Universidade Federal de Itajubá
Programa de Pós-Graduação em Matemática

Um Estudo Sobre a Entropia Topológica de Sistemas Dinâmicos

Jefferson Fernando Zambrano Sánchez

Orientador: Prof. Fernando Pereira Micena

Durante o desenvolvimento deste trabalho o autor recebeu auxílio financeiro da CAPES

Itajubá, 15 de dezembro de 2023

UNIVERSIDADE FEDERAL DE ITAJUBÁ
PROGRAMA DE PÓS-GRADUAÇÃO EM MATEMÁTICA

Um Estudo Sobre a Entropia Topológica de Sistemas Dinâmicos

Jefferson Ferando Zambrano Sánchez

Orientador: Prof. Fernando Pereira Micena

Dissertação submetida ao Programa de Pós-Graduação em Matemática como parte dos requisitos para obtenção do Título de Mestre em Ciências em Matemática

Areá de Concentração:

Sistemas Dinâmicos

ITAJUBÁ - MG
15 DE DEZEMBRO DE 2023

Dedico este trabalho aos meus pais, Marco Tulio Zambrano e Cruz Dalia Sánchez

Agradecimentos

Agradeço a Deus pela oportunidade que outros não tem. A meus pais, Marcos e Cruz Dalia, minha vida é por vocês.

*A meu orientador Fernando Pereira Micena pela amizade, conselhos, disposição, paciência que foi muita e ensinamento **é um verdadeiro maestro**.*

A minha namorada Jenifer Barajas que sempre me deu o apoio necessário e incondicional, também a meu companheiro William Clavijo que me guiou em tudo e brindo toda a ajuda necessária em meus estudos.

Ao professor Fabio Scalco Dias quem no primeiro semestre não me permitiu desistir do mestrado.

Finalmente, Agradeço ao pais de Brasil que ajuda a melhorar a vida das pessoas, e que oferece a oportunidade de cumprir sonhos.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES).

"Não importa se você vai devagar, o importante é sempre seguir para frente"

Fernando Pereira Micena.

Resumo

Neste trabalho daremos uma prova detalhada do Principio Variacional que estabelece que a entropia topológica de uma aplicação contínua definida num espaço métrico compacto é igual ao supremo das entropias de medidas invariantes. Também estabeleceremos resultados da entropia topológica, métrica e condicional de maneira mais explicita. Todo o anterior foi feito tendo como referencia [11]. Ademais mostraremos um exemplo sobre o Principio Variacional o qual mostra que a entropia dos conjunto de pontos não errantes e o mesmo que a entropia do espaço, ou seja a entropia de um conjunto esta carregada sobre o conjunto de pontos errantes.

Palavras–clave: Entropia Métrica, Entropia Topológica, Principio Variacional, Entropia Condicional.

Abstract

In this paper we will give a detailed proof of the Variational Principle which states that the topological entropy of a continuous application defined in a compact metric space is equal to the supremum of the entropies of invariant measures. We will also establish topological, metric and conditional entropies in a more explicit way. All of was done with reference to [11]. We will also show an example of the Variational Principle which shows that the entropy of the set of non-errant points is the same as the entropy of the set of non-errant points is the same as the entropy of space, that is, the entropy of a set is loaded on the set of wandering points.

Keywords: Metric Entropy, Topological Entropy, Variational Principle, Conditional Entropy.

Sumário

Introdução

O pioneiro e fundador da termodinâmica o alemão Rudolf Clausius (1822-1888), inventou a palavra **entropia** em 1865. Em física, entropia é uma magnitude que mede a parte da energia que não pode ser usada para realizar um trabalho e que em consequência se perde. O termo entropia é utilizado em diversas disciplinas do conhecimento , tais como a física, a química, a matemática, a astrofísica, a linguística, a computação, para fazer referência a **medida de "desordem" que tem um sistema**.

Em 1948 em [4] se estabelece a primeira noção de entropia por parte de Claude E. Shannon (1916-2001), que foi matemático, engeniero e criptógrafo estado-unidense, ele define a entropia de probabilidade de um vector $P = (p_1, ..., p_n)$ de valores não negativos cuja soma é 1 e que pode descrever-se como:

$$H(P) = -\sum_{i=1}^{n} p_i \log p_i$$

Com respeito a entropia topológica, definimos um conjunto $X \neq \emptyset$ e uma função continua $f : X \to X$, para cada $x \in X$ se define sua orbita $\{f^i(x)\}_{i=0}^{\infty}$. Se toma-se dois pontos $x, y \in X$, então se pode definir uma métrica assim:

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

Esta nova distância mede a distância das orbitas de x e y.

Neste contexto a entropia topológica definida como:

$$h_{top}(f) = \lim_{\varepsilon \to 0} \lim_{n \to \infty} \sup \frac{1}{n} \log N_d(f, \varepsilon, n),$$

Onde $N_d(f, \varepsilon, n)$ é o máximo cardinal de um conjunto (n, ε)-separado contido em X, onde X é um espaço métrico compacto, é o invariante mais importante, pois representa a taxa de crescimento exponencial do número de segmentos de orbita.

Por outro lado, na década de 60, com a contribuição conjunta entre Goodwyn[8], Dinaburg [7], Goodman[9], mostraram resultados muito importantes relacionado a entropia, este resultado é conhecido como **Principio Variacional**. Faremos brevemente uma apresentação do principio variacional, considere o espaço métrico compacto (X, d), é possível associar uma álgebra de Borel isto é, a σ-álgebra gerada por todas a bolas abertas $B_d(x, \varepsilon)$ com $x \in X$ e $\varepsilon > 0$, considere também $\mathcal{M}(X)$ o espaço de todas as medidas de probabilidade sobre esta σ-álgebra e $T : X \to X$ uma transformação contínua mensurável.Definamos a entropia de Kolmogorov-Sinai (que é a antropia com medida ou entropia métrica) ou seja

$$h_\mu(T) = \sup_{\mathcal{P}:finita} h_\mu(T, \mathcal{P})$$

onde $\mathcal{P}$ é uma partição mensurável do espaço e o principio variacional nos diz que:

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

onde $h_{top}(T)$ é uma outra entropia, que não usa medida, apenas a topologia do espaço e a dinâmica de T, conhecida como entropia topológica.

Capítulo 1

Preliminares

Nosse primeiro capítulo, daremos uma breve introdução dos conceitos que serão de muita importância para entender os demais capítulos que aqui apresentaremos, os temas que trataremos são: espaços métricos, topologia e teoria da medida. Muitos dos conceitos foram estudados de [12], [13] e [14].

1.1 Espaços Métricos

Definição 1.1.1. *Seja X um conjunto. Uma métrica sobre X é uma função $d : X \times X \to \mathbb{R}$, tal que:*

1. $d(x, y) \geq 0$,
2. $d(x, y) = 0$ *só, e somente se,* $x = y$,
3. $d(x, y) = d(y, x)$ *e*
4. $d(x, y) \leq d(x, z) + d(z, y)$ *(desigualdade triangular),*

para todo $x, y, z \in X$.

Se d é uma métrica sobre X dizemos que (X, d) é um espaço métrico. A função d é conhecida como **função distância** e para cada x, y de X, se diz $d(x, y)$ é a distância entre x e y.

Exemplos:

- Se X é um conjunto qualquer, a função $d : X \times X \to \mathbb{R}$ definida por:

$$d(x,y) = \begin{cases} 0 & : se \quad x = y \\ 1 & : se \quad x \neq y \end{cases}$$

é chamada métrica discreta.

- O plano $\mathbb{R}^2$, junto com a função d_1 definida por:

$$d_1((x_1, y_1), (x_2, y_2)) = |x_1 - x_2| + |y_1 - y_2|.$$

é chamada a métrica da soma.

Definição 1.1.2. *Seja (X, d) um espaço métrico, $x \in X$ e $\varepsilon > 0$. O conjunto $B(x, \varepsilon) = \{y \in X : d(x, y) < \varepsilon\}$ chama-se a **bola aberta** centrada em x com raio ε. O conjunto $B[x, \varepsilon] = \{y \in X : d(x, y) \leq \varepsilon\}$ chama-se a **bola fechada** centrada em x com raio ε.*

Definição 1.1.3. *Seja (X, d) um espaço métrico. Se A é um subconjunto não vazio de X então o diâmetro de A, $diam(A)$ é $\sup\{d(x, y) : x, y \in A\}$. Se $diam(A) < \infty$ então dizemos que o conjunto A e limitado. Por convenção, o $diam(\emptyset) = 0$.*

Observação: Da definição se tem que o espaço métrico X é limitado se a função d é limitada.

Definição 1.1.4. *Seja (X, d) um espaço métrico. Se A é um conjunto não vazio de X então a distância $d(x, A)$ de um ponto $x \in X$ ao conjunto A é definido como o ínfimo dos números $d(x, a)$ onde $a \in A$; isto é,*

$$d(x, A) = inf\{d(x, a) : a \in A\}.$$

Definição 1.1.5. *Sejam (X, d) e (Y, d') espaços métricos. Uma função $f : X \to Y$ é contínua em um ponto $x_0 \in X$ se dado $\varepsilon > 0, \exists \delta > 0$ tal que se $d(x, x_0) < \delta$ então $d'(f(x), f(x_0)) < \varepsilon$. Se a função f é contínua em cada ponto de X, dizemos simplesmente que f é contínua.*

Proposição 1.1.1. *Sejam (X, d) um espaço métrico e $A \subset X, A \neq \emptyset$. Então, a função de X nos números reais definida por $x \longmapsto d(x, A)$ é contínua.*

Prova: Seja $x_0 \in X$ e $\varepsilon > 0$. Se $d(x, x_0) < \varepsilon$, então para cada $a \in A$, tem-se que:

$$\begin{aligned} d(x, A) &\leq d(x, a) \\ &\leq d(x, x_0) + d(x_0, a) \\ &< \varepsilon + d(x_0, a). \end{aligned}$$

Então,$d(x, A) - \varepsilon < d(x_0, a), \forall\ a \in A$. Então $d(x, A) - \varepsilon \leq d(x_0, A)$, obtendo $d(x, A) - d(x_0, A) \leq \varepsilon$. Usando o mesmo raciocínio podemos obter

$$d(x_0, A) - d(x, A) \leq \varepsilon.$$

Concluindo, $|d(x, A) - d(x_0, A)| \leq \varepsilon$, mostrando o desejado.

$\square$

1.2 Topologia

Definição 1.2.1. *Seja X um conjunto. Uma coleção τ de subconjuntos de X tal que*

1. *$\emptyset$ e X pertencem a τ;*
2. *Se $A_1, ..., A_n \in \tau$, então $A_1 \cap ... \cap A_n \in \tau$;*
3. *Para qualquer família arbitraria de abertos $\{A_\lambda\}_{\lambda \in L}$ com $A_\lambda \in \tau$ para cada $\lambda \in L$, temos $\displaystyle\bigcup_{\lambda \in L} A_\lambda \in \tau$.*

é uma topologia sobre X.

Um espaço topológico é um par (X, τ), onde X é um conjunto e τ uma topologia de X. Observemos que todo espaço métrico (X, d) pode ser considerado um espaço topológico, onde τ é formada pelos subconjuntos abertos de X.

Definição 1.2.2. *Sejam X e Y dois espaços topológicos. Uma função $f : X \to Y$ é contínua se $f^{-1}(A)$ é um conjunto aberto em X para cada conjunto aberto A de Y.*

Definição 1.2.3. *Sejam X e Y dois espaços topológicos. Uma função $f : X \to Y$ é contínua em um ponto $x \in X$ se para cada aberto V de $f(x)$ em Y existe um aberto U de x em X tal que $f(U) \subset V$.*

Definição 1.2.4. *Uma coleção A de subconjuntos do espaço X se diz que* ***cobre*** *X,ou é uma cobertura de X, se a união dos elementos de A coincide com X.Dizemos que A é uma* ***cobertura aberta*** *de X se é uma cobertura de X formado por conjuntos abertos de X.*

Definição 1.2.5. *Dizemos que um espaço X é compacto se de cada cobertura aberta A de X podemos extrair uma subcoleção finita que também cobre X.*

Definição 1.2.6. *Seja X um espaço topológico . Um subconjunto F de X é um conjunto fechado em X se seu complementar F^c é um conjunto aberto.*

Teorema 1.2.1. *Seja X um espaço métrico e $K \subseteq X$ compacto. Toda sequência $(x_n)_{n\in\mathbb{N}} \subseteq K$ tem uma sub-sequência $(x_{n_k})_k \in \mathbb{N}$ que converge para algum $x_0 \in K$.*

Prova: Seja $(x_n)_{n\in\mathbb{N}}$ uma sequência em K. Suponha que para cada $x \in K, \exists\ \varepsilon_x > 0$ para o qual existe $m_x \in \mathbb{N}$ tal que $x_n \notin B(x, \varepsilon_x), \forall\ n \geq m_x$.

Note que a família de conjuntos $\{B(x, \varepsilon_x) : x \in K, \varepsilon_x > 0\}$ formam uma cobertura aberta de K, como K é compacto existem $x_1, x_2, ..., x_k$ tais que

$$K \subset B(x_1, \varepsilon_{x_1}) \cup B(x_2, \varepsilon_{x_2}) \cup ... \cup B(x_k, \varepsilon_{x_k}),$$

então, $x_n \notin K$, para $n \geq \max\{m_{x_1}, m_{x_2}, ..., m_{x_k}\}$ o qual é uma contradição.

Portanto, $\exists\ x_0 \in K$ tal que $\forall\ \varepsilon > 0$, existe $x_k \in B(x_0, \varepsilon)$.

Agora para cada $j \in \mathbb{N}$, existe $x_{n_j} \in B(x_0, 1/j)$ tal que $n_j > n_{j-1}$ a sequência x_{n_j} é uma sub-sequência de x_n, note que $d(x_{n_j}, x_0) < 1/j$, para todo x_{n_j} é uma sub-sequência que converge para x_0.

$\square$

Teorema 1.2.2. *Sejam X um espaço topológico e $K \subset X$ compacto, se $F \subset K$ fechado, então F é compacto.*

Prova: Seja $\{U_\alpha\}_{\alpha \in I}$ uma cobertura aberta de F, como F é fechado, então F^c é aberto em X. Suponha que existe $x \in K \setminus F$, então $x \in F^c$, isto prova que

$$K \subset F^c \cup \bigcup_{\alpha \in I} U_\alpha$$

Assim, temos que $F^c \cup \bigcup_{\alpha \in I} U_\alpha$ é uma cobertura aberta para K, pela compacidade de K existem $U_1, ..., U_n$ tais que $K \subset F^c \cup \bigcup_{i=1}^{n} U_i$, como $F \subset K$, então

$$F \subset F^c \cup \bigcup_{\alpha \in I} U_\alpha$$

como $F \not\subset F^c$, então $F \subset \bigcup_{i=1}^{n} U_i$. Assim, para toda cobertura aberta de F, existe uma sub-cobertura aberta finita tal que F está contido. Portanto, F é compacto.

$\square$

1.3 Teoria da Medida

Definição 1.3.1. *Uma coleção Σ de subconjuntos de um conjunto X é uma **σ-álgebra** em X se;*

- $X, \emptyset \in \Sigma$
- *Se $A \in \Sigma$ então $A^c \in \Sigma$.*
- *Se $A_j \in \Sigma$, $j = 1, 2, 3, ...,$ então $\left(\bigcup_{j=1}^{\infty} A_j\right) \in \Sigma$.*

Definição 1.3.2. *Se Σ é uma σ-álgebra em X, então X é chamado **espaço mensurável** e os elementos de Σ são chamados **conjuntos mesuráveis** de X.*

Definição 1.3.3. *Sejam (X, B, μ) e (Y, C, ρ) espaços de probabilidade. Dizemos que uma aplicação $T : X \longrightarrow Y$ é uma transformação mensurável que preserva a medida de probabilidade μ em X se $\mu(T^{-1}(C)) = \rho(C)$.*

Definição 1.3.4. *A σ-álgebra de Borel $\mathcal{B}$ ou boreliana de um espaço topológico é a σ-álgebra gerada pela topologia, isto é, a menor σ-álgebra que contém todos os subconjuntos abertos de X.*

Seja (X, Σ) um espaço mensurável. Uma medida em (X, Σ) é uma função $\mu : \Sigma \to \mathbb{R}^+$ tal que $\mu(\emptyset) = 0$ e

$$\mu\left(\bigcup_{j=1}^{\infty} A_j\right) = \sum_{j=1}^{\infty} \mu(A_j)$$

para quaisquer conjunto $A_j \in \Sigma$ disjuntos dois a dois. A segunda propriedade é chamada de σ-aditividade.

A tripla (X, Σ, μ) é chamado de espaço de medida. Além disso quando temos $\mu(X) < \infty$, dizemos que μ é finita e, se $\mu(X) = 1$ então dizemos que μ é uma medida de probabilidade e (X, Σ, μ) é um espaço de probabilidade.

Proposição 1.3.1. *Se E é um conjunto mensurável e se*

$$\chi_E(x) = \begin{cases} 1; & se \quad x \in E \\ \\ 0; & se \quad x \notin E \end{cases}$$

então χ_E é uma função mensurável.

Prova: Seja A um aberto de $\mathbb{R}$, então pode ocorrer:

1. Se $0 \notin A$ e $1 \notin A$, então $\chi_E^{-1}(A) = \emptyset$ que é mensurável em X.

2. Se $0 \in A$ e $1 \notin A$, então $\chi_E^{-1}(A) = E^c$ que é mensurável em X.

3. Se $0 \notin A$ e $1 \in A$, então $\chi_E^{-1}(A) = E$ que é mensurável em X.

4. Se $0 \in A$ e $1 \in A$, então $\chi_E^{-1}(A) = X$ que é mensurável.

Portanto χ_E é una função mensurável.

$\square$

Definição 1.3.5. *Uma função $s : (X, \Sigma) \to \mathbb{R}$ é uma **função simples** se assume um número finito de valores, ou seja, existem $E_1, .., E_n$ com $E_1 \cup ... \cup E_n = X$ e $a_1, ..., a_n$ em $\mathbb{R}$ de modo que*

$$s(x) = \sum_{j=1}^{n} a_j \chi_{E_j}(x).$$

Proposição 1.3.2. *O limite pontual (supondo que exista) de uma sequência f_n de funções mensuráveis é mensurável.*

Teorema 1.3.1 (Teorema de Convergência Monótona). *Seja $0 \leq f_1 \leq f_2 \leq f_3 \leq ... \leq \infty$ uma sequência crescente de funções mensuráveis em (X, Σ, μ) e suponha $f_n(x) \to f(x); n \to \infty, \forall\ x \in X$. Então $f = \lim\limits_{n\to\infty} f_n$ é mensurável e $\displaystyle\int_X f d\mu = \lim_{n\to\infty} \int_X f_n d\mu$.*

Prova: Temos pela proposição (1.3.2) que f é mensurável. Como $f_n \leq f_{n+1} \leq ... \leq f$ temos que:

$$\int_X f_n d\mu \leq \int_X f_{n+1} d\mu \leq ... \leq \int_X f d\mu, \forall n.$$

Então, $\alpha := \lim \displaystyle\int_X f_n d\mu$ existe e

$$\alpha \leq \int_X f d\mu. \tag{1.1}$$

Seja $0 \leq s \leq f$ uma função simples mensurável, para $0 < c < 1$ considere

$$E_n := \{x; f_n(x) \geq cs(x)\}.$$

Note que $E_n \subset E_{n+1} \subset ...$ e $\displaystyle\bigcup_{n=1}^{\infty} E_n = X$ ou seja, dado $x \in X$, se $f(x) = 0$, então $s(x) = 0$, assim $x \in E_1$. Se $f(x) > 0$ então $cs(x) < f(x)$ visto que $c < 1$, portanto, $x \in E_n$ para algum n.

Assim, $\displaystyle\int_X f_n d\mu \geq \int_{E_n} f_n d\mu \geq \int_{E_n} cs d\mu = c \int_{E_n} s d\mu$, então,

$$\lim_{n\to\infty} \int_X f_n d\mu \geq c \lim_{n\to\infty} \int_{E_n} s d\mu \geq c \lim_{n\to\infty} \varphi_s(E_n) = c\ \varphi_s(X) = c \int_X s d\mu.$$

Pois φ_s é uma medida.

Tomando $\sup\limits_{0\leq s\leq f}$ na desigualdade de acima, obtemos,

$$\lim_{n\to\infty}\int_X f_n d\mu \geq c\int_X f d\mu.$$

Como $0 < c < 1$ é arbitrário, logo

$$\alpha = \lim_{n\to\infty}\int_X f_n d\mu \geq \int_X f d\mu. \tag{1.2}$$

Logo de (1.1) e (1.2) temos que:

$$\int_X f d\mu = \lim_{n\to\infty}\int_X f_n d\mu.$$

$\square$

Lema 1.3.1 (Lema de Fatou). *Se $f_n : (X, \Sigma, \mu) \to [0, \infty]$ mensuráveis, então,*

$$\lim_{n\to\infty}\inf\int_X f_n d\mu \geq \int_X (\lim_{n\to\infty}\inf f_n) d\mu.$$

Teorema 1.3.2 (Convergência dominada). *Seja $f_n : (X, \Sigma, \mu) \to \mathbb{C}$ uma sequência de funções mensuráveis e $g : (X, \Sigma, \mu) \to [0, \infty]$ em $L^1_\mu(X)$ de forma que $|f_n(x)| \leq g(x); n = 1, 2, ..., x \in X$ e $f(x) = \lim_{n\to\infty} f_n(x)$ existe $\forall\ x \in X$. Então,*

1. $f \in L^1_X(X)$
2. $\lim_{n\to\infty}\int_X |f_n - f| d\mu = 0$
3. $\lim_{n\to\infty}\int_X f_n d\mu = \int_X f d\mu$

Prova:

1. Pela hipótese temos que $|f_n| \leq g$ então $|f| \leq g$, logo $\int_X |f| d\mu \leq \int_X g d\mu < \infty$. Assim, $f \in L^1_\mu(X)$.

2. Como temos que $|f_n| \leq g\ \forall n$, então $|f| \leq g$, logo,

$$|f_n - f| \leq |f_n| + |f| \leq g + g = 2g; \forall\ n \text{ então,}$$

$2g - |f_n - f| \geq 0$, pelo Lema de Fatou (1.3.1),

$$\begin{aligned}
\int_X 2g d\mu &= \int_X \lim_{n\to\infty} \inf(2g - |f_n - f|) d\mu \\
&\leq \lim_{n\to\infty} \inf \int_X (2g - |f_n - f|) d\mu \\
&= \int_X 2g d\mu - \lim_{n\to\infty} sup \int_X |f_n - f| d\mu
\end{aligned}$$

Portanto,

$$0 \geq \lim_{n\to\infty} \sup \int_X |f_n - f| d\mu$$

Tomaremos $W_n = \displaystyle\int_X |f_n - f| d\mu$. Então, $\lim \displaystyle\int_X |f_n - f| d\mu = 0$. Pois se uma sequência de números reais não negativa neste caso W_n não converge para 0, então o $\displaystyle\lim_{n\to\infty} \sup W_n > 0$.

3. $|\displaystyle\int_X f_n d\mu - \int_X f d\mu| = |\int_X (f_n - f) d\mu| \leq \int_X |f_n - f| d\mu \to 0$ quando $n \to \infty$, pelo item anterior.

□

Definição 1.3.6 (Medidas Invariantes). *Sejam (X, Σ, μ) um espaço de medida e $f : X \to X$ uma transformação mensurável. Dizemos que a medida μ e f-invariante se*

$$\mu(E) = \mu(f^{-1}(E)).$$

Para todo conjunto mensurável $E \subset X$. Também dizemos que f preserva μ.

Teorema 1.3.3. *Seja (X, d) um espaço métrico compacto e $f : X \to X$ contínua. Então f admite ao menos uma medida de probabilidade boreliana $f-$invariante.*

Definição 1.3.7. *Um ponto $x \in X$ é **não-errante** se para toda vizinhança U de x existe $n \in \mathbb{N}$ tal que $U \cap f^n(U) \neq \emptyset$. Caso contrario diremos que x é **errante**. Denotamos por $\Omega(f)$ o conjunto de todos os pontos não-errantes.*

Teorema 1.3.4 (Teorema de Recorrência de Poincaré versão mensurável). *Seja $f : X \to X$ uma transformação mensurável é μ uma medida finita invariante por f. Seja $E \subset X$ qualquer conjunto mensurável com $\mu(E) > 0$. Então, para $\mu - qtp$ $x \in E$ existem infinitos de valores de n para os quais $f^n(x)$ também está em E.*

Prova: Seja $(X, \sum, \mu, f)$ e $E \subset X$ um conjunto mensurável. Considere agora $(\bigcap_{n\in\mathbb{N}} f^{-n}(E^c)) \cap E = \tilde{E} = \{x \in E | f^n(x) \notin E; \forall\, n \in \mathbb{N}\}$. Mostraremos que $\mu(\tilde{E}) = 0$.

Afirmação: As pre-imagens f^{-n} são disjuntas duas a duas.

Prova da afirmação: Suponha que existem $m, n \geq 1$ e $m > n$ tais que $f^{-n}(\tilde{E}) \cap f^{-m}(\tilde{E}) \neq \emptyset$. Seja $x \in (f^{-n}(\tilde{E}) \cap f^{-m}(\tilde{E}))$ e seja $y = f^n(x)$. Então $y \in \tilde{E}$, $f^{m-n}(y) = f^m(x)$ ou seja $f^{m-n}(y) = f^m(f^{-n}(f^n(x)) = f^m(x) \in E$, dado que $\tilde{E} \subset E$.

O argumento anterior implica que y volta a E pelo menos uma vez. Este fato contradiz a definição de $\tilde{E}$, mostrando o desejado.

Como temos pela hipótese que μ é invariante por f, então,

$$\mu(f^{-n}(\tilde{E})) = \mu(\tilde{E}), \forall\, n \in \mathbb{N},$$

Conclui-se que se $\mu(\tilde{E}) > 0$, logo,

$$\infty > \mu(X) = 1 \geq \mu\big(\bigcup_{n=1}^{\infty} f^{-n}(\tilde{E})\big) = \sum_{n=1}^{\infty} \mu(f^{-n}(\tilde{E})) = \sum_{n=1}^{\infty} \mu(\tilde{E}) = \infty$$

Isto força que $\mu(\tilde{E}) = 0$;, mostrando o desejado.

Seja $F = \{x \in E | f^n(x) \in E, \ para\ finitos\ valores\ de\ n\}$. Logo como consequência da definição inicial temos que todo ponto $x \in F$ tem algum iterado $f^k(x) \in E$ e $f^n(x)$ não pertence a E para $n > k$. Ou seja

$$F \subset \bigcup_{k=0}^{\infty} f^{-k}(\tilde{E}). \tag{1.3}$$

Logo, tomando μ em (1.3) temos que $\mu(F) = 0$, como se desejava.

□

Seja X um espaço topológico munido com sua σ-álgebra de Borel $\mathcal{B}$.

Definição 1.3.8. *Dizemos que um ponto $x \in X$ chama-se recorrente para uma transformação $f : X \to X$, se existe uma sequência $n_j \to \infty$ em $\mathbb{N}$ tal que $f^{n_j}(x) \to x$.*

Teorema 1.3.5 (Teorema de Recorrência de Poincaré versão Topológica). *Suponha que X admite uma base enumerável de abertos. Seja $f : X \to X$ uma transformação mensurável e seja μ uma medida finita em X e invariante por f. Então, μ-qtp $x \in X$ é recurrente para f.*

Prova: Consideremos uma base enumerável de abertos U_k de X, onde $k \in \mathbb{N}$. Para cada k, representamos por $\tilde{U}_k$ ao conjunto de pontos $x \in U_k$ que não regressam a U_k. Pelo teorema de Poincaré versão mensurável sabemos que

$$\mu(\tilde{U}_k) = 0,$$

consequentemente,

$$\tilde{U} = \bigcup_{k \in \mathbb{N}} \tilde{U}_k, \text{ logo se deduze que } \mu(\tilde{U}) = 0.$$

Afirmação: Todo ponto $x \notin \tilde{U}$ é recurrente.

Seja $x \in (\tilde{U})^c$ e seja U uma vizinhança qualquer de x. Por definição da topologia gerada por básicos, existe algum elemento U_k da base de abertos tal que $x \in U_k$ e $U_k \subset U$. Como x não está em $\tilde{U}$, também temos que $x \notin \tilde{U}_k$, $\forall\ k$. Em outras palavras, $\exists n \geq 1$ tal que $f^n(x)$ está em U_k. Em particular, $f^n(x)$ também está em U.

Como a vizinhança U é arbitrária, isto prova que x é um ponto recurrente. Mostrando o desejado.

$\square$

1.4 Variedades Topológicas

Definição 1.4.1. *Seja M um espaço topológico.* ***Um sistema de coordenadas locais*** *ou uma* ***carta local*** *em M é um homeomorfismo $\varphi : U \to \varphi(U)$ de um subconjunto aberto $U \in M$ sobre um aberto de $\mathbb{R}^m$. Diz-se que $m = m(U)$ é a dimensão de $\varphi : U \to \varphi(U)$. Para cada $x \in U$ tem-se $\varphi(x) = (\varphi^1(x), ..., \varphi^m(x))$. Os números $\varphi^i = \varphi^i(x), i = 1, ..., m$ são chamados de coordenadas do ponto $x \in M$ no sistema φ.*

Definição 1.4.2. *Um atlas de dimensão m sobre um espaço topológico M é uma coleção $\mathcal{A}$ de sistemas de coordenadas locais $\varphi : U \to \mathbb{R}^m$ em M, cujos domínios U cobrem M. Os domínios U dos sistemas de coordenadas $\varphi \in \mathcal{A}$ são chamados as* ***vizinhanças coordenadas*** *de $\mathcal{A}$.*

Definição 1.4.3. *Um espaço topológico M é* **localmente euclidiano** *de dimensão m se cada ponto x em M tiver uma vizinhança U tal que exista um homeomorfismo φ de U para um subconjunto aberto de $\mathbb{R}^m$. Chamamos o par $(U, \varphi : U \to \mathbb{R}^m)$ de carta, U de vizinhança coordenada ou conjunto aberto coordenado, e φ de função de coordenadas ou sistema de coordenadas em U. Dizemos que uma carta (U, φ) é centrado em $x \in U$ se $\varphi(x) = 0$.*

Definição 1.4.4. *Uma* ***variedade topológica*** *é um espaço de Hausdorff, segundo contável, localmente euclidiano. Diz-se que é de dimensão n se for localmente euclidiano de dimensão n.*

Dados os sistemas de coordenadas $\varphi : U \to \mathbb{R}^m$ e $\psi : V \to \mathbb{R}^m$ no espaço topológico M, tais que $U \cap V \neq \emptyset$, cada ponto $x \in U \cap V$ tem coordenadas $\varphi_i = \varphi_i(x)$ no sistema φ e coordenadas $\psi_i = \psi_i(x)$ no sistema ψ.

A correspondência $(\varphi^1(x), ..., \varphi^m(x)) \leftrightarrow (\psi^1(x), ..., \psi^m(x))$ estabelece um homeomorfismo $\phi_{\varphi\psi} = \psi \circ \varphi^{-1} : \varphi(U \cap V) \to \psi(U \cap V)$ que é chamado **mudança de coordenadas**.

1.5 Variedades diferenciáveis

Definição 1.5.1. *Duas cartas locais $(U, \varphi), (\psi, V)$ num espaço topológico M são ditas* ***compatíveis*** *se $\varphi(U \cap V)$ e $\psi(U \cap V)$ são abertos em $\mathbb{R}^m$ e aplicação de transição $\psi \circ \varphi^{-1}$ é um difeomorfismo.*

Note que a condição de $\psi \circ \varphi^{-1}$ ser um difeomorfismo implica que $\varphi \circ \psi^{-1}$ também seja um difeomorfismo.

Observação: A noção de compatibilidade para cartas locais (U, φ) e (V, ψ) faria sentido também na situação mais geral em que $\varphi(U)$ é um aberto de $\mathbb{R}^m$ e $\psi(V)$ é um aberto

$\mathbb{R}^n$ onde, ao principio, m não precisa ser igual a n mais se $U \cap V \neq \emptyset$, tal compatibilidade implicaria na existência de um difeomorfismo de um aberto não-vazio de $\mathbb{R}^m$ sobre um aberto de $\mathbb{R}^n$ o que implicaria que $m = n$.

Definição 1.5.2. *Um* ***atlas*** $\mathcal{A}$ *de dimensão* n *num espaço topológico* M *é uma coleção* $\mathcal{A} = \{(U_\alpha, \varphi_\alpha); \alpha \in I\}$ *de cartas locais em* M*, duas a duas compatíveis, onde cada* $\varphi_\alpha(U_\alpha)$ *é aberto em* $\mathbb{R}^m$*, e tal que* $M = \bigcup_{\alpha \in I} U_\alpha$.

Um atlas $\mathcal{A}$ sobre um espaço topológico M é um atlas diferenciável de classe $C^r, r \geq 1$, se todas as mudanças de coordenadas $\phi_{\varphi\psi}$, com $\varphi, \psi \in \mathcal{A}$, são aplicações C^r. Escreve-se então $\mathcal{A} \in C^r$.

Como $\phi_{\varphi\psi} = (\phi_{\psi\varphi})^{-1}$, os $\phi_{\varphi\psi}$ são, de fato difeomorfismo de classe C^r. Seja $\mathcal{A}$ um atlas de dimensão m e classe C^r num espaço topológico M. Um sistema de coordenadas $\phi : W \to \mathbb{R}^m$ em M diz-se **admissível** relativamente ao atlas $\mathcal{A}$ se, para todo sistema de coordenadas locais $\varphi : U \to \mathbb{R}^m$, pertence a $\mathcal{A}$, com $U \cap W \neq \emptyset$, as mudanças de coordenadas $\phi_{\varphi\psi}$ e $\phi_{\psi\varphi}$ são C^r em M. Em outras palavras, se $\mathcal{A} \cup \{\phi\}$ é ainda um atlas de classe C^r em M.

Um atlas $\mathcal{A}$ de dimensão m e classe C^r sobre M diz-se **maximal** quando contém todos os sistemas de coordenadas locais que são admissíveis em relação a $\mathcal{A}$.

Todo atlas de classe C^r em M pode ser ampliado, de modo único, até se tornar um atlas máximo de classe C^r: basta acrescentar-lhe todos os sistemas de coordenadas admissíveis.

Definição 1.5.3. *Uma* ***variedade diferenciável****, de dimensão* m *e classe* C^r *é um par ordenado* $(M, \mathcal{A})$*, onde* M *é um espaço topológico de Hausdorff, com base enumerável e* $\mathcal{A}$ *é um atlas máximo de dimensão* m *e classe* C^r *sobre* M.

Capítulo 2

Entropia Topológica

Neste capítulo vamos definir a entropia topológica de um sistema dinâmico. É necessário deixar claro que de agora em diante todo espaço métrico X será considerado compacto, a menos que se mencione o contrário. As demonstrações tem como referência [1], [2], [3]. Além disso, para contextualizar o leitor, vamos a inserir algumas definições sobre os sistemas dinâmicos que serão usados ao longo do desenvolvimento deste capítulo.

Um sistema dinâmico em um tempo discreto, consiste de um conjunto $X \neq \emptyset$ e uma função $f : X \to X$. Para $n \in \mathbb{N}$, a $n-$ésima iterada de f é a $n-upla$ composição $f^n = f \circ ... \circ f$ $n-$vezes ou $n-$tempos; definiremos f^0 como a função identidade, denotaremos por $f^0 = Id$. Se f é invertível, então $f^{-n} = f^{-1} \circ ... \circ f^{-1}$ ($n-$vezes). Como $f^{n+m} = f^n \circ f^m$, estas iterações formam um grupo se f for invertível e um semigrupo caso contrário.

Definição 2.0.1. *Um ponto $p \in X$ é **fixo** se $f(p) = p$. O conjunto de todos os pontos fixos será denotado por **Fix(f)**.*

Definição 2.0.2. *Um ponto $y \in X$ é **periódico** se existe $n \in \mathbb{Z}^+$ tais que $f^n(y) = y$. Se $n = \min\{k : f^k(y) = y; k \in \mathbb{Z}^+\}$, n é chamado período e, neste caso y é chamado periódico de período n. O conjunto de todos os pontos periódicos será denotados por **Per(f)**.*

Definição 2.0.3. *Para uma transformação $f : X \to X$, escreveremos $P_n(f)$ como o número de pontos periódicos de f com período n (não necessariamente mínimo), ou seja, o número de pontos fixos de f^n.*

A medida mais natural do crescimento assintótico do número de pontos periódicos é a taxa de crescimento exponencial $p(f)$ para a sequência $P_n(f)$:

$$p(f) = \lim_{n\to\infty} \sup \frac{\log\left(\max\left(P_n(f), 1\right)\right)}{n}. \tag{2.1}$$

Escrevemos $\max\left(P_n(f), 1\right)$ em vez $P_n(f)$ para evitar tomar $\log 0$.

Iniciando propriamente o assunto da entropia, seja $f : \mathrm{X} \longrightarrow \mathrm{X}$ uma função contínua de um espaço métrico compacto X, com distância d. Para cada $n \in \mathbb{N}$, introduzimos uma nova distância em X por:

$$d_n^f(x, y) = \max\{d(f^k(x), f^k(y)) : 0 \le k \le n-1\} \tag{2.2}$$

Em outras palavras, d_n^f mede a distância entre o segmento de órbita $I_x^f = \{x, ...f^{n-1}(x)\}$ e $I_y^f = \{y, ..., f^{n-1}(y)\}$. Denota-se a bola aberta $\{y \in X : d_n^f(x, y) < \varepsilon\}$ por $B_f(x, \varepsilon, n)$.

Um subconjunto $A \subset X$ é $(n, \varepsilon) -$ *separado* se dois pontos distintos quaisquer de A, estão separados por pelos menos ε na métrica d_n. Seja $N_d(n, \varepsilon)$ o maior cardinal de um conjunto $(n, \varepsilon) -$ *separado*, que é finito pela compacidade.

Definição 2.0.4. *Suponha que $f : X \to X$ é uma transformação contínua num espaço métrico compacto X. Dado $x \in X, n \ge 1$ e $\varepsilon > 0$, chamamos **bola dinâmica** de $n-$nível e raio ε em torno de x o conjunto:*

$$\begin{aligned} B_n(x, \varepsilon) &= \{y \in X|\ d(f^i(x), f^i(y)) < \varepsilon;\ 0 \le i \le n-1\} \\ &= \bigcap_{i=0}^{n-1} f^{-i}\big(B(f^i(x), \varepsilon)\big). \end{aligned}$$

Observação: Para fins de esclarecimento, apresentamos as seguintes notações para uma d_n-bola dinâmica.

$$B_f(x, \varepsilon, n) = B_n(x, \varepsilon) = B_n^f(x, \varepsilon) = B(x, \varepsilon, n).$$

Definição 2.0.5. *A **entropia topológica** de f é definida por:*

$$h_{top}(f) = \lim_{\varepsilon\to 0} \lim_{n\to\infty} \sup \frac{1}{n} \log N_d(n, \varepsilon). \tag{2.3}$$

Notemos que a função

$$\varepsilon \mapsto \lim_{n\to\infty} sup\tfrac{1}{n}\log N_d(f,\varepsilon,n)$$

é não decrescente, o limite em (2.3), quando $\varepsilon \to 0$ está bem definido.

Outras definições de entropia

Agora vamos descrever outras caracterizações alternativas importantes da entropia e algumas equivalências.

Para iniciar, é necessário falar de alguns conjuntos de muita importância;

- Um subconjunto $E \subset X$ é chamado de $(n,\varepsilon)-expansivo$ se $X \subset \bigcup_{x\in E} B_f(x,\varepsilon,n)$. Seja $S_d(f,\varepsilon,n)$ o cardinal mínimo de um $(n,\varepsilon)-expansivo$.

- Seja $S \subset X$ é um $(n,\varepsilon)-gerador$ se para cada $x \in X$ existe $y \in S$, tal que $d_n(x,y) \leq \varepsilon$. Seja $G_d(n,\varepsilon)$ o menor número de pontos de um $(n,\varepsilon)-gerador$.

- Seja $M_d(f,\varepsilon,n)$ o mínimo número de pontos $p_1,...,p_m$ de X tal que as bolas $B_d(p_1,\varepsilon),...,B_d(p_m,\varepsilon)$ cobrem X.

Podemos também definir a entropia topológica da seguinte forma:

$$h_{top}(f) = \lim_{\varepsilon\to 0}\lim_{n\to\infty}\sup\frac{1}{n}\log S_d(f,\varepsilon,n); \tag{2.4}$$

$$h_{top}(f) = \lim_{\varepsilon\to 0}\lim_{n\to\infty}\sup\frac{1}{n}\log G_d(f,\varepsilon,n); \tag{2.5}$$

$$h_{top}(f) = \lim_{\varepsilon\to 0}\lim_{n\to\infty}\sup\frac{1}{n}\log M_d(f,\varepsilon,n). \tag{2.6}$$

Proposição 2.0.1. *Para todo $n \in \mathbb{N}$ e para todo $\varepsilon > 0$, o número $M_d(f,n,\varepsilon)$ é igual a quantidade mínima de bolas dinâmicas $B_f(p,\varepsilon,n)$ necessárias para cobrir X.*

Prova: Seja uma cobertura $\{B(n,\varepsilon,p_1),...,B(n,\varepsilon,p_m)\}$ com uma quantidade mínima m de bolas dinâmicas de raio ε e centros $p_1,...,p_m$ respectivamente. Por ser um cobertura, dado qualquer ponto $x \in X$, existe p_i tais que $x \in B(p_i,\varepsilon,n)$. Logo, por definição de bola dinâmica, $d_n(x,p_i) < \varepsilon$, dizemos que $\{p_1,...,p_m\}$ é um $(n,\varepsilon)-$gerador. Então, a quantidade de pontos $M_d(f,\varepsilon,n)$ de um $(n,\varepsilon)-$gerador minimal e menor ou igual que m. Revelando que

$$M_d(f,\varepsilon,n) \leq m. \tag{2.7}$$

Agora provaremos a desigualdade oposta. Seja $p_1, ..., p_k$ um $(n,\varepsilon)-$gerador minimal. Então, $M_d(f,\varepsilon,n) = k$; onde k é a quantidades mínima de pontos do conjunto $(n,\varepsilon)-$gerador.

Consideremos as bolas dinâmicas $B(n,\varepsilon,p_1), ..., B(n,\varepsilon,p_k)$. Como qualquer ponto $x \in X$ está a uma d_n-distância menor que ε de algum p_i, resulta que $x \in B(n,\varepsilon,p_i)$ para algum i. Donde $X = \bigcup_{i=1}^{k} B(p_i,\varepsilon,n)$, ou seja $\{B(n,\varepsilon,p_1), ..., B(n,\varepsilon,p_k)\}$ é um cobertura para X.

Portanto, k é maior ou igual que a quantidade mínima m de bolas dinâmicas de raio ε necessárias para cobrir X. Assim , obtemos

$$M_d(f,\varepsilon,n) \geq m. \tag{2.8}$$

De (2.7) e (2.8) obtemos que $M_d(f,\varepsilon,n) = m$.

□

Teorema 2.0.1. *Seja $f : X \to X$ uma função contínua. Se d' é outra métrica de X que define a mesma topologia que d, então $h_{d'}(f)$=$h_d(f)$.*

Demostração: Denotamos a função identidade como:

$$I_d : X \to X,$$

que I_d é um homeomorfismo. Logo da compacidade de X, também é uniformemente contínua. Dado $\varepsilon > 0$, existe δ_ε tal que, se $d'(x_1,x_2) < \delta_\varepsilon$, então $d(x_1,x_2) < \varepsilon$, para $x_1, x_2 \in X$, isto significa que; $B_{d'}(x_1,\delta_\varepsilon) \subset B_d(x_1,\varepsilon)$. Pela continuidade uniforme temos que o mesmo vale também para ${d'}_n^f$ e d_n^f, ou seja,

$$B_{d'}(f^i(x_1),\delta_\varepsilon) \subset B_d(f^i(x_1),\varepsilon); 0 \leq i \leq n-1.$$

Como X é compacto, então podemos recobrir o espaço por uma quantidade finita de bolas $B_{d'}$ e B_d, o que significa que:

$$S_{d'}(f,\delta_\varepsilon,n) \geq S_d(f,\varepsilon,n).$$

Logo,

$$\begin{aligned}
\lim_{n\to\infty}\sup\frac{1}{n}\log S_{d'}(f,\delta_\varepsilon,n) &\geq \lim_{n\to\infty}\sup\frac{1}{n}\log S_d(f,\varepsilon,n)\\
h_{d'}(f,\delta_\varepsilon) &\geq h_d(f,\varepsilon)\\
\lim_{\delta_\varepsilon\to 0} h_{d'}(f,\delta_\varepsilon) &\geq \lim_{\delta_\varepsilon\to 0} h_d(f,\varepsilon)\\
h_{d'}(f) &\geq h_d(f,\varepsilon)\\
\lim_{\varepsilon\to 0} h_{d'}(f) &\geq \lim_{\varepsilon\to 0} h_d(f,\varepsilon)\\
h_{d'}(f) &\geq h_d(f).
\end{aligned}$$

Assim obtivemos uma das desigualdades. A outra desigualdade segue de modo análogo.

□

Definição 2.0.6. *Sejam X e Y espaços métricos. Dadas duas funções contínuas $f : X \to X$ e $g : Y \to Y$ dizemos que são **topologicamente conjugadas** se existe um homeomorfismo $h : X \to Y$ tal que para todo $x \in X$ tem-se que $h(f(x)) = g(h(x))$. Onde h é chamada conjugação topológica.*

Corolário 2.0.1. *Sejam f e g funções contínuas. A entropia topológica é um invariante por conjugação topológica.*

Demostração: Seja $f : X \to X$ e $g : Y \to Y$ topologicamente conjugadas através de um homeomorfismo $h : X \to Y$. Fixemos uma métrica d em X e definimos d' em Y, isto é, $d'(y_1, y_2) = d(h^{-1}(y_1), h^{-1}(y_2))$, então h é uma isometria e aplicando o Teorema (2.0.1) temos: $h_{d'}(f) = h_d(g)$.

Proposição 2.0.2. *Sejam X um espaço métrico compacto e $f : X \to X$ uma função contínua.*

1. *Se $\Lambda \subset X$ é um conjunto fechado f-invariante, então $h_{top}(f|_\Lambda) \leq h_{top}(f)$.*

2. *Se* $X = \bigcup_{i=1}^{n} \Lambda_i$, *onde* Λ_i; $i = 1, ..., n$, *são conjuntos fechados f-invariantes então,* $h_{top}(f) = \max_{1 \leq i \leq n} h_{top}(f|_{\Lambda_i})$.

3. $h_{top}(f^m) = m h_{top}(f)$.

Prova:

1. Em primeiro lugar, é evidente que Λ é compacto, assim temos que $\Lambda \subset \bigcup_{z \in \Lambda} B_{f|_{\Lambda}}(z, \varepsilon, n)$ e que $X \subset \bigcup_{x \in X} B_f(x, \varepsilon, n)$, assim,

$$S_d(f|_{\Lambda}, \varepsilon, n) \leq S_d(f, \varepsilon, n)$$

e, aplicando (2.4) temos

$$\begin{aligned} \lim_{\varepsilon \to 0} \lim_{n \to \infty} \sup \frac{1}{n} \log S_d(f|_{\Lambda}, \varepsilon, n) &\leq \lim_{\varepsilon \to 0} \lim_{n \to \infty} \sup \frac{1}{n} \log S_d(f, \varepsilon, n) \\ h_{top}(f|_{\Lambda}) &\leq h_{top}(f). \end{aligned}$$

Como se desejava.

2. Mostraremos primeiro que: $h_{top}(f) \leq \max_{1 \leq i \leq n} h_{top}(f|_{\Lambda_i})$. A união de conjuntos $(n, \varepsilon)-$expansivos para $\Lambda_i' s$ é um conjunto $(n, \varepsilon)-$expansivo para X. Portanto, $S_d(f|_{\Lambda_i}, \varepsilon, n)$ e o cardinal mínimo de um $(n, \varepsilon)-$expansivo de Λ_i, então

$$S_d(n, \varepsilon, f) \leq \sum_{i=1}^{m} S_d(f|_{\Lambda_i}, \varepsilon, n) \leq m \max_{1 \leq i \leq m} S_d(f|_{\Lambda_i}, \varepsilon, n).$$

Assim temos que: $S_d(n, \varepsilon, f) \leq m \max_{1 \leq i \leq m} S_d(f|_{\Lambda_i}, \varepsilon, n)$, logo aplicando (2.4), obtemos

$$\begin{aligned}
\lim_{\varepsilon\to 0}\lim_{n\to\infty}\sup\frac{1}{n}\log S_d(f,\varepsilon,n) &\leq \lim_{\varepsilon\to 0}\lim_{n\to\infty}\sup\frac{1}{n}\log m\max_{1\leq i\leq m} S_d(f|_{\Lambda_i},\varepsilon,n)\\
&\leq \lim_{\varepsilon\to 0}\lim_{n\to\infty}\sup\frac{1}{n}\log m+\lim_{\varepsilon\to 0}\lim_{n\to\infty}\sup\frac{1}{n}\log\max_{1\leq i\leq m} S_d(f|_{\Lambda_i},\varepsilon,n)\\
\lim_{\varepsilon\to 0} h(f,\varepsilon) &\leq 0+\lim_{\varepsilon\to 0}\max_{1\leq i\leq m} h(f|_{\Lambda_i},\varepsilon)\\
h_{top}(f) &\leq \max_{1\leq i\leq m} h_{top}(f|_{\Lambda_i}). \qquad (2.9)
\end{aligned}$$

A segunda desigualdade é evidente, ou seja como cada Λ_i são subconjuntos de X, então particularmente,

$$\max_{1\leq i\leq m} h_{top}(f|_{\Lambda_i}) \leq h_{top}(f). \qquad (2.10)$$

Mostrando o desejado. Agora de (2.9) e (2.10) obtemos que:

$$h_{top}(f)=\max_{1\leq i\leq m} h_{top}(f|_{\Lambda_i}).$$

3. Mostraremos que $h_{top}(f^m)\leq m h_{top}(f)$. Temos que:

$$\begin{aligned}
d_n^{f^m}(x,y) &= \max\{d(f^{im}(x),d(f^{im}(y)));0\leq i\leq n-1\}\\
&\leq \max\{d(f^{i}(x),d(f^{i}(y)));0\leq i\leq mn-1\}\\
&= d_{nm}^{f}(x,y).
\end{aligned}$$

Portanto, qualquer $d_n^{f^m}$ $\varepsilon-bola$ contém uma d_{nm}^{f} $\varepsilon-bola$.

Ou seja $B(f,\varepsilon,nm)\subset B(f^m,\varepsilon,n)$, assim temos que, $S_d(f^m,\varepsilon,n)\leq S_d(f,\varepsilon,mn)$, aplicando (2.4), obtemos,

$$\begin{aligned}
\lim_{\varepsilon\to 0}\lim_{n\to\infty}\sup\frac{1}{n}\log S_d(f^m,\varepsilon,n) &\leq \lim_{\varepsilon\to 0}\lim_{n\to\infty}\sup\frac{1}{n}\log S_d(f,\varepsilon,mn)\\
&\leq m\lim_{\varepsilon\to 0}\lim_{n\to\infty}\sup\frac{1}{mn}\log S_d(f,\varepsilon,mn)\\
h_{top}(f^m) &\leq mh_{top}(f). \qquad (2.11)
\end{aligned}$$

Mostrando assim a desigualdade desejada. Agora mostraremos a desigualdade oposta, ou seja, $h_{top}(f^m)\geq mh_{top}(f)$.

Dado que f é uniformemente contínua, para cada $\varepsilon>0$ existe $\delta(\varepsilon)>0$ com $\delta(\varepsilon)\in(0,\varepsilon)$ tal que $d_m^f(x,y)<\varepsilon$ quando $d(x,y)<\delta(\varepsilon)$ ou seja, $B(x,\delta(\varepsilon))\subset B_f(x,\varepsilon,m)$.

Dado que $d_{mn}^f(x,y)=\max\{d(f^{im}(x),f^{im}(y));\ 0\leq i\leq n-1\}$ assim, $d_{mn}^f(x,y)<\varepsilon$ quando $d_n^{f^m}(x,y)<\delta(\varepsilon)$ ou seja

$$\begin{aligned}
B_{f^m}(x,\delta(\varepsilon),n) &= \bigcap_{i=0}^{n-1} f^{-im}B(f^{im}(x),\delta(\varepsilon))\\
&\subset \bigcap_{i=0}^{n-1} f^{-im}B_f(f^{im}(x),\varepsilon,m)\\
&= B_f(x,\varepsilon,mn),
\end{aligned}$$

isto é, $B_{f^m}(x,\delta(\varepsilon),n)\subset B_f(x,\varepsilon,mn)$.

Assim,

$$S_d(f,\varepsilon,mn)\leq S_d(f^m,\delta(\varepsilon),n).$$

Aplicando (2.4), temos que

$$\lim_{\varepsilon\to 0}\lim_{n\to\infty}\sup\frac{1}{n}\log S_d(f,\varepsilon,mn) \leq \lim_{\varepsilon\to 0}\lim_{n\to\infty}\sup\frac{1}{n}\log S_d(f^m,\delta(\varepsilon),n)$$

$$m \lim_{\varepsilon\to 0}\lim_{n\to\infty}\sup \frac{1}{mn}\log S_d(f,\varepsilon,mn) \quad \leq \quad \lim_{\varepsilon\to 0}\lim_{n\to\infty}\sup \frac{1}{n}\log S_d(f^m,\delta(\varepsilon),n)$$

$$mh_{top}(f) \leq h_{top}(f^m). \tag{2.12}$$

De (2.11) e (2.12) se conclui que $h_{top}(f^m) = mh_{top}(f)$. Concluindo a prova da proposição. □

Lembre-se que o $d - diametro$ de um conjunto $A \subset X$ é o número

$$\sup\{d(x,y); x, y \in A\}.$$

Dado $\varepsilon > 0$, seja $D_d(f,n,\varepsilon)$ o menor número de conjuntos de d-diâmetro menor que ε necessários para cobrir X.

Proposição 2.0.3. *Para cada $n \in \mathbb{N}$ e $\varepsilon > 0$, temos que:*

$$D_d(f,2\varepsilon,n) \leq M_d(f,\varepsilon,n) \leq N_d(f,\varepsilon,n) \leq M_d(f,\varepsilon/2,n) \leq D_d(f,\varepsilon/2,n).$$

Prova:

- Provaremos que: $D_d(f,2\varepsilon,n) \leq M_d(f,\varepsilon,n)$. Tomemos os pontos $p_1,...,p_m \in X$ tal que para cada $x \in X$ temos que $d_n^f(x,p_i) < \varepsilon$ para algum i,. Claramente, as $d_n^f -$ *bolas* abertas $B_f(p_i,\varepsilon,n)$ cobrem todo X. Logo como $B_f(p_i,\varepsilon,n)$ tem d_n^f-diâmetro menor ou igual a 2ε, então podemos concluir que, $D_d(f,2\varepsilon,n) \leq M_d(f,\varepsilon,n)$.

- Provaremos que: $M_d(f,\varepsilon,n) \leq N_d(f,\varepsilon,n)$. Agora tomemos um conjunto de pontos $p_1,...,p_m \in X$ tais que $d_n(p_i,p_j) \geq \varepsilon$ para $i \neq j$. Note que $X - \{p_1,...,p_m\}$ satisfaz que $d_n(x,p_i) < \varepsilon$ para algum i, logo, $M_d(f,\varepsilon,n) \leq N_d(f,\varepsilon,n)$.

- Provaremos que: $N_d(f,\varepsilon,n) \leq M_d(f,\varepsilon/2,n)$.

  Nesta desigualdade, podemos ver que nenhuma d_n-bola de raio $\varepsilon/2$ pode conter dois pontos a uma d_n-distância de ε, portanto, é evidente que $N_d(f,\varepsilon,n) \leq M_d(f,\varepsilon/2,n)$.

- Finalmente, provaremos que: $M_d(f,\varepsilon/2) \leq D_d(f,\varepsilon/2,n)$. Tomemos agora uma cobertura para X, digamos que $A_1,...,A_n$ é tal cobertura por conjuntos de d_n-diâmetro menor que 2ε. Agora tomemos um ponto $p_i \in A_i$, para cada i. Logo, $A_i \subset B_n(p_i,\varepsilon/2)$ e $\displaystyle\bigcup_{i=1}^{m} B_n(p_i,\varepsilon/2) \supset X$. Assim, $M_d(f,\varepsilon/2,n) \leq D_d(f,\varepsilon/2,n)$.

$\square$

Lema 2.0.1. *Se $(a_n)_{n\in\mathbb{N}}$ é uma sequência de números positivos tal que $a_{n+m} \leq a_n + a_m$ para cada $n,m \in \mathbb{N}$, então*

$$\lim_{k\to\infty} \frac{a_k}{k} = \inf\left\{\frac{a_k}{k} : k \in \mathbb{N}\right\} \geq 0. \tag{2.13}$$

Prova: Como a sequência é uma sequência de números positivos, logo é limitada inferiormente por zero.

Definamos $c = \inf\{\frac{a_n}{n}\}$; quando $n \geq 1$, a ideia é mostrar que $\lim_{n\to\infty} \frac{a_n}{n} = c$.

Seja $\epsilon > 0$ qualquer e $k \in \mathbb{N}$ tais que $c \leq \frac{a_k}{k} < c+\frac{\epsilon}{2}$. Seja $R = \max\{a_i; 1 \leq i \leq k-1\}$.

Dado $n \in \mathbb{N}$ existem $q \geq 0$ e $r \in \{0,...,k-1\}$ tal que $n = qk + r$.

$$\begin{aligned}
\frac{a_n}{n} &= \frac{a_{qk+r}}{qk+r} \\
&\leq \frac{a_{qk}+a_r}{qk+r} \\
&\leq \frac{qa_k+a_r}{qk+r} \\
&\leq \frac{qa_k}{qk+r} + \frac{R}{n} \\
&= \left(\frac{qk}{qk+r}\right)\left(\frac{a_k}{k}\right)+\frac{R}{n}.
\end{aligned}$$

Dado que $n = qk + r$, podemos obter $qk = n - r$, assim;

$$= \left(\frac{n-r}{n}\right)\left(\frac{a_k}{k}\right)+\frac{R}{n}. \text{ Onde,}$$

$$
\begin{aligned}
\limsup_{n\to\infty} \frac{a_n}{n} &\leq \limsup_{n\to\infty} \left\{ \left(\frac{n-r}{n}\right)\left(\frac{a_k}{k}\right) + \frac{R}{n} \right\} \\
&= \frac{a_k}{k} < c + \epsilon.
\end{aligned}
$$

Assim,

$$\limsup_{n\to\infty} \frac{a_n}{n} \leq c, \text{ ou seja}$$

$$\limsup_{n\to\infty} \frac{a_n}{n} \leq \liminf_{n\to\infty} \frac{a_n}{n}$$

Por outro lado, sempre vale que

$$\liminf_{n\to\infty} \frac{a_n}{n} \leq \limsup_{n\to\infty} \frac{a_n}{n}.$$

Concluindo finalmente que:

$$\limsup_{n\to\infty} \frac{a_n}{n} = \liminf_{n\to\infty} \frac{a_n}{n},$$

ou seja,

$$\lim_{n\to\infty} \frac{a_n}{n} = c.$$

Lema 2.0.2. *Dados $m, n \in \mathbb{N}$ e $\varepsilon > 0$, temos que,*

$$D_d(f, \varepsilon, m+n) \leq D_d(f, \varepsilon, n) D_d(f, \varepsilon, m).$$

Prova: Seja U uma cobertura para X por $D_d(f, \varepsilon, n)$ conjuntos de d_n^f-diâmetro menor que ε, ou seja, $\sup\{d_n^f(x,y) < \varepsilon; x, y \in U_i; U_i \in U\}$ e W outra cobertura de X por $D_d(f, \varepsilon, m)$ conjuntos de d_m^f-diâmetro, ou seja, $\sup\{d_m^f(z,r) < \varepsilon; z, r \in W_j; W_j \in W\}$.

Logo o conjunto,

$$V = U_i \cup T^{-n}(W_j); U_i \in U \text{ e } W_j \in W$$

tem $d_{m+n}-$diâmetro menor que ε.

Isto é $d_{m+n}(x,y) = \sup\{d_n(x,y), d_m(f^n(x), f^n(y))\} < \varepsilon$.

Logo V contém no máximo $D_d(f,\varepsilon,n)D_d(f,\varepsilon,m)$ conjuntos e V também é uma cobertura para X.

Portanto,

$$D_d(f,\varepsilon,m+n) \leq D_d(f,\varepsilon,n)D_d(f,\varepsilon,m).$$

$\square$

Exemplo 1:

Seja $\widetilde{A}$ uma matriz com entradas enteiras de 2×2 com determinante igual a 1 e valores próprios λ, λ^{-1} com $|\lambda| > 1$; e seja $A : \mathbb{T}^2 \to \mathbb{T}^2$ o automorfismo toral hiperbólico associado, então $h_{top}(A) = \log(\lambda)$.

Solução:

- Definamos a projeção natural: $\pi : \mathbb{R}^2 \to \mathbb{T}^2$ que é um homeomorfismo local satisfazendo $\pi\widetilde{A} = A\pi$.

- Qualquer métrica $\widetilde{d}$ em $\mathbb{R}^2$ invariante sobre traslações inteiras, induz uma métrica d em $\mathbb{T}^2$, onde $d(x,y)$ é a $\widetilde{d}$-distância entre os conjuntos $\pi^{-1}(x), \pi^{-1}(y)$, para estas métricas π é uma isometria local.

- Sejam v_1 e v_2 os vetores próprios de A de comprimento 1 correspondentes ao valores próprios λ, λ^{-1}.

- Para $x, y \in \mathbb{R}^2$, escrevemos $x - y = a_1v_1 + a_2v_2$ e definimos $\widetilde{d}(x,y) = \max(|a_1|, |a_2|)$.

- Uma $\widetilde{d}$-bola de raio ε é um paralelogramo cujos lados são de comprimento 2ε e paralelos a v_1 e v_2.

As bolas abertas são;

- $\left\{y \in \mathbb{T}^2 / \widetilde{d}(A^i(x), A^i(y)) < \varepsilon; i = \{0, ..., n-1\}\right\}$.

- Na métrica $\widetilde{d}_n$ (definida por $\widetilde{A}$) é um paralelogramo de raio ε com comprimento $2\varepsilon|\lambda|^{-n}$ na direção v_1 e 2ε na direção v_2.

- Qual é o volume de uma bola aberta?

  Observe que o volume desta bola aberta é

$$V_n = \frac{4\varepsilon^2}{|\lambda|^n} \tag{2.14}$$

- Logo precisamos de pelo menos:

$$k = \frac{|\lambda|^n}{4\varepsilon^2} \tag{2.15}$$

  bolas, para cobrir o toro.

- Para garantir que todo o toro seja coberto, vamos definir um pequeno deslocamento da forma $(x + \frac{1}{5}, y + \frac{1}{5})$ onde (x, y) são os centros das bolas no toro.

- Com o anterior deslocamento obtivemos outro recobrimento com o mesmo cardinal que o primeiro.

  Portanto temos que para fazer um recobrimento total do toro precisamos de k_1 bolas, de forma que

$$k_1 = \frac{2|\lambda|^n}{4\varepsilon^2} = \frac{|\lambda|^n}{2\varepsilon^2} \tag{2.16}$$

- Assim temos de (2.14) e (2.15) que:

$$k = \frac{|\lambda|^n}{4\varepsilon^2} \leqslant N_{\widetilde{d}}(A, \varepsilon, n) \leqslant \frac{|\lambda|^n}{2\varepsilon^2} = k_1. \tag{2.17}$$

Finalmente aplicando: $\lim\limits_{\varepsilon \to 0} \lim\limits_{n \to \infty} \sup \frac{1}{n} \log$ encontramos o desejado:

$$h_{top}(A) = \log(\lambda).$$

Agora daremos uma demostração de um caso mais geral do automorfismo hiperbólico do toro.

Proposição 2.0.4. *Seja $\widetilde{A}$ uma matriz inteira de $m \times m$, com polinômio caraterístico irredutível sobre $\mathbb{Q}$ e com todas raízes reais simples, tal que $det(\widetilde{A}) = \pm 1$. Considere $\lambda_1, \ldots, \lambda_s$ os autovalores de $\widetilde{A}$, tais que $|\lambda_i| > 1, i = 1, \ldots, s$. Considere A o automorfismo de $\mathbb{T}^m$ induzido de $\widetilde{A}$, então $h_{top}(A) = \sum\limits_{i=1}^{s} \log(|\lambda_i|)$.*

Prova: Seja a projeção $\pi : \mathbb{R}^m \to \mathbb{T}^m$ que é um homeomorfismo local e satisfaz $\pi\widetilde{A} = A\pi$. Qualquer métrica $\widetilde{d}$ definida em $\mathbb{R}^m$, induz uma métrica d em $\mathbb{T}^m$, dado que é um invariante por traslações, onde $\widetilde{d}(x, y)$ é a $\widetilde{d}-$distância entre os conjuntos $\pi^{-1}(x)$ e $\pi^{-1}(y)$.

Sejam $v_1, ..., v_s$ os autovetores onde; $|v_i| = 1; i = 1, ..., s$, correspondentemente a os autovalores $\lambda_1, ..., \lambda_s$. Definamos agora uma d_N-bola dinâmica assim, $\{y \in \mathbb{T}^m; d(A^j(x), A^j(y)) < \varepsilon; j = 0, ..., N-1\}$ e sejam $x = (x_1, ..., x_m)$ e $y = (y_1, ..., y_m)$, onde x e y estão em coordenadas dadas por uma base de autovetores de norma 1. Logo faremos as iterações da seguinte forma:

$$\begin{aligned}
d(A^0(x), A^0(y)) < \varepsilon &\iff \max\{|y_1 - x_1|, |y_2 - x_2|, ..., |y_m - x_m|\} \\
d(A^1(x), A^1(y)) < \varepsilon &\iff \max\{A|y_1 - x_1|, A|y_2 - x_2|, ..., A|y_m - x_m|\} \\
\vdots &\iff \vdots \\
d(A^{N-1}(x), A^{N-1}(y)) < \varepsilon &\iff \max\{A^{N-1}|y_1 - x_1|, A^{N-1}|y_2 - x_2|, ..., A^{N-1}|y_m - x_m|\}
\end{aligned}$$

Do anterior temos que:

$$\begin{aligned}
|y_1 - x_1| &< \varepsilon \\
&\vdots \\
|y_m - x_m| &< \varepsilon
\end{aligned}$$

Como $\lambda_1, ..., \lambda_s > 1$, então,

$$\begin{aligned}
|Av_1| &= |\lambda_1||v_1| > v_1 \\
&\vdots \\
|Av_s| &= |\lambda_s||v_s| > v_s
\end{aligned} \tag{2.18}$$

Assim temos que:

$$\begin{aligned}
A|y_1 - x_1| < \varepsilon &\iff |\lambda_1||y_1 - x_1| < \varepsilon \iff |y_1 - x_1| < \frac{\varepsilon}{|\lambda_1|} \\
&\vdots \\
A^{N-1}|y_m - x_m| < \varepsilon &\iff |\lambda_s|^{N-1}|y_m - x_m| < \varepsilon \iff |y_m - x_m| < \frac{\varepsilon}{|\lambda_s|^{N-1}}
\end{aligned}$$

Generalizando o raciocínio

$$\begin{aligned}
|y_1 - x_1| &< \frac{\varepsilon}{|\lambda_1|^{N-1}} \\
&\vdots \\
|y_m - x_m| &< \frac{\varepsilon}{|\lambda_s|^{N-1}}.
\end{aligned}$$

Então o comprimento de cada aresta da bola dinâmica paralela a v_i; $i = 1, ..., m$ tem comprimento dado respectivamente por

$$\frac{2\varepsilon}{|\lambda_1|^{N-1}}, ..., \frac{2\varepsilon}{|\lambda_s|^{N-1}}.$$

Calculando o volume de uma bola dinâmica $V_{d_N} = |\lambda_1| \cdot |\lambda_2| \cdot ... \cdot |\lambda_s| = \dfrac{2^N \varepsilon^N}{|\lambda_1|^N \cdot ... \cdot |\lambda_s|^N}$. Agora, calculando a quantidade de bolas dinâmicas para realizar o primeiro recobrimento temos $k \in \mathbb{R}$ tal que

$$\begin{aligned}
k\left(\frac{2^N \varepsilon^N}{|\lambda_1|^N \cdot ... \cdot |\lambda_s|^N}\right) &\geq 1. \\
k &\geq \frac{|\lambda_1|^N \cdot ... \cdot |\lambda_s|^N}{2^N \varepsilon^N}
\end{aligned}$$

Como cada bola dinâmica é uma bola aberta, então parte de $\mathbb{T}^n$ não fica totalmente coberto, para isso é preciso fazer um novo recobrimento que consiste em uma pequena translação, assim a cardinalidade do novo recobrimento é:

$$\begin{aligned}
2k &\geq \left(\frac{2\lambda_1|^N \cdot ... \cdot |\lambda_s|^N}{2^N \varepsilon^N}\right) \\
&\geq \frac{|\lambda_1|^N \cdot ... \cdot |\lambda_s|^N}{2^{N-1} \varepsilon^N}.
\end{aligned}$$

Agora, seja $S_{\widetilde{d}}(A, N, \varepsilon)$ a cardinalidade de um conjunto $(N, \varepsilon)-$expansivo, então temos

$$\frac{|\lambda_1|^N \cdot ... \cdot |\lambda_s|^N}{2^N \varepsilon^N} \leq S_{\widetilde{d}}(A, N, \varepsilon) \leq \frac{|\lambda_1|^N \cdot ... \cdot |\lambda_s|^N}{2^{N-1} \varepsilon^N}.$$

Logo, fazendo uso da definição da entropia obtemos que: $h_{top}(A) = \sum_{i=0}^{s} \log|\lambda_i|$ mostrando o desejado. $\square$

2.1 Finitude da entropia das aplicações de Lipschitz

Uma classe importante das funções contínuas são as funções de Lipschitz. De fato toda transformação C^1 entre variedades compactas é Lipschitz.

Definição 2.1.1. *Seja (X,d) um espaço métrico, $f : X \to X$ uma função contínua Lipschitz. A constante de Lipschitz $L(f)$ que é definida por*

$$L(f) := \sup_{x \neq y} \frac{d(f(x), f(y))}{d(x,y)}.$$

Definição 2.1.2. *Sejam (X,d) um espaço métrico compacto e $b(\varepsilon)$ o menor cardinal de uma cobertura de X por $\varepsilon - bolas$. Então,*

$$D(X) := \limsup_{\varepsilon \to 0} \frac{\log b(\varepsilon)}{|\log(\varepsilon)|} \in \mathbb{R} \cup \{\infty\}. \tag{2.19}$$

*é chamado **dimensão esférica de X.***

Teorema 2.1.1. *Sejam (X,d) um espaço métrico compacto de dimensão esférica finita $D(X)$ e $f : X \to X$ uma aplicação de Lipschitz contínua. Então,*

$$h_{top}(f) \leq D(X) \max(0, \log L(f)).$$

Prova: Seja $1 < \max\{(1, L(f))\}$, então $L > 1$ para qualquer caso, então $d(f(x), f(y)) \leq Ld(x,y) < \varepsilon$ e realizando as iterações temos,

$$d(f^m(x), f^m(y)) \leq ... \leq L^n d(x,y) < \varepsilon, \text{ então,}$$

$$d(f^m(x), f^m(y)) \leq ... \leq d(x,y) \leq L^{-n}\varepsilon < \varepsilon.$$

Em particular, pela continuidade temos que $f^m(B(x, L^n\varepsilon)) \subset B(f^m(x), \varepsilon)$, portanto $B(x, L^n\varepsilon)) \subset B(f(x), \varepsilon)$, quando $0 \leq m \leq n$.

Logo temos que, $S(f, n, \varepsilon) \leq b(L^n\varepsilon)$.

A prova desta última desigualdade segue agora.

Seja $f^m(B((x), \varepsilon L^{-n})) \subset B(f^m, \varepsilon L^{-n+m}) \subset B(f^m(x), \varepsilon)$; $0 \leq m \leq n$. Seja $K \subset X$ é $(n, \varepsilon) - expansivo$ se, e somente se, $\forall y \in X, \exists k \in K$ tal que $d(f^i(x), f^i(y)) < \varepsilon$; $0 \leq i \leq n-1$, então $S(f, \varepsilon, n) = \min\{\#k; k \in X\}$.

Logo, $B(x, L^n \varepsilon)$ é uma cobertura para X por $L^{-n}\varepsilon$ bolas, tal que os centros das bolas formam um conjunto $(n, \varepsilon) - expasivo$, portanto,

$$S(f, n, \varepsilon) \leq b(n, L^n \varepsilon).$$

Por outra parte, analisaremos $|\log(L^{-n}\varepsilon)|$, então,

$$\begin{aligned}
|\log(L^{-n}\varepsilon)| &= |\log(L^{-n}) + \log(\varepsilon))| \\
&= |-n\log(L) + \log(\varepsilon)| \\
&= |n\log(L) - \log(\varepsilon)|.
\end{aligned}$$

Logo se n é um valor muito grande, temos que $n\log(L) - \log(\varepsilon) > 0$, portanto, isto, é o mesmo que:

$$|n\log(L) - \log(\varepsilon)| = n\log(L) - \log(\varepsilon).$$

Assim,

$$|\log(L^{-n}\varepsilon)| = n\log(L) - \log(\varepsilon).$$

Despejando n temos:

$$\begin{aligned}
n &= \frac{|\log(L^{-n}\varepsilon)| + \log(\varepsilon)}{\log(L)} \\
n &= \frac{|\log(L^{-n}\varepsilon)| + \log(\varepsilon)}{\log(L)} \cdot \frac{|\log(L^{-n}\varepsilon)|}{|\log(L^{-n}\varepsilon)|} \\
n &= \frac{|\log(L^{-n}\varepsilon)|}{|\log(L^{-n}\varepsilon)|} + \frac{\log(\varepsilon)}{|\log(L^{-n}\varepsilon)|} \cdot \frac{|\log(L^{-n}\varepsilon)|}{|\log(L^{-n}\varepsilon)|}
\end{aligned}$$

$$\begin{aligned}
n &= 1+\frac{\log(\varepsilon)}{n\log(L)-\log(\varepsilon)}\cdot\frac{|\log(L^{-n}\varepsilon)|}{|\log(L)|}\\
n &= 1+\left(\frac{1}{n}\right)\frac{\log(\varepsilon)}{\log(L)-\frac{1}{n}\log(\varepsilon)}\cdot\frac{|\log(L^{-n}\varepsilon)|}{|\log(L)|}\\
n &= \left(1+O\left(\frac{1}{n}\right)\right)\cdot\frac{|\log(L^{-n}\varepsilon)|}{|\log(L)|}
\end{aligned}$$

Lembre que :

$$\begin{aligned}
\frac{\log S(f,\varepsilon,n)}{n} &\leq \frac{\log b(L^{-n}\varepsilon)}{n}\\
\frac{\log S(f,\varepsilon,n)}{n} &\leq \frac{\log b(L^{-n}\varepsilon)}{\left(1+O\left(\frac{1}{n}\right)\right)\cdot\frac{|\log(L^{-n}\varepsilon)|}{|\log(L)|}}.
\end{aligned}$$

Aplicando limites temos:

$$\begin{aligned}
\lim_{\varepsilon\to 0}\sup\lim_{n\to\infty}\frac{1}{n}\log S(f,\varepsilon,n) &\leq \lim_{\varepsilon\to 0}\sup\lim_{n\to\infty}\frac{\log(L)\cdot\log b(L^{-n}\varepsilon)}{\left(1+O\left(\frac{1}{n}\right)\right)|\log(L^{-n}\varepsilon)|}\\
&\leq \lim_{\varepsilon\to 0}\sup\lim_{n\to\infty}\frac{\log(L)}{\left(1+O\left(\frac{1}{n}\right)\right)}\cdot\lim_{\varepsilon\to 0}\sup\lim_{n\to\infty}\frac{\log b(L^{-n}\varepsilon)}{|\log(L^{-n}\varepsilon)|}\\
h_{top}(f) &\leq \log(L)\cdot D(X).
\end{aligned}$$

$\square$.

No seguinte exemplo faze uma aplicação das funções de Lipschitz e mostrará a dimensão de um retângulo em $\mathbb{R}^2$.

Exemplo 2: Seja Δ um retângulo em $\mathbb{R}^2$ e tome um recobrimento por bolas abertas com a métrica do máximo, onde $x=(x_1,x_2), y=(y_1,y_2)\in\mathbb{R}^2$.

$$d_{\max}(x,y) := \max\{|x_1-x_2|,|y_1-y_2|\}.$$

Tome o recobrimento de cardinalidade $n\times n=n^2$ bolas (ver Figura 2.1). Como as fronteiras de cada bola é aberta, então, precisamos de um novo recobrimento. Basta

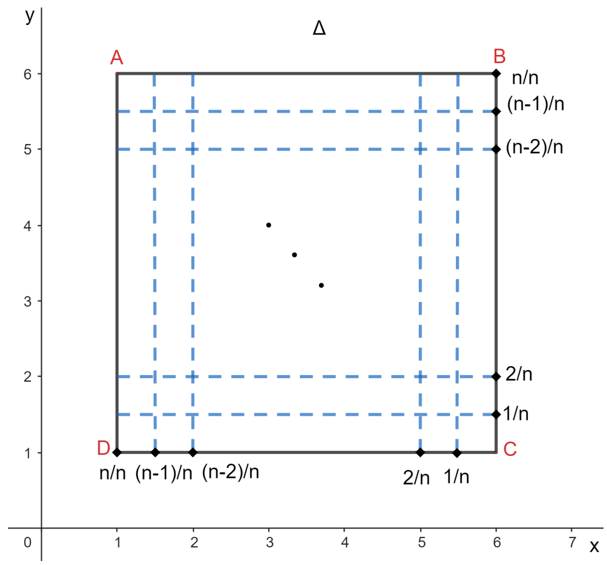

Figura 2.1: Retângulo $\Delta \subset \mathbb{R}^2$.

tomar o dobro do primeiro, logo o cardinal do novo recobrimento é $2n^2$. Assim se pode estabelecer a seguinte desigualdade

$$n^2 \leq b(\varepsilon) \leq 2n^2. \tag{2.20}$$

Por outra parte temos que os raios das bolas tem comprimento $\varepsilon = \dfrac{1}{2n}$, então

$$\begin{aligned}
\frac{\log(b(\varepsilon))}{\log(\varepsilon)} &= \frac{\log(b(\varepsilon))}{\log(\frac{1}{2n})} \\
&= \frac{\log(b(\varepsilon))}{-|\log(\frac{1}{2n})|} \\
&= \frac{\log(b(\varepsilon))}{|\log(2n)|} \\
&= \frac{\log(b(\varepsilon))}{\log(2n)}.
\end{aligned}$$

A última igualdade se tem dado que para n suficientemente grande o $\log(2n) > 0$. Logo aplicando $\limsup_{\varepsilon\to 0}\lim_{n\to\infty}$ em (2.20) temos

$$\limsup_{\varepsilon\to 0}\lim_{n\to\infty}\frac{\log(n^2)}{\log(\varepsilon)} \leq \limsup_{\varepsilon\to 0}\lim_{n\to\infty}\frac{\log(b(\varepsilon))}{\log(\varepsilon)} \leq \limsup_{\varepsilon\to 0}\lim_{n\to\infty}\frac{\log(2n^2)}{\log(\varepsilon)}$$

$$\limsup_{\varepsilon\to 0}\lim_{n\to\infty}\frac{\log(n^2)}{\log(2)+\log(n)} \leq \limsup_{\varepsilon\to 0}\lim_{n\to\infty}\frac{\log(b(\varepsilon))}{\log(\varepsilon)} \leq \limsup_{\varepsilon\to 0}\lim_{n\to\infty}\frac{\log(2n^2)}{\log(2)+\log(n)}.$$

Aplicando limites infinitos temos

$$2 \leq \limsup_{\varepsilon\to 0}\frac{\log(b(\varepsilon))}{\log(\varepsilon)} \leq 2$$

$$2 \leq D(\Delta) \leq 2.$$

Portanto, a $D(\Delta)$ é 2.

2.2 Funções expansivas

Definição 2.2.1. *Um homeomorfismo (correspondentemente uma função contínua) $f : X \to X$ se chama expansivo se existe uma constante $\delta > 0$ tal que se, $d(f^n(x), f^n(y)) < \delta$, para todo $n \in \mathbb{Z}$(correspondentemente, $n \in \mathbb{N}_0$), então, $x = y$.*

O número máximo δ_0 que satisfaz esta propriedade é chamada **constante de expansividade** do sistema dinâmico. Pela compacidade, a propriedade de ser expansivo não depende da escolha particular de uma métrica sobre X que define uma topologia dada, e portanto é um invariante por conjugação topológica. Entretanto, a constante de expansividade depende da escolha da métrica.

Lema 2.2.1. *Seja X um espaço métrico compacto e $f : X \to X$ um homeomorfismo expansivo com constante de expansividade δ_0. Então para $0 < \varepsilon < \dfrac{\delta_0}{2}$ e $\delta > 0$, existe $C_{\varepsilon,\delta}$ tal que para todo $n \in \mathbb{N}$, temos*

$$N_d(f,\delta,n) \leq C_{\delta,\varepsilon} N_d(f,\varepsilon,n).$$

Prova: Para $0 < \varepsilon < \dfrac{\delta_0}{2}$, seja $N \in \mathbb{N}$ e $\alpha > 0$ tal que $d^f_{2N+1}(f^{-N}(x), f^{-N}(y)) < 2\varepsilon \Longrightarrow d(x,y) < \delta$ e $d(x,y) \leq \alpha \Longrightarrow d^f_{2N+1}(f^{-N}(x), f^{-N}(y)) \leq \delta$.

Para mostrar as implicações anteriores é necessário demostrar o seguinte lema:

Lema 2.2.2. *Seja $f : X \to X$ um homeomorfismo δ_0- expansivo. Sejam $0 < \varepsilon < \dfrac{\delta_0}{2}$, $\delta > 0$ então existe $N = N(\varepsilon,\delta); N \geq 0$ tal que $d_{|i|\leq N}(f^i(x), f^i(y)) < 2\varepsilon$, então $d(x,y) < \delta_0$.*

Prova: Como X é compacto então existem subsequências x_{n_k} e y_{n_k} convergentes em X a pontos $\bar{x}$ e $\bar{y}$ respetivamente. Suponha por absurdo que para cada $n \in \mathbb{N}_0$, existe x_n, y_n tais que $d_{|i|\leq N}(f^i(x_n), f^i(y_n)) \leq 2\varepsilon$, então $d(x,y) \geq \delta$. Como X é compacto então existem subsequências x_{n_k}, y_{n_k} convergentes em X a pontos $\bar{x}$ e $\bar{y}$ respetivamente.

Então

$$d(x_{n_k}, y_{n_k}) \geq 2\delta \Longrightarrow d(\bar{x},\bar{y}) \geq 2\delta > 0. \tag{2.21}$$

Assim, $\bar{x} \neq \bar{y}$.

O anterior é possível dado que aplicamos limite quando $k \to \infty$ e a continuidade da função distância.

Escolha $m \in \mathbb{Z}$, $m \geq 0$. Seja k_0 tal que $n_{k_0} \geq |m|$. Considere $k \geq k_0$, temos $d(f^i(x_{n_k}), f^i(y_{n_k})) \leq 2\varepsilon, \forall k \geq k_0$, então

$$\begin{aligned} d(f^i(x_{n_k}), f^i(y_{n_k}) \leq 2\varepsilon \quad &\Longrightarrow \quad d(f^m(x_{n_k}), f^m(y_{n_k}) \leq 2\varepsilon, \forall k \geq k_0 \\ &\Longrightarrow \quad d(f^m(\bar{x}), f^m(\bar{y}) \leq 2\varepsilon < \delta_0. \end{aligned} \tag{2.22}$$

De (2.21) e (2.22) temos que $\bar{x} \neq \bar{y}$ então $d(f^m(\bar{x}), f^m(\bar{y})) < \delta_0, \forall m \in \mathbb{Z}$ o que é um absurdo! Logo f não é expansivo.

□

Portanto,

$$d^f_{2N+1}(f^{-N}(x), f^{-N}(y)) \leq 2\varepsilon \Longrightarrow d(x,y) < \delta \text{ e } d(x,y) \leq \alpha \Longrightarrow d^f_{2N+1}(f^{-N}(x), f^{-N}(y)) < \delta.$$

A existência de N é dada pela expansividade e a existência de α é dada pela continuidade uniforme.

Seja E um conjunto $(n, \varepsilon) - separado$ maximal e F um conjunto $(n, \delta) - separado$ maximal. Para todo $x \in E$ existe $z(x) \in F$ tal que $d_n^f(x, z(x)) < \varepsilon$.

De fato, se não acontece isto, então, para um certo $\bar{x} \in E$ ocorre que $d_n^f(\bar{x}, z(\bar{x})) \geq \varepsilon, \forall z \in F \Longrightarrow F \cup \{\bar{x}\}$ é $(n, \varepsilon) - separado$ o que é um absurdo, pois contradiz a maximalidade de F.

Portanto, $Card(E) \leq \sum_{z \in F} Card(E_z)$, onde $E_z := \{x \in E \mid z(x) = z\}$ e a afirmação se deduz uma vez que encontremos um limite em $Card(E_z)$ que não dependa de n.

Observe que, para $n > 2N$ vale que $N < n - N < n$, então $n > 2N$. Para $x, y \in E_z$ tem-se que $d_n^f(x, y) \leq 2\varepsilon$ pela definição de E_z. Portanto, $d(f^i(x), f^i(y)) \leq \delta$, para $i \in [N, n-N]$ pela escolha de N e portanto, pela escolha de α e dado que $\{x, y\}$ é (n, δ)- separado, $d(x, y) > \alpha$ ou $d(f^n(x), f^n(y)) > \alpha$.

Portanto,

$$
\begin{aligned}
card(E_z) &= card\{(x, f^n(x)) | x \in E_z\} \\
&\leq \max\{cardA | A \subset X \times X \ e \ \forall \ a, b \in A, d(a, b) > \alpha\} := M.
\end{aligned}
$$

desde que o $(x, f^n(x))$ forma exatamente um conjunto α-separado.

$\square$

Corolário 2.2.1. *Se $f : X \to X$ é expansivo com constante de expansividade δ_0, e $0 < \delta < \dfrac{\delta_0}{2}$, então $h_d(f, \delta) = h_{top}(f)$.*

Prova: Sejam $\varepsilon, \delta > 0$ tal que $0 < \delta < \dfrac{\delta_0}{2}$ onde δ_0 é a constante de expansividade. Existem $C^1_{\delta,\varepsilon}, C^2_{\delta,\varepsilon}$, constantes. Aplicando o Lema (2.2.1), temos que $0 < \delta < \dfrac{\delta_0}{2}$, então existe $C^1_{\delta,\varepsilon}$ tal que:

$$N_d(f, \delta, n) \leq C^1_{\delta,\varepsilon} N_d(f, \varepsilon, n).$$

Logo aplicando definição de entropia temos:

$$
\begin{aligned}
\lim_{n\to\infty}\sup\frac{1}{n}\log(N_d(f,\delta,n)) &\leq \lim_{n\to\infty}\sup\frac{1}{n}\log(C^1_{\delta,\varepsilon}N_d(f,\varepsilon,n)) \\
h_d(f,\delta) &\leq h_d(f,\varepsilon)
\end{aligned}
$$

De maneira análoga e usando $C^2_{\delta,\varepsilon}$ obtemos

$$h_d(f,\varepsilon) \leq h_d(f,\delta).$$

Revelando que $h_d(f,\varepsilon) = h_d(f,\delta)$. Finalmente,

$$
\begin{aligned}
\lim_{\varepsilon\to\ 0} h_d(f,\varepsilon) &= \lim_{\varepsilon\to\ 0} h_d(f,\delta) \\
h_{top}(f) &= h(f,\delta).
\end{aligned}
$$

□

Proposição 2.2.1. *$p(f) \leq h_{top}(f)$ para um homeomorfismo expansivo f de um espaço métrico compacto.*

Prova: Seja δ_0 a constante de expansividade, logo $Fix(f^n)$ é $(n,\delta_0)-separado$, para $n \in \mathbb{N}; x \neq y \in Fix(f^n)$.

Seja $S\ :\ \max\{d(f^i(x),f^i(y))\ |\ 0 \leq i < n\}; x \neq y$. Observe que $\delta > \delta_0$, pois $d(f^i(x),f^i(y)) \geq \delta_0$ porém dado que δ é o máximo $d(f^i(x),f^i(y)) \leq \delta$.

Seja $\varepsilon > 0$ tal que $\delta_0 \geq \varepsilon$ então temos que $d(f^i(x),f^i(y)) \geq \varepsilon, x \neq y$, logo temos um conjunto $N_d(f,n,\varepsilon) - separado$. Assim, temos que

$$N(f,n,\delta_0) \leq N(f,n,\varepsilon).$$

aplicando a definição de entropia concluímos que:

$$p(f) \leq h_{top}(f).$$

□

Proposição 2.2.2. *Se $f : X \to X$ é expansivo, onde δ_0 é a constante de expansividade. $h : X \to X$ uma função contínua tal que $d(x, h(x)) < \delta_0, \forall x$ e $f \circ h = h \circ f$; então h é igual a função identidade.*

Prova: Considere que existem f e h como no enunciado , mas que $h(x) = y \neq x$ ou seja h não é a identidade. Pela definição de expansividade temos então que, $d(f^n(x), f^n(y)) \geq \delta_0$.

Logo temos que: $f^n(y) = f^n(h(x)) = h(f^n(x))$, dado que $f \circ h = h \circ f$. Chamaremos $f^n(x) = z$, então $d(z, h(z)) \geq \delta_0$, isto contradiz que f é expansiva. Revelando que, h é a identidade.

□

Teorema 2.2.1 (Teorema de Misiurewicz-Przytycki). *Se M é uma variedade suave compacta orientável e $f : M \to M$ de classe C^1, então $h_{top}(f) \geq \log |grau(f)|$.*

Algumas definições antes de iniciar a demostração.

Definição 2.2.2 (Grau de f). *Seja M um conjunto compacto e conexo. Seja $f : M \to M$ uma aplicação de recobrimento, então dado $y \in M$ temos que:*

$$\# f^{-1}(\{y\}) = n, \forall y \in M. \textit{Definimos } n = grau(f).$$

Definição 2.2.3 (Ponto Regular). *Seja $f : N \to M$ de classe C^1,onde M e N são variedades suaves. Um ponto $x \in N$ é chamado um **ponto regular** (para f) se D_f é invertível em x. $x \in X$ é chamado **valor regular** (de f) se $f^{-1}(\{x\})$ consistir em pontos regulares e caso contrario, um valor singular.*

Outra forma de definir o **grau de f**, é:

Um elemento ω de volume positivo em M é uma $n - forma$ contínua. Se diz que ω esta normalizado se:

$$\int_M \omega = 1$$

Pode-se mostrar que $\displaystyle\int f^* \omega = k$ onde $k = grau_x(f); x$ é regular.

As definições de grau com mais detalhes estão nas paginas 310-313 de nossa referência [1].

Definição 2.2.4. ***(Teorema de Sard).*** *Se $f \in C^1(M,N)$, então o conjunto de valores singulares de f tem medida de Lebesgue zero.*

Prova: Teorema de Misiurewicz-Przytycki

Seja ω uma forma de volumem em M, $\alpha \in (0,1)$ arbitrário . Seja

$$L := \sup_{x \in X} Jf(x) < \infty, \text{ pois } f \text{ é de classe } C^1.$$

Agora sejam $\varepsilon = L^{-\alpha/\alpha-1}$ e $B := \{x \mid |Jf(x)| \geq \varepsilon\}$. Cubra B por abertos, sobre os quais f é injetiva e tome δ o número de Lebesgue da cobertura.

Fixemos $n \in \mathbb{N}$ e seja

$$A = \{x \in M \mid \#(B \cap \{x, f(x), f^2(x), ..., f^n(x)\}) \leq \alpha n\}.$$

Note que, se $x \in A$,

$$Jf^n(x) = \prod_{j=0}^{n-1} Jf(f^i(x)) < \varepsilon^{(1-\alpha)n} L^{\alpha n} = 1.$$

Segue também que, se $x \in A$, então,

$$\#\big(\{x, f(x), f^2(x), ..., f^{n-1}(x)\} \cap B^c\big) > (1-\alpha)n. \tag{2.23}$$

onde $B^c = \{x \in M \mid |Jf(x)| < \varepsilon\}$.

Por (2.22) podemos selecionar $(1-\alpha)n$ valores de i tais que $Jf(f^i(x)) < \varepsilon$, e para os outros valores αn de i, $Jf(f^i(x)) < L$.

Portanto, o volume $f^n(A)$ é menor que o volume de M e pelo **Teorema de Sard** (2.2.4), existe um valor regular $x \in M \ \setminus f^n(A)$ de f^n. Seja S o conjunto dos pontos singulares de f, então $vol(S) = 0$ e como f é C^1 então $vol(f^i(S)) = 0$, portanto, $vol\big(\bigcup_{i=1}^{\infty} f^i(S)\big) = 0$. Como x é ponto regular de f^n, então x é regular em f, esta afirmação é dada pela aplicação da regra da cadeia.

Dado $f^{-n}(\{x\})$ escolhemos o maior subconjunto $(n,\delta) - separado$.

Uma vez que x é regular, então $\#f^{-1}(\{x\})$ é pelo menos N, onde $N = grau(f)$ ou seja, N é o número de pré-imagens. Se N pré-imagens de x estão em B, tomemos Q_1 exatamente N pre-imagens de x em B (transição boa). Caso contrário tome Q_1 como uma única pré-imagem fora de B (transição ruins).

Afirmação: Q_n é $(n,\delta)-$*separado*.

Prova: Sejam $y_1, y_2 \in Q_n$ tais que $d(f^i(y_1), f^i(y_2)) \leq \delta; i = 0, ..., n-1$. Onde δ é o número de Lebesgue

Tomando, $i = n-1$ temos:

$$\begin{aligned} z_1 &= f^{n-1}(y_1) \\ z_2 &= f^{n-1}(y_2) \end{aligned}$$

1. Se $z_1, z_2 \in B^c$, então $z_1 = z_2$ e $f^{n-1}(y_1) = f^{n-1}(y_2)$.

2. se $z_1, z_2 \in B$ e $d(z_1, z_2) < \delta$, $f(z_1) = f(z_2) = x$, pois como f é injetiva em B, então $z_1 = z_2$.

Em qualquer caso para a escolha de quaisquer z_n com $n \in \mathbb{N}$, se tem que $z_1 = z_2 = z_3 = ... = z_n = z$ Agora bem, se tomamos $i = n-2$ temos:

$$\begin{aligned} w_1 &= f^{n-2}(y_1) \\ w_2 &= f^{n-2}(y_2). \end{aligned}$$

Pelo mesmo raciocínio conclui-se que $w_1 = w_2$.

Aplicando recursivamente este argumento, concluímos que $y_1 = y_2$. Ou seja, se $y_1 \neq y_2$ então $d(f^i(y_1), f^i(y_2)) \geq \delta$; para algum $i = 0, ..., n-1$, portanto, Q_n é $(n,\delta)-$*separado*.

$\square$

Agora, $Q_n \subset f^{-n}(\{x\}) \subset f^{-n}(M \setminus f^n(A)) \subset M \setminus A$, logo, $Q_n \cap A = \emptyset$, ou seja $Q_n \subset A^c$.

Por definição existem no máximo de αn valores de k, onde $k \in \{0, ..., n-1\}$ tais que $f^k(y) \in B$, para passar de x a $y \in Q_n$, existem pelo menos $m := \alpha n + 1$ "boas transições", pois as "transições ruins"passan fora de B.

Finalmente, $\#Q_n \geq N^m \geq N^{\alpha n}$ implicando que $N_d(f, \delta, n) \geq N^{\alpha n}$, logo

$$\begin{aligned}
\lim_{n\to\infty}\sup\frac{1}{n}\log N_d(f,\delta,n) &\geq \lim_{n\to\infty}\sup\frac{1}{n}\log N^{\alpha n}\\
h_{top}(f) &\geq \lim_{n\to\infty}\sup\frac{1}{n}\log \alpha n N\\
&\geq \alpha\log N\\
&\geq \log N\\
&= \log|grau(f)|. \qquad \square
\end{aligned}$$

Exemplo 3. Seja $f:\overline{\mathbb{C}}\to\overline{\mathbb{C}}$ uma função racional da esfera de Riemann. Seja $f(z)=P(z)/Q(z)$ onde P,Q polinômios primos relativos. Assim $\overline{\mathbb{C}}\cong S^2$, então $f:S^2\to S^2$, onde f é C^1 e S^2 compacta.

Nestas condições o $grau(R_w(z))=\max(grauP, grauQ)$. Agora, considere a equação

$$\frac{P(z)}{Q(z)}=w.$$

Escolha w tal que $R(w)=0$ não tenha raízes múltiplas.

Logo

$$\begin{aligned}
f(z) &= w\\
\frac{P(z)}{Q(z)} &= w\\
P(z) &= Q(z)w\\
P(z)-Q(z)w &= 0.
\end{aligned}$$

Logo $R_w(z) = P(z) - wQ(z) = 0$. Agora temos que escolher w tal que o termo líder de $wQ(z)$ não anule o termo líder de P (os w que satisfazem isso tem medida total).

Afirmação: w é tal que $R_w(z) = 0$ tem raiz múltipla, então w é imagens de um valor singular de f; $w = f(z_0)$ onde z_0 é valor singular de f ou seja $f'(z_0) = 0$.

Prova: suponha $R_w(z_0) = 0$ e z_0 uma raiz múltipla. Agora tomando

$$\begin{cases} P(z_0) - wQ(z_0) & = 0 \quad (1) \\ P'(z_0) - wQ'(z_0) & = 0. \quad (2) \end{cases}$$

Logo, multiplicando a equação (1) e (2) por $Q'(z_0)$ e $Q(z_0)$ respectivamente

$$\begin{cases} P(z_0) - wQ(z_0) & = 0(Q'(z_0)) \\ P'(z_0) - wQ'(z_0) & = 0(Q(z_0)) \end{cases}$$

fazendo a substração se tem,

$$P'(z_0)Q(z_0) - P(z_0)Q'(z_0) = 0.$$

Então $\left(\dfrac{P}{Q}\right)'(z_0) = 0$, ou seja, $f'(z_0) = 0$. Do anterior temos que

$$P(z_0) - wQ(z_0) = 0 \Longrightarrow \left(\frac{P}{Q}\right)(z_0) = w \Longrightarrow f(z_0) = w$$

mostrando que de fato w é imagem de um valor singular. Portanto, o grau de f é igual ao máximo dos graus algébricos de P e Q. Pelo Teorema (2.2.1) temos que

$$h_{top}(f) \geq \log \max(grauP, grauQ).$$

De fato, a desigualdade inversa $h_{top}(f) \geq \log \max(grauP, grauQ)$ também se tem, logo concluímos que

$$h_{top}(f) = \log \max(grauP, grauQ).$$

Definição 2.2.5. *Uma função contínua $f : M \to M$ é chamada expansora se para algum $\mu > 1$, $\varepsilon_0 > 0$ e cada $x, y \in X$ com $x \neq y$ e $d(x, y) < \varepsilon_0$ se tem $d(f(x), f(y)) > \mu d(x, y)$.*

Exemplo 4. Seja M um variedade riemanniana compacta e $f : M \to M$ uma função expansora de classe C^1. Mostrar que:

$$h_{top}(f) = \log |grau(f)|.$$

Prova: Seja $f : M \to M$, C^1, expansora. Tomemos $D = |grau(f)| = n$ de pre-imagens de qualquer ponto. Pelo Teorema (2.2.1) temos que

$$h_{top}(f) \geq \log |grau(f)|. \tag{2.24}$$

Então, somente temos que mostrar que: $h_{top}(f) \leq \log |grau(f)|$.

Seja $\varepsilon > 0, n > 0$ inteiro, dado $S_0 \in M$ um conjunto $(0, \varepsilon)-$gerador ou seja dado $y \in M$, $\exists\ x \in S_0$ tal que $d(x, y) < \varepsilon$.

Seja $S_n = f^{-n}(S_0)$.

Afirmação: S_n é um conjunto $(n, \varepsilon)-$gerador.

Prova da afirmação: Como f^n é sobrejetora, dado $y \in M$, existe x tal que $f^n(x) = y$.

Seja $z \in S_0$ tal que $d(z, y) < \varepsilon$, então existe $w \in S_n$ tal que $f^n(w) = z$. Portanto, $d(f^n(w), f^n(x)) < \varepsilon$.

Usando a hipótese que f é expansora, temos que:

- Seja $\lambda > 1$ o fator de expansão de f, existe $w_{n-1} \in f^{-1}(f^n(w))$ tal que $d(w_{n-1}, f^{n-1}(x)) < \dfrac{\varepsilon}{\lambda}$.

- Seja $\lambda > 1$ o fator de expansão de f, existe $w_{n-2} \in f^{-2}(f^n(w))$ tal que $d(w_{n-2}, f^{n-2}(x)) < \dfrac{\varepsilon}{\lambda^2}$.

  Continuando de maneira indutiva temos,

- Seja $\lambda > 1$ o fator de expansão de f, existe $w_{n-n} = w_0 \in f^{-n}(z)$ tal que $d(w_0, x) < \dfrac{\varepsilon}{\lambda^n} < \varepsilon$.

Logo assim, S_n é $(n, \varepsilon)-$ gerador, mostrando a prova da afirmação.

Seja α o cardinal de conjuntos $(0, \varepsilon)-$gerador de S_0, então o cardinal de $S_n = D^n\alpha$. Logo tomando $M(n, \varepsilon)$=min$\{\#K | K$ é $(0,\ \varepsilon)$- gerador$\} \leq D^n\alpha$.

Agora aplicando a definição de entropia

$$
\begin{aligned}
\lim_{\varepsilon\to 0}\lim_{n\to\infty}\sup\frac{1}{n}\log M(n,\varepsilon) &\leq \lim_{\varepsilon\to 0}\lim_{n\to\infty}\sup\frac{1}{n}\log(D^n\alpha)\\
h_{top}(f) &\leq \lim_{\varepsilon\to 0}\lim_{n\to\infty}\sup\frac{1}{n}\log(D^n)+\lim_{\varepsilon\to 0}\lim_{n\to\infty}\sup\frac{1}{n}\log(\alpha)\\
&= \lim_{\varepsilon\to 0}\lim_{n\to\infty}\sup\frac{1}{n}\log(D^n)\\
&= \lim_{\varepsilon\to 0}\lim_{n\to\infty}\sup\frac{1}{n}n\log(D)\\
&= \lim_{\varepsilon\to 0}\lim_{n\to\infty}\sup\log(D)\\
&= \log(D)\\
&= \log|grau(f)|.
\end{aligned}
$$

Assim obtemos que:

$$h_{top}(f)\leq \log|grau(f)|. \tag{2.25}$$

Logo de (2.24) e (2.25) obtemos o desejado.

Capítulo 3

Entropia Métrica

No Capítulo 3, definiremos inicialmente que uma partição de X, é uma coleção finita de conjuntos (ζ) essencialmente disjuntos mensuráveis C_i (chamados elementos de ζ), cuja união cobre todo X a menos de medida zero. Dizemos que uma partição ζ' é um refinamento de ζ e escrevemos $\zeta \leq \zeta'$ (ou $\zeta' \leq \zeta$).

- **Entropia métrica**

Seja (X, A, μ) um espaço mensurável com $\mu(X) = 1$.

Definição 3.0.1. *Dizemos que uma família finita $\xi \subset A$ é um partição de X (com respeito de μ) se:*

1. $\mu(\bigcup_{C \in \xi} C) = 1.$
2. $\mu(C \cap D) = 0$ *para quaisquer* $C, D \in \xi$ *com* $D \neq C$.

Definição 3.0.2. *A entropia de uma partição mensurável ξ de X (com respeito de μ) é dada por:*

$$H_\mu(\xi) = -\sum_{C \in \xi} \mu(C) \log \mu(C). \tag{3.1}$$

com a convenção que $0 \log(0) = 0$.

Pois se $f(x) = x\log x$, então,

$$\lim_{x\to 0} x\log x \quad = \lim_{x\to 0}\frac{\log x}{\frac{1}{x}}.$$

Dado que obtemos uma indeterminação, usando a regra de L'Hopital, obtendo assim;

$$\lim_{x\to 0}[-x] = 0.$$

Obtendo a convenção.

Exemplo 2: Considere o espaço $X = [0,1]$ com a medida de Lebesgue μ. Ademais, para cada $k \in \mathbb{N}$, considere a partição mensurável,

$$[0,1) = \{C_j := \left[\frac{j}{k}, \frac{(j+1)}{k}\right); j = 0, ..., k-1\}$$

de X, que consiste de k intervalos de igual medida $\mu(C_j) = \dfrac{1}{k}$.

Temos:

$$\begin{aligned}
H_\mu(\xi_k) &= -\sum_{j=0}^{k-1}\mu(C_j)\log\mu(C_j) \\
&= -\sum_{j=0}^{k-1}\frac{1}{k}\log\big(\frac{1}{k}\big) \\
&= -\sum_{j=0}^{k-1}\frac{1}{k}(\log(1)-\log(k)) \\
&= -\big(k\frac{1}{k}(\log(1)-\log(k)\big) \\
&= -(-\log(k)) \\
&= \log(k).
\end{aligned}$$

Seja $\psi : [0,1] \to \mathbb{R}$ uma função contínua definida por:

$$\psi(x) = \begin{cases} -x\log x & ; se \quad 0 < x \leq 1 \\ 0 & ; se \quad x = 0. \end{cases} \tag{3.2}$$

Calculando a segunda derivada obtemos: $\psi'' = -\frac{1}{x} < 0$, para $x > 0$, então a função ψ é estritamente concava e portanto;

$$\sum_{i=1}^{p} a_i \psi(x_i) \leq \psi\Big(\sum_{i=1}^{p} a_i x_i\Big). \tag{3.3}$$

Para quaisquer número $x_1, ..., x_p, a_1, ..., a_p \in [0,1]$ tal que $\sum_{i=1}^{p} a_i = 1$.

Em particular, a desigualdade anterior permite dar um limite superior para a entropia em termos da cardinalidade de ξ. Se $k = card\ \xi$, obtemos:

$$\begin{aligned}
H_\mu(\xi) &= k \sum_{C \in \xi} \frac{1}{k} \psi\big(\mu(C)\big) \\
&\leq k\psi\Big(\sum_{C \in \xi} \frac{\mu(C)}{k}\Big) \\
&= k\psi\Big(\sum_{i=1}^{\infty} \frac{\mu(C_i)}{k}\Big) \\
&= k\psi\Big[\frac{1}{k}\Big(\sum_{i=1}^{\infty} \mu(C_i)\Big)\Big] \\
&= k\psi\Big[\frac{1}{k}\Big(\mu\big(\bigcup_{i=1}^{\infty} C_i\big)\Big)\Big] \\
&= k\psi\Big[\frac{1}{k}(1)\Big] \\
&= \log k \\
&= \log(card\ \xi).
\end{aligned}$$

Finalmente, $H_\mu(\xi) \leq \log(card\ \xi)$.

Definição 3.0.3. *Sejam duas partições mensuráveis ξ e η de X, dizemos que η é um* ***refinamento*** *de ξ se para cada $D \in \eta$ existe $C \in \xi$ tal que $\mu(D \setminus C) = 0$.*

Proposição 3.0.1. *Seja ξ e η partições mensuráveis de X. Se η é um refinamento de ξ, então:*

$$H_\mu(\xi) \leq H_\mu(\eta). \tag{3.4}$$

Prova: Como η é um refinamento, podemos escrever:

$$H_\mu(\eta) = \sum_{D \in \eta} \psi(\mu(D)).$$

Então dado tal refinamento, temos que $D \subset C$ logo escrevemos

$$H_\mu(\eta) = \sum_{C \in \xi} \sum_{D \subset C} \psi(\mu(D))$$

Para cada $C \in \xi$, temos

$$\sum_{D \subset C} (\mu(D)) = \mu(C)$$

$$\begin{aligned}
\psi(\mu(C)) &= \psi\left(\sum_{D \subset C} (\mu(D))\right) \\
&= -\sum_{D \subset C} \mu(D) \log \sum_{D \subset C} \mu(D) \\
&\leq -\sum_{D \subset C} \mu(D) \log \mu(D) \\
&= \sum_{D \subset C} \psi(\mu(D)).
\end{aligned}$$

Logo temos

$$\psi(\mu(C)) \leq \sum_{D \subset C} \psi(\mu(D)).$$

Finalmente aplicando somatório as desigualdades temos

$$\begin{aligned} \sum_{C \in \xi} \psi(\mu(C)) &\leq \sum_{C \in \xi} \sum_{D \subset C} \psi(\mu(D)) \\ H_\mu(\xi) &\leq H_\mu(\eta). \end{aligned}$$

Mostrando o desejado.

□

Dadas duas partições mensuráveis ξ e η de X, definimos a partição mensurável:

$$\xi \vee \eta = \{C \cap D : C \in \xi, D \in \eta\}. \tag{3.5}$$

Podemos supor que os elementos da partição $\xi \vee \eta$ tem medida positiva.

Esta partição é um refinamento de ambas partições e $\xi \vee \eta$ é a partição mais pequena em termos de cardinalidades.

Proposição 3.0.2. *Se $T : X \longrightarrow X$ é uma transformação mensurável que preserva a medida de probabilidade μ em X e ξ é uma partição mensurável de X, então:*

$$\lim_{n\to\infty} \frac{1}{n} H_\mu(\xi_n) = \inf_{n\in\mathbb{N}} \frac{1}{n} H_\mu(\xi_n) \tag{3.6}$$

existe, quando;

$$\xi_n = \bigvee_{k=0}^{n-1} T^{-k}(\xi) = \xi \vee T^{-1}(\xi) \vee ... \vee T^{-(n-1)}(\xi), \tag{3.7}$$

para todo $n \in \mathbb{N}$.

Para mostrar esta proposição é preciso provar os seguintes lemas:

Lema 3.0.1. *Se ξ e η são partições mensuráveis de X, então*

$$H_\mu(\xi \vee \eta) \leq H_\mu(\xi) + H_\mu(\eta). \tag{3.8}$$

Prova:

$$\begin{aligned}
H_\mu(\xi \vee \eta) &= \sum_{C\in\xi, D\in\eta} \psi(\mu(C\cap D)) \\
&= -\sum_{C\in\xi, D\in\eta} \mu(C\cap D)\log \mu(C\cap D) \\
&= -\sum_{C\in\xi, D\in\eta} \mu(C\cap D)\log\left(\mu(C\cap D)\frac{\mu(C)}{\mu(C)}\right) \\
&= -\sum_{C\in\xi, D\in\eta} \mu(C\cap D)\left(\log\frac{\mu(C\cap D)}{\mu(C)} + \log\mu(C)\right) \\
&= -\sum_{C\in\xi, D\in\eta} \mu(C\cap D)\log\left(\frac{\mu(C\cap D)}{\mu(C)}\right) - \sum_{C\in\xi, D\in\eta} \mu(C\cap D)\log\mu(C) \\
&= \sum_{C\in\xi, D\in\eta} \mu(C)\psi\left(\frac{\mu(C\cap D)}{\mu(C)}\right) - \sum_{C\in\xi, D\in\eta} \mu(C\cap D)\log\mu(C) \\
&= \sum_{D\in\eta}\sum_{C\in\xi} \mu(C)\psi\left(\frac{\mu(C\cap D)}{\mu(C)}\right) - \sum_{C\in\xi}\sum_{D\in\eta} \mu(C\cap D)\log\mu(C).
\end{aligned}$$

Fixemos $D \in \eta$ no primeiro termo e fixemos também $C \in \xi$ no segundo termo

$$\begin{aligned}
&\leq \sum_{D\in\eta} \psi\left(\sum_{C\in\xi} \mu(C)\frac{\mu(C\cap D)}{\mu(C)}\right) - \sum_{C\in\xi} \mu(C)\log\mu(C) \\
&= \sum_{D\in\eta} \psi\left(\sum_{C\in\xi} \mu(C\cap D)\right) - \sum_{C\in\xi} \mu(C)\log\mu(C)) \\
&= \sum_{D\in\eta} \psi(\mu(D)) - \sum_{C\in\xi} \psi(\mu(C).
\end{aligned}$$

Mostrando finalmente que: $H_\mu(\xi \vee \eta) \leq H_\mu(\eta) + H_\mu(\xi)$.

Agora consideremos a partição mensurável

$$T^{-n}(\xi) = \left\{T^{-n}C : C \in \xi\right\}. \tag{3.9}$$

Lema 3.0.2. *Se ξ é uma partição mensurável de X, então*

$$H_\mu(T^{-n}\xi) = H_\mu(\xi), \tag{3.10}$$

para todo $n \in \mathbb{N}$.

Prova: A prova sera feita por indução.

- Para $n = 1$.

$$\begin{aligned}
H_\mu(T^{-1}(\xi)) &= \sum_{C\in\xi} \psi(\mu(T^{-1}(C))) \\
&= -\sum_{C\in\xi} \mu(T^{-1}C)\log\mu(T^{-1}C) \\
&= -\sum_{C\in\xi} \mu(C)\log\mu(C); \text{ pela invariância de } T. \\
&= -\sum_{C\in\xi} \psi(\mu(C)) \\
&= H_\mu(\xi).
\end{aligned}$$

- Suponha válido para $n = k$, então temos que:

$$H_\mu(T^{-k}(\xi)) = H_\mu(\xi).$$

- Provaremos para $n = k + 1$.

$$
\begin{aligned}
H_\mu(T^{-1}(\xi)) &= \sum_{C\in\xi} \psi(\mu(T^{-(k+1)}(\xi))) \\
&= -\sum_{C\in\xi} \mu(T^{-k-1}(\xi)) \log(\mu(T^{-k-1}(\xi))) \\
&= -\sum_{C\in\xi} \mu(T^{-1}(T^{-k}(\xi))) \log(\mu(T^{-1}(T^{-k}(\xi)))) \\
&= -\sum_{C\in\xi} \mu(T^{-k}(\xi)) \log(\mu(T^{-k}(\xi))); \text{ pela invarianza de T.} \\
&= \sum_{C\in\xi} \psi(\mu(T^{-k}(\xi))) \\
&= H_\mu(T^{-k}(\xi)) \\
&= H_\mu(\xi).
\end{aligned}
$$

O que conclui a prova do lema.

Agora vamos realizar uma observação mais profunda sobre as partições da forma descrita em (3.7).

Dados $n, m \in \mathbb{N}$, temos

$$\xi_{n+m} = \xi_n \vee T^{-n}\xi_m.$$

Podemos escrever;

$$\begin{aligned}
\xi_{n+m} &= \left\{\bigcap_{j=0}^{n+m-1} T^{-j}C_j : C_j \in \xi \textit{ para cada } j.\right\}\\
&= \left\{\left(\bigcap_{j=0}^{n-1} T^{-j}C_j\right)\cap\left(\bigcap_{j=n}^{n+m-1} T^{-j}C_j\right): C_j \in \xi \textit{ para cada } j\right\}\\
&= \left\{\left(\bigcap_{j=0}^{n-1} T^{-j}C_j\right)\cap T^{-n}\left(\bigcap_{k=0}^{m-1} T^{-k}D_k\right): C_j, D_k \in \xi \textit{ para cada } j,k\right\}\\
&= \xi_n \vee T^{-n}\xi_m.
\end{aligned}$$

Usando as observações e os lemas anteriores, obtemos,

$$\begin{aligned}
H_\mu(\xi_{n+m}) &= H_\mu(\xi_n \vee T^{-n}(\xi_n))\\
&= H_\mu(\xi_n) + H_\mu(T^{-n}(\xi_m))\\
&= H_\mu(\xi_n) + H_\mu(\xi_n), \text{ para todo } n, m \in \mathbb{N}.
\end{aligned}$$

Definição 3.0.4. *Seja $T : X \longrightarrow X$ é uma transformação mensurável que preserva a medida de probabilidade μ em X. Definimos a **Entropia Métrica** de T com respeito a μ e a partição mensurável ξ de X por:*

$$h_\mu(T,\xi) = \lim_{n\to\infty} \frac{1}{n} H_\mu(\xi_n), \tag{3.11}$$

com ξ_n como em (3.7). Finalmente, definimos a entropia métrica de T com respeito de μ por

$$h_\mu(T) = \sup\{h_\mu(T,\xi) : \xi \textit{ é uma partição mensurável de } X\}.$$

Proposição 3.0.3. *Temos $h_\mu(T^k) = kh_\mu(T)$ para cada $k \in \mathbb{N}$.*

Prova: Primeiro mostraremos que $h_\mu(T^k) \leq kh_\mu(T)$, para cada $k \in \mathbb{N}$. É claro que

$\bigvee_{i=0}^{nk-1} T^{-i}\xi$ é um refinamento de $\bigvee_{i=0}^{n-1} T^{-ki}\xi$. Portanto, pela Proposição (3.0.1) temos,

$$H_\mu\Big(\bigvee_{i=0}^{n-1} T^{-ki}\xi\Big) \;\leq\; H_\mu\Big(\bigvee_{i=0}^{nk-1} T^{-i}\xi\Big)$$

$$\lim_{n\to\infty}\frac{1}{n}H_\mu\Big(\bigvee_{i=0}^{n-1} T^{-ki}\xi\Big) \;\leq\; \lim_{n\to\infty}\frac{1}{n}H_\mu\Big(\bigvee_{i=0}^{nk-1} T^{-i}\xi\Big).$$

Por (3.7) temos que,

$$\lim_{n\to\infty}\frac{1}{n}H_\mu\Big(\xi_n\Big) \;\leq\; \lim_{n\to\infty}\frac{1}{nk}H_\mu\Big(\bigvee_{i=0}^{nk-1} T^{-i}\xi\Big).$$

Pela Definição (3.0.4) temos;

$$h_\mu(T^k,\xi) \;\leq\; \lim_{n\to\infty}\frac{1}{n}H_\mu\Big(\bigvee_{i=0}^{nk-1} T^{-i}\xi\Big).$$

Da expressão do lado direito temos nk iterações, implicando,

$$h_\mu(T^k,\xi) \leq \lim_{n\to\infty}\frac{1}{n}H_\mu\Big(\bigvee_{i=0}^{n-1} T^{-i}\xi\Big) = kh_\mu(T,\xi).$$

Assim,

$$h_\mu(T^k,\xi) \leq kh_\mu(T,\xi).$$

Tomando o sup em ξ, implica que $h_\mu(T^k) \leq kh_\mu(T,\xi)$, provando o desejado.

Agora provaremos que $h_\mu(T^k) \geq kh_\mu(T,\xi)$.

Observe que $\bigvee_{\ell=0}^{n-1} T^{-k\ell}\xi_k = \bigvee_{i=0}^{nk-1} T^{-i}\xi$, logo fazendo $n\to\infty$,

$$\begin{aligned}
\lim_{n\to\infty}\frac{1}{n}H_\mu\left(\bigvee_{i=0}^{n-1}T^{-k\ell}\xi_k\right) &= \lim_{n\to\infty}\frac{1}{n}H_\mu\left(\bigvee_{i=0}^{nk-1}T^{-i}\xi\right)\\
h_\mu(T^k,\xi_k) &= kh_\mu(T,\xi)\\
h_\mu(T^k) &= \sup_\xi h_\mu(T^k,\xi)\\
&\geq \sup_\xi h_\mu(T^k,\xi_k)\\
&= k\sup_\xi h_\mu(T,\xi)\\
&= kh_\mu(T).
\end{aligned}$$

Mostrando que $h_\mu(T^k) \geq kh_\mu(T,\xi)$.

Portanto,

$$h_\mu(T^k) = kh_\mu(T,\xi).$$

$\square$

Capítulo 4

Entropia Condicional

A ideia neste capítulo é dar a conhecer o concito de entropia condicional. Com a noção que a entropia condicional é a entropia de uma partição com respeito de uma medida de probabilidade da partição digamos ξ de X condicionada a outra partição de X digamos η, se pode estabelecer uma caracterização da entropia métrica. Estas noções sará usadas no próximo capitulo.

Definição 4.0.1. *Dadas duas partições mensuráveis ξ e η de X, definimos a* ***entropia condicional*** *de ξ com respeito de η por:*

$$H_\mu(\xi|\eta) = -\sum_{C\in\xi, D\in\eta} \mu(C\cap D)\log\frac{\mu(C\cap D)}{\mu(D)}. \tag{4.1}$$

A noção de entropia condicional pode ser usada para dar a seguinte caracterização da entropia métrica.

Teorema 4.0.1. *Se $T : X \to X$ é uma transformação mensurável que preserva a medida de probabilidade μ em X e ξ é uma partição mensurável de X então:*

$$h_\mu(T,\xi) = \lim_{n\to\infty} H_\mu(\xi|\bigvee_{i=0}^{n-1} T^{-i}(\xi)). \tag{4.2}$$

Para estabelecer a existência do limite em (4.2), usaremos os seguintes resultados auxiliares mostrando que a sequência $H_\mu(\xi|\bigvee_{i=0}^{n-1} T^{-i}(\xi))$ é não decrescente.

Lema 4.0.1. *Se η é um refinamento de ζ, então*

$$H_\mu(\xi|\eta) \leq H_\mu(\xi|\zeta), \text{ para qualquer partição } \xi.$$

Demostração:

$$\begin{aligned}
H_\mu(\xi|\zeta) &= -\sum_{C\in\xi, D\in\zeta} \mu(C\cap D)\log\frac{\mu(C\cap D)}{\mu(D)} \\
&= -\sum_{C\in\xi, E\in\eta} \mu(E)\sum_{D\in\zeta}\frac{\mu(D)}{\mu(E)}\frac{\mu(C\cap D)}{\mu(D)}\log\frac{\mu(C\cap D)}{\mu(D)} \\
&= \sum_{C\in\xi, E\in\eta} \mu(E)\sum_{D\in\zeta}\frac{\mu(D)}{\mu(E)}\psi\left(\frac{\mu(C\cap D)}{\mu(D)}\right) \\
&\leq \sum_{C\in\xi, E\in\eta} \mu(E)\psi\left(\sum_{D\in\zeta}\frac{\mu(D)}{\mu(E)}\frac{\mu(C\cap D)}{\mu(D)}\right) \\
&= \sum_{C\in\xi, E\in\eta} \mu(E)\psi\left(\sum_{D\in\zeta}\frac{\mu(C\cap D)}{\mu(E)}\right).
\end{aligned}$$

Agora como temos que ζ é um refinamento de η, então:

$$\begin{aligned}
\sum_{D\in\zeta}\frac{\mu(C\cap D)}{\mu(E)} &= \sum_{D\in\zeta, D\subset E}\frac{\mu(C\cap D)}{\mu(E)} \\
&= \frac{\mu(C\cap E)}{\mu(E)}.
\end{aligned}$$

Dai obtemos:

$$\begin{aligned}
\sum_{C\in\xi, E\in\eta} \mu(E)\psi\left(\sum_{D\in\zeta}\frac{\mu(C\cap D)}{\mu(E)}\right) &= \sum_{C\in\xi, E\in\eta} \mu(E)\psi\left(\frac{\mu(C\cap E)}{\mu(E)}\right) \\
&= -\sum_{C\in\xi, E\in\eta} \mu(E)\frac{\mu(C\cap E)}{\mu(E)}\log\frac{\mu(C\cap E)}{\mu(E)}
\end{aligned}$$

$$= -\sum_{C\in\xi,E\in\eta}\mu(C\cap E)\log\frac{\mu(C\cap E)}{\mu(E)}$$

$$= H_\mu(\xi|\eta).$$

Para finalizar a prova do Teorema (4.0.1), precisamos do seguinte resultado.

Lema 4.0.2. *Se ξ, η e ζ são partições mensuráveis de X, então*

$$H_\mu(\xi\vee\zeta|\eta) = H_\mu(\xi|\zeta\vee\eta) + H_\mu(\zeta|\eta).$$

Demostração:

$$\begin{aligned}
H_\mu(\xi\vee\zeta|\eta) &= -\sum_{C\in\xi,D\in\zeta,E\in\eta}\mu(C\cap D\cap E)\log\frac{\mu(C\cap D\cap E)}{\mu(E)}\\
&= -\sum_{C\in\xi,D\in\zeta,E\in\eta}\mu(C\cap D\cap E)\log\left(\frac{\mu(C\cap D\cap E)}{\mu(E)}\left(\frac{\mu(D\cap E)}{\mu(D\cap E)}\right)\right)\\
&= -\sum_{C\in\xi,D\in\zeta,E\in\eta}\mu(C\cap D\cap E)\log\left(\frac{\mu(C\cap D\cap E)}{\mu(D\cap E)}\left(\frac{\mu(D\cap E)}{\mu(E)}\right)\right)\\
&= -\sum_{C\in\xi,D\in\zeta,E\in\eta}\mu(C\cap D\cap E)\left(\log\left(\frac{\mu(C\cap D\cap E)}{\mu(D\cap E)}\right)+\left(\log\frac{\mu(D\cap E)}{\mu(E)}\right)\right)\\
&= -\sum_{C\in\xi,D\in\zeta,E\in\eta}\mu(C\cap D\cap E)\log\left(\frac{\mu(C\cap D\cap E)}{\mu(D\cap E)}\right)\\
&- \sum_{C\in\xi,D\in\zeta,E\in\eta}\mu(C\cap D\cap E)\log\left(\frac{\mu(D\cap E)}{\mu(E)}\right)\\
&= -\sum_{C\in\xi,D\in\zeta,E\in\eta}\mu(C\cap D\cap E)\log\left(\frac{\mu(C\cap D\cap E}{\mu(D\cap E)}\right)-\sum_{D\in\zeta,E\in\eta}\mu(D\cap E)\log\left(\frac{\mu(D\cap E}{\mu(E)}\right)\\
&= H_\mu(\xi|\zeta\vee\eta) + H_\mu(\zeta|\eta).
\end{aligned}$$

Em particular, para a partição trivial $\eta = \{X\}$ segue-se do Lema (3.0.1) que:

$$H_\mu(\xi \vee \zeta) = H_\mu(\xi|\zeta) + H_\mu(\zeta). \tag{4.3}$$

Agora usaremos indução para provar que:

$$H_\mu(\xi_n) = H_\mu(\xi) + \sum_{j=1}^{n-1} H_\mu(\xi|\bigvee_{i=0}^{j} T^{-i}\xi). \tag{4.4}$$

- *Para $n = 1$ é evidente.*

- *Para $n = 2$.*

$$\begin{aligned} H_\mu(\xi_2) &= H_\mu(\xi \vee T^{-1}\xi) \\ &= H_\mu(\xi) + H_\mu(\xi|T^{-1}\xi). \end{aligned}$$

- *Para n=3.*

$$H_\mu(\xi_3) = H_\mu(\xi \vee \bigvee_{j=1}^{2} T^{-j}\xi).$$

Agora tomando $\alpha_2 = T^{-1}(\xi_2)$ ou seja $\alpha_2 = T^{-1}\xi \vee T^{-2}\xi$,

$$\begin{aligned} H_\mu(\xi_3) &= H_\mu(\xi \vee \alpha_2) \\ &= H_\mu(\alpha_2) + H_\mu(\xi|\alpha_2) \\ &= H_\mu(T^{-1}\xi_2) + H_\mu(\xi|\alpha_2) \\ &= H_\mu(\xi_2) + H_\mu(\xi|T^{-1}\xi \vee T^{-1}) \\ &= H_\mu(\xi) + H_\mu(\xi|T^{-1}\xi) + H_\mu(\xi|T^{-1}\xi \vee T^{-1}) \\ &= H_\mu(\xi) + \sum_{j=1}^{2} H_\mu(\xi|\bigvee_{i=1}^{j} T^{-i}\xi). \end{aligned}$$

- *Suponha para $n = k$, então:*

$$H_\mu(\xi_k) = H_\mu(\xi) + \sum_{j=1}^{k-1} H_\mu(\xi | \bigvee_{i=0}^{j} T^{-i}\xi).$$

- *Provaremos para $n = k + 1$*

Tomando $\theta_k = T^{-1}(\xi_k)$ ou seja $\theta_k = T^{-1}\xi \vee T^{-2}\xi \vee ... \vee T^{-k}\xi$. Então:

$$\begin{aligned}
H_\mu(\xi_{k+1}) &= H_\mu(\xi \vee \theta_k) \\
&= H_\mu(\theta_k) + H_\mu(\xi|\theta_k) \\
&= H_\mu(T^{-1}(\xi_k)) + H_\mu(\xi|\theta_k) \\
&= H_\mu(\xi_k) + H_\mu(\xi|\theta_k) \\
&= H_\mu(\xi) + \sum_{j=1}^{k-1} H_\mu(\xi | \bigvee_{i=1}^{j} T^{-i}\xi) + H_\mu(\xi | \bigvee_{i=1}^{j} T^{-i}\xi) \\
&= H_\mu(\xi) + \sum_{j=1}^{k} H_\mu(\xi | \bigvee_{i=1}^{j} T^{-i}\xi).
\end{aligned}$$

Dado que a sequência $H_\mu\big(\xi | \bigvee_{i=1}^{n} T^{-i}\xi\big)$ converge, se deduz de: $H_\mu(\xi_n) = H_\mu(\xi) + \sum_{j=1}^{n-1} H_\mu\big(\xi | \bigvee_{i=1}^{j} T^{-i}\xi\big)$ que:

$$\begin{aligned}
h_\mu(T,\xi) &= \lim_{n\to\infty} \frac{1}{n} H_\mu(\xi_n), \text{ pela Definição (3.0.4).} \\
&= \lim_{n\to\infty} \frac{1}{n}\Bigg(H_\mu(\xi) + \sum_{j=1}^{n-1} H_\mu\big(\xi | \bigvee_{i=1}^{j} T^{-i}\xi\big)\Bigg) \\
&= \lim_{n\to\infty} \frac{1}{n}(H_\mu(\xi) + \lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n-1} H_\mu\Bigg(\xi | \bigvee_{i=1}^{j} T^{-i}\xi\Bigg)
\end{aligned}$$

$$= \lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n-1} H_\mu\Big(\xi| \bigvee_{i=1}^{j} T^{-i}\xi\Big).$$

Como $\lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n-1} H_\mu\Big(\xi| \bigvee_{i=1}^{j} T^{-i}\xi\Big) = \lim_{n\to\infty} \frac{1}{n} H_\mu\Big(\xi| \bigvee_{i=1}^{n-1} T^{-i}\xi\Big)$, então, temos finalmente que:

$$h_\mu(T,\xi) = \lim_{n\to\infty} \frac{1}{n} H_\mu\Big(\xi| \bigvee_{i=1}^{n-1} T^{-i}\xi\Big).$$

□

Dada uma partição mensurável ξ de X, denotaremos por $A(\xi)$ a σ-álgebra gerada por ξ, que é a mais pequena σ-álgebra que contém todos os elementos de ξ. Ademais, denotaremos por $\bigvee_{n\in I} A(\xi_n)$ a σ-álgebra gerada pela partição mensurável ξ_n para $n \in I$.

Teorema 4.0.2. *Seja $T : X \to X$ uma transformação que preserva a medida de probabilidade μ em X. Se $(\alpha_n)_{n\in\mathbb{N}}$ é uma sequência de partições mensuráveis de X com $\bigvee_{n=1}^{\infty} A(\alpha_n) = A$ tal que α_{n+1} é um refinamento de α_n para cada $n \in \mathbb{N}$, então;*

$$h_\mu(T) = \lim_{n\to\infty} h_\mu(T,\alpha_n) = \sup_{n\in\mathbb{N}} h_\mu(T,\alpha_n).$$

Demostração: Seja $\bigvee_{i=0}^{n-1} T^{-1}\alpha_{m+1}$ um refinamento de $\bigvee_{i=0}^{n-1} T^{-1}\alpha_m$. Pela Proposição (3.0.1) temos:

$$H_\mu\Big(\bigvee_{i=0}^{n-1} T^{-i}\alpha_m\Big) \;\leq\; H_\mu\Big(\bigvee_{i=0}^{n-1} T^{-i}\alpha_{m+1}\Big) \text{ ; Para qualquer } m,n \in \mathbb{N}.$$

$$\lim_{n\to\infty} \frac{1}{n} H_\mu\Big(\bigvee_{i=0}^{n-1} T^{-1}\alpha_m\Big) \;\leq\; \lim_{n\to\infty} \frac{1}{n} H_\mu\Big(\bigvee_{i=0}^{n-1} T^{-1}\alpha_{m+1}\Big).$$

Pelo Teorema (4.0.1), obtemos;

$h_\mu(T,\alpha_m) \leq h_\mu(T,\alpha_{m+1})$; para cada $m \in \mathbb{N}$.

Uma vez que a sequência $h_\mu(T,\alpha_n)$ é não decrescente temos que

$\lim_{n\to\infty} h_\mu(T,\alpha_n) = \sup_{n\in\mathbb{N}} h_\mu(T,\alpha_n)$.

Agora estabeleceremos três resultados auxiliares, que combinados dará os resultado desejado.

Lema 4.0.3. *(Lema indutivo).*

Fixada uma partição β vale que:

$$H_\mu(\bigvee_{i=0}^{n-1} T^{-i}\alpha|\beta) \leq \sum_{i=0}^{n-1} H_\mu(T^{-i}\alpha|\beta); \tag{4.5}$$

$\forall n \geq 2$ e qualquer partição α.

Demonstração:

- Provaremos para n=2.

$$\begin{aligned} H_\mu(\bigvee_{i=0}^{1} T^{-i}\alpha|\beta) &= H_\mu(T^{-1}\alpha \vee \alpha|\beta) \\ &= H_\mu(T^{-1}\alpha|\alpha \vee \beta) + H_\mu(\alpha|\beta) \\ &\leq H_\mu(T^{-1}\alpha|\beta) + H_\mu(\alpha|\beta) \\ &= \sum_{i=0}^{1} H_\mu(T^{-i}\alpha|\beta). \end{aligned}$$

- Suponha para $n = k$, então

$$H_\mu(\bigvee_{i=0}^{k-1} T^{-i}\alpha|\beta) \leq \sum_{i=0}^{k-1} H_\mu(T^{-i}\alpha|\beta).$$

- Provemos para $n = k + 1$

$$\begin{aligned} H_\mu(\bigvee_{i=0}^{k} T^{-i}\alpha|\beta) &= H_\mu(\bigvee_{i=1}^{k} T^{-i}\alpha \vee \alpha|\beta) \\ &= H_\mu(\bigvee_{i=1}^{k} T^{-i}\alpha|\alpha \vee \beta) + H_\mu(\alpha|\beta) \\ &\leq H_\mu(\bigvee_{i=1}^{k} T^{-j}\alpha|\beta) + H_\mu(\alpha|\beta). \end{aligned}$$

Agora tomando;

$$\begin{aligned}\theta &= T^{-1}\alpha,\\ T^{-1}\theta &= T^{-2}\alpha,\\ \vdots &= \vdots\\ T^{-k+1}\theta &= T^{-k}\alpha.\end{aligned}$$

Então temos:

$$\begin{aligned}H_\mu(\bigvee_{i=1}^{k} T^{-i}\alpha|\beta) + H_\mu(\alpha|\beta) &= H_\mu(\bigvee_{i=0}^{k-1} T^{-i}\theta|\beta) + H_\mu(\alpha|\beta)\\ &\leq \sum_{i=0}^{k-1} H_\mu(T^{-i}\theta|\beta) + H_\mu(\alpha|\beta)\\ &= \sum_{i=0}^{k-1} H_\mu(T^{-i}(T^{-1}\alpha)|\beta) + H_\mu(\alpha|\beta)\\ &= \sum_{i=0}^{k} H_\mu(T^{-i}\alpha|\beta) + H_\mu(\alpha|\beta).\end{aligned}$$

Mostrando o desejado.

Lema 4.0.4. *Se η e ζ são partições mensuráveis de X, então*

$$h_\mu(T,\eta) \leq h_\mu(T,\zeta) + H_\mu(\eta|\zeta). \tag{4.6}$$

Demonstração: Seja,

$$\begin{aligned}H_\mu(\eta_n) &\leq H_\mu(\eta_n) + H_\mu(\zeta_n)\\ &= H_\mu(\eta_n \vee \zeta_n),\end{aligned}$$

ou seja;

$$H_\mu(\eta_n) \leq H_\mu(\eta_n \vee \zeta_n) = H_\mu(\eta_n|\zeta_n) + H_\mu(\zeta_n), \tag{4.7}$$

onde,

$$\eta_n = \bigvee_{i=0}^{n-1} T^{-i}\eta \quad e \quad \zeta_n = \bigvee_{i=0}^{n-1} T^{-i}\zeta \tag{4.8}$$

Agora dadas α, β e γ partições qualquer de X pelo Lema (3.0.1) temos

$$\begin{aligned} H_\mu(\alpha \vee \beta) &= H_\mu(\alpha|\beta \vee \gamma) + H_\mu(\beta|\gamma) \\ &\leq H_\mu(\alpha|\gamma) + H_\mu(\beta|\gamma). \end{aligned}$$

Para qualquer partição mensurável α, β e γ de X. Portanto, por (4.7) e (4.8) e Lema (3.0.1) obtemos;

$$\begin{aligned} H_\mu(\eta_n) &\leq H_\mu(\eta_n) + H_\mu(\zeta_n) \\ &\leq H_\mu(\eta_n) + \sum_{i=0}^{n-1} H_\mu(T^{-i}\eta|\zeta_n) \\ &\leq H_\mu(\eta_n) + \sum_{i=0}^{n-1} H_\mu(T^{-i}\eta|T^{-i}\zeta) \\ &= H_\mu(\eta_n) + nH_\mu(\eta|\zeta). \end{aligned}$$

Usando a invariância de μ para a última igualdade, finalmente obtemos:

$$\begin{aligned} \lim_{n\to\infty} \frac{1}{n} H_\mu(\eta_n) &\leq \lim_{n\to\infty} \frac{1}{n}(H_\mu(\eta_n) + nH_\mu(\eta|\zeta)) \\ &= \lim_{n\to\infty} \frac{1}{n}(H_\mu(\zeta_n)) + \lim_{n\to\infty} \frac{1}{n} n(H_\mu(\eta|\zeta)) \\ &= \lim_{n\to\infty} \frac{1}{n}(H_\mu(\zeta_n)) + H_\mu(\eta|\zeta) \\ h_\mu(T, \eta) &\leq h_\mu(T, \zeta) + H_\mu(\eta|\zeta). \end{aligned}$$

$\square$

Lema 4.0.5. *Para qualquer partição mensurável η de X, temos $H_\mu(\eta|\alpha_n) \to 0$, quando $n \to \infty$.*

Demostração: Considere $\eta = \{C_1, C_2, ..., C_k\}$ uma partição mensurável de X. Para $n \in \mathbb{N}$ suficientemente grande e $i = 1, ..., k$, seja $D_i^n \subset C_i$ um conjunto em $A(\alpha_n)$ tal que $\mu(C_i \setminus D_i^n) \to 0$ quando $n \to \infty$. Seja $\beta_n = \{D_0^n, D_1^n, ..., D_k^n\}$ onde $D_0^n = X \setminus \bigcup_{i=1}^{k} D_i^n$. Claramente, $\mu\big(C_i \cup D_i^n\big) \to 0$ quando $n \to \infty$. Desde que

$$C_i \cap D_i^n = D_i^n \text{ e } \mu(C_i \cap D_i^n) = 0$$

para $i = 1, ..., k$ e $i \neq j$, obtemos

$$\begin{aligned}
H_\mu(\eta|\beta_n) &= -\sum_{i=1}^{k}\sum_{j=0}^{k} \mu\big(C_i \cap D_j^n\big) \log \frac{\mu(C_i \cap D_j^n)}{\mu(D_j^n)} \\
&= -\sum_{i=1}^{k} \mu(D_i^n) \log \frac{\mu(D_i^n)}{\mu(D_i^n)} - \sum_{i=1}^{k} \mu\big(C_i \cap D_0^n\big) \log \frac{\mu(C_i \cap D_0^n)}{\mu(D_0^n)} \\
&= -\sum_{i=1}^{k} \mu(C_i \cap D_0^n) \log \frac{\mu(C_i \cap D_0^n)}{\mu(D_0^n)} \\
&= -\sum_{i=1}^{k} \mu(C_i \cap D_0^n)\big(\log \mu(C_i \cap D_0^n) - \log \mu(D_0^n)\big) \\
&= -\sum_{i=1}^{k} \mu(C_i \cap D_0^n) \log \mu(C_i \cap D_0^n) + \sum_{i=1}^{k} \mu(C_i \cap D_0^n) \log \mu(D_0^n) \\
&= \sum_{i=1}^{k} \psi\big(\mu(C_i \cap D_0^n)\big) + \sum_{i=1}^{k} \mu(C_i \cap D_0^n) \log \mu(D_0^n) \\
&= \sum_{i=1}^{k} \Big(\psi\big(\mu(C_i \cap D_0^n)\big) + \mu(C_i \cap D_0^n) \log \mu(D_0^n)\Big) \\
&= \sum_{i=1}^{k} \Big(\psi\big(\mu(C_i \cap D_0^n)\big) + \mu(D_0^n) \log \mu(D_0^n)\Big) \\
&= \sum_{i=1}^{k} \Big(\psi\big(\mu(C_i \cap D_0^n)\big) - \psi\big(\mu(D_0^n)\big)\Big).
\end{aligned}$$

Agora temos: $H_\mu(\eta|\beta_n) = \sum_{i=1}^{k}\Big(\psi\big(\mu(C_i \cap D_0^n)\big) - \psi\big(\mu(D_0^n)\big)\Big) \to 0$ quando $n \to \infty$. Finalmente, dado que α_n é um refinamento de β_n. Então finalmente temos

$$H_\mu(\eta|\alpha_n) \leq H_\mu(\eta|\beta_n) \to 0 \text{ quando } n \to \infty.$$

Mostrando o desejado.

Da Equação (4.6) temos, $h_\mu(T,\eta) \leq h_\mu(T,\zeta) + H_\mu(\eta|\zeta)$, então tomando $\zeta = \alpha_n$ temos que,

$$\begin{aligned} h_\mu(T,\eta) &\leq h_\mu(T,\alpha_n) + H_\mu(\eta|\alpha_n) \\ \lim_{n\to\infty} h_\mu(T,\eta) &\leq \lim_{n\to\infty} h_\mu(T,\alpha_n) + \lim_{n\to\infty} H_\mu(\eta|\alpha_n) \\ h_\mu(T,\eta) &\leq \lim_{n\to\infty} h_\mu(T,\alpha_n). \end{aligned}$$

Agora, $h_\mu(T) = \sup\{h_\mu(T,\eta)\} = E$ e $h_\mu(T,\eta) \leq \lim_{n\to\infty} h_\mu(T,\alpha_n) = a$, ou seja,

$$h_\mu(T,\eta) \leq a \leq E. \tag{4.9}$$

Por outra parte temos que, $h_\mu(T,\eta) \leq a : \forall \eta$, então $h_\mu(T) \leq a$, ou seja,

$$E \leq a. \tag{4.10}$$

Finalmente, de (4.9) e (4.10) conclui-se que

$$h_\mu(T,\eta) = \lim_{n\to\infty} h_\mu(T,\alpha_n), \text{ como desejávamos.}$$

$\square$

Capítulo 5

Princípio Variacional

Neste capítulo final daremos uma prova detalhada do principio variacional para a entropia. Provaremos que a entropia topológica é igual ao supremo das entropias métricas sobre todas as medidas $T-$invariantes de probabilidade de borel, tal conjunto de medidas as notaremos por $\mathcal{M}(X)$.

Ademais, daremos um exemplo da aplicação do principio variacional. O objetivo do exemplo é mostrar que o conjunto de puntos não errantes que será denotado por $\Omega(f)$, carrega com toda a complexidade do sistema.

Teorema 5.0.1. *Se $T : X \to X$ é uma transformação contínua de um espaço métrico compacto, então*

$h_{top}(T)=$ sup$\{h_\mu(T)$*: μ é uma medida T-invariante em $\mathcal{M}(X)$, onde $\mathcal{M}(X)$ é conjunto de todas as medidas de probabilidade borelianas e T-invariantes*$\}$

Demonstração: Primeiro provaremos que $h_\mu(T) \leq h_{top}(T)$ (a entropia métrica é menor ou igual que a entropia topológica) para qualquer medida de probabilidade T-invariante μ em X. Seja

$$\eta = \{C_1, C_2, ..., C_k\}$$

uma partição mensurável de X. Dado $\delta > 0$, para cada $i = 1, ..., k$, seja $D_i \subset C_i$, um conjunto compacto tal que:

$$\mu(C_i \setminus D_i) < \delta \ \text{ e } \ D_i \cap D_j = \emptyset, \text{ para } i \neq j.$$

Agora seja $\beta = \{D_0, D_1, ..., D_k\}$ onde $D_0 = X \setminus \bigcup_{i=1}^{k} D_i$, claramente β é uma partição mensurável de X. Procedendo da mesma forma que no Lema (4.0.5) e substituindo D_i^n por D_i, obtemos

$$H_\mu(\eta|\beta) = \sum_{j=1}^{n-1} \psi(\mu(C_i \cap D_0)) - \psi(\mu(D_0)).$$

Por construção, temos que $\mu(C_i \cap D_0) < \delta$, pois

$$\begin{aligned} C_i \cap D_0 &= C_i \cap (X \setminus \bigcup_{i=1}^{k} D_i) \\ &= C_i \cap \bigcap_{i=1}^{k} D_i^c \ \subset \ C_i \cap D_i^c \ = C_i \setminus D_i \\ \implies & \ C_i \cap D_0 \subset C_i \setminus D_i. \end{aligned}$$

Logo tomando a medida temos que:

$\mu(C_i \cap D_0) \leq \mu(C_i \setminus D_i) < \delta$ e dado que ψ é não decrescente perto de zero, concluímos que $H_\mu(\eta|\beta) < 1$, para qualquer δ muito pequeno, portanto, pelo Lema (4.0.4), obtemos que

$$h_\mu(T, \eta) \leq h_\mu(T, \beta) + H_\mu(\eta|\beta) < h_\mu(T, \beta) + 1,$$

este resultado é dado, pois η e β são partições mensuráveis de X.

Agora construiremos uma sequência de coberturas abertas, de tal forma que nos permita relacionar a entropia métrica e a entropia topológica.

Considere as coberturas abertas de X dadas por

$$\mathcal{U} = \{D_0 \cup D_1, D_0 \cup D_2, ..., D_0 \cup D_k\}.$$

e

$$\mathcal{U}_n = \left\{ \bigcap_{i=0}^{n-1} T^{-i} U_i : U_0, U_1, ..., U_{n-1} \in \mathcal{U} \right\},$$

para cada $n \in \mathbb{N}$. Seja $\mathcal{N}(\mathcal{U}_n)$ a miníma cardinalidade de uma subcobertura de $\mathcal{U}_n$. Desejamos relacionar $\mathcal{N}(\mathcal{U}_n)$ com a cardinalidade das partições $\beta_n = \bigvee_{i=0}^{n-1} T^{-i}\beta$. Primeiro perceba que

$$T^{-i}(D_0 \cup D_j) = T^{-i}(D_0) \cup T^{-i}(D_j).$$

Para $j = 1, ..., k$ e $i \in \mathbb{N}$, cada elemento de $\mathcal{U}_n$ é união de máximo 2^n elementos de β_n. Daremos uma demostração do argumento anterior, tal demostração sera feita por indução, então, para $n = 1$, se tem

$$\begin{aligned} \mathcal{U}_1 &= T^0(U_0) \\ &= U_0 \\ &= D_0 \cup D_1; D_0 \in \beta_1 \ e \ D_1 \in \beta_1. \end{aligned}$$

Portanto, cada elemento de $\mathcal{U}_1$ é união de máximo $2^1 = 2$ elementos de β_1.

Suponha para $n = k$, assim o conjunto pode ser escrito como união de máximo 2^{k-1} elementos de β_k.

Agora mostraremos para $n = k + 1$. Seja V um aberto de $\mathcal{U}_{k+1}$, então,

$$\bigcap_{i=0}^{k} T^{-i}(U_i) = \left(\bigcap_{i=0}^{k-1} T^{-i}(U_i)\right) \cap T^{-k}(U_k)$$

pela hipótese temos que: $\bigcap_{i=0}^{k-1} T^{-i}(U_i) = \bigcup_{i=1}^{2^k} A_i;\ A_i \in \beta_k$ então,

$$\begin{aligned} \bigcap_{i=0}^{k} T^{-i}(U_i) &= \left(\bigcup_{i=1}^{2^k} A_i\right) \cap T^{-k}(U_k) \\ &= \left(\bigcup_{i=1}^{2^k} A_i\right) \cap T^{-k}(D_0 \cup D_k) \\ &= \left(\bigcup_{i=1}^{2^k} A_i\right) \cap \left(T^{-k}D_0 \cup T^{-k}D_k\right) \\ &= \left(\bigcup_{i=1}^{2^k} A_i \cap T^{-k}D_0\right) \cup \left(\bigcup_{i=1}^{2^k} A_i \cap T^{-k}D_k\right). \end{aligned}$$

Logo, $\bigcup_{i=1}^{2^k} A_i \cap T^{-k}D_0$ é união do máximo 2^k elementos de β_{k+1} e o mesmo argumento e valido para $\left(\bigcup_{i=1}^{2^k} A_i \cap T^{-k}D_k\right)$. Portanto, temos que esta união tem no máximo $2^k + 2^k$ elementos de β_{k+1}, ou seja 2^{k+1} elementos. Assim obtemos o desejado.

Além disso, como os elementos de β_n são disjuntos, cada elemento de β_n está contido em algum elemento de qualquer subcobertura de $\mathcal{U}_n$. Portanto, $card\beta_n \leq 2^n\mathcal{N}(\mathcal{U}_n)$

Logo temos,

$$\begin{aligned} H_\mu(\beta_n) &\leq \log card(\beta_n) \\ &\leq \log(2^n(\mathcal{N}(\mathcal{U}_n)) \\ &= \log(2^n) + \log\mathcal{N}(\mathcal{U}_n) \\ &= n\log(2) + \log\mathcal{N}(\mathcal{U}_n). \end{aligned}$$

Por outro lado, pela teoria de espaços métricos sabemos que como X é compacto, existe $\varepsilon > 0$ tal que toda bola de raio máximo ε está contida em algum elemento de $\mathcal{U}$. Ademais, se pode verificar facilmente que se ε é o número de Lebesgue de $\mathcal{U}$, também é o número de Lebesgue de $\mathcal{U}_n$ com respeito da distância d_n.

De fato pela definição de d_n temos,

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

E como ε é o número de Lebesgue de $\mathcal{U}$, cada $B_{d_n}(T^i(x),\varepsilon)$ está contida em algum elemento U_i de $\mathcal{U}$, portanto,

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

e a contenção anterior é obtida dado que:

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

Agora tome pontos $x_1, x_2, x_3, ..., x_p \in X$ tal que

$$\left\{B_{d_n}(x_1,\varepsilon), B_{d_n}(x_2,\varepsilon), ..., B_{d_n}(x_p,\varepsilon)\right\},$$

forme uma cobertura aberta de X com $p = M_d(f,\varepsilon,n)$. Como ε é o número de Lebesgue de $\mathcal{U}_n$ com respeito a distância d_n, para cada $i = 1,...,p$ se pode escolher um elemento $U_i \in \mathcal{U}$ tal que $B_{d_n}(x_i,\varepsilon) \subset U_i$.

É claro que a coleção de conjuntos abertos $\{U_i : 1 \leq i \leq p\}$ formam uma subcobertura de $\mathcal{U}_n$ que cobre X e portanto,

$$\mathcal{N}(\mathcal{U}_n) \leq card\{U_i : 1 \leq i \leq p\} \leq p = M_d(f,\varepsilon,n)$$

Como $H_\mu(\beta_n) \leq \log card(\beta_n) \leq n\log 2 + \log \mathcal{N}(\mathcal{U}_n)$, pela Definição (3.0.5) temos que:

$$h_\mu(T,\beta) \;=\; \lim_{n\to\infty} \frac{1}{n} H_\mu(\beta_n) \;\leq\; \lim_{n\to\infty} \log(2) + \lim_{n\to\infty} \frac{1}{n} \log M_d(f,\varepsilon,n)$$

Ou seja;

$$h_\mu(T,\beta) \;\leq\; \lim_{n\to\infty} \log(2) + \lim_{n\to\infty} \sup \frac{1}{n} \log M_d(n,\varepsilon)$$

Tomando $\lim\limits_{\varepsilon\to 0}$ obtemos;

$$h_\mu(T,\beta) \;\leq\; \log(2) + h_{top}(T).$$

Ademais como temos que

$h_\mu(T,\eta) \leq h_\mu(T,\beta) + 1$, então $h_\mu(T,\eta) \leq h_\mu(T,\beta) + h_{top}(T) + 1$.

Tomando, $h_\mu(T) = \sup\{h_\mu(T,\eta); \eta$ é uma partição mensurável de X$\}$, então,

$$h_\mu(T) \leq \log(2) + h_{top}(T) + 1. \tag{5.1}$$

Para obter esta desigualdade, não precisamos nada da transformação T mais que o fato que seja contínua e que a medida μ seja T-invariante.

Portanto, a Desigualdade (5.1) segue sendo valida quando se substitui T por T^m, então,

$$h_\mu(T^m) \leq h_{top}(T^m) + \log(2) + 1, \text{ para todo } m \in \mathbb{N}.$$

Pela Proposição (2.0.2) e a Proposição (3.0.3) obtemos,

$$\begin{aligned}
h_\mu(T) &= \frac{1}{m} h_{top}(T^m);\ \forall\ m \in \mathbb{N} \\
&\leq \frac{1}{m}(h_{top}(T^m) + \log(2) + 1) \\
&= \frac{1}{m} h_{top}(T^m) + \frac{1}{m}(\log(2)+1) \\
&= h_{top}(T) + \frac{1}{m}(\log(2)+1).
\end{aligned}$$

Tomando, $\lim_{m\to\infty}$ obtemos que $h_\mu(T) \leq h_{top}(T)$ que é a desigualdade desejada.

Agora mostraremos a desigualdade contraria ou seja, $h_{top}(T) \leq \sup h_\mu(T)$, para isto, mostraremos que $h_{top}(T) \leq \sup_\nu h_\nu(T)$.

Dado $\varepsilon > 0$, para cada $n \in \mathbb{N}$, consideremos o conjunto E_n de pontos a uma d_n-distância pelo menos ε tal que $card(E_n) = N_d(f, \varepsilon, n)$.

Definamos as seguintes medidas,

$$\nu_n = \frac{1}{cardE_n} \sum_{x \in E_n} \delta_x \quad \text{e} \quad \mu_n = \frac{1}{n} \sum_{i=0}^{n-1} T_*^i \nu_n$$

onde δ_x é a medida delta de Dirac e $T_*^i \nu_n$ é definida como segue:

$$(T_*^i \nu_n)(A) = \nu_n(T^{-i} A).$$

Dado que o conjunto de todas as medidas de probabilidade de Borel sobre um espaço métrico compacto é um espaço metrizável compacto, portanto, para cada sequência de medidas tem uma subsequência convergente. Além disso, uma sequência de medidas μ_n converge a uma medida μ se e somente se

$$\int_X \varphi \, d\mu_n \to \int_X \varphi \, d\mu \tag{5.2}$$

quando $n \to \infty$, para toda função contínua φ.

Dada uma sequência $(k_n)_{n\in\mathbb{N}} \subset \mathbb{N}$ tal que

$$\lim_{n\to\infty} \frac{1}{k_n} \log cardE_{k_n} = \lim_{n\to\infty} \sup \frac{1}{n} \log cardE_n. \tag{5.3}$$

Seja $(\ell_n)_{n\in\mathbb{N}}$ uma subsequência de $(k_n)_{n\in\mathbb{N}}$ tal que $(\mu_{\ell_n})_{n\in\mathbb{N}}$ converge para alguma medida de probabilidade μ.

Provaremos que μ é T- invariante. Primeiro estabeleceremos um resultado auxiliar

Lema 5.0.1. *A transformação $T_* : \mathcal{M}(X) \to \mathcal{M}(X)$ definida por*

$$(T_*\mu)(A) = \mu(T^{-1}A),$$

é contínua.

Prova: Demostraremos que se $\mu_n \to \mu$, então $T_*\mu_n \to T_*\mu$, estabelecendo a continuidade de T_*. Primeiro demostraremos a seguinte afirmação

Afirmação: $\displaystyle\int_X \varphi\, dT_*\mu = \int_X (\varphi \circ T)\, d\mu_n;\ \forall\varphi \in C(X)$ e $\mu \in \mathcal{M}(X)$.

Prova da afirmação: Provaremos inicialmente a continuidade, assuma,

$\displaystyle\int_X \varphi\, dT_*\mu = \int_X \psi\, d\mu_n\ ;\varphi \in C(X)$, onde $\psi = \varphi \circ T$.

Então, assim temos que,

$\displaystyle\int_X \varphi\, dT_*\mu = \int_X \psi\, d\mu_n$, logo por (5.2) temos que

$$\int_X \psi\, d\mu_n \to \int_X \psi\, d\mu \Longleftrightarrow \int_X \varphi\, dT_*\mu_n \to \int_X \varphi\, dT_*\mu.$$

Provando o desejado.

Agora por outro lado, provaremos que a afirmação é valida para funções simples

$$\begin{aligned}
\int_X 1_A\, dT_*\mu_n &= (T_*\mu)(A)\\
&= \mu(T^{-1}A)\\
&= \int_X 1_{T^{-1}A}\, d\mu\\
&= \int_X (1_A \circ T)\, d\mu.
\end{aligned}$$

Dada $\varphi \in C(X)$, existem $\varphi_1, \varphi_2, ..., \varphi_n...$ funções simples tais que $\varphi_n \to \varphi$ pontualmente, e $m \leq \varphi \leq M$. Pelo Teorema de Convergência Dominada (T.C.D) (1.3.2),

$\int_X \varphi_n \, d\mu \to \int_X \varphi \, d\mu$. Logo,

$$\int_X \varphi_n \, d(T_*\mu) \to \int_X \varphi \, d(T_*\mu), \tag{5.4}$$

e

$$\int_X \varphi_n \, d(T_*\mu) \to \int_X (\varphi_n \circ T) \, d\mu. \tag{5.5}$$

Dado $\varphi_n \to \varphi$, como T é contínua então $\varphi_n \circ T \to \varphi \circ T$ pontualmente. Assim, $m \leq \varphi \circ T \leq M$. Pelo (T.C.D)

$$\int_X (\varphi_n \circ T) \, d\mu \to \int_X (\varphi \circ T) \, d\mu.$$

Agora pela unicidade do limite de (5.4) e (5.5) obtemos que:

$$\int_X \varphi \, d(T_*\mu) = \int_X (\varphi \circ T) \, d\mu.$$

Finalmente, quando $n \to \infty$, temos que $(T_*\mu_n)_{n\in\mathbb{N}}$ converge para $T_*\mu$. Isto mostra que a transformação T_* é contínua. Assim finaliza a demostração da afirmação.

Agora provaremos que μ é um ponto fixo de T_*, e portanto, T-invariante.

Podemos escrever;

$$\begin{aligned}
T_*(\mu_{\ell_n}) &= T_*\left(\frac{1}{\ell_n}\sum_{i=o}^{\ell_n-1} T_*^i \nu_{\ell_n}\right) \\
&= T_*\left(\frac{1}{\ell_n}\left(\nu_{\ell_n} + T_*\nu_{\ell_n} + T_*^2\nu_{\ell_n} + \ldots + T_*^{\ell_{n-1}}\nu_{\ell_n}\right)\right) \\
&= \frac{1}{\ell_n}\left(T_*\nu_{\ell_n} + T_*^2\nu_{\ell_n} + T_*^3\nu_{\ell_n} + \ldots + T_*^{\ell_n}\nu_{\ell_n}\right) \\
&= \frac{1}{\ell_n}\sum_{i=0}^{\ell_n-1} T_*^{i+1}\nu_{\ell_n}
\end{aligned}$$

$$= \frac{1}{\ell_n}\left(\sum_{i=0}^{\ell_n-1} T_*^i \nu_{\ell_n} + T_*^{\ell_n}\nu_{\ell_n} - \nu_{\ell_n}\right)$$

$$= \frac{1}{\ell_n}\sum_{i=0}^{\ell_n-1} T_*^i \nu_{\ell_n} + \frac{T_*^{\ell_n}\nu_{\ell_n}}{\ell_n} - \frac{\nu_{\ell_n}}{\ell_n}.$$

Logo, fazendo $n \to \infty$,

$$\begin{aligned}
\lim_{n\to\infty}(T_*\mu_{\ell_n}) &= \lim_{n\to\infty}\left(\mu_{\ell_n} + \frac{T_*^{\ell_n}\nu_{\ell_n}}{\ell_n} - \frac{\nu_{\ell_n}}{\ell_n}\right) \\
&= \lim_{n\to\infty}\Big(\mu_{\ell_n}\Big) + \lim_{n\to\infty}\left(\frac{T_*^{\ell_n}\nu_{\ell_n}}{\ell_n}\right) - \lim_{n\to\infty}\left(\frac{\nu_{\ell_n}}{\ell_n}\right) \\
&= \mu.
\end{aligned}$$

Finalmente, obtemos que $T_*\mu = \mu$ como se desejava.

Agora vamos considerar alguma partição mensurável particular. Seja $\{B_1, ..., B_k\}$ uma cobertura de X por bolas de raio inferior a $\dfrac{\varepsilon}{2}$ tal que $\mu(\partial B_i) = 0$ para $i = 1, .., k$, onde ∂B_i denota a fronteira de B_i. Isto é possível sempre, dado que para cada $x \in X$ há no máximo um número contável de valores de $r > 0$, tal que $\mu\big(\partial B(x, r)\big) > 0$. Definimos una partição $\xi = \{C_1, ..., C_k\}$ de X dada por

$$C_1 = \overline{B_1} \text{ e } C_i = \overline{B_i} \setminus \bigcup_{j=1}^{i-1} \overline{B_j} \text{ para } i = 2, ..., k.$$

Lema 5.0.2. *Seja X um espaço métrico compacto e $\mu \in \mathcal{M}(X)$ uma medida de probabilidade.*

1. *Se $x \in X$ e $\delta > 0$, então existe $\delta' < \delta$ tal que $\mu\big(\partial B(x, \delta')\big) = 0$.*

2. *Se $\delta > 0$, então existe uma partição finita $\xi = \{A_1, ..., A_k\}$ de $(X, \mathcal{M})$ tal que $diam(A_j) < \delta$ e $\mu(\partial A_j) = 0$, para cada j.*

Prova: **1**.A prova de o Item 1, será feita por contra-positiva. Vamos supor que $\mu\big(\partial B(x,\delta')\big) \neq 0$ para todo $\delta' < \delta$ com $\delta' > 0$.

Consideremos a coleção $\mathcal{C} = \{\partial\big(B(x,\delta')\big); \delta' < \delta \wedge \delta' > 0\}$. Note que $\mathcal{C}$ é uma coleção disjunta de conjuntos de medida positiva. Sabendo que toda coleção disjunta de conjuntos com medida positiva com relação a uma medida finita do espaço X é no máximo enumerável, ou seja, suponha que para cada $x \in X$ existem incontáveis $\partial\big(B(x,\delta')\big)$ tais que $\partial\big(B(x,\delta')\big) \geq \dfrac{1}{n}$ para cada $n \in \mathbb{N}$.

Observemos que $\sum \dfrac{1}{n} \leq \sum \mu\big(B(x,\delta')\big)$, também temos que $B(x,\delta) = \bigcup\limits_{\delta'<\delta} \partial B(x,\delta')$ então,

$$\begin{aligned}
\mu\big(B(x,\delta)\big) &= \mu\Bigg(\bigcup_{\delta'<\delta} \partial B(x,\delta')\Bigg) \\
\\
&= \sum \mu\big(\partial B(x,\delta')\big).
\end{aligned}$$

Finalmente,

$$\infty = \sum \frac{1}{n} \leq \sum \mu\big(B(x,\delta')\big) = \mu\big(B(x,\delta)\big) < \mu(X) < \infty.$$

logo, temos feita a prova de 1.

2. Dado que X é um espaço métrico compacto, então existe uma sub-cobertura finita por bolas abertas, digamos que $\beta = \{B_1, ..., B_r\}$ é tal cobertura finita formada por bolas de raio menor que $\dfrac{\delta}{2}$ e por o item (1), temos que $\mu(\partial B_j) = 0$ para $j = \{0, ..., r\}$.

Seja $A_1 = \overline{B_1}$ e para $n > 1$, seja $A_n = \overline{B_n} \setminus \bigcup\limits_{i=1}^{n-1} \overline{B_i}$, então $\xi = \{A_1, ..., A_r\}$ é uma partição de $(X, \mathcal{M})$ com $diam(A_n) < \delta$. Agora como $\partial A_n \subset \bigcup\limits_{i=1}^{n} \partial B_i$, temos que $0 \leq \mu(\partial A_n) \leq \sum \mu \partial B_i = 0$ então, $\mu(\partial A_n) = 0$ para todo $n \in \mathbb{N}$. Provando assim o item (2) e finalmente o lema.

Então $diam(C_i) < \varepsilon$ e $\mu(\partial C_i) = 0$ para cada i desde $\partial C_i \subset \bigcup\limits_{j=1}^{k} \partial B_j$. Além disso, dado que $diam(C_i) < \varepsilon$, cada elemento $\xi_n = \bigvee\limits_{i=0}^{n-1} T^{-1}\xi$ tem d_n-diâmetro menor que ε, e portanto cada elemento de ξ_n contém no máximo um ponto em E_n. Portanto, existe exatamente, um número $cardE_n$ de elementos de ξ_n com ν_n- medida igual a $\dfrac{1}{cardE_n}$ e portanto,

$$
\begin{aligned}
H_{\nu_n}(\xi_n) &= -\sum_{C\in\xi_n} \nu_n(C)\log\big(\nu_n(C)\big)\\
&= \log\,(cardE_n).
\end{aligned}
$$

Dado $m, n \in \mathbb{N}$, podemos escrever $n = qm + r$, onde $q \geq 0$ e $0 \leq r < m$. Temos então

$$
\xi_n = \xi_{qm+r} = \bigvee_{j=0}^{q-1} T^{-jm}\xi_m \vee \bigvee_{j=qm}^{qm+r-1} T^{-j}\xi
$$

e portanto, para cada $i = 0, ..., m-1$ a partição mensurável

$$
\eta_i = \bigvee_{j=0}^{q-1} T^{-jm-i}\xi_m \vee \left(\bigvee_{j=qm}^{qm+r-1} T^{-j}\xi \vee \xi_i\right)
$$

é um refinamento de ξ_n. Note que

$$
card\left(\bigvee_{j=qm}^{qm+r-1} T^{-j}\xi \vee \xi_i\right) \leq card(\xi)^{2m}.
$$

A última desigualdade se verifica da seguinte forma:

Seja $\alpha_1, ..., \alpha_r$ partições tais que $\alpha_i = \{A_1^i, ..., A_\theta^i\}$, cada α_i tem tantos elementos como a partição ξ, então $card(\alpha_1 \vee ... \vee \alpha_r) \leq (card\xi)^r \leq (card\xi)^m; r < m$. de maneira análoga como $\xi_i = \bigvee_{k=0}^{i-1} T^{-k}\xi$, então, $card(\xi) \leq (card\xi)^i$, logo

$$
\bigvee_{j=qm}^{qm+r-1} T^{-j}\xi \vee \xi_i \quad = \quad \left(\bigvee_{j=qm}^{qm+r-1} T^{-j}\xi\right) \vee \xi_i
$$

Note que,

$$
\begin{aligned}
card\left(\left(\bigvee_{j=qm}^{qm+r-1} T^{-j}\xi\right) \vee \xi_i\right) &\leq \big(card\xi\big)^{m+i}\\
&\leq \big(card\xi\big)^{m+m}\\
&= \big(card\xi\big)^{2m}.
\end{aligned}
$$

O seguinte segue da Proposição (3.0.1), Lema (3.0.1) e Exemplo (2) que

$$\begin{aligned}
H_{\nu_n}(\xi_n) &\leq H_{\nu_n}(\eta_i) \\
&= H_{\nu_n}\left(\bigvee_{j=0}^{q-1} T^{-jm-i}\xi_m \vee \bigvee_{j=qm}^{qm+r-1} T^{-j}\xi \vee \xi_i\right) \\
&= H_{\nu_n}\left(\bigvee_{j=0}^{q-1} T^{-jm-i}\xi_m\right) + H_{\nu_n}\left(\bigvee_{j=qm}^{qm+r-1} T^{-j}\xi \vee \xi_i\right) \\
&\leq H_{\nu_n}\left(\bigvee_{j=0}^{q-1} T^{-jm-i}\xi_m\right) + \log(card\ \xi)^{2m} \\
&= H_{\nu_n}\left(\bigvee_{j=0}^{q-1} T^{-jm-i}\xi_m\right) + \ 2m\log(card\xi) \qquad (5.6) \\
&\leq \sum_{j=0}^{q-1} H_{\nu_n}\left(T^{-jm-i}\xi_m\right) + \ 2m\log(card\xi).
\end{aligned}$$

Agora observemos que dado que a função ψ é concava, temos que

$$\begin{aligned}
\frac{1}{n}\sum_{i=0}^{m-1}\sum_{j=0}^{q-1} H_{\nu_n}\big(T^{-jm-i}\xi_m\big) &\leq \frac{1}{n}\sum_{\ell=0}^{n-1} H_{\nu_n}(T^{-\ell}\xi_m) \\
&= \sum_{A\in\xi_m}\sum_{\ell=0}^{n-1}\frac{1}{n}\psi\big(\nu_n(T^{-\ell}(A))\big) \\
&\leq \sum_{A\in\xi_m}\psi\left(\sum_{\ell=0}^{n-1}\left(\frac{1}{n}\nu_n(T^{-\ell}(A))\right)\right) \\
&= \sum_{A\in\xi_m}\psi\left(\frac{1}{n}\sum_{\ell=0}^{n-1}(T_*\nu_n)(A)\right) \\
&= \sum_{A\in\xi_m}\psi\Big(\mu_n(A)\Big) \\
&= H_{\mu_n}(\xi_m).
\end{aligned}$$

E dado que em (5.3) se cumpre para $i = 0, ..., m-1$, obtemos

$$\frac{m}{n} H_{\nu_n}(\xi_n) \leq H_{\mu_n}(\xi_m) + \frac{2m^2}{n} \log card\ \xi,$$

isto é

$$\frac{1}{n} H_{\nu_n}(\xi_n) \leq \frac{H_{\mu_n}(\xi_m)}{m} + \frac{2m}{n} \log card\ \xi.$$

Por outro lado temos

$$\frac{1}{n} \log card\ E_n = \frac{1}{n} H_{\nu_n}(\xi_n).$$

Substituindo n por ℓ_n temos

$$\frac{1}{\ell_n} \log card\ E_{\ell_n} = \frac{1}{\ell_n} H_{\nu_{\ell_n}}(\xi_{\ell_n}) \leq \frac{H_{\mu_{\ell_n}}(\xi_m)}{m} + \frac{2m}{\ell_n} \log card\ \xi$$

Por outro lado, por (4.3), tomando $n \to \infty$ obtemos

$$\begin{aligned} \lim_{n\to\infty} \sup \frac{1}{n} \log card\ E_n &= \lim_{n\to\infty} \sup \frac{1}{n} \log N_d(n, \varepsilon) \\ &= \lim_{n\to\infty} \sup \frac{1}{\ell_n} \log card\ E_{\ell_n} \qquad (5.7) \\ &\leq \frac{1}{m} \lim_{n\to\infty} \sup H_{\mu_{\ell_n}}(\xi_m). \end{aligned}$$

Agora faremos uso do seguinte lema para mostrar que o limite da última linha de (5.7) converge para $H_\mu(\xi_m)$ e usaremos o requisito de que $\mu(\partial B_i) = 0$ para $i = 1, ..., k$.

Lema 5.0.3. *Seja $(\mu_n)_{n\in\mathbb{N}}$ uma sequência de medidas convergentes para μ e seja ξ uma partição mensurável tal que $\mu(\partial A) = 0; \forall\ A \in \xi$, então*

$$\lim_{n\to\infty} \sup H_{\mu_n}(\xi) = \lim_{n\to\infty} H_{\mu_n}(\xi) = H_\mu(\xi).$$

Prova: Seja A um conjunto mensurável de ξ. Dado que $(\mu_n)_{n\in\mathbb{N}}$ converge para μ, se $\varphi_k : X \to [0, 1]$ é uma sequência de funções contínuas decrescendo para $\chi_{\overline{A}}$ quando $k \to \infty$, então

$$\begin{aligned} \lim_{n\to\infty} \sup H_{\mu_n}(\overline{A}) &\leq \lim_{n\to\infty} \sup \int_X \varphi_k\, d\mu_n \\ &= \int_X \varphi_k\, d\mu. \end{aligned}$$

Por outra parte e usando o Teorema da Convergência Monótona (T.C.M)(1.3.1),

$$\begin{aligned}\lim_{k\to\infty}\int_X \varphi_k\, d\mu &= \int_X \lim_{k\to\infty}\varphi_k\, d\mu\\ &= \int_X \chi_{\overline{A}}\, d\mu\\ &= \mu(\overline{A}).\end{aligned}$$

Dado que $\mu(\partial A)=0$; isto é,

$$\begin{aligned}0 &= \mu(\partial A)\\ &= \mu(\overline{A}\backslash intA)\\ &= \mu(\overline{A})-\mu(intA)\\ &= \mu(\overline{A})=\mu(intA)=\mu(A).\end{aligned}$$

Assim,

$$\lim_{n\to\infty}\sup\mu_n(A)\le\lim_{n\to\infty}\sup\mu_n(\overline{A})\le\mu(\overline{A})=\mu(intA)=\mu(A). \tag{5.8}$$

Concluindo que

$$\lim_{n\to\infty}\sup\mu_n(A)\le\mu(A).$$

Por outro lado, sabendo que $\partial(X\setminus A)=\partial(A)$, também temos

$$\lim_{n\to\infty}\sup\mu_n(X\setminus A)\le\mu(X\setminus A).$$

e finalmente obtemos que

$$\lim_{n\to\infty}\inf\mu_n(A)\ge\mu(A). \tag{5.9}$$

Então, de (5.8) e (5.9) concluísse que,

$$\lim_{n\to\infty} \mu_n(A) = \mu(A). \tag{5.10}$$

Agora usando a continuidade de ψ temos o seguinte;

$$\begin{aligned}
\lim_{n\to\infty} H_{\mu_n}(\xi) &= \lim_{n\to\infty} \sum_{A\in\xi} \psi\big(\mu_n(A)\big) \\
&= \sum_{A\in\xi} \psi\big(\lim_{n\to\infty} \mu_n(A)\big) \\
&= \sum_{A\in\xi} \psi\big(\mu_n(A)\big) \\
&= H_\mu(\xi).
\end{aligned}$$

Para finalizar a demostração, temos que se $\mu(\partial C) = 0; \forall\ C \in \xi$, temos que

$$\mu(\partial C) = 0;\ para\ C \in \xi_n.$$

Logo, usando o Lema (5.0.3) em (5.7), obtemos,

$$\begin{aligned}
\lim_{n\to\infty} \sup \frac{1}{n} \log N_d(n,\varepsilon) &\leq \frac{1}{m} H_\mu(\xi_m) \\
\lim_{m\to\infty} \lim_{n\to\infty} \sup \frac{1}{n} \log N_d(n,\varepsilon) &\leq \lim_{m\to\infty} \frac{1}{m} H_\mu(\xi_m) \\
\lim_{n\to\infty} \sup \frac{1}{n} \log N_d(n,\varepsilon) &\leq h_\mu(T,\xi) \\
&\leq h_\mu(T) \\
&= \sup\{h_\mu(T,\xi); \xi \text{ é partição de } X\ \}.
\end{aligned} \tag{5.12}$$

Portanto,

$$\begin{aligned}
\lim_{n\to\infty}\sup\frac{1}{n}\log N(d_n,\varepsilon) &\leq \sup_{\nu} h_\nu(T)\\
\lim_{\varepsilon\to 0}\lim_{n\to\infty}\sup\frac{1}{n}\log N(d_n,\varepsilon) &\leq \sup_{\nu} h_\nu(T)\\
h_{top}(T) &\leq \sup_{\nu} h_\nu(T).
\end{aligned}$$

Assim, finalmente obtemos o resultado esperado

$$h_{top}(T) = \sup\{h_\mu(T) : \mu \text{ é uma medida } T\text{-invariante em } \mathcal{M}(X)\}.$$

$\square$

Teorema 5.0.2. *Seja (X, d) um espaço métrico compacto e $f : X \to X$ contínua. Então $h_{top}(f) = h_{top}(f|_{\Omega(f)})$*

Antes de mostrar o Teorema (5.0.2) mostraremos alguns lemas que serão de utilidade para provar o Teorema.

Lema 5.0.4. *Seja $f : X \to X$ uma aplicação contínua e X um espacio métrico compacto. O conjunto $\Omega(f)$ é fechado.*

Prova: Basta mostrar que o conjunto dos pontos errantes é aberto.

Seja x um ponto errante arbitrário, então existe uma vizinhança U de x tal que $U \cap f^n(U) = \emptyset, \forall\ n \in \mathbb{N}$.

Dado $z \in U$, tomemos W uma vizinhança de z, de tal forma que $W \subset U$, como temos que $f^n(W) \subset f^n(U)$, então $f^n(W) \cap W \subset f^n(U) \cap U = \emptyset$, logo z é um ponto interior de $f^n(U) \cap U$. Assim, $\Omega(f)$ é fechado. $\square$

Observação: Do lema imediatamente anterior, se mostra facilmente que o conjunto $\Omega(f)$ é compacto.

Lema 5.0.5. *O conjunto $\Omega(f)$ é invariante.*

Prova: Seja $x \in \Omega(f)$. Considere V uma vizinhança de $f(x)$. Portanto, $U = f^{-1}(V)$ é uma vizinhança de x, assim existe $m \in \mathbb{N}$ tal que $f^m(U) \cap U \neq \emptyset$. Como

$$\begin{aligned} \emptyset \neq f(f^m(U) \cap U) &\subset f^m(f(U)) \cap f(U) \\ &= f^m(f(f^{-1}(V)) \cap f(f^{-1}(V)) \\ &= f^m(V) \cap V. \end{aligned}$$

Segue que $f(x) \in \Omega(f)$. $\square$

Observação: Se f é um homeomorfismo $\Omega(f)$ é completamente invariante.

Lema 5.0.6. *O conjunto $\Omega(f)$ é não vazio.*

Para mostrar este lema, daremos em primeiro lugar a definição de Ômega limite $(\omega_f(x))$.

Definição 5.0.1. *Para um ponto $x \in X$ o seu conjunto $\omega - limite$ é o conjunto de todos os pontos de acumulação da sua órbita positiva,*

$$\omega_f(x) = \{y \in X | \exists n_k \to \infty \ tal\ que\ f^{n_k} \to y\}.$$

Lema 5.0.7. *O conjunto $\omega_f(x) \subset \Omega(f)$.*

Prova: É claro que se X é compacto, então $\omega_f(x) \neq \emptyset$ ou seja, dada $(x_n)_{n\in\mathbb{N}}$ uma sequência, existe uma subsequência $(x_{n_k})_{k\in\mathbb{N}}$ convergente a $y \in X$, logo $f^{n_k}(x) \to y$, portanto $y \in \omega_f(x)$.

Seja $x \in X$ fixo e $y \in \omega_f(x)$. Consideremos U uma vizinhança de y. Basta encontrar um $k \in \mathbb{N}$ tal que $f^k(U) \cap U \neq \emptyset$ ou em outras palavras devemos encontrar um $z \in U$ tal que $f^k(z) \in U$.

Seja $k_n \to \infty$ tal que $f^{k_n}(x) \to y$, portanto, existem $k_0 < k_1$ tal que
$f^{k_i}(x) \in U; i = 0, 1$. Tomando $z = f^{k_0}(x)$ e $k = k_1 - k_0$ assim conseguimos que $f^k(z) \in U$, logo $f^k(U) \cap U \neq \emptyset$; implicando que $\omega_f(x) \subset \Omega(f)$. Portanto $\Omega(f)$ é não vazio.

$\square$

Corolário 5.0.1. *Para qualquer medida f invariante de probabilidade boreliana μ, vale que $\mu(\Omega(f)) = 1$.*

Prova: Primeiro definiremos o conjunto de pontos errantes como sigue:

$$(\Omega(f))^c = \{x \in X | \exists\ U \text{ aberto que contem } x, \forall\ n > 0 \text{ tal que } f^n(U) \cap U = \emptyset\}.$$

Agora podemos escrever $X = \Omega(f) \cup (\Omega(f))^c$, logo aplicando μ temos:

$$\begin{aligned}\mu(X) &= \mu(\Omega(f) \cup (\Omega(f))^c) \\ \\ 1 &= \mu(\Omega(f)) + \mu((\Omega(f))^c).\end{aligned}$$

Pelo teorema de Recorrência versão topológica temos que

$$1 = \mu(\Omega(f)) + 0.$$

Logo $\mu(\Omega(f)) = 1$, para qualquer μ invariante.

$\square$

Prova do Teorema (5.0.2): Seja $f : X \to X$ é uma função e X um espaço métrico compacto. Pelo Corolário (5.0.1) temos que $\mu(\Omega(f)) = 1, \forall \mu$ medida de probabilidade f-invariante.

Seja $A \subset X$ então,

$$\begin{aligned}\mu(A) &= \mu(A \cap \Omega(f)) + \mu(A \cap (\Omega(f))^c) \\ \\ &= \mu(A \cap \Omega(f)) + 0 \\ \\ &= \mu(A \cap \Omega(f)).\end{aligned}$$

Seja $g = f|_{\Omega(f)} : \Omega(f) \to \Omega(f)$. Dado que μ é f- invariante e definindo $\nu_\mu(A) = \mu(A \cap \Omega(f))$, obtemos que ν_μ é g- invariante, ou seja $\mu \longmapsto \nu_\mu$. Logo $\mu(A) = \nu_\mu(A)$.

Seja P uma partição de X, $P = \{p_i\}_{i=1}^n$, então,

$$\sum_{i=1}^{n} \mu(p_i)\log(\mu(p_i)) = \sum_{i=1}^{n} \nu_\mu(p_i)\log(\nu_\mu(p_i)).$$

Obtendo que $h_\mu(f) = h_{\nu_\mu}(g)$,então,

$$h_{top}(f) = \sup_{\mu\in\mathcal{M}(f)}\{h_\mu(f)\} = \sup_{\mu\in\mathcal{M}(f)}\{h_{\nu_\mu}(g)\} \leq \sup_{\nu\in\mathcal{M}(g)}\{h_\nu(g)\} = h_{top}(g).$$

Logo, $h_{top}(f) \leq h_{top}(g)$.

Também temos que:

$$h_{top}(f) \leq h_{top}(g) = h_{top}(f|_{\Omega(f)}) \leq h_{top}(f).$$

Logo, $h_{top}(g) \leq h_{top}(f)$, obtendo que :

$$h_{top}(f) = h_{top}(g).$$

$\square$

Referências Bibliográficas

[1] A. Katok, B. Hosselblatt, **Introduction to the Modern Theory of Dinâmical Systems.** Cambrigde University Press (1996).

[2] M. Brin, G. Stuck, **Introduction to Dynamical Systems.** Cambrigde University Press (2002).

[3] P. Walters, **Introduction to Ergodic Theory**. Graduate Text In Mathematics. Springer Verlag, New York-Berlin, (1982).

[4] C. Shannon, **A Mathematical Theory of comunication**, Bell Systems Technical Journal (1948).

[5] K. Oliveira, M. Viana. **Foudations of Ergodic Theory**, Cambridge University Press, 2016.

[6] Adler, R., Konheim, A. e Mcandrew, H. **"Topological Entropy."** Transactions of the American Mathematical Society, vol. 114, no. 2, 1965, pp. 309–319.

[7] Dinaburg, E. I. **A correlation between topological entropy and metric entropy.** Doklady Akademii Nauk SSSR, v. 190, n. 1, p. 19-22, 1970.

[8] Goodwyn, L. W. **Topological entropy bounds measure-theoretic entopy.** Proceedings of the American Mathematical Society, v. 23, n. 3, p. 679-688, 1969.

[9] Goodman, T. N. T. **Relating topological entropy and measure entropy.** Bulletin of the London Mathematical Society, v. 3, n. 3, p. 176-180, jul. 1971.

[10] Bowen, R. **Entropy for group endomorphism and homogeneous spaces.** Transactions of the American Mathematical Society, n. 153, p. 401-414, Jan 1971.

[11] Casimiro, J.P. **Metric Entropy and Topological Entropy: The Variational Principle**. Dissertação (Mestrado em matemáticas e aplacações)-Departamento de matemática. Universidade Técnica de Lisboa, (2014).

[12] Melo, K. **Difeomorfismos de Anosov**. Dissertação (Mestrado em matemáticas). Uberlândia-MG, (2016).

[13] Neira, C. **Topología general**. Bogotá, Colombia: Universidad Nacional de Colombia. 2011.

[14] Munkres,J. R. **Topology a first course**, Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1975.

[15] Lima, E. L. **Curso de análise, volume 1.** 15ª Edição. Rio de Janeiro: IMPA, 2022.