ALEXANDRE DE SOUZA FERNANDES

UMA RELAÇÃO ENTRE EXPOENTES DE LYAPUNOV E ENTROPIA

UNIVERSIDADE FEDERAL DE UBERLÂNDIA
FACULDADE DE MATEMÁTICA
2017

ALEXANDRE DE SOUZA FERNANDES

UMA RELAÇÃO ENTRE EXPOENTES DE LYAPUNOV E ENTROPIA

Dissertação apresentada ao Programa de Pós-Graduação em Matemática da Universidade Federal de Uberlândia, como parte dos requisitos para obtenção do título de **MESTRE EM MATEMÁTICA.**

Área de Concentração: Matemática.
Linha de Pesquisa: Sistemas Dinâmicos.

Orientador: Prof. Dr. Thiago Aparecido Catalan.

UBERLÂNDIA - MG
2017

Dados Internacionais de Catalogação na Publicação (CPI)
Sistema de Bibliotecas da UFU, MG, Brasil.

F363r
2017

Fernandes, Alexandre de Souza, 1990 -
Uma relação entre expoentes de Lyapunov e entropia / Alexandre de Souza Fernandes. - 2017
68 f. : il.

Orientador: Thiago Aparecido Catalan.
Dissertação (mestrado) - Universidade Federal de Uberlândia. Programa de Pós-Graduação em Matemática.
Inclui bibliografia.

1. Matemática - Teses. 2. Entropia - Teses. 3. Teoria ergódica - Teses. 4. Lyapunov, Funções de - Teses. I. Catalan, Thiago Aparecido. II. Universidade Federal de Uberlândia. Programa de Pós-Graduação em Matemática. III. Título.

CDU: 51

UNIVERSIDADE FEDERAL DE UBERLÂNDIA
FACULDADE DE MATEMÁTICA
PROGRAMA DE PÓS-GRADUAçÃO EM MATEMÁTICA
Av. João Naves de Ávila, 2121, Bloco 1F, Sala 1F 152
Campus Santa Mônica, Uberlândia - MG, CEP 38400-902

ALUNO: Alexandre de Souza Fernandes.

NÚMERO DE MATRÍCULA: 11512MAT001.

ÁREA DE CONCENTRAÇÃO: Matemática.

LINHA DE PESQUISA: Sistemas Dinâmicos.

PÓS-GRADUAÇÃO EM MATEMÁTICA: Nível Mestrado.

TÍTULO DA DISSERTAÇÃO: Uma relação entre expoentes de Lyapunov e entropia.

ORIENTADOR: Prof. Dr. Thiago Aparecido Catalan.

Esta dissertação foi **APROVADA** em reunião pública realizada na Sala 1F119, Bloco 1F, Campus Santa Mônica, em 24 de março de 2017, às 13h30min, pela seguinte Banca Examinadora:

NOME	ASSINATURA
Prof. Dr. Thiago Aparecido Catalan UFU - Universidade Federal de Uberlândia (orientador)	
Prof. Dr. Alexander Eduardo Arbieto Mendoza UFRJ - Universidade Federal do Rio de Janeiro	
Prof. Dr. Jean Venato Santos UFU - Universidade Federal de Uberlândia	

Uberlândia-MG, 24 de Março de 2017.

Dedicatória

Esta obra é dedicada à determinação: determinação é o que difere do usual ao genial, do imaginário ao real, do abstrato ao concreto. Determinação não é um estado de espírito, é uma vontade impetuosa, concebida porventura de uma complexa epifania ou de uma singela palavra. Não confunda determinação com força . . . determinação é o que nos torna iluminados.

Agradecimentos

Meus agradecimentos serão específicos, porém objetivos:

Agradeço a minha primeira família, "a grande farofa". Aquela que me transborda. Aquela que viu meu prelúdio, meu meio, e estarão comigo até meu desfecho.

Agradeço ao meu pai, José Rubens, por nunca desistir de mim, agradeço minha mãe, Maria Cristina, por mostrar o exemplo mais intenso, mais simples e mais absurdo possível de amor e carinho que já presenciei. Agradeço a minha irmã, Aliucha, por ser um exemplo de força e companheirismo, mais ainda, mostrar que apesar das diferenças (e brigas) entre irmãos, sempre poderão se amar, se respeitar e cuidar um do outro.

Agradeço minha segunda família, os "brabos", meus amigos/irmãos Agostinho, Lucas, Gian, Felipe, Rafael, Fernando, Ítalo, Valdecir. Por toda nossa história. Por todas as aventuras e desventuras que passamos juntos.

Agradeço aos meus amigos que fiz em Uberlândia, em especial dois grandes, Douglas e Ueslei, um sendo meu conterrâneo e outro um "verdadeiro" exemplo de mineiro. São grandes amigos pois ambos me ensinaram as maiores lições de vida, a de sempre prezar pelos outros não esperando absolutamente nada em troca e, de ser autêntico, a de não ser um total "clichê".

Agradeço aos meus professores e amigos que fiz na graduação e mestrado em matemática da Universidade Federal de Uberlândia. Aos meus amigos de Augusto, Danilo, Suélen, Wagner, Guilherme, José Lucas, Edmilson, Magna, Davidson, Javier, Julian, Aloísio, Paulo Victor e Ana Maria pelos momentos de assistência e auxílio nos estudos e pelos momentos de alegria e descontração. Aos professores, por todos os ensinamentos acadêmicos e sociais, além de apresentarem a maior das ciências, a do conhecimento. Em particular, agradeço ao meu orientador Thiago Aparecido Catalan pela conclusão deste trabalho.

Agradeço à Capes pelo auxílio financeiro durante todo o curso de mestrado.

Agradeço também aos professores Jean Venato Santos e Alexander Eduardo Arbieto Mendoza por terem aceito o convite para fazerem parte da minha banca.

FERNANDES, A. S. *Uma relação entre expoentes de Lyapunov e entropia.* 2016. - 66p. Dissertação de Mestrado, Universidade Federal de Uberlândia, Uberlândia-MG.

Resumo

Nesta dissertação, vamos estabelecer uma relação entre os expoentes de Lyapunov, dados pelo teorema Ergódico Multiplicativo de Oseledets, e a entropia métrica, dada pela definição de Komolgorov-Sinai. Para tal utilizaremos como ferramentas o estudo de teoria ergódica. Pelo meio, buscamos analisar e demonstrar o teorema multiplicativo de Oseledets, tendo em vista caracterizar os expoentes de Lyapunov em uma variedade Riemmaniana compacta e por fim, apresentaremos a desigualdade de Ruelle, resultado final que faz esta relação entre os expoentes positivos e a entropia métrica de um sistema ergódico.

Palavras-chave: Entropia; Teorema Multiplicativo de Oseledets; Expoente de Lyapunov; Cociclo Linear; Desigualdade de Ruelle.

FERNANDES, A. S. *A relation between Lyapunov exponents and entropy.* 2016. - 66p. M. Sc. Dissertation, Federal University of Uberlândia, Uberlândia-MG.

Abstract

In this dissertation, we are going to stabilish a relation between Lyapunov exponents, given by Oseledets Multiplicative Ergodic Theorem, and metric entropy, owing to the definition given by Komolgorov-Sinai. For this, we are going to use as tools the study of ergodic's theory. By the middle, we want to analise and demonstrate the Oseledets multiplicative theorem, that categoryzes the Lyapunov's expoents of a compact Rimmenian manyfold and finally, introduce the Ruelle's inequallity, the final result that stabilished the relation between the positive expoents and the metric entropy of a ergodic sistem.

Keywords: Entropy; Oseledets Multiplicative Theorem; Lyapunov exponents; Linear Cocyclic; Ruelle's Inequallity.

SUMÁRIO

INTRODUÇÃO

Com seus estudos sobre estabilidade de soluções de equações diferenciais em sistemas dinâmicos, a parte fundamental do trabalho do matemático russo Aleksandr Mikhailovich Lyapunov (1857-1918) foi desenvolver um método que estuda o comportamento assintótico da contração ou expansão de um ponto ao longo de sua órbita. O expoente de Lyapunov (caso não for nulo) visa estabelecer uma previsibilidade deste comportamento assintótico.

Sobre os trabalhos envolvendo leis da termodinâmica, a conservação de energia e entropia, o matemático russo Yakov Sinai, observou que os sistemas conservativos da mecânica clássica verificam a lei da conservação da energia e estão ligados à hipótese da ergodicidade. Mais ainda, notou que a divergência exponencial de órbitas originada em pontos próximos está relacionada a entropia positiva nos sistemas dinâmicos. Foi daí então que Valery Oseledets, aluno de pós-graduação orientando de Sinai, se interessou pelo problema de divergência exponencial.

Aqui, propomos em primeiro lugar medir a complexidade de um sistema, a saber calcular sua *entropia*, onde utilizamos a definição dada pelos matemáticos soviéticos Andrey Komolgorov e Sinai que buscavam definir de maneira adequada o cálculo da entropia de um sistema na Teoria Ergódica. Uma outra boa forma de medir a complexidade de um sistema é através dos expoentes de Lyapunov. Assim sendo, é natural buscar uma relação entre expoentes de Lyapunov e entropia.

Seguidamente, vamos primeiro demonstrar o teorema Multiplicativo de Oseledets, um importante resultado que irá garantir, sob certas condições, a existência do expoente de Lyapunov, para quase todo ponto em relação à uma medida invariante para a dinâmica. Isto foi provado exatamente por Valery Oseledets em 1967 e publicado em 1968 , usando a teoria de *cociclos lineares*, teoria clássica dentro da teoria de sistemas dinâmicos e teoria ergódica ([9]). Feito isso, demonstraremos o resultado obtido pelo físico teórico

David Ruelle, um dos criadores da teoria ergódica, que conseguiu relacionar o cálculo da entropia e os expoentes de Lyapunov ([8]).

CAPÍTULO 1

TEORIA ERGÓDICA

Neste capítulo vamos relembrar algumas definições e resultados da teoria ergódica, necessárias para introduzir de maneira mais apropriada os resultados principais da dissertação. Tais definições e resultados podem ser vistas na referência bibliográfica [7].

1.1 Medida Invariante e o teorema ergódico de Birkhoff

Apresentaremos nesta seção, definições e resultados importantes necessárias para o conceito de ergodicidade. Será considerado no espaço mensurável X, medidas boreleanas μ que no nosso caso serão medidas de probabilidade, isto é, $\mu(X) = 1$.

Definição 1.1. *Seja* $(X, \mathcal{A}, \mu)$ *um espaço de medida. Dizemos que uma medida* μ *é* ***invariante*** *por uma função* $f : X \longrightarrow X$ *se*

$$\mu(E) = \mu(f^{-1}(E)), \text{ para todo conjunto mensurável } E \subset X.$$

Dizemos também que f ***preserva*** μ *e, em particular, que o subconjunto mensurável* $E \subset X$ *é* ***invariante*** *por* f *(ou* f*-invariante) com respeito à medida* μ.

Definição 1.2. *Seja* $(X, \mathcal{A}, \mu)$ *um espaço de medida. Considerando* $f : X \longrightarrow X$ *uma função mensurável e* E *um conjunto mensurável de* X*, definimos pot*

$$\tau_n(E, x) = \frac{1}{n}\#\left\{j \in \{0,\ 1, \ldots, n-1\};\ f^j(x) \in E\right\}$$

o tempo médio de permanência dentre os n-iterados de x *no conjunto* E.

Podemos escrever também por

$$\tau_n(E,x) = \frac{1}{n}\sum_{j=0}^{n-1}\chi_E(f^j(x)),$$

sendo χ_E *a função característica em* E. *Dizemos que o* ***tempo médio de permanência da órbita de*** x ***em*** E *é dado por*

$$\tau(E,x) = \lim_{n\longrightarrow\infty}\tau_n(E,x).$$

Definição 1.3. *Uma função* $\psi : X \longrightarrow \mathbb{R}$ *é dita* ***invariante*** *por uma função* $f : X \longrightarrow X$ *com respeito a uma medida* μ *(ou simplesmente* f***-invariante****)se para* μ*-quase todo ponto* $x \in X$ *vale* $\psi(f(x)) = \psi(x)$. *Além disso, dizemos que um conjunto mensurável* $E \subset X$ *é* ***invariante*** *se sua função característica* χ_E *é uma função invariante.*

Teorema 1.4. *[Birkhoff] Seja* $f : X \longrightarrow X$ *uma função mensurável e* μ *uma medida de probabilidade invariante por* f. *Dada qualquer função integrável* $\varphi : X \longrightarrow \mathbb{R}$ *o limite*

$$\tilde{\varphi}(x) = \lim_{n\longrightarrow\infty}\frac{1}{n}\sum_{j=0}^{n-1}\varphi(f^j(x))$$

existe para μ*-quase todo ponto* $x \in X$. *Além disso, função* $\tilde{\varphi}$ *definida desta forma é integrável e*

$$\int \tilde{\varphi}(x)d\mu(x) = \int \varphi(x)d\mu(x).$$

Corolário 1.5. *Seja* $f : X \longrightarrow X$ *uma função mensurável e* μ *uma medida de probabilidade invariante por* f. *Dada qualquer função integrável* $\varphi : X \longrightarrow \mathbb{R}$*, temos que*

$$\tilde{\varphi}(x) = \tilde{\varphi}(f(x))$$

para μ*-q.t.p.* $x \in X$.

Antes de demonstrar o corolário, precisamos do seguinte lema:

Lema 1.6. *Seja* $f : X \longrightarrow X$ *uma aplicação mensurável e considere o espaço de probabilidade* $(X, \mathcal{A}, \mu)$ *com* μ *invariante sobre* f. *Considerando a função* $C : X \longrightarrow \mathbb{R}$ *uma função mensurável tal que* $C \circ f - C$ *é* μ*-integrável. Então*

$$\lim_{n\longrightarrow+\infty}\frac{1}{n}C\circ f^n = 0, \text{ para } \mu\text{-q.t.p } x \in X.$$

Demonstração: Como $C \circ f - C$ é μ-integrável por hipótese, o limite dado no Teorema 1.4 definido pela função $\tilde{\varphi} : X \longrightarrow \mathbb{R}$

$$\tilde{\varphi}(x) = \lim_{n\longrightarrow+\infty}\frac{1}{n}\sum_{j=0}^{n-1}(C\circ f - C)(f^j(x))$$

existe para μ-q.t.p. $x \in X$. Assim,

$$\begin{aligned}
\tilde{\varphi}(x) &= \lim_{n \longrightarrow +\infty} \frac{1}{n} \sum_{j=0}^{n-1} (C \circ f - C)(f^j(x)) \\
&= \lim_{n \longrightarrow +\infty} \frac{1}{n} \sum_{j=0}^{n-1} (C(f^{j+1}(x)) - C(f^j(x)) \\
&= \lim_{n \longrightarrow +\infty} \frac{1}{n} \left(\sum_{j=0}^{n-2} C(f^{j+1}(x))) + C(f^n(x)) - \sum_{j=0}^{n-1} C(f^j(x)) \right) \\
&= \lim_{n \longrightarrow +\infty} \frac{1}{n} \left(\sum_{j=0}^{n-2} C(f^{j+1}(x))) + C(f^n(x)) - \sum_{j=1}^{n-1} C(f^j(x)) - C(x) \right) \\
&= \lim_{n \longrightarrow +\infty} \frac{1}{n} \left(\sum_{j=1}^{n-1} C(f^{j}(x))) + C(f^n(x)) - \sum_{j=1}^{n-1} C(f^j(x)) - C(x) \right) \\
&= \lim_{n \longrightarrow +\infty} \frac{1}{n} C(f^n(x)) - \lim_{n \longrightarrow +\infty} \frac{1}{n} C(x) \\
&= \lim_{n \longrightarrow +\infty} \frac{1}{n} (C \circ f^n)(x).
\end{aligned}$$

Agora para cada $\delta > 0$, e olhando para o conjunto $\left\{ x \in X;\ \frac{1}{n} |C \circ f^n(x)| \geq \delta \right\}$, obtemos

$$\begin{aligned}
\mu \left(\left\{ x \in X;\ \frac{1}{n} |C \circ f^n(x)| \geq \delta \right\} \right) &= \mu \left(\left\{ x \in X;\ \frac{1}{n} C \circ f^n(x) \geq \delta \ \text{ ou } \ \frac{1}{n} C \circ f^n(x) \leq -\delta \right\} \right) \\
&= \mu \left(\{ x \in X;\ C \circ f^n(x) \geq n\delta \ \text{ ou } \ C \circ f^n(x) \leq -n\delta \} \right) \\
&= \mu \left(\left\{ x \in X;\ C \circ f^n(x) \in (-n\delta, n\delta)^C \right\} \right) \\
&= \mu \left(f^{-n} \circ C^{-1}(-n\delta, n\delta)^C \right) \\
&= \mu(C^{-1}(-n\delta, n\delta)^C).
\end{aligned}$$

Assim, fazendo n tender à infinito, temos que $\mu(C^{-1}(-n\delta, n\delta)^C)$ tenderá a zero. Como $\delta > 0$ foi tomado de modo arbitrário, podemos concluir que $(C \circ f^n)/n$ converge para zero (em medida), quando n tender a infinito.

Como o limite $((C \circ f^n)(x))/n$ existe para μ-q.t.p. $x \in X$, temos que este limite deverá ser zero. Isto é,

$$\lim_{n \longrightarrow +\infty} \frac{1}{n} C \circ f^n = 0,$$

para μ-q.t.p. $x \in X$. ■

Demonstração (Corolário 1.5): Por definição,

$$\begin{aligned}
\tilde{\varphi}(f(x)) &= \lim_{n\longrightarrow\infty} \frac{1}{n}\sum_{j=1}^{n} \varphi(f^j(x)) \\
&= \lim_{n\longrightarrow\infty} \frac{1}{n}\sum_{j=0}^{n-1} \varphi(f^j(x)) + \frac{1}{n}\left[\varphi(f^n(x)) - \varphi(x)\right] \\
&= \tilde{\varphi}(x) + \lim_{n\longrightarrow\infty} \frac{1}{n}\left[\varphi(f^n(x)) - \varphi(x)\right].
\end{aligned}$$

Pelo Lema 1.6, tomando $C = \varphi$, temos que $(\varphi(f^n(x)) - \varphi(x))/n$ tende à 0, quando n tende a infinito, para μ-q.t.p. $x \in X$. Isto encerra a demonstração. ■

Definição 1.7. *Seja* $(X, \mathcal{A}, \mu)$ *um espaço de probabilidade. A medida de probabilidade* μ *diz-se ergódica para a função* $f : X \longrightarrow X$ *(ou que* f *diz-se ergódica relativamente a* μ*, ou que o sistema* (f, μ) *é ergódico) se as médias temporais dadas pelo Teorema de Birkhoff coincidem em quase todo ponto com as médias espaciais, isto é,*

$$\tilde{\varphi}(x) = \lim_{n\longrightarrow\infty} \frac{1}{n}\sum_{j=0}^{n-1} \varphi(f^j(x)) = \int \varphi d\mu$$

para toda função integrável $\varphi : X \longrightarrow \mathbb{R}$ *em* μ*-q.t.p.* $x \in X$.

Veremos a seguir um importante resultado que nos diz um pouco mais sobre ergodicidade, tanto nas várias maneiras equivalentes que podemos defini-lá quanto a entender o seu significado.

Proposição 1.8. *Seja* μ *uma probabilidade invariante de uma aplicação* $f : X \longrightarrow X$ *mensurável. As seguintes condições são equivalentes:*

(1) Para todo conjunto mensurável $E \subset X$ *tem se* $\tau(E, x) = \mu(B)$ *para* μ*-q.t.p* $x \in X$.

(2) Para todo conjunto mensurável $E \subset X$*, a função* $\tau(E, \cdot)$ *é constante para* μ*-q.t.p.* $x \in X$.

(3) Para toda função integrável $\varphi : X \longrightarrow \mathbb{R}$ *tem-se* $\tilde{\varphi}(x) = \int \varphi d\mu$ *para* μ*-q.t.p.* $x \in X$.

(4) Para toda função integrável $\varphi : X \longrightarrow \mathbb{R}$*, a média temporal* $\tilde{\varphi} : X \longrightarrow \mathbb{R}$ *é constante para* μ*-q.t.p.* $x \in X$.

(5) Toda função integrável f*-invariante* $\psi : X \longrightarrow \mathbb{R}$ *tem-se* $\psi(x) = \int \psi d\mu$ *para* μ*-q.t.p.* $x \in X$.

(6) Toda função integrável f*-invariante* $\psi : X \longrightarrow \mathbb{R}$ *é constante em* μ*-q.t.p.* $x \in X$.

(7) Para todo subconjunto invariante A *de* X *tem se que* $\mu(A) = 0$ *ou* $\mu(A) = 1$.

Demonstração: **(1)** $\Rightarrow$ **(2)** Imediato!

(2) $\Rightarrow$ **(3)** Teorema 1.4.

(3) $\Rightarrow$ **(4)** Consequência do Teorema 1.4.

(4) $\Rightarrow$ **(5)** Como ψ é integrável em particular, temos da hipótese em **(4)** que

$$\tilde{\psi}(x) = \lim_{n \longrightarrow \infty} \frac{1}{n} \sum_{j=0}^{n-1} \psi(f^j(x)) = \int \psi d\mu, \tag{1.1}$$

em μ-q.t.p. $x \in X$. Pela invariância de ψ por f, obtemos $\psi(f^j(x)) = \psi(x)$ em μ-q.t.p. $x \in X$, para todo $j \in \mathbb{N}$. Assim, da equação (1.2) temos que

$$\tilde{\psi}(x) = \lim_{n \longrightarrow \infty} \frac{1}{n} \sum_{j=0}^{n-1} \psi(f^j(x)) = \lim_{n \longrightarrow \infty} \frac{1}{n} \sum_{j=0}^{n-1} \psi(x) = \lim_{n \longrightarrow \infty} \frac{1}{n}(n\psi(x)) = \psi(x),$$

concluindo que

$$\psi(x) = \int \psi d\mu \tag{1.2}$$

em μ-q.t.p. $x \in X$.

(5) $\Rightarrow$ **(6)** Imediato!

(6) $\Rightarrow$ **(7)** Considere χ_A a função característica no subconjunto invariante A de X, isto é, $\chi_A(x) = 1$ se $x \in A$, ou $\chi_A(x) = 0$ se $x \notin A$, e $\mu(f^{-1}(A)) = \mu(A)$ para μ-q.t.p. $x \in X$.

Em particular, considerando que $\phi = \chi_A$, temos da hipótese em **(6)** que χ_A é invariante por f, é integrável e χ_A é constante em μ-q.t.p. $x \in X$. Mais ainda, temos que $\tau(A, \cdot)$ é constante em μ-q.t.p. $x \in X$, isto é, $\mu(A) = 0$ ou $\mu(A) = 1$.

(7) $\Rightarrow$ **(5)** Seja $\psi : X \longrightarrow \mathbb{R}$ uma função integrável e f-invariante. Considere todo conjunto

$$E_c = \{x \in X;\ \psi(x) \leq c\}. \tag{1.3}$$

Temos que E_c é um conjunto invariante. Logo, a hipótese implica que $\mu(E_c) \in \{0, 1\}$ para todo $c \in \mathbb{R}$. Como a função $c \mapsto \mu(E_c)$ é não-decrescente, isto é, existe $c' \in \mathbb{R}$ tal que $\mu(E_{c'}) = 0$, para todo $c < c'$ e $\mu(E_c) = 1$ para todo $c \geq c'$. Então $\psi = c'$ em μ-q.t.p. Logo $\displaystyle\int \phi d\mu = c'$ e, portanto, $\psi(x) = \displaystyle\int \psi d\mu$ em μ-q.t.p. $x \in X$.

(5) $\Rightarrow$ (3) Como temos que a média temporal é invariante por f (Corolário 1.5), a implicação segue de imediato aplicando o Teorema 1.4.

(3) $\Rightarrow$ (1) A média de visita é uma média temporal (da função característica em E), provando assim a implicação.

Como **(7)** implica **(5)**, **(5)** implica **(3)** e **(3)** implica **(1)**, temos que **(7)** implica **(1)**, finalizando a demonstração. ∎

A proposição que apresentaremos a seguir é uma consequência do Teorema 1.4 que garantirá um resultado importante, necessário para o teorema de Oseledets que veremos no próximo capítulo.

1.2 Entropia

1.2.1 Entropia Métrica

Nesta seção, vamos introduzir o conceito de **entropia métrica** (com respeito à uma medida μ invariante) de uma função mensurável $f : X \longrightarrow X$, a qual será denotada por $h_\mu(f)$.

Antes disto, vamos introduzir o espaço das sequências de d símbolos. Alguns sistemas modelam sequências de experimentos aleatórios em que o resultado de cada experimento é independente dos demais. Supõe-se que em cada experimento há um número finito de resultados possíveis, designados por $1,\ 2,\ \ldots,\ d$, com probabilidades $p(1),\ p(2),\ \ldots,\ p(d)$ de ocorrerem, sendo

$$p(1) + p(2) + \cdots + p(d) = 1.$$

Consideremos Σ_d o conjunto das sequências $\underline{\alpha} = (\alpha_n)_{n\in\mathbb{Z}}$, onde cada $\alpha_n \in [1,\ 2, \ldots,\ d]$. Definiremos a seguir subconjuntos particulares de Σ_d, assim como uma importante função nesse conjunto, a famosa função *shift*.

Definição 1.9. *Seja Σ_d o conjunto das sequências $\underline{\alpha} = (\alpha_n)_{n\in\mathbb{Z}}$ com d símbolos. Será chamado de* ***cilindro*** *o subconjunto*

$$[k,\ \ldots,\ l;\ a_k,\ \ldots,\ a_l] = \{\underline{\alpha};\ \alpha_k = a_k,\ \ldots,\ \alpha_l = a_l\},$$

onde $k, l \in \mathbb{Z}$, com $k \leq l$, e cada $a_n \in [1,\ 2, \ldots,\ d]$.

Definimos a seguinte medida sobre cilindros:

$$\mu([k,\ \ldots,\ l;\ a_k,\ \ldots,\ a_l]) = p(a_k).\ldots.p(a_l),$$

sendo $a_i \in [1,\ 2, \ldots,\ d]$, para todo $i = k, \ldots, l$. Heurísticamente, a probabilidade de um evento composto

$$\alpha_k = a_k,\ \alpha_{k+1} = a_{k+1},\ \ldots,\ \alpha_l = a_l$$

é o produto das probabilidades de cada um deles. Em outras palavras, os resultados sucessivos independem entre si.

Vamos considerar agora em Σ_d, a σ-álgebra $\mathcal{B}$ gerada pelos cilindros. A família $\mathcal{B}_0$ das uniões disjuntas finitas dos cilindros é uma álgebra.

Definição 1.10. *Seja* $E \in \mathcal{B}_0$ *uma união disjunta finita de cilindros* $C_1, \ldots, C_N$*, definimos*

$$\mu(E) = \mu(C_1) + \cdots + \mu(C_N).$$

Assim temos uma medida de probabilidade na σ*-álgebra* $\mathcal{B}$ *gerada por* $\mathcal{B}_0$ *que é uma extensão de* μ*. Chamaremos* μ *a* ***medida de Bernoulli****, definida por* $p(1),\ \ldots,\ p(d)$*.*

Definição 1.11. *Seja* Σ_d *o conjunto das sequências* $\underline{\alpha} = (\alpha_n)_{n\in\mathbb{Z}}$ *com* d *símbolos. A aplicação* $\nu : \Sigma_d \longrightarrow \Sigma_d$ *dada por*

$$\nu((\alpha_n)_{n\in\mathbb{Z}}) = (\alpha_{n+1})_{n\in\mathbb{Z}}$$

é chamada de ***deslocamento*** *(ou "shift")* ***à esquerda****, que corresponde à translação no tempo.*

Além da família dos cilindros gerar a σ-álgebra $\mathcal{B}$, note que a pré-imagem de um cilindro ainda é um cilindro. De fato, se $C = [k,\ \ldots,\ l; a_k,\ \ldots,\ a_l]$, então

$$\nu^{-1}(C) = [k+1,\ \ldots,\ l+1;\ a_k,\ \ldots,\ a_l] = \{\underline{\alpha};\ \alpha_{k+1} = a_k, \ldots,\ \alpha_{l+1} = a_l\}.$$

Assim, temos

$$\mu(\nu^{-1}(C)) = p(a_k).\ldots.p(a_l) = \mu(C).$$

Pela igualdade acima, junto com o fato da medida de Bernoulli em $\mathcal{B}$ ser única, σ-aditiva em $\mathcal{B}_0$ e, como definimos μ, pelo lema a seguir teremos que ela é invariante para ν.

Lema 1.12. *Seja* $f : X \longrightarrow X$ *uma transformação mensurável e* μ *uma medida finita em* X*. Suponha que exista uma sub-álgebra geradora* $\mathcal{I}$ *da* σ*-álgebra de* X *tal que para todo* $E \in \mathcal{I}$*, temos* $\mu(E) = \mu(f^{-1}(E))$*. Então o mesmo vale para todo conjunto mensurável* E*, isto é, a medida* μ *é invariante por* f*.*

Agora daremos um pequeno exemplo que motivará a definição de entropia métrica.

Motivação: Representando o lançamento de uma moeda, podemos considerar ν e Σ_2 tal que $\nu : \Sigma_2 \longrightarrow \Sigma_2$, com

$$\alpha_n = \begin{cases} 0, & \text{se o lançamento der cara.} \\ 1, & \text{se o lançamento der coroa.} \end{cases}$$

Desta forma, consideremos:

n = número de lançamentos.

p = possibilidade de sair cara.

1-p = possibilidade de sair coroa.

Constatamos que **p.n** e **(1-p).n** representa a média de caras e corôas em **n** lançamentos, respectivamente. Assim a sequência típica de caras e coroas de cada termo é

$$(p^{p.n}.(1-p)^{(1-p).n}) = e^{(p.\log p + (1-p)\log(1-p)).n}.$$

Assim, o número $p \log p + (1-p)\log(1-p)$ representa a **taxa exponencial** desta sequencia típica com respeito à n.

Definição 1.13. *Seja A um conjunto não vazio. Uma* ***partição*** *de um conjunto A é qualquer coleção P de subconjuntos não vazios de A dotada da seguinte propriedade: todo elemento de A pertence a um e apenas um dos elementos de P.*

Assim uma coleção de conjuntos $P = A_1, A_2, \ldots, A_n$ é uma ***partição*** *(finita) do conjunto A, se as seguintes condições forem simultaneamente satisfeitas:*

(1) $A_i \neq \emptyset$, *para* $i = 1, 2, \ldots, n$;

(2) $A_i \subset A$, *para* $i = 1, 2, \ldots, n$

(3) $A = A_1 \cup A_2 \cup \cdots \cup A_n$;

(4) $A_1, A_2, \ldots, A_n$ *são mutuamente disjuntos, isto é,* $A_i \cap A_j = \emptyset$, *para* $i \neq j$, *com* $i, j = 1, 2, \ldots, n$.

Definição 1.14. *Seja $(X, \mathcal{A}, \mu)$ um espaço de probabilidade e $f : X \longrightarrow X$ uma transformação que preserva a medida de probabilidade μ no espaço mensurável X. Tomando P uma partição finita de X, a* ***entropia de*** P ***com respeito à*** μ *é dada por:*

$$H_\mu(P) = -\sum_{C\in P} \mu(C)\log(\mu(C))$$

Exemplo: (Shift de Bernoulli) Considere o espaço de símbolos Σ_d e a partição P_n por cilindros de tamanho n, ou seja, se $C \in P_n$, então

$$C = [0,\dots,n-1;\ a_0,\dots,a_{n-1}].$$

Consideremos μ a medida de Bernoulli para o vetor de possibilidades $p = (p_1,\dots,p_d)$. Mostremos a seguir que

$$H_\mu(P_n) = -n.\left(\sum_{i=1}^{d} p_i \log p_i\right).$$

Antes de tudo, dada uma partição $P = \{C_1,\dots,C_d\}$ de um conjunto X, vamos mostrar que $H_\mu(P) \le \log d$. Melhor dizendo, a entropia de P assumirá valor máximo quando ela for perfeita, isto é, $\mu(C_i) = 1/d$, para todo $i = 1,\dots,d$.

Proposição 1.15. $H_\mu(P)$ *assume valor máximo quando* P *é perfeita, isto é,*

$$P = \{C_1,\dots,C_d\} \text{ com } \mu(C_i) = \frac{1}{d},$$

para todo $i = 1,\dots,d$.

Demonstração: Seja $\phi : [0,\infty) \longrightarrow \mathbb{R}$ tal que

$$\phi(x) = \begin{cases} x\log x, & se\ x > 0 \\ 0, & se\ x = 0 \end{cases}$$

e $f : \mathbb{R}^d \longrightarrow \mathbb{R}$ com $f(x) = \sum_{i=1}^{d} \phi(x_i)$, $x = (x_1,\dots,x_d)$ e $\sum_{i=1}^{d} x_i = 1$.

Vamos encontrar o mínimo da função f. Defina $g : \mathbb{R}^d \longrightarrow \mathbb{R}$ com

$$g(x_1,\dots,x_d) = 1 - \sum_{i=1}^{d} x_i = 0.$$

Pelo Teorema do Multiplicador de Lagrange, existe $\lambda \in \mathbb{R}$ tal que

$$\nabla f(x_1,\dots,x_d) = \lambda \nabla g(x_1,\dots,x_d).$$

Desta forma,

$$(\log x_1 + 1,\dots,\log x_d + 1) = \lambda(-1,\dots,-1),$$

isto é,

$$\log x_i + 1 = -\lambda, \text{ então } x_i = \exp(-1-\lambda),$$

para todo $i = 1, \ldots, d$. Substituindo o valor de x_i em $\sum_{i=1}^{d} x_i = 1$, temos

$$\begin{aligned}
\sum_{i=1}^{d} \exp(-1-\lambda) = 1 \Leftrightarrow \quad & \exp(-1-\lambda) = \frac{1}{d} \\
& \Leftrightarrow -\lambda = \log\left(\frac{1}{d}\right) + 1 \\
& \Leftrightarrow x_i = \exp\left(\log\left(\frac{1}{d}\right) + 1 - 1\right) = \frac{1}{d},
\end{aligned}$$

para todo $i = 1, \ldots, d$. Então, a função $-f(x_1, \ldots, x_d) = -\sum_{i=1}^{d} \phi(x_i)$ assumirá valor máximo quando $x_i = 1/d$, para todo $i = 1, \ldots, d$.

Em outras palavras, a entropia de P assumirá valor máximo quando P for perfeita. ■

1.2.2 Medida e Entropia Condicional

Nesta parte, trabalharemos com medida e entropia de duas partições mensuráveis de um espaço X, tal que seus elementos se relacionam do ponto de vista de conjuntos. Desta forma, podemos definir o conceito de medida e entropia condicional, que relaciona uma partição com a outra. Também definiremos uma outra partição que refinará ambas, facilitando o cálculo da entropia de uma partição relacionada com a outra.

No mais, isso ajudará definirmos entropia métrica de uma função que preserva uma medida invariante. Em vista disto, introduziremos a definição da *função informação* de uma partição de X, que dependerá de uma medida invariante μ. Isso nos dará uma outra definição de entropia de uma partição sobre esta função, além do que, outra importante finalidade da função informação é de poder relacionar uma partição à outra, facilitando na definição de entropia condicional.

Definição 1.16. *Seja* I *e* J *conjuntos finitos, consideremos duas partições do espaço* X, $P = \{C_\alpha;\ \alpha \in I\}$ *e* $Q = \{D_\beta;\ \beta \in J\}$. *Dizemos que a partição* ***Q refina P*** *(ou que* P *é menos fina que* Q*), denotando por* $P \leq Q$*, se todo elemento de* Q *está contido em um elemento de* P*, a menos de medida nula. Isto é, dado* $\beta \in J$*, existe* $\alpha \in I$ *tal que* $D_\beta \subset C_\alpha$.

Definição 1.17. *A* ***medida condicional*** *do conjunto mensurável* A *em relação à* B, *denotada por* $\mu(A|B)$, *é dada por*

$$\mu(A|B) = \frac{\mu(A \cap B)}{\mu(B)}.$$

Seja $(X, \mathcal{A}, \mu)$ um espaço de probabilidade e $P = \{C_\alpha;\ \alpha \in I\}$ e $Q = \{D_\beta;\ \beta \in J\}$ duas partições mensuráveis de X. Denotaremos a partição $P \vee Q$ por

$$P \vee Q = \{C_\alpha \cap D_\beta;\ \text{para todo } \alpha \in I \text{ e } \beta \in J\}.$$

Note que a partição $P \vee Q$ **refina** ambas as partições P e Q.

Definição 1.18. **(a)** *A função* $I_P : X \longrightarrow \mathbb{R}$ *dada por*

$$I_P(x) = -\log(\mu(C_\alpha(x))),$$

*é chamada **função informação da partição** P de X, com $C_\alpha(x)$ sendo o elemento da partição $P = \{C_\alpha;\ \alpha \in I\}$ que contém x.*

(b) *A função* $I_{P,Q} : X \longrightarrow \mathbb{R}$ *definida por*

$$I_{P,Q}(x) = -\log(\mu(C_\alpha(x)|D_\beta(x))),$$

*é chamada **função informação da partição** $P|Q$ de X, onde $C_\alpha(x)$ e $D_\beta(x)$ são os elementos das partições $P = \{C_\alpha;\ \alpha \in I\}$ e $Q = \{D_\beta;\ \beta \in J\}$, os quais contém x.*

Através da função informação I_P de uma partição P de X, definimos

$$H_\mu(P) = \int_X I_P d\mu.$$

Em vista disto, definimos

$$H_\mu(P|Q) = \int_X I_{P,Q} d\mu,$$

como sendo a **entropia condicional** da partição P com respeito à Q. Notemos que se considerarmos a partição $Q = \{X\}$, então $H_\mu(P|Q) = H_\mu(P)$.

Seja $P_{D_\beta} = \{C_\alpha \cap D_\beta;\ \alpha \in I\}$ a partição de um conjunto $D_\beta \in Q$ e μ_β a medida condicional de P em relação à D_β, isto é, $\mu_\beta(C_\alpha) = \mu(C_\alpha|D_\beta)$, com $\alpha \in I$.

Se $x \in D_\beta$, então $x \in C_\alpha \cap D_\beta$ para algum $\alpha \in I$. Assim,

$$H_{\mu_\beta}(P_{D_\beta}) = -\sum_{\alpha \in I} \mu_\beta(C_\alpha)\log(\mu_\beta(C_\alpha)) = -\sum_{\alpha \in I} \mu(C_\alpha|D_\beta)\log(\mu(C_\alpha|D_\beta))$$

e pela definição de $H_\mu(P|Q)$, temos

$$\sum_{\beta\in J}\mu(D_\beta)H_{\mu_\beta}(P_{D_\beta}) = -\sum_{\beta\in J}\mu(D_\beta)\sum_{\alpha\in I}\mu(C_\alpha|D_\beta)\log(\mu(C_\alpha|D_\beta)) = H_\mu(P|Q),$$

concluindo que

$$H_\mu(P|Q) = \sum_{\beta\in J}\mu(D_\beta)H_{\mu_\beta}(P_{D_\beta}).$$

Definição 1.19. *Dizemos que duas partições* P *e* Q *são* ***independentes*** *se*

$$\mu(C_\alpha \cap D_\beta) = \mu(C_\alpha).\mu(D_\beta), \text{ para todo } C_\alpha \in P \text{ e } D_\beta \in Q.$$

Proposição 1.20. *Seja* $(X, \mathcal{A}, \mu)$ *um espaço de probabilidade. Seja* $P = \{C_\alpha;\ \alpha \in I\}$, $Q = \{D_\beta;\ \beta \in J\}$ *e* $R = \{E_\gamma;\ \gamma \in K\}$ *partições finitas e mensuráveis de* X*. Então*

(a) $0 < -\log\left(\sup_{\alpha\in I}\{\ \mu(C_\alpha)\ \}\right) \leq H_\mu(P) \leq \log(\#P)$*. Além disso,* $H_\mu(P) = \log(\#P)$ *se, e somente se,* P *é perfeita, ou seja,* $\mu(C_{\alpha'}) = \mu(C_\alpha)$*, para todo* α, $\alpha' \in I$.

(b) P *e* Q *são independentes se, e somente se,*

$$0 \leq H_\mu(P|Q) \leq H_\mu(P).$$

(c) $H_\mu(P|Q) = 0$ *se, e somente se,* $P < Q$.

(d) *Se* $Q \leq R$*, então* $H_\mu(P|R) \leq H_\mu(P|Q)$.

(e) $H_\mu(P \vee Q|R) = H_\mu(P|R) + H_\mu(Q|P \vee R)$*. Em particular, se* $R = \{X\}$*, então*

$$H_\mu(P \vee Q) = H_\mu(P) + H_\mu(Q|P).$$

(f) $H_\mu(P \vee Q|R) \leq H_\mu(P|R) + H_\mu(Q|R)$*. Em particular, se* $R = \{X\}$*, então*

$$H_\mu(P \vee Q) \leq H_\mu(P) + H_\mu(Q).$$

(g) $H_\mu(P|R) \leq H_\mu(P|Q) + H_\mu(Q|R)$.

(h) *Seja* λ *uma outra medida de* X*, então para toda partição mensurável* P *de* X *em relação à* μ *e* λ*, e para todo* $p \in [0, 1]$*, temos*

$$pH_\mu(P) + (1-p)H_\lambda(P) \leq H_{p\mu+(1-p)\lambda}(P).$$

Demonstração: **(a)** Se $H_\mu(P) = 0$, então $P = \{X\}$ (mod 0) e nada temos o que fazer. Portanto, suponhamos que $H_\mu(P) > 0$, então existem $\alpha, \alpha' \in I$ tais que $\mu(C_\alpha)$, $\mu(C_{\alpha'}) \in (0, 1)$.

Mostremos primeiramente que $-\log\left(\sup\limits_{\alpha\in I} \{\ \mu(C_\alpha)\ \}\right) \leq H_\mu(P)$.

Da definição de I_P, temos

$$\begin{aligned}
\inf I_P &= \inf\{I_P(x);\ x \in X\} \\
&= \inf\{-\log\ \mu(C_\alpha);\ \alpha \in I\ \} \\
&= -\sup\{-(-\log\ \mu(C_\alpha)\ );\ \alpha \in I\ \} \\
&= -\log\left(\sup_{\alpha\in I}\ \{\ \mu(C_\alpha)\ \}\right).
\end{aligned}$$

Assim,

$$H_\mu(P) = \int_X I_P d\mu \geq \int_X \inf I_P d\mu = \inf I_P.\int_X d\mu = \inf I_P.\mu(X) = -\log\left(\sup_{\alpha\in I}\ \{\ \mu(C_\alpha)\ \}\ \right). \tag{1.4}$$

Para mostrar que $H_\mu(P) \leq \log(\#P)$, lembramos que a entropia assume valor máximo quando P é perfeita (Proposição 1.15). Seja $P' = \{C_1, \ldots, C_k\}$ uma partição perfeita, isto é, quando $\mu(C_i) = \mu(C_j)$, para todo $i,\ j = 1, \ldots, k$, tal que $\log(\#P) = \log(\#P')$. Assim,

$$H_\mu(P) \leq H_\mu(P') = \log(\#P') = \log(\#P), \tag{1.5}$$

para toda partição P de X.

Segue de (1.4) e (1.5) que

$$0 < -\log\left(\sup_{\alpha\in I}\ \{\ \mu(C_\alpha)\ \}\ \right) \leq H_\mu(P) \leq \log(\#P).$$

(b) Antes de demonstrarmos este item, precisamos de dois resultados:

Proposição 1.21. *Seja* $(a_j)_{j\in\mathbb{N}}, (b_j)_{j\in\mathbb{N}} \subset \mathbb{R}$ *duas sequências tais que* $0 \leq b_j \leq a_j$ *e* $\sum\limits_{j=1}^{\infty} a_j = \sum\limits_{j=1}^{\infty} b_j$, *então* $a_j = b_j$, *para todo* $j \in \mathbb{N}$.

Demonstração: Suponha que exista $j_0 \in \mathbb{N}$ tal que $b_{j0} < a_{j0}$, então

$$\sum_{j=1}^{\infty} a_j = \left(\sum_{j\neq j_0} a_j\right) + a_{j_0} > \left(\sum_{j\neq j_0} a_j\right) + b_{j_0} \geq \left(\sum_{j\neq j_0} b_j\right) + b_{j_0} = \sum_{j=1}^{\infty} b_j,$$

o que é absurdo! Logo, $a_j = b_j$, para todo $j \in \mathbb{N}$. ■

Proposição 1.22. *Seja* $a_j > 0$*, com* $j = 1, \ldots, n$ *tal que* $\sum_{j=1}^{n} a_j = 1$ *e* $\phi : (0, \infty) \longrightarrow \mathbb{R}$ *uma função estritamente convexa. Se*

$$\phi\left(\sum_{j=1}^{n} a_j x_j\right) = \sum_{j=1}^{n} a_j \phi(x_j), \text{ então } x_1 = x_2 = \cdots = x_n.$$

Demonstração: Demonstremos por indução.

$(k = 2)$ Suponha que exista x_1 e x_2 com $x_1 \neq x_2$. Como $a_2 = 1 - a_1$ e por ϕ ser estritamente convexa, temos

$$\begin{aligned} \phi(a_1x_1 + a_2x_2) &= \phi(a_1x_1 + (1 - a_1)x_2) \\ &< a_1\phi(x_1) + (1 - a_1)(\phi(x_2)) \\ &= a_1\phi(x_1) + a_2\phi(x_2) \end{aligned}$$

o que contradiz a hipótese. Logo $x_1 = x_2$.

$(k = n + 1)$ Suponha a hipótese seja válida para $k = n$, então:

$$\phi\left(\sum_{j=1}^{n+1} a_j x_j\right) = \phi\left((1 - a_{n-1})\left(\sum_{j=1}^{n} \frac{a_j}{1 - a_{n+1}} x_j\right) + a_{n+1}x_{n+1}\right).$$

Tome $y = \sum_{j=1}^{n} (a_j/(1 - a_n + 1))x_j$. Se $y \neq x_{n+1}$, então pela convexidade de ϕ, temos

$$\phi\left(\sum_{j=1}^{n+1} a_j x_j\right) = \phi\left((1 - a_{n+1})y + a_{n+1}x_{n+1}\right) < (1 - a_{n+1})\phi(y) + a_{n+1}x_{n+1}$$

contradizendo a hipótese. Se $y = x_{n+1}$, então pela hipótese de indução para y, temos que $x_1 = \cdots = x_n = y = x_{n+1}$, como queríamos demonstrar. ■

Voltando a demonstração de (b), e utilizando a função $\phi : (0, \infty) \longrightarrow \mathbb{R}$ da Proposição 1.22, temos

$$\begin{aligned} 0 \leq H_\mu(P|Q) &= -\sum_{\beta \in J} \mu(D_\beta) \sum_{\alpha \in I} \phi(\mu(C_\alpha|D_\beta)) \\ &= -\sum_{\alpha \in I} \sum_{\beta \in J} \mu(D_\beta)\phi(\mu(C_\alpha|D_\beta)) \end{aligned}$$

$$
\begin{aligned}
&\leq -\sum_{\alpha\in I}\phi\left(\sum_{\beta\in J}\mu(D_\beta)(\mu(C_\alpha|D_\beta)\right)\\
&= -\sum_{\alpha\in I}\phi\left(\sum_{\beta\in J}\mu(C_\alpha\cap D_\beta)\right)\\
&= -\sum_{\alpha\in I}\phi\left(\mu\left(\bigcup_{\beta\in J}D_\beta\cap C_\alpha\right)\right)\\
&= -\sum_{\alpha\in I}\phi(\mu(C_\alpha))\\
&= H_\mu(P).
\end{aligned}
$$

Pela Proposição 1.21, temos

$$\sum_{\alpha\in I}\sum_{\beta\in J}\mu(D_\beta)\phi\left(\mu(C_\alpha|D_\beta)\right) = \sum_{\alpha\in I}\phi\left(\sum_{\beta\in J}\mu(D_\beta)\mu(C_\alpha|D_\beta)\right)$$

se, e somente se,

$$\sum_{\beta\in J}\mu(D_\beta)\phi\left(\mu(C_\alpha|D_\beta)\right) = \phi\left(\sum_{\beta\in J}\mu(D_\beta)\mu(C_\alpha|D_\beta)\right),$$

para todo $\alpha \in I$. Agora pela Proposição 1.22, a igualdade acima é verdadeira se, e somente se,

$$\mu(C_\alpha|D_\beta) = \mu(C_\alpha|D_{\beta'}),$$

para todo β, $\beta' \in J$. Isto implica que

$$\mu(C_\alpha|D_{\beta'}) = \sum_{\beta\in J}\mu(D_\beta)\mu(C_\alpha|D_\beta) = \mu(C_\alpha),$$

isto é,

$$\mu(C_\alpha\cap D_{\beta'}) = \mu(C_\alpha).\mu(D_{\beta'}),$$

para todo $\beta' \in J$.

(c) Como $\phi(x) \leq 0$, para todo $x \in X$ e pela injetividade da função log, temos

$$H_\mu(P|Q) = 0 \ \Leftrightarrow\ -\sum_{\alpha\in I}\sum_{\beta\in J}\mu(D_\beta)\phi(\mu(C_\alpha|D_\beta)) = 0$$

$$
\begin{aligned}
&\Leftrightarrow \phi(\mu(C_\alpha|D_\beta)) = 0\\
&\Leftrightarrow \log(\mu(C_\alpha \cap D_\beta)) = \log(\mu(D_\beta))\\
&\Leftrightarrow \mu(C_\alpha \cap D_\beta) = \mu(D_\beta)\\
&\Leftrightarrow P \leq Q \text{ (mod } 0).
\end{aligned}
$$

(d) Como $Q \leq R$, temos que $R \vee Q = R$, assim para concluirmos o resultado basta mostrar que $H_\mu(P|Q \vee R) \leq H_\mu(P|Q)$. Agora, observe que

$$H_\mu(P|Q) = \sum_{\beta \in J} \mu(D_\beta) H_{\mu_\beta}(P)$$

com $H_{\mu_\beta}(P) = -\sum_{\alpha\in I} \mu_\beta(C_\alpha)\log(\mu_\beta(C_\alpha)) = -\sum_{\alpha \in I}\mu(C_\alpha|D_\beta)\log(\mu(C_\alpha|D_\beta))$. Além disso, $H_{\mu_\beta}(P|R) \leq H_{\mu_\beta}(P)$, pela primeira parte deste item, o resultado segue da seguinte afirmação.

Afirmação: $H_\mu(P|Q \vee R) = \sum_{\beta\in J} \mu(D_\beta) H_{\mu_\beta}(P|R)$.

Demonstração: Temos que

$$
\begin{aligned}
\sum_{\beta\in J} \mu(D_\beta) H_{\mu_\beta}(P|R) &= \sum_{\beta \in J}\mu(D_\beta)\left[\sum_{\gamma\in K}\mu_\beta(E_\gamma) H_{(\mu_\beta)_\gamma}(P)\right]\\
&= -\sum_{\beta\in J}\mu(D_\beta)\sum_{\gamma\in K}\mu_\beta(E_\gamma)\sum_{\alpha\in I}\mu_{(\beta)_\gamma}(C_\alpha)\log(\mu_{(\beta)_\gamma}(C_\alpha))\\
&= -\sum_{\beta\in J}\mu(D_\beta)\sum_{\gamma\in K}\mu_\beta(E_\gamma)\sum_{\alpha\in I}\mu_\beta(C_\alpha|E_\gamma)\log(\mu_\beta(C_\alpha|E_\gamma))\\
&= -\sum_{\beta\in J}\mu(D_\beta)\sum_{\gamma\in K}\mu_\beta(E_\gamma)\sum_{\alpha\in I}\frac{\mu_\beta(C_\alpha\cap E_\gamma)}{\mu_\beta(E_\gamma)}\log\left(\frac{\mu_\beta(C_\alpha \cap E_\gamma)}{\mu_\beta(E_\gamma)}\right)\\
&= -\sum_{\beta\in J}\mu(D_\beta)\sum_{\gamma\in K}\mu_\beta(E_\gamma)\sum_{\alpha\in I}\frac{\mu_\beta(C_\alpha\cap E_\gamma)}{\mu_\beta(E_\gamma)}\log\left(\frac{\frac{\mu(C_\alpha\cap D_\beta\cap E_\gamma)}{\mu(D_\beta)}}{\frac{\mu(E_\gamma\cap D_\beta)}{\mu(D_\beta)}}\right)\\
&= -\sum_{\beta\in J}\mu(D_\beta)\sum_{\substack{\alpha\in I\\ \gamma\in K}}\frac{\mu(C_\alpha\cap D_\beta\cap E_\gamma)}{\mu(D_\beta)}\log\left(\frac{\mu(C_\alpha\cap D_\beta\cap E_\gamma)}{\mu(E_\gamma\cap D_\beta)}\right)\\
&= -\sum_{\substack{\alpha\in I\\ \beta\in J\\ \gamma\in K}}\mu(E_\gamma\cap D_\gamma)\frac{\mu(C_\alpha\cap D_\beta\cap E_\gamma)}{\mu(E_\gamma\cap D_\beta)}\log\left(\frac{\mu(C_\alpha\cap D_\beta\cap E_\gamma)}{\mu(D_\beta\cap E_\gamma)}\right)\\
&= -\sum_{\substack{\beta\in J\\ \gamma\in K}}\mu(E_\gamma\cap D_\gamma)\sum_{\alpha\in I}\mu(C_\alpha|E_\gamma\cap D_\beta)\log\left(\mu(C_\alpha|E_\gamma\cap D_\beta)\right)
\end{aligned}
$$

$$= H_\mu(P|Q \vee R).$$

(e) Por definição,

$$\begin{aligned}
H_\mu(P \vee Q|R) &= -\sum_{\substack{\alpha\in I\\ \beta\in J\\ \gamma\in K}} \mu(E_\gamma)\phi(\mu(C_\alpha \cap D_\beta|E_\gamma))\\
&= -\sum_{\substack{\alpha\in I\\ \beta\in J\\ \gamma\in K}} \mu(C_\alpha \cap D_\beta \cap E_\gamma)\log\left(\frac{\mu(C_\alpha \cap D_\beta \cap E_\gamma)}{\mu(E_\gamma)}\right)\\
&= -\sum_{\substack{\alpha\in I\\ \beta\in J\\ \gamma\in K}} \mu(C_\alpha \cap D_\beta \cap E_\gamma)\log\left(\frac{\mu(C_\alpha \cap D_\beta \cap E_\gamma)}{\mu(E_\gamma)}.\frac{\mu(C_\alpha \cap E_\gamma)}{\mu(C_\alpha \cap E_\gamma)}\right)\\
&= -\sum_{\substack{\alpha\in I\\ \beta\in J\\ \gamma\in K}} \mu(C_\alpha \cap D_\beta \cap E_\gamma)\log\left(\tfrac{\mu(C_\alpha \cap D_\beta \cap E_\gamma)}{\mu(C_\alpha \cap E_\gamma)}\right) - \sum_{\substack{\alpha\in I\\ \beta\in J\\ \gamma\in K}} \mu(C_\alpha \cap D_\beta \cap E_\gamma)\log\left(\tfrac{\mu(C_\alpha \cap E_\gamma)}{\mu(E_\gamma)}\right)\\
&= -\sum_{\substack{\alpha\in I\\ \beta\in J\\ \gamma\in K}} \mu(C_\alpha \cap D_\beta)\phi(\mu(D_\beta|C_\alpha \vee E_\gamma)) - \sum_{\substack{\alpha\in I\\ \gamma\in K}} \mu(C_\alpha \cap E_\gamma)\log\left(\tfrac{\mu(C_\alpha \cap E_\gamma)}{\mu(E_\gamma)}\right)\\
&= H_\mu(Q|P \vee R) + H_\mu(P|R).
\end{aligned}$$

(f) Como $P \vee R$ refina R ($R \leq P \vee R$), temos por **(b)** e **(e)** que

$$H_\mu(Q|P \vee R) \leq H_\mu(Q|R)$$

e

$$H_\mu(P \vee Q|R) = H_\mu(Q|P \vee R) + H_\mu(P|R),$$

respectivamente. Desta forma, temos

$$H_\mu(P \vee Q|R) = H_\mu(Q|P \vee R) + H_\mu(P|R) \leq H_\mu(Q|R) + H_\mu(P|R).$$

(g) Pelos itens **(e)** e **(f)**, temos que

$$H_\mu(R|P \vee Q) = H_\mu(P \vee R|Q) - H_\mu(P|Q) \leq H_\mu(R|Q).$$

Em vista disto,

$$\begin{aligned}
H_\mu(P|Q) + H_\mu(Q|R) &= (H_\mu(P \vee Q) - H_\mu(Q)) + (H_\mu(Q \vee R) - H_\mu(R)) \\
&= H_\mu(P \vee Q) + H_\mu(R|Q) - H_\mu(R) \\
&= H_\mu(P \vee Q \vee R) - H_\mu(R|P \vee Q)) + H_\mu(R|Q) - H_\mu(R) \\
&\geq H_\mu(P \vee Q \vee R) - H_\mu(R) \\
&\geq H_\mu(P \vee R) - H_\mu(R) \\
&= H_\mu(P|R).
\end{aligned}$$

(h) Pela convexidade da função ϕ, temos

$$\begin{aligned}
pH_\mu(P) + (1-p)H_\lambda(P) &= -p\left(\sum_{\alpha \in I} \phi(\mu(C_\alpha))\right) - (1-p)\left(\sum_{\alpha \in I} \phi(\lambda(C_\alpha))\right) \\
&\leq -\sum_{\alpha \in I} \phi((p\mu + (1-p)\lambda)(C_\alpha)) \\
&= H_{p\mu+(1-p)\lambda}(P).
\end{aligned}$$

Isto termina a demonstração da Proposição 1.20. ■

1.2.3 Partição Geradora

Definição 1.23. *Dada uma função* $f : X \longrightarrow X$ *e uma partição* P *de* X.

(a) *Denotemos por*

$$P^n = \bigvee_{i=0}^{n-1} f^{-i}(P) = \left\{ \bigcap_{i=0}^{n-1} f^{-i}(C);\ \textit{para todo } C \in P \right\}.$$

(b) *Uma partição* P *é dita uma* ***partição geradora*** *se* $\bigvee_{i=-\infty}^{+\infty} f^{-i}(P)$ *gera a* σ*-álgebra de* $\mathcal{A}$*, com* $(X, \mathcal{A}, \mu)$ *um espaço de medida e* f *uma função invertível. Caso,* f *não seja invertível, então pede-se que* $\bigvee_{i=0}^{+\infty} f^{-i}(P)$ *gere a* σ*-álgebra.*

Exemplo 1.24. *Considere o espaço* Σ_d *das sequências de* d *símbolos,* ν^+ *a função deslocamento à direita (isto é,* $\nu^+(\alpha_n) = \alpha_{n-1}$*, para todo* $n \in \mathbb{Z}$*) e* $P = \{C_1, \ldots, C_d\}$ *a partição dos cilindros de tamanho* 1*. A partição* P *é geradora, pois dado um cilindro* C *de tamanho* n*, obtemos*

$$C = [0, \dots, n-1; a_0, \dots, a_{n-1}] = \left\{ \bigcap_{i=0}^{n-1} f^{-i}(C_k);\ k \in \{1,\ \dots, d\} \right\}.$$

Então, temos que $C \in P^n$ *e todo cilindro de tamanho* n *é exatamente* n*-interseções finitas de cilindros de tamanho 1, isto é,* P *é geradora.*

1.2.4 Entropia de uma função com respeito a uma medida invariante

Aqui nesta seção, relembremos o conceito de uma medida invariante por uma função mensurável f em X. Uma medida de probabilidade μ é invariante por f se para todo conjunto mensurável E de X, temos

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

Denotemos por $\mathcal{M}(f)$ o conjunto das medidas (de probabilidade) invariantes pela transformação f.

Na sessão anterior, definimos entropia de uma partição mensurável P. Se μ for invariante para f, não é difícil notar que

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

Antes de definirmos entropia de uma partição P com respeito à uma função f e, entropia de f com relação à uma medida invariante μ, precisamos provar um importante resultado.

Proposição 1.25. *Sejam* $\{X,\ \mathcal{A},\ \mu\}$ *um espaço de probabilidade e* P *uma partição* X*. Temos que o limite*

$$\lim_{n \longrightarrow \infty} \frac{H_\mu(P^n)}{n}$$

existe.

Para demonstrarmos a proposição, provaremos que a sequência $(a_n/n)_{n\in\mathbb{N}}$ é convergente, se $(a_n)_{n\in\mathbb{N}}$ é subaditiva (isto é, dados $m, n \in \mathbb{N}$ então $a_{m+n} \leq a_m + a_n$). Em seguida, mostraremos que a sequência $(H_\mu(P^n))_{n\in\mathbb{N}}$ será também subaditiva, concluindo assim a demonstração da Proposição 1.25.

Lema 1.26. *Toda* $\left(\dfrac{a_n}{n}\right)_{n\in\mathbb{N}}$ *é convergente, se* $(a_n)_{n\in\mathbb{N}}$ *é subaditiva.*

Demonstração: Dados m, $n \in \mathbb{N}$, podemos tomar $q \in \mathbb{N}$ e $r \in \{0, \dots, m-1\}$ de forma que $n = m.q + r$. Pela definição de subaditividade, temos

$$a_n = a_{m.q+r} \leq a_{q.m} + a_r \leq q.a_m + a_r.$$

Assim,

$$\frac{a_n}{n} \leq \frac{a_m}{\frac{n}{q}} + \frac{a_r}{n}.$$

Fixando m, temos

$$\limsup_{n \longrightarrow \infty} \frac{a_n}{n} \leq \limsup_{n \longrightarrow \infty} \frac{a_m}{\frac{n}{q}} + \frac{a_r}{n} = \limsup_{n \longrightarrow \infty} \frac{a_m}{\frac{n}{q}},$$

pois r também está fixado. Note que quando $n \longrightarrow +\infty$, então $q \longrightarrow +\infty$, então

$$\lim_{n \longrightarrow \infty} \frac{n}{q} = \lim_{q \longrightarrow \infty} \frac{m.q + r}{q} = m,$$

de modo que

$$\limsup_{n \longrightarrow \infty} \frac{a_m}{\frac{n}{q}} = \frac{a_m}{m}.$$

Conclui-se então que

$$\limsup_{n \longrightarrow \infty} \frac{a_n}{n} \leq \frac{a_m}{m}$$

e, em particular,

$$\limsup_{n \longrightarrow \infty} \frac{a_n}{n} \leq \liminf_{m \longrightarrow \infty} \frac{a_m}{m}.$$

Por definição, temos que $(a_n/n)_{n \in \mathbb{N}}$ é convergente. ■

Demonstracao **(Proposição 1.25)**: Pelo lema anterior, basta provarmos que a sequência $(H_\mu(P^n))_{n \in \mathbb{N}}$ é subaditiva.

Afirmação: A sequência $(H_\mu(P^n))_{n \in \mathbb{N}}$ é subaditiva.

Por definição de função invariante e pelo item **(f)** da Proposição 1.20, temos

$$\begin{aligned} H_\mu(P^{m+n}) &= H_\mu\left(\bigvee_{i=0}^{m+n-1} f^{-i}(P)\right) \\ &= H_\mu\left(\bigvee_{i=0}^{m-1} f^{-i}(P) \vee \bigvee_{i=m}^{n+m-1} f^{-i}(P)\right) \end{aligned}$$

$$= H_\mu\left(\bigvee_{i=0}^{m-1} f^{-i}(P) \vee f^{-m}\left(\bigvee_{i=0}^{n-1} f^{-i}(P)\right)\right)$$
$$= H_\mu(P^m \vee f^{-m}(P^n))$$
$$\leq H_\mu(P^m) + H_\mu(f^{-m}(P^n))$$
$$= H_\mu(P^m) + H_\mu(P^n). \blacksquare$$

Agora faz sentido definirmos a entropia de uma partição P finita com respeito à f e, a entropia de uma função f em relação a uma medida invariante μ.

Definição 1.27. *A **entropia** de uma função **f** com respeito à partição P finita e uma medida μ é*

$$h_\mu(f, P) = \lim_{n \longrightarrow \infty} \frac{H_\mu(P^n)}{n}.$$

Definição 1.28. *A **entropia** de uma função **f** com respeito à uma medida μ é*

$$h_\mu(f) = \sup_P h_\mu(f, P),$$

com P uma partição finita de X.

A seguir, veremos exemplos de como calcular esta entropia, a de uma função com respeito à uma partição e de uma função com respeito a uma medida invariante.

Exemplo 1.29. *Neste exemplo vamos fixar μ a medida de Bernoulli com respeito ao vetor de probabilidade $(p_1, \ldots, p_d)$, considerando Σ_d o espaços das sequências com d símbolos, v^+ a função deslocamento à direita e P a partição dos cilindros C de tamanho 1. Observe que este exemplo está atrelado a se provar um resultado anterior.*

Como já foi visto,

$$P^n = \bigvee_{i=0}^{n-1} f^{-i}(P) = \left\{ \bigcap_{i=0}^{n-1} f(C);\ C \in P \right\}$$

é uma partição por cilindros de tamanho n e (como já calculamos anteriormente)

$$H_\mu(P^n) = -n. \sum_{i=1}^{d} p_i \log p_i.$$

Dessa maneira,

$$h_\mu(f, P) = \lim_{n \longrightarrow \infty} \frac{H_\mu(P^n)}{n} = \lim_{n \longrightarrow \infty} \frac{-n \sum_{i=1}^{d} p_i \log p_i}{n} = -\sum_{i=1}^{d} p_i \log p_i = H_\mu(P).$$

Exemplo 1.30. *Seja* $R_\alpha : S^1 \longrightarrow S^1$ *a rotação de ângulo* α *e* μ *a medida de Lebesgue em* S^1. *Repare que a medida de Lebesgue é invariante sobre a rotação* R_α *e que uma partição* P *do círculo com* k *elementos é determinada por uma sequência* $p_1, \ldots, p_k$ *de pontos de* S^1.

Em vista disto, note também que se denotarmos por $p_i^j = R_\alpha^{-j}(p_i)$ *então a partição* P^n *é determinada pelo conjunto de pontos* $C_n = \left\{(p_i^j) \in S^1;\ i = 1, \ldots, k \text{ e } j = 0, \ldots, n-1\right\}$.

Note que a interseção entre C_n *e* C_{n-1} *pode ser não vazia. Então*

$$\#C_n \leq k + \#C_{n-1} \leq \cdots \leq n.k.$$

Pela continuidade da função log *(e pelo item* **(a)** *da Proposição 1.20), temos*

$$h_\mu(f, P) = \lim_{n\longrightarrow\infty} \frac{H_\mu(P^n)}{n} \leq \lim_{n\longrightarrow\infty} \frac{1}{n} \log(\#P^n) \leq \lim_{n\longrightarrow\infty} \frac{1}{n} \log(k.n) = 0.$$

Logo, a entropia de f *em relação a medida de Lebesgue é sempre nula, pois* P *é arbitrária.*

Observação: Note que a igualdade $h_\mu(f, P) = H_\mu(P)$, obtida no exemplo **(1)** não é um caso particular. De fato, sempre que a partição for geradora tal fato acontece.

Veremos à frente alguns resultados e propriedades boas sobre entropia que, por exemplo, completarão o cálculo da entropia dos deslocamentos de Bernoulli. Em particular, serão necessários na demonstração da desigualdade de Ruelle, resultado que veremos no terceiro e último capítulo desta dissertação.

Proposição 1.31. *Para todo* $k \in \mathbb{N}$, *temos* $h_\mu(f^k) = k h_\mu(f)$. *Caso* f *for invertível, temos* $h_\mu(f^k) = |k| h_\mu(f)$, *para todo* $k \in \mathbb{Z}$.

Antes de demonstrar a proposição, provemos um lema importante que possui boas propriedades.

Lema 1.32. *Sejam* P *uma partição finita de* X, $f : X \longrightarrow X$ *uma função e* $\mu \in \mathcal{M}(f)$. *Então,*

$$h_\mu(f, P) = \lim_{n\longrightarrow\infty} H_\mu(P \mid \bigvee_{i=1}^{n} f^{-i}(P)).$$

Observação: Caso $f : X \longrightarrow X$ for invertível e considerando $P^{\pm n} = \bigvee_{i=-n}^{n-1} f^{-j}(P)$, podemos obter o mesmo resultado do lema acima para $h_\mu(f, P^{\pm n})$, para todo $n \in \mathbb{N}$ e $\mu \in \mathcal{M}(f)$.

Demonstração: Usando o item **(e)** da Proposição 1.20 junto com o fato de μ ser invariante, temos

$$H_\mu(P^n) = H_\mu\left(\bigvee_{i=1}^{n-1} f^{-i}(P)\right) + H_\mu\left(P \mid \bigvee_{i=1}^{n-1} f^{-i}(P)\right) = H_\mu(P^{n-1}) + H_\mu\left(P \mid \bigvee_{i=1}^{n-1} f^{-i}(P)\right),$$

para todo $n \in \mathbb{N}$. Desta forma, repetindo o processo anterior indutivamente, temos

$$H_\mu(P^n) = H_\mu(P) + \sum_{k=1}^{n-1} H_\mu\left(P \mid \bigvee_{i=1}^{k} f^{-i}(P)\right).$$

Por consequência,

$$h_\mu(f,P) = \lim_{n\longrightarrow\infty} \frac{1}{n} H_\mu(P^n) = \lim_{n\longrightarrow\infty} \frac{1}{n} H_\mu(P) + \frac{1}{n}\sum_{k=1}^{n-1} H_\mu\left(P \mid \bigvee_{i=1}^{k} f^{-i}(P)\right).$$

Pelo item **(d)** da Proposição 1.20, a sequência $\left(H_\mu\left(P \mid \bigvee_{i=1}^{n} f^{-i}(P)\right)\right)_{n\in\mathbb{N}}$ é decrescente, e portanto, $\lim_{n\longrightarrow\infty} H_\mu\left(P \mid \bigvee_{i=1}^{n} f^{-i}(P)\right)$ existe.

Isto garante que a média dos termos $a_n = H_\mu\left(P \mid \bigvee_{i=1}^{n} f^{-i}(P)\right)$ também converge a este limite quando $n \longrightarrow +\infty$. Portanto, temos

$$h_\mu(f,P) = \lim_{n\longrightarrow\infty} \frac{1}{n}\sum_{k=1}^{n-1} H_\mu\left(P \mid \bigvee_{i=1}^{k} f^{-i}(P)\right) = \lim_{n\longrightarrow\infty} H_\mu\left(P \mid \bigvee_{i=1}^{n} f^{-j}(P)\right). \blacksquare$$

Demonstração **(Proposição 1.31)**: Tomemos $g = f^k$ e seja P uma partição finita de X. Temos que

$$P^{kn} = \bigvee_{i=0}^{kn-1} f^{-i}(P) = \bigvee_{i=0}^{n-1} f^{-ki}\left(\bigvee_{i=0}^{k-1} f^{-i}(P)\right) = \bigvee_{i=0}^{n-1} g^{-i}(P^k).$$

Pelo Lema 1.32, segue que

$$kh_\mu(f,P) = \lim_{n\longrightarrow\infty} \frac{1}{n} H_\mu(P^{kn}) = h_\mu(g,P^k). \tag{1.6}$$

Como $P \leq P^k$, temos que $h_\mu(g,P) \leq kh_\mu(f,P) \leq h_\mu(g)$, para P qualquer, em particular, para o supremo das partições. Assim,

$$h_\mu(g) \leq kh_\mu(f) \leq h_\mu(g).$$

Agora, suponha que f seja invertível. Para qualquer $n \geq 1$

$$H_\mu(P^n) = H_\mu\left(f^{n-1}\left(\bigvee_{i=0}^{n-1} f^i(P)\right)\right) = H_\mu\left(\bigvee_{i=0}^{n-1} f^i(P)\right)$$

uma vez que μ é invariante. Desta forma, é fácil de ver que $h_\mu(f,P) = h_\mu(f^{-1},P)$ e como P é uma partição qualquer, temos

$$h_\mu(f) = h_\mu(f^{-1}). \tag{1.7}$$

Substituindo f por $g = f^k$, segue de (1.6) e (1.7) que

$$h_\mu(f^{-k}) = h_\mu(f^k) = kh_\mu(f),$$

para todo $k \in \mathbb{N}$. ■

Teorema 1.33. *[Komolgorov - Sinai] Seja $(X, \mathcal{A}, \mu)$ um espaço de probabilidade. Seja P uma partição finita geradora para $f : X \longrightarrow X$ uma função invariante sobre a medida de probabilidade μ, então*

$$h_\mu(f,P) = h_\mu(f).$$

1.2.5 Equivalência Ergódica

Aqui introduziremos o conceito de equivalência do ponto de vista ergódico entre duas transformações que preservam medida.

Definição 1.34. *Seja $f_1 : X_1 \longrightarrow X_1$ e $f_2 : X_2 \longrightarrow X_2$, duas transformações que preservam a medida μ_1 e μ_2, respectivamente, sobre espaços métricos compactos X_1 e X_2. Dizemos que o sistema (f_1, μ_1) é* ***equivalente*** *ao sistema (f_2, μ_2) se podemos escolher Y_1 e Y_2 tal que $\mu_1(X_1 - Y_1) = 0$ e $\mu_2(X_2 - Y_2) = 0$, e uma função bijetora e mensurável $\phi : Y_1 \longrightarrow Y_2$, com inversa mensurável, tal que*

(1) $\mu_1(\phi^{-1}(A)) = \mu_2(A)$, *para todo $A \subset Y_2$ mensurável.*

(2) $\phi \circ f_1 = f_2 \circ \phi$.

A relação de equivalência definida acima é chamada de **conjugação ergódica** entre (f_1, μ_1) e (f_2, μ_2) e ϕ é chamada de **conjugação**.

Observação: Note que Y_1 e Y_2 podem ser tomados de modo que $f_1(Y_1) \subset Y_1$ e $f_2(Y_2) \subset Y_2$. De fato! Tome $B_1 = X_1 - Y_1$. Se B_1 for invariante por f_1, tome $\overline{B_1} = \bigcup_{i=0}^{+\infty} f_1^{-i}(B_1)$. Como

$\mu(B_1) = 0$, temos que $\mu(f_1^{-1}(B_1)) = 0$, o que implica $\mu(\overline{B_1}) = 0$. Analogamente, o resultado vale para $B_2 = X_2 - Y_2$.

Proposição 1.35. *A conjugação ergódica é uma relação de equivalência.*

Demonstração:

(Reflexiva) Basta tomar a função ϕ como identidade, que é bijetora e mensurável.

(Simétrica) Basta tomar a função $\psi = \phi^{-1} : Y_2 \longrightarrow Y_1$ (obviamente bijetora e mensurável). Em vista disto,

(1) $\mu_2(\psi^{-1}(B)) = \mu_2(\phi(B)) = \mu_1(\phi^{-1}(\phi(B))) = \mu_1(B)$, $\forall B \subset Y_1$.

(2) $\psi \circ f_2 = \phi^{-1} \circ f_2 = (f_2^{-1} \circ \phi)^{-1} = (\phi \circ f_1^{-1})^{-1} = f_1 \circ \phi^{-1} = f_1 \circ \psi$.

(Transitiva) Se $(f_1, \mu_1) \sim (f_2, \mu_2)$ e $(f_2, \mu_2) \sim (f_3, \mu_3)$, então existe Y_i mensurável tal que $\mu_i(X_i - Y_i) = 0$, (para $i = 1, 2, 3$) funções $\phi : Y_1 \longrightarrow Y_2$ e $\psi : Y_2 \longrightarrow Y_3$, ambas mensuráveis e bijetoras, com suas respectivas inversas mensuráveis, tais que

(1) $\mu_1(\phi^{-1}(A)) = \mu_2(A)$ e $\mu_2(\psi^{-1}(B)) = \mu_3(B)$, $\forall A \subset Y_2$ e $B = \psi(A) \subset Y_3$ mensuráveis.

(2) $\phi \circ f_1 = f_2 \circ \phi$ e $\psi \circ f_2 = f_3 \circ \psi$.

Tomando $\xi = \psi \circ \phi : Y_1 \longrightarrow Y_3$ (que é uma bijeção mensurável, pois ψ e ϕ também são). Então,

(1) $\mu_3(B) = \mu_2(\psi^{-1}(B)) = \mu_2(A) = \mu_1(\phi^{-1}(A)) = \mu_2(\phi^{-1}(\psi^{-1}(B))) = \mu_1(\xi^{-1}(B))$, $\forall B \subset Y_3$ mensurável e $\psi^{-1}(B) = A$ com $A \subset Y_2$ também mensurável.

(2) $\xi \circ f_1 = (\psi \circ \phi) \circ f_1 = \psi \circ (\phi \circ f_1) = \psi \circ (f_2 \circ \phi) = (\psi \circ f_2) \circ \phi = (f_3 \circ \psi) \circ \phi = f_3 \circ (\psi \circ \phi) = f_3 \circ \xi$.

Provando assim que a equivalência ergódica é uma relação de equivalência. ■

1.2.6 A entropia como um invariante

Como vimos na seção anterior que a equivalência ergódica é uma relação de equivalência, vejamos agora que a entropia é um invariante para esta relação. Lembremos que a entropia visa medir a complexidade da dinâmica. Assim sendo, é razoável esperarmos que a entropia seja a mesma para transformações equivalentes.

Proposição 1.36. *Se* (f_1, μ_1) *e* (f_2, μ_2) *são ergodicamente equivalentes, então*

$$h_{\mu_1}(f_1) = h_{\mu_2}(f_2).$$

Demonstração: Seja P_1 uma partição de X_1. Desprezando os conjuntos de medida nula para μ_1, podemos supor, sem perda de generalidade, que P_1 é partição de Y_1. Agora, defina

$$P_2 = \phi(P_1) = \{\phi(P) \subset Y_2;\ P \in P_1\}.$$

Como ϕ é bijeção, temos que P_2 é partição de X_2. Por hipótese, $\phi \circ f_1 = f_2 \circ \phi$, então $\phi({P_1}^n) = \phi^n(P_1) = {P_2}^n$, e como $\mu_1(P) = \mu_2(\phi(P))$, para todo $P \in P_1$, segue que

$$\begin{aligned} H_{\mu_1}({P_1}^n) &= -\sum_{P \in {P_1}^n} \mu_1(P) \log(\mu_1(P)) \\ &= -\sum_{P \in {P_1}^n} \mu_2(\phi(P)) \log(\mu_2(\phi(P))) \\ &= -\sum_{P' \in {P_2}^n} \mu_2(P') \log(\mu_2(P')) \\ &= H_{\mu_2}({P_2}^n), \end{aligned}$$

com $P' = \phi(P) \in {P_2}^n$. A fortiori,

$$h_{\mu_1}(f_1) \leq h_{\mu_2}(f_2).$$

Como P_1 foi escolhida de modo arbitrário e pelo fato da função ϕ ser bijetora, basta tomar ϕ^{-1} e tomar P_2 de modo arbitrário, para obter de forma análoga que

$$h_{\mu_1}(f_1) \geq h_{\mu_2}(f_2),$$

concluindo a igualdade. ■

Observação: A entropia métrica **não** é um invariante completo. Um exemplo claro disto é a função R_α (a função rotação sobre um ângulo α) que preserva a medida de Lebesgue, qual já vimos anteriormente possuir entropia nula. Ou seja, a proposição nos mostra quando duas entropias não são equivalentes. Pois, a rotação irracional não é equivalente a função racional. De fato, toda rotação racional é periódica e rotações irracionais não possuem pontos periódicos. Já para shifts de Bernoulli, a entropia é um invariante completo, resultado do Teorema de Orsteins que enunciaremos, mas não será demonstrado.

Teorema 1.37. *Dois shifts de Bernoulli com a mesma entropia são (necessariamente) equivalentes.*

CAPÍTULO 2

EXPOENTES DE LYAPUNOV E O TEOREMA DE OSELEDETS

Neste capítulo, daremos a motivação e a definição dos expoentes de Lyapunov para difeomorfismos definidos em variedades diferenciáveis, além da definição de pontos regulares. Em seguida, apresentaremos uma demonstração do teorema de Oseledets segundo [1], resultado que mostra que quase todo ponto $x \in X$ é regular.

2.1 Pontos regulares e o expoente de Lyapunov

Nesta seção, apresentaremos a definição de pontos regulares que viabiliza a caracterização dos expoentes de Lyapunov, lembrando que X é uma variedade Riemmaniana compacta, conexa e sem bordo.

Definição 2.1. *Seja $f : X \longrightarrow X$ um difeomorfismo de uma variedade compacta. Dizemos que $x \in X$ é um* ***ponto regular*** *de f se existem $\lambda_1(x) > \cdots > \lambda_j(x)$ e uma decomposição*

$$T_xX = E_1(x) \oplus E_2(x) \oplus \cdots \oplus E_j(x) \tag{2.1}$$

tais que

$$\lim_{n \longrightarrow \pm\infty} \frac{1}{n} \log \|D_x f^n v\| = \lambda_i(x), \tag{2.2}$$

para todo $v \in E_i(x)$ não-nulo e $1 \leq i \leq j$.

*Para cada $x \in X$, os números $\lambda_i(x)$ são chamados de **expoentes de Lyapunov** de x e $E_i(x)$ são os **subespaços próprios de f** no ponto regular x. Denotaremos por Λ o conjuntos dos pontos regulares de X.*

Na sequência, veremos o teorema de Oseledets, que garantirá a existência de pontos regulares. Mais ainda, sob o ponto de vista de medidas invariantes, Λ é quase todo X.

Teorema 2.2. *[Oseledets] Seja X uma variedade compacta e Riemmanniana de dimensão finita. Dado um difeomorfismo $f : X \longrightarrow X$ de classe C^1, o conjunto dos pontos regulares Λ possui medida total para qualquer medida ergódica invariante μ sobre X.*

Observemos que o teorema é um pouco mais forte. De fato, Λ possui medida total para qualquer medida invariante. Para ver a versão geral e demonstração veja [6].

Vamos demonstrar uma versão mais geral do Teorema 2.2, para cociclos lineares. Mas antes de tudo, apresentaremos definições básicas para que o teorema em questão fique claro.

Definição 2.3. *Dado um espaço topológico X, definimos um **fibrado vetorial** F de $\mathbb{R}^d$ sobre X por:*

$$F = \left\{(x, v);\ x \in X \text{ e } v \in \mathbb{R}^d\right\}.$$

*Para cada $x \in X$, o conjunto $F_x = \left\{(x, v);\ v \in \mathbb{R}^d\right\}$ é chamado de **fibra** do fibrado F, tal que $F = \bigcup\limits_{x \in X} F_x$.*

Definição 2.4. *Um **isomorfismo** $T : F \longrightarrow F$ é uma família de isomorfismos lineares $T_x = T(x) : F_x \longrightarrow F_{f(x)}$, para $x \in X$.*

Definição 2.5. *Seja X um espaço métrico compacto, $f : X \longrightarrow X$ um homeomorfismo e F um fibrado vetorial de dimensão finita sobre X dotado de uma métrica Riemanniana contínua. Seja $\Pi : F \longrightarrow X$ a projeção e $T : F \longrightarrow F$ um isomorfismo que cobre f (isto é, $\Pi \circ T = f \circ \Pi$), onde $T(x, v) = (f(x), T_x(v))$, para todo $x \in X$, $v \in F_x$ e com T e T^{-1} com normas limitadas. Então T é chamada de **cociclo linear**. Para facilitar a notação, definimos o seguinte:*

$$T_n(x) = T(f^{n-1}(x)) \circ \cdots \circ T(f(x)) \circ T(x)$$

e

$$T_{-n}(x) = T^{-1}(f^{-n}(x)) \circ \cdots \circ T^{-1}(f^{-1}(x)),$$

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

Aqui iremos definir o conjunto de pontos regulares para cociclos lineares e, de maneira análoga, os expoentes de Lyapunov e seus respectivos espaços próprios de um fibrado F.

Definição 2.6. *Seja* F *um fibrado sobre um espaço topológico* X. *Denotamos por* Λ *o conjunto de pontos* $x \in X$ *tais que cada fibra* F_x *de* F *admite uma decomposição*

$$F_x = E_1(x) \oplus E_2(x) \oplus \cdots \oplus E_j(x),$$

e existem números reais $\lambda_1(x) > \cdots > \lambda_j(x)$ *satisfazendo*

$$\lim_{n \longrightarrow \pm\infty} \frac{1}{n} \log \|T_n(x)v\| = \lambda_i(x),$$

para μ*-q.t.p.* $x \in X$ *e* $v \in E_i(x) \setminus \{0\}$, *com* $1 \leq i \leq j$.

Neste caso, dizemos que Λ *é o conjunto de pontos regulares do isomorfismo* T, *e que* $\lambda_i(x)$ *e* $E_i(x)$ *são, respectivamente, os* ***expoentes de Lyapunov*** *e os* ***subespaços próprios de F*** *em* x, ***ponto regular*** *de* T.

Agora de fato, poderemos apresentar a versão do teorema para cociclos lineares e note que ele de fato, implicará no Teorema 2.2.

Teorema 2.7. *Considerando o cociclo* $T(x,v) = (f(x), T_x(v))$ *sobre* (X, F) *onde* F *é um fibrado sobre* X. *Então temos que o conjunto* Λ *de todos os pontos regulares de* T *é mensurável e tem medida total para toda medida ergódica* μ *para* f *em* X.

Pelas hipóteses do Teorema 2.2, considerando o cociclo $Df(x,v) = (f(x), Df_x(v))$ tal que F seja o fibrado tangente de X (isto é, $F_x = T_xX$, para cada $x \in X$) e pelo fato de $f \in C^1$, $T = Df$ é o isomorfismo contínuo em F (logo possui norma limitada). Em outras palavras, o Teorema 2.2 será um caso particular do Teorema 2.7.

2.2 Demonstração do Teorema de Oseledets para cociclos

Nesta seção, vamos demonstrar o Teorema 2.7. Isto será feito em duas etapas:

- $1^{\underline{a}}$ **Etapa:** Mostrar que Λ possui medida total para toda medida μ-ergódica.

- $2^{\underline{a}}$ **Etapa:** Mostrar que Λ é mensurável.

2.2.1 Demonstração da primeira etapa: Medida Total.

Primeiramente, vejamos algumas propriedades e definições necessárias.

Definição 2.8. *Seja* $T(x,v) = (f(x), T_x(v))$ *o cociclo sobre* (X, F), *com* T *um isomorfismo sobre o fibrado* F. *Dizemos que o* ***subfibrado*** E *de* F *(isto é, para cada* $x \in X$, *a fibra* $E_x \subset F_x$*) é* ***T-invariante*** *se* $T(x)(E_x) = E_{f(x)}$ *para* μ*-q.t.p.* $x \in X$.

Definição 2.9. *Dado* T *um isomorfismo de um fibrado* F. *Definimos*

$$\lambda_1(T, x) = \limsup_{n \longrightarrow +\infty} \frac{1}{n} \log \|T_n(x)\|.$$

Afirmação 1: A função λ_1 é invariante por f, isto é, $\lambda_1(T, x) = \lambda_1(T, f(x))$.

Lembrando que os isomorfismos T e T^{-1} possuem norma limitada, temos

$$\begin{aligned}
\lambda_1(T, x) &= \limsup_{n \longrightarrow +\infty} \frac{1}{n} \log \|T_{n-1}(f(x)) \circ T(x)\| \\
&\leq \limsup_{n \longrightarrow +\infty} \frac{1}{n} \log \|T_{n-1}(f(x))\| + \limsup_{n \longrightarrow +\infty} \frac{1}{n} \log \|T(x)\| \\
&= \limsup_{n \longrightarrow +\infty} \left(\frac{n-1}{n}\right) \left(\frac{1}{n-1}\right) \log \|T_{n-1}(f(x))\| \\
&= \lambda_1(T, f(x)).
\end{aligned}$$

Por outro lado,

$$\begin{aligned}
\lambda_1(T, f(x)) &= \limsup_{n \longrightarrow +\infty} \frac{1}{n} \log \|T_n(f(x))\| \\
&= \limsup_{n \longrightarrow +\infty} \frac{1}{n} \log \left\|T_n(f(x)) T(x) T^{-1}(x)\right\| \\
&\leq \limsup_{n \longrightarrow +\infty} \frac{1}{n} \log \|T_n(f(x)) T(x)\| + \limsup_{n \longrightarrow +\infty} \frac{1}{n} \log \left\|T^{-1}(x)\right\| \\
&= \limsup_{n \longrightarrow +\infty} \left(\frac{n+1}{n}\right) \left(\frac{1}{n+1}\right) \log \|T_{n+1}(x)\| \\
&= \lambda_1(T, x).
\end{aligned}$$

Portanto, como $\lambda_1(T, x)$ é uma função f-invariante, temos que $\lambda_1(T, x)$ é constante em quase todo ponto de X. Isto é, $\lambda_1(T, x) = \lambda_1(T)$ para μ-q.t.p. $x \in X$, dado que μ é ergódica.

O lema que apresentaremos a seguir, mostra a existência de um subfibrado G mensurável T-invariante de F, tal que as fibras de G possuem dimensão positiva, para μ-q.t.p. $x \in X$. Além disso, G_x será um espaço próprio de x associado ao expoente de Lyapunov $\lambda_1(T)$.

Lema 2.10. *[Subfibrado Invariante] Seja* $T(x,v) = (f(x), T_x(v))$ *o cociclo sobre* (X, F)*, com* T *um isomorfismo sobre o fibrado* F*. Dado* $x \in X$*, considere o conjunto*

$$G_x = \left\{ v \in F_x;\ \limsup_{n\longrightarrow+\infty} \frac{1}{n} \log \|T_{-n}(x)v\| \leq -\lambda_1(T) \right\} \cup \{0\}.$$

Então $G = \bigcup_{x\in X} G_x$ *é um subfibrado mensurável* T*-invariante de* F *tal que* $\dim G \neq 0$ *e*

$$\lim_{n\longrightarrow\pm\infty} \frac{1}{n} \log \|T_n(x)v\| = \lambda_1(T)$$

para μ*-q.t.p.* $x \in X$ *e* $v \in G_x \setminus \{0\}$*.*

Pelo Lema 2.10, podemos considerar o conjunto $\Lambda_1 \subset X$ de medida total para o qual temos bem definido G_x. Caso $F \neq G$, consideremos $G^\perp$ o complemento ortogonal de G, tal que $F_x = G_x^\perp \oplus G_x$.

Assim, seja $T^{-1}|_G$ o isomorfismo inverso restrito à G. Podemos afirmar o seguinte resultado:

Afirmação 2: $\lambda_1(T^{-1}|_G) = -\lambda_1(T)$**.**

De fato, para todo $v \in G_x$, temos

$$\begin{aligned}
\lambda_1(T) &= \lim_{n\longrightarrow+\infty} -\frac{1}{n} \log \left\| T_{-n}(x) \frac{v}{\|v\|} \right\| \\
&= -\lim_{n\longrightarrow+\infty} \frac{1}{n} \log \left\| T_n^{-1}(x) \frac{v}{\|v\|} \right\|.
\end{aligned}$$

Assim,

$$\begin{aligned}
-\lambda_1(T) &= \sup_{v\in G_x\setminus\{0\}} \left(\lim_{n\longrightarrow+\infty} \frac{1}{n} \log \left\| T_n^{-1}(x) \frac{v}{\|v\|} \right\| \right) \\
&= \lim_{n\longrightarrow+\infty} \frac{1}{n} \log \sup \left(\left\| T_n^{-1}(x) \frac{v}{\|v\|} \right\| \right) \\
&= \lim_{n\longrightarrow+\infty} \frac{1}{n} \log \left\| (T_n^{-1}|_G)(x) \right\| \\
&= \lambda_1(T^{-1}|_G).
\end{aligned}$$

Considerando $\pi : F \longrightarrow G^\perp$ a projeção ortogonal, definimos por $\widehat{T}$ o isomorfismo de $G^\perp$ tal que $\widehat{T}(x) = \pi \circ T(x)|_{G_x^\perp}$. Desta forma, temos por indução a seguinte afirmação:

Afirmação 3: $\widehat{T}_n(x) = \pi(T_n(x))$**, para todo** $n \in \mathbb{N}$**.**

Para $n = 1$, segue diretamente da definição.

Suponha que a hipótese seja válida para $k = n$. Dado $v \in F_x$ não-nulo, temos

$$\widehat{T}_{n+1}(x)v = \widehat{T}(f^n(x)) \circ \widehat{T}_n(x)v$$

e

$$T_{n+1}(x)v = T(f^n(x)) \circ T_n(x)v.$$

Como $F_x = G_x \oplus G_x^\perp$ e G é T-invariante, pela hipótese de indução temos que

$$T_n(x)v = \pi(T_n(x)v) + u = \widehat{T}_n(x)v + u, \text{ para algum } u \in G_{f^n(x)}.$$

Desta forma,

$$T_{n+1}(x)v = T(f^n(x)) \circ T_n(x)v = T(f^n(x)) \circ (\widehat{T}_n(x)v + u) = T(f^n(x)) \circ (\widehat{T}_n(x)v) + T(f^n(x))u.$$

Aplicando π no resultado acima, obtemos

$$\begin{aligned}
\pi(T_{n+1}(x)v) &= \pi(T(f^n(x)) \circ (\widehat{T}_n(x)v)) + \pi(T(f^n(x))u) \\
&= \pi(T(f^n(x)))\widehat{T}(x)v \\
&= \widehat{T}(f^n(x)) \circ \widehat{T}_n(x)v \\
&= \widehat{T}_{n+1}(x)v,
\end{aligned}$$

para todo $v \in F_x \setminus \{0\}$.

Afirmação 4: $\lambda_1(T) > \lambda_1(\widehat{T})$.

Se considerarmos o fibrado $G^\perp$ sobre X e o isomorfismo $\widehat{T}$ sobre $G^\perp$, podemos aplicar o Lema 2.10 para encontrar $\widehat{G}$, um subfibrado mensurável $\widehat{T}$-invariante de $G^\perp$ tal que

$$\lim_{n \longrightarrow \pm\infty} \frac{1}{n} \log \left\| \widehat{T}_n(x)v \right\| = \lambda_1(\widehat{T}),$$

para μ-q.t.p. $x \in X$ e $v \in \widehat{G}_x \subset G_x^\perp$ não-nulo.

Como $\|\pi\| \leq 1$, pela **Afirmação 3** teremos que

$$\left\| \widehat{T}_n(x)v \right\| = \|(\pi \circ T_n)(x)v\| \leq \|T_n(x)v\|,$$

para todo $v \in \widehat{G}_x \setminus \{0\}$. Consequentemente,

$$\lambda_1(\widehat{T}) = \lim_{n \longrightarrow +\infty} \frac{1}{n} \log \left\| \widehat{T}_n(x)v \right\| \leq \limsup_{n \longrightarrow +\infty} \frac{1}{n} \log \|T_n(x)v\| \leq \lambda_1(T).$$

Suponhamos $\lambda_1(\widehat{T}) = \lambda_1(T)$. Pela definição de G_x, dado $v \in G_x \setminus \{0\}$, temos

$$\limsup_{n \longrightarrow \pm\infty} \frac{1}{n} \log \|T_{-n}(x)v\| \leq -\lambda_1(T).$$

Assim,

$$\begin{aligned} \limsup_{n \longrightarrow \pm\infty} \frac{1}{n} \log \left\|\widehat{T}_{-n}(x)v\right\| &\leq \limsup_{n \longrightarrow \pm\infty} \frac{1}{n} \log \|T_{-n}(x)v\| \\ &\leq -\lambda_1(T) \\ &\leq -\lambda_1(\widehat{T}). \end{aligned}$$

Portanto, pela definição de $\widehat{G}_x$, temos que $v \in \widehat{G}_x \setminus \{0\}$. Isto é uma contradição desde que $v \in \widehat{G}_x \subset G_x^{\perp}$. À fortiori, $\lambda_1(\widehat{T}) < \lambda_1(T)$.

Para concluir a demonstração do Teorema 2.7, usaremos o seguinte lema.

Lema 2.11. *[Subfibrado Complementar] Sejam $T(x, v) = (f(x), T_x(v))$ o cociclo sobre (X, F), com T um isomorfismo sobre o fibrado F, e G um subfibrado T-invariante de F tal que $G \neq F$. Considerando $\widehat{T}$ o isomorfismo do complemento ortogonal de G, $G^{\perp}$, se*

$$\lambda_1(\widehat{T}) + \lambda_1(T^{-1}|_G) < 0,$$

então existe um subfibrado H mensurável T-invariante de F tal que

$$F_x = H_x \oplus G_x,$$

com $\lambda_1(T|_H) = \lambda_1(\widehat{T})$ e $\lambda_1(T|_H) < \lambda_1(T)$, para μ-q.t.p. $x \in X$.

Pelas **Afirmações 2** e **4**, temos de fato que $\lambda_1(\widehat{T}) + \lambda_1(T^{-1}|_G) < 0$. Assim, podemos usar o Lema 2.11 para garantir a existência do subfibrado H.

Vejamos agora, como concluir a demonstração do Teorema 2.7.

Pelo Lema 2.10 para $(f(x), T_x v)$, consideremos $E_1(x) = G_x$ e $\lambda_1 = \lambda_1(T)$. Denotaremos agora por $H_1 = H$, dado pelo Lema 2.11, de modo que

$$F_x = H_1(x) \oplus E_1(x)$$

para μ-q.t.p. $x \in X$. Notemos também que para todo $v \in E_1(x) \setminus \{0\}$, obtemos

$$\lim_{n \longrightarrow \pm\infty} \frac{1}{n} \log \|T_n(x)v\| = \lambda_1.$$

Denotaremos por $\Lambda_1 \subset X$ o conjunto de medida total no qual o Lema 2.10 e Lema 2.11 são satisfeitos como acima.

Considere agora a restrição $T|_{H_1}$ do isomorfismo T em H_1 e denotemos $\lambda_2 = \lambda_1(T|_{H_1})$. Podemos aplicar novamente o Lema 2.10 e, se necessário, o Lema 2.11 e encontrarmos subfibrados mensuráveis T-invariantes, E_2 e H_2 de H_1, tais que

$$F_x = H_2(x) \oplus E_2(x) \oplus E_1(x)$$

e

$$\lim_{n \longrightarrow \pm\infty} \frac{1}{n} \log \|T_n(x)v\| = \lambda_2,$$

para μ-q.t.p. $x \in X$ e $v \in E_2(x) \setminus \{0\}$. Como antes, vale que $\lambda_2 > \lambda_1(T|_{H_2})$.

Analogamente, denotemos por $\Lambda_2 \subset \Lambda_1$ o conjunto de medida total onde as propriedades acima são satisfeitas.

Lembrando que o fibrado F possui dimensão finita, temos que o processo acima pode ser realizado finitas vezes, afim de encontrar números reais $\lambda_1 > \lambda_2 > \cdots > \lambda_j$ e subfibrados $E_1, \ldots, E_j$, tais que

$$F_x = E_j(x) \oplus \cdots \oplus E_2(x) \oplus E_1(x),$$

e

$$\lim_{n \longrightarrow \pm\infty} \frac{1}{n} \log \|T_n(x)v\| = \lambda_i$$

para todo $v \in E_i(x) \setminus \{0\}$, com $1 \leq i \leq j$, e todo $x \in \Lambda_j$, onde $\Lambda_j \subset X$ possui medida total.

Pela construção dos subfibrados E_i de F, podemos afirmar que eles são mensuráveis e T-invariantes e, por definição, λ_i é uma função mensurável, para todo $1 \leq i \leq j$. Mais ainda, temos que Λ_j é quase todo X. No mais, teremos que Λ possuirá medida total se Λ for mensurável, como veremos à seguir.

2.2.2 Demonstração da segunda etapa: Mensurabilidade.

Mostraremos nesta etapa que Λ é mensurável através de conjuntos que possuem propriedades pertinentes. Por este fato, precisamos de resultados que nos ajude a entender tais conjuntos.

Pode ser verificado na demonstração do Lema 2.10 (veja as **Afirmações 1** e **4** da 1ª etapa da seção 2.2.3) que, dado $\lambda \in \mathbb{R}$, os conjuntos

$$G_x(\lambda) = \left\{ v \in F_x \setminus \{0\};\ \limsup_{n \longrightarrow +\infty} \frac{1}{n} \log \|T_{-n}(x)v\| \leq -\lambda \right\} \cup \{0\}$$

e

$$H_x(\lambda) = \left\{ v \in F_x \setminus \{0\};\ \limsup_{n \longrightarrow -\infty} \frac{1}{n} \log \|T_{n}(x)v\| \leq \lambda \right\} \cup \{0\},$$

são subespaços não-nulos e variam mensuravelmente em X.

Considere assim, as seguintes funções mensuráveis:

$$G(\lambda) : X \longrightarrow F, \text{ tal que } G(\lambda)(x) = G_x(\lambda)$$

e

$$H(\lambda) : X \longrightarrow F, \text{ tal que } H(\lambda)(x) = H_x(\lambda).$$

Considere d a dimensão da variedade X e $n_1, \dots, n_j$ naturais fixados tais que $n_1 \cdots + n_j = d$. Dado $k \in \mathbb{N}$, definimos por

$$\Lambda^k(\alpha_1, \dots, \beta_j; n_1, \dots, n_j)$$

o conjunto tal que, para cada $x \in X$ existam racionais $\alpha_1 < \beta_1 < \alpha_2 < \cdots < \beta_j$, com $\beta_i - \alpha_i < 1/k$ e

$$\tilde{E}_i(x) = G_x(\alpha_i) \cap H_x(\beta_i),$$

tal que

$$\dim \tilde{E}_i(x) = n_i,$$

para todo $1 \leq i \leq j$.

Pela maneira que foi definido, note que $\tilde{E}_i(x)$ é um subespaço, pois $G_x(\alpha_i)$ e $H_x(\beta_i)$ também são, para todo $1 \leq i \leq j$.

Afirmação 1: $\tilde{E}_i \cap \tilde{E}_p = \{0\}$, para todo i, $p \in \{1, \dots, j\}$ com $i \neq p$.

Sem perda de generalidade, suponha por absurdo que exista $v \in \tilde{E}_1(x) \cap \tilde{E}_2(x)$ não-nulo, isto é, $v \in G_x(\alpha_1) \cap H_x(\beta_1)$ e $v \in G_x(\alpha_2) \cap H_x(\beta_2)$.

À fortiori,

$$\begin{aligned}
\limsup_{n \longrightarrow +\infty} \frac{1}{n} \log \|T_{-n}(x)v\| \leq -\alpha_2 &\Leftrightarrow -\limsup_{n \longrightarrow +\infty} \frac{1}{n} \log \|T_{-n}(x)v\| \geq \alpha_2 \\
&\Leftrightarrow \liminf_{n \longrightarrow +\infty} \frac{1}{-n} \log \|T_{-n}(x)v\| \geq \alpha_2 \\
&\Leftrightarrow \liminf_{n \longrightarrow -\infty} \frac{1}{n} \log \|T_{n}(x)v\| \geq \alpha_2.
\end{aligned}$$

Como

$$\limsup_{n \longrightarrow -\infty} \frac{1}{n} \log \|T_n(x)v\| \geq \liminf_{n \longrightarrow -\infty} \frac{1}{n} \log \|T_n(x)v\|$$

e $v \in H_x(\beta_1)$, podemos concluir que

$$\beta_1 \geq \limsup_{n \longrightarrow -\infty} \frac{1}{n} \log \|T_n(x)v\| \geq \alpha_2.$$

Isto é uma contradição, pois $\beta_1 < \alpha_2$. Em vista disto, temos $\tilde{E}_1 \cap \tilde{E}_2 = \{0\}$.

De maneira geral, isto conclui que os subespaços $\tilde{E}_i$ são disjuntos dois a dois, com $1 \leq i \leq j$.

Afirmação 2: $\Lambda^k(\alpha_1, \ldots, \beta_j; n_1, \ldots, n_j)$ é mensurável.

Consideremos α, $\beta \in \mathbb{Q}$, com $\alpha < \beta$. Definimos as seguintes funções:

$$\eta_{\alpha,\beta} : X \longrightarrow F, \text{ com } \eta_{\alpha,\beta}(x) = \tilde{E}_x = G_x(\alpha) \cap H_x(\beta)$$

e

$$\tilde{\eta}_{\alpha,\beta} : X \longrightarrow \mathbb{N}, \text{ com } \tilde{\eta}_{\alpha,\beta}(x) = \dim(G_x(\alpha) \cap H_x(\beta)),$$

que naturalmente, serão funções mensuráveis.

Em vista disto, chegaremos na seguinte conclusão:

$$\Lambda^k(\alpha_1, \beta_1, \ldots, \beta_j; n_1, \ldots, n_j) = \bigcap_{i=1}^{j} (\tilde{\eta}(\alpha_i, \beta_i))^{-1}(n_i).$$

Em outras palavras, temos que $\Lambda^k(\alpha_1, \beta_1, \ldots, \beta_j; n_1, \ldots, n_j)$ é pré-imagem de conjuntos mensuráveis, provando assim a afirmação.

Afirmação 3: $\Lambda = \bigcap_{k \geq 1} \bigcup_{\substack{j \in \mathbb{N} \\ n_i \in \mathbb{N} \\ \alpha_i, \beta_i \in \mathbb{Q}}} \Lambda^k(\alpha_1, \ldots, \beta_j; n_1, \ldots, n_j)$.

Seja $x \in \Lambda$. Assim, existem $m_1, \ldots, m_j$ naturais tais que $m_1 + \cdots + m_j = d$, a fibra F_x admite uma decomposição tal que

$$F_x = E_1(x) \oplus E_2(x) \oplus \cdots \oplus E_j(x),$$

com $\dim E_i(x) = m_i$ e existem números reais $\lambda_1 < \cdots < \lambda_j$ tais que

$$\lim_{n \longrightarrow \pm\infty} \frac{1}{n} \log \|T_n(x)v\| = \lambda_i,$$

para todo $v \in E_i(x) \setminus \{0\}$, com $1 \leq i \leq j$.

Fixado $k \in \mathbb{N}$, para cada λ_i podemos tomar racionais α_i e β_i, tais que

$$\alpha_i \leq \lambda_i \leq \beta_i, \text{ com } \beta_i - \alpha_i \leq \frac{1}{k}.$$

Desta forma, para todo $v \in E_i(x) \setminus \{0\}$, temos

$$\begin{aligned}
\lim_{n \longrightarrow \pm\infty} \frac{1}{n} \log \|T_n(x)v\| = \lambda_i &\Leftrightarrow - \lim_{n \longrightarrow \pm\infty} \frac{1}{n} \log \|T_n(x)v\| = -\lambda_i \\
&\Leftrightarrow - \lim_{n \longrightarrow \pm\infty} \frac{1}{n} \log \|T_n(x)v\| \leq -\alpha_i \\
&\Leftrightarrow - \liminf_{n \longrightarrow +\infty} \frac{1}{-n} \log \|T_{-n}(x)v\| \leq -\alpha_i \\
&\Leftrightarrow \limsup_{n \longrightarrow +\infty} - \left(\frac{1}{-n} \right) \log \|T_{-n}(x)v\| \leq -\alpha_i \\
&\Leftrightarrow \limsup_{n \longrightarrow +\infty} \frac{1}{n} \log \|T_{-n}(x)v\| \leq -\alpha_i
\end{aligned} \tag{2.3}$$

e

$$\begin{aligned}
\lim_{n \longrightarrow \pm\infty} \frac{1}{n} \log \|T_n(x)v\| = \lambda_i &\Leftrightarrow \lim_{n \longrightarrow -\infty} \frac{1}{n} \log \|T_n(x)v\| = \lambda_i \\
&\Leftrightarrow \lim_{n \longrightarrow -\infty} \frac{1}{n} \log \|T_n(x)v\| \leq \beta_i \\
&\Leftrightarrow \limsup_{n \longrightarrow -\infty} \frac{1}{n} \log \|T_n(x)v\| \leq \beta_i.
\end{aligned} \tag{2.4}$$

Segue de (2.3) e (2.4) que $v \in G_x(\alpha_i) \cap H_x(\beta_i)$, isto é, $v \in \tilde{E}_i(x)$. Mais ainda, temos que $E_i(x) \subset \tilde{E}_i(x)$ e, consequentemente, $\dim(E_i(x)) \leq \dim(\tilde{E}_i(x))$, para todo $1 \leq i \leq j$.

Considere então $n_1, \ldots, n_j$ naturais tais que $\dim \tilde{E}_i(x) = n_i$, com $n_1 + \cdots + n_j = d$. Sem perda de generalidade, suponha que $\dim E_1(x) < \dim \tilde{E}_1(x)$. Então,

$$d = m_1 + m_2 + \cdots + m_j < n_1 + m_2 + \cdots + m_j \leq n_1 + n_2 + \cdots + n_j = d,$$

o que é uma contradição. Logo, para todo $1 \leq i \leq j$, temos que $\dim(E_i(x)) = \dim(\tilde{E}_i(x)) = m_i$ e, por conseguinte, $E_i(x) = \tilde{E}_i(x)$.

Como resultado, temos que

$$F_x = \tilde{E}_1(x) \oplus \cdots \oplus \tilde{E}_j(x),$$

com

$$\tilde{E}_i(x) = G_x(\alpha_i) \cap H_x(\beta_i) \text{ e } \dim(\tilde{E}_i(x)) = n_i,$$

para todo $1 \leq i \leq j$.

Em outras palavras, $x \in \Lambda^k(\alpha_1, \ldots, \beta_j; m_1, \ldots, m_j)$ e como $k \in \mathbb{N}$ foi escolhido de modo arbitrário, concluímos que $x \in \bigcap\limits_{k \geq 1} \bigcup\limits_{\substack{j \in \mathbb{N} \\ n_i \in \mathbb{N} \\ \alpha_i, \beta_i \in \mathbb{Q}}} \Lambda^k(\alpha_1, \ldots, \beta_j; n_1, \ldots, n_j)$. Assim,

$$\Lambda \subset \bigcap_{k \geq 1} \bigcup_{\substack{j \in \mathbb{N} \\ n_i \in \mathbb{N} \\ \alpha_i, \beta_i \in \mathbb{Q}}} \Lambda^k(\alpha_1, \ldots, \beta_j; n_1, \ldots, n_j).$$

Agora, suponha $x \in \bigcap\limits_{k \geq 1} \bigcup\limits_{\substack{j \in \mathbb{N} \\ n_i \in \mathbb{N} \\ \alpha_i, \beta_i \in \mathbb{Q}}} \Lambda^k(\alpha_1, \ldots, \beta_j; n_1, \ldots, n_j)$, isto é, dado $k \in \mathbb{N}$, consideremos $x \in \Lambda^k(\alpha_1^k, \ldots, \beta_{j_k}^k; n_1^k, \ldots, n_{j_k}^k)$.

Como $n_1^k + \cdots + n_{j_k}^k = d$, a menos de uma subsequência, podemos assumir n_i^k constante igual a n_i, para todo $k \in \mathbb{N}$.

Considerando a sequência limitada $(j_k)_{k \in \mathbb{N}}$, a menos de uma subsequência podemos obter j_k constante, isto é, $j_k = j$ para um k suficientemente grande. Assim,

$$x \in \Lambda^k(\alpha_1^k, \ldots, \beta_j^k; n_1, \ldots, n_j).$$

Considere então o subsespaço $\tilde{E}_i^k(x) = G_x(\alpha_i^k) \cap H_x(\beta_i^k)$, tal que $\dim \tilde{E}_i^k(x) = n_i$ e $n_1 + \cdots + n_j = d$, para todo $1 \leq i \leq j$. Observe que a notação pelo símbolo k serve para mostrar a dependência do mesmo, de modo que $(\alpha_i^k - \beta_i^k) < 1/k$.

Dadas as sequências $(\alpha_i^k)_{k \in \mathbb{N}}$ e $(\beta_i^k)_{k \in \mathbb{N}}$, note que ambas são monótonas e limitadas, logo convergem. Mais ainda, como $\beta_i^k - \alpha_i^k < 1/k$, as sequências possuem o mesmo limite, digamos λ_i, isto é $\alpha_i^k \longrightarrow \lambda_i$ e $\beta_i^k \longrightarrow \lambda_i$, quando $k \longrightarrow +\infty$.

Além disso, repare que

$$G_x(\alpha_i^{k+1}) \subset G_x(\alpha_i^k) \text{ e } H_x(\beta_i^{k+1}) \subset H_x(\beta_i^k),$$

logo

$$\tilde{E}_i^{k+1} = G_x(\alpha_i^{k+1}) \cap H_x(\beta_i^{k+1}) \subset G_x(\alpha_i^k) \cap H_x(\beta_i^{k+1}) = \tilde{E}_i^k.$$

Dado $\epsilon > 0$, tomemos $v \in G_x(\alpha_i^k) \setminus \{0\}$. Pelo fato de $\alpha_i^k \longrightarrow \lambda_i$, temos

$$\limsup_{n \longrightarrow +\infty} \frac{1}{n} \log \|T_{-n}(x)v\| \leq -\alpha_i^k \leq -\lambda_i + \epsilon,$$

isto é

$$\begin{aligned} \lambda_i - \epsilon \geq \alpha_i^k &\geq -\limsup_{n \longrightarrow +\infty} \frac{1}{n} \log \|T_{-n}(x)v\| \\ &= \liminf_{n \longrightarrow +\infty} \frac{1}{-n} \log \|T_{-n}(x)v\| \\ &= \liminf_{n \longrightarrow -\infty} \frac{1}{n} \log \|T_n(x)v\| . \end{aligned} \tag{2.5}$$

Por outro lado, se considerarmos $v \in H_x(\beta_i^k) \setminus \{0\}$ e como $\beta_i^k \longrightarrow \lambda_i$, temos

$$\limsup_{n \longrightarrow -\infty} \frac{1}{n} \log \|T_n(x)v\| \leq \beta_i^k \leq \lambda_i + \epsilon. \tag{2.6}$$

Como $\epsilon > 0$ é arbitrário, de (2.5) e (2.6) obtemos

$$\limsup_{n \longrightarrow -\infty} \frac{1}{n} \log \|T_n(x)v\| \leq \lambda_i \leq \liminf_{n \longrightarrow -\infty} \frac{1}{n} \log \|T_n(x)v\| ,$$

isto é,

$$\lim_{n \longrightarrow -\infty} \frac{1}{n} \log \|T_n(x)v\| = \lambda_i, \tag{2.7}$$

para todo $v \in \tilde{E}_i^k(x) = G_x(\alpha_i^k) \cap H_x(\beta_i^k)$ não-nulo, com $1 \leq i \leq j$.

Pelas circunstâncias, temos que $\tilde{E}_i = \bigcap_{k=1}^{+\infty} \tilde{E}_i^k$ e relembrando que $\dim \tilde{E}_i^k = n_i$, concluímos que $\dim \tilde{E}_i = n_i$, para todo $1 \leq i \leq j$.

Assim, dado $\epsilon > 0$ e $v \in \tilde{E}_i(x) \setminus \{0\}$, definimos as seguintes funções:

$$\tilde{A}_\epsilon(x) = \sup_{v \in \tilde{E}_i(x) \setminus \{0\}} \tilde{A}_\epsilon(x, v)$$

e

$$\tilde{C}_\epsilon(x) = \inf_{v \in \tilde{E}_i(x) \setminus \{0\}} \tilde{C}_\epsilon(x, v),$$

com

$$\tilde{A}_\epsilon(x, v) = \sup_{n\in\mathbb{N}} \left\{ \|T_{-n}(x)v\| \exp(n(\lambda_i + \epsilon)) \|v\|^{-1} \right\}$$

e

$$\tilde{C}_\epsilon(x, v) = \inf_{n\in\mathbb{N}} \left\{ \|T_{-n}(x)v\| \exp(n(-\lambda_i + \epsilon)) \|v\|^{-1} \right\}.$$

Poderemos ver mais a frente que $\tilde{A}_\epsilon(x)$ e $\tilde{C}_\epsilon(x)$ estão bem definidas e são mensuráveis. Mais ainda, as funções $\log(\tilde{A}_\epsilon \circ f) - \log \tilde{A}_\epsilon$ e $\log(\tilde{C}_\epsilon \circ f) - \log \tilde{C}_\epsilon$ são μ-integráveis. Assim, pelo Lema 1.6, teremos que

$$\lim_{n\longrightarrow\pm\infty} \frac{1}{n} \log \tilde{A}_\epsilon \circ f^{-n}(x) = \lim_{n\longrightarrow\pm\infty} \frac{1}{n} \log \tilde{C}_\epsilon \circ f^{n}(x) = 0 \tag{2.8}$$

(veja a **Afirmação 7** da 3ª Etapa da seção 2.2.3). Assim sendo, repare que

$$\begin{aligned}
\|v\| &= \|T_{-n}(f^n(x))T_n(x)v\| \\
&= \|T_{-n}(f^n(x))T_n(x)v\| \exp(-n(\lambda_i + \epsilon)) \exp(n(\lambda_i + \epsilon)) \|T_n(x)v\| \|T_n(x)v\|^{-1} \\
&\leq \tilde{A}_\epsilon(f^n(x)) \exp(-n(\lambda_i + \epsilon)) \|T_n(x)v\|.
\end{aligned}$$

para todo $v \in \tilde{E}_i(x) \setminus \{0\}$. Aplicando a função log e dividindo por n na desigualdade acima, obtemos que

$$\frac{1}{n} \log \|v\| \leq \frac{1}{n} \log \tilde{A}_\epsilon(f^n(x)) - (\lambda_i + \epsilon) + \frac{1}{n} \log \|T_n(x)v\|,$$

isto é,

$$\frac{1}{-n} \log \|v\| \geq \frac{1}{-n} \log \tilde{A}_\epsilon(f^n(x)) + (\lambda_i + \epsilon) + \frac{1}{-n} \log \|T_n(x)v\|,$$

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

Por (2.8), chegamos ao seguinte resultado:

$$\begin{aligned}
\liminf_{n\longrightarrow+\infty} \frac{1}{-n} \log \|T_n(x)v\| \leq -(\lambda_i + \epsilon) &\Leftrightarrow -\limsup_{n\longrightarrow+\infty} \frac{1}{n} \log \|T_n(x)v\| \geq -(\lambda_i + \epsilon) \\
&\Leftrightarrow \limsup_{n\longrightarrow+\infty} \frac{1}{n} \log \|T_n(x)v\| \leq \lambda_i + \epsilon.
\end{aligned} \tag{2.9}$$

Por outro lado,

$$
\begin{aligned}
\|v\| &= \|T_n(f^{-n}(x))T_{-n}(x)v\| \\
&= \|T_n(f^{-n}(x))T_{-n}(x)v\| \exp(n(-\lambda_i+\epsilon)) \exp(-n(-\lambda_1+\epsilon)) \|T_{-n}(x)v\| \, \|T_{-n}(x)v\|^{-1} \\
&\geq \tilde{C}_\epsilon(f^{-n}(x)) \exp(-n(-\lambda_i+\epsilon)) \|T_{-n}(x)v\| .
\end{aligned}
$$

Aplicando a função log e dividindo por $-n$ a desigualdade acima, temos

$$\frac{1}{-n}\log\|v\| \leq \frac{1}{-n}\log\tilde{C}_\epsilon(f^{-n}(x)) + (-\lambda_i+\epsilon) + \frac{1}{-n}\log\|T_{-n}(x)v\| ,$$

para todo $n \in \mathbb{N}$. Assim, novamente por (2.8), concluímos que

$$\liminf_{n\longrightarrow+\infty}\frac{1}{n}\log\|T_n(x)v\| \geq \lambda_i - \epsilon. \tag{2.10}$$

Como $\epsilon > 0$ é arbitrário, de (2.9) e (2.10) que

$$\liminf_{n\longrightarrow+\infty}\frac{1}{n}\log\|T_n(x)v\| \geq \lambda_i \geq \limsup_{n\longrightarrow+\infty}\frac{1}{n}\log\|T_n(x)v\| ,$$

implicando em

$$\lim_{n\longrightarrow+\infty}\frac{1}{n}\log\|T_n(x)v\| = \lambda_i. \tag{2.11}$$

para todo $v \in \tilde{E}_i(x) \setminus \{0\}$, com $1 \leq i \leq j$.

Por (2.7) e (2.11), podemos concluir o seguinte resultado:

$$\lim_{n\longrightarrow\pm\infty}\frac{1}{n}\log\|T_n(x)v\| = \lambda_i,$$

para todo $v \in \tilde{E}_i(x) \setminus \{0\}$, com $1 \leq i \leq j$.

Pela **Afirmação 1**, os subespaços próprios $\tilde{E}_i(x)$ são disjuntos dois a dois. Além disso, temos que $\dim\tilde{E}_i(x) = n_i$, com $n_1 + \cdots + n_j = d$. Assim, pela construção dos números reais λ_i, temos que $\lambda_1 < \cdots < \lambda_j$ e o fibrado F_x admite uma decomposição tal que

$$F_x = \tilde{E}_i(x) \oplus \cdots \oplus \tilde{E}_j(x).$$

Em outras palavras, temos por definição que x é ponto regular, isto é, $x \in \Lambda$. Isto conclui a **Afirmação 3** e pela **Afirmação 2**, segue que Λ é mensurável.

2.2.3 Demonstração do Lema 2.10

Faremos aqui, a demonstração do Lema 2.10 em três etapas. No primeiro momento, mostraremos que G é um subfibrado mensurável T-invariante de F. Em seguida, teremos para μ-q.t.p $x \in X$ que $\dim G_x > 0$ e por fim, provaremos a existência do limite citado no enunciado do lema.

Previamente, vamos provar que G_x é um subespaço de F_x, para todo $x \in X$.

• Dados u, $v \in G_x$, temos

$$\begin{aligned}
\limsup_{n \longrightarrow +\infty} \frac{1}{n} \log \|T_{-n}(x)(u+v)\| &= \limsup_{n \longrightarrow +\infty} \frac{1}{n} \log \|T_{-n}(x)u + T_{-n}(x)v\| \\
&\leq \limsup_{n \longrightarrow +\infty} \frac{1}{n} \log \left(\|T_{-n}(x)u\| + \|T_{-n}(x)v\|\right) \\
&\leq \limsup_{n \longrightarrow +\infty} \frac{1}{n} \log \left(2 \max\{\|T_{-n}(x)u\|, \|T_{-n}(x)v\|\}\right) \\
&= \limsup_{n \longrightarrow +\infty} \frac{1}{n} \log 2 + \limsup_{n \longrightarrow +\infty} \frac{1}{n} \log \left(\max\{\|T_{-n}(x)u\|, \|T_{-n}(x)v\|\}\right) \\
&\leq -\lambda_1.
\end{aligned}$$

Desta forma, $u + v \in G_x$.

• Seja $\alpha \in \mathbb{R}$, então

$$\begin{aligned}
\limsup_{n \longrightarrow +\infty} \frac{1}{n} \log \|T_{-n}(x)(\alpha v)\| &= \limsup_{n \longrightarrow +\infty} \frac{1}{n} \log \left(|\alpha| \, \|T_{-n}(x)(v)\|\right) \\
&= \limsup_{n \longrightarrow +\infty} \frac{1}{n} \log |\alpha| + \limsup_{n \longrightarrow +\infty} \frac{1}{n} \log \|T_{-n}(x)v\| \\
&\leq -\lambda_1.
\end{aligned}$$

Desta forma, $\alpha v \in G_x$. Isso conclui que G_x é subespaço de F_x.

$1^{\underline{a}}$ Etapa: G é um subfibrado mensurável T-invariante de F.

Dado $x \in X$, mostremos que $T(x)(G_x) = G_{f(x)}$. Para tal, seja $v \in G_x$ não-nulo e notemos que

$$\begin{aligned}
\limsup_{n \longrightarrow +\infty} \frac{1}{n} \log \|T_{-n}(f(x)) \circ T(x)v\| &= \limsup_{n \longrightarrow +\infty} \frac{1}{n} \log \left\|T_{-(n-1)}(x)v\right\| \\
&= \limsup_{n \longrightarrow +\infty} \left(\frac{1}{n-1}\right)\left(\frac{n-1}{n}\right) \log \|T_{-n+1}(x)v\| \\
&\leq \limsup_{n \longrightarrow +\infty} \frac{1}{n-1} \log \left\|T_{-(n-1)}(x)v\right\|
\end{aligned}$$

$$\leq -\lambda_1.$$

Isto mostra que $T(x)v \in G_{f(x)}$, para $v \in G_x \setminus \{0\}$, isto é, $T(x)(G_x) \subset G_{f(x)}$. De maneira análoga, temos que $G_{f(x)} \subset T(x)(G_x)$, garantindo que $T(x)(G_x) = G_{f(x)}$. Observe que como T é um isomorfismo, podemos concluir que $\dim T(x)(G_x) = \dim G_{f(x)}$.

Mostremos à seguir que G_x é um subfibrado mensurável. Para tal, seja $k \in \mathbb{N}$, definimos

$$G_x^k = \left\{ v \in F_x;\ \limsup_{n \longrightarrow +\infty} \frac{1}{n} \log \|T_{-n}(x)v\| \leq -\lambda_1 + \tfrac{1}{k} \right\}.$$

Da mesma forma que G_x é um subespaço vetorial de F_x, G_x^k também é. Note que $\bigcap_{k=1}^{+\infty} G_x^k = G_x$, isto é, existe $k_0 \in \mathbb{N}$ tal que $G_x^k = G_x$, para todo $k \geq k_0$, pelo fato de F ser um espaço de dimensão finita. Considere então

$$X_k = \left\{ x \in X;\ G_x^k = G_x \right\},$$

observando que $X = \bigcup_{k=1}^{+\infty} X_k$.

Afirmação 1: X_k é mensurável, para todo k.

Considere as funções

$$\lambda_1 : X \longrightarrow \mathbb{R}, \text{ com } \lambda_1(x) = \limsup_{n \longrightarrow +\infty} \frac{1}{n} \log \|T_n(x)\|$$

e

$$\phi : F \longrightarrow \mathbb{R}, \text{ com } \phi(x,v) = \limsup_{n \longrightarrow +\infty} \frac{1}{n} \log \|T_{-n}(x)v\|.$$

Pelo fato das sequências $((1/n)\log\|T_n(x)\|)_{n\in\mathbb{N}}$ e $((1/n)\log\|T_{-n}(x)v\|)_{n\in\mathbb{N}}$ serem limitadas e compostas de funções mensuráveis, temos que λ_1 e ϕ são funções mensuráveis.

Considere a projeção $\pi : F \longrightarrow X$ e suponha que $x \notin X_k$, isto é, $G_x^k \neq G_x$. Então existe $v \in G_x^k$ com $v \notin G_x$. Isto é,

$$\begin{aligned}
-\lambda_1(x) < \limsup_{n \longrightarrow +\infty} \frac{1}{n} \log \|T_{-n}(x)v\| \leq -\lambda_1(x) + \frac{1}{k} \quad &\Leftrightarrow \quad -\lambda_1(x) < \phi(x,v) \leq -\lambda_1(x) + \frac{1}{k} \\
&\Leftrightarrow \quad 0 < \phi(x,v) + \lambda_1 \circ \pi(x,v) \leq \frac{1}{k} \\
&\Leftrightarrow \quad (x,v) \in (\phi + \lambda_1 \circ \pi)^{-1}((0,k]) \\
&\Leftrightarrow \quad \pi(x,v) = x \in \pi(\phi + \lambda_1 \circ \pi)^{-1}((0,k]),
\end{aligned}$$

Assim podemos concluir que $X_k = X - \pi(\phi + \lambda_1 \circ \pi)^{-1}((0,k])$. À fortiori, X_k é mensurável para todo k, pelo fato de π, ϕ e λ_1 serem mensuráveis.

Fixemos $k \in \mathbb{N}$. Dados $x \in X_k$ e $m \geq 1$, consideremos

$$G_x^k(m) = \left\{ v \in F_x ; \|T_{-n}(x)v\| \leq m \exp(-n(\lambda_1 - \tfrac{1}{k})) \|v\| \text{, para todo } n \in \mathbb{N} \right\}.$$

Afirmação 2: $G_x^k = \bigcup_{m \geq 1} G_x^k(m)$.

Dado $v \in \bigcup_{m \geq 1} G_x^k(m)$, existe $m_0 \geq 1$ tal que $v \in G_x^k(m_0)$. Pela maneira que definimos $G_x^k(m_0)$, temos

$$\|T_{-n}(x)v\| \leq m_0 \exp(-n(\lambda_1 - \tfrac{1}{k})) \|v\|.$$

Assim, concluímos

$$\begin{aligned} -\frac{1}{n}\log\|T_{-n}(x)v\| \geq -\frac{1}{n}(\log(m_0) + \log\|v\|) + \lambda_1 - \frac{1}{k} &\Leftrightarrow \liminf_{n \longrightarrow +\infty} -\frac{1}{n}\log\|T_{-n}(x)v\| \geq \lambda_1 - \frac{1}{k} \\ &\Leftrightarrow -\liminf_{n \longrightarrow +\infty} -\frac{1}{n}\log\|T_{-n}(x)v\| \leq -\lambda_1 + \frac{1}{k} \\ &\Leftrightarrow \limsup_{n \longrightarrow +\infty} \frac{1}{n}\log\|T_{-n}(x)v\| \leq -\lambda_1 + \frac{1}{k}, \end{aligned}$$

e portanto $v \in G_x^k$.

Seja w um vetor não-nulo em G_x^k então $w/\|w\| \in G_x^k$. Assim

$$\limsup_{n \longrightarrow +\infty} \frac{1}{n}\log\left\|T_{-n}(x)\tfrac{w}{\|w\|}\right\| \leq -\lambda_1 + \frac{1}{k},$$

então existe $n_0 \in \mathbb{N}$ tal que para todo $n \geq n_0$

$$\frac{1}{n}\log\left\|T_{-n}(x)\tfrac{w}{\|w\|}\right\| = \frac{1}{n}\log\|T_{-n}(x)w\| - \frac{1}{n}\log\|w\| \leq -\lambda_1 + \frac{1}{k}. \tag{2.12}$$

Lembrando que T possui norma limitada, podemos escolher $m_1 \geq 1$, tal que

$$\frac{1}{n}\log\|T_{-n}(x)w\| - \frac{1}{n}\log\|w\| \leq \frac{1}{n}\log m_1 - \lambda_1 + \frac{1}{k}, \tag{2.13}$$

para $n = 1, \ldots, n_0 - 1$. Das equações (2.12) e (2.13), concluímos que para todo $n \in \mathbb{N}$

$$\|T_{-n}(x)w\| \leq m_1 \exp(-n(\lambda_1 - \tfrac{1}{k})) \|w\|.$$

Assim $v \in G_x^k(m_1) \subset \bigcup_{m \geq 1} G_x^k(m)$. Portanto, temos $G_x^k = \bigcup_{m \geq 1} G_x^k(m)$.

Definimos agora

$$X_k^m = \left\{x \in X_k;\ G_x^k = G_x^k(m)\right\} = \left\{x \in X_k;\ G_x^k(m) = G_x^k(\tilde{m}) \text{ para todo } \tilde{m} \geq n\right\}.$$

Afirmação 3: $X_k \subset \bigcup_{m \geq 1} X_k^m$, para todo $k \in \mathbb{N}$. Além disso, X_k^m é mensurável para cada $m \geq 1$.

Considere a sequência de funções $(\gamma_k)_{k \in \mathbb{N}}$ tal que $\gamma_k : \pi^{-1}(X_k) \longrightarrow \mathbb{R}$, tal que

$$\gamma_k(x, v) = \|T_{-n}(x)v\| \exp(n(\lambda_1 - \tfrac{1}{k})) \, \|v\|^{-1}.$$

Caso $x \notin X_k^m$, então existe $v \in F_x$ e $\tilde{m} \geq m$ tal que

$$\begin{aligned} m < \gamma_k(x, v) \leq \tilde{m} &\Leftrightarrow \gamma_k(x, v) \in (m, \tilde{m}] \\ &\Leftrightarrow (x, v) \in \gamma_k^{-1}((m, \tilde{m}]) \\ &\Leftrightarrow \pi(x, v) = x \in \pi \circ \gamma_k^{-1}((m, \tilde{m}]) \end{aligned}$$

para todo $\tilde{m} \geq m$. Desta forma $X_k^m = X_k \setminus \bigcup_{\tilde{m}=m}^{+\infty} \pi \circ \gamma_k^{-1}((m, \tilde{m}])$ e como m foi escolhido arbitrariamente, temos que X_k^m é mensurável, para todo $m \geq 1$.

Agora, seja $x \in X_k$ e suponhamos que $\dim G_x^k = d > 0$. Considere $\{v_1, \ldots, v_d\}$ uma base ortogonal do subespaço G_x^k. Como $G_x^k = \bigcup_{m \geq 1} G_x^k(m)$, existem $m_1, \ldots, m_d \geq 1$ tais que $v_j \in G_x^k(m_j)$, para $1 \leq j \leq d$.

Assim, dado $v \in G_x^k$, seja $\alpha_1, \ldots, \alpha_d \in \mathbb{R}$ tais que $v = \sum_{j=1}^{d} \alpha_j v_j \in G_x^k$. Considerando $M = \max\{m_1, \ldots, m_d\}$, para $n \geq 1$ temos:

$$\begin{aligned} \|T_{-n}(x)v\| &= \left\| T_{-n}(x) \left(\sum_{j=1}^{d} \alpha_j v_j \right) \right\| \\ &\leq \sum_{j=1}^{d} |\alpha_j| \, \|T_{-n}(x)v_j\| \leq \sum_{j=1}^{d} |\alpha_j| m_j \exp(-n(\lambda_1 - \tfrac{1}{k})) \\ &\leq \sum_{j=1}^{d} |\alpha_j| M \exp(-n(\lambda_1 - \tfrac{1}{k})) \end{aligned}$$

$$
\begin{aligned}
&\leq M \exp(-n(\lambda_1 - \tfrac{1}{k})) \sum_{j=1}^{d} |\alpha_j| \\
&\leq M \exp(-n(\lambda_1 - \tfrac{1}{k})) \|v\|_2 \\
&\leq M \exp(-n(\lambda_1 - \tfrac{1}{k}))\beta \|v\|
\end{aligned}
$$

onde $\|\cdot\|_2$ é a norma de v com respeito à base $\{v_1, \ldots, v_d\}$ e $\beta > 0$ a constante da equivalência de normas em espaços de dimensão finita. Desta forma, temos que $v \in G_x^k(\beta M)$ o que implica $G_x^k \subset G_x^k(\beta M)$ e portanto $x \in X_k^{\beta M}$.

Afirmação 4: G_x varia semicontinuamente inferior restrito a cada X_k^m.

Lembrando que para cada $x \in X_k^m$, temos que $G_x^k = G_x^k(m)$ e considerando um k suficientemente grande, vale $G_x = G_x^k = G_x^k(m)$.

Considere $(x_i)_{i\in\mathbb{N}}$ uma sequência de X_k^m tal que $x_i \longrightarrow x$, com $x \in X_k^m$. Agora, considere a sequência de vetores $(v_i)_{i\in\mathbb{N}}$ com $v_i \longrightarrow v$, com $v_i \in G_{x_i}$ e $v \in F_x$.

Fixando $n \in \mathbb{N}$, como $v_i \in G_{x_i}$, para todo $i \in \mathbb{N}$ e $x_i \in X_k^m$, temos

$$\|T_{-n}(x_i)v_i\| \leq m \exp(-n(\lambda_1 - \tfrac{1}{k})) \|v_i\| .$$

Assim, fazendo $i \longrightarrow +\infty$, temos

$$\|T_{-n}(x)v\| \leq m \exp(-n(\lambda_1 - \tfrac{1}{k})) \|v\| .$$

Como n foi fixado arbitrariamente, temos que $v \in G_x^k(m)$, ou seja, $v \in G_x^k$ desde que $x \in X_k^m$. Portanto G_x varia semicontinuamente inferior quando restrito a X_k^m.

Definimos agora o seguinte conjunto

$$X_k^m(j) = \{x \in X_k^m;\ \dim G_x \geq j\}.$$

Uma consequência direta da **Afirmação 4** é que os conjuntos $X_k^m(j)$ são fechados em X_k^m. Mais ainda, podemos observar que G_x varia continuamente quando restrito a $X_k^m(j) \setminus X_k^m(j+1)$, pois de fato a dimensão de G_x neste conjunto é constante.

Assim, podemos concluir que $X_k^m = X_k^m(j) \setminus X_k^m(j+1)$ forma uma partição mensurável de X_k^m, e como $(X_k)_{k\in\mathbb{N}}$ é uma cobertura mensurável de X (**Afirmação 1**) e, $(X_k^m)_{m\in\mathbb{N}}$ é uma cobertura mensurável de X_k (**Afirmação 3**), temos assim que $(X_k^m(j) \setminus X_k^m(j+1))_{j,m,k\in\mathbb{N}}$ é uma cobertura mensurável de X.

Definindo

$$\widehat{\Lambda_j} = \{x \in X;\ \dim G_x = j\},$$

temos que $\widehat{\Lambda_j} = X_k^m(j) \setminus X_k^m(j+1)$, e portanto é um conjunto mensurável. Mais ainda, usando o fato que G_x varia continuamente em $X_k^m(j) \setminus X_k^m(j+1)$, temos que G_x varia mensuravelmente em $\widehat{\Lambda_j}$.

Até aqui, mostramos que G é um subfibrado mensurável T-invariante de F. No entanto, o mesmo ainda poderia ser o subfibrado trivial nulo. O que mostraremos a seguir é que G_x de fato, tem dimensão positiva.

2ª Etapa: $\dim G_x > 0$, **para** μ**-q-t.p.** $x \in X$

Fixemos inicialmente $k \in \mathbb{N}$. Assim, para $m \geq 1$ definimos os conjuntos

$$Y_k^m = \left\{x \in X;\ \text{existe } v \in F_x \setminus \{0\} \text{ tal que } \|T_{-n}(x)v\| \leq \exp(-n(\lambda_1 - \tfrac{1}{k}))\,\|v\|\,,\ 1 \leq n \leq m\right\}$$

e

$$Y_k = \bigcap_{m \geq 1} Y_k^m.$$

Afirmação 5: O conjunto Y_k possui medida positiva.

Note que $Y_k^1 \supset Y_k^2 \supset Y_k^3 \supset \ldots$, desta forma se existe $\delta > 0$ tal que $\mu(Y_k^m) \geq \delta$ para todo $m \geq 1$, então teremos que $\mu(Y_k) \geq \delta$.

Assim sendo, fixemos $m \in \mathbb{N}$. Garantiremos a existência de $\delta > 0$ usando o clássico lema de Pliss, o qual enunciaremos abaixo.

Lema 2.12. *[Pliss] Dado* $\lambda \in \mathbb{R}$, $\epsilon > 0$ *e* $A > 0$, *existe* $\delta = \delta(\lambda, \epsilon, A) > 0$ *tal que para todo* $a_0, \ldots, a_{N-1} \in \mathbb{R}$ *com* $\sum_{i=0}^{N-1} a_i \leq N\lambda$ *e* $|a_i| \leq A$, *para* $0 \leq i \leq N-1$, *então existe* $l \geq N\delta$ *e* $0 < n_1 \leq \cdots \leq n_l \leq N-1$ *tal que*

$$\sum_{i=n}^{n_j-1} a_i \leq (n_j - n)(\lambda + \epsilon)$$

para todo $0 \leq n < n_j$ *e* $1 \leq j \leq l$.

Demonstração: Ver [2], página 365.

Consideremos $\lambda = -\lambda_1 + 1/2k$, $\epsilon = 1/2k > 0$ e $A = \log \left\|T^{-1}\right\|$ no lema de Pliss, e seja $\delta > 0$ dado pelo lema.

Sabemos que $\lambda_1(T, x) = \lambda_1(T)$, para μ-q.t.p. x em X. Assim, seja $x \in X$ tal que $\lambda_1(T, x) = \lambda_1(T)$. Pela definição de $\lambda_1(T, x)$, dado $\epsilon = 1/2k > 0$, existe $N \in \mathbb{N}$ arbitrariamente grande tal que

$$\frac{1}{N}\log\|T_N(x)\| \geq \lambda_1 - \frac{1}{2k}$$

então

$$\|T_N(x)\| \geq \exp(N(\lambda_1 - \tfrac{1}{2k})).$$

Por definição de norma, existe $v \in F_x \setminus \{0\}$ tal que

$$\left\|T_N(x)\frac{v}{\|v\|}\right\| \geq \exp(N(\lambda_1 - \tfrac{1}{2k})).$$

E assim,

$$\|T_N(x)v\| \geq \exp(N(\lambda_1 - \epsilon))\,\|v\|\,. \tag{2.14}$$

Agora tomemos $a_i = \log\left\|T^{-1}(f^{i+1}(x))\right\|(T_{i+1}(x)v/\|T_{i+1}(x)v\|)$, com $0 \leq i \leq N-1$. Uma vez que

$$a_i \leq \log\left\|T^{-1}(f^{i+1}(x))\right\|\frac{\|T_{i+1}(x)v\|}{\|T_{i+1}(x)v\|} = \log\left\|T^{-1}(f^{i+1}(x))\right\| \leq \log\left\|T^{-1}\right\|,$$

temos

$$\begin{aligned}
\sum_{i=0}^{N-1} a_i &= \sum_{i=0}^{N-1}\log\left\|T^{-1}(f^{i+1}(x))\frac{T_{i+1}(x)v}{\|T_{i+1}(x)v\|}\right\| = \sum_{i=0}^{N-1}\log\left\|T^{-1}(f^{i+1}(x))\circ T(f^{i+1}(x))\frac{T_i(x)v}{\|T_{i+1}(x)v\|}\right\| \\
&= \sum_{i=0}^{N-1}\log\left\|\frac{T_i(x)v}{\|T_{i+1}(x)v\|}\right\| = \log\left\|\frac{v}{T(x)v}\right\| + \log\left\|\frac{T(x)v}{T_2(x)v}\right\| + \cdots + \log\left\|\frac{T_{N-1}v}{T_N(x)v}\right\| \\
&= \log\|v\| - \log\|T(x)v\| + \log\|T(x)v\| - \log\|T_2(x)v\| + \cdots + \log\|T_{N-1}(x)v\| - \log\|T_N(x)v\| \\
&= \log\|v\| - \log\|T_N(x)v\| = \log\left\|\frac{v}{T_N(x)v}\right\|.
\end{aligned}$$

Portanto, pela equação (2.14), temos

$$\left\|\frac{v}{T_N(x)v}\right\| \leq \exp(-N(\lambda_1 - \tfrac{1}{2k})),$$

o que implica

$$\sum_{i=0}^{N-1} a_i = \log\left\|\frac{v}{T_N(x)v}\right\| \leq -N(\lambda_1 - \frac{1}{2k}) = N(-\lambda_1 + \frac{1}{2k}).$$

Portanto, pelo Lema 2.12, existem $n_1, \ldots, n_l$ com $0 < n_1 \leq \cdots \leq n_l \leq N-1$ tais que

$$\sum_{i=n}^{n_j-1} a_i = \log\left\|\frac{T_n(x)v}{\|T_{n_j}(x)v\|}\right\| \leq (n_j - n)\left(-\lambda_1 + \frac{1}{k}\right) \tag{2.15}$$

para todo $0 \le n < n_j$ e $1 \le j \le l$. Agora, observe que

$$T_n(x)v = T_{n-n_j}(f^{n_j}(x)) \circ T_{n_j}(x)v. \tag{2.16}$$

Substituindo (2.15) em (2.16) e tomando $w_j = T_{n_j}(x)v / \left\| T_{n_j}(x) \right\|$, obtemos

$$\begin{aligned} \sum_{i=n}^{n_j-1} a_i = \log \left\| \frac{T_n(x)v}{\left\| T_{n_j}(x) \right\|} \right\| = \log \left\| \frac{T_{n-n_j}(f^{n_j}(x)) \circ T_{n_j}(x)v}{\left\| T_{n_j}(x)v \right\|} \right\| &= \log \left\| T_{n-n_j}(f^{n_j}(x))w_j \right\| \\ &\le (n_j - n)\left(-\lambda_1 + \frac{1}{k}\right). \end{aligned} \tag{2.17}$$

para todo $0 \le n < n_j$ e $1 \le j \le l$. Da última desigualdade da equação (2.17), segue que para $1 \le j \le l$, temos

$$\left\| T_{-(n_j-n)}(f^{n_j}(x))w_j \right\| \le \exp(-(n_j - n)(\lambda_1 - \tfrac{1}{k})), \text{ para todo } 0 \le n < n_j,$$

isto é, $f^{n_j}(x) \in Y_k^{n_j}$.

Agora, observe que se $n_j \ge m$, então $Y_k^{n_j} \subset Y_k^m$. Pelo Lema 2.12, sabemos que $l > N\delta$. Também, escolhemos $N \in \mathbb{N}$ arbitrariamente grande tal que $N\delta > m$. Assim, com certeza $n_j \ge m$ para todo $l - m + 1 \le j \le l$.

Desta forma, podemos concluir que $f^{n_j}(x) \in Y_k^m$, para todo $l - m + 1 \le j \le l$, o que implica que

$$\# \left\{ 0 \le i \le N;\ f^i(x) \in Y_k^m \right\} \ge l - m. \tag{2.18}$$

Como x foi tomado arbitrariamente μ-q.t.p., podemos assumir pelo Teorema 1.4 que o limite

$$\tau(x, Y_k^m) = \lim_{n \longrightarrow +\infty} \frac{1}{n} \# \left\{ 0 \le i \le n;\ f^i(x) \in Y_k^m \right\}.$$

existe. Por outro lado, como N pôde ser tomado arbitrariamente grande, temos em particular que

$$\tau(x, Y_k^m) = \lim_{N \longrightarrow +\infty} \frac{1}{N} \# \left\{ 0 \le i \le N;\ f^i(x) \in Y_k^m \right\}, \tag{2.19}$$

Usando a equação (2.18) e o fato que $N\delta < l$, temos

$$\lim_{N \longrightarrow +\infty} \frac{1}{N} \# \left\{ 0 \le i \le N;\ f^i(x) \in Y_k^m \right\} \ge \lim_{N \longrightarrow +\infty} \frac{1}{N}(l - m) \ge \lim_{N \longrightarrow =\infty} \frac{1}{N}(N\delta - m) = \delta.$$

Desta forma, mostramos que $\tau(x, Y_k^m) \ge \delta$, para μ-q.t.p. $x \in X$. Como μ é ergódica, temos $\delta \leq \tau(x, Y_k^m) = \mu(Y_k^m)$ e como m foi escolhido arbitrariamente, temos que

$$\mu(Y_k) = \mu\left(\bigcap_{m\geq 1} Y_k^m\right) \geq \delta.$$

Afirmação 6: Dado $x \in Y_k$, $\dim G_x^k > 0$

Lembrando o fato que G_x é T-invariante, temos que G_x^k também será, para todo $k \in \mathbb{N}$. Isto é, se $\dim G_x^k > 0$, então $\dim G_{f^n(x)}^k > 0$, para todo $n \in \mathbb{N}$.

Assim, dado $x \in Y_k$, existe $(v_m)_{m\in\mathbb{N}}$ de F_x uma sequência limitada de vetores não-nulos tais que $\|v_m\| = 1$ e

$$\|T_{-n}(x)v_m\| \leq \exp(-n(\lambda_1 - \tfrac{1}{k}))\,\|v_m\|, \text{ para } 1 \leq n \leq m.$$

Assim, fixando $n \in \mathbb{N}$, tomemos $(v_{m_j})_{j\in\mathbb{N}}$ a subsequência de $(v_m)_m$ tal que $v_{m_j} \longrightarrow v$ com $\|v\| = 1$ (pois a sequência de vetores esta compreendida uma esfera, assim como o limite da subsequência convergente, pela compacidade), então existe $j_0 \in \mathbb{N}$ tal que $m_j > n$, para todo $j \geq j_0$. Para as escolhas dos v_{m_j}, por continuidade temos que

$$\|T_{-n}(x)v\| \leq \exp(-n(\lambda_1 - \tfrac{1}{k})\,\|v\|\,, \tag{2.20}$$

quando $j \longrightarrow +\infty$.

Como n foi fixado arbitrariamente, a desigualdade (2.20) vale para todo $n \in \mathbb{N}$. Em vista disto, $v \in G_x^k(1) \subset G_x^k$.

Definimos agora

$$\tilde{Y}_k = \bigcup_{n\in\mathbb{Z}} f^n(Y_k).$$

Por construção, note que $\tilde{Y}_k$ é invariante por f e $\mu(\tilde{Y}_k) \geq \mu(Y_k) \geq \delta$. Como μ é ergódica, segue da Proposição 1.8 que $\mu(\tilde{Y}_k) = 1$, para todo $k \in \mathbb{N}$. Isto é, concluímos que $\dim G_x^k > 0$ para μ-q.t.p. $x \in X$.

Pelas afirmações, mostraremos agora que $\dim G_x > 0$, para μ-q.t.p. $x \in X$.

Definindo $\tilde{Y} = \bigcap_{k=1}^{\infty} \tilde{Y}_k$ e como $\mu(\tilde{Y}_k) = 1$, para todo $k \in \mathbb{N}$, segue que $\mu(\tilde{Y}) = 1$.

Seja $x \in \tilde{Y}$, temos que existe $\tilde{k} \in \mathbb{N}$ tal que $G_x = G_x^k$, para todo $k \geq \tilde{k}$. Como $\dim G_x^k > 0$, segue que $\dim G_x > 0$, para μ-q.t.p. $x \in X$.

3ª Etapa: $\lim\limits_{n\longrightarrow\pm\infty} \dfrac{1}{n} \log\|T_n(x)v\| = \lambda_1$, para μ-q.t.p. e todo $v \in G_x$.

Utilizando a definição da função $\lambda_1(T)$, dado $\epsilon > 0$ temos que existe $n_0 \in \mathbb{N}$ tal que

$$\frac{1}{n} \log \|T_n(x)\| \leq \lambda_1 + \epsilon$$

para todo $n \geq n_0$. Assim,

$$\|T_n(x)v\| \exp(-n(\lambda_1 + \epsilon)) \|v\|^{-1} \leq 1$$

para todo $v \in F_x \setminus \{0\}$ (em particular para $v \in G_x \setminus \{0\}$) e $n \geq n_0$.

Logo, para μ-q.t.p. $x \in X$, temos que a função

$$A_\epsilon(x, v) = \sup_{n \in \mathbb{N}} \left\{ \|T_n(x)v\| \exp(-n(\lambda_1 + \epsilon)) \|v\|^{-1} \right\}$$

está bem definida e é mensurável. Em particular

$$A_\epsilon(x) = \sup_{v \in G_x} A_\epsilon(x, v).$$

também está bem definida. Dado $\alpha \in \mathbb{R}$, note que $A_\epsilon(x, v) = A_\epsilon(x, \alpha v)$. Agora, usando a definição de G_x, temos para μ-q.t.p. $x \in X$ que a função

$$C_\epsilon(x, v) = \sup_{n \in \mathbb{N}} \left\{ \|T_n(x)v\| \exp(-n(-\lambda_1 + \epsilon)) \|v\|^{-1} \right\}$$

está bem definida e é mensurável. Assim como antes, definimos

$$C_\epsilon(x) = \sup_{v \in G_x} C_\epsilon(x, v).$$

Afirmação 7: $\displaystyle\lim_{n \longrightarrow \pm\infty} \frac{1}{n} \log A_\epsilon \circ f^{-n}(x) = \lim_{n \longrightarrow \pm\infty} \frac{1}{n} \log C_\epsilon \circ f^n(x) = 0.$

Vamos mostrar que $\displaystyle\lim_{n \longrightarrow \pm\infty} \frac{1}{n} \log C_\epsilon \circ f^n(x) = 0$, pois o outro limite é demonstrado de maneira análoga.

Sem perda de generalidade, podemos supor que $C_\epsilon(x) \geq 1$. Agora, dado $n \in \mathbb{N}$, temos

$$\begin{aligned}
&\|T_{n+1}(x)v\| \exp(-(n+1)(\lambda_1(T) + \epsilon)) \|v\|^{-1} \\
&= \|T_n(f(x))T(x)v\| \exp(-n(\lambda_1(T) + \epsilon)) \exp(-(\lambda_1(T) + \epsilon)) \|v\|^{-1} \|T(x)v\|^{-1} \|T(x)v\| \\
&= C_\epsilon(f(x), T(x)v) \exp(-(\lambda_1(T) + \epsilon)) \|v\|^{-1} \|T(x)v\|.
\end{aligned}$$

Como a igualdade acima vale para todo $n \in \mathbb{N}$, vale em particular para o supremo. Assim,

$$C_\epsilon(x, v) = C_\epsilon(f(x), T(x)v) \exp(-(\lambda_1(T) + \epsilon)) \|v\|^{-1} \|T(x)v\| . \tag{2.21}$$

Como T possui norma limitada, seja $\tilde{\alpha}$, $\tilde{\beta} \geq 0$ tais que $\tilde{\alpha} \leq \|T(x)v\| / \|v\| \leq \tilde{\beta}$. Desta forma,

$$\exp(-(\lambda_1 + \epsilon))\tilde{\alpha} \leq \exp(-(\lambda_1 + \epsilon))\frac{\|T(x)v\|}{\|v\|} \leq \exp(-(\lambda_1 + \epsilon))\tilde{\beta}.$$

Tomando $\alpha = \exp(-(\lambda_1 + \epsilon))\tilde{\alpha}$ e $\beta = \exp(-(\lambda_1 + \epsilon))\tilde{\beta}$, temos portanto

$$\alpha C_\epsilon(f(x), T(x)v) \leq C_\epsilon(f(x), T(x)v) \exp(-(\lambda_1(T) + \epsilon)) \|v\|^{-1} \|T(x)v\| \leq C_\epsilon(f(x), T(x)v)\beta.$$

Ou seja, da equação (2.21), obtemos

$$\alpha C_\epsilon(f(x), T(x)v) \leq C_\epsilon(x, v) \leq C_\epsilon(f(x), T(x)v)\beta. \tag{2.22}$$

Como

$$C_\epsilon(f(x)) = \sup_{v \in G_x} C_\epsilon(f(x), T(x)v),$$

na hipótese de T ser isomorfismo, a equação (2.22) vale para todo $v \in G_x \setminus \{0\}$, em particular, para o supremo. Assim

$$\alpha C_\epsilon(f(x)) \leq C_\epsilon(x) \leq C_\epsilon(f(x))\beta,$$

o que implica

$$\alpha \leq \frac{C_\epsilon(x)}{C_\epsilon(f(x))} \leq \beta.$$

Aplicando a função log na expressão acima, é fácil de ver que a função $\log(C_\epsilon \circ f) - \log C_\epsilon$ é limitada, em particular, integrável. Pelo Lema 1.6, segue que

$$\lim_{n \longrightarrow \pm\infty} \frac{1}{n} \log C_\epsilon \circ f^n(x) = 0, \tag{2.23}$$

para μ-q.t.p. $x \in X$.

Afirmação 8: $\limsup_{n \longrightarrow \pm\infty} \frac{1}{n} \log \|T_n(x)v\| \leq \lambda_1$**, para** μ**-q.t.p.** $x \in X$ **e** $v \in G_x$**.**

Note que

$$\limsup_{n \longrightarrow +\infty} \frac{1}{n} \log \|T_n(x)v\| \leq \lambda_1$$

sai direto da definição da função λ_1, pois

$$\frac{1}{n} \log \|T_n(x)v\| \leq \frac{1}{n} \log \|T_n(x)\| + \frac{1}{n} \log \|v\|, \text{ para todo } n \in \mathbb{N}.$$

Vejamos o $\limsup$, quando $n \longrightarrow -\infty$.

Dado $\epsilon > 0$, para μ-q.t.p. $x \in X$ tome $v \in G_x \setminus \{0\}$ e $n \in \mathbb{N}$. Então,

$$\begin{aligned}
\|v\| &= \|T_n(f^{-n}(x))T_{-n}(x)v\| \\
&= \|T_n(f^{-n}(x))T_{-n}(x)v\| \exp(-n(\lambda_1 + \epsilon)) \exp(n(\lambda_1 + \epsilon)) \|T_{-n}(x)v\| \|T_{-n}(x)v\|^{-1} \\
&\leq A_\epsilon(f^{-n}(x)) \exp(n(\lambda_1 + \epsilon)) \|T_{-n}(x)v\|.
\end{aligned}$$

Aplicando a função log e dividindo por n a desigualdade acima, temos

$$\frac{1}{n} \log \|v\| \leq \frac{1}{n} \log A_\epsilon(f^{-n}(x)) + (\lambda_1 + \epsilon) + \frac{1}{n} \log \|T_{-n}(x)v\|, \tag{2.24}$$

para todo $n \in \mathbb{N}$. Assim, pela **Afirmação 7**, concluímos que

$$\begin{aligned}
\liminf_{n \longrightarrow +\infty} \frac{1}{n} \log \|T_{-n}(x)v\| \geq -(\lambda_1 + \epsilon) &\Leftrightarrow \limsup_{n \longrightarrow +\infty} -\frac{1}{n} \log \|T_{-n}(x)v\| \leq \lambda_1 + \epsilon \\
&\Leftrightarrow \limsup_{n \longrightarrow -\infty} \frac{1}{n} \log \|T_n(x)v\| \leq \lambda_1 + \epsilon
\end{aligned}$$

e como ϵ foi escolhido de forma arbitrária, segue a afirmação.

Afirmação 9: $\liminf_{n \longrightarrow \pm\infty} \frac{1}{n} \log \|T_n(x)v\| \geq \lambda_1$**, para μ-q.t.p.** $x \in X$ **e** $v \in G_x$**.**

Fica plausível a afirmação acima para $\liminf$ quando $n \longrightarrow -\infty$.

Dado $v \in G_x$, temos por definição que

$$\limsup_{n \longrightarrow +\infty} \frac{1}{n} \log \|T_{-n}(x)v\| \leq -\lambda_1.$$

Assim,

$$-\limsup_{n \longrightarrow +\infty} \frac{1}{n} \log \|T_{-n}(x)v\| = \liminf_{n \longrightarrow +\infty} -\frac{1}{n} \log \|T_{-n}(x)v\| \geq \lambda_1$$

e portanto

$$\liminf_{n \longrightarrow -\infty} \frac{1}{n} \log \|T_n(x)v\| \geq \lambda_1. \tag{2.25}$$

Para o outro caso, seja $\epsilon > 0$ e para μ-q.t.p. $x \in X$ tome $v \in G_x \setminus \{0\}$ e $n \in \mathbb{N}$. Então,

$$\begin{aligned}
\|v\| &= \|T_{-n}(f^n(x))T_n(x)v\| \\
&= \|T_{-n}(f^n(x))T_n(x)v\| \exp(-n(-\lambda_1+\epsilon))\exp(n(-\lambda_1+\epsilon))\,\|T_n(x)v\|\,\|T_n(x)v\|^{-1} \\
&\leq C_\epsilon(f^n(x))\exp(n(-\lambda_1+\epsilon))\,\|T_n(x)v\|\,.
\end{aligned}$$

Aplicando a função log e dividindo por n a desigualdade acima, temos

$$\frac{1}{n}\log\|v\| \leq \frac{1}{n}\log C_\epsilon(f^n(x)) - (\lambda_1-\epsilon) + \frac{1}{n}\log\|T_n(x)v\|\,, \tag{2.26}$$

para todo $n \in \mathbb{N}$. Assim, novamente pela **Afirmação 7**, concluímos que

$$\liminf_{n\longrightarrow+\infty}\frac{1}{n}\log\|T_n(x)v\| \geq \lambda_1 - \epsilon$$

e como ϵ foi escolhido de forma arbitrária, segue a afirmação.

Desta forma, a terceira e última etapa esta concluída pelas **Afirmação 7**, **8** e **9**, encerrando a demonstração do Lema 2.10. ■

2.2.4 Demonstração do Lema 2.11

Consideremos inicialmente,

$$\Sigma = \left\{A : G^\perp \longrightarrow G;\ A_x : G_x^\perp \longrightarrow G_x \text{ é uma aplicação linear, para cada } x \in X\ \right\}.$$

Neste espaço de funções, consideremos o operador

$$\phi : \Sigma \longrightarrow \Sigma, \text{ tal que } \phi(A) = T^{-1}\circ A \circ \widehat{T},$$

isto é, para cada $x \in X$, temos

$$\phi(A)(x) = T^{-1}(f(x)) \circ A(f(x)) \circ \widehat{T}(x).$$

Seja a aplicação

$$P = (T|_{G^\perp} - \widehat{T}) : G^\perp \longrightarrow G, \text{ com } P(x)v = T(x)v - \widehat{T}(x)v,$$

notemos que para $v \in G_x^\perp$, temos $P(x)v \in G_{f(x)}$ e, pela invariância de G por T, segue que $B = -T^{-1}\circ P$ pertence à Σ. Por fim, consideremos a sequência de funções $\{\phi^n(B)\}_{n\in\mathbb{N}}$.

Afirmação 1: $\sum_{n=0}^{+\infty} \phi^n(B)$ **converge para uma função em** Σ **μ-q.t.p. .**

Para tal afirmação, se mostrarmos que existe uma função $C : X \longrightarrow \mathbb{R}$ mensurável e $\lambda < 0$ tal que

$$\|\phi^n(B)(x)\| \leq C(x)\exp(n\lambda),$$

para todo $n \geq 0$ e μ-q.t.p. $x \in X$, pelo critério da comparação, a afirmação estará provada.

Note que a série $\sum_{n=0}^{+\infty} C(x)\exp(n\lambda)$ é convergente, pois $\lambda < 0$.

Dado $\epsilon > 0$, defina

$$\widehat{A}_\epsilon(x) = \sup_{n\in\mathbb{N}} \left\{ \left\|\widehat{T}_n(x)\right\| \exp(-n(\lambda_1(\widehat{T}) + \epsilon)) \right\}$$

e

$$\widehat{C}_\epsilon(x) = \sup_{n\in\mathbb{N}} \left\{ \left\|(T^{-1}|_G)_n(x)\right\| \exp(-n(\lambda_1(T^{-1}|_G) + \epsilon)) \right\}.$$

Utilizando os mesmo argumentos das funções A_ϵ e C_ϵ (dadas na demonstração da 3ª etapa do Lema 2.10) também estão bem definidas. Mais ainda, pelo mesmo procedimento da **Afirmação 7** do Lema 2.10, temos para μ-q.t.p. $x \in X$ que

$$\lim_{n\longrightarrow\pm\infty} \frac{1}{n}\log(\widehat{C}_\epsilon \circ f^n(x)) = \lim_{n\longrightarrow\pm\infty} \frac{1}{n}\log(\widehat{A}_\epsilon \circ f^n(x)) = 0.$$

Notemos também que $\|B(x)\|$ é mensurável, com $\|\log B \circ f\| - \|\log B\|$ é integrável. Pelo Lema 1.6, teremos também que

$$\lim_{n\longrightarrow\pm\infty} \frac{1}{n}\log(\|B \circ f^n(x)\|) = 0.$$

Para cada $n \in \mathbb{N}$, temos $\phi^n(B)(x) = (T^{-1}|_G)_n(f^n(x)) \circ B(f^n(x)) \circ \widehat{T}_n(x)$ e, munido de todas estas afirmações acima, teremos que

$$\begin{aligned}
\|\phi^n(B)(x)\| &= \left\|(T^{-1}|_G)_n(f^n(x)) \circ B(f^n(x)) \circ \widehat{T}_n(x)\right\| \\
&\leq \left\|(T^{-1}|_G)_n(f^n(x))\right\| \|B(f^n(x)\| \left\|\widehat{T}_n(x)\right\| \\
&\leq \widehat{C}_\epsilon(f^n(x)) \|B(f^n(x))\| \widehat{A}_\epsilon(x)\exp(n(\lambda + 2\epsilon)). \qquad (2.27)
\end{aligned}$$

Agora, como $\lambda_1(T^{-1}|_G) + \lambda_1(\widehat{T}) < 0$ por hipótese, podemos tomar $\epsilon > 0$ tal que $\lambda = \lambda_1(T^{-1}|_G) + \lambda_1(\widehat{T}) + \epsilon < 0$. Da mesma, por escolha de $\widehat{C}_\epsilon(f^n(x))$ e $\|B(f^n(x))\|$, temos que a função

$$C(x) = \sup_{n\in\mathbb{N}} \widehat{C}_\epsilon(f^n(x)) \, \|B(f^n(x))\| \, \widehat{A}_\epsilon(x)$$

está bem definida. Por (2.27), temos

$$\|\phi^n(B)(x)\| \leq C(x)\exp(n\lambda), \text{ para todo } n \in \mathbb{N} \text{ e } \mu\text{-q.t.p. } x \in X,$$

o que conclui portanto a afirmação.

Tomando $A = \displaystyle\sum_{n=0}^{+\infty} \phi^n(B)$, observe que

$$A - \phi(A) = \sum_{n=0}^{+\infty} \phi^n(B) - \sum_{n=0}^{+\infty} \phi^{n+1}(B) = B.$$

Assim, temos que $-T^{-1} \circ P = A - \phi(A)$.

Definimos

$$H = \text{ graf } A = \left\{v \in G^\perp;\ v + Av \right\}.$$

Claramente $H \subset F$ é um subfibrado mensurável com $\dim H_x = \dim G_x^\perp$, para todo $x \in X$.

Afirmação 2: H é um subfibrado T-invariante de F.

Seja $v \in G^\perp$, então

$$\begin{aligned}
T(x)(v + Av) &= T(x)v + T(x)(A(x)v) \\
&= \widehat{T}(x)v + P(x)v + T(x)(A(x)v) \\
&= \widehat{T}(x)v + P(x)v + T(x)(A(x)v) + A(f(x))\widehat{T}(x)v - A(f(x))\widehat{T}(x)v \\
&= \widehat{T}(x)v + A(f(x))\widehat{T}(x)v - T(x)(\phi(A)(x)v) + T(x)(A(x)v) + P(x)v \\
&= \widehat{T}(x)v + A(f(x))\widehat{T}(x)v + T(x)(A(x)v - \phi(A)(x)v) + P(x)v \\
&= \widehat{T}(x)v + A(f(x))\widehat{T}(x)v + T(x)(B(x)v) + P(x)v \\
&= \widehat{T}(x)v + A(f(x))\widehat{T}(x)v - P(x)v + P(x)v \\
&= \widehat{T}(x)v + A(f(x))\widehat{T}(x)v,
\end{aligned}$$

com $\widehat{T}(x)v + A(f(x))\widehat{T}(x)v \in H_{f(x)}$, concluindo assim o fato que H é um subfibrado mensurável e T-invariante de F.

Terminaremos a prova do Lema 2.11 pela afirmação seguinte, lembrando que $\lambda_1(\widehat{T}) < \lambda_1(T^{-1}|_G)$ e que $-\lambda_1(T) = \lambda_1(T^{-1}|_G)$ (veja a **Afirmação 2** do Teorema 2.7).

Afirmação 3: $\lambda_1(T|_H) = \lambda_1(T)$.

Para tal, definimos

$$S : G^{\perp} \longrightarrow H, \text{ com } S(x)v = v + A(x)v, \; v \in G^{\perp}.$$

Observe que para todo $v \in G^{\perp}$, temos

$$(T \circ S)v = T(v + Av) = Tv + T \circ Av = \widehat{T}v + A \circ \widehat{T}v = S \circ \widehat{T}v.$$

Em particular,

$$(T_n \circ S)(x)v = S \circ \widehat{T}_n(x)v, \text{ para todo } v \in G^{\perp}.$$

Assim como a função B, S também é mensurável e

$$\lim_{n \longrightarrow \pm\infty} \frac{1}{n} \log \|(S(f^n(x))\| = 0.$$

Assim, temos

$$\begin{aligned}
\lambda_1(T|_H) &= \limsup_{n \longrightarrow +\infty} \frac{1}{n} \log \|(T|_H)_n(x)\| \\
&= \limsup_{n \longrightarrow +\infty} \frac{1}{n} \log \|S(f^n(x))\| \left\|\widehat{T}_n(x)\right\| \left\|S^{-1}(x)\right\| \\
&\leq \limsup_{n \longrightarrow +\infty} \frac{1}{n} \log(\|S(f^n(x))\| \left\|\widehat{T}_n(x)\right\| + \left\|S^{-1}(x)\right\| \\
&= \limsup_{n \longrightarrow +\infty} \frac{1}{n} \log \left\|\widehat{T}_n(x)\right\| \\
&= \lambda_1(\widehat{T}).
\end{aligned} \tag{2.28}$$

Utilizando a função S^{-1}, teremos que

$$\begin{aligned}
\lambda_1(\widehat{T}) &= \limsup_{n \longrightarrow +\infty} \frac{1}{n} \log \left\|\widehat{T}_n(x)\right\| \\
&= \limsup_{n \longrightarrow +\infty} \frac{1}{n} \log \left\|(S^{-1})(f(x))(T|_H)(x)S(x)\right\|
\end{aligned}$$

$$
\begin{aligned}
&\leq \limsup_{n \longrightarrow +\infty} \frac{1}{n} \log(\left\|(S^{-1})(f^n(x))\right\| + \|(T|_H)_n(x)\| + \|S(x)\|) \\
&= \limsup_{n \longrightarrow +\infty} \frac{1}{n} \log \|(T|_H)_n(x)\| \\
&= \lambda_1(T|_H). \qquad (2.29)
\end{aligned}
$$

Desta forma, por (2.28) e (2.29), respectivamente, temos que

$\lambda_1(\widehat{T}) \geq \lambda_1(T|_H)$ e $\lambda_1(\widehat{T}) \leq \lambda_1(T|_H)$, implicando que $\lambda_1(\widehat{T}) = \lambda_1(T|_H)$. ■

CAPÍTULO 3

A DESIGUALDADE DE RUELLE

Neste capítulo, vamos trabalhar um resultado básico da teoria ergódica, a desigualdade de Ruelle. Mais precisamente iremos relacionar a entropia com os expoentes de Lyapunov. Tal resultado foi publicado por David Ruelle, no boletim da Sociedade Brasileira de Matemática, em 1978, e tem o objetivo de relacionar o cálculo da entropia de um sistema (f, μ) invariante de classe C^1 com os expoentes de Lyapunov positivos (dados pelo teorema de Oseledets) de uma variedade compacta C^∞.

Definição 3.1. *Se* f *é um difeomorfismo de uma variedade compacta* X *sem bordo e* Λ *o conjunto de seus pontos regulares, definimos a função* $\chi : X \longrightarrow \mathbb{R}$ *por:*

$$\chi(x) = \sum_{\lambda_i(x) \geq 0} \lambda_i(x) \dim E_i(x).$$

Para o caso onde os expoentes de Lyapunov em x *são negativos, definimos por* $\chi(x) = 0$. *Quando* μ *é ergódica, temos que os expoentes de Lyapunov são constantes em quase todo ponto de* X, *assim como a dimensão dos autoespaços destes.*

Observação: Para facilitar a notação, tomaremos $\chi_i(x) = \lambda_i(x) \dim E_i(x)$.

Teorema 3.2. *[Desigualdade de Ruelle] Seja* f *um difeomorfismo de classe* C^1 *e* X *uma variedade Riemaniana compacta e* μ *ergódica. Então*

$$h_\mu(f) \leq \sum_{\lambda_i \geq 0} \lambda_i \dim E_i.$$

Demonstração: Para não complicarmos a notação e conseguirmos exprimir uma boa ideia da demonstração da desigualdade de Ruelle, vamos assumir que

$$h_\mu(f) = \sup_{\epsilon > 0} h_\mu(f, P_\epsilon),$$

onde P_ϵ são as partições de X formada por cubos com lados menores que $\epsilon > 0$.

Este fato pode ser verificado usando de triangulações e propriedades de entropia de partições.

Temos que existem constantes K, α e β tais que, para todo $A \in P_\epsilon$, vale

$$\text{diam } A \leq K\epsilon$$

e

$$\alpha\epsilon^d \leq \text{Vol } A \leq \beta\epsilon^d.$$

Consideremos a seguinte afirmação que fará uma relação com o isomorfismo Df com a partição P_ϵ.

Afirmação 1: Sejam $B \in P_\epsilon$, $\|Df\| = \sup_{x\in X}\{\|D_x f\|\}$ e $N \in \mathbb{N}$ suficientemente grande, tais que:

(a) Caso $\|Df\| \leq 1$, existe uma constante $C > 0$ tal que

$$\#\{A \in P_\epsilon;\ f^N(B) \cap A \neq \emptyset\} \leq C.$$

(b) Caso $\|Df\| > 1$, existe uma constante $C' > 0$, tal que

$$\#\{A \in P_\epsilon;\ f^N(B) \cap A \neq \emptyset\} \leq C' \|Df\|^{Nd},$$

Pela definição de P_ϵ e considerando $N \in \mathbb{N}$ suficientemente grande, temos que existe uma $K\epsilon$-vizinhança de $f^N(B)$, $V_{K\epsilon}^{f^N(B)}$, tal que se $f^N(B) \cap A \neq \emptyset$, com $A \in P_\epsilon$, então A está contido em $V_{K\epsilon}^{f^N(B)}$.

Pelo fato de f ser um difeomorfismo de classe C^1, pela desigualdade do Valor Médio temos que

$$\text{diam } (f^n(B)) \leq \tilde{K} \|Df\|^n \text{diam } B, \tag{3.1}$$

para todo $n \in \mathbb{N}$, onde $\tilde{K}$ é uma constante que depende apenas de X. Assim, em particular, podemos concluir que se $x, y \in V_{K\epsilon}^{f^N(B)}$, temos

$$\begin{aligned} d(x,y) \leq 2K\epsilon + \text{diam } f^N(B) &\leq 2K\epsilon + K\epsilon(\tilde{K} \|Df\|^N) \\ &\leq K\epsilon(2 + \tilde{K} \|Df\|^N), \end{aligned}$$

isto é,

$$\text{diam } V_{K\epsilon}^{f^N(B)} \leq K\epsilon(2+\tilde{K}\,\|Df\|^N).$$

Observe que existe $l>0$ dependendo apenas de X tal que

$$\text{Vol } D \leq (l \text{ diam } D)^d, \tag{3.2}$$

para todo conjunto $D \subset X$ mensurável. À fortiori,

$$\begin{aligned}
\#\{A \in P_\epsilon; A\cap\ f^N(B) \neq \emptyset\}\alpha\epsilon^d &\leq \sum_{A\in P_\epsilon;\ A\cap f^N(B)\neq\emptyset} \text{Vol } A \\
&\leq \text{Vol }\left(V_{K\epsilon}^{f^N(B)}\right) \\
&\leq \left(l \text{ diam } V_{K\epsilon}^{f^N(B)}\right)^d \\
&\leq (lK\epsilon(2+\tilde{K}\,\|Df\|^N))^d.
\end{aligned}$$

Portanto,

$$\#\{A \in P_\epsilon; A\cap\ f^N(B) \neq \emptyset\} \leq \frac{1}{\alpha}\left(lK\left(2+\tilde{K}\,\|Df\|^N\right)\right)^d. \tag{3.3}$$

Para o caso em que $\|Df\| \leq 1$, segue de (3.3) que

$$\#\{A \in P_\epsilon;\ f^N(B)\cap A \neq \emptyset\} \leq \frac{(2lK + lK\tilde{K})^d}{\alpha}, \tag{3.4}$$

concluindo assim o item **(a)** da **Afirmação 1**, se tomarmos $C = (2lK + lK\tilde{K})^d/\alpha$.

Caso $\|Df\| > 1$, existe $N(f) = ([\ln(2)/\ln\|Df\|] + 1) \in \mathbb{N}$ tal que $\|Df\|^N \geq 2$, para todo $N \geq N(f)$ e

$$\#\{A \in P_\epsilon;\ f^N(B)\cap A \neq \emptyset\} \leq \frac{\left(2K\,\|Df\|^N\right)^d}{\alpha},$$

sendo $[\ln(2)/\ln\|Df\|]$ a parte inteira de $\ln(2)/\ln\|Df\|$.
Tomando $C' = (2K)^d/\alpha$, concluímos o item **(b)** da **Afirmação 1**.

Se considerarmos a invariância de μ por f (e pelo fato de f ser invertível), a partição P_ϵ de X e $H_\mu(P_\epsilon|_{f^n(B)})$ a entropia de P_ϵ com respeito a medida condicional da n-ésima imagem de $B \in P_\epsilon$, utilizando o Lema 1.32 e a Proposição 1.20, obtemos o seguinte resultado:

$$
\begin{aligned}
h_\mu(f^n, P_\epsilon) &= \lim_{k\longrightarrow\infty} H_\mu\left(P_\epsilon \mid \bigvee_{i=1}^{k} f^{-ni}(P_\epsilon)\right)\\
&\leq H_\mu(P_\epsilon|f^n(P_\epsilon))\\
&= \sum_{B\in P_\epsilon} H_\mu(P_\epsilon|_{f^n(B)})\mu(f^n(B))\\
&= \sum_{B\in P_\epsilon} H_\mu(P_\epsilon|_{f^n(B)})\mu(B)\\
&\leq \sum_{B\in P_\epsilon} \log(\#\{A\in P_\epsilon;\ A\cap f^n(B)\neq\emptyset\})\mu(B), \qquad (3.5)
\end{aligned}
$$

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

Pelo item **(a)** da **Afirmação 1** e da desigualdade em (3.5), temos que existe $N \in \mathbb{N}$ suficientemente grande tal que

$$
\begin{aligned}
h_\mu(f^N, P_\epsilon) &\leq \sum_{B\in P_\epsilon} \log(\#\left\{A\in P_\epsilon;\ A\cap f^N(B)\neq\emptyset\right\})\mu(B)\\
&\leq \sum_{B\in P_\epsilon} \log(C)\mu(B)\\
&= \log C.
\end{aligned}
$$

No mais, pela Proposição 1.31, temos

$$h_\mu(f, P_\epsilon) \leq \frac{\log C}{N}.$$

Quando fazemos N tender a infinito, teremos que $h_\mu(f, P_\epsilon) = 0$. Assim, dado $\delta > 0$ e escolhendo $\epsilon(N)$ de modo que $h_\mu(f) = h_\mu(f, P_\epsilon) + \delta$, com $N \in \mathbb{N}$ suficientemente grande, teremos que analogamente que $h_\mu(f) = 0$.

Em vista disto, vamos considerar apenas o caso em que $\|Df\| > 1$.

Afirmação 2: Seja $B \in P_\epsilon$ e, considerando $N \in \mathbb{N}$ suficientemente grande tal que sua N-ésima imagem por f contém um ponto $x \in \Lambda$, existe uma constante $C'' > 0$ e $\tilde{\delta} > 0$ tais que

$$\#\left\{A\in P_\epsilon;\ A\cap f^N(B)\neq\emptyset\right\} \leq C'' \prod_{\chi_i>0} \exp(N(\chi_i + \tilde{\delta})).$$

Antes de mais nada, fixado $N \in \mathbb{N}$ grande, observemos pelo Lema 2.10 que dado $\delta > 0$, existe $\tilde{\delta} > 0$ e $\tilde{\Lambda} \subset \Lambda$, com $\mu(\tilde{\Lambda}) = 1 - \delta$ tal que se $x \in \tilde{\Lambda}$, temos

$$\left\| D_x f^N(x) v_i \right\| \leq m \exp(N(\chi_i + \tilde{\delta})) \left\| v_i \right\|, \tag{3.6}$$

para todo $v_i \in E_i$, onde $T_x X = E_1 \oplus \cdots \oplus E_j$ é a decomposição de Oseledets para $x \in \tilde{\Lambda} \subset \Lambda$.

Agora, sem perda de generalidade, suponha que $x \in B \cap \tilde{\Lambda}$, com $B \in P_\epsilon$. Observe que $B \subset \mathcal{B}(x, K\epsilon)$ e que por escolha de P_ϵ, temos $A \subset V_{K\epsilon}(f^N(\mathcal{B}(x, K\epsilon)))$, para todo $A \in P_\epsilon$ tal que $A \cap f^N(B) \neq \emptyset$. Desta forma,

$$\begin{aligned} \# \left\{ A \in P_\epsilon;\ A \cap f^N(B) \neq \emptyset \right\} \alpha \epsilon^d &\leq \sum_{A \in P_\epsilon;\ A \cap f^N(B) \neq \emptyset} \text{Vol } A \\ &\leq \text{Vol } V_{K\epsilon}(f^N(\mathcal{B}(x, K\epsilon))). \end{aligned} \tag{3.7}$$

Escolhendo $\epsilon > 0$ menor, se necessário, podemos assumir que f^N e Df^N possuem um comportamento parecido na vizinhança, mais precisamente,

$$\begin{aligned} \text{Vol } V_{K\epsilon}(f^N(\mathcal{B}(x, K\epsilon))) &\leq K' \text{ Vol } (f^N(x) + V_{K\epsilon}(D_x f^N(\mathcal{B}(0, K\epsilon)))) \\ &= K' \text{ Vol } V_{K\epsilon}(D_x f^N(\mathcal{B}(0, K\epsilon))), \end{aligned}$$

onde $\mathcal{B}(0, K\epsilon) \subset T_x X$.

Considerando a decomposição $T_x X = E_i \oplus \cdots \oplus E_j$ dado pela decomposição de Oseledets e como $T_{f^N(x)} X = Df^N(E_1) \oplus \cdots \oplus Df^N(E_j)$, existe uma constante $K'' > 0$ tal que

$$\text{Vol } V_{K\epsilon}(D_x f^N(\mathcal{B}(0, K\epsilon))) \leq K'' \prod_{i=1}^{j} \left\| D_x f^N v_i \right\|, \tag{3.8}$$

para $v_i \in E_i$, com $\|v_i\| = K\epsilon$ e $\{v_1, \ldots, v_d\}$ a base de $T_x X$.

Como N é suficientemente grande, podemos considerar que

$$\left\| D_x f^N v_i \right\| \leq \tilde{\delta},$$

se $\chi_i \leq 0$. Portanto, por (3.6) e (3.8) temos

$$\text{Vol } V_{K\epsilon}(D_x f^N(\mathcal{B}(0, K\epsilon))) \leq K'' \prod_{\chi_i > 0} m \exp(N(\chi_i + \tilde{\delta})) \|v_i\|$$

$$\leq K''m^d(K\epsilon)^d \prod_{\chi_i>0} \exp(N(\chi_i + \tilde{\delta})).$$

Tomando $C'' = K''m^d(K\epsilon)^d$, temos a **Afirmação 2** provada.

De (3.5), podemos chegar no seguinte resultado:

$$h_\mu(f^N, P) \leq \sum_{B\in P_\epsilon} \log(\#\left\{A \in P_\epsilon;\ A \cap f^N(B) \neq \emptyset\right\})\mu(B)$$
$$= \sum_{B\cap\Lambda\neq\emptyset} \log(\#\left\{A \in P_\epsilon;\ A \cap f^N(B) \neq \emptyset\right\})\mu(B) + \sum_{B\cap\Lambda=\emptyset} \log(\#\left\{A \in P_\epsilon;\ A \cap f^N(B) \neq \emptyset\right\})\mu(B). \tag{3.9}$$

Pela Proposição 1.31 e pelas **Afirmações 1** e **2**, concluímos

$$\begin{aligned} Nh_\mu(f) - \delta &= h_\mu(f^N) - \delta \\ &< h_\mu(f^N, P_\epsilon) \\ &\leq \sum_{B\cap\Lambda\neq\emptyset} \log(C'' \prod_{\chi_i>0} \exp(N(\chi_i + \tilde{\delta})))\mu(B) + \sum_{B\cap\Lambda=\emptyset} \log(C' \|Df\|^{Nd})\mu(B) \\ &\leq \log C'' + N \sum_{\chi_i>0} (\chi_i + \tilde{\delta}) + (\log C' + dN \log \|Df\|)\, \mu(X - \Lambda), \end{aligned} \tag{3.10}$$

para algum $N \in \mathbb{N}$ suficientemente grande.

Pelo Teorema 2.2, podemos concluir que $\mu(X - \Lambda) = 0$ e, pelos resultados em (3.9) e (3.10), temos

$$h_\mu(f) \leq \delta + \sum_{\chi_i>0} \chi_i,$$

quando $N \longrightarrow \infty$. A demonstração do Teorema 3.2 se encerra quando fazemos $\delta \longrightarrow 0$.
$\blacksquare$

REFERÊNCIAS BIBLIOGRÁFICAS

[1] ALVES, F. F. *Teorema Ergódico Multiplicativo de Oseledets.* 71 p. Dissertação (Mestrado em Matemática) — UNESP - Campus São José do Rio Preto - Instituto de Biociências, Letras e Ciências Exatas, São José do Rio Preto, São Paulo, 2010.

[2] ALVES, J. F.; BONATTI, C.; VIANA, M. Srb measures for partially hyperbolic systems whose central direction is mostly expanding. *Inventiones Mathematicae*, v. 140, p. 351-398, 2000.

[3] BARTLE, R. G. *The Elements of Integration and Lebesgue Measure.* 1. ed. New York: John Wiley & Sons, INC., 1995.

[4] KATOK, A.; HASSELBLATT, B. *Introduction to the modern theory of dynamical systems.* Cambridge University Press, 1995. v. 54. (Encyclopedia of Mathematics and its Applications, v. 54). With a supplementary chapter by Katok and Leonardo Mendoza.

[5] LIMA, E. L. *Variedades Diferenciáveis.* Rio de Janeiro: IMPA, 2014. Publicações Matemáticas.

[6] MAÑÉ, R. *Introdução à Teoria Ergódica.* Rio de Janeiro: Instituto de Matemática Pura e Aplicada, CNPq, 1983.

[7] OLIVEIRA, K.; VIANA, M. *Fundamentos da Teoria Ergódica.* 1. ed. Rio de Janeiro: SBM, 2014.

[8] RUELLE, D. *An inequality for the entropy of differentiable maps.* Bol. Soc. Bras. Mat., 1978, v. 9, p. 83-87. Disponível em

<http://www.ihes.fr/~/ruelle/PUBLICATIONS/%5B51%5D.pdf>. Acesso em 01 abr. 2017.

[9] OSELEDETS, V. I. *A multiplicative ergodic theorem. Characteristic Ljapunov, exponents of dynamical systems.* Tr. Mosk. Mat. Obs., **19**, MSU, M., 1968, 179-210. Disponível em <http://www.mathnet.ru/links/08b04fd2d575969b661d22259794057e/mmo214.pdf>. Acesso em 01 abr. 2017.