamma} \sup_{y\in U} e^{\phi_m(y)} \mid \gamma \overset{\text{cover}}{\subseteq} \alpha_{-n}^{m}\} \\
&\leq \inf\{\sum_{U\in\gamma} \sup_{x\in U} e^{\phi_n(x)} \mid \gamma \overset{\text{cover}}{\subseteq} \alpha_0^{n}\} \inf\{\sum_{U\in\gamma} \sup_{y\in U} e^{\phi_m(y)} \mid \gamma \overset{\text{cover}}{\subseteq} \alpha_0^{m}\} \\
&= P_n P_m.
\end{aligned}$$

□

Now that the subadditivity of P_n is proved, we can define the pressure of a map and a potential with respect to a cover:

Definition 6.1.1. *Pressure of a homeomorphism $T\colon X \to X$ and a potential $\phi\colon X \to \mathbb{R}$ with respect to a cover α is defined as:*

$$P(T,\phi,\alpha) := \lim_{n\to\infty} \frac{1}{n} \log P_n(T,\phi,\alpha)$$

Definition 6.1.2. *Pressure of a homeomorphism $T\colon X \to X$ and a potential $\phi\colon X \to \mathbb{R}$, is defined as:*

$$P(T,\phi) := \lim_{\text{diam }\alpha\to 0} P(T,\phi) \tag{6.2}$$

Note that according to Remark (6.1.1), pressure in (6.2) is well-defined. In the above definitions if we use 'inf' instead of 'sup', we don't reach to the subaditivity property in Proposition (6.1.2), So we have to take 'lim sup' or 'lim inf' instead of 'lim'; Although, we can have the same result in (6.2) by considering either 'lim sup' or 'lim inf'.
For showing these concepts, let us begin with introducing some notation:

$$\begin{aligned}
Q_n(T,\phi,\alpha) &:= \inf\{\sum_{U\in\gamma} \inf_{x\in U} e^{\phi_n(x)} \mid \gamma \overset{\text{cover}}{\subseteq} \alpha_0^n\} \\
Q_n^+(T,\phi,\alpha) &:= \limsup_{n\to\infty} \frac{1}{n} \log Q_n(T,\phi,\alpha) \\
Q_n^-(T,\phi,\alpha) &:= \liminf_{n\to\infty} \frac{1}{n} \log Q_n(T,\phi,\alpha)
\end{aligned}$$

Proposition 6.1.3. *For a TDS (X,T) and a potential $\phi\colon X \to \mathbb{R}$,*

$$P(T,\phi) = \lim_{\text{diam }\alpha\to 0} Q_n^+(T,\phi,\alpha) = \lim_{\text{diam }\alpha\to 0} Q_n^-(T,\phi,\alpha) \tag{6.3}$$

Proof. The potential ϕ is continues map on a compact space, so it's uniformly continues; i.e.

$$\forall\epsilon>0\colon \exists\delta>0 \text{ s.t. } \forall x,y\in X\colon d(x,y)<\delta \implies |\phi(x)-\phi(y)|<\epsilon$$

So for every $\epsilon>0$, choose $\delta>0$, so for every $C\subseteq X$ that diam $C<\delta$ we have:

$$\inf_{x\in C}\phi_n(x) \le \sup_{x\in C}\phi_n(x) \le n\epsilon + \inf_{x\in C}\phi_n(x)$$

Thus,

$$Q_n(T,\phi,\alpha) \le P_n(T,\phi,\alpha) \le e^{n\epsilon} Q_n(T,\phi,\alpha)$$

as a result,

$$\begin{aligned}\frac{1}{n}\log Q_n(T,\phi,\alpha) &\leq \frac{1}{n}\log P_n(T,\phi,\alpha)\\ &= \frac{1}{n}\log P_n(T,\phi,\alpha) \leq \frac{1}{n}\log Q_n(T,\phi,\alpha)+\epsilon\end{aligned}$$

by taking $\limsup_{n\to\infty}$ we have:

$$\begin{aligned}\limsup_{n\to\infty}\frac{1}{n}\log Q_n(T,\phi,\alpha) &\leq \limsup_{n\to\infty}\frac{1}{n}\log P_n(T,\phi,\alpha)\\ &= \liminf_{n\to\infty}\frac{1}{n}\log P_n(T,\phi,\alpha) \leq \liminf_{n\to\infty}\frac{1}{n}\log Q_n(T,\phi,\alpha)+\epsilon\end{aligned}$$

now as diam $\alpha \to 0$, we have:

$$\lim_{\text{diam }\alpha\to 0} Q_n^+(T,\phi,\alpha) \leq P(T,\phi) \leq \lim_{\text{diam }\alpha\to 0} Q_n^-(T,\phi,\alpha)+\epsilon$$

now we can take $\lim_{\epsilon\to 0}$ and use the fact that $Q_n^-(T,\phi,\alpha) \leq Q_n^+(T,\phi,\alpha)$.

$$\lim_{\text{diam }\alpha\to 0} Q_n^+(T,\phi,\alpha) \leq P(T,\phi) \leq \lim_{\text{diam }\alpha\to 0} Q_n^-(T,\phi,\alpha) \leq \lim_{\text{diam }\alpha\to 0} Q_n^+(T,\phi,\alpha)$$

so all the inequalities are equality. □

For completing our tools for proofing the variational principle, we are going to introduce equivalence definitions of pressure.

Definition 6.1.3. *Let (X,T) be a dynamical system. We call $E \subseteq X$, a (n,ϵ)-generating set of $K \subseteq X$, when the collection $\{B(x;n,\epsilon) \mid x \in E\}$ is a cover of K; where*

$$B(x;n,\epsilon) = \{y \in X \mid \forall 0 \leq i \leq n-1\colon d(T^i x, T^i y) < \epsilon\}$$

is called a dynamical ball of center x, length n, and radius ϵ.

Definition 6.1.4. *Let (X,T) be a dynamical system. We call $E \subseteq X$ a (n,ϵ)-separated* *set, when*

$$\forall x, y \in E \colon y \notin B(x; n, \epsilon)$$

or equivalently,

$$\forall x \in X \colon E \cap B(x; n, \epsilon) = \{x\}$$

where $B(x; n, \epsilon)$ is a dynamical ball mentioned before.

Example 6.1.1. *Consider the dynamical system (X, T),*

- *For every $n \in \mathbb{N}$, and every $\epsilon > 0$, $K \subseteq X$ is a trivial (n, ϵ)-generating set for K.*
- *For every $n \in \mathbb{N}$, every $\epsilon > 0$, and every $x \in X$, $\{x\}$ is a trivial (n, ϵ)-separated set.*

Remark 6.1.2. *Every maximal (n, ϵ)-separated set is a (n, ϵ)-generating set of X.*

Proof. Let $E \subseteq X$ be a maximal (n, ϵ)-separated set. The maximality means that

$$\forall z \in X \setminus E \colon E \cap B(z, n, \epsilon) \neq \emptyset$$

so there exists $y \in E \cap B(z, n, \epsilon)$.

Notice that $y \in B(z, n, \epsilon)$, so by definition and the symmetric property of metric, we have: $z \in B(y, n, \epsilon)$ and $y \in E$, Thus

$$\forall z \in X \colon z \in \bigcup_{x \in E} B(x; n, \epsilon)$$

$\square$

Now consider some notations:

$$
\begin{aligned}
G_n(T,\phi,\epsilon) &:= \inf\{\sum_{x\in E} e^{\phi_n(x)} \mid E\subseteq X \text{ be a } (n.\epsilon)\text{-generating set for } X\}\\
S_n(T,\phi,\epsilon) &:= \sup\{\sum_{x\in E} e^{\phi_n(x)} \mid E\subseteq X \text{ be a } (n.\epsilon)\text{-separated set}\}\\
G(T,\phi,\epsilon) &:= \overline{\lim_{n\to\infty}} \frac{1}{n}\log G_n(T,\phi,\epsilon)\\
S(T,\phi,\epsilon) &:= \overline{\lim_{n\to\infty}} \frac{1}{n}\log S_n(T,\phi,\epsilon)\\
G(T,\phi) &:= \lim_{\epsilon\to 0} G(T,\phi,\epsilon)\\
S(T,\phi) &:= \lim_{\epsilon\to 0} S(T,\phi,\epsilon)
\end{aligned}
$$

Proposition 6.1.4. *Let (X,T) be a TDS and $\phi\colon X\to\mathbb{R}$ be a potential.*

$$P(T,\phi) = S(T,\phi) = G(T,\phi)$$

Proof. 1. $S(T,\phi)\geq G(T,\phi)$

For every $n\in\mathbb{N}$, $\epsilon>0$, and F an arbitrary maximal $(n.\epsilon)$-separated set:

$$
\begin{aligned}
S_n(T,\phi,\epsilon) &= \sup\{\sum_{x\in E} e^{\phi_n(x)} \mid E\subseteq X \text{ be a } (n.\epsilon)\text{-separated set}\}\\
&\geq \sup\{\sum_{x\in E} e^{\phi_n(x)} \mid E\subseteq X \text{ be a Maximal } (n.\epsilon)\text{-separated set}\}\\
&\geq \sum_{x\in F} e^{\phi_n(x)}\\
&\geq \inf\{\sum_{x\in E} e^{\phi_n(x)} \mid E\subseteq X \text{ be a } (n.\epsilon)\text{-generating set}\} \qquad \text{(According to Remark(6.1.2))}\\
&= G_n(T,\phi,\epsilon)
\end{aligned}
$$

Thus, $S(T,\phi)\geq G(T,\phi)$.

2. $P(T,\phi)\geq S(T,\phi)$

Since ϕ is uniformly continues,

$$\forall \epsilon > 0\colon \exists \delta > 0 \text{ s.t. } \forall x, y \in X\colon d(x,y) < \delta \implies |\phi(x) - \phi(y)| < \epsilon$$

Now for arbitrary $\epsilon > 0$, let α be an open cover of X that $\operatorname{diam}(\alpha) < \delta$ and $E \subseteq X$ be a (n, δ)-separating set.
For every $\gamma \overset{\text{cover}}{\subseteq} \alpha_0^{n-1}$, every member of E is in at least one member of γ and since E is (n, δ)-separated set, no two members of E could be in one cell of our subcover γ (notice that each open set in γ is in $B(x, n, \delta)$); So, each member of E is located in exactly one member of γ. So we have:

$$\sum_{x \in E} e^{\phi_n(x)} \leq \sum_{U \in \gamma} \sup_{y \in U} e^{\phi_n(y)}$$

remember our fundamental definition of pressure (6.1).
Thus $S_n(T, \phi, \delta) \leq P_n(T, \phi, \alpha)$; Therefore $S(T, \phi, \delta) \leq P(T, \phi, \alpha)$.
By letting $\delta \to 0$, we have $S(T, \phi) \leq P(T, \phi)$.

3. $G(T, \phi) \geq P(T, \phi)$

   For arbitrary $\epsilon > 0$, let α be an open cover of X that $\operatorname{diam}(\alpha) < \delta$, let $\rho > 0$ be the Lebesgue cover number of α, and $E \subseteq X$ be a (n, ρ)-generating set of X.
   A_x^i denotes the member of α containing $B(T^i(x), \rho) := B(T^i(x), n = 0, \rho)$. Set

$$\gamma(x) := \bigcap_{i=0}^{n-1} T^{-i}(A_x^i) \in \alpha_0^{n-1}$$

$B(x,n,\rho) \subseteq \gamma(x)$ since

$$\begin{aligned}
\forall y \in B(x,n,\rho)\colon \mathrm{d}(x,y) < \rho \quad & \implies y \in B(x,\rho) \\
\mathrm{d}(Tx,Ty) < \rho \quad & \implies Ty \in B(Tx,\rho) \\
& \implies y \in T^{-1}B(Tx,\rho) \\
\mathrm{d}(T^2x,T^2y) < \rho \quad & \implies T^2y \in B(T^2x,\rho) \\
& \implies y \in T^{-2}B(T^{-2}x,\rho) \\
\vdots \quad & \\
\mathrm{d}(T^{n-1}x,T^{n-1}y) < \rho \quad & \implies T^{n-1}y \in B(T^{n-1}x,\rho) \\
& \implies y \in T^{-(n-1)}B(T^{n-1}x,\rho)
\end{aligned}$$

so

$$y \in \bigcap_{i=0}^{n-1} T^{-i}B(T^ix,\rho) \subseteq \bigcap_{i=0}^{n-1} T^{-i}A_x^i = \gamma(x)$$

E was (n,ρ)-generating set of X; So

$$\gamma := \{\gamma(x) \mid x \in E\} \overset{\text{cover}}{\subseteq} \alpha_0^{n-1}$$

Again by uniformly continuity of ϕ we have:

$$\forall \epsilon > 0\colon \exists \delta > 0 \text{ s.t. } \forall x,y \in X\colon d(x,y) < \delta \implies |\phi(x) - \phi(y)| < \epsilon$$

So $\forall x \in E\colon \sup_{y\in\gamma(x)} \phi_n(y) \leq n\epsilon + \phi_n(x)$. Therefore,

$$\sum_{U\in\gamma} \sup_{y\in U} e^{\phi_n(y)} \leq e^{n\epsilon} \sum_{x\in E} e^{\phi_n(x)}$$

Thus,

$$P(T,\phi < \alpha) \leq \epsilon + \varliminf_{n\to\infty} \frac{1}{n} \log G_n(T,\phi,\rho) \leq \epsilon + G(T,\phi,\rho)$$

now let $\rho \to 0$. we obtain: $P(T,\phi,\alpha) \leq \epsilon + G(T,\phi)$.

Let $\epsilon \to 0$: $P(T,\phi) \leq G(T,\phi)$.

All in all,

$$G(T,\phi) \geq P(T,\phi) \geq S(T,\phi) \geq G(T,\phi)$$

all the inequalities above are equalities. □

Proposition 6.1.5. *Let* $\pi\colon (X,T) \to (Y,S)$ *be a factor map between two TDS,* $\phi : Y \to \mathbb{R}$ *be a potential, and* $\mathcal{U} \in C_Y^{\circ}$*. Then*

$$P(T,\phi \circ \pi, \pi^{-1}\mathcal{U}) = P(S,\phi,\mathcal{U})$$

To finish this section it is worth mentioning that all the propositions and definitions mentioned before, have equivalent phenomenon in entropy theory we studied before, by letting the potential *phi* equals to the zero function. It is easy to observe by letting zero function be our potential, the phenomenon pressure become the concept of topological entropy. Thus we can obtain some equivalent definition of topological entropy by using above propositions and zero function as potential.

6.2 Variational principle

In this section we are going to prove the extension of Variational Principle of Entropy i.e.

Theorem 6.2.1. *(Variational Principle of Entropy) For a TDS* (X,T) *let* $M_T(X)$ *be the set of all Borel invariant probability measures then,*

$$h_{\text{top}}(T) = \sup\{h_\nu(T) \mid \nu \in M_T(X)\}$$

Instead of proving that theorem we are going to see the Variational Principle of Pressure i.e. theorem (6.2.2). As you can guess, by letting zero function be our potential ϕ, we could reach the Variational Principle of Entropy 6.2.1.
Before proving the main theorem, we need some lemma.

Lemma 6.2.1. *Any probability measure μ on a metric space is regular.*

Proof. We introduce the sketch of proof.

Let

$$\mathcal{B} := \{B \subseteq X \mid \forall \epsilon > 0\colon \exists \text{ a closed set } F, \text{ an open set A s.t. } F \subseteq B \subseteq A \text{ and } \mu(A \backslash F) < \epsilon\}$$

$\mathcal{B}$ contain all the closed set. and it's indeed a σ-algebra. Thus $\mathcal{B}$ contains all the Borel subsets of X. $\square$

Lemma 6.2.2. *For every $k \geq 1$ and $\epsilon > 0$, there exists a $\delta > 0$ such that for every two finite partition $\mathcal{P} = \{P_1, P_2, \ldots, P_k\}$ and $\mathcal{Q} = \{Q_1, Q_2, \ldots, Q_k\}$ we have:*

$$\forall i \in \{1, 2, \ldots, k\}\colon \mu(P_i \triangle Q_i) < \delta \implies H_\mu(Q \mid P) < \epsilon$$

Proof. Fix $\epsilon > 0$ and $k \geq 1$.

The function $\begin{cases} \phi\colon [0,1] \to \mathbb{R} \\ \phi(x) = -x \log x \end{cases}$ is continues, there exist $\rho > 0$ such that for each $x \in [0, \rho) \cap (1 - \rho, 1]$: $\phi(x) < \frac{\epsilon}{k^2}$. Put $\delta := \frac{\rho}{k}$.

Now we have:

$$\begin{aligned}
\forall i \neq j \in \{1, 2, \ldots, k\}\colon \mu(P_i \cap Q_j) &= \mu\Bigg(\Big((P_i \setminus Q_i) \cap Q_j\Big) \cup \Big((Q_i \setminus P_i) \cap Q_j\Big)\Bigg) \\
&= \mu\Big((P_i \triangle Q_i) \cap Q_j\Big) \\
&\leq \mu(P_i \triangle Q_i) < \delta \leq \rho
\end{aligned}$$

also we have:

$$
\begin{aligned}
\mu\left(\bigcup_{i=1}^{k}(P_i \cap Q_i)\right) &= \sum_{i=1}^{k} \mu\Big((P_i \cap Q_i)\Big) \\
&= \sum_{i=1}^{k} \mu\Big((P_i \cup Q_i) \setminus (P_i \triangle Q_i)\Big) \\
&= \sum_{i=1}^{k} \Big(\mu(P_i \cup Q_i) - \mu(P_i \triangle Q_i)\Big) \\
&> \sum_{i=1}^{k} \Big(\mu(P_i) - \delta\Big) = 1 - k\delta = 1 - \rho
\end{aligned}
$$

All in all we have:

$$
\mu(P_i \cap Q_j) \begin{cases} < \rho & i \neq j \\ > 1 - \rho & i = j \end{cases}
$$

So by considering the partition

$$
\mathcal{R} := \{P_i \cap Q_j \mid i \neq j \in \{1, 2, \ldots, k\}\} \cup \{\bigcup_{i=1}^{k}(P_i \cap Q_i)\}
$$

we have:

$$
H_\mu(\mathcal{R}) = \sum_{R \in \mathcal{R}} \phi\Big(\mu(R)\Big) < \#\mathcal{R}\ \frac{\epsilon}{k^2} \leq \epsilon
$$

Note that by definition of $\mathcal{R}$, $P \vee \mathcal{R} = P \vee Q$; Thus

$$
\begin{aligned}
H_\mu(Q \mid P) = H_\mu(P \vee Q) - H_\mu(P) &= H_\mu(\mathcal{R} \mid \mathcal{P}) \\
&< H_\mu(\mathcal{R}) \leq \epsilon
\end{aligned}
$$

□

Lemma 6.2.3. *If* $a_1, a_2, \ldots, a_k \in \mathbb{R}$, $p_1, p_2, \ldots, p_k \in \mathbb{R}^+ \cup \{0\}$, $\sum_{i=1}^{k} p_i = 1$, *and* $A := \sum_{i=1}^{k} e^{a_i}$, *then*

$$
\sum_{i=1}^{k} p_i(a_i - \log p_i) \leq \log A
$$

Proof. Let $t_i := \frac{e^{a_i}}{A}$ and $x_i := \frac{p_i}{e^{a_i}}$. Observe that $\sum_{i=1}^{k} t_i = 1$.

the function $\begin{cases} \phi\colon [0,1] \to \mathbb{R} \\ \phi(x) = -x \log x \end{cases}$ is concave, i.e.

$$\forall x, y \in [0,1]\colon \forall 0 \leq \theta \leq 1\colon \theta\phi(x) + (1-\theta)\phi(y) \leq \phi\Big(\theta x + (1-\theta)y\Big)^{1}$$

Consequently we have:

$$\sum_{i=1}^{k} t_i \phi(x_i) \leq \phi\Big(\sum_{i=1}^{k} t_i x_i\Big) \tag{6.4}$$

and

$$\begin{aligned} t_i\phi(x_i) &= -\frac{e^{a_i}}{A}\left(\frac{p_i}{e^{a_i}} \log \frac{p_i}{e^{a_i}}\right) &&= \frac{p_i}{A}\Big(a_i - \log p_i\Big) \\ \sum_{i=1}^{k} t_i x_i &= \sum_{i=1}^{k}\left(\frac{e^{a_i}}{A}\frac{p_i}{e^{a_i}}\right) &&= \frac{1}{A} \end{aligned}$$

Thus by (6.4) we reach to:

$$\sum_{i=1}^{k} \frac{p_i}{A}\Big(a_i - \log p_i\Big) \leq \frac{1}{A} \log A$$

□

Theorem 6.2.2 (Variational Principle of Pressure)**.** *If X be a compact metric space, $T\colon X \to X$ be a continues function, and $M_T(X)$ denotes the set of all T-invariant probability measures on X, then for every potential $\phi\colon X \to \mathbb{R}$ we have*

$$P(T, \phi) = \sup\{h_\nu(T) + \int \phi \; \mathrm{d}\nu \mid \nu \in M_T(X)\}$$

Proof.

- $\forall \nu \in M_T(X)\colon h_\nu(T) + \int \phi \; \mathrm{d}\nu \leq P(T, \phi)$

For proving the fact above, fix a finite partition $\mathcal{P} = \{P_1, P_2, \ldots, P_s\}$ of space X.

[1] For more detail see ‘Convex Optimization’ [3]

Due to Lemma (6.2.1), ν is regular, so by the definition of regularity:

$$\forall \epsilon > 0\colon \forall 1 \leq i \leq s\colon \exists Q_i \overset{\text{compact}}{\subseteq} P_i \text{ s.t. } \nu(P_i \setminus Q_i) < \epsilon$$

Define $Q_0 := \left(\bigcup_{i=1}^{s} Q_i\right)^c$ and $P_0 := \emptyset$.

$\mathcal{Q} = \{Q_0, \ldots, Q_s\}$ and $\mathcal{P} = \{P_0, \ldots, P_s\}$ become partitions of X that $\nu(P_i \triangle Q_i) < s\epsilon$. By Lemma (6.2.2), we can find $\delta_{\log 2}$ such that

$$\nu(P_i \triangle Q_i) < \delta_{\log 2} \implies H_\nu(\mathcal{Q} \mid \mathcal{P}) < \log 2$$

By fixing $\epsilon > 0$ in a way that $s\epsilon < \delta_{\log 2}$ we have $H_\nu(\mathcal{Q} \mid \mathcal{P}) < \log 2$.

As we saw in theorem (5.1.2) part (4), for two cover with finite entropy we have:

$$h_\nu(T, \mathcal{P}) \leq h_\nu(T, \mathcal{Q}) + H_\nu(\mathcal{P} \mid \mathcal{Q}) \leq h_\nu(T, \mathcal{Q}) + \log 2$$

Thus,

$$H_\nu(\mathcal{Q}_0^{n-1}) + \int \phi_n \,\mathrm{d}\nu \leq \sum_{Q \in \mathcal{Q}_0^{n-1}} \nu(Q)\Big(-\log \nu(Q) + \sup_{x \in Q} \phi_n(x)\Big)$$

Since $\sum_{Q \in \mathcal{Q}_0^{n-1}} \nu(Q) = 1$, by Lemma(6.2.3) we have:

$$H_\nu(\mathcal{Q}_0^{n-1}) + \int \phi_n \,\mathrm{d}\nu \leq \log\Bigg(\sum_{Q \in \mathcal{Q}_0^{n-1}} \sup_{x \in Q} e^{\phi_n(x)}\Bigg)$$

Now choose a cover α with

$$\operatorname{diam}(\alpha) < \min\{d(Q_i, Q_j) \mid 1 \leq i < j \leq s\}$$

Let $\gamma \overset{\text{cover}}{\subseteq} \alpha_0^{n-1}$.

Since $\operatorname{diam}(\alpha) < \min_{i \neq j \geq 1}\{d(Q_i, Q_j)\}$, and α was an open cover, each member of α intersects with at most closure of two members of $\mathcal{Q}$ -at most with one Q_i for $i \geq 1$

and with $\overline{Q_0}$. Therefore each member of α_0^{n-1} intersects with at most closure of 2^n members of $\mathcal{Q}_0^{n-1}$. For each $Q \in \mathcal{Q}_0^{n-1}$, consider the point $x_Q \in \overline{Q}$ such that

$$\phi_n(x_Q) = \sup_{x \in Q} \phi_n(x)$$

and let $U_Q \in \gamma$ be the open set containing x_Q. So that we have:

$$\sup_{x \in Q} \phi_n(x) \leq \sup_{y \in U_Q} \phi_n(y)$$

In particular for each $U \in \gamma$, there is at most 2^n sets $Q \in \mathcal{Q}_0^{n-1}$ such that $U_Q = U$ Thus:

$$\sum_{Q \in \mathcal{Q}_0^{n-1}} \sup_{x \in Q} e^{\phi_n(x)} \leq 2^n \sum_{U \in \gamma} \sup_{y \in U} e^{\phi_n(y)}$$

All in all:

$$H_\nu(\mathcal{Q}_0^{n-1}) + \int \phi_n \, \mathrm{d}\nu \leq n \log 2 + \log P_n(T, \phi, \alpha)$$
$$h_\nu(T, \mathcal{Q}) + \int \phi \, \mathrm{d}\nu \leq \log 2 + \log P(T, \phi, \alpha)$$

Since the above inequality is true for every continues function $T \colon X \to X$ and every potential $\phi \colon X \to \mathbb{R}$, replace T with T^k and ϕ with ϕ_k. Observe that $\int \phi_k \, \mathrm{d}\nu = k \int \phi \, \mathrm{d}\nu$ and $P(T^k, \phi_k) = kP(f, \phi)$ and $h_\nu(T^k, \mathcal{Q}) = kh_\nu(T, \mathcal{Q})$. So by dividing the sides of inequality by k and taking $k \to \infty$ we reach to the inequality:

$$\forall \nu \in M_T(X) \colon h_\nu(T) + \int \phi \, \mathrm{d}\nu \leq P(T, \phi)$$

- $\forall \epsilon > 0 \colon \exists \mu \in M_T(X) \ s.t. \ h_\mu(T) + \int \phi \, \mathrm{d}\nu \geq S(T, \phi, \epsilon)$

  By definition

$$S_n(T, \phi, \epsilon) = \sup\{\sum_{x \in E} e^{\phi_n(x)} \mid E \subseteq X \text{ be a } (n.\epsilon)\text{-separated set}\}$$

So we can find $E \subseteq X$ in a way that

$$A := \sum_{y \in E} e^{\phi_n(y)} \geq \frac{1}{2} S_n(T, \phi, \epsilon)$$

Let $\delta_x \in M_T(x)$ be a Dirac measure i.e. for every measurable set B,

$$\delta_x(B) := \begin{cases} 1 & x \in B \\ 0 & x \notin B \end{cases}$$

Notice that for a function f, $\int f \, \mathrm{d}\delta_x = f(x)$ now define

$$\nu_n := \frac{1}{A} \sum_{x \in E} e^{\phi_n(x)} \delta_x$$
$$\mu_n := \frac{1}{n} \sum_{i=0}^{n-1} T_*^i \nu_n$$

where $\begin{cases} T_* \colon M_T(x) \to M_T(x) \\ T_*(\nu)(B) := \nu(T^{-1}B), \ \forall B \overset{\text{measurable}}{\subseteq} X \end{cases}$

Let $\{n_j\}_{j \in \mathbb{N}}$ be a subsequence that

$$\lim_{j \to \infty} \frac{1}{n_j} \log S_{n_j}(T, \phi, \epsilon) = S(T, \phi, \epsilon)$$

and Since the space of probability measures is compact in weak* topology, let μ be a measure that

$$\lim_{j \to \infty} \mu_{n_j} = \mu$$

By all the definitions and notations above we end the proof in three steps:

1. μ is T-invariant. So $\mu \in M_T(X)$

By proposition (**??**) we have:

$$\begin{aligned}
\int \varphi \ \mathrm{d}(T_*\mu_n) &= \int \varphi \circ T \ \mathrm{d}\mu_n \\
&= \int \varphi \circ T \ \mathrm{d}\Big(\frac{1}{n}\sum_{i=0}^{n-1} T_*^i \nu_n\Big) \\
&= \frac{1}{n}\sum_{i=0}^{n-1}\int \varphi \circ T \ \mathrm{d}T_*^i\nu_n \\
&= \frac{1}{n}\sum_{i=0}^{n-1}\int \varphi \circ T^{i+1} \ \mathrm{d}\nu_n \\
&= \frac{1}{n}\sum_{i=1}^{n}\int \varphi \circ T^{i} \ \mathrm{d}\nu_n \\
&= \frac{1}{n}\sum_{i=0}^{n-1}\int \varphi \circ T^{i} \ \mathrm{d}\nu_n + \frac{1}{n}\int \varphi \circ T^n \ \mathrm{d}\nu_n - \frac{1}{n}\int \varphi \ \mathrm{d}\nu_n \\
&= \int \varphi \ \mathrm{d}\mu_n + \frac{1}{n}\Big(\int \varphi \circ T^n \ \mathrm{d}\nu_n - \int \varphi \ \mathrm{d}\nu_n\Big) \\
&\Longrightarrow \left|\int \varphi \ \mathrm{d}T_*\mu_n - \int \varphi \ \mathrm{d}\mu_n\right| \leq \frac{2}{n}\sup|\varphi|
\end{aligned}$$

now by restricting $\{n\}$ to $\{n_j\}$, we have:

$$\lim_{j\to\infty}\int \varphi \ \mathrm{d}T_*\mu_{n_j} = \int \varphi \ \mathrm{d}\mu$$

2. Finding the entropy $H_{\nu_n}(\mathcal{P})$ for a partition $\mathcal{P}$ with $\mathrm{diam}(\mathcal{P}) < \epsilon$ and since μ is regular, we can choose $\mathcal{P}$ such that $\mu(\partial\mathcal{P}) = 0$.

   Since E was (n, ϵ)-separated set, each member of $\mathcal{P}_0^{n-1}$ containes at most one member of E and every member of E is contained in one member of $\mathcal{P}_0^{n-1}$. Let $\mathcal{P}(x)$ denotes the element of $\mathcal{P}$ that contain x.

   Thus,

$$
\begin{aligned}
H_{\nu_n}(\mathcal{P}_0^{n-1}) &= \sum_{P\in\mathcal{P}_0^{n-1}} -\nu_n(P)\log\nu_n(P) \\
&= \sum_{x\in E} -\nu_n(\mathcal{P}(x))\log\nu_n(\mathcal{P}(x)) \\
&= \sum_{x\in E} -\left(\frac{1}{A}\sum_{y\in E} e^{\phi_n(y)}\delta_y(\mathcal{P}(x))\right)\left(\log\left(\frac{1}{A}\sum_{y\in E} e^{\phi_n(y)}\delta_y(\mathcal{P}(x))\right)\right) \\
&\left(\delta_y(\mathcal{P}(x)) = \begin{cases} 1 & x=y \\ 0 & x\neq y \end{cases}\right) \\
&= \sum_{x\in E} -\frac{1}{A}e^{\phi_n(x)}\log\left(\frac{1}{A}e^{\phi_n(x)}\right) \\
&= \sum_{x\in E} -\frac{1}{A}e^{\phi_n(x)}\Big(-\log A + \phi_n(x)\Big) \\
&= \sum_{x\in E} \frac{1}{A}\big(\log A\big)e^{\phi_n(x)} - \frac{1}{A}\phi_n(x)e^{\phi_n(x)} \\
&= \log A - \frac{1}{A}\sum_{x\in E}\phi_n(x)e^{\phi_n(x)} \\
&= \log A - \int \phi_n \,\mathrm{d}\nu_n
\end{aligned}
$$

3. Finding the entropy related to μ_i.

   Consider a number $1 \le k \le n-1$. for every $0 \le r \le k-1$, let q_r be the greatest number the $kq_r + r \leq n$, then we have:

$$\begin{aligned}
\mathcal{P}_0^{n-1} &= \bigvee_{i=0}^{n-1} T^{-i}\mathcal{P} = \bigvee_{i=0}^{r-1} T^{-1}\mathcal{P} \vee \bigvee_{i=r}^{kq_r+r-1} T^{-i}\mathcal{P} \vee \bigvee_{i=kq_r+r}^{n-1} T^{-i}\mathcal{P} \\
\bigvee_{i=0}^{r-1} T^{-1}\mathcal{P} &= \mathcal{P}_0^{r-1} \\
\bigvee_{i=r}^{kq_r+r-1} T^{-i}\mathcal{P} &= T^{-r}\Big(\mathcal{P} \vee T^{-1}\mathcal{P} \vee T^{-2}\mathcal{P} \vee \cdots \vee T^{-(kq_r-1)}\mathcal{P}\Big) \\
&= T^{-r}\Big(\mathcal{P} \vee T^{-k}\mathcal{P} \vee T^{-2k}\mathcal{P} \vee \cdots \vee T^{-(q_r-1)k}\mathcal{P}\Big) \\
&\vee T^{-r}\Big(T^{-(k+1)}\mathcal{P} \vee T^{-(2k+1)}\mathcal{P} \vee \cdots \vee T^{-((q_r-1)k+1)}\mathcal{P}\Big) \\
&\vee \\
&\vdots \\
&\vee T^{-r}\Big(T^{-(k+(k-1))}\mathcal{P} \vee T^{-(2k+(k-1))}\mathcal{P} \vee \cdots \vee T^{-((q_r-1)k+(k-1))}\mathcal{P}\Big) \\
&= T^{-r}\left(\bigvee_{j=0}^{q_r-1} T^{-kj}\mathcal{P}\right) \vee T^{-r}\left(\bigvee_{j=0}^{q_r-1} T^{-(kj+1)}\mathcal{P}\right) \\
&\vee \cdots \vee T^{-r}\left(\bigvee_{j=0}^{q_r-1} T^{-(kj+(k-1))}\mathcal{P}\right) \\
&= \bigvee_{i=0}^{k-1} \bigvee_{j=0}^{q_r-1} T^{-(kj+i+r)}\mathcal{P} \\
&= \bigvee_{j=0}^{q_r-1} T^{-(kj+r)}\left(\bigvee_{i=0}^{k-1} T^{-i}\mathcal{P}\right) \\
&= \bigvee_{j=0}^{q_r-1} T^{-(kj+r)}\mathcal{P}_0^{k-1} \\
\bigvee_{i=kq_r+r}^{n-1} T^{-i}\mathcal{P} &= T^{-(kq_r+r)}\left(\bigvee_{i=0}^{n-(kq_r+r)-1} T^{-i}\mathcal{P}\right) = T^{-(kq_r+r)}\mathcal{P}_0^{n-(kq_r+r)-1}
\end{aligned}$$

So by considering all the computations above we reach to the equality:

$$\mathcal{P}_0^{n-1} = \mathcal{P}_0^{r-1} \vee \bigvee_{j=0}^{q_r-1} T^{-(kj+r)}\mathcal{P}_0^{k-1} \vee T^{-(kq_r+r)}\mathcal{P}_0^{n-(kq_r+r)-1}$$

So

$$
\begin{aligned}
H_{\nu_n}(\mathcal{P}_0^{n-1}) &\le H_{\nu_n}(\mathcal{P}_0^{r-1}) + \sum_{j=0}^{q_r-1} H_{\nu_n}(T^{-(kj+r)}\mathcal{P}_0^{k-1}) + H_{\nu_n}(T^{-(kq_r+r)}\mathcal{P}_0^{n-(kq_r+r)-1}) \\
&= H_{\nu_n}(\mathcal{P}_0^{r-1}) + \sum_{j=0}^{q_r-1} H_{T_*^{kj+r}\nu_n}(\mathcal{P}_0^{k-1}) + H_{T_*^{kq_r+r}\nu_n}(\mathcal{P}_0^{n-(kq_r+r)-1}) \\
&\le r\log\#\mathcal{P} + \sum_{j=0}^{q_r-1} H_{T_*^{kj+r}\nu_n}(\mathcal{P}_0^{k-1}) + (n-(kq_r+r))\log\#\mathcal{P} \\
&\le k\log\#\mathcal{P} + \sum_{j=0}^{q_r-1} H_{T_*^{kj+r}\nu_n}(\mathcal{P}_0^{k-1}) + k\log\#\mathcal{P} \\
&= \sum_{j=0}^{q_r-1} H_{T_*^{kj+r}\nu_n}(\mathcal{P}_0^{k-1}) + 2k\log\#\mathcal{P}
\end{aligned}
$$

As the result above is true for every $r \in \{0, 1, \dots, k-1\}$, we have:

$$
\begin{aligned}
\sum_{r=0}^{k-1} H_{\nu_n}(\mathcal{P}_0^{n-1}) &\le \sum_{r=0}^{k-1}\sum_{j=0}^{q_r-1} H_{T_*^{kj+r}\nu_n}(\mathcal{P}_0^{k-1}) + \sum_{r=0}^{k-1} 2k\log\#\mathcal{P} \\
\implies kH_{\nu_n}(\mathcal{P}_0^{n-1}) &\le \sum_{i=0}^{n-1} H_{T_*^{i}\nu_n}(\mathcal{P}_0^{k-1}) + 2k^2\log\#\mathcal{P}
\end{aligned}
$$

Since the function $\phi(x) = -x\log x$ is concave, for each $A \in \mathcal{P}_0^{n-1}$ we have:

$$
\frac{1}{n}\sum_{i=0}^{n-1}\phi\Big(T_*^i\nu_n(A)\Big) \le \phi\left(\frac{1}{n}\sum_{i=0}^{n-1}T_*^i\nu_n(A)\right) = \phi\Big(\mu_n(A)\Big)
$$

so

$$
\begin{aligned}
\frac{1}{n}H_{\nu_n}(\mathcal{P}_0^{n-1}) &\le \frac{1}{nk}\sum_{i=0}^{n-1} H_{T_*^{i}\nu_n}(\mathcal{P}_0^{k-1}) + \frac{1}{n}2k\log\#\mathcal{P} \\
&\le \frac{1}{k}H_{\mu_n}(\mathcal{P}_0^{k-1}) + \frac{2k}{n}\log\#\mathcal{P}
\end{aligned}
$$

Notice that $\frac{1}{n}\int \phi_n \ \mathrm{d}\nu_n = \frac{1}{n}\sum_{i=0}^{n-1}\int \phi\circ T^i \ \mathrm{d}\nu_n = \int \phi \ \mathrm{d}\mu_n$. so,

$$\frac{1}{n}H_{\nu_n}(\mathcal{P}_0^{n-1}) + \frac{1}{n}\int \phi_n \ \mathrm{d}\nu_n \leq \frac{1}{k}H_{\mu_n}(\mathcal{P}_0^{k-1}) + \int \phi \ \mathrm{d}\mu_n + \frac{2k}{n}\log\#\mathcal{P}$$

Let us combine the result with step 2, Since $H_{\nu_n}(\mathcal{P}_0^{n-1}) = \log A - \int \phi_n \ \mathrm{d}\nu_n$, we have:

$$\frac{1}{n}\log A - \frac{1}{n}\int \phi_n \ \mathrm{d}\nu_n + \frac{1}{n}\int \phi_n \ \mathrm{d}\nu_n \leq \frac{1}{k}H_{\mu_n}(\mathcal{P}_0^{k-1}) + \int \phi \ \mathrm{d}\mu_n + \frac{2k}{n}\log\#\mathcal{P}$$

4. All in all,

$$\begin{aligned}\frac{1}{k}H_{\mu_n}(\mathcal{P}_0^{k-1}) + \int \phi \ \mathrm{d}\mu &\geq \frac{1}{n}\log A - \frac{2k}{n}\log\#\mathcal{P} \\ &\geq \frac{1}{n}\log S_n(T,\phi,\epsilon) - \frac{1}{n}\log 2 - \frac{2k}{n}\log\#\mathcal{P}\end{aligned}$$

by restriction $\{n\}$ to $\{n_j\}$ and taking $\lim_{j\to\infty}$ we have:

$$\frac{1}{k}H_{\mu}(\mathcal{P}_0^{k-1}) + \int \phi \ \mathrm{d}\mu \geq S(T,\phi,\epsilon)$$

and by taking $k\to\infty$ we reach to the result:

$$h_\mu(T,\mathcal{P}) + \int \phi \ \mathrm{d}\mu \geq S(T,\phi) = P(T,\phi)$$

□

6.3 Local Variational Principle

The concept of variational principle adapts a local approach. and as we can expect, it has a projection: **Local Variational Principle for pressure**. This concept gets attentions from the time the entropy pairs and their importance has been introduced to dynamical systems.

in this section we are going to prove the Variations Principle for Pressure (**??**). By adopting zero function as our potential we reach to the Variational Principle of Entropy, i.e.

Theorem 6.3.1 (Local Variational Principle of Entropy)**.** *For a given TDS (X,T) and an open cover $\mathcal{U}$:*

$$h_{\text{top}}(T,\mathcal{U}) = \sup\{h_\mu(T,\mathcal{U})\colon \mu \in M_T(X)\} \tag{6.5}$$

The very first proofs of the theorem(6.3.1), are due to Banchard,Glasner and Weiss. Although, in this section our guideline is the article by Wen Huang and Yingfei Yi [11], who generalize the Local Variational Principle for pressure.

For this, we use the extension of the concept of measure entropy of a partition to the measure entropy of an open cover. we'll see some lemmas that will help us proving the theorem.

Definition 6.3.1. *A topological space is zero-dimensional, if it has a base consisting of clopen sets.*

Lemma 6.3.1. *For any $\mathcal{U} = \{U_1, U_2, \ldots, U_n\} \in C_X$ the family of partitions in*

$$\mathcal{U}^* = \left\{\alpha = \{A_1, A_2, \ldots, A_n\} \in \mathcal{P}_X \;\middle|\; \forall i \in \{1,2,\ldots,n\} : A_i \subseteq U_i\right\}$$

consisting of clopen sets is countable.

Lemma 6.3.2. *Let (X,T) be a zero-dimensional TDS, $\mu \in M_T(X)$, $\phi\colon X \to \mathbb{R}$ be a potential, and $\mathcal{U} \in C_X^\circ$. Assume that for some $K \in \mathbb{N}$, $\{\alpha_l\}_{l=1}^K$ is a sequence of finite clopen partitions of X which are finer than U. Then for each $N \in \mathbb{N}$, there exists a finite subset B_N of X such that for all $l \in \{1,2,\ldots,K\}$, each atom of $(\alpha_l)_0^{N-1}$ contains at most one element of B_N and*

$$\sum_{x\in B_N} e^{\phi_n(x)} \geq \frac{P_N(T,\phi,\mathcal{U})}{K}$$

Proposition 6.3.1. *If (X,T) is an invertible zero-dimensional TDS, $\phi\colon X \to \mathbb{R}$ is a potential, and $\mathcal{U} \in C_X^\circ$, then*

$$\exists \mu \in M_T(X) \text{ s.t. } h_\mu^+(T,\mathcal{U}) + \int_X \phi \, \mathrm{d}\mu \geq P(T,\phi,\mathcal{U})$$

Theorem 6.3.2 (Local Variational Principle of pressure). *If (X,T) is a TDS and $M_T(X)$ denotes the set of all T-invariant probability measures on X, then for every potential $\phi\colon X \to \mathbb{R}$ and every open cover $\mathcal{U} \in C_X^\circ$ we have*

$$P(T,\phi,\mathcal{U}) = \sup\{h_\nu^-(T,\mathcal{U}) + \int_X \phi \, \mathrm{d}\nu \colon \nu \in M_T(X)\} \tag{6.6}$$

Proof. • $\forall \mu \in M_T(X)\colon P(T,\phi,\mathcal{U}) \geq h_\nu^-(T,\mathcal{U}) + \int_X \phi \, \mathrm{d}\nu$ Let $\mu \in M_T(X)$. By the corollary (6.1.1),

$$\forall n \in \mathbb{N}\colon \exists \beta \in \mathcal{P}^*(\mathcal{U}_0^{n-1}) \text{ s.t. } P_n(T,\phi,\mathcal{U}) = \sum_{B\in\beta} \sup_{x\in B} e^{\phi_n(x)}$$

so by lemma (6.2.3), we have

$$\begin{aligned}
\log P_n(T,\phi,\mathcal{U}) &= \log\left(\sum_{B\in\beta} \sup_{x\in B} e^{\phi_n(x)}\right) \\
&\geq \sum_{B\in\beta}\left(\mu(B)(\sup_{x\in B}\phi_n(x) - \log\mu(B))\right) \\
&= H_\mu(\beta) + \sum_{B\in\beta}\mu(B)\sup_{x\in B}\phi_n(x) \\
&\geq H_\mu(\beta) + \sum_{B\in\beta}\int_B \phi_n \, \mathrm{d}\mu \\
&= H_\mu(\beta) + \int_X \phi_n \, \mathrm{d}\mu \\
&= H_\mu(\beta) + n\int_X \phi \, \mathrm{d}\mu \qquad \text{by corollary (4.0.2) since T is } \mu\text{-invariant} \\
&\geq H_\mu(\mathcal{U}) + n\int_X \phi \, \mathrm{d}\mu
\end{aligned}$$

now by dividing both sides by n and taking the limit $\lim_{n\to\infty}$ we have

$$\lim_{n\to\infty}\frac{1}{n}\log P_n(T,\phi,\mathcal{U}) \geq \lim_{n\to\infty}\frac{1}{n}H_\mu(\mathcal{U}) + \int_X \phi \ \mathrm{d}\mu$$
$$\implies P(T,\phi,\mathcal{U}) \geq h_\mu^-(T,\mathcal{U}) + \int_X \phi \ \mathrm{d}\mu$$

- $\exists \mu \in M_T(X)$ s.t. $P(T,\phi,\mathcal{U}) \leq h_\mu^-(T,\mathcal{U}) + \int_X \phi \ \mathrm{d}\mu$

  We prove the above inequality in three steps:

  1. The inequality is true if our system is invertible and zero-dimensional.

     Let $\mathcal{U} = \{U_1, U_2, \ldots, U_n\} \in C_X^\circ$. By lemma(6.3.1), the collection of clopen sets in

     $$\mathcal{U}^* = \left\{\alpha = \{A_1, A_2, \ldots, A_n\} \in \mathcal{P}_X \;\middle|\; \forall i \in \{1,2,\ldots,n\} : A_i \subseteq U_i\right\}$$

     is countable. Therefore let $\{\alpha_l\}_{l\in\mathbb{N}}$ be an enumeration of this collection.

     By definition, $h_\mu^+(T,\mathcal{U}) = \inf_{\beta\succeq\mathcal{U}} h_\mu(T,\beta)$ and by theorem (5.1.2) part (3), it's sufficient to take the inf over n-elements partitions finer than $\mathcal{U}$.

     Thus for any $k \in \mathbb{N}$ and $\mu \in M_T(X)$, we have

     $$h_\mu^+\left(T^k, \bigvee_{i=0}^{k-1} T^{-i}\mathcal{U}\right) = \inf_{s_k\in\mathbb{N}^k} h_\mu\left(T^k, \bigvee_{i=0}^{k-1} T^{-i}\alpha_{s_k(i)}\right)$$

     Let

     $$M(k,s_k) = \left\{\mu \in M_T(X) \qquad : \frac{1}{k}\left(h_\mu\left(T^k, \bigvee_{i=0}^{k-1} T^{-i}\alpha_{s_k(i)}\right) + \int_X \phi_k \ \mathrm{d}\mu\right) \geq \frac{1}{k}P_k(T^k, \phi_k, \mathcal{U}_0^{k-1})\right\}$$

     (a) Let us prove that $M(k,s_k) \neq \emptyset$:

       i. Note that by proposition (6.3.1),

       $$\exists \mu_k \in M_{T^k}(X) \text{ s.t. } h_{\mu_k}^+(T^k,\mathcal{U}_0^{k-1}) + \int_X \phi_k \ \mathrm{d}\mu_k \geq P(T^k, \phi_k, \mathcal{U}_0^{k-1})$$

And since $\bigvee_{i=0}^{k-1} T^{-i}\alpha_{s_k(i)} \succeq \mathcal{U}_0^{k-1}$ we have

$$h_{\mu_k}(T^k, \bigvee_{i=0}^{k-1} T^{-i}\alpha_{s_k(i)}) + \int_X \phi_k \, \mathrm{d}\mu_k \geq P(T^k, \phi_k, \mathcal{U}_0^{k-1})$$

ii. Define $\nu_k := \frac{\mu_k + T_*\mu_k + \cdots + T_*^{k-1}\mu_k}{k}$

– $\forall i \in \{0, 1, \ldots, k-1\}\colon T_*^i\mu_k \in M_{T^k}(X)$ since $\mu_k \in M_{T^k}(X)$ and

$$\forall A \overset{\text{measurable}}{\subseteq} X\colon T_*^i\mu_k(T^{-k}A) = \mu_k(T^{-(i+k)}A) = \mu_k(T^{-i}A) = T_*^i\mu_k(A)$$

– $\nu_k \in M_T(X)$ since $\forall i \in \{0, 1, \ldots, k-1\}\colon T_*^i\mu_k \in M_{T^k}(X)$ and

$$\begin{aligned}
\forall A \overset{\text{measurable}}{\subseteq} X\colon \nu_k(T^{-1}A) &= \frac{\mu_k(T^{-1}A) + T_*\mu_k(T^{-1}A) + \cdots + T_*^{k-1}\mu_k(T^{-1}A)}{k} \\
&= \frac{\mu_k(T^{-1}A) + \cdots + \mu_k(T^{-(k-1)}A) + \mu_k(T^{-k}A)}{k} \\
&= \frac{\mu_k(T^{-1}A) + \cdots + \mu_k(T^{-(k-1)}A) + \mu_k(A)}{k} \\
&= \frac{\mu_k(A) + \mu_k(T^{-1}A) + \cdots + \mu_k(T^{-(k-1)}A)}{k} \\
&= \frac{\mu_k(A) + T_*\mu_k(A) + \cdots + T_*^{k-1}\mu_k(A)}{k} \\
&= \nu_k(A)
\end{aligned}$$

iii. Let for each $s_k \in \mathbb{N}^k$, $\begin{cases} \sigma\colon \mathbb{N}^k \to \mathbb{N}^k \\ \sigma\Big((s_1, \ldots, s_{k-1}, s_k)\Big) = (s_k, s_1, \ldots, s_{k-1}) \end{cases}$ be cyclic shift map.

for each $j \in \{0, 1, \ldots, k-1\}$ we have

– $h_{T_*^j\mu_k}\Big(T^k, \bigvee_{i=0}^{k-1} T^{-i}\alpha_{s_k(i)}\Big) = h_{\mu_k}\Big(T^k, \bigvee_{i=0}^{k-1} T^{-i}\alpha_{\sigma^j s_k(i)}\Big)$ Since for every $n \in \mathbb{N}$ By theorem (5.1.1), part (13)

$$\begin{aligned}
H_{T^j_*\mu_k}\left(\left(\bigvee_{i=0}^{k-1} T^{-i}\alpha_{s_k(i)}\right)_0^{n-1}\right) &= H_{\mu_k}\left(T^{-j}\left(\bigvee_{i=0}^{k-1} T^{-i}\alpha_{s_k(i)}\right)_0^{n-1}\right)\\
&= H_{\mu_k}\left(T^{-j}\left(\bigvee_{r=0}^{n-1} T^{-rk}\bigvee_{i=0}^{k-1} T^{-i}\alpha_{s_k(i)}\right)\right)\\
&= H_{\mu_k}\left(\bigvee_{r=0}^{n-1} T^{-rk}\bigvee_{i=0}^{k-1} T^{-(i+j)}\alpha_{s_k(i)}\right)\\
&= H_{\mu_k}\left(\bigvee_{r=0}^{n-1} T^{-rk}\bigvee_{i=j}^{j+k-1} T^{-i}\alpha_{s_k(i-j)}\right)\\
&= H_{\mu_k}\left(\bigvee_{r=0}^{n-1} T^{-rk}\bigvee_{i=j}^{j+k-1} T^{-i}\alpha_{\sigma^j s_k(i)}\right)\\
&= H_{\mu_k}\left(T^{-j}\left(\bigvee_{r=0}^{n-1} T^{-rk}\bigvee_{i=0}^{k-1} T^{-i}\alpha_{\sigma^j s_k(i)}\right)\right)\\
&= H_{\mu_k}\left(T^{-j}\left(\bigvee_{i=0}^{k-1} T^{-i}\alpha_{\sigma^j s_k(i)}\right)_0^{n-1}\right)
\end{aligned}$$

$$\begin{aligned}
\implies h_{T^j_*\mu_k}\left(T^k, \bigvee_{i=0}^{k-1} T^{-i}\alpha_{s_k(i)}\right) &= \lim_{n\to\infty}\frac{1}{n}H_{T^j_*\mu_k}\left(\left(\bigvee_{i=0}^{k-1} T^{-i}\alpha_{s_k(i)}\right)_0^{n-1}\right)\\
&= \lim_{n\to\infty}\frac{1}{n}H_{\mu_k}\left(T^{-j}\left(\bigvee_{i=0}^{k-1} T^{-i}\alpha_{\sigma^j s_k(i)}\right)_0^{n-1}\right)\\
&= h_{\mu_k}\left(T^k, \bigvee_{i=0}^{k-1} T^{-i}\alpha_{\sigma^j s_k(i)}\right)
\end{aligned}$$

– $\forall k \in \mathbb{N}$: $\int_X \phi_k \, \mathrm{d}T^j_*\mu_k = \int_X \phi_k \, \mathrm{d}\mu_k$ since

$$\int_X \phi_k \, \mathrm{d}T^j_*\mu_k = \int_X \phi_k \circ T^j \, \mathrm{d}\mu_k$$

Thus,

$$
\begin{aligned}
&h_{T_*^j\mu_k}(T^k, \bigvee_{i=0}^{k-1} T^{-i}\alpha_{s_k(i)}) + \int_X \phi_k \ \mathrm{d}T_*^j\mu_k \\
&\qquad = h_{\mu_k}\left(T^k, \bigvee_{i=0}^{k-1} T^{-i}\alpha_{\sigma^j s_k(i)}\right) + \int_X \phi_k \ \mathrm{d}\mu_k \\
&\qquad \geq P(T^k, \phi_k, \mathcal{U}_0^{k-1})
\end{aligned}
$$

iv. $\forall s_k \in \mathbb{N}^k$:

$$
\begin{aligned}
&h_{\nu_k}(T^k, \bigvee_{i=0}^{k-1} T^{-i}\alpha_{s_k(i)}) + \int_X \phi_k \ \mathrm{d}\nu_k \\
&\qquad = \frac{1}{k}\sum_{i=0}^{k-1}\left(h_{T_*^j\mu_k}(T^k, \bigvee_{i=0}^{k-1} T^{-i}\alpha_{s_k(i)}) + \int_X \phi_k \ \mathrm{d}T_*^j\mu_k\right) \\
&\qquad \geq P(T^k, \phi_k, \mathcal{U}_0^{k-1})
\end{aligned}
$$

Ergo

$$\nu_k \in \bigcap_{s_k \in \mathbb{N}^k} M(k, s_k)$$

Let $M(k) := \bigcap_{s_k \in \mathbb{N}^k} M(k, s_k) \neq \emptyset$.

(b) The next step is to show that $M := \bigcap_{k\in\mathbb{N}} M(k) \neq \emptyset$.

(c) Consider $\nu \in \bigcap_{k\in\mathbb{N}} M(k)$ and $k \in \mathbb{N}$. we have

$$
\begin{aligned}
&\frac{1}{k}h_\nu^+(T^k, \mathcal{U}_0^{k-1}) + \int_X \phi \ \mathrm{d}\nu \\
&\qquad = \frac{1}{k}\left(h_\nu^+(T^k, \mathcal{U}_0^{k-1}) + \int_X \phi_k \ \mathrm{d}\nu\right) \\
&\qquad = \inf_{s_k \in \mathbb{N}^k} \frac{1}{k}\left(h_\nu(T^k, \bigvee_{i=0}^{k-1} T^{-i}\alpha_{s_k(i)})\right) \\
&\qquad \geq P(T, \phi, \mathcal{U})
\end{aligned}
$$

By letting $k \to \infty$ we have the desired result:

$$P(T,\phi,\mathcal{U}) \leq h_\nu^-(T,\mathcal{U}) + \int_X \phi \, \mathrm{d}\nu$$

2. If our TDS is invertible, we know that by corollary (2.3.2), each system is a factor of an invertible zero-dimensional system. Let us call that system (Z, R) and the factor map, $\pi\colon Z \to X$; so by step 1,

$$\exists \mu_Z \in M_S(Z) \text{ s.t. } P(R, \phi \circ \pi, \pi^{-1}\mathcal{U}) \leq h_{\mu_Z}^-(R, \pi^{-1}\mathcal{U}) + \int_Z \phi \circ \pi \, \mathrm{d}\mu_Z$$

   Define $\mu = \pi_* \mu_Z$. Then

   – $\mu \in M_T(X)$.

   – By proposition (4.0.6), $\int_Z \phi \circ \pi \, \mathrm{d}\mu_Z = \int_X \phi \, \mathrm{d}\mu$

   – By theorem (5.2.3), $h_{\mu_Z}^-(R, \pi^{-1}\mathcal{U}) = h_\mu^-(T,\mathcal{U})$.

   – By proposition (6.1.5), $P(R, \phi \circ \pi, \pi^{-1}\mathcal{U}) = P(T,\phi,\mathcal{U})$.

   Thus,

$$\exists \mu \in M_T(X) \text{ s.t. } P(T,\phi,\mathcal{U}) \leq h_\mu^-(T,\mathcal{U}) + \int_X \phi \, \mathrm{d}\mu$$

3. In general case that our system does not need to be invertible, we remind the definition (2.3.5)

   let $d\colon X \to \mathbb{R}$ be our distance function on X. Define

$$\tilde{X} := \{(x_1, x_2, \dots)\colon T(x_{i+1}) = x_i, x_i \in X, i \in \mathbb{N}\}$$

   and consider the metric $\begin{cases} d_T : \tilde{X} \to \mathbb{R} \\ d_T\Big((x_1, x_2, \dots), (y_1, y_2, \dots)\Big) = \sum_{i=1}^{\infty} \frac{d(x_i,y_i)}{2^i} \end{cases}$ and let $\begin{cases} \sigma_T : \tilde{X} \to \tilde{X} \\ \sigma_T\Big((x_1, x_2, \dots)\Big) = (T(x_1), x_1, x_2, \dots) \end{cases}$ be shift homeomorphism.

It's easy to check that $\begin{cases} \pi : (\tilde{X}, \sigma_T) \to (X, T) \\ \pi\Big((x_1, x_2, \dots)\Big) = x_1 \end{cases}$ is a homomorphism between dynamical system and we call $(\tilde{X}, \sigma_T)$ the natural extension of dynamical system (X, T).

Now give any TDS (X, T), call it's natural extension $(\tilde{X}, \sigma_T)$. Since $(\tilde{X}, \sigma_T)$ is invertible, by the previous steps,

$$\exists \nu \in M_{\sigma_T}(\tilde{X}) \text{ s.t. } h_\nu(\sigma_T, \pi^{-1}\mathcal{U}) + \int_{\tilde{X}} \phi \circ \pi \, \mathrm{d}\nu = P(\sigma_T, \phi \circ \pi, \pi^{-1}\mathcal{U})$$

Let $\mu := \pi_* \nu$ Due to previous lemmas and propositions,

$$\begin{aligned} h_\mu(T, \mathcal{U}) + \int_X \phi \, \mathrm{d}\mu &= h_\nu(\sigma_T, \pi^{-1}\mathcal{U}) + \int_{\tilde{X}} \phi \circ \pi \, \mathrm{d}\nu \\ &= P(\sigma_T, \phi \circ \pi, \pi^{-1}\mathcal{U}) \\ &= P(T, \phi, \mathcal{U}) \end{aligned}$$

□

Corollary 6.3.1 (the topological variational principle of pressure (6.2.2))**.**

Proof. By proposition (5.2.4), it's sufficient to take $\sup_{\mathcal{U} \in C_X^\circ}$ from the equality (6.6). □

Chapter 7

Entropy Pairs

In this chapter, will learn about entropy pair that find out useful in specifying zero entropy systems.In this chapter we'll see either a upe TDS disjoint from all minimal systems with zero entropy or not. As a tool for treating this question Blanchard introduced the notion of topological entropy pairs.

7.1 Topological Entropy Pairs

Definition 7.1.1. *Let (X,T) be a TDS. A pair $(x_1,x_2) \in X \times X$ is an entropy pair if $x_1 \neq x_2$ and for every disjoint closed neighborhoods U_1 and U_2 of x_1 and x_2, respectively,*

$$h_{\text{top}}(T, \{U_1^c, U_2^c\}) > 0$$

We denote the set of all entropy pairs if (X,T) by $E(X,T)$ and call the cover $\{U_1^c, U_2^c\}$ a cover distinguishing x_1, x_2 i.e. an open cover $\{U,V\}$ that $x_1 \in (U^c)^\circ, x_2 \in (V^c)^\circ$.

Remark 7.1.1. *The definition above is equivalent to:*

A pair $(x_1,x_2) \in X \times X$ is an entropy pair if $x_1 \neq x_2$ and for every disjoint neighborhoods U_1 and U_2 of x_1 and x_2, respectively,

$$h_{\text{top}}(T, \{\overline{U_1}^c, \overline{U_2}^c\}) > 0$$

and equivalently,

Non-diagonal element (x, x'), is said to be an entropy pair if for any standard cover $\{U, V\}$ of X distinguishing x and x' one has $h_{\text{top}}(T, \{U, V\}) > 0$

Example 7.1.1. *If (X, T), as a TDS, has U.P.E. then every non-diagonal element of $X \times X$, i.e. $(x, y) \in X \times X, x \neq y$, are entropy pairs.*

Lemma 7.1.1. *For a TDS (X, T), If there's a finite standard open cover with positive entropy, then there exist a two-element standard open cover with positive entropy.*

Proof. Let assume there's a finite open cover $\alpha = \{U_1, U_2, \ldots, U_n\}$ such that $h_{\text{top}}(T, \alpha) > 0$.

Since X is metric space, let for every $i \in \{1, 2, \ldots, n\}$ and every $x \in U_i$, $B_i(x)$ be a ball that

$$x \in B_i(x) \subseteq \overline{B_i(x)} \subseteq U_i$$

Since X is compact and $\{B_i(x) \colon 1 \leq i \leq n, x \in X\}$ form a cover of X, there are finite elements of the set mentioned that form a cover for X. Let call it

$$\{B_i(x_j) \colon i \in I, j \in J\}$$

where $I = \{1, 2, \ldots, n\}, \#J < \infty$.

Let consider

$$F_{ij} := \overline{B_i(x_j)} \text{ and } V_i := \bigcap_{j \in J} F_{ij}^c$$

For F_{ij} we have:

- $F_{ij} \subseteq U_i$
- F_ij is closed

For V_i we have:

- $U_i^c \subseteq V_i$ so $U_i \cup V_i = X$
- V_i is open
- $\bigcap_{i\in I} V_i = \bigcap_{\substack{i\in I\\ j\in J}} F_{ij}^c = \bigcap_{\substack{i\in I\\ j\in J}} \left(\overline{B_i(x_j)}\right)^c = \left(\bigcup_{\substack{i\in I\\ j\in J}} \overline{B_i(x_j)}\right)^c = X \setminus X = \emptyset$

Consider open cover $\alpha_i := \{U_i, V_i\}$ and then the open cover $\bigvee_{i\in I} \alpha_i$. Since $\bigcap_{i\in I} V_i = \emptyset$, every element of $\bigvee_{i\in I} \alpha_i$ is in at least one element of α; Therefore

$$\alpha \preceq \bigvee_{i\in I} \alpha_i$$

Thus,

$$0 < h_{\text{top}}(T, \alpha) \le h_{\text{top}}\left(T, \bigvee_{i\in I} \alpha_i\right) \le \sum_{i\in I} h(T, \alpha_i)$$

so there must be $i \in I$ such that $h(T, \alpha_i) > 0$. Let consider $\alpha_i = \{U, V\}$.

If one of them is dense, e.g. U, then for a arbitrary element of V, consider a ball $B(x)$ such that

$$x \in B(x) \subseteq \overline{B(x)} \subseteq V$$

Let $F := \overline{B(x)}$ and $U' := U \setminus F$. U' is open. If V is dense, by the above process find V'. now consider the open cover $\beta := \{U', V'\}$.

Since $\beta \succeq \alpha_i$ we have $h_{\text{top}}(T, \beta) > 0$ and now none of the elements are dense. and our two-element open cover with positive entropy will be β. □

Lemma 7.1.2. *Let consider a TDS (X, T) and an standard open cover $\beta = \{U, V\}$ with positive entropy. Then there is an open cover $\beta_1 = \{U_1, V_1\}$ that*

- *$\beta_1 \preceq \beta$.*
- $\text{diam}(U_1^c) \le \frac{1}{2}\text{diam}(U^c)$
- $\text{diam}(V_1^c) \le \frac{1}{2}\text{diam}(V^c)$
- $h_{\text{top}}(T, \beta_1) > 0$

Proof. The proof has two steps:

1. First let T be a homeomorphism.

   Then we wish to show that $\#U^c \geq 2$:

   If not, assuming that $U^c = \{x\}$, by definition.

$$h_{\text{top}}(T, \beta) = \lim_{n\to\infty} \frac{1}{n} \log \mathcal{N}(\beta_0^{n-1}) = \lim_{n\to\infty} \frac{1}{n} \log \mathcal{N}(\bigvee_{i=0}^{n-1} T^{-i}\beta)$$

   Note that

$$U \cap T^{-1}U \cap \cdots \cap T^{-(n-1)} \in \bigvee_{i=0}^{n-1} T^{-i}\beta$$

   and since T is homeomorphism,

$$U \cap T^{-1}U \cap \cdots \cap T^{-(n-1)}U = X \setminus \{x, T^{-1}x, \ldots, T^{-(n-1)}x\}$$

   So taking $U \cap T^{-1}U \cap \cdots \cap T^{-(n-1)}U$ from $\bigvee_{i=0}^{n-1} T^{-i}\beta$ with at most n other members each containing one of $T^i x, 0 \leq i \leq n-1$, give us a subcover of $\bigvee_{i=0}^{n-1} T^{-i}\beta$; Thus $\mathcal{N}(\beta_0^{n-1}) \leq n+1$. Consequently,

$$h_{\text{top}}(T, \beta) = \lim_{n\to\infty} \frac{1}{n} \log \mathcal{N}(\beta_0^{n-1}) = 0$$

   This contradiction shows that $\#U^c \geq 2$.

   Now, let d be the distance function on X and $y, y' \in U^c$ such that $d(y, y') > 0$. Choose $\epsilon > 0$ such that $0 < \epsilon < \frac{1}{2}d(y, y')$.

   Let $B_\epsilon(x)$ denote the ball of radius ϵ and centre x. Consider

$$\beta'_\epsilon := \{B_\epsilon(x) \colon x \in U^c\}$$

   as a cover of the compact set U^c. Let

$$\beta_\epsilon := \{B_\epsilon(x_i) \colon x_i \in U^c, i \in \{1, 2, \ldots, m\}\}$$

   be a finite subcover of β'_ϵ. Notice that each subcover of β'_ϵ has at least 2 member since $\epsilon < \frac{1}{2}d(y, y')$.

Define for each $i \in \{1, 2, \dots, m\}$

$$F_i' := \overline{B_\epsilon(x_i)} \text{ and } F_i := U^c \cap F_i'$$

So that

$$U^c = \bigcup_{i=1}^{m} F_i \text{ and } U = \bigcap_{i=1}^{m} F_i^c$$

Thus for all $i \in \{1, 2, \dots, m\}$, $\{F_i^c, V\}$ is cover of X and for each $i \in \{1, 2, \dots, m\}$

$$\{F_i^c, V\} \preceq \beta = \{U, V\} \preceq \bigvee_{i=1}^{m} \{F_i^c, V\}$$

Therefore we have

$$0 < h_{\text{top}}(T, \beta) \leq h_{\text{top}}(T, \bigvee_{i=1}^{m_1} \{F_i^c, V\}) \leq \sum_{i=1}^{m} h_{\text{top}}(T, \{F_i^c, V\})$$

So there must be a $j \in \{1, 2, \dots, m\}$ such that $h_{\text{top}}(T, \{F_j^c, V\}) > 0$. let $U_1 := F_j^c$. By the same process for V, we find a V_1. Let $\beta_1 := \{U_1, V_1\}$

- $\beta_1 \preceq \beta$
- $\text{diam}(U_1^c) \leq \frac{1}{2}\text{diam}(U^c)$
- $\text{diam}(V_1^c) \leq \frac{1}{2}\text{diam}(V^c)$
- $h_{\text{top}}(T, \beta_1) > 0$

2. In general case Let $(\tilde{X}, \sigma_T)$ be the natural extension of (X, T) with the factor map $\pi\colon \tilde{X} \to X$, just like the definition (2.3.5).

   Since $\pi^{-1}\beta$ is an open cover of $\tilde{X}$ by step 1, we have that there exists an open cover $\beta' = \{U', V'\}$ with the properties

   - $\beta' \preceq \pi^{-1}\beta = \{U \times \prod_{i\in\mathbb{N}} X, V \times \prod_{i\in\mathbb{N}} X\}$

     Thus,

     $$\beta' = \{A \times \prod_{i\in\mathbb{N}} X, B \times \prod_{i\in\mathbb{N}} X\}$$

     for $A, B \subseteq X$ such that $A \cup B = X$.

- $\operatorname{diam}(U'^c) \leq \frac{1}{2}\operatorname{diam}(\pi^{-1}U^c)$. Thus,

$$\operatorname{diam}(A^c \times \prod_{i\in\mathbb{N}} X) \leq \frac{1}{2}\operatorname{diam}(U^c \times \prod_{i\in\mathbb{N}} X)$$

  and therefore:

$$\operatorname{diam}(A^c) \leq \frac{1}{2}\operatorname{diam}(U^c)$$

- $\operatorname{diam}(V'^c) \leq \frac{1}{2}\operatorname{diam}(\pi^{-1}V^c)$ and therefore:

$$\operatorname{diam}(B^c) \leq \frac{1}{2}\operatorname{diam}(V^c)$$

- $h_{\text{top}}(\sigma_T, \beta') > 0$

Thus by letting $\beta_1 := \pi\beta' = \{\pi U', \pi V'\} = \{A, B\}$, we have:

- β_1 is an open cover since π is open map.
- $\beta_1 \preceq \beta$ since $\beta' \preceq \pi^{-1}\beta \implies \pi\beta' \preceq \beta$.
- $\operatorname{diam}((\pi U')^c) = \operatorname{diam}(A^c) \leq \frac{1}{2}\operatorname{diam}(U^c)$.
- $\operatorname{diam}((\pi V')^c) = \operatorname{diam}(B^c) \leq \frac{1}{2}\operatorname{diam}(V^c)$.
- $h_{\text{top}}(T, \beta_1) = h_{\text{top}}(T, \pi\beta') = h_{\text{top}}(\sigma_T, \pi^{-1}\pi\beta') = h_{\text{top}}(\sigma_T, \beta') > 0$, by (8), (2), and the fact that $\beta' = \{A \times \prod_{i\in\mathbb{N}} X, B \times \prod_{i\in\mathbb{N}} X\}$.

□

Proposition 7.1.1. *In a TDS (X, T), for any standard cover $\beta = \{U, V\}$ with positive entropy, there exists $x \in U^c, x' \in V^c$ that $(x, x') \in E(X, T)$.*

Proof. Lemma (7.1.2) gives us a partition $\beta_1 := \{U_1, V_1\}$ such that

- $\beta_1 \preceq \beta$.
- $\operatorname{diam}(U_1^c) \leq \frac{1}{2}\operatorname{diam}(U^c)$
- $\operatorname{diam}(V_1^c) \leq \frac{1}{2}\operatorname{diam}(V^c)$

again, by lemma (7.1.2), since $\beta_1 = \{U_1, V_1\}$ has positive entropy, we could find an open cover β_2 coarser than β_1 with the properties mentioned. Thus, inductively, we find two decreasing sequences of closed sets $\{U_i^c\}_{i\in\mathbb{N}}$ and $\{V_i^c\}_{i\in\mathbb{N}}$. Therefore,

$$\bigcap_{i\in\mathbb{N}} U_i^c \neq \emptyset \text{ and } \bigcap_{i\in\mathbb{N}} V_i^c \neq \emptyset$$

So

$$\exists x \in \bigcap_{i\in\mathbb{N}} U_i^c \text{ and } \exists x' \in \bigcap_{i\in\mathbb{N}} V_i^c$$

Also since $\mathrm{diam}(U_{i+1}^c) \leq \frac{1}{2}\mathrm{diam}(U_i^c)$ and $\mathrm{diam}(V_{i+1}^c) \leq \frac{1}{2}\mathrm{diam}(V_i^c)$ we have that

$$\lim_{i\to\infty} \mathrm{diam}(U_i^c) = \lim_{i\to\infty} \mathrm{diam}(V_i^c) = 0$$

So each of the intersections mentioned, just has one element. Thus

$$\{x\} = \bigcap_{i\in\mathbb{N}} U_i^c \text{ and } \{x'\} = \bigcap_{i\in\mathbb{N}} V_i^c$$

And since $x \in U_i^c$, $x' \in V_i^c$, and $U_i \cup V_i = X$ so that $U_i^c \cap V_i^c = \emptyset$, we have that $x \neq x'$.
The last step is to show that $(x, x') \in E(X, T)$:
For each closed neighbourhoods G and G' of x and x', respectively - so that $x \in G, x' \in G'$ - we shall prove that $h_{\mathrm{top}}(T, \{G^c, G'^c\}) > 0$.
Since X is a metric space, we have

$$\exists\epsilon > 0 \text{ s.t. } x \in \overline{B_\epsilon(x)} \subseteq G \text{ and } x' \in \overline{B_\epsilon(x')} \subseteq G'$$

Let $\epsilon_i := \max\{\mathrm{diam}(U_i^c), \mathrm{diam}(V_i^c)\}$ and $I := \{i \in \mathbb{N}\colon \epsilon_i < \frac{1}{2}\epsilon\}$ -Notice that since $\lim_{i\to\infty} \epsilon_i = 0$, $I \neq \emptyset$ - so that for every $i \in I$

$$U_i^c \subseteq \overline{B_\epsilon(x)} \subseteq G \text{ and } V_i^c \subseteq \overline{B_\epsilon(x')} \subseteq G'$$

Therefore

$$G^c \subseteq U_i \text{ and } G'^c \subseteq V_i$$

So

$$\{G^c, G'^c\} \succeq \{U_i, V_i\}$$

Thus,

$$0 < h_{\text{top}}(T, \{U_i, V_i\}) \leq h_{\text{top}}(T, \{G, G'\})$$

Ergo, $(x, x') \in E(X, T)$. □

Corollary 7.1.1. *For a TDS (X, T) and a standard cover $\beta = \{U, V\}$, we have:*

$$\exists x \in U^c, x' \in V^c \text{ s.t. } (x, x') \in E(X, T) \Longleftrightarrow h_{\text{top}}(T, \beta) > 0$$

Proof. if $h_{\text{top}}(T, \beta) > 0$ by proposition (7.1.1) the result is obtained.
Conversely, by definition β is a cover distinguishing (x, x') and $(x, x') \in E(X, T)$, Thus $h_{\text{top}}(T, \beta) > 0$ □

Theorem 7.1.1. *Let (X, T) be a TDS.*

$$E(X, T) \neq \emptyset \Longleftrightarrow h_{\text{top}}(T) > 0$$

Proof. Let assume that $\sup_{\alpha \in C_X^\circ} h_{\text{top}}(T, \alpha) = h_{\text{top}}(T) > 0$. So there exist a finite open cover α such that $h_{\text{top}}(T, \alpha) > 0$. By lemma (7.1.1), let $\beta = \{U, V\}$ be a standard open cover with positive entropy.
By proposition (7.1.1), there exists $(x, x') \in E(X, T)$ so that

$$E(X, T) \neq \emptyset$$

Conversely, if $E(X, T) \neq \emptyset$, assuming $(x, x') \in E(X, T)$,

$$h_{\text{top}}(T) = \sup_{\alpha \in C_X^\circ} h_{\text{top}}(T, \alpha) > 0$$

Since we have that for every open cover distinguishing x, x' it's entropy is positive. $\square$

Theorem 7.1.2. *Consider* $\pi \colon (X,T) \to (Y,S)$ *as a factor map between two TDSs.*

1. $(x,x') \in E(X,T), \pi(x) \neq \pi(x') \implies \big(\pi(x), \pi(x')\big) \in E(Y,S)$
2. $(y,y') \in E(Y,S) \implies \exists (x,x') \in E(X,T)$ s.t. $\pi(x) = y, \pi(x') = y'$

Proof. 1. Consider an open cover $\beta := \{U, V\}$ distinguishing $\pi(x), \pi(x')$. Since π is continues, $\pi^{-1}\beta = \{\pi^{-1}U, \pi^{-1}V\}$ is an open cover distinguishing (x,x').

Since $(x,x') \in E(X,T)$ we have

$$h_{\text{top}}(T, \pi^{-1}\beta)$$

by the fact that π is a factor map, by proposition (3.2.1), we have

$$0 < h_{\text{top}}(T, \pi^{-1}\beta) = h_{\text{top}}(S, \beta)$$

and by the assumption that $\pi(x) \neq \pi(x')$ we have $\big(\pi(x), \pi(x')\big) \in E(Y,S)$.

Thus, by letting $\Delta := \{(x_1, x_2) \in X \times X \colon \pi(x_1) = \pi(x_2)\}$ we have the diagrams below:

$$\begin{tikzcd} X \arrow[r, "T"] \arrow[d, "\pi"] \arrow[dr, dashed] \arrow[dr, dashed, bend left] \arrow[dr, phantom, "="] & X \arrow[d, "\pi"] \\ Y \arrow[r, "S"] & Y \end{tikzcd} \qquad \begin{tikzcd} (x,x') \arrow[d, "\pi(x)\neq\pi(x')"] & \in & E(X,T)\setminus\Delta \arrow[d, "\pi"] \\ \big(\pi(x), \pi(x')\big) & \in & E(Y,S) \end{tikzcd}$$

2. If $(y,y') \in E(X,T)$, let $\alpha := \{\{y\}^c, \{y'\}^c\}$ be the standard open cover distinguishing y, y'. By hypothesis,

$$0 < h_{\text{top}}(S, \beta)$$

By proposition (3.2.1), since π is a factor map, we have

$$0 < h_{\text{top}}(S, \beta) = h_{\text{top}}(T, \pi^{-1}\beta)$$

and since $\pi^{-1}\beta$ is an open cover of X, by proposition (7.1.1), we have

$$\exists (x,x') \in E(X,T) \text{ s.t. } x \in \left(\pi^{-1}\{y\}^c\right)^c, x' \in \left(\pi^{-1}\{y'\}^c\right)^c$$

Since $\left(\pi^{-1}\{y\}^c\right)^c = \pi^{-1}(y)$ and $\left(\pi^{-1}\{y'\}^c\right)^c = \pi^{-1}(y')$ we have

$$\exists (x,x') \in E(X,T) \text{ s.t. } x \in \pi^{-1}(y), x' \in \pi^{-1}(y')$$

Thus, we have the diagrams below:

$$\begin{tikzcd} X \arrow[r, "T"] \arrow[d, "\pi"'] \arrow[dr, dashed] \arrow[dr, dashed, bend left] \arrow[dr, phantom, "="] & X \arrow[d, "\pi"] \\ Y \arrow[r, "S"] & Y \end{tikzcd} \qquad \begin{tikzcd} \exists (x,x') & \in & E(X,T) \cap (\pi^{-1}y \times \pi^{-1}y') \\ (y,y') \arrow[u, "\pi^{-1}"] & \in & E(Y,S) \arrow[u, "\pi^{-1}"] \end{tikzcd}$$

$\square$

Proposition 7.1.2. *Consider a TDS (X,T). If $(W, T{\restriction}_W)$ be a subsystem of (X,T), then $E(W,T) \subseteq E(X,T)$*

Proof. Let $(x,x') \in E(W, T{\restriction}_W)$. Take two closed set $U, V \subseteq X$ with empty intersection that $x \in U, x' \in V$. We should show that $h_{\text{top}}(T, \{U^c, V^c\}) > 0$.
Since $U_W := U \cap W$ and $V_W := V \cap W$ are closed subset of W without intersection, and $x \in U_W, x' \in V_W$, and also $(x,x') \in E(W, T{\restriction}_W)$ we have:

$$h_{\text{top}}(T{\restriction}_W, \{W \setminus U_W, W \setminus V_W\}) = h_{\text{top}}(T{\restriction}_W, \{W \setminus U, W \setminus V\}) > 0$$

Since for each $n \in \mathbb{N}$ and for any subcover $\mathcal{R}$ of $\{X \setminus U, X \setminus V\}_0^{n-1}$, covers $\mathcal{R}{\restriction}_W := \{R \cap W \colon R \in \mathcal{R}\}$ is a subcover of $\{W \setminus U, W \setminus V\}_0^{n-1}$ we have:

$$\mathcal{N}(\{W \setminus U, W \setminus V\}_0^{n-1}) \leq \mathcal{N}(\{X \setminus U, X \setminus V\}_0^{n-1})$$

So that we have

$$0 < h_{\text{top}}(T\restriction_W, \{W \setminus U, W \setminus V\}) \leq h_{\text{top}}(T, \{X \setminus U, X \setminus V\})$$

□

Proposition 7.1.3. *Let* (X,T) *be a TDS.*
(X,T) *has U.P.E. iff* $E(X,T) \cup \Delta = X \times X$. *Where* $\Delta := \{(x,x)\colon x \in X\}$.

Proof. If (X,T) has U.P.E. then each non-diagonal elements of $X \times X$ is entropy pair. So $E(X,T) \cup \Delta = X \times X$.
Conversely, if $E(X,T) \cup \Delta = X \times X$, each non-diagonal element is an entropy pair so considering any non-trivial open cover $\beta = \{U, V\}$, Since $U^c \cap V^c = \emptyset$, let $(x,x') \in U^c \times V^c$ be our non-diagonal element so that by assumption, $(x,x') \in E(X,T)$. Thus,

$$h_{\text{top}}(T, \beta) > 0$$

Since β is standard cover distinguishing x, x'. □

Proposition 7.1.4. *Considering* (X,T) *as a TDS,* $E(X,T) \cup \Delta \subseteq X^2$ *is a closed subset of* X^2 *and it's* $T \times T$*-invariant.*
Also, $E(X,T) \cup \Delta = X$ *iff* (X,T) *has U.P.E. .*

Proof.
- First let show that it's closed.
  Consider any element $(x,x') \in \overline{E(X,T) \cup \Delta}$.
  if $x = x'$, then $(x,x') \in \Delta$. if not, there's a sequence $\{(x_i, x_i')\}_{i\in\mathbb{N}}$ in $E(X,T)$ such that $\lim_{i\to\infty} x_i = x$ and $\lim_{i\to\infty} x_i' = x'$.
  Now let $U, V \subseteq X$ be any closed subset of X with empty intersection, that $x \in U$ and $x' \in V$.

Since $\lim_{i\to\infty}(x_i, x_i') = (x, x')$ we have:

$$\exists N \in \mathbb{N} \text{ s.t. } \forall n > N\colon x_n \in U, x_n' \in V$$

For any $n > N$, $(x_n, x_n') \in E(X,T)$ so we have

$$h_{\text{top}}(T, \{U^c, V^c\}) > 0$$

Thus, $(x, x') \in E(X,T)$.

- $E(X,T) \cup \Delta$ is $T \times T$-invariant, since

  For any $(x, x') \in E(X,T) \cup \Delta$, if $Tx = Tx'$, then $T \times T(x, x') \in \Delta$. if not, for any two closed subset $U, V \subseteq X$ with empty intersection that $Tx \in U, Tx' \in V$, $T^{-1}U, T^{-1}V$ are closet subset of X, containing x and x' respectively. Thus, by proposition (3.2.2),

$$0 < h_{\text{top}}(T, \{T^{-1}U^c, T^{-1}V^c\}) = h_{\text{top}}(T, T^{-1}\{U^c, V^c\}) = h_{\text{top}}(T, \{U^c, V^c\})$$

  Therefore, $(Tx, Ty) \in E(X,T)$.

  All in all, $E(X,T) \cup \Delta$ is $T \times T$-invariant and closed subset of X^2.

$\square$

7.2 Measure Theoretical Entropy Pairs

Definition 7.2.1. *Consider the MDS $(X, \mathscr{B}, \mu, T)$ where $\mathscr{B}$ denote the Borel σ-algebra and a partition $P_A := \{A, A^c\}$. P_A is called replete, if A is measurable, $A^\circ \neq \emptyset$, and $(A^c)^\circ \neq \emptyset$.*

Proposition 7.2.1. *Let $(X, \mathscr{B}, \mu, T)$ be a MDS and $P_A := \{A, A^c\}$ be a partition. Then, P_A is replete iff there is a standard cover $\mathcal{U}$ of X such that $P_A \succeq \mathcal{U}$.*

Proof. If P_A is replete, since $A^\circ \neq \emptyset$ we have $\overline{A} \neq X$ ($\overline{A^c} = (A^\circ)^c$). Likewise, $\overline{A^c} \neq X$. That means A, A^c are not dense subsets.

For a fixed $\epsilon > 0$, let $B_\epsilon(x)$ be open ball with radius ϵ and center x. Define

$$B_1 := \bigcup_{x \in A} B_\epsilon(x) \text{ and } B_2 := \bigcup_{x \in A^c} B_\epsilon(x)$$

Notice that $A \subseteq B_1$ and $A^c \subseteq B_2$; So that $\{B_1, B_2\}$ is an open cover of X which is coarser than P_A.

If B_1 is dense, since $(A^c)^\circ \neq \emptyset$, Let

$$U := B_1 \setminus (A^c)^\circ$$

and if B_2 is dense, let

$$V := B_2 \setminus A^\circ$$

Considering $\mathcal{U} := \{U, V\}$, $\mathcal{U}$ is a standard open cover coarser than P_A.

Conversely, if $\mathcal{U} = \{U, V\}$ is a standard open cover coarser than $P_A = \{A, A^c\}$ -means $A \subseteq U$ and $A^c \subseteq V$- we have that $\overline{U}, \overline{V} \neq X$ so that $(U^c)^\circ, (V^c)^\circ \neq \emptyset$.

Since $U^c \subseteq A^c$, it implies $(U^c)^\circ \subseteq (A^c)^\circ$. Thus, $(A^c)^\circ \neq \emptyset$; By the same logic, $A^\circ \neq \emptyset$. □

Definition 7.2.2. *For MDS $(X, \mathscr{B}, \mu, T)$, the replete partition $P_A = \{A, A^c\}$ distinguish $x, x' \in X$, if $x \in A^\circ$ and $x' \in (A^c)^\circ$.*

Definition 7.2.3. *Let $(X, \mathscr{B}, \mu, T)$ be a MDS and $(x, x') \in X \times X$ be a non-diagonal pair. (x, x') is called a μ-entropy pair if for every replete partition P_A distinguishing x, x', we have*

$$h_\mu(T, \{A, A^c\}) > 0$$

$E_\mu(X, T)$ is represented the set of all μ-entropy pairs of X.

Proposition 7.2.2. *Let μ be T-invariant measure on the space (X, T) and $\{A, A^c\}$ be a partition of X such that $h_\mu(T, \{A, A^c\}) > 0$. Then, there exist $(x, x') \in E_\mu(X, T)$ such* *that* $x \in A^\circ$ *and* $x' \in (A^c)^\circ$.

Corollary 7.2.1. *Considering MDS* $(X, \mathscr{B}, \mu, T)$, *we have*

$$h_\mu(X,T) = 0 \iff E_\mu(X,T) = \emptyset$$

Proposition 7.2.3. $E_\mu(X,T) \cup \Delta$ *is a closed* $T \times T$*-invariant set.*

Lemma 7.2.1. *Let* $(X, \mathscr{B}, \mu, T)$ *be a MDS and* $\mathcal{U} = \{U, V\}$ *be a measurable cover of* X *such that for any partition* $P_A = \{A, A^c\}$ *finer than* $\mathcal{U}$, $h_\mu(T, P_A) > 0$. *Then,* $h_{\mathrm{top}}(T, \mathcal{U}) > 0$.

Theorem 7.2.1. $E_\mu(X,T) \subseteq E(X,T)$.

Proof. Let $(x, x') \in E_\mu(X,T)$. For any closed neighbourhoods U, V of x, x' respectively, let $\mathcal{U} := \{U, V\}$. For every partition $P := \{A, A^c\}$ that $\mathcal{U} \preceq P$, we have $h_\mu(T, P)$. Otherwise, if there is $P = \{A, A^c\}$ that $h_\mu(T,P) = 0$ and $\mathcal{U} \preceq P$, then x, x', both, must belong to A or must belong to A^c. since $A \subseteq U^c$, if $x, x' \in A$ then $x, x' \in U^c$, but we knew that $x \in U^\circ$. This contradiction shows that every replete partition finer than $\mathcal{U}$, have positive μ-entropy. Thus by lemma above, $h_{\mathrm{top}}(T, \mathcal{U}) > 0$ and $(x, x') \in E(X,T)$. $\square$

Lemma 7.2.2. *Consider TDS* (X, T) *and a closed set* $K \subseteq X$.

$$\exists \mu \in M_T^e(X) \ s.t. \ h_\mu(T) > 0, \mu(K) > 0$$

iff

$$\exists \text{ an open cover of } K, \text{ named } \mathcal{U} \ s.t. \ h_{\mathrm{top}}(\mathcal{U}, K^c) > 0$$

Theorem 7.2.2. *If* (X, T) *is a TDS, then there is a measure* $\mu \in M_T(X)$ *such that* $E(X,T) = E_\mu(X,T)$.

Bibliography

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

[2] F Blanchard and Y Lacroix. Zero entropy factors of topological flows. *Proceedings of the American mathematical society*, 119(3):985–992, 1993.

[3] Stephen Boyd, Stephen P Boyd, and Lieven Vandenberghe. *Convex optimization*. Cambridge university press, 2004.

[4] Manfred Einsiedler and Thomas Ward. *Ergodic theory*. Springer London, 2013.

[5] Nobuoki Eshima et al. *Statistical data analysis and entropy*. Springer, 2020.

[6] Harry Furstenberg. Disjointness in ergodic theory, minimal sets, and a problem in diophantine approximation. *Mathematical systems theory*, 1(1):1–49, 1967.

[7] Eli Glasner. *Ergodic theory via joinings*. Number 101. American Mathematical Soc., 2003.

[8] Eli Glasner and Xiangdong Ye. Local entropy theory. *Ergodic Theory and Dynamical Systems*, 29(2):321–356, 2009.

[9] Tim NT Goodman. Relating topological entropy and measure entropy. *Bulletin of the London Mathematical Society*, 3(2):176–180, 1971.

[10] L Wayne Goodwyn. Topological entropy bounds measure-theoretic entropy. *Proceedings of the American Mathematical Society*, 23(3):679–688, 1969.

[11] Wen Huang and Yingfei Yi. A local variational principle of pressure and its applications to equilibrium states. *Israel Journal of Mathematics*, 161(1):29–74, 2007.

[12] Michal Misiurewicz. Topological conditional entropy. *Studia Mathematica*, 55(2):175–200, 1976.

[13] Karl E Petersen. Disjointness and weak mixing of minimal sets. *Proceedings of the American Mathematical Society*, 24(2):278–280, 1970.

[14] Pierre-Paul Romagnoli. A local variational principle for the topological entropy. *Ergodic Theory and Dynamical Systems*, 23(5):1601–1610, 2003.

[15] Marcelo Viana and Krerley Oliveira. *Foundations of ergodic theory*. Number 151. Cambridge University Press, 2016.

[16] Peter Walters. A variational principle for the pressure of continuous transformations. *American Journal of Mathematics*, 97(4):937–971, 1975.

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